STOCHASTIC DOMINANCE IN PORTFOLIO EFFICIENCY TESTING

Miloš Kopa

HABILITATION THESIS
Mathematics–Probability and Mathematical Statistics

Charles University in Prague, Faculty of Mathematics and Physics

May 2013

Preface

The habilitation thesis deals with portfolio efficiency testing with respect to stochastic dominance criteria. The main author's results published in attached papers are summarized in a unified manner in Chapters 2-5 what makes this publication more compact. Chapter 1 serves as an introduction to the decision making theory under risk. All author's results related to stochastic dominance in portfolio efficiency may be found in the papers attached at CD. The main body of this habilitation thesis is divided into four chapters corresponding, respectively, to the most important areas of portfolio efficiency testing with respect to stochastic dominance criteria:

1. Chapter 2: Portfolio efficiency tests with respect to the first-order stochastic dominance (FSD) criterion

2. Chapter 3: Portfolio efficiency tests with respect to the second-order stochastic dominance (SSD) criterion

3. Chapter 4: Portfolio efficiency tests with respect to higher order stochastic dominance (NSD) criteria

4. Chapter 5: Robustness and contamination in portfolio efficiency tests with respect to the second-order stochastic dominance criterion

The thesis is mainly based on five papers published by the author and few coauthors during the years 2008–2013:

- The first paper, **Kopa, M. and Post, T. (2009): A portfolio efficiency test based on FSD optimality, Journal of Financial and Quantitative Analysis, 44, 5, 1103 – 1124**, presents a new approach to portfolio efficiency testing based on the first-order stochastic dominance optimality. It introduces a necessary and sufficient condition for FSD optimality of a given portfolio that is based on a set of representative utility functions. The suggested mixed-integer linear programming algorithm for FSD portfolio optimality testing and some necessary conditions for FSD portfolio optimality have been recognized as a crucial step for applications, especially in finance. In the paper, it is applied to US market portfolio efficiency analysis. Surprisingly, using this new testing procedures, the market portfolio is classified as FSD non-optimal. Moreover, nine more attractive portfolios are identified. It means that every decision maker prefers at least one of them to the US market portfolio what is an important and unexpected result.

- The second paper, **Kopa, M. and Chovanec, P. (2008): A second-order stochastic dominance portfolio efficiency measure, Kybernetika, 44, 2, 243 - 258**, deals with portfolio efficiency tests with respect to the second-order stochastic dominance (SSD). It presents a new SSD portfolio efficiency test, that is based on a dual stochastic dominance representation. The test also gives the information about possible improvement of investment strategy such that every risk averse decision maker would prefer it to the tested portfolio. Moreover, the paper suggests a new tool for measuring the degree of SSD portfolio inefficiency of a given portfolio which is consistent with the second-order stochastic dominance relation.

- The third paper, **Post, T. and Kopa, M. (2013): General Linear Formulations of Stochastic Dominance Criteria, European Journal of Operational Research, Online first**, applies the higher order stochastic dominance criteria to the portfolio efficiency testing. The approach is based on a piece-wise polynomial representation of utility functions and their derivatives and can be implemented by solving a relatively small system of linear inequalities. Moreover, linear dual formulations in terms of lower partial moments and co-lower partial moments are derived. Comparing to all previous papers dealing with portfolio efficiency with respect to stochastic dominance the new tests applies for any order ($N \geq 2$) of stochastic dominance and also for convex stochastic dominance approach.

- Finally, the fourth paper, **Kopa (2010): Measuring of second-order stochastic dominance portfolio efficiency, Kybernetika,**

46, 3, 488 - 500, presents some possible robustness extensions of existing SSD portfolio efficiency tests. Stability of SSD efficiency classification with respect to changes in probability measure is analysed. Contrary to all previous SSD portfolio efficiency tests focused on SSD inefficiency measuring, the paper introduces a new measure of SSD efficiency applicable for SSD efficient portfolios. Moreover, another new SSD portfolio efficiency tests allowing small changes or contaminations of the original probability distribution are derived in the last paper: **Dupačová, J. and Kopa, M. (2012): Robustness in stochastic programs with risk constraints, Annals of Operations Research 200, 1, 55 - 74**.

The support of my family, my co-authors, my colleagues, and grants MSM0021620839, GA201/07/P107, GA402/10/1610, P402/12/0558 is gratefully acknowledged.

Prague, May 2013 *Miloš Kopa*

Contents

Chapter 1

Introduction to decision making under risk

The theory of decision making under risk is one of the most appealing issues within financial mathematics. It is based on the basic economical principles, however it also exploits optimization techniques and statistical tools. Mathematical formulations of decision making problems under risk lead to stochastic programming models which are searching for the optimal solution with respect to a chosen objective (criterion) and feasibility constraints. In financial applications, they turn out to so called portfolio selection problems. These problems basically capture two fundamental principles: non-satiation and risk attitude. While the non-satiation axiom is easy to implement and generally accepted in economics and finance, the risk attitude can be understood or expressed in various ways.

The first portfolio selection problem was introduced by Markowitz (1952). The model jointly focuses on maximizing expected return and minimizing variance of the portfolio, where variance serves as a measure of risk. Consider N risky assets with returns modeled by random vector $\boldsymbol{\varrho}$. Let $\boldsymbol{\lambda} \in \mathbb{R}^M$ be a vector of weights determining the way how the initial wealth is invested. Following Markowitz (1952), short sales are allowed, that is, the set of all feasible portfolios Λ is given by condition: $\mathbf{1}'\boldsymbol{\lambda} = 1$. Consequently, the return and variance of portfolio (with weights) $\boldsymbol{\lambda}$ are $E(\boldsymbol{\varrho}'\boldsymbol{\lambda})$ and $\mathrm{var}(\boldsymbol{\varrho}'\boldsymbol{\lambda})$. Hence, Markowitz portfolio selection problem may be formulated as follows:

$$\max_{\boldsymbol{\lambda} \in \Lambda} E(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \vartheta \mathrm{var}(\boldsymbol{\varrho}'\boldsymbol{\lambda})$$

where $\vartheta \geq 0$ is a risk aversion parameter.

In the last 60 years, measuring of portfolio risks has become very important, especially during the crises periods. The suitability of variance as a measure of risk was analyzed. It was shown that variance suffered from several shortcomings unless the Gaussian distribution of returns was assumed. Therefore some other measures of risks were introduced, for example: semi-variance, mean absolute deviation, mean absolute semi-deviation, Gini mean,... In the last 15 years Value at Risk (VaR) and Condition Value at Risk (CVaR) have proved to be the most popular risk measures because of their meaningful financial interpretation and nice theoretical properties. Let $F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)$ be the cumulative probability distribution function of portfolio $\boldsymbol{\lambda}$. Then Value at Risk of portfolio $\boldsymbol{\lambda}$ at level α is defined as the α-quantile of the portfolio loss, that is,

$$\mathrm{VaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) = F^{(-1)}_{-\boldsymbol{\varrho}'\boldsymbol{\lambda}}(v) = \min\{u : F_{-\boldsymbol{\varrho}'\boldsymbol{\lambda}}(u) \geq v\}. \tag{1.1}$$

Unfortunately, Value of Risk does not fulfill all generally accepted coherence assumptions (monotonicity, translation invariance, positive homogeneity and subadditivity, cf. Artzner, et al. (1999)). Moreover, it does not take into account the losses that occur with probability smaller than $1 - \alpha$. Therefore Rockafellar and Uryasev (2000, 2002) introduced Conditional Value at Risk as an example of a coherent risk measure derived from VaR. It is also known under the names Average Value at Risk or Tail Value at Risk. It can be computed as follows:

$$\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \;\; = \;\; \min_{a\in\mathbb{R}} \;\; a + \frac{1}{1-\alpha} E(-\boldsymbol{\varrho}'\boldsymbol{\lambda} - a)^+ \tag{1.2}$$

Economically, $\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$ equals the mean of losses that are higher than $\mathrm{VaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$. See Rockafellar and Uryasev (2000, 2002) for more details. Alternatively, a risk may be expressed as a deviation from the expected value of the portfolio return. Recently, Rockafellar, Uryasev and Zabarankin (2006) introduced general deviation measures as functionals $D(\boldsymbol{\varrho}'\boldsymbol{\lambda})$ that are translation equivariant, positively homogeneous, subadditive and monotone. Contrary to the coherent risk measures $R(\boldsymbol{\varrho}'\boldsymbol{\lambda})$, deviation measures are not affected by the expected value. As shown in Rockafellar, Uryasev and Zabarankin (2006), deviation measures correspond one-to-one with strictly expectation bounded risk measures (i.e. translation invariant, positively homogeneous, subadditive risk measures satisfying $R(\boldsymbol{\varrho}'\boldsymbol{\lambda}) > -E\boldsymbol{\varrho}'\boldsymbol{\lambda}$) under the relations: $R(\boldsymbol{\varrho}'\boldsymbol{\lambda}) = D(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - E\boldsymbol{\varrho}'\boldsymbol{\lambda}$ or $D(\boldsymbol{\varrho}'\boldsymbol{\lambda}) = R(\boldsymbol{\varrho}'\boldsymbol{\lambda} - E\boldsymbol{\varrho}'\boldsymbol{\lambda})$.

Using risk measure $R(\boldsymbol{\varrho}'\boldsymbol{\lambda})$ and mean return $E\boldsymbol{\varrho}'\boldsymbol{\lambda}$, one can view the portfolio selection problem as a bi-objective optimization model that can be formulated as:

$$\max_{\boldsymbol{\lambda}\in\Lambda} E(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \vartheta R(\boldsymbol{\varrho}'\boldsymbol{\lambda})$$

for some set of feasible portfolios Λ, where $\vartheta \geq 0$ is a risk aversion parameter. However, finding an appropriate parameter value ϑ may be a difficult task in some empirical applications. If so, one can use an alternative mean-risk formulation, for example:

$$\min_{\boldsymbol{\lambda}\in\Lambda, E\boldsymbol{\varrho}'\boldsymbol{\lambda}\geq\mu} R(\boldsymbol{\varrho}'\boldsymbol{\lambda}).$$

which focuses on risk minimization under condition on mean return.

Alternatively, one may model the risk attitude using utility functions introduced in von Neumann and Morgenstern (1944). Contrary to the bi-objective mean-risk models, application of utility function leads to classical nonlinear programming problem - to maximize expected utility of the final wealth. Unfortunately, identifying the particular utility function of the decision maker is usually very difficult. When utility function is not precisely known, one can consider at least some suitable, economically meaningful classes of utility functions, for example the class of all nondecreasing and concave utility functions. In this case, the optimal solution of the maximizing expected utility problem can not be exactly found. However, one can at least compare two portfolios between each other. If one of them gives higher or equal expected utility of final wealth (or returns) than the other one for all considered utility functions then a relation of stochastic dominance (SD) between them exists. The notion of stochastic dominance was introduced in statistics more than 50 years ago and it was firstly applied to economics and finance in Quirk and Saposnik (1962), Hadar and Russell (1969) and Hanoch and Levy (1969). The basic principles of stochastic dominance theory and applications to portfolio efficiency are summarized in the following sections.

1.1 First-order stochastic dominance relation

If only non-satiation of decision maker is assumed, that is, the set of all non-decreasing utility functions (U_1) is considered, the first-order stochastic dominance (FSD) relation allows to compare two portfolios between each other. For two portfolios $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ with respective cumulative probability distribution functions of their returns $F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)$, $F_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(x)$ we say that portfolio $\boldsymbol{\lambda}$ *dominates* portfolio $\boldsymbol{\tau}$ *by the first-order stochastic dominance*: $\boldsymbol{\lambda} \succ_{FSD} \boldsymbol{\tau}$ if

$$Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0$$

for all utility functions $u \in U_1$, such that these expected values exist, with strict inequality for at least one $u \in U_1$.[1] In the investment theory, it means that at least one non-satiated decision maker prefers portfolio $\boldsymbol{\lambda}$ to $\boldsymbol{\tau}$ and the others are indifferent between them or prefer $\boldsymbol{\lambda}$ to $\boldsymbol{\tau}$, too. The corresponding weak first-order stochastic dominance relation ($\boldsymbol{\lambda} \succeq_{FSD} \boldsymbol{\tau}$) does not require existence of a utility function $u \in U_1$ that satisfies:

$$Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) > 0.$$

According to Russel and Seo (1989), U_1 may be represented by a set of simple utility functions in the following sense:

$$Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0 \quad \forall u \in U_1 \iff Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0 \quad \forall u \in V$$

where

$$V = \{u_{\eta,\nu}(x) : u_{\eta,\nu}(x) = \max\{\nu, \min\{x - \eta, 0\}\}, \eta \in \mathbb{R}, \nu \in \mathbb{R}^-\}. \tag{1.3}$$

The first computable necessary and sufficient conditions for the first-order stochastic dominance relation were proposed in Hanoch and Levy (1969):

- $\boldsymbol{\lambda} \succeq_{FSD} \boldsymbol{\tau} \quad \Leftrightarrow \quad F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x) \leq F_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(x) \quad \forall x \in \mathbb{R}$
- $\boldsymbol{\lambda} \succ_{FSD} \boldsymbol{\tau} \quad \Leftrightarrow \quad F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x) \leq F_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(x) \quad \forall x \in \mathbb{R}$ where at least one strict inequality holds.

Consider now the *quantile model of stochastic dominance*, see Ogryczak and Ruszczyński (2002). Applying a quantile function and Value at Risk (1.1), one can reformulate the necessary and sufficient conditions for FSD relation:

$$\begin{aligned} \boldsymbol{\lambda} \succeq_{FSD} \boldsymbol{\tau} \quad &\Leftrightarrow \quad F^{(-1)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(p) \geq F^{(-1)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(p) \quad \forall p \in (0, 1\rangle. &(1.4)\\ &\Leftrightarrow \quad \mathrm{VaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \leq \mathrm{VaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \quad \forall \alpha \in (0, 1\rangle. &(1.5) \end{aligned}$$

[1]This relation is sometimes called "strict FSD", see Levy (2006).

1.2 Second-order stochastic dominance relation

The second-order stochastic dominance relation is generated by the set U_2 of all concave utility functions. These functions express the risk attitude of risk averse decision makers. For two portfolios $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ with respective cumulative probability distributions functions we say that $\boldsymbol{\lambda}$ *dominates* $\boldsymbol{\tau}$ *by the second-order stochastic dominance*: $\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\tau}$ if

$$Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0$$

for every $u \in U_2$ with strict inequality for at least one $u \in U_2$. That is, every non-satiated and risk averse investor prefers $\boldsymbol{\lambda}$ to $\boldsymbol{\tau}$ or is indifferent between them and at least one investor (strictly) prefers $\boldsymbol{\lambda}$ to $\boldsymbol{\tau}$. Similarly to the FSD case, the weak SSD ($\boldsymbol{\lambda} \succeq_{SSD} \boldsymbol{\tau}$) occurs if the strict preference is not required. Russel and Seo (1989) proved that U_2 may be replaced by a set of concave and two-piece linear utility functions:

$$Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0 \quad \forall u \in U_2 \iff Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0 \quad \forall u \in V_2$$

where $V_2 = \{u_\eta(x) : u_\eta(x) = \min\{x - \eta, 0\}, \eta \in \mathbb{R}\}$.

Consider now a *twice cumulated distribution function* of returns of portfolio $\boldsymbol{\lambda}$:

$$F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t) = \int_{-\infty}^{t} F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)dx$$

and the corresponding *second quantile function*:

$$\begin{aligned} F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(p) &= \int_{-\infty}^{p} F^{(-1)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t)dt && \text{for } 0 < p \leq 1 \\ &= 0 && \text{for } p = 0 \\ &= +\infty && \text{otherwise.} \end{aligned} \tag{1.6}$$

As shown in Bawa (1975), the twice cumulated distribution function is equal to the first-order lower partial moment, that is:

$$F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t) = E\max(t - \boldsymbol{\varrho}'\boldsymbol{\lambda}, 0)$$

and Ogryczak and Ruszczyński (2002) proved that the second quantile function is its Fenchel dual term. Moreover, CVaR can be computed from the second quantile function as follows:

$$\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) = -\frac{F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(1-\alpha)}{1-\alpha}$$

for all $\alpha \in \langle 0,1)$, see Ogryczak and Ruszczyński (2002) and Kopa and Chovanec (2008).

Similarly as in the FSD case, several necessary and sufficient conditions for the second-order stochastic dominance relation can be derived, see e.g. Hanoch and Levy (1969), Levy (2006) or Ogryczak and Ruszczyński (2002):

(i) $\boldsymbol{\lambda} \succeq_{SSD} \boldsymbol{\tau}$ if and only if $F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t) \leq F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(t), \quad \forall t \in \mathbb{R}$,

(ii) $\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\tau}$ if and only if $F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t) \leq F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(t), \quad \forall t \in \mathbb{R}$ with strict inequality for at least one $t \in \mathbb{R}$,

(iii) $\boldsymbol{\lambda} \succeq_{SSD} \boldsymbol{\tau}$ if and only if $F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(p) \geq F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(p)$ for all $p \in \langle 0,1\rangle$,

(iv) $\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\tau}$ if and only if $F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(p) \geq F^{(-2)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(p)$ for all $p \in \langle 0,1\rangle$ with strict inequality for at least one p,

(v) $\boldsymbol{\lambda} \succeq_{SSD} \boldsymbol{\tau}$ if and only if $\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \leq \mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau})$ for all $\alpha \in \langle 0,1\rangle$,

(vi) $\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\tau}$ if and only if $\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \leq \mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau})$ for all $\alpha \in \langle 0,1\rangle$ with strict inequality for at least one α.

1.3 Other stochastic dominance relations

In previous sections, stochastic dominance relations were generated by U_1 or U_2. In general, one can consider any subset of U_1. Let $U_N \subset U_1$ be the set of N times differentiable utility functions such that: $(-1)^k u^{(k)} \leq 0$ for all $k = 1, 2, ..., N$. As the limiting case, $U_\infty \subset U_1$ is the set of infinitely differentiable utility functions with alternating signs of the derivatives. These functions are also called *completely monotonic utility functions*. More details about completely monotonic utility functions can be found in Whitmore (1989) and references therein. For example, all hyperbolic absolute risk aversion (HARA) utility functions are completely monotonic. Especially, exponential, logarithmic or power utility functions, as special elements of HARA class, are very popular in financial applications. In general, utility function $u(x)$ belongs to HARA class if

$$-\frac{u(x)^{(2)}}{u(x)^{(1)}} = \frac{1}{ax+b}, \quad ax+b>0.$$

For each set $U_N \subset U_1$, $N = 3, 4, ...$ we can define weak Nth-order stochastic dominance relation between portfolios $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ as a generalization of weak first-order SD relation: $\boldsymbol{\lambda} \succeq_{NSD} \boldsymbol{\tau}$ if $\quad Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0$ for every utility function $u \in U_N$. If strict inequality holds true for at least one $u \in U_N$ then the relation is called Nth-order stochastic dominance.

For random variables with bounded support (a, b) Levy (2006) presents a necessary and sufficient condition for weak Nth-order stochastic dominance relation in terms of N times cumulated distribution functions

$$F^{(N)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(t) = \int_{-\infty}^{t} F^{(N-1)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)dx$$

as follows: $\boldsymbol{\lambda} \succeq_{NSD} \boldsymbol{\tau}$ if and only if $F^{(k)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(b) \leq F^{(k)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(b)$, for all $k = 2, 3, ..., N-1$ and $F^{(N)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x) \leq F^{(N)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(x)$ for all $x \in \langle a, b\rangle$. Again, in the case of NSD relation, at least one strict inequality is required.

The infinite-order stochastic dominance relation (ISD) is defined as a limit case of NSD when $N \to \infty$: $\boldsymbol{\lambda} \succeq_{ISD} \boldsymbol{\tau}$ if $\quad Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0$ for every utility function $u \in U_\infty$ and $\boldsymbol{\lambda} \succ_{ISD} \boldsymbol{\tau}$ if $\quad Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) \geq 0$ for every utility function $u \in U_\infty$ and strict inequality holds for at least one $u \in U_\infty$.

For infinite-order of stochastic dominance, Whitmore (1989) derived a necessary and sufficient condition based on the Bernstein Theorem: $\boldsymbol{\lambda} \succ_{ISD} \boldsymbol{\tau}$ if and only if

$$E\left(e^{-a\boldsymbol{\varrho}'\boldsymbol{\lambda}} - e^{-a\boldsymbol{\varrho}'\boldsymbol{\tau}}\right) \leq 0 \text{ for all } a \geq 0.$$

The recursive nature of the definition of the N-times cumulated distribution function makes it clear that (weak) Nth-order stochastic dominance implies (weak) $\overline{N}$th-order stochastic dominance for $\overline{N} > N$. Moreover, any Nth-order stochastic dominance implies infinite-order stochastic dominance. Thistle (1993) also proved that if $\boldsymbol{\lambda} \succ_{ISD} \boldsymbol{\tau}$ then $\boldsymbol{\lambda} \succ_{NSD} \boldsymbol{\tau}$ for some finite N.

1.4 Portfolio efficiency with respect to stochastic dominance criteria

From a theoretical perspective, stochastic dominance analysis is an appealing approach to analyzing investment and portfolio choice problems when the utility function of decision maker is not precisely known. Most notably, the most popular criterion of the second-order stochastic dominance (SSD) assumes only non-satiation and risk aversion for the investor's preferences and it in effect considers the entire return distribution rather than a finite set of moments. Until recently, practical applications of stochastic dominance relations were hampered by the absence of tractable algorithms to deal with diversification across multiple assets. The criteria for pairwise comparisons have only a limited use when looking for efficient portfolios with respect to SD relations. In the last decade, under assumptions of discrete distribution of returns, Post (2003), Kuosmanen (2004) and Kopa and Chovanec (2008) developed linear programming tests to analyze if a portfolio is efficient relative to all possible portfolios formed from the considered assets. The tests are formulated for the case when no short sales are allowed, that is :

$$\Lambda = \{\boldsymbol{\lambda} \in R^M | \mathbf{1}'\boldsymbol{\lambda} = 1, \ \lambda_m \geq 0, \ m = 1, 2, \ldots, M\} \tag{1.7}$$

however, one can easily modify the results for any nonempty bounded polytope set Λ.

Definition 1.1:

A given portfolio $\boldsymbol{\tau}$ is SSD efficient if there is no portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\tau}$.

Following Kopa and Chovanec (2008) one can view SSD portfolio efficiency tests as a multi-objective generalization of mean-risk models, where multiple CVaRs (at different α levels) are employed. Later on, Kopa and Post (2013) introduced a general SSD efficiency test which is a generalization of all previous SSD efficiency tests what makes the theory of SSD portfolio efficiency more compact. Moreover, they presented also a new SSD portfolio efficiency test that is much less computationally demanding and applicable even for thousands of scenarios. Although the test does not identify a dominating portfolio, the computational tractability allows for further sensitivity, robustness or bootstrap techniques.

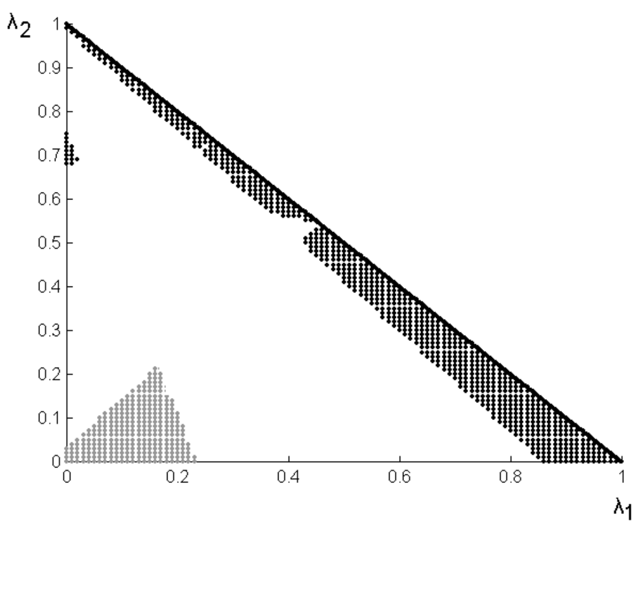
When the first-order stochastic dominance relation is used one must distinguish between two concepts of efficiency: " FSD admissibility" and " FSD optimality".

Definition 1.2:

A given portfolio $\boldsymbol{\tau}$ is FSD admissible if there is no portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\boldsymbol{\lambda} \succ_{FSD} \boldsymbol{\tau}$. A given portfolio $\boldsymbol{\tau}$ is FSD optimal if there exists at least one utility function $u \in U_1$ such that $Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) \geq 0$ for all $\boldsymbol{\lambda} \in \Lambda$ with strict inequality for at least one $\boldsymbol{\lambda} \in \Lambda$.

There is a subtle difference between FSD admissibility and FSD optimality. For example, if the optimal solution of the maximizing expected utility problem is unique, the FSD admissibility is a necessary condition for FSD optimality. The illustration of this situation is captured at Figure 1.1.

Figure 1.1: *FSD admissibility and FSD optimality.*
The figure shows the efficiency classification according to the FSD admissibility test and our FSD optimality test for $M = 3$. We applied these tests to all portfolios $\boldsymbol{\tau} \in \Lambda \cap \{0, 0.01, \ldots, 1\}^3$, that is, when using a grid with step size 0.01 for the portfolio weights. The FSD optimal set is represented by the black dots. The FSD admissible set is the union of the black dots and the grey dots.



Kuosmanen (2004) proposed FSD portfolio admissibility test that identifies also a dominating portfolio (if the tested one is FSD inadmissible). Kopa and Post (2009) enriched it by the FSD optimality test which searches for an "optimal utility function".

For $N \geq 2$ the notion of "NSD admissibility" is equivalent to the "NSD optimality", hence there is no need to consider both definitions. Therefore, the general definition of NSD efficiency for $N \geq 2$ can be seen as an extension of SSD efficiency and, following Post and Kopa (2013), we formulate it in the "NSD optimality" form.

Definition 1.3:

A given portfolio $\boldsymbol{\tau}$ is NSD efficient ($N \geq 2$), if there exists at least one utility function $u \in U_N$ such that $Eu(\boldsymbol{\varrho}'\boldsymbol{\tau}) - Eu(\boldsymbol{\varrho}'\boldsymbol{\lambda}) \geq 0$ for all $\boldsymbol{\lambda} \in \Lambda$ with strict inequality for at least one $\boldsymbol{\lambda} \in \Lambda$.

The following three chapters of the thesis present in a succinct form the algorithms for testing whether a given portfolio is FSD optimal (Kopa and Post (2009)), SSD efficient (Post (2003), Kuosmanen (2004), Kopa and Chovanec (2008), Kopa and Post (2013)) or NSD efficient (Post and Kopa (2013)). In all cases the computational tractability of the algorithms is a desired feature. While NSD portfolio efficiency tests for $N \geq 2$ can be formulated using linear programming problems, FSD admissibility and FSD optimality would call for solving mixed-integer linear programming problems what significantly increase the complexity of the algorithms. In all these portfolio efficiency tests, a scenario approach is assumed, that is, asset returns have a discrete probability distribution. The disadvantage of these tests is the fact that a small perturbation in the probability distribution can completely change the efficiency classification. Therefore some robustness extensions of portfolio efficiency tests are presented in Chapter 5. For example, Kopa (2010, 2012) considered a "neighborhood approach"; Dupačová and Kopa (2012, 2013) applied contamination techniques. These robustness extensions make portfolio efficiency testing more computationally demanding, however they give more valuable results in real-life applications.

Chapter 2

Portfolio efficiency test with respect to the first-order stochastic dominance criterion

In this chapter, we present the new test derived in Kopa and Post (2009). It is the first test that allows for full diversification across the assets and uses a new concept of FSD optimality of a given portfolio relative to all possible portfolios. It enables to identify the efficient portfolios with respect to the class of all nondecreasing utility functions.

2.1 Preliminaries

Consider again M choice alternatives with random returns $\boldsymbol{\varrho}$. In order to derive a tractable FSD optimality test we assume that $\boldsymbol{\varrho}$ is carried by T scenarios with equal probabilities. The outcomes of the choice alternatives for various scenarios are given by

$$X = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^T \end{pmatrix}$$

where $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_M^t)$ is the t-th row of matrix X. Without loss of generality we can assume that the columns of X are linearly independent.

The evaluated portfolio is denoted by $\boldsymbol{\tau} \in \Lambda$ and is assumed to be risky (the return of portfolio $\boldsymbol{\tau}$ is not deterministic). Let $y^{[k]}$ be the k-th smallest element among $y^1, y^2, ..., y^T$, that is, $y^{[1]} \leq y^{[2]} \leq \ldots \leq y^{[T]}$. Let

$$\underline{m} = \min_{t,m} x_m^t, \quad \overline{m} = \max_{t,m} x_m^t \quad \text{and} \quad k(\boldsymbol{\tau}) = \min\{t : (X\boldsymbol{\tau})^{[t]} > (X\boldsymbol{\tau})^{[1]}\}.$$

The constants $\underline{m}$ and $\overline{m}$ denote the minimum and maximum possible returns. After ordering the returns of the tested portfolio $\boldsymbol{\tau}$ from the smallest to the largest one, $k(\boldsymbol{\tau})$ determines the order of the second smallest return. Without ties, we have $k(\boldsymbol{\tau}) = 2$, but if the smallest value occurs multiple times, then $k(\boldsymbol{\tau}) > 2$.

Since utility functions are unique up to the level of a positive linear transformation, without loss of generality, we may focus on the following set of standardized utility functions:

$$U_1(\boldsymbol{\tau}) \;=\; \{u \in U_1 : u(\underline{m}) = 0; \;\; u((X\boldsymbol{\tau})^{[T]}) - u((X\boldsymbol{\tau})^{[k(\boldsymbol{\tau})]}) = 1\}. \tag{2.1}$$

Note that the standardization depends on the evaluated portfolio and hence it will differ when evaluating different portfolios. Furthermore, the standardization requires utility to be strictly increasing at least somewhere in the interior of the range for the evaluated portfolio. This requirement is natural, because, testing optimality relative to all $u \in U_1$ is trivial. Specifically, every portfolio $\boldsymbol{\lambda} \in \Lambda$ is an optimal solution for $u_0 = I(x \geq (X\boldsymbol{\tau})^{[1]})$, that is, two-piece constant utility function. Thus $U_1(\boldsymbol{\tau})$ is the largest subset of U_1 for which testing optimality is non-trivial and Definition 1.2 may be reformulated as follows:

Definition 2.1:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if there exists $u \in U_1(\boldsymbol{\tau})$ such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\lambda}) \geq 0 \quad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD non-optimal.

2.2 Representative utility functions

For pairwise FSD comparisons, Russell and Seo (1989) show that the set of three-piece linear utility functions is representative for all admissible utility functions, see (1.3). In our portfolio context, with diversification allowed, the following class of piecewise constant utility functions matters:

$$R_1(\boldsymbol{\tau}) = \{u \in U_1 | u(y) = \sum_{t=1}^{T} a_t I(y \geq (X\boldsymbol{\tau})^{[t]}), \quad \mathbf{a} \in A(\boldsymbol{\tau})\} \tag{2.2}$$

$$\begin{aligned} A(\boldsymbol{\tau}) = \{\mathbf{a} \in \mathbb{R}_+^T : \sum_{t=k(\boldsymbol{\tau})}^{T} a_t = 1, \quad (X\boldsymbol{\tau})^{[t]} = (X\boldsymbol{\tau})^{[s]} \\ \wedge \ t < s \Rightarrow a_s = 0, \quad t, s = 1, 2, \ldots, T\} \end{aligned} \tag{2.3}$$

where

$$\begin{aligned} I(y \geq y_0) &= 1 \text{ for } y \geq y_0 \\ &= 0 \text{ otherwise.} \end{aligned}$$

This class consists of at most $(T+1)$ - piecewise constant, upper semi-continuous utility functions. It extends the representative utility functions used by Russell and Seo (1989) to testing pairwise FSD relationship. In fact, our utility functions can be obtained as a sum of the first derivatives of the Russell and Seo (1989) representative utility functions on the relevant interval $(\underline{m}, \overline{m})$.[1] The utility functions are also reminiscent of the piecewise linear functions used by Post (2003) to test SSD portfolio efficiency.

Theorem 2.1:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if and only if there exists $u \in R_1(\boldsymbol{\tau})$ such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\lambda}) \geq 0 \quad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD non-optimal.

[1]Russell and Seo (1989) functions are continuous three-piece functions that consist of two constant pieces and one linear, increasing piece in between. Choose T such functions with increasing pieces with slopes $a_1, a_2, ..., a_T$ for the intervals $((X\tau)^{[1]}, (X\tau)^{[2]})$, $((X\tau)^{[2]}, (X\tau)^{[3]})$,...,$((X\tau)^{[T-1]}, (X\tau)^{[T]})$, $((X\tau)^{[T]}, \overline{m})$. Our piecewise constant utility function is the sum of the first derivatives on these intervals.

Apart from replacing $U_1(\boldsymbol{\tau})$ with $R_1(\boldsymbol{\tau})$, we may also replace Λ in Theorem 2.1 with a reduced portfolio set that considers only portfolios with a higher minimum than that of the evaluated portfolio:

$$\Lambda(\boldsymbol{\tau}) = \left\{\boldsymbol{\lambda} \in \Lambda : (X\boldsymbol{\tau})^{[1]} \leq (X\boldsymbol{\lambda})^{[1]}\right\}.$$

Using the representative utility functions and the reduced portfolio set, we can construct the following FSD non-optimality measure for any $\Lambda_0 \subseteq \Lambda(\boldsymbol{\tau})$:

$$\xi(\boldsymbol{\tau}, \Lambda_0) = \frac{1}{T} \min_{u \in R_1(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda_0} \sum_{t=1}^{T} \left(u(\mathbf{x}^t\boldsymbol{\lambda}) - u(\mathbf{x}^t\boldsymbol{\tau})\right). \tag{2.4}$$

Replacing Λ with $\Lambda(\boldsymbol{\tau})$ reduces the parameter space but it causes no harm, because

$$\max_{\boldsymbol{\lambda} \in \Lambda} \sum_{t=1}^{T} \left(u(\mathbf{x}^t\boldsymbol{\lambda}) - u(\mathbf{x}^t\boldsymbol{\tau})\right) = \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{t=1}^{T} \left(u(\mathbf{x}^t\boldsymbol{\lambda}) - u(\mathbf{x}^t\boldsymbol{\tau})\right)$$

for all $u \in R_1(\boldsymbol{\tau})$ with sufficiently large a_1 and we minimize the maximum of expected utility differences. If the evaluated portfolio has the highest minimum then we can directly conclude that $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) = 0$, that is, the evaluated portfolio is FSD optimal (see the following Corollary).

Corollary 2.1:

(i) Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) = 0$. Otherwise, $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) > 0$.

(ii) If $\Lambda_0 \subseteq \Lambda(\boldsymbol{\tau})$ then $\xi(\boldsymbol{\tau}, \Lambda_0) \leq \xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))$.

2.3 Mathematical Programming Formulations

Let

$$\begin{aligned} h_s(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= \sum_{t=1}^{T} I(\mathbf{x}^t\boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]}), \qquad s = 1, \ldots, T & (2.5)\\ \mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= (h_1(\boldsymbol{\lambda}, \boldsymbol{\tau}), \ldots, h_T(\boldsymbol{\lambda}, \boldsymbol{\tau})) & (2.6)\\ H(\boldsymbol{\tau}) &= \{\mathbf{h} \in \{0, \ldots, T\}^T : \mathbf{h} = \mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}), \;\; \boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})\}. & (2.7) \end{aligned}$$

Since $h_s(\boldsymbol{\lambda}, \boldsymbol{\tau})$ represents the number of returns of portfolio $\boldsymbol{\lambda}$ exceeding the s-th smallest return of portfolio $\boldsymbol{\tau}$, it can take at most $T+1$ values $(0, 1, \dots, T)$ for any $s = 1, \dots, T$. Thus $H(\boldsymbol{\tau})$ is a finite set. For small-scale applications, identifying all its elements is quite trivial task. However, for large-scale applications, the task is more challenging and can become computationally demanding. Some computational strategies to identifying the elements of $H(\boldsymbol{\tau})$ are discussed below. Perhaps interestingly, given $H(\boldsymbol{\tau})$, the FSD non-optimality measure $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))$ can be computed using simple linear programming.

Theorem 2.2:

Let $H_0 \subseteq H(\boldsymbol{\tau})$. Let

$$\delta^*(H_0) = \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \quad \delta \tag{2.8}$$

$$\text{s.t.} \qquad \sum_{s=k(\boldsymbol{\tau})}^{T} a_s(\overline{h}_s - h_s(\boldsymbol{\tau}, \boldsymbol{\tau})) \;\; \leq \;\; \delta \qquad \forall \overline{\mathbf{h}} \in H_0. \tag{2.9}$$

Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\delta^*(H(\boldsymbol{\tau})) = 0$. If $\delta^*(H_0) > 0$ for some $H_0 \subseteq H(\boldsymbol{\tau})$ then $\boldsymbol{\tau}$ is FSD non-optimal.

The idea of this result is to find a representative utility function for which $\boldsymbol{\tau}$ maximizes expected utility. Note that $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) = \delta^*/T$. Since $a \in A(\boldsymbol{\tau})$ and $\mathbf{h} \in \{0, \dots, T\}^T$ for all $\mathbf{h} \in H(\boldsymbol{\tau})$, using Corollary 2.1(i), we have $0 \leq \xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) \leq 1$.

Among other things, Theorem 2.2 implies the following result about the relationship between the efficiency concepts of optimality and admissibility.

Corollary 2.2:

If $(T \leq 4)$ then FSD optimality is equivalent to FSD admissibility.

The remaining problem is to identify the elements of the set $H(\boldsymbol{\tau})$. We may adopt several strategies for this task. The next section provides a mixed integer linear programming (MILP) algorithm that identifies a set of candidate vectors $\widetilde{H}(\boldsymbol{\tau}) \supseteq H(\boldsymbol{\tau})$, and checks if $\mathbf{h} \in H(\boldsymbol{\tau})$ for every candidate $\mathbf{h} \in \widetilde{H}(\boldsymbol{\tau})$. A drawback of this approach is that the number of candidates increases exponentially with the number of scenarios (T). Hence, some sort of approximation may be needed.

For example, we may form a sample $H_s(\boldsymbol{\tau})$ of elements $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ by using a sample $\Lambda_s \in \Lambda(\boldsymbol{\tau})$ and constructing the associated values for $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$. The test procedure is then applied to the sample $H_s(\boldsymbol{\tau})$ instead of the complete set $H(\boldsymbol{\tau})$. According to Corollary 2.1(ii), this will lead to a necessary condition for FSD optimality. There exist various sampling techniques, including a regular grid, Monte Carlo methods or Quasi-Monte Carlo methods; see, for example, Jackel (2002) and Glasserman (2004).

2.4 Mathematical programming Algorithm

This section provides a mixed-integer linear programming algorithm (MILP) for identifying the elements of $H(\boldsymbol{\tau})$, applies Theorem 2.2 and suggests some additional stopping rules for testing FSD optimality.

STEP 1: *Perform a FSD admissibility test*

Use, for example, the MILP FSD admissibility test of Kuosmanen (2004). If $\boldsymbol{\tau}$ is FSD inadmissible then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal.

STEP 2: *Identify initial candidates for* $H(\boldsymbol{\tau})$

For all $j = k(\boldsymbol{\tau}), ..., T$ solve the following MILP problem:

$$\begin{array}{rrcll}
 & \max \quad h_j & + & \frac{1}{T^2}\sum_{t=k(\boldsymbol{\tau})}^{T} h_t & \\
\text{s.t.} & (v_{s,t}-1)(\overline{m}-\underline{m}) & \leq & \mathbf{x}^s\boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} \leq v_{s,t}(\overline{m}-\underline{m}) & \\
 & & & & s = 1,\ldots,T;\ t = k(\boldsymbol{\tau}),\ldots,T \\
 & h_t & = & \sum_{s=1}^{T} v_{s,t} & t = k(\boldsymbol{\tau}),\ldots,T \\
 & v_{s,t} & \in & \{0,1\} & s = 1,\ldots,T;\ t = k(\boldsymbol{\tau}),\ldots,T \\
 & \boldsymbol{\lambda} & \in & \Lambda(\boldsymbol{\tau}) &
\end{array} \quad (2.11)$$

The problem is solved only for $j \geq k(\boldsymbol{\tau})$; solving it for $j < k(\boldsymbol{\tau})$ will identify no new candidates, because the optimal solutions of (2.11) for any $j < k(\boldsymbol{\tau})$ are equal to that for $j = k(\boldsymbol{\tau})$.

Use $(h_t^{*j}, \lambda_t^{*j}, v_{s,t}^{*j})$ for the optimal solution of this problem. Let $\Lambda_1 \in \Lambda(\boldsymbol{\tau})$ be a set of pairwise different $\boldsymbol{\lambda}^{*j}$ (all redundancies are removed). Set

$$\begin{aligned} h_t^{max} &= \max_j h_t^{*j} \\ H_1 &= \{\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}) : \quad \boldsymbol{\lambda} \in \Lambda_1\} \end{aligned}$$

STEP 3: *Stopping rules*

Consider $\mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau})$ as defined by (2.5)-(2.6). If there exists $t \in \{k(\boldsymbol{\tau}), \ldots, T\}$ such that $h_t^{max} \leq h_t(\boldsymbol{\tau}, \boldsymbol{\tau})$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD optimal. Otherwise, solve problem (2.8)-(2.9) for $H_0 = H_1$. If $\delta^*(H_1) > 0$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal.

STEP 4: *Construct and reduce the candidate set* $\overline{H}$

Let $\overline{H}_t = \{0, 1, \ldots, h_t^{max}\}$. Denote by $\overline{H}$ the Cartesian product $\overline{H} = \bigotimes_{k(\boldsymbol{\tau})}^T \overline{H}_t$. Clearly $H(\boldsymbol{\tau}) \subseteq \overline{H}$, and hence $\overline{H}$ is a candidate set. Exclude the candidates $\widetilde{H} = \widetilde{H}_1 \cup \widetilde{H}_2 \cup \widetilde{H}_3 \cup \widetilde{H}_4$, where

$$\begin{aligned} \widetilde{H}_1 &= \left\{\mathbf{h} \in \overline{H} | h_{t_1} < h_{t_2} \ \text{ for some } \ t_1 < t_2\right\} \\ \widetilde{H}_2 &= \left\{\mathbf{h} \in \overline{H} | h_t \geq h_t(\boldsymbol{\tau}, \boldsymbol{\tau}) \ \ \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\}\right\} \\ \widetilde{H}_3 &= \Big\{\mathbf{h} \in \overline{H} | \exists \widehat{\mathbf{h}} \in H_1 : h_t \geq \widehat{h}_t \ \ \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\} \ \text{ with at least} \\ &\qquad \text{one strict inequality}\} \\ \widetilde{H}_4 &= \left\{\mathbf{h} \in \overline{H} | \exists \mathbf{b} = (b_0, b_{k(\boldsymbol{\tau})}, \ldots, b_T) : h_t \leq \sum_{j=k(\boldsymbol{\tau})}^T b_j h_t^{*j} + b_0 h_t(\boldsymbol{\tau}, \boldsymbol{\tau}), \right. \\ &\qquad \left. b_0 + \sum_{l=k(\boldsymbol{\tau})}^T b_l = 1, \quad \mathbf{b} \geq 0, \quad \mathbf{h}^{*j} \in H_1, \ \ \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\} \right\}. \end{aligned}$$

The elements of $\widetilde{H}_1 \cup \widetilde{H}_2 \cup \widetilde{H}_3$ are not feasible, that is, there exist no corresponding portfolios: the elements of $\widetilde{H}_1$ contradict the definition of vector $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$, see (2.5)-(2.6); feasibility of an element of $\widetilde{H}_2$ implies FSD inadmissibility of $\boldsymbol{\tau}$ (in step 1, we have found that $\boldsymbol{\tau}$ is FSD admissible); every element of $\widetilde{H}_3$ gives a strictly higher value of the objective function in (2.11) than at least one initial candidate, hence it cannot be a feasible candidate. Adding the elements of $\widetilde{H}_4$ to H_1 does not affect the solution of (2.8)-(2.9).

Set $p = 1$.

STEP 5: *Check feasibility of the remaining candidates*

If $\overline{H} \setminus \widetilde{H}$ is empty, that is, all possible $\mathbf{h} \in \overline{H}$ have been considered, then stop the algorithm; portfolio $\boldsymbol{\tau}$ is FSD optimal. Otherwise, choose $\mathbf{h} \in \overline{H} \setminus \widetilde{H}$ and add it to $\widetilde{H}$. Let $p = p + 1$, $H_p = H_{p-1} \cup \mathbf{h}$ and go to the next step if there exists a feasible solution of the system:

$$\begin{aligned}
(v_{s,t} - 1)(\overline{m} - \underline{m}) \leq \mathbf{x}^s\boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} &\leq v_{s,t}(\overline{m} - \underline{m}) && s = 1, \ldots, T; \\
& && t = t_1, \ldots, T \\
h_t &= \textstyle\sum_{s=1}^{T} v_{s,t} && t = t_1, \ldots, T \\
v_{s,t} &\in \{0, 1\} && s = 1, \ldots, T; \\
& && t = t_1, \ldots, T \\
\boldsymbol{\lambda} &\in \Lambda(\boldsymbol{\tau}).
\end{aligned} \tag{2.12}$$

Otherwise, repeat Step 5.

STEP 6: *Test optimality using the feasible candidates*

Solve problem (2.8)-(2.9) for $H_0 = H_p$. If $\delta^*(H_p) > 0$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal. Otherwise, for $\delta^*(H_p) = 0$, go to Step 5.

Chapter 3

Portfolio efficiency tests with respect to the second-order stochastic dominance criterion

This chapter recalls LP tests for testing if a given portfolio is SSD efficient as derived in Post (2003), Kuosmanen (2004) and Kopa and Chovanec (2008). Moreover, following Kopa and Post (2013) it presents the more general test based on primal-dual formulations with its interpretation and properties. Finally, it enriches the theory of SSD portfolio efficiency testing by a reduced version of the general test introduced in Kopa and Post (2013). In all these tests the discrete distribution of returns with equiprobable scenarios is assumed and the notation from Chapter 2 is preserved.

3.1 The Post test

The Post test uses a small, computationally efficient linear program, derived from the first-order necessary condition for portfolio optimization for an increasing and concave utility function. From the managerial perspective, a limitation of the test is that it focuses exclusively on the efficiency classification of the evaluated portfolio and gives only a minimal information about directions for improvement if the portfolio is classified as inefficient. In addition, whereas the test gives a general necessary condition for portfolio efficiency, it gives a sufficient condition for data generated by a continuous population distribution.

Before using the test, it is necessary to reorder the scenarios according to the returns of the evaluated portfolio $\boldsymbol{\tau}$, such that $\mathbf{x}^1\boldsymbol{\tau} \leq \mathbf{x}^2\boldsymbol{\tau} \leq ... \leq \mathbf{x}^T\boldsymbol{\tau}$.

Theorem 3.1:

Let

$$\theta^* = \min_{\theta,\beta_t} \theta \tag{3.1}$$

$$\begin{array}{llcll} s.t. & \displaystyle\sum_{t=1}^{T} \beta_t(\mathbf{x}^t\boldsymbol{\tau} - x_m^t) + T\theta & \geq & 0, & m = 1, 2, ..., M \\ & \beta_t - \beta_{t+1} & \geq & 0, & t = 1, 2, ..., T-1 \\ & \beta_t & \geq & 0, & t = 1, 2, ..., T-1 \\ & \beta_T & = & 1. & \end{array}$$

The portfolio $\boldsymbol{\tau}$ *is strictly SSD efficient if and only if* $\theta^* = 0$*. Moreover, if the portfolio* $\boldsymbol{\tau}$ *is SSD efficient then* $\theta^* = 0$*.*

The constraints need to be modified if some ties in elements of $X\boldsymbol{\tau}$ occur. See Post (2003) for more details and for the dual formulation.

3.2 The Kuosmanen test

Applying Hardy, Littlewood and Polya's (1934) majorization theorem, Kuosmanen (2004) derived an alternative test that also identifies a SSD dominating portfolio (if the tested portfolio is classified as SSD inefficient). This test involves solving two linear programs. These problems generally are very large and more computationally demanding than the Post test.

The exact formulation depends on ties in $X\boldsymbol{\tau}$. We say that a k-way tie occurs if k elements of $X\boldsymbol{\tau}$ are equal.

Theorem 3.2:

Let

$$\theta^* = \max_{W,\boldsymbol{\lambda}} \sum_{t=1}^{T} \left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) \tag{3.2}$$

$$
\begin{aligned}
s.t. \qquad X\boldsymbol{\lambda} &\geq WX\boldsymbol{\tau} \\
\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} &= 1, \quad w_{ij} \geq 0 \qquad i,j = 1,2,...,T \\
\boldsymbol{\lambda} &\in \Lambda
\end{aligned}
$$

and

$$
\theta^{**} = \min_{W,\boldsymbol{\lambda},S^{+},S^{-}} \sum_{j=1}^{T}\sum_{i=1}^{T}(s_{ij}^{+} + s_{ij}^{-}) \tag{3.3}
$$

$$
\begin{aligned}
s.t. \qquad X\boldsymbol{\lambda} &= WX\boldsymbol{\tau} \\
s_{ij}^{+} - s_{ij}^{-} &= w_{ij} - \frac{1}{2} \qquad i,j = 1,2,...,T \\
s_{ij}^{+}, s_{ij}^{-}, w_{ij} &\geq 0 \qquad i,j = 1,2,...,T \\
\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} &= 1 \qquad i,j = 1,2,...,T \\
\boldsymbol{\lambda} &\in \Lambda
\end{aligned}
$$

where $S^{+} = \{s_{ij}^{+}\}_{i,j=1}^{T}$, $S^{-} = \{s_{ij}^{-}\}_{i,j=1}^{T}$ *and* $W = \{w_{ij}\}_{i,j=1}^{T}$*. Let* ϵ_k *denote the number of* k*-way ties in* $X\boldsymbol{\tau}$*. Then portfolio* $\boldsymbol{\tau}$ *is SSD efficient if and only if*

$$
\theta^{*} = 0 \quad \wedge \; \theta^{**} = \frac{T^2}{2} - \sum_{k=1}^{T} k\epsilon_k.
$$

Let $\boldsymbol{\lambda}^{*}$ and $\boldsymbol{\lambda}^{**}$ be the optimal solution of (3.2) and (3.3), respectively. If $\theta^{*} > 0$ then $\boldsymbol{\varrho}'\boldsymbol{\lambda}^{*} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$. If $\theta^{*} = 0$ and $\theta^{**} < \frac{T^2}{2} - \sum_{k=1}^{T} k\epsilon_k$ then $\boldsymbol{\varrho}'\boldsymbol{\lambda}^{**} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$.

If $\theta^{*} > 0$ then problem (3.3) need not to be solved, because portfolio $\boldsymbol{\tau}$ is SSD inefficient and the optimal solution $\boldsymbol{\lambda}^{*}$ is an SSD dominating portfolio, see Kuosmanen (2004) for more details.

If a given portfolio $\boldsymbol{\tau}$ is SSD inefficient then, from the entire set of SSD dominating portfolios, the Kuosmanen test identifies that with the highest mean return.

3.3 The Kopa and Chovanec test

We start with reformulation of SSD criterion in terms of CVaR for scenario approach because the Kopa and Chovanec test, contrary to the previous tests, is based on second quantile functions (1.6).

Lemma 3.1:

A portfolio $\boldsymbol{\lambda} \in \Lambda$ *dominates portfolio* $\boldsymbol{\tau} \in \Lambda$ *if and only if*

$$\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}\boldsymbol{\lambda}) \leq \mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}\boldsymbol{\tau}) \qquad \forall \alpha \in \{0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}\}.$$

Applying

$$\begin{array}{rcl} \mathrm{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{\varrho}\boldsymbol{\lambda}) & = & \displaystyle\min_{b_k, w_k^t} b_k + \frac{1}{T(1-\frac{k-1}{T})}\sum_{t=1}^{T} w_k^t \\ s.t. & & w_k^t \geq 0, w_k^t + b_k \geq -\mathbf{x}^t\boldsymbol{\lambda}, \quad \forall t, k = 1, 2, ..., T \end{array}$$

in Lemma 3.1 for each $k = 1, 2, ..., T$, Kopa and Chovanec (2008) derived the following SSD portfolio efficiency test.

Theorem 3.3:

Let

$$D^*(\boldsymbol{\tau}) = \max_{D_k, \lambda_n, b_k, w_k^t} \sum_{k=1}^{T} D_k \tag{3.4}$$

$$\begin{array}{rrcll} \text{s.t.} \quad \mathrm{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) - b_k - \displaystyle\frac{1}{(1-\frac{k-1}{T})T}\sum_{t=1}^{T} & w_k^t & \geq & D_k, & k = 1, \ldots, T \\ & w_k^t + b_k & \geq & -\mathbf{x}^t\boldsymbol{\lambda}, & t, k = 1, \ldots, T \\ & w_k^t & \geq & 0, & t, k = 1, \ldots, T \\ & D_k & \geq & 0, & k = 1, \ldots, T \\ & \boldsymbol{\lambda} & \in & \Lambda. & \end{array}$$

If $D^*(\boldsymbol{\tau}) > 0$ *then* $\boldsymbol{\tau}$ *is SSD inefficient and* $\boldsymbol{\varrho}'\boldsymbol{\lambda}^* \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$*. Otherwise,* $D^*(\boldsymbol{\tau}) = 0$ *and* $\boldsymbol{\tau}$ *is SSD efficient.*

If a tested portfolio $\boldsymbol{\tau}$ is SSD inefficient the test identifies SSD dominating portfolio $\boldsymbol{\lambda}^*$ that is SSD efficient. In general, the dominating portfolio is very different from that of the Kuosmanen test. While the Kuosmanen test identifies a dominating portfolio with the highest mean, the Kopa and Chovanec test focuses more on possible risk reductions. In this case, the risk is expressed by the sum of CVaRs, contrary to a single CVaR application in e.g. mean-risk models.

3.4 A general test for SSD portfolio efficiency

In this section, following Kopa and Post (2013), we present a generalization of all three tests presented in Sections 3.1-3.3. Contrary to the previous tests, the portfolio possibilities are represented by a polytope of general form:

$$\Lambda \;=\; \{\boldsymbol{\lambda} \in R^M | A\boldsymbol{\lambda} \leq \mathbf{b}\} \tag{3.5}$$

with $(L \times M)$ matrix A of coefficients for L linear inequality restrictions placed on M assets, and $\mathbf{b}$ for a vector of right-hand side coefficients. To guarantee the existence of optimal solutions for our test, we assume that Λ is nonempty and bounded polytope.

We propose the following measure for testing SSD relation:

$$\begin{aligned}
\omega(\boldsymbol{\lambda}, \boldsymbol{\tau}|\mathbf{w}) \;&=\; \min_{\beta_t} T^{-1} \sum_{t=1}^{T} \beta_t \left((X\boldsymbol{\lambda})^{[t]} - \mathbf{x}^t \boldsymbol{\tau} \right) \\
\text{s.t.} \quad & \beta_t - \beta_{t+1} \geq w_t, \quad t = 1, 2, ..., T-1 \\
& \beta_T \geq w_T.
\end{aligned} \tag{3.6}$$

The variables $\beta_t,\ t = 1, 2, ..., T$ represent subgradients of feasible utility functions. The weights $\mathbf{w} = (w_1, w_2, ..., w_T)$ are lower bounds for the decrements of the subgradients. To be consistent with strict risk aversion, the weights must be strictly positive, i.e. $\mathbf{w} > \mathbf{0}$. Apart from these requirements, the exact values of the weights have no effect on the sign of the dominance measure (3.6) and the dominance classification. Still, the weights do affect the measured degree of dominance and can be used to control the desired degree measure. Section 3.5 will elaborate further on the specification of weights.

Lemma 3.2:

Portfolio $\boldsymbol{\lambda}$ SSD dominates portfolio $\boldsymbol{\tau}$ if and only if $\omega(\boldsymbol{\lambda}, \boldsymbol{\tau}|\mathbf{w}) > 0$ where $\mathbf{w} > \mathbf{0}$.

To test SSD portfolio efficiency, Kopa and Post (2013) propose to solve the following problem:

$$\xi(\boldsymbol{\tau}|\mathbf{w}) \quad = \quad \min_{\gamma_{t,s},\theta_i} \sum_{i=1}^{L} b_i\theta_i - T^{-1}\sum_{t=1}^{T}\sum_{s=t}^{T}\frac{1}{s}\sum_{k=1}^{T}\gamma_{k,s}\mathbf{x}^t\boldsymbol{\tau} \tag{3.7}$$

$$\begin{aligned}
\text{s.t.} \quad \textstyle\sum_{i=1}^{L} a_{ij}\theta_i - T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\gamma_{t,s}x_j^t &= 0, & j &= 1,2,...,M\\
\tfrac{1}{s}\textstyle\sum_{t=1}^{T}\gamma_{t,s} &\geq w_s, & s &= 1,2,...,T\\
\gamma_{t,s} - \tfrac{1}{s}\textstyle\sum_{k=1}^{T}\gamma_{k,s} &\leq 0, & t,s &= 1,2,...,T\\
\gamma_{t,s} &\geq 0, & t,s &= 1,2,...,T\\
\theta_i &\geq 0, & i &= 1,2,...,L
\end{aligned}$$

Theorem 3.4:

A given portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if $\xi(\boldsymbol{\tau}|\mathbf{w}) = 0$ where $\mathbf{w} > \mathbf{0}$.

The objective value $\xi(\boldsymbol{\tau}|\mathbf{w})$ can be seen as a measure of inefficiency of portfolio $\boldsymbol{\tau}$ when using weights $\mathbf{w}$. Moreover, if portfolio $\boldsymbol{\tau}$ is SSD efficient then the optimal values $\gamma_{t,s}^*$ determine the derivatives of the utility function that maximizes expected utility in $\boldsymbol{\tau}$. Specifically,

$$\beta_t^* = \sum_{s=t}^{T}\frac{1}{s}\sum_{k=1}^{T}\gamma_{k,s}^*.$$

The variables θ_i give information about the sensitivity of $\xi(\boldsymbol{\tau}|\mathbf{w})$ to changes in b_i of portfolio set (3.5).

Consider now the exact LP dual formulation of program (3.7):

$$\xi(\boldsymbol{\tau}|\mathbf{w}) = \max_{d_s,v_{t,s},\boldsymbol{\lambda}} \sum_{s=1}^{T} w_s d_s \tag{3.8}$$

$$
\begin{aligned}
\text{s.t.}\quad -T^{-1}\mathbf{x}^t\boldsymbol{\lambda}+\tfrac{1}{s}d_s-v_{t,s}+\tfrac{1}{s}\textstyle\sum_{k=1}^{T}v_{k,s} &\le -T^{-1}\tfrac{1}{s}\textstyle\sum_{k=1}^{s}\mathbf{x}^k\boldsymbol{\tau}\\
& \quad t,s=1,2,...,T\\
\textstyle\sum_{j=1}^{M}a_{ij}\lambda_j &\le b_i,\quad i=1,2,...,L\\
v_{t,s} &\ge 0,\quad t,s=1,2,...,T\\
d_s &\ge 0,\quad s=1,2,...,T.
\end{aligned}
$$

Whereas the primal model (3.7) has a compelling interpretation in terms of utility functions, the dual formulation (3.8) is probably best understood in terms of cumulative returns:

$$
\Omega(\boldsymbol{\lambda},s)=T^{-1}\sum_{k=1}^{s}(X\boldsymbol{\lambda})^{[k]}. \tag{3.9}
$$

For the evaluated portfolio $\boldsymbol{\tau}$, we have:

$$
\Omega(\boldsymbol{\lambda},s)=T^{-1}\sum_{k=1}^{s}\mathbf{x}^k\boldsymbol{\tau}. \tag{3.10}
$$

Theorem 3.5:

The inefficiency measure can be expressed in terms of cumulative returns as follows:

$$
\begin{aligned}
\xi(\boldsymbol{\tau}|\mathbf{w}) = \max_{d_s,\boldsymbol{\lambda}} \sum_{s=1}^{T} w_s d_s & \qquad\qquad (3.11)\\
\text{s.t.}\quad \Omega(\boldsymbol{\lambda},s)-\Omega(\boldsymbol{\tau},s) &\ge d_s,\quad s=1,2,...,T\\
d_s &\ge 0,\quad s=1,2,...,T\\
\boldsymbol{\lambda} &\in \Lambda.
\end{aligned}
$$

The dual thus maximizes the weighted sum of differences in cumulative returns between the evaluated portfolio $\boldsymbol{\tau}$ and a second portfolio $\boldsymbol{\lambda}$ that SSD dominates $\boldsymbol{\tau}$.

Theorem 3.6:

A given portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if $\xi(\boldsymbol{\tau}|\mathbf{w})=0$ where $\mathbf{w}>\mathbf{0}$. If $\xi(\boldsymbol{\tau}|\mathbf{w})>0$ then solution portfolio $\boldsymbol{\lambda}^*$ of (3.8) is SSD efficient and it dominates $\boldsymbol{\tau}$ by SSD.

3.5 Specification of the weights in the general test

The primary function of the weights is to avoid negative values for marginal utility, or risk seeking segments of the utility function. To get a proper necessary and sufficient condition, the weights should be strictly positive, i.e. $\mathbf{w} > \mathbf{0}$. Allowing for zero weights, or $\mathbf{w} \geq \mathbf{0}$ amounts to assuming weak risk aversion rather than strict risk aversion. This difference is unlikely to have a noticeable effect in empirical applications with data generated by a (approximately) continuous probability distribution. Still, from a theoretical perspective, it is important to exclude zero weights to eliminate weakly concave utility functions.

The relative values of the weights do not affect the efficiency classification, but they do affect the optimal objective value as a measure of the degree of inefficiency. In the primal model, weight w_s gives a lower bound for the marginal utility decrement $(u'(z_s) - u'(z_{s+1}))$. In the dual model, the weight determines the relevance of the s-th cumulated return $\Omega(\cdot, s)$. Clearly, a higher value for the low-rank weights has the effect of focusing more on risk reduction, while a higher value for the high-rank weights has the effect of focusing more on improvements in central tendency.

To get an economically meaningful degree measure, it seems useful to remain consistent with the maintained economic assumptions. It seems natural to measure deviations from efficiency by using

$$w_s^u = \left\{ \begin{array}{rl} (u'(\mathbf{x}^s\boldsymbol{\tau}) - u'(\mathbf{x}^{s+1}\boldsymbol{\tau}))/u'(\mathbf{x}^1\boldsymbol{\tau}), & s = 1, 2, ..., T-1 \\ u'(\mathbf{x}^s\boldsymbol{\tau})/u'(\mathbf{x}^1\boldsymbol{\tau}), & s = T. \end{array} \right\} \quad (3.12)$$

for an explicit utility function $u \in U_2$. One possible choice is the one-parameter exponential function, i.e.

$$\begin{aligned} u(x) &= \frac{1 - \exp(-\alpha(100 + x))}{\alpha} \\ u'(x) &= \exp(-\alpha(100 + x)). \end{aligned} \quad (3.13)$$

In these expressions, $\alpha > 0$ is an index of risk aversion and $(100 + x)$ is gross return measured in percentages. If we assume near risk neutrality, or $\alpha \to 0$, the weights converge to $w_1^v = w_2^v = ... = w_{T-1}^v = 0$ and $w_T^v = 1$. In this case, the SSD inefficiency measure gives an information about the largest possible increase in average return achieved by a portfolio that obeys the model restrictions: $\Omega(\boldsymbol{\lambda}, T) = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{x}^t\boldsymbol{\lambda})$. This (implicit) weighting scheme was used in Post (2003) and Kuosmanen (2004). Both tests focus on improvements in central tendency by assigning a very large weight to the highest-rank scenario and a very small weight to the other scenarios. Increasing the level of risk aversion increases the weight for lower-ranked scenarios shifts the focus to risk reduction. In the case of extreme risk aversion, $\alpha \to \infty$, the weights converge to $w_1^v = 1$ and $w_2^v = ... = w_T^v = 0$. In this case, the focus is on the potential increase in minimum return: $T\Omega(\boldsymbol{\lambda}, 1) = (X\boldsymbol{\lambda})^{[1]}$.

Using utility-based weights w_s^u implies that the weights generally depend on the return levels $\mathbf{x}^s\boldsymbol{\tau}$ and we cannot fix the weighting scheme without reference to the data set. By contrast, Kopa and Chovanec (2008) used the fixed set of weights

$$w_s = \left(s\sum_{t=1}^{T}\frac{1}{t}\right)^{-1}, s = 1, 2, ..., T.$$

3.6 Reduced SSD portfolio efficiency test

Following Kopa and Post (2013), we will derive a small version of the general test if an explicit dominating portfolio is not required. In this reduced form, ties in returns of portfolio $\boldsymbol{\tau}$ play an important role, and therefore, we will consider a partition of scenarios $\{1, 2, ..., T\} = \bigcup_{k=1}^{K}\Phi_k$ where Φ_k contains all scenarios with the k-th smallest return of portfolio $\boldsymbol{\tau}$. Put differently, $X\boldsymbol{\tau}$ takes K different values, Φ_1 contains the smallest ones, Φ_K the largest ones, and all $1 < t_1$, $t_1 + 1$,...,$t_2 < T$ such that $\mathbf{x}^{t_1-1}\boldsymbol{\tau} < \mathbf{x}^{t_1}\boldsymbol{\tau} = \mathbf{x}^{t_1+1}\boldsymbol{\tau} = ... = \mathbf{x}^{t_2}\boldsymbol{\tau} < \mathbf{x}^{t_2+1}\boldsymbol{\tau}$ are collected in one Φ_k, $k = 2$,...,$K-1$. Clearly, without ties, we have $K = T$ and $\Phi_k = \{k\}$. Moreover, for each Φ_k, we consider a second partition $\Phi_k = \bigcup_{l_k=1}^{L_k}\Phi_k^{l_k}$ based on ties in the entire return vector (a common situation when using sub-scenarios and pseudo-samples), that is, $t_1, t_2 \in \Phi_k^{l_k}$ if and only if $\mathbf{x}^{t_1} = \mathbf{x}^{t_2}$. Consider:

$$v(\boldsymbol{\tau}|\mathbf{w}) \quad = \quad \min_{\theta_i, \beta_k^l}\sum_{i=1}^{L} b_i\theta_i - T^{-1}\sum_{k=1}^{K}\sum_{l_k=1}^{L_k}\beta_k^{l_k}\sum_{t\in\Phi_k^{l_k}}\mathbf{x}^t\boldsymbol{\tau} \qquad (3.14)$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{i=1}^{L} a_{ij}\theta_i - T^{-1}\sum_{k=1}^{K}\sum_{l_k=1}^{L_k}\beta_k^{l_k}\sum_{t\in\Phi_k^{l_k}} x_j^t = 0, \quad j = 1, ..., M \\
& \beta_k^{l_k} - \beta_{k+1}^{l_{k+1}} \geq w_k, \quad k = 1, ..., K-1, l_k = 1, ..., L_k, l_{k+1} = 1, ..., L_{k+1} \\
& \beta_K^{l_K} \geq w_K, \quad l_K = 1, ..., L_K \\
& \theta_i \geq 0, \quad i = 1, 2, ..., L.
\end{aligned}
$$

Theorem 3.7 (Reduced test):

A given portfolio $\boldsymbol{\tau}$ *is SSD efficient if and only if* $\upsilon(\boldsymbol{\tau}|\mathbf{w}) = 0$ *for* $\mathbf{w} > \mathbf{0}$.

The dual LP program to (3.14) is

$$
\begin{aligned}
\upsilon(\boldsymbol{\tau}|\mathbf{w}) \quad &= \max_{\boldsymbol{\lambda},\rho_{l_k l_{k+1}},\rho_{t_K}} \sum_{k=1}^{K-1} w_k \sum_{l_k=1}^{L_k}\sum_{l_{k+1}=1}^{L_{k+1}} \rho_{l_k l_{k+1}} + w_K \sum_{l_K=1}^{L_K} \rho_{l_K} \qquad (3.15) \\
\text{s.t.} \quad \sum_{l_2=1}^{L_2} \rho_{l_1 l_2} \quad &\leq \quad T^{-1}\sum_{t\in\Phi_1^{l_1}} (\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}), \quad l_1 = 1, ..., L_1 \\
\sum_{l_{k+1}=1}^{L_{k+1}} \rho_{l_k l_{k+1}} - \sum_{l_{k-1}=1}^{L_{k-1}} \rho_{l_{k-1} l_k} \quad &\leq \quad T^{-1}\sum_{t\in\Phi_k^{l_k}} (\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}), \quad l_k = 1, ..., L_k, \\
& \qquad\qquad k = 2, 3, ..., K-1 \\
\rho_{l_K} - \sum_{l_{K-1}=1}^{L_{K-1}} \rho_{l_{K-1} l_K} \quad &\leq \quad T^{-1}\sum_{t\in\Phi_K^{l_K}} (\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}), \quad l_K = 1, ..., L_K \\
\rho_{l_k l_{k+1}}, \rho_{l_K} \quad &\geq \quad 0, \quad l_k = 1, ..., L_k, \quad k = 1, ..., K-1 \\
\boldsymbol{\lambda} \quad &\in \quad \Lambda.
\end{aligned}
$$

The comparison of computational complexity of all considered tests is presented in Table 3.1.

Table 3.1: *Mathematical programming problem dimensions.*

The table shows the Linear Programming problem dimensions for various SSD portfolio efficiency tests. Included are the Post (2003) test, the necessary test of Kuosmanen (2004), the sufficient test of Kuosmanen (2004), the Kopa and Chovanec (2008) test, the full dual test and the reduced dual test of Kopa and Post (2013). The table shows the general dimensions as a function of the number of scenarios (T) and the number of base assets (M). It is assumed that the portfolio set is the simplex (1.7) and that no ties occur ($T = K$). For the sake of comparison with Kuosmanen (2004) and Kopa and Chovanec (2008) tests (which have no known utility representation), we show the dimensions of the dual formulation (in terms of cumulative returns) of the Post test and our tests. Panel A includes non-negativity constraints; Panel B excludes these constraints.

Panel A: Including the non-negativity constraints

Test	Constrains × Variables		
	T scenarios, M assets	T=120, M =12	T=480, M = 12
Post (2003) dual test	$(2T + M) \times (T + M - 1)$	252×131	972×491
Kuosmanen (2004) necessary condition	$(T^2 + 3T + M + 1) \times (T^2 + M)$	14773×14412	231853×230412
Kuosmanen (2004) sufficient condition	$(4T^2 + 3T + M + 1) \times (3T^2 + M)$	57973×43212	923053×691212
Kopa and Chovanec (2008) test	$(2T^2 + 2T + M + 1) \times (T^2 + 2T + M)$	29053×14652	461773×231372
Kopa and Post (2013) dual test	$(2T^2 + T + M + 1) \times (T^2 + T + M)$	28933×14532	461293×230892
Kopa and Post (2013) reduced dual test	$(2T + M + 1) \times (T + M)$	253×132	973×492

Panel B: Excluding the non-negativity constraints

Test	Constrains × Variables		
	T scenarios, M assets	T=120, M =12	T=480, M = 12
Post (2003) dual test	$(T + 1) \times (T + M - 1)$	121×131	481×491
Kuosmanen (2004) necessary condition	$(3T + 1) \times (T^2 + M)$	361×14412	1441×230412
Kuosmanen (2004) sufficient condition	$(T^2 + 3T + 1) \times (3T^2 + M)$	14761×43212	231841×691212
Kopa and Chovanec (2008) test	$(T^2 + T + 1) \times (T^2 + 2T + M)$	14521×14652	230881×231372
Proposed dual test	$(T^2 + 1) \times (T^2 + T + M)$	14401×14532	230401×230892
Reduced dual test	$(T + 1) \times (T + M)$	121×132	481×492

Chapter 4

Portfolio efficiency test with respect to N-th order stochastic dominance

This chapter summarizes the NSD portfolio efficiency tests for $N \geq 2$ derived in Post and Kopa (2013). It exploits Definition 1.3 of N-th order stochastic dominance relation. Similarly to the general SSD portfolio efficiency test presented in Section 3.4, the NSD portfolio efficiency test is formulated in the primal-dual way. In this chapter the discrete probability distribution of returns of M assets is still assumed. However, the scenarios $\mathbf{x}^t$, $t = 1, 2, ..., T$ need not be equiprobable, they occur with probability p_t. The results are derived for Λ given by (1.7). However, they can be easily modified for any nonempty bounded polytope set.

4.1 Linear formulation in terms of piece-wise polynomial utility

Post and Kopa (2013) consider the following reformulation of Definition 1.3.

Lemma 4.1:

An evaluated portfolio $\boldsymbol{\tau} \in \Lambda$ is efficient in terms of N-th order stochastic dominance, $N \geq 2$ relative to all feasible portfolios $\boldsymbol{\lambda} \in \Lambda$ if it is an optimal solution of the maximizing expected utility problem for some admissible utility *function $u \in U_N$, that is:*

$$\sum_{t=1}^{T} p_t u(\mathbf{x}^t \boldsymbol{\tau}) \geq \sum_{t=1}^{T} p_t u(\mathbf{x}^t \boldsymbol{\lambda}) \quad \forall \boldsymbol{\lambda} \in \Lambda \Longleftrightarrow$$
$$\sum_{t=1}^{T} p_t u'(\mathbf{x}^t \boldsymbol{\tau})(\mathbf{x}^t \boldsymbol{\tau} - x_j^t) \geq 0, \quad j = 1, 2, ..., M.$$

The lemma follows from the Karush-Kuhn-Tucker first-order condition for selecting the optimal combination of assets: $\max_{\boldsymbol{\lambda} \in \Lambda} \sum_{t=1}^{T} p_t u\left(\mathbf{x}^t \boldsymbol{\lambda}\right)$. An admissible function $u \in U_N$ is now chosen as any function from U_N which is not constant on the relevant interval $\left\langle (X\boldsymbol{\tau})^{[1]}, (X\boldsymbol{\tau})^{[T]} \right\rangle$ in order to avoid some trivial cases. This reformulation was first introduced by Post (2003) for SSD ($N = 2$) and applies also for higher-order criteria ($N > 2$), but it does not apply for FSD ($N = 1$). Kopa and Post (2009) presented a different utility-based formulation for this case, see Section 2.2.

Lemma 4.2:

For any utility function $u \in U_N$, $N \geq 2$, and a discrete set of outcomes $z_1 \leq \cdots \leq z_T$, we represent the levels of utility and its derivatives by piecewise polynomial functions:

$$\begin{aligned}
u\left(z_t\right) &= \sum_{n=0}^{N-2} \beta_n \left(z_t - z_T\right)^n + \sum_{k=t}^{T} \gamma_k (z_t - z_k)^{N-1} \\
u^q\left(z_t\right) &= \sum_{n=q}^{N-2} \frac{n!}{(n-q)!} \beta_n \left(z_t - z_T\right)^{n-q} + \frac{(N-1)!}{(N-q-1)!} \sum_{k=t}^{T} \gamma_k (z_t - z_k)^{N-q-1}, \\
& \quad q = 1, \ldots, N-1
\end{aligned}$$

where

$$\begin{aligned}
\beta_n &= \frac{u^n(z_T)}{n!}, \quad n = 0, 1, \ldots, N-2 \\
\gamma_k &= \frac{u^{N-1}\left(z_{k-1}^*\right) - u^{N-1}\left(z_k^*\right)}{(N-1)!}, \quad k = 2, \ldots, T-1 \ , \ \gamma_T = \frac{u^{N-1}\left(z_{T-1}^*\right)}{(N-1)!}
\end{aligned}$$

for some values $z_k^* \in \langle z_k, z_{k+1} \rangle, \quad k = 1, \dots, T-1$ *such that*

$$\begin{aligned} (-1)^n \beta_n &\leq 0, \ n = 1, \dots, N-2 & (4.1) \\ (-1)^{N-1} \gamma_k &\leq 0, \ k = 1, 2, \dots, T. & (4.2) \end{aligned}$$

Moreover, for all parameters satisfying (4.1)-(4.2) *we can construct an admissible utility function* $u \in U_N$.

Combining Lemma 1 with Lemma 2, Post and Kopa (2013) proposed to solve the following linear program:

$$\theta^* = \min_{\beta_n, \gamma_k, \theta} \theta \tag{4.3}$$

$$\text{s.t.} \quad \left(\mathbf{x}^t \boldsymbol{\tau} - x_j^t\right) p_r \sum_{t=1}^{T} \left(\sum_{n=1}^{N-2} n \beta_n \left(\mathbf{x}^t \boldsymbol{\tau} - \mathbf{x}^T \boldsymbol{\tau}\right)^{n-1} + \right.$$

$$\left. (N-1) \sum_{k=t}^{T} \gamma_k \left(\mathbf{x}^t \boldsymbol{\tau} - \mathbf{x}^k \boldsymbol{\tau}\right)^{N-2} \right) + \theta \geq 0, \ j = 1, \cdots, M$$

$$(-1)^n \beta_n \leq 0, \ n = 1, \dots, N-2$$

$$(-1)^{N-1} \gamma_k \leq 0, \ k = 1, 2, \dots, T$$

$$\sum_{t=1}^{T} \left(\sum_{n=1}^{N-2} n \beta_n \left(\mathbf{x}^t \boldsymbol{\tau} - \mathbf{x}^T \boldsymbol{\tau}\right)^{n-1} + (N-1) \sum_{k=t}^{T} \gamma_k \left(\mathbf{x}^t \boldsymbol{\tau} - \mathbf{x}^k \boldsymbol{\tau}\right)^{N-2} \right) p_r = 1,$$

where the last restriction on average marginal utility is a harmless standardization to avoid the trivial solution of an indifferent decision maker.

Theorem 4.1:

A portfolio $\boldsymbol{\tau}$ *is NSD efficient if and only if* θ^* *given by (4.3) is equal to zero.*

4.2 Dual formulation in terms of lower partial moments

Consider the following definition for the n-th order lower partial moment (Bawa (1975)) for portfolio $\boldsymbol{\tau}$ and threshold value w:

$$\mathrm{LPM}^n_{\boldsymbol{\tau}}(w) = \sum_{t=1}^{T} p_t \left(w - \mathbf{x}^t \boldsymbol{\tau}\right)^n 1_{(\mathbf{x}^t\boldsymbol{\tau} \le w)}.$$

For analyzing the NSD efficiency, we extend the lower partial moments to the n-th order co-lower partial moments between a portfolio $\boldsymbol{\lambda}$ and portfolio $\boldsymbol{\tau}$ as follows (Bawa and Lindenberg (1977)):

$$\mathrm{coLPM}^n_{\tau,\lambda}(w) = \sum_{t=1}^{T} \left(\mathbf{x}^t \boldsymbol{\lambda}\right) p_t \left(w - \mathbf{x}^t \boldsymbol{\tau}\right)^n 1_{(\mathbf{x}^t\boldsymbol{\tau} \le w)}.$$

Using n-th order co-lower partial moments between a portfolio $\boldsymbol{\lambda}$ and portfolio $\boldsymbol{\tau}$, a dual program to (4.3) with slightly different standardization $((-1)^N(N-1)\gamma_T = 1)$ is:

$$\begin{aligned} & & \eta^* = \max_{\boldsymbol{\lambda},\eta} \eta & & \quad (4.4) \\ \text{s.t. } \mathrm{coLPM}^n_{\tau,\lambda}\left(\mathbf{x}^T\boldsymbol{\tau}\right) &\le \mathrm{coLPM}^n_{\tau,\tau}\left(\mathbf{x}^T\boldsymbol{\tau}\right), & & n = 0, \ldots, N-3 \\ \mathrm{coLPM}^{N-2}_{\tau,\lambda}\left(\mathbf{x}^t\boldsymbol{\tau}\right) &\le \mathrm{coLPM}^{N-2}_{\tau,\tau}\left(\mathbf{x}^t\boldsymbol{\tau}\right), & & t = 1, \ldots, T-1 \\ \mathrm{coLPM}^{N-2}_{\tau,\lambda}\left(\mathbf{x}^T\boldsymbol{\tau}\right) + \eta &\le \mathrm{coLPM}^{N-2}_{\tau,\tau}\left(\mathbf{x}^T\boldsymbol{\tau}\right). \\ \lambda &\in \Lambda. \end{aligned}$$

Theorem 4.2:

A portfolio $\boldsymbol{\tau}$ is NSD efficient if and only if η^* given by (4.4) is equal to zero.

Chapter 5

Robustness in portfolio efficiency tests with respect to SSD criterion

In all SSD portfolio efficiency tests, the scenario approach is assumed, the returns of assets are modeled by discrete distribution. Therefore, especially for SSD efficient portfolios, one can ask how sensitive are the SSD efficiency tests with respect to changes in scenarios, or how the original scenarios can be changed such that a given SSD efficient portfolio remains SSD efficient for perturbed scenarios, too. To circumvent this problem, we present in this chapter several approaches to robustness in SSD portfolio efficiency testing. In this chapter, we follow the notation from Chapter 3.

5.1 δ-SSD portfolio efficiency test

In this section we follow Kopa (2010) in introducing a δ-SSD portfolio efficiency as a new type of portfolio efficiency with respect to the second-order stochastic dominance criteria. Fixing the number of equiprobable scenarios, we identify the maximal perturbation of original scenarios satisfying δ-SSD portfolio efficiency condition for a given portfolio. The magnitude of this maximal perturbation, expressed in terms of a distance between original and perturbed scenarios, can be considered as a measure of δ-SSD efficiency and the limiting case for $\delta \to 0$ leads to a new SSD efficiency measure. We consider only special perturbations where all scenarios are equiprobable and the number of scenarios is fixed. Contrary to the SSD inefficiency measures discussed in Chapter 3, δ-SSD and SSD portfolio efficiency measures are defined as measures of stability. While all tests mentioned in Chapter 3 give an information about a measure (degree) of inefficiency (if a tested portfolio is SSD inefficient), the new measures based on δ-SSD portfolio efficiency assign to each SSD efficient portfolio a number that can be seen as a measure of stability or a measure of SSD portfolio efficiency. We start with definition of δ-SSD relation and δ-SSD portfolio efficiency for arbitrary $\delta > 0$.

Definition 5.1:

Let $\delta > 0$. Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by the δ-second-order stochastic dominance ($\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{\delta-SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$) if there exists a double stochastic matrix $W = \{w\}_{ij}$ such that $X\boldsymbol{\lambda} \geq WX\boldsymbol{\tau}$ and $\mathbf{1}'X\boldsymbol{\lambda} - \mathbf{1}'X\boldsymbol{\tau} \geq \delta$.

The strictly positive parameter δ in Definition 5.1 is chosen sufficiently small, that is, such that if $X\boldsymbol{\lambda} \geq WX\boldsymbol{\tau}$ and $\mathbf{1}'X\boldsymbol{\lambda} - \mathbf{1}'X\boldsymbol{\tau} < \delta$ then vectors $X\boldsymbol{\lambda}$ and $WX\boldsymbol{\tau}$ are empirically indistinguishable. Moreover, as it was shown in Kuosmanen (2004) and references therein, the existence of a double stochastic matrix $W = \{w\}_{ij}$ such that $X\boldsymbol{\lambda} \geq WX\boldsymbol{\tau}$ is equivalent to weak SSD relation between $\boldsymbol{\varrho}'\boldsymbol{\lambda}$ and $\boldsymbol{\varrho}'\boldsymbol{\tau}$ as defined in Section 1.2. Therefore, if $\boldsymbol{\lambda}$ δ-SSD dominates portfolio $\boldsymbol{\tau}$ for some $\delta > 0$ then $\boldsymbol{\lambda}$ SSD dominates $\boldsymbol{\tau}$.

Definition 5.2:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is δ-SSD inefficient if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{\delta-SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is δ-SSD efficient.

We modify the Kuosmanen test, see Section 3.2, to δ-SSD portfolio efficiency test.

Lemma 5.1:

Let

$$\theta_\delta^* = \max_{W,\boldsymbol{\lambda}} \sum_{t=1}^{T} \left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) \tag{5.1}$$

$$\begin{array}{rcl} s.t. \qquad X\boldsymbol{\lambda} & \geq & WX\boldsymbol{\tau} \\ \sum_{t=1}^{T}\left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) & \geq & \delta \\ \sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} & = & 1, \quad w_{ij} \geq 0, \qquad i,j = 1,2,\ldots,T \\ \boldsymbol{\lambda} & \in & \Lambda. \end{array}$$

If an optimal solution of (5.1) exists then portfolio $\boldsymbol{\tau}$ is δ-SSD inefficient and $\boldsymbol{\varrho}'\boldsymbol{\lambda}^* \succ_{\delta-SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$. Otherwise, $\boldsymbol{\tau}$ is δ-SSD efficient portfolio.

Until now a fixed scenario matrix was considered and all portfolio efficiency tests were done for this scenario matrix. Unfortunately, usually we do not have perfect information about the distribution of returns. Therefore, the stability of SSD portfolio efficiency and δ-SSD portfolio efficiency with respect to changes in scenario matrix is of interest.

Since the SSD portfolio efficiency tests and the δ-SSD portfolio efficiency test are derived under the assumption of equiprobable scenarios collected in matrix X we will consider only perturbation matrices X_p of the original matrix X which have exactly T rows, that is, we admit only approximations with T equiprobable scenarios. Let $\mathcal{X}_p$ be the set of all such perturbation matrices. and let matrix $\Upsilon = \{\upsilon_{ij}\}_{i,j=1}^{T}$ be defined as $\Upsilon = X_p - X$. Finally, let $D(X, X_p) = \max_{i,j} |\upsilon_{ij}|$ denote a distance between matrices X and X_p on $\mathcal{X}_p$.

Definition 5.3:

The δ-SSD portfolio efficiency measure γ_δ of δ-SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ is defined as the optimal value of the following optimization problem:

$$\gamma_\delta(\boldsymbol{\tau}) = \max \quad \varepsilon \tag{5.2}$$
$$s.t. \;\; \boldsymbol{\tau} \;\; is \;\; \delta - SSD \;\; efficient\ for\ all \;\; X_p \in \mathcal{X}_p \;\; such\ that \;\; D(X, X_p) \leq \varepsilon.$$

This measure gives us an information how large is the neighborhood of X such that the portfolio $\boldsymbol{\tau}$ remains classified as δ-SSD efficient for all matrices from this neighborhood. The problem (5.2) consists of infinitely many δ-SSD efficiency constraints. Moreover, according to the Lemma 5.1, each constraint involves a maximization problem what makes problem (5.2) practically unsolvable. Therefore we reinterpret the δ-SSD portfolio efficiency measure for a given δ-SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ as the minimal distance between the original matrix X and any other matrix X_p that makes portfolio $\boldsymbol{\tau}$ δ-SSD inefficient, that is,

$$\begin{aligned} \gamma_\delta(\boldsymbol{\tau}) = \min_{X_p \in \mathcal{X}_p} \quad & D(X, X_p) \\ \text{s.t.} \ \ \boldsymbol{\tau} \ \text{ is } \ \delta - SSD \ & \text{ inefficient for } \ X_p. \end{aligned} \tag{5.3}$$

Using Lemma 5.1, $\Upsilon = X_p - X$ and $D(X, X_p) = \max_{i,j} |\upsilon_{ij}|$, the SSD portfolio efficiency measure $\gamma_\delta(\boldsymbol{\tau})$ can be computed using the following non-linear program:

$$\gamma_\delta(\boldsymbol{\tau}) = \min_{\boldsymbol{\lambda} \in \Lambda, \Upsilon, \varepsilon} \varepsilon \tag{5.4}$$

$$\begin{array}{rcl} \text{s.t.} \qquad (X + \Upsilon)\boldsymbol{\lambda} - W(X + \Upsilon)\boldsymbol{\tau} & \geq & 0 \\ \displaystyle\sum_{t=1}^{T} \left((\mathbf{x}^t + \boldsymbol{\upsilon}^t)\boldsymbol{\lambda} - (\mathbf{x}^t + \boldsymbol{\upsilon}^t)\boldsymbol{\tau}\right) & \geq & \delta \\ \displaystyle\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} = 1, \quad w_{ij} & \geq & 0 \quad i,j = 1,2,\ldots,T \\ -\varepsilon \ \leq \ \upsilon_{ij} & \leq & \varepsilon \quad i,j = 1,2,\ldots,T, \end{array}$$

where $\boldsymbol{\upsilon}^t = (\upsilon_{t1}, \upsilon_{t2}, \ldots, \upsilon_{tT})$ is the t-th row of matrix Υ. For a given portfolio $\boldsymbol{\tau}$ we have $\gamma_\delta(\boldsymbol{\tau}) \geq 0$ for all $\delta > 0$. Moreover, if $\delta_1 < \delta_2$ then the set of feasible solutions of (5.5) is larger for δ_1 than for δ_2 and consequently $\gamma_{\delta_1}(\boldsymbol{\tau}) \leq \gamma_{\delta_2}(\boldsymbol{\tau})$. Therefore, we can define a measure of SSD efficiency in the following way.

Definition 5.4:

The SSD portfolio efficiency measure γ of SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ is defined as: $\gamma(\boldsymbol{\tau}) = \lim_{\delta \to 0^+} \gamma_\delta(\boldsymbol{\tau}) = \inf_{\delta > 0} \gamma_\delta(\boldsymbol{\tau})$.

5.2 ϵ-SSD portfolio efficiency test with unequal probabilities

Contrary to the previous section, Dupačová and Kopa (2012) assume that random vector of returns takes scenarios with probabilities $\mathbf{p} = (p_1, p_2, ..., p_T)$. As in Section 2.1, for any portfolio $\boldsymbol{\lambda} \in \Lambda$, let $(-X\boldsymbol{\lambda})^{[k]}$ be the k-th smallest element of $(-X\boldsymbol{\lambda})$, i.e. $(-X\boldsymbol{\lambda})^{[1]} \leq (-X\boldsymbol{\lambda})^{[2]} \leq \ldots \leq (-X\boldsymbol{\lambda})^{[T]}$ and let $I(\boldsymbol{\lambda})$ be a permutation of the index set $I = \{1, 2, ..., T\}$ such that $-\mathbf{x}^{i(\boldsymbol{\lambda})}\boldsymbol{\lambda} = (-X\boldsymbol{\lambda})^{[i]}$. Accordingly, we can order the corresponding probabilities and we denote $p_i^{\boldsymbol{\lambda}} = p_{i(\boldsymbol{\lambda})}$. Hence, $p_i^{\boldsymbol{\lambda}} = P(-\boldsymbol{\varrho}'\boldsymbol{\lambda} = (-X\boldsymbol{\lambda})^{[i]})$. Consider also the cumulative probabilities: $q_0^{\boldsymbol{\lambda}} = 0$ and $q_s^{\boldsymbol{\lambda}} = \sum_{t=1}^{s} p_i^{\boldsymbol{\lambda}}$, $s = 1, 2, ...T$. The same notation is applied for the tested portfolio $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_M)'$.

Allowing unequal scenario probabilities, Dupačová and Kopa (2012) modified the Kopa and Chovanec test (2008) in the following way:

Theorem 5.1:

Let

$$\begin{aligned} \xi(\boldsymbol{\tau}, X, \mathbf{p}) &= \min_{a_s, \boldsymbol{\lambda}} \sum_{s=0}^{T} a_s \qquad (5.5) \\ \text{s.t.} \quad & \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \leq a_s, \quad s = 0, 1, ..., T \\ & a_s \leq 0, \quad s = 0, 1, ..., T \\ & \boldsymbol{\lambda} \in \Lambda. \end{aligned}$$

A given portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if $\xi(\boldsymbol{\tau}, X, \mathbf{p}) = 0$. If $\xi(\boldsymbol{\tau}, X, \mathbf{p}) < 0$ then the optimal portfolio $\boldsymbol{\lambda}^*$ in (5.5) is SSD efficient and it dominates portfolio $\boldsymbol{\tau}$ by SSD.

The objective function of (5.5) represents the sum of differences between CVaRs of a portfolio $\boldsymbol{\lambda}$ and CVaRs of the tested portfolio $\boldsymbol{\tau}$. The differences are considered in points $q_s^{\boldsymbol{\lambda}}$, $s = 0, 1, ..., T$. All differences must be nonpositive and at least one negative to guarantee that portfolio $\boldsymbol{\lambda}$ dominates portfolio $\boldsymbol{\tau}$.

Assume now, that random returns $\bar{\boldsymbol{\varrho}}$ have the probability distribution $\bar{P}$ and take again values $\mathbf{x}^t$, $t = 1, 2, ..., T$ but with other probabilities $\bar{\mathbf{p}} = (\bar{p}_1, \bar{p}_2, ..., \bar{p}_T)$. We define the distance between P and $\bar{P}$ as $d(\bar{P}, P) = \max_i |\bar{p}_i - p_i|$.

Definition 5.5:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is ϵ-SSD inefficient if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ and $\bar{P}$ such that $d(\bar{P}, P) \leq \epsilon$ with $\bar{\boldsymbol{\varrho}}'\boldsymbol{\lambda} \succ_{SSD} \bar{\boldsymbol{\varrho}}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is ϵ-SSD efficient.

The introduced ϵ-SSD efficiency is a robustification of the classical SSD portfolio efficiency. It guarantees stability of the SSD efficiency classification with respect to small changes (prescribed by parameter ϵ) in probability vector $\mathbf{p}$. A given portfolio $\boldsymbol{\tau}$ is ϵ-SSD efficient if and only if no portfolio $\boldsymbol{\lambda}$ SSD dominates $\boldsymbol{\tau}$ neither for the original probabilities $\mathbf{p}$ nor for arbitrary probabilities $\bar{\mathbf{p}}$ from ϵ-neighborhood of the original vector $\mathbf{p}$. For testing ϵ-SSD efficiency of a given portfolio $\boldsymbol{\tau}$ we modify (5.5) in order to introduce a new measure of ϵ-SSD efficiency:

$$
\begin{aligned}
\xi_\epsilon(\boldsymbol{\tau}, X, \mathbf{p}) \quad = \quad & \min_{a_s, \boldsymbol{\lambda}, \bar{\mathbf{p}}} \sum_{s=0}^{T} a_s && (5.6)\\
\text{s.t.} \quad & \mathrm{CVaR}_{\bar{q}_s^{\boldsymbol{\lambda}}}(-\bar{\boldsymbol{\varrho}}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\bar{q}_s^{\boldsymbol{\lambda}}}(-\bar{\boldsymbol{\varrho}}'\boldsymbol{\tau}) \leq a_s, \quad s = 0, 1, ..., T \\
& \bar{q}_s^{\boldsymbol{\lambda}} = \sum_{i=1}^{s} \bar{p}_i^{\boldsymbol{\lambda}}, \quad s = 1, ..., T \\
& \bar{q}_0^{\boldsymbol{\lambda}} = 0 \\
& \sum_{i=1}^{S} \bar{p}_i = 1 \\
& -\epsilon \leq \bar{p}_i - p_i \leq \epsilon, \quad i = 1, 2, ..., T \\
& \bar{p}_i \geq 0, \quad i = 1, 2, ..., T \\
& a_s \leq 0, \quad s = 0, 1, ..., T \\
& \boldsymbol{\lambda} \in \Lambda.
\end{aligned}
$$

Theorem 5.2:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is ϵ-SSD efficient if and only if $\xi_\epsilon(\boldsymbol{\tau}, X, \mathbf{p})$ given by (5.6) is equal to zero.

5.3 Resistance of SSD portfolio efficiency with respect to additional scenarios

In the previous sections, we assumed a fixed set of scenarios. In many practical applications, an additional scenario may be of interest, e.g. for stress testing. Therefore, the aim of this section is to analyze the robustness of SSD portfolio efficiency with respect to an additional scenario denoted by $\mathbf{x}^{T+1}$. For a contamination parameter $\kappa \in [0,1]$, we assume that the random return $\tilde{\boldsymbol{\varrho}}(\kappa)$ takes values $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^{T+1}$ with probabilities $\tilde{\mathbf{p}}(\kappa) = ((1-\kappa)p_1, (1-\kappa)p_2, ..., (1-\kappa)p_T, \kappa)$. More details about contamination techniques can be found in e.g. Dupačová (1990, 1996, 1998, 2006), Dupačová and Polívka (2007) or Dupačová and Kopa (2012). The cumulative probabilities for portfolio $\boldsymbol{\lambda}$ are

$$\tilde{q}_s^{\boldsymbol{\lambda}}(\kappa) = \sum_{i=1}^{s} \tilde{p}_i^{\boldsymbol{\lambda}}(\kappa) = \sum_{i=1}^{s} P(-\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\lambda} = (-\tilde{X}\boldsymbol{\lambda})^{[i]}), \quad s = 1, 2, ..., T+1$$

$$\tilde{q}_0^{\boldsymbol{\lambda}}(\kappa) = 0$$

where $\tilde{X}$ is the extended scenario matrix, that is,

$$\tilde{X} = \begin{pmatrix} X \\ \mathbf{x}^{T+1} \end{pmatrix}$$

and the same notation is used for portfolio $\boldsymbol{\tau}$.

Definition 5.6:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is directionally SSD inefficient with respect to $\mathbf{x}^{T+1}$ if it exists $\kappa_0 > 0$ such that for every $\kappa \in [0, \kappa_0]$ there is a portfolio $\boldsymbol{\lambda}(\kappa) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\lambda}(\kappa) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\tau}$.

Definition 5.7:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is directionally SSD efficient with respect to $\mathbf{x}^{T+1}$ if there does not exist $\kappa_0 > 0$ such that for every $\kappa \in [0, \kappa_0]$ there is a portfolio $\boldsymbol{\lambda}(\kappa) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\lambda}(\kappa) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\tau}$.

According to these definitions, a given portfolio is classified as directionally SSD efficient (inefficient) with respect to scenario $\mathbf{x}^{T+1}$ if it is SSD efficient (inefficient) and a sufficiently small contamination of the original probability distribution of returns by the additional scenario does not change the SSD efficiency classification, that is, the SSD efficient (inefficient) portfolio remains SSD efficient (inefficient). Applying (5.5) to contaminated data, portfolio $\boldsymbol{\lambda}(\kappa) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\lambda}(\kappa) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\tau}$ exists if and only if $\xi(\boldsymbol{\tau}, \tilde{X}, \tilde{\mathbf{p}}(\kappa)) < 0$, where

$$\xi(\boldsymbol{\tau}, \tilde{X}, \tilde{\mathbf{p}}(\kappa)) = \min_{a_s,\boldsymbol{\lambda}} \sum_{s=0}^{T} a_s \tag{5.7}$$

$$\begin{aligned} \text{s.t.} \quad & \mathrm{CVaR}_{\tilde{q}_s^{\boldsymbol{\lambda}}}(-\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\tilde{q}_s^{\boldsymbol{\lambda}}}(-\tilde{\boldsymbol{\varrho}}(\kappa)'\boldsymbol{\tau}) \le a_s, \qquad s = 0, 1, ..., T \\ & a_s \le 0, \qquad s = 0, 1, ..., T \\ & \boldsymbol{\lambda} \in \Lambda. \end{aligned}$$

Using contamination bounds Dupačová and Kopa (2012) derive a sufficient condition for directional SSD efficiency and directional SSD inefficiency with respect to additional scenario $\mathbf{x}^{T+1}$.

Theorem 5.3:

Let $\boldsymbol{\tau} \in \Lambda$ be an SSD efficient portfolio for the noncontaminated distribution P. Let

$$\mathbf{x}^{T+1}\boldsymbol{\tau} \ge \mathbf{x}^{T+1}\boldsymbol{\lambda} \quad \text{for all} \quad \boldsymbol{\lambda} \in \Lambda. \tag{5.8}$$

Then $\boldsymbol{\tau} \in \Lambda$ is directionally SSD efficient with respect to $\mathbf{x}^{T+1}$.

Theorem 5.4:

Let $\boldsymbol{\tau} \in \Lambda$ be an SSD inefficient portfolio for the noncontaminated distribution P. If there exists a portfolio $\boldsymbol{\lambda} \in \Lambda$ such that

$$\mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \;<\; 0, \qquad s = 0, 1, ..., T \tag{5.9}$$

$$\mathbf{x}^{T+1}\boldsymbol{\lambda} \;\ge\; \min((X\boldsymbol{\tau})^{[1]}, \mathbf{x}^{T+1}\boldsymbol{\tau}) \tag{5.10}$$

then $\boldsymbol{\tau}$ is directionally SSD inefficient with respect to $\mathbf{x}^{T+1}$.

Condition (5.10) is needed to guarantee that even in the contaminated case the smallest return of portfolio $\boldsymbol{\lambda}$ is larger than or equal to that of portfolio $\boldsymbol{\tau}$ what is a necessary condition of SSD relation. The proofs and more details can be found in Dupačová and Kopa (2012).

Bibliography

[1] P. Artzner, J. Delbaen, J. Eber, D. Heath: Coherent measures of risk. *Matematical Finance*, **9** (1999), 202–228.

[2] V. S. Bawa: Optimal rules for ordering uncertain prospects, *Journal of Financial Economics*, **2**, 1 (1975), 95-121.

[3] V. S. Bawa, E. B. Lindenberg: Capital market equilibrium in a mean-lower partial moment framework, *Journal of Financial Economics*, **5** (1977), 189- 200.

[4] J. Dupačová: Stability and sensitivity analysis in stochastic programming, *Annals of Operations Research*, **27** (1990), 115–142.

[5] J. Dupačová: Scenario based stochastic programs: Resistance with respect to sample. *Annals of Operations Research*, **64** (1996), 21–38.

[6] J. Dupačová: Reflections on robust optimization, in: Marti, K., Kall, P. (eds.), *Stochastic Programming Methods and Technical Applications*, pp. 111–127, LNEMS 437, 1998, Springer, Berlin.

[7] J. Dupačová: Stress testing via contamination. In: K. Marti et al. (Eds.), *Coping with Uncertainty, Modeling and Policy Issues*, pp. 29–46, LNEMS 581, 2006, Springer, Berlin.

[8] J. Dupačová, M. Kopa: Robustness in stochastic programs with risk constraints, *Annals of Operations Research*, **200**, 1 (2012), 55 – 74.

[9] J. Dupačová, M. Kopa: Robustness of optimal portfolios under risk and stochastic dominance constraints, submitted in *European Journal of Operational Research.*

[10] J. Dupačová, J. Polívka: Stress testing for VaR and CVaR, *Quantitative Finance*, **7** (2007), 411–421.

[11] P. Glasserman: *Monte Carlo Methods in Financial Engineering*, Springer Verlag, 2004.

[12] J. Hadar and W. R. Russell: Rules for Ordering Uncertain Prospects, *American Economic Review*, **59**, 1 (1969), 25–34.

[13] G. Hanoch, H. Levy: The efficiency analysis of choices involving risk, *Review of Economic Studies* **36** (1969), 335–346.

[14] P. Jackel: *Monte Carlo Methods in Finance*, John Wiley & Sons, 2002.

[15] M. Kopa: Measuring of second-order stochastic dominance portfolio efficiency, *Kybernetika*, **46**, 3 (2010), 488 – 500.

[16] M. Kopa: Robustness in SSD portfolio efficiency testing, *Proceedings of International Conference on Operations Research*, pp. 527 - 532 Springer, Berlin, 2012.

[17] M. Kopa, P. Chovanec: A second-order stochastic dominance portfolio efficiency measure, Kybernetika **44**, 2 (2008), 243-258.

[18] M. Kopa, T. Post: A portfolio optimality test based on the first-order stochastic dominance criterion, *Journal of Financial and Quantitative Analysis* **44**, 5 (2009), 1103-1124.

[19] M. Kopa, T. Post: A general test for portfolio efficiency, working paper 2013, available at www.ssrn.com.

[20] T. Kuosmanen: Efficient diversification according to stochastic dominance criteria, *Management Science* **50**, 10 (2004), 1390–1406.

[21] H. Levy: *Stochastic Dominance: Investment Decision Making Under Uncertainty*. Second edition. Springer Science, New York 2006.

[22] J. von Neumann, O. Morgenstern: *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

[23] H. M. Markowitz: Portfolio Selection, *Journal of Finance*, **7**, 1 (1952), 77–91.

[24] W. Ogryczak, A. Ruszczyński: Dual stochastic dominance and related mean-risk models, *SIAM J. Optim.*, **13** (2002), 60–78.

[25] T. Post: Empirical tests for stochastic dominance efficiency, *Journal of Finance*, **58** (2003), 1905–1932.

[26] T. Post, M. Kopa: General Linear Formulations of Stochastic Dominance Criteria, *European Journal of Operational Research*, Online first 2013.

[27] J. P. Quirk, R. Saposnik: Admissibility and measurable utility functions, *Review of Economic Studies*, **29** (1962), 140-146.

[28] W. R. Russell, T. K. Seo: Representative sets for stochastic dominance rules, Studies in the Economics of Uncertainty: In honor of Josef Hadar (Thomas B. Fomby and Tae Kun Seo eds.), Springer Verlag, New York, (1989), 59–76.

[29] R. T. Rockafellar, S. Uryasev: Optimization of Conditional Value-At-Risk. *Journal of Risk* **2**, 3, (2000), 21–41.

[30] R. T. Rockafellar, S. Uryasev: Conditional value-at-risk for general loss distributions, *Journal of Banking* & *Finance* **26** (2002), 1443–1471.

[31] R. T. Rockafellar, S. Uryasev, M. Zabarankin: Generalized deviations in risk analysis, *Finance Stochastics*, **10** (2006), 51–74.

[32] P. D. Thistle: Negative moments, risk aversion, and stochastic dominance, *Journal of Financial and Quantitative Analysis*, **2** (1993), 301–311.

[33] G. A. Whitmore: Stochastic dominance for the class of completely monotonic utility functions, *Studies in the Economics of Uncertainty* (Thomas B. Fomby and Tae Kun Seo eds.), Springer Verlag, New York, (1989), 77–88.

JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 44, No. 5, Oct. 2009, pp. 1103–1124

doi:10.1017/S0022109009990251

A Portfolio Optimality Test Based on the First-Order Stochastic Dominance Criterion

Miloš Kopa and Thierry Post*

Abstract

Existing approaches to testing for the efficiency of a given portfolio make strong parametric assumptions about investor preferences and return distributions. Stochastic dominance-based procedures promise a useful nonparametric alternative. However, these procedures have been limited to considering binary choices. In this paper we take a new approach that considers all diversified portfolios and thereby introduce a new concept of first-order stochastic dominance (FSD) optimality of a given portfolio relative to all possible portfolios. Using our new test, we show that the U.S. stock market portfolio is significantly FSD nonoptimal relative to benchmark portfolios formed on market capitalization and book-to-market equity ratios. Without appealing to parametric assumptions about the return distribution, we conclude that no nonsatiable investor would hold the market portfolio in the face of the attractive premia of small caps and value stocks.

I. Introduction

Portfolio analysis and asset pricing tests typically focus on the mean-variance criterion. It is well-known that this criterion implicitly assumes a quadratic utility function or a normal probability distribution, which is quite restrictive in many cases. A good illustration of the limitations of the mean-variance criterion comes from Levy ((1998), p. 2):

> [Consider] two alternative investments: x providing \$1 or \$2 with equal probability and y providing \$2 or \$4 with equal probability, with an identical

*Kopa, kopa@karlin.mff.cuni.cz, Charles University in Prague, Faculty of Mathematics and Physics, Sokolovska 83, 186 75 Prague, Czech Republic; Post, thierrypost@hotmail.com, HAPO Center for Financial Research, Spoorstraat 38, Deventer 7412VE, The Netherlands. We appreciate the comments and suggestions from two anonymous referees, Stephen Brown (the editor), and participants at the stochastic dominance symposium at London School of Economics, November 4–5, 2005. This work was partially supported by grant MSMTCR 0021620839 and the Grant Agency of the Czech Republic (grants 201/05/H007, 402/05/0115, and 201/07/P107). Financial support by Tinbergen Institute, Erasmus Research Institute of Management, and Erasmus Center of Financial Research is gratefully acknowledged. Any remaining errors are the authors' responsibility.

> investment of, say, \$1.1. A simple calculation shows that both the mean and the variance of y are greater than the corresponding parameters of x; hence the mean-variance rule remains silent regarding the choice between x and y. Yet, any rational investor would (and should) select y, because the lowest return on y is equal to the largest return on x.

The criteria of stochastic dominance (SD) are useful nonparametric alternatives. Most notably, first-order stochastic dominance (FSD) is one of the basic concepts of decision making under uncertainty, relying only on the assumption of nonsatiation, or increasing utility. It does not require further specification of the shape of the utility function or the shape of the probability distribution. FSD analysis is generally more difficult to implement than mean-variance analysis. There exist well-known simple tests for establishing FSD relationships between a pair of choice alternatives (see, e.g., Levy (1998), sect. 5.2). Unfortunately, these tests have limited use for portfolio analysis and asset pricing tests, because investors generally can form a large number of portfolios by diversifying across individual assets. Therefore, there is a need to develop a test for establishing if a given portfolio is "FSD efficient" relative to all possible portfolios. Such a test would be a useful alternative for existing mean-variance portfolio efficiency tests (e.g., Gibbons, Ross, and Shanken (1989)), especially if the return distribution is skewed and fat-tailed.

A complication in testing FSD portfolio efficiency is that we must distinguish between efficiency criteria based on "admissibility" and "optimality." There is a subtle difference between these two concepts. A choice alternative is FSD admissible if and only if no other alternative is preferred by all nonsatiable decision-makers. A choice alternative is FSD optimal if and only if it is the optimal choice for at least some nonsatiable decision-maker. For pairwise comparison, the two concepts are identical; alternative x_1 is FSD undominated by alternative x_2 if and only if some nonsatiable decision-maker prefers x_1 to x_2. However, more generally, when multiple choice alternatives are available, FSD admissibility is a necessary but not sufficient condition for FSD optimality. In other words, a choice alternative may be admissible even if it is not optimal for any increasing utility function.

Bawa, Bodurtha, Rao, and Suri (1985) and Kuosmanen (2004) propose FSD tests that apply under more general conditions than a pairwise test does. The two tests differ in a subtle way. While Bawa et al. (1985) consider all convex combinations of the distribution functions of a given set of choice alternatives, Kuosmanen (2004) considers the distribution function for all convex combinations of a given set of choice alternatives. Each of these two tests captures an important aspect of portfolio choice that is not captured by a pairwise FSD test. Still, both tests miss some key aspect of a proper FSD portfolio optimality test, and both tests generally give a necessary but not sufficient condition. The linear programming (LP) test of Bawa et al. (1985) is based on optimality, but it does not account for full diversification across the choice alternatives. Bawa et al. (1985) use a set of undiversified base assets as the choice alternatives. In principle, diversification can enter through the back door by including combinations of the base assets as additional choice alternatives. However, since the number of possible combinations is infinitely large, this approach generally gives only a necessary condition, and potentially it yields a very large computation load. The mixed integer linear programming (MILP) test of Kuosmanen (2004) does account for full diversification, but it relies on admissibility rather than optimality.

In this study, we derive a proper test for FSD optimality of a given portfolio relative to all portfolios formed from a set of choice alternatives and apply that test to analyze the U.S. stock market portfolio. In contrast to Bawa et al. (1985), our test considers all diversified portfolios in addition to the individual undiversified choice alternatives, and in contrast to Kuosmanen (2004), it relies on optimality rather than admissibility. Both features lead to a more powerful FSD test, based on a necessary and sufficient condition, than is currently available.

The new test contributes to recent methodological developments that make the SD methodology more applicable to problems in financial economics by improving the statistical power and providing more efficient computation algorithms. Our test is a natural complement to the second-order stochastic dominance (SSD) efficiency test of Post (2003). Due to concavity of utility, the analysis of SSD is generally simpler than that of FSD. First, SSD admissibility and SSD optimality are equivalent in a portfolio context, and the definition of "SSD efficiency" is less ambiguous than that of "FSD efficiency."[1] Second, SSD efficiency can be tested by simply evaluating the first-order optimality condition for all individual, undiversified choice alternatives. Third, the representative utility functions have a piecewise-linear shape, and the first-order optimality condition can be checked by searching over these functions using a single small-scale LP problem.

We apply our test to U.S. stock market data in order to analyze the FSD optimality of the market portfolio relative to portfolios formed on market capitalization and book-to-market (BM) equity ratio. This application seems relevant because a large class of capital market equilibrium models predicts that the market portfolio is FSD optimal. Surprisingly, we find that the market portfolio is significantly FSD nonoptimal. Without appealing to parametric assumptions about the return distribution, we conclude that no nonsatiable investor would hold the market portfolio in the face of the attractive premia of small caps and value stocks.

The remainder of this paper is structured as follows. Section II introduces preliminary notation, assumptions, and definitions. Next, Section III reformulates the FSD optimality criterion in terms of piecewise-constant representative utility functions, in the spirit of the representative utility functions used by Russell and Seo (1989). Section IV develops an LP test for searching over all representative utility functions in order to test portfolio optimality and suggests several approaches to identifying the input to this test. Section V uses a numerical example to illustrate our test and compare it with the two existing tests. Section VI discusses our empirical analysis of the U.S. stock market portfolio. Finally, Section VII presents concluding remarks and suggestions for further research.

[1]Theorem 1 of Post (2003) shows the equivalence using Sion's (1958) minimax theorem. Other treatments of SSD admissibility and optimality include Peleg and Yaari (1975), Dybvig and Ross (1982), and Bawa and Goroff (1982), (1983).

II. Preliminaries

Consider N choice alternatives and T scenarios with equal probability. The outcomes of the choice alternatives in the various scenarios are given by

$$X = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^T \end{pmatrix},$$

where $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_N^t)$ is the tth row of matrix X. Without loss of generality, we can assume that the columns of X are linearly independent. In addition to the individual choice alternatives, the decision-maker may also combine the choice alternatives into a portfolio. We will use $\boldsymbol{\lambda} \in R^N$ for a vector of portfolio weights, and the portfolio possibilities are given by $\Lambda = \{\boldsymbol{\lambda} \in R^N | 1'\boldsymbol{\lambda} = 1, \lambda_n \geq 0, n = 1, 2, \ldots, N\}$.[2] The evaluated portfolio is denoted by $\boldsymbol{\tau} \in \Lambda$ and is assumed to be risky.[3] Let $y^{[k]}$ be the kth smallest element among $y^1, y^2, \ldots, y^N$; that is, $y^{[1]} \leq y^{[2]} \leq \cdots \leq y^{[N]}$. Let

$$\underline{m} = \min_{t,n} x_n^t, \quad \overline{m} = \max_{t,n} x_n^t, \quad \text{and} \quad k(\boldsymbol{\tau}) = \min\left\{t : (X\boldsymbol{\tau})^{[t]} > (X\boldsymbol{\tau})^{[1]}\right\}.$$

The constants $\underline{m}$ and $\overline{m}$ are the minimum and maximum possible returns. After ordering the returns of the tested portfolio $\boldsymbol{\tau}$ from the smallest to the largest one, $k(\boldsymbol{\tau})$ determines the order of the second smallest return. Without ties, we have $k(\boldsymbol{\tau}) = 2$, but if the smallest value occurs multiple times, then $k(\boldsymbol{\tau}) > 2$.

Decision-makers obey the rules of expected utility theory. Their preferences belong to the class of weakly increasing utility functions U_1, and their decision-making problem can be represented as

$$\max_{\boldsymbol{\lambda} \in \Lambda} \sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\lambda}). \tag{1}$$

Since utility functions are unique up to the level of a positive linear transformation, without loss of generality we may focus on the following set of standardized utility functions:

$$U_1(\boldsymbol{\tau}) = \left\{u \in U_1 : u(\underline{m}) = 0\,;\ u\left((X\boldsymbol{\tau})^{[T]}\right) - u\left((X\boldsymbol{\tau})^{[k(\boldsymbol{\tau})]}\right) = 1\right\}. \tag{2}$$

[2]By using the simplex Λ, we exclude short selling. Short selling typically is difficult to implement in practice due to margin requirements and explicit or implicit restrictions on short selling for institutional investors. Still, we may generalize our analysis to include (bounded) short selling. In fact, the analysis applies to any portfolio set that takes the form of a polytope (roughly speaking, a nonempty and closed set that is defined by linear restrictions) if we replace the N choice alternatives with the set of M extreme points of the polytope.

[3]Testing optimality for a riskless portfolio is trivial, because we then only need to check if there exists some portfolio that achieves a higher minimum return than the riskless rate. If no such portfolio exists, the riskless alternative is the optimal solution for extreme risk averters and hence FSD optimal.

Note that the standardization depends on the evaluated portfolio and hence will differ when evaluating different portfolios. Furthermore, the standardization requires utility to be strictly increasing at least somewhere in the interior of the range for the evaluated portfolio. This requirement is natural, because testing optimality relative to all $u \in U_1$ is trivial. Specifically, every portfolio $\boldsymbol{\lambda} \in \Lambda$ is an optimal solution for $u_0 = I(x \geq (X\boldsymbol{\tau})^{[1]})$, that is, a two-piece constant utility function. Thus $U_1(\boldsymbol{\tau})$ is the largest subset of U_1 for which testing optimality is nontrivial.

Definition 1. Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if and only if it is the optimal solution of expression (1) for at least some utility function $u \in U_1(\boldsymbol{\tau})$; that is, there exists $u \in U_1(\boldsymbol{\tau})$ such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\lambda}) \quad \geq \quad 0, \qquad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD nonoptimal.

The intuition behind FSD optimality is that the evaluated portfolio is of potential interest to investors if it achieves a higher expected utility than all other portfolios for some increasing utility function. This concept allows for several variations. Most notably, we can choose between weakly and strictly increasing utility, and we can choose between weakly and strongly higher expected utility. Empirically, these variations are often not distinguishable. A weakly increasing utility function $u(x)$ generally is empirically indistinguishable from the strictly increasing function $u(x) + ax$ for some infinitely small value $a > 0$. Similarly, infinitely small data perturbations generally suffice to change a weak inequality to a strong one. In addition, it can be shown that requiring strictly increasing utility and strong inequality is the same as weakly increasing utility and weak inequality. This study will not try to answer the question of which type of utility function or inequality is most relevant. Rather, we will focus on accounting for all possible portfolios in an optimality test that is based on weakly increasing utility and weak inequality.

III. Representative Utility Functions

This section reformulates the optimality criterion in terms of a set of elementary representative utility functions. For pairwise FSD comparisons, Russell and Seo (1989) show that the set of three-piece linear utility functions is representative for all admissible utility functions. In our portfolio context, with diversification allowed, a class of piecewise-constant utility functions is relevant:

$$\text{(3)} \qquad R_1(\boldsymbol{\tau}) \quad = \quad \left\{ u \in U_1 | u(y) = \sum_{t=1}^{T} a_t I\left(y \geq (X\boldsymbol{\tau})^{[t]}\right), \ \mathbf{a} \in A(\boldsymbol{\tau}) \right\} \quad \text{and}$$

$$\text{(4)} \qquad A(\boldsymbol{\tau}) \quad = \quad \left\{ \mathbf{a} \in \mathbb{R}_+^T : \sum_{t=k(\boldsymbol{\tau})}^{T} a_t = 1, \quad (X\boldsymbol{\tau})^{[t]} \ = \ (X\boldsymbol{\tau})^{[s]} \right.$$

$$\wedge \ t < s \quad \Rightarrow \quad a_s = 0, \quad t,s = 1,2,\ldots,T \Big\},$$

where

$$\begin{aligned} I(y \geq y_0) &= 1, \quad \text{for } y \geq y_0 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

This class consists of at most $(T+1)$-piece constant, upper semicontinuous utility functions. This class is reminiscent of the representative utility functions used by Russell and Seo (1989) to test pairwise FSD relationship. In fact, our utility functions can be obtained as a sum of the first derivatives of the Russell and Seo (1989) representative utility functions on the relevant interval $(\underline{m}, \overline{m})$.[4] The utility functions are also reminiscent of the piecewise-linear functions used by Post (2003) to test SSD portfolio efficiency.

Theorem 1. Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if and only if it is the optimal solution of expression (1) for at least some utility function $u \in R_1(\boldsymbol{\tau})$; that is, there exists $u \in R_1(\boldsymbol{\tau})$ such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\lambda}) \quad \geq \quad 0, \qquad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD nonoptimal.

Proof. The sufficient condition follows directly from $R_1(\boldsymbol{\tau}) \subset U_1(\boldsymbol{\tau})$. To establish the necessary condition, suppose that $\boldsymbol{\tau}$ is optimal for $u(y) \in U_1(\boldsymbol{\tau})$ and let

$$u_R(y) \quad = \quad \sum_{t=1}^{T} a_t I \left(y \geq (X\boldsymbol{\tau})^{[t]} \right),$$

with $a_1 = u(X\boldsymbol{\tau})^{[1]}$, $a_t = 0$, $t = 2,\ldots,k(\boldsymbol{\tau}) - 1$, and $a_t = u(X\boldsymbol{\tau})^{[t]} - u(X\boldsymbol{\tau})^{[t-1]}$, $t = k(\boldsymbol{\tau}),\ldots,T$. By construction, $u_R(y) \in R_1(\boldsymbol{\tau})$. Furthermore, $u_R(y) \leq u(y)$, $\forall y \in \langle \underline{m}, \overline{m} \rangle$, and $u_R(y) = u(y)$, for $y = (X\boldsymbol{\tau})^{[1]}$, $(X\boldsymbol{\tau})^{[2]},\ldots,(X\boldsymbol{\tau})^{[T]}$. Therefore,

$$\sum_{t=1}^{T} u_R(\mathbf{x}^t \boldsymbol{\tau}) - \sum_{t=1}^{T} u_R(\mathbf{x}^t \boldsymbol{\lambda}) \quad \geq \quad \sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t \boldsymbol{\lambda}), \qquad \forall \boldsymbol{\lambda} \in \Lambda.$$

Since $\boldsymbol{\tau}$ is optimal for $u(y) \in U_1(\boldsymbol{\tau})$, the right-hand side (RHS) is nonnegative for all $\boldsymbol{\lambda} \in \Lambda$, and hence $\boldsymbol{\tau}$ is also optimal for $u_R(y) \in R_1(\boldsymbol{\tau})$, which completes the proof. $\square$

[4]Russell and Seo (1989) functions are continuous three-piece functions that consist of two constant pieces and one linear, increasing piece in between. Choose T such functions with increasing pieces with slopes $a_1, a_2, \ldots, a_T$ for the intervals $((X\tau)^{[1]}, (X\tau)^{[2]}), ((X\tau)^{[2]}, (X\tau)^{[3]}), \ldots, ((X\tau)^{[T-1]}, (X\tau)^{[T]})$, $((X\tau)^{[T]}, \overline{m})$. Our piecewise-constant utility function is the sum of the first derivatives on these intervals.

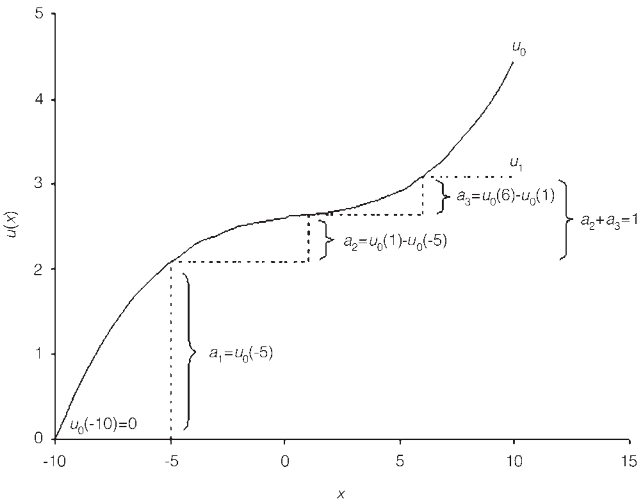
The proof makes use of the fact that any utility function can be transformed into a piecewise-constant function with increments only at $\mathbf{x}^t\boldsymbol{\tau}$, $t = 1, \ldots, T$. This transformation does not affect the expected utility for the evaluated portfolio, but it may lower the expected utility of other portfolios. Since the objective is to analyze whether the evaluated portfolio is optimal for some utility function, only the representative utility functions need to be checked; all other utility functions are known to put the evaluated portfolio in a worse perspective than any representative utility function.

To illustrate the representation theorem, consider the cubic utility function $u(y) = 10 + y - 0.1y^2 + 0.05y^3$ and a portfolio with returns $(X\boldsymbol{\tau})^{[1]} = -5$, $(X\boldsymbol{\tau})^{[2]} = 1$, and $(X\boldsymbol{\tau})^{[3]} = 6$. Figure 1 shows a version of this function that is transformed such that it belongs to $U_1(\boldsymbol{\tau})$: $u_0(y) = 2.6 + 0.04y - 0.004y^2 + 0.002y^3$ (the solid line). Since the latter function is obtained after a positive linear transformation, it yields the same results as the former function. The dashed line represents the piecewise-constant function $u_R(y) = 2.087I(y \geq -5) + 0.546I(y \geq 1) + 0.454I(y \geq 6)$. This function is constructed such that it yields exactly the same utility levels for the evaluated portfolio as $u_0(y)$ does. Furthermore, the utility levels for all other portfolios are smaller than or equal to those for $u_0(y)$. Thus, if the evaluated portfolio is optimal for $u_0(y)$, then it is also optimal for $u_R(y)$. A similar analysis applies for every admissible utility function $u(y) \in U_1(\boldsymbol{\tau})$.

FIGURE 1

Representative Utility Function

Figure 1 shows the original utility function u_0 and the associated representative utility function u_1.



Apart from replacing $U_1(\boldsymbol{\tau})$ with $R_1(\boldsymbol{\tau})$, we may also replace Λ with a reduced portfolio set that only considers portfolios with a higher minimum than the evaluated portfolio:

$$\Lambda(\boldsymbol{\tau}) \quad = \quad \left\{\boldsymbol{\lambda} \in \Lambda : (X\boldsymbol{\tau})^{[1]} \leq (X\boldsymbol{\lambda})^{[1]}\right\}.$$

Using the representative utility functions and the reduced portfolio set, we can construct the following FSD nonoptimality measure for any $\Lambda_0 \subseteq \Lambda(\boldsymbol{\tau})$:

$$\text{(5)} \qquad \xi(\boldsymbol{\tau}, \Lambda_0) \quad = \quad \frac{1}{T} \min_{u \in R_1(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda_0} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right).$$

Replacing Λ with $\Lambda(\boldsymbol{\tau})$ reduces the parameter space, but it causes no harm, because

$$\max_{\boldsymbol{\lambda} \in \Lambda} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right) \quad = \quad \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right)$$

for all $u \in R_1(\boldsymbol{\tau})$ with sufficiently large a_1, and we minimize the maximum of expected utility differences. If the evaluated portfolio has the highest minimum, then we can directly conclude that $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) = 0$; that is, the evaluated portfolio is FSD optimal (see the following corollary).

Corollary 1. i) Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))=0$. Otherwise, $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) > 0$. ii) If $\Lambda_0 \subseteq \Lambda(\boldsymbol{\tau})$, then $\xi(\boldsymbol{\tau}, \Lambda_0) \leq \xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))$.

The next section will show that $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))$ can be computed by solving an LP problem.

IV. Mathematical Programming Algorithm

There exist well-known simple algorithms for establishing FSD-dominance relationships between a pair of choice alternatives (see, e.g., Levy (1998), sect. 5.2). Bawa et al. (1985) derive an LP algorithm for FSD optimality relative to a discrete set of choice alternatives. Kuosmanen's (2004) test for FSD admissibility in a portfolio context is computationally more demanding, because we need to account for changes to the ranking of the portfolio returns as the portfolio weights change, a task that requires integer programming. A similar complication arises for testing FSD optimality in a portfolio context. This section develops an LP test for testing portfolio optimality. However, the input to the LP test may require an initial phase of MILP or subsampling.

Before presenting the algorithm, we stress that in some cases, simple necessary or sufficient conditions will suffice to classify the evaluated portfolio as FSD optimal or FSD nonoptimal. For example, a pairwise dominance relationship or a nonoptimality classification by the Bawa et al. (1985) test suffices to conclude that the portfolio is FSD nonoptimal. Similarly, if the evaluated portfolio is classified as efficient according to a mean-variance test or an SSD test, we can conclude that the portfolio is FSD optimal.

Let

$$\text{(6)} \qquad h_s(\boldsymbol{\lambda}, \boldsymbol{\tau}) \quad = \quad \sum_{t=1}^{T} I\left(\mathbf{x}^t \boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]} \right), \qquad s = 1, \ldots, T,$$

$$\text{(7)} \qquad \mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}) \quad = \quad (h_1(\boldsymbol{\lambda}, \boldsymbol{\tau}), \ldots, h_T(\boldsymbol{\lambda}, \boldsymbol{\tau})), \qquad \text{and}$$

$$\text{(8)} \qquad H(\boldsymbol{\tau}) \quad = \quad \{\mathbf{h} \in \{0, \ldots, T\}^T : \mathbf{h} = \mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}), \quad \boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})\}.$$

Since $h_s(\boldsymbol{\lambda}, \boldsymbol{\tau})$ represents the number of returns of portfolio $\boldsymbol{\lambda}$ exceeding the sth smallest return of portfolio $\boldsymbol{\tau}$, it can take at most $T+1$ values $(0, 1, \ldots, T)$ for any $s = 1, \ldots, T$. Thus, the set $H(\boldsymbol{\tau})$ has a finite number of elements. For small-scale applications, identifying all elements is a fairly trivial task. However, for large-scale applications, the task is more challenging and can become computationally demanding. Some computational strategies that identify the elements of $H(\boldsymbol{\tau})$ are discussed below. Interestingly, given $H(\boldsymbol{\tau})$, the test statistic $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))$ can be computed using simple LP problem. To see this, consider the following chain of equalities:

$$
\begin{aligned}
&\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau}))\\
&= \frac{1}{T} \min_{u \in R_1(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau})\right)\\
&= \frac{1}{T} \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{t=1}^{T} \sum_{s=1}^{T} a_s \left(I\left(\mathbf{x}^t \boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]}\right) - I\left(\mathbf{x}^t \boldsymbol{\tau} \geq (X\boldsymbol{\tau})^{[s]}\right)\right)\\
&= \frac{1}{T} \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{t=1}^{T} \sum_{s=k(\boldsymbol{\tau})}^{T} a_s \left(I\left(\mathbf{x}^t \boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]}\right) - I\left(\mathbf{x}^t \boldsymbol{\tau} \geq (X\boldsymbol{\tau})^{[s]}\right)\right)\\
&= \frac{1}{T} \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{s=k(\boldsymbol{\tau})}^{T} a_s \left(\sum_{t=1}^{T} I\left(\mathbf{x}^t \boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]}\right) - \sum_{t=1}^{T} I\left(\mathbf{x}^t \boldsymbol{\tau} \geq (X\boldsymbol{\tau})^{[s]}\right)\right)\\
&= \frac{1}{T} \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{\tau})} \sum_{s=k(\boldsymbol{\tau})}^{T} a_s (h_s(\boldsymbol{\lambda}, \boldsymbol{\tau}) - h_s(\boldsymbol{\tau}, \boldsymbol{\tau}))\\
&= \frac{1}{T} \min_{\mathbf{a} \in A(\boldsymbol{\tau}), \delta} \left\{\delta : \sum_{s=k(\boldsymbol{\tau})}^{T} a_s (\overline{h}_s - h_s(\boldsymbol{\tau}, \boldsymbol{\tau})) \leq \delta, \quad \forall \overline{\mathbf{h}} \in H(\boldsymbol{\tau})\right\}.
\end{aligned}
$$

The RHS of the final equality involves the minimization of a linear objective under a finite set of linear constraints. Thus, testing FSD optimality requires solving a simple LP problem, and Corollary 1.i) implies the following sufficient and necessary condition for FSD optimality:

Theorem 2. Let $H_0 \subseteq H(\boldsymbol{\tau})$. Let

$$
\delta^*(H_0) = \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \delta \tag{9}
$$

$$
\text{s.t.} \quad \sum_{s=k(\boldsymbol{\tau})}^{T} a_s \left(\overline{h}_s - h_s(\boldsymbol{\tau}, \boldsymbol{\tau})\right) \leq \delta, \quad \forall \overline{\mathbf{h}} \in H_0. \tag{10}
$$

Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\delta^*(H(\boldsymbol{\tau})) = 0$. If $\delta^*(H_0) > 0$ for some $H_0 \subseteq H(\boldsymbol{\tau})$, then $\boldsymbol{\tau}$ is FSD nonoptimal.

The idea of this result is to find a representative utility function for which $\boldsymbol{\tau}$ maximizes expected utility. Note that $\xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) = \delta^*/T$. Since $a \in A(\boldsymbol{\tau})$ and $\mathbf{h} \in \{0, \ldots, T\}^T$ for all $\mathbf{h} \in H(\boldsymbol{\tau})$, using Corollary 1.i), we have $0 \leq \xi(\boldsymbol{\tau}, \Lambda(\boldsymbol{\tau})) \leq 1$.

Among other things, the theorem implies the following about the relationship between the efficiency concepts of optimality and admissibility.

Corollary 2. If $(T \leq 4)$, then FSD optimality is equivalent to FSD admissibility.

Proof. Without loss of generality, let $T=4$, and let $\boldsymbol{\tau}$ be FSD admissible. Consider all possible $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ that are not dominated by each other:[5] $\mathbf{h}^1(\boldsymbol{\lambda}, \boldsymbol{\tau})=(4,2,2,2)$, $\mathbf{h}^2(\boldsymbol{\lambda}, \boldsymbol{\tau})=(4,3,3,0)$, $\mathbf{h}^3(\boldsymbol{\lambda}, \boldsymbol{\tau})=(4,4,2,0)$, and $\mathbf{h}^4(\boldsymbol{\lambda}, \boldsymbol{\tau})=(4,4,1,1)$. Entering these candidates in the LP test in Theorem 2, we can see that $\boldsymbol{\tau}$ is the optimal portfolio for a representative utility function with $a_2 = a_3 = a_4 = 1/3$, and hence $\boldsymbol{\tau}$ is FSD optimal. □

The numerical example in the next section shows that the two efficiency concepts diverge for $T \geq 5$.

A remaining problem is identifying the elements of the set $H(\boldsymbol{\tau})$. We may adopt several strategies for this task. The Appendix provides an MILP algorithm that identifies a set of candidate vectors $\widetilde{H}(\boldsymbol{\tau}) \supseteq H(\boldsymbol{\tau})$ and checks if $\mathbf{h} \in H(\boldsymbol{\tau})$ for every candidate $\mathbf{h} \in \widetilde{H}(\boldsymbol{\tau})$. A drawback of this approach is that the number of candidates increases exponentially with the number of scenarios (T). Hence, for large numbers of scenarios, this strategy may become computationally prohibitive, and some sort of approximation may then be required.

For example, we may form a sample $H_s(\boldsymbol{\tau})$ of elements $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ by using a sample $\Lambda_s \in \Lambda(\boldsymbol{\tau})$ and constructing the associated values for $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$. The test procedure is then applied to the sample $H_s(\boldsymbol{\tau})$ instead of the complete set $H(\boldsymbol{\tau})$.[6] According to Corollary 1.ii), this will lead to a necessary condition for FSD optimality. Various techniques, including a regular grid, Monte Carlo methods, or quasi-Monte Carlo methods, exist for performing the sampling task (see, e.g., Jackel (2002), Glasserman (2004)).

While the MILP algorithm starts from a large set of candidate vectors and checks feasibility for every candidate, sampling from the portfolio space avoids searching over infeasible candidates. Of course, the limitation of this strategy is that the critical sample size needed to obtain an accurate approximation increases exponentially as the number of individual choice alternatives (N) increases. Still, this approach can yield an accurate approximation in an efficient manner if N is low. This is especially true when the correlation between the individual choice alternatives is high and hence small changes in the portfolio weights do not lead to large changes in the values of $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$.

An alternative approach is to enrich the Bawa et al. (1985) test by including the same sample of diversified portfolios Λ_s as additional choice alternatives. This will lead to a more powerful necessary condition for FSD optimality than considering the undiversified choice alternatives only. However, using the sample Λ_s in our test generally leads to a more favorable trade-off between computation time and numerical accuracy.

Specifically, if we apply the Bawa et al. (1985) test to a grid with step size s, the relevant linear program has $M \cdot T$ columns and M rows (see the LP problem

[5]A dominated $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ cannot change the solution of equations (9)–(10).

[6]Since every $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ is known to be feasible, we can skip Steps 2–5 of the algorithm and take only Steps 1 and 6. Step 1 in this case boils down to performing pairwise dominance tests between every sampled portfolio and the evaluated portfolio. The computational burden of the step can be ignored.

in Bawa et al. (1985), sect. IC, p. 423), or dimensions $M \cdot T \times M$, while the dimensions of our linear program in equations (9)–(10) are $T \times M$, where

$$M = \prod_{i=1}^{N-1} \left(1 + \frac{1}{si}\right)$$

is the number of portfolios from the grid. For example, if we use $T = 120$ time-series observations, $N = 10$ base assets, and grid step size $s = 0.1$, the Bawa et al. (1985) test has dimensions $1.11 \cdot 10^7 \times 9.24 \cdot 10^4$, while our program has dimensions $120 \times 9.24 \cdot 10^4$.

V. Numerical Example

A numerical example can illustrate our test and the difference with the Bawa et al. (1985) test and Kuosmanen (2004) test. We focus on an example with five scenarios $(T = 5)$, because FSD optimality is equivalent to FSD admissibility for $(T \leq 4)$ (see Corollary 2).

Table 1 shows the returns to three choice alternatives (X_1, X_2, X_3) and the tested portfolio $Z = 0.16X_1 + 0.21X_2 + 0.63X_3$ in the five scenarios $(1, 2, 3, 4, 5)$.

TABLE 1

Example Showing that the Bawa et al. and Kuosmanen Tests Do Not Give a Sufficient Condition for FSD Optimality

Table 1 shows the returns in five scenarios to three choice individual alternatives (X_1, X_2, and X_3) and the tested portfolio $Z = 0.16X_1 + 0.21X_2 + 0.63X_3$. No convex combination of X_1, X_2, and X_3 FSD dominates Z, and hence Z is FSD admissible.

t	X_1	X_2	X_3	Z
1	−1	6	−4	−1.42
2	−2	5.90	2	2.18
3	3.50	2.20	3	2.91
4	8.70	2	5	4.96
5	10	7	7.50	7.80

One can immediately see that no individual choice alternative (X_1, X_2, and X_3) FSD dominates Z; no other alternative involves a 100% chance of a return above −2% and a 20% chance of a return above 7%. However, this does not mean that Z is an optimal portfolio. Therefore, it is interesting to employ the three efficiency tests.

To implement the Kuosmanen (2004) test, we need to solve the following LP problem for each of the $5! = 120$ permutations of Z, say $\mathbf{y}_j = (y_j^1, y_j^2, y_j^3, y_j^4, y_j^5)$, $j = 1, 2, \ldots, 120$, or an equivalent mixed integer linear problem:

$$\Psi_j = \max_{\lambda_1,\lambda_2,\lambda_3} \frac{1}{5} \sum_{t=1}^{5} \left(\lambda_1 x_1^t + \lambda_2 x_2^t + \lambda_3 x_3^t - y_j^t\right)$$

$$\text{s.t.} \quad \lambda_1 x_1^t + \lambda_2 x_2^t + \lambda_3 x_3^t \geq y_j^t, \qquad t = 1, 2, 3, 4, 5,$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \qquad \text{and}$$

$$\lambda_1, \lambda_2, \lambda_3 \geq 0.$$

We find $\Psi_j^* = 0$ for every $j = 1, 2, \ldots, 120$, and hence Z is in the FSD admissible set (not FSD dominated by any convex combination of X_1, X_2, and X_3).

To test FSD optimality according to Bawa et al. (1985), we need to establish whether some convex combination of the cumulative distribution functions (CDFs) of X_1, X_2, and X_3 dominates the CDF of Z (see Bawa et al. (1985), eq. 5, p. 421). Table 2 shows the CDFs of the three choice alternatives (Φ_{X_1}, Φ_{X_2}, Φ_{X_3}) and the CDF of Z (Φ_Z). Note that these CDFs need to be evaluated only at the observed return levels: $\{z_j\}_{j=1}^{19}$.

TABLE 2

Example Showing that the Bawa et al. and Kuosmanen Tests Do Not Give a Sufficient Condition for FSD Optimality—Continued

Table 2 shows the CDFs of the three individual choice alternatives (X_1, X_2, X_3) and the tested portfolio Z for all observed return levels. No convex combination of Φ_{X_1}, Φ_{X_2}, Φ_{X_3} dominates Φ_Z, and hence Z is classified as optimal.

j	z_j	Φ_{X_1}	Φ_{X_2}	Φ_{X_3}	Φ_Z
1	−4	0	0	1/5	0
2	−2	1/5	0	1/5	0
3	−1.42	1/5	0	1/5	1/5
4	−1	2/5	0	1/5	1/5
5	2	2/5	1/5	2/5	1/5
6	2.18	2/5	1/5	2/5	2/5
7	2.2	2/5	2/5	2/5	2/5
8	2.91	2/5	2/5	2/5	3/5
9	3	2/5	2/5	3/5	3/5
10	3.5	3/5	2/5	3/5	3/5
11	4.962	3/5	2/5	3/5	4/5
12	5	3/5	2/5	4/5	4/5
13	5.9	3/5	3/5	4/5	4/5
14	6	3/5	4/5	4/5	4/5
15	7	3/5	1	4/5	4/5
16	7.5	3/5	1	1	4/5
17	7.795	3/5	1	1	1
18	8.7	4/5	1	1	1
19	10	1	1	1	1

To implement the test, we need to solve the following LP problem (see the LP problem in Bawa et al. (1985), sect. IC, p. 423):

$$\eta = \max_{\lambda_1,\lambda_2,\lambda_3} \sum_{j=1}^{19} (\Phi_Z(z_j) - \lambda_1\Phi_{X_1}(z_j) - \lambda_2\Phi_{X_2}(z_j) - \lambda_3\Phi_{X_3}(z_j))$$

$$\begin{aligned} \text{s.t.} \quad \lambda_1\Phi_{X_1}(z_j) + \lambda_2\Phi_{X_2}(z_j) + \lambda_3\Phi_{X_3}(z_j) &\leq \Phi_Z(z_j), \qquad j = 1, \ldots, 19, \\ \lambda_1 + \lambda_2 + \lambda_3 &= 1, \qquad \text{and} \\ \lambda_1, \lambda_2, \lambda_3 &\geq 0. \end{aligned}$$

Solving this problem, we find $\eta^* = 0$, and hence Z is classified as optimal; not every nonsatiable decision-maker will prefer X_1, X_2, or X_3 to Z. Based on the positive outcomes of the two tests, we may be tempted to conclude that Z is the optimal portfolio for some increasing utility function. Perhaps surprisingly, this conclusion is wrong, as demonstrated by the application of our MILP algorithm. We will follow the steps outlined in the Appendix.

Since we have already tested FSD admissibility, we start with the second step of identifying the initial candidates for $H(\boldsymbol{\tau})$. For $j = 2, 3, 4, 5$, we solve (A-1), where $k(\boldsymbol{\tau}) = 2$, $T = 5$, $\underline{m} = -4$, $\overline{m} = 10$, and $X\boldsymbol{\tau} = Z$. (Recall that the constants $\underline{m}$ and $\overline{m}$ are the minimal and maximal possible returns, and $k(\boldsymbol{\tau})$ is the order of the second smallest return of $\boldsymbol{\tau}$.) Table 3 shows the optimal solutions for $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ and $\boldsymbol{\lambda}$. It follows that $\mathbf{h}^{\max} = (5, 5, 4, 3, 2)$. In this example, we find $\Lambda_1 = \{(0.1483, 0.8517, 0), (0.1187, 0.8813, 0), (0.9266, 0.0734, 0)\}$, and $H_1 = \{(5, 5, 4, 2, 0), (5, 5, 3, 3, 0), (5, 3, 3, 2, 2)\}$ for the set of corresponding values of $\mathbf{h}^*$.

TABLE 3

Initial Candidates

Table 3 presents the initial candidates H_1 and the associated $\Lambda_1(\boldsymbol{\tau})$ obtained in Step 2 of our algorithm.

j	h_1^*	h_2^*	h_3^*	h_4^*	h_5^*	λ_1^*	λ_2^*	λ_3^*
2	5	5	4	2	0	0.1483	0.8517	0
3	5	5	4	2	0	0.1483	0.8517	0
4	5	5	3	3	0	0.1187	0.8813	0
5	5	3	3	2	2	0.9266	0.0734	0

In the third step, we apply the stopping rules for the initial candidates. Since $\mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau}) = (5, 4, 3, 2, 1)$, $h_t^{\max} > h_t(\boldsymbol{\tau}, \boldsymbol{\tau})$ for all $t = k(\boldsymbol{\tau}), \ldots, T$, the sufficient condition of FSD optimality is not fulfilled. Since $\xi(\boldsymbol{\tau}, \Lambda_1) = 0$, the necessary condition of FSD optimality is also not fulfilled, and there exists a decision-maker who prefers $\boldsymbol{\tau}$ to all portfolios in Λ_1.

Thus, we proceed with the fourth step of constructing and reducing the candidate set $\overline{H}$. Since $\mathbf{h}^{\max} = (5, 5, 4, 3, 2)$, the candidate set consists of $6*6*5*4*3 = 2{,}160$ elements. We exclude candidates for which a corresponding portfolio cannot exist, that is, the members of the sets $\widetilde{H} = \widetilde{H}_1 \cup \widetilde{H}_2 \cup \widetilde{H}_3 \cup \widetilde{H}_4$. The remaining candidates are $\mathbf{h}_c^1 = (5, 5, 4, 1, 1)$, $\mathbf{h}_c^2 = (5, 5, 2, 2, 2)$, $\mathbf{h}_c^3 = (5, 5, 2, 2, 1)$, $\mathbf{h}_c^4 = (5, 5, 2, 1, 1)$, $\mathbf{h}_c^5 = (5, 5, 1, 1, 1)$, $\mathbf{h}_c^6 = (5, 4, 4, 1, 1)$, $\mathbf{h}_c^7 = (5, 4, 2, 2, 2)$, and $\mathbf{h}_c^8 = (5, 3, 3, 3, 1)$.

Finally, we employ the last two steps of our algorithm. Step 5 tests the feasibility of a remaining candidate using (A-2). If the candidate is infeasible, then we choose the next one. If the candidate is feasible, then we add it to H_1 and recompute $\xi(\boldsymbol{\tau}, H_1)$. Let us start with $\mathbf{h}_c^1 = (5, 5, 4, 1, 1)$. This candidate is feasible, as it corresponds to $\boldsymbol{\lambda} = (0.265, 0.735, 0)$. Adding this candidate, we consider $\Lambda_2 = \Lambda_1 \cup (0.265, 0.735, 0)$ and $H_2 = H_1 \cup (5, 5, 4, 1, 1)$. Applying Theorem 2, we solve the following linear problem:

$$
\begin{array}{llllllll}
 & & & \min \delta & & & & \\
\text{s.t.} & a_2 & +a_3 & & -a_5 & \leq & \delta \\
 & a_2 & & +a_4 & -a_5 & \leq & \delta \\
 & -a_2 & & & +a_5 & \leq & \delta \\
 & a_2 & +a_3 & -a_4 & & \leq & \delta \\
 & a_2 & +a_3 & +a_4 & +a_5 & = & 1.
\end{array}
$$

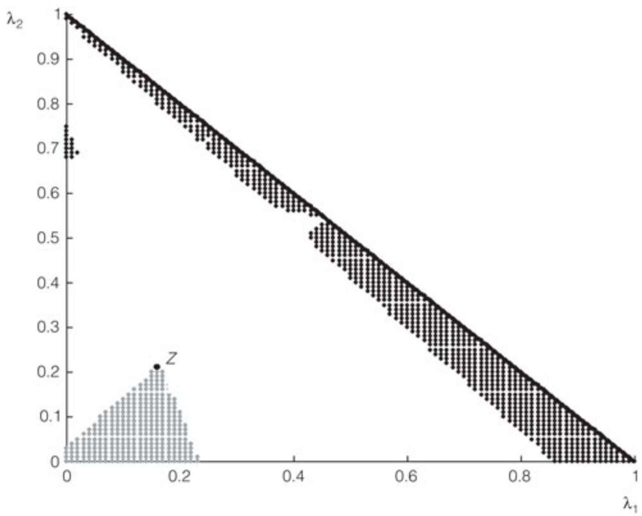
We find $\delta^* = 1/9$, $\xi(\boldsymbol{\tau}, \Lambda_2) = \delta^*/5 = 1/45 > 0$. This means that we cannot find a representative utility function that rationalizes the evaluated portfolio. Thus, adding portfolio $(0.265, 0.735, 0)$ to Λ_1 suffices to demonstrate nonoptimality in this case. Note that this portfolio does not dominate the evaluated portfolio, as the evaluated portfolio is FSD admissible. However, we do know that every well-behaved investor will prefer $(0.265, 0.735, 0)$ or an element of Λ_1 to the evaluated portfolio. Since the evaluated portfolio is classified as FSD nonoptimal, the algorithm is complete. Thus, in this example, Z is classified as optimal according to the Bawa et al. (1985) and Kuosmanen (2004) tests. Still, it can be demonstrated to be nonoptimal for any increasing utility function.

We may repeat this exercise for more portfolios $\boldsymbol{\tau} \in \Lambda \cap \{0, 0.01, \ldots, 1\}^3$, that is, when using a grid with step size 0.01 for the portfolio weights. Figure 2 illustrates the comparison between FSD admissibility and FSD optimality.

FIGURE 2

Admissibility and Optimality

Figure 2 shows the efficiency classification according to the FSD admissibility test and our FSD optimality test. We applied these tests to all portfolios $\boldsymbol{\tau} \in \Lambda \cap \{0, 0.01, \ldots, 1\}^3$, that is, when using a grid with step size 0.01 for the portfolio weights. Our optimal set is represented by the black dots. The admissible set is the union of the black dots and the grey dots.



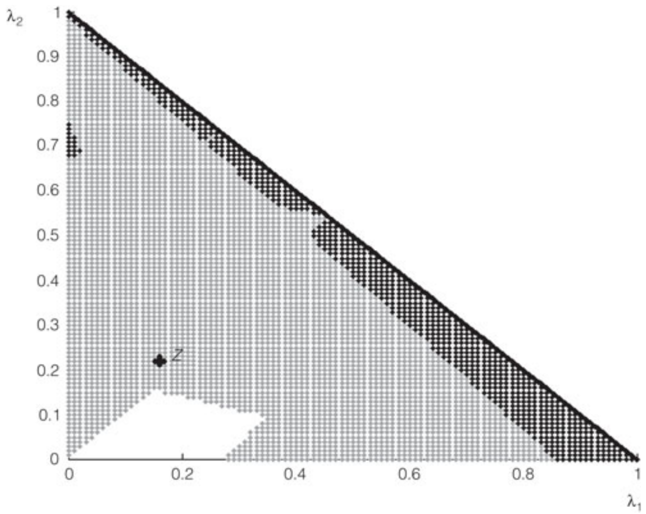
The Kuosmanen (2004) test recognizes that many diversified portfolios are FSD dominated by other diversified portfolios, most notably those that assign a high weight to X_3. In this example, only 22% of the considered portfolios are FSD admissible (the union of the grey and black dots). The FSD optimal set is even smaller than the admissible set. The set of grey dots, including Z, is now excluded, leaving only the black dots. The reduction in the efficient set to 16% of all considered portfolios (a 26% reduction) is possible because the optimality test acknowledges that a choice alternative may not be optimal for all investors even if no single other choice is preferred by all. Note that the efficient regions are not convex; witness, for example, the small isolated optimal area near $\boldsymbol{\lambda} = (0, 0.7, 0.3)$.

A similar analysis can be done for FSD optimality, according to Bawa et al. (1985). Figure 3 shows that 93% of all portfolios are classified as optimal. Only 17% of these portfolios are FSD optimal. The optimal set is substantially larger than ours, because the Bawa et al. (1985) optimality test does not account for full diversification.

FIGURE 3

Bawa et al. Optimality and FSD Optimality

Figure 3 shows the optimality classification according to the Bawa et al. (1985) test and our test for FSD optimality. Our optimal set is represented by the black dots. The Bawa et al. (1985) optimal set is the union of the black dots and the grey dots.



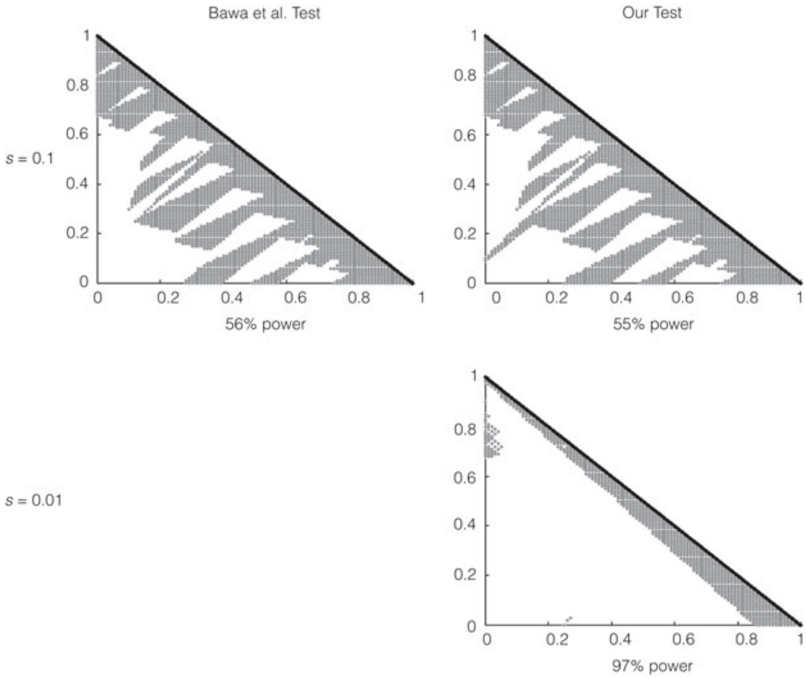
As discussed in Section III, we can increase the power of the Bawa et al. (1985) test by adding a grid of diversified portfolios to the individual choice alternatives. Of course, this approach will still yield only a necessary condition, because it is computationally impossible to include all infinitely many relevant portfolios. In addition, using the same grid of diversified portfolios in our test will lead to a smaller linear program. Figure 4 shows the set of portfolios that are not classified as FSD nonoptimal using the enriched Bawa et al. (1985) test and our test using the same grid step size.

There are only small differences in the power of the two tests for $s = 0.1$. However, our test is roughly 120 times faster than the enriched Bawa et al. (1985) test. For $s = 0.01$, our test is very powerful: 97% of nonoptimal portfolios are correctly classified as nonoptimal. Unfortunately, we were unable to implement the enriched Bawa et al. (1985) test for this step size due to the excessive computation

FIGURE 4

Subsampling Approach

Figure 4 shows the outcomes of the Bawa et al. (1985) test and our test when applied to a grid of portfolios with step size $s = 0.1$ or $s = 0.01$. The grey dots are portfolios that passed the necessary test; the other portfolios failed the test and are classified as FSD nonoptimal. The percentages of FSD nonoptimal portfolios that are detected using the necessary tests are given below each graph.



load. The differences in computation load will be even larger for real-life applications with higher dimensions.

VI. Empirical Application

To further illustrate our test, we apply it to U.S. stock market data in order to analyze FSD optimality of the market portfolio relative to portfolios formed on market capitalization of equity (size) and BM equity ratio. This test seems relevant for asset pricing theory, because all single-period portfolio-oriented representative-investor models of capital market equilibrium predict that the market portfolio is optimal for a representative investor with well-behaved preferences.

The investment universe of stocks is proxied by the well-known six value-weighted Fama and French portfolios constructed as the intersection of two groups formed on size (small caps and large caps) and three groups formed on BM (growth, neutral, and value stocks). We proxy the market portfolio by the Center for Research in Security Prices (CRSP) all-share index, a value-weighted average of common stocks listed on NYSE, AMEX, and NASDAQ, and the riskless asset by the one-year U.S. government bond index from Ibbotson Associates. We consider yearly (January–December) excess returns from 1963 to 2002 (40 annual observations).[7,8] Excess returns are computed by subtracting the riskless rate from the nominal returns; that is, the riskless asset always has a return of 0.

Table 4 shows some descriptive statistics for our data set. Particularly puzzling is the value premium in the small cap segment. The small value stocks earned an average annual excess return of 13.86%, 8.55 percentage points in excess of the 5.31% for small growth stocks. It seems difficult to explain away this premium with risk because the small growth stocks actually have a higher standard deviation than the small value stocks. Indeed, the market portfolio is SSD inefficient, as shown by Post (2003). This means that in the face of attractive premiums from investing in small cap stocks and value stocks, investing in the market portfolio seems nonoptimal for any risk-averse investor.

TABLE 4

Descriptive Statistics

Table 4 shows descriptive statistics for the annual (January–December) excess returns of the six Fama and French stock portfolios formed on market capitalization of equity and book-to-market equity ratio (SG = small growth, SN = small neutral, SV = small value, BG = big growth, BN = big neutral, and BV = big value) and the CRSP all-equity index (CRSP). Excess returns are computed by subtracting the return to the one-year U.S. government bond from the nominal returns. The sample period is from 1963 to 2002 (40 annual observations). Equity data are from Kenneth French and bond data are from Ibbotson Associates.

	Mean	Std. Dev.	Skew.	Kurt.	Min.	Max.
SG	5.309	28.520	0.323	0.175	−49.28	83.68
SN	11.301	22.728	−0.308	0.062	−37.38	65.48
SV	13.861	23.158	−0.373	−0.222	−33.86	61.14
BG	5.303	18.820	−0.317	−0.537	−40.49	34.67
BN	6.340	16.120	−0.241	−0.090	−34.13	34.73
BV	8.946	17.723	−0.690	−0.026	−34.24	40.34
CRSP	5.536	17.191	−0.602	−0.404	−39.19	31.89

Still, the market portfolio may be FSD optimal; for example, it may be optimal for investors who are risk-seeking for losses and risk-averse for gains. Our first step in analyzing FSD optimality is to apply the Bawa et al. (1985) test. This test classifies the market portfolio as optimal, meaning that some investors prefer the market portfolio to all seven benchmark portfolios (six Fama and French and the riskless asset). However, as discussed before, the test does not account for diversification between the seven portfolios. To analyze the effect of diversification, we can enrich the Bawa et al. (1985) test by adding diversified portfolios, or we can apply the Kuosmanen (2004) test. Using the grid $\Lambda_g = \Lambda(\boldsymbol{\tau}) \cap \{0, 0.1, \ldots, 1\}^7$, the enriched Bawa et al. (1985) test already leads to a

[7] As discussed in Benartzi and Thaler ((1995), p. 83), one year is a plausible choice for the investor's evaluation period, because "individual investors file taxes annually, receive their most comprehensive reports from their brokers, mutual funds, and retirement accounts once a year, and institutional investors also take the annual reports most seriously." Excess returns are computed by subtracting the riskless rate from the nominal returns.

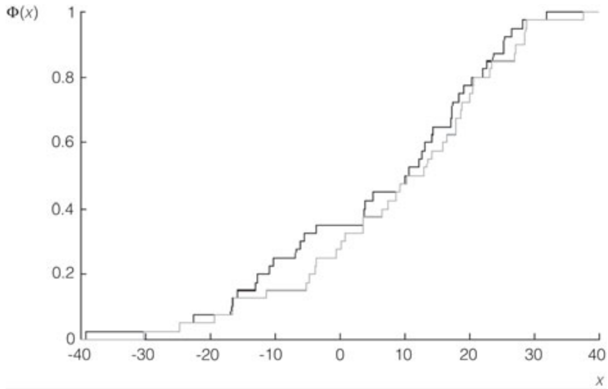
[8] There are two reasons for starting in 1963 and omitting the pre-1963 data. First, prior to 1963, the Compustat database is affected by survivorship bias caused by the backfilling procedure excluding delisted firms, which typically are less successful (Kothari, Shanken, and Sloan (1995)). Further, from June 1962, AMEX-listed stocks are added to the CRSP database, which includes only NYSE-listed stocks before this month. Since AMEX stocks generally are smaller than NYSE stocks, the relative number of small caps in the analysis increases from June 1962. Since the value effect is most pronounced in the small-cap segment, the post-1962 data set is most challenging.

linear program with more than 320,000 constrains and 8,000 variables. We therefore apply the Kuosmanen (2004) test, which involves solving a mixed-integer program with 1,600 integer variables. Interestingly, this test classifies the market portfolio as FSD inadmissible and identifies the dominating portfolio shown in Figure 5.

FIGURE 5

Pairwise FSD Dominance

Figure 5 shows the CDF of the stock market portfolio (black line) and the dominating portfolio (grey line): $\boldsymbol{\lambda}_d$ = (0, 0.04, 0.43, 0.37, 0.04, 0, 0.13). Since the dominating portfolio is preferred by all investors, the market portfolio is FSD inadmissible.



Since FSD inadmissibility implies FSD nonoptimality, there is no need to apply our test in this case. Still, it is useful to apply our test for the purpose of illustration and comparison of the complexity of these three tests.

Since the number of choice alternatives (7) is small in comparison to the number of scenarios (40), we apply the method of sampling portfolios using the grid: $\Lambda_g = \Lambda(\boldsymbol{\tau}) \cap \{0, 0.1, \ldots, 1\}^7$. The associating vectors $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ are collected in H_g, and H_g is used to proxy for $H(\boldsymbol{\tau})$ in the LP problem in equations (9)–(10). This linear program has only 8,000 constrains and 40 variables. Therefore, our test is much faster than both the Kuosmanen (2004) and the enriched Bawa et al. (1985) tests for the same grid. Interestingly, the nonoptimality measure is strictly positive; $\xi(\boldsymbol{\tau}, \Lambda_g) = 0.00275$. According to Corollary 1.ii), this implies that the market portfolio is not optimal for any increasing utility function.

Table 5 illustrates the nonoptimality classification. It shows nine combinations of the seven benchmark portfolios. For the vectors $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ associated with these combinations, the restrictions in equations (9)–(10) are binding. This means that the value of the nonoptimality measure critically depends on these vectors. By contrast, the other vectors can be excluded without affecting the nonoptimality measure. None of these nine combinations FSD dominates the evaluated portfolio. Still, for every increasing utility function, at least one of these combinations is better than the market portfolio. Not surprisingly, each of these portfolios assigns a substantial weight to small cap stocks and/or value stocks.

The above analysis focuses on sample optimality. It is desirable to account for sampling error and establish the statistical confidence we have in population

TABLE 5

Nine Combinations Showing FSD Nonoptimality of the Market Portfolio

Table 5 shows the portfolio weights of nine combinations of the six Fama and French stock portfolios formed on size and BM (SG = small growth, SN = small neutral, SV = small value, BG = big growth, BN = big neutral, and BV = big value) and the riskless alternative (RL). For every increasing utility function, at least one of these nine combinations is preferred to the market portfolio, and hence the market portfolio is FSD nonoptimal.

Combination	SG	SN	SV	BG	BN	BV	RL
1	0	0	0.1	0.3	0.1	0.4	0.1
2	0	0	0.3	0.2	0	0.2	0.3
3	0	0	0.4	0	0	0.3	0.3
4	0	0	0.4	0	0.1	0.3	0.2
5	0	0	0.6	0	0.1	0	0.3
6	0	0	0.6	0.2	0	0.1	0.1
7	0	0.1	0.5	0.1	0.1	0.1	0.1
8	0	0.1	0.9	0	0	0	0
9	0	0.2	0.8	0	0	0	0

optimality. For mean-variance efficiency tests, the sampling distribution is well-known (see, e.g., Gibbons et al. (1989)). The sampling distribution for SD tests is more difficult to derive because the shape of the population return distribution is not restricted. We therefore resort to the bootstrap method, a well-established tool to analyze the sensitivity of empirical estimators to sampling variation in situations where the sampling distribution is difficult to obtain analytically.

Under the assumption of serially i.i.d. returns, the empirical return distribution is a consistent estimator of the population return distribution, and bootstrapping samples can be obtained simply by randomly sampling with replacements from the empirical return distribution. Nelson and Pope (1991) demonstrate in a convincing way that this approach can quantify the sensitivity of the empirical return distribution to sampling variation, and that SD analysis based on the bootstrapped return distribution is more powerful than analysis based on the original empirical return distribution. We implement this method by generating 10,000 random pseudo-samples and apply our tests for each pseudo-sample. We do not apply the enriched Bawa et al. (1985) test or the Kuosmanen (2004) test, because of the associated computational burden. Rather, we apply our LP necessary test in equations (9)–(10) using the nine combinations from Table 5. In 97.9% of the pseudo-samples, the market portfolio did not pass this necessary test. Then, for the remaining 2.1% of the pseudo-samples, we apply our LP necessary test in equations (9)–(10) using the grid $\Lambda_g = \Lambda(\boldsymbol{\tau}) \cap \{0, 0.1, \ldots, 1\}^7$. In 0.8% of the pseudo-samples, the market portfolio failed this necessary test. For the remaining 1.3% of the pseudo-samples, we applied our necessary and sufficient test. The market portfolio was classified as FSD optimal in all of these pseudo-samples. Thus, the bootstrap p-value is 1.3%, and the market portfolio can be classified as significantly FSD nonoptimal with 98.7% confidence.

The classification of the market portfolio as FSD nonoptimal reinforces Post's (2003) finding that the market portfolio is SSD inefficient. This finding is potentially important for asset pricing theory. All single-period, portfolio-oriented, representative-investor models predict FSD optimality. FSD nonoptimality would contradict all these models and may call for multiperiod models, consumption-oriented models or heterogeneous-investor models. However, we stress that this application only serves to illustrate our nonoptimality test. Among other things, the choice of the benchmark portfolios and market portfolio, investment horizon, and sample period requires more analysis than is possible here.

VII. Conclusions

We have developed a test for "FSD efficiency" of a given portfolio that is more powerful than those currently available. In contrast to Bawa et al. (1985), our test compares the evaluated portfolio not only with the finite set of individual choice alternatives, but also with all portfolios formed by combining the individual alternatives. In contrast to Kuosmanen (2004), our efficiency test is based on the criterion of FSD optimality rather than the weaker criterion of FSD admissibility.

The test can be performed by solving a simple linear programming (LP) problem. However, the input to the LP problem may require an initial phase of mixed integer linear programming (MILP). For large numbers of scenarios, this strategy may become computationally prohibitive, and we may have to resort to an approximation based on sampling portfolios from the portfolio possibilities set. This subsampling approach improves the trade-off between computational complexity and numerical accuracy compared with enriching the Bawa et al. (1985) test with diversified choice alternatives.

Using our new test, we show that the U.S. stock market portfolio is significantly FSD nonoptimal relative to benchmark portfolios formed on market capitalization and book-to-market equity ratio; no nonsatiable investor would hold the market portfolio in the face of the small cap premium and the value stock premium. FSD nonoptimality would contradict all single-period, portfolio-oriented, and representative-investor models of capital market equilibrium and would call for multiperiod, consumption-oriented, or heterogeneous-investor models. The focus of our study is, however, on methodology; a rejection of market portfolio optimality requires a more rigorous empirical analysis than is possible in this study.

Appendix

This appendix provides an MILP algorithm for identifying the elements of $H(\boldsymbol{\tau})$ and suggests some stopping rules for testing FSD optimality.

Step 1. Perform an FSD admissibility test.

As an initial stopping rule, test FSD admissibility of $\boldsymbol{\tau}$, for example, using the MILP test of Kuosmanen (2004). If $\boldsymbol{\tau}$ is FSD inadmissible, then stop the algorithm; $\boldsymbol{\tau}$ is FSD nonoptimal.

Step 2. Identify initial candidates for $H(\boldsymbol{\tau})$.

For all $j = k(\boldsymbol{\tau}), \ldots, T$, solve the following MILP problem:

$$\text{(A-1)} \qquad \max\ h_j + \frac{1}{T^2}\sum_{t=k(\boldsymbol{\tau})}^{T} h_t$$

$$\text{s.t.} \quad (v_{s,t}-1)(\overline{m}-\underline{m}) \quad \leq \quad \mathbf{x}^s\boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} \quad \leq \quad v_{s,t}(\overline{m}-\underline{m}),$$
$$s = 1, \ldots, T, \quad t = k(\boldsymbol{\tau}), \ldots, T\,;$$

$$\begin{aligned} h_t &= \sum_{s=1}^{T} v_{s,t}, \\ &\qquad t = k(\boldsymbol{\tau}), \ldots, T\,; \\ v_{s,t} &\in \{0,1\}, \\ &\qquad s = 1, \ldots, T, \quad t = k(\boldsymbol{\tau}), \ldots, T\,; \\ \boldsymbol{\lambda} &\in \Lambda(\boldsymbol{\tau}). \end{aligned}$$

The problem is solved only for $j \geq k(\boldsymbol{\tau})$; solving it for $j < k(\boldsymbol{\tau})$ will identify no new candidates, because the optimal solution of (A-1) for any $j < k(\boldsymbol{\tau})$ is equal to that for $j = k(\boldsymbol{\tau})$.

Use $(h_t^{*j}, \lambda_t^{*j}, v_{s,t}^{*j})$ for the optimal solution of this problem. Let $\Lambda_1 \in \Lambda(\boldsymbol{\tau})$ be a set of pairwise different $\boldsymbol{\lambda}^{*j}$ (all redundancies are removed). Set

$$h_t^{\max} = \max_j h_t^{*j} \qquad \text{and} \qquad H_1 = \{\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}) :\ \boldsymbol{\lambda} \in \Lambda_1\}.$$

Step 3. Apply stopping rules.

Consider $\mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau})$ as defined by equations (6)–(7). If there exists $t \in \{k(\boldsymbol{\tau}), \ldots, T\}$ such that $h_t^{\max} \leq h_t(\boldsymbol{\tau}, \boldsymbol{\tau})$, then stop the algorithm; $\boldsymbol{\tau}$ is FSD optimal. Otherwise, solve the problem in equations (9)–(10) for $H_0 = H_1$. If $\delta^*(H_1) > 0$, then stop the algorithm; $\boldsymbol{\tau}$ is FSD nonoptimal.

Step 4. Construct and reduce the candidate set $\overline{H}$.

Let $\overline{H}_t = \{0, 1, \ldots, h_t^{\max}\}$. Use $\overline{H}$ for the Cartesian product $\overline{H} = \bigotimes_{k(\boldsymbol{\tau})}^{T} \overline{H}_t$. Clearly, $H(\boldsymbol{\tau}) \subseteq \overline{H}$, and hence $\overline{H}$ is a candidate set. Exclude the candidates $\widetilde{H} = \widetilde{H}_1 \cup \widetilde{H}_2 \cup \widetilde{H}_3 \cup \widetilde{H}_4$, where

$$\begin{aligned} \widetilde{H}_1 &= \left\{\mathbf{h} \in \overline{H} | h_{t_1} < h_{t_2}, \qquad \text{for some } \ t_1 < t_2\right\}, \\ \widetilde{H}_2 &= \left\{\mathbf{h} \in \overline{H} | h_t \geq h_t(\boldsymbol{\tau}, \boldsymbol{\tau}), \quad \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\}\right\}, \\ \widetilde{H}_3 &= \Big\{\mathbf{h} \in \overline{H} | \exists \widehat{\mathbf{h}} \in H_1 : h_t \geq \widehat{h}_t, \quad \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\} \\ &\qquad\qquad \text{with at least one strict inequality}\Big\}, \qquad \text{and} \\ \widetilde{H}_4 &= \Bigg\{\mathbf{h} \in \overline{H} | \exists \mathbf{b} = (b_0, b_{k(\boldsymbol{\tau})}, \ldots, b_T) : h_t \leq \sum_{j=k(\boldsymbol{\tau})}^{T} b_j h_t^{*j} + b_0 h_t(\boldsymbol{\tau}, \boldsymbol{\tau}), \\ &\qquad b_0 + \sum_{l=k(\boldsymbol{\tau})}^{T} b_l = 1, \quad \mathbf{b} \geq 0, \ \ \mathbf{h}^{*j} \in H_1, \quad \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\}\Bigg\}. \end{aligned}$$

The elements of $\widetilde{H}_1 \cup \widetilde{H}_2 \cup \widetilde{H}_3$ are not feasible; that is, there exist no corresponding portfolios. The elements of $\widetilde{H}_1$ contradict the definition of vector $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ (see equations (6)–(7)). In Step 1, we have found that $\boldsymbol{\tau}$ is FSD admissible. Feasibility of an element of $\widetilde{H}_2$ implies FSD inadmissibility of $\boldsymbol{\tau}$. Every element of $\widetilde{H}_3$ gives a value of the objective function in (A-1) that is strictly higher than at least one initial candidate. Thus, it cannot be a feasible candidate. Adding the elements of $\widetilde{H}_4$ to H_1 does not affect the solution of equations (9)–(10).

Set $p = 1$.

Step 5. Check the feasibility of the remaining candidates.

If $\overline{H} \setminus \widetilde{H}$ is empty, that is, all possible $\mathbf{h} \in \overline{H}$ have been considered, then stop the algorithm; portfolio $\boldsymbol{\tau}$ is FSD optimal. Otherwise, choose $\mathbf{h} \in \overline{H} \setminus \widetilde{H}$ and add it to $\widetilde{H}$. Let $p = p + 1$, $H_p = H_{p-1} \cup \mathbf{h}$ and go to the next step if there exists a feasible solution of the system:

$$
\text{(A-2)} \qquad \begin{aligned}
(v_{s,t} - 1)(\overline{m} - \underline{m}) \quad &\leq \quad \mathbf{x}^s \boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} \quad \leq \quad v_{s,t}(\overline{m} - \underline{m}) \\
& \qquad\qquad s = 1, \ldots, T, \quad t = t_1, \ldots, T\,; \\
h_t \quad &= \quad \sum_{s=1}^{T} v_{s,t} \\
& \qquad\qquad t = t_1, \ldots, T\,; \\
v_{s,t} \quad &\in \quad \{0, 1\} \\
& \qquad\qquad s = 1, \ldots, T, \quad t = t_1, \ldots, T\,; \\
\boldsymbol{\lambda} \quad &\in \quad \Lambda(\boldsymbol{\tau}).
\end{aligned}
$$

Otherwise, repeat this step.

Step 6. Test optimality using the feasible candidates.

Solve the problem in equations (9)–(10) for $H_0 = H_p$. If $\delta^*(H_p) > 0$, then stop the algorithm; $\boldsymbol{\tau}$ is FSD nonoptimal. Otherwise, go to Step 5.

References

Bawa, V. S.; J. N. Bodurtha, Jr.; M. R. Rao; and H. L. Suri. "On Determination of Stochastic Dominance Optimal Sets." *Journal of Finance*, 40 (1985), 417–431.

Bawa, V. S., and D. L. Goroff. "Admissible, Best and Efficient Choices under Uncertainty." Working Paper, University of Texas (1982).

Bawa, V. S., and D. L. Goroff. "Stochastic Dominance, Efficiency and Separation in Financial Markets." *Journal of Economic Theory*, 30 (1983), 410–414.

Benartzi, S., and R. H. Thaler. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, 110 (1995), 73–92.

Dybvig P. H., and S. Ross. "Portfolio Efficient Sets." *Econometrica*, 50 (1982), 1525–1546.

Gibbons M. R.; S. A. Ross; and J. Shanken. "A Test of the Efficiency of a Given Portfolio." *Econometrica*, 57 (1989), 1121–1152.

Glasserman, P. *Monte Carlo Methods in Financial Engineering.* New York, NY: Springer Verlag (2004).

Jackel, P. *Monte Carlo Methods in Finance.* Chichester, UK: John Wiley & Sons (2002).

Kothari, S.; J. Shanken; and R. Sloan. "Another Look at the Cross-Section of Expected Returns." *Journal of Finance*, 50 (1995), 185–224.

Kuosmanen, T. "Efficient Diversification According to Stochastic Dominance Criteria." *Management Science*, 50 (2004), 1390–1406.

Levy, H. *Stochastic Dominance.* Norwell, MA: Kluwer Academic Publishers (1998).

Nelson, R. D., and R. Pope. "Bootstrapped Insights into Empirical Applications of Stochastic Dominance." *Management Science*, 37 (1991), 1182–1194.

Peleg, B., and M. Yaari. "A Price Characterization of Efficient Random Variables." *Econometrica*, 43 (1975), 283–292.

Post, T. "Empirical Tests for Stochastic Dominance Efficiency." *Journal of Finance*, 58 (2003), 1905–1931.

Russell, W. R., and T. K. Seo. "Representative Sets for Stochastic Dominance Rules." In *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, T. B. Fomby and T. K. Seo, eds. New York, NY: Springer Verlag (1989), 59–76.

Sion M. "On General Minimax Theorems." *Pacific Journal of Mathematics*, 8 (1958), 171–176.

Kybernetika

VOLUME *44* (2008), NUMBER 2

The Journal of the Czech Society for Cybernetics and Information Sciences

Published by:

Institute of Information Theory and Automation of the AS CR

Editorial Office:

Pod Vodárenskou věží 4, 182 08 Praha 8

Kybernetika is a bi-monthly international journal dedicated for rapid publication of high-quality, peer-reviewed research articles in fields covered by its title.

Kybernetika traditionally publishes research results in the fields of Control Sciences, Information Sciences, System Sciences, Statistical Decision Making, Applied Probability Theory, Random Processes, Fuzziness and Uncertainty Theories, Operations Research and Theoretical Computer Science, as well as in the topics closely related to the above fields.

The Journal has been monitored in the Science Citation Index since 1977 and it is abstracted/indexed in databases of Mathematical Reviews, Zentralblatt für Mathematik, Current Mathematical Publications, Current Contents ISI Engineering and Computing Technology.

K y b e r n e t i k a . Volume *44* (2008) ISSN 0023-5954, MK ČR E 4902.

Published bimonthly by the Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. — Address of the Editor: P. O. Box 18, 182 08 Prague 8, e-mail: kybernetika@utia.cas.cz. — Printed by PV Press, Pod vrstevnicí 5, 140 00 Prague 4. — Orders and subscriptions should be placed with: MYRIS TRADE Ltd., P. O. Box 2, V Štíhlách 1311, 142 01 Prague 4, Czech Republic, e-mail: myris@myris.cz. — Sole agent for all “western” countries: Kubon & Sagner, P. O. Box 34 01 08, D-8 000 München 34, F.R.G.

Published in May 2008.

KYBERNETIKA — VOLUME 44 (2008), NUMBER 2, PAGES 243 – 258

A SECOND–ORDER STOCHASTIC DOMINANCE PORTFOLIO EFFICIENCY MEASURE

Miloš Kopa and Petr Chovanec

In this paper, we introduce a new linear programming second-order stochastic dominance (SSD) portfolio efficiency test for portfolios with scenario approach for distribution of outcomes and a new SSD portfolio inefficiency measure. The test utilizes the relationship between CVaR and dual second-order stochastic dominance, and contrary to tests in Post [14] and Kuosmanen [7], our test detects a dominating portfolio which is SSD efficient. We derive also a necessary condition for SSD efficiency using convexity property of CVaR to speed up the computation. The efficiency measure represents a distance between the tested portfolio and its least risky dominating SSD efficient portfolio. We show that this measure is consistent with the second-order stochastic dominance relation. We find out that this measure is convex and we use this result to describe the set of SSD efficient portfolios. Finally, we illustrate our results on a numerical example.

Keywords: stochastic dominance, CVaR, SSD portfolio efficiency measure

AMS Subject Classification: 91B28, 91B30

1. INTRODUCTION

The questions how to maximize profit and how to diversify risk has been around for centuries; however, both these questions took another dimension with the work of Markowitz [10]. In his work, Markowitz identified two main components of portfolio performance, mean reward and risk represented by variance, and by applying a simple parametric optimization model found the optimal trade-off between these two components. Unfortunately, these optimal portfolios are not consistent with expected utility maximization unless the utility is quadratic or returns are normally distributed; because of this Markowitz [11] suggested as more plausible the semivariance instead of variance. Decades later Ogryczak and Ruszczyński [12] proved that the optimal mean-semivariance portfolio is also optimal in second-order stochastic dominance sense and vice-versa.

Stochastic dominance is another possible approach to portfolio selection. In economics and finance it was introduced independently in Hadar and Russel [4], Hanoch and Levy [5], Rothschild and Stiglitz [15] and Whitmore [19].[1]

[1] for more information see Levy [8] or Levy [9].

The usual definition of stochastic dominance uses cumulative distribution function[2], but the following alternative definition fits our questions better – it has nice financial consequences with risk-averse agents, and is easier to understand in our context. We say that risky asset X stochastically dominates (in the first-order) risky asset Y, if and only if $\mathbb{E}\,u(X) \geq \mathbb{E}\,u(Y)$ for every utility function (i. e. for every non-decreasing function). If this holds only for the concave utility function (for every risk-averter), we say that X stochastically dominates Y in the second-order. This can be applied in the portfolio selection problem as a search for a portfolio that no risk-averse agent would want to choose.

In this paper, we propose a new test of second-order stochastic dominance of a portfolio relative to all portfolios created from a set of assets with discrete distributions. Until 2003, stochastic dominance tests considered only pairs of assets and not the sets of assets; however, especially in finance we would like to know whether our portfolio is the best one, or whether for any risk-averse agent there exists another better portfolio. Therefore, a test for stochastic dominance efficiency was needed. In 2003, Post [14] published a linear programming procedure for testing the second-order stochastic dominance of a given portfolio relative to a given set of assets and he discussed its statistical properties. Post used a primal approach and a representative characterization of concave utility functions. Therefore, his algorithm does not identify the SSD efficient portfolio. On the other hand, linear programming algorithm works in linear space in both numbers of assets and scenarios. Our approach is thus slower, but it identifies the dominating SSD efficient portfolio. In the same year, Ruszczyński and Vanderbei [17] developed a parametric linear programming procedure for computing all efficient portfolios in the dual mean risk space (in the second-order stochastic dominance sense). They used dual approach, the same as we did, but they used another identity. Our test procedure should generate more sparse matrix and, therefore, should be quicker. Furthermore, another linear programming test for second-order stochastic dominance was presented in Kuosmanen [7]. This test is based on comparisons of cumulated returns. It identifies a dominating portfolio but this dominating portfolio need not to be SSD efficient. Moreover, the Kuosmanen test is computationally more demanding than our test.

Our approach is based on second-order stochastic dominance consistence with Conditional Value-at-Risk (shown in Ogryczak and Ruszczyński [12]), and because CVaR has a linear programming representation explored by Uryasev and Rockafellar [18], it is sufficient to solve a linear program. Another connection of risk measures and SSD relation was analyzed in [1] or [2]. We derive a LP sufficient and necessary condition for SSD efficiency. Moreover, using convexity of CVaR, a necessary condition is presented. In addition, our test identifies the dominating portfolio which is already SSD efficient. With the help of this test, we introduce a SSD portfolio inefficiency measure in the dual risk (CVaR) space. Our measure is consistent with the second-order stochastic dominance relation and it is represented by a distance between the tested portfolio and its dominating SSD efficient portfolio. If there exist more dominating SSD efficient portfolios then the least risky portfolio is considered. Since the set of SSD efficient portfolios can be non-convex, see Dybvig and Ross [3],

[2]and can be found in e. g. Levy [8] or Levy [9]

we explore the convexity of this measure. We prove that all portfolios dominated by a given portfolio form a convex set and the measure is convex on these sets.

The rest of the paper is organized as follows. A Preliminaries section with precise assumptions and definitions for a stochastic dominance relation is followed by a section dealing with CVaR. In Section 4, we state our main theorems, allowing us to test the SSD efficiency of a given portfolio and to identify a dominating portfolio which is already SSD efficient. Subsequent section defines the inefficiency measure in dual risk (CVaR) space and presents the convexity results. This section is followed by the numerical illustration in Section 6.

2. PRELIMINARIES

For two random variables X_1 and X_2 with respective cumulative probability distributions functions $F_1(x)$, $F_2(x)$ we say that X_1 *dominates* X_2 *by second-order stochastic dominance*: $X_1 \succeq_{\mathrm{SSD}} X_2$ if

$$\mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0$$

for every $u \in U_2$ where U_2 denotes the set of all concave utility functions such that these expected values exist. The corresponding strict dominance relation $\succ_{\mathrm{SSD}}$ is defined in the usual way: $X_1 \succ_{\mathrm{SSD}} X_2$ if and only if $X_1 \succeq_{\mathrm{SSD}} X_2$ and $X_2 \not\succeq_{\mathrm{SSD}} X_1$. According to Russel and Seo [16], $u \in U_2$ may be represented by simple utility functions in the following sense:

$$\mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0 \quad \forall\, u \in U_2 \Longleftrightarrow \mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0 \quad \forall\, u \in V$$

where $V = \{u_\eta(x) : \eta \in \mathbb{R}\}$ and $u_\eta(x) = \min\{x - \eta, 0\}$.

Set

$$F_i^{(2)}(t) = \int_{-\infty}^{t} F_i(x)\,\mathrm{d}x \qquad i = 1, 2.$$

The following necessary and sufficient conditions for the second-order stochastic dominance relation were proved in Hanoch and Levy [5].

Lemma 1. Let $F_1(x)$ and $F_2(x)$ be the cumulative distribution functions of X_1 and X_2. Then

- $X_1 \succeq_{\mathrm{SSD}} X_2 \Longleftrightarrow F_1^{(2)}(t) \leq F_2^{(2)}(t)\ \forall\, t \in \mathbb{R}$
- $X_1 \succ_{\mathrm{SSD}} X_2 \Longleftrightarrow F_1^{(2)}(t) \leq F_2^{(2)}(t)\ \forall\, t \in \mathbb{R}$ with at least one strict inequality.

Lemma 1 can be used as an alternative definition of the second-order stochastic dominance relation.

Consider now the quantile model of stochastic dominance as in Ogryczak and Ruszczyński [12]. The *first quantile function* $F_X^{(-1)}$ corresponding to a real random variable X is defined as the left continuous inverse of its cumulative probability distribution function F_X:

$$F_X^{(-1)}(v) = \min\{u : F_X(u) \geq v\}. \tag{1}$$

The *second quantile function* $F_X^{(-2)}$ is defined as

$$\begin{aligned} F_X^{(-2)}(p) &= \int_{-\infty}^{p} F_X^{(-1)}(t)\,\mathrm{d}t && \text{for } 0 < p \leq 1 \\ &= 0 && \text{for } p = 0 \\ &= +\infty && \text{otherwise.} \end{aligned}$$

The function $F_X^{(-2)}$ is convex and it is well defined for any random variable X satisfying the condition $\mathbb{E}\,|X| < \infty$. An interpretation of this function will be given in Section 3.

Lemma 2. For every random variable X with $\mathbb{E}\,|X| < \infty$ we have:

(i) $$F_X^{(-2)}(p) = \sup_{\nu}\{\nu p - \mathbb{E}\,\max(\nu - X, 0)\}$$

(ii) $$X_1 \succeq_{\mathrm{SSD}} X_2 \Longleftrightarrow \frac{F_1^{(-2)}(p)}{p} \geq \frac{F_2^{(-2)}(p)}{p} \qquad \forall\, p \in \langle 0, 1\rangle.$$

These properties follow from the Fenchel duality relation between $F_X^{(2)}$ and $F_X^{(-2)}$. For the entire proof of Lemma 2 and more details about dual stochastic dominance see Ogryczak and Ruszczyński [12].

3. CVaR FOR SCENARIO APPROACH

Let Y be a random loss variable corresponding to the return described by random variable X, i.e. $Y = -X$. We assume that $\mathbb{E}\,|Y| < \infty$. For a fixed level α, the *value-at-risk* (VaR) is defined as the α-quantile of the cumulative distribution function F_Y:

$$\mathrm{VaR}_\alpha(Y) = F_Y^{(-1)}(\alpha). \tag{2}$$

We follow Pflug [13] in defining *conditional value-at-risk* (CVaR) as the solution of the optimization problem

$$\mathrm{CVaR}_\alpha(Y) = \min_{a \in \mathbb{R}} \left\{ a + \frac{1}{1-\alpha}\mathbb{E}\,[Y - a]^+ \right\} \tag{3}$$

where $[x]^+ = \max(x, 0)$. This problem has always a solution and one of the minimizers is $\mathrm{VaR}_\alpha(Y)$, see Pflug [13] for the proof and more details. It was shown in

Uryasev and Rockafellar [18] that the CVaR can be also defined as the conditional expectation of Y, given that $Y > \mathrm{VaR}_\alpha(Y)$, i. e.

$$\mathrm{CVaR}_\alpha(Y) = \mathbb{E}\,(Y \mid Y > \mathrm{VaR}_\alpha(Y)). \tag{4}$$

If we use $-Y$ and $1-\alpha$ instead of X and p, respectively, we can directly from the definition of CVaR and Lemma 2 derive:

$$\frac{F_X^{(-2)}(p)}{p} = -\mathrm{CVaR}_\alpha(Y),$$

and consequently

$$X_1 \succeq_{\mathrm{SSD}} X_2 \Longleftrightarrow \mathrm{CVaR}_\alpha(Y_1) \le \mathrm{CVaR}_\alpha(Y_2) \quad \forall\, \alpha \in \langle 0, 1 \rangle. \tag{5}$$

From now on, let us assume that Y is a discrete random variable which takes scenarios $y^t, t = 1, \dots, T$ with equal probabilities. Following Rockafellar and Uryasev [18] and Pflug [13], (3) can be rewritten as a linear programming problem:

$$\begin{aligned} \mathrm{CVaR}_\alpha(Y) \quad &= \quad \min_{a, w_t} \quad a + \frac{1}{(1-\alpha)T} \sum_{t=1}^{T} w_t \\ \text{s. t.} \quad w_t \quad &\ge \quad y_t - a \\ w_t \quad &\ge \quad 0. \end{aligned} \tag{6}$$

Let $y^{[k]}$ be the kth smallest element among $y^1, y^2, \dots, y^T$, i. e. $y^{[1]} \le y^{[2]} \le \dots \le y^{[T]}$. The optimal solution of (6) is derived in the following theorem.

Theorem 3. If $\alpha \in \left\langle \frac{k}{T}, \frac{k+1}{T} \right\rangle$ and $\alpha \ne 1$ then

$$\mathrm{CVaR}_\alpha(Y) = y^{[k+1]} + \frac{1}{(1-\alpha)T} \sum_{i=k+1}^{T} (y^{[i]} - y^{[k+1]}) \tag{7}$$

for $k = 0, 1, \dots, T-1$ and $\mathrm{CVaR}_1(Y) = y^{[T]}$.

Proof. Consider a random variable Y which takes values $y^t, t = 1, \dots, T$ with probabilities $p_1, p_2, \dots, p_T$. For a chosen α define j_α such that

$$\alpha \in \left\langle \sum_{j=1}^{j_\alpha - 1} p_j, \sum_{j=1}^{j_\alpha} p_j \right\rangle.$$

Then the following formula was proved in Rockafellar and Uryasev [18]:

$$\mathrm{CVaR}_\alpha(Y) = \frac{1}{1-\alpha} \left[\left(\sum_{j=1}^{j_\alpha} p_j - \alpha \right) y^{[j_\alpha]} + \sum_{j=j_\alpha+1}^{T} p_j y^{[j]} \right].$$

Since $p_t = 1/T$, $t = 1, \dots, T$ we set: $j_\alpha = k + 1$ and the theorem follows. □

Combining Theorem 3 with (5) we obtain the necessary and sufficient condition for the second-order stochastic dominance. This conditions can be more easily verified than the general conditions in Lemma 1, Lemma 2 or (5).

Theorem 4. Let $Y_1 = -X_1$ and $Y_2 = -X_2$ be discrete random variables which take values y_1^t and y_2^t, $t = 1, \dots, T$, respectively, with equal probabilities. Then

$$X_1 \succeq_{\mathrm{SSD}} X_2 \Longleftrightarrow \mathrm{CVaR}_\alpha(Y_1) \le \mathrm{CVaR}_\alpha(Y_2) \quad \forall \alpha \in \left\{0, \frac{1}{T}, \frac{2}{T}, \dots, \frac{T-1}{T}\right\}. \tag{8}$$

Proof. Let $\alpha_k = k/T$, $k = 0, 1, \dots, T-2$. Lemma 1 implies:

$$\mathrm{CVaR}_{\beta_1}(Y_i) = \mathrm{CVaR}_{\beta_2}(Y_i), \quad i = 1, 2 \quad \text{for all} \quad \beta_1, \beta_2 \in \left\langle \frac{T-1}{T}, 1 \right\rangle.$$

Thus it suffices to show that if

$$\mathrm{CVaR}_{\alpha_k}(Y_1) \le \mathrm{CVaR}_{\alpha_k}(Y_2) \tag{9}$$

and

$$\mathrm{CVaR}_{\alpha_{k+1}}(Y_1) \le \mathrm{CVaR}_{\alpha_{k+1}}(Y_2) \tag{10}$$

then it holds for all $\alpha \in \langle \alpha_k, \alpha_{k+1} \rangle$. To obtain a contradiction, suppose that (9) and (10) holds and there exists $\widetilde{\alpha} \in \langle \alpha_k, \alpha_{k+1} \rangle$ such that $\mathrm{CVaR}_{\widetilde{\alpha}}(Y_1) > \mathrm{CVaR}_{\widetilde{\alpha}}(Y_2)$. From continuity of CVaR in α there exists $\alpha^1 \in \langle \alpha_k, \alpha_{k+1} \rangle$ and $\alpha^2 \in \langle \alpha_k, \alpha_{k+1} \rangle$, $\alpha^1 \ne \alpha^2$ such that

$$\mathrm{CVaR}_{\alpha^1}(Y_1) = \mathrm{CVaR}_{\alpha^1}(Y_2) \tag{11}$$

$$\mathrm{CVaR}_{\alpha^2}(Y_1) = \mathrm{CVaR}_{\alpha^2}(Y_2). \tag{12}$$

Substituting (7) into (11) and (12) we conclude that $\alpha^1 = \alpha^2$, contrary to $\alpha^1 \ne \alpha^2$, and the proof is complete. □

4. SSD PORTFOLIO EFFICIENCY CRITERIA

Consider a random vector $\boldsymbol{r} = (r_1, r_2, \dots, r_N)'$ of returns of N assets and T equiprobable scenarios. The returns of the assets for the various scenarios are given by

$$X = \begin{pmatrix} \boldsymbol{x}^1 \\ \boldsymbol{x}^2 \\ \vdots \\ \boldsymbol{x}^T \end{pmatrix}$$

where $\boldsymbol{x}^t = (x_1^t, x_2^t, \dots, x_N^t)$ is the t-th row of matrix X. Without loss of generality we can assume that the columns of X are linearly independent. In addition to the individual choice alternatives, the decision maker may also combine the alternatives into a portfolio. We will use $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)'$ for a vector of portfolio weights and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in R^N \,|\, \mathbf{1}'\boldsymbol{\lambda} = 1, \quad \lambda_n \geq 0, \; n = 1, 2, \ldots, N\}.$$

The tested portfolio is denoted by $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_N)'$.

Definition 1. A given portfolio $\boldsymbol{\tau} \in \Lambda$ is *SSD inefficient* if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\boldsymbol{r}'\boldsymbol{\lambda} \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is *SSD efficient.*

This definition classifies portfolio as SSD efficient if and only if no other portfolio is better for all risk averse and risk neutral decision makers.

In Post [14] and Kuosmanen [7], the SSD portfolio efficiency tests based on applications of Lemma 1 were introduced. We will derive sufficient and necessary conditions for SSD efficiency of $\boldsymbol{\tau}$ based on quantile model of second order stochastic dominance, in particular the relationship between CVaR and SSD will be used.

We start with necessary condition using the following theorem. To simplify the notation, set $\Gamma = \left\{0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}\right\}$.

Theorem 5. Let $\alpha_k \in \Gamma$ and

$$d^* = \max_{\lambda_n} \sum_{k=0}^{T-1} \sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \tag{13}$$

$$\text{s. t.} \sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \geq 0, \quad k = 0, 1, \ldots, T-1, \; \lambda \in \Lambda.$$

If $d^* > 0$ then $\boldsymbol{\tau}$ is SSD inefficient. Optimal solution $\boldsymbol{\lambda}^*$ of (13) is an SSD efficient portfolio such that $\boldsymbol{r}'\boldsymbol{\lambda}^* \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}$.

Proof. If $d^* > 0$ then there is a feasible solution $\boldsymbol{\lambda}$ of problem (13) satisfying

$$\sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \geq 0, \qquad \forall\, \alpha_k \in \Gamma$$

where at least one strict inequality holds. For this $\boldsymbol{\lambda}$ we have

$$\sum_{n=1}^{N} \lambda_n \mathrm{CVaR}_{\alpha_k}(-r_n) \leq \mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}), \qquad \forall\, \alpha_k \in \Gamma$$

with at least one strict inequality. From the convexity of CVaR we obtain

$$\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\lambda}) \leq \sum_{n=1}^{N} \lambda_n \mathrm{CVaR}_{\alpha_k}(-r_n) \qquad \forall\, \alpha_k \in \Gamma.$$

Hence
$$\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\lambda}) \le \mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) \quad \forall\, \alpha_k \in \Gamma$$
with at least one strict inequality and the rest of the proof follows from Theorem 4. □

The power of necessary condition in Theorem 5 depends on correlation between random variables r_n, $n = 1, 2, \dots, N$ and portfolio $\boldsymbol{\tau}$ can be SSD inefficient even if (13) has no feasible solution or $d^* = 0$. If $d^* = 0$ then two possibilities may occur:

(1) Problem (13) has a unique solution $\boldsymbol{\lambda}^* = \boldsymbol{\tau}$. If this is the case then $\boldsymbol{\tau}$ is SSD efficient.

(2) Problem (13) has an optimal solution $\boldsymbol{\lambda}^* \neq \boldsymbol{\tau}$. In this case, $\boldsymbol{\tau}$ is SSD inefficient and $\boldsymbol{r}'\boldsymbol{\lambda}^* \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}$. Moreover, $\boldsymbol{\lambda}^*$ is an SSD efficient portfolio.

The situation when $d^* = 0$, $\boldsymbol{\lambda}^* \neq \boldsymbol{\tau}$ and $\boldsymbol{\tau}$ is SSD efficient would imply $X\boldsymbol{\lambda}^* = X\boldsymbol{\tau}$ which contradicts the assumption of linearly independent columns of X.

If problem (13) has no feasible solution then we can employ the following necessary and sufficient condition for SSD efficiency.

Theorem 6. Let $\alpha_k \in \Gamma$ and
$$D^*(\boldsymbol{\tau}) = \max_{D_k, \lambda_n, b_k} \sum_{k=0}^{T-1} D_k \tag{14}$$
s. t.
$$\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) - b_k - \frac{1}{1-\alpha_k}\mathbb{E}\max(-\boldsymbol{r}'\boldsymbol{\lambda} - b_k, 0) \ \ge\ D_k,\ k = 0, 1, \dots, T-1$$
$$D_k \ \ge\ 0, \quad k = 0, 1, \dots, T-1$$
$$\boldsymbol{\lambda} \ \in\ \Lambda.$$
If $D^*(\boldsymbol{\tau}) > 0$ then $\boldsymbol{\tau}$ is SSD inefficient and $\boldsymbol{r}'\boldsymbol{\lambda}^* \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}$. Otherwise, $D^*(\boldsymbol{\tau}) = 0$ and $\boldsymbol{\tau}$ is SSD efficient.

Proof. Let $\boldsymbol{\lambda}^*$, b_k^*, $k = 0, 1, \dots, T-1$ be an optimal solution of (14). If $D^*(\boldsymbol{\tau}) > 0$ then
$$b_k^* + \frac{1}{1-\alpha_k}\mathbb{E}\max(-\boldsymbol{r}'\boldsymbol{\lambda}^* - b_k^*, 0) \le \mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau}) \ \ \forall\, \alpha_k \in \Gamma \tag{15}$$
where at least one inequality holds strict. Since from the definition of CVaR we have
$$\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\lambda}^*) = \min_{b_k}\left\{b_k + \frac{1}{1-\alpha_k}\mathbb{E}\max(-\boldsymbol{r}'\boldsymbol{\lambda}^* - b_k, 0)\right\}$$
we conclude from (15) that
$$\mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\lambda}^*) \le \mathrm{CVaR}_{\alpha_k}(-\boldsymbol{r}'\boldsymbol{\tau})$$
with at least one strict inequality. Hence $\boldsymbol{r}'\boldsymbol{\lambda}^* \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}$ and $\boldsymbol{\tau}$ is SSD inefficient.

If $D^*(\boldsymbol{\tau}) = 0$ then problem (14) has unique optimal solution: $\boldsymbol{\lambda}^* = \boldsymbol{\tau}$, because the presence of another optimal solution contradicts the assumption of linearly independent columns of X. Thus there is no strictly dominating portfolio and hence $\boldsymbol{\tau}$ is SSD efficient. Since $\boldsymbol{\tau}$ is always a feasible solution of (14), D^* can not be negative and the proof is complete. $\square$

Nonlinear program (14) has $N+2T+1$ constraints and $N+2T$ variables. Inspired by (6) and following Pflug [13], Rockafellar and Uryasev [18], it can be rewritten as a linear programming problem with $2T(T+1)+N+1$ constraints and $T(T+2)+N$ variables:

$$D^*(\boldsymbol{\tau}) = \max_{D_k,\lambda_n,b_k,w_k^t} \sum_{k=1}^{T} D_k \tag{16}$$

s. t.

$$\begin{aligned}
\mathrm{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}) - b_k - \frac{1}{(1-\frac{k-1}{T})T}\sum_{t=1}^{T} w_k^t &\geq D_k, & k &= 1,2,\dots,T\\
w_k^t &\geq -\mathbf{x}^t\boldsymbol{\lambda} - b_k, & t,k &= 1,2,\dots,T\\
w_k^t &\geq 0, & t,k &= 1,2,\dots,T\\
D_k &\geq 0, & k &= 1,2,\dots,T\\
\boldsymbol{\lambda} &\in \Lambda. &&
\end{aligned}$$

Using (16) instead of (14) in Theorem 6 we obtain a linear programming criterion for SSD efficiency.

This sufficient and necessary condition requires to solve a smaller linear program than it is in the Kuosmanen test. Furthermore, contrary to the Post and the Kuosmanen test, it identifies SSD efficient dominating portfolio as a by-product.

5. A SSD PORTFOLIO INEFFICIENCY MEASURE

Inspired by Post [14] and Kopa and Post [6], $D^*(\boldsymbol{\tau})$ from (14) or (16) can be considered as a measure of inefficiency of portfolio $\boldsymbol{\tau}$, because it expresses the distance between a given tested portfolio and its dominating SSD efficient portfolio. If there exist more dominating SSD efficient portfolios then the least risky portfolio, measured by CVaR, is considered. To be able to compare SSD inefficiency of two portfolios we need to consider such a measure, which is "consistent" with SSD relation.

Definition 2. Let ξ be a measure of SSD portfolio inefficiency. We say that ξ is consistent with SSD if and only if

$$\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2 \Rightarrow \xi(\boldsymbol{\tau}_2) \geq \xi(\boldsymbol{\tau}_1)$$

for any $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \Lambda$.

The property of consistency guarantees that if a given portfolio is worse than the other one for every risk averse investor then it has larger measure of inefficiency. Let $\Lambda^*(\boldsymbol{\tau}) \in \Lambda$ be a set of optimal solutions $\boldsymbol{\lambda}^*$ of (14) or (16).

Theorem 7.

(i) The measure of SSD portfolio inefficiency D^* given by either (14) or (16) is consistent with SSD.

(ii) If $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\text{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$ and both $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ are SSD inefficient then

$$D^*(\boldsymbol{\tau}_2) = D^*(\boldsymbol{\tau}_1) + \sum_{k=1}^{T} \left[\text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \right].$$

(iii) If $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\text{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$ then

$$D^*(\boldsymbol{\tau}_2) \geq D^*(\boldsymbol{\tau}_1) + \sum_{k=1}^{T} \left[\text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \right].$$

Proof. Applying Theorem 4, if $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\text{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$ then

$$\sum_{k=1}^{T} \left[\text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k-1}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \right] \geq 0.$$

Hence it suffices to prove (ii) and (iii).

Let $\boldsymbol{r}'\boldsymbol{\tau}_1$ be SSD inefficient. It is easily seen that (14) can be rewritten in the following way:

$$\begin{aligned} D^*(\boldsymbol{\tau}) = \max_{\lambda_n} \sum_{k=0}^{T-1} & \left[\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}) \right] \\ \text{s.t.} \;\; & \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}) \geq 0, \quad k = 0,1,\ldots,T-1 \\ & \boldsymbol{\lambda} \in \Lambda. \end{aligned} \tag{17}$$

Let $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$, $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_2) \in \Lambda^*(\boldsymbol{\tau}_2)$. Using Theorem 4 and $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\text{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$,

$$\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \geq 0 \quad k = 0,1,\ldots,T-1.$$

Since the sum of these differences does not depend on the choice of $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)$, the dominating portfolio $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)$ is also an optimal solution of (14) when deriving $D^*(\boldsymbol{\tau}_2)$, i.e. $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$. Hence

$$\begin{aligned} D^*(\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \left[\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_2)) \right] \\ &= \sum_{k=0}^{T-1} \left[\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \right] \\ &\quad + \sum_{k=0}^{T-1} \left[\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \right] \\ &= D^*(\boldsymbol{\tau}_1) + \sum_{k=0}^{T-1} \left[\text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \text{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) \right] \end{aligned}$$

what completes the proof of (ii).

Let $\boldsymbol{r}'\boldsymbol{\tau}_1$ be SSD efficient. From Theorem 6, we have $D^*(\boldsymbol{\tau}_1) = 0$. According to (17),

$$\begin{aligned} D^*(\boldsymbol{\tau}_2) = \max_{\lambda_n} \sum_{k=0}^{T-1} \Big[\mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda})\Big] & \\ \text{s.t.}\ \ \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}) \ \ \geq \ \ 0, \quad k = 0, 1, \ldots, T-1 & \\ \boldsymbol{\lambda} \ \ \in \ \ \Lambda. & \end{aligned}$$

Since $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$, portfolio $\boldsymbol{\tau}_1$ is a feasible solution of (17). Hence

$$D^*(\boldsymbol{\tau}_2) \geq \sum_{k=0}^{T-1} \Big[\mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1)\Big]$$

and combining it with (ii), the proof is complete. □

Since SSD relation is not complete, i.e. there exist incomparable pairs of portfolios, the strict inequality of values of any portfolio inefficiency measure can not imply SSD relation. Also for the measure D^* some pair of portfolios $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$ can be found such that $D^*(\boldsymbol{\tau}_2) \geq D^*(\boldsymbol{\tau}_1)$ and $\boldsymbol{r}'\boldsymbol{\tau}_1 \not\succeq_{SSD} \boldsymbol{r}'\boldsymbol{\tau}_2$.

In the following theorem, a convexity property of portfolio inefficiency measure D^* is analyzed.

Theorem 8. Let $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3 \in \Lambda$.

(i) If $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$ then

$$D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) \leq \eta D^*(\boldsymbol{\tau}_1) + (1-\eta)D^*(\boldsymbol{\tau}_2)$$

for any $\eta \in \langle 0, 1\rangle$.

(ii) If $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$ and $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_3$ then

$$\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3)$$

and

$$D^*(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3) \leq \eta D^*(\boldsymbol{\tau}_2) + (1-\eta)D^*(\boldsymbol{\tau}_3)$$

for any $\eta \in \langle 0, 1\rangle$.

P r o o f. (i) Applying Lemma 1 for equiprobable scenario approach, we obtain

$$\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2 \Rightarrow \boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\left(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2\right) \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$$

for any $\eta \in \langle 0, 1\rangle$. By analogy to the proof of previous theorem, if $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$ then $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$ and $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2)$. Hence

$$\begin{aligned}
D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\boldsymbol{r}'[\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2]\right) \\
&= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
D^*(\boldsymbol{\tau}_1) &= \sum_{k=0}^{T-1} \left[\mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) - \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1))\right] \\
D^*(\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \left[\mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1))\right].
\end{aligned}$$

Combining it with convexity of CVaR, we obtain

$$\begin{aligned}
D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\boldsymbol{r}'[\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2]\right) \\
&\quad - \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
&\leq \eta \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_1) + (1-\eta)\sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\tau}_2) \\
&\quad -\eta \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
&\quad -(1-\eta) \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\boldsymbol{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
&\leq \eta D^*(\boldsymbol{\tau}_1) + (1-\eta) D^*(\boldsymbol{\tau}_2).
\end{aligned}$$

(ii) Applying Lemma 1 for scenario approach, we obtain:

$$\boldsymbol{r}'\boldsymbol{\tau} \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\lambda} \Longleftrightarrow \sum_{t=1}^{T}(\boldsymbol{x}^t\boldsymbol{\tau} - \boldsymbol{x}^t\boldsymbol{\lambda}) \geq 0 \quad \forall\, t = 1, 2, \ldots, T. \tag{18}$$

Hence

$$\begin{aligned}
\sum_{t=1}^{T}(\boldsymbol{x}^t\boldsymbol{\tau}_1 - \boldsymbol{x}^t\boldsymbol{\tau}_2) &\geq 0 \quad \forall\, t = 1, 2, \ldots, T \\
\sum_{t=1}^{T}(\boldsymbol{x}^t\boldsymbol{\tau}_1 - \boldsymbol{x}^t\boldsymbol{\tau}_3) &\geq 0 \quad \forall\, t = 1, 2, \ldots, T
\end{aligned}$$

and therefore

$$\sum_{t=1}^{T}(\boldsymbol{x}^t\boldsymbol{\tau}_1 - \eta\boldsymbol{x}^t\boldsymbol{\tau}_2 - (1-\eta)\boldsymbol{x}^t\boldsymbol{\tau}_3) \geq 0 \quad \forall\, t = 1, 2, \ldots, T$$

for any $\eta \in \langle 0,1\rangle$. Thus, according to Lemma 1,

$$\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3) \quad \text{for any } \eta \in \langle 0,1\rangle.$$

Similarly to the proof of previous theorem, if $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$ then $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$, $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_3)$ and $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3)$ for any $\eta \in \langle 0,1\rangle$ and the rest of the proof follows by analogy to (i). □

Let $I(\boldsymbol{\tau})$ be a set of all portfolios whose returns are SSD dominated by return of $\boldsymbol{\tau}$, i.e.

$$I(\boldsymbol{\tau}) = \{\boldsymbol{\lambda} \in \Lambda \,|\, \boldsymbol{r}'\boldsymbol{\tau} \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\lambda}\}.$$

Theorem 8 shows that $I(\boldsymbol{\tau})$ is convex and D^* is convex on $I(\boldsymbol{\tau})$ for any $\boldsymbol{\tau} \in \Lambda$. Both these properties are consequences of convexity of CVaR. The following example illustrates these results and we stress the fact that the set of SSD efficient portfolios is not convex.

6. NUMERICAL EXAMPLE

Consider three assets with three scenarios:

$$X = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & 7 & 5 \end{pmatrix}.$$

It is easy to check that $\boldsymbol{\lambda}_1 = (1,0,0)'$, $\boldsymbol{\lambda}_2 = (0,1,0)'$ and $\boldsymbol{\lambda}_3 = (0,0,1)'$ are SSD efficient. Let $\boldsymbol{\tau}_1 = \boldsymbol{\lambda}_3$, $\boldsymbol{\tau}_2 = (\frac{1}{2},\frac{1}{2},0)'$ and let $\boldsymbol{\tau}_3 = (\frac{1}{3},\frac{2}{3},0)'$. Then $X\boldsymbol{\tau}_2 = (-\frac{1}{2},\frac{1}{2},\frac{9}{2})$ and according to (18), $\boldsymbol{r}'\boldsymbol{\tau}_1 \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_2$. Hence the set of SSD efficient portfolios is not convex. Similarly, $\boldsymbol{r}'\boldsymbol{\tau}_1 \succ_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_3$ and $\boldsymbol{r}'\boldsymbol{\tau}_1 \succeq_{\mathrm{SSD}} \boldsymbol{r}'\boldsymbol{\tau}_1$. Applying Theorem 8, a set of convex combinations of $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$ is a subset of $I(\boldsymbol{\tau}_1)$. We will show that $I(\boldsymbol{\tau}_1)$ consists only of convex combinations of $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ and $\boldsymbol{\tau}_3$, i.e.

$$I(\boldsymbol{\tau}_1) = \left\{\boldsymbol{\lambda} \in \Lambda \,|\, \lambda = \eta_1\boldsymbol{\tau}_1 + \eta_2\boldsymbol{\tau}_2 + \eta_3\boldsymbol{\tau}_3,\ \eta_i \geq 0,\ i = 1,2,3, \sum_{i=1}^{3}\eta_i = 1\right\}.$$

Substituting into (18) we can see that only portfolios $\boldsymbol{\lambda} \in \Lambda$ satisfying the following system of inequalities can be included in $I(\boldsymbol{\tau}_1)$:

$$\begin{aligned} -\lambda_2 &\leq 0 \\ \lambda_1 - \lambda_2 &\leq 0 \\ 3\lambda_1 + 6\lambda_2 + 5(1-\lambda_1-\lambda_2) &\leq 5. \end{aligned}$$

The graphical solution of this system is illustrated in the following Figure 1 and we can see that the set of portfolios which returns are SSD dominated by return of portfolio $\boldsymbol{\tau}_1$ is equal to the set of all convex combinations of portfolios $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$. Points A, B and C correspond to portfolios $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$, $\boldsymbol{\tau}_1$, respectively.

As was shown in Theorem 8 (ii), SSD portfolio inefficiency measure D^* is convex on $I(\boldsymbol{\tau}_1)$. Figure 2 shows the graph of D^* on $I(\boldsymbol{\tau}_1)$. Since $\boldsymbol{\tau}_1$ is SSD efficient, $D^*(\boldsymbol{\tau}_1) = 0$ and $D^*(\boldsymbol{\tau}) > 0$ for all $\boldsymbol{\tau} \in I(\boldsymbol{\tau}_1)\backslash\{\boldsymbol{\tau}_1\}$.

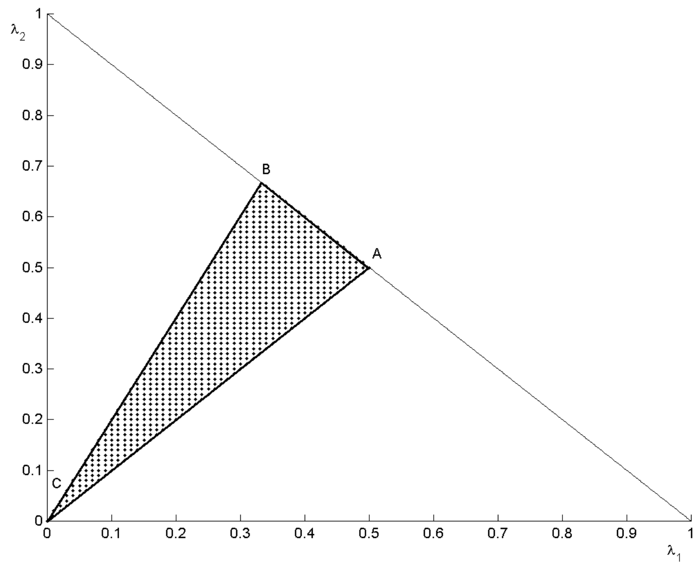


Fig. 1. The set $I(\boldsymbol{\tau}_1)$ of portfolios whose returns are SSD dominated by return of portfolio $\boldsymbol{\tau}_1 = (0, 0, 1)$.

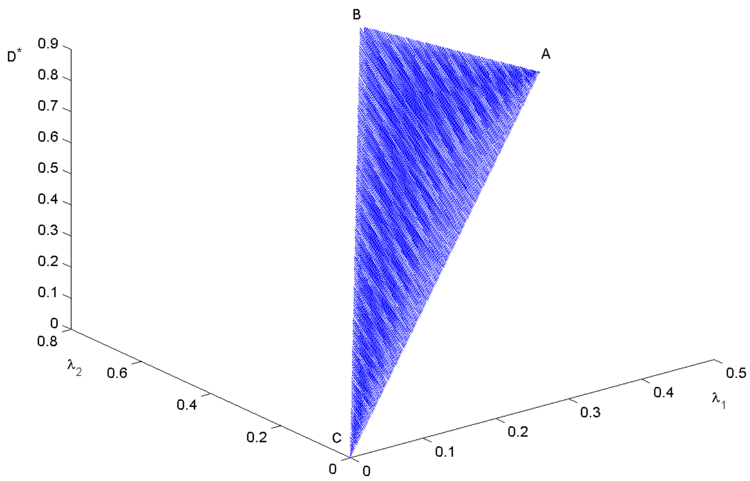


Fig. 2. The graph of D^* on $I(\boldsymbol{\tau}_1)$.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic under grant 0021620839 and the Czech Science Foundation (grant 201/05/H007 and grant 201/07/P107). This research was conducted while Petr Chovanec was Junior Researcher at Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic (Department of Econometrics). He wants to thank for support the Czech Science Foundation (grants 402/03/H057, and 402/04/1294).

(Received November 5, 2007.)

REFERENCES

[1] E. De Giorgi: Reward-risk portfolio selection and stochastic dominance. J. Banking Finance *29* (2005), 895–926.
[2] E. De Giorgi and T. Post: Second order stochastic dominance, reward-risk portfolio selection and the CAPM. J. Financial Quantitative Analysis, to appear.
[3] P. H. Dybvig and S. A. Ross: Portfolio efficient sets. Econometrica *50* (1982), 6, 1525–1546.
[4] J. Hadar and W. R. Russell: Rules for ordering uncertain prospects. Amer. Econom. Rev. *59* (1969), 1, 25–34.
[5] G. Hanoch and H. Levy: The efficiency analysis of choices involving risk. Rev. Econom. Stud. *36* (1969), 335–346.
[6] M. Kopa and T. Post: A portfolio optimality test based on the first-order stochastic dominance criterion. J. Financial Quantitative Analysis, to appear.
[7] T. Kuosmanen: Efficient diversification according to stochastic dominance criteria. Management Sci. *50* (2004), 10, 1390–1406.
[8] H. Levy: Stochastic dominance and expected utility: Survey and analysis. Management Sci. *38* (1992), 4, 555–593.
[9] H. Levy: Stochastic Dominance: Investment Decision Making Under Uncertainty. Second edition. Springer Science, New York 2006.
[10] H. M. Markowitz: Portfolio Selection. J. Finance *7* (1952), 1, 77–91.
[11] H. M. Markowitz: Portfolio Selection: Efficient Diversification in Investments. Wiley, New York 1959.
[12] W. Ogryczak and A. Ruszczyński: Dual stochastic dominance and related mean-risk models. SIAM J. Optim. *13* (2002), 60–78.
[13] G. Ch. Pflug: Some remarks on the value-at-risk and the conditional value-at-risk. In: Probabilistic Constrained Optimization: Methodology and Applications (S. Uryasev, ed.), Kluwer Academic Publishers, Norwell MA 2000, pp. 278–287.
[14] T. Post: Empirical tests for stochastic dominance efficiency. J. Finance *58* (2003), 1905–1932.
[15] M. Rothschild and J. E. Stiglitz: Rules for ordering uncertain prospects. J. Econom. Theory *2* (1969), 225–243.
[16] W. R. Russell and T. K. Seo: Representative sets for stochastic dominance rules. In: Studies in the Economics of Uncertainty (T. B. Fomby and T. K. Seo, eds.), Springer-Verlag, New York 1989, pp. 59–76.
[17] A. Ruszczyński and R. J. Vanderbei: Frontiers of stochastically nondominated portfolios. Econometrica *71* (2003), 4, 1287–1297.
[18] S. Uryasev and R. T. Rockafellar: Conditional value-at-risk for general loss distributions. J. Banking Finance *26* (2002), 1443–1471.

[19] G. A. Whitmore: Third degree stochastic dominance. Amer. Econom. Rev. *60* (1970), 457–459.

Miloš Kopa, Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics – Charles University, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: kopa@karlin.mff.cuni.cz

Petr Chovanec, Suez LNG NA, LLC, One Liberty Square, 10th Floor, Boston, MA 02109. U. S. A.
e-mail: petr.chovanec@suezenergyna.com

European Journal of Operational Research xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

General linear formulations of stochastic dominance criteria

Thierry Post [a,*], Miloš Kopa [b]

[a] Koç University's Graduate School of Business, Rumelifeneri Yolu, 34450 Sariyer, Istanbul, Turkey
[b] Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Czech Republic

ARTICLE INFO

Article history:
Received 11 September 2012
Accepted 8 April 2013
Available online xxxx

Keywords:
Stochastic dominance
Utility theory
Non-satiation
Risk aversion
Prudence
Temperance

ABSTRACT

We develop and implement linear formulations of general *N*th order stochastic dominance criteria for discrete probability distributions. Our approach is based on a piece-wise polynomial representation of utility and its derivatives and can be implemented by solving a relatively small system of linear inequalities. This approach allows for comparing a given prospect with a discrete set of alternative prospects as well as for comparison with a polyhedral set of linear combinations of prospects. We also derive a linear dual formulation in terms of lower partial moments and co-lower partial moments. An empirical application to historical stock market data suggests that the passive stock market portfolio is highly inefficient relative to actively managed portfolios for all investment horizons and for nearly all investors. The results also illustrate that the mean–variance rule and second-order stochastic dominance rule may not detect market portfolio inefficiency because of non-trivial violations of non-satiation and prudence.

1. Introduction

Stochastic dominance (SD), first introduced in Quirk and Saposnik (1962), Hadar and Russell (1969) and Hanoch and Levy (1969), is a useful concept for analyzing risky decision making when only partial information about the decision maker's risk preferences is available. The concept is used in numerous empirical studies and practical applications, ranging from agriculture and health care to financial management and public policy making; see, for example, the extensive survey in the text book of Levy (2006). A selection of recent studies in OR/MS journals includes Post (2008), Lozano and Gutiérrez (2008), Blavatskyy (2010), Dupačová and Kopa (2012), Lizyayev and Ruszczyński (2012), Lizyayev (2012) and Brown et al. (2012).

SD imposes general preference restrictions without assuming a functional form for the decision maker's utility function. The SD rules of order one to four are particularly interesting, because they impose (in a cumulative way) the standard assumptions of non-satiation, risk-aversion, prudence and temperance, which are necessary conditions for standard risk aversion (Kimball, 1993). This approach is theoretically appealing but not always easy to implement. In some special cases, a closed-form analytical solution exists, as is true, for example, for the textbook case of a pair-wise comparison of two given prospects based on the second-order stochastic dominance (SSD) rule.

However, more generally, a closed-form solution does not exist and numerical optimization is required. For example, Meyer's (1977a,b) stochastic dominance with respect to a function (SDWRF) requires solving an (small and standard) optimal control problem. The rules of convex stochastic dominance (Fishburn, 1974) for comparing more than two prospects simultaneously also require optimization. For example, Bawa et al. (1985) develop Linear Programming tests for convex first-order stochastic dominance (FSD), convex SSD and an approximation for convex third-order stochastic dominance (TSD). Shalit and Yitzhaki (1994), Post (2003), Kuosmanen (2004) and Kopa and Chovanec (2008) develop Linear Programming tests that compare a given prospect using SSD with a polyhedral set of linear combinations of a discrete set of prospects.

Unfortunately, a general algorithm is not available. How can we test, for example, whether a given medical treatment is dominated by convex fourth-order stochastic dominance (FOSD) relative to a set of alternative treatments? How can we test whether a given investment portfolio is FOSD efficient relative to a polyhedral set of portfolios formed from a set of base assets? Without an algorithm for these specific cases, we may be forced to use known tests for less discriminating decision criteria. For example, we could use a set of pair-wise FOSD tests to compare the evaluated medical treatment with every alternative treatment. Similarly, we could use pair-wise tests to compare the evaluated investment portfolio with a large number of alternative portfolios, for example, using a grid search or random search over the possibilities set. However, pair-wise comparisons generally are less powerful than convex SD, because a prospect can be non-optimal for all admissible utility functions without being dominated by any alternative prospect. A

* Corresponding author. Tel.: +90 212 338 13 75; fax: +90 212 338 16 53.
E-mail address: thierrypost@hotmail.com (T. Post).

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.04.015

further possible loss of power stems from using a discrete approximation to a continuous choice set.

Section 2 of this study develops linear formulations of general stochastic dominance rules. Our approach is based on a piece-wise polynomial representation of utility and its derivatives. This representation applies generally for higher-order SD rules (*N*th order SD), comparing a given prospect with a discrete set of alternative prospects (convex NSD analysis), and comparing a given prospect with a polyhedral set of linear combinations of prospects (NSD efficiency analysis). Our analysis therefore represents a generalization of the lower-order tests of Bawa et al. (1985) and Post (2003). We can also deal with additional preference restrictions such as the bounds on the level of risk aversion of Meyer (1977a,b) and the bounds on utility curvature by Leshno and Levy (2002). The use of piece-wise polynomial functions also generalizes results by Hadar and Seo (1988) and Russell and Seo (1989) on simple representative utility functions for pairwise comparison based on lower-order SD rules.

To arrive at a finite optimization problem, we focus on discrete probability distributions. In empirical studies, we usually face discrete sample distributions, and experimental studies generally use prospects with a discrete population distribution. In addition, many continuous distributions can be approximated accurately with a discrete distribution. Our approach can be implemented by solving a relatively small system of linear inequalities. The linear structure seems particularly convenient for the application of statistical re-sampling methods in the spirit of Nelson and Pope (1991) and Barrett and Donald (2003).

Our focus is on utility and its derivatives and on restrictions that follow from utility theory. Still, Section 3 also derives linear dual formulations that are formulated in terms of lower partial moments (Bawa, 1975) and co-lower partial moments (Bawa and Lindenberg, 1977) of the probability distribution. We focus on the dominance classification of a given prospect and we do not attempt to identify an alternative prospect that dominates the evaluated prospect. In the case of a discrete choice set, a non-admissible prospect need not be dominated by any alternative prospect. In addition, a prospect that dominates the choice of a given decision maker need not be optimal for that decision maker, and, moreover, the optimum need not dominate the current choice. Finally, the dominance relation between a pair of prospects generally is less robust than the classification of a given prospect. For these reasons, the search for a dominant prospect seems irrelevant for our purposes. Still, the dual formulations are useful for computational efficiency and robustness analysis.

Section 4 applies a range of SD tests to historical stock return data to compare the broad stock market portfolio with alternative portfolios formed from a set of risky benchmark stock portfolios and riskless Treasury bills. We analyze horizons ranging from 1 month to 10 years and consider the decision criteria of SSD, TSD, FOSD, SDWRF, ASSD and mean–variance (M–V) analysis. The analysis is relevant because a large class of capital market equilibrium models predict that the market portfolio is efficient. Another reason for expecting market portfolio efficiency is the popularity of passive mutual funds and exchange traded funds that passively track broad stock market indices.

Our empirical analysis shows that the market portfolio is highly and significantly inefficient by the TSD, FOSD, SDWRF and ASSD criteria for every horizon. Few rational risk averters would hold the broad market portfolio in the face of the historical return premiums to active strategies. The appeal of active strategies only increases with the horizon. Our results also illustrate that pair-wise dominance comparisons and the SSD and M–V rules have limited discriminating power and can generate misleading results in relevant applications. The SSD criterion may fail to detect market portfolio inefficiency for short horizons, because it penalizes small-cap stocks for having a relatively high positive systematic skewness, violating prudence. M–V analysis underestimates the level of market portfolio inefficiency for long horizons, because it assigns negative weights to large positive market returns, placing a penalty on outperformance during bull markets. In our application, these phenomena lead to a non-trivial underestimation of the alphas for small-cap stocks.

2. Linear formulation in terms of piece-wise polynomial utility

We consider M prospects with risky outcomes $x_1,\ldots,x_M$. A prospect is defined here in a general way as an available choice alternative and it could be a given combination of multiple base alternatives, for example, a combination of production methods, financial assets or marketing instruments. Depending on the application, the outcomes may be total wealth, consumption, income, or any variable that can reasonably be assumed to enter as an argument to a utility function that obeys the maintained assumptions. The outcomes are treated as random variables with a discrete, state-dependent, joint probability distribution characterized by R mutually exclusive and exhaustive scenarios with probabilities $p_r > 0$, $r = 1,\ldots,R$. We use $x_{i,r}$ for the outcome of prospect i in scenario r. We collect all possible outcomes across prospects and states in $Y = \{y{:}\ y = x_{i,r}\ i = 1,\ldots,M;\ r = 1,\ldots,R\}$, rank these values in ascending order $y_1 \leqslant \cdots \leqslant y_S$ and use $q_{i,s} = \Pr[x_i = y_s] = \sum_{r:x_{i,r}=y_s} p_r$.

Decision makers' preferences are described by N-times continuously differentiable, von Neumann–Morgenstern utility functions $u(x): D \to \mathbb{R}$. We use $u^n(x)$ for the nth order derivative, $n = 1,\ldots,N$, and $u^0(x) = u(x)$. To implement stochastic dominance of order $N \geqslant 1$, we will consider the following set of admissible utility functions:

$$U_N = \{u \in C^N:\ (-1)^{n-1}u^n(x) \geqslant 0 \quad \forall x \in D,\ n = 1,\ldots,N\}. \tag{1}$$

Thus, first-order dominance assumes non-satiation ($u^1(x) \geqslant 0$, $\forall x \in D$); second-order dominance assumes also risk aversion ($u^2(x) \leqslant 0$, $\forall x \in D$); the third-order criterion adds prudence ($u^3(x) \geqslant 0$, $\forall x \in D$) and fourth-order SD also assumes temperance ($u^4(x) \leqslant 0$, $\forall x \in D$). In some applications, zero values for the derivatives may not be allowed, for example, in the cases of strict non-satiation ($u^1(x) > 0$, $\forall\ x \in D$) and strict risk aversion ($u^2(x) < 0$, $\forall x \in D$). The needed adjustments to our Linear Programming tests are obvious substitutions of weak and strict inequalities. In our experience, these adjustments have a negligible effect in empirical applications. For the sake of brevity, we therefore ignore this issue here.

For practical reasons, it is often useful to assume some sort of standardization, such as $u^1(y_1) = 1$, in order to avoid numerical problems when evaluating utility functions that approximate $u^1(x) = 0\ \forall x \in D$, or the indifferent decision maker. Since utility analysis is invariant to positive linear transformations, such standardizations are harmless.

We distinguish between three types of SD relations: pair-wise dominance relations, discrete convex dominance relations and continuous convex dominance relations, or efficiency relations. These relations differ regarding to the assumed choice possibilities: a single prospect, a discrete set of prospects, or all convex combinations of a discrete set of prospects. Consider first the case of pair-wise comparison between two given prospects:

Definition 1 (*Pair-wise Comparison*). An evaluated prospect $i \in \{1,\ldots,M\}$ is not dominated in terms of Nth order stochastic dominance, $N \geqslant 1$, by an alternative prospect $j \in \{1,\ldots,M\}$ if there exists an admissible utility function $u \in U_N$ for which it is preferred to the alternative:

$$\sum_{r=1}^{R} p_r u(x_{i,r}) \geqslant \sum_{r=1}^{R} p_r u(x_{j,r}) \Longleftrightarrow \sum_{s=1}^{S} u(y_s)(q_{i,s} - q_{j,s}) \geqslant 0. \tag{2}$$

This formulation uses a weak inequality and hence does not require a strict preference relation. We can alternatively use strict inequality to require strict dominance. There generally is no robust difference between the two definitions. For example, if we compare prospect x_1 with a mean-preserving anti-spread x_2, then (2) will apply with equality for a risk neutral decision maker. In this case, there exists a weak preference for some utility functions but no strict preference for any utility function. However, the violation of strict preference is infinitely small and the difference between the two definitions is not robust. The same consideration applies below for convex SD. For the sake of brevity, we use weak inequalities here.

If there are $M > 2$ prospects, we could perform $(M - 1)$ pair-wise dominance tests for any given prospect. However, a prospect can be non-optimal for all admissible utility functions without being dominated by any individual alternative prospect. All decision makers may agree that the evaluated prospect does not maximize their expected utility even if they do not agree on which specific alternatives achieve a higher expected utility.

Definition 2 (*Convex Stochastic Dominance*). An evaluated prospect $i \in \{1, \ldots, M\}$ is admissible in terms of Nth order stochastic dominance, $N \geqslant 1$, relative to the set of prospects $\{1, \ldots, M\}$ if there exists an admissible utility function $u \in U_N$ for which it is preferred to every alternative prospect:

$$\begin{aligned} &\sum_{r=1}^{R} p_r u(x_{i,r}) \geqslant \sum_{r=1}^{R} p_r u(x_{j,r}) \quad j = 1, \ldots, M \\ &\Longleftrightarrow \sum_{s=1}^{S} u(y_s)(q_{i,s} - q_{j,s}) \geqslant 0 \quad j = 1, \ldots, M. \end{aligned} \tag{3}$$

Since admissibility can be violated without a pair-wise dominance relation, this criterion generally is more powerful than pair-wise dominance tests.

Despite the adjective 'convex', convex SD does not consider convex combinations of the prospects. Rather, the terminology reflects that the convex SD criterion can equivalently be formulated by considering convex combinations of the cumulative distribution function (CDF) of the prospects. By contrast, the analysis of Shalit and Yitzhaki (1994), Post (2003) and Kuosmanen (2004), among others, assumes that convex combinations of the prospects are feasible:

$$X = \left\{ \sum_{j=1}^{M} \lambda_j x_j : \ \sum_{j=1}^{M} \lambda_j = 1; \quad \lambda_j \geqslant 0 \ j = 1, \ldots, M \right\}. \tag{4}$$

This situation is relevant if, for example, the decision maker can create a mixture of production methods, financial assets or marketing instruments. Diversification is especially relevant for risk averters ($N \geqslant 2$).

The formulation is not restricted to convex combinations but can also be applied in the more general case that general linear combinations of the prospects can be made subject to a set of general linear restrictions. The Minkowski–Weyl Theorem says that any polytope can be represented as the convex hull of its vertices (vertex representation). Therefore, the prospects should be considered more generally as the vertices of a polyhedral choice set.

We use $x^* \in X$ for the evaluated combination of prospects. The ordering of the scenarios is inconsequential in our analysis and we are free to label the scenarios by their ranking with respect to the evaluated combination: $x_1^* \leqslant \cdots \leqslant x_T^*$.

Definition 3 (*Stochastic Dominance Efficiency*). An evaluated combination of prospects $x^* \in X$ is efficient in terms of Nth order dominance, $N \geqslant 2$, relative to all feasible combinations $x \in X$ if it is the optimum for some admissible utility function $u \in U_N$:

$$\begin{aligned} &\sum_{r=1}^{R} p_r u(x_r^*) \geqslant \sum_{r=1}^{R} p_r u(x_r) \quad \forall x \in X \\ &\Longleftrightarrow \sum_{r=1}^{R} p_r u'(x_r^*)(x_r^* - x_{j,r}) \geqslant 0 \ \ j = 1, \ldots, M. \end{aligned} \tag{5}$$

The definition follows from the Karush–Kuhn–Tucker first-order condition for selecting the optimal combination of prospects: $\max_{x \in X} \sum_{r=1}^{R} p_r u(x_r)$. This formulation was first introduced by Post (2003) for SSD ($N = 2$) and applies also for higher-order criteria ($N > 2$), but it does not apply for FSD ($N = 1$). Kopa and Post (2009) present a different utility-based formulation for this case.

All three criteria (pairwise dominance, convex dominance, efficiency) seek an admissible utility function or marginal utility function that solves a finite set of inequality conditions. If utility and marginal utility can be expressed as linear functions of a finite set of parameters, then the criteria reduce to solving a set of linear inequalities, a task that can be performed using Linear Programming.

Theorem 1 (*Linearization of Utility and its Derivatives*). *For any utility function $u \in U_N$, $N \geqslant 1$, and a discrete set of outcomes $z_1 \leqslant \cdots \leqslant z_T$, we can represent the levels of utility and its derivatives by piece-wise polynomial functions:*

$$u(z_t) = \sum_{n=0}^{N-2} \beta_n (z_t - z_T)^n + \sum_{k=t}^{T} \gamma_k (z_t - z_k)^{N-1}, \tag{6}$$

$$\begin{aligned} u^q(z_t) &= \sum_{n=q}^{N-2} \frac{n!}{(n-q)!} \beta_n (z_t - z_T)^{n-q} + \frac{(N-1)!}{(N-q-1)!} \sum_{k=t}^{T} \gamma_k (z_t - z_k)^{N-q-1}, \\ &\quad q = 1, \ldots, N-1, \end{aligned} \tag{7}$$

where

$$\beta_n = \frac{u^n(z_T)}{n!}, \quad n = 0, 1, \ldots, N-2, \tag{8}$$

$$\begin{aligned} \gamma_k &= \frac{u^{N-1}(z_{k-1}^*) - u^{N-1}(z_k^*)}{(N-1)!}, \\ k &= 2, \ldots, T-1, \quad \gamma_T = \frac{u^{N-1}(z_{T-1}^*)}{(N-1)!}, \end{aligned} \tag{9}$$

for some values $z_k^* \in [z_k, z_{k+1}]$, $k = 1, \ldots, T-1$ such that

$$(-1)^n \beta_n \leqslant 0, \quad n = 1, \ldots, N-2, \tag{10}$$

$$(-1)^{N-1} \gamma_k \leqslant 0, \quad k = 1, 2, \ldots, T-1, \tag{11a}$$

$$(-1)^{N-1} \gamma_T \leqslant 0 \text{ for } N \geqslant 2. \tag{11b}$$

Moreover, for all parameters satisfying (10) and (11) *we can construct an admissible utility function $u \in U_N$.* (Proof in the Appendix)

The theorem uses a piecewise-constant representation to the $(N-1)$th derivative of the utility function, or, equivalently, a piecewise-linear representation of the $(N-2)$th derivative. The lower-order derivatives are obtained by integrating over the higher-order derivatives and take the shape of piecewise higher-order polynomials. The $(N-2)$th order derivative will be a kinked, piecewise linear function; the $(N-3)$th order derivative will be a smooth, piecewise quadratic function; and so forth. The piece-wise polynomial representation generalizes results by Hadar and Seo (1988) and Russell and Seo (1989) on simple representative utility functions for pairwise comparison based on lower-order SD rules.

In the special case of a constant $(N-1)$th order derivative, or $u^{N-1}(z_t) = c$ for $t = 1,\ldots,T$, we could set $\gamma_t = \beta_{N-1}$, for $t = 1,\ldots,T-1$, and find a $(N-1)$th order polynomial utility function

$$u(z_t) = \sum_{n=0}^{N-2}\beta_n(z_t - z_T)^n + \beta_{N-1}(z_t - z_T)^{N-1}. \quad (12)$$

Our approach differs from using this polynomial function in an important way. A polynomial generally does not obey the regularity conditions (restrictions on the signs of the derivatives) over the entire range. In addition, if we restrict the polynomial to obey the regularity conditions, then it generally loses its flexibility. Our approach in effect uses a local (rather than global) polynomial representation for the utility function. Since the NSD criterion restricts only the sign (and not the shape) of the Nth derivative, it allows a piecewise constant representation of the $(N-1)$th derivative and the piecewise polynomial representation of lower-order derivatives (but not for the global polynomial approximation).

The utility levels and marginal utility levels are linear in the $(N+T-1)$ parameters β_n, $n = 0,1,\ldots,N-2$, and γ_t, $t = 1,\ldots,T$. This finding implies that the inequalities (2), (3) and (5) are linear in these parameters, allowing for Linear Programming. Applying our theorem to the admissibility conditions (3), we find:

Corollary 1 (*Convex Stochastic Dominance*). *An evaluated prospect $i \in \{1,\ldots,M\}$ is admissible in terms of Nth order stochastic dominance, $N \geqslant 1$, relative to the set of prospects $\{1,\ldots,M\}$ if there exists a non-zero solution for the following system of inequalities:*

$$\sum_{s=1}^{S}\left(\sum_{n=0}^{N-2}\beta_n(y_s - y_S)^n + \sum_{k=s}^{S}\gamma_k(y_s - y_k)^{N-1}\right)(q_{i,s} - q_{j,s}) \geqslant 0,$$
$$j = 1,\ldots,M, \quad (13.1)$$

$$(-1)^n\beta_n \leqslant 0, \quad n = 1,\ldots,N-2, \quad (13.2)$$

$$(-1)^{N-1}\gamma_k \leqslant 0, \quad k = 1,2,\ldots,S-1, \quad (13.3a)$$
$$(-1)^{N-1}\gamma_S \leqslant 0 \text{ for } N \geqslant 2. \quad (13.3b)$$

We exclude zero solutions to exclude the trivial utility function $u(x) = 0\ \forall x \in D$, or an indifferent decision maker. A pair-wise NSD test arises as a special case with $M = 2$. Similarly, applying the theorem to the efficiency conditions (5), we find:

Corollary 2 (*Stochastic Dominance Efficiency*). *An evaluated combination of prospects $x^* \in X$ is efficient in terms of Nth order dominance, $N \geqslant 2$, relative to all feasible combinations $x \in X$ if and only if there exists a non-zero solution for the following system of inequalities:*

$$\sum_{r=1}^{R}\left(\sum_{n=1}^{N-2}n\beta_n\left(x_r^* - x_R^*\right)^{n-1} + (N-1)\sum_{k=r}^{R}\gamma_k\left(x_r^* - x_k^*\right)^{N-2}\right)(x_r^* - x_{j,r})p_r \geqslant 0,$$
$$j = 1,\ldots,M, \quad (14.1)$$

$$(-1)^n\beta_n \leqslant 0, \quad n = 1,\ldots,N-2, \quad (14.2)$$
$$(-1)^{N-1}\gamma_k \leqslant 0, \quad k = 1,2,\ldots,R. \quad (14.3)$$

For $N = 2$, these inequalities reduce to those underlying the Post (2003, Thm. 2) SSD test.

We can specify LP problems to test the systems of inequalities in (13) and (14). The specific objective function and standardization of the variables would depend on the specific application. Our empirical application in Section 4 will use the following LP tests for convex NSD ($N \geqslant 2$):

$$\theta^* = \min_{\beta_n,\gamma_k,\theta} \theta, \quad (15)$$

$$\text{s.t. } \sum_{s=1}^{S}\left(\sum_{n=0}^{N-2}\beta_n(y_s - y_S)^n + \sum_{k=s}^{S}\gamma_k(y_s - y_k)^{N-1}\right)(q_{i,s} - q_{j,s}) + \theta \geqslant 0,$$
$$j = 1,\ldots,M,$$

$$(-1)^n\beta_n \leqslant 0, \; n = 1,\ldots,N-2,$$
$$(-1)^{N-1}\gamma_k \leqslant 0, \; k = 1,2,\ldots,S,$$
$$\sum_{s=1}^{S}\left(\sum_{n=1}^{N-2}n\beta_n(y_s - y_S)^{n-1} + (N-1)\sum_{k=s}^{S}\gamma_k(y_s - y_k)^{N-2}\right)q_{i,s} = 1.$$

Parameter θ is an upper bound on the violations of the admissibility conditions. A value of $\theta^* \leqslant 0$ means that the evaluated prospect is admissible; a value of $\theta^* > 0$ means that it is non-admissible.[1] The last restriction on average marginal utility is a harmless standardization to avoid the trivial solution of an indifferent decision maker.

Similarly, our empirical application will use the following LP test for NSD efficiency:

$$\theta^* = \min_{\beta_n,\gamma_k,\theta} \theta, \quad (16)$$

$$\text{s.t. } \sum_{r=1}^{R}\left(\sum_{n=1}^{N-2}n\beta_n\left(x_r^* - x_R^*\right)^{n-1} + (N-1)\sum_{k=r}^{R}\gamma_k\left(x_r^* - x_k^*\right)^{N-2}\right)(x_r^* - x_{j,r})p_r$$
$$+ \theta \geqslant 0, \quad j = 1,\ldots,M,$$

$$(-1)^n\beta_n \leqslant 0, \; n = 1,\ldots,N-2,$$
$$(-1)^{N-1}\gamma_k \leqslant 0, \; k = 1,2,\ldots,R,$$
$$\sum_{r=1}^{R}\left(\sum_{n=1}^{N-2}n\beta_n\left(x_r^* - x_R^*\right)^{n-1} + (N-1)\sum_{k=r}^{R}\gamma_k\left(x_r^* - x_k^*\right)^{N-2}\right)p_r = 1.$$

Parameter θ now is an upper bound on the violations of the first-order conditions. A value of $\theta^* = 0$ means that the evaluated combination of prospects is efficient; a value of $\theta^* > 0$ means that it is inefficient.

The higher-order derivatives ($q \geqslant 2$) do not enter explicitly in convex NSD tests and NSD efficiency tests (regardless of the order N). Still, the higher-order derivatives may be useful to impose additional structure on the utility function, such as in Meyer (1977a,b) and Leshno and Levy (2002). For example, Meyer (1977a,b) bounds the coefficient of absolute risk aversion $a(x) = -u^2(x)/u^1(x)$ from below by a given function $f(x) \geqslant 0$ and from above by another given function $g(x) \geqslant 0$. We can impose these bounds by means of the following restrictions:

$$u^2(z_t) + f(z_t)u^1(z_t) \leqslant 0, \quad t = 1,\ldots,T, \quad (17)$$
$$u^2(z_t) + g(z_t)u^1(z_t) \geqslant 0, \quad t = 1,\ldots,T. \quad (18)$$

Since the first-order and second-order derivatives are linear functions of the β_n and γ_k parameters and $f(x)$ is exogenous, the restrictions are linear in the parameters. In a similar way, we could impose restrictions on, for example, the coefficient of relative prudence $b(x) = -u^3(x)x/u^2(x)$.

Similarly, the Almost Second-order Stochastic Dominance (ASSD) rule bounds the relative range of the second-order derivative $u^2(x)$ from above by the constant $\left[\frac{1}{\varepsilon} - 1\right]$, $\varepsilon \in (0, \frac{1}{2})$. Leshno and Levy (2002) present a closed-form solution for pair-wise comparison. The same restriction can be implemented for convex SD and SD efficiency tests by using the following linear restrictions:

[1] Values of $\theta^* < 0$ are not possible if the evaluated prospect is one of the M prospects. In this case, admissibility implies $\theta^* = 0$.

$$\delta \leqslant -u^2(z_t) \leqslant \delta\left[\frac{1}{\varepsilon}-1\right], \quad t=1,\ldots,T, \tag{19}$$

$$\delta \geqslant 0. \tag{20}$$

3. Dual formulation in terms of lower partial moments

The focus of this study is on utility functions and their derivatives and on restrictions that follow from utility theory. It is well known that SD criteria can also be formulated in terms of lower partial moments or related statistics such as cumulated distribution functions, quantiles and Gini coefficients. This section develops linear dual formulations of our utility-based tests in terms of lower partial moments (Bawa, 1975) and co-lower partial moments (Bawa and Lindenberg, 1977). The dual representation is less economically appealing than the utility representation, but it is often adopted in OR/MS for computational efficiency and robustness analysis. The analysis in this section allows for a direct comparison with and generalization of a range of earlier studies based on lower partial moments or related statistics.

We use the following definition for the nth order lower partial moment for prospect i and threshold value w:

$$LPM_i^n(w) = \sum_{r=1}^{R} p_r (w - x_{i,r})^n 1(x_{i,r} \leqslant w). \tag{21}$$

For analyzing combinations of prospects, we use the following definition of the nth order co-lower partial moment between a given combination $x \in X$ with weights $\lambda = (\lambda_1 \cdots \lambda_M)$ and another combination $x^* \in X$ with weights $\tau = (\tau_1 \cdots \tau_M)$:

$$coLPM_{\tau,\lambda}^n(w) = \sum_{r=1}^{R}\left(\sum_{j=1}^{M} x_{j,r}\lambda_j\right) p_r \left(w - \left(\sum_{j=1}^{M} x_{j,r}\tau_j\right)\right)^n 1\left(\sum_{j=1}^{M} x_{j,r}\tau_j \leqslant w\right). \tag{22}$$

Theorem 2 (*Dual Convex NSD Test*). *An evaluated prospect $i \in \{1,\ldots,M\}$ is non-admissible in terms of Nth order stochastic dominance, $N \geqslant 1$, relative to the set of prospects $\{1,\ldots,M\}$ if and only if there exists a solution for the following system of inequalities:*

$$\sum_{j=1}^{M}\lambda_j LPM_j^n(y_S) \leqslant LPM_i^n(y_S), \quad n=1,\ldots,N-2, \tag{23.1}$$

$$\sum_{j=1}^{M}\lambda_j LPM_j^{N-1}(y_k) \leqslant LPM_i^{N-1}(y_k), \quad k=1,2,\ldots,S-1, \tag{23.2}$$

$$\sum_{j=1}^{M}\lambda_j LPM_j^{N-1}(y_S) < LPM_i^{N-1}(y_S) \tag{23.3}$$

$$\sum_{j=1}^{M}\lambda_j = 1, \tag{23.4}$$

$$\lambda_j \geqslant 0, \quad j=1,\ldots,M. \tag{23.5}$$

(Proof in the Appendix)

Since the nth order LPM equals the n-times cumulative distribution function, this test is a direct generalization of Fishburn (1974, Thm. 2), Bawa et al. (1985, p. 423) and Levy (2006, p. 131).

For $N = 2$, we arrive at the linear, slack-variable formulation of convex SSD of Bawa et al. (1985, p. 423), safe some trivial differences between weak and strict inequalities related to our utility functions (1) allowing for zero values for the utility derivatives. Similarly, for $N = 3$, we arrive at the linear, super-convexity approximation to convex TSD of Bawa et al. (1985, Section B) for a given perturbation parameter $\varepsilon > 0$, after modifying (23.4) to $\sum_{j=1}^{M}\lambda_j = 1+\varepsilon$.

Finally, for $M = 2$, our test (23) reduces to the pair-wise NSD test presented in Levy (2006, p. 131). Our test however applies more generally for convex NSD, including, for example, convex FOSD for $M > 2$.

Theorem 3 (*Dual NSD Efficiency Test*). *An evaluated combination of prospects $x^* \in X$ with weights $\tau = (\tau_1 \cdots \tau_M)$ is inefficient in terms of Nth order dominance, $N \geqslant 2$, relative to all feasible combinations $x \in X$ if and only if there exists a solution for the following system of inequalities:*

$$coLPM_{\tau,\lambda}^n\left(\sum_{j=1}^{M} x_{j,R}\tau_j\right) \leqslant coLPM_{\tau,\tau}^n\left(\sum_{j=1}^{M} x_{j,R}\tau_j\right), \quad n=0,\ldots,N-3, \tag{24.1}$$

$$coLPM_{\tau,\lambda}^{N-2}\left(\sum_{j=1}^{M} x_{j,k}\tau_j\right) \leqslant coLPM_{\tau,\tau}^{N-2}\left(\sum_{j=1}^{M} x_{j,k}\tau_j\right), k=1,\ldots,R-1, \tag{24.2}$$

$$coLPM_{\tau,\lambda}^{N-2}\left(\sum_{j=1}^{M} x_{j,R}\tau_j\right) < coLPM_{\tau,\tau}^{N-2}\left(\sum_{j=1}^{M} x_{j,R}\tau_j\right), \tag{24.3}$$

$$\sum_{j=1}^{M}\lambda_j = 1, \tag{24.4}$$

$$\lambda_j \geqslant 0, \quad j=1,\ldots,M. \tag{24.5}$$

(Proof in the Appendix)

For $N = 2$, (24.1) disappears and (24.2)–(24.4), coincides with Post (2003, p. 1929) dual SSD test in terms of zero-order co-lower partial moments or 'ordered mean differences'.

We can specify LP problems to test the systems of inequalities in (23) and (24) by analogy to problems (15) and (16). As discussed in the introduction, we are rather skeptical about attempts to identify an alternative prospect that dominates the evaluated prospect. Areas where the dual formulation clearly adds value to the primal formulation are computational efficiency and robustness analysis. Fábián et al. (2011) and Gollmer et al. (2011) suggest algorithmic improvements for stochastic optimization problems with SSD ($N = 2$) constraints based on dual problem formulations. Applying their insights may also reduce the computational burden of our tests for large sample applications, which is particularly relevant when statistical resampling methods are used. Similarly, robustness analysis of dominance relationships traditionally focuses on the dual formulation; see, for example, Dentcheva and Ruszczyński (2010), Dupačová and Kopa (2012) or Liu and Xu (2013) for the case of SSD ($N = 2$).

4. Empirical application to US stock market data

We will now use a range of SD tests to analyze the efficiency of the broad stock market portfolio for various investment horizons. Our stock market portfolio is constructed as a value-weighted average of all NYSE, AMEX and NASDAQ stocks. It is compared with a standard set of 10 active benchmark stock portfolios that are formed, and annually rebalanced, based on individual stocks' market capitalization of equity (or 'size'), each representing a decile of the cross-section of stocks in a given year. Furthermore, we include the 1-month US Treasury bill as a riskless asset. We use data on monthly value-weighted returns (month-end to month-end) from July 1926 to December 2011 (1026 months) obtained from Kenneth French' data library. The size portfolios are of particular interest because a wealth of empirical research, starting with Banz (1981), suggests that small-cap stocks earn a return premium that defies rational explanation.[2]

SD analysis is invariant for positive linear transformations of the returns and does not require a specification of the initial wealth level. Nevertheless, it is useful for the SDWRF criterion to use gross holding period returns (HPRs), which are positive ($x_{min} > 0$) and proportional to final wealth (rather than the increase in wealth); the relevant restrictions on the ARA coefficient depend on the return definition. Our analysis does not use continuously compounded or log returns, because log returns generally are not proportional to final wealth, and, in addition, do not combine linearly in the cross-section. We do not object to assuming a log-normal distribution for some assets and horizons and we also do not object to logarithmic utility for some investors. SD analysis simply does not require such parametric specifications.

Monthly returns are commonplace in the empirical finance literature. However, a 1-month period may not be appropriate as the horizon of the representative investor. We therefore also consider returns for periods of 1 and 10 years. As discussed in Benartzi and Thaler (1995), a period of 1 year seems most plausible as the relevant evaluation horizon, because most financial reporting takes place on an annual basis (for example, financial statements, tax files and updates of retirement accounts). To represent long-term investors, our analysis also includes 10-year returns. Whereas the benchmark portfolios are annually rebalanced, our analysis fixes the investor's allocation across the benchmark portfolios during the investment period. For a long-term investor who periodically adjusts her asset allocation and style mix, a dynamic programming model may therefore be more appropriate.

Common problems in the analysis of long-term return are a limited number of non-overlapping return intervals and a possible sensitivity to the specification of the starting month and year. We therefore focus on the HPRs for all 1015 sub-periods of 12 sequential months and all 907 sub-periods of 120 sequential months. This approach preserves possible auto-correlation in the monthly data and therefore leads to more realistic long-term HPR scenarios than random simulation. Nevertheless, our conclusions are robust to using non-overlapping long-term HPRs for any starting month and year, and for simulated long-term HPRs based on independent random draws of 12 or 120 monthly returns (with or without replacement).

Table 1 shows descriptive statistics for the excess returns of the relevant portfolios. We generally recommend to measure systematic risk based on co-lower partial moments rather than co-variance. Still, the descriptives include the classical market beta, because of its familiarity, and because it helps to interpret the M–V results. Not surprisingly, small-cap stocks tend to have a higher standard deviation, market beta and skewness than large-cap stocks. Interestingly, the market portfolio has less skewness than nine of the 10 benchmark stock portfolios. Apparently, broad diversification yields a relatively small reduction in downside risk (relative to the more concentrated size-decile portfolios) at the cost of a relatively large reduction in upside potential. However, the differences become smaller for long horizons. The market skewness increases with the horizon, reflecting the effect of compounding returns. By contrast, the returns to small caps become less skewed and have lower betas for longer horizons, presumably reflecting the effect of long-term mean reversion.

In this application, it seems natural to test whether the market portfolio is efficient relative to all convex combinations of the 11 base assets, so as to allow for portfolio diversification (but without short selling). Nevertheless, it is insightful to also apply pair-wise dominance and convex dominance relative to the undiversified base assets. For each of the 11 base assets, we will apply LP test (15) to analyze whether the market portfolio is pair-wise dominated by the base asset ($M = 2$). A test for admissibility applies LP test (15) to compare the market portfolio with the 11 base assets simultaneously ($M = 11$). To test whether the market portfolio is efficient, or not dominated by any convex combination of the base assets, we use LP test (16).

All tests use equal weights for the T historical observations ($p_t = 1/T$). Since we analyze the market portfolio, the marginal utility function can be interpreted as a pricing kernel and the violations of the first-order conditions as pricing errors, or 'alphas'. The objective function is the largest positive alpha, because this term represents a deviation from optimality even in case of binding short-sales constraints, making the test more general than one based on, for example, the sum of squared alphas or the mean absolute alpha. After all, a large negative alpha generally offers only limited profit opportunity without short selling. The standardization in (15) and (16) follows a convention in the asset pricing literature to set the average value of the pricing kernel equal to unity.

We implement the SSD, TSD and FOSD criteria by setting $N = 2$, 3, 4, respectively, in (15) and (16). We also perform SDWRF tests based on ARA restrictions (17) and (18) with lower bound $f(x) = 0$ and upper bound $g(x) = 3x^{-1}$. This means that the coefficient of relative risk aversion (RRA) $r(x) = a(x)x$ is bounded from above by the value of 3. For the utility of wealth, as opposed to the utility of consumption, there exist compelling arguments for an average RRA value close to one and slightly increasing; see for example, Meyer and Meyer (2005). Since we analyze gross returns, which are proportional to final wealth, a value of $r(x) = 3$ seems relatively high.[3] We also implement ASSD tests using restrictions (19) and (20) on utility curvature. Following Levy et al. (2010) and Bali et al. (2009), we use critical value of $\varepsilon = 0.032$. Finally, we include tests based on the mean–variance (M–V) criterion, using a linear marginal utility function (or quadratic utility). For the sake of brevity, we will focus on the primal problem solution and results here and omit the dual problem solution.

Since we use empirical returns that are generated by an unobserved population distribution, we must account for the effect of sampling error on our test results. Unfortunately, the sampling distribution appears analytically intractable due to the large number of inequality restrictions involved. Fortunately, statistical re-sampling methods can overcome analytical intractability using brute computational force. An early study by Nelson and Pope (1991) demonstrated that SD analysis based on a bootstrapped return distribution is more powerful than analysis based on the original empirical return distribution. More recently, Barrett and Donald (2003) and Linton et al. (2005) derive powerful consistent bootstrap and sub-sampling tests for pair-wise comparisons.

Interestingly, the tractable LP structure of (15) and (16) suggests that the computational burden of re-sampling is manageable also for convex SD tests and SD efficiency tests. Under the assumption of identically and independently distributed time-series returns, the empirical return distribution is a consistent estimator

[2] We arrive at similar conclusions using benchmark portfolios formed on stocks' book-to-market equity ratio or prior 12 – 2 month returns (momentum) and/or using a sample that starts in July 1963, a popular breakpoint in the empirical asset pricing literature. An exception occurs for 10-year returns to the size portfolios (but not for book-to-market and R12 – 2 portfolios). The recent sample of 10-year returns assigns a relatively large weight to the extraordinary late-1990s large-cap stock market rally, which has the effect of increasing the market beta of large-caps above that of small-caps. Therefore, risk aversion in effect penalizes large caps and the lowest alphas are obtained with a (nearly) risk-neutral pricing kernel for every efficiency criterion.

[3] For the utility of consumption rather than the utility of wealth, higher RRA values may be required to be consistent with the historical equity premium; see, for example, Mehra and Prescott (1985).

Table 1

Descriptive statistics size portfolios. The table shows descriptive statistics for gross holding-period returns to the 10 size benchmark portfolios and the market portfolio. The benchmark portfolios are based on individual stocks' market capitalization of equity, and each represent a value-weighted average of a decile of the cross-section of stocks (using NYSE size break points). The stock market portfolio is a value-weighted average of all NYSE, AMEX and NASDAQ stocks. The sample period ranges from July 1926 to December 2011 (1026 months), and net returns are computed in excess of the 1-month T-bill. The raw month-end-to-month-end returns are taken from Kenneth French' data library. The long-term returns are multiplicative returns for all holding periods of 12 or 120 sequential monthly returns in the original sample. We compute the total HPR for an investment period of H months up to and including month t from the individual monthly gross returns in excess of the riskless rate: $x_t^{(H)} = \prod_{s=1}^{H}(1 + x_{t-H+s} - r_{t-H+s})$. These HPRs express final investment wealth (after subtracting funding costs) at the end of the investment period as a percentage of initial wealth.

Portf	1 month				12 months				120 months			
	Avg.	Stdev.	Skew	Beta	Avg.	Stdev.	Skew	Beta	Avg.	Stdev.	Skew	Beta
Small	101.14	10.23	3.70	1.44	116.45	49.50	3.41	1.86	333.30	284.90	2.59	1.53
2	100.96	8.94	2.24	1.39	113.38	40.88	3.08	1.63	275.70	173.00	1.47	1.22
3	100.96	8.18	1.92	1.34	112.91	36.28	3.14	1.49	275.60	152.60	1.26	1.13
4	100.92	7.57	1.54	1.26	112.14	32.30	2.30	1.36	268.60	145.80	1.14	1.09
5	100.87	7.26	1.14	1.24	111.46	29.80	2.00	1.29	258.30	123.50	0.83	1.02
6	100.87	6.94	1.01	1.21	111.41	28.80	1.77	1.26	252.20	125.20	0.85	1.07
7	100.82	6.57	0.80	1.16	110.84	27.15	1.35	1.21	248.60	117.20	0.86	1.06
8	100.76	6.22	0.74	1.11	109.85	25.03	1.47	1.11	230.80	104.00	0.78	1.00
9	100.70	5.93	0.56	1.07	109.20	23.58	0.84	1.08	226.20	108.70	0.69	1.05
Large	100.57	5.13	0.09	0.93	107.31	20.13	0.12	0.92	201.90	104.90	0.69	1.01
Mkt	100.62	5.46	0.17	1.00	108.03	21.53	0.37	1.00	207.90	101.70	0.77	1.00

of the population return distribution, and bootstrap samples can simply be obtained by random sampling with replacement from the empirical return distribution. To ensure that the population distribution of the bootstrap samples obeys the null hypothesis of market portfolio efficiency, we first re-center the empirical distribution by correcting the original time-series of returns for a given base asset by subtracting the estimated alpha of that base asset. While this adjustment aligns the assets' means with the null hypothesis, it does not affect the general risk levels and the dependence structure between the assets. We implement this bootstrap method by generating 10,000 pseudo-samples of the same size as the original sample through random draws with replacement from the re-centered original sample, and test market portfolio efficiency in every pseudo-sample. Since the long-term HPRs were constructed from historical sequences of 12 or 120 monthly returns, the method in effect is a block bootstrap applied to random blocks of the same length. Finally, we compute the critical values for the original test-statistics from the percentiles of the bootstrap distribution.

Table 2 shows the test results for the different decision criteria and investment horizons. For the sake of comparability, we multiply the monthly results by 12 and divide the 10-year results by 10, to arrive at 'annualized' results. This method obviously does not account for compounding effects, but it does preserve the relative differences between the portfolios.

Testing whether the market portfolio is dominated by one of the base assets in a pair-wise fashion using any of the six decision criteria has limited discriminating power in this study. The market portfolio has a relatively low risk level and is not dominated by any of the more risky size deciles 1–9 using any of the criteria and horizons. At the same time, the market portfolio has a higher average return than the large-cap stock portfolios and the T-bill and is also not dominated by these alternatives. As discussed above, the pair-wise approach ignores that the market portfolio can be non-admissible without being dominated. Different investors may agree that the market portfolio is not optimal, even if they do not agree on which base assets are better than the market portfolio; they may see improvement possibilities in different base assets. The violations of admissibility are however also relatively small and not statistically significant. Presumably, the admissibility test has limited power in this application, because it overlooks the effects of diversification across the base assets. Indeed, the efficiency test detects economically and statistically significant violations of market portfolio efficiency. In the remainder of this section, we focus on these tests results in more detail. Table 3 shows the alphas for the individual size decile portfolios and Fig. 1 shows the associated pricing kernels.

For monthly returns, M–V analysis classifies the market portfolio as significantly inefficiently due to a substantial undervaluation of small-cap stocks, confirming known empirical results. For the first decile portfolio, the alpha is 2.05% per annum, and highly significant. The SSD criterion leads to a large reduction of the alphas by using a step function for the pricing kernel (see Fig. 1); the alpha for small-cap stocks falls to 0.71% per annum. The kernel has large concave (!) segments and it penalizes small-cap stocks for having a relatively high positive skewness. Clearly, this is not consistent with prudence (or skewness preference). The higher-order SD tests do not allow for this pattern and their results cannot be distinguished from the M–V results (a linear kernel) in this case. The SDWRF kernel is convex and places a reward on positive skewness; the SDWRF alpha for small-cap stocks increases to 2.40% per annum. By contrast, the ASSD test does not remedy the problem of skewness aversion and yields similar results as the SSD test. Arguably, the parameter value $\varepsilon = 0.032$ is too high for the relatively narrow range of monthly returns.

The situation is remarkably different for annual returns. In this case, the M–V kernel becomes negative (!) for gross returns in excess of about 160% (or net return of 60% per annum). While such cases represent only a few percent of the total number of annual observations, they do represent a substantial part of the total return of in the sample period and can have a large effect on the estimated alphas. The M–V alpha for the first decile portfolio is only 0.99% per annum and is not statistically significant. This number seems to underestimate the appeal of small-cap stocks, because M–V analysis penalizes these stocks for having more systematic upside potential than large-cap stocks do.

The SD efficiency tests impose non-satiation and avoid negative weights. However, the SSD test again penalizes small-cap stocks for their relatively high skewness. The higher-order tests impose both non-satiation (violated by M–V analysis) and prudence (violated by the SSD test). The TSD kernel is a convex two-piece linear function with a kink at a net market return of about 50% and it generates an alpha of 1.62% per annum for the first decile portfolio. This value still seems to underestimate the appeal of small-cap stocks, because the kernel displays a discontinuous drop in its slope for high-return levels, violating the assumption of temperance. The FOSD efficiency test corrects for

Table 2

Dominance, admissibility and efficiency test statistics. The table shows the (annualized) test statistics obtained by testing for dominance, admissibility and efficiency of the value-weighted market portfolio relative to the 10 size-decile portfolios and the 1-month T-bill. We use six different efficiency criteria: (i) mean–variance (M–V) analysis, (ii) second-order SD, (iii) third-order SD, (iv) fourth-order SD, (v) SD With Respect to a Function (SDWRF) and (vi) Almost SSD. Separate results are shown for monthly, annual and 10-year gross holding-period returns in the sample period from July 1926 to December 2011. For each of the 11 base assets, we analyze whether the market portfolio is pair-wise dominated using test (9) for the market portfolio and the base asset ($M = 2$). We test admissibility using (9) by including all 11 base assets simultaneously ($M = 12$). Test (10) is used to determine whether the market portfolio is efficient relative to all convex combinations of the 11 base assets.

Horizon	Criterion	Dominance by individual benchmark portfolios										Admissibility	Efficiency
		Smll	2	3	4	5	6	7	8	9	Lrg		
1 month	M–V	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.05**
	SSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71
	TSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.05**
	FOSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.51**
	SDWRF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.40**
	ASSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.26*
12 months	M–V	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99
	SSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.08
	TSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	1.62*
	FOSD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	1.99**
	SDWRF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.29**
	ASSD	0.00	0.00	0.04	0.07	0.06	0.09	0.07	0.07	0.04	0.00	0.09	1.51*
120 months	M–V	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.75**
	SSD	0.05	0.00	0.00	0.05	0.00	0.05	0.00	0.05	0.00	0.00	0.05	2.51**
	TSD	0.03	0.00	0.00	0.03	0.00	0.03	0.00	0.03	0.00	0.00	0.03	2.71**
	FOSD	0.02	0.00	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.00	0.02	2.71**
	SDWRF	0.05	0.00	0.00	0.05	0.00	0.05	0.00	0.05	0.00	0.00	0.05	6.07**
	ASSD	1.78*	1.07*	1.10*	1.00	0.85	0.74	0.69	0.40	0.31	0.00	1.78*	7.82***

* Indicates that the bootstrap p-value is smaller than 10%.
** Indicates that the bootstrap p-value is smaller than 5%.
*** Indicates that the bootstrap p-value is smaller than 1%.

Table 3

Pricing errors. The table shows the (annualized) alphas obtained by testing efficiency of the value-weighted market portfolio relative to the 10 size-decile portfolios and the 1-month T-bill. We use six different efficiency criteria: (i) mean–variance (M–V) analysis, (ii) second-order SD, (iii) third-order SD, (iv) fourth-order SD, (v) SD With Respect to a Function (SDWRF) and (vi) Almost SSD. Separate results are shown for monthly, annual and 10-year gross holding-period returns in the sample period from July 1926 to December 2011. The results for a monthly period are based on a re-centered bootstrap that generates 10,000 pseudo-samples of T returns by means of random sampling with replacement from the re-centered original sample of monthly returns. The results for an annual (10-year) period are based on a re-centered block bootstrap that generates 10,000 pseudo-samples of T random blocks of 12 (120) sequential monthly returns from the re-centered original sample of monthly returns. T is the size of the original sample of monthly, annual or 10-year returns.

Portfolio											
Horizon	Criterion	Small	2	3	4	5	6	7	8	9	Large
1 month	M–V	2.05**	0.44	0.92	1.16	0.79	1.05	0.92	0.65	0.39	0.09
	SSD	0.71	−0.44	0.26	0.71	0.37	0.71	0.71	0.61	0.30	0.12
	TSD	2.05**	0.44	0.92	1.16	0.80	1.05	0.93	0.66	0.39	0.09
	FOSD	2.51**	0.61	1.10	1.32	0.89	1.17	1.01	0.74	0.46	0.10
	SDWRF	2.40**	0.52	1.05	1.25*	0.86	1.13	0.98	0.71	0.46	0.11
	ASSD	1.26*	−0.05	0.59	0.96	0.57	0.90	0.85	0.66	0.37	0.08
12 months	M–V	0.67	−0.37	0.45	0.86	0.83	0.99	0.94	0.79	0.49	0.01
	SSD	0.89	−0.10	0.94	0.99	1.08	1.08	1.02	1.02	0.66	0.00
	TSD	1.62*	0.36	1.18*	1.34*	1.23*	1.30*	1.13*	1.06	0.61	−0.02
	FOSD	1.99**	0.57	1.29*	1.46*	1.25*	1.36*	1.15	1.06	0.61	−0.02
	SDWRF	2.29**	0.76	1.42*	1.53*	1.27*	1.41*	1.14*	1.11	0.72	−0.01
	ASSD	1.51*	0.30	1.17*	1.26*	1.24*	1.25*	1.09	1.07	0.68	−0.04
120 months	M–V	4.28***	3.41***	4.75***	4.65***	4.68***	3.33***	3.08***	2.37**	1.00	−0.76
	SSD	2.51**	0.10	0.73	2.08*	0.78	1.21	0.48	0.95	0.12	0.16
	TSD	2.71**	0.50	1.22	2.61	1.39	1.57	0.60	1.22	−0.09	0.08
	FOSD	2.71**	0.50	1.22	2.61	1.39	1.57	0.60	1.22	−0.09	0.08
	SDWRF	6.07***	3.06***	3.69***	3.81***	3.21***	2.63**	2.33**	1.92**	0.87	−0.64
	ASSD	7.82***	4.24***	4.78***	4.66***	4.01***	3.25***	2.96**	2.13**	1.13	−0.73

* Indicates that the bootstrap p-value is smaller than 10%.
** Indicates that the bootstrap p-value is smaller than 5%.
*** Indicates that the bootstrap p-value is smaller than 1%.

this problem and yields a kernel that is consistent with all four assumptions: non-satiation, risk aversion, prudence and temperance. The FOSD alpha for the first decile portfolio is 1.99% per annum, far exceeding the M–V estimate in terms of economic and statistical significance. The SDWRF and ASSD tests yield similar results as the FOSD and TSD tests, respectively.

Due to the multiplicative nature of HPRs, the average return and standard deviation of small caps increase faster than those of large

Please cite this article in press as: Post, T., Kopa, M. General linear formulations of stochastic dominance criteria. European Journal of Operational Research (2013), http://dx.doi.org/10.1016/j.ejor.2013.04.015

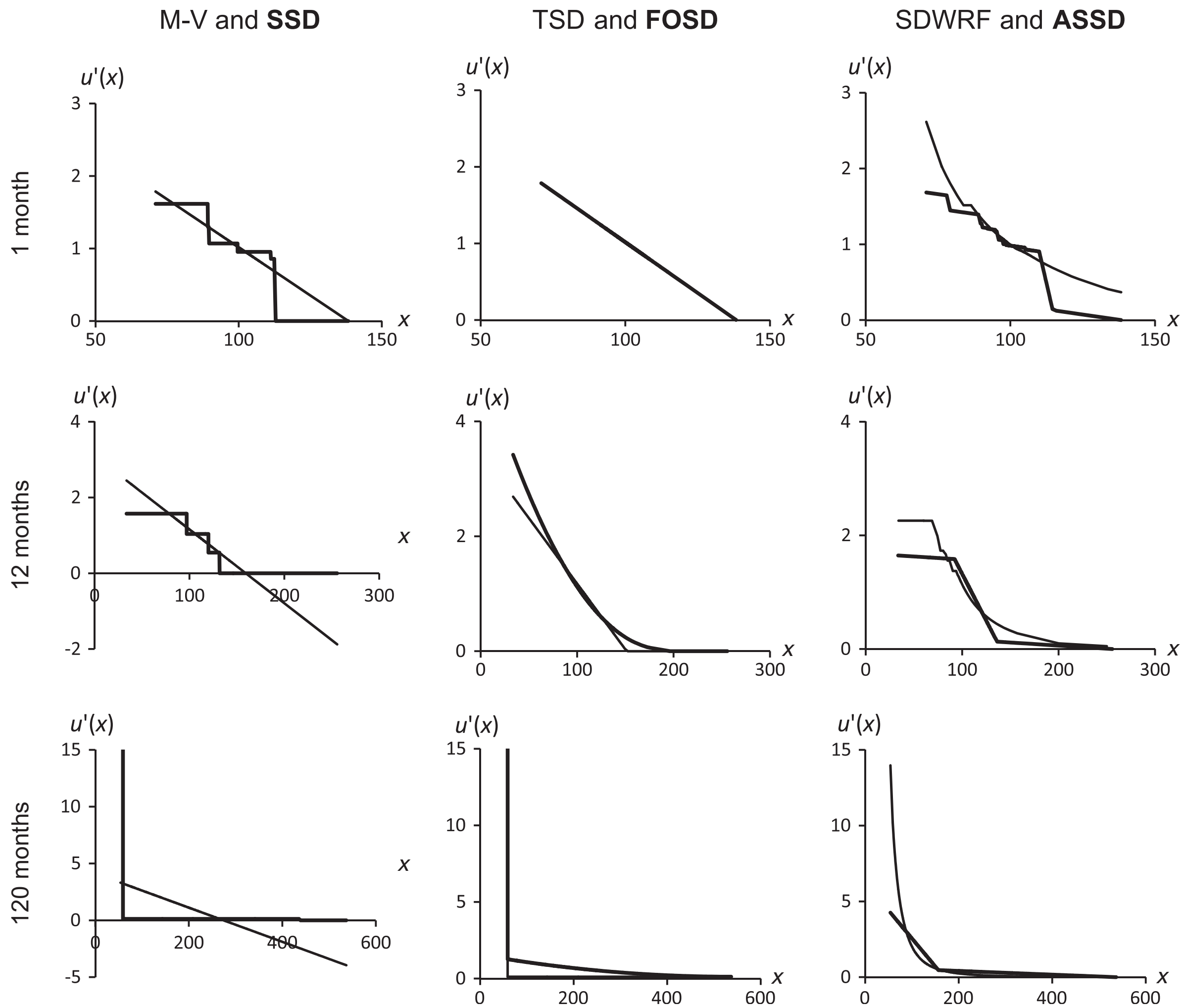


Fig. 1. Pricing kernels. The graphs shows the optimal marginal utility functions, or pricing kernels, obtained by testing for efficiency of the market portfolio relative to the 10 size-decile portfolios and the riskless 1-month Treasury bill. Results are shown for gross holding period returns for all 1-month, 1-year and 10-year holding periods in the original sample period from July 1926 to December 2011. The SSD, TSD and FOSD results are based on LP problem (16) using $N = 2, 3, 4$, respectively. The mean–variance (M–V) efficiency test assumes a quadratic utility function. SDWRF results use ARA restrictions based on lower bound $f(x) = 0$ and upper bound $g(x) = 3x^{-1}$; the ASSD results use $\varepsilon = 0.032$. The tests identify marginal utility level for a large number of different market return levels. A continuous graph is obtained by means of linear interpolation. The graphs are shown only for the observed range of market returns.

caps as the investment horizon increases.[4] However, the market beta of small-caps actually decreases at long horizons, presumably reflecting long-term mean-reversion. Related to this, the M–V alpha for the small-caps portfolio increases to 4.75% per annum for a 10-year horizon, economically and statistically highly significant. Fig. 1 shows that the M–V pricing kernel takes an alarming shape for this horizon. Specifically, the kernel becomes negative already for gross 10-year returns in excess of about 270% (or net return of 17% per annum) and its values range from −3.95 to 3.31 in this sample. This means that the M–V criterion in effect penalizes small caps for their systematic upside potential and the M–V alpha may still underestimate the true long-term appeal of these stocks.

The SD tests yield different results. The SSD kernel assigns extremely large weights (in excess of 250!) to the largest negative market returns. This weighting scheme penalizes small caps for their downside risk and reduces their alphas to economically and statistically less significant levels. Specifically, the average return difference between small caps and large caps during bear markets is smaller than the unconditional average difference, presumably due an overall increase in correlation between stocks during bear markets. The SSD criterion reduces the alphas by assigning almost all weight to the worst market returns. The SSD kernel in this case is almost convex and the TSD and FOSD criteria yield similar results.

The assumed preference structure however seems not representative for most risk averters. The SDWRF criterion ignores these preferences and uses a more moderate weighting scheme by requiring the RRA coefficient to take values between 0 and 3. The SDWRF alpha for small caps is as high as 6.07% per annum, substantially higher than the M–V and SSD values. These large differences arise because SDWRF avoids negative weights for the right tail of the market return distribution and large negative weights for the left tail. The 10-year SDWRF results are confirmed by the ASSD test, which places a cap on the relative range of utility curvature. In fact, the ASSD alphas are so large in this case that we may question whether the parameter value

[4] Investment risk increases at a slower rate when using annualized log returns, confirming known results by Fama and French (1988) and Poterba and Summers (1988), among others. This approach would however lead to spurious 'time-diversification' effects in the context of our study, as we assume that the investor maximizes the expected utility of her wealth at the end of a 10-year period.

Please cite this article in press as: Post, T., Kopa, M. General linear formulations of stochastic dominance criteria. European Journal of Operational Research (2013), http://dx.doi.org/10.1016/j.ejor.2013.04.015

of $\varepsilon = 0.032$ may be too low for the wide return range of long-term returns.

5. Concluding remarks

We may formulate stochastic dominance criteria for discrete probability distributions using a piece-wise polynomial representation of utility and its derivatives. This approach applies generally for higher-order SD rules and can also deal with additional preference restrictions such as the SDWRF bounds on the level of risk aversion and the ASSD bounds on utility curvature. The approach allows for comparing a given prospect (or combination of prospects) with a discrete set of prospects but also for comparison with all linear combinations of a set of prospects. The approach can be implemented by solving a relatively small system of linear inequalities by means of Linear Programming. A linear dual formulation uses lower partial moments or co-lower partial moments.

Our empirical application suggests that the passive stock market portfolio is highly inefficient relative to actively managed portfolios for all horizons for nearly all investors. It appears impossible to rationalize the market portfolio for any investment horizon without allowing for implausible shapes of the utility function.

Pair-wise dominance tests appear too weak to generate plausible results in this study. Despite their very large alphas, small-cap stocks do not dominate the market portfolio, for the simple reason that they are more risky than the market portfolio. A joint test for convex dominance relative to a small caps, large caps and T-bills also has limited power, as it overlooks the benefits of diversification across these market segments. The SSD criterion appears too weak to generate plausible results. For short-term returns, it penalizes small-cap stocks for having a relatively high positive skewness, violating prudence. The other decision criteria avoid this pattern and show that properly accounting for skewness and kurtosis lowers the level of market portfolio efficiency. An investor who looks for short-term downside protection and/or upside potential will find the market portfolio less appealing than a more concentrated position in small-cap stocks (possibly combined with T-bills to achieve the same standard deviation).

The results also support the hypothesis that M–V analysis gives a good second-order approximation for any well-be, haved utility function on the typical range of short-term return for diversified portfolios (see, for example, Levy and Markowitz, 1979). The approximation however breaks down for longer investment horizons. In this case, the M–V criterion assigns negative weights to large positive market returns, hence placing a penalty on the systematic upside potential of active investment strategies. As a result, it can underestimate market portfolio inefficiency and the appeal of active strategies for longer investment horizons.

Overall, the TSD and FOSD results appear more plausible than the M–V and SSD results. However, for long-term returns, SSD, TSD and FOSD all assign extremely large positive weights to large negative market returns, hence placing a large penalty on the systematic downside risk of active strategies. The emphasis on the left tail seems to reflect the elevated correlation between stocks during bear markets rather than an extreme aversion to tail risk. The SDWRF and ASSD rules can avoid this type of over-fitting by limiting the level of risk aversion.

Undoubtedly, parts of our alphas reflect market micro-structure issues regarding liquidity and transactions costs, and the appeal of active strategies diminishes without the professional trading facilities available to specialized investment companies. Notwithstanding these possible effects, we conclude that risk definitions and risk preferences are unlikely explanations for the high average returns of small caps (and similar results are found for value stocks and past winners). The M–V and SSD criteria can place implausible weights on the systematic downside risk and systematic upside potential of active investment strategies. However, our results show that using positive and moderate weights for all scenarios inevitably leads to the conclusion that the market portfolio is not optimal for all horizons and nearly all investors.

Acknowledgments

The research was partially supported by Czech Science Foundation (Grant P402/12/G097). We thank Turan Bali, Garry Barrett, Enrico De Giorgi, Haim Levy, Jack Meyer, Valerio Poti and Pim van Vliet for providing valuable comments and suggestions to this study. Any remaining errors are our own.

Appendix A

Proof of Theorem 1. [Linearization of Marginal Utility] For $N = 1$, set $\gamma_k = u(z_k) - u(z_{k+1})$, $k = 1, 2, \ldots, T-1$, $\gamma_T = u(z_T)$, and the theorem follows directly. We therefore focus on $N \geqslant 2$ below. Since the first N derivatives of $u \in U_N$ exist, we may apply a Taylor expansion about z_T at point z_t:

$$u(z_t) = \sum_{n=0}^{N-2} \frac{u^n(z_T)}{n!}(z_t - z_T)^n + RT,$$

with the following reminder term (RT) in the integral form:

$$RT = -\int_{z_t}^{z_T} \frac{u^{N-1}(t)}{(N-2)!}(z_t - t)^{N-2}\, dt.$$

Splitting the integral from z_t to z_T in smaller integrals and using the first mean value theorem for integration we find:

$$\begin{aligned}
RT &= -\sum_{k=t}^{T-1} \int_{z_k}^{z_{k+1}} \frac{u^{N-1}(t)}{(N-2)!}(z_t - t)^{N-2}\, dt \\
&= -\sum_{k=t}^{T-1} \frac{u^{N-1}(z_k^*)}{(N-2)!} \int_{z_k}^{z_{k+1}} (z_t - t)^{N-2}\, dt = \\
&= \sum_{k=t}^{T-1} \frac{u^{N-1}(z_k^*)}{(N-1)!}\left[(z_t - z_{k+1})^{N-1} - (z_t - z_k)^{N-1}\right] \\
&= \sum_{k=t+1}^{T-1}\left[\frac{u^{N-1}(z_{k-1}^*) - u^{N-1}(z_k^*)}{(N-1)!}\right](z_t - z_k)^{N-1} + \frac{u^{N-1}(z_{T-1}^*)}{(N-1)!}(z_t - z_T)^{N-1}
\end{aligned}$$

for some $z_k^* \in [z_k, z_{k+1}]$. Setting

$$\begin{aligned}
\beta_n &= \frac{u^n(z_T)}{n!}, \quad n = 0, 1, \ldots, N-2 \\
\gamma_k &= \frac{u^{N-1}(z_{k-1}^*) - u^{N-1}(z_k^*)}{(N-1)!}, \quad k = 2, \ldots, T-1, \quad \gamma_T = \frac{u^{N-1}(z_{T-1}^*)}{(N-1)!},
\end{aligned}$$

we obtain (6) and (8), (9). The conditions (10) and (11) on parameters β_n, γ_k follow directly from the definition of set U_N. Finally, for given parameters β_n, γ_k satisfying (10) and (11), one can easily verify that the piece-wise polynomial utility function:

$$u(x) = \sum_{n=0}^{N-2} \beta_n (x - z_T)^n + \sum_{k=t+1}^{T} \gamma_k (x - z_k)^{N-1}, \quad x \in [z_t, z_{t+1}],$$

$$t = 1, 2, \ldots, T-1$$

is admissible, that is, $u(x) \in U_N$. □

Proof of Theorem 2. [Dual Convex NSD Test] Using $b_n = -(-1)^n \beta_n$, $n = 1, 2, \ldots, N-2$ and $c_k = -(-1)^{N-1} \gamma_k$, $k = 1, 2, \ldots, S$, the system (13.1)–(13.3) can be rewritten as follows:

$$\sum_{s=1}^{S}(q_{j,s}-q_{i,s})\left[\sum_{n=1}^{N-2}b_n(y_S-y_s)^n+\sum_{k=s}^{S}c_k(y_k-y_s)^{N-1}\right]\geqslant 0,$$
$$j=1,\ldots,M, \tag{25.1}$$

$$b_n \geqslant 0, \quad n=1,2,\ldots,N-2, \tag{25.2}$$
$$c_k \geqslant 0, \quad k=1,2,\ldots,S-1, \tag{25.3}$$
$$c_S = 1, \tag{25.4}$$

where (25.4) is a harmless standardization to exclude zero solutions. Since

$$\sum_{s=1}^{S}(q_{j,s}-q_{i,s})\sum_{k=s}^{S}c_k(y_k-y_s)^{N-1}=\sum_{k=1}^{S}c_k\sum_{s=1}^{k}(q_{j,s}-q_{i,s})(y_k-y_s)^{N-1},$$

the system (25.1)–(25.4) can be tested using the following linear programming problem:

$$\theta^*=\min_{b_n,c_k,\theta}\theta,$$

$$\text{s.t.}\ \sum_{s=1}^{S}(q_{j,s}-q_{i,s})\sum_{n=1}^{N-2}b_n(y_S-y_s)^n+\sum_{k=1}^{S}c_k\sum_{s=1}^{k}(q_{j,s}-q_{i,s})(y_k-y_s)^{N-1}$$
$$+\theta\geqslant 0, \quad j=1,\ldots,M,$$
$$b_n \geqslant 0, \quad n=1,2,\ldots,N-2,$$
$$c_k \geqslant 0, \quad k=1,2,\ldots,S-1,$$
$$c_S=1,$$

where the ith prospect is NSD admissible if and only if the optimal value of objective function is equal to zero. The linear programming dual follows:

$$\max_{\lambda_j,\upsilon}\ \upsilon, \tag{26}$$

$$\sum_{j=1}^{M}\sum_{s=1}^{S}(\lambda_j q_{j,s}-q_{i,s})(y_S-y_s)^n\leqslant 0, \quad n=1,2,\ldots,N-2,$$
$$\sum_{j=1}^{M}\sum_{s=1}^{k}(\lambda_j q_{j,s}-q_{i,s})(y_k-y_s)^{N-1}\leqslant 0, \quad k=1,2,\ldots,S-1,$$
$$\sum_{j=1}^{M}\sum_{s=1}^{S}(\lambda_j q_{j,s}-q_{i,s})(y_S-y_s)^{N-1}+\upsilon\leqslant 0,$$
$$\sum_{j=1}^{M}\lambda_j=1, \quad \lambda_j\geqslant 0, \quad j=1,2,\ldots,M.$$

Using (21) for $w = y_k$, $k = 1,2,\ldots,S$, we can rewrite (26) in the form of the system of inequalities (23.1)–(23.5), because only the sign of the optimal objective value is of interest. □

Proof of Theorem 3. [Dual NSD Efficiency Test] Similar to the proof of Theorem 2, setting $b_n = n(-1)^n\beta_n$, $n = 1,2,\ldots,N-2$ and $c_k = (N-1)(-1)^{N-1}\gamma_{N-1}$, $k = 1,2,\ldots,R$, we can test the system (14.1)–(14.3) using the following linear programming problem:

$$\theta^*=\min_{b_n,c_k,\theta}\theta,$$

$$\text{s.t.}\ \sum_{n=1}^{N-2}b_n\sum_{r=1}^{R}p_r\left(\sum_{j=1}^{M}x_{j,r}\tau_j-x_{j,r}\right)\left(\sum_{j=1}^{M}x_{j,R}\tau_j-\sum_{j=1}^{M}x_{j,r}\tau_j\right)^{n-1},$$
$$+\sum_{k=1}^{R}c_k\sum_{r=1}^{k}p_r\left(\sum_{j=1}^{M}x_{j,r}\tau_j-x_{j,r}\right)\left(\sum_{j=1}^{M}x_{j,k}\tau_j-\sum_{j=1}^{M}x_{j,r}\tau_j\right)^{N-2}+\theta\geqslant 0,$$
$$j=1,\ldots,M,$$

$$b_n \geqslant 0, \quad n=1,2,\ldots,N-2,$$
$$c_k \geqslant 0, \quad k=1,2,\ldots,R-1,$$
$$c_R=1,$$

where the evaluated prospect is NSD efficient if and only if the optimal value of objective function is equal to zero. The linear programming dual follows:

$$\max_{\lambda_j,\upsilon}\upsilon,$$

$$\text{s.t.}\sum_{r=1}^{R}p_r\left(\sum_{j=1}^{M}x_{j,r}\tau_j-\sum_{j=1}^{M}x_{j,r}\lambda_j\right)\left(\sum_{j=1}^{M}x_{j,R}\tau_j-\sum_{j=1}^{M}x_{j,r}\tau_j\right)^{n-1}\leqslant 0,$$
$$n=1,2,\ldots,N-2,$$
$$\sum_{r=1}^{k}p_r\left(\sum_{j=1}^{M}x_{j,r}\tau_j-\sum_{j=1}^{M}x_{j,r}\lambda_j\right)\left(\sum_{j=1}^{M}x_{j,k}\tau_j-\sum_{j=1}^{M}x_{j,r}\tau_j\right)^{N-2}\leqslant 0,$$
$$k=1,2,\ldots,R-1,$$
$$\sum_{r=1}^{R}p_r\left(\sum_{j=1}^{M}x_{j,r}\tau_j-\sum_{j=1}^{M}x_{j,r}\lambda_j\right)\left(\sum_{j=1}^{M}x_{j,R}\tau_j-\sum_{j=1}^{M}x_{j,r}\tau_j\right)^{N-2}+\upsilon\leqslant 0,$$
$$\sum_{j=1}^{M}\lambda_j=1, \quad \lambda_j\geqslant 0, \quad j=1,2,\ldots,M.$$

Similarly to the proof of Theorem 2, Using (22) for $w=\sum_{j=1}^{M}x_{j,r}\tau_j$, $r=1,\ldots,R$, we can rewrite the dual problem in the form of the system of inequalities (24.1)–(24.5). □

References

Bali, Turan G., Ozgur Demirtas, K., Levy, Haim, Wolf, Avner, 2009. Bonds versus stocks: investors' age and risk taking. Journal of Monetary Economics 56, 817–830.
Banz, Rolf W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9 (1), 3–18.
Bawa, Vijay, 1975. Optimal rules for ordering uncertain prospects. Journal of Financial Economics 2 (1), 95–121.
Bawa, Vijay S., Lindenberg, Eric B., 1977. Capital market equilibrium in a mean-lower partial moment framework. Journal of Financial Economics 5, 189–200.
Bawa, Vijay S., Bodurtha Jr., James N., Rao, M.R., Suri, Hira L., 1985. On determination of stochastic dominance optimal sets. Journal of Finance 40, 417–431.
Barrett, Garry, Donald, Stephen G., 2003. Consistent tests for stochastic dominance. Econometrica 71, 71–104.
Benartzi, Shlomo, Thaler, Richard H., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110 (1), 73–92.
Blavatskyy, Pavlo R., 2010. Modifying the mean–variance approach to avoid violations of stochastic dominance. Management Science 56, 2050–2057.
Brown, David B., De Giorgi, Enrico, Sim, Melvyn, 2012. Aspirational preferences and their representation by risk measures. Management Science 58, 2095–2113.
Dentcheva, Darinka, Ruszczyński, Andrzej, 2010. Robust stochastic dominance and its application to risk-averse optimization. Mathematical Programming A 123 (1), 85–100.
Dupačová, Jitka, Kopa, Miloš, 2012. Robustness in stochastic programs with risk constraints. Annals of Operations Research 200 (1), 55–74.
Fábián, Csaba I., Mitra, Gautam, Roman, Diana, 2011. Processing second-order stochastic dominance models using cutting-plane representations. Mathematical Programming A 130 (1), 33–57.
Fama, Eugene F., French, Kenneth R., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.
Fishburn, Peter C., 1974. Convex stochastic dominance with continuous distribution functions. Journal of Economic Theory 7 (2), 143–158.
Gollmer, Ralf, Gotzes, Uwe, Schultz, Rüdiger, 2011. A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. Mathematical Programming A 126 (1), 179–190.
Hadar, Josef, Russell, William R., 1969. Rules for ordering uncertain prospects. American Economic Review 59, 25–34.
Hadar, Josef, Seo, Tae Kun, 1988. Asset proportions in optimal portfolios. The Review of Economic Studies 55 (3), 459–468.
Hanoch, Giora, Levy, Haim, 1969. The efficiency analysis of choices involving risk. Review of Economic Studies 36, 335–346.
Kimball, Miles S., 1993. Standard risk aversion. Econometrica 61 (3), 589–611.

Kopa, Miloš, Chovanec, Petr, 2008. A second-order stochastic dominance portfolio efficiency measure. Kybernetika 44 (2), 243–258.
Kopa, Miloš, Post, Thierry, 2009. A portfolio optimality test based on the first-order stochastic dominance criterion. Journal of Financial and Quantitative Analysis 44 (5), 1103–1124.
Kuosmanen, Timo, 2004. Efficient diversification according to stochastic dominance criteria. Management Science 50 (10), 1390–1406.
Leshno, Moshe, Levy, Haim, 2002. Preferred by all and preferred by most decision makers: almost stochastic dominance. Management Science 48, 1074–1085.
Levy, Haim, 2006. Stochastic Dominance: Investment Decision Making under Uncertainty. Springer, ISBN: 038729302.
Levy, Haim, Leshno, Moshe, Leibovitch, Boaz, 2010. Economically relevant preferences for all observed epsilon. Annals of Operations Research 176 (1), 153–178.
Levy, Haim, Markowitz, Harry, 1979. Approximating expected utility by a function of mean and variance. American Economic Review 69 (3), 308–317.
Linton, O., Maasoumi, E., Whang, Y.-J., 2005. Consistent testing for stochastic dominance under general sampling schemes. Review of Economic Studies 72, 735–765.
Liu, Yongchao, Xu, Huifu, 2013. Stability analysis of stochastic programs with second order dominance constraints. Mathematical Programming A, forthcoming. http://dx.doi.org/10.1007/s10107-012-0585-0.
Lozano, Sebastián, Gutiérrez, Ester, 2008. Data envelopment analysis of mutual funds based on second-order stochastic dominance. European Journal of Operational Research 189 (1), 230–244.
Lizyayev, Andrey, Ruszczyński, Andrzej, 2012. Tractable almost stochastic dominance. European Journal of Operational Research 218 (2), 448–455.
Lizyayev, Andrey, 2012. Stochastic dominance efficiency analysis of diversified portfolios: classification, comparison and refinements. Annals of Operations Research 196 (1), 391–410.
Mehra, Rajnish, Prescott, Edward C., 1985. The equity premium: a puzzle. Journal of Monetary Economics 15 (2), 145–161.
Meyer, Jack, 1977a. Second degree stochastic dominance with respect to a function. International Economic Review 18, 477–487.
Meyer, Jack, 1977b. Choice among distributions. Journal of Economic Theory 14, 326–336.
Meyer, Donald, Meyer, Jack, 2005. Relative risk aversion: what do we know? Journal of Risk and Uncertainty 3, 243–262.
Nelson, Ray D., Pope, Rulon, 1991. Bootstrapped insights into empirical applications of stochastic dominance. Management Science 37, 1182–1194.
Post, Thierry, 2003. Empirical tests for stochastic dominance efficiency. Journal of Finance 58, 1905–1932.
Post, Thierry, 2008. On the dual test for SSD efficiency: with an application to momentum investment strategies. European Journal of Operational Research 185 (3), 1564–1573.
Poterba, James M., Summers, Lawrence H., 1988. Mean reversion in stock returns: evidence and implications. Journal of Financial Economics 22, 27–60.
Quirk, James P., Saposnik, Rubin, 1962. Admissibility and measurable utility functions. Review of Economic Studies 29, 140–146.
Russell, W.R., Seo, T.K., 1989. Representative sets for stochastic dominance rules. In: Fomby, T.B., Seo, T.K. (Eds.), Studies in the Economics of Uncertainty: In Honor of Josef Hadar. Springer Verlag, New York, pp. 59–76.
Shalit, Haim, Yitzhaki, Shlomo, 1994. Marginal conditional stochastic dominance. Management Science 40, 670–684.

Please cite this article in press as: Post, T., Kopa, M. General linear formulations of stochastic dominance criteria. European Journal of Operational Research (2013), http://dx.doi.org/10.1016/j.ejor.2013.04.015

KYBERNETIKA — VOLUME 46 (2010), NUMBER 3, PAGES 488–500

MEASURING OF SECOND–ORDER STOCHASTIC DOMINANCE PORTFOLIO EFFICIENCY

Miloš Kopa

In this paper, we deal with second-order stochastic dominance (SSD) portfolio efficiency with respect to all portfolios that can be created from a considered set of assets. Assuming scenario approach for distribution of returns several SSD portfolio efficiency tests were proposed. We introduce a δ-SSD portfolio efficiency approach and we analyze the stability of SSD portfolio efficiency and δ-SSD portfolio efficiency classification with respect to changes in scenarios of returns. We propose new SSD and δ-SSD portfolio efficiency measures as measures of the stability. We derive a non-linear and mixed-integer non-linear programs for evaluating these measures. Contrary to all existing SSD portfolio inefficiency measures, these new measures allow us to compare any two δ-SSD efficient or SSD efficient portfolios. Finally, using historical US stock market data, we compute δ-SSD and SSD portfolio efficiency measures of several SSD efficient portfolios.

Keywords: stochastic dominance, stability, SSD portfolio efficiency measure

Classification: 91B28, 91B30

1. INTRODUCTION

When solving portfolio selection problem several approaches can be used: mean-risk models, maximising expected utility problems, stochastic dominance criteria, etc. If the information about the risk attitude of a decision maker is not known one may adopt stochastic dominance approach to test an efficiency of a given portfolio with respect to a considered set of utility functions. If only non-satiation and risk aversion of decision maker is assumed, that is, concave utility functions are considered, second-order stochastic dominance (SSD) relation allows comparison of any two portfolios.

Stochastic dominance was introduced independently in Hadar & Russel [6], Hanoch & Levy [7], Rothschild & Stiglitz [20] and Whitmore [23].[1]

The definition of second-order stochastic dominance relation uses comparisons of either twice cumulative distribution functions, or expected utilities (see for example Levy [13]). Alternatively, one can define SSD relation using cumulative quantile functions or conditional value at risk (see Ogryczak & Ruszczyński [15] or Kopa & Chovanec [9]).

[1]For more information see Levy [13].

Similarly to well-known mean-variance criterion, second-order stochastic dominance relation can be used in portfolio efficiency analysis. A given portfolio is called SSD efficient if there exists no other portfolio preferred by all risk-averse and risk-neutral decision makers (see for example Ruszczyński & Vanderbei [21], Kuosmanen [12] or Kopa & Chovanec [9]).

To test SSD portfolio efficiency of a given portfolio relative to all portfolios created from a set of assets Post [17], Kuosmanen [12] and Kopa & Chovanec [9] proposed several linear programming algorithms. While the Post test is based on representative utility functions and strict SSD efficiency criterion, the Kuosmanen and the Kopa-Chovanec test focuses on identifying a SSD dominating portfolio. The last two tests can be formulated as optimization problems with SSD constraints. Similar types of problems were discussed in Dentcheva & Ruszczyński [2, 3, 4], Rudolf & Ruszczyński [5] and Luedtke [14]. In these papers, weak stochastic dominance relation is used, contrary to SSD portfolio tests where strict stochastic dominance relation is considered.

For SSD inefficient portfolios, several SSD portfolio inefficiency measures were introduced in Post [17], Kuosmanen [12] and Kopa & Chovanec [9]. These measures are based on a "distance" between a tested portfolio and some other portfolio identified by a SSD portfolio efficiency test.

In all SSD portfolio efficiency tests, the scenario approach is assumed, that is, the returns of assets are modeled by discrete distribution with equiprobable scenarios. Therefore, especially for SSD efficient portfolios, one can ask how the original scenarios can be changed such that a given SSD efficient portfolio remains SSD efficient for perturbed scenarios, too. To circumvent this problem, Kopa & Post [10] suggested bootstrap techniques for first-order stochastic dominance (FSD) portfolio efficiency and Kopa [11] for SSD portfolio efficiency. In both cases, the inefficiency of a US stock market portfolio was detected with more than 95% significance. Alternatively, Dentcheva, Henrion and Ruszczyński [1] used a general stability results in stochastic programming (see Rőmisch [19]) for optimization problems with weak FSD constraints.

In this paper, we introduce a δ-SSD portfolio efficiency as a new type of portfolio efficiency with respect to second-order stochastic dominance criteria.

Fixing the number of equiprobable scenarios, we identify the maximal perturbation of original scenarios satisfying δ-SSD portfolio efficiency condition for a given portfolio. The magnitude of this maximal perturbation, expressed in terms of a distance between original and perturbed scenarios, can be considered as a measure of δ-SSD efficiency and the limiting case for $\delta \to 0$ leads to a new SSD efficiency measure. We consider only special perturbations where all scenarios are equiprobable and the number of scenarios is fixed. The more general approach can not be used because all SSD portfolio efficiency tests were developed only for equiprobable scenarios.

Contrary to the SSD inefficiency measures discussed above, δ-SSD and SSD portfolio efficiency measures are defined as measures of stability. In comparison with bootstrap techniques suggested by Kopa & Post [10] and Kopa [11], this new stability approach is more robust because it is not based only on a subsampling of given scenarios. The results reached in Dentcheva, Henrion and Ruszczyński [1] for optimization problems with weak FSD constraints can probably be extended for weak SSD constraints. However, this extension would be too technically and computationally demanding for SSD portfolio efficiency testing based on scenario approach and strict SSD relation. Moreover, the general stability results do not deal with any measure of stability.

We apply our stability analysis to the historical US stock market data (six Famma and French portfolios and a riskless asset) in order to compute the values of our δ-SSD and SSD portfolio efficiency measures for two SSD efficient portfolios. As the first portfolio, we choose the portfolio with the highest mean return. Since CVaR is consistent with SSD relation we find the second portfolio by solving mean-CVaR problem. For more details about the consistency see Ogryczak & Ruszczyński [15]. Another way of identifying a SSD efficient portfolio satisfying some required properties was presented in Roman, Darby-Dowman, and Mitra [18].

The remainder of the paper is organized as follows. The Preliminaries section starts with notation, assumptions and definitions for the SSD relation and SSD portfolio efficiency. We introduce a δ-SSD relation and δ-SSD portfolio efficiency as a new type of SSD relation and SSD portfolio efficiency. It is followed by a section dealing with SSD portfolio efficiency test derived in Kuosmanen [12] and it's modification for δ-SSD portfolio efficiency. In Section 4, we state our main stability ideas and we introduce new measures of SSD portfolio efficiency and δ-SSD portfolio efficiency as measures of stability. Using US stock market data, the final section presents a numerical illustration where we compute the δ-SSD and SSD portfolio efficiency measures for two SSD efficient portfolios .

2. PRELIMINARIES

We consider a random vector $\mathbf{r} = (r_1, r_2, \ldots, r_N)'$ of returns of N assets with a discrete probability distribution described by T equiprobable scenarios. The returns of the assets for the various scenarios are given by

$$X = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^T \end{pmatrix}$$

where $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_N^t)$ is the tth row of matrix X. We will use $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)'$ for a vector of portfolio weights and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^N | \mathbf{1}'\boldsymbol{\lambda} = 1, \ \ \lambda_n \geq 0, \ \ n = 1, 2, \ldots, N\}.$$

Alternatively, one can consider any bounded polytope:

$$\Lambda' = \{\boldsymbol{\lambda} \in \mathbb{R}^N | A\boldsymbol{\lambda} \geq \mathbf{b}\}.$$

The tested portfolio is denoted by $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_N)'$. Following Ruszczyński and Vanderbei [21], Kuosmanen [12], Kopa and Chovanec [9], we define second-order stochastic dominance relation in the strict form in the context of SSD portfolio efficiency. Let $F_{\mathbf{r}'\boldsymbol{\lambda}}(x)$ denote the cumulative probability distribution function of returns of portfolio $\boldsymbol{\lambda}$. The twice cumulative probability distribution function of returns of portfolio $\boldsymbol{\lambda}$ is defined as:

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) \quad = \quad \int_{-\infty}^{t} F_{\mathbf{r}'\boldsymbol{\lambda}}(x)\,\mathrm{d}x. \tag{1}$$

Definition 2.1. Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by second-order stochastic dominance ($\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$) if and only if

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) \le F^{(2)}_{\mathbf{r}'\boldsymbol{\tau}}(t) \qquad \forall\, t \in \mathbb{R}$$

with strict inequality[2] for at least one $t \in \mathbb{R}$.

The following SSD criteria can be used as alternative definitions of the SSD relation:

(i) $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ if and only if $\mathrm{E}u(\mathbf{r}'\boldsymbol{\lambda}) \ge \mathrm{E}u(\mathbf{r}'\boldsymbol{\tau})$ for all concave utility functions u provided the expected values above are finite and strict inequality is fulfilled for at least some concave utility function, see for example Levy [13].

(ii) $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ if and only if $F^{(-2)}_{\mathbf{r}'\boldsymbol{\lambda}}(\frac{p}{T}) \ge F^{(-2)}_{\mathbf{r}'\boldsymbol{\tau}}(\frac{p}{T})$ for all $p = 1, 2, \ldots, T$ with strict inequality for at least some p where the second quantile function $F^{(-2)}_{\mathbf{r}'\boldsymbol{\lambda}}$ is the convex conjugate function of $F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}$ in the sense of Fenchel duality, see Ogryczak & Ruszczyński [15]. Let $k = T - p$. Since

$$\mathrm{CVaR}_{1-\frac{p}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) = -\frac{F^{(-2)}_{\mathbf{r}'\boldsymbol{\tau}}(\frac{p}{T})}{\frac{p}{T}}$$

for all $p = 1, 2, \ldots, T$, where conditional value at risk (CVaR) can be defined via the optimization problem:

$$\begin{aligned} \mathrm{CVaR}_{\frac{k}{T}}(Y) \quad &= \quad \min_{a, w_t} \quad a + \frac{1}{(1-\frac{k}{T})T} \sum_{t=1}^{T} w_t \\ \text{s.t.} \quad w_t \quad &\ge \quad y_t - a \\ w_t \quad &\ge \quad 0, \end{aligned} \tag{2}$$

we can alternatively formulate the criterion in the following way: $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ if and only if $\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) \le \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau})$ for all $k = 0, 1, \ldots, T-1$ with strict inequality for at least some k. See Kopa and Chovanec [9], Uryasev & Rockafellar [22] and Pflug [16] for details.

[2]This type of SSD relation is sometimes referred to as the strict second-order stochastic dominance. If no strict inequality is required then the relation can be called the weak second-order stochastic dominance.

(iii) $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ if and only if there exists a double stochastic matrix $W = \{w\}_{ij}$ such that ($WX\boldsymbol{\tau} \leq X\boldsymbol{\lambda}$ and $\mathbf{1}'WX\boldsymbol{\tau} < \mathbf{1}'X\boldsymbol{\lambda}$) or ($WX\boldsymbol{\tau} = X\boldsymbol{\lambda}$ and $\sum_{i=1}^{T} w_{ii} < T$) where $\mathbf{1}' = (1, 1, \ldots, 1)$. See Kuosmanen [12] and Hardy, Littlewood & Polya [8] (Theorem 46) for details.

Since $\mathbf{1}'W = \mathbf{1}'$ for all double stochastic matrices W, using criterion (iii) we define a new type of SSD relation.

Definition 2.2. Let $\delta > 0$. Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by the δ-second-order stochastic dominance ($\mathbf{r}'\boldsymbol{\lambda} \succ_{\delta-SSD} \mathbf{r}'\boldsymbol{\tau}$) if there exists a double stochastic matrix $W = \{w\}_{ij}$ such that $X\boldsymbol{\lambda} \geq WX\boldsymbol{\tau}$ and $\mathbf{1}'X\boldsymbol{\lambda} - \mathbf{1}'X\boldsymbol{\tau} \geq \delta$.

The strictly positive parameter δ in Definition 2 is chosen sufficiently small, that is, such that if $X\boldsymbol{\lambda} \geq WX\boldsymbol{\tau}$ and $\mathbf{1}'X\boldsymbol{\lambda} - \mathbf{1}'X\boldsymbol{\tau} < \delta$ then vectors $X\boldsymbol{\lambda}$ and $WX\boldsymbol{\tau}$ are empirically indistinguishable.[3] It is easily seen that if portfolio $\boldsymbol{\lambda}$ δ-SSD dominates portfolio $\boldsymbol{\tau}$ for some $\delta > 0$ then $\boldsymbol{\lambda}$ SSD dominates $\boldsymbol{\tau}$. On the other hand, SSD relation need not imply δ-SSD relation for any $\delta > 0$. Hence, δ-SSD relation for some $\delta > 0$ is only sufficient condition of SSD relation.

Definition 2.3. A given portfolio $\boldsymbol{\tau} \in \Lambda$ is *SSD inefficient* if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is *SSD efficient*.

This definition classifies portfolio $\boldsymbol{\tau} \in \Lambda$ as SSD efficient if and only if no other portfolio is better (in the sense of the SSD relation) for all risk averse and risk neutral decision makers. Another definition of SSD efficiency was presented in Post [17]. Based on Definition 2, we can similarly define δ-SSD portfolio efficiency.

Definition 2.4. A given portfolio $\boldsymbol{\tau} \in \Lambda$ is *δ-SSD inefficient* if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\mathbf{r}'\boldsymbol{\lambda} \succ_{\delta-SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is δ-SSD efficient.

Since δ-SSD relation implies SSD relation, δ-SSD portfolio efficiency is a necessary condition of SSD portfolio efficiency, that is, every SSD efficient portfolio is δ-SSD efficient for all strictly positive δ.

3. SSD AND δ–SSD PORTFOLIO EFFICIENCY TEST

In this section we present the linear programming test of SSD portfolio efficiency in the form of necessary and sufficient condition derived in Kuosmanen [12]. From the three SSD efficiency tests: the Post test [17], the Kopa–Chovanec test [9] and the Kuosmanen test [12], we choose the last one, because the Kuosmanen test can be easily modify to a new δ-SSD efficiency test. The Kuosmanen test is based on criterion (iii) and it tries to identify a portfolio $\boldsymbol{\lambda} \in \Lambda$ that SSD dominates the given portfolio $\boldsymbol{\tau}$.

[3] This kind of approximation is sometimes used in empirical finance.

Lemma 3.1. (The Kuosmanen test) Let

$$\theta^* = \max_{W,\boldsymbol{\lambda}} \sum_{t=1}^{T} \left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) \tag{3}$$
$$\begin{aligned} \text{s.t.} \quad X\boldsymbol{\lambda} &\geq WX\boldsymbol{\tau} \\ \sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} &= 1, \quad w_{ij} \geq 0 \quad i,j = 1,2,\ldots,T \\ \boldsymbol{\lambda} &\in \Lambda \end{aligned}$$

and

$$\theta^{**} = \min_{W,\boldsymbol{\lambda},S^+,S^-} \sum_{j=1}^{T}\sum_{i=1}^{T}(s_{ij}^+ + s_{ij}^-) \tag{4}$$
$$\begin{aligned} \text{s.t.} \quad X\boldsymbol{\lambda} &= WX\boldsymbol{\tau} \\ s_{ij}^+ - s_{ij}^- &= w_{ij} - \frac{1}{2} \quad i,j = 1,2,\ldots,T \\ s_{ij}^+, s_{ij}^-, w_{ij} &\geq 0 \quad i,j = 1,2,\ldots,T \\ \sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} &= 1 \quad i,j = 1,2,\ldots,T \\ \boldsymbol{\lambda} &\in \Lambda \end{aligned}$$

where $S^+ = \{s_{ij}^+\}_{i,j=1}^T$, $S^- = \{s_{ij}^-\}_{i,j=1}^T$ and $W = \{w_{ij}\}_{i,j=1}^T$. Let ϵ_k denote the number of k-way ties in $X\boldsymbol{\tau}$.[4] Then portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if

$$\theta^* = 0 \quad \wedge \ \theta^{**} = \frac{T^2}{2} - \sum_{k=1}^{T} k\epsilon_k.$$

Let $\boldsymbol{\lambda}^*$ and $\boldsymbol{\lambda}^{**}$ be the optimal solution of (3) and (4), respectively. If $\theta^* > 0$ then $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. If $\theta^* = 0$ and $\theta^{**} < \frac{T^2}{2} - \sum_{k=1}^{T} k\epsilon_k$ then $\mathbf{r}'\boldsymbol{\lambda}^{**} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$.

If $\theta^* > 0$ then problem (4) need not to be solved, because portfolio $\boldsymbol{\tau}$ is SSD inefficient and the optimal solution $\boldsymbol{\lambda}^*$ is a SSD dominating portfolio, see Kuosmanen [12] for more details.

If a given portfolio $\boldsymbol{\tau}$ is SSD inefficient then, from the entire set of SSD dominating portfolios, the Kuosmanen test identifies that with the highest mean return. That is, (3) and (4) can be reformulated in the following way:

$$\max_{\boldsymbol{\lambda}\in\Lambda} \quad f(\boldsymbol{\lambda},\boldsymbol{\tau}) \tag{5}$$
$$\text{s.t.} \quad \mathbf{r}'\boldsymbol{\lambda} \quad \succ_{SSD} \quad \mathbf{r}'\boldsymbol{\tau}, \tag{6}$$

where $f(\boldsymbol{\lambda},\boldsymbol{\tau}) = T\left(\mathrm{E}(\mathbf{r}'\boldsymbol{\lambda}) - \mathrm{E}(\mathbf{r}'\boldsymbol{\tau})\right) = \sum_{t=1}^{T} \mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}$.

[4] We say that a k-way tie occurs if k elements of $X\boldsymbol{\tau}$ are equal to each other.

Problem (5) – (6) is an optimization problem with a stochastic dominance constraint. Contrary to problems discussed in Dentcheva & Ruszczyński [2, 3, 4], Rudolf & Ruszczyński [5] and Luedtke [14], the stochastic dominance constraint (6) is in the strict form.

The optimal value θ^* of (3) can be considered as a measure of SSD portfolio inefficiency. It gives us information about the maximal possible difference, expressed in mean return (or sum of returns), between the tested portfolio and a SSD dominating portfolio. The alternative SSD portfolio inefficiency measures arise from the Post test and the Kopa–Chovanec test. All these three measures allow comparison of two SSD inefficient portfolios. Unfortunately, these measures are not suitable for SSD efficiency measuring, because all these measures are equal to zero for all SSD efficient portfolios. Therefore, for SSD portfolio efficiency measuring, we suggest another approach, based on the δ-SSD portfolio efficiency and stability of δ-SSD portfolio efficiency classification. Firstly, we modify the Kuosmanen test to δ-SSD portfolio efficiency test.

Lemma 3.2. (The δ-SSD portfolio efficiency test) Let

$$\begin{aligned} \theta^*_\delta = \max_{W,\boldsymbol{\lambda}} \sum_{t=1}^{T} \left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) & & & \qquad (7)\\ \text{s.t.} \qquad X\boldsymbol{\lambda} &\geq WX\boldsymbol{\tau} \\ \sum_{t=1}^{T}\left(\mathbf{x}^t\boldsymbol{\lambda} - \mathbf{x}^t\boldsymbol{\tau}\right) &\geq \delta \\ \sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} &= 1, \quad w_{ij} \geq 0 \qquad i,j = 1,2,\ldots,T \\ \boldsymbol{\lambda} &\in \Lambda. \end{aligned}$$

If an optimal solution of (7) exists then portfolio $\boldsymbol{\tau}$ is δ-SSD inefficient and $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{\delta-SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, $\boldsymbol{\tau}$ is δ-SSD efficient portfolio.

The proof of Lemma 3.2 directly follows from Lemma 3.1, criterion (iii), Definition 2.2 and Definition 2.4.

4. STABILITY OF SSD AND δ-SSD PORTFOLIO EFFICIENCY CLASSIFICATION

In previous sections a fixed scenario matrix was considered and all portfolio efficiency tests were done for this scenario matrix. Unfortunately, usually we do not have perfect information about the distribution of returns. Therefore, the stability of SSD portfolio efficiency and δ-SSD portfolio efficiency with respect to changes in the scenario matrix is investigated in this section.

Since the SSD portfolio efficiency tests and the δ-SSD portfolio efficiency test are derived under the assumption of equiprobable scenarios collected in matrix X we will consider only perturbation matrices X_p of the original matrix X which have exactly T rows, that is, we admit only approximations with T equiprobable scenarios. Let $\mathcal{X}_p$ be the set of all such perturbation matrices. In this section we analyze how the results of the SSD and δ-SSD portfolio efficiency test for a given portfolio depend on the original scenario matrix X and which other matrices X_p from a neighbourhood[5] of X guarantee the SSD or δ-SSD portfolio efficiency of the given portfolio.

Let matrix $\Upsilon = \{\upsilon_{ij}\}_{i,j=1}^{T}$ be defined as $\Upsilon = X_p - X$. Let $D(X, X_p) = \max_{i,j} |\upsilon_{ij}|$ denote a distance between matrices X and X_p on $\mathcal{X}_p$. We introduce a new measure of δ-SSD portfolio efficiency as a measure of stability.

Definition 4.1. The δ-SSD portfolio efficiency measure γ_δ of δ-SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ is defined as the optimal value of the following optimization problem:

$$\begin{aligned} \gamma_\delta(\boldsymbol{\tau}) = \max \quad & \varepsilon \\ \text{s.t.} \quad \boldsymbol{\tau} \ \text{ is } \ \delta - SSD \ \text{ efficient for all } \quad & X_p \in \mathcal{X}_p \quad \text{such that} \quad D(X, X_p) \leq \varepsilon. \end{aligned} \tag{8}$$

This measure gives us information how large is the neighborhood of X such that for all matrices from this neighborhood the portfolio $\boldsymbol{\tau}$ is classified as δ-SSD efficient. The problem (8) consists of infinitely many δ-SSD efficiency constraints. Moreover, according to the Lemma 3.2, each constraint involves a maximization problem what makes problem (8) practically unsolvable. Therefore we reinterpret the δ-SSD portfolio efficiency measure for a given δ-SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ as the minimal distance between the original matrix X and any other matrix X_p that makes portfolio $\boldsymbol{\tau}$ δ-SSD inefficient, that is,

$$\begin{aligned} \gamma_\delta(\boldsymbol{\tau}) = \min_{X_p \in \mathcal{X}_p} \quad & D(X, X_p) \\ \text{s.t.} \quad \boldsymbol{\tau} \ \text{ is } \ \delta - SSD \ \text{ inefficient for } \quad & X_p. \end{aligned} \tag{9}$$

Using Lemma 3.2, the SSD portfolio efficiency measure $\gamma_\delta(\boldsymbol{\tau})$ can be computed in a much less computationally demanding way:

$$\gamma_\delta(\boldsymbol{\tau}) = \min_{\boldsymbol{\lambda} \in \Lambda, X_p \in \mathcal{X}_p} \quad D(X, X_p) \tag{10}$$

$$\begin{aligned} \text{s.t.} \qquad X\boldsymbol{\lambda} - WX\boldsymbol{\tau} \quad & \geq \quad 0 \\ \sum_{t=1}^{T} \left(\mathbf{x}^t \boldsymbol{\lambda} - \mathbf{x}^t \boldsymbol{\tau}\right) \quad & \geq \quad \delta \\ \sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} = 1, \quad w_{ij} \quad & \geq \quad 0 \qquad i, j = 1, 2, \ldots, T. \end{aligned}$$

Since $\Upsilon = X_p - X$ and $D(X, X_p) = \max_{i,j} |\upsilon_{ij}|$ the measure $\gamma_\delta(\boldsymbol{\tau})$ can be computed using the following non-linear program.

[5] for a given metric on $\mathcal{X}_p$

$$\gamma_\delta(\boldsymbol{\tau}) = \min_{\boldsymbol{\lambda}\in\Lambda,\Upsilon,\varepsilon} \varepsilon \tag{11}$$

$$\begin{array}{rcll}
\text{s.t.} \quad (X+\Upsilon)\boldsymbol{\lambda} - W(X+\Upsilon)\boldsymbol{\tau} & \geq & 0 & \\
\sum_{t=1}^{T}\left((\mathbf{x}^t+\boldsymbol{\upsilon}^t)\boldsymbol{\lambda} - (\mathbf{x}^t+\boldsymbol{\upsilon}^t)\boldsymbol{\tau}\right) & \geq & \delta & \\
\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} = 1, \quad w_{ij} & \geq & 0 & i,j = 1,2,\dots,T \\
-\varepsilon \leq \upsilon_{ij} & \leq & \varepsilon & i,j = 1,2,\dots,T,
\end{array}$$

where $\boldsymbol{\upsilon}^t = (\upsilon_{t1}, \upsilon_{t2}, \dots, \upsilon_{tT})$ is the tth row of matrix Υ. For a given portfolio $\boldsymbol{\tau}$ we have $\gamma_\delta(\boldsymbol{\tau}) \geq 0$ for all $\delta > 0$. Moreover, if $\delta_1 < \delta_2$ then the set of feasible solutions of (11) is larger for δ_1 than for δ_2 and consequently $\gamma_{\delta_1}(\boldsymbol{\tau}) \leq \gamma_{\delta_2}(\boldsymbol{\tau})$. Therefore, we can define a measure of SSD efficiency in the following way.

Definition 4.2. The SSD portfolio efficiency measure γ of SSD efficient portfolio $\boldsymbol{\tau} \in \Lambda$ is defined as: $\gamma(\boldsymbol{\tau}) = \lim_{\delta\to 0^+} \gamma_\delta(\boldsymbol{\tau}) = \inf_{\delta>0} \gamma_\delta(\boldsymbol{\tau})$.

4.1. One scenarion perturbation – A given scenario

Assume that only the tth scenario can be changed, that is $\upsilon_{ij} = 0$ for all $i \neq t$. Then $D(X, X_p) = \max_j |\upsilon_{tj}|$ and the corresponding δ-SSD efficiency measure γ_δ^t is defined as

$$\gamma_\delta^t(\boldsymbol{\tau}) = \min_{\boldsymbol{\lambda}\in\Lambda,\Upsilon,\varepsilon} \varepsilon \tag{12}$$

$$\begin{array}{rcll}
\text{s.t.} \quad (X+\Upsilon)\boldsymbol{\lambda} - W(X+\Upsilon)\boldsymbol{\tau} & \geq & 0 & \\
\sum_{t=1}^{T}\left((\mathbf{x}^t+\boldsymbol{\upsilon}^t)\boldsymbol{\lambda} - (\mathbf{x}^t+\boldsymbol{\upsilon}^t)\boldsymbol{\tau}\right) & \geq & \delta & \\
\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} = 1, \quad w_{ij} & \geq & 0 & i,j = 1,2,\dots,T \\
-\varepsilon \leq \upsilon_{tj} & \leq & \varepsilon & j = 1,2,\dots,T \\
\upsilon_{ij} & = & 0 & i \neq t \;\; j = 1,2,\dots,T.
\end{array}$$

Similarly to the complete scenario perturbation case, the SSD efficiency measure for one scenario perturbation is: $\gamma^t(\boldsymbol{\tau}) = \lim_{\delta\to 0^+} \gamma_\delta(\boldsymbol{\tau}) = \inf_{\delta>0} \gamma_\delta(\boldsymbol{\tau})$.

4.2. One scenarion perturbation – An arbitrary scenario

In this section we still assume that only one scenario can be changed. Contrary to the previous case, now we do not prescribe which scenario it is. Therefore we again consider $D(X, X_p) = \max_{i,j} |\upsilon_{ij}|$ as in the general case and the δ-SSD portfolio efficiency measure for this situation is defined as:

$$\overline{\gamma}_\delta(\boldsymbol{\tau}) = \min_{\boldsymbol{\lambda}\in\Lambda,\Upsilon,\varepsilon} \varepsilon \tag{13}$$

$$\begin{array}{rcl}
\text{s.t.} \qquad (X+\Upsilon)\boldsymbol{\lambda} - W(X+\Upsilon)\boldsymbol{\tau} & \geq & 0 \\
\sum_{t=1}^{T} \left((\mathbf{x}^t + \boldsymbol{\upsilon}^t)\boldsymbol{\lambda} - (\mathbf{x}^t + \boldsymbol{\upsilon}^t)\boldsymbol{\tau}\right) & \geq & \delta \\
\sum_{j=1}^{T} w_{ij} = 1, \quad \sum_{i=1}^{T} w_{ij} = 1, \quad w_{ij} & \geq & 0 \quad i,j = 1,2,\ldots,T \\
-\varepsilon \;\leq\; \upsilon_{ij} & \leq & \varepsilon \quad i,j = 1,2,\ldots,T \\
\upsilon_{ij} & \leq & My_i \;\; j = 1,2,\ldots,T \\
\sum_{i=1}^{T} y_i = 1, \quad y_i & \in & \{0,1\},
\end{array}$$

where M is a sufficiently large constant, for example $M = 2\sum_{i,j=1}^{T} |x_{ij}|$. Problem (13) is more computationally demanding than (12) because T binary variables are added. The corresponding SSD efficiency measure is again defined as the limiting case: $\overline{\gamma}(\boldsymbol{\tau}) = \lim_{\delta\to 0^+} \gamma_\delta(\boldsymbol{\tau}) = \inf_{\delta>0} \gamma_\delta(\boldsymbol{\tau})$.

5. EMPIRICAL APPLICATION

To illustrate our portfolio efficiency measuring, we apply it to the US stock market data in order to compute the δ-SSD portfolio efficiency measure γ_δ, and SSD portfolio efficiency measure γ of two SSD efficient portfolios. The investment universe of stocks is proxied by the well-known six value-weighted Fama and French portfolios. The last considered asset is the riskless asset that is proxied by the one-year US government bond index from Ibbotson Associates. We consider yearly excess returns from 1963 to 2002 (40 annual observations). Excess returns are computed by subtracting the riskless rate from the nominal returns, that is, the riskless asset always has a return of zero. Table 1 shows descriptive statistics for our data set.

We start with identifying two SSD efficient portfolios. Since short sales are not allowed and no two assets have the same mean, the portfolio consisting only of the asset with the highest mean $\boldsymbol{\tau}_1 = (0,0,1,0,0,0,0)$ is obviously SSD efficient. Ogryczak & Ruszczyński [15] proved that several mean-risk models are consistent with SSD relation, e. g., for CVaR as a measure of risk. Therefore, if mean-CVaR model has an unique optimal solution then it is a SSD efficient portfolio. Solving mean-CVaR model with $\alpha = 0.95$ we identified the second SSD efficient portfolio $\boldsymbol{\tau}_2 = (0,0,0.385,0.016,0,0.013,0.586)$ were the minimal required mean was equal to the mean of market portfolio proxied by the CRSP all-share index. We solve problems (11) for both SSD efficient portfolios and five levels $\delta = 1, 0.1, 0.01, 0.001, 0.0001$ using GAMS system (solver COINIPOPT). The results are presented in Table 2.

Table 1. Descriptive statistics for 6 Famma and French portfolios formed on market capitalization of equity and book-to-market equity ratio (SG = small growth, SN = small neutral, SV = small value, BG = big growth, BN = big neutral and BV = big value).

	Mean	St.dev.	Skew.	Kurt.	Min.	Max.
SG	5.309	28.520	0.323	0.175	–49.28	83.68
SN	11.301	22.728	–0.308	0.062	–37.38	65.48
SV	13.861	23.158	–0.373	–0.222	–33.86	61.14
BG	5.303	18.820	–0.317	–0.537	–40.49	34.67
BN	6.340	16.120	–0.241	–0.090	–34.13	34.73
BV	8.946	17.723	–0.690	–0.026	–34.24	40.34

Table 2. δ-SSD efficiency measures for portfolio $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$.

	$\delta = 1$	$\delta = 0.1$	$\delta = 0.01$	$\delta = 0.001$	$\delta = 0.0001$
$\gamma_\delta(\boldsymbol{\tau}_1)$	1.369	1.369	1.369	1.369	1.369
$\gamma_\delta(\boldsymbol{\tau}_2)$	0.937	0.412	0.393	0.388	0.388

From Table 2 we can see that $\gamma_\delta(\boldsymbol{\tau}_1) = 1.369$ for all $\delta \in \langle 1, 0.0001\rangle$ and therefore we can expect that $\gamma(\boldsymbol{\tau}_1) = \inf_{\delta>0} \gamma_\delta(\boldsymbol{\tau}_1) = 1.369$. To prove it, we apply the modified Kousmanen test where we use $X + \Upsilon$ instead of X and we include the additional constraints:[6]

$$-(1.369 - \xi) \quad \leq \quad v_{ij} \quad \leq \quad (1.369 - \xi) \qquad i, j = 1, 2, \ldots, T, \tag{14}$$

where ξ is a sufficiently small number,[7] in our case we choose $\xi = 0.0005$. This modified test tries to identify a SSD dominating portfolio for any feasible perturbed scenario matrix. We can find that the test fails to identify a SSD dominating portfolio for completely perturbed scenario matrices X_p with $D(X, X_p) \leq 1.3685$. Therefore, we can conclude that the SSD portfolio efficiency measure of portfolio $\boldsymbol{\tau}_1$ is equal to 1.369. By analogy, we can easily check that $\gamma(\boldsymbol{\tau}_2) = 0.388$ where we use $\xi = 0.0005$ and

$$-(0.388 - \xi) \quad \leq \quad v_{ij} \quad \leq \quad (0.388 - \xi) \qquad i, j = 1, 2, \ldots, T.$$

instead of (14).

ACKNOWLEDGEMENT

This work was partially supported by the Czech Science Foundation (grant 201/07/P107 and grant P402/10/1610).

(Received April 14, 2010)

[6] Note that including matrix of variables Υ makes the test non-linear.

[7] The choice of ξ depends on a prescribed accuracy level. In our example we round all values to three decimal digits accuracy.

REFERENCES

[1] D. Dentcheva, R. Henrion, and A. Ruszczyński: Stability and sensitivity of optimization problems with first order stochastic dominance constraints. SIAM J. Optim. *18* (2007), 322–333.

[2] D. Dentcheva and A. Ruszczyński: Optimization with stochastic dominance constraints. SIAM J. Optim. *14* (2003), 548–566.

[3] D. Dentcheva and A. Ruszczyński: Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. Math. Programming *99* (2004), 329–350.

[4] D. Dentcheva and A. Ruszczyński: Portfolio optimization with stochastic dominance constraints. J. Banking and Finance *30* (2006), 2, 433–451.

[5] G. Rudolf and A. Ruszczyński: Optimization problems with second order stochastic dominance constraints: duality, compact formulations, and cut generation methods. SIAM J. Optim. *19* (2008), 3, 1326–1343.

[6] J. Hadar and W. R. Russell: Rules for ordering uncertain prospects. Amer. Econom. Rev. *59* (1969), 1, 25–34.

[7] G. Hanoch and H. Levy: The efficiency analysis of choices involving risk. Rev. Econom. Stud. *36* (1969), 335–346.

[8] G. H. Hardy, J. E. Littlewood, and G. Polya: Inequalities. Cambridge University Press, Cambridge 1934.

[9] M. Kopa and P. Chovanec: A second-order stochastic dominance portfolio efficiency measure. Kybernetika *44* (2008), 2, 243–258.

[10] M. Kopa and T. Post: A portfolio optimality test based on the first-order stochastic dominance criterion. J. Financial and Quantitative Analysis *44* (2009), 5, 1103–1124.

[11] M. Kopa: An efficient LP test for SSD portfolio efficiency. Working paper, available at: http://ssrn.com/abstract=1340863.

[12] T. Kuosmanen: Efficient diversification according to stochastic dominance criteria. Management Sci. *50* (2004), 10, 1390–1406.

[13] H. Levy: Stochastic Dominance: Investment Decision Making Under Uncertainty. Second edition. Springer Science, New York 2006.

[14] J. Luedtke: New formulations for optimization under stochastic dominance constraints. SIAM J. Optim. *19* (2008), 3, 1433–1450.

[15] W. Ogryczak and A. Ruszczyński: Dual stochastic dominance and related mean-risk models. SIAM J. Optim. *13* (2002), 60–78.

[16] G. Ch. Pflug: Some remarks on the value-at-risk and the conditional value-at-risk. In: Probabilistic Constrained Optimization: Methodology and Applications (S. Uryasev, ed.), Kluwer Academic Publishers, Norwell MA 2000, pp. 278–287.

[17] T. Post: Empirical tests for stochastic dominance efficiency. J. Finance *58* (2003), 1905–1932.

[18] D. Roman, K. Darby-Dowman, and G. Mitra: Portfolio construction based on stochastic dominance and target return distributions. Math. Programming, Series B *108* (2006), 541–569.

[19] W. Römisch: Stability of stochastic programming problems. In: Stochastic Programming. Handbooks in Operations Research and Management Science *10* (A. Ruszczyński and A. Shapiro, eds.), Elsevier, Amsterdam 2003, pp. 483–554.

[20] M. Rothschild and J. E. Stiglitz: Rules for ordering uncertain prospects. J. Economic Theory *2* (1969), 225–243.

[21] A. Ruszczyński and R. J. Vanderbei: Frontiers of stochastically nondominated portfolios. Econometrica *71* (2003), 4, 1287–1297.

[22] S. Uryasev and R. T. Rockafellar: Conditional value-at-risk for general loss distributions. J. Banking and Finance *26* (2002), 1443–1471.

[23] G. A. Whitmore: Third degree stochastic dominance. Amer. Econom. Rev. *60* (1970), 457–459.

Miloš Kopa, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8 and Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Prague. Czech Republic.

e-mail: kopa@karlin.mff.cuni.cz

Ann Oper Res (2012) 200:55–74
DOI 10.1007/s10479-010-0824-9

Robustness in stochastic programs with risk constraints

Jitka Dupačová · Miloš Kopa

Published online: 5 January 2011

Abstract This paper is a contribution to the robustness analysis for stochastic programs whose set of feasible solutions depends on the probability distribution P. For various reasons, probability distribution P may not be precisely specified and we study robustness of results with respect to perturbations of P. The main tool is the contamination technique. For the optimal value, local contamination bounds are derived and applied to robustness analysis of the optimal value of a portfolio performance under risk-shaping CVaR constraints. A new robust portfolio efficiency test with respect to the second order stochastic dominance criterion is suggested and the contamination methodology is exploited to analyze its resistance with respect to additional scenarios.

Keywords Expectation type constraints · Robustness analysis · Contamination technique · Risk-shaping with CVaR · Second order stochastic dominance · Robust SSD portfolio efficiency test

1 Introduction

In this paper we shall deal with robustness properties of risk constrained stochastic programs of the form

$$\min_{\mathbf{x}\in\mathcal{X}} F_0(\mathbf{x}, P)$$

subject to

$$F_j(\mathbf{x}, P) \le 0, \quad j = 1, \dots, J, \tag{1}$$

where

J. Dupačová (✉) · M. Kopa
Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics,
Charles University in Prague, Sokolovská 83, 18675 Prague, Czech Republic
e-mail: dupacova@karlin.mff.cuni.cz

M. Kopa
e-mail: kopa@karlin.mff.cuni.cz

– P is the probability distribution of a random vector ω with range $\Omega \subset \mathbb{R}^M$,
– $\mathcal{X} \subset \mathbb{R}^N$ is a fixed nonempty convex set,
– functions $F_j(\mathbf{x}, P)$, $j = 0, \ldots, J$ may depend on P.

We shall denote $\mathcal{X}(P)$ the set of feasible solutions, $\mathcal{X}^*(P)$ the set of optimal solutions and $\varphi(P)$ the optimal value of the objective function in (1).

Probably the first paper formulating and analyzing risk constrained stochastic programs is due to Prékopa (1973) which includes joint probability constraints and constraints in the form of conditional expectations; see also Wets (1989) for the problem formulation and for properties of expectation functionals. Notice that *chance or probability constraints* are a special case of (1), however the set of feasible solutions $\mathcal{X}(P)$ is then convex only under special distributional and structural assumptions; consult Prékopa (2003).

Due to the tendency of an adequate treatment of risk, a growing interest in the risk constrained problems can be observed since 2000. It turns out that among others, the Sample Average Approximation technique, see e.g. Shapiro (2003), Pagoncelli et al. (2009), Wang and Ahmed (2008), and its asymptotics can be applied. This assumes that i.i.d. samples are drawn from a *fixed* (known, preselected) probability distribution P.

The wish is to apply reliable, robust or efficient decisions of (1) even in situations when the true probability distribution P has been approximated or when it is known only partly. Partial knowledge of P can be included into the model formulation, see e.g. Dentcheva and Ruszczyński (2010) for robust stochastic dominance constraints or Pflug and Wozabal (2007) for an inclusion of ambiguity of P into the model. In a similar vein a robust portfolio efficiency test will be developed in Sect. 3.2. A special case of robust portfolio efficiency was analyzed in Kopa (2010). Contrary to that, our new test allows probability distributions with nonequiprobable scenarios.

Another possibility is to rely on general quantitative stability results valid under suitable continuity assumptions for $F_j(\mathbf{x}, P)$, $j = 0, \ldots, J$. Such results were proved by Römisch (2003) without convexity requirements and were detailed e.g. for chance constraints of a special structure and formulated also for risk measures *nonlinear* in P. Under modest assumptions they apply to the convex problem (1).

In Sect. 2, we shall follow the relatively simple ideas of output analysis based on the contamination technique, cf. Dupačová (1996, 2006), Dupačová and Polívka (2007). The considered special type of perturbations gets on with needs for what-if-analysis or stress testing. Robustness results with respect to contamination of P by another fixed probability distribution have been mainly developed for convex stochastic programs whose set of feasible decisions does not depend on P, an assumption which does not apply to problem (1), and for the objective function $F_0(\mathbf{x}, P)$ convex in $\mathbf{x}$ and linear or concave in P. To elaborate special techniques for stress testing and robustness analysis for problem (1) it is necessary to relax the assumption of a fixed set of feasible decisions and to allow its dependence on P. To this purpose, it is convenient if the constraints are *linear* in P being expectations of random convex functions. Even with the expectation type constraints the problem formulation (1) covers various known examples, e.g. CVaR constraints from Rockafellar and Uryasev (2002), Krokhmal et al. (2002) or the second order stochastic dominance constraints. This is the class of problems for which we shall detail our robustness analysis and provide numerical illustrations. The next example introduces the prototype form of the problem.

Example 1 (Risk-shaping with CVaR; Rockafellar and Uryasev 2002) Let $f(\mathbf{x}, \omega)$ denote the random loss caused by the decision $\mathbf{x} \in \mathcal{X}$ and $\alpha \in (0, 1)$ the selected confidence level. The *Conditional Value at Risk* at the confidence level α, CVaR_α, is defined as the mean of the α-tail distribution of $f(\mathbf{x}, \omega)$. According to the fundamental minimization formula by Rockafellar and Uryasev (2002) it can be evaluated by minimization of the auxiliary function

$$\Phi_\alpha(\mathbf{x}, v, P) := v + \frac{1}{1-\alpha} E_P(f(\mathbf{x}, \omega) - v)^+$$

with respect to $v \in \mathbb{R}$.

The auxiliary function $\Phi_\alpha(\mathbf{x}, v, P)$ is evidently linear in P and convex in v. Moreover, if $f(\mathbf{x}, \omega)$ is a *convex* function of $\mathbf{x}$, $\Phi_\alpha(\mathbf{x}, v, P)$ is convex jointly in $(v, \mathbf{x})$.

If P is a *discrete probability distribution* concentrated on $\omega^1, \dots, \omega^S$, with probabilities $p_s > 0$, $s = 1, \dots, S$, and $\mathbf{x}$ a *fixed* element of $\mathcal{X}$, then the optimization problem $\mathrm{CVaR}_\alpha(\mathbf{x}, P) = \min_v \Phi_\alpha(\mathbf{x}, v, P)$ has the form

$$\mathrm{CVaR}_\alpha(\mathbf{x}, P) = \min_v \left\{ v + \frac{1}{1-\alpha} \sum_{s=1}^{S} p_s (f(\mathbf{x}, \omega^s) - v)^+ \right\} \tag{2}$$

and can be written as

$$\mathrm{CVaR}_\alpha(\mathbf{x}, P) = \min_{v, z_1, \dots, z_S} \left\{ v + \frac{1}{1-\alpha} \sum_{s=1}^{S} p_s z_s \,|\, z_s \geq 0,\ z_s + v \geq f(\mathbf{x}, \omega^s)\ \forall s \right\}. \tag{3}$$

Risk-shaping with CVaR handles several probability thresholds $\alpha_1, \dots, \alpha_J$ and loss tolerances b_j, $j = 1, \dots, J$. The problem is to minimize a performance function $F(\mathbf{x})$ subject to $\mathbf{x} \in \mathcal{X}$ and constraints $\mathrm{CVaR}_{\alpha_j}(\mathbf{x}, P) \leq b_j$, $j = 1, \dots, J$. According to Theorem 16 of Rockafellar and Uryasev (2002), this problem is equivalent to

$$\min_{\mathbf{x}, v_1, \dots, v_J} \{F(\mathbf{x}) \,|\, \mathbf{x} \in \mathcal{X},\ \Phi_{\alpha_j}(\mathbf{x}, v_j, P) \leq b_j,\ j = 1, \dots, J\},$$

i.e. it is a problem of the form (1) with expectation type constraints.

2 Contamination bounds

Contamination means to model the perturbations of P by its contamination by another fixed probability distribution Q, i.e. to use $P_t := (1-t)P + tQ$, $t \in [0, 1]$ in stochastic program (1) at the place of P. Then the set of feasible solutions of (1) for the contaminated probability distribution P_t equals

$$\mathcal{X}(P_t) = \mathcal{X} \cap \{\mathbf{x} \,|\, F_j(\mathbf{x}, P_t) \leq 0, j = 1, \dots, J\}. \tag{4}$$

We denote $\mathcal{X}(t)$, $\varphi(t)$, $\mathcal{X}^*(t)$ the set of feasible solutions, the optimal value $\varphi(P_t)$ and the set of optimal solutions $\mathcal{X}^*(P_t)$ of the contaminated problem

$$\text{minimize } F_0(\mathbf{x}, P_t) \text{ on the set } \mathcal{X}(P_t). \tag{5}$$

This is a *nonlinear parametric program* with a scalar parameter $t \in [0, 1]$ and a *parameter dependent set of feasible solutions* $\mathcal{X}(t) := \{\mathbf{x} \in \mathcal{X} \,|\, F_j(\mathbf{x}, t) \leq 0, j = 1, \dots, J\}$.

The task is to construct computable lower and upper bounds for $\varphi(t)$. Such bounds were obtained for $\mathcal{X}$ fixed, independent of P and for objective function $F_0(\mathbf{x}, P)$ linear or concave in P, cf. Dupačová (1996, 1998). In this case, one can exploit the fact that the optimal value function $\varphi(t)$ is a concave function of the contamination parameter t. The derived bounds proved to be useful for testing the resistance with respect to a sample for scenario-based stochastic programs, e.g. Dupačová (1996), in stress testing of CVaR optimization problems, cf. Dupačová (2006), Dupačová and Polívka (2007), or for problems with polyhedral risk objectives, cf. Dupačová (2008). For the parameter dependent sets of feasible solutions the optimal value function $\varphi(t)$ is concave only under rather strict assumptions such as $F_j(\mathbf{x}, t),\ j = 1, \dots, J$ jointly concave on $\mathcal{X} \times [0, 1]$ (cf. Corollary 3.2 of Kyparisis and Fiacco 1987) which is not in agreement with our problem formulation.

We shall examine how to construct contamination bounds for SP of the type (5) whose constraints depend on the probability distribution. These bounds will be then applied in robustness analysis for risk-shaping with CVaR or for a stochastic dominance test with respect to inclusion of additional scenarios. We shall see that thanks to the assumed structure of perturbations the lower bound can be derived for $F_j(\mathbf{x}, P),\ j = 0, \dots, J$ linear or concave with respect to P without any smoothness or convexity assumptions with respect to $\mathbf{x}$. Convexity of the stochastic program (1) is essential for directional differentiability of the optimal value function, and further assumptions are needed for derivation of an upper bound.

2.1 Lower bound

Consider first only one constraint dependent on probability distribution P and an objective F_0 *independent* of P, i.e. the problem is

$$\min_{\mathbf{x} \in \mathcal{X}} F_0(\mathbf{x}) \quad \text{subject to} \quad F(\mathbf{x}, P) \le 0. \tag{6}$$

For probability distribution P contaminated by another fixed probability distribution Q, i.e. for $P_t := (1 - t)P + tQ$, $t \in (0, 1)$ we get

$$\min_{\mathbf{x} \in \mathcal{X}} F_0(\mathbf{x}) \quad \text{subject to} \quad F(\mathbf{x}, t) := F(\mathbf{x}, P_t) \le 0. \tag{7}$$

Theorem 1 *Let $F(\mathbf{x}, t)$ be a concave function of $t \in [0, 1]$. Then the optimal value function of* (7)

$$\varphi(t) := \min_{\mathbf{x} \in \mathcal{X}} F_0(\mathbf{x}) \quad \textit{subject to} \quad F(\mathbf{x}, t) \le 0$$

is quasiconcave in $t \in [0, 1]$ with the lower bound

$$\varphi(t) \ge \min\{\varphi(0), \varphi(1)\}. \tag{8}$$

Proof For arbitrary $t_1,\ t_2 \in [0, 1]$ and $0 \le \lambda \le 1$ we have

$$\mathcal{X}((1 - \lambda)t_1 + \lambda t_2) \subset \{\mathbf{x} \in \mathcal{X} \mid (1 - \lambda)F(\mathbf{x}, t_1) + \lambda F(\mathbf{x}, t_2) \le 0\} \subset \mathcal{X}(t_1) \cup \mathcal{X}(t_2). \tag{9}$$

Hence, similarly as in Proposition 3.11 of Kyparisis and Fiacco (1987), the optimal value $\varphi(t)$ of (7) is quasiconcave which results in the lower bound (8). □

When also the objective function *depends* on the probability distribution, i.e. on the contamination parameter t, the problem is

$$\min_{\mathbf{x} \in \mathcal{X}} F_0(\mathbf{x}, t) := F_0(\mathbf{x}, P_t) \quad \text{subject to} \quad F(\mathbf{x}, t) \le 0. \tag{10}$$

For $F_0(\mathbf{x}, P)$ linear or concave in P, a lower bound can be obtained by application of the bound (8) separately to $F_0(\mathbf{x}, P)$ and $F_0(\mathbf{x}, Q)$:

$$\begin{aligned}\varphi(t) &= \min_{\mathbf{x}\in\mathcal{X}(t)} F_0(\mathbf{x}, (1-t)P + tQ) \geq \min_{\mathbf{x}\in\mathcal{X}(t)} [(1-t)F_0(\mathbf{x}, P) + tF_0(\mathbf{x}, Q)] \\ &\geq (1-t)\min\Big\{\varphi(0), \min_{\mathcal{X}(Q)} F_0(\mathbf{x}, P)\Big\} + t\min\Big\{\varphi(1), \min_{\mathcal{X}(P)} F_0(\mathbf{x}, Q)\Big\}. \end{aligned} \tag{11}$$

The bound is more complicated but still computable. It requires solution of 4 problems two of which are the non-contaminated programs for probability distributions P, Q and the other ones use both P and Q alternating in the objective function and constraints.

2.1.1 Comment

Of course, the lower bounds (8), (11) are loose, but for small values of t, say $t \leq t_0$ they can be improved to $\varphi(t) \geq \min\{\varphi(0), \varphi(t_0)\}$ when applied to P and to $\tilde{Q} := (1-t_0)P + t_0 Q$. Notice that no convexity assumption with respect to $\mathbf{x}$ is needed.

For *multiple constraints* and contaminated probability distribution it would be necessary to prove first the inclusion $\mathcal{X}(t) \subset \mathcal{X}(0) \cup \mathcal{X}(1)$ and then the lower bound (8) for the optimal value $\varphi(t) = \min_{\mathbf{x}\in\mathcal{X}(t)} F_0(\mathbf{x}, P_t)$ can be obtained as in the case of one constraint. As we shall see in Sect. 3.3, such inclusion holds true under special circumstances, otherwise we get only the following:

Denote $\mathcal{X}_j(t) = \{\mathbf{x}F \mid F_j(\mathbf{x}, P_t) \leq 0\}$. Then according to (9), $\mathcal{X}_j(t) \subset \mathcal{X}_j(0) \cup \mathcal{X}_j(1)$, hence

$$\mathcal{X}(t) \subset \mathcal{X} \cap \bigcap_j [\mathcal{X}_j(P) \cup \mathcal{X}_j(Q)] := \mathcal{X}_0.$$

To evaluate the corresponding lower bound $\min_{\mathbf{x}\in\mathcal{X}_0} F_0(\mathbf{x})$ would mean to solve a facial disjunctive program.

2.2 Directional derivative

Assume now that $F_j(\mathbf{x}, P)$, $j = 0, 1, \ldots, J$ in (1) are *convex* functions of $\mathbf{x}$. The directional derivative of the optimal value function can be obtained by the formula of Gol'shtein (1970), Theorem 17 applied to the Lagrange function

$$L(\mathbf{x}, \mathbf{u}, t) = F_0(\mathbf{x}, t) + \sum_j u_j F_j(\mathbf{x}, t)$$

provided that both the set of optimal solutions $\mathcal{X}^*(P) = \mathcal{X}^*(0)$ and the set of Lagrange multipliers $\mathcal{U}^*(P) = \mathcal{U}^*(0)$ are nonempty and bounded. If the functions F_j are *linear in* P, i.e. functions $F_j(\mathbf{x}, t)\ \forall j$ are linear in the contamination parameter t, then

$$\varphi'(0^+) = \min_{\mathbf{x}\in\mathcal{X}^*(0)} \max_{\mathbf{u}\in\mathcal{U}^*(0)} \frac{\partial}{\partial t} L(\mathbf{x}, \mathbf{u}, 0) = \min_{\mathbf{x}\in\mathcal{X}^*(0)} \max_{\mathbf{u}\in\mathcal{U}^*(0)} (L(\mathbf{x}, \mathbf{u}, Q) - L(\mathbf{x}, \mathbf{u}, P)). \tag{12}$$

Formula (12) simplifies substantially when $\mathcal{U}^*(0)$ is a singleton. When the constraints do not depend on P we get

$$\begin{aligned}\varphi'(0^+) &= \min_{\mathbf{x}\in\mathcal{X}^*(0)} \frac{\partial}{\partial t} F_0(\mathbf{x}, 0^+) = \min_{\mathbf{x}\in\mathcal{X}^*(0)} (F_0(\mathbf{x}, Q) - F_0(\mathbf{x}, P)) \\ &= \min_{\mathbf{x}\in\mathcal{X}^*(0)} F_0(\mathbf{x}, Q) - \varphi(0). \end{aligned} \tag{13}$$

These formulas can be exploited to construct an upper bound for the optimal value function $\varphi(t)$ of the form

$$\varphi(t) \le \varphi(0) + t\varphi'(0^+) \quad \forall t \in [0, 1] \tag{14}$$

provided that $\varphi(t)$ is *concave*; see e.g. Dupačová (1996, 2006), Dupačová and Polívka (2007). The contaminated probability distribution P_t may also be understood as a result of contaminating Q by P and an alternative upper bound may be constructed in a similar way.

Under additional assumptions, Theorem 17 of Gol'shtein (1970) provides a formula for derivative of the optimal value function also in case of nonlinear dependence of functions F_j on t. See Dupačová (1990, 1996, 1998) for details and applications for problems with a fixed set $\mathcal{X}$ of feasible solutions. The general nonconvex case is treated e.g. in Theorems 4.25 and 4.26 of Bonnans and Shapiro (2000).

Example 2 (Upper contamination bound for CVaR) With reference to Rockafellar and Uryasev (2002), Example 1 and Dupačová (2006), Dupačová and Polívka (2007) we shall use the formula

$$\mathrm{CVaR}_\alpha(\mathbf{x}, P) = \min_v \Phi_\alpha(\mathbf{x}, v, P) := v + \frac{1}{1-\alpha} E_P(f(\mathbf{x}, \omega) - v)^+$$

and apply the contamination technique to get an upper bound. It is an unconstrained optimization problem, the set $\mathcal{V}^*(\mathbf{x}, P)$ of its optimal solutions is a nonempty compact interval of $\mathbb{R}$, for a fixed $\mathbf{x}$ the objective function is convex in v and linear in P. Formula (13) for $\mathrm{CVaR}_\alpha(\mathbf{x}, (1-t)P + tQ)$ reduces to

$$\frac{\partial}{\partial t}\mathrm{CVaR}_\alpha(\mathbf{x}, 0^+) = \min_{v \in \mathcal{V}^*(\mathbf{x}, P)} \Phi_\alpha(\mathbf{x}, v, Q) - \mathrm{CVaR}_\alpha(\mathbf{x}, P). \tag{15}$$

The optimal value function, now $\mathrm{CVaR}_\alpha(\mathbf{x}, t) := \mathrm{CVaR}_\alpha(\mathbf{x}, (1-t)P + tQ)$ is a concave function of t, hence, its lower bound is $(1-t)\mathrm{CVaR}_\alpha(\mathbf{x}, P) + t\mathrm{CVaR}_\alpha(\mathbf{x}, Q)$. For an arbitrary optimal solution $v^*(\mathbf{x}, P) \in \mathcal{V}^*(\mathbf{x}, P)$, the upper bound for the contaminated CVaR value at $\mathbf{x}$ follows by substitution to (14):

$$\mathrm{CVaR}_\alpha(\mathbf{x}, (1-t)P + tQ) \le (1-t)\mathrm{CVaR}_\alpha(\mathbf{x}, P) + t\Phi_\alpha(\mathbf{x}, v^*(\mathbf{x}, P), Q). \tag{16}$$

2.3 Upper bound

To derive an upper bound for the optimal value of the contaminated problem with probability dependent constraints we shall confine ourselves mostly to the *expectation type of the objective function and constraints*. Hence, all functions $F_j(\mathbf{x}, t)$, $j = 0, \dots, J$, are linear in t on the interval $[0, 1]$. Denote $F(\mathbf{x}, P_t) = F(\mathbf{x}, t) := \max_j F_j(\mathbf{x}, t)$. For convex $F_j(\bullet, P)\ \forall j$ the "max" function $F(\bullet, P)$ is convex as well. This allows to rewrite the set $\mathcal{X}(t)$ of feasible solutions of (5) in the form

$$\mathcal{X}(t) = \mathcal{X} \cap \{\mathbf{x} : F(\mathbf{x}, t) \le 0\}$$

with one linearly perturbed convex constraint.

Assume first that $F(\mathbf{x}^*(0), P) = 0$ for an optimal solution $\mathbf{x}^*(0) := \mathbf{x}^*(P)$ of (1) and $F(\mathbf{x}^*(0), Q) \le 0$. Then at least one of the constraints is active at the optimal solution. Moreover, $\mathbf{x}^*(0) \in \mathcal{X}(t), \forall t \in [0, 1]$:

$$\begin{aligned} F(\mathbf{x}^*(0), t) &= \max_j [(1-t)F_j(\mathbf{x}^*(0), P) + tF_j(\mathbf{x}^*(0), Q)] \\ &\le (1-t)F(\mathbf{x}^*(0), P) + tF(\mathbf{x}^*(0), Q) \le 0. \end{aligned}$$

It means that there is a *trivial global upper bound*

$$\varphi(t) \le F_0(\mathbf{x}^*(0), t) \quad \forall t \in [0, 1]. \tag{17}$$

When $F_0(\mathbf{x}, \bullet)$ is linear, a more convenient form of (17) follows:

$$\varphi(t) \le F_0(\mathbf{x}^*(0), t) = (1-t)\varphi(0) + tF_0(\mathbf{x}^*(0), Q) \quad \forall t \in [0, 1] \tag{18}$$

otherwise one may apply suitable numerically tractable upper bounds for $F_0(\mathbf{x}^*(0), t)$; see Example 3.

If the above assumption $F(\mathbf{x}^*(0), P) = 0$ and $F(\mathbf{x}^*(0), Q) \le 0$ is not fulfilled, to get at least a local upper bound for $\varphi(t)$ valid for small t we shall switch to stability results for nonlinear parametric programming. Let $J_0 := \{j : F_j(\mathbf{x}^*(0), P) = 0\}$ be the set of indexes of active constraints of (1) at $\mathbf{x}^*(0)$.

In the *convex case*, it is possible to analyze the optimal value function by the first order methods. Various results in this direction can be mentioned: For example, according to Robinson (1987) the perturbed problem with a fixed convex polyhedral set $\mathcal{X}$ in (4) reduces locally to a problem with a parameter independent set of feasible solutions if $\mathbf{x}^*(0)$ is a nondegenerate point and the strict complementarity conditions hold true. In particular, $\mathbf{x}^*(0)$ is a nondegenerate point of (1) iff gradients $\nabla_x F_j(\mathbf{x}^*(0), P),\ j \in J_0$ are linearly independent, i.e. under the linear independence condition; cf. Bonnans and Shapiro (2000), Example 4.78. Then for t small enough, $t \le t_0$, $t_0 > 0$, the optimal value function $\varphi(t)$ is concave and its upper bound equals

$$\varphi(t) \le \varphi(0) + t\varphi'(0^+) \quad \forall t \in [0, t_0]. \tag{19}$$

A more detailed insight can be obtained if there is a continuous trajectory $[\mathbf{x}^*(t), \mathbf{u}^*(t)]$ of optimal solutions and Lagrange multipliers of the perturbed problem (5) emanating from the unique optimal solution $\mathbf{x}^*(0)$ and unique Lagrange multipliers $u_j^*(0), j = 1, \dots, J$ of (1). Such result follows usually by the implicit function theorem applied to the first order necessary conditions. In addition to the nondegeneracy and the strict complementarity conditions it requires also nonsingularity of the Hessian matrix of the Lagrange function on the tangent space to the active constraints, i.e. the second order sufficient condition valid at $\mathbf{x}^*(0), \mathbf{u}^*(0)$; see e.g. Bonnans and Shapiro (2000) or Fiacco (1983). At this point, convexity with respect to $\mathbf{x}$ is not needed and the trajectory $[\mathbf{x}^*(t), \mathbf{u}^*(t)]$ satisfies the first order optimality conditions also for $0 < t \le t_0$:

$$\begin{aligned} &F_j(\mathbf{x}^*(t), P_t) \le 0,\ u_j^*(t) \ge 0,\ F_j(\mathbf{x}^*(t), P_t)u_j^*(t) = 0, \quad j = 1, 2, \dots, J \\ &\nabla_x F_0(\mathbf{x}^*(t), P_t) + \sum_j u_j^*(t)\nabla_x F_j(\mathbf{x}^*(t), P_t) = 0. \end{aligned}$$

Moreover, for convex expectation type functionals F_j, $j = 0, \ldots, J$, the derivative (12) of the optimal value function reduces to

$$\begin{aligned}\varphi'(0^+) &= \frac{\partial}{\partial t} L(\mathbf{x}^*(0), \mathbf{u}^*(0), 0) = L(\mathbf{x}^*(0), \mathbf{u}^*(0), Q) - L(\mathbf{x}^*(0), \mathbf{u}^*(0), P) \\ &= F_0(\mathbf{x}^*(0), Q) + \sum_j u_j^*(0) F_j(\mathbf{x}^*(0), Q) - F_0(\mathbf{x}^*(0), P). \end{aligned} \tag{20}$$

If no constraint is active at $\mathbf{x}^*(0)$, we face a locally unconstrained optimization problem and the optimal value function $\varphi(t)$ is concave on a right neighborhood of 0, say for $t \in [0, t_0]$, $t_0 > 0$, hence, for $t \le t_0$, the upper bound (19) applies.

In the opposite case, the strict complementarity conditions imply that for small $t \in [0, t_0]$, $t_0 > 0$ the set J_0 of indexes of active constraints remains fixed and for a local analysis, constraints $F_j(\mathbf{x}, P) \le 0$ with $j \notin J_0$ need not be considered. Then $\mathcal{X}(t)$ reduces locally to the set of solutions of the system of equations $F_j(\mathbf{x}, t) = 0$, $j \in J_0$ which can be replaced locally by a parameter independent set.

To summarize – there exists $t_0 > 0$ such that for $0 \le t \le t_0$ the optimal value $\varphi(t)$ of the contaminated problem (5) can be obtained as $\varphi(t) = \min_{\mathbf{x} \in \mathcal{X}_0} F_0(\mathbf{x}, t)$ where the set of feasible solutions $\mathcal{X}_0$ does not depend on t. Hence, $\varphi(t)$ is concave on $[0, t_0]$, $t_0 > 0$ which opens the possibility of constructing local upper contamination bounds (19). Accordingly, the following theorem holds true:

Theorem 2 *Let* (1) *be a twice differentiable program*, $\mathbf{x}^*(P) = \mathbf{x}^*(0)$ *its optimal solution and* $\varphi(P) = \varphi(0)$ *its optimal value. Assume that at* $\mathbf{x}^*(0)$ *linear independence*, *the strict complementarity and the second order sufficient conditions are satisfied. Then there exists* $t_0 > 0$ *such that for all* $t \in [0, t_0]$ *the optimal value function* $\varphi(t)$ *is concave and the* local *upper contamination bound is given by*

$$\varphi(t) \le \varphi(0) + t\varphi'(0^+) \quad \forall t \in [0, t_0]. \tag{21}$$

Moreover, for convex expectation type problems (1) *the directional derivative is given by* (20).

2.3.1 Comment

Except for the form of the directional derivative, Theorem 2 applies also to problems with nonconvex functions $F_j(\bullet, P)\ \forall j$.

2.4 Illustrative examples

Consider $S = 50$ equiprobable scenarios of monthly returns $\boldsymbol{\varrho}$ of $N = 9$ assets (8 European stock market indexes: AEX, ATX, FCHI, GDAXI, OSEAX, OMXSPI, SSMI, FTSE and a risk free asset) in period June 2004–August 2008. The scenarios can be collected in the matrix

$$R = \begin{pmatrix} \mathbf{r}^1 \\ \mathbf{r}^2 \\ \vdots \\ \mathbf{r}^S \end{pmatrix}$$

Table 1 Descriptive statistics and the additional scenario of returns of 8 European stock indexes and of the risk free asset

Index	Country	Mean	Max	Min	A.S.
AEX	Netherlands	0.00456	0.07488	−0.14433	−0.19715
ATX	Austria	0.01358	0.13247	−0.14869	−0.23401
FCHI	France	0.0044	0.0615	−0.13258	−0.1005
GDAXI	Germany	0.01014	0.07111	−0.15068	−0.09207
OSEAX	Norway	0.01872	0.12176	−0.19505	−0.23934
OMXSPI	Sweden	0.00651	0.08225	−0.14154	−0.12459
SSMI	Switzerland	0.00563	0.05857	−0.09595	−0.08065
FTSE	England	0.00512	0.06755	−0.08938	−0.13024
Risk free		0.002	0.002	0.002	0.002

where $\mathbf{r}^s = (r_1^s, r_2^s, \ldots, r_N^s)$ is the s-th scenario. We will use $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)'$ for the vector of portfolio weights and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^N | \mathbf{1}'\boldsymbol{\lambda} = 1,\ \lambda_n \ \geq\ 0,\ n = 1, 2, \ldots, N\},$$

that is, the short sales are not allowed. The historical data comes from pre-crisis period. The data is contaminated by a scenario $\mathbf{r}^{S+1}$ from September 2008 when all indexes strongly fell down. The additional scenario can be understood as a stress scenario or the worst-case scenario. It can be seen in Table 1 presenting basic descriptive statistics of the original data and the additional scenario (A.S.).

We will apply the contamination bounds to mean-risk models with CVaR as a measure of risk. Two formulations are considered: In the first one, we are searching for a portfolio with minimal CVaR and at least the prescribed expected return, see e.g. Dupačová (2006) or Kilianová and Pflug (2009). Secondly, we minimize the expected loss of the portfolio under the condition that CVaR is below a given level, a special case of Example 1.

Example 3 (Minimizing CVaR) Mean-CVaR model with CVaR minimization is a special case of the general formulation (1) when $F_0(\mathbf{x}, P) = \mathrm{CVaR}(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$ and $F_1(\mathbf{x}, P) = E_P(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mu(P)$; $\mu(P)$ is the maximal allowable expected loss. We choose

$$\mu(P) = -E_P\boldsymbol{\varrho}'\left(\frac{1}{9}, \frac{1}{9}, \ldots, \frac{1}{9}\right)' = \frac{1}{50}\sum_{s=1}^{50} -\mathbf{r}^s\left(\frac{1}{9}, \frac{1}{9}, \ldots, \frac{1}{9}\right)'.$$

It means that the minimal required expected return is equal to the average return of the equally diversified portfolio. The significance level $\alpha = 0.95$ and Λ is a fixed convex polyhedral set representing constraints that do not depend on P. Since P is a discrete distribution with equiprobable scenarios $\mathbf{r}^1, \mathbf{r}^2, \ldots, \mathbf{r}^{50}$, using (3), the mean-CVaR model can be formulated as the following linear program:

$$\begin{aligned}\varphi(0) = &\min_{\lambda\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} v + \frac{1}{50 * 0.05}\sum_{s=1}^{50} z_s \\ \text{s.t.}\quad & z_s \geq -\mathbf{r}^s\boldsymbol{\lambda} - v, \quad s = 1, 2, \ldots, 50 \\ & \frac{1}{50}\sum_{s=1}^{50} -\mathbf{r}^s\boldsymbol{\lambda} - \mu(P) \leq 0.\end{aligned} \tag{22}$$

By analogy, for the additional scenario we have:

$$\begin{aligned} \varphi(1) &= \min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z\in\mathbb{R}^+} v + \frac{1}{0.05} z \\ \text{s.t.} \quad & z \geq -\mathbf{r}^{51}\boldsymbol{\lambda} - v, \ -\mathbf{r}^{51}\boldsymbol{\lambda} - \mu(Q) \leq 0 \end{aligned} \tag{23}$$

or, equivalently:

$$\varphi(1) = \min_{\boldsymbol{\lambda}\in\Lambda} \{-\mathbf{r}^{51}\boldsymbol{\lambda} \mid -\mathbf{r}^{51}\boldsymbol{\lambda} - \mu(Q) \leq 0\} \tag{24}$$

where $\mu(Q) = -\mathbf{r}^{51}(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9})'$.

First, we compute for $t \in [0, 1]$ the optimal value function of the contaminated problem.

$$\begin{aligned} \varphi(t) &= \min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} v + \frac{1}{0.05}\left(\sum_{s=1}^{50} \frac{1}{50}(1-t) z_s + t z_{51}\right) \\ \text{s.t.} \quad & z_s \geq -\mathbf{r}^s\boldsymbol{\lambda} - v, \quad s = 1, 2, \dots, 51 \\ & -\sum_{s=1}^{50} \frac{1}{50}(1-t)\mathbf{r}^s\boldsymbol{\lambda} - t\mathbf{r}^{51}\boldsymbol{\lambda} - \mu((1-t)P + tQ) \leq 0 \end{aligned} \tag{25}$$

where $\mu((1-t)P + tQ) = -\sum_{s=1}^{50} \frac{1}{50}(1-t)\mathbf{r}^s(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9})' - t\mathbf{r}^{51}(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9})'$.

Secondly, applying (11), we derive a *lower bound* for $\varphi(t)$. Note that now

$$\begin{aligned} \min_{\mathcal{X}(Q)} F_0(\mathbf{x}, P) &= \min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} v + \frac{1}{50 * 0.05} \sum_{s=1}^{50} z_s \\ \text{s.t.} \quad & z_s \geq -\mathbf{r}^s\boldsymbol{\lambda} - v, \quad s = 1, 2, \dots, 50 \\ & -\mathbf{r}^{51}\boldsymbol{\lambda} - \mu(Q) \leq 0 \end{aligned}$$

and

$$\min_{\mathcal{X}(P)} F_0(\mathbf{x}, Q) = \min_{\boldsymbol{\lambda}\in\Lambda} \left\{ -\mathbf{r}^{51}\boldsymbol{\lambda} \mid \frac{1}{50} \sum_{s=1}^{50} -\mathbf{r}^s\boldsymbol{\lambda} - \mu(P) \leq 0 \right\}.$$

Finally, we construct an *upper bound* for $\varphi(t)$. Since the optimal solution $\boldsymbol{\lambda}^*$ of (22) is a feasible solution of (23) we can apply (17) with $\mathbf{x}^*(0) = \boldsymbol{\lambda}^*$ as a trivial upper bound for all $t \in [0, 1]$:

$$\begin{aligned} \varphi(t) \leq F_0(\mathbf{x}^*(0), t) &= \min_{v\in\mathbb{R}, z_s\in\mathbb{R}^+} v + \frac{1}{0.05}\left(\sum_{s=1}^{50} \frac{1}{50}(1-t) z_s + t z_{51}\right) \\ \text{s.t.} \quad & z_s \geq -\mathbf{r}^s\boldsymbol{\lambda}^* - v, \quad s = 1, 2, \dots, 51. \end{aligned}$$

The disadvantage of this trivial bound is the fact, that it would require evaluation of the CVaR for $\boldsymbol{\lambda}^*$ for each t. Linearity with respect to t does not hold true, but we may apply the bound (16). This yields an upper estimate for $F_0(\mathbf{x}^*(0), t)$ which is a convex combination of $\varphi(0)$ and $\Phi_\alpha(\mathbf{x}^*(0), v^*(\mathbf{x}^*(0), P), Q)$. The optimal value $\varphi(0)$ is given by (22) and

$$\Phi_\alpha(\mathbf{x}^*(0), v^*(\mathbf{x}^*(0), P), Q) = v^* + \frac{1}{0.05}(-\mathbf{r}^{51}\boldsymbol{\lambda}^* - v^*)^+$$

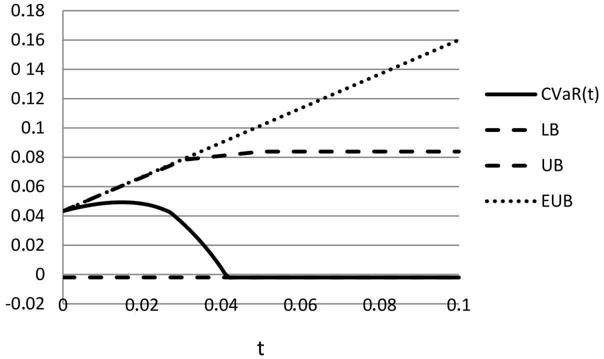


Fig. 1 Comparison of minimal (CVaR(t)) value of mean-CVaR model with lower bound (LB), upper bound (UB) and the estimated upper bound (EUB)

where v^* and $\boldsymbol{\lambda}^*$ are optimal solutions of (22). The graphs of $\varphi(t)$, its lower bound and two upper bounds (trivial one and its upper estimate) for small contamination $t \in [0, 0.1]$ are presented in Fig. 1. Since all original scenarios have probability 0.02, the performance for $t > 0.1$ is not of much interest. For $t > 0.04$, $\varphi(t)$ in (25) coincides with its lower bound because the optimal portfolios consist only of risk free asset. The upper bound is piecewise linear in t and for small values of t it coincides with the estimated upper bound.

Example 4 (Minimizing expected loss) As the second example, consider the mean-CVaR model minimizing the expected loss subject to a constraint on CVaR. This corresponds to (1) with $F_0(\mathbf{x}, P) = E_P(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$ and $F_1(\mathbf{x}, P) = \text{CVaR}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - c$ where $c = 0.19$ is the maximal accepted level of CVaR. For simplicity, this level does not depend on the probability distribution. Similarly to the previous example, we compute the optimal value $\varphi(t)$ and its lower and upper bound. Using Theorem 16 of Rockafellar and Uryasev (2002), the minimal CVaR-constrained expected loss is obtained for $t \in [0, 1]$ as

$$\varphi(t) = \min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}} -\sum_{s=1}^{50} \frac{1}{50}(1-t)\mathbf{r}^s\boldsymbol{\lambda} - t\mathbf{r}^{51}\boldsymbol{\lambda} \tag{26}$$

$$\text{s.t.} \quad v + \frac{1}{0.05}\left(-\sum_{s=1}^{50} \frac{1}{50}(1-t)\mathbf{r}^s\boldsymbol{\lambda} - t\mathbf{r}^{51}\boldsymbol{\lambda} - v\right)^+ - c \le 0 \tag{27}$$

and equals thus the optimal value function of the parametric linear program

$$\begin{aligned} \varphi(t) = &\min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} -\sum_{s=1}^{50} \frac{1}{50}(1-t)\mathbf{r}^s\boldsymbol{\lambda} - t\mathbf{r}^{51}\boldsymbol{\lambda} \\ &\text{s.t.} \quad v + \frac{1}{0.05}\left(\sum_{s=1}^{50} \frac{1}{50}(1-t)z_s + tz_{51}\right) - c \le 0 \\ &\qquad z_s \ge -\mathbf{r}^s\boldsymbol{\lambda} - v, \quad s = 1, 2, \ldots, 51 \end{aligned} \tag{28}$$

for $t \in [0, 1]$. In particular, for $t = 1$ we have

$$\begin{aligned} \varphi(1) = &\min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} -\mathbf{r}^{51}\boldsymbol{\lambda} \\ &\text{s.t.} \quad v + \frac{1}{0.05}z_{51} - c \le 0, \; z_{51} + v \ge -\mathbf{r}^{51}\boldsymbol{\lambda}, \end{aligned}$$

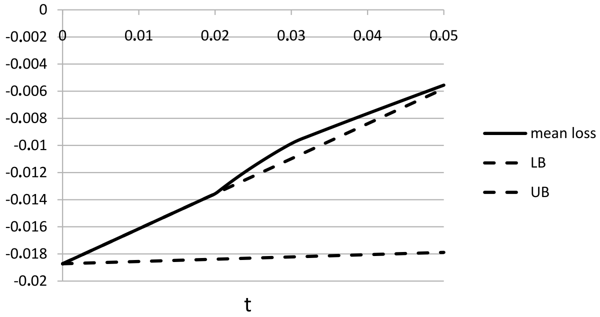


Fig. 2 Comparison of minimal mean loss value with its lower bound (LB) and upper bound (UB)

what is equivalent to

$$\varphi(1) = \min_{\boldsymbol{\lambda}\in\Lambda}\{-\mathbf{r}^{51}\boldsymbol{\lambda} \mid -\mathbf{r}^{51}\boldsymbol{\lambda} - c \le 0\};$$

compare with (24). Using (11), we can evaluate the *lower bound* for $\varphi(t)$ with

$$\min_{\mathcal{X}(Q)} F_0(\mathbf{x}, P) = \min_{\boldsymbol{\lambda}\in\Lambda}\left\{-\sum_{s=1}^{50}\frac{1}{50}\mathbf{r}^s\boldsymbol{\lambda} \mid -\mathbf{r}^{51}\boldsymbol{\lambda} - c \le 0\right\}$$

and

$$\min_{\mathcal{X}(P)} F_0(\mathbf{x}, Q) = \min_{\boldsymbol{\lambda}\in\Lambda, v\in\mathbb{R}, z_s\in\mathbb{R}^+} -\mathbf{r}^{51}\boldsymbol{\lambda}$$

$$\text{s.t.} \quad v + \frac{1}{0.05}\sum_{s=1}^{50}\frac{1}{50}z_s - c \le 0, \quad z_s \ge -\mathbf{r}^s\boldsymbol{\lambda} - v, \quad s = 1, 2, \dots, 50.$$

Finally, we compute an *upper bound* for $\varphi(t)$. Contrary to the previous example, the optimal solution $\mathbf{x}^*(0)$ of the noncontaminated problem is not a feasible solution of the fully contaminated problem. Therefore, the trivial global upper bound (17) cannot be used. We apply instead the local upper bound (21) with the directional derivative (20). In this example, the value of multiplier $\mathbf{u}^*(0)$ corresponding to (27) for $t = 0$ is equal to zero, the CVaR constraint (27) is not active and for sufficiently small t, the upper bound reduces to:

$$\varphi(t) \le (1-t)\varphi(0) + tF_0(\mathbf{x}^*(0), Q). \tag{29}$$

Figure 2 depicts the graph of $\varphi(t)$ given by (28) and its lower and upper bound. The upper bound coincides with $\varphi(t)$ for $t \le 0.02$. It illustrates the fact that the local upper bound is meaningful if the probability of the additional scenario is not too large, i.e. no more than probabilities of the original scenarios for our example.

3 Robustness in portfolio efficiency testing

3.1 Portfolio efficiency test

In this section, we shall study robustness of portfolio efficiency tests with respect to the second-order stochastic dominance relation. Consider N assets and a random vector of their returns $\boldsymbol{\varrho}$. Since all existing portfolio efficiency tests have been derived for a discrete probability distribution P of returns we assume that $\boldsymbol{\varrho}$ takes S values $\mathbf{r}^s = (r_1^s, r_2^s, \dots, r_N^s)$, called scenarios, with probabilities $p_1, p_2, \dots, p_S$. Contrary to all former tests, e.g. Kopa and Chovanec (2008) or Kopa (2010), *we do not assume equiprobable scenarios.* Again, the scenarios are collected in the matrix

$$R = \begin{pmatrix} \mathbf{r}^1 \\ \mathbf{r}^2 \\ \vdots \\ \mathbf{r}^S \end{pmatrix}$$

and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^N | \mathbf{1}'\boldsymbol{\lambda} = 1,\ \lambda_n \geq 0,\ n = 1, 2, \dots, N\}.$$

Alternatively, one can consider any bounded polytope: $\Lambda' = \{\boldsymbol{\lambda} \in R^N | A\boldsymbol{\lambda} \geq \mathbf{b}\}$.

For any portfolio $\boldsymbol{\lambda} \in \Lambda$, let $(-R\boldsymbol{\lambda})^{[k]}$ be the k-th smallest element of $(-R\boldsymbol{\lambda})$, i.e. $(-R\boldsymbol{\lambda})^{[1]} \leq (-R\boldsymbol{\lambda})^{[2]} \leq \cdots \leq (-R\boldsymbol{\lambda})^{[S]}$ and let $I(\boldsymbol{\lambda})$ be a permutation of the index set $I = \{1, 2, \dots, S\}$ such that $-\mathbf{r}^{i(\boldsymbol{\lambda})}\boldsymbol{\lambda} = (-R\boldsymbol{\lambda})^{[i]}$. Accordingly, we can order the corresponding probabilities and we denote $p_i^{\boldsymbol{\lambda}} = p_{i(\boldsymbol{\lambda})}$. Hence, $p_i^{\boldsymbol{\lambda}} = P(-\boldsymbol{\varrho}\boldsymbol{\lambda} = (-R\boldsymbol{\lambda})^{[i]})$. The same notation is applied for the tested portfolio $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_N)'$.

Let $F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)$ denote the cumulative probability distribution function of returns of portfolio $\boldsymbol{\lambda}$. The twice cumulative probability distribution function of returns of portfolio $\boldsymbol{\lambda}$ is defined as

$$F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(y) = \int_{-\infty}^{y} F_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(x)\, dx. \tag{30}$$

Following Ruszczyński and Vanderbei (2003), Kuosmanen (2004), Kopa and Chovanec (2008) and Kopa (2010), we define the second-order stochastic dominance relation in the strict form in the context of SSD portfolio efficiency.

Definition 1 Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by the second-order stochastic dominance ($\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$) if and only if

$$F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\lambda}}(y) \leq F^{(2)}_{\boldsymbol{\varrho}'\boldsymbol{\tau}}(y) \quad \forall y \in \mathbb{R}$$

with strict inequality[1] for at least one $y \in \mathbb{R}$.

As in Ogryczak and Ruszczyński (2002) or Kopa and Chovanec (2008), we express the SSD relation using the conditional value at risk (CVaR).

Lemma 1 *Let $\boldsymbol{\lambda}, \boldsymbol{\tau} \in \Lambda$. Then $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$ if and only if*

$$\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \leq \mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \quad \textit{for all } \alpha \in [0, 1] \tag{31}$$

with strict inequality for at least one α.

[1] This type of SSD relation is sometimes referred to as the strict second-order stochastic dominance. If no strict inequality is required then the relation can be called the weak second-order stochastic dominance.

Since we limit our attention to a discrete probability distribution of returns, the inequality of CVaRs need not be verified in all $\alpha \in [0, 1]$, but only in at most $S+1$ particular points.

Theorem 3 *Let* $q_s^{\boldsymbol{\lambda}} = \sum_{i=1}^{s} p_i^{\boldsymbol{\lambda}}$ *and* $q_s^{\boldsymbol{\tau}} = \sum_{i=1}^{s} p_i^{\boldsymbol{\tau}}$, $s = 1, 2, \dots, S$. *Let* $q_0^{\boldsymbol{\lambda}} = q_0^{\boldsymbol{\tau}} = 0$. *Then* $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$ *if and only if* $\mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \le \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\tau})$ *for all* $s = 0, 1, 2, \dots, S$ *with strict inequality for at least one* $q_s^{\boldsymbol{\lambda}}$.

Proof Assume $\alpha > 0$. Following Rockafellar and Uryasev (2002), Proposition 8, let $s(\alpha)$ be the unique index such that $q_{s(\alpha)}^{\boldsymbol{\lambda}} \ge \alpha > q_{s(\alpha)-1}^{\boldsymbol{\lambda}}$. Then

$$\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) = \frac{1}{1-\alpha}\left[(q_{s(\alpha)}^{\boldsymbol{\lambda}} - \alpha)(-R\boldsymbol{\lambda})^{[s(\alpha)]} + \sum_{i=s(\alpha)+1}^{S} p_i^{\boldsymbol{\lambda}}(-R\boldsymbol{\lambda})^{[i]}\right].$$

Consider $\mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) := (1-\alpha)\mathrm{CVaR}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$. Since $1 - q_{s(\alpha)}^{\boldsymbol{\lambda}} = \sum_{i=s(\alpha)+1}^{S} p_i^{\boldsymbol{\lambda}}$ we have:

$$\begin{aligned}
\mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) &= q_{s(\alpha)}^{\boldsymbol{\lambda}}(-R\boldsymbol{\lambda})^{[s(\alpha)]} - \alpha(-R\boldsymbol{\lambda})^{[s(\alpha)]} + \sum_{i=s(\alpha)+1}^{S} p_i^{\boldsymbol{\lambda}}(-R\boldsymbol{\lambda})^{[i]} \\
&= (1-\alpha)(-R\boldsymbol{\lambda})^{[s(\alpha)]} - (-R\boldsymbol{\lambda})^{[s(\alpha)]}(1 - q_{s(\alpha)}^{\boldsymbol{\lambda}}) + \sum_{i=s(\alpha)+1}^{S} p_i^{\boldsymbol{\lambda}}(-R\boldsymbol{\lambda})^{[i]} \\
&= (1-\alpha)(-R\boldsymbol{\lambda})^{[s(\alpha)]} + \sum_{i=s(\alpha)+1}^{S} p_i^{\boldsymbol{\lambda}}\big((-R\boldsymbol{\lambda})^{[i]} - (-R\boldsymbol{\lambda})^{[s(\alpha)]}\big).
\end{aligned}$$

A similar analysis can be done for portfolio $\boldsymbol{\tau}$. Since both $\mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda})$ and $\mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau})$ are concave piecewise linear functions in α, Lemma 1 implies that $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$ if and only if $\mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) \le \mathrm{LC}_\alpha(-\boldsymbol{\varrho}'\boldsymbol{\tau})$ for all $\alpha = q_s^{\boldsymbol{\lambda}}$, $s = 0, 1, \dots, S$, with strict inequality for at least one $q_s^{\boldsymbol{\lambda}}$. Passing back to CVaR expressions completes the proof. □

Following Ruszczyński and Vanderbei (2003), Kuosmanen (2004), Kopa and Chovanec (2008) and Kopa (2010) we define portfolio efficiency with respect to the second order stochastic dominance.

Definition 2 A given portfolio $\boldsymbol{\tau} \in \Lambda$ is SSD inefficient if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is SSD efficient.

This definition classifies portfolio $\boldsymbol{\tau} \in \Lambda$ as SSD efficient if and only if no other portfolio is better (in the sense of the SSD relation) for all risk averse and risk neutral decision makers. Inspired by Kopa and Chovanec (2008) we consider the following measure:

$$\begin{aligned}
&\xi(\boldsymbol{\tau}, R, \mathbf{p}) = \min_{a_s, \boldsymbol{\lambda}} \sum_{s=0}^{S} a_s \\
&\text{s.t.} \quad \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{q_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \le a_s, \quad s = 0, 1, \dots, S \\
&\qquad a_s \le 0, \quad s = 0, 1, \dots, S \\
&\qquad \boldsymbol{\lambda} \in \Lambda.
\end{aligned} \tag{32}$$

The objective function of (32) represents the sum of differences between CVaRs of a portfolio $\boldsymbol{\lambda}$ and CVaRs of the tested portfolio $\boldsymbol{\tau}$. The differences are considered in points $q_s^{\boldsymbol{\lambda}}$, $s = 0, 1, \dots, S$. All differences must be non-positive and at least one negative to guarantee that portfolio $\boldsymbol{\lambda}$ dominates portfolio $\boldsymbol{\tau}$. Moreover, minimizing these differences, we find portfolio $\boldsymbol{\lambda}^*$ that cannot be dominated by any other one. On the other hand, if no dominating portfolio exists, that is, portfolio $\boldsymbol{\tau}$ is SSD efficient, then $\xi(\boldsymbol{\tau}, R, \mathbf{p}) = 0$ because the only feasible solutions of (32) are $\boldsymbol{\tau}$ and portfolios $\overline{\boldsymbol{\lambda}}$ satisfying $R\overline{\boldsymbol{\lambda}} = R\boldsymbol{\tau}$. Summarizing, Theorem 3 implies the following necessary and sufficient SSD portfolio efficiency test:

Theorem 4 *A given portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if $\xi(\boldsymbol{\tau}, R, \mathbf{p}) = 0$. If $\xi(\boldsymbol{\tau}, R, \mathbf{p}) < 0$ then the optimal portfolio $\boldsymbol{\lambda}^*$ in (32) is SSD efficient and it dominates portfolio $\boldsymbol{\tau}$ by SSD.*

Until now, perfect information about the probability distribution of returns was assumed and portfolio $\boldsymbol{\tau}$ was tested with respect to this distribution. However, in many practical applications, the probability distribution of returns is not perfectly known. And therefore, we will study robust versions of SSD efficiency.

3.2 Portfolio efficiency with respect to ϵ-SSD relation

Assume that the probability distribution $\bar{P}$ of random returns $\bar{\boldsymbol{\varrho}}$ takes again values $\mathbf{r}^s$, $s = 1, 2, \dots, S$ but with other probabilities $\bar{\mathbf{p}} = (\bar{p}_1, \bar{p}_2, \dots, \bar{p}_S)$. We define the distance between P and $\bar{P}$ as $d(\bar{P}, P) = \max_i |\bar{p}_i - p_i|$.

Definition 3 A given portfolio $\boldsymbol{\tau} \in \Lambda$ is ϵ-SSD inefficient if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ and $\bar{P}$ such that $d(\bar{P}, P) \le \epsilon$ with $\bar{\boldsymbol{\varrho}}'\boldsymbol{\lambda} \succ_{SSD} \bar{\boldsymbol{\varrho}}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is ϵ-SSD efficient.

The introduced ϵ-SSD efficiency is a robustification of the classical SSD portfolio efficiency. It guarantees stability of the SSD efficiency classification with respect to small changes (prescribed by parameter ϵ) in probability vector $\mathbf{p}$. A given portfolio $\boldsymbol{\tau}$ is ϵ-SSD efficient if and only if no portfolio $\boldsymbol{\lambda}$ SSD dominates $\boldsymbol{\tau}$ neither for the original probabilities $\mathbf{p}$ nor for arbitrary probabilities $\bar{\mathbf{p}}$ from ϵ-neighborhood of the original vector $\mathbf{p}$. For testing ϵ-SSD efficiency of a given portfolio $\boldsymbol{\tau}$ we modify (32) in order to introduce a new measure of ϵ-SSD efficiency:

$$
\begin{aligned}
\xi_\epsilon(\boldsymbol{\tau}, R, \mathbf{p}) = {} & \min_{a_s, \boldsymbol{\lambda}, \bar{\mathbf{p}}} \sum_{s=0}^{S} a_s \\
\text{s.t.} \quad & \mathrm{CVaR}_{\bar{q}_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\bar{q}_s^{\boldsymbol{\lambda}}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) \le a_s, \quad s = 0, 1, \dots, S \\
& \bar{q}_s^{\boldsymbol{\lambda}} = \sum_{i=1}^{s} \bar{p}_i^{\boldsymbol{\lambda}}, \quad s = 1, \dots, S \\
& \bar{q}_0^{\boldsymbol{\lambda}} = 0 \\
& \sum_{i=1}^{S} \bar{p}_i = 1 \\
& -\epsilon \le \bar{p}_i - p_i \le \epsilon, \quad i = 1, 2, \dots, S \\
& \bar{p}_i \ge 0, \quad i = 1, 2, \dots, S
\end{aligned}
$$

$$a_s \le 0, \quad s = 0, 1, \ldots, S$$
$$\boldsymbol{\lambda} \in \Lambda. \tag{33}$$

Theorem 5 *Portfolio $\boldsymbol{\tau} \in \Lambda$ is ϵ-SSD efficient if and only if $\xi_\epsilon(\boldsymbol{\tau}, R, \mathbf{p})$ given by (33) is equal to zero.*

Proof The proof directly follows from Theorem 4 because (33) is obtained from (32) by an additional minimization over $\bar{\mathbf{p}}$ from ϵ-neighborhood of the original probability vector $\mathbf{p}$. □

3.3 Resistance of SSD portfolio efficiency with respect to additional scenarios

In the previous sections, we assumed a fixed set of scenarios. In many practical applications, an additional scenario may be of interest. Therefore, the aim of this section is to analyze the robustness of SSD portfolio efficiency with respect to the additional scenario denoted by $\mathbf{r}^{S+1}$. For a contamination parameter $t \in [0, 1]$, we assume that the random return $\tilde{\boldsymbol{\varrho}}(t)$ takes values $\mathbf{r}^1, \mathbf{r}^2, \ldots, \mathbf{r}^{S+1}$ with probabilities $\tilde{\mathbf{p}}(t) = ((1-t)p_1, (1-t)p_2, \ldots, (1-t)p_S, t)$. The cumulative probabilities for portfolio $\boldsymbol{\lambda}$ are

$$\tilde{q}_s^{\boldsymbol{\lambda}} = \sum_{i=1}^{s} \tilde{p}_i^{\boldsymbol{\lambda}} = \sum_{i=1}^{s} P(-\tilde{\boldsymbol{\varrho}}(t)\boldsymbol{\lambda} = (-\tilde{R}\boldsymbol{\lambda})^{[i]}), \quad s = 1, 2, \ldots, S+1, \quad \tilde{q}_0^{\boldsymbol{\lambda}} = 0$$

and the same notation is used for portfolio $\boldsymbol{\tau}$. We denote the extended scenario matrix by $\tilde{R}$, that is,

$$\tilde{R} = \begin{pmatrix} R \\ \mathbf{r}^{S+1} \end{pmatrix}.$$

Definition 4 A given portfolio $\boldsymbol{\tau} \in \Lambda$ is directionally SSD inefficient with respect to $\mathbf{r}^{S+1}$ if it exists $t_0 > 0$ such that for every $t \in [0, t_0]$ there is a portfolio $\boldsymbol{\lambda}(t) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}(t) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$.

Definition 5 A given portfolio $\boldsymbol{\tau} \in \Lambda$ is directionally SSD efficient with respect to $\mathbf{r}^{S+1}$ if there exists $t_0 > 0$ such that for arbitrary $t \in [0, t_0]$ there is no portfolio $\boldsymbol{\lambda}(t) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}(t) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$.

According to these definitions, a given portfolio is classified as directionally SSD efficient (inefficient) with respect to scenario $\mathbf{r}^{S+1}$ if it is SSD efficient (inefficient) and a sufficiently small contamination of the original probability distribution of returns by the additional scenario does not change the SSD efficiency classification, that is, the SSD efficient (inefficient) portfolio remains SSD efficient (inefficient). Applying (32) to contaminated data, portfolio $\boldsymbol{\lambda}(t) \in \Lambda$ satisfying $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}(t) \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$ exists if and only if $\xi(\boldsymbol{\tau}, \tilde{R}, \tilde{\mathbf{p}}(t)) < 0$, where

$$\xi(\boldsymbol{\tau}, \tilde{R}, \tilde{\mathbf{p}}(t)) = \min_{a_s, \boldsymbol{\lambda}} \sum_{s=0}^{S} a_s$$
$$\text{s.t.} \quad \mathrm{CVaR}_{\tilde{q}_s^{\boldsymbol{\lambda}}}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\tilde{q}_s^{\boldsymbol{\lambda}}}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}) \le a_s, \quad s = 0, 1, \ldots, S$$
$$a_s \le 0, \quad s = 0, 1, \ldots, S$$
$$\boldsymbol{\lambda} \in \Lambda. \tag{34}$$

Example 5 (a) Consider the following three assets and three scenarios example:

$$R = \begin{pmatrix} 0 & 3 & 2 \\ 2 & 2 & 2 \\ 4 & 1 & 2 \end{pmatrix}.$$

Assume that scenarios are equiprobable. It can be shown that portfolio $\boldsymbol{\tau} = (\frac{1}{3}, \frac{2}{3}, 0)$ is SSD efficient. Let the additional scenario $\mathbf{r}^4 = (0, 0, 2)$ and consider portfolio $\boldsymbol{\lambda} = (0, 0, 1)$. Then $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda} \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$ for any contamination parameter $t > 0$. Hence, portfolio $\boldsymbol{\tau}$ is SSD efficient but not directionally SSD efficient with respect to scenario $\mathbf{r}^4$.

(b) Consider another three assets and three scenarios example:

$$R = \begin{pmatrix} 0 & 3 & 2 \\ 2 & 2 & 3 \\ 4 & 1 & 2 \end{pmatrix}.$$

Assume again that scenarios are equiprobable. It can be shown that portfolio $\boldsymbol{\tau} = (\frac{1}{3}, \frac{2}{3}, 0)$ is SSD inefficient, because portfolio $\boldsymbol{\lambda} = (0, 0, 1)$ SSD dominates portfolio $\boldsymbol{\tau}$. Let the additional scenario $\mathbf{r}^4 = (2, 2, 0)$. Then no portfolio SSD dominates $\boldsymbol{\tau} = (\frac{1}{3}, \frac{2}{3}, 0)$ for any contamination parameter $t > 0$. Hence, portfolio $\boldsymbol{\tau}$ is SSD inefficient but not directionally SSD inefficient with respect to scenario $\mathbf{r}^4$.

Example 5 shows that there are situations where an arbitrarily small contamination of the original probability distribution of returns leads to the opposite SSD classification. Using contamination bounds we will derive a sufficient condition for directional SSD efficiency and directional SSD inefficiency with respect to additional scenario $\mathbf{r}^{S+1}$.

Theorem 6 *Let $\boldsymbol{\tau} \in \Lambda$ be an SSD efficient portfolio for the noncontaminated distribution P. Let*

$$\mathbf{r}^{S+1}\boldsymbol{\tau} \geq \mathbf{r}^{S+1}\boldsymbol{\lambda} \quad \text{for all } \boldsymbol{\lambda} \in \Lambda. \tag{35}$$

Then $\boldsymbol{\tau} \in \Lambda$ is directionally SSD efficient with respect to $\mathbf{r}^{S+1}$.

Proof The SSD efficiency of $\boldsymbol{\tau}$ implies that $\xi(\boldsymbol{\tau}, R, \mathbf{p}) = 0$. Condition (35) gives $\xi(\boldsymbol{\tau}, \mathbf{r}^{S+1}, 1) = 0$. Since the objective function of (32) does not depend on probability distribution, verification of (9) for $t_1 = 0$, $t_2 = 1$ will imply the lower bound (8). Consequently, $\xi(\boldsymbol{\tau}, \tilde{R}, \tilde{\mathbf{p}}(t))$ will necessarily be equal to zero for all $t \in [0, 1]$ what yields directional SSD efficiency with respect to $\mathbf{r}^{S+1}$ of $\boldsymbol{\tau}$. Hence, it suffices to show, that any feasible solution $\boldsymbol{\lambda}$ of (34) with an arbitrary parameter $t \in (0, 1)$ is a feasible solution of (32). Let $F^{(2)}_{\boldsymbol{\lambda},S+1}(z)$ be a cumulative distribution function of returns of portfolio $\boldsymbol{\lambda}$ for the contaminated distribution taking $S + 1$ scenarios with probabilities $\tilde{\mathbf{p}} = ((1-t)p_1, (1-t)p_2, \ldots, (1-t)p_S, t)$. Similarly, let $F^{(2)}_{\boldsymbol{\lambda},S}(z)$ correspond to the original distribution with S scenarios. Then

$$\begin{aligned} F^{(2)}_{\boldsymbol{\lambda},S+1}(z) &= \int_{-\infty}^{z} F_{\boldsymbol{\lambda}}(y)\,dy = \int_{-\infty}^{z} \sum_{s=1}^{S+1} \tilde{p}_s 1_{(\mathbf{r}^s\boldsymbol{\lambda} \leq y)}\,dy \\ &= \sum_{s=1}^{S+1} \tilde{p}_s (z - \mathbf{r}^s\boldsymbol{\lambda}) 1_{(\mathbf{r}^s\boldsymbol{\lambda} \leq z)} = \sum_{s=1}^{S+1} \tilde{p}_s (z - \mathbf{r}^s\boldsymbol{\lambda})^+. \end{aligned} \tag{36}$$

The same notation and analysis is applied to portfolio $\boldsymbol{\tau}$.

Since $\boldsymbol{\lambda}$ is a feasible solution of (34), Theorem 3 implies that $\boldsymbol{\varrho}(t)'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}(t)'\boldsymbol{\tau}$. Hence, directly from Definition 1, one obtains

$$F^{(2)}_{\boldsymbol{\lambda},S+1}(z) \leq F^{(2)}_{\boldsymbol{\tau},S+1}(z) \quad \forall z \in \mathbb{R}. \tag{37}$$

Applying (36) to (37)

$$\sum_{s=1}^{S}(1-t)p_s(z-\mathbf{r}^s\boldsymbol{\lambda})^+ + t(z-\mathbf{r}^{S+1}\boldsymbol{\lambda})^+ \leq \sum_{s=1}^{S}(1-t)p_s(z-\mathbf{r}^s\boldsymbol{\tau})^+ + t(z-\mathbf{r}^{S+1}\boldsymbol{\tau})^+. \tag{38}$$

Note that according to (35) $(z-\mathbf{r}^{S+1}\boldsymbol{\lambda})^+ \geq (z-\mathbf{r}^{S+1}\boldsymbol{\tau})^+$. Combining it with (38) implies that $\sum_{s=1}^{S} p_s(z-\mathbf{r}^s\boldsymbol{\lambda})^+ \leq \sum_{s=1}^{S} p_s(z-\mathbf{r}^s\boldsymbol{\tau})^+$. Therefore

$$F^{(2)}_{\boldsymbol{\lambda},S}(z) \leq F^{(2)}_{\boldsymbol{\tau},S}(z) \quad \forall z \in \mathbb{R}.$$

According to Definition 1, $\boldsymbol{\varrho}'\boldsymbol{\lambda} \succ_{SSD} \boldsymbol{\varrho}'\boldsymbol{\tau}$ and the rest of the proof directly follows from Theorem 3. □

In Example 5(a), $\xi(\boldsymbol{\tau}, \mathbf{r}^4, 1) = -2$ and $\xi(\boldsymbol{\tau}, \tilde{R}(t), \tilde{\mathbf{p}}(t)) < 0$ for all $t \in (0,1]$ because $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda} \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$ for all $t \in (0,1]$.

Theorem 7 *Let $\boldsymbol{\tau} \in \Lambda$ be an SSD inefficient portfolio for the noncontaminated distribution P. If there exists a portfolio $\boldsymbol{\lambda} \in \Lambda$ such that*

$$\mathrm{CVaR}_{q_s^{\lambda}}(-\boldsymbol{\varrho}'\boldsymbol{\lambda}) - \mathrm{CVaR}_{q_s^{\lambda}}(-\boldsymbol{\varrho}'\boldsymbol{\tau}) < 0, \quad s = 0, 1, \ldots, S \tag{39}$$

$$\mathbf{r}^{S+1}\boldsymbol{\lambda} \geq \min((R\boldsymbol{\tau})^{[1]}, \mathbf{r}^{S+1}\boldsymbol{\tau}) \tag{40}$$

then $\boldsymbol{\tau}$ is directionally SSD inefficient with respect to $\mathbf{r}^{S+1}$.

Proof Let $j(\boldsymbol{\tau})$ be such index that $(-\tilde{R}\boldsymbol{\tau})^{[j(\boldsymbol{\tau})]} = -\mathbf{r}^{S+1}\boldsymbol{\tau}$ and similarly let $j(\boldsymbol{\lambda})$ be such that $(-\tilde{R}\boldsymbol{\lambda})^{[j(\boldsymbol{\lambda})]} = -\mathbf{r}^{S+1}\boldsymbol{\lambda}$. If $j(\boldsymbol{\lambda}) \geq 2$ then continuity of CVaR and assumptions (39) imply that there exists a sufficiently small t_0 such that for all $t \in [0, t_0]$

$$\mathrm{CVaR}_{\tilde{q}_s^{\lambda}(t)}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\tilde{q}_s^{\lambda}(t)}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}) < 0, \quad s = 0, 1, \ldots, S$$
$$\mathrm{CVaR}_{\tilde{q}_s^{\tau}(t)}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda}) - \mathrm{CVaR}_{\tilde{q}_s^{\tau}(t)}(-\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}) < 0, \quad s = 0, 1, \ldots, S$$

holds true. Hence, $\tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\lambda} \succ_{SSD} \tilde{\boldsymbol{\varrho}}(t)'\boldsymbol{\tau}$ and therefore $\boldsymbol{\lambda}$ is a feasible solution of (34) for all $t \in [0, t_0]$. The directional SSD inefficiency with respect to $\mathbf{r}^{S+1}$ of $\boldsymbol{\tau}$ follows.

If $j(\boldsymbol{\lambda}) = 1$ then (40) implies that $(\tilde{R}\boldsymbol{\lambda})^{[1]} \geq (\tilde{R}\boldsymbol{\tau})^{[1]}$ and the rest of the proof is similar to the previous case. □

Condition (40) is needed to guarantee that even in the contaminated case the smallest return of portfolio $\boldsymbol{\lambda}$ is larger than or equal to that of portfolio $\boldsymbol{\tau}$ what is a necessary condition of SSD relation. For data in Example 5(b), none of the conditions (39)–(40) is fulfilled.

4 Conclusions

The contamination technique was extended to construction of bounds for the optimal value function of perturbed stochastic programs whose set of feasible solutions depends on the probability distribution. In spite of the local character of these bounds their usefulness was illustrated for analysis of resistance with respect to additional scenarios in stochastic programs with risk constraints and in a new SSD portfolio efficiency test. Unlike the former portfolio efficiency tests, neither this test nor its robust version assume equiprobable scenarios.

Acknowledgements We would like to thank D. Dentcheva and an anonymous referee for the suggested simplification of Theorem 3 which leads to simplification of the SSD portfolio efficiency tests. The research was partly supported by the project "Methods of modern mathematics and their applications" —MSM 0021620839 and by the Czech Science Foundation (grants 201/09/0133, 402/10/1610).

References

Bonnans, J. S., & Shapiro, A. (2000). *Perturbation analysis of optimization problems*. *Springer series in OR*. New York.

Dentcheva, D., & Ruszczyński, A. (2010). Robust stochastic dominance and its application to risk-averse optimization. *Mathematical Programming*, *123*, 85–100.

Dupačová, J. (1990). Stability and sensitivity analysis for stochastic programming. *Annals of Operations Research*, *27*, 115–142.

Dupačová, J. (1996). Scenario-based stochastic programs: resistance with respect to sample. *Annals of Operations Research*, *64*, 21–38.

Dupačová, J. (1998). Reflections on robust optimization. In K. Marti & P. Kall (Eds.), *LNEMS: Vol.* 458. *Stochastic programming methods and technical applications* (pp. 111–127). Berlin: Springer.

Dupačová, J. (2006). Stress testing via contamination. In K. Marti et al. (Ed.), *LNEMS: Vol.* 581. *Coping with uncertainty, modeling and policy issues* (pp. 29–46). Berlin: Springer.

Dupačová, J. (2008). Risk objectives in two-stage stochastic programming models. *Kybernetika*, *44*, 227–242.

Dupačová, J., & Polívka, J. (2007). Stress testing for VaR and CVaR. *Quantitative Finance*, *7*, 411–421.

Fiacco, A. (1983). *Introduction to sensitivity and stability analysis in nonlinear programming*. New York: Academic Press.

Gol'shtein, E. G. (1970). *Vypukloje programmirovanije. Elementy teoriji*. Moscow: Nauka. [*Theory of convex programming, translations of mathematical monographs*, *Vol.* 36, American Mathematical Society, Providence, 1972].

Kilianová, S., & Pflug, G. (2009). Optimal pension fund management under multi-period risk minimization. *Annals of Operations Research*, *166*, 261–270.

Kopa, M. (2010). Measuring of second-order stochastic dominance portfolio efficiency. *Kybernetika*, *46*, 488–500.

Kopa, M., & Chovanec, P. (2008). A second-order stochastic dominance portfolio efficiency measure. *Kybernetika*, *44*, 243–258.

Krokhmal, P., Palmquist, J., & Uryasev, S. (2002). Portfolio optimization with conditional-value-at-risk objective and constraints. *The Journal of Risk*, *2*, 11–27.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, *50*, 1390–1406.

Kyparisis, J., & Fiacco, A. (1987). Generalized convexity and concavity of the optimal value function in nonlinear programming. *Mathematical Programming*, *39*, 285–304.

Ogryczak, W., & Ruszczyński, A. (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, *13*, 60–78.

Pagoncelli, B. K., Ahmed, S., & Shapiro, A. (2009). Sample average approximation method for chance constrained programming: theory and applications. *Journal of Optimization Theory and Applications*, *142*, 399–416.

Pflug, G., & Wozabal, D. (2007). Ambiguity in portfolio selection. *Quantitative Finance*, *7*, 435–442.

Prékopa, A. (1973). Contributions to the theory of stochastic programming. *Mathematical Programming*, *4*, 202–221.

Prékopa, A. (2003), Probabilistic programming, Chap. 5 in Ruszczyński and Shapiro (2003) (pp. 267–351).
Robinson, S. M. (1987). Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity. *Mathematical Programming Studies*, *30*, 45–66.
Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, *26*, 1443–1471.
Römisch, W. (2003). Stability of stochastic programming problems. Chap. 8 in Ruszczyński and Shapiro (2003) (pp. 483–554).
Ruszczyński, A., & Shapiro, A. (Eds.) (2003). *Stochastic programming*. *Handbooks in OR & MS*, Vol. 10. Amsterdam: Elsevier.
Ruszczyński, A., & Vanderbei, R. J. (2003). Frontiers of stochastically nondominated portfolios. *Econometrica*, *71*, 1287–1297.
Shapiro, A. (2003). Monte Carlo sampling methods, Chap. 6 in Ruszczyński and Shapiro (2003) (pp. 353–425).
Wang, W., & Ahmed, S. (2008). Sample average approximation for expected value constrained stochastic programs. *OR Letters*, *36*, 515–519.
Wets, R. J.-B. (1989). Stochastic programming. In G. L. Nemhauser et al. (Ed.), *Handbooks in OR & MS: Vol.* 1. *Optimization* (pp. 573–629). Amsterdam: Elsevier. Chap. VIII.