An enhanced model for portfolio choice with SSD criteria: a constructive approach

Csaba I. Fábián[*] Gautam Mitra[†] Diana Roman[‡] Victor Zverovich[§]

Abstract

We formulate a portfolio planning model which is based on Second-order Stochastic Dominance as the choice criterion. This model is an enhanced version of the multi-objective model proposed by Roman, Darby-Dowman, and Mitra (2006); the model compares the scaled values of the different objectives, representing tails at different confidence levels of the resulting distribution. The proposed model can be formulated as risk minimisation model where the objective function is a convex risk measure; we characterise this risk measure and the resulting optimisation problem. Moreover, our formulation offers a natural generalisation of the SSD-constrained model of Dentcheva and Ruszczyński (2006). A cutting plane-based solution method for the proposed model is outlined. We present a computational study showing: (a) the effectiveness of the solution methods and (b) the improved modelling capabilities: the resulting portfolios have superior return distributions.

[*]Institute of Informatics, Kecskemét College, 10 Izsáki út, 6000 Kecskemét, Hungary ; and Dept of OR, Loránd Eötvös Univ. E-mail: fabian.csaba@gamf.kefo.hu.

[†]CARISMA: The Centre for the Analysis of Risk and Optimisation Modelling Applications, School of Information Systems, Computing and Mathematics, Brunel University, UK.

[‡]CARISMA: The Centre for the Analysis of Risk and Optimisation Modelling Applications, School of Information Systems, Computing and Mathematics, Brunel University, UK.

[§]CARISMA: The Centre for the Analysis of Risk and Optimisation Modelling Applications, School of Information Systems, Computing and Mathematics, Brunel University, UK.

1 Introduction

1.1 Second-order Stochastic Dominance

Let R and R' denote random returns.

For a precise description, let $(\Omega, \mathcal{M}, P)$ be an appropriate probability space. (The set Ω is equipped with the probability measure P, the field of measurable sets being $\mathcal{M}$.) We consider integrable random variables, formally, let $R, R' \in \mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{M}, P)$.

Second-order Stochastic Dominance (SSD) is defined by the following equivalent conditions:

(a) $\mathrm{E}\left(U(R)\right) \geq \mathrm{E}\left(U(R')\right)$ holds for any nondecreasing and concave utility function U for which these expected values exist and are finite.

(b) $\mathrm{E}\left([t-R]_+\right) \leq \mathrm{E}\left([t-R']_+\right)$ holds for each $t \in \mathrm{I\!R}$.

(c) $\mathrm{Tail}_\alpha(R) \geq \mathrm{Tail}_\alpha(R')$ holds for each $0 < \alpha \leq 1$, where $\mathrm{Tail}_\alpha(R)$ denotes the unconditional expectation of the least $\alpha * 100\%$ of the outcomes of R.

For the equivalence of *(a)* and *(b)* see for example Whitmore and Findlay (1978), Theorem 2.2, page 65. The equivalence of *(b)* and *(c)* is shown in Ogryczak and Ruszczyński (2002): they consider $\mathrm{Tail}_\alpha(R)$ as a function of α, and $\mathrm{E}\left([t-R]_+\right)$ as a function of t ; and observe that these functions are convex conjugates.

If *(a b c)* above hold, we say that R dominates R' with respect to SSD, and use the notation $R \succeq_{_{SSD}} R'$. The corresponding strict dominance relation is defined in the usual way: $R \succ_{_{SSD}} R'$ means that $R \succeq_{_{SSD}} R'$ and $R' \not\succeq_{_{SSD}} R$.

In this paper we deal with portfolio returns. Let n denote the number of the assets into which we may invest at the beginning of a fixed time period. A portfolio $\boldsymbol{x} = (x_1, \dots x_n) \in \mathrm{I\!R}^n$ represents the proportions of the portfolio value invested in the different assets. Let the n-dimensional random vector $\boldsymbol{R}$ denote the returns of the different assets at the end of the investment period.

We assume that the components of $\boldsymbol{R}$ belong to $\mathcal{L}^1$.

It is usual to consider the distribution of $\boldsymbol{R}$ as discrete, described by the realisations under various scenarios. The random return of portfolio $\boldsymbol{x}$ will be $R_{\boldsymbol{x}} := x_1 R_1 + \dots x_n R_n$.

Let $X \subset \mathrm{I\!R}^n$ denote the set of the feasible portfolios. We assume that X is a bounded convex polyhedron. A portfolio $\boldsymbol{x}^\star$ is said to be SSD-efficient if there is no feasible portfolio $\boldsymbol{x} \in X$ such that $R_{\boldsymbol{x}} \succ_{_{SSD}} R_{\boldsymbol{x}^\star}$.

The importance of SSD as a choice criterion in portfolio selection, as well as the difficulty in applying it in practice have been widely recognised (Hadar and Russell 1969, Whitmore and Findlay 1978, Kroll and Levy 1980, Ogryczak and Ruszczyński 2001, 2002). SSD is a meaningful choice criterion, due to its relation to risk averse behaviour (as stated at (a)), which is the general assumption about investment behaviour. The computational difficulty of the SSD-based models arises from the fact that, finding the set of SSD efficient portfolios is a model with a continuum of objectives (as stated at (c)). Only recently, SSD-based portfolio models based have been proposed (Dentcheva and Ruszczyński 2003, 2006, Roman et al. 2006, Fábián et al. 2009).

1.2 Portfolio models based on the SSD criteria

The SSD-based models reviewed here assume that a reference random return $\widehat{R}$, with a known (discrete) distribution, is available; $\widehat{R}$ could be for example the return of a stock index or of a benchmark portfolio.

Dentcheva and Ruszczyński (2006) propose an SSD constrained portfolio optimisation model:

$$\begin{gathered} \max\ f(\boldsymbol{x}) \\ \text{such that} \quad \boldsymbol{x} \in X, \\ R_{\boldsymbol{x}} \succeq_{_{SSD}} \widehat{R}, \end{gathered} \tag{1}$$

where f is a concave function. In particular, they consider $f(\boldsymbol{x}) = \mathrm{E}(R_{\boldsymbol{x}})$. They formulate the problem using criterion *(b)* (section 1.1) and prove that, in case of finite discrete distributions, the SSD relation can be characterised by a finite system of inequalities from those in *(b)*. The authors developed a solution method based on a dual formulation and the Regularized Decomposition method of Ruszczyński (1986). The authors implemented this method, using a dataset with 719 real-world assets and 616 possible realizations of their joint return rates; favorable performance is reported.

Roman, Darby-Dowman, and Mitra (2006) use criterion *(c)* (section 1.1). They assume finite discrete distributions with equally probable outcomes, and prove that, in this case, the SSD relation can be characterised by a finite system of inequalities. Namely, $R_{\boldsymbol{x}} \succeq_{_{SSD}} \widehat{R}$ is equivalent to $\mathrm{Tail}_{\frac{i}{S}}(R_{\boldsymbol{x}}) \geq \mathrm{Tail}_{\frac{i}{S}}(\widehat{R}) \quad (i = 1, \ldots, S)$, where S is the number of (equally probable) scenarios. The authors propose a multi-objective model whose Pareto optimal solutions are SSD-efficient portfolios.
A specific solution is chosen whose return distribution comes close to, or emulates, the reference return $\widehat{R}$ in a uniform sense. Uniformity is meant in terms of differences among tails; the "worst" tail difference $\vartheta = \min\limits_{i=1\ldots S} \left(\mathrm{Tail}_{\frac{i}{S}}(R_{\boldsymbol{x}}) - \mathrm{Tail}_{\frac{i}{S}}(\widehat{R})\right)$ is maximised:

$$\begin{gathered} \max\ \vartheta \\ \text{such that} \quad \vartheta \in \mathrm{I\!R}, \quad \boldsymbol{x} \in X, \\ \mathrm{Tail}_{\frac{i}{S}}(R_{\boldsymbol{x}}) \ \geq\ \mathrm{Tail}_{\frac{i}{S}}(\widehat{R}) + \vartheta \qquad (i = 1, \ldots, S). \end{gathered} \tag{2}$$

The return distribution of the chosen portfolio comes "close or better" than the reference distribution: if the reference distribution is not SSD efficient (which is often the case), the model improves on it until SSD-efficiency is reached.

The authors implemented the model outlined above, and made extensive testing on problems with 76 real-world assets using 132 possible realisations of their joint return rates. Powerful modelling capabilities were demonstrated by in-sample and out-of-sample analysis of the return distributions of the optimal portfolios.

1.3 An enhanced model

In this paper we propose a scaled version of the model (2) of Roman, Darby-Dowman, and Mitra (2006). The new model is formulated in the compact form

$$\begin{gathered} \max\ \vartheta \\ \text{such that} \quad \vartheta \in \mathrm{I\!R}, \quad \boldsymbol{x} \in X, \\ R_{\boldsymbol{x}} \ \succeq_{_{SSD}} \widehat{R} + \vartheta. \end{gathered} \tag{3}$$

Obviously we have $\mathrm{Tail}_{\frac{i}{S}}(\widehat{R} + \vartheta) = \mathrm{Tail}_{\frac{i}{S}}(\widehat{R}) + \frac{i}{S}\vartheta \quad (i = 1, \ldots, S)$ with $\vartheta \in \mathrm{I\!R}$. Hence, using criterion *(c)* (section 1.1), the model (3) can be equivalently formulated as

$$\begin{gathered} \max\ \vartheta \\ \text{such that} \quad \vartheta \in \mathrm{I\!R}, \quad \boldsymbol{x} \in X, \\ \mathrm{Tail}_{\frac{i}{S}}(R_{\boldsymbol{x}}) \ \geq\ \mathrm{Tail}_{\frac{i}{S}}(\widehat{R}) + \tfrac{i}{S}\vartheta \qquad (i = 1, \ldots, S). \end{gathered} \tag{4}$$

The difference between (2) and (4) is that the tails are scaled in the latter model.

This approach has several advantages; one of them lies in the connection with the risk minimisation theory.

The quantity ϑ in (4) measures the preferability of the portfolio return $R_{\boldsymbol{x}}$ relative to the reference return $\widehat{R}$.

The relation $R_{\boldsymbol{x}} \succeq_{SSD} \widehat{R}+\vartheta$ means that we prefer the return distribution of portfolio $\boldsymbol{x}$ to the combination of the reference return and a riskless return ϑ (usually cash).

We can introduce an opposite measure as

$$\widehat{\rho}\big(R\big) := \min\left\{ \varrho \in \mathbb{R} \,\middle|\, R + \varrho \succeq_{SSD} \widehat{R} \right\} \quad \text{for any return } R. \tag{5}$$

In words, $\widehat{\rho}\big(R\big)$ measures the amount of riskless return whose addition makes R preferable to $\widehat{R}$.

Using this, the problem (3) can be formulated as

$$\min_{\boldsymbol{x}\in X} \; \widehat{\rho}\big(R_{\boldsymbol{x}}\big). \tag{6}$$

We show that $\widehat{\rho}$ is a *convex risk measure.* We also develop a cutting-plane representation of $\widehat{\rho}$ by adapting the approach presented in Fábián, Mitra, and Roman (2009). This gives a solution method for problem (6).

Using the risk measure $\widehat{\rho}$, an extension of the SSD-constrained model (1) of Dentcheva and Ruszczyński can be formulated as

$$\begin{aligned} &\max \; f(\boldsymbol{x}) \\ &\text{such that} \quad \boldsymbol{x} \in X, \\ &\widehat{\rho}\big(R_{\boldsymbol{x}}\big) \le \gamma, \end{aligned} \tag{7}$$

where $\gamma \in \mathbb{R}$ is a parameter. In an application, the setting of the parameter γ is usually the responsibility of the decision makers. We can help them by constructing the efficient frontier. Suppose that values and subgradients can be computed to the function f. The efficient frontier of problem (7) can be approximated by solving Lagrangian problems

$$\max_{\boldsymbol{x}\in X} \; f(\boldsymbol{x}) - \lambda\, \widehat{\rho}\big(R_{\boldsymbol{x}}\big) \tag{8}$$

with different values of $\lambda \ge 0$. Once the right-hand-side parameter γ is tuned by the decision maker, the problem can be solved by a constrained convex method.

The paper is organised as follows:

In Section 2, we overview coherent and convex risk measures, and show that $\widehat{\rho}$ defined as (5) is a convex risk measure. We also present the dual representation of $\widehat{\rho}$.

In Section 3, we compare different formulations of the enhanced portfolio choice problem (3).

In Section 4, we describe a cutting-plane approach for the enhanced portfolio choice problem (3), and study its convergence. We also sketch a solution method for the problem (7).

In Section 5, we present a computational study. We compare the return distributions of the respective optimal portfolios belonging to the multi-objective problem of Roman, Darby-Dowman, and Mitra (2) on the one hand, and to the scaled version (4) on the other hand.

The results are summarised in Section 6.

2 Convex risk measures

2.1 Overview of risk measures

Let $\mathcal{R}$ denote the space of legitimate returns. In most cases, we can consider $\mathcal{R} = \mathcal{L}^1$; but for the dual representation described in section 2.2, finite variance is also required, and hence $\mathcal{R} = \mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{M}, P)$ is considered.

A *risk measure* is mapping $\rho : \mathcal{R} \to [-\infty, +\infty]$. The *acceptance set* of a risk measure ρ is defined as

$$\mathcal{A}_\rho := \{ R \in \mathcal{R} \,|\, \rho(R) \le 0 \}\,. \tag{9}$$

Conversely, an *acceptance set* $\mathcal{A}$ defines a risk measure $\rho_{\mathcal{A}}$:

$$\rho_{\mathcal{A}}(R) := \inf \{ \varrho \in \mathrm{I\!R} \,|\, R + \varrho \in \mathcal{A} \} \quad (R \in \mathcal{R}). \tag{10}$$

The concept of *coherent risk measures* was developed by Artzner et al. (1999) and Delbaen (2002).

These are mappings that satisfy the following four criteria:

Subadditivity: $\rho(R + R') \leq \rho(R) + \rho(R')$ holds for $R, R' \in \mathcal{R}$.

Positive homogeneity: $\rho(\lambda R) = \lambda \rho(R)$ holds for $R \in \mathcal{R}$ and $\lambda \geq 0$.

Monotonicity: $\rho(R) \leq \rho(R')$ holds for $R, R' \in \mathcal{R},\ R \geq R'$.

Translation equivariance: $\rho(R + \varrho) = \rho(R) - \varrho$ holds for $R \in \mathcal{R},\ \varrho \in \mathrm{I\!R}$.

Artzner et al. show that the acceptance set of a coherent risk measure is a convex cone, and that a convex cone defines a coherent risk measure.

Another axiomatic description of risk was proposed by Kijima and Ohnishi (1993). They consider mappings σ that satisfy subadditivity, positive homogeneity, and the following two criteria:

Nonnegativity: $\sigma(R) \geq 0$ holds for $R \in \mathcal{R}$.

Translation invariance: $\sigma(R + \varrho) = \sigma(R)$ holds for $R \in \mathcal{R},\ \varrho \geq 0$.

Mappings satisfying these (or similar) criteria are called *deviation measures*. Rockafellar, Uryasev, and Zabarankin (2002, 2006) present a systematic study of general deviation measures, and explore the connection between deviation measures and coherent risk measures.

A well-known example for a coherent risk measure is *Conditional Value-at-Risk* (CVaR), charactarised by Rockafellar and Uryasev (2000-2002). In words, $\mathrm{CVaR}_\alpha(R)$ is the conditional expectation of the upper α-tail of $-R$. (In our setting, R represents gain, hence $-R$ represents loss.) We have the relation

$$\mathrm{CVaR}_\alpha(R) = -\frac{1}{\alpha}\,\mathrm{Tail}_\alpha(R) \quad (0 < \alpha \leq 1). \tag{11}$$

(Rockafellar and Uryasev define CVaR_α using $(1-\alpha)$-tails, making the definition compatible with that of VaR. Since we do not use VaR in this paper, we define CVaR_α using α-tails, also a customary way.)

The concept of *convex risk measures* is a natural generalisation of coherency, which allows more general convex sets as acceptance sets. The concept was introduced by Heath (2000), Carr, Geman, and Madan (2001), and Föllmer and Schied (2002). A mapping ρ is said to be a convex risk measure if it satisfies monotonicity, translation equivariance, and

Convexity: $\rho\big(\lambda R + (1-\lambda)R'\big) \leq \lambda\rho(R) + (1-\lambda)\rho(R')$ holds for $R, R' \in \mathcal{R}$ and $0 \leq \lambda \leq 1$.

Rockafellar, Uryasev, and Zabarankin (2002, 2006) develop another generalisation of the coherency concept, which also includes the *scalability criterion*: $\rho(R_\varrho) = -\varrho$ for each $\varrho \in \mathrm{I\!R}$ (here $R_\varrho \in \mathcal{R}$ denotes the return of constant ϱ). An overview can be found in Rockafellar (2007).

The above cited works also develop dual representations of risk measures. A coherent risk measure can be represented as

$$\rho(R) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q(-R), \tag{12}$$

where $\mathcal{Q}$ is a set of probability measures on Ω. A risk measure that is convex or coherent in the extended sense of Rockafellar et al., can be represented as

$$\rho(R) = \sup_{Q \in \mathcal{Q}} \Big\{ \mathrm{E}_Q(-R) \;-\; \nu(Q) \Big\}, \tag{13}$$

where $\mathcal{Q}$ is a set of probability measures, and ν is a mapping from the set of the probability measures to $(-\infty, +\infty]$. (Properties of $\mathcal{Q}$ and ν depend on the type of risk measure, and also on space $\mathcal{R}$, and the topology used.) On the basis of these dual representations, an optimisation problem that involves a risk measure can be interpreted as a robust optimisation problem.

Ruszczyński and Shapiro (2006) develop optimality and duality theory for problems with convex risk functions.

2.2 Convexity and dual representation of the risk measure $\widehat{\rho}$

The risk measure $\widehat{\rho}$ defined in (5) derives from the acceptance set $\widehat{\mathcal{A}} := \left\{ R \in \mathcal{R} \,\middle|\, R \succeq_{SSD} \widehat{R} \right\}$ in the manner of (10). We prove convexity of $\widehat{\rho}$ using the following proposition from Föllmer and Schied (2002):

Proposition 1 *Let $\mathcal{A}$ be a convex subset of $\mathcal{R}$, such that*

$$\mathcal{A} \neq \emptyset \quad \textit{and} \quad \rho_{\mathcal{A}}(R_0) > -\infty, \tag{14}$$

where $R_0 \in \mathcal{R}$ denotes the return of constant 0. Suppose that $\mathcal{A}$ has the following property:

$$\textit{if} \quad R \in \mathcal{A} \quad \textit{and} \quad R' \in \mathcal{R},\ R' \geq R, \quad \textit{then} \quad R' \in \mathcal{A}. \tag{15}$$

Then $\rho_{\mathcal{A}}$, defined as (10), is a convex risk measure. (If, moreover, $\mathcal{A}$ satisfies a certain closedness criterion, then $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$ holds.)

Proof of convexity and closedness of $\widehat{\mathcal{A}}$ can be found in Dentcheva and Ruszczyński (2003). (They use criterion *(b)* cited in section 1.1, and the fact that the mapping $R \mapsto \mathrm{E}\left([t-R]_+\right)$ is convex and continuous for any t.)

$\widehat{\mathcal{A}}$ evidently satisfies (14), and (15) is a consequence of the monotonicity of our utility functions. Hence we see that the risk measure $\widehat{\rho}$ is convex.

In order to construct a dual representation of $\widehat{\rho}$, we follow the approach of Rockafellar, Uryasev, and Zabarankin. The space $\mathcal{R}$ of returns is $\mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{M}, P)$, i.e., the measurable functions for which the mean and variance exist. As for probability measures, Rockafellar at al. focus on those that can be described by density functions with respect to P. Moreover the density functions need to fall into $\mathcal{L}^2$. Let Q be a legitimate probability measure with density function d_Q. Under these conditions $\mathrm{E}_Q(R) = E(R\, d_Q)$ holds.

Rockafellar at al. show that the dual representation of CVaR_α in the form of (12) is

$$\mathrm{CVaR}_\alpha(R) = \sup_{d_Q \leq \alpha^{-1}} \mathrm{E}_Q(-R) \qquad (0 < \alpha \leq 1). \tag{16}$$

Based on this result, we construct a dual representation of $\widehat{\rho}$. According to the definition *(c)* of SSD in Section 1.1, we have

$$\widehat{\mathcal{A}} = \bigcap_{0<\alpha\leq 1} \widehat{\mathcal{B}}_\alpha \tag{17}$$

with

$$\widehat{\mathcal{B}}_\alpha := \left\{ R \,\middle|\, \mathrm{Tail}_\alpha(R) \geq \mathrm{Tail}_\alpha(\widehat{R}) \right\} = \left\{ R \,\middle|\, \mathrm{CVaR}_\alpha(R) \leq \mathrm{CVaR}_\alpha(\widehat{R}) \right\} \qquad (0 < \alpha \leq 1).$$

(The above equality is an obvious consequence of (11).) Substituting (16) we get

$$\widehat{\mathcal{B}}_\alpha = \left\{ R \,\middle|\, \mathrm{E}_Q(-R) \leq \mathrm{CVaR}_\alpha(\widehat{R}) \text{ holds for each } Q \text{ having } d_Q \leq \alpha^{-1} \right\} \qquad (0 < \alpha \leq 1).$$

Substituting this into (17), we get

$$\widehat{\mathcal{A}} = \left\{ R \,\middle|\, \mathrm{E}_Q(-R) \leq \mathrm{CVaR}_\alpha(\widehat{R}) \text{ holds for each } (Q, \alpha) \text{ having } d_Q \leq \alpha^{-1} \right\}. \tag{18}$$

Let us define

$$\mathcal{Q} := \{ Q \mid \sup d_Q < +\infty \} \quad \text{and} \quad s(Q) := \left(\sup d_Q\right)^{-1} \quad (Q \in \mathcal{Q}).$$

(We have $s(Q) \leq 1$ for each legitimate Q.) Equality (18) can be continued as

$$\begin{aligned} \widehat{\mathcal{A}} &= \left\{ R \,\middle|\, \mathrm{E}_Q(-R) \leq \mathrm{CVaR}_\alpha(\widehat{R}) \text{ holds for each } Q \in \mathcal{Q}, \alpha \leq s(Q) \right\} \\ &= \left\{ R \,\middle|\, \mathrm{E}_Q(-R) \leq \mathrm{CVaR}_{s(Q)}(\widehat{R}) \text{ holds for each } Q \in \mathcal{Q} \right\}, \end{aligned}$$

since $\mathrm{CVaR}_\alpha(\widehat{R})$ is decreasing function of α. We have $\widehat{\rho} = \rho_{\widehat{\mathcal{A}}}$ hence by (10)

$$\begin{aligned} \widehat{\rho}(R) &= \inf \left\{ \varrho \in \mathrm{I\!R} \,\middle|\, \mathrm{E}_Q(-R-\varrho) \leq \mathrm{CVaR}_{s(Q)}(\widehat{R}) \text{ holds for each } Q \in \mathcal{Q} \right\} \\ &= \sup_{Q \in \mathcal{Q}} \left\{ \mathrm{E}_Q(-R) \,-\, \mathrm{CVaR}_{s(Q)}(\widehat{R}) \right\} \end{aligned} \tag{19}$$

holds for each $R \in \mathcal{R}$, and this has the form of the dual representation (13) with $\nu(Q) = \mathrm{CVaR}_{s(Q)}(\widehat{R})$.

In the remaining part of this paper we focus on discrete finite distributions with equiprobable outcomes.

Let us specialise the above dual representation to this case. Let $S \in \mathrm{I\!N}$ denote the cardinality of the sample space. The probability measure P assigns the weight $1/S$ to each of the S elements of the sample space. A measure $Q \in \mathcal{Q}$ is defined by the weights (probabilities) $q_1, \ldots, q_S$. The density d_Q of Q with respect to P takes the values $\frac{q_1}{1/S}, \ldots, \frac{q_S}{1/S}$. We have $s(Q) = 1/\left(S \max_{1 \leq j \leq S} q_j\right)$, and hence (19) can be written as

$$\widehat{\rho}(R) = \sup_{\substack{q_1, \ldots, q_s \geq 0 \\ q_1 + \cdots + q_s = 1}} \left\{ \sum_{j=1}^{S} -q_j \, r^{(j)} \,-\, \mathrm{CVaR}_{1/(S \,\max q_j)}(\widehat{R}) \right\}, \tag{20}$$

where $r^{(1)}, \ldots, r^{(S)}$ denote the realisations of the random return R.

Since $\widehat{\rho}$ is invariant to permutations of the realisations, the acceptance set $\widehat{\mathcal{A}}$ has a highly symmetric structure.

3 Problem formulation

We compare different formulations of the enhanced model (3). The dominance relation can be formulated by either tails or integrated chance constraints, according to criterions *(b)* or *(c)* in Section 1.1.

We assume that the joint distribution of $\boldsymbol{R}$ and $\widehat{R}$ is discrete finite, having S equally probable outcomes. Let $\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(S)}$ denote the realisations of the random $\boldsymbol{R}$ vector of asset returns. Similarly, let $\widehat{r}^{(1)}, \ldots, \widehat{r}^{(S)}$ denote realisations of the reference return $\widehat{R}$. For the reference tails, we will use the brief notation $\widehat{\tau}_i := \mathrm{Tail}_{\frac{i}{S}}\left(\widehat{R}\right) \quad (i = 1, \ldots S)$.

3.1 Formulation using tails

In their multi-objective model (2), Roman, Darby-Dowman, and Mitra computed tails in the following form, by adapting the CVaR-optimisation formula of Rockafellar and Uryasev (2000, 2002):

$$\mathrm{Tail}_{\frac{i}{S}}\left(R\boldsymbol{x}\right) = \max_{t_i \in \mathrm{I\!R}} \left\{ \frac{i}{S} t_i - \frac{1}{S} \sum_{j=1}^{S} \left[t_i - \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \right]_+ \right\}.$$

Roman et al. then formulated (2) in linear programming form, introducing new variables for the positive parts $\left[\, t_i - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \,\right]_+$. The resulting problems were found to be computationally demanding, though. Instead of introducing new variables, Fábián, Mitra, and Roman (2009) used the following cutting-plane representation, adapting the approach of Künzi-Bay and Mayer (2006):

$$\begin{aligned} \mathrm{Tail}_{\frac{i}{S}}\left(R\boldsymbol{x}\right) = \quad & \min \ \frac{1}{S} \sum_{j \in \mathcal{J}_i} \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \\ & \text{such that} \quad \mathcal{J}_i \subset \{1, \ldots, S\}, \ \ |\mathcal{J}_i| = i. \end{aligned} \tag{21}$$

In its original form, the Künzi-Bay – Mayer representation includes a cut for each subset of $\{1, \ldots, S\}$. However it can be seen that, in the equiprobable case, only cuts of cardinality i are required for the computation of $\frac{i}{S}$-tails.

The Künzi-Bay – Mayer representation is the CVaR-analogue of the Klein Haneveld – Van der Vlerk (2006) representation for integrated chance constraints. The two approaches employ the same idea, originally developed by Klein Haneveld and Van der Vlerk. Künzi-Bay and Mayer obtained their representation independently.

Formula (21) enables a cutting-plane approach to the multi-objective model (2). Fábián et al. implemented this cutting-plane approach, and found it highly effective.

Applying (21) to the present, scaled-tail model (4) results the following cutting-plane representation:

$$\begin{array}{lll} \max & \vartheta & \\ \text{such that} & \vartheta \in \mathbb{R}, \quad \boldsymbol{x} \in X, & \\ & \frac{i}{S}\vartheta + \widehat{\tau}_i \ \le \ \frac{1}{S} \sum_{j \in \mathcal{J}_i} \boldsymbol{r}^{(j)\,T}\boldsymbol{x} & \text{for each} \quad \mathcal{J}_i \subset \{1, \ldots, S\}, \ \ |\mathcal{J}_i| = i, \\ & & \text{where} \quad i = 1, \ldots, S. \end{array} \tag{22}$$

3.2 Formulation using integrated chance constraints

In their SSD-constrained model (1), Dentcheva and Ruszczyński characterise stochastic dominance with criterion *(b)* in Section 1.1. They prove that if $\widehat{R}$ has a discrete finite distribution with realisations $\widehat{r}^{(1)}, \ \ldots, \ \widehat{r}^{(S)}$, then $R\boldsymbol{x} \succeq_{SSD} \widehat{R}$ is equivalent to a finite system of inequalities

$$\mathrm{E}\left(\left[\,\widehat{r}^{(i)} - R\boldsymbol{x}\,\right]_+\right) \ \le \ \mathrm{E}\left(\left[\,\widehat{r}^{(i)} - \widehat{R}\,\right]_+\right) \qquad (i = 1, \ldots, S). \tag{23}$$

The right-hand sides of the above constraints do not depend on $\boldsymbol{x}$. In case of scenario i, the left-hand side represents expected shortfall of portfolio return $R\boldsymbol{x}$ relative to the constant $\widehat{r}^{(i)}$. Such a constraint is called integrated chance constraint after Klein Haneveld (1986).

Dentcheva and Ruszczyński (2006) transform (23) into a set of linear constraints by introducing new variables to represent positive parts. The resulting linear programming problem has a specific structure. For such specific problems, Dentcheva and Ruszczyński develop a duality theory in which the dual objects are utility functions. Based on this duality theory, they construct a dual problem that consists of the minimization of a weighted sum of polyhedral convex functions. Dentcheva and Ruszczyński adapt the Regularized Decomposition method of Ruszczyński (1986) to these special dual problems.

Klein Haneveld and Van der Vlerk (2006) present a cutting-plane representation for integrated chance constraints in case of discrete finite distributions. Based on this representation, they develop a cutting-plane method for integrated chance constrained optimisation problems. In this method, cuts are generated in the space of the variables. Klein Haneveld and Van der Vlerk present a computational study demonstrating the effectiveness of their cutting-plane approach. Fábián, Mitra, and Roman (2009) propose handling SSD-constraints by applying the Klein Haneveld – Van der Vlerk method to the finite system (23). Though Fábián et al. view the method in a somewhat different perspective: instead of generating cuts in the variable space, they construct a constraint function, and generate cuts to the graph of this function. Such view enables regularisation.

Independently, Rudolf and Ruszczyński (2008) also develop cutting-plane methods for the solution of SSD-constrained problems. They propose an extension of the Klein Haneveld – Van der Vlerk representation to integrable random variables. They formulate a primal cutting-plane method which, in case of finite distributions, is equivalent to the application of the Klein Haneveld – Van der Vlerk method. Rudolf and Ruszczyński also develop a duality theory for SSD-constrained problems, and propose a dual column-generation method for the solution of such problems. Their computational study demonstrates that the primal cutting-plane method is computationally efficient.

The constraint $R\boldsymbol{x} \succeq_{SSD} \widehat{R} + \vartheta$ of the present enhanced model (3) can be formulated in the manner of (23), as the set of integrated chance constraints

$$\sum_{j=1}^{S} \frac{1}{S} \left[\widehat{r}^{(i)} + \vartheta - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \right]_+ \leq \sum_{j=1}^{S} \frac{1}{S} \left[\widehat{r}^{(i)} - \widehat{r}^{(j)} \right]_+ \qquad (i = 1, \ldots, S). \tag{24}$$

(Under our equiprobability assumption, the probability of each scenario is $1/S$.) The Klein Haneveld – Van der Vlerk cutting-plane representation of the ith constraint from the system (24) is

$$\sum_{j\in\mathcal{J}_i} \frac{1}{S} \left\{ \widehat{r}^{(i)} + \vartheta - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \right\} \leq \sum_{j=1}^{S} \frac{1}{S} \left[\widehat{r}^{(i)} - \widehat{r}^{(j)} \right]_+ \quad \text{for each } \mathcal{J}_i \subset \{1, \ldots, S\}.$$

Using the above representation, the enhanced model (3) can be formulated as

$$\begin{aligned}
&\max && \vartheta \\
&\text{such that} && \vartheta \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\
& && \sum_{j\in\mathcal{J}_i} \tfrac{1}{S} \left\{ \widehat{r}^{(i)} + \vartheta - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \right\} \leq \sum_{j=1}^{S} \tfrac{1}{S} \left[\widehat{r}^{(i)} - \widehat{r}^{(j)} \right]_+ \qquad \text{for each } \mathcal{J}_i \subset \{1, \ldots, S\}, \\
& && \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{where } i = 1, \ldots, S.
\end{aligned} \tag{25}$$

3.3 Comparison of formulations and connection with risk measures

The problems (25) and (22) are equivalent since they are different formulations of the enhanced model (3). To be specific, we can find a mapping between the two constraint sets: For the sake of simplicity, assume that the realisations of the reference return are ordered: $\widehat{r}^{(1)} \leq \ldots \leq \widehat{r}^{(S)}$. It is easily seen that the constraint belonging to a set $\mathcal{J}_i \subset \{1, \ldots, S\}$ in (25) is

– equivalent to the constraint belonging to $\mathcal{J}_i$ in (22), if $|\mathcal{J}_i| = i$;

– redundant if $|\mathcal{J}_i| \neq i$.

Hence the formulation (22) is more economic under our equiprobability assumption.

Remark 2 *If we drop the equiprobability assumption, but keep the discrete finite assumption, then the formulation (25) will be more convenient.*

New formulations of stochastic dominance constraints are proposed by Luedtke (2008). His computational study demonstrates that his schemes are effective, but do not surpass the efficiency of the Klein Haneveld – Van der Vlerk formulation.

In the remaining part of this paper we will use the formulation (22). Changing the scope of optimisation, (22) can be stated as minimisation of a polyhedral convex function:

$$\min_{\boldsymbol{x}\in X}\ \varphi(\boldsymbol{x}) \qquad \text{where} \quad \varphi(\boldsymbol{x}) := \begin{array}[t]{l} \max \left\{ -\frac{1}{i} \sum\limits_{j\in\mathcal{J}_i} \boldsymbol{r}^{(j)\,T}\boldsymbol{x} \ + \frac{S}{i}\widehat{\tau}_i \right\} \\ \text{such that} \quad \mathcal{J}_i \subset \{1,\dots,S\}, \ \ |\mathcal{J}_i| = i, \\ \text{where} \quad i = 1,\dots,S. \end{array} \tag{26}$$

We have $\frac{S}{i}\widehat{\tau}_i = -\mathrm{CVaR}_{\frac{i}{S}}(\widehat{R})$ according to (11). Using this, the definition of φ can be written as

$$\varphi(\boldsymbol{x}) = \max_{\substack{\mathcal{J}\subset\{1,\dots,S\} \\ \mathcal{J}\neq\emptyset}} \left\{ \sum_{j\in\mathcal{J}} -\frac{1}{|\mathcal{J}|}\boldsymbol{r}^{(j)\,T}\boldsymbol{x} \quad - \mathrm{CVaR}_{\frac{|\mathcal{J}|}{S}}(\widehat{R}) \right\}. \tag{27}$$

In the present discrete, finite, equiprobable case, the above formulation is equivalent to the enhanced model (3). The latter model is in turn equivalent to the risk minimisation problem (6), with the risk measure $\widehat{\rho}$ having dual representation (20). This equivalence holds regardless of the selection of the feasible domain X. Hence we have $\varphi(\boldsymbol{x}) = \widehat{\rho}\big(R_{\boldsymbol{x}}\big)$ ($\boldsymbol{x}\in\mathbb{R}^n$).

It is easily seen that a set $\mathcal{J}$ in (27) maps to the measure Q in (20) having weights

$$q_j = \begin{cases} 1/|\mathcal{J}| & \text{if } j\in\mathcal{J}, \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

Other measures in (20) are redundant from a cutting-plane point of view. It means that if we consider (20) as a robust optimisation model, then all critical probability measures will have the form (28).

4 Solution methods

We solve the portfolio optimisation problem in the form (26).

4.1 Pure cutting plane method

Applied to a convex programming problem $\min\limits_{\boldsymbol{x}\in\mathcal{X}} \ \phi(\boldsymbol{x})$, the cutting plane method generates a sequence of iterates from $\mathcal{X}$. At each iterate, supporting linear functions (cuts) are constructed to the objective function. A cutting-plane model of the objective function is then maintained as the upper cover of known cuts. The next iterate is obtained by minimising the current model function over $\mathcal{X}$.

Cutting plane methods are considered fairly efficient for quite general problem classes. However, according to our knowledge, efficiency estimates only exist for the continuously differentiable, strictly convex case. An overview can be found in Dempster and Merkovsky (1995), who present a geometrically convergent version.

Supporting linear functions to our present objective function φ in (26) are easily constructed:

- Given $\boldsymbol{x}^\star\in X$, let $r_{\boldsymbol{x}^\star}^{(j_1^\star)} \leq \ldots \leq r_{\boldsymbol{x}^\star}^{(j_S^\star)}$ denote the ordered outcomes of $R_{\boldsymbol{x}^\star} = \boldsymbol{R}^T\boldsymbol{x}^\star$.
- Using this ordering of the scenarios, let us construct the sets $\mathcal{J}_i^\star := \{j_1^\star,\dots,j_i^\star\}$ $\quad(i=1,\dots,S)$.
- Let us select $i^\star \in \arg\max\limits_{1\leq i\leq S} \left\{ -\frac{1}{i}\sum\limits_{j\in\mathcal{J}_i^\star} \boldsymbol{r}^{(j)\,T}\boldsymbol{x}^\star \ + \frac{S}{i}\widehat{\tau}_i \right\}$.

– A supporting linear function at $\boldsymbol{x}^\star$ is then $l(\boldsymbol{x}) := -\frac{1}{i^\star} \sum_{j \in \mathcal{J}^\star_{i^\star}} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} + \frac{S}{i^\star} \widehat{\tau}_{i^\star}$.

Klein Haneveld and Van der Vlerk (2006), Künzi-Bay and Mayer (2006), Fábián, Mitra, and Roman (2009) report application of the cutting plane method to stochastic programming problems similar to the present one. The method proved very effective for these problems. When a problem was solved with increasing numbers of scenarios, the iteration count increased very slowly.

4.2 Level method

The level method is a regularised cutting plane method, proposed by Lemaréchal, Nemirovskii, and Nesterov (1995). A cutting-plane model function is maintained as in the pure cutting plane method. The next iterate is obtained by minimising the current model function over the feasible domain, and then projecting the minimiser to a certain level set of the current model function. (Projecting requires the solution of a convex quadratic programming problem.)

Lemaréchal, Nemirovskii, and Nesterov prove the following efficiency estimate: Suppose that the objective function is Lipschitz-continuous with the parameter L, and let D denote the diameter of the feasible domain. To obtain an ϵ-optimal solution, it suffices to perform $c\left(\frac{LD}{\epsilon}\right)^2$ iterations, where c is a constant. Lemaréchal, Nemirovskii, and Nesterov also implemented the method, and report much better practical behaviour then the theoretical efficiency estimate.

Fábián and Szőke (2007) used the level method for the solution of stochastic programming problems. They also found practical behaviour to be much better then the cited theoretical efficiency estimate. When a problem was solved with increasing numbers of scenarios, the iteration count stabilised early, i.e., iteration count proved independent of the number of the scenarios. When a problem was solved with increasing accuracy, i.e., with decreasing ϵ tolerance, iteration count was found to increase in proportion with $\ln \frac{1}{\epsilon}$.

5 Computational study

The purpose of this study is to compare the proposed model (4), based on comparison of scaled tails, with the model (2) proposed by Roman, Darby-Dowman and Mitra (based on comparison of unscaled tails) with respect to:

1. The computational behaviour: we solved the two models using their cutting plane representations (section 3.1) with the methods described in Section 4, i.e. the pure cutting plane method and the level method. We compared the number of iterations required in order to reach ϵ optimality.
2. The modelling aspect: we analyse the return distributions of the portfolio solutions of (2) and (4) respectively.

5.1 Implementation issues

The methods were implemented using the AMPL modelling system (Fourer, Gay and Kernighan 1989) and the AMPL COM Component Library (Sadki 2005), integrated with C functions. Under AMPL we use the FortMP solver. FortMP was developed at Brunel University and NAG Ltd by Ellison et al. (1999), the project being co-ordinated by E.F.D. Ellison.

In our cutting-plane system, cut generation is implemented in C, and cutting-plane model problem data are forwarded to AMPL in each iteration. Hence the bulk of the arithmetic computations is done in C, since the number of the scenarios is typically large as compared to the number of the assets. Moreover, our test results imply that acceptable accuracy can be achieved by a relatively small number of cuts in the master problem. Hence the sizes of the model problems do not directly depend on the number of scenarios.

The methods were terminated when the difference between the upper and lower bounds on the objective function $\varphi(\boldsymbol{x})$ became less or equal to the specified absolute tolerance ϵ.

scenarios	pure cutting plane iterations original	scaled	regularised iterations original	scaled
5,000	60	74	23	39
7,000	84	79	27	45
10,000	73	97	28	45
15,000	91	74	24	39
20,000	120	97	27	45
30,000	92	97	27	48

Table 1: Iteration counts

Even though the implementation of the methods leaves many possibilities for speed-up, the performance of the methods was reasonably good. Even the largest problems with 30000 scenarios were solved within 1 minute on a computer with 1.73 GHz Intel Core Duo CPU and 2 GB of RAM running Windows XP.

5.2 Test problems

We generated scenario sets using Geometric Brownian Motion (GBM), which is standard in finance for modelling asset prices, see e.g., Ross (2002). Parameters for scenario generation were derived from a data set of 132 historical monthly returns of 76 stocks (all the stocks that belonged to the FTSE 100 during the period January 1993 - December 2003).

For reference return $\widehat{R}$, we used the FTSE 100 index. Scenario sets for the FTSE 100 monthly return were generated in the same way (using GBM and and historical returns of the index between January 1993 - December 2003).

We tested with different scenario sets, each containing up to 30000 scenarios. (A single scenario consists of 77 return values: one for each of the 76 component stocks, and one for the FTSE 100 index.)

5.3 Analysis of test results

In the first experiment, we compared the solution methods. We counted the iterations the different methods made until reaching ϵ-optimal solutions. Typical iteration counts are cited in Table 1. (They were obtained with stopping tolerance set to $\epsilon = 1e{-}7$, and the level parameter set to 0.5.) It can be seen that regularisation substantially decreases the number of iterations.

In the second experiment, given a scenario set, we solved both problems (the "original" model (2) and the "scaled" model (4)) and compared the return distributions of the optimal portfolios. We made several comparisons with similar results. Basically, the proposed scaled tail model results in a return distribution that is mostly shifted to the right, as compared to the return distribution of the "original" model (i.e. the model proposed by Roman, Darby-Dowman and Mitra).

Figure 1 depicts the histograms of the return distributions obtained, for the case of the 30000-scenario problem, using the historical dataset described at the previous section (January 1993 - December 2003). The "original" distribution (in blue) is the return distribution of the portfolio obtained with model (2) of Roman, Darby-Dowman and Mitra). The "scaled" distribution (in red) is the return distribution of the portfolio obtained with the presently proposed model (4). The yellow histogram represents the reference distribution, of the FTSE 100 index.
The first two distributions clearly dominate the reference FTSE 100 distribution.
The "scaled" distribution is mostly "shifted to the right", as compared to the "original" distribution.
We underline that none of these distributions ("scaled" and "original") dominates the other; they are both

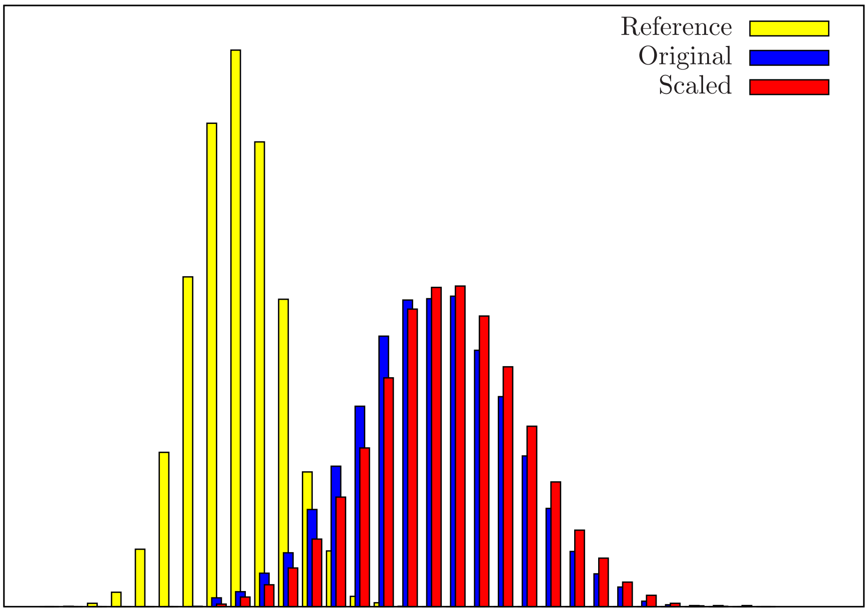

Figure 1: Dataset Jan 1993 - Dec 2003. Histograms for the return distributions of the optimal portfolios of SSD based models (”original” and ”scaled”) and for the FTSE100 Index (”reference”)

non-dominated with respect to SSD. The ”original” distribution (in blue) has a slightly better ”worst-case outcome” than the ”scaled” distribution (there is an invisible red bin situated at the left of the blue histogram). However, the ”scaled” distribution has in most cases larger numbers of outcomes in the bins situated at the right.
Figure 2 depicts the cumulative distribution functions of the return distributions described above. The cumulative distribution function of the ”scaled” distribution is generally ”lower”, indicating overall higher outcomes. It is clear that both ”original” and ”scaled” distributions dominate the index. The scaled distribution has a higher mean at no expense on the standard deviation - Table 2 presents statistics of the three return distributions considered.

We repeated the tests using scenario sets constructed from different historical datasets. A 30,000 scenarios set has been created using a dataset of 70 stocks from FTSE 100 (plus the FTSE 100 index itself) with prices monitored monthly from Dec 1992 to Apr 2000 (the histograms of the resulting distributions are displayed in Figure 3 and the cumulative distribution functions in Figure 4).
A similar dataset, but with prices monitored monthly from May 2000 to Sep 2007, was used for generating a further 30,000 scenarios set (the histograms of the resulting distributions are displayed in Figure 5 and the cumulative distribution functions in Figure 6).
In all cases the ”scaled” distribution is mostly shifted to the right, and its cumulative distribution is ”smaller”; exception makes only a very small part of the left tails. Although the worst-scenario outcomes are slightly better in the case of the ”original” distribution, the vast majority of outcomes (and the mean value) are better for the ”scaled” distribution.
We believe that an investor would prefer the ”scaled” distribution (red histogram).

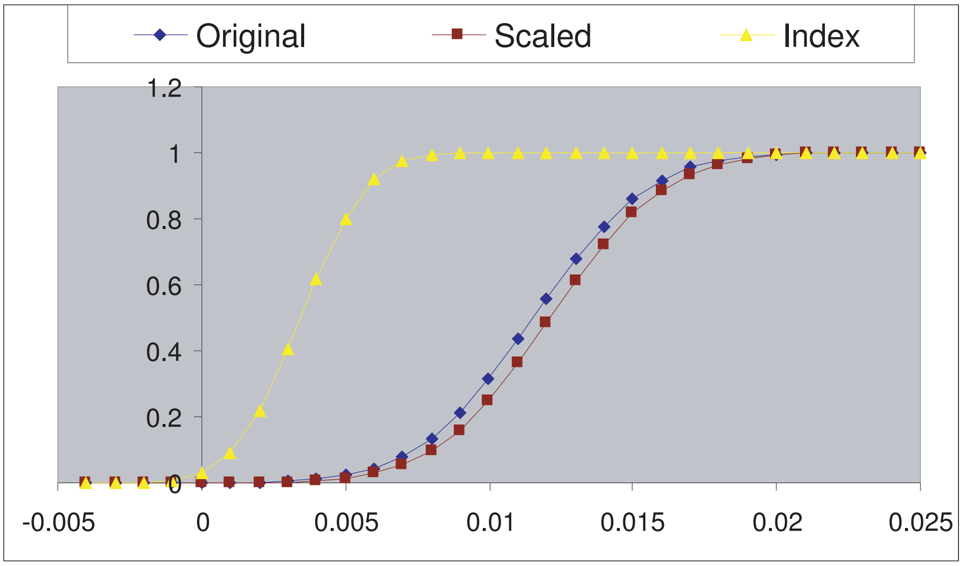

Figure 2: Dataset Jan 1993 - Dec 2003. Cumulative distribution functions for the return distributions of the optimal portfolios of SSD based models (”original” and ”scaled”) and for the FTSE100 Index (”reference”)

	original	scaled	index
Mean	0.0115	0.0121	0.0034
Median	0.0115	0.0121	0.0034
Standard Deviation	0.0032	0.0032	0.0018
Excess Kurtosis	-0.0441	-0.0147	-0.0267
Skewness	0.0318	0.0236	0.0100
Range	0.0215	0.0232	0.0136
Minimum	0.0023	0.0017	-0.0034
Maximum	0.0238	0.0250	0.0102

Table 2: Statistics of the three return distributions considered

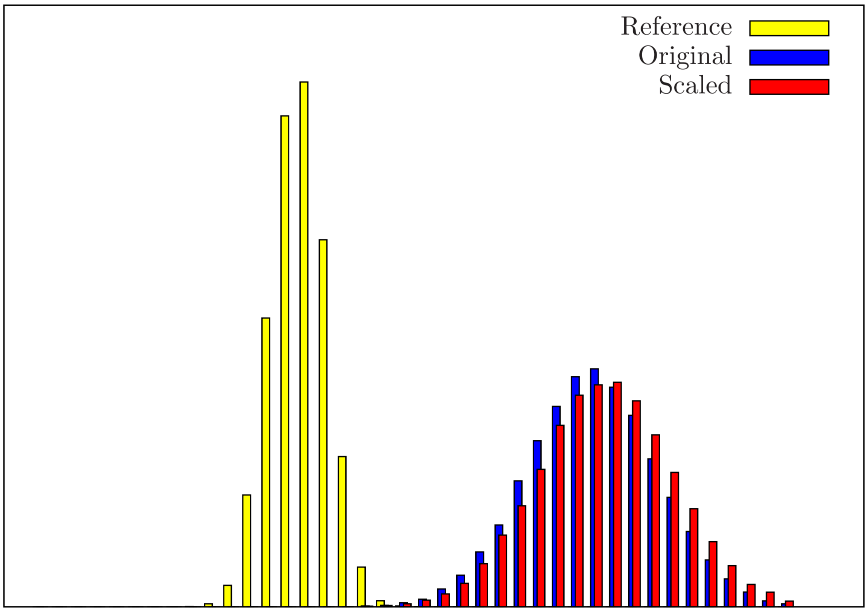

Figure 3: Dataset Dec 1992 - Apr 2000. Histograms for the return distributions of the optimal portfolios of "original" and "scaled" models and for the FTSE100 Index ("reference")

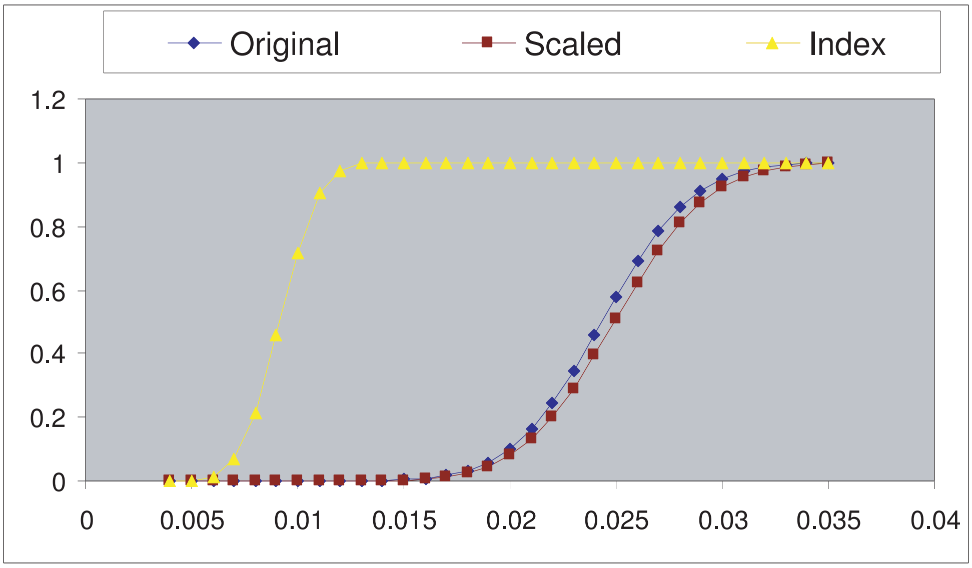

Figure 4: Dataset Dec 1992 - Apr 2000. Cumulative distribution functions for the return distributions of the optimal portfolios of SSD based models ("original" and "scaled") and for the FTSE100 Index ("reference")

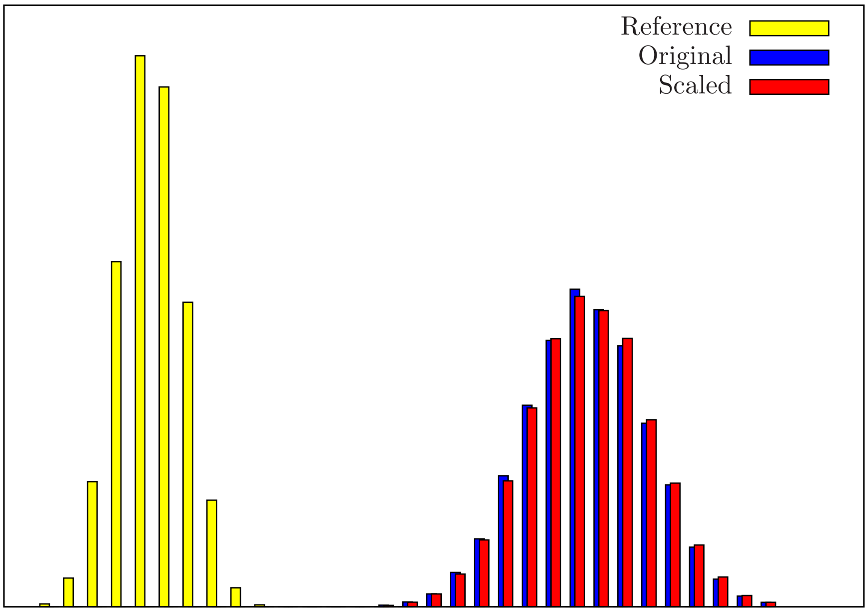

Figure 5: Dataset May 2000 - Sep 2007. Histograms for the return distributions of the optimal portfolios of "original" and "scaled" models and for the FTSE100 Index ("reference")

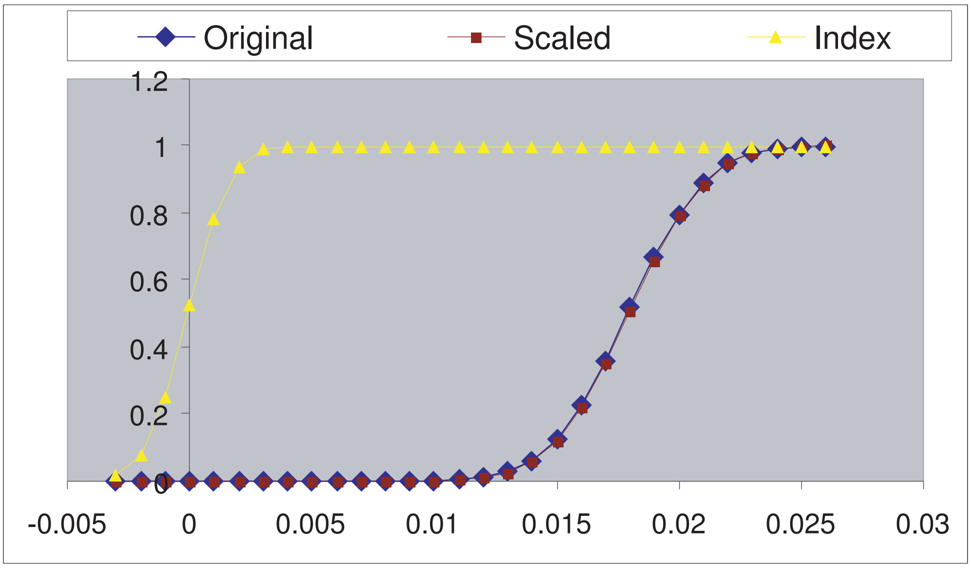

Figure 6: Dataset May 2000 - Sep 2007. Cumulative distribution functions for the return distributions of the optimal portfolios of SSD based models ("original" and "scaled") and for the FTSE100 Index ("reference")

6 Discussion and conclusions

In this paper we have proposed an enhanced version of the SSD-based portfolio-selection model of Roman, Darby-Dowman and Mitra (2006). The present approach is based on comparison of the scaled tails of the distributions. This approach has advantages both from a theoretical and a practical perspective.

The enhanced model can be also formulated as a risk minimisation model using a convex risk measure. Moreover, the new model offers a natural generalisation of the SSD-constrained formulation of Dentcheva and Ruszczyński (2006).

Our empirical study reveals that the enhanced model leads to portfolios with return distributions superior to those obtained by the model of Roman, Darby-Dowman and Mitra. The salient feature of the enhanced model is that the new return distributions are mostly shifted to the right, in relation to the original model.

The enhanced model is formulated constructively using a cutting plane representation and we have compared solution methods. The level method has proven more effective than the pure cutting-plane method, and the former method showed better scale-up properties. We have solved problems with tens of thousands of scenarios; in all cases, the solution time was less than one minute.

Acknowledgements

The authors would like to acknowledge the support for this research from several sources.
Professor Csaba Fábián's research has been partly supported by OTKA, Hungarian National Fund for Scientific Research, project 47340 and by Mobile Innovation Centre, Budapest University of Technology, project 2.2; his visiting academic position to CARISMA has been supported by OptiRisk Systems, Uxbridge, UK and by BRIEF (Brunel University Research Innovation and Enterprise Fund).
Dr Diana Roman's contribution to this work was supported by BRIEF (Brunel University Research Innovation and Enterprise Fund). The PhD studies of Victor Zverovich have been supported by OptiRisk Systems. These sources of support are gratefully acknowledged.

References

[1] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance* **9** 203-227.

[2] Carr, P., H. Geman, and D. Madan (2001). Pricing and hedging in incomplete markets. *Journal of Financial Economics* **62** 131-167.

[3] Delbaen, F. (2002) Coherent risk measures on general probability spaces. *Essays in Honour of Dieter Sondermann.* Springer-Verlag, Berlin, Germany.

[4] Dempster, M.A.H. and R.R. Merkovsky (1995). A practical geometrically convergent cutting plane algorithm. *SIAM Journal on Numerical Analysis* **32**, 631-644.

[5] Dentcheva, D. and A. Ruszczyński (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization* **14**, 548-566.

[6] Dentcheva, D. and A. Ruszczyński (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance* **30**, 433-451.

[7] Ellison, E.F.D., M. Hajian, R. Levkovitz, I. Maros, and G. Mitra (1999). A Fortan based mathematical programming system FortMP. Brunel University, Uxbridge, UK and NAG Ltd, Oxford UK.

[8] Fábián, C.I., G. Mitra, and D. Roman (2009). Processing Second-Order Stochastic Dominance models using cutting-plane representations. *Mathematical Programming* DOI 10.1007/s10107-009-0326-1.

[9] Fábián, C.I. and Z. Szőke (2007). Solving two-stage stochastic programming problems with level decomposition, *Computational Management Science* **4**, 313-353.

[10] Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and Stochastics* **6** 429-447.

[11] Fourer, R., D. M. Gay, and B. Kernighan. (1989). *AMPL: A Mathematical Programming Language.*

[12] Hadar, J. and W. Russel (1969). Rules for ordering uncertain prospects. *The American Economic Review* **59** 25-34.

[13] Heath, D. (2000). Back to the future. *Plenary Lecture at the First World Congress of the Bachelier Society,* Paris, June 2000.

[14] Kijima, M. and M. Ohnishi (1993). Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities. *Annals of Operations Research* **45**, 147-163.

[15] Klein Haneveld, W.K. (1986). Duality in Stochastic Linear and Dynamic Programming. *Lecture Notes in Economics and Math. Systems* **274**. Springer-Verlag, New York.

[16] Klein Haneveld, W.K. and M.H. van der Vlerk (2006). Integrated chance constraints: reduced forms and an algorithm. *Computational Management Science* **3**, 245-269.

[17] Kroll, Y. and H. Levy (1980). Stochastic dominance: A review and some new evidence. *Research in Finance* **2**, 163-227.

[18] Künzi-Bay, A. and J. Mayer (2006). Computational aspects of minimizing conditional value-at-risk. *Computational Management Science* **3**, 3-27.

[19] Lemaréchal, C., A. Nemirovskii, and Yu. Nesterov (1995). New variants of bundle methods. *Mathematical Programming* **69**, 111-147.

[20] Luedtke, J. (2008). New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization* **19**, 1433-1450.

[21] Ogryczak, W. and A. Ruszczyński (2001). On consistency of stochastic dominance and mean-semideviations models. *Mathematical Programming* **89**, 217-232.

[22] Ogryczak, W. and A. Ruszczyński (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization* **13**, 60-78.

[23] Rockafellar, R.T. (2007). Coherent approaches to risk in optimization under uncertainty. *Tutorials in Operations Research* INFORMS 2007, 38-61.

[24] Rockafellar, R.T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk* **2**, 21-41.

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

[26] Rockafellar, R. T., S. Uryasev, and M. Zabarankin (2002). Deviation measures in risk analysis and optimization. *Research Report* 2002-7, Department of Industrial and Systems Engineering, University of Florida.

[27] Rockafellar, R. T., S. Uryasev, and M. Zabarankin (2006). Generalised deviations in risk analysis. *Finance and Stochastics* **10**, 51-74.

[28] Roman, D., K. Darby-Dowman, and G. Mitra (2006). Portfolio construction based on stochastic dominance and target return distributions. *Mathematical Programming* Series B **108**, 541-569.

[29] Ross, S.M. (2002). *An Elementary Introduction to Mathematical Finance.* Cambridge University Press.

[30] Rudolf, G. and A. Ruszczyński (2008). Optimization problems with second order stochastic dominance constraints: duality, compact formulations, and cut generation methods. *SIAM Journal on Optimization* **19**, 1326-1343.

[31] Ruszczyński, A. and A. Shapiro (2006). Optimization of convex risk functions. *Mathematics of Operations Research* **31**, 433-452.

[32] Sadki, A.M. (2005). AMPL COM Component Library, User's Guide Version 1.6. *Internal T. Report.* See also http://www.optirik-systems.com/products/AMPLCOM.

[33] Whitmore, G.A. and M.C. Findlay (1978). *Stochastic Dominance: An Approach to Decision-Making Under Risk.* D.C.Heath, Lexington, MA.