CARISMA

A multi-disciplinary research centre focussed on understanding, modelling, quantification, management and control of RISK

- School of Information Systems, Computing and Mathematics
- Brunel Business School
- School of Engineering
- School of Social Sciences

TECHNICAL REPORT

CTR/80-08

An Enhanced Model for Portfolio Choice with SSD Criteria: a Constructive Approach.

Csaba I. Fabian, Gautam Mitra, Diana Roman, Victor Zviarovich

CARISMA

Centre for the Analysis of Risk and Optimisation Modelling Applications

Brunel UNIVERSITY WEST LONDON

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

Csaba I. Fábián* Gautam Mitra† Diana Roman‡ Victor Zviarovich§

Contents

*Institute of Informatics, Kecskemét College and Department of Operations Research, Loránd Eötvös University of Budapest. Postal address: P.O.Box 120, Budapest, H-1518 Hungary. E-mail: fabian@cs.elte.hu. This research has been supported by OptiRisk Systems, Uxbridge, UK ; by OTKA, Hungarian National Fund for Scientific Research, project No. 47340 ; and by the Mobile Innovation Centre of the Budapest University of Technology, integrated project No. 2.2.

†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 Abstract

We formulate a portfolio choice model applying Second-order Stochastic Dominance. This model is an enhanced version of the multi-objective model proposed by Roman, Darby-Dowman, and Mitra (2006). We use a special scaling of the different objectives, representing tails at different confidence levels of the resulting distribution. The new model offers a natural generalisation of the SSD-constrained model of Dentcheva and Ruszczyński (2006). Moreover, it can be formulated as risk minimisation model with the use of a convex risk measure. We characterise this risk measure and the resulting optimisation problem. We outline solution methods for the proposed model, based on cutting plane representations. 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.

2 Introduction

2.1 Second-order Stochastic Dominance

Let R and R' denote random returns. Second-order Stochastic Dominance is defined by the following equivalent criteria:

(a) $\mathrm{E}\,(U(R)) \geq \mathrm{E}\,(U(R'))$ holds for any monotonic and concave (integrable) utility function U.

(b) $\mathrm{E}\,([t-R]_+) \;\leq\; \mathrm{E}\,([t-R']_+)$ holds for each $t \in \mathbb{R}$.

(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). 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 $\mathrm{E}\,([t-R]_+)$ 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', 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 $\boldsymbol{x} \in \mathbb{R}^n$ represents the amounts of money invested in the different assets. Let the n-dimensional random vector $\boldsymbol{R}$ denote the returns of the different assets. The random yield of portfolio $\boldsymbol{x}$ will be $R_{\boldsymbol{x}} := \boldsymbol{R}^T\boldsymbol{x}$.

Let $X \subset \mathbb{R}^n$ 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 $R_{\boldsymbol{x}} \succ_{_{SSD}} R_{\boldsymbol{x}^\star}$.

Let $\widehat{R}$ denote some reference return, possibly a stock index, or the return of a benchmark portfolio. In this paper, excepting Section 3, we assume that the joint distribution of $\boldsymbol{R}$ and $\widehat{R}$ is discrete finite, having S equally probable outcomes.

2.2 Portfolio optimisation models

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

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

where f is a concave function. In particular, they consider $f(\boldsymbol{x}) = \mathrm{E}\left(R_{\boldsymbol{x}}\right)$. They formulate the problem using criterion *(b)* ; 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)*.

Roman, Darby-Dowman, and Mitra (2006) use criterion *(c)*. 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 $\mathrm{Tail}_{\frac{i}{S}}\big(R_{\boldsymbol{x}}\big) \;\geq\; \mathrm{Tail}_{\frac{i}{S}}\big(\widehat{R}\big) \quad (i = 1, \ldots, S)$. Roman et al. 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 in a uniform sense. Uniformity is meant in terms of differences among tails:

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

Roman et al. 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.

2.3 Contribution of present paper

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

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

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

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

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

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 portfolio $\boldsymbol{x}$ to the combination of the reference portfolio and ϑ in 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 cash 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 (2008). 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

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

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

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

with different values of $\lambda \geq 0$. 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 3, 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 4, we compare different formulations of the enhanced portfolio choice problem (3).

In Section 5, 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 6, 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 7.

3 Convex risk measures

3.1 Overview of risk measures

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

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

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

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

The concept of *coherent risk measures* was developed by Artzner et al. (1999) and Delbaen (2002). These measures satisfy the four criteria of sub-additivity, positive homogeneity, monotonicity, and translation equivariance. The acceptance set of a coherent measure is a convex cone.

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

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

(In some papers, CVaR_α is defined using $(1-\alpha)$-tails.)

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 risk measure ρ is said to be convex if it satisfies the following criteria:

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

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

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

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 $\varrho \in \mathbb{R}$ (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

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

where $\mathcal{Q}$ is a set of probability measures, and α is a mapping from the set of the probability measures to $(-\infty, +\infty]$. (Properties of $\mathcal{Q}$ and α 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.

3.2 Convexity of the risk measure $\widehat{\rho}$

The risk measure $\widehat{\rho}$ defined in (5) derives from the acceptance set $\widehat{\mathcal{A}} := \Big\{ R \in \mathcal{R} \,\Big|\, R \succeq_{SSD} \widehat{R} \Big\}$ 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*

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

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

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

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.)

In the remaining part of this subsection, we just show that $\widehat{\mathcal{A}}$ satisfies the criteria of Proposition 1.

In order to prove that $\widehat{\mathcal{A}}$ is convex, let $R, R' \in \widehat{\mathcal{A}}$, i.e., $R, R' \succeq_{SSD} \widehat{R}$. We show that $\lambda R + (1-\lambda)R' \in \widehat{\mathcal{A}}$ for $0 \le \lambda \le 1$. Let us first observe that the dominance criterion *(a)* of Section 2.1 can be reformulated as

(a′) $\mathrm{E}U(R) \ge \mathrm{E}U(R')$ for any monotonic and concave (integrable) utility function U having $U(R_0) = 0$.

Let U be a such a utility function. Expected utility inherits concavity of U, hence we have

$$\mathrm{E}U\big(\lambda R + (1-\lambda)R'\big) \;\ge\; \mathrm{E}U(\lambda R) + \mathrm{E}U\big((1-\lambda)R'\big). \tag{16}$$

The function $\lambda \mapsto \mathrm{E}U(\lambda R)$ is obviously concave, and for $\lambda = 0$ or 1 we have $\mathrm{E}U(\lambda R) = \lambda \mathrm{E}U(R)$. Hence

$$\mathrm{E}U(\lambda R) \;\ge\; \lambda \mathrm{E}U(R) \qquad \text{and} \qquad \mathrm{E}U\big((1-\lambda)R'\big) \;\ge\; (1-\lambda)\mathrm{E}U(R'). \tag{17}$$

From our assumptions, we have

$$\mathrm{E}U(R),\ \mathrm{E}U(R')\ \geq \mathrm{E}U(\widehat{R}). \tag{18}$$

Putting (16), (17), (18) together, we get

$$\mathrm{E}U\big(\lambda R + (1-\lambda)R'\big)\ \geq\ \mathrm{E}U(\widehat{R}).$$

According to *(a′)* above, this implies $\lambda R + (1-\lambda)R' \succeq_{SSD} \widehat{R}$, which proves convexity of $\widehat{\mathcal{A}}$.

$\widehat{\mathcal{A}}$ evidently satisfies (14), and (15) is a consequence of the monotonicity of our utility functions.

3.3 Dual representation of $\widehat{\rho}$

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 exists. (The set Ω is equipped with the probability measure P, the field of measurable sets being $\mathcal{M}$.) 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

$$\mathrm{CVaR}_\alpha(R) = \sup_{d_Q \leq \alpha^{-1}} \mathrm{E}_Q(-R) \qquad (0 < \alpha \leq 1). \tag{19}$$

Based on this result, we construct a dual representation of $\widehat{\rho}$. According to the definition *(c)* of SSD in Section 2.1, we have

$$\widehat{\mathcal{A}} = \bigcap_{0<\alpha\leq 1} \widehat{\mathcal{B}}_\alpha \tag{20}$$

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 (19) 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 (20), we get

$$\widehat{\mathcal{A}} = \left\{ R \,\middle|\, \mathrm{E}_Q(-R) \leq \mathrm{CVaR}_\alpha(\widehat{R}) \text{ holds for each } (Q, \alpha) \text{ having } d_Q \leq \alpha^{-1} \right\}. \tag{21}$$

Let us define

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

(We have $s(Q) \leq 1$ for each legitimate Q.) Equality (21) 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)

$$\begin{aligned} \widehat{\rho}(R) &= \inf \left\{ \varrho \in \mathrm{I\!R} \,\middle|\, \mathrm{E}_Q(-R - \varrho) \leq \mathrm{CVaR}_{s(Q)}(\widehat{R}) \text{ holds for each } Q \in \mathcal{Q} \right\} \\ &= \sup_{Q\in\mathcal{Q}} \left\{ \mathrm{E}_Q(-R)\ -\ \mathrm{CVaR}_{s(Q)}(\widehat{R}) \right\} \end{aligned} \tag{22}$$

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

In the remaining part of this paper we focus on discrete finite distributions with equiprobable outcomes. In this case $\mathcal{R}$ is a finite dimensional space, and the acceptance set $\widehat{\mathcal{A}}$ is a polyhedron with a highly symmetric structure.

4 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 2.1.

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 $\widehat{\tau}_i := \mathrm{Tail}_{\frac{i}{S}}\left(\widehat{R}\right) \quad (i = 1, \ldots S)$.

4.1 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):

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

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 (2008) used the following cutting-plane representation, adapting the approach of Künzi-Bay and Mayer (2006):

$$\begin{aligned} \mathrm{Tail}_{\frac{i}{S}}\left(R_{\boldsymbol{x}}\right) = \quad & \min \ \tfrac{1}{S} \sum_{j \in \mathcal{J}_i} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \\ & \text{such that} \quad \mathcal{J}_i \subset \{1, \ldots, S\}, \ \ |\mathcal{J}_i| = i. \end{aligned} \tag{23}$$

The above formula 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 (23) to the present, scaled-tail model (4) results the following cutting-plane representation:

$$\begin{aligned} & \max && \vartheta \\ & \text{such that} && \vartheta \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\ & && \tfrac{i}{S}\vartheta + \widehat{\tau}_i \ \le \ \tfrac{1}{S} \sum_{j \in \mathcal{J}_i} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \qquad \text{for each} \quad \mathcal{J}_i \subset \{1, \ldots, S\}, \ \ |\mathcal{J}_i| = i, \\ & && \qquad\qquad\qquad\qquad\qquad\qquad \text{where} \quad i = 1, \ldots, S. \end{aligned} \tag{24}$$

4.2 Formulation using integrated chance constraints

In their SSD-constrained model (1), Dentcheva and Ruszczyński characterise stochastic dominance with criterion *(b)* in Section 2.1. 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

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

The constraint $R_{\boldsymbol{x}} \succeq_{SSD} \widehat{R} + \vartheta$ of the present enhanced model (3) can be formulated in the same manner:

$$\sum_{j=1}^{S} \frac{1}{S} \left[\widehat{r}^{(i)} + \vartheta - \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \right]_+ \leq \sum_{j=1}^{S} \frac{1}{S} \left[\widehat{r}^{(i)} - \widehat{r}^{(j)} \right]_+ \qquad (i = 1, \ldots, S). \tag{25}$$

The inequalities in (25) are integrated chance constraints, and can be formulated as finite sets of linear inequalities using the cutting-plane representation of Klein Haneveld and Van der Vlerk (2006). These authors also present a computational study demonstrating the effectiveness of their cutting-plane approach for optimisation problems. The cutting-plane representation of the ith constraint from our system (25) 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 cutting-plane representation for each integrated chance constraint, the enhanced model (3) can be formulated as

$$\begin{aligned}
&\max \quad \vartheta \\
&\text{such that} \quad \vartheta \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\
&\sum_{j \in \mathcal{J}_i} \tfrac{1}{S} \left\{ \widehat{r}^{(i)} + \vartheta - \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \right\} \leq \sum_{j=1}^{S} \tfrac{1}{S} \left[\widehat{r}^{(i)} - \widehat{r}^{(j)} \right]_+ \quad \text{for each } \mathcal{J}_i \subset \{1, \ldots, S\}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{where } i = 1, \ldots, S.
\end{aligned} \tag{26}$$

The problems (26) and (24) 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 (26) is

– equivalent to the constraint belonging to $\mathcal{J}_i$ in (24), if $|\mathcal{J}_i| = i$;

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

In the remaining part of this paper we will use the formulation (24) as it is more economic under our assumptions.

Remark 2 *If we drop the equiprobability assumption, but keep the discrete finite assumption, then the formulation (26) will be more convenient.*

4.3 Connection with risk measures

Changing the scope of optimisation, problem (24) can be formulated as minimisation of a polyhedral convex function:

$$\min_{\boldsymbol{x} \in X} \varphi(\boldsymbol{x}) \quad \text{where} \quad \varphi(\boldsymbol{x}) := \max \left\{ -\tfrac{1}{i} \sum_{j \in \mathcal{J}_i} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} + \tfrac{S}{i} \widehat{\tau}_i \right\}$$
$$\text{such that} \quad \mathcal{J}_i \subset \{1, \ldots, S\}, \ |\mathcal{J}_i| = i, \tag{27}$$
$$\text{where} \quad i = 1, \ldots, S.$$

We have $\frac{S}{i}\widehat{\tau}_i = -\mathrm{CVaR}_{\frac{i}{S}}(\widehat{R})$ according to (11). The above definition of φ is just the specialisation of the dual representation (22) to the present discrete, finite, equiprobable case. Hence we have $\varphi(\boldsymbol{x}) = \widehat{\rho}(R_{\boldsymbol{x}})$ with the convex risk measure introduced in Section 3.

5 Solution methods

We solve the portfolio optimisation problem in the form (27).

5.1 Pure cutting plane method

Applied to a convex programming problem $\min_{\boldsymbol{x} \in \mathcal{X}} \phi(\boldsymbol{x})$, 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 (27) 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 $\mathcal{J}_i^\star := \{j_1^\star, \ldots, j_i^\star\} \quad (i = 1, \ldots, S)$.
- Let us select $i^\star \in \arg\max_{1 \leq i \leq S} \left\{ -\frac{1}{i} \sum_{j \in \mathcal{J}_i^\star} \boldsymbol{r}^{(j)\,T} \boldsymbol{x}^\star + \frac{S}{i} \widehat{\tau}_i \right\}$.
- A supporting linear function at $\boldsymbol{x}^\star$ is then $l(\boldsymbol{x}) := -\frac{1}{i^\star} \sum_{j \in \mathcal{J}_{i^\star}^\star} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} + \frac{S}{i^\star} \widehat{\tau}_{i^\star}$.

Klein Haneveld and Van der Vlerk (2006), Künzi-Bay and Mayer (2006), Fábián, Mitra, and Roman (2008) 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.

5.2 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 $\ln \frac{1}{\epsilon}$.

6 Computational study

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

1. The computational behaviour: we solved the two models using their cutting plane representations (section 4.1) with the methods described in Section 5. 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.

6.1 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.

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 min on a computer with 1.73 GHz Intel Core Duo CPU and 2 GB of RAM running Windows XP.

6.2 Test problems

We generated scenario sets using Geometric Brownian Motion, 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, 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.)

6.3 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, 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. The "original" distribution is the return distribution of the portfolio obtained with model (2) of Roman, Darby-Dowman and Mitra). The "scaled" distribution is the return distribution of the portfolio obtained with the presently proposed model (4). We also depict the return distribution of the FTSE 100 index, which is the reference return.

scenarios	pure cutting plane iterations		regularised iterations	
	original	scaled	original	scaled
5,000	71	74	23	39
7,000	83	79	27	45
10,000	73	97	28	45
15,000	91	74	24	39
20,000	119	97	27	45
30,000	96	97	27	48

Table 1: Iteration counts

	original	scaled	index
Mean	0.0116	0.0122	0.0034
Median	0.0116	0.0122	0.0034
Standard Deviation	0.0032	0.0032	0.0018
Excess Kurtosis	0.0013	0.0079	-0.0267
Skewness	-0.0090	-0.0028	0.0100
Range	0.0257	0.0254	0.0136
Minimum	-0.0013	-0.0003	-0.0034
Maximum	0.0244	0.0251	0.0102

Table 2: Statistics of the three return distributions considered

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 non-dominated with respect to SSD. However, the "scaled" distribution has in most cases larger numbers of outcomes in the bins situated at the right.

We think that a decision maker would prefer the "scaled" distribution.

Table 6.3below presents statistics of the three return distributions considered above. The "scaled" return distribution has better statistics (higher expected return, higher minimum, etc.)
The situation is similar in the case of 10,000 and 15,000 scenarios.

7 Discussion and conclusion

In this paper we 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 from both a theoretical and a practical point of view.

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

Our numerical results showed that the proposed model results in return distributions superior to those resulted from the model of Roman et al.

The proposed model was formulated using a cutting plane representation and we compared solution methods. The level method proved more effective than the pure cutting-plane method, and the former method showed better scale-up properties. We solved problems with tens of thousands of scenarios; in all

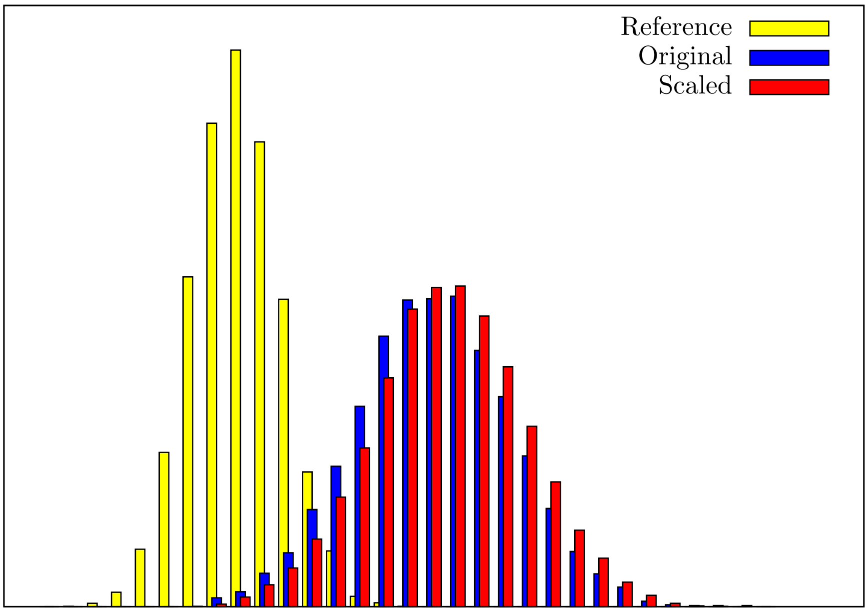

Figure 1: Histograms for the return distributions of the optimal portfolios of SSD based models (”original” and ”scaled) and for the FTSE100 Index (”reference”)

cases, the solution time was less than one minute.

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 (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance* **30**, 433-451.

[6] 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.

[7] Fábián, C.I., G. Mitra, and D. Roman (2008). Processing Second-Order Stochastic Dominance models using cutting-plane representations. *Stochastic Programming E-Print Series* 2008-10.

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

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

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

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

[12] 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.

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

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

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

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

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

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

[19] 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.

[20] Rockafellar, R. T., *S. Uryasev*, and *M. Zabarankin* (2006). Generalised deviations in risk analysis. *Finance and Stochastics* **10** 5174.

[21] 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.

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

[23] Ruszczyński, A. and A. Shapiro (2006). Optimization of convex risk functions. *Mathematics of Operations Research* **31** 433452.

[24] 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.

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