Intl. Trans. in Op. Res. 9 (2002) 661–680

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Dual stochastic dominance and quantile risk measures

Włodzimierz Ogryczak and Andrzej Ruszczyński[a]

Warsaw University of Technology, Institute of Control and Computation Engineering, Nowowiejska 15/19, 00665 Warsaw, Poland

[a]*Rutgers University, RUTCOR and Department of Management Science and Information Systems, Piscataway, NJ 08854, USA*

Received 18 June 2001; received in revised form 19 December 2001; accepted 23 January 2002

Abstract

Following the seminal work by Markowitz, the portfolio selection problem is usually modeled as a bicriteria optimization problem where a reasonable trade-off between expected rate of return and risk is sought. In the classical Markowitz model, the risk is measured with variance. Several other risk measures have been later considered thus creating the entire family of mean-risk (Markowitz type) models. In this paper, we analyze mean-risk models using quantiles and tail characteristics of the distribution. Value at risk (VAR), defined as the maximum loss at a specified confidence level, is a widely used quantile risk measure. The corresponding second order quantile measure, called the worst conditional expectation or Tail VAR, represents the mean shortfall at a specified confidence level. It has more attractive theoretical properties and it leads to LP solvable portfolio optimization models in the case of discrete random variables, i.e., in the case of returns defined by their realizations under the specified scenarios. We show that the mean-risk models using the worst conditional expectation or some of its extensions are in harmony with the stochastic dominance order. For this purpose, we exploit duality relations of convex analysis to develop the quantile model of stochastic dominance for general distributions.

Keywords: Portfolio optimization, stochastic dominance, mean-risk, value at risk, linear programming

1. Introduction

The relation of stochastic dominance is one of the fundamental concepts of decision theory (cf. Whitmore and Findlay, 1978; Levy, 1992). It introduces a partial order in the space of real random variables. The first-degree relation carries over to expectations of monotone utility functions, and the second-degree relation to expectations of concave non-decreasing utility functions. While theoretically attractive, stochastic dominance order is computationally very difficult, as a multi-objective model with a continuum of objectives. Following the seminal work by Markowitz (1952), the portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem. The classical Markowitz model uses the variance as the risk measure. Since then many authors have pointed out that the mean–variance model is, in general, not consistent with stochastic dominance rules.

When applied to portfolio selection or similar optimization problems with polyhedral feasible sets, the mean-variance approach results in a quadratic programming problem. Following Sharpe's (1971a) work on linear programming (LP) approximation to the mean-variance model, many attempts have been made to linearize the portfolio optimization problem. This resulted in the consideration of various risk measures which were LP-computable in the case of finite discrete random variables. Yitzhaki (1982) introduced the mean-risk model using the Gini's mean (absolute) difference as a risk measure. Konno and Yamazaki (1991) analyzed the model where risk is measured by the (mean) absolute deviation. Young (1998) considered the minimax approach (the worst-case performances) to measure the risk. If the rates of return are multivariate normally distributed, then most of these models are equivalent to the Markowitz' mean-variance model. However, they do not require any specific type of return distributions and, opposite to the mean-variance approach, they can be applied to general (possibly non-symmetric) random variables (Michalowski and Ogryczak, 2001). In the case of finite discrete random variables all these mean-risk models have LP formulations and are special cases of the multiple criteria LP model (Ogryczak, 2000) based on the majorization theory (Marshall and Olkin, 1979) and Lorenz-type orders (Shorrocks, 1983).

In this paper, we analyze mean-risk models using quantiles and tail characteristics of the distribution. Value at risk (VAR), defined as the maximum loss at a specified confidence level, is a widely-used quantile risk measure (Jorion, 1997). The corresponding second order quantile measure, called the worst conditional expectation or Tail VAR, represents the mean shortfall at a specified confidence level. It leads to LP-solvable portfolio optimization models in the case of discrete random variables, i.e., in the case of returns defined by their realizations under the specified scenarios. The worst conditional expectation itself and some of its extensions have attractive theoretical properties. We show that the corresponding mean-risk models are in harmony with the stochastic dominance order. For this purpose, we exploit duality relations of convex analysis to develop the *dual* (quantile) model of the stochastic dominance for general distributions. The only restriction we impose is that all the random returns R under consideration satisfy the condition $\mathbb{E}|R| < \infty$. The results are graphically illustrated within the framework of the absolute Lorenz curves (dual SSD) as well as the Outcome–Risk diagram (primal SSD).

The paper is organized as follows. In Section 2 we formally define the concepts of consistency of the mean-risk models with the stochastic dominance relations. Section 3 introduces primal and dual shortfall risk measures and exploits the duality theory to characterize the stochastic dominance relations in terms of quantile performance funtions. This allows us to relate the quantile risk measures with the stochastic dominance. In Section 4, the LP solvability of the corresponding mean-risk models is discussed, and some extensions are considered.

2. Stochastic dominance and mean-risk models

The portfolio optimization problem considered in this paper is based on a single-period model of investment. Let $J = \{1, 2, \ldots, n\}$ denote a set of securities considered for an investment. For each security $j \in J$, its rate of return is represented by a random variable R_j with a given mean $\mu_j = \mathbb{E}\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,\ldots,n}$ denote a vector of decision variables x_j expressing the weights defining a portfolio. To represent a portfolio, the weights must satisfy a set of constraints that form a feasible set $\mathscr{P}$. The simplest way of defining a feasible set is by a requirement that the weights must sum to one, i.e. $\sum_{j=1}^{n} x_j = 1$ and $x_j \geqslant 0$ for $j = 1, \ldots, n$. An investor usually needs to consider some other requirements expressed as a set of additional side constraints. Hereafter, it is assumed that $\mathscr{P}$ is a general LP-feasible set given in a canonical form as a system of linear equations with non-negative variables.

Each portfolio $\mathbf{x}$ defines a corresponding random variable $R(\mathbf{x}) = \sum_{j=1}^{n} R_j x_j$ that represents a portfolio return. The mean return for portfolio $\mathbf{x}$ is given as $\mu(\mathbf{x}) = \mathbb{E}\{R(\mathbf{x})\} = \sum_{j=1}^{n} \mu_j x_j$. Following Markowitz (1952), the portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem

$$\max\{[\mu(\mathbf{x}), -\varrho(\mathbf{x})]: \quad \mathbf{x} \in \mathscr{P}\} \tag{1}$$

where the mean $\mu(\mathbf{x})$ is maximized and the risk measure $\varrho(\mathbf{x})$ is minimized. A feasible portfolio $\mathbf{x}^0 \in \mathscr{P}$ is called the efficient solution of problem (1) or the μ/ϱ*–efficient* portfolio if there is no $\mathbf{x} \in \mathscr{P}$ such that $\mu(\mathbf{x}) \geqslant \mu(\mathbf{x}^0)$ and $\varrho(\mathbf{x}) \leqslant \varrho(\mathbf{x}^0)$ with at least one inequality strict. We restrict our analysis to the class of Markowitz-type mean-risk models where risk measures, similar to the standard deviation, are translation invariant and risk relevant dispersion parameters. Thus the risk measures we consider are not affected by any shift of the outcome scale and they are equal to 0 in the case of a risk-free portfolio, while taking positive values for any risky portfolio. Moreover, in order to model possible taking advantages of a portfolio diversification, risk measure $\varrho(\mathbf{x})$ should be a convex function of $\mathbf{x}$.

The Markowitz model is frequently criticized as not consistent with axiomatic models of preferences for choice under risk. Namely, except for the case of returns meeting the multivariate normal distribution, the mean-variance model may lead to inferior conclusions with respect to the stochastic dominance order. The concept of stochastic dominance order (Whitmore and Findlay, 1978) corresponds to an expected utility axiomatic model of risk-averse preferences (Rothschild and Stiglitz, 1969; Levy, 1992). In stochastic dominance, uncertain returns (random variables) are compared by pointwise comparison of some performance functions constructed from their distribution functions. The first performance function $F^{(1)}$ is defined as the right-continuous cumulative distribution function: $F_R^{(1)}(\eta) = F_R(\eta) = \mathbb{P}\{R \leqslant \eta\}$. The weak relation of the first degree stochastic dominance (FSD) is defined as follows:

$$R' \succeq_{FSD} R'' \Leftrightarrow F_{R'}(\eta) \leqslant F_{R''}(\eta) \quad \text{for all } \eta.$$

The second function is derived from the first as $F_R^{(2)}(\eta) = \int_{-\infty}^{\eta} F_R(\xi) d\xi$ for real numbers η, and it defines the (weak) relation of *second degree stochastic dominance* (SSD):

$$R' \succeq_{SSD} R'' \Leftrightarrow F_{R'}^{(2)}(\eta) \leqslant F_{R''}^{(2)}(\eta) \quad \text{for all } \eta.$$

We say that portfolio $\mathbf{x}'$ dominates $\mathbf{x}''$ under the SSD ($R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$), if $F_{R(\mathbf{x}')}^{(2)}(\eta) \leqslant F_{R(\mathbf{x}'')}^{(2)}(\eta)$ for all η, with at least one strict inequality. A feasible portfolio $\mathbf{x}^0 \in \mathscr{P}$ is called SSD efficient if there is no $\mathbf{x} \in \mathscr{P}$ such that $R(\mathbf{x}) \succ_{SSD} R(\mathbf{x}^0)$. If $R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$, then $R(\mathbf{x}')$ is preferred to $R(\mathbf{x}'')$ within all risk-averse preference models where larger outcomes are preferred. It is therefore a matter of primary importance that a model for portfolio optimization be consistent with the SSD relation, which implies that the optimal portfolio is SSD-efficient.

The Markowitz model is not SSD consistent, since its efficient set may contain portfolios characterized by a small risk, but also very low return (Porter and Gaumnitz, 1972). Unfortunately, it is a common flaw of all Markowitz-type mean-risk models where risk is measured with some dispersion measures. Although, the necessary condition for the SSD relation is (Fishburn, 1980)

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow \mu(\mathbf{x}') \geqslant \mu(\mathbf{x}'') \tag{2}$$

this is not enough to guarantee the μ/ϱ dominance. For dispersion-type risk measures $\varrho(\mathbf{x})$, it may occur that $R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$ and simultaneously $\varrho(\mathbf{x}') > \varrho(\mathbf{x}'')$. This can be illustrated by two portfolios $\mathbf{x}'$ and $\mathbf{x}''$ with returns given as follows:

$$\mathbb{P}\{R(\mathbf{x}') = \xi\} = \begin{cases} 1 & \xi = 1.0 \\ 0, & \text{otherwise} \end{cases} \quad \mathbb{P}\{R(\mathbf{x}'') = \xi\} = \begin{cases} 1/2, & \xi = 3.0 \\ 1/2, & \xi = 5.0 \\ 0, & \text{otherwise} \end{cases}$$

Note that the risk-free portfolio $\mathbf{x}'$ with the guaranteed result 1.0 is obviously worse than the risky portfolio $\mathbf{x}''$ giving 3.0 or 5.0. In all preference models based on the risk aversion axioms (Artzner et al., 1999; Levy, 1992; Whitmore and Findlay, 1978) portfolio $\mathbf{x}'$ is dominated by $\mathbf{x}''$, in particular $R(\mathbf{x}'') \succ_{SSD} R(\mathbf{x}')$. On the other hand, when a dispersion-type risk measure $\varrho(\mathbf{x})$ is used, then both the portfolios are efficient in the corresponding mean-risk model since for each such a measure $\varrho(\mathbf{x}'') > 0$ while $\varrho(\mathbf{x}') = 0$.

In order to overcome this flaw of the Markowitz model, Baumol (1964) suggested a safety measure he called the expected gain-confidence limit criterion, $\mu(\mathbf{x}) - \lambda\sigma(\mathbf{x})$ to be maximized instead of the minimization of $\sigma(\mathbf{x})$ itself. Similarly, Yitzhaki (1982) considered maximization of the safety measure $\mu(\mathbf{x}) - \Gamma(\mathbf{x})$ defined by the risk measure of Gini's mean difference $\Gamma(\mathbf{x})$, and he demonstrated its SSD consistency. Markowitz (1959) introduced two downside risk measures to replace the variance: the below-target downside semivariance and the below-mean downside semivariance. For the former, its SSD consistency was shown by Porter (1994). Recently, similar consistency results have been introduced (Ogryczak and Ruszczyński, 1999, 2001) for safety measures corresponding to the below-mean downside standard semideviation and to the below-mean downside mean semideviation (half of the mean absolute deviation). Hereafter, for any dispersion type risk measure $\varrho(\mathbf{x})$, the function $s(\mathbf{x}) = \mu(\mathbf{x}) - \varrho(\mathbf{x})$ will be referred to as the corresponding *safety* measure. Note that risk measures, we consider, are defined as translation-invariant and risk-relevant dispersion parameters. Hence, the corresponding safety measures are translation equivariant in the sense that any shift of the outcome scale results in an equivalent change of the safety measure value (with opposite sign as safety measures are maximized). In other words, the safety measures distinguish (and order) various risk-free portfolios (outcomes) according to their values. The safety measures, we consider, are risk relevant, but in the sense that the value of a safety measure for any risky portfolio is less than the value for the risk-free portfolio with the same expected return. Moreover, when risk measure $\varrho(\mathbf{x})$ is a convex function of $\mathbf{x}$, then the corresponding safety measure $s(\mathbf{x})$ is concave.

The SSD consistency of the safety measures may be formalized as follows. We say that the safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$ is SSD consistent or that the risk measure $\varrho(\mathbf{x})$ is SSD safety consistent if

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow \mu(\mathbf{x}') - \varrho(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \varrho(\mathbf{x}'') \tag{3}$$

The relation of SSD (safety) consistency is called strong if, in addition to (3), the following holds

$$R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'') \Rightarrow \mu(\mathbf{x}') - \varrho(\mathbf{x}') > \mu(\mathbf{x}'') - \varrho(\mathbf{x}'') \tag{4}$$

Theorem 1. *If the risk measure $\varrho(\mathbf{x})$ is SSD safety consistent (3), then except for portfolios with identical values of $\mu(\mathbf{x})$ and $\varrho(\mathbf{x})$, every efficient solution of the bicriteria problem*

$$\max\{[\mu(\mathbf{x}), \mu(\mathbf{x}) - \varrho(\mathbf{x})]: \quad \mathbf{x} \in \mathcal{P}\} \tag{5}$$

is an SSD efficient portfolio. In the case of strong SSD safety consistency (4), every portfolio $\mathbf{x} \in \mathcal{P}$ efficient to (5) is, unconditionally, SSD efficient.

Proof. Let $\mathbf{x}^0 \in \mathcal{P}$ be an efficient solution of (5). Suppose that $\mathbf{x}^0$ is not SSD efficient. This means, there exists $\mathbf{x} \in \mathcal{P}$ such that $R(\mathbf{x}) \succ_{SSD} R(\mathbf{x}^0)$. Then, from (2) it follows $\mu(\mathbf{x}) \geqslant \mu(\mathbf{x}^0)$, and simultaneously $\mu(\mathbf{x}) - \varrho(\mathbf{x}) \geqslant \mu(\mathbf{x}^0) - \varrho(\mathbf{x}^0)$, by virtue of the SSD safety consistency (3). Since $\mathbf{x}^0$ is efficient to (5) no inequality can be strict, which implies $\mu(\mathbf{x}) = \mu(\mathbf{x}^0)$ and $\varrho(\mathbf{x}) = \varrho(\mathbf{x}^0)$.

In the case of the strong SSD safety consistency (4), the supposition $R(\mathbf{x}) \succ_{SSD} R(\mathbf{x}^0)$ implies $\mu(\mathbf{x}) \geqslant \mu(\mathbf{x}^0)$ and $\mu(\mathbf{x}) - \varrho(\mathbf{x}) > \mu(\mathbf{x}^0) - \varrho(\mathbf{x}^0)$ which contradicts the efficiency of $\mathbf{x}^0$ with respect to (5). Hence, $\mathbf{x}^0$ is SSD efficient.

Following Theorem 1, one may consider the mean-safety bicriteria model (5) as a reasonable alternative to the corresponding mean-risk model (1). Note that having $\mu(\mathbf{x}') \geqslant \mu(\mathbf{x}'')$ and $\varrho(\mathbf{x}') \leqslant \varrho(\mathbf{x}'')$ with at least one inequality strict, one gets $\mu(\mathbf{x}') - \varrho(\mathbf{x}') > \mu(\mathbf{x}'') - \varrho(\mathbf{x}'')$. Hence, a portfolio dominated in the mean-risk model (1) is also dominated in the corresponding mean-safety model (5). In other words, the efficient portfolios of problem (5) form a subset of the entire μ/ϱ–efficient set. We illustrate this in the μ/ϱ image space in Fig. 1. Due to the convexity of $\varrho(\mathbf{x})$ and linearity of $\mu(\mathbf{x})$, the portfolios $\mathbf{x} \in \mathcal{P}$ form in the μ/ϱ image space a set with the convex boundary from the side of μ–axis (i.e., the set $\{(\mu, \varrho): \mu = \mu(\mathbf{x}), \varrho \geqslant \varrho(\mathbf{x}), \mathbf{x} \in \mathcal{P}\}$ is convex). This boundary represents a curve of the relative minimum risk portfolios spanning from the best expectation portfolio (BEP) to the worst expectation portfolio (WEP). The minimum risk portfolio (MRP), defined as the

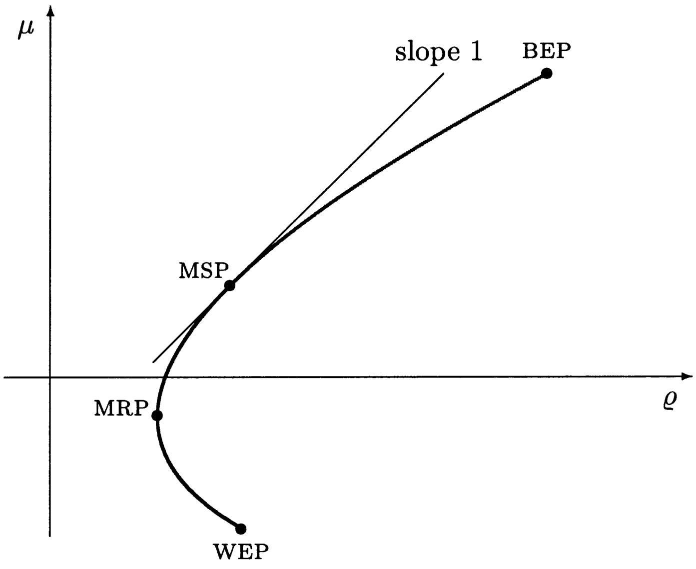

Fig. 1. The mean-risk analysis.

solution of $\min_{\mathbf{x} \in \mathscr{P}} \varrho(\mathbf{x})$, limits the curve to the mean-risk efficient frontier from BEP to MRP. Similar, the maximum safety portfolio (MSP), defined as the solution of $\max_{\mathbf{x} \in \mathscr{P}} [\mu(\mathbf{x}) - \varrho(\mathbf{x})]$, distinguishes a part of the mean-risk efficient frontier, from BEP to MSP, which is also mean-safety efficient. By virtue of Theorem 1, in the case of an SSD safety consistent risk measure, this part of the efficient frontier represents portfolios which are SSD efficient. Exactly, if a point (μ, ϱ) located at the efficient frontier between BEP and MSP is generated by the unique portfolio, then this portfolio is SSD efficient. In the case of multiple portfolios generating the same point (μ, ϱ), at least one of them is SSD efficient, but some other portfolios may be SSD dominated (by that efficient). The strong SSD consistency guarantees that all possible alternative portfolios are SSD efficient.

For specific types of return distributions or specific feasible sets, the subset of portfolios with guaranteed SSD efficiency may be larger (Ogryczak and Ruszczyński, 1999), exceeding the limit of the MSP. Hence, the mean-safety model (5) may be too restrictive in some practical investment decisions and it may be important to keep the full modeling capabilities of the original mean-risk approach (1). On the other hand, some mean-risk models, like the minimax (Young, 1998), were originally introduced in the form of corresponding mean-safety models (5). Therefore, the only way to provide a platform for the analysis of various Markowitz-type models is, in our opinion, to consider both mean-risk and mean-safety forms for all the measures.

An implementation of the Markowitz-type mean-risk model may take advantage of the efficient frontier convexity to perform the trade-off analysis. Having assumed a trade-off coefficient λ between the risk and the mean, the so-called *risk aversion coefficient*, one may directly compare real values $\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x})$ and find the best portfolio by solving the optimization problem:

$$\max\{\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x}): \quad \mathbf{x} \in \mathscr{P}\}. \tag{6}$$

Various positive values of parameter λ allow us to generate various efficient portfolios. By solving a parametric problem equation (6) with changing $\lambda > 0$ one gets the so-called *critical line approach* (Markowitz, 1959). Due to convexity of risk measures $\varrho(\mathbf{x})$ with respect to $\mathbf{x}$, $\lambda > 0$ provides a parameterization of the entire set of the μ/ϱ–efficient portfolios (except of its two ends BEP and MRP which are the limiting cases). Note that $(1 - \lambda)\mu(\mathbf{x}) + \lambda(\mu(\mathbf{x}) - \varrho(\mathbf{x})) = \mu(\mathbf{x}) - \lambda\varrho(\mathbf{x})$. Hence, bounded trade-off $0 < \lambda < 1$ in the Markowitz-type mean-risk model (1) corresponds to the complete weighting parameterization of the mean-safety model (5). This allows us to use Theorem 1 to derive the following SSD consistency results for the trade-off approach.

Corollary 1. If the risk measure $\varrho(\mathbf{x})$ is SSD safety consistent (3), then except for portfolios with identical values of $\mu(\mathbf{x})$ and $\varrho(\mathbf{x})$, every optimal solution of problem (6) with $0 < \lambda < 1$ is an SSD efficient portfolio. In the case of strong SSD safety consistency (4), every portfolio $\mathbf{x} \in \mathscr{P}$ optimal to (6) with $0 < \lambda < 1$ is, unconditionally, SSD efficient.

Recall that for specific types of return distributions or specific feasible sets the subset of portfolios with guaranteed SSD efficiency may be larger and this can easily be represented by the larger upper limit on λ. For instance, in the case of risk measures defined as the mean semideviation or the Gini's mean difference, this upper limit may be doubled still guaranteeing the SSD efficiency for solutions with $0 < \lambda < 2$, provided that the return distributions are symmetric with respect to their means (Ogryczak and Ruszczyński, 1999, 2002). Thus the trade-off model (6) offers a universal tool covering both the standard mean-risk and the corresponding mean-safety approaches. It provides easy modeling of the risk aversion and control of the SSD efficiency.

3. Dual stochastic dominance and shortfall risk measures

The notion of risk is related to a possible failure of achieving some targets and was formalized as the so-called shortfall criteria. The simplest shortfall criterion for a target value η is the probability of underachievement $\mathbb{P}\{R \leqslant \eta\}$, which is used in the FSD relation. Function $F^{(2)}$, used to define the SSD relation, can be presented as follows (Ogryczak and Ruszczyński, 1999):

$$F_R^{(2)}(\eta) = \mathbb{P}\{R \leqslant \eta\}\mathbb{E}\{\eta - R | R \leqslant \eta\} = \mathbb{E}\{\max\{\eta - R, 0\}\}.$$

Hence, the SSD relation is the Pareto dominance for mean below-target deviations from infinite number (continuum) of targets. The function $F_R^{(2)}$ is well-defined for any random variable R satisfying the condition $\mathbb{E}|R| < \infty$. As shown by Ogryczak and Ruszczyński (1999), it is a continuous, convex, non-negative, and non-decreasing function of η. The graph $F_{R(\mathbf{x})}^{(2)}(\eta)$, referred to as the Outcome–Risk (O–R) diagram, has two asymptotes which intersect at the point $(\mu(\mathbf{x}), 0)$. Specifically, the η-axis is the left asymptote and the line $\eta - \mu(\mathbf{x})$ is the right asymptote (Fig. 2). In the case of a deterministic (risk-free) return ($R(\mathbf{x}) = \mu(\mathbf{x})$), the graph of $F_{R(\mathbf{x})}^{(2)}(\eta)$ coincides with the asymptotes, whereas any uncertain return with the same expected value $\mu(\mathbf{x})$ yields a graph above (precisely, not below) the asymptotes. The space between the curve $(\eta, F_{R(\mathbf{x})}^{(2)}(\eta))$, and its asymptotes represents the dispersion (and thereby the riskiness) of $R(\mathbf{x})$ in comparison to the deterministic return $\mu(\mathbf{x})$. Therefore, it is called the *dispersion space*. While the variance $\sigma^2(\mathbf{x})$ represents the doubled area of the dispersion space, the mean semideviation (from the mean) $\bar{\delta}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R(\mathbf{x}), 0\}\} = F_{R(\mathbf{x})}^{(2)}(\mu(\mathbf{x}))$ turns out to be its largest vertical diameter. Therefore (Ogryczak and Ruszczyński, 1999), the mean semideviation is SSD safety consistent, i.e., $R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'')$ implies $\mu(\mathbf{x}') - \bar{\delta}(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \bar{\delta}(\mathbf{x}'')$. The mean semideviation is a half of the mean absolute deviation $\delta(\mathbf{x}) = \mathbb{E}\{|\mu(\mathbf{x}) - R(\mathbf{x})|\} = 2\bar{\delta}(\mathbf{x})$. Hence, the corresponding mean-risk model is equivalent to the MAD model (Konno and Yamazaki, 1991).

The mean below-target deviation from a specific target represents a single criterion of the SSD

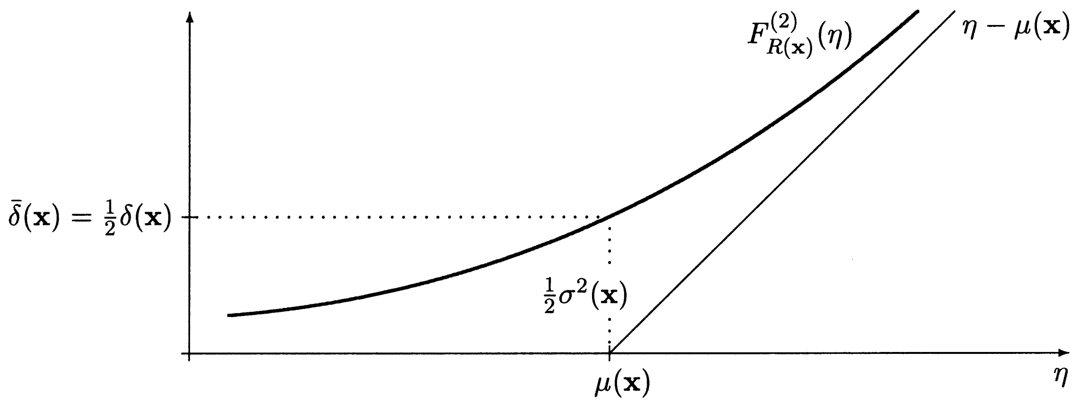

Fig. 2. The O–R diagram.

relation. Such a shortfall measure is very useful in investment situations with clearly defined minimum acceptable returns (e.g. bankruptcy level). However, in general cases, the quantile shortfall risk measures are more commonly used and accepted. In banking it was formalized with the Basle Committee on Banking Supervision recommending the measure of *value at risk* (VAR) defined as the maximum loss at a specified confidence level (c.f. Jorion, 1997; and references therein). VAR is now a widely used quantile measure, although the VAR calculations are usually based on the normal distribution assumption (Morgan, 1996). In this paper we do consider VAR measures for arbitrary distribution, following the original definition of the VAR as a high quantile (typically 95th or 99th percentile) of a distribution of losses.

Given $p \in [0, 1]$, the number $q_R(p)$ is called a *p-quantile* of the random variable R if

$$\mathbb{P}\{R < q_R(p)\} \leqslant p \leqslant \mathbb{P}\{R \leqslant q_R(p)\}.$$

For $p \in (0, 1)$ the set of such p-quantiles is a closed interval (Embrechts, Klüppelberg and Mikosch, 1997). Recall that value at risk depicts the worst (maximum) loss within a given confidence interval (Jorion, 1997). Hence, for a given portfolio $\mathbf{x}$ generating random returns $R(\mathbf{x})$ its value at risk for a confidence level c is a c-quantile of the random variable $-R(\mathbf{x})$, where the minus sign is due to counting losses while $R(\mathbf{x})$ represents the portfolio returns. In order to formalize the quantile measures, we introduce the first quantile function $F_R^{(-1)}$ corresponding to a real random variable R as the left-continuous inverse of the cumulative distribution function F_R:

$$F_R^{(-1)}(p) = \inf\{\eta : F_R(\eta) \geqslant p\} \quad \text{for } 0 < p \leqslant 1.$$

The value $F_R^{(-1)}(p)$ represents the p-quantile of R and, in the case of a non-unique quantile, it is the left end of the entire interval of quantiles. For a given portfolio $\mathbf{x}$, its VAR for a confidence level c may now be written as

$$\begin{aligned}\mathrm{VAR}_c(\mathbf{x}) &= \sup\{\xi : \mathbb{P}\{-R(\mathbf{x}) < \xi\} \leqslant c\}\\ &= -\inf\{\eta : \mathbb{P}\{R(\mathbf{x}) \leqslant \eta\} \geqslant 1 - c\} = -F_{R(\mathbf{x})}^{(-1)}(1 - c).\end{aligned} \tag{7}$$

The above formula defines the so-called absolute VAR related to the absolute loss. One may also consider the loss relative to the mean return (Jorion, 1997) which results in the following formula $\mathrm{RVAR}_c(\mathbf{x}) = \mu(\mathbf{x}) - F_{R(\mathbf{x})}^{(-1)}(1 - c)$. Note that $\mathrm{VAR}_c(\mathbf{x}) = \mathrm{RVAR}_c(\mathbf{x}) - \mu(\mathbf{x})$. In the case of a risk-free portfolio $R(\mathbf{x}) = \mu(\mathbf{x})$, for any confidence level $0 < c < 1$ one has $\mathrm{RVAR}_c(\mathbf{x}) = 0$ and $\mathrm{VAR}_c(\mathbf{x}) = -\mu(\mathbf{x})$. Hence, in terms of the Markowitz-type mean-risk analysis, we consider, the relative value at risk is the risk measure $\mathrm{RVAR}_c(\mathbf{x}) = \varrho(\mathbf{x})$ while the absolute VAR is the negative of the corresponding safety measure $\mathrm{VAR}_c(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(\mathbf{x}) = -s(\mathbf{x})$.

Directly from the definition of FSD we see that

$$R' \succeq_{FSD} R'' \Leftrightarrow F_{R'}^{(-1)}(p) \geqslant F_{R''}^{(-1)}(p) \quad \text{for all } 0 < p \leqslant 1.$$

Thus, the function $F^{(-1)}$ can be considered as a continuum-dimensional safety measure within the FSD. Using any specific (left) quantile as a scalar safety measure is consistent with the FSD. This justifies $\mathrm{VAR}_c(\mathbf{x})$ for any $0 \leqslant c < 1$, when minimized, as an FSD-consistent measure in the sense that $R(\mathbf{x}') \succeq_{FSD} R(\mathbf{x}'')$ implies $\mathrm{VAR}_c(\mathbf{x}') \leqslant \mathrm{VAR}_c(\mathbf{x}'')$. This does not guarantee, however, any consistency with the SSD, because it may happen that $R' \succeq_{SSD} R''$, but $F_{R'}^{(-1)}(p) < F_{R''}^{(-1)}(p)$ for some p. The lack of the SSD consistency we illustrate by two portfolios $\mathbf{x}'$ and $\mathbf{x}''$ with returns given as follows:

$$\mathbb{P}\{R(\mathbf{x}') = \xi\} = \begin{cases} 0.01, & \xi = -10 \\ 0.05, & \xi = -6 \\ 0.94, & \xi = 10 \\ 0, & \text{otherwise} \end{cases} \qquad \mathbb{P}\{R(\mathbf{x}'') = \xi\} = \begin{cases} 0.03, & \xi = -10 \\ 0.05, & \xi = -4 \\ 0.90, & \xi = 10 \\ 0.02, & \xi = 25 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Note that $\mu(\mathbf{x}') = \mu(\mathbf{x}'') = 9$ and $\sigma(\mathbf{x}') < \sigma(\mathbf{x}'')$ as well as $\delta(\mathbf{x}') < \delta(\mathbf{x}'')$. Actually, $R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$ and in all preference models based on the risk-aversion axioms portfolio $\mathbf{x}'$ is preferred to $\mathbf{x}''$. On the other hand, when calculating the VAR measures for the 95% confidence level one gets $\text{VAR}_{.95}(\mathbf{x}') = 6$ and $\text{RVAR}_{.95}(\mathbf{x}') = 15$ while $\text{VAR}_{.95}(\mathbf{x}'') = 4$ and $\text{RVAR}_{.95}(\mathbf{x}'') = 13$. Thus, portfolio $\mathbf{x}'$ is considered as more risky with respect to both the absolute and the relative VAR measures. There are also known examples showing directly that increasing diversification of a portfolio results in the larger value of the VAR measure (Artzner et al., 1999).

To obtain quantile measures consistent with the SSD we introduce the second quantile function defined for a random variable R as:

$$F_R^{(-2)}(p) = \int_0^p F_R^{(-1)}(\alpha)d\alpha \quad \text{for } 0 < p \leqslant 1 \quad \text{and} \quad F_R^{(-2)}(0) = 0. \tag{9}$$

Similarly to $F_R^{(2)}$, the function $F_R^{(-2)}$ is well-defined for any random variable R satisfying the condition $\mathbb{E}\,|R| < \infty$. By construction, it is convex. The graph of $F_R^{(-2)}$ is called the absolute Lorenz curve or ALC diagram for short. The pointwise comparison of the second quantile functions defines the so-called absolute (or general) Lorenz order (Shorrocks, 1983).

Recently, an intriguing duality relation between the second quantile function $F_R^{(-2)}$ and the second performance function $F_R^{(2)}$ has been shown (Ogryczak and Ruszczyński, 2002). Namely, function $F_R^{(-2)}$ is a conjugent (dual function) (Rockafellar, 1970) of $F_R^{(2)}$, i.e., for every $p \in [0, 1]$,

$$F_R^{(-2)}(p) = \sup_\eta \{\eta p - F_R^{(2)}(\eta)\}. \tag{10}$$

It follows from the duality theory (Rockafellar, 1970) that we may fully characterize the SSD relation by using the conjugate function $F^{(-2)}$:

$$R' \succeq_{SSD} R'' \Leftrightarrow F_{R'}^{(-2)}(p) \geqslant F_{R''}^{(-2)}(p) \quad \text{for all } 0 \leqslant p \leqslant 1. \tag{11}$$

In other words, the absolute Lorenz order is equivalent to the SSD order.

By the properties of the O–R diagram and equation (10), for any portfolio $\mathbf{x}$ one gets $F_{R(\mathbf{x})}^{(-2)}(1) = \mu(\mathbf{x})$. Thus the ALC diagram is a continuous convex curve connecting points (0, 0) and (1, $\mu(\mathbf{x})$), whereas a deterministic outcome with the same expected value $\mu(\mathbf{x})$, yields the chord (straight line) connecting the same points. Hence, the space between the curve $(p, F_{R(\mathbf{x})}^{(-2)}(p))$, $0 \leqslant p \leqslant 1$, and its chord represents the dispersion (and thereby the riskiness) of $R(\mathbf{x})$ in comparison to the deterministic outcome of $\mu(\mathbf{x})$ (Fig. 3). We shall call it the dual dispersion space. Both size and shape of the dual dispersion space are important for complete description of the riskiness of a portfolio. Nevertheless, it is quite natural to consider some size parameters as summary characteristics of riskiness.

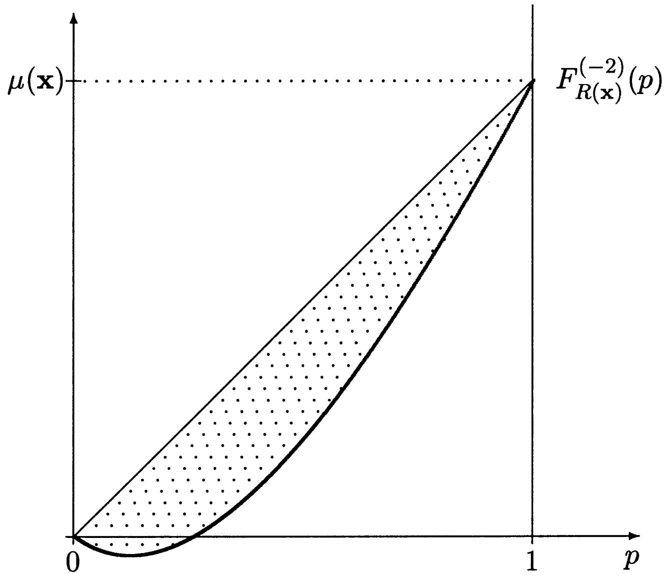

Fig. 3. The absolute Lorenz curve and the dual dispersion space.

Let us analyze the vertical diameter of the dual dispersion space defined as:

$$d_\beta(\mathbf{x}) = \beta\mu(\mathbf{x}) - F^{(-2)}_{R(\mathbf{x})}(\beta) \quad \text{for } 0 \leqslant \beta \leqslant 1. \tag{12}$$

Certainly, for any portfolio $\mathbf{x}$, $d_0(\mathbf{x}) = d_1(\mathbf{x}) = 0$. For every $\beta \in (0, 1)$, it follows from (10) that

$$d_\beta(\mathbf{x}) = \min_\xi(\beta(\mu(\mathbf{x}) - \xi) + F^{(2)}_{R(\mathbf{x})}(\xi)) = \min_\xi(\beta\mathbb{E}\{R(\mathbf{x}) - \xi\} + \mathbb{E}\{\max\{0, \xi - R(\mathbf{x})\}\})$$

$$= \min_\xi(\beta\mathbb{E}\{\max(0, R(\mathbf{x}) - \xi)\} + (1-\beta)\mathbb{E}\{\max(0, \xi - R(\mathbf{x}))\}).$$

Hence,

$$d_\beta(\mathbf{x}) = \min_\xi \mathbb{E}\{\max(\beta(R(\mathbf{x}) - \xi), (1-\beta)(\xi - R(\mathbf{x})))\} \tag{13}$$

and the minimum in the expression above is attained at any β-quantile (Ogryczak and Ruszczyński, 2002). Note that, for $\beta = 0.5$, the diameter $d_{0.5}(\mathbf{x})$ is a half of the mean absolute deviation from the median, the MAD-type risk measure recommended by Sharpe (1971b).

For a direct interpretation of the vertical diameters $d_\beta(\mathbf{x})$ we can provide another, scaled representation of the second quantile function and the diameter itself (Fig. 4). For any real parameter $0 < \beta \leqslant 1$, function $F^{(-2)}$ allows to define the worst conditional expectation M_β and the worst conditional semideviation Δ_β as follows:

$$M_\beta(\mathbf{x}) = \frac{1}{\beta}F^{(-2)}_{R(\mathbf{x})}(\beta) \quad \text{and} \quad \Delta_\beta(\mathbf{x}) = \frac{1}{\beta}d_\beta(\mathbf{x}) = \mu(\mathbf{x}) - M_\beta(\mathbf{x}) \quad \text{for } 0 < \beta \leqslant 1. \tag{14}$$

Note that in the case of a risk-free portfolio $R(\mathbf{x}) = \mu(\mathbf{x})$, for any value $0 < \beta \leqslant 1$ one has

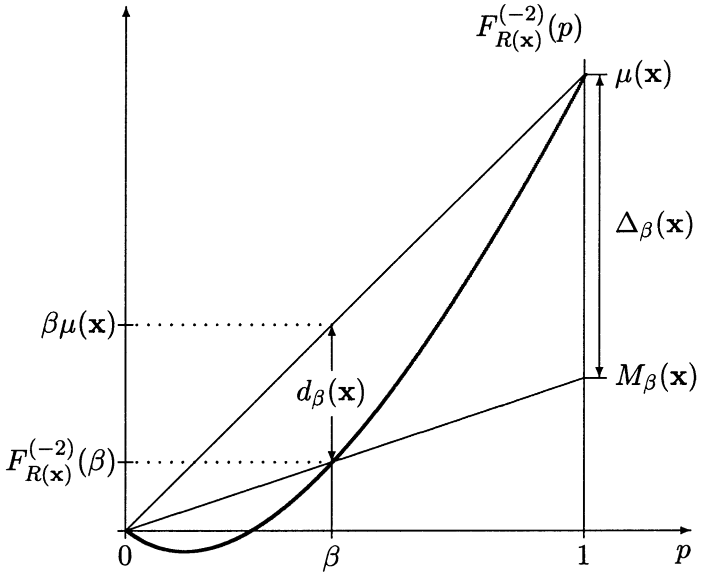

Fig. 4. The absolute Lorenz curve and the conditional worst realizations.

$M_\beta(\mathbf{x}) = \mu(\mathbf{x})$ and $\Delta_\beta(\mathbf{x}) = 0$. Hence, in terms of the Markowitz-type mean-risk analysis, we consider, the worst conditional semideviation is the risk measure $\Delta_\beta(\mathbf{x}) = \varrho(\mathbf{x})$, while the worst conditional expectation itself is the corresponding safety measure $M_\beta(\mathbf{x}) = \mu(\mathbf{x}) - \varrho(\mathbf{x}) = s(\mathbf{x})$. By the convexity of $F^{(-2)}$, the worst conditional expectation $M_\beta(\mathbf{x})$ is non-decreasing and continuous function of β. Note that $M_1(\mathbf{x}) = \mu(\mathbf{x})$ and $\Delta_1(\mathbf{x}) = 0$. In the case of a random variable with lower bounded support, the value of $M_\beta(\mathbf{x})$ tends to the minimum outcome $M(\mathbf{x})$ when β approaches 0 ($\beta \to 0_+$). Hence, the minimax portfolio selection rule of Young (1998) is a limiting case of the mean-risk model using the worst conditional expectation.

The relation (11) can be rewritten in the form

$$R' \succeq_{SSD} R'' \Leftrightarrow F^{(-2)}_{R'}(p)/p \geqslant F^{(-2)}_{R''}(p)/p \quad \text{for all } 0 < p \leqslant 1, \tag{15}$$

thus justifying the SSD consistency of the worst conditional expectation $M_\beta(\mathbf{x})$ as a safety measure.

Proposition 1. *For any $0 < \beta \leqslant 1$, the worst conditional semideviation $\Delta_\beta(\mathbf{x})$ is an SSD safety consistent risk measure, in the sense that*

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow M_\beta(\mathbf{x}') = \mu(\mathbf{x}') - \Delta_\beta(\mathbf{x}') \geqslant M_\beta(\mathbf{x}'') = \mu(\mathbf{x}'') - \Delta_\beta(\mathbf{x}''). \tag{16}$$

Let $p \in (0, 1)$ and suppose that η is such that $\mathbb{P}\{R(\mathbf{x}) \leqslant \eta\} = p$. Then, by the theory of conjugate functions (Rockafellar, 1970, Thm.23.5), the equality $F^{(-2)}_{R(\mathbf{x})}(p) + F^{(2)}_{R(\mathbf{x})}(\eta) = p\eta$ holds.

Hence, in the case of $\mathbb{P}\{R(\mathbf{x}) \leqslant -\mathrm{VAR}_c(\mathbf{x})\} = 1 - c$, one gets

$$M_{(1-c)}(\mathbf{x}) = -\mathrm{VAR}_c(\mathbf{x}) - \frac{1}{1-c} F^{(2)}_{R(\mathbf{x})}(-\mathrm{VAR}_c(\mathbf{x}))$$
$$= -\mathrm{VAR}_c(\mathbf{x}) + \mathbb{E}\{R(\mathbf{x}) + \mathrm{VAR}_c(\mathbf{x}) | R(\mathbf{x}) \leqslant -\mathrm{VAR}_c(\mathbf{x})\}$$
$$= \mathbb{E}\{R(\mathbf{x}) | R(\mathbf{x}) \leqslant -\mathrm{VAR}_c(\mathbf{x})\}.$$

Hence, the quantity $M_{(1-c)}(\mathbf{x})$ may be interpreted as the negative to the expected loss size, given that the losses on the level of $\mathrm{VAR}_c(\mathbf{x})$ or greater are considered. The latter is called the expected shortfall $\mathrm{ES}_c(\mathbf{x}) = \mathbb{E}\{-R(\mathbf{x}) | -R(\mathbf{x}) \geqslant \mathrm{VAR}_c(\mathbf{x})\}$ (Embrechts et al., 1997), the Tail Conditional Expectation (Tail VAR) (Artzner et al., 1999) or the conditional value at risk (CVAR) (Andersson et al., 2001). Similar characteristics called absolute concentration curves were earlier considered by Shalit and Yitzhaki (1994). They compared absolute concentration curves for individual securities R_j in their analysis if there are securities whose share will be increased by every risk-averse investor. The relation

$$M_{(1-c)}(\mathbf{x}) = -\mathbb{E}\{-R(\mathbf{x}) | -R(\mathbf{x}) \geqslant \mathrm{VAR}_c(\mathbf{x})\} = -\mathrm{ES}_c(\mathbf{x}) \tag{17}$$

facilitates the understanding of the measure of the worst conditional expectation (Artzner et al., 1999) and the nature of the second quantile function (Shalit and Yitzhaki, 1994). Although valid for many continuous distributions (Andersson et al., 2001) equation (17), in general, cannot serve as a definition of the worst conditional expectation because η such that $\mathbb{P}\{R(\mathbf{x}) \leqslant \eta\} = p$ need not exist.

In general, $\mathbb{P}\{R(\mathbf{x}) \leqslant -\mathrm{VAR}_c(\mathbf{x})\} = \beta' \geqslant \beta = 1 - c$ and $M_\beta(\mathbf{x}) \leqslant M_{\beta'}(\mathbf{x}) = \mathbb{E}\{R(\mathbf{x}) | R(\mathbf{x}) \leqslant -\mathrm{VAR}_c(\mathbf{x})\}$. Moreover, while according to Proposition 1 the worst conditional expectation is always SSD consistent, this is not true for the expected shortfall (or equivalent measures). We illustrate this by two portfolios $\mathbf{x}'$ and $\mathbf{x}''$ with returns given in (8). Recall that $\mu(\mathbf{x}') = \mu(\mathbf{x}'') = 9$ and $R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$. When calculating the VAR measures for the 95% confidence level we got $\mathrm{VAR}_{.95}(\mathbf{x}') = 6$ and $\mathrm{VAR}_{.95}(\mathbf{x}'') = 4$. Now, computing the worst conditional expectations we get

$$M_{.05}(\mathbf{x}') = (-10 \cdot 0.01 - 6 \cdot 0.04)/0.05 = -6.8 > M_{.05}(\mathbf{x}'') = (-10 \cdot 0.03 - 4 \cdot 0.02)/0.05 = -7.6,$$

thus preferring portfolio $\mathbf{x}'$ as less risky. On the other hand, when calculating the expected shortfalls for the 95% confidence level, one gets

$$\mathrm{ES}_{.95}(\mathbf{x}') = (10 \cdot 0.01 + 6 \cdot 0.05)/0.06 = 40/6 > \mathrm{ES}_{.95}(\mathbf{x}'') = (10 \cdot 0.03 + 4 \cdot 0.05)/0.08 = 6.26.$$

Thus, despite $R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'')$, the portfolio $\mathbf{x}'$ would be considered as more risky with respect to the expected shortfall measures.

Dual risk characteristics can also be presented in the (primal) O–R diagram (Fig. 5). Recall that $F^{(-2)}$ is the conjugate function of $F^{(2)}$ and, therefore, $F^{(-2)}$ describes the affine functions majorized by $F^{(2)}$ (Rockafellar, 1970). For any $\beta \in (0, 1)$, the line with slope β supports the graph of $F^{(2)}$ at every β-quantile thus at $-\mathrm{VAR}_c(\mathbf{x})$ for $c = 1 - \beta$. The line is given analytically as $S^c_{R(\mathbf{x})}(\eta) = (1 - c)(\eta + \mathrm{VAR}_c(\mathbf{x})) + F^{(2)}_{R(\mathbf{x})}(-\mathrm{VAR}_c(\mathbf{x}))$. The value of the absolute Lorenz curve is given by the intersection of the tangent line with the vertical (risk) axis, $F^{(-2)}_{R(\mathbf{x})}(\beta) = -S^{(1-\beta)}_{R(\mathbf{x})}(0)$. For any $\beta \in (0, 1)$, the tangent line intersects both asymptotes of $F^{(2)}$. It intersects the outcome axis (the left asymptote) at the point of the corresponding worst conditional expectation: $\eta = F^{(-2)}_{R(\mathbf{x})}(\beta)/\beta = M_\beta(\mathbf{x})$.

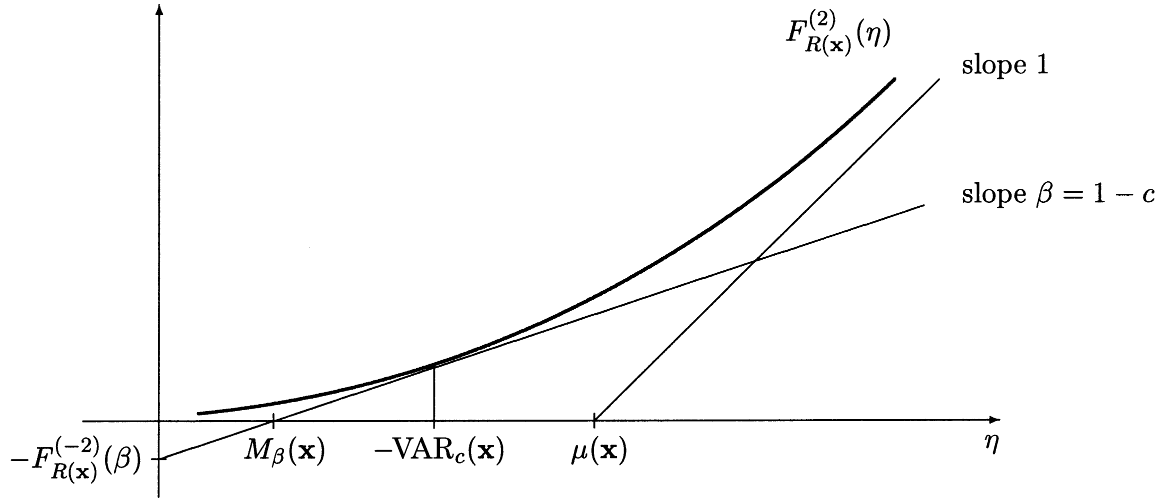

Fig. 5. Quantile safety measures in the O–R diagram.

4. LP computability

While the original Markowitz model with risk measured by the variance forms a quadratic programming problem, following Sharpe (1971a), many attempts have been made to linearize the portfolio optimization procedure (c.f., Speranza, 1993, and references therein). Certainly, to model advantages of a diversification, risk measures cannot be a linear function of $\mathbf{x}$. Nevertheless, the risk measure can be LP-computable in the case of discrete random variables, i.e., in the case of returns defined by their realizations under the specified scenarios. We will consider T scenarios with probabilities p_t (where $t = 1, \ldots, T$). We will assume that for each random variable R_j there is known its realization r_{jt} under the scenario t. Typically, the realizations are derived from historical data treating T historical periods as equally probable scenarios ($p_t = 1/t$). The realizations of the portfolio returns $R(\mathbf{x})$ are given as $y_t = \sum_{j=1}^n r_{jt}x_j$ and the expected value can be computed as $\mu(\mathbf{x}) = \sum_{t=1}^T y_t p_t = \sum_{t=1}^T [p_t \sum_{j=1}^n r_{jt}x_j]$.

Several risk measures can be LP-computable with respect to the realizations y_t. Konno and Yamazaki (1991) presented and analyzed the complete portfolio optimization model (MAD model) based on the risk measure defined as the mean absolute deviation from the mean $\delta(\mathbf{x}) = \mathbb{E}\{|R(\mathbf{x}) - \mu(\mathbf{x})|\}$. For a discrete random variable represented by its realizations y_t, the mean absolute deviation, when minimized, is LP-computable as

$$\delta(\mathbf{x}) = \min \sum_{t=1}^T (d_t^- + d_t^+) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \mu(\mathbf{x}) - y_t,\ d_t^-,\ d_t^+ \geqslant 0 \quad \text{for } t = 1, \ldots, T \tag{18}$$

where d_t^- and d_t^+ are nonnegative variables introduced to represent, respectively, the downside and the upside deviation from the mean return under the scenario t. This can be simplified by taking advantages of the symmetry $\delta(\mathbf{x}) = 2\bar{\delta}(\mathbf{x})$ as

$$\bar{\delta}(\mathbf{x}) = \min \sum_{t=1}^{T} d_t^- p_t \quad \text{s.t.} \quad d_t^- \geqslant \mu(\mathbf{x}) - y_t,\ d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T.$$

Yitzhaki (1982) introduced the mean-risk model using Gini's mean (absolute) difference as the risk measure. For a discrete random variable represented by its realizations y_t, the Gini's mean difference

$$\Gamma(\mathbf{x}) = \frac{1}{2} \sum_{t'=1}^{T} \sum_{t''=1}^{T} |y_{t'} - y_{t''}| p_{t'} p_{t''} \tag{19}$$

is obviously LP computable (when minimized).

For a discrete random variable represented by its realizations y_t, the worst realization

$$M(\mathbf{x}) = \min_{t=1,\ldots,T} y_t \tag{20}$$

is a well-appealing safety measure, while the *maximum (downside) semideviation*

$$\Delta(\mathbf{x}) = \mu(\mathbf{x}) - M(\mathbf{x}) = \max_{t=1,\ldots,T} (\mu(\mathbf{x}) - y_t) \tag{21}$$

represents the corresponding (dispersion) risk measure. The latter is well defined in the O–R diagram (Fig. 2) as it represents the maximum horizontal diameter of the dispersion space. The measure $M(\mathbf{x})$ was applied to portfolio optimization by Young (1998).

Actually, all the classical LP-computable risk measures are well-defined size characteristics of the dual dispersion space (Fig. 6). The mean semideviation $\bar{\delta}(\mathbf{x})$ turns out to be the maximal vertical diameter of the dual dispersion space (Ogryczak and Ruszczyński, 2002). It is commonly known that the Gini's mean difference $\Gamma(\mathbf{x})$ represents the doubled area of the space or its average vertical

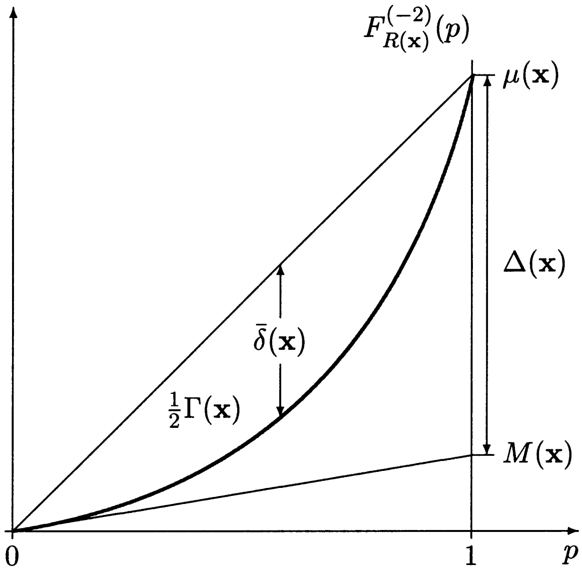

Fig. 6. The absolute Lorenz curve and LP-computable risk measures.

diameter. The maximum semideviation $\Delta(\mathbf{x})$ measures the external envelope of the dispersion space (Ogryczak, 2000). However, both the mean semideviation and the maximum semideviation are rather rough measures when comparing to the Gini's mean difference. The latter is SSD safety consistent (3), as shown already by Yitzhaki (1982), but it satisfies also the requirements (4) of the strong SSD safety consistency (Ogryczak and Ruszczyński, 2002).

Proposition 2. *The Gini's mean difference* $\Gamma(\mathbf{x})$ *is a strongly SSD safety consistent risk measure, in the sense that both following implications are valid:*

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow \mu(\mathbf{x}') - \Gamma(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \Gamma(\mathbf{x}'')$$
$$R(\mathbf{x}') \succ_{SSD} R(\mathbf{x}'') \Rightarrow \mu(\mathbf{x}') - \Gamma(\mathbf{x}') > \mu(\mathbf{x}'') - \Gamma(\mathbf{x}'')$$

The SSD safety consistency was also shown for the maximum semideviation (Ogryczak, 2000) and for the mean semideviation (Ogryczak and Ruszczyński, 1999, 2002) since both these measures are related to the dual dispersion space.

Proposition 3. *The maximum semideviation* $\Delta(\mathbf{x})$ *is an SSD safety consistent risk measure, in the sense that*

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow M(\mathbf{x}') = \mu(\mathbf{x}') - \Delta(\mathbf{x}') \geqslant M(\mathbf{x}'') = \mu(\mathbf{x}'') - \Delta(\mathbf{x}'').$$

Proposition 4. *The mean semideviation* $\bar{\delta}(\mathbf{x})$ *is an SSD safety consistent risk measure.*

The dual approach of the absolute Lorenz order, similar to the primal approach of the SSD relation, is based on a continuum-dimensional risk measurement. However, in the case of (discrete) random variables defined by their realizations for T equally probable ($p_t = 1/T$) scenarios (historical data), the dual approach takes a form of multiple criteria optimization with only T criteria $F^{(-2)}(k/T)$, $k = 1, \ldots, T$ (Ogryczak, 2000). The corresponding worst conditional expectations $M_{k/T}(\mathbf{x})$ express then the mean returns under the k worst scenarios and represent a natural generalization of the (absolute) worst realization $M(\mathbf{x})$.

Due to the duality relation (10), the absolute Lorenz curve is defined by optimization (Ogryczak and Ruszczyński, 2002):

$$F^{(-2)}_{R(\mathbf{x})}(\beta) = \max_{\eta}[\beta\eta - F^{(2)}_{R(\mathbf{x})}(\eta)] = \max_{\eta}[\beta\eta - \mathbb{E}\{\max\{\eta - R(\mathbf{x}), 0\}\}] \tag{22}$$

where η is a real variable taking the value of β–quantile at the optimum. Hence, the absolute Lorenz curve and the corresponding risk measures express the results of the O–R diagram analysis according to a slant direction (Fig. 5 above). For a discrete random variable represented by its realizations y_t, (22) becomes an LP problem. Thus the worst conditional expectation is LP-computable as:

$$M_\beta(\mathbf{x}) = \max\left[\eta - \frac{1}{\beta}\sum_{t=1}^{T} d_t^- p_t\right] \quad \text{s.t.} \quad d_t^- \geqslant \eta - y_t,\ d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T \tag{23}$$

where η is an auxiliary (unbounded) variable. The worst conditional semideviations may be computed as the corresponding differences from the mean ($\Delta_\beta(\mathbf{x}) = \mu(\mathbf{x}) - M_\beta(\mathbf{x})$) or directly as:

$$\Delta_\beta(\mathbf{x}) = \min \sum_{t=1}^{T} \left(d_t^+ + \frac{1-\beta}{\beta} d_t^- \right) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \eta - y_t,\ d_t^+,\ d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T.$$

Note that for $\beta = 0.5$ one has $1 - \beta = \beta$. Hence, the LP problem for computing the mean absolute deviation from the median

$$\Delta_{.5}(\mathbf{x}) = \min \sum_{t=1}^{T} (d_t^- + d_t^+) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \eta - y_t, \quad d_t^+,\ d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T$$

differs from that for the mean absolute deviation from the mean (18) only by a single variable η replacing the $\mu(\mathbf{x})$.

The risk measures we considered, although all derived from the SSD shortfall criteria, are quite different in modeling of the downside risk aversion. Definitely the strongest with this respect is the maximum semideviation $\Delta(\mathbf{x})$. It is the strict worst-case measure where only the worst scenario is taken into account. The measure of worst conditional semideviation $\Delta_\beta(\mathbf{x})$ allows us to extend the approach to a specified β quantile of the worst returns, which results in a continuum of models evolving from the strongest downside risk-aversion (β close to 0) to the complete risk neutrality ($\beta = 1$). These measures may be further extended (or rather combined) to enhance the risk-aversion modeling capabilities. We analyze a combination of risk measures by the weighted sum which allows us to generate various mixed measures.

Consider a set, say m, risk measures $\varrho_k(\mathbf{x})$ and their linear combination:

$$\varrho_{\mathbf{w}}^{(m)}(\mathbf{x}) = \sum_{k=1}^{m} w_k \varrho_k(\mathbf{x}), \qquad \sum_{k=1}^{m} w_k \leqslant 1, \quad w_k \geqslant 0 \quad \text{for } k = 1, \ldots, m \tag{24}$$

Note that

$$\mu(\mathbf{x}) - \varrho_{\mathbf{w}}^{(m)}(\mathbf{x}) = w_0 \mu(\mathbf{x}) + \sum_{k=1}^{m} w_k (\mu(\mathbf{x}) - \varrho_k(\mathbf{x}))$$

where $w_0 = 1 - \sum_{k=1}^{m} w_k \geqslant 0$. Hence, the following assertion is valid.

Theorem 2. *If all risk measures ϱ_k are SSD safety consistent, then every combined risk measure (24) is also SSD safety consistent in the sense that*

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \Rightarrow \mu(\mathbf{x}') - \varrho_{\mathbf{w}}^{(m)}(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \varrho_{\mathbf{w}}^{(m)}(\mathbf{x}'')$$

Moreover, if at least one measure ϱ_{k_o} is strongly SSD safety consistent, then every combined risk measure (24) with a strictly positive corresponding weight ($w_{k_o} > 0$) is strongly SSD safety consistent.

Theorem 2 allows us to combine various risk measures, preserving their SSD consistency properties. In particular, one may consider several, say m, levels $0 < \beta_1 < \beta_2 < \ldots < \beta_m \leqslant 1$ and use weighted sum of the worst conditional semideviations $\Delta_{\beta_k}(\mathbf{x})$ to define a new risk measure, the weighted conditional semideviation:

$$\Delta_{\mathbf{w}}^{(m)}(\mathbf{x}) = \sum_{k=1}^{m} w_k \Delta_{\beta_k}(\mathbf{x}), \qquad \sum_{k=1}^{m} w_k = 1, \quad w_k > 0 \quad \text{for } k = 1, \ldots, m \tag{25}$$

with the corresponding safety measure of the weighted conditional expectation:

$$M_{\mathbf{w}}^{(m)}(\mathbf{x}) = \mu(\mathbf{x}) - \Delta_{\mathbf{w}}^{(m)}(\mathbf{x}) = \sum_{k=1}^{m} w_k M_{\beta_k}(\mathbf{x}). \tag{26}$$

As linear non-negative combinations of LP-computable risk measures, the weighted conditional measures remain LP computable. Certainly, the auxiliary constraints and variables (23) must be used for each level β_k thus resulting in an m times larger LP formulation. As mentioned, by virtue of Theorem 2, the following assertion is valid.

Proposition 5. *For any set of levels* $0 < \beta_1 < \beta_2 < \ldots < \beta_m \leqslant 1$*, the weighted conditional semideviations* $\Delta_{\mathbf{w}}^{(m)}(\mathbf{x})$ *(25) is an SSD safety consistent risk measure, in the sense that*

$$R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'') \Rightarrow M_{\mathbf{w}}^{(m)}(\mathbf{x}') = \mu(\mathbf{x}') - \Delta_{\mathbf{w}}^{(m)}(\mathbf{x}') \geqslant M_{\mathbf{w}}^{(m)}(\mathbf{x}'') = \mu(\mathbf{x}'') - \Delta_{\mathbf{w}}^{(m)}(\mathbf{x}'').$$

In the case of equally probable T scenarios with $p_t = 1/T$ (historical data for T periods), the safety measure $M_{\mathbf{w}}^{(T)}(\mathbf{x})$ defined with $m = T$ levels $\beta_k = k/T$ for $k = 1, 2, \ldots, T$ represents the standard weighting approach to the multiple criteria LP portfolio optimization model with criteria $F^{(-2)}(k/T)$ (Ogryczak, 2000). The use of weights $w_k = (2k)/T^2$ for $k = 1, 2, \ldots, T-1$ and $w_T = 1/T$ while defining $\Delta_{\mathbf{w}}^{(T)}(\mathbf{x})$ results then in the Gini's mean difference as

$$\Delta_{\mathbf{w}}^{(T)}(\mathbf{x}) = \sum_{k=1}^{T-1} \frac{2k}{T^2} \Delta_{k/T}(\mathbf{x}) = \frac{2}{T} \sum_{k=1}^{T-1} d_{k/T}(\mathbf{x}) = \Gamma(\mathbf{x}).$$

Moreover, for any set of strictly positive weights $w_k > 0$, the weighted conditional deviation (25) can be rewritten as

$$\Delta_{\mathbf{w}}^{(T)}(\mathbf{x}) = \sum_{k=1}^{T} w_k \Delta_{k/T}(\mathbf{x}) = \overline{w}_0 \Gamma(\mathbf{x}) + \sum_{k=1}^{T} \overline{w}_k \Delta_{k/T}(\mathbf{x})$$

where $\overline{w}_k > 0$ for $k = 0, 1, \ldots, T$ and $\sum_{k=0}^{T} \overline{w}_k = 1$. Hence, by virtue of Theorem 2, the weighted conditional deviation is then strongly SSD safety consistent.

In order to model downside risk aversion one may consider, instead of the Gini's mean difference, the tail Gini's measure defined for any $\beta \in (0, 1]$ by averaging the vertical diameters $d_p(\mathbf{x})$ within the tail interval $p \leqslant \beta$ as:

$$G_\beta(\mathbf{x}) = \frac{2}{\beta^2} \int_0^\beta (\mu(\mathbf{x})\alpha - F_{R(\mathbf{x})}^{(-2)}(\alpha)) d\alpha. \tag{27}$$

The measure is an area characteristic of the dual dispersion space but, similar to the worst conditional semideviation, it is restricted to the tail performances within the β-quantile. A simple analysis of the ALC diagram leads to the following assertion (Ogryczak and Ruszczyński, 2002).

Proposition 6. *For any* $0 < \beta \leqslant 1$, *the tail Gini's measure* $G_\beta(\mathbf{x})$ *is an SSD safety consistent risk measure, in the sense that* $R(\mathbf{x}') \succeq_{SSD} R(\mathbf{x}'')$ *implies* $\mu(\mathbf{x}') - G_\beta(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - G_\beta(\mathbf{x}'')$.

Again, in the simplest case of equally probable T scenarios with $p_t = 1/T$ (historical data for T periods), the tail Gini's measure for $\beta = \kappa/T$ may be expressed as the weighted conditional semideviation $\Delta_{\mathbf{w}}^{(\kappa)}(\mathbf{x})$ with levels $\beta_k = k/T$ for $k = 1, 2, \ldots, \kappa$ and properly defined weights. Exactly, while using the weights $w_k = (2k)/\kappa^2$ for $k = 1, 2, \ldots, \kappa - 1$ and $w_\kappa = 1/\kappa$, one gets

$$\Delta_{\mathbf{w}}^{(\kappa)}(\mathbf{x}) = \sum_{k=1}^{\kappa-1} \frac{2k}{\kappa^2} \Delta_{k/T}(\mathbf{x}) + \frac{\kappa}{\kappa^2} \Delta_{\kappa/T}(\mathbf{x}) = \frac{1}{\beta^2 T} \left[2 \sum_{k=1}^{\kappa-1} d_{k/T}(\mathbf{x}) + d_{\kappa/T}(\mathbf{x}) \right] = G_\beta(\mathbf{x}).$$

In a general case of T scenarios with arbitrary probabilities p_t, a reduction of the tail Gini's measure to the LP computational case is harder. One possibility is to introduce such a grid of levels β_k that contains all possible break points of the Lorenz curves, but it may be unnecessarily numerous. Another possibility is to resort to an approximation with $\Delta_{\mathbf{w}}^{(m)}(\mathbf{x})$ based on some reasonably-chosen grid β_k, $k = 1, \ldots, m$. For any $0 < \beta \leqslant 1$, while using levels $\beta_k = (k\beta)/m$ for $k = 1, 2, \ldots, m$ and weights defined as $w_k = (2k)/m^2$ for $k = 1, 2, \ldots, m - 1$ and $w_m = 1/m$, one gets the weighted conditional semideviation $\Delta_{\mathbf{w}}^{(m)}(\mathbf{x})$, expressing the trapezoidal approximation to the tail Gini's measure $G_\beta(\mathbf{x})$. It must be emphasized that despite being only an approximation to (27), quantity $\Delta_{\mathbf{w}}^{(m)}(\mathbf{x})$ itself is a well-defined LP-computable risk measure with guaranteed SSD safety consistency (Proposition 5).

5. Concluding remarks

VAR, defined as the maximum loss at a specified confidence level, is a widely used and commonly accepted quantile risk measure. In banking it was formalized with the recommendations of the Basle Committee on Banking Supervision. Although the VAR calculations are usually based on the normal distribution assumption (Morgan, 1996), in this paper we have considered VAR measures for arbitrary distribution, following the original definition of the VAR as a high quantile of a distribution of losses. We have shown that the minimization of VAR (for any confidence level) is consistent with the first degree stochastic dominance (FSD). This does not guarantee, however, any consistency with the SSD.

The SSD relation is crucial for decision-making under risk. An SSD-dominating portfolio is preferred within all risk-averse preference models where larger outcomes are preferred. The SSD relation covers increasing and concave utility functions, while the first stochastic dominance is less specific as it covers all increasing utility functions, thus neglecting a risk-averse attitude. It is therefore a matter of primary importance that a risk measure be consistent with the SSD relation. This can be achieved with the second-order quantile measure, called the worst conditional expectation or Tail VAR, representing the mean shortfall at a specified confidence level. By exploiting duality relations of convex analysis to develop the quantile model of the stochastic dominance, we have shown that the worst conditional expectation is SSD-consistent for general distributions. Moreover, it has attractive computational properties as it leads to LP-solvable portfolio optimization models. The introduced methodology has enabled the development of Tail Gini's measures for more precise modeling of the downside risk aversion.

Theoretical properties, although crucial for understanding the modeling concepts, provide only a very limited background for comparison of the final optimization models. Recently-released initial experiments show that the second-order quantile risk measures may effectively help to reduce credit risk of the portfolio of bounds (Andersson et al., 2001) as well as to build well-diversified and well-performing portfolios of stocks (see Mansini et al., 2001 for the Milano Stock Exchange results). It shows a need for further comprehensive experimental studies analyzing practical performances of the higher-order quantile risk measures within specific areas of financial applications.

Acknowledgements

Research conducted by W. Ogryczak was supported by the grant PBZ-KBN-016/P03/99 from The State Committee for Scientific Research (Poland).

References

Andersson, F., Mausser, H., Rosen, D., Uryasev, S., 2001. Credit risk optimization with Conditional Value at Risk criterion. *Mathematical Programming* 89, 273–291.
Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance* 9, 203–228.
Baumol, W.J., 1964. An expected gain-confidence limit criterion for portfolio selection. *Management Science* 10, 174–182.
Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
Fishburn, P.C., 1980. Stochastic dominance and moments of distributions. *Mathematics of Operations Research* 5, 94–100.
Jorion, P., 1997. *Value at Risk: The New Benchmark for Controlling Market Risk*. McGraw-Hill, New York.
Konno, H., Yamazaki, H., 1991. Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science* 37, 519–531.
Levy, H., 1992. Stochastic dominance and expected utility: survey and analysis. *Management Science* 38, 555–593.
Mansini, R., Ogryczak, W., Speranza, M.G., 2001. LP Solvable Models for Portfolio Optimization: A Classification and Computational Comparison. Technical Report 2001, Dept. Quantitative Methods, University of Brescia.
Markowitz, H.M., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
Markowitz, H.M., 1959. *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York.
Marshall, A.W., Olkin, I., 1979. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.
Michalowski, W., Ogryczak, W., 2001. Extending the MAD portfolio optimization model to incorporate downside risk aversion. *Naval Research Logistics* 48, 185–200.
Morgan, J.P., Inc., 1996. RiskMetrics: Technical document. J.P. Morgan, New York.
Ogryczak, W., 2000. Multiple criteria linear programming model for portfolio selection. *Annals of Operations Research* 97, 143–162.
Ogryczak, W., Ruszczyński, A., 1999. From stochastic dominance to mean-risk models: semideviations as risk measures. *European Journal of Operational Research* 116, 33–50.
Ogryczak, W., Ruszczyński, A., 2001. On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming* 89, 217–232.
Ogryczak, W., Ruszczyński, A., 2002. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, in print (Report RRR 10–01, RUTCOR, Rutgers Univ., 2001).
Porter, R.B., 1974. Semivariance and stochastic dominance: A comparison. *The American Economic Review* 64, 200–204.
Porter, R.B., Gaumnitz, J.E., 1972. Stochastic dominance versus mean–variance portfolio analysis. *The American Economic Review* 62, 438–446.
Rockafellar, R.T., 1970. *Convex Analysis*. Princeton University Press, Princeton.
Rothschild, M., Stiglitz, J.E., 1969. Increasing risk: I. A definition. *Journal of Economic Theory* 2, 225–243.
Shalit, H., Yitzhaki, S., 1994. Marginal conditional stochastic dominance. *Management Science* 40, 670–684.

Sharpe, W.F., 1971a. A linear programming approximation for the general portfolio analysis problem. *Journal of Financial and Quantitative Analysis* 6, 1263–1275.

Sharpe, W.F., 1971b. Mean-absolute deviation characteristic lines for securities and portfolios. *Management Science* 18, B1–B13.

Shorrocks, A.F., 1983. Ranking income distributions. *Economica* 50, 3–17.

Speranza, M.G., 1993. Linear programming models for portfolio optimization. *Finance* 14, 107–123.

Whitmore, G.A, Findlay, M.C., 1978. *Stochastic Dominance: An Approach to Decision–Making Under Risk*. D.C. Heath, Lexington.

Yitzhaki, S., 1982. Stochastic dominance, mean variance, and Gini's mean difference. *American Economic Revue* 72, 178–185.

Young, M.R., 1998. A minimax portfolio selection rule with linear programming solution. *Management Science* 44, 673–683.