Stochastic Dominance Relation and Linear Programming Mean–Risk Models*

Włodzimierz Ogryczak

Warsaw University, Institute of Informatics

Banacha 2, 02–097 Warsaw, Poland

ogryczak@mimuw.edu.pl

Abstract

Two methods are frequently used for modeling the choice among uncertain prospects: stochastic dominance relation and mean–risk approaches. The former is based on an axiomatic model of risk-averse preferences but does not provide a convenient computational recipe. The latter quantifies the problem in a lucid form of two criteria with possible trade-off analysis, but cannot model all risk-averse preferences. The seminal Markowitz model uses the variance as the risk measure in the mean–risk analysis which results in a formulation of a quadratic programming model. Following the pioneering work of Sharpe, many attempts have been made to linearize the mean–risk approach. There were introduced risk measures which lead to linear programming mean–risk models. This paper focuses on two such risk measures: the Gini's mean (absolute) difference and the mean absolute deviation. Consistency of the corresponding mean–risk models with the second degree stochastic dominance (SSD) is reexamined. Both the models are in some manner consistent with the SSD rules, provided that the trade-off coefficient is bounded by a certain constant. However, for the Gini's mean difference the consistency turns out to be much stronger than that for the mean absolute deviation. The analysis is graphically illustrated within the framework of the absolute Lorenz curves.

*Presented at the 23rd Meeting of the EURO Working Group on Financial Modelling, October 8–10, 1998, Cracow, Poland and published as "Stochastic Dominance Relation and Linear Risk Measures, in: Financial Modelling - Proceedings of the 23rd Meeting of the EURO Working Group on Financial Modelling, A.M.J. Skulimowski (Ed.), Progress & Business Publishers, Cracow 1999, 191-212."

1 Introduction

Comparing uncertain outcomes is one of fundamental interests of decision theory. Our objective is to analyze relations between the existing approaches and to provide some tools to facilitate the analysis. We consider decisions with real-valued outcomes, such as return, net profit or number of lives saved. A leading example focusing our attention, originating from finance, is the problem of choice among investment opportunities or portfolios having uncertain returns. Although we discuss in details the consequences of our analysis in the portfolio optimization context, the theoretical analysis itself is valid for the general problem of comparing real-valued random variables (distributions), assuming that larger outcomes are preferred. We describe a random variable X by the probability measure P_X induced by it on the real line R. It is a general framework: the random variables considered may be discrete, continuous, or mixed (Pratt et al., 1995). The only restriction we impose is that all the random variables X under consideration satisfy the condition $E\{|X|\} < \infty$ (which is certainly true in the portfolio optimization context). Owing to that, our analysis covers a variety of problems of choosing among uncertain prospects that occur in economics and management.

Two methods are frequently used for modeling choice among uncertain prospects: *stochastic dominance* (Whitmore and Findlay, 1978; Levy, 1992), and *mean–risk analysis* (Markowitz, 1987). The former is based on an axiomatic model of risk-averse preferences: it leads to conclusions which are consistent with the axioms. Unfortunately, the stochastic dominance approach does not provide us with a simple computational recipe—it is, in fact, a multiple criteria model with a continuum of criteria. The mean–risk approach quantifies the problem in a lucid form of only two criteria: the mean, representing the expected outcome, and the risk: a scalar measure of the variability of outcomes. The mean–risk model is appealing to decision makers and allows a simple trade-off analysis, analytical or geometrical. On the other hand, mean–risk approaches are not capable of modeling the entire gamut of risk-averse preferences. Moreover, for typical dispersion statistics used as risk measures, the mean–risk approach may lead to inferior conclusions.

The seminal Markowitz (1952) portfolio optimization model uses the variance as the risk measure in the mean–risk analysis. The mean–variance approach applied to the portfolio optimization results in a formulation of a quadratic programming model. Following Sharpe's (1971) work on linear approximation to the mean–variance model, many attempts have been made to linearize the portfolio optimization problem. Yitzhaki (1982) introduced the mean–risk model using Gini's mean (absolute) difference as the risk measure (hereafter referred to as GMD model). Konno and Yamazaki (1991) proposed the MAD portfolio optimization model where risk is measured by (mean) absolute deviation instead of variance. Both these models are computationally attractive as (for discrete random variables) they result in solving linear programming (LP) problems. Capital assets pricing models, similar to the mean–variance CAPM model (c.f., Elton and Gruber, 1987), were developed for both GMD (Shalit and Yitzhaki, 1984) and MAD (Konno, 1990) models. If the rates of return are multivariate normally distributed, then the GMD model as well as the MAD are equivalent to the Markowitz mean–variance model. However, both the linear mean–risk models do not require any specific type of return distributions. Opposite to the mean–variance approach, for general random variables some (partial) consistency with the stochastic dominance rules was shown for the GMD model (Yitzhaki, 1982) and for the MAD model (Ogryczak and Ruszczyński, 1997).

The MAD model was validated with the Tokyo stock exchange data (Konno and Yamazaki, 1991) and it was applied to portfolio optimization for mortgage-backed securities (Zenios and Kang, 1993) where distribution of rate of return is known to be not symmetric. The GMD model seems to get much less recognition from applied studies. In this paper we examine consistency of the MAD and GMD models with the second degree stochastic dominance rules. It turns out that the consistency results for the GMD model are much stronger than that for the MAD model.

2 Mean–risk models

The portfolio optimization problem considered in this paper follows the original Markowitz formulation and is based on a single period model of investment. At the beginning of a period, an investor allocates capital among various securities. Assuming that each security is represented by a variable, this is equivalent to assigning a nonnegative weight to each of the variables. During the investment period, a security generates a certain (random) rate of return. The change of capital invested observed at the end of the period is measured by the weighted average of the individual rates of return.

Let $J = \{1, 2, \ldots, n\}$ denotes 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 $\mu_j = E\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,2,\ldots,n}$ denote a vector of securities weights (decision variables) defining a portfolio. To represent a portfolio, the weights must satisfy a set of constraints which form a feasible set W. The simplest way of defining a feasible set is by a requirement that the weights must sum to one, i.e.:

$$\{\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \ : \ \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n\} \tag{1}$$

An investor usually needs to consider some other requirements expressed as a set of additional side constraints. Hereafter, it is assumed that W is a general polyhedral set given in a LP canonical form as a system of linear equations with nonnegative variables:

$$W = \{\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \ : \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geqq \mathbf{0}\} \tag{2}$$

where $\mathbf{A}$ is a given $p \times n$ matrix and $\mathbf{b} = (b_1, \ldots, b_p)^T$ is a given RHS vector. A vector $\mathbf{x} \in W$ defines a corresponding random variable $X = \sum_{j=1}^{n} R_j x_j$ which is called a *portfolio*. Thus a portfolio X belongs to the (convex) attainable set of random variables defined as

$$Q = \{X = \sum_{j=1}^{n} R_j x_j \ : \quad \mathbf{x} \in W\}.$$

The mean rate of return for portfolio X is given as:

$$\mu_X = E\{X\} = \sum_{j=1}^{n} \mu_j x_j$$

Following Markowitz (1952), the portfolio optimization problem is modeled as a mean–risk optimization problem where μ_X is maximized and some risk measure ϱ_X is minimized. An important advantage of mean–risk approach is a possibility of trade-off analysis. Having assumed a trade-off coefficient λ between the risk and the mean, one may directly compare real values $\mu_X - \lambda\varrho_X$ and find the best portfolio by solving the optimization problem:

$$\max \{\mu_X - \lambda\varrho_X \ : \quad X \in Q\} \tag{3}$$

This analysis is conducted with a so-called *critical line approach* (Markowitz, 1987), by solving parametric problem (3) with changing $\lambda > 0$. Such an approach allows to select appropriate value of the trade-off coefficient λ and the corresponding optimal portfolio through a graphical analysis in the mean-risk image space.

If the risk is measured by variance $\sigma_X^2 = E\{(\mu_X - X)^2\}$ (Markowitz model), then problem (3) results in having a quadratic objective function. The Markowitz model is widely recognized as a starting point for the portfolio theory (c.f., Elton and Gruber, 1987). On the other hand, it is seldom used as a tool for optimizing large portfolios, due to following reasons (Konno and Yamazaki, 1991): (a) a necessity to solve a large scale quadratic programming problem; (b) investor's reluctance to rely on variance as a measure of risk (the Markowitz model is known to be valid and consistent with the stochastic dominance in the case of normal distribution of returns but it becomes doubtful in case of other return distributions, especially nonsymmetric ones).

One may consider an alternative risk measure defined as the *(mean) absolute deviation* from the mean:

$$\delta_X = E\{|X - \mu_X|\} = \int_{-\infty}^{+\infty} |\mu_X - \xi| \; P_X(d\xi). \tag{4}$$

The absolute deviation was considered in the portfolio analysis (Sharpe, 1971a, and references therein) and has been given official status as a recommended measure of dispersion by the Bank Administration Institute (1968). Konno and Yamazaki (1991) presented the complete portfolio optimization model based on the absolute deviation as a risk measure, so-called MAD model, and they validated this model using the Tokyo stock exchange data.

Note that absolute deviation δ_X equals twice the (downside) absolute semideviation:

$$\begin{aligned} \bar{\delta}_X &= E\{\max\{\mu_X - X, 0\}\} = E\{\mu_X - X | X \leq \mu_X\} P\{X \leq \mu_X\} \\ &= \int_{-\infty}^{\mu_X} (\mu_X - \xi) \; P_X(d\xi) = \frac{1}{2} \int_{-\infty}^{\infty} |\mu_X - \xi| \; P_X(d\xi) = \frac{1}{2}\delta_X \end{aligned} \tag{5}$$

The absolute semideviation $\bar{\delta}_X$ is well defined for any random variable X satisfying the condition $E\{|X|\} < \infty$. The following parametric optimization problem we refer to as the MAD model:

$$\max \; \{\mu_X - \lambda \bar{\delta}_X \;\; : \quad X \in Q\} \tag{6}$$

Simplicity and computational robustness are perceived as the most important advantages of the MAD model. Let r_{jt} be the realization of random variable R_j during period t (where $t = 1, \ldots, T$) which is assumed to be available from, for example, the historical data. It is also assumed that the expected value of R_j can be approximated by the average derived from these realizations. Thus:

$$\mu_j = \frac{1}{T} \sum_{t=1}^{T} r_{jt}$$

Therefore, model (6) can be rewritten (Feinstein and Thapa, 1993) as the following LP:

$$\max \; \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda}{T} \sum_{t=1}^{T} d_t \tag{7}$$

subject to

$$\mathbf{x} \in W \tag{8}$$

$$d_t \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \ldots, T \tag{9}$$

$$d_t \geq 0 \quad \text{for } t = 1, \ldots, T \tag{10}$$

The LP formulation (7)–(10) can be effectively solved even for large number of securities. Moreover, a number of securities included in the optimal portfolio (i.e. a number of weights with nonzero values) is controlled by number T. In the case when feasible set W is given by (1), no more than $T+1$ securities will be included in the optimal portfolio.

Yitzhaki (1982) considered the mean–risk model with risk measured by the *Gini's mean (absolute) difference*:

$$\Gamma_X = \frac{1}{2} \int\!\!\int |\xi - \eta| \; P_X(d\xi) P_X(d\eta). \tag{11}$$

The Gini's mean difference Γ_X is well defined for any random variable X satisfying the condition $E\{|X|\} < \infty$. The following parametric optimization problem we refer to as the GMD model:

$$\max \{\mu_X - \lambda\Gamma_X \; : \quad X \in Q\} \tag{12}$$

Lets return to the case when the mean rates of return of securities are derived from r_{jt} (for $j = 1, \ldots, n$ and $t = 1, \ldots, T$). The GMD model can be then rewritten as the following LP:

$$\max \; \frac{1}{T}\sum_{t=1}^{T} y_t - \frac{\lambda}{T^2}\sum_{t'=1}^{T-1}\;\sum_{t''=t'+1}^{T} d_{t't''} \tag{13}$$

subject to

$$\mathbf{x} \in W \tag{14}$$

$$y_t = \sum_{j=1}^{n} r_{jt}x_j \quad \text{for } t = 1, \ldots, T \tag{15}$$

$$d_{t't''} \geq y_{t''} - y_{t'} \quad \text{for } t' = 1, \ldots, T-1;\; t'' = t'+1, \ldots, T \tag{16}$$

$$d_{t't''} \geq y_{t'} - y_{t''} \quad \text{for } t' = 1, \ldots, T-1;\; t'' = t'+1, \ldots, T \tag{17}$$

Note that the LP formulation (13)–(17) contains T^2 linear inequalities whereas for the MAD model (7)–(10) we have introduced only T inequalities. Nevertheless, the optimization problem (13)–(17) can be effectively solved even for large number of securities provided that the number T is not too large.

3 Stochastic dominance and absolute Lorenz curves

Stochastic dominance is based on an axiomatic model of risk-averse preferences (Fishburn, 1964). It originated in the majorization theory (Hardy, Littlewood and Pólya, 1934) for the discrete case and was later extended to general distributions (Hanoch and Levy, 1969; Rothschild and Stiglitz, 1969). Since that time it has been widely used in economics and finance (see Bawa, 1982; Levy, 1992 for numerous references). Detailed and comprehensive discussion of a stochastic dominance and its relation to the downside risk measures is given in Ogryczak and Ruszczyński (1997, 1998).

In the stochastic dominance approach random variables are compared by pointwise comparison of some performance functions constructed from their distribution functions. Let X be a random variable with the probability measure P_X. The first performance function $F_X^{(1)}$ is defined as the right-continuous cumulative distribution function itself:

$$F_X^{(1)}(\eta) = F_X(\eta) = P\{X \leq \eta\} \quad \text{for } \eta \in R.$$

The weak relation of the *first degree stochastic dominance* (FSD) is defined as follows

$$X \succeq_{FSD} Y \quad \Leftrightarrow \quad F_X(\eta) \leq F_Y(\eta) \quad \text{for all } \eta \in R.$$

The second performance function $F_X^{(2)}$ is given by areas below the distribution function F_X:

$$F_X^{(2)}(\eta) = \int_{-\infty}^{\eta} F_X(\xi)\, d\xi \quad \text{for } \eta \in R,$$

and defines the weak relation of the *second degree stochastic dominance* (SSD):

$$X \succeq_{SSD} Y \quad \Leftrightarrow \quad F_X^{(2)}(\eta) \leq F_Y^{(2)}(\eta) \quad \text{for all } \eta \in R. \tag{18}$$

The corresponding strict dominance relations $\succ_{FSD}$ and $\succ_{SSD}$ are defined by the standard rule

$$X \succ Y \quad \Leftrightarrow \quad X \succeq Y \quad \text{and} \quad Y \not\succeq X. \tag{19}$$

Thus, we say that X *dominates* Y *under the FSD rules* ($X \succ_{FSD} Y$), if $F_X(\eta) \le F_Y(\eta)$ for all $\eta \in R$, where at least one strict inequality holds. Similarly, we say that X *dominates* Y *under the SSD rules* ($X \succ_{SSD} Y$), if $F_X^{(2)}(\eta) \le F_Y^{(2)}(\eta)$ for all $\eta \in R$, with at least one inequality strict. A feasible portfolio $X \in Q$ is called *efficient under the SSD (FSD) rules* if the is no $Y \in Q$ such that $Y \succ_{SSD} X$ ($Y \succ_{FSD} X$.

The SSD relation is crucial for decision making under risk. If $X \succ_{SSD} Y$, then X is preferred to Y within all risk-averse preference models that prefer larger outcomes. 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 efficient under the SSD rules.

Function $F_X^{(2)}$, used to define the SSD relation can also be presented as (Ogryczak and Ruszczyński, 1997):

$$F_X^{(2)}(\eta) = \int_{-\infty}^{\eta} (\eta - \xi)\ P_X(d\xi) = P\{X \le \eta\} E\{\eta - X | X \le \eta\} = E\{\max\{\eta - X, 0\}\} \tag{20}$$

thus expressing the expected shortage for each target outcome η. Hence, in addition to being the most general dominance relation for all risk-averse preferences, SSD is also intuitive multidimensional (continuum-dimensional) risk measure. As shown by Ogryczak and Ruszczyński (1997), the graph of $F_X^{(2)}$, referred to as the Outcome–Risk (O–R) diagram, appears to be particularly useful for comparing uncertain returns, since the function $F_X^{(2)}$ is continuous, convex, nonnegative and nondecreasing. The O–R diagram can be used to justify partial consistency of the MAD model with the SSD rules, since $\bar{\delta}_X = F_X^{(2)}(\mu_X)$ is a well established geometrical characteristic of the diagram. However, the Gini's mean difference cannot be placed within the O–R diagram. Therefore, we consider the quantile model of stochastic dominance (Levy and Kroll, 1978).

The left-continuous inverse of the cumulative distribution function F_X is defined as follows (Gastwirth, 1971):

$$F_X^{(-1)}(p) = \inf\ \{\eta\ :\quad F_X(\eta) \ge p\}, \quad 0 \le p \le 1.$$

This definition of the inverse function agrees with the standard one when F_X has a continuous derivative (density) but is general enough to cover all types of random variables. Function $F_X^{(-1)}$ allows the alternative quantile characteristic of the FSD relation, since

$$X \succeq_{FSD} Y \quad \Leftrightarrow \quad F_X^{(-1)}(p) \ge F_Y^{(-1)}(p) \quad \text{for all } 0 \le p \le 1.$$

The second quantile function corresponding to random variable X is defined to be

$$F_X^{(-2)}(p) = \int_0^p F_X^{(-1)}(\alpha) d\alpha \quad \text{for } 0 \le p \le 1.$$

Similar to $F_X^{(2)}$, function $F_X^{(-2)}$ is well defined for any random variable X satisfying the condition $E\{|X|\} < \infty$. The graph of $F_X^{(-2)}$ is referred to as *absolute Lorenz curve* or *ALC diagram* (for short).

It follows from the Young inequality (Young, 1912; and later generalizations) that:

$$F_X^{(-2)}(p) = pF_X^{(-1)}(p) - \int_{-\infty}^{F_X^{(-1)}(p)} F_X(\xi) d\xi = pF_X^{(-1)}(p) - F_X^{(2)}(F_X^{(-1)}(p)) \tag{21}$$

and

$$F_X^{(-2)}(p) = \sup_{\eta}\ \{p\eta - F_X^{(2)}(\eta)\} = \sup_{\eta,\xi}\ \{p\eta - \xi\ :\quad \xi \ge F_X^{(2)}(\eta)\}.$$

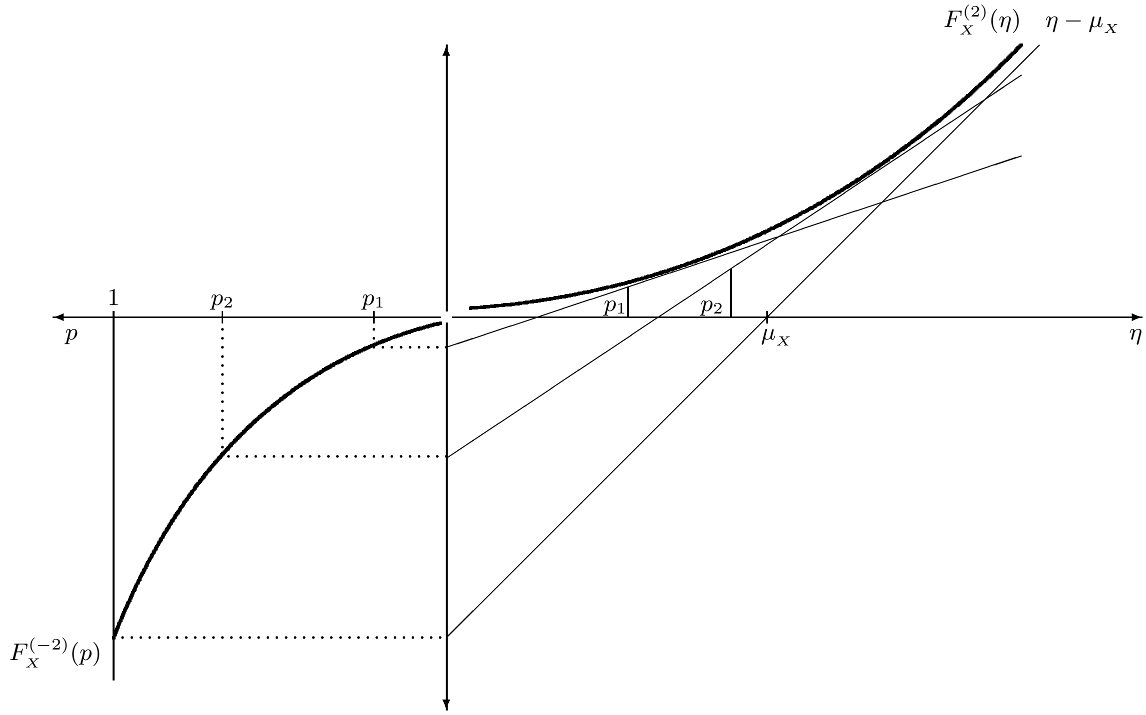

Figure 1: $F_X^{(-2)}(p)$ as the conjugate of $F_X^{(2)}(\eta)$

Hence, $F_X^{(-2)}(p)$ is the conjugate (Rockafellar, 1970) of $F_X^{(2)}(\eta)$ as illustrated in Figure 1. Thus the absolute Lorenz curves provides the dual characteristic of the SSD relation:

$$X \succeq_{SSD} Y \quad \Leftrightarrow \quad F_X^{(-2)}(p) \geq F_Y^{(-2)}(p) \quad \text{for all } 0 \leq p \leq 1. \tag{22}$$

Note that $F_X^{(-2)}(0) = 0$ and $F_X^{(-2)}(1) = \mu_X$, while for $0 < p < 1$, due to (21) and (20), one gets:

$$\begin{aligned} F_X^{(-2)}(p) &= pF_X^{(-1)}(p) - \int_{-\infty}^{F_X^{(-1)}(p)} (F_X^{(-1)}(p) - \xi) P_X(d\xi) \\ &= \int_{-\infty}^{F_X^{(-1)}(p)} \xi P_X(d\xi) - F_X^{(-1)}(p)(F(F_X^{(-1)}(p)) - p) \\ &= E\{X | X \leq F_X^{(-1)}(p)\} P\{X \leq F_X^{(-1)}(p)\} - F_X^{(-1)}(p)(P\{X \leq F_X^{(-1)}(p)\} - p). \end{aligned}$$

In particular, $F_X^{(-2)}(F_X(\eta)) = E\{X|X \leq \eta\}P\{X \leq \eta\}$ thus expressing the (downside) partial mean, while for p not representing any value $F_X(\eta)$, the value $F_X^{(-2)}(p)$ is given by the linear interpolation. Hence, similar to (18), the dual characteristic of the SSD relation (22) is based on a continuum-dimensional risk measurement. However, in the case of (discrete) random variables defined by their realizations for T periods (historical data), the dual approach allows us to consider only T criteria: $F^{(-2)}(i/T)$ for $i = 1, \ldots, T$. This opens an opportunity to employ standard techniques of multiple criteria optimization to portfolio optimization (Ogryczak, 1997).

For any uncertain outcome X, its absolute Lorenz curve $F_X^{(-2)}$ is a continuous convex curve connecting points $(0, 0)$ and $(1, \mu_X)$, whereas a deterministic outcome with the same expected value μ_X, yields the chord (straight line) connecting the same points. Hence, the space between the curve $(p, F_X^{(-2)}(p))$, $0 \leq p \leq 1$, and its chord represents the dispersion (and thereby the riskiness) of X in comparison to the deterministic outcome of μ_X (Fig. 2). We shall call it the *dispersion space*. Note that from ALC diagram one can easily derive the following, commonly known, necessary condition for

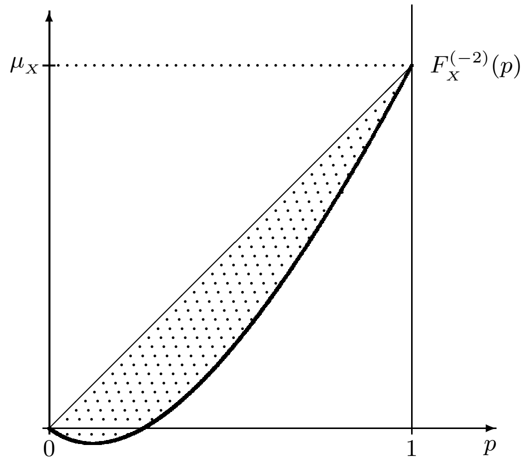

Figure 2: The absolute Lorenz curve and the dispersion space

the SSD relation (e.g., Levy, 1992):

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X \geq \mu_Y. \tag{23}$$

Both size and shape of the dispersion space are important for complete description of the riskiness. Nevertheless, it is quite natural to consider some size parameters as summary characteristics of riskiness. The vertical diameter of the dispersion space at point p is given as:

$$h_X(p) = \mu_X p - F_X^{(-2)}(p) = \int_{-\infty}^{F_X^{(-1)}(p)} (\mu_X - \xi) P_X(d\xi) - (F_X(F_X^{(-1)}(p)) - p)(\mu_X - F_X^{(-1)}(p))$$

It is commonly known that the Gini's mean difference can be expressed as (Atkinson, 1970):

$$\Gamma_X = 2 \int_0^1 (\mu_X p - F_X^{(-2)}(p)) dp \tag{24}$$

thus representing the doubled area of the dispersion space.

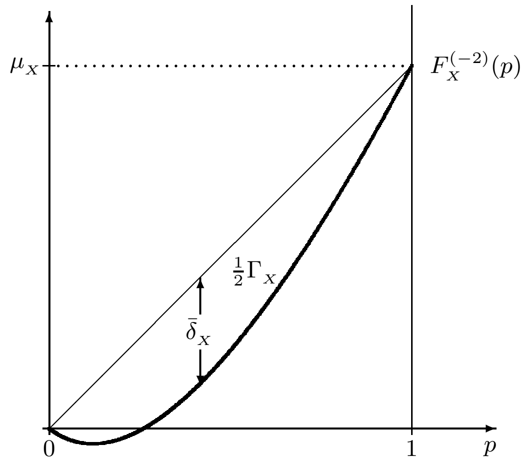

Figure 3: The absolute Lorenz curve and risk measures

The relation between the absolute deviation and the absolute Lorenz curve seems to be less known or at least not widely used in the literature. Note that for any $\bar{p}$ such that $P\{X < \mu_X\} \leq \bar{p} \leq P\{X \leq$

$\mu_X\}$ the following holds:

$$h_X(\bar{p}) = \int_{-\infty}^{\mu_X} (\mu_X - \xi) P_X(d\xi) = \bar{\delta}_X \geq h_X(p) \quad \text{for all } 0 \leq p \leq 1 \tag{25}$$

Hence, the absolute semideviation $\bar{\delta}_X$ turns out to be the maximal vertical diameter of the dispersion space. Thus both Γ and $\bar{\delta}$ are well defined size characteristics of the dispersion space (Fig. 3). However, the absolute semideviation is rather rough measure when comparing to the Gini's mean difference. Note that $\bar{\delta}_X/2$ may be also interpreted in the ALC diagram as the area of the triangle given by vertices: $(0,0)$, $(1,\mu_X)$ and $(\bar{p}, F_X^{(-2)}(\bar{p}))$, which is a triangular approximation of the dispersion space. In fact, $\bar{\delta}_X$ is the Gini's mean difference of a two-point distribution approximating random variable X. Nevertheless, for a given value of $\bar{\delta}_X$, the Gini's mean difference Γ_X can take various values within the interval $[\bar{\delta}_X, 2\bar{\delta}_X)$.

4 Mean–risk models and stochastic dominance

In this section we use the absolute Lorenz curves to analyze the consistency of the MAD and GMD models with the SSD efficiency. Recall that $\bar{\delta}$ represents the largest vertical diameter of the dispersion space while $\Gamma/2$ measures its area. Hence, both $\bar{\delta}$ and Γ are well defined geometrical characteristics in the ALC diagram.

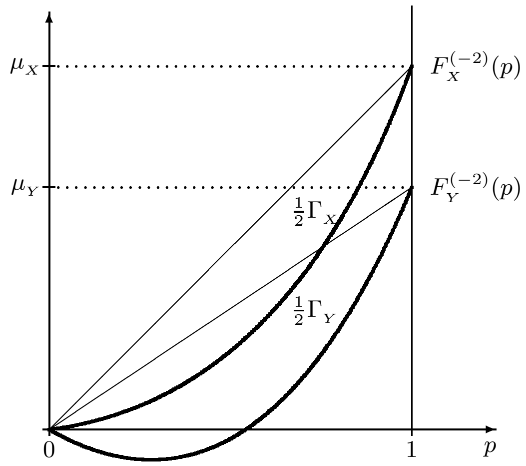

Figure 4: $X \succeq_{SSD} Y \;\Rightarrow\; \frac{1}{2}\Gamma_X \leq \frac{1}{2}\Gamma_Y + \frac{1}{2}(\mu_X - \mu_Y)$

Consider two random variables X and Y in the common ALC diagram (Figure 4). If $X \succeq_{SSD} Y$, then, due to (22), $F_X^{(-2)}$ is bounded from below by $F_Y^{(-2)}$, and $\mu_X \geq \mu_Y$ from (23). Thus the area of the dispersion space for X is (upper) bounded by the area of the dispersion for Y plus the area of the triangle between the chords (with vertices: $(0,0)$, $(1,\mu_X)$ and $(1,\mu_Y)$). Hence, $\frac{1}{2}\Gamma_X \leq \frac{1}{2}\Gamma_Y + \frac{1}{2}(\mu_X - \mu_Y)$ and, due to continuity of the Lorenz curves, this inequality becomes strict whenever $X \succ_{SSD} Y$. This simple analysis of the ALC diagram allows us to derive the following necessary conditions for the second degree stochastic dominance.

Proposition 1 *For random variables X and Y the following implications hold:*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - \Gamma_X \geq \mu_Y - \Gamma_Y, \tag{26}$$

$$X \succ_{SSD} Y \quad \Rightarrow \quad \mu_X - \Gamma_X > \mu_Y - \Gamma_Y. \tag{27}$$

Condition (26) was first shown by Yitzhaki (1982) for the bounded distributions. For portfolio optimization problems, the stronger condition (27) allows us to prove the following theorem.

Theorem 1 *Every portfolio $X \in Q$ that maximizes $\mu_X - \lambda\Gamma_X$ with $0 < \lambda \le 1$ is efficient under the SSD rules.*

Proof. Let $0 < \lambda \le 1$ and $X \in Q$ be maximal by $\mu - \lambda\Gamma$. This means that $\mu_X - \lambda\Gamma_X \ge \mu_Y - \lambda\Gamma_Y$ for all $Y \in Q$. Suppose that there exists $Z \in Q$ such that $Z \succ_{SSD} X$. Then, from (23) and (27),

$$\mu_Z \ge \mu_X \quad \text{and} \quad \mu_Z - \Gamma_Z > \mu_X - \Gamma_X.$$

Adding these inequalities multiplied by $(1-\lambda)$ and λ, respectively, we obtain $\mu_Z - \lambda\Gamma_Z > \mu_X - \lambda\Gamma_X$ which contradicts the maximality of $\mu_X - \lambda\Gamma_X$. □

Theorem 1 justifies the critical line approach to the GMD model in the sense that by solving parametric problem (12) with varying $0 < \lambda \le 1$ only SSD efficient are generated. The upper bound on the trade-off coefficient λ in Theorem 1 cannot be increased for general distributions. For any $\varepsilon > 0$ there exist random variables $X \succ_{SSD} Y$ such that $\mu_X > \mu_Y$ and $\mu_X - (1+\varepsilon)\bar{\delta}_X = \mu_Y - (1+\varepsilon)\bar{\delta}_Y$. As an example one may consider two finite random variables X and Y defined as:

$$P\{X = \xi\} = \begin{cases} 1/(1+\varepsilon), & \xi = 0 \\ \varepsilon/(1+\varepsilon), & \xi = 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad P\{Y = \xi\} = \begin{cases} 1, & \xi = 0 \\ 0, & \text{otherwise} \end{cases} \tag{28}$$

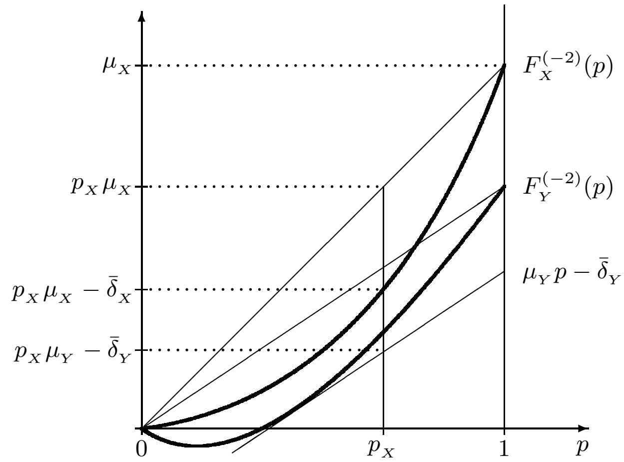

Figure 5: $X \succeq_{SSD} Y \;\Rightarrow\; p_X\mu_X - \bar{\delta}_X \ge p_X\mu_Y - \bar{\delta}_Y$ where $p_X = P\{X < \mu_X\}$

In order to analyze the MAD model, consider again two random variables X and Y in the common ALC diagram (Figure 5). Recall that $X \succeq_{SSD} Y$ implies that $F_X^{(-2)}$ is bounded from below by $F_Y^{(-2)}$ and $\mu_X \ge \mu_Y$. Since $\bar{\delta}_Y$ represents the maximal vertical diameter of the dispersion space for variable Y, its absolute Lorenz curve $F_Y^{(-2)}(p)$ is bounded from below by the straight line $\mu_Y p - \bar{\delta}_Y$. Focusing on point $p_X = P\{X < \mu_X\}$ one gets:

$$\mu_X p_X - \bar{\delta}_X = F_X^{(-2)}(p_X) \ge F_Y^{(-2)}(p_X) \ge \mu_Y p_X - \bar{\delta}_Y.$$

Since $p_X = P\{X < \mu_X\} < 1$, this simple analysis of the ALC diagram allows us to derive the following necessary condition for the second degree stochastic dominance.

Proposition 2 *If $X \succeq_{SSD} Y$, then $\mu_X \ge \mu_Y$ and $\mu_X - \bar{\delta}_X \ge \mu_Y - \bar{\delta}_Y$, where the second inequality is strict whenever $\mu_X > \mu_Y$.*

Proposition 2 was first shown by Ogryczak and Ruszczyński (1997) with the use O–R diagram. Here, by placing the considerations within the (dual) ALC diagram we make transparent that Proposition 2 is based on comparison of the absolute Lorenz curves at only one point p_X whereas the entire curves were taken into account to derive Proposition 1 for the Gini's mean difference. For portfolio optimization problem, Proposition 2 allows us to prove the following theorem.

Theorem 2 *Except for portfolios with identical mean and absolute semideviation, every portfolio $X \in Q$ that maximizes $\mu_X - \lambda\bar{\delta}_X$ with $0 < \lambda \leq 1$ is efficient under the SSD rules.*

Proof. Let $0 < \lambda \leq 1$ and $X \in Q$ be maximal by $\mu - \lambda\bar{\delta}$. This means that $\mu_X - \lambda\bar{\delta}_X \geq \mu_Y - \lambda\bar{\delta}_Y$ for all $Y \in Q$. Suppose that there exists $Z \in Q$ such that $Z \succ_{SSD} X$. Then, from Proposition 2,

$$\mu_Z \geq \mu_X \quad \text{and} \quad \mu_Z - \bar{\delta}_Z \geq \mu_X - \bar{\delta}_X.$$

Adding these inequalities multiplied by $(1-\lambda)$ and λ, respectively, we obtain: $\mu_Z - \lambda\bar{\delta}_Z \geq \mu_X - \lambda\bar{\delta}_X$. The latter together with the fact that X is optimal implies $\mu_Z - \lambda\bar{\delta}_Z = \mu_X - \lambda\bar{\delta}_X$ which means that Z must be also an optimal solution. If $\mu_Z = \mu_X$, then obviously $\bar{\delta}_Z = \bar{\delta}_X$. Otherwise, by Proposition 2,

$$\mu_Z - \lambda\bar{\delta}_Z = (1-\lambda)\mu_Z + \lambda(\mu_Z - \bar{\delta}_Z) > (1-\lambda)\mu_X + \lambda(\mu_X - \bar{\delta}_X) = \mu_X - \lambda\bar{\delta}_X$$

which contradicts maximality of $\mu_X - \lambda\bar{\delta}_X$. □

It follows from Theorem 2 that the unique optimal solution of the MAD problem (model (6)) with the trade-off coefficient $0 < \lambda \leq 1$ is efficient under the SSD rules. In the case of multiple optimal solutions (which is common in linear programming), one of them is efficient under SSD rules, but also some of them may be SSD dominated. Due to Theorem 2, a MAD optimal portfolio $X \in Q$ can be SSD dominated only by another MAD optimal portfolio $Y \in Q$ such that $\mu_Y = \mu_X$ and $\bar{\delta}_Y = \bar{\delta}_X$. However, two random variables with equal expected values and absolute semideviations can be quite different. For instance, for finite random variables X and Y defined as:

$$P\{X = \xi\} = \begin{cases} 0.5, & \xi = -20 \\ 0.5, & \xi = 20 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad P\{Y = \xi\} = \begin{cases} 0.01, & \xi = -1000 \\ 0.98, & \xi = 0 \\ 0.01, & \xi = 1000 \\ 0, & \text{otherwise} \end{cases}$$

one gets $\mu_X = \mu_Y = 0$ and $\bar{\delta}_X = \bar{\delta}_Y = 10$, but $X \succ_{SSD} Y$ and $F_X^{(-2)}(p) < F_Y^{(-2)}(p)$ for all $0 < p < 1$ except for $p = 0.5$. Note that $\Gamma_Y = 19.9$ is almost two times greater than $\Gamma_X = 10$.

Theorem 2 partially justifies the critical line approach to the MAD model in the sense that by solving parametric problem (6) with varying $0 < \lambda \leq 1$ the corresponding image in the $\mu/\bar{\delta}$ space represents SSD efficient solutions. Thus it can be used as the mean–risk map to seek a satisfactory $\mu/\bar{\delta}$ compromise. It does not mean, however, that the solutions generated during the parametric optimization (6) are SSD efficient. Therefore, having decided on some values of μ and $\bar{\delta}$ one should apply additional specification to select a specific portfolio which is SSD efficient. This can be implemented with additional minimization of the Gini's mean difference, because it follows from Proposition 1 that the portfolio that minimizes Γ within the set of portfolios with the same mean value is efficient under the SSD rules.

The upper bound on the trade-off coefficient λ in Theorem 2 cannot be increased for general distributions. For any $\varepsilon > 0$ there exist random variables $X \succ_{SSD} Y$ such that $\mu_X > \mu_Y$ and $\mu_X - (1+\varepsilon)\bar{\delta}_X = \mu_Y - (1+\varepsilon)\bar{\delta}_Y$. As an example one may consider two finite random variables X and Y defined as (28).

Although the upper bound on the trade-off coefficient λ in the MAD and GMD models cannot be increased for general distribution, it can be doubled in the case of symmetric random variables. Note that for a symmetric random variable X the following holds: $h_X(p) = h_X(1-p)$ and $\bar{\delta}_X = h_X(0.5)$. Hence, for symmetric random variables one may restrict the analysis of absolute Lorenz curves to the interval $[0, 0.5]$. Considering again two random variables $X \succeq_{SSD} Y$ in the common ALC diagram (Figure 6) for $0 \leq p \leq 0.5$, one can easily derive the following assertions.

Proposition 3 *For symmetric random variables X and Y the following implications hold:*

$$\begin{aligned} X \succeq_{SSD} Y \quad &\Rightarrow \quad \mu_X - 2\Gamma_X \geq \mu_Y - 2\Gamma_Y, \\ X \succ_{SSD} Y \quad &\Rightarrow \quad \mu_X - 2\Gamma_X > \mu_Y - 2\Gamma_Y. \end{aligned}$$

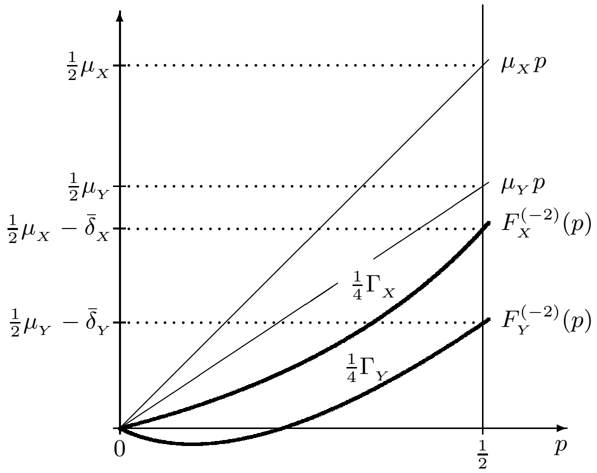

Figure 6: Symmetric case: $X \succeq_{SSD} Y \;\Rightarrow\; \frac{1}{2}\mu_X - \bar{\delta}_X \geq \frac{1}{2}\mu_Y - \bar{\delta}_Y$ and $\frac{1}{4}\Gamma_X \leq \frac{1}{4}\Gamma_Y + \frac{1}{8}(\mu_X - \mu_Y)$

Proposition 4 *For symmetric random variables X and Y the following implication holds:*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - 2\bar{\delta}_X \geq \mu_Y - 2\bar{\delta}_Y.$$

Corollary 1 *Within the class of symmetric random variables every random variable $X \in Q$ that maximizes $\mu_X - \lambda\Gamma_X$ with $0 < \lambda \leq 2$, is efficient under the SSD rules.*

Corollary 2 *Within the class of symmetric random variables, except for random variables with identical mean and absolute semideviation, every random variable $X \in Q$ that maximizes $\mu_X - \lambda\bar{\delta}_X$ with $0 < \lambda < 2$, is efficient under the SSD rules.*

5 Concluding remarks

The mean–risk approach quantifies the problem in a lucid form of only two criteria: the mean, representing the expected outcome, and the risk: a scalar measure of the variability of outcomes. The mean–risk model is appealing to decision makers and allows a simple trade-off analysis, analytical or geometrical. The seminal Markowitz (1952) portfolio optimization model uses the variance as the risk measure in the mean–risk analysis. The mean–variance approach applied to the portfolio optimization results in a formulation of a quadratic programming model. Yitzhaki (1982) introduced the GMD model using Gini's mean difference as the risk measure. Konno and Yamazaki (1991) proposed the MAD portfolio optimization model where risk is measured by (mean) absolute deviation instead of variance. Both these models are computationally attractive as for discrete random variables defined by their realizations for T periods (historical data) they result in solving linear programming problems. The LP formulation for the MAD model is much simpler than that for the GMD model as the former contains only T linear inequalities whereas T^2 linear inequalities is necessary for the latter.

Opposite to the mean–variance approach, both GMD and MAD models are at least partially consistent with the SSD rules. Every optimal solution to the GMD model with the trade-off coefficient bounded by 1 is efficient under the SSD rules. The consistency of the MAD model depends on the assertion that except for portfolios with identical mean and absolute deviation, every optimal solution to the MAD model with the trade-off coefficient bounded by 1 is efficient under the SSD rules. This means, the unique optimal solution to the MAD model (with bounded trade-off) is efficient under the SSD rules, but in the case of multiple optimal solutions, some of them may be SSD dominated. It is a serious weakness of the MAD model since large linear programming problems usually have multiple optimal solutions and typical LP solvers generate one of them (essentially at random). Thus, the MAD model although much simpler than the GMD one, it requires additional specification if one wants to maintain the SSD efficiency for every optimal portfolio.

The mean absolute deviation may be considered an approximation to the Gini's mean difference. Actually, it is the Gini's mean difference for a two-point distribution approximating the original random variable. Thus this approximation is very rough. Therefore, further work on additional refinement of the MAD model (Konno, 1990; Michalowski and Ogryczak, 1998) seems to be a very promising direction of research on linear programming models for portfolio optimization.

References

Atkinson, A.B., (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, **2**, 244–263.

Bank Administration Institute (1968), *Measuring the Investment Performance of Pension Funds for the Purpose of Inter-Fund Comparison*, Bank Administration Institute, Park Ridge, Ill.

Bawa, V.S. (1982), "Stochastic Dominance: A Research Bibliography", *Management Science*, **28**, 698–712.

Elton, E.J., Gruber, M.J. (1987), *Modern Portfolio Theory and Investment Analysis* (Third Edition), John Wiley & Sons, New York.

Feinstein, C.D., Thapa, M.N. (1993), "A Reformulation of a Mean–absolute Deviation Portfolio Optimization Model", *Management Science*, **39**, 1552–1553.

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

Gastwirth, J.L. (1971), "A General Definition of the Lorenz Curve", *Econometrica*, **39**, 1037–1039.

Hanoch, G., Levy, H. (1969), "The Efficiency Analysis of Choices Involving Risk", *Rev. Economic Studies*, **36**, 335–346.

Hardy, G.H., Littlewood, J.E., Pólya, G. (1934), *Inequalities*, Cambridge University Press, Cambridge, MA.

Konno, H. (1990), "Piecewise Linear Risk Function and Portfolio Optimization", *Journal of the Operations Research Society of Japan*, **33**, 139–156.

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.

Levy, H., Kroll, Y. (1978), "Ordering Uncertain Options with Borrowing and Lending", *Journal of Finance*, **33**, 553–573.

Markowitz, H.M. (1952), "Portfolio Selection", *Journal of Finance*, **7**, 77–91.

Markowitz, H.M. (1987), *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford.

Michalowski, W., Ogryczak, W. (1998), "Extending the MAD Portfolio Optimization Model to Incorporate Downside Risk Aversion", Interim Report IR–98–041, IIASA, Laxenburg.

Ogryczak, W. (1997), "Multiple Criteria Linear Programming Model for Portfolio Selection", Technical Report TR 97–06 (243), Institute of Informatics, Warsaw University (*Annals of Operations Research*, in print).

Ogryczak, W., Ruszczyński, A. (1997), "From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures", Interim Report IR–97–027, IIASA, Laxenburg (*European Journal of Operational Research*, in print).

Ogryczak, W., Ruszczyński, A. (1998), "On Stochastic Dominance and Mean–Semideviation Models", Research Report RRR 7–98, RUTCOR, Rutgers University, Piscataway NJ.

Pratt, J.W., Raiffa, H., Schlaifer, R. (1995), *Introduction to Statistical Decision Theory*, MIT Press, Cambridge, MA.

Rockafellar, R.T. (1970), *Convex Analysis*, Princeton Univ. Press, Princeton, NJ.

Rothschild, M., Stiglitz, J.E. (1969), "Increasing Risk: I. A Definition", *Journal of Economic Theory*, **2**, 225–243.

Shalit, H., Yitzhaki, S. (1984), "Mean–Gini, Portfolio Theory, and the Pricing of Risky Assets", *Jornal of Finance*, **39**, 1449–1468.

Sharpe W.F. (1971), "A Linear Programming Approximation for the General Portfolio Analysis Problem", *Journal of Financial and Quantitative Analysis*, **6**, 1263–1275.

Sharpe, W.F. (1971a), "Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios", *Management Science*, **18**, B1–B13.

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

Yitzhaki, S. (1982), "Stochastic Dominance, Mean Variance, and Gini's Mean Difference", *American Economic Review*, **72**, 178–185.

Young, W.H. (1912), "On Classes of Summable Functions and Their Fourier Series", *Proceedings of the Royal Society of London*, **A87**, 235–239.

Zenios, S.A., Kang, P. (1993), "Mean–Absolute Deviation Portfolio Optimization for Mortgage-Backed Securities", *Annals of Operations Research*, **45**, 433–450.