DUAL STOCHASTIC DOMINANCE AND RELATED MEAN–RISK MODELS*

WŁODZIMIERZ OGRYCZAK† AND ANDRZEJ RUSZCZYŃSKI‡

Abstract. We consider the problem of constructing mean–risk models which are consistent with the second degree stochastic dominance relation. By exploiting duality relations of convex analysis we develop the quantile model of stochastic dominance for general distributions. This allows us to show that several models using quantiles and tail characteristics of the distribution are in harmony with the stochastic dominance relation. We also provide stochastic linear programming formulations of these models.

Key words. Decisions under uncertainty, stochastic dominance, Fenchel duality, mean–risk analysis, quantile risk measures, stochastic programming.

AMS subject classifications. Primary, 90A05, 90A46, 52A41; Secondary, 90A09, 90C15

1. Introduction. The relation of *stochastic dominance* is one of the fundamental concepts of the decision theory (cf. [26, 12]). 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 nondecreasing utility functions. While theoretically attractive, stochastic dominance order is computationally very difficult, as a multiobjective model with a continuum of objectives.

The practice of decision making under uncertainty frequently resorts to *mean–risk models* (cf. [15]). The mean–risk approach uses only two criteria: the *mean*, representing the expected outcome, and the *risk*: a scalar measure of the variability of outcomes. This allows a simple trade-off analysis, analytical or geometrical. However, for typical dispersion statistics used as risk measures, the mean–risk approach may lead to inferior conclusions, that is, some efficient (in the mean–risk sense) solutions may be stochastically dominated by other feasible solutions. It is of primary importance to construct mean–risk models which are in harmony with stochastic dominance relations.

The classical Markowitz [14] model uses the variance as the risk measure in the mean–risk analysis. Since then many authors have pointed out that the mean–variance model is, in general, not consistent with stochastic dominance rules. In our preceding paper [18] we have proved that the standard semideviation (square root of the semivariance) or the mean absolute deviation (from the mean) as the risk measures make the corresponding mean–risk models consistent with the second degree stochastic dominance, provided that the trade-off coefficient is bounded by a certain constant. These results were further generalized in [7, 19] where it was shown that mean–risk models using higher order central semideviations as risk measures are in harmony with the stochastic dominance relations of the corresponding degree.

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 [25] work on linear programming (LP) approximation to

*July 10, 2000; submitted for publication in *SIAM Journal on Optimization*.

†Warsaw University of Technology, Institute of Control and Computation Engineering, 00–665 Warsaw, Poland (W.Ogryczak@ia.pw.edu.pl).

‡Rutgers University, RUTCOR and Department of Management Science and Information Systems, Piscataway, NJ 08854, USA (rusz@rutcor.rutgers.edu).

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 [27] introduced the mean–risk model using the Gini's mean (absolute) difference as a risk measure. Konno and Yamazaki [11] analyzed the model where risk is measured by the (mean) absolute deviation. Young [28] 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. 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 [17] based on the majorization theory [9] and Lorenz type orders [1].

In this paper we analyze a *dual* model of the stochastic dominance by exploiting duality relations of convex analysis (see, e.g., [21]). These transformations allow us to show consistency with stochastic dominance of mean–risk models using quantiles and tail characteristics of the distribution as risk measures. We also show that these models are equivalent to certain stochastic linear programming problems, thus opening a new area of applications of stochastic programming.

The paper is organized as follows. In §2 we formally define stochastic dominance relations and the concept of consistency of mean–risk models with these relations. Section 3 introduces dual formulations of stochastic dominance and exploits Fenchel duality to characterize dominance in terms of quantile performance functions. In §4 we consider several risk measures based on quantiles and tail characteristics of the distribution and we analyse their relation to stochastic dominance. Section 5 is devoted to the analysis of mean–risk models using these risk measures. In §6 we present stochastic linear programming formulations of these models. Finally, we have a conclusions section.

We use $(\Omega, \mathcal{B}, \mathbb{P})$ to denote an abstract probability space. For a random variable $X : \Omega \to \mathbb{R}$ we denote by P_X the measure induced by it on the real line. For a convex function $F : \mathbb{R} \to \bar{\mathbb{R}}$ we denote by F^* its convex conjugate [21], $F^*(p) = \sup_\xi \{p\xi - F(\xi)\}$.

2. Stochastic Dominance and Mean–Risk Models. Stochastic dominance is based on an axiomatic model of risk-averse preferences [5]. It originated in the majorization theory [9] for the discrete case and was later extended to general distributions [8, 23]. Since that time it has been widely used in economics and finance (see [3, 12] for numerous references). Detailed and comprehensive discussion of a stochastic dominance and its relation to downside risk measures is given in [18, 19].

In the stochastic dominance approach random variables are compared by pointwise comparison of some performance functions constructed from their distribution functions. For a real random variable X, its first performance function $F_X^{(1)} : \mathbb{R} \to [0, 1]$ is defined as the right-continuous cumulative distribution function itself:

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

In the definition below, and elsewhere in this paper, we assume that larger outcomes are preferred to smaller.

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

follows

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

The second performance function $F_X^{(2)} : \mathbb{R} \to \mathbb{R}_+$ is given by areas below the distribution function F_X,

$$F_X^{(2)}(\eta) = \int_{-\infty}^{\eta} F_X(\xi)\, d\xi \quad \text{for } \eta \in \mathbb{R}, \tag{2.1}$$

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

$$X \succeq_{SSD} Y \quad \Leftrightarrow \quad F_X^{(2)}(\eta) \leq F_Y^{(2)}(\eta) \quad \text{for all } \eta \in \mathbb{R}. \tag{2.2}$$

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

$$X \succ Y \quad \Leftrightarrow \quad X \succeq Y \quad \text{and} \quad Y \not\succeq X. \tag{2.3}$$

Thus, we say that X *dominates* Y *under the FSD rules* ($X \succ_{FSD} Y$), if $F_X(\eta) \leq F_Y(\eta)$ for all $\eta \in \mathbb{R}$, where at least one strict inequality holds. Similarly, we say that X *dominates* Y *under the SSD rules* ($X \succ_{SSD} Y$), if $F_X^{(2)}(\eta) \leq F_Y^{(2)}(\eta)$ for all $\eta \in \mathbb{R}$, with at least one inequality strict.

For a set Q of random variables, a variable $X \in Q$ is called *SSD-efficient* (or FSD-efficient) in Q if there is no $Y \in Q$ such that $Y \succ_{SSD} X$ (or $Y \succ_{FSD} X$).

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. The function $F_X^{(2)}$ can also be expressed as the expected shortfall [18]: for each target value η we have

$$\begin{aligned} F_X^{(2)}(\eta) &= \int_{-\infty}^{\eta} (\eta - \xi)\, P_X(d\xi) \\ &= \mathbb{E}\{\max(\eta - X, 0)\} = \mathbb{P}\{X \leq \eta\}\, \mathbb{E}\{\eta - X | X \leq \eta\}. \end{aligned} \tag{2.4}$$

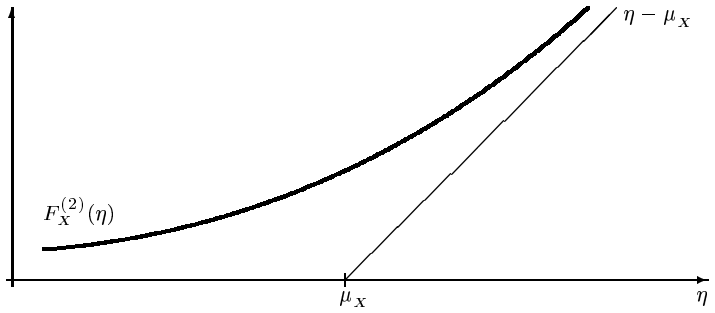

FIG. 2.1. *The O–R diagram*

The function $F_X^{(2)}$ is continuous, convex, nonnegative and nondecreasing. Its graph, referred to as the Outcome–Risk (O–R) diagram and illustrated in Figure 2.1, has two asymptotes which intersect at the point $(\mu_X, 0)$: the horizontal axis, and the line $\eta - \mu_X$. In the case of a deterministic outcome ($X = \mu_X$), the graph of

$F_X^{(2)}$ coincides with the asymptotes, whereas any uncertain outcome with the same expected value μ_X yields a graph above (precisely, not below) the asymptotes. Hence, the space between the curve $(\eta, F_X^{(2)}(\eta))$, $\eta \in \mathbb{R}$, and its asymptotes represents the dispersion (and thereby the riskiness) of X in comparison to the deterministic outcome of μ_X. It is refered to as the *dispersion space*.

It is convenient to introduce also the distance to the right asymptote,

$$\overline{F}_X^{(2)}(\eta) = F_X^{(2)}(\eta) - (\eta - \mu_X), \tag{2.5}$$

which can be rewritten as

$$\begin{aligned}\overline{F}_X^{(2)}(\eta) &= \int_\eta^\infty (\xi - \eta)\ P_X(d\xi) \\ &= \mathbb{E}\{\max(X - \eta, 0)\} = \mathbb{P}\{X \geq \eta\}\, \mathbb{E}\{X - \eta | X \geq \eta\}\end{aligned} \tag{2.6}$$

thus expressing the expected surplus for each target outcome η (see [18]). The vertical diameter of the dispersion space at a point η is given as:

$$d_X(\eta) = \min(F_X^{(2)}(\eta),\ \overline{F}_X^{(2)}(\eta)) \tag{2.7}$$

While SSD is a sound theoretical concept, its application to real world decision problems is difficult, because it requires a pairwise comparison of all possible outcome distributions. We would prefer to use simple mean–risk models, and deduce from them whether a particular outcome distribution is dominated or not.

In general, considering a mean–risk model with the risk of a random outcome X measured by some functional r_X, we can introduce the following definition.

DEFINITION 2.1. *We say that the mean–risk model* (μ_X, r_X) *is* consistent *with SSD, if the following relation holds*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X \geq \mu_Y \ \text{ and } r_X \leq r_Y.$$

It is well known that the first inequality at the right hand side is true: $X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X \geq \mu_Y$ (see [12]). The inequality for the risk term, though, is not true for some popular risk measures, like the variance or absolute deviation.

Directly from (2.4) we see that the mean–risk model with the risk functional defined as the expected shortfall below some fixed target t,

$$r_X^t = \mathbb{E}\{\max(t - X, 0)\},$$

is consistent with the SSD. Integrating the inequality $r_X^t \leq r_Y^t$ with respect to some probability measure P_T we conclude that the expected shortfall from a *random* target T distributed according to P_T,

$$r_X = \int \mathbb{E}\{\max(t - X, 0)\}\ P_T(dt) = \mathbb{E}\{\max(T - X, 0)\}, \tag{2.8}$$

is consistent with the SSD.

While the use of consistent mean–risk models is quite straightforward, there are some reasonable risk measures which do not enjoy the consistency property of Definition 2.1. Therefore, following [19], we relax it a little.

DEFINITION 2.2. *We say that the mean–risk model* (μ_X, r_X) *is* α-consistent *with SSD, where* $\alpha > 0$, *if the following relation is true*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - \alpha r_X \geq \mu_Y - \alpha r_Y .$$

It is clear that α-consistency implies λ-consistency for all $0 \leq \lambda \leq \alpha$.

The concept of α-consistency turned out to be fruitful. In [18] we have proved that the mean–risk model in which the risk is defined as the *absolute semideviation*,

$$\bar{\delta}_X = \mathbb{E}\{\max(\mu_X - X, 0)\} = \int_{-\infty}^{\mu_X} (\mu_X - \xi)\ P_X(d\xi) \tag{2.9}$$

is 1-consistent with SSD. An identical result (under the condition of finite second moments) has been obtained in [18] for the *standard semideviation*,

$$\bar{\sigma}_X = \Big(\mathbb{E}\{(\max(\mu_X - X, 0))^2\}\Big)^{1/2} = \Big(\int_{-\infty}^{\mu_X} (\mu_X - \xi)^2\ P_X(d\xi)\Big)^{1/2}. \tag{2.10}$$

These results have been further extended in [19] to central semideviations of higher orders and stochastic dominance relations of higher degrees.

REMARK 1. In [2] a class of *coherent risk measures* has been defined by means of several axioms. In our terms, these measures correspond to composite objectives of form $\rho(X) = -\mu_X + \alpha r_X$ (note the sign change), where $\alpha > 0$. The axioms are: translation invariance, positive homogeneity, subadditivity, 'monotonicity' ($X \geq Y$ a.s. $\Rightarrow \rho(X) \leq \rho(Y)$), and 'relevance' ($X \leq 0, X \neq 0 \Rightarrow \rho(X) < 0$).

Both $\bar{\delta}_X$ and $\bar{\sigma}_X$, as seminorms in $\mathcal{L}_1$ and $\mathcal{L}_2$, are convex and positively homogeneous. Therefore the composite objectives $-\mu_X + \alpha\bar{\delta}_X$ and $-\mu_X + \alpha\bar{\sigma}_X$ do satisfy the first three axioms (contrary to the statement in [2, Rem. 2.10]). For $\alpha \in (0, 1]$, owing to the consistency with stochastic dominance in the sense of Definition 2.2, they also satisfy 'monotonicity' and 'relevance', because $X \geq Y$ a.s. $\quad \Rightarrow \quad X \succeq_{SSD} Y$.

Our objective is to analyze the consistency with stochastic dominance of risk measures using the quantiles of the distribution of X.

3. Quantile Dominance and the Lorenz Curve. Let us consider the *quantile model* of stochastic dominance [13]. The first quantile function $F_X^{(-1)} : (0, 1] \to \overline{\mathbb{R}}$ corresponding to a real random variable X is defined as the left-continuous inverse of the cumulative distribution function F_X [6]:

$$F_X^{(-1)}(p) = \inf\ \{\eta : F_X(\eta) \geq p\} \quad \text{for } 0 < p \leq 1.$$

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

$$\mathbb{P}\{X < q_X(p)\} \leq p \leq \mathbb{P}\{X \leq q_X(p)\}.$$

For $p \in (0, 1)$ the set of such p-quantiles is a closed interval and $F_X^{(-1)}(p)$ represents its left end [4].

Directly from the definition of FSD we see that

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

Thus, the function $F^{(-1)}$ can be considered as a continuum-dimensional safety measure (negative of a risk measure) within the FSD; using any specific (left) p-quantile

as a scalar safety measure is consistent with the FSD. It is not, however, consistent with the SSD, because it may happen that $X \succeq_{SSD} Y$, but $F_X^{(-1)}(p) < F_Y^{(-1)}(p)$ for some p.

REMARK 2. *Value-at-Risk* (VaR) defined as the maximum loss at a specified confidence level p is a widely used quantile risk measure [20]. It corresponds to the right p-quantile of the random variable X representing gains [2] whereas our dual stochastic dominance model uses the left p-quantile. Nevertheless, the FSD consistency results can be also shown for the right quantile $q_X^r(p) = \sup\{\eta : F_X(\eta) \leq p\}$ (where $p \in [0,1)$), thus justifying the VaR measures.

To obtain quantile measures consistent with the SSD we introduce the *second quantile function* $F_X^{(-2)} : \mathbb{R} \to \bar{\mathbb{R}}$, defined as

$$F_X^{(-2)}(p) = \int_0^p F_X^{(-1)}(\alpha)d\alpha \quad \text{for} \quad 0 < p \leq 1, \tag{3.1}$$

$F_X^{(-2)}(0) = 0$. For completeness, we also set $F_X^{(-2)}(p) = +\infty$ for $p \notin [0,1]$.

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

REMARK 3. The Lorenz curves are used for inequality ordering [1, 6, 16] of positive random variables, relative to their (positive) expectations. Such a Lorenz curve, $L_X(p) = F_X^{(-2)}(p)/\mu_X$, is convex and increasing. The absolute Lorenz curves, though, are not monotone, when negative outcomes occur.

There is an intriguing duality relation between the second quantile function $F_X^{(-2)}$ and the second performance function $F_X^{(2)}$.

THEOREM 3.1. *For every random variable X with $\mathbb{E}|X| < \infty$ we have*

(i) $F_X^{(-2)} = [F_X^{(2)}]^*$*; and*

(ii) $F_X^{(2)} = [F_X^{(-2)}]^*$.

Proof. By the definition of the conjugate function, for every $p \in [0,1]$,

$$[F_X^{(2)}]^*(p) = \sup_\eta\{\eta p - F_X^{(2)}(\eta)\}. \tag{3.2}$$

Thus, by (2.4) and (2.6), $[F_X^{(2)}]^*(0) = 0$ and $[F_X^{(2)}]^*(1) = \mu_X$. For $p \in (0,1)$ the supremum in (3.2) is attained at any η for which $p \in \partial F_X^{(2)}(\eta)$. Recalling the definition (2.1) we see that η is a p-quantile of X, and we can choose $\eta = F_X^{(-1)}(p)$. Therefore, by [21, Thm.23.5(iv)],

$$F_X^{(-1)}(p) \in \partial[F_X^{(2)}]^*(p).$$

This yields the representation

$$[F_X^{(2)}]^*(p) = \int_0^p F_X^{(-1)}(\alpha)\, d\alpha \quad \text{for} \quad p \in (0,1].$$

If $p = 0$, (3.2) yields 0, and for $p \notin [0,1]$ we obtain $+\infty$, as can be seen from Figure 2.1. This proves (i). Assertion (ii) is the consequence of the closedness of $F_X^{(2)}$ and [21, Thm.12.2]. □

While the above result can also be obtained from the Young inequality ([29] and later generalizations), we hope that connections to the convex analysis may prove fruitful.

It follows from Theorem 3.1 that we may fully characterize the SSD relation by using the conjugate function $F_X^{(-2)}$.

THEOREM 3.2. $X \succeq_{SSD} Y \quad \Leftrightarrow \quad F_X^{(-2)}(p) \geq F_Y^{(-2)}(p)$ *for all* $0 \leq p \leq 1$.

Therefore, the properties of $F^{(-2)}$ are of profound importance for stochastic dominance relations.

COROLLARY 3.3. *The following statements are equivalent:*

(i) η *is a p-quantile of* X;
(ii) $\sup_\xi(\xi p - F_X^{(2)}(\xi))$ *is attained at* η;
(iii) $\sup_\alpha(\eta\alpha - F_X^{(-2)}(\alpha))$ *is attained at* p;
(iv) $F_X^{(-2)}(p) + F_X^{(2)}(\eta) = p\eta$.

Proof. Directly from the definitions (2.4) and (3.1), assertion (i) is equivalent to (v) $p \in \partial F_X^{(2)}(\eta)$, and (vi) $\eta \in \partial F_X^{(-2)}(p)$. The equivalence of (ii)–(vi) follows from Theorem 3.1 and [21, Thm.23.5]. □

We can now provide another representation of the second quantile function. Let $p \in (0,1)$ and suppose that η is such that $\mathbb{P}\{X \leq \eta\} = p$. Then by Corollary 3.3(iv) and (2.4),

$$\begin{aligned} F_X^{(-2)}(p) &= p\eta - F_X^{(2)}(\eta) \\ &= p\eta + p\,\mathbb{E}\{X - \eta | X \leq \eta\} = p\,\mathbb{E}\{X | X \leq \eta\}. \end{aligned} \tag{3.3}$$

It facilitates the understanding of the nature of the second quantile function, but cannot serve as a definition because η such that $\mathbb{P}\{X \leq \eta\} = p$ need not exist; (3.1) and Theorem 3.1(i) are precise descriptions.

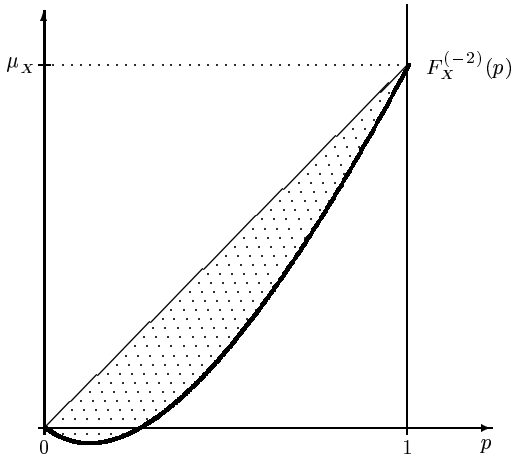

FIG. 3.1. *The absolute Lorenz curve and the dual dispersion space*

Graphical interpretation provides an additional insight into the properties of the second quantile function. 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 corresponds to the chord connecting these points. Hence, the space between the curve $(p, F_X^{(-2)}(p))$, $0 \leq p \leq 1$, and its chord is related to the riskiness of X in comparison to the deterministic outcome of μ_X (Fig. 3.1). 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 X. Nevertheless, it is quite natural to consider some size parameters as summary characteristics of riskiness.

Let us start from the vertical diameter of the dual dispersion space defined as

$$h_X(p) = \mu_X p - F_X^{(-2)}(p). \tag{3.4}$$

LEMMA 3.4. *For every* $p \in (0,1)$

$$h_X(p) = \min_{\xi \in \mathbb{R}} \mathbb{E}\{\max(p(X-\xi), (1-p)(\xi - X))\}, \tag{3.5}$$

and the minimum in the expression above is attained at any p-quantile.

Proof. By Theorem 3.1(i),

$$h_X(p) = \inf_{\xi}((\mu_X - \xi)p + F_X^{(2)}(\xi)).$$

Subdifferentiating with respect to p and using (2.1), we see that the infimum is attained at any p-quantile. From (2.5) we obtain

$$h_X(p) = \min_{\xi}(p\overline{F}_X^{(2)}(\xi) + (1-p)F_X^{(2)}(\xi)).$$

With a view to (2.4) and (2.6),

$$h_X(p) = \min_{\xi}(p\,\mathbb{E}\{\max(0, X-\xi)\} + (1-p)\,\mathbb{E}\{\max(0, \xi - X)\}),$$

which completes the proof. □

The above result reveals a close relation between the vertical dimension of the dual dispersion space and the absolute deviation from the median,

$$\Delta_X = \mathbb{E}\left|X - F_X^{(-1)}(\tfrac{1}{2})\right|.$$

COROLLARY 3.5. $h_X(\frac{1}{2}) = \frac{1}{2}\Delta_X$.

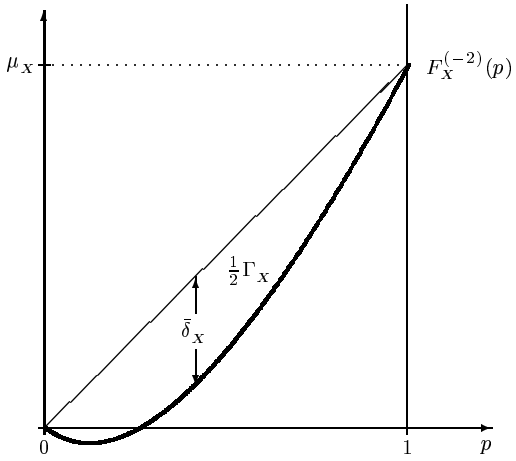

FIG. 3.2. *The absolute Lorenz curve and risk measures*

The maximum vertical diameter of the dual dispersion space (which exists by compactness and continuity) turns out to be the absolute semideviation of X.

LEMMA 3.6. $\max_{p\in[0,1]} h_X(p) = \bar{\delta}_X$
and the maximum is attained at any p_X *for which* $\mathbb{P}\{X < \mu_X\} \le p_X \le \mathbb{P}\{X \le \mu_X\}$.

Proof. By Theorem 3.1(ii),

$$\max_{p\in[0,1]} h_X(p) = \max_{p\in[0,1]} (\mu_X p - F_X^{(-2)}(p)) = F_X^{(2)}(\mu_X)$$

and the the first assertion follows from (2.4) and (2.9). By Corollary 3.3, μ_X is a p_X-quantile. □

It is known that the doubled area of the dual dispersion space,

$$\Gamma_X = 2\int_0^1 (\mu_X p - F_X^{(-2)}(p))\ dp \tag{3.6}$$

is equal to the Gini's mean difference [16]:

$$\Gamma_X = \frac{1}{2}\iint |\eta - \xi|\ P_X(d\xi)\ P_X(d\eta). \tag{3.7}$$

The equality can be verified by calculating the integral in (3.7) over $\xi \le \eta$ (by symmetry) as follows:

$$\begin{aligned}\Gamma_X &= \iint_{\xi\le\eta} \eta\ P_X(d\xi)\ P_X(d\eta) - \iint_{\xi\le\eta} \xi\ P_X(d\xi)\ P_X(d\eta) = 2\int_{-\infty}^{\infty} \eta F(\eta)\ P_X(d\eta) - \mu_X \\ &= 2\int_0^1 pF^{(-1)}(p)\ dp - \mu_X = -2\int_0^1 F^{(-2)}(p)\ dp + \mu_X,\end{aligned}$$

where in the last transformation we employed the integration by parts.

The Gini's mean difference (3.7) may be also expressed as the integral of $F_X^{(2)}$ with respect to the probability measure P_X

$$\Gamma_X = \iint_{\xi\le\eta} (\eta - \xi)\ P_X(d\xi)\ P_X(d\eta) = \int \mathbb{E}\{\max(\eta - X, 0)\}\ P_X(d\eta).$$

Thus, similar to (2.8), it represents the expected shortfall from a random target distributed according to P_X but this distribution is a function of X. Therefore, the corresponding SSD-consistency results (cf. (2.8)) cannot be applied directly to the Gini's mean difference. Alternatively, Γ_X can be expressed with the integral of the vertical diameter of the dispersion space d_X (2.7) with respect to the probability measure P_X

$$\Gamma_X = \bar{\delta}_X + \int d_X(\eta)\ P_X(d\eta).$$

Both Γ and $\bar{\delta}$ are well defined size characteristics of the dual dispersion space (Fig. 3.2). However, the absolute semideviation is a rather rough measure compared 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 $(p_X, F_X^{(-2)}(p_X))$, where $\mathbb{P}\{X < \mu_X\} \le p_X \le \mathbb{P}\{X < \mu_X\}$ (see Lemma 3.6). In fact, $\bar{\delta}_X$ is the Gini's

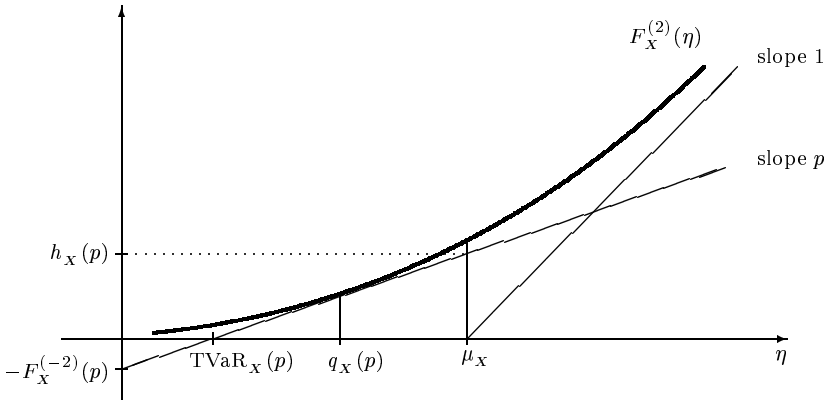

FIG. 3.3. *Dual quantities in the O-R diagram*

mean difference of a two-point distribution approximating X in such a way that μ_X and δ_X remain unchanged.

Dual risk characteristics can also be presented in the (primal) O–R diagram (Fig. 3.3). Recall that $F^{(-2)}$ is the conjugate function of $F^{(2)}$ and therefore, $F^{(-2)}$ describes the affine functions majorized by $F^{(2)}$ [21]. For any $p \in (0.1)$, the line with slope p supports the graph of $F^{(2)}$ at every p-quantile (Corollary 3.3(i),(ii)). It is given analytically as

$$S_X^p(\eta) = p(\eta - q_X(p)) + F_X^{(2)}(q_X(p)),$$

where $q_X(p)$ denotes a p-quantile of X.

From Corollary 3.3(iv) it follows that $F_X^{(-2)}(p) = -S_X^p(0)$, so the value of the absolute Lorenz curve is given by the intersection of the tangent line S_X^p with the vertical (risk) axis. For any $p \in (0,1)$, the tangent line intersects both asymptotes of $F^{(2)}$. It intersects the outcome axis (the left asymptote) at the point $\eta = F_X^{(-2)}(p)/p = \mu_X - h_X(p)/p$ (see (3.4)). In Fig. 3.3 this point is marked as $\mathrm{TVaR}_X(p)$ due to its interpretation discussed in the next section. The intersection with the right asymptote takes place at $\eta = \mu_X + h_X(p)/(1-p)$ (by simple geometry).

Fig. 3.3 provides also an interesting interpretation of Lemma 3.6. By elementary geometry, the tangent line S_X^p intersects the vertical line at $\eta = \mu_X$ at the value $S_X^p(\mu_X) = h_X(p)$ thus defining the vertical diameter of the dual dispersion space at p. This justifies $\bar{\delta}_X = F^{(2)}(\mu_X)$ as the *maximum* vertical diameter of the dual dispersion space.

It might be worth noting that the segment of the line S_X^p restricted to the area between the asymptotes, together with the asymptotes themselves, represents an O–R diagram (the function $F_Y^{(2)}$) for a random variable Y which has a two-point distribution, the same mean as X, and dominates X in the sense of the SSD. It has the largest variance among such variables (with probabilities p and $1-p$ fixed).

4. Dual Risk Measures. From the ALC diagram one can easily derive the following, commonly known, necessary condition for the SSD relation (cf. [12]):

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X \geq \mu_Y. \tag{4.1}$$

But we can get much more.

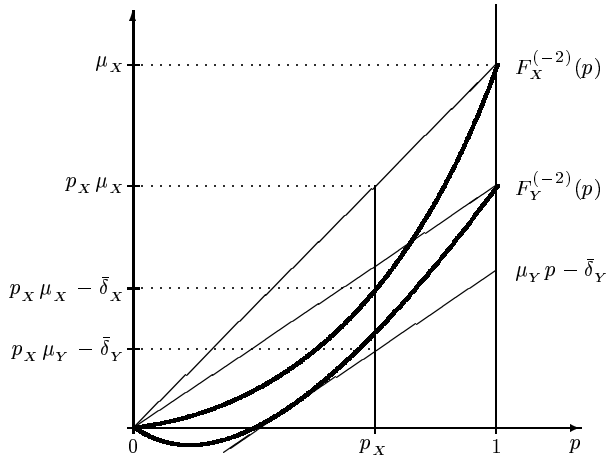

FIG. 4.1. $X \succeq_{SSD} Y \;\Rightarrow\; p_X\,\mu_X - \bar{\delta}_X \geq p_X\,\mu_Y - \bar{\delta}_Y$, *where* $p_X = \mathbb{P}\{X < \mu_X\} < 1$

Consider two random variables X and Y in the common ALC diagram (Fig. 4.1). Since $\bar{\delta}_Y$ represents the maximal vertical diameter of the dual dispersion space for the variable Y, its absolute Lorenz curve $F_Y^{(-2)}(p)$ is bounded from below by the straight line $\mu_Y p - \bar{\delta}_Y$. At the point $p_X = \mathbb{P}\{X < \mu_X\}$ at which $h_X(p_X) = \bar{\delta}_X$ (cf. Lemma 3.6) one gets:

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

This simple analysis of the ALC diagram allows us to derive the following necessary condition for the second degree stochastic dominance.

PROPOSITION 4.1. *If* $X \succeq_{SSD} Y$, *then* $\mu_X \geq \mu_Y$ *and* $\mu_X - \bar{\delta}_X \geq \mu_Y - \bar{\delta}_Y$, *where the second inequality is strict whenever* $\mu_X > \mu_Y$.

Proposition 4.1 was first shown in [18] with the use of the O–R diagram. Here, by placing the considerations within the (dual) ALC diagram we make it transparent that the result is based on the comparison of the absolute Lorenz curves at only one point, p_X. For symmetric random variables we have $p_X \leq 1/2$ and the coefficient in front of $\bar{\delta}$ in Proposition 4.1 can be increased to 2.

The main application of the ALC diagram, though, is the analysis of risk and safety measures using quantiles of the distribution of the random outcome.

Tail Value-at-Risk

The relation in Theorem 3.2 can be rewritten in the form

$$X \succeq_{SSD} Y \quad \Leftrightarrow \quad F_X^{(-2)}(p)/p \geq F_Y^{(-2)}(p)/p \quad \text{for all } 0 < p \leq 1, \tag{4.2}$$

thus justifying the safety measure

$$\mathrm{TVaR}_X(p) = F_X^{(-2)}(p)/p. \tag{4.3}$$

From Theorem 3.2 we immediately obtain the following observation.

PROPOSITION 4.2. *The mean–risk model* $(\mu_X, -\mathrm{TVaR}_X)$ *is consistent with the second degree stochastic dominance relation.*

With a view to (3.3), the quantity $\mathrm{TVaR}_X(p)$ may be interpreted the expected (or tail) VaR measure (see [2, Def. 5.1] and [22])

$$\mathrm{TVaR}_X(F_X(\eta)) = \mathbb{E}\{X|X \le \eta\}.$$

By the convexity of $F^{(-2)}$, the function $\mathrm{TVaR}_X : (0,1] \to \mathbb{R}$ is nondecreasing, continuous, and $\mathrm{TVaR}_X(1) = \mu_X$. In the case of a random variable with lower bounded support, the value of $\mathrm{TVaR}_X(p)$ tends to the minimum outcome when $p \to 0_+$. Hence, the max–min selection rule of [28] is a limiting case of the $(\mu_X, -\mathrm{TVaR}_X)$ model.

It follows from Lemma 3.4 that for every $p \in (0,1)$ the corresponding value $\mathrm{TVaR}_X(p)$ can be computed as

$$\mathrm{TVaR}_X(p) = \mathbb{E}\{X\} - \min_{\xi\in\mathbb{R}} \mathbb{E}\{\max(X-\xi, \frac{1-p}{p}(\xi - X))\}. \tag{4.4}$$

This formula may be transformed into

$$\mathrm{TVaR}_X(p) = \max_{\xi\in\mathbb{R}}(\xi - \frac{1}{p}\mathbb{E}\{\max(0, \xi - X)\}), \tag{4.5}$$

which corresponds to the direct representation of $F^{(-2)}$ as the conjugate function to $F^{(2)}$ (c.f. (3.2)). By Corollary 3.3, the maximum above is attained at any p-quantile. Interestingly, (4.5) appears also in [22] in so-called *Conditional VaR* models; our analysis puts them into the context of stochastic dominance.

Mean absolute deviation from a quantile

Proposition 4.2 allows us to identify an interesting α-consistent risk measure following from the dual characterization of the SSD. Recalling the vertical diameter $h_X(p)$ of the dual dispersion space we have the following result.

PROPOSITION 4.3. *For any $p \in (0,1)$, the mean–risk model $(\mu_X, h_X(p)/p)$ is 1-consistent with the second degree stochastic dominance relation.*

Proof. By Proposition 4.2 and (3.4) we have

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X \ge \mu_Y \quad \text{and} \quad \mu_X - h_X(p)/p \ge \mu_Y - h_Y(p)/p,$$

as required. □

Owing to Lemma 3.4 we may interpret the risk measure $h_X(p)/p$ as the weighted mean absolute deviation from the p-quantile.

For $p = 1/2$, recalling Corollary 3.5, we obtain the following observation (illustrated graphically in Fig. 4.2).

COROLLARY 4.4. *For any $p \in (0,1)$, the mean–risk model (μ_X, Δ_X) is 1-consistent with the second degree stochastic dominance relation.*

Comparing this to Proposition 4.1 we see that we are able to cover both the general and the symmetric case with a higher weight put on the risk term. Indeed, in the symmetric case one has $\Delta_X = 2\bar{\delta}_X$.

Tail Gini's mean difference

Let us now pass to risk measures based on area characteristics of the dual dispersion space. Consider two random variables X and Y in the common ALC diagram (Fig. 4.3). If $X \succeq_{SSD} Y$, then, due to Theorem 3.2, $F_X^{(-2)}$ is bounded from below

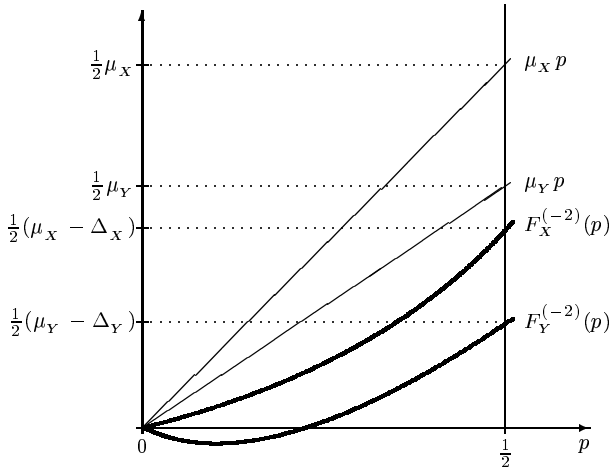

FIG. 4.2. *Median case:* $X \succeq_{SSD} Y \;\Rightarrow\; \frac{1}{2}(\mu_X - \Delta_X) \geq \frac{1}{2}(\mu_Y - \Delta_Y)$

by $F_Y^{(-2)}$, and $\mu_X \geq \mu_Y$ from (4.1). Thus the area of the dual dispersion space for X is (upper) bounded by the area of the dual dispersion space 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 the continuity of the Lorenz curves, this inequality becomes strict whenever $X \succ_{SSD} Y$. This allows us to derive the following necessary conditions for the second degree stochastic dominance.

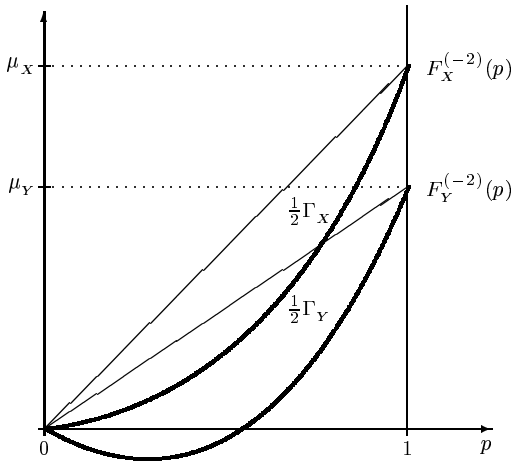

FIG. 4.3. $X \succeq_{SSD} Y \;\Rightarrow\; \frac{1}{2}\Gamma_X \leq \frac{1}{2}\Gamma_Y + \frac{1}{2}(\mu_X - \mu_Y)$

PROPOSITION 4.5. *For integrable random variables X and Y the following implications hold:*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - \Gamma_X \geq \mu_Y - \Gamma_Y, \tag{4.6}$$

$$X \succ_{SSD} Y \quad \Rightarrow \quad \mu_X - \Gamma_X > \mu_Y - \Gamma_Y. \tag{4.7}$$

Condition (4.6) was first shown by Yitzhaki [27] for bounded distributions.

Similarly, for $p \in (0,1]$ one may consider the *tail Gini's* measure:

$$G_X(p) = \frac{2}{p^2} \int_0^p (\mu_X \alpha - F_X^{(-2)}(\alpha)) d\alpha. \tag{4.8}$$

The next result is an obvious extension of Proposition 4.5.

PROPOSITION 4.6. *For every* $p \in (0,1]$,

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - G_X(p) \geq \mu_Y - G_Y(p). \tag{4.9}$$

In other words, the mean–risk model $(\mu_X, G_X(p))$ is 1-consistent with the SSD.

By convexity, $G_X(p) \geq h_X(p)/p$ for all $p \in (0,1]$, so Proposition 4.6 is stronger than Proposition 4.3.

The coefficient 1 in front of $G_X(p)$ (and $G_Y(p)$) cannot be increased for general distributions, but it can be doubled in the case of *symmetric* random variables (and $p = 1$). Indeed, for a symmetric random variable X one has $h_X(p) = h_X(1-p)$, so $G_X(\frac{1}{2}) = 2\Gamma_X$, which leads to the following result.

PROPOSITION 4.7. *For symmetric random variables* X *and* Y *the following implications hold:*

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

5. Mean–Risk Models with Dual Risk Measures. Given a certain set Q of integrable random variables X, let us analyze in more detail the mean–risk optimization problems of form

$$\max_{X \in Q} (\mu_X - \lambda r_X), \tag{5.1}$$

with $\lambda > 0$ and with the risk functional r_X defined as one of our dual (quantile) measures. We assume that the set Q is convex, closed and bounded in $\mathcal{L}_q$ for some $q > 1$.

The first issue that needs to be clarified is the convexity of problem (5.1). This will help to establish the existence of solutions and to formulate computationally tractable models.

LEMMA 5.1. *For every* $p \in [0,1]$ *the functional* $X \to h_X(p)$ *given by (3.4) is convex and positively homogeneous on* $\mathcal{L}_1$.

Proof. Let $\beta \in (0,1)$, $X, Y \in Q$, and let m_X and m_Y be the p-quantiles of X and Y. By Lemma 3.4,

$$\begin{aligned} h_{\beta X + (1-\beta)Y}(p) &= \min_t \mathbb{E} \max \Big\{ p(\beta X + (1-\beta)Y - t), (1-p)(t - \beta X - (1-\beta)Y) \Big\} \\ &\leq \mathbb{E} \max \Big\{ p(\beta(X - m_X) + (1-\beta)(Y - m_Y)), (1-p)(\beta(m_X - X) + (1-\beta)(m_Y - Y)) \Big\}. \end{aligned}$$

Using the inequality $\max(a+b, c+d) \leq \max(a,c) + \max(b,d)$ and Lemma 3.4 again

we obtain

$$\begin{aligned} h_{\beta X+(1-\beta)Y}(p) &\le \beta\mathbb{E}\,\max\Big\{p(X-m_X),(1-p)(m_X-X)\Big\} \\ &\quad+(1-\beta)\mathbb{E}\,\max\Big\{p(Y-m_Y),(1-p)(m_Y-Y)\Big\} \\ &= \beta h_X(p)+(1-\beta)h_Y(p), \end{aligned}$$

because m_X and m_Y are p-quantiles. This proves the convexity. The positive homogeneity follows directly from (3.5). □

For the tail Gini's mean difference used as a risk measure we have a similar result.

LEMMA 5.2. *For every $p\in(0,1]$ the functional $X\to G_X(p)$ given by (4.8) is convex and positively homogeneous on $\mathcal{L}_1$.*

Proof. We have

$$G_X(p)=\frac{2}{p^2}\int_0^p h_X(\alpha)\,d\alpha,$$

and the result follows from Lemma 5.1. □

REMARK 4. Again, the composite objectives of form $\rho(X)=-\mu_X+\alpha r_X$, where $\alpha\in(0,1]$ and r_X is defined as $h_X(p)/p$ or $G_X(p)$, satisfy all axioms of so-called *coherent risk measures* discussed in [2] (cf. Remark 1). The convexity and the positive homogeneity have just been proved, the translation invariance is trivial, and the monotonicity follows from Propositions 4.3 and 4.6, respectively. Indeed, as in Remark 1, $X\ge Y$ a.s. $\Rightarrow$ $X\succeq_{SSD}Y$, and these propositions apply.

Having established the convexity, we can pass now to the analysis of the SSD-efficiency of the solutions to problem (5.1). We start from the case of Gini's mean difference $\Gamma_X=G_X(1)$.

THEOREM 5.3. *Assume that the set Q is convex, bounded and closed in $\mathcal{L}_q$ for some $q>1$, and $r_X=\Gamma_X$. Then for every $\lambda\in(0,1]$, the set of optimal solutions of (5.1) is nonempty and each its element is SSD-efficient in Q.*

Proof. Let us show that the optimal set of (5.1) is nonempty. By Lemma 5.2 the objective functional is concave. In the reflexive Banach space $\mathcal{L}_q$, the set Q is weakly compact (as convex, bounded and closed [10, Thm. 6, p. 179]), and the functional $\mu_X-\lambda\Gamma_X$ is weakly upper semicontinuous (as concave and bounded). Therefore the set of optimal solutions of (5.1) is nonempty.

Let $X\in Q$ be an optimal solution and suppose that X is not SSD-efficient. Then there exists $Z\in Q$ such that $Z\succ_{SSD}X$. From (4.1) and (4.7) we obtain

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

Adding these inequalities multiplied by $(1-\lambda)$ and λ, respectively, we obtain the sharp ($\lambda>0$) inequality $\mu_Z-\lambda\Gamma_Z>\mu_X-\lambda\Gamma_X$. This contradicts the maximality of $\mu_X-\lambda\Gamma_X$. □

Let us now consider the risk measure $r_X=h_X(p)/p$. Recall that, owing to (3.4) and (4.3), the objective in (5.1) can be equivalently expressed as

$$\mu_X-\lambda h_X(p)/p=(1-\lambda)\mu_X+\lambda\,\mathrm{TVaR}_X(p).$$

THEOREM 5.4. *Assume that the set Q is convex, bounded and closed in $\mathcal{L}_q$ for some $q>1$, and $r_X=h_X(p)/p$ with $p\in(0,1)$. Then for every $\lambda\in(0,1]$, the set Q^**

of optimal solutions of (5.1) is nonempty and for each $X \in Q^*$ there exists a point $X^* \in Q^*$ which is SSD-efficient in Q and with $\mu_{X^*} = \mu_X$ and $h_{X^*}(p) = h_X(p)$.

Proof. The proof that the optimal set Q^* of (5.1) is nonempty is the same as in Theorem 5.3. By the convexity of the set Q and the concavity of the objective functional, the set Q^* is convex, closed, and bounded.

Suppose that $X \in Q^*$ is not SSD-efficient. Then there exists $Z \in Q$ such that $Z \succ_{SSD} X$. From (4.1) and Proposition 4.3 we obtain

$$\mu_Z \geq \mu_X \quad \text{and} \quad \mu_Z - h_Z(p)/p \geq \mu_X - h_X(p)/p.$$

Adding these inequalities multiplied by $(1-\lambda)$ and λ, respectively, we obtain

$$\mu_Z - \lambda h_Z(p)/p \geq \mu_X - \lambda h_X(p)/p.$$

Since $Z \in Q$ we must have $Z \in Q^*$ and an equality above. Thus $\mu_Z = \mu_X$ and $h_Z(p) = h_X(p)$.

Define the set $Q^*(X) = \{Z \in Q^* : \mu_Z = \mu_X\}$, and consider the problem

$$\min_{Z \in Q^*(X)} \Gamma_Z. \tag{5.2}$$

The set $Q^*(X)$ is convex, closed and bounded, and (5.2) is equivalent to maximizing $\mu_Z - \lambda\Gamma_Z$. By Theorem 5.3, a solution X^* of (5.2) exists and is SSD-efficient in $Q^*(X)$. It is also SSD-efficient in Q, because we have proved in the preceding paragraph that it cannot be dominated by a point $Z \in Q \setminus Q^*(X)$. By construction, $\mu_{X^*} = \mu_X$ and $h_{X^*}(p) = h_X(p)$, as required. □

Let us now consider the risk measure in the form of the tail Gini's mean difference. Analogously to Theorem 5.4 we obtain the following result.

THEOREM 5.5. *Assume that the set Q is convex, bounded and closed in $\mathcal{L}_q$ for some $q > 1$, and let $r_X = G_X(p)$ with $p \in (0,1)$. Then for every $\lambda \in (0,1]$, the set Q^* of optimal solutions of (5.1) is nonempty and for each $X \in Q^*$ there exists an SSD-efficient point $X^* \in Q^*$ with $\mu_{X^*} = \mu_X$ and $G_{X^*}(p) = G_X(p)$.*

REMARK 5. For symmetric random variables and $p \geq 1/2$, since $h_X(p) = h_X(1-p)$, all optimal solutions are SSD-efficient, as follows from Theorem 5.3. Also, since $G_X(\frac{1}{2}) = 2\Gamma_X$, the coefficient λ in (5.1) can be chosen from $(0,2]$.

6. Stochastic Programming Formulations. Let us formulate a more explicit convex optimization problem which is equivalent to (5.1) with $r_X = h_X(p)/p$:

$$\max \quad \mathbb{E}X - \frac{\lambda}{p}\mathbb{E}V \tag{6.1a}$$

$$\text{subject to} \quad V(\omega) \geq p(X(\omega) - t), \qquad \text{a.s.} \tag{6.1b}$$

$$V(\omega) \geq (1-p)(t - X(\omega)), \quad \text{a.s.} \tag{6.1c}$$

$$X \in Q, \quad V \in \mathcal{L}_1(\Omega), \quad t \in \mathbb{R}. \tag{6.1d}$$

The next result follows from Lemma 3.4.

PROPOSITION 6.1. *Problem (6.1) is equivalent to problem (5.1) with $r_X = h_X(p)/p$ in the following sense:*

(i) *for every solution $\hat{X}$ of (5.1), the triple:*

$$\hat{X}, \quad \hat{t} = F^{(-1)}_{\hat{X}}(p), \quad \hat{V}(\omega) = \max(p(\hat{X}(\omega) - \hat{t}), (1-p)(\hat{t} - \hat{X}(\omega)))$$

is an optimal solution of (6.1);

(ii) *for every optimal solution* $(\hat{X},\hat{t},\hat{V})$ *of (6.1),* $\hat{X}$ *is an optimal solution of (5.1),* $\hat{t}$ *is a p-quantile of* $\hat{X}$*, and* $\mathbb{E}\hat{V}(\omega) = h_{\hat{X}}(p)$.

In particular, if

$$Q = \Big\{ \sum_{i=1}^{n} d_i X_i \,:\, (d_1,\dots,d_n) \in D \Big\}, \tag{6.2}$$

where D is a convex closed polyhedron in $\mathbb{R}^n$ and $X_1,\dots,X_n$ are integrable random variables, we recognize a linear two-stage problem of stochastic programming. In this problem $d \in D$ and $t \in \mathbb{R}$ are first stage variables, while V is the second stage variable. In the case of finitely many realizations $(x_1^j,\dots,x_n^j)$, $j=1,\dots,N$, of $(X_1,\dots,X_n)$, attained with probabilities $\pi_1,\dots,\pi_N$ we obtain the problem:

$$\begin{aligned}
\max \quad & \sum_{j=1}^{N} \pi_j \Big(\sum_{i=1}^{n} d_i x_i^j - \frac{\lambda}{p} v^j \Big) \\
\text{subject to} \quad & v^j \ge p\Big(\sum_{i=1}^{n} d_i x_i^j - t \Big), \qquad j=1,\dots,N, \\
& v^j \ge (1-p)\Big(t - \sum_{i=1}^{n} d_i x_i^j \Big), \quad j=1,\dots,N, \\
& d \in D, \quad v \in \mathbb{R}^N, \quad t \in \mathbb{R}.
\end{aligned}$$

Representing $\sum_{i=1}^{n} d_i x_i^j - t$ as a difference of its positive part u_j and its negative part w_j and eliminating the expectation from the objective, we can transform the last problem to a simple recourse formulation:

$$\begin{aligned}
\max \quad & \Big[t + \sum_{j=1}^{N} \pi_j \Big((1-\lambda) u_j - (1-\lambda+\frac{\lambda}{p}) w_j \Big) \Big] \\
\text{subject to} \quad & \sum_{i=1}^{n} d_i x_i^j - t = u_j - w_j, \quad j=1,\dots,N, \\
& d \in D, \quad u \in \mathbb{R}^N_+, \quad w \in \mathbb{R}^N_+, \quad t \in \mathbb{R}.
\end{aligned}$$

Let us now formulate a stochastic programming problem which is equivalent to (5.1) with $r_X = G_X(p)$:

$$\max \quad \mathbb{E}X - \frac{2\lambda}{p^2} \int_0^p \int V(\alpha,\omega)\ \mathbb{P}(d\omega)\ d\alpha \tag{6.3a}$$

$$\text{subject to} \quad V(\alpha,\omega) \ge \alpha(X(\omega) - t(\alpha)), \qquad \text{a.s. in } [0,p]\times\Omega, \tag{6.3b}$$

$$V(\alpha,\omega) \ge (1-\alpha)(t(\alpha) - X(\omega)), \quad \text{a.s. in } [0,p]\times\Omega, \tag{6.3c}$$

$$X \in Q, \quad V \in \mathcal{L}_1([0,p]\times\Omega), \quad t \in \mathcal{L}_1([0,p]). \tag{6.3d}$$

The product space $[0,p]\times\Omega$ is assumed to be equipped with the product measure of the Lebesgue measure and $\mathbb{P}$.

PROPOSITION 6.2. *Problem (6.3) is equivalent to problem (5.1) with* $r_X = G_X(p)$ *in the following sense:*

(i) *for every solution* $\hat{X}$ *of (5.1), the triple:*

$$\hat{X}, \quad \hat{t}(\alpha) = F^{(-1)}_{\hat{X}}(\alpha), \quad \hat{V}(\alpha,\omega) = \max(\alpha(\hat{X}(\omega) - \hat{t}(\alpha)), (1-\alpha)(\hat{t}(\alpha) - \hat{X}(\omega)))$$

is an optimal solution of (6.3);

(ii) *for every optimal solution* $(\hat{X}, \hat{t}, \hat{V})$ *of (6.3),* $\hat{X}$ *is an optimal solution of (5.1),* $\hat{t}(\alpha)$ *is an* α*-quantile of* $\hat{X}$ *for almost all* $\alpha \in (0,p]$*, and* $\mathbb{E}\hat{V}(\alpha,\omega) = h_{\hat{X}}(\alpha)$ *for almost all* $\alpha \in (0,p]$.

Proof. For $X \in Q$ the quantile $F^{(-1)}_X(\cdot)$ is integrable in $(0,p]$, so restricting t to $\mathcal{L}_1([0,p])$ is allowed. The rest of the proof follows from Lemma 3.4, as in Proposition 6.1. □

In particular, if Q is defined by (6.2) and $(X_1,\dots,X_n)$ is a discrete random vector with N *equally probable* realizations $(x_1^j,\dots,x_n^j)$, $j=1,\dots,N$, we can further simplify this problem. We notice first that $h_X(\alpha)$ is a piecewice linear concave function with break points at k/N, $k=0,\dots,N$. Thus the inequalities (6.3b)–(6.3c) need to be enforced only at the break points. Moreover, the integral in the objective of (6.3) can be calculated exactly by using the values at the break points, by the method of trapezoids.

To be more specific, let m be the smallest integer for which $m/N \geq p$ and let $\alpha_k = k/N$, $k=0,\dots,m-1$, $\alpha_m = p$. We obtain the following two-stage stochastic program:

$$\begin{aligned}
\max \quad & \sum_{j=1}^{N} \pi_j \Big(\sum_{i=1}^{n} d_i x_i^j - \frac{\lambda}{p^2} \sum_{k=0}^{m} (\alpha_{k+1} - \alpha_k)(v_{k+1}^j + v_k^j) \Big) \\
\text{subject to} \quad & v_k^j \geq \alpha_k \Big(\sum_{i=1}^{n} d_i x_i^j - t_k \Big), \qquad j=1,\dots,N, \quad k=0,\dots,m, \\
& v_k^j \geq (1-\alpha_k) \Big(t_k - \sum_{i=1}^{n} d_i x_i^j \Big), \quad j=1,\dots,N, \quad k=0,\dots,m, \\
& d \in D, \quad v \in \mathbb{R}^N \times \mathbb{R}^{m+1}, \quad t \in \mathbb{R}^{m+1}.
\end{aligned}$$

In the above problem v_k^j represents the value of $V(\alpha_k)$ in the jth realization, and $t_k = t(\alpha_k)$. Similarly to problem (6.1), the last problem can also be transformed to a simple recourse formulation.

If the probabilities π_j of realizations of $(X_1,\dots,X_n)$ are *not* equal, though, the break points may depend on our decisions, and the reduction to the finite dimensional case is harder. One possibility is to introduce such a grid that contains *all* possible break points, but it may be unnecessarily numerous. Another possibility is to resort to an approximation with some reasonably chosen grid α_k, $k=1,\dots,m$. It will be a relaxation because $h(\cdot)$ is a concave function.

For $p = 1$ all these complications disappear, because the alternative defini-

tion (3.7) of Γ_X has an obvious linear programming representation:

$$\begin{aligned}
\max \quad & \left[\sum_{j=1}^{N}\pi_j\sum_{i=1}^{n}d_i x_i^j - \lambda\sum_{j=1}^{N}\sum_{l=j+1}^{N}\pi_j\pi_l v^{jl}\right] \\
\text{subject to} \quad & v^{jl} \geq \sum_{i=1}^{n} d_i(x_i^j - x_i^l), \quad j=1,\ldots,N, \quad l=j+1,\ldots,N, \\
& v^{jl} \geq \sum_{i=1}^{n} d_i(x_i^l - x_i^j), \quad j=1,\ldots,N, \quad l=j+1,\ldots,N, \\
& d \in D, \quad v \in \mathbb{R}^{N(N-1)/2}.
\end{aligned}$$

It has a much larger number of variables and constraints, though.

All finite dimensional stochastic programing models of this section can be solved by specialized decomposition methods [24].

7. Conclusions. We have defined dual relations of stochastic dominance for arbitrary random variables with finite expectations. The second degree stochastic dominance can be expressed as a relation of conjugate functions to second order performance functions.

By using concepts and methods of convex analysis and optimization theory, we have identified several security and risk measures, which can be employed in mean–risk decision models: *tail Value-at-Risk*,

$$\mathrm{TVaR}_X(p) = q_X(p) - \frac{1}{p}\mathbb{E}\{\max(0, q_X(p) - X)\},$$

where $q_X(p)$ is a p-quantile, *weighted mean deviation from a quantile*,

$$h_X(p) = \mathbb{E}\{\max(p(X - q_X(p)), (1-p)(q_X(p) - X))\},$$

and *tail Gini's mean difference*,

$$G_X(p) = \frac{2}{p^2}\int_0^p h_X(\alpha)\, d\alpha.$$

We have shown that the mean–risk models using these measures: $(\mu_X, -\mathrm{TVaR}_X(p))$, $(\mu_X, h_X(p))$, and $(\mu_X, G_X(p))$ are consistent with the second degree stochastic dominance relation (in the sense of Definition 2.1 for $\mathrm{TVaR}_X(p)$ and Definition 2.2 for the other two measures). In particular, the optimal solutions of the corresponding mean–risk models, if unique, are efficient under the second degree stochastic dominance relation.

Finally, we have found stochastic linear programming formulations of these models. This opens a new area of applications of the theory and methods of stochastic programming.

REFERENCES

[1] B. C. Arnold, *Majorization and the Lorenz Order: A Brief Introduction*, Lecture Notes in Statistics 43, Springer-Verlag, Berlin, 1980.

[2] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, *Coherent measures of risk*, Mathematical Finance, 9 (1999), pp. 203–228.

[3] V. S. Bawa, *Stochastic dominance: a research bibliography*, Management Science, 28 (1982), pp. 698–712.

[4] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, New York, 1997.

[5] P. C. Fishburn, *Decision and Value Theory*, John Wiley & Sons, New York, 1964.

[6] J. L. Gastwirth, *A general definition of the Lorenz curve*, Econometrica, 39 (1971), pp. 1037–1039.

[7] J. Y. Gotoh and H. Konno, *Third degree stochastic dominance and mean–risk analysis*, Management Science, 46 (2000), pp. 289–301.

[8] G. Hanoch and H. Levy, *The efficiency analysis of choices involving risk*, Rev. Economic Studies, 36 (1969), pp. 335–346.

[9] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, MA, 1934.

[10] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.

[11] H. Konno and H. Yamazaki, *Mean–absolute deviation portfolio optimization model and its application to Tokyo stock market*, Management Science, 37 (1991), pp. 519–531.

[12] H. Levy, *Stochastic dominance and expected utility: survey and analysis*, Management Science, 38 (1992), pp. 555–593.

[13] H. Levy and Y. Kroll, *Ordering uncertain options with borrowing and lending*, Journal of Finance, 33 (1978), pp. 553–573.

[14] H. M. Markowitz, *Portfolio selection*, Journal of Finance, 7 (1952), pp. 77–91.

[15] ———, *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford, 1987.

[16] P. Muliere and M. Scarsini, *A note on stochastic dominance and inequality measures*, Journal of Economic Theory, 49 (1989), pp. 314–323.

[17] W. Ogryczak, *Multiple criteria linear programming model for portfolio selection*, Annals of Operations Research, to appear. (Technical Report TR 97-06 (243), Institute of Informatics, Warsaw University, 1997).

[18] W. Ogryczak and A. Ruszczyński, *From stochastic dominance to mean–risk models: semideviations as risk measures*, European Journal of Operational Research, 116 (1999), pp. 33–50.

[19] ———, *On consistency of stochastic dominance and mean–semideviation models*, Mathematical Programming, 89 (2001), pp. 217–232.

[20] RiskMetrics, Technical document, J.P.Morgan Inc., New York, 1996.

[21] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, NJ, 1970.

[22] R. T. Rockafellar and S. Uryasev, *Optimization of Conditional Value-at-Risk*, Journal of Risk, 2 (2000), pp. 21–41.

[23] M. Rothschild and J. E. Stiglitz, *Increasing risk: I. A definition*, Journal of Economic Theory, 2 (1969), pp. 225–243.

[24] A. Ruszczyński, *Decomposition methods in stochastic programming*, Mathematical Programming, 79 (1997), pp. 333–353.

[25] W. F. Sharpe, *A linear programming approximation for the general portfolio analysis problem*, Journal of Financial and Quantitative Analysis, 6 (1971), pp. 1263–1275.

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

[27] S. Yitzhaki, *Stochastic dominance, mean variance, and Gini's mean difference*, American Economic Review, 72 (1982), pp. 178–185.

[28] M. R. Young, *A minimax portfolio selection rule with linear programming solution*, Management Science, 44 (1998), pp. 673–683.

[29] W. H. Young, *On classes of summable functions and their Fourier series*, Proceedings of the Royal Society of London, A87 (1912), pp. 235–239.