

## DUAL STOCHASTIC DOMINANCE AND RELATED MEAN-RISK MODELS\*

WŁODZIMIERZ OGRYCZAK<sup>†</sup> AND ANDRZEJ RUSZCZYŃSKI<sup>‡</sup>

**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

**PII.** S1052623400375075

**1. Introduction.** The relation of *stochastic dominance* is one of the fundamental concepts of decision theory (cf. [32, 14]). It introduces a partial order in the space of real random variables. While theoretically attractive, stochastic dominance order is computationally very difficult, as it involves a multiobjective model with a continuum of objectives.

The practice of decision making under uncertainty frequently resorts to *mean-risk models* (cf. [18]). 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 model [17] uses the variance as the risk measure in the mean-risk analysis. Since its introduction, many authors have pointed out that the mean-variance model is, in general, not consistent with stochastic dominance rules. In our preceding paper [22] we have proved that the standard semideviation (square root of the semivariance) or the mean absolute deviation (from the mean) as 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, 23], 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 [31] work on a linear programming (LP) approxima-

---

\*Received by the editors July 10, 2000; accepted for publication (in revised form) November 22, 2001; published electronically May 15, 2002.

<http://www.siam.org/journals/siopt/13-1/37507.html>

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

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

tion to the mean-variance model, many attempts have been made to linearize the portfolio optimization problem. This resulted in the consideration of various risk measures which were LP computable in the case of finite discrete random variables. Yitzhaki [33] introduced a mean-risk model using the Gini mean (absolute) difference as a risk measure. Konno and Yamazaki [12] analyzed a model in which risk is measured by the (mean) absolute deviation. Young [34] considered the minimax approach (the worst case performances) to measure risk. If the rates of return are multivariate and 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, as opposed to the mean-variance approach, they can be applied to general (possibly nonsymmetric) 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 [21] based on majorization theory [10, 19] and Lorenz-type orders [16, 1].

In this paper we analyze a *dual* model of stochastic dominance by exploiting duality relations of convex analysis (see, e.g., [27]). These transformations allow us to show the 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 LP problems, thus opening a new area of applications for stochastic programming.

The paper is organized as follows. In section 2 we formally define stochastic dominance relations and the concept of the 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 section 4 we consider several risk measures based on quantiles and tail characteristics of the distribution, and we analyze their relation to stochastic dominance. Section 5 is devoted to the analysis of mean-risk models using these risk measures. In section 6 we present stochastic LP formulations of these models. Finally, we draw some conclusions in section 7.

We use  $(\Omega, \mathcal{B}, \mathbb{P})$  to denote an abstract probability space. For a random variable  $X : \Omega \rightarrow \mathbb{R}$ , we denote by  $P_X$  the measure induced by it on the real line. For a convex function  $F : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ , we denote by  $F^*$  its convex conjugate (see [27]),  $F^*(p) = \sup_{\xi} \{p\xi - F(\xi)\}$ .

**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 [10, 19] for the discrete case and was later extended to general distributions [25, 8, 9, 29]. Since that time it has been widely used in economics and finance (see [3, 14] for numerous references).

In the stochastic dominance approach, random variables are compared by point-wise comparison of some performance functions constructed from their distribution functions. For a real random variable  $X$ , its first performance function is defined as the right-continuous cumulative distribution function itself:

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

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 (see [13, 25]):

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

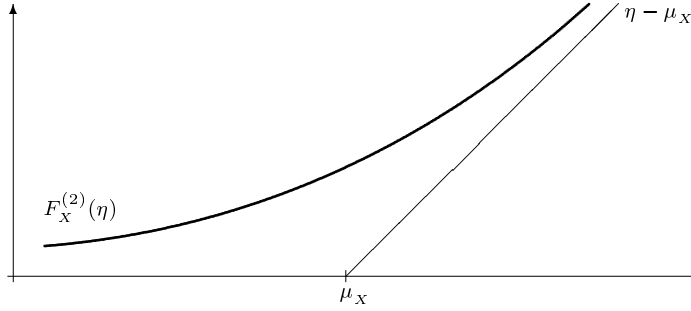


FIG. 2.1. The O-R diagram.

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

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

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

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

(see [8, 9]). The corresponding strict dominance relations  $\succ_{FSD}$  and  $\succ_{SSD}$  are defined by the standard rule

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

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.

Stochastic dominance relations are of crucial importance for decision theory. It is known that  $X \succeq_{FSD} Y$  if and only if  $\mathbb{E}U(X) \geq \mathbb{E}U(Y)$  for any nondecreasing function  $U(\cdot)$  for which these expected values are finite. Also,  $X \succeq_{SSD} Y$  if and only if  $\mathbb{E}U(X) \geq \mathbb{E}U(Y)$  for every nondecreasing and concave  $U(\cdot)$  for which these expected values are finite (see, e.g., [14]).

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. As mentioned above, 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 (see [22]): for each target value  $\eta$  we have

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

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  $\mu_x$  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  $\mu_x$ . It is referred to as the *primal dispersion space*.

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

$$(2.5) \quad \bar{F}_x^{(2)}(\eta) = F_x^{(2)}(\eta) - (\eta - \mu_x),$$

which can be rewritten as

$$(2.6) \quad \begin{aligned} \bar{F}_x^{(2)}(\eta) &= \int_{\eta}^{\infty} (\xi - \eta) P_x(d\xi) \\ &= \mathbb{E}\{\max(X - \eta, 0)\} = \mathbb{P}\{X \geq \eta\} \mathbb{E}\{X - \eta | X \geq \eta\}, \end{aligned}$$

thus expressing the expected surplus for each target outcome  $\eta$  (see [22]). The vertical diameter of the primal dispersion space at a point  $\eta$  is given as

$$(2.7) \quad d_x(\eta) = \min(F_x^{(2)}(\eta), \bar{F}_x^{(2)}(\eta)).$$

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 nonnegative functional  $r_x$ , we can introduce the following definition.

DEFINITION 2.1. *We say that the mean-risk model  $(\mu_x, r_x)$  is consistent with SSD if the following relation holds:*

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_x \geq \mu_y \quad \text{and} \quad r_x \leq r_y.$$

It is known that the first inequality on the right-hand side is true:  $X \succeq_{SSD} Y \Rightarrow \mu_x \geq \mu_y$  (see [14]). 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  $P_T$ ,

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

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 [23], we relax it by considering a scalarization of the

partial order in the  $(\mu_x, r_x)$  space. This will allow us to derive new necessary conditions of dominance, which will make searching for an SSD-efficient solution a more tractable task.

DEFINITION 2.2. *We say that the mean-risk model  $(\mu_x, r_x)$  is  $\alpha$ -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  $\alpha$ -consistency implies  $\lambda$ -consistency for all  $0 \leq \lambda \leq \alpha$ .

The concept of  $\alpha$ -consistency turned out to be fruitful. In [22] we have proved that the mean-risk model in which the risk is defined as the *absolute semideviation*,

$$(2.9) \quad \bar{\delta}_x = r_x^{\mu_x} = \mathbb{E}\{\max(\mu_x - X, 0)\} = \int_{-\infty}^{\mu_x} (\mu_x - \xi) P_x(d\xi),$$

is 1-consistent with SSD. An identical result (under the condition of finite second moments) has been obtained in [22] for the *standard semideviation*,

$$(2.10) \quad \bar{\sigma}_x = \left( \mathbb{E}\{(\max(\mu_x - X, 0))^2\} \right)^{1/2} = \left( \int_{-\infty}^{\mu_x} (\mu_x - \xi)^2 P_x(d\xi) \right)^{1/2}.$$

These results have been further extended in [23] 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 the form  $\rho(X) = -\mu_x + \alpha r_x$  (note the opposite scalarization via 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.  $\Rightarrow X \succeq_{SSD} Y$ .

Our objective is to analyze risk measures using the quantiles of the distribution of  $X$  which are consistent with stochastic dominance.

**3. Quantile dominance and the Lorenz curve.** Let us consider the *quantile model* of stochastic dominance [15]. The first quantile function  $F_x^{(-1)} : (0, 1] \rightarrow \overline{\mathbb{R}}$ , corresponding to a real random variable  $X$ , is defined as the left-continuous inverse of the cumulative distribution function  $F_x$  (see [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  $q = q_x(p)$  is called a *p-quantile* of the random variable  $X$  if

$$\mathbb{P}\{X < q\} \leq p \leq \mathbb{P}\{X \leq q\}.$$

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

$$(3.1) \quad X \succeq_{FSD} Y \iff F_X^{(-1)}(p) \geq F_Y^{(-1)}(p) \quad \text{for all } 0 < p \leq 1.$$

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 [26]. 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*  $F_X^{(-2)} : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ , defined as

$$(3.2) \quad F_X^{(-2)}(p) = \int_0^p F_X^{(-1)}(\alpha) d\alpha \quad \text{for } 0 < p \leq 1,$$

$F_X^{(-2)}(0) = 0$ . 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 (ALC) diagram*.

*Remark 3.* The Lorenz curves are used for inequality ordering [1, 6, 20] 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 ALCs, though, are not monotone when negative outcomes occur.

Directly from (2.4), using the right-continuity of  $F(\cdot)$ , we obtain

$$(3.3) \quad \partial F_X^{(2)}(\eta) = [\mathbb{P}\{X < \eta\}, \mathbb{P}\{X \leq \eta\}].$$

This allows us to develop a Fenchel 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  $p \in [0, 1]$ ,

$$(3.4) \quad [F_X^{(2)}]^*(p) = \sup_{\eta} \{ \eta p - F_X^{(2)}(\eta) \}.$$

From (2.4) it is evident that  $[F_X^{(2)}]^*(0) = 0$  and  $[F_X^{(2)}]^*(1) = \mu_X$ . For  $p \in (0, 1)$  the supremum in (3.4) is attained at any  $\eta$  for which  $p \in \partial F_X^{(2)}(\eta)$ . By (3.3),  $\eta$  is a  $p$ -quantile of  $X$ , and we can choose  $\eta = F_X^{(-1)}(p)$ . Therefore, by [27, Thm. 23.5(iv)],

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

This yields the representation

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

If  $p = 0$ , then (3.4) 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 [27, Thm. 12.2].  $\square$

While the above result can also be obtained from the Young inequality ([35] and later generalizations), we hope that connections to convex analysis may prove fruitful.

It follows from Theorem 3.1 that we may fully characterize the SSD relation by using the conjugate function  $F_x^{(-2)}$ , similarly to the relation (3.1) for FSD.

**THEOREM 3.2.**  $X \succeq_{SSD} Y \Leftrightarrow F_x^{(-2)}(p) \geq F_y^{(-2)}(p)$  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)  $\eta$  is a  $p$ -quantile of  $X$ ;
- (ii)  $\sup_{\xi}(\xi p - F_x^{(2)}(\xi))$  is attained at  $\eta$ ;
- (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 definitions (2.4) and (3.2), 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 [27, Thm. 23.5].  $\square$

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

$$(3.5) \quad \begin{aligned} F_x^{(-2)}(p) &= p\eta - F_x^{(2)}(\eta) \\ &= p\eta + p\mathbb{E}\{X - \eta | X \leq \eta\} = p\mathbb{E}\{X | X \leq \eta\}. \end{aligned}$$

The last relation facilitates the understanding of the nature of the second quantile function, but cannot serve as a definition because  $\eta$  such that  $\mathbb{P}\{X \leq \eta\} = p$  need not exist; (3.2) and Theorem 3.1(i) are precise descriptions.

Graphical interpretation provides an additional insight into the properties of the second quantile function. For any uncertain outcome  $X$ , its ALC  $F_x^{(-2)}$  is a continuous convex curve connecting points  $(0, 0)$  and  $(1, \mu_x)$ , whereas a deterministic outcome with the same expected value  $\mu_x$  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  $\mu_x$  (Figure 3.1). We shall call it the *dual dispersion space*.

Both the size and the shape of the dual dispersion space are important for complete description of the riskiness of  $X$ . We shall use its size parameters as summary characteristics of riskiness.

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

$$(3.6) \quad h_x(p) = \mu_x p - F_x^{(-2)}(p).$$

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

$$(3.7) \quad h_x(p) = \min_{\xi \in \mathbb{R}} \mathbb{E}\{\max(p(X - \xi), (1 - p)(\xi - X))\},$$

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

*Proof.* By Theorem 3.1(i),

$$h_x(p) = \inf_{\xi} ((\mu_x - \xi)p + F_x^{(2)}(\xi)).$$

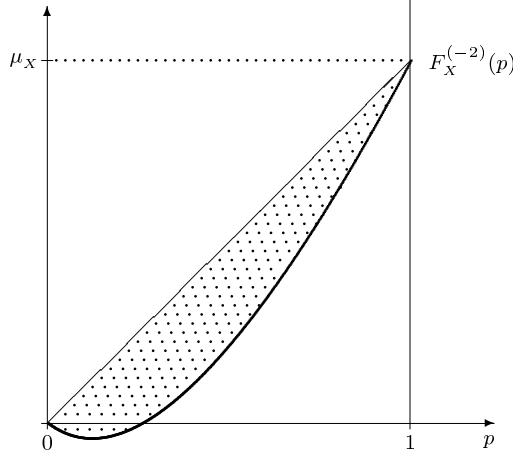


FIG. 3.1. The ALC and the dual dispersion space.

Subdifferentiating with respect to  $\xi$  and using (3.3), we see that the infimum is attained at any  $p$ -quantile. From (2.5) we obtain

$$h_x(p) = \min_{\xi} \left( p \bar{F}_x^{(2)}(\xi) + (1-p) F_x^{(2)}(\xi) \right).$$

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.  $\square$

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

$$\Delta_x = \mathbb{E} \left| X - F_x^{(-1)} \left( \frac{1}{2} \right) \right|.$$

COROLLARY 3.5.  $h_x(\frac{1}{2}) = \frac{1}{2} \Delta_x$ .

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.  $\max_{p \in [0,1]} h_x(p) = \delta_x$ , and the maximum is attained at any  $p_x$  for which  $\mathbb{P}\{X < \mu_x\} \leq p_x \leq \mathbb{P}\{X \leq \mu_x\}$ .

*Proof.* By Theorem 3.1(ii),

$$\max_{p \in [0,1]} h_x(p) = \max_{p \in [0,1]} (\mu_x p - F_x^{(-2)}(p)) = F_x^{(2)}(\mu_x),$$

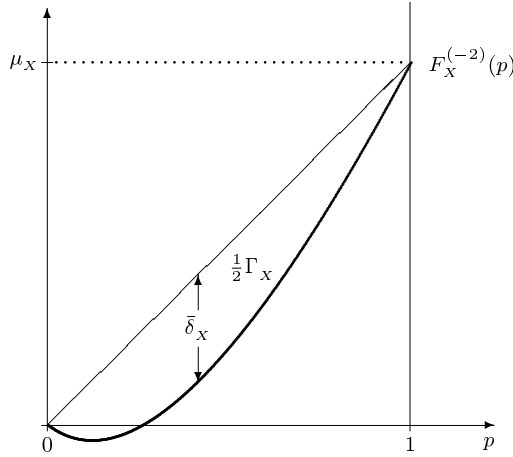
and the first assertion follows from (2.4) and (2.9). Now by Corollary 3.3,  $\mu_x$  is a  $p_x$ -quantile.  $\square$

If the distribution is symmetric, then  $p_x = 1/2$  is a maximizer, and we have  $\max_{p \in [0,1]} h_x(p) = h_x(\frac{1}{2}) = \frac{1}{2} \Delta_x = \bar{\delta}_x$ .

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

$$(3.8) \quad \Gamma_x = 2 \int_0^1 (\mu_x p - F_x^{(-2)}(p)) dp,$$



FIG. 3.2. *The ALC and risk measures.*

is equal to the Gini mean difference (see [20]):

$$(3.9) \quad \Gamma_x = \frac{1}{2} \iint |\eta - \xi| P_x(d\xi) P_x(d\eta).$$

The Gini mean difference (3.9) may be also expressed as the integral of  $F_x^{(2)}$  with respect to the probability measure  $P_x$ :

$$\Gamma_x = \iint_{\xi \leq \eta} (\eta - \xi) P_x(d\xi) P_x(d\eta) = \int \mathbb{E}\{\max(\eta - X, 0)\} P_x(d\eta).$$

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 mean difference.

Both  $\Gamma$  and  $\bar{\delta}$  are well-defined size characteristics of the dual dispersion space (Figure 3.2). However, the absolute semideviation is a rather rough measure compared to the Gini 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  $\mathbb{P}\{X < \mu_x\} \leq p_x \leq \mathbb{P}\{X \leq \mu_x\}$  (see Lemma 3.6). In fact,  $\bar{\delta}_x$  is the Gini mean difference of a two-point distribution approximating  $X$  in such a way that  $\mu_x$  and  $\delta_x$  remain unchanged.

Dual risk characteristics can also be presented in the (primal) O-R diagram (Figure 3.3). Recall that  $F^{(-2)}$  is the conjugate function of  $F^{(2)}$ , and therefore  $F^{(-2)}$  describes the affine functions majorized by  $F^{(2)}$  [27]. 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  $F_x^{(-2)}(p) = -S_x^p(0)$ , and thus the value of the ALC is given by the intersection of the tangent line  $S_x^p$  with the vertical (risk)

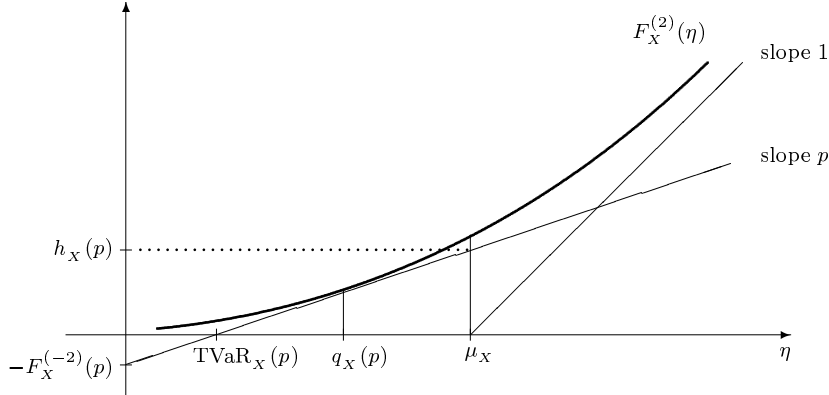


FIG. 3.3. Dual quantities in the O-R diagram.

axis. For any  $p \in (0, 1)$ , the tangent line intersects the outcome axis at the point  $\eta = F_X^{(-2)}(p)/p = \mu_X - h_X(p)/p$  (see (3.6)). In Figure 3.3 this point is marked as  $\text{TVaR}_X(p)$  due to its interpretation discussed in the next section.

Figure 3.3 also provides 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.

**4. Dual risk measures.** From the ALC diagram one can easily derive the following, commonly known, necessary condition for the SSD relation (cf. [14]):

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

But we can get much more.

Consider two random variables  $X$  and  $Y$  such that  $X \succeq_{SSD} Y$  in the common ALC diagram (Figure 4.1). Since  $\bar{\delta}_Y$  represents the maximal vertical diameter of the

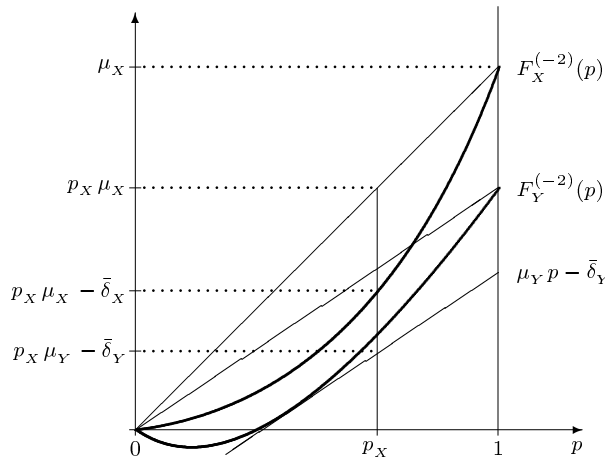


FIG. 4.1.  $X \succeq_{SSD} Y \Rightarrow p_X \mu_X - \bar{\delta}_X \geq p_X \mu_Y - \bar{\delta}_Y$ , where  $p_X = \mathbb{P}\{X < \mu_X\} < 1$ .

dual dispersion space for the variable  $Y$ , its ALC  $F_Y^{(-2)}(p)$  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 SSD.

**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 [22] with the use of an 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 ALCs 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 VaR.* The relation in Theorem 3.2 can be rewritten in the form

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

thus justifying the safety measure

$$(4.3) \quad \text{TVaR}_x(p) = F_X^{(-2)}(p)/p.$$

From Theorem 3.2 we immediately obtain the following observation.

**PROPOSITION 4.2.** *The mean-risk model  $(\mu_X, -\text{TVaR}_X)$  is consistent with the SSD relation.*

In light of (3.5), the quantity  $\text{TVaR}_x(p)$  may be interpreted as the expected (or tail) VaR measure (see [2, Def. 5.1] and [28]):

$$\text{TVaR}_x(F_x(\eta)) = \mathbb{E}\{X|X \leq \eta\}.$$

By the convexity of  $F^{(-2)}$ , the function  $\text{TVaR}_x : (0, 1] \rightarrow \mathbb{R}$  is nondecreasing, continuous, and  $\text{TVaR}_x(1) = \mu_x$ . In the case of a lower bounded random variable, the value of  $\text{TVaR}_x(p)$  tends to the infimum of the outcomes when  $p \rightarrow 0_+$ . Hence, the max-min selection rule of [34] is a limiting case of the  $(\mu_x, -\text{TVaR}_x)$  model.

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

$$(4.4) \quad \text{TVaR}_x(p) = \mathbb{E}\{X\} - \min_{\xi \in \mathbb{R}} \mathbb{E} \left\{ \max \left( X - \xi, \frac{1-p}{p}(\xi - X) \right) \right\}.$$

This formula may be transformed into

$$(4.5) \quad \text{TVaR}_x(p) = \max_{\xi \in \mathbb{R}} \left( \xi - \frac{1}{p} \mathbb{E} \{ \max(0, \xi - X) \} \right),$$

which corresponds to the direct representation of  $F^{(-2)}$  as the conjugate function to  $F^{(2)}$  (c.f. (3.4)). By Corollary 3.3, the maximum above is attained at any  $p$ -quantile.

Interestingly, (4.5) also appears in [28] in so-called *conditional VaR* models; our analysis puts them into the context of dual stochastic dominance. An alternative proof of the consistency of conditional VaR with SSD has been given in [24].

*Mean absolute deviation from a quantile.* Proposition 4.2 allows us to identify an interesting  $\alpha$ -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 SSD relation.*

*Proof.* By Proposition 4.2 and (3.6) we have

$$X \succeq_{SSD} Y \Rightarrow \mu_x \geq \mu_y \quad \text{and} \quad \mu_x - h_x(p)/p \geq \mu_y - h_y(p)/p,$$

as required.  $\square$

Because of 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 Figure 4.2).

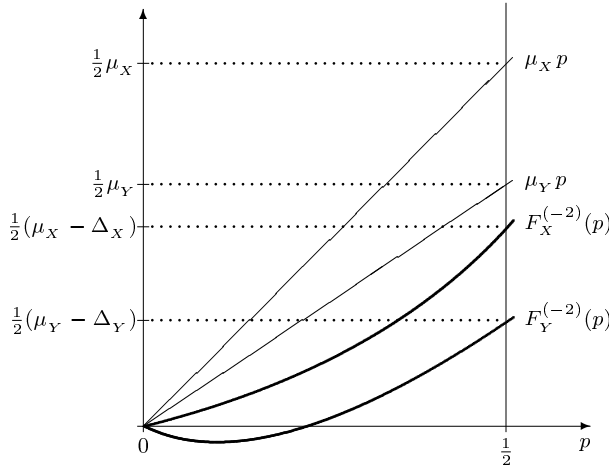


FIG. 4.2. *Median case:*  $X \succeq_{SSD} Y \Rightarrow \frac{1}{2}(\mu_x - \Delta_x) \geq \frac{1}{2}(\mu_y - \Delta_y)$ .

COROLLARY 4.4. *The mean-risk model  $(\mu_x, \Delta_x)$  is 1-consistent with the SSD 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 we have  $\Delta_x = 2\bar{\delta}_x$ .

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

PROPOSITION 4.5. *For integrable random variables  $X$  and  $Y$  the following im-*

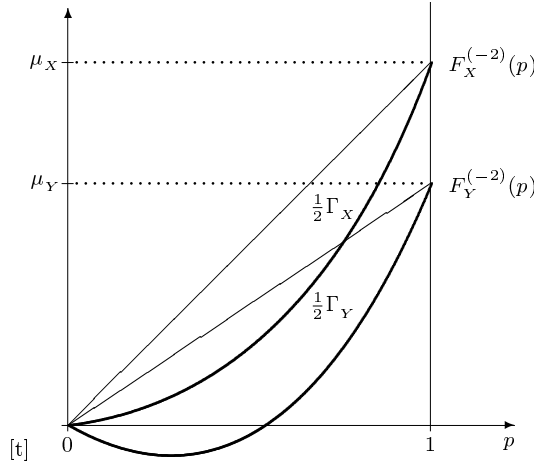


FIG. 4.3.  $X \succeq_{SSD} Y \Rightarrow \frac{1}{2}\Gamma_X \leq \frac{1}{2}\Gamma_Y + \frac{1}{2}(\mu_X - \mu_Y)$ .

lications hold:

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

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

Condition (4.6) was first shown by Yitzhaki [33] for bounded distributions. Similarly, for  $p \in (0, 1]$  one may consider the *tail Gini* measure:

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

The next result is an obvious extension of Proposition 4.5.

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

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

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  $h_X(p) = h_X(1 - p)$ , and thus  $G_X(\frac{1}{2}) = 2\Gamma_X$ , which leads to the following result.

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

$$X \succeq_{SSD} Y \quad \Rightarrow \quad \mu_X - 2\Gamma_X \geq \mu_Y - 2\Gamma_Y,$$

$$X \succ_{SSD} Y \quad \Rightarrow \quad \mu_X - 2\Gamma_X > \mu_Y - 2\Gamma_Y.$$

**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

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

with  $\lambda > 0$  and with 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  $X \rightarrow h_x(p)$  given by (3.6) 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,

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

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

$$\begin{aligned} h_{\beta X + (1-\beta)Y}(p) &\leq \beta \mathbb{E} \max \{p(X - m_x), (1-p)(m_x - X)\} \\ &\quad + (1-\beta) \mathbb{E} \max \{p(Y - m_y), (1-p)(m_y - Y)\} \\ &= \beta h_x(p) + (1-\beta) h_y(p), \end{aligned}$$

because  $m_x$  and  $m_y$  are  $p$ -quantiles. This proves the convexity. The positive homogeneity follows directly from (3.7).  $\square$

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

LEMMA 5.2. *For every  $p \in (0, 1]$  the functional  $X \rightarrow G_x(p)$  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.  $\square$

*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 the so-called *coherent risk measures* discussed in [2] (cf. Remark 1). The convexity and 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,  $X \geq Y$  a.s.  $\Rightarrow X \succeq_{SSD} Y$ , and these propositions apply.

Having established 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 the Gini 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 of its elements 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 [11, 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

$$\mu_z \geq \mu_x \quad \text{and} \quad \mu_z - \Gamma_z > \mu_x - \Gamma_x.$$

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

Let us now consider the risk measure  $r_X = h_X(p)/p$ . Recall that, owing to (3.6) and (4.3), the objective in (5.1) can be equivalently expressed as

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

**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 that 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  $\lambda$ , 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

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

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.  $\square$

Let us now consider the risk measure in the form of the tail Gini 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  $\lambda$  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  $r_X = h_X(p)/p$ :

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

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

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

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

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*

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

*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  $\mathbb{E}\hat{V}(\omega) = h_x(p)$ .*

In particular, if

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

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

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

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:

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

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

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

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

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

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



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*

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

*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  $\alpha$ -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  $F_x^{(-1)}(\cdot)$  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.  $\square$

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 piecewise 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:

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

In the above problem,  $v_k^j$  represents the value of  $V(\alpha_k)$  in the  $j$ th 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  $\pi_j$  of the realizations of  $(X_1, \dots, X_n)$  are *not* equal, however, the break points may depend on our decisions, and the reduction to the finite dimensional case is harder. One way around this difficulty is to repeat the outcomes (as many times as needed) to ensure this property (in the case of rational probabilities). This, however, may dramatically increase the size of the problem. Another possibility is to introduce such a grid that contains *all* possible break points, but it may be unnecessarily large. Yet another possibility is to resort to an approximation with some reasonably chosen grid  $\alpha_k$ ,  $k = 1, \dots, m$ . This would be a relaxation because  $h(\cdot)$  is a concave function.

For  $p = 1$  all these complications disappear, because the alternative definition (3.9) of  $\Gamma_x$  has an obvious LP representation:

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

This has a much larger number of variables and constraints, however.

All finite dimensional stochastic programming models of this section can be solved by specialized decomposition methods [30].

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

By using the 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*,

$$\text{TVaR}_x(p) = q_x(p) - \frac{1}{p} \mathbb{E} \{ \max(0, q_x(p) - X) \},$$

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 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— $(\mu_x, -\text{TVaR}_x(p))$ ,  $(\mu_x, h_x(p))$ , and  $(\mu_x, G_x(p))$ —are consistent with the SSD relation (in the sense of Definition 2.1 for  $\text{TVaR}_x(p)$ , 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 SSD relation.

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

REFERENCES

[1] B. C. ARNOLD, *Majorization and the Lorenz Order: A Brief Introduction*, Lecture Notes in Statist. 43, Springer-Verlag, Berlin, 1980.  
 [2] P. ARTZNER, F. DELBAEN, J.-M. EBER, AND D. HEATH, *Coherent measures of risk*, Math. 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. GOTOH AND H. KONNO, *Third degree stochastic dominance and mean-risk analysis*, *Management Science*, 46 (2000), pp. 289–301.
- [8] J. HADAR AND W. RUSSELL, *Rules for ordering uncertain prospects*, *American Economic Review*, 59 (1969), pp. 25–34.
- [9] G. HANOCHE AND H. LEVY, *The efficiency analysis of choices involving risk*, *Rev. Econom. Stud.*, 36 (1969), pp. 335–346.
- [10] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge University Press, Cambridge, MA, 1934.
- [11] L. V. KANTOROVICH AND G. P. AKILOV, *Functional Analysis*, Pergamon Press, Oxford, UK, 1982.
- [12] 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.
- [13] E. LEHMANN, *Ordered families of distributions*, *Annals of Mathematical Statistics*, 26 (1955), pp. 399–419.
- [14] H. LEVY, *Stochastic dominance and expected utility: survey and analysis*, *Management Science*, 38 (1992), pp. 555–593.
- [15] H. LEVY AND Y. KROLL, *Ordering uncertain options with borrowing and lending*, *J. Finance*, 33 (1978), pp. 553–573.
- [16] M. O. LORENZ, *Methods of measuring concentration of wealth*, *J. Amer. Statist. Assoc.*, 9 (1905), pp. 209–219.
- [17] H. M. MARKOWITZ, *Portfolio selection*, *J. Finance*, 7 (1952), pp. 77–91.
- [18] H. M. MARKOWITZ, *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford, UK, 1987.
- [19] A. W. MARSHALL AND I. OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, San Diego, 1979.
- [20] P. MULIERE AND M. SCARSINI, *A note on stochastic dominance and inequality measures*, *J. Econom. Theory*, 49 (1989), pp. 314–323.
- [21] W. OGRYCZAK, *Multiple criteria linear programming model for portfolio selection*, *Ann. Oper. Res.*, 97 (2000), pp. 143–162.
- [22] W. OGRYCZAK AND A. RUSZCZYŃSKI, *From stochastic dominance to mean-risk models: Semideviations as risk measures*, *Eur. J. Oper. Res.*, 116 (1999), pp. 33–50.
- [23] W. OGRYCZAK AND A. RUSZCZYŃSKI, *On consistency of stochastic dominance and mean-semideviation models*, *Math. Program.*, 89 (2001), pp. 217–232.
- [24] G. C. PFLUG, *Some remarks on the value-at-risk and the conditional value-at-risk*, in *Probabilistic Constrained Optimization: Methodology and Applications*, S. Uryasev, ed., Kluwer Academic Publishers, Norwell, MA, 2000, pp. 278–287.
- [25] J. P. QUIRK AND R. SAPOSNIK, *Admissibility and measurable utility functions*, *Rev. Econom. Stud.*, 29 (1962), pp. 140–146.
- [26] *RiskMetrics*, Technical document, J. P. Morgan, New York, 1996.
- [27] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [28] R. T. ROCKAFELLAR AND S. URYASEV, *Optimization of conditional value-at-risk*, *J. Risk*, 2 (2000), pp. 21–41.
- [29] M. ROTHSCHILD AND J. E. STIGLITZ, *Increasing risk: I. A definition*, *J. Econom. Theory*, 2 (1969), pp. 225–243.
- [30] A. RUSZCZYŃSKI, *Decomposition methods in stochastic programming*, *Math. Program.*, 79 (1997), pp. 333–353.
- [31] W. F. SHARPE, *A linear programming approximation for the general portfolio analysis problem*, *J. Financial and Quantitative Analysis*, 6 (1971), pp. 1263–1275.
- [32] G. A. WHITMORE AND M. C. FINDLAY, EDs., *Stochastic Dominance: An Approach to Decision-Making Under Risk*, D. C. Heath, Lexington, MA, 1978.
- [33] S. YITZHAKI, *Stochastic dominance, mean variance, and Gini's mean difference*, *American Economic Review*, 72 (1982), pp. 178–185.
- [34] M. R. YOUNG, *A minimax portfolio selection rule with linear programming solution*, *Management Science*, 44 (1998), pp. 673–683.
- [35] W. H. YOUNG, *On classes of summable functions and their Fourier series*, *Proc. Roy. Soc. London*, A87 (1912), pp. 235–239.