Math. Program., Ser. B 89: 217–232 (2001)

Digital Object Identifier (DOI) 10.1007/s101070000203

Włodzimierz Ogryczak · Andrzej Ruszczyński

On consistency of stochastic dominance and mean–semideviation models*

Received: September 15, 1998 / Accepted: October 1, 2000

Abstract. We analyze relations between two methods frequently used for modeling the choice among uncertain outcomes: stochastic dominance and mean–risk approaches. New necessary conditions for stochastic dominance are developed. These conditions compare values of a certain functional, which contains two components: the expected value of a random outcome and a risk term represented by the central semideviation of the corresponding degree. If the weight of the semideviation in the composite objective does not exceed the weight of the expected value, maximization of such a functional yields solutions which are efficient in terms of stochastic dominance. The results are illustrated graphically.

Key words. decisions under risk – stochastic dominance – mean–risk models – portfolio optimization

1. Introduction

Uncertainty is the key ingredient in many decision problems. Financial planning, cancer screening and airline scheduling are just few examples of areas in which ignoring uncertainty may lead to inferior or simply wrong decisions. There are many ways to model uncertainty; one that proved particularly fruitful is to use probabilistic models.

We consider decisions with real–valued outcomes, such as return, net profit or number of lives saved. Although we sometimes discuss implications of our analysis in the portfolio selection context, we do not assume any specificity related to this or any other application.

Whatever the application, the fundamental question is how to compare uncertain outcomes. This has been the concern of many authors and will remain our concern in this paper. The general assumption that we make is that larger outcomes are preferred over smaller outcomes.

Two methods are frequently used for modeling the choice among uncertain prospects: *stochastic dominance* [20, 11] and *mean–risk* approaches [14]. The first one is based on an axiomatic model of risk averse preferences, but does not provide a simple computational recipe. It is, in fact, a multiple criteria model with a continuum of criteria.

W. Ogryczak: Warsaw University of Technology, Institute of Control and Computation Engineering, 00–665 Warsaw, Poland, e-mail: ogryczak@ia.pw.edu.pl

A. Ruszczyński: Rutgers University, Department of Management Science and Information Systems, Piscataway, NJ 08854, USA, e-mail: rusz@rutcor.rutgers.edu

Mathematics Subject Classification (2000): 91B06, 90C15, 91B28

* This work was partially supported by the International Institute for Applied Systems Analysis, Laxenburg, Austria, and was distributed on line as Interim Report IR-97-043

The second approach quantifies the problem in a lucid form of two criteria: the *mean*, which is the expected outcome, and the *risk* – a scalar measure of the variability of outcomes. The mean–risk model is appealing to decision makers and allows a simple trade-off analysis, analytical or geometrical. More specifically, one may maximize scalarized objectives of the form

$$\mu_X - \lambda r_X, \tag{1}$$

where μ_X is the mean and r_X is the risk associated with the random variable X, and $\lambda > 0$ is a *trade-off* coefficient.

On the other hand, the mean–risk approach is unable to model the entire gamut of risk-averse preferences. Moreover, for typical dispersion statistics used as risk measures, the mean–risk approach may lead to obviously inferior solutions.

The seminal portfolio optimization model of Markowitz [12] uses variance as the risk measure. It is, in general, not consistent with stochastic dominance rules; the use of semivariance rather than variance was already recommended by Markowitz himself [13]. Porter [18] showed that a fixed target semivariance as the risk measure makes the mean–risk model consistent with the stochastic dominance. This approach was extended by Fishburn [5] to more general risk measures associated with outcomes below some fixed target.

Our aim is to develop relations between the stochastic dominance and mean–risk approaches that use more natural measures of risk, associated with all underachievements below the mean. Therefore, we focus our analysis on the central semideviations:

$$\begin{aligned}\bar{\delta}_X^{(k)} &= \Big(\mathbb{E}\big\{(\mu_X - X)^k \mathbb{1}_{X \le \mu_X}\big\}\Big)^{1/k} \\ &= \left(\int_{-\infty}^{\mu_X} (\mu_X - \xi)^k \, P_X(d\xi)\right)^{1/k}, \quad k = 1, 2, \dots ,\end{aligned} \tag{2}$$

where P_X denotes the probability measure induced by the random variable X on the real line,

$$\mu_X = \mathbb{E}\{X\} = \int_{-\infty}^{\infty} \xi \, P_X(d\xi), \tag{3}$$

and $\mathbb{1}_{X \le \mu_X}$ denotes the indicator function of the event $\{X \le \mu_X\}$. In particular, (2) for $k = 1$ represents the *absolute semideviation*

$$\bar{\delta}_X^{(1)} = \bar{\delta}_X = \int_{-\infty}^{\mu_X} (\mu_X - \xi) \, P_X(d\xi) = \frac{1}{2} \int_{-\infty}^{\infty} |\xi - \mu_X| \, P_X(d\xi) \tag{4}$$

(with the last equality following from straightforward calculation), and for $k = 2$ it represents the *standard semideviation*:

$$\bar{\delta}_X^{(2)} = \bar{\sigma}_X = \left(\int_{-\infty}^{\mu_X} (\mu_X - \xi)^2 \, P_X(d\xi)\right)^{1/2}. \tag{5}$$

We shall show that mean–risk models using semideviations as risk measures are consistent with stochastic dominance orders, if the mean–risk trade-off coefficient in (1) is bounded by one. This will imply, roughly speaking, that maximization of functionals of form (1), yields solutions which are efficient in terms of stochastic dominance.

In Sect. 2 we recall the notion of stochastic dominance and establish its basic properties. Section 3 develops new necessary conditions for stochastic dominance and Sect. 4 contains their graphical interpretation. In Sect. 5 we use these conditions to establish relations between stochastic dominance and mean–risk models, and in Sect. 6 we present simple sufficient conditions for stochastic efficiency.

2. Stochastic dominance

Stochastic dominance is based on an axiomatic model of risk averse preferences [3]. It originated from the majorization theory for the discrete case [9,15] and was later extended to general distributions [8,19]. Since that time it has been widely used in economics and finance (see [1,11] for numerous references). In the stochastic dominance approach random variables are compared by the pointwise comparison of their distribution functions $F^{(k)}$.

For a real random variable X the first function $F_X^{(1)}$ is the right–continuous cumulative distribution function

$$F_X^{(1)}(\eta) = F_X(\eta) = \int_{-\infty}^{\eta} P_X(d\xi) = \mathbb{P}\{X \le \eta\} \quad \text{for } \eta \in \mathbb{R}. \tag{6}$$

The kth function $F_X^{(k)}$ (for $k = 2, 3, \dots$) is defined recursively as

$$F_X^{(k)}(\eta) = \int_{-\infty}^{\eta} F_X^{(k-1)}(\xi)\, d\xi \quad \text{for } \eta \in \mathbb{R}. \tag{7}$$

The *kth degree stochastic dominance* (kSD) is understood in the following way:

$$X \succeq_{(k)} Y \quad \Leftrightarrow \quad F_X^{(k)}(\eta) \le F_Y^{(k)}(\eta) \quad \text{for all } \eta \in \mathbb{R}. \tag{8}$$

The corresponding strict dominance relation $\succ_{(k)}$ is defined by the standard rule

$$X \succ_{(k)} Y \quad \Leftrightarrow \quad X \succeq_{(k)} Y \quad \text{and} \quad Y \not\succeq_{(k)} X. \tag{9}$$

Thus, we say that X *dominates* Y *by the kSD rule* ($X \succ_{(k)} Y$), if $F_X^{(k)}(\eta) \le F_Y^{(k)}(\eta)$ for all $\eta \in R$, with strict inequality holding for at least one η. In (8) and (9) we implicitly assume that the functions $F_X^{(k)}$ and $F_Y^{(k)}$ are well defined; this is guaranteed when $\mathbb{E}|X|^{k-1} < \infty$ and $\mathbb{E}|Y|^{k-1} < \infty$ (see Proposition 1 below).

Clearly, $X \succeq_{(k-1)} Y$ implies $X \succeq_{(k)} Y$ and $X \succ_{(k-1)} Y$ implies $X \succ_{(k)} Y$, provided that the kth degree function $F_X^{(k)}$ is well defined.

We shall employ a slightly more general approach to the topic. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space, and let $\mathbb{E}X = \int X(\omega)\, \mathbb{P}(d\omega)$ denote the expected value of the random variable X. The space of real random variables X such that $\mathbb{E}\{|X|^k\} < \infty$ is denoted, as usual, $\mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})$ (we frequently write simply $\mathcal{L}_k$). The norm in $\mathcal{L}_k$ is defined as

$$\|X\|_k = \left(\mathbb{E}\{|X|^k\}\right)^{1/k}.$$

The distribution functions (7) are closely related to the norms in the spaces $\mathcal{L}_k$ [4,6].

Proposition 1. *Let $k \geq 1$ and $X \in \mathcal{L}_k$. Then for all $\eta \in \mathbb{R}$*

$$F_X^{(k+1)}(\eta) = \frac{1}{k!}\int_{-\infty}^{\eta} (\eta - \xi)^k \, P_X(d\xi) = \frac{1}{k!}\|\max(0, \eta - X)\|_k^k.$$

Proof. If $k = 0$, the left equation follows directly from (6) (with the convention $0! = 1$). Assuming that it holds for $k - 1$ we shall show it for k. We have

$$\begin{aligned} F_X^{(k+1)}(\zeta) &= \frac{1}{(k-1)!}\int_{-\infty}^{\zeta}\left(\int_{-\infty}^{\eta}(\eta-\xi)^{k-1}\,P_X(d\xi)\right)d\eta \\ &= \frac{1}{(k-1)!}\int_{-\infty}^{\zeta}\left(\int_{\xi}^{\zeta}(\eta-\xi)^{k-1}\,d\eta\right)P_X(d\xi), \end{aligned}$$

where the order of integration could be changed by Fubini's theorem (see, e.g., [2]). Evaluation of the integral with respect to η gives the result for k.

□

Remark 1. Proposition 1 allows to define the functions $F_X^{(\kappa)}$ and the corresponding dominance relations for arbitrary real $\kappa > 0$:

$$F_X^{(\kappa+1)}(\eta) = \frac{1}{\Gamma(\kappa+1)}\|\max(0, \eta - X)\|_\kappa^\kappa,$$

where $\Gamma(\cdot)$ denotes Euler's gamma function. In the sequel, however, we shall consider only integer κ.

Equation (2) and Proposition 1 imply the following observation.

Corollary 1. *Let $k \geq 1$ and $X \in \mathcal{L}_k$. Then $\bar{\delta}_X^{(k)} = \left(k! F_X^{(k+1)}(\mu_X)\right)^{1/k}$.*

It is also clear that the functions $F^{(k)}$ are nondecreasing for $k \geq 1$ and convex for $k \geq 2$, but the convexity property can be strengthened substantially.

Proposition 2. *Let $k \geq 1$ and $X \in \mathcal{L}_k$. Then for all $a, b \in \mathbb{R}$ and all $t \in [0, 1]$ one has*

$$F_X^{(k+1)}((1-t)a + tb) \leq \left((1-t)\left(F_X^{(k+1)}(a)\right)^{1/k} + t\left(F_X^{(k+1)}(b)\right)^{1/k}\right)^k. \quad (10)$$

Proof. Let $t \in [0, 1]$. Consider the random variables $A = \max(0, a - X)$, $B = \max(0, b - X)$, and $U = \max(0, (1-t)a + tb - X)$. By the convexity of the function $z \to \max(0, z - X)$, with probability one

$$0 \leq U \leq (1-t)A + tB.$$

Therefore,

$$\|U\|_k \leq \|(1-t)A + tB\|_k \leq (1-t)\|A\|_k + t\|B\|_k,$$

where we used the triangle inequality for $\|\cdot\|_k$.

By Proposition 1, $k!F^{(k+1)}(a) = \|A\|_k^k$. Similarly, $k!F^{(k+1)}(b) = \|B\|_k^k$, and $k!F^{(k+1)}((1-t)a + tb) = \|U\|_k^k$. Substitution into the last inequality yields the required result.

□

In a similar way we can prove the following properties.

Proposition 3. *Let* $k \geq 1$ *and* $X, Y \in \mathcal{L}_k$. *Then for all* $\eta \in \mathbb{R}$ *and all* $t \in [0,1]$ *one has*

$$F^{(k+1)}_{(1-t)X+tY}(\eta) \leq \left((1-t)\left(F_X^{(k+1)}(\eta)\right)^{1/k} + t\left(F_Y^{(k+1)}(\eta)\right)^{1/k}\right)^k. \tag{11}$$

Proof. We define the random variables $A = \max(0, \eta - X)$, $B = \max(0, \eta - Y)$, and $U = \max(0, \eta - (1-t)X - tY)$, and proceed exactly as in the proof of Proposition 2.

□

Proposition 4. *Let* $k \geq 1$ *and* $X, Y \in \mathcal{L}_k$. *Then for all* $t \in [0,1]$ *one has*

$$\bar{\delta}^{(k)}_{(1-t)X+tY} \leq (1-t)\bar{\delta}^{(k)}_X + t\bar{\delta}^{(k)}_Y. \tag{12}$$

Proof. Define $A = \max(0, \mu_X - X)$, $B = \max(0, \mu_Y - Y)$, and $U = \max(0, (1-t)\mu_X + t\mu_Y - (1-t)X - tY)$, and proceed as in the proof of Proposition 2.

□

3. Necessary conditions for stochastic dominance

The simplest necessary condition for kth degree stochastic dominance is the corresponding inequality for expected values [7].

Proposition 5. *Let* $k \geq 1$ *and* $X, Y \in \mathcal{L}_k$. *If* $X \succeq_{(k+1)} Y$, *then* $\mu_X \geq \mu_Y$.

Our objective is to develop stronger necessary conditions that involve central semideviations. At first we establish some technical results.

Lemma 1. *Let* $k \geq 1$ *and* $X \in \mathcal{L}_k$. *Then*

$$\left(i!F_X^{(i+1)}(\eta)\right)^{1/i} \leq \left(k!F_X^{(k+1)}(\eta)\right)^{1/k}\left(\mathbb{P}\{X < \eta\}\right)^{1/i-1/k} \quad \text{for } i = 1, \dots, k.$$

Proof. We have

$$i!F_X^{(i+1)}(\eta) = \mathbb{E}\left\{(\max(0, \eta - X))^i\right\} = \mathbb{E}\left\{(\max(0, \eta - X))^i \cdot \mathbb{1}_{X<\eta}\right\},$$

where $\mathbb{1}_{X<\eta}$ denotes the indicator function of the event $\{X < \eta\}$.

Define $A = (\max(0, \eta - X))^i$, $B = \mathbb{1}_{X<\eta}$, $p = k/i$ and $q = k/(k-i)$. From Hölder's inequality $\mathbb{E}\{AB\} \leq \|A\|_p \|B\|_q$ (see, e.g., [2]) we obtain

$$\begin{aligned} i!F_X^{(i+1)}(\eta) &\leq \left\|(\max(0, \eta - X))^i\right\|_{k/i} \cdot \left\|\mathbb{1}_{X<\eta}\right\|_{k/(k-i)} \\ &= \|\max(0, \eta - X)\|_k^i \left(\mathbb{P}\{X < \eta\}\right)^{(k-i)/k}. \end{aligned}$$

Raising both sides to the power $1/i$ we obtain the result.

□

Lemma 2. *Let* $k \geq 1$, $X, Y \in \mathcal{L}_k$ *and let* $X \succeq_{(k+1)} Y$. *Then*

(i) $\left(i!F_X^{(i+1)}(\mu_Y)\right)^{1/i} \leq \bar{\delta}_Y^{(k)}\left(\mathbb{P}\{X < \mu_Y\}\right)^{1/i-1/k}$ *for all* $i = 1, \ldots, k$;

(ii) *if* $\bar{\delta}_Y^{(k)} > 0$, *then* $\left(i!F_X^{(i+1)}(\mu_Y)\right)^{1/i} < \bar{\delta}_Y^{(k)}$ *for all* $i = 1, \ldots, k-1$.

Proof. By Lemma 1 and the postulated dominance,

$$\begin{aligned}\left(i!F_X^{(i+1)}(\mu_Y)\right)^{1/i} &\leq \left(k!F_X^{(k+1)}(\mu_Y)\right)^{1/k}\left(\mathbb{P}\{X < \mu_Y\}\right)^{1/i-1/k}\\ &\leq \left(k!F_Y^{(k+1)}(\mu_Y)\right)^{1/k}\left(\mathbb{P}\{X < \mu_Y\}\right)^{1/i-1/k}\\ &= \bar{\delta}_Y^{(k)}\left(\mathbb{P}\{X < \mu_Y\}\right)^{1/i-1/k},\end{aligned} \tag{13}$$

for $i = 1, \ldots, k$, which completes the proof of (i). To prove (ii), note that Proposition 5 implies that $\mathbb{P}\{X < \mu_Y\} \leq \mathbb{P}\{X < \mu_X\} < 1$.

□

We are now ready to state the main result of this section.

Theorem 1. *Let* $k \geq 1$ *and* $X, Y \in \mathcal{L}_k$. *If* $X \succeq_{(k+1)} Y$ *then* $\mu_X \geq \mu_Y$ *and*

$$\mu_X - \bar{\delta}_X^{(k)} \geq \mu_Y - \bar{\delta}_Y^{(k)},$$

where the last inequality is strict whenever $\mu_X > \mu_Y$.

Proof. By (7) and (8),

$$F_X^{(k+1)}(\mu_X) = F_X^{(k+1)}(\mu_Y) + \int_{\mu_Y}^{\mu_X} F_X^{(k)}(\xi)\, d\xi \leq F_Y^{(k+1)}(\mu_Y) + \int_{\mu_Y}^{\mu_X} F_X^{(k)}(\xi)\, d\xi. \tag{14}$$

Let $k > 1$. Owing to Proposition 5, $\mu_X \geq \mu_Y$, and the assertion needs to be proved only in the case of $\bar{\delta}_X^{(k)} > \bar{\delta}_Y^{(k)}$. The integral on the right hand side of (14) can be estimated by Proposition 2:

$$\begin{aligned}&\int_{\mu_Y}^{\mu_X} F_X^{(k)}(\xi)\, d\xi = (\mu_X - \mu_Y)\int_0^1 F_X^{(k)}((1-t)\mu_X + t\mu_Y)\, dt\\ &\quad \leq (\mu_X - \mu_Y)\int_0^1 \left((1-t)\left(F_X^{(k)}(\mu_X)\right)^{1/(k-1)} + t\left(F_X^{(k)}(\mu_Y)\right)^{1/(k-1)}\right)^{k-1} dt.\end{aligned}$$

Using Lemmas 1 and 2 (with $i = k-1$), and integrating we obtain:

$$\int_{\mu_Y}^{\mu_X} F_X^{(k)}(\xi)\, d\xi \le \frac{\mu_X - \mu_Y}{(k-1)!} \int_0^1 \Big((1-t)\bar{\delta}_X^{(k)} (\mathbb{P}\{X < \mu_X\})^{1/k(k-1)} + t\bar{\delta}_Y^{(k)} (\mathbb{P}\{X < \mu_Y\})^{1/k(k-1)} \Big)^{k-1} dt$$

$$\le \frac{\mu_X - \mu_Y}{(k-1)!} \int_0^1 \Big((1-t)\bar{\delta}_X^{(k)} + t\bar{\delta}_Y^{(k)} \Big)^{k-1} dt$$

$$= \frac{\mu_X - \mu_Y}{k!} \cdot \frac{(\bar{\delta}_X^{(k)})^k - (\bar{\delta}_Y^{(k)})^k}{\bar{\delta}_X^{(k)} - \bar{\delta}_Y^{(k)}}. \tag{15}$$

Substitution into (14) and simplification with the use of Corollary 1 yield

$$\bar{\delta}_X^{(k)} - \bar{\delta}_Y^{(k)} \le \mu_X - \mu_Y, \tag{16}$$

which was set out to prove.

We shall now prove that (16) is strict, if $\mu_X > \mu_Y$. Suppose that $\bar{\delta}_Y^{(k)} > 0$. By virtue of Lemma 2(ii), inequality (15) is strict, which makes (16) strict, too.

If $\mu_X > \mu_Y$ and $\bar{\delta}_Y^{(k)} = 0$, we must have $\mathbb{P}\{X < \mu_Y\} = 0$, so

$$\bar{\delta}_X^{(k)} \le \mathbb{P}\{X < \mu_X\}^{1/k}(\mu_X - \mu_Y) < \mu_X - \mu_Y$$

and (16) is strict again.

If $k = 1$ the integral on the right hand side of (14) can be simply bounded by $\mu_X - \mu_Y$, and we get (16) in this case, too. Moreover, $F_X(\xi) < 1$ for $\xi < \mu_X$, and the inequality is strict whenever $\mu_Y < \mu_X$.

□

Since the dominance relation $X \succeq_{(k+1)} Y$ implies $X \succeq_{(m)} Y$ for all $m \ge k+1$ such that $F_X^{(m)}$ is well-defined, we obtain the following corollary.

Corollary 2. *If $X \succeq_{(k+1)} Y$ for some $k \ge 1$, then $\mu_X \ge \mu_Y$ and $\mu_X - \bar{\delta}_X^{(m)} \ge \mu_Y - \bar{\delta}_Y^{(m)}$ for all $m \ge k$ such that $\mathbb{E}\{|X|^m\} < \infty$.*

A careful analysis of the proof of Theorem 1 reveals that its assertion can be slightly strengthened. Indeed, estimating in (15) the quantities $\mathbb{P}\{X < \mu_Y\}$ and $\mathbb{P}\{X < \mu_X\}$ from above by some constant ρ,

$$\mathbb{P}\{X < \mu_X\} \le \rho \le 1,$$

we obtain the necessary condition

$$\rho^{1/k}\mu_X - \bar{\delta}_X^{(k)} \ge \rho^{1/k}\mu_Y - \bar{\delta}_Y^{(k)}.$$

Unfortunately, it does not possess the separability properties of the assertion of Theorem 1, because the right hand side contains a factor dependent on X. In the special case of *symmetric* distributions, however, we can use $\rho = 1/2$. We can also use central deviations

$$\delta_X^{(k)} = \left(\int_{-\infty}^{\infty} |\mu_X - \xi|^k \, P_X(d\xi) \right)^{1/k} = 2^{1/k} \bar{\delta}_X^{(k)},$$

and obtain a stronger necessary condition.

Corollary 3. *If X and Y are symmetric random variables and $X \succeq_{(k+1)} Y$ for some $k \geq 1$, then $\mu_X \geq \mu_Y$ and $\mu_X - \delta_X^{(m)} \geq \mu_Y - \delta_Y^{(m)}$ for all $m \geq k$ such that $\mathbb{E}\{|X|^m\} < \infty$.*

4. The Outcome–Risk diagram

Our results have a useful graphical interpretation. Let us start from the special case of the second degree stochastic dominance.

For a random outcome X having a bounded variance we consider the graph of the function $F_X^{(2)}$: the Outcome-Risk (O-R) diagram (Fig. 1). By Corollary 1, the first two semimoments are easily identified in the O-R diagram: the absolute semideviation $\bar{\delta}_X = \bar{\delta}_X^{(1)}$ is the value $F_X^{(2)}(\mu_X)$, and the semivariance $\bar{\sigma}_X^2 = (\bar{\delta}_X^{(2)})^2$ is twice the area below the graph from $-\infty$ to μ_X. We also have a manifestation of the Lyapunov inequality $\bar{\sigma}_X \geq \bar{\delta}_X$ (Lemma 1 with $\eta = \mu_X$, $k = 2$ and $i = 1$), because the shaded area contains the triangle with vertices $(\mu_X, 0)$, $(\mu_X, \bar{\delta}_X)$ and $(\mu_X - \bar{\delta}_X, 0)$.

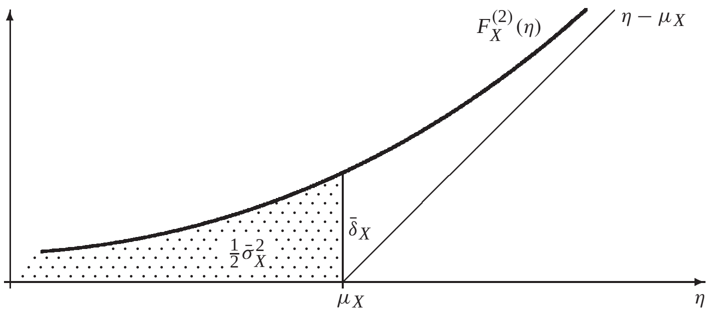

Fig. 1. The O-R diagram and semimoments

Now consider two random variables X and Y in a common O-R diagram. If $X \succeq_{(2)} Y$ then Theorem 1 implies that $\mu_X - \bar{\delta}_X^{(1)} \geq \mu_Y - \bar{\delta}_Y^{(1)}$. This is obvious from Fig. 2, because function $F_X^{(2)}$ is bounded from below by the linear function $\bar{\delta}_X + p_X(\eta - \mu_X)$ where $p_X = \mathbb{P}\{X < \mu_X\} < 1$.

But we also have a more refined relation illustrated in Fig. 3. The area below $F_X^{(2)}$ to the left of μ_X, equal to $\frac{1}{2}\bar{\sigma}_X^2$, is not larger than the area below $F_Y^{(2)}$ to the left of

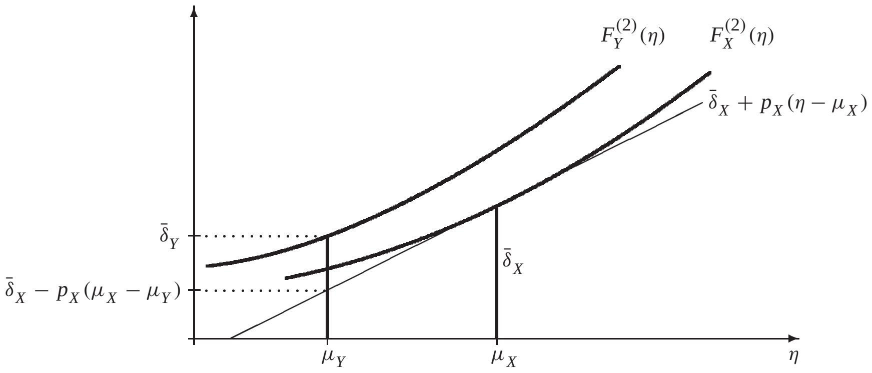

Fig. 2. The first necessary condition: $X \succeq_{(2)} Y \Rightarrow \bar{\delta}_X - p_X(\mu_X - \mu_Y) \leq \bar{\delta}_Y$ where $p_X = \mathbb{P}\{X < \mu_X\}$

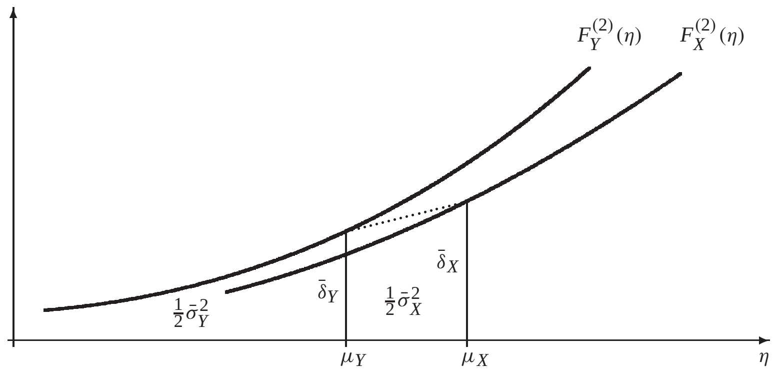

Fig. 3. The second necessary condition: $X \succeq_{(2)} Y \Rightarrow \frac{1}{2}\bar{\sigma}_X^2 \leq \frac{1}{2}\bar{\sigma}_Y^2 + \frac{1}{2}(\mu_X - \mu_Y)(\bar{\delta}_X + \bar{\delta}_Y)$

μ_Y, increased by the area of the trapezoid with the vertices: $(\mu_Y, 0)$, $(\mu_Y, \bar{\delta}_Y)$, $(\mu_X, 0)$, and $(\mu_X, \bar{\delta}_X)$. Employing the Lyapunov inequalities $\bar{\delta}_X \leq \bar{\sigma}_X$ and $\bar{\delta}_Y \leq \bar{\sigma}_Y$, we obtain a graphical proof of Corollary 2 for $k = 1$ and $m = 2$. For more details on the properties of the O-R diagram in the case of the second degree dominance, the reader is referred to [17].

For a higher degree $k > 1$ it is more convenient to analyze the graph of the function

$$G_X^{(k)}(\eta) = \left(k! F_X^{(k+1)}(\eta)\right)^{1/k} = \|\max(0, \eta - X)\|_k, \tag{17}$$

instead of $F_X^{(k+1)}$ itself. It has the following properties.

Proposition 6. *Let $k > 1$ and $X \in \mathcal{L}_k$. Then*

(i) $\lim_{\eta \to -\infty} G_X^{(k)}(\eta) = 0$;

(ii) $G_X^{(k)}(\eta) \geq \eta - \mu_X$ *for all* $\eta \in \mathbb{R}$, *and* $\lim_{\eta \to \infty} \left(G_X^{(k)}(\eta) - \eta + \mu_X\right) = 0$;

(iii) *the function* $G_X^{(k)}(\cdot)$ *is convex;*

(iv) *at each* $\eta \in \mathbb{R}$ *such that* $\mathbb{P}\{X < \eta\} > 0$, *the function* $G_X^{(k)}(\eta)$ *is continuously differentiable and*

$$\frac{dG_X^{(k)}(\eta)}{d\eta} \le \big(\mathbb{P}\{X < \eta\}\big)^{1/k}.$$

Proof. Assertions (i) and (iii) follow directly from the definition. To prove (ii) let us recall (17):

$$\begin{aligned}\big(G_X^{(k)}(\eta)\big)^k &= \mathbb{E}\big\{(\max(0, \eta - X))^k\big\}\\ &\le \mathbb{E}\big\{(\eta - X)^k\big\} + \mathbb{E}\big\{\max(0, X - \eta)^k\big\}.\end{aligned} \tag{18}$$

The second component at the right hand side of the last inequality tends to zero, as $\eta \to \infty$. We shall estimate the first component:

$$\mathbb{E}\big\{(\eta - X)^k\big\} = \mathbb{E}\big\{(\eta - \mu_X + \mu_X - X)^k\big\} = (\eta - \mu_X)^k\, \mathbb{E}\left\{\left(1 + \frac{\mu_X - X}{\eta - \mu_X}\right)^k\right\}. \tag{19}$$

Let us denote

$$\Delta = \frac{\mu_X - X}{\eta - \mu_X}.$$

Since $\mathbb{E}\Delta = 0$, we have

$$\begin{aligned}\mathbb{E}\big\{(1 + \Delta)^k\big\} &= \mathbb{E}\left\{\sum_{i=0}^{k}\binom{k}{i}\Delta^i\right\} = 1 + \mathbb{E}\left\{\sum_{i=2}^{k}\binom{k}{i}\Delta^i\right\}\\ &\le 1 + C_1\mathbb{E}\big\{\Delta^2(1+\Delta)^{k-2}\big\}\\ &= 1 + C_1(\eta - \mu_X)^{-2}\mathbb{E}\big\{(\mu_X - X)^2(1+\Delta)^{k-2}\big\},\end{aligned} \tag{20}$$

where C_1 is a constant independent on η. Since $X \in \mathcal{L}_k$, there exists a constant $C_2 > 0$ such that

$$\mathbb{E}\big\{(\mu_X - X)^2(1+\Delta)^{k-2}\big\} \le C_2 \tag{21}$$

for all sufficiently large η. Putting together (18)–(21) we obtain

$$\limsup_{\eta\to\infty}\Big(G_X^{(k)}(\eta) - \eta + \mu_X\Big)$$
$$\limsup_{\eta\to\infty}\left((\eta - \mu_X)\left[\left(1 + \frac{C_1C_2}{(\eta - \mu_X)^2}\right)^{1/k} - 1\right]\right) = 0.$$

On the other hand, by Jensen's inequality,

$$\big(G_X^{(k)}(\eta)\big)^k = \mathbb{E}\big\{(\max(0, \eta - X))^k\big\} \ge (\max(0, \eta - \mu_X))^k \ge (\eta - \mu_X)^k,$$

which completes the proof of (ii). To prove (iv) we first note that the differentiability of $G_X^{(k)}(\cdot)$ for $k > 1$ follows from the Lebesgue theorem. Since we only need to consider the case when $\mathbb{P}\{X < \eta\} > 0$, we simply differentiate:

$$\frac{dG_X^{(k)}(\eta)}{d\eta} = \left(\frac{\|\max(0, \eta - X)\|_{k-1}}{\|\max(0, \eta - X)\|_k}\right)^{k-1}.$$

The application of Lemma 1 with $i = k - 1$ to the right hand side of the last equation completes the proof.

□

We see that for all $k > 1$ the function $G_X^{(k)}$ has properties similar to $G_X^{(1)} = F_X^{(2)}$: convexity and two asymptotes. These are the horizontal axis and the line $\eta - \mu_X$ (cf. Fig. 1). For a deterministic outcome $X = \mu_X$ the graph of $G^{(k)}(\eta)$ coincides with the asymptotes. For a non-deterministic outcome the area between the graph of $G_X^{(k)}$ and its asymptotes is a measure of risk. In particular, it follows from Proposition 6(i)–(iii) that the maximum vertical diameter (the distance to the asymptotes) equals to $G_X^{(k)}(\mu_X)$, that is $\bar{\delta}_X^{(k)}$.

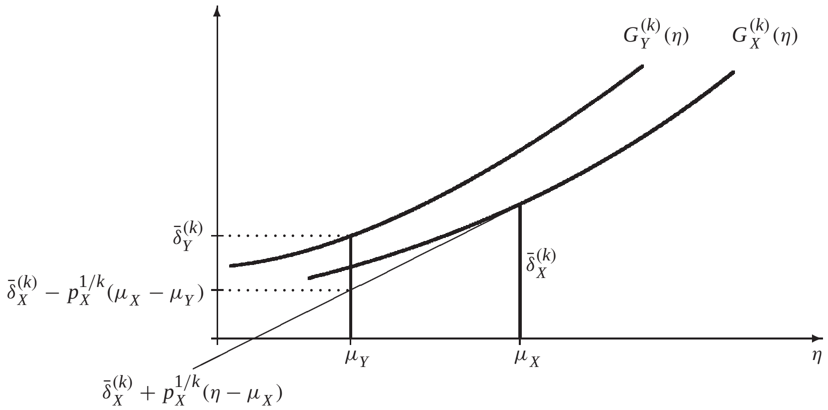

Fig. 4. The O-R diagram for the $(k+1)$st degree stochastic dominance: $X \succeq_{(k+1)} Y \Rightarrow \bar{\delta}_X^{(k)} - p_X^{1/k}(\mu_X - \mu_Y) \leq \bar{\delta}_Y^{(k)}$, where $p_X = \mathbb{P}\{X < \mu_X\}$

We can graphically interpret Theorem 1 for $k > 1$ in a way similar to the second degree case, as illustrated in Fig. 4. It follows from Proposition 6 (iii)–(iv) that for $\eta \leq \mu_X$ function $G_X^{(k)}(\eta)$ is bounded from below by the linear function $\bar{\delta}_X^{(k)} + p_X^{1/k}(\eta - \mu_X)$ where $p_X = \mathbb{P}\{X < \mu_X\} < 1$.

5. Mean–semideviation models

Mean–risk approaches are based on comparing two scalar characteristics (summary statistics) of each outcome: the expected value μ and some measure of risk r. The weak relation of mean–risk dominance is defined as follows

$$X \succeq_{\mu/r} Y \quad \Leftrightarrow \quad \mu_X \geq \mu_Y \quad \text{and} \quad r_X \leq r_Y.$$

The corresponding strict dominance relation $\succ_{\mu/r}$ is defined in the standard way (9). Thus we say that X *dominates* Y *by the* μ/r *rules* ($X \succ_{\mu/r} Y$), if $\mu_X \geq \mu_Y$ and $r_X \leq r_Y$ where at least one inequality is strict.

An important advantage of mean–risk approaches is the possibility to perform a pictorial trade-off analysis. Having assumed a trade-off coefficient $\lambda \geq 0$ one may directly compare real values of $\mu_X - \lambda r_X$ and $\mu_Y - \lambda r_Y$. This approach is consistent with mean–risk dominance in the sense that

$$X \succeq_{\mu/r} Y \quad \Rightarrow \quad \mu_X - \lambda r_X \geq \mu_Y - \lambda r_Y \quad \text{for all } \lambda \geq 0. \tag{22}$$

Therefore, an outcome that is inferior in terms of $\mu - \lambda r$ for some $\lambda \geq 0$ cannot be superior by the mean–risk dominance relation.

A mean–risk model is said to be *consistent* with the stochastic dominance relation of degree k if

$$X \succeq_{(k)} Y \quad \Rightarrow \quad X \succeq_{\mu/r} Y. \tag{23}$$

Such a consistency would be highly desirable, because it would allow us to search for stochastically non-dominated solutions (that is, solutions X for which there is no Y satisfying $Y \succeq_{(k)} X$) by the rule:

$$X \succ_{\mu/r} Y \quad \Rightarrow \quad Y \not\succeq_{(k)} X.$$

Moreover, negating (22) and (23) we would get a rule involving the scalarized objective:

$$\mu_X - \lambda r_X > \mu_Y - \lambda r_Y \quad \text{for some } \lambda \geq 0 \quad \Rightarrow \quad Y \not\succeq_{(k)} X.$$

We would then know that using simplified aggregate measures of the form $\mu - \lambda r$ would not lead to solutions that are inferior in terms of stochastic dominance.

A natural question arises: can mean–risk models be consistent with a stochastic dominance relation?

The most commonly used risk measure is the variance [14]. Unfortunately, the resulting mean–risk model is not, in general, consistent with stochastic dominance. The use of fixed-target risk measures is a possible remedy, because stochastic dominance relations are based on norms of fixed-target underachievements (Proposition 1).

We shall try to address the question in a different way. We modify the concept of consistency to accommodate scalarizations, and we use central semideviations as risk measures.

Definition 1. For a non-negative constant α, we say that a mean–risk model is α-*consistent* with the kth degree stochastic dominance relation if

$$X \succeq_{(k)} Y \quad \Rightarrow \quad \mu_X \geq \mu_Y \quad \text{and} \quad \mu_X - \alpha\, r_X \geq \mu_Y - \alpha\, r_Y. \tag{24}$$

By virtue of (22), consistency in the sense of (23) implies α-consistency for all $\alpha \geq 0$. Moreover,

$$\mu_X \geq \mu_Y \text{ and } \mu_X - \alpha\, r_X \geq \mu_Y - \alpha\, r_Y \quad \Rightarrow \quad \mu_X - \lambda r_X \geq \mu_Y - \lambda r_Y \text{ for all } 0 \leq \lambda \leq \alpha \tag{25}$$

(combine the inequalities at the left hand side with the weights $1 - \lambda/\alpha$ and λ/α). Thus α-consistency implies λ-consistency for all $\lambda \in [0, \alpha]$. It still guarantees that the mean–risk analysis leads us to non-dominated results in the sense that

$$\mu_X - \lambda r_X > \mu_Y - \lambda r_Y \text{ for some } 0 \leq \lambda \leq \alpha \quad \Rightarrow \quad Y \not\succeq_{(k)} X.$$

With these general definitions we can return to our main question: how can the risk measure be defined to maintain α-consistency with the corresponding stochastic dominance order? The answer follows immediately from Theorem 1.

Theorem 2. *In the space* $\mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})$ *the mean–risk model with* $r = \bar{\delta}^{(k)}$ *is* 1*-consistent with all stochastic dominance relations of degrees* $1, \ldots, k+1$.

Proof. By Theorem 1,

$$\mu_X - \bar{\delta}_X^{(k)} > \mu_Y - \bar{\delta}_Y^{(k)} \quad \Rightarrow \quad Y \not\succeq_{(k+1)} X.$$

The implication $Y \not\succeq_{(i+1)} X \Rightarrow Y \not\succeq_{(i)} X$, $i = k, \ldots, 1$, completes the proof.

□

In the special case of $k = 1$ we conclude that the mean–absolute deviation model of Konno and Yamazaki [10] is $\frac{1}{2}$-consistent with the first and the second degree stochastic dominance. Indeed, the absolute deviation satisfies $\delta^{(1)} = 2\bar{\delta}^{(1)}$, and Theorem 2 implies the result.

For $k = 2$ we see that the use of the central semideviation as the risk measure (instead of the variance in the Markowitz model) guarantees 1-consistency with stochastic dominance relations of degrees one, two and three.

The constant $\alpha = 1$ in Theorem 2 cannot be increased for general distributions, as the following example shows: $\mathbb{P}\{X = 0\} = (1+\varepsilon)^{-k}$, $\mathbb{P}\{X = 1\} = 1 - (1+\varepsilon)^{-k}$, and $Y = 0$. Obviously $X \succ_{(k+1)} Y$, but for each $\alpha > 1$ we can find $\varepsilon > 0$ for which $\mu_X - \alpha\bar{\delta}_X^{(k)} < 0 = \mu_Y - \alpha\bar{\delta}_Y^{(k)}$.

For symmetric distributions we can use Corollary 3 to get a wider range of trade-offs for semideviations, which allows us to replace semideviations with the corresponding deviations.

Corollary 4. *In the class of symmetric random variables in* $\mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})$ *the mean–risk model with* $r = \delta^{(k)}$ *is* 1*-consistent with all stochastic dominance relations of degrees* $1, \ldots, k+1$.

6. Stochastic efficiency in a set

Comparison of random variables is usually related to the problem of choice among risky alternatives in a given feasible (attainable) set Q. For example, in the simplest problem of portfolio selection [14] the feasible set of random variables is defined as all convex combinations of a given collection of investment opportunities (securities). A feasible random variable $X \in Q$ is called *efficient* under the relation $\succeq$ if there is no $Y \in Q$ such that $Y \succ X$. Consistency (24) leads to the following result.

Proposition 7. *If the mean–risk model is α-consistent for some $\alpha > 0$ with the kth degree stochastic dominance relation, then, except for random variables with identical μ and r, every random variable that maximizes $\mu - \lambda r$ for some $0 < \lambda < \alpha$ is efficient under the kth degree stochastic dominance rules.*

Proof. Let $0 < \lambda < \alpha$ and $X \in Q$ be maximal by $\mu - \lambda r$. Suppose that there exists $Z \in Q$ such that $Z \succ_{(k)} X$. Then from (24) we obtain

$$\mu_Z \geq \mu_X,$$

and

$$\mu_Z - \alpha r_Z \geq \mu_X - \alpha r_X.$$

Due to the maximality of X,

$$\mu_Z - \lambda r_Z \leq \mu_X - \lambda r_X.$$

All these relations may be true only if they are satisfied as equations; otherwise, combining the first two with weights $1 - \lambda/\alpha$ and λ/α we obtain a contradiction with the third one. Consequently, $\mu_Z = \mu_X$ and $r_Z = r_X$.

□

It follows from Proposition 7 that for mean–risk models satisfying (24), an optimal solution of problem

$$\max\{\mu_X - \lambda\, r_X \ : \ X \in Q\}$$

with any $0 < \lambda < \alpha$, is efficient under the kSD rules, provided that it has a unique pair (μ_X, r_X) among all optimal solutions.

Combining Proposition 7 and Theorem 2 we obtain the following sufficient condition for stochastic efficiency.

Theorem 3. *If $\widehat{X}$ is the unique solution of the problem*

$$\max\left\{\mu_X - \lambda\, \bar{\delta}_X^{(k)} \ : \ X \in Q\right\} \tag{26}$$

for some $\lambda \in (0, 1]$ and $k \geq 1$, then it is efficient under the rules of stochastic dominance of degrees $1, \ldots, k+1$.

Proof. It remains to consider the case $\lambda = 1$. Suppose that there exists $Z \in Q$ such that $Z \succ_{(k)} \widehat{X}$. Then from Theorem 2, $\mu_Z \geq \mu_X$ and

$$\mu_Z - \lambda\, \bar{\delta}_Z^{(k)} \geq \mu_{\widehat{X}} - \lambda\, \bar{\delta}_{\widehat{X}}^{(k)}.$$

Since $\widehat{X}$ is the unique maximizer of (26), $Z = \widehat{X}$.

□

Theorem 3 can be extended to risk measures defined as convex combinations of semideviations of various degrees. By applying Theorem 2 for $i = k, \ldots, m$ the following sufficient condition for stochastic efficiency can be obtained.

Corollary 5. *If $\widehat{X}$ is the unique solution of the problem*

$$\max \Big\{ \mu_X - \sum_{i=k}^{m} \lambda_i \bar{\delta}_X^{(i)} \ :\ X \in Q \Big\} \tag{27}$$

for some $m \geq k \geq 1$ such that $\mathbb{E}\{|X|^m\} < \infty$, and some sequence $\lambda_i \geq 0$ satisfying $0 < \sum_{i=k}^{m} \lambda_i \leq 1$, then it is efficient under the rules of stochastic dominance of degrees $1, \ldots, k+1$.

Owing to Corollaries 3 and 4, for symmetric distributions we can use the central deviation $\delta_X^{(i)}$ as the risk measures in (27), and still use the set $0 < \sum_{i=k}^{m} \lambda_i \leq 1$ of the coefficients in (27).

It is of practical importance that the simplified objective functionals in (26) and (27) are concave in X (see Proposition 4). In the case of multiple maximizers of the simplified objective functionals, some of them may be stochastically dominated, but only by other maximizers with the same mean value (Proposition 7). Then, an SD efficient maximizer may be identified by an additional (second-level) minimization of the corresponding central deviation.

7. Conclusions

The stochastic dominance relation $X \succeq_{(k+1)} Y$ is rather strong and difficult to verify: it is an inequality of two distribution functions, $F_X^{(k+1)} \leq F_Y^{(k+1)}$. The necessary conditions of Section 3 establish useful relations:

$$\mu_X - \lambda\bar{\delta}_X^{(k)} \geq \mu_Y - \lambda\bar{\delta}_Y^{(k)}, \quad \text{for all } \lambda \in [0, 1],$$

that follow from the dominance (μ_X and $\bar{\delta}_X^{(k)}$, defined in (3) and (2), denote the expectation and the kth central semideviation of X).

This allows us to relate stochastic dominance to mean–risk models with risk represented by the kth central semideviation $\bar{\delta}_X^{(k)}$. The key observation is that maximization of simplified objective functionals of the form

$$\mu_X - \lambda\bar{\delta}_X^{(k)},$$

where $\lambda \in (0, 1]$, yields solutions which are efficient in terms of stochastic dominance. In particular, maximization of $\mu_X - \lambda\bar{\delta}_X$ yields solutions efficient in terms of the first and the second degree stochastic dominance whereas maximization of $\mu_X - \lambda\bar{\sigma}_X$ generates efficient solutions for the first, the second and the third degree stochastic dominance. This may help to quickly identify promising candidates in complex decision problems under uncertainty.

References

1. Bawa, V.S. (1982): Stochastic dominance: a research bibliography. Manage. Sci. **28**, 698–712
2. Billingsley, P. (1995): Probability and Measure. John Wiley & Sons, New York
3. Fishburn, P.C. (1964): Decision and Value Theory. John Wiley & Sons, New York
4. Fishburn, P.C. (1976): Continua of stochastic dominance relations for bounded probability distributions. J. Math. Econ. **3**, 295–311
5. Fishburn, P.C. (1977): Mean–Risk Analysis with Risk Associated with Below Target Returns. Amer. Econ. Rev. **67**, 116–126
6. Fishburn, P.C. (1980): Continua of stochastic dominance relations for unbounded probability distributions. J. Math. Econ. **7**, 271–285
7. Fishburn, P.C. (1980): Stochastic dominance and moments of distributions. Math. Oper. Res. **5**, 94–100
8. Hanoch, G., Levy, H. (1969): The efficiency analysis of choices involving risk. Rev. Econ. Stud. **36**, 335–346
9. Hardy, G.H., Littlewood, J.E., Pólya, G. (1934): Inequalities. Cambridge University Press, Cambridge, MA
10. Konno, H., Yamazaki, H. (1991): Mean Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market. Manage. Sci. **37**, 519–531
11. Levy, H. (1992): Stochastic dominance and expected utility: survey and analysis. Manage. Sci. **38**, 555–593
12. Markowitz, H.M. (1952): Portfolio selection. J. Finance **7**, 77–91
13. Markowitz, H.M. (1959): Portfolio Selection. John Wiley & Sons, New York
14. Markowitz, H.M. (1987): Mean–Variance Analysis in Portfolio Choice and Capital Markets. Blackwell, Oxford
15. Marshall, A.W., Olkin, I. (1979): Inequalities: Theory of Majorization and its Applications. Academic Press, New York
16. Ogryczak, W., Ruszczyński, A. (1999): From stochastic dominance to mean–risk models: semideviations as risk measures. Eur. J. Oper. Res. **116**, 33–50
17. Ogryczak, W., Ruszczyński, A. (1997): On stochastic dominance and mean–semideviation models. Interim Report IR-97-043, International Institute for Applied Systems Analysis, Laxenburg
18. Porter, R.B. (1974): Semivariance and stochastic dominance: a comparison. Amer. Econ. Rev. **64**, 200–204
19. Rothschild, M., Stiglitz, J.E. (1970): Increasing risk: I. a definition. J. Econ. Theory **2**, 225–243
20. Whitmore, G.A., Findlay, M.C., eds. (1978): Stochastic Dominance: An Approach to Decision-Making Under Risk. Heath, D.C., Lexington, MA