Engineering and Process Economics, 4 (1979) 257–268

STATISTICAL DECISIONS BETWEEN PORTFOLIOS : MEAN-VARIANCE ANALYSIS REVISITED

K. Mosler
Fachbereich Wirtschafts- und Organisationswissenschaften
der Hochschule der Bundeswehr Hamburg
Postfach 700822 D-2000 Hamburg 70
(Germany)

The problem of choice from a given set of portfolios is a problem of ordering the proper set F of probability distribution functions. In this paper the relations between (first and second degree) stochastic dominance rules and certain mean-variance rules are explored. Conditions are given under which stochastic dominance efficient sets are contained in mean-variance efficient sets as well as conditions under which the sets coincide. In this light some recent empirical work on efficient sets of portfolios is reconsidered; theoretical and practical reasons are discussed which limit the applicability of stochastic dominance rules in portfolio analysis.

0. Introduction

Rules for ordering a set of risky assets have been a major object of the modern theory of choice. If an individual's utility is not fully known or may change over time only some rough features such as risk awareness and decreasing risk aversion can be employed in a decision model which is set up to forecast future decisions of the individual or to be a guide for the decision making individual himself.

The theory of semi-orderings on the set of distribution functions serves as an elegant tool since an efficient set contains all distribution functions of those risky assets which are most desired by rational decision makers of a certain class. The semi-orderings which belong to the class of risk averters and risk lovers respectively are well-known stochastic dominance relations; however, decisions based on them have the great drawback that they afford knowledge of the whole probability distribution functions.

Besides of pure gambling situations the decision maker is not able to know the distribution F of the returns completely; in nearly all economic situations he can only acquire some information on F or make a subjective guess at F or do both. Since empirical investigations are costly and in portfolio decisions the data base cannot be expanded without reducing its predictive power and since a subjective guess will not extend to the whole distribution but be restricted on a few intuitive parameters, he will decide by limited information only, as well on the possible alternatives as on their return distributions. The latter means that the decision will rely on a few distribution parameters which have been either taken empirically from past data or estimated intuitively by introspection. The question is : on what parameters should the decision be based, and is this decision rational in the sense of maximizing utility. Decision rules which are based on the first two moments alone have been in most popular use. But since the time the von Neumann-Morgenstern axioms had become fully accepted mean-variance analysis (MVA) has been questioned by a series of modern authors ([3],[4],[16]). Others like [11] and [17] built their theories on MVA. The shortcomings of MVA are amply known: if the distributions are arbitrary the utilities must be quadratic. On the other hand, if a wider class of utilities is taken into account MVA works only on classes of distributions which can be parametrised by mean and variance alone. In addition, if arbitrarily aggregated portfolios are allowed and the securities are independent all distributions must be normal. If all portfolios are normally distributed, mean-variance (MV) and stochastic dominance (SD) criteria coincide, i.e. MVA is rational.

The first question of this paper will be: Are there other classes of portfolios on which MVA gives rational decisions? On arbitrary classes a SD criterion is superior to the proper MV criterion in theory, of course; however, when the alternative criteria are applied to empirical instead of theoretical distribution functions, will the same hold true? This will be a second question discussed below.

In MVA a portfolio F is preferred to another portfolio G if its mean is greater and its varianceis smaller than that of G (the case of risk awareness) or, alternatively, if its mean and its variance exceed that of G (the case of risk proneness). It will be shown below that on certain classes of portfolios these MV preference relations imply expected utility preference by all individuals who have a concave (resp. convex) utility function. This is not restricted to sets of normal distributions but applies as well to sets which contain simple alternatives, rectangle or triangle distributions or others; in general to sets of portfolios which are taken from a common (μ,σ)-family. Further we will prove that on a set of gamma-distributed portfolios MV preference implies expected utility preference for risk averters. A similar result will be shown for risk lovers on lognormally distributed portfolios.
Recent empirical investigations have compared mean-variance efficient subsets with those generated by stochastic dominance relations ([1],[8],[9],[12],[13]). As their results are ambigous we will discuss them in section 3.

If the set of portfolios is not "very dense" with respect to μ and σ then the SD efficient sets enlarge considerably with the number of observations; this comes out to be a general property of SD efficient sets which are computed from empirical distribution functions. In that case the direct application of SD decision criteria to empirical distributions makes no sense and is in practice inferior to MVA.

In section 1 we define the different concepts of ordering a set of portfolios; section 2 presents theorems on the theoretical relations between the alternative ordering criteria and includes complete proofs. In section 3 we discuss empirical work and give some conclusions.

1. Alternative preference criteria

Let F be a set of portfolios, each of which is identified with the distribution function F of its (relative) return X. Define by F_k and F^k (k=0,1,2,...) the following successive integrals of F; $F_o = F^o = F$,

(1.1) $F_k(y) = \int_{-\infty}^{y} F_{k-1}(x)dx,$ (k=1,2,...)

(1.2) $\begin{cases} F^1(y) = \int_{y}^{\infty} [1-F(x)]dx, \\ F^k(y) = \int_{y}^{\infty} F^{k-1}(x)dx & (k=2,3,...) \end{cases}$

$F_1(y)$ exists for all y if the condition

(1.3) $\int_{-\infty}^{0} F(t)dt < \infty$ holds;

$F^1(y)$ exists for any y if

(1.4) $\int_{0}^{\infty} [1-F(t)]dt < \infty$ is valid.

Let Y be the return on another portfolio $G \in F$. Y is said to be not less preferred than X

- in general ($Y\ D_g\ X$), iff $F(y) \geq G(y)$ for all real y,
- by risk averters ($Y\ D_a\ X$), iff (1.3) and $F_1(y) \geq G_1(y)$ for all real y,
- by risk lovers ($Y\ D_l\ X$), iff (1.4) and $F^1(y) \leq G^1(y)$ for all real y.

If both X and Y have finite means μ_X, μ_Y and variances σ_X^2, σ_Y^2, Y is said to be not less preferred than X

- on mean and variance by risk averters ($Y\ D_{mva}\ X$), iff $\mu_X \leq \mu_Y$ and $\sigma_X^2 \geq \sigma_Y^2$.
- on mean and variance by risk lovers ($Y\ D_{mvl}\ X$), iff $\mu_X \leq \mu_Y$ and $\sigma_X^2 \leq \sigma_Y^2$.

The set

$E_g := \{F_Y \in F \mid \text{for any } F_X \in F\ (X\ D_g\ Y \Rightarrow Y\ D_g\ X)\}$

is called the g-efficient set of F; in the same way we define E_a, E_l, E_{mva}, and E_{mvl}.

The three relations D_g, D_a, and D_l are named stochastic dominance relations in the literature;see [5] and [1]. Each of them is reflexive, transitive, and antisymmetric, i.e. a semiordering on F, while D_{mva} and D_{mvl} are reflexive and transitive, but in general not antisymmetric. As it is easily derived from the definitions

(1.5) $Y\ D_g\ X \Rightarrow Y\ D_a\ X$ and $Y\ D_l\ X$;

(1.6) $Y\ D_a\ X \Leftrightarrow -X\ D_l\ -Y$.

The following lemma establishes the connection between the stochastic dominance relations and utility theory by specifying for each of the relations the proper class of utility functions u under which the expected utility of X does not exceed that of Y :

(1.7) $E(u(X)) \leq E(u(Y))$

For given X and Y let $U = U(X,Y)$ be the set of real functions for which both expectations exist and are finite.

1.1. Lemma Let X and Y have finite expectations.

(i) $Y\, D_g\, X$ iff (1.7) holds for every non-decreasing $u \in U$,

(ii) $Y\, D_a\, X$ iff (1.7) holds for every $u \in U$ which is non-decreasing and concave

(iii) $Y\, D_l\, X$ iff (1.7) holds for every $u \in U$ which is non-decreasing and convex.

Note that u is not restricted to differentiable functions. 1.1 (i) and (ii) have been proved in [6]; (iii) may be derived from (ii) by (1.6). Lemma 1.1 yields immediately : if $Y\, D\, X$ by one of the three stochastic dominance relations, then $EX \leq EY$ (provided they are finite). Further, under proper transformations of X and Y each of the relations is preserved:

1.2 Lemma Let f be a non-decreasing real function.

(i) If $Y\, D_g\, X$, then $f(Y)\, D_g\, f(X)$,

(ii) If $Y\, D_a\, X$, f concave, then $f(Y)\, D_a\, f(X)$,

(iii) If $Y\, D_l\, X$, f convex, then $f(Y)\, D_l\, f(X)$.

Proof: Assume, e.g., $Y\, D_a\, X$; if f and g are non-decreasing and concave functions then also $g \circ f$. If $g \in U(f(X),f(Y))$, then $g \circ f \in U(X,Y)$, hence $E((g \circ f)(X)) \leq E((g \circ f)(Y))$, and by Lemma 1.1 (ii) $f(Y)\, D_a\, f(X)$.

Higher degree stochastic dominance relations can be easily defined in terms of the F_k and F^k. They have simular properties which shall not be discussed here.

2. Theoretical relations between the criteria

As it is well known in theory, without taking further assumptions the preference criterion D_a of risk averters does not coincide with the preference criterion D_{mva} of mean-variance risk averters, as well as D_l does not coincide with D_{mvl}. If we impose no restrictions on the distribution law of the portfolios considered, the two MV-criteria are rational in terms of expected utility, if and only if no other than quadratic utility functions are admitted. In other words if $\mathcal{F}$ is the set of all distribution functions, then U must be restricted to the set of quadratic functions; since this restriction seems to be very unsatisfactory, we may do the opposite: reduce $\mathcal{F}$ to a class, which makes economic sense or by which real data can be fitted, and retain U as the set of all utilities which are integrable with respect to any $F \in \mathcal{F}$.

We will show that in certain linear (μ,σ)-families the stochastic dominance relations coincide with the respective mean-variance relations. This will be an extension of work by Schneeweiß [16], Hanoch/Levy [6], and Ali [2]. For risk averters on gamma-distributed portfolios the implication

$Y\ D_{mva}\ X \Rightarrow Y\ D_a\ X$ will be proved. Some findings on lognormal distributions and other transformed families will conclude this section.

2.1 Lemma

Let X and Y be random variables with finite expectations μ_X and μ_Y; let F_X and F_Y be their respective distribution functions. If there exists $y_0 \in [-\infty,\infty]$ such that:

$$(2.1)\quad \begin{cases} F_X(y) \geq F_Y(y) & \text{for all } y < y_0 \text{ and} \\ F_X(y) \leq F_Y(y) & \text{for all } y \geq y_0 \end{cases}$$

then

(i) $\mu_X \leq \mu_Y$ iff $Y\ D_a\ X$,

(ii) $\mu_Y \leq \mu_X$ iff $X\ D_1\ Y$

The "if " parts follow from Lemma 1.1 above. Part (i) for finite y_0 is due to Hanoch/Levy [6, p. 314]; (ii) can be shown by similar means. The case $y_0 = \infty$ means $Y\ D_g\ X$, hence $\mu_X \leq \mu_Y$, $Y\ D_a\ X$ and $Y\ D_1\ X$; from $\mu_Y \leq \mu_X$ follows $\mu_X = \mu_Y$ and (because $F \geq G$) $F = G$, $X\ D_1\ Y$. The case $y_0 = -\infty$ is similar.

Any criterion which is based on mean and variance alone, of course, can discriminate only between distributions which differ at least in mean or variance. This leads to the following definition:

Let F be a family of distribution functions, and let F_X denote the distribution function of a random variable X. F is called a (μ,σ)-family, if for any $F \in F$ the second moment exists and if F can be parametrized by μ and σ, i.e. if for any two F_X, F_Y $\mu_X = \mu_Y$ and $\sigma_X = \sigma_Y$ together imply $F_X = F_Y$. F is called linear, if for any two F_X, $F_Y \in F$ there are numbers α and β, $\alpha \geq 0$, such that $F_X = F_{\alpha Y+\beta}$ or, alternatively, $F_Y = F_{\alpha X+\beta}$.

2.2 Lemma

F is a linear (μ,σ)-family, iff the second moments exist and there exists a distribution function H (not necessarily in F) such that for any $F_X \in F$ the variance of which does not vanish there holds

$$(2.2)\quad F_{\frac{X-\mu_X}{\sigma_X}} = H$$

Instead of (2.2) we can write

$$(2.3)\quad F_X(y) = H\left(\frac{y-\mu_X}{\sigma_X}\right) \quad \text{for all } y.$$

The proof is obvious. Note that F needs not be complete in the sense that it contains distributions for any $\mu \in \mathbb{R}$ and $\sigma \geq 0$; further, there may be degenerate distributions in F.

2.3 Theorem

Let F be a linear (μ,σ)-family of distribution functions. F_X and $F_Y \in F$.

(i) If $\sigma_X > \sigma_Y$, then $\mu_X \leq \mu_Y \iff Y\ D_a\ X$

$\mu_Y \leq \mu_X \iff X\ D_1\ Y$

(ii) If $\sigma_X < \sigma_Y$, then $\mu_X \le \mu_Y \Leftrightarrow Y\ D_1\ X$

$\mu_Y \le \mu_X \Leftrightarrow X\ D_a\ Y$

(iii) If $\sigma_X = \sigma_Y$, then $\mu_X \le \mu_Y \Leftrightarrow Y\ D_g\ X$

(iv) If $\mu_X = \mu_Y$, then $\sigma_X \le \sigma_Y \Leftrightarrow$

$Y\ D_1\ X$ and $X\ D_a\ Y$

From this theorem we see immediately

2.4. Corollary

If F is a linear (μ, σ)-family then

(i) $Y\ D_{mva}\ X \Rightarrow Y\ D_a\ X$

(ii) $Y\ D_{mv1}\ X \Rightarrow Y\ D_1\ X$

(iii) $Y\ D_{mva}\ X$ and $Y\ D_{mv1}\ X \Rightarrow Y\ D_g\ X$

(iv) $E_a \subset E_{mva}$

(v) $E_1 \subset E_{mv1}$

Proof of 2.3: Define H as in lemma 2.2; thus if $\sigma_X > 0$ and $\sigma_Y > 0$

(+) $F_X(y) = H(\frac{y-\mu_X}{\sigma_X})$ and $F_Y(y) = H(\frac{y-\mu_Y}{\sigma_Y})$

(i): If $\sigma_X > \sigma_Y = 0$ F_Y is the Hamilton function jumping at μ_Y; thus F_X cuts F_Y according to (2.1); hence (i) by lemma 2.1. If $\sigma_X > \sigma_Y > 0$ define $y_0 := (\mu_X\sigma_Y - \mu_Y\sigma_X)/(\sigma_Y - \sigma_X)$; then

$$\frac{y-\mu_X}{\sigma_X} < \frac{y-\mu_Y}{\sigma_Y} \quad \text{for } y > y_0 \quad \text{and}$$

$$\frac{y-\mu_X}{\sigma_X} > \frac{y-\mu_Y}{\sigma_Y} \quad \text{for } y < y_0 \ ;$$

when applying the nondecreasing function H on these two inequalities and using (+) we conclude

$$F_X(y) \ge F_Y(y) \quad \text{for } y < y_0$$

$$F_X(y) \le F_Y(y) \quad \text{for } y \ge y_0,$$

which yields (i) by Lemma 2.1 .

(ii) is seen from (i), when X and Y are interchanged.

(iii): Assume $\sigma_X = \sigma_Y = 0$; then F_X and F_Y are Hamilton functions. $\mu_X \le \mu_Y \Rightarrow F_X \ge F_Y \Rightarrow Y\ D_g\ X$. Now assume $\sigma_X = \sigma_Y > 0$. $\mu_X \le \mu_Y$ implies

$$\frac{y-\mu_X}{\sigma_X} \ge \frac{y-\mu_Y}{\sigma_Y} \quad \text{for all } y;$$

when we apply H on both sides and insert (+), then $F_X(y) \ge F_Y(y)$ for all y, hence $Y\ D_g\ X$. By interchanging X and Y we see: $\mu_Y \le \mu_X \Rightarrow X\ D_g\ Y$. - The reverse is always true.

(iv): Assume $\mu_X = \mu_Y$. From $\sigma_X \le \sigma_Y$ follow $Y\ D_1\ X$ and $X\ D_a\ Y$ by (ii) and (iii). For the reverse direction, assume $Y\ D_1\ X$, $X\ D_a\ Y$, and $\sigma_X > \sigma_Y$; from $\mu_X = \mu_Y$ deduce $Y\ D_a\ X$ and $X\ D_1\ Y$ by (i), whence $F_X = F_Y$ which implies $\sigma_X = \sigma_Y$; contradiction.

2.5 Theorem

Let $\mathcal{F}$ be a linear (μ,σ)-family, F_X and $F_Y \in \mathcal{F}$, and H defined as above.

(i) When $\sigma_X > 0$ and $H(y) > 0$ for every y, $Y\ D_{mva}\ X \Leftrightarrow Y\ D_a\ X$

(ii) When $\sigma_Y > 0$ and $H(y) < 1$ for every y, $Y\ D_{mvl}\ X \Leftrightarrow Y\ D_l\ X$

(iii) When $\sigma_X > 0$ and $\sigma_Y > 0$ and $0 < H(y) < 1$ for every y, then $Y\ D_{mva}\ X$ and $Y\ D_{mvl}\ X \Leftrightarrow Y\ D_g\ X$

Proof: The assertions of necessity have already been shown.
(i): From $Y\ D_a\ X$ there follows $\mu_X \le \mu_Y$ immediately; we have still to show $\sigma_X \ge \sigma_Y$. Recall the notation of (1.1). By definition, $Y\ D_a\ X$ means $(F_X)_1(y) \ge (F_Y)_1(y)$ for all y, which by inserting (2.3) and substitution of variables is equivalent to (provided σ_Y is positive)

(2.4)
$$\sigma_X H_1\left(\frac{y-\mu_X}{\sigma_X}\right) \ge \sigma_Y H_1\left(\frac{y-\mu_Y}{\sigma_Y}\right) \quad \text{for all } y.$$

If we assume $\sigma_X < \sigma_Y$, (2.4) implies

$$H_1\left(\frac{y-\mu_X}{\sigma_X}\right) \ge H_1\left(\frac{y-\mu_Y}{\sigma_Y}\right) \quad \text{for all } y.$$

As $H(y) > 0$ for all y, H_1 is a strictly increasing function; we conclude $(y-\mu_X)/\sigma_X \ge (y-\mu_Y)/\sigma_Y$ for all y, which yields necessarily $\sigma_X = \sigma_Y$; contradiction.

Thus $\sigma_X \ge \sigma_Y$. - If $\sigma_Y = 0$, $\sigma_X \ge \sigma_Y$ holds trivially.

(ii) can be shown by an analogous argument or can be derived from (i) by the use of (1.6).

(iii) $Y\ D_g\ X$ gives $Y\ D_a\ X$ and $Y\ D_l\ X$, hence by (i) and (ii) $Y\ D_{mva}\ X$ and $Y\ D_{mvl}\ X$. Q.e.d.

$H(y) > 0$ for all y means that the random variables belonging to $\mathcal{F}$ are not bounded below or are constant (i.e. a safe asset).

2.6 Corollary

Let $\mathcal{F}$ be a linear (μ,σ)-family.

(i) When $\mathcal{F}$ contains no random variables bounded below (i.e. $0 < F(y)$ for all $y \in \mathbb{R}$, $F \in \mathcal{F}$), then $\mathcal{E}_{mva} = \mathcal{E}_a$.

(ii) When $\mathcal{F}$ contains no random variables bounded above (i.e. $F(y) < 1$ for all $y \in \mathbb{R}$, $F \in \mathcal{F}$), then $\mathcal{E}_{mvl} = \mathcal{E}_l$.

Remarks: 1) The assumption of nonvanishing variances in 2.5 may not be dropped: E.g., let be $X = x_0$ a.s., hence F_X a Hamilton function jumping at x_0 and $\sigma_X = 0$, let F_Y be a distribution function with $\sigma_Y > 0$ and $F_Y(y) = 0$ for $y < x_0$. Then $Y\ D_g\ X$, therefore $Y\ D_a\ X$, but $Y\ D_{mva}\ X$ is not true.

2) Parts (iii) and (iv) of theorem 2.3 have been proved by Schneeweiß for linear (μ,σ)-families which are complete [16,p.125]; his further discussion applies only to situations in which there exist indifference curves in the (μ,σ)-plane.

3) Hanoch and Levy have stated a similar theorem [6,p.343], from which they conclude that $Y\ D_a\ X \Leftrightarrow Y\ D_{mva}\ X$ if $\mathcal{F}$ is a "family of distributions with two parameters which are independent functions of μ and σ, respectively" and if "the two distributions

intersect". The latter proposition is not correct, but has been repeated in subsequent work ([9],[10],[15],and [1]).

4) The two-stage criterion as introduced by [10] coincides with D_a and D_{mva} when the (μ,σ)-family is linear and contains no random variables bounded below. Then

$$E_{\text{two stage}} := \{F_X \in E_g \mid \text{not } Y \, D_{mva} \, X \quad \forall F_Y \in E_g, \; F_X \neq F_Y\}$$

$$= E_{mva} = E_a.$$

Before we shall go beyond the linear (μ,σ)-families, we give some examples. The family of normal distributions is a prominent linear (μ,σ)-family but not at all the single one to be applied. It may be supplemented by the class of one-point distributions which forms a linear (μ,σ)-family by itself. Other examples are the family of simple alternatives and the families of triangle and rectangle distributions or any subset of them. The members of a linear (μ,σ)-family need not to be symmetric distributions; for any unsymmetric distribution can generate such a family (e.g., define $X := \alpha\chi^2_r + \beta$, $\alpha > 0$, $\beta \in \mathbb{R}$, where χ^2_r is a chi-square variable with $r \geq 3$ degrees of freedom).

Other (μ,σ)-families such as the class of gamma and the class of lognormal distributions are not linear, but of great practical interest. Testing the changes of German stock market prices by several goodness-of-fit tests Ronning [14] found out that gamma and lognormal distributions fit the data much better than normal distributions do. This is partly due to the fact that the cash flows cannot be less than zero.

Let F_Γ be the family of gamma distribution functions defined by

$$(2.5) \quad F_{b,p}(y) = \frac{b^p}{\Gamma(p)} \int_0^y e^{-bx} x^{p-1} dx \quad \text{for } y \geq 0,$$

$F_{b,p}(y) = 0$ else, with parameters $b > 0$ and $p > 0$. Mean and variance are given by $\mu = p/b$ and $\sigma^2 = p/b^2$, hence $\mu > 0$ and $\sigma > 0$ holds. On the other hand b and p can be recomputed from μ and σ by the formulae $b = \mu/\sigma^2$ and $p = \mu^2/\sigma^2$. As there is a one-to-one correspondence between (μ,σ) and (b,p), F_Γ is a (μ,σ)-family, although no linear one. Nevertheless, the D_a-efficient set of F_Γ is included in the D_{mva}-efficient set; thus, when using the simple mva-relation on F_Γ no stochastic dominance efficient portfolio is eliminated.

2.7 Theorem

If F_X and $F_Y \in F_\Gamma$, then

(i) $\mu_X \leq \mu_Y$, $\sigma_X \geq \sigma_Y \Rightarrow Y \, D_a \, X$

(ii) $\mu_X = \mu_Y$, $\sigma_X \leq \sigma_Y \Rightarrow Y \, D_1 \, X$

Proof: Denote F_X by $F_{b,p}$ and F_Y by $F_{c,q}$. First we show: if $p \leq q$, then the graph of $F_{b,p}$ cuts the graph of $F_{c,q}$ at most once, and from above; i.e. F_X and F_Y meet the assumption (2.1) of Lemma 2.1.

Restrict $F_{b,p}$ on the nonnegative axis $\mathbb{R}_+$ and let $F^{-1}_{b,p}$ be its inverse function. Van Zwet has shown [18,p.60 f] that for $p \leq q$ the function $F^{-1}_{1,q} \circ F_{1,p}$ is concave on $\mathbb{R}_+$. Hence $F^{-1}_{1,p} \circ F_{1,q}$ is convex and, because of $F_{b,p}(y) = F_{1,p}(by)$, for arbitrary b and c> 0 the function $h := F^{-1}_{b,p} \circ F_{c,q}$ is convex, too. As $h(0) = 0$, the graph of the identity function g, $g(x) := x$, cuts the graph of h at most once, and from above; since $F_{b,p}$ is a nondecreasing function, the same applies to $F_{b,p} \circ g = F_{b,p}$ and $F_{b,p} \circ h = F_{c,q}$.

(i): If $\mu_X \leq \mu_Y$ and $\sigma_X \geq \sigma_Y$, then $\mu_X^2/\sigma_X^2 \leq \mu_Y^2/\sigma_Y^2$, i.e. $p \leq q$. By the above and Lemma 2.1 we conclude $Y\ D_a\ X$.

(ii):If $\mu_X = \mu_Y$ and $\sigma_X \leq \sigma_Y$, then $p \geq q$. F_X and F_Y meet the assumptions of Lemma 2.1 with X and Y being interchanged. Hence $Y\ D_1\ X$. Q.e.d.

Remark on proof: Ali [2] has proved by other means that $Y\ D_a\ X \iff q/p \geq \mathrm{Max}\{1,c/b\}$, wherefrom (i) can be derived, too.

The previous theorem shows that a (μ,σ)-family F, on which D_{mva} implies D_a, needs not be linear. If, in addition, D_{mvl} implies D_1, it may be conjectured that F must necessarily be a linear family.

The assumption of <u>lognormally distributed</u> portfolios looks sound not only for fitting the data but by theoretical considerations, too, as the return on a portfolio can be considered as a product of many independent multiplicative effects. For risk averters, Levy [7] has demonstrated that $Y\ D_{mva}\ X \Rightarrow Y\ D_a\ X$, when X and Y are lognormal variables; moreover, $Y\ D_a\ X \iff \mu_X \leq \mu_Y$ and $\tilde{\sigma}_X \geq \tilde{\sigma}_Y$, where $\tilde{\sigma}_X$ and $\tilde{\sigma}_Y$ denote the standard deviation of $\log X$ and $\log Y$, respectively. For risk lovers, a modified criterion is rational:

<u>2.8 Theorem</u>

Let F_X and F_Y be lognormal distribution functions, i.e. let

$\tilde{X} := \log X$ and $\tilde{Y} := \log Y$

be normally distributed, $\tilde{X} \in N(\tilde{\mu}_X,\tilde{\sigma}_X)$, $\tilde{Y} \in N(\tilde{\mu}_Y,\tilde{\sigma}_Y)$. Then

$$\tilde{\mu}_X \leq \tilde{\mu}_Y,\quad \tilde{\sigma}_X \leq \tilde{\sigma}_Y \implies Y\ D_1\ X.$$

<u>Proof:</u> Using theorem 2.5 (ii) we derive $\tilde{Y}\ D_1\ \tilde{X}$, hence by Lemma 1.2 (iii) $Y\ D_1\ X$. - More generally, there holds:

<u>2.9 Theorem</u>

Let G be a linear (μ,σ)-family, f a given non-decreasing real function, $F_f := \{F_X | X = f(\tilde{X}),\ F_{\tilde{X}} \in G\}$, and $F_X, F_Y \in F_f$.

(i) If f is convex, then

$$\tilde{Y}\ D_{mvl}\ \tilde{X} \implies Y\ D_1\ X.$$

(ii) If f is concave, then

$$\tilde{Y}\ D_{mva}\ \tilde{X} \implies Y\ D_a\ X.$$

<u>3. Empirical aspects</u>

Let $X_1, X_2, \ldots, X_n$ be the subsequent returns on a given portfolio. When the random walk hypothesis (for references see [14]) applies, the X_i are independent random variables; if in addition the distributions can be considered to be stable in time, hypotheses on the distribution of $X = X_i$ may be tested by the usual chi-quare or Kolmogorov methods [14]. In particular, in order to justify the use of MVA on a set F of portfolios the elements of F may be tested whether they belong to a given class of distributions like one of the (μ,σ)-families mentioned in the theorems 2.4 to 2.9 above.

When in addition the returns on different portfolios can be taken as stochastically independent (which is of course not often found in reality) we may test directly the hypothesis that all $F_X \in F$ belong to the same linear (μ,σ)-family or, equivalenty by lemma 2.2, that the standardized variables $(X - \mu_X)/\sigma_X$ have identical distributions for all $F_X \in F$. In order to test this we may employ standard nonparametric tests on the identity of distributions either by testing pairs of variables or by testing the portfolios simultaneously, provided their number is moderate. If the hypothesis is rejected we may transform all return variables X by a nondecreasing convex transformation and test again; if there is no rejection we may use the criterion of theorem 2.9 (ii).

In the remainder we shall discuss some recent empirical findings on efficient sets of portfolios, i.e. the empirical distribution functions $\hat{F}_X$ of their returns. Table 1 synopsises five empirical studies: The number of portfolios in F and the sizes of the various efficient sets E_g, E_a, and E_{mva} relative to F as well as the size of the set $E_a \cap E_{mva}$ of portfolios which are both a- and mva-efficient. Moreover, for each computation the number of observations is given.

The studies [9], [15], and [1] investigate the annual returns of U.S. open end mutual funds, which in general are composed of many diverse stocks. Hence, the distribution of their annual rates of return can be supposed to be not too far from normal.

In fact, Sarnat [15] when applying a Kolmogorov-Smirnov test cannot reject the hypothesis that the rates of return on these funds are normally distributed, at a five percent level. In spite of this and theorem 2.5 above Sarnat's a-efficient sets come out to be much larger than his mva-efficient ones. In the studies [9] and [1], which are similar, E_a and E_{mva} differ more or less, too, while in most cases E_a is the larger one. These unsatisfactory outcomes are presumably due to the small number of portfolios considered.

In another paper [13] Porter and Gaumnitz demonstrate that if F is large and sufficiently dense with respect to μ and σ (at least at the border) the two efficient sets come very close. Beginning with the real data of 925 stocks they generate at random a set F of 893 portfolios. For this set the following approximately holds:

(i) any portfolio which is a-efficient is also mva-efficient, i.e. $E_a \subset E_{mva}$.

(ii) the portfolios which are mva-efficient but not a-efficient are those with rather low mean and variance.

More precisely, when being transformed into the (μ,σ)-plane the set E_a is contained in a close neighbourhood $U(E_{mva})$ of E_{mva}, and $U(E_{mva}) \setminus E_a$ is contained in a small circle around the origin.

Looking at study [12] in table 1 we see that the sets E_g and E_a enlarge considerably when the number n of observations is increased by taking quarterly instead of semi-annual data; E_{mva} increases only slightly when n is increased. The same

TABLE 1 Relative sizes of efficient sets in various empirical studies

		No. of observations[1)]	No.of portfolios in F	Percentage[2)] of portfolios in E_g	E_a	E_{mva}	$E_a \cap E_{mva}$
Levy/Sarnat [9]	1946-67	22	58	ca.85	29,3	20,7	15,5
	1958-67	12	149	ca.60	12,1	14,1	7,4
Sarnat [15]	1946-57	12	56		21,4	16,1	
	1958-69	12	56		23,2	16,1	
Porter/Gaumnitz [13]	1960-65	24	893	22,2	4,5	4,4	2,7
Porter [12]	1960-65 (quarterly data)	24		75,7	14,2	7,3	
	1960-65 (semi-annual data)	12		45,2	9,2	6,3	
Aharony/Loeb [1]	1947-72	26	56	91,1	33,9	25,0	21,4
	1947-59	13	56	71,4	25,0	26,8	19,6
	1960-72	13	56	87,5	21,4	17,9	16,1

1) annual data except of [13] and [12].
2) in % of the number of portfolios in F.

effect can be observed in [9] and [1]; it is a general property of stochastic dominance efficient sets which are computed from empirical distribution functions:

Let F_X be a g-efficient portfolio, i.e. there is no $F_Y \in F$ such that $F_Y \neq F_X$ and $Y\ D_g\ X$, i.e. for any $F_Y \in F$, $F_Y \neq F_X$, there is a $y \in \mathbb{R}$ with $F_X(y) < F_Y(y)$. With the empirical $\hat{F}_X$ and $\hat{F}_Y$ instead of F_X and F_Y this inequality will occur more probably when n is enlarged. By an analogous argument E_a and E_1 increase with n.

Hence, if we base a stochastic dominance criterion on a broad data set the efficient set will be large; in particular, E_a will be much larger than E_{mva}. However, a large efficient set does not give much help in practice, and computing it may be not worthwile since it does not eliminate enough portfolios from further considerations.

But anyway, in practice n cannot be large. If we go far back into history the distribution of the portfolio will not be the same as today and if we reduce distances of observations down to daily observations or below, the observations will not be independent. As it has been shown in the literature (see [14]) the distribution of X can be considered as stationary only over short spans of time. So the efficient set under any ot the above criteria is not stable in time, too, which is confirmed by [9] and [15]. Whether there are differences in the degree of stability between E_a and E_{mva} needs further investigation.

Other problems to be investigated are: a more complete description of the classes of portfolio distributions on which mean-variance and stochastic-dominance criteria imply each other, and the statistical properties of dominance decisions based on empirical distribution functions.

References

[1] Aharony J. and Loeb M, 1977. Mean-variance vs. stochastic dominance: some empirical findings on efficient sets, J.Banking and Finance 1, pp 95-102

[2] Ali M.M., 1975. Stochastic Dominance and Portfolio Analysis. J.Financial Economics 2, pp 205-229.

[3] Borch, K., 1969. A note on uncertainty and indifference curves, Rev.Econ.Studies 36,pp 1-4.

[4] Feldstein M.S., 1969. Mean-variance analysis in the theory of liquidity preference and portfolio selection, Rev.Econ.Studies 36, pp 5-12.

[5] Hadar J. and Russel W.R., 1969. Rules for ordering uncertain propects, Amer.Econ. Review 59, pp 25-34.

[6] Hanoch, G. and Levy H.,1969. The efficiency analysis of choices involving risk, Rev.Econ.Studies 36, pp 335-346.

[7] Levy H., 1973. Stochastic dominance among log-normal prospects, Internat.Econ.Review 14, pp 601-614.

[8] Levy H. and Hanoch G., 1970. Relative effectiveness of efficiency criteria for portfolio selection, J.Financial Quant.Analysis 5, pp 63-76.

[9] Levy H. and Sarnat M., 1970. Alternative efficiency criteria: an empirical analysis, J.Finance 25, pp 1153-1158.

[10] Levy H. and Sarnat M., 1972. Investment and portfolio analysis, Wiley, New York, 604 pp.

[11] Markowitz H.M., 1959. Portfolio selection, Wiley, New York,

[12] Porter R.B., 1973. An empirical comparison of stochastic dominance and mean-variance portfolio choice criteria, J.Financial Quant. Analysis 8, pp 587-608.

[13] Porter, R.B. and Gaumnitz J.E., 1972. Stochastic dominance vs. mean-variance portfolio analysis: an empirical evaluation, Amer.Econ.Review 62, pp 438-446.

[14] Ronning G., 1974. Das Verhalten von Aktienkursveränderungen. Eine Überprüfung von Unabhängigkeits- und Verteilungshypothesen anhand von nichtparametrischen Testverfahren, Allg.Statist. Archiv 58, pp 272-302.

[15] Sarnat M., 1972. A note on the prediction of portfolio performance from ex post data, J.Finance 27, pp 903-906.

[16] Schneeweiß, H. 1967. Entscheidungskriterien bei Risiko, Springer,Berlin, 214 pp.

[17] Tobin J., 1957/8. Liquidity preference as behaviour towards risk, Rev.Econ.Studies 25, pp 65-86.

[18] Van Zwet W.R., 1964. Convex transformations of random variables. Diss.math.-nat., Mathematisch Centrum, Amsterdam.

[19] Whitmore G.A., 1970. Third-degree stochastic dominance, Amer.Econ.Review 60, pp 457-459.