Stochastic Orders and Decision under Risk
IMS Lecture Notes – Monograph Series (1991)

SOME THEORY OF STOCHASTIC DOMINANCE

BY KARL MOSLER* AND MARCO SCARSINI*†

Universität der Bundeswehr, Hamburg, and Università D'Annunzio

Three different types of stochastic dominance relations are considered: set dominance, kernel dominance and higher degree dominance. The connections between these definitions are examined. Preservation results are given and implications between joint and marginal dominance are studied in the finite and infinite dimensional setting.

1. Introduction. For distributions on the real line, two basic stochastic orderings have been of interest to researchers in many fields: stochastic dominance with respect to all increasing functions and stochastic dominance with respect to all convex functions. These orderings can be characterized by shifts and dilations, respectively, or, in the first case, by inequalities for distribution functions, and, in the second case, by inequalities for integrals of the distribution functions. Beginning with these characterizations several attempts have been made to unify the theory of stochastic dominance relations in d-dimensional and more general spaces (Brumelle and Vickson (1975), Fishburn and Vickson (1978), Stoyan (1977) (1983), Mosler (1982)). Based on this tradition, the primary aim of this paper is to investigate three different ways by which stochastic dominance relations on several spaces may be characterized. The first one is the characterization via probability inequalities for certain families of sets. Orderings which allow for this kind of characterization are named set dominance orderings. The second approach employs Markov kernels to define a stochastic ordering. These orderings are called kernel dominance orderings. Third, inequalities on integrals of distribution functions are

* The authors thank Yosi Rinott for processing this paper and Bruno Bassan and two anonymous referees for their valuable comments.

† This paper was written while the second author was a member of GNAFA-CNR.

AMS 1980 Subject Classification: 90A05, 90A10, 60E15

Key words and phrases: Set dominance, kernel dominance, higher degree stochastic dominance, preservation of stochastic orderings, copula, dominance for stochastic processes

used to characterize ordering realations which we call higher degree stochastic dominance orderings. The second part of the paper concentrates on three more special questions: under which transformations of the underlying random variables are the above orderings preserved? Under which circumstances is an ordering of distributions on a product space implied by the same ordering of all marginal distributions? How does the ordering of the finite marginals of a process extend to an ordering of the process?

The paper is mainly expository and certain caveats apply. No complete survey of the literature is intended. The results presented strongly reflect the taste of the authors and their past work. Most of the results are not new. Where applicable, proofs are omitted and references to the literature are provided. Where proofs are given they serve partly to illustrate the exposition and partly to support new results.

The paper is organized as follows. Section 2 treats set dominance, Section 3 kernel dominance and Section 4 higher degree dominance. Section 5 is devoted to some preservation results and Section 6 contains a relatively comprehensive presentation of marginal vs. global dominance when the copula is fixed. Section 7 sketches dominance for stochastic processes.

2. Set Dominance. Consider a set $\mathcal{P}$ of probability measures on a measurable space $(\Omega, \mathcal{S})$. The space $\mathcal{P}$ can be endowed with a (partial) preorder $\precsim_{\mathcal{A}}$ defined as follows: for $P_1, P_2 \in \mathcal{P}$, $P_1 \precsim_{\mathcal{A}} P_2$ if and only if $P_1(A) \leq P_2(A)$ $\forall A \in \mathcal{A}$, where $\mathcal{A} \subset \mathcal{S}$.

Some of the most usually encountered orders on spaces of probability measures are of this form (which we will call set dominance).

For any probability measure $P \in \mathcal{P}$, $P(A) = \int_\Omega I_A(\omega)\ P(d\omega)$. Therefore, if $P_1(A) \leq P_2(A), \forall A \in \mathcal{A}$, then

$$\int_\Omega \sum_i \alpha_i I_{A_i}(\omega)\ P_1(d\omega) \leq \int_\Omega \sum_i \alpha_i I_{A_i}(\omega)\ P_2(d\omega), \qquad \alpha_i \geq 0,\ A_i \in \mathcal{A},$$

by linearity of the expectation and

$$\int_\Omega \phi\ dP_1 \leq \int_\Omega \phi\ dP_2 \tag{2.1}$$

for all ϕ such that there exists a sequence $\phi_n \nearrow \phi$, with $\phi_n(\cdot) = \sum_{i=1}^n \alpha_{i,n} I_{A_{i,n}}(\cdot)$, $\alpha_{i,n} \geq 0$, $A_{i,n} \in \mathcal{A}$, by the Lebesgue monotone convergence theorem. In other words, (2.1) holds for all functions ϕ in the closed convex cone generated by the indicator functions of sets $A \in \mathcal{A}$.

In the sequel, given a class $\mathcal{F}$ of measurable functions $\phi : \Omega \to \mathbb{R}$, for $P_1, P_2 \in \mathcal{P}$, we will write $P_1 \precsim^{\mathcal{F}} P_2$ if and only if $\int_\Omega \phi \, dP_1 \leq \int_\Omega \phi \, dP_2 \quad \forall \phi \in \mathcal{F}$ for which both integrals exist. Again $\precsim^{\mathcal{F}}$ is a (partial) pre-order on $\mathcal{P}$. It is usually called stochastic dominance with respect to $\mathcal{F}$. Provided the class $\mathcal{A}$ (resp. $\mathcal{F}$) is rich enough, the pre-order $\precsim_{\mathcal{A}}$ (resp. $\precsim^{\mathcal{F}}$) is actually an order (see Alfsen (1971, p. 22), Mosler (1982, Theorem 4.1)).

We consider some examples. In all the examples the space Ω will always be a Polish space, and $\mathcal{S}$ the Borel σ-field generated by the open sets. When further structure is necessary, it will be specified. In the examples $\mathcal{A}_i$ will be a class of sets in $\mathcal{S}$ and $\mathcal{F}_i$ will be the convex cone of functions generated by indicators of sets $A \in \mathcal{A}_i$.

Example 2.1. Let $(\Omega, \leq)$ be a partially ordered Polish space (POPS), where $\leq$ is closed. A subset A of the partially ordered space is called upper if $x \in A$ and $x \leq y$ imply $y \in A$. An upper set of a POPS is measurable. Let $\mathcal{A}_1$ be the class of upper sets of Ω. Then $\mathcal{F}_1$ is the cone of increasing functions. In the whole paper "increasing" means "nondecreasing". The order $\precsim_{\mathcal{A}_1}$ is the usual stochastic dominance order (Lehmann (1955)), or first degree stochastic dominance. In this generality it has been studied by Kamae, Krengel and O'Brien (1977).

Example 2.2. Let Ω be as in Example 2.1. For $x \in \Omega$ the set $A_x = \{y : x \leq y\}$ is an upper interval. Call $\mathcal{A}_2$ the class of upper intervals of Ω. When $\Omega = \mathbb{R}^d$, the order $\precsim_{\mathcal{A}_2}$ is equivalent to the order obtained by comparing the survival functions corresponding to P_1 and P_2. If we define $\Delta_{s=x}^{y} \phi(\ldots, s, \ldots) = \phi(\ldots, y, \ldots) - \phi(\ldots, x, \ldots)$, then $\mathcal{F}_2$ is the cone of functions ϕ such that

$$\mathop{\Delta}_{s_{i_1}=x_{i_1}}^{y_{i_1}} \cdots \mathop{\Delta}_{s_{i_k}=x_{i_k}}^{y_{i_k}} \phi(\ldots, s_{i_1}, \ldots, s_{i_k}, \ldots) \geq 0$$

$$\forall \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\}, \ \forall x_{i_j} \leq y_{i_j}, \ j = 1, \ldots, k$$

and for all the values of the other arguments of ϕ whose index is not in $\{i_1, \ldots, i_k\}$. This order has been studied under different conditions by Cambanis, Simons and Stout (1976), Tchen (1980), Rüschendorf (1980), Mosler (1984), Scarsini (1988a).

Example 2.3. Endow a POPS Ω with a linear space structure and define $\mathcal{A}_3$ as the class of upper convex sets. We note that $\mathcal{F}_3$ contains all the increasing quasi-concave functions (by definition, a function ϕ is quasi-concave if the set $\{x : \phi(x) \geq \alpha\}$ is convex for all $\alpha \in \mathbb{R}$). The set of increasing quasi-concave functions is not convex, though: in general the sum of two increasing quasi-concave functions is not quasi-concave. The case $\Omega = \mathbb{R}^d$ has been studied by Levhari, Paroush and Peleg (1975) and Bergmann (1991).

EXAMPLE 2.4. Let Ω be a POPS with an inner product structure. A set $A_{a,\alpha} \subset \Omega$ is a half-space if it has the form $A_{a,\alpha} = \{x : \langle a, x\rangle \geq \alpha\}$. Call $\mathcal{A}_4$ the class of upper half-spaces. If

$$\phi(x) = v(\langle a, x\rangle), \quad \text{with } v : \mathbb{R} \to \mathbb{R} \text{ increasing, and } a \geq 0,\ a \neq 0, \tag{2.2}$$

then $\phi \in \mathcal{F}_4$, but the set of functions defined in (2.2) is not convex. This order has been studied by Scarsini (1986) and Muliere and Scarsini (1989), and has been used to compare random cash flows or bundles of commodities, when the price vector is not fixed.

EXAMPLE 2.5. Consider $\Omega = \mathbb{R}^d$ with the Schur ordering $\leq_S$, which is a closed pre-order, and let $\mathcal{A}_5$ be the class of Schur-convex sets (a set A is Schur-convex if I_A is a Schur convex function). Then $\mathcal{F}_5$ is the class of Schur-convex functions (see Nevius, Proschan and Sethuraman (1977), Marshall and Olkin (1979)).

When Ω is a linear space, another relation can be defined in terms of a dual family of functions as follows. For given $\mathcal{F}, \mathcal{A}$, define

$$\mathcal{F}^{\text{dual}} = \{\psi : \forall x \in \Omega, \quad \psi(x) = -\phi(-x), \quad \phi \in \mathcal{F}\},$$

$$\mathcal{A}^{\text{dual}} = \{-\overline{A} : A \in \mathcal{A}\}, \quad \text{where } \overline{A} = \Omega \setminus A, \quad \text{and } -A = \{x : -x \in A\}.$$

Then,

$$\mathcal{F}_1^{\text{dual}} = \mathcal{F}_1, \quad \mathcal{F}_4^{\text{dual}} = \mathcal{F}_4, \quad \mathcal{F}_5^{\text{dual}} = \mathcal{F}_5.$$

$$\mathcal{F}_2^{\text{dual}} = \{\psi : (-1)^{k+1} \mathop{\Delta}_{s_{i_1}=x_{i_1}}^{y_{i_1}} \cdots \mathop{\Delta}_{s_{i_k}=x_{i_k}}^{y_{i_k}} \psi(\ldots, s_{i_1}, \ldots, s_{i_k}, \ldots) \geq 0$$
$$\forall \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\},\ \forall x_{i_j} \leq y_{i_j}\},$$

$$\mathcal{F}_3^{\text{dual}} \supset \{\psi : \psi \text{ is increasing and quasi-convex}\},$$

The class $\mathcal{F}_2^{\text{dual}}$ has interesting economic applications, since it contains the utility functions that represent multivariate risk aversion (see Section 4 for details).

Again, the class $\mathcal{F}_i^{\text{dual}}$ is generated by the indicator functions of sets in $\mathcal{A}_i^{\text{dual}}$, where

$$\mathcal{A}_1^{\text{dual}} = \mathcal{A}_1, \qquad \mathcal{A}_4^{\text{dual}} = \mathcal{A}_4, \qquad \mathcal{A}_5^{\text{dual}} = \mathcal{A}_5,$$

and

$$\mathcal{A}_2^{\text{dual}} = \{B : \Omega \setminus B \text{ is a lower interval}\},$$

$$\mathcal{A}_3^{\text{dual}} = \{B : \Omega \setminus B \text{ is a lower convex set}\}.$$

EXAMPLE 2.6. Let Ω be a linear space, and let $\mathcal{A}_6$ be the class of convex sets symmetric about the origin. $\mathcal{F}_6$ is the cone of central unimodal functions (see Anderson (1955), Sherman (1955), and for recent results Dharmadhikari and Joag-Dev (1988), Bergmann (1991), Eaton and Perlman (1991)).

EXAMPLE 2.7. Let Ω be a normed space and let $\mathcal{A}_7$ be the class of sets $A_\alpha = \{x : \|x\| \leq \alpha\}$. Then $\mathcal{F}_7$ is the cone of decreasing in norm functions (Rüschendorf (1981)).

EXAMPLE 2.8. Let $\Omega = \Omega_1 \times \Omega_1$, where (Ω_1, ρ) is a metric space. Let $\mathcal{A}_8$ be the class of sets $A_\delta = \{\mathbf{x} : \rho(x_1, x_2) \leq \delta\}$. Then $\mathcal{F}_8 \ni \phi : \Omega \to \mathbb{R}$, if and only if $\phi(\mathbf{x}) = v(\rho(x_1, x_2))$, with v decreasing. This ordering has some useful applications in the study of variability and is connected to some variability measures like Gini's mean difference.

EXAMPLE 2.9. Let Ω be arbitrary and ψ a given function $\Omega \to \mathbb{R}$. Consider the class $\mathcal{A}_9$ of sets $\{x : \psi(x) \geq \alpha\}$, $\alpha \in \mathbb{R}$. Then $\mathcal{F}_9$ is the family of all functions $v \circ \psi$ with $v : \mathbb{R} \to \mathbb{R}$ increasing. Obviously Examples 2.7 and 2.8 are special cases of this ordering. The ordering has applications in multivariate choice under risk with deterministic preferences (Kihlstrom and Mirman (1974), Levy and Levy (1984)).

If $P_1 \precsim_{\mathcal{B}} P_2$, implies $P_1 \precsim_{\mathcal{A}} P_2$, we write $\precsim_{\mathcal{A}} \supset \precsim_{\mathcal{B}}$. In general, if $\mathcal{A} \subset \mathcal{B}$, then $\precsim_{\mathcal{A}} \supset \precsim_{\mathcal{B}}$. Thus $\precsim_{\mathcal{A}_j} \supset \precsim_{\mathcal{A}_1}$ for $j = 2, 3, 4$. Furthermore $\precsim_{\mathcal{A}_4} \supset \precsim_{\mathcal{A}_3}$, $\precsim_{\mathcal{A}_2} \supset \precsim_{\mathcal{A}_3}$ and $\precsim_{\mathcal{A}_7} \supset \precsim_{\mathcal{A}_6}$.

For a broad exposition about set dominance relations we refer to Bergmann (1991).

3. Kernel Dominance. Consider a measurable space $(\Omega, \mathcal{S})$ as above. We investigate situations where a probability distribution dominates another one if and only if it is the other one's transform through a proper Markov kernel. For $P_1 \in \mathcal{P}$ and a Markov kernel M,

$$MP_1(A) = \int M(x, A)\, dP_1(x) \qquad A \in \mathcal{S},$$

defines a probability distribution $MP_1 \in \mathcal{P}$. The discrete case is simple: Let $P_1 \in \mathcal{P}$ and S be a finite or countable set containing the support of P_1,

$S = \{s_1, s_2, \ldots\}$. Define

$$M(x, \{y\}) = \begin{cases} \pi_{ij}, & \text{if } x = s_i \text{ and } y = s_j, \\ 1, & \text{if } x \notin S \text{ and } y = x, \\ 0, & \text{otherwise,} \end{cases}$$

with

$$\pi_{ij} \geq 0 \qquad \forall i, j, \tag{3.1}$$

$$\sum_j \pi_{ij} = 1 \qquad \forall i. \tag{3.2}$$

Then M is a Markov kernel,

$$MP_1(\{s_j\}) = \sum_i \pi_{ij} P_1(\{s_i\}) \qquad \forall j, \tag{3.3}$$

and MP_1 has support contained in S, again. On the other hand, if M is a Markov kernel and P_1 and MP_1 have finite or countable support contained in S, there must be numbers π_{ij} such that (3.1), (3.2) and (3.3) hold for M.

For two Markov kernels N, M denote

$$(N \circ M)(x, A) = \int N(y, A)\, M(x, dy),$$

hence $(N \circ M)P = N(MP)$, and let E denote the unit kernel, $EP = P, \quad \forall P$. We consider a family $\mathcal{M}$ of Markov kernels which is closed under $\circ$ and contains E. We say that $P_1 \overset{\mathcal{M}}{\precsim} P_2$ if and only if

$$P_2 = MP_1 \quad \text{for some } M \in \mathcal{M}.$$

The relation $\overset{\mathcal{M}}{\precsim}$ defines a pre-order in $\mathcal{P}$. For instance, any Markov semigroup $\mathcal{M} = \{M_t : t \in \mathbb{R}_+\}$ containing E may serve.

In the following examples, $(\Omega, \mathcal{S})$ is a Polish space bearing some additional structure as specified.

EXAMPLE 3.1. Let Ω be endowed with a closed partial order $\leq$ and let $\mathcal{M}_1$ be the family of all upward Markov kernels, i.e., the kernels such that $M(x, A_x) = 1$, $\forall x$, with $A_x = \{y : x \leq y\}$. This again yields the usual stochastic dominance order which, in our framework, is both a set dominance and a kernel dominance order. See Example 2.1 for references.

EXAMPLE 3.2. Let $\mathcal{M}_5$ denote the set of upward kernels on an Euclidean d-space, endowed with the Schur-ordering $\leq_S$, which is a closed pre-order. See Example 2.5.

EXAMPLE 3.3. Let Ω be a linear space and let $\mathcal{M}_{10}$ be the set of kernels M such that

$$x = \int y \; M(x, dy) \qquad \forall x \in \Omega. \tag{3.4}$$

The ordering $\overset{\mathcal{M}_{10}}{\precsim}$ is called dilation ordering.

More generally, let $\mathcal{F}$ be a family of measurable functions $\Omega \to \mathbb{R}$. Let ϵ_x be the degenerate probability measure at x. A Markov kernel M is called an $\mathcal{F}$-diffusion if, for all x,

$$\epsilon_x \precsim^{\mathcal{F}} M\epsilon_x = M(x, \cdot), \quad \text{i.e.,}$$

$$\phi(x) \leq \int \phi(y) \; M(x, dy) \qquad \forall \phi \in \mathcal{F}. \tag{3.5}$$

We denote by $\mathcal{M}_{\mathcal{F}}$ the family of all $\mathcal{F}$-diffusions; we say that $\mathcal{M}_{\mathcal{F}}$ is generated by $\mathcal{F}$. It is easily seen that E is in $\mathcal{M}_{\mathcal{F}}$, and $N \circ M$ belongs to $\mathcal{M}_{\mathcal{F}}$ whenever N and M do. In the sequel, for the sake of brevity, we will write $\mathcal{M}_i$ for $\mathcal{M}_{\mathcal{F}_i}$.

It is easy to show that $\mathcal{M}_1, \mathcal{M}_5$ and $\mathcal{M}_{10}$ are special cases of $\mathcal{F}$-diffusions: for $\mathcal{M}_1$ choose $\mathcal{F}_1' = \{I_A : A \text{ upper set}\}$ or $\mathcal{F}_1 = \{\phi : \phi \text{ increasing}\}$, analogously for $\mathcal{M}_5$. The class $\mathcal{M}_{10}$ is generated by $\mathcal{F}_{10} = \{\phi : \phi \text{ convex}\}$ as well as by $\mathcal{F}_{10}' = \{\phi : \phi \text{ affine}\}$. We continue with two more examples, which are related to the dilation ordering $\overset{\mathcal{M}_{10}}{\precsim}$.

EXAMPLE 3.4. Endow Ω with a closed partial order and a linear space structure and let $\mathcal{F}_{11} = \{\phi : \phi \text{ increasing and affine}\}$. Observe that $\mathcal{F}_{11} \subset \mathcal{F}_{10}$, hence $\mathcal{M}_{10} \subset \mathcal{M}_{11}$; the dilation ordering $\overset{\mathcal{M}_{10}}{\precsim}$ is contained in $\overset{\mathcal{M}_{11}}{\precsim}$.

EXAMPLE 3.5. Consider the set $\mathcal{F}_{12}$ of functions on $\mathbb{R}^d$ which are convex and permutation symmetric, i.e. $\phi(x) = \phi(x_\rho)$ for every $x \in \mathbb{R}^d$ and every permutation ρ. Obviously $P_1 \overset{\mathcal{M}_{10}}{\precsim} P_2$ implies $P_1 \overset{\mathcal{M}_{12}}{\precsim} P_2$.

The following theorem exhibits the equivalence between kernel dominance $\overset{\mathcal{M}}{\precsim}$ and stochastic dominance $\precsim^{\mathcal{F}}$ for the above cases when $\mathcal{M}$ is an $\mathcal{F}$-diffusion. The result is due to Kamae, Krengel and O'Brien (1977) for $k = 1$, Nevius, Proschan and Sethuraman (1977) for $k = 5$, Strassen (1965), who generalized a famous theorem by Hardy, Littlewood and Polya (1967) and others, for $k = 10$ and 11, Rüschendorf (1981) for $k = 12$. The cases $k = 7, 8$ and 9 are similar.

THEOREM 3.1. *Let $k \in \{1, 5, 7, 8, 9, 10, 11, 12\}$ and assume $\Omega = \mathbb{R}^d$, when $k \neq 1$. Then $P_1 \precsim^{\mathcal{F}_k} P_2$ if and only if $P_1 \overset{\mathcal{M}_k}{\precsim} P_2$,* PROOF. Sufficiency is proved as follows. Let $P_1 \overset{\mathcal{M}_k}{\precsim} P_2$, $\phi \in \mathcal{F}_k$. Then for some $M \in \mathcal{M}_k$,

$$\int \phi \, dP_2 = \int \phi \, dMP_1 = \int \int \phi(y) \, M(x, dy) \, P_1(dx) \geq \int \phi \, dP_1,$$

because of (3.5). Necessity can be shown by proper application of Strassen's (1965, Theorem 3) result. ∎

When Ω is compact and $\mathcal{F}$ contains only continuous functions, Theorem 3.1 follows from a general result proved in Meyer (1966, chapter XI, T53): Let Ω be compact, let $\mathcal{F}$ be a cone of continuous functions closed under maximum formation and containing the positive constants. Then $P_1 \precsim^{\mathcal{F}} P_2$ if and only if $P_1 \overset{\mathcal{M}_{\mathcal{F}}}{\precsim} P_2$.

We observe that the cones $\mathcal{F}_2$ to $\mathcal{F}_4$ and $\mathcal{F}_6$ of Section 2 are not closed under the maximum operation, whereas $\mathcal{F}_1$, $\mathcal{F}_5$, $\mathcal{F}_7$ and $\mathcal{F}_8$ are. As we see, $\overset{\mathcal{M}_1}{\precsim}$, $\overset{\mathcal{M}_5}{\precsim}$, $\overset{\mathcal{M}_7}{\precsim}$ and $\overset{\mathcal{M}_8}{\precsim}$ are orderings which are both set dominance and kernel dominance relations, whereas $\overset{\mathcal{M}_{10}}{\precsim}$ is an example which is not a set dominance relation, since $\mathcal{F}_{10}$, the cone of convex functions, cannot be generated by any set of indicator functions. On the other hand, there exist examples of set dominance which are not kernel dominance relations, e.g. $\precsim^{\mathcal{F}_3}$.

In several important cases an $\mathcal{F}$-diffusion allows for a pointwise characterization of $\mathcal{M}_{\mathcal{F}}$-dominance (for proofs, see, e.g., Rüschendorf (1981) when $k = 5, 10, 11, 12$, Kamae, Krengel and O'Brien (1977) when $k = 1$. The cases $k = 7, 8$ and 9 are proved similarly.).

THEOREM 3.2. *Let $k \in \{1, 5, 7, 8, 9, 10, 11, 12\}$. Assume that $\Omega = \mathbb{R}^d$ when $k \neq 1$ and that P_1, P_2 have first moments when $k = 10, 11, 12$. Then $P_1 \overset{\mathcal{M}_k}{\precsim} P_2$ if and only if there exist two Ω-valued random variables X, Y such that $\mathcal{L}(X) = P_1$, $\mathcal{L}(Y) = P_2$ and, P-a.s.,*

$$X \leq Y \qquad \textit{if} \quad k = 1,$$

$$X \leq_S Y \qquad \textit{if} \quad k = 5,$$

$$\|X\| \geq \|Y\| \qquad \textit{if} \quad k = 7,$$

$$\rho(X_1, X_2) \geq \rho(Y_1, Y_2), \quad \textit{where } X = (X_1, X_2) \textit{ and } Y = (Y_1, Y_2), \qquad \textit{if} \quad k = 8,$$

$$\psi(X) \leq \psi(Y) \qquad \textit{if} \quad k = 9,$$

$$X = E(Y|X) \qquad \textit{if} \quad k = 10,$$

$$\phi(X) \le E(\phi(Y)|X) \quad \textit{for all } \phi \textit{ increasing affine, if} \quad k = 11,$$
$$X \le_{\mathrm{S}} E(Y_{(\,)}|X) \qquad \textit{if} \quad k = 12,$$

where $Y_{(\,)}$ denotes the downward ordered random vector.

For given discrete distributions it is possible to check whether $P_1 \overset{\mathcal{M}_i}{\precsim} P_2$, i.e. $P_2 = MP_1$ for some $M \in \mathcal{M}_i$, as follows. In view of (3.1), (3.2), (3.3) we have to search for numbers $\pi_{ij} \ge 0$ which fulfill

$$\sum_j \pi_{ij} = 1 \quad \forall i,$$

$$P_2(\{s_j\}) = \sum_i \pi_{ij} P_1(\{s_i\}) \quad \forall j,$$

plus a condition from Theorem 3.2 which corresponds to (3.4). For instance, when $k = 1$, this condition reads

$$\sum_{\{j: s_i \le s_j\}} \pi_{ij} = 1 \quad \forall i,$$

similarly ($\le$ being the Schur ordering $\le_{\mathrm{S}}$ in $\mathbb{R}^d$) with $k = 5$. With $k = 10$ and 11, the additional conditions are

$$\sum_j \pi_{ij} s_j = s_i$$

and

$$\sum_j \pi_{ij} \psi(s_j) \ge \psi(s_i) \quad \text{for all increasing affine } \psi$$

respectively. When $\Omega = \mathbb{R}^d$, these two cases can be easily solved by linear programming methods. For $k = 10$, Shaked (1980) gives a numerical example while Kemperman (1973) determines the π_{ij}'s when P_2 is a binomial and P_1 is a hypergeometric distribution.

4. Higher Degree Dominances. When $\Omega = \mathbb{R}$ the orders $\precsim_{\mathcal{A}_1}$, $\precsim_{\mathcal{A}_2}$, $\precsim_{\mathcal{A}_3}$, $\precsim_{\mathcal{A}_4}$, are equivalent and assume a very simple form: $P_1 \precsim_{\mathcal{A}_1} P_2$ if and only if

$$F_{P_1}(x) \ge F_{P_2}(x) \qquad \forall x \in \mathbb{R},$$

or, equivalently,

$$\overline{F}_{P_1}(x) \le \overline{F}_{P_2}(x) \qquad \forall x \in \mathbb{R},$$

where, for $i = 1, 2$, F_{P_i} is the distribution function and $\overline{F}_{P_i}$ is the survival function associated with P_i: $F_{P_i}(x) = P_i\{(-\infty, x]\}$, $\overline{F}_{P_i}(x) = 1 - F_{P_i}(x)$.

This is not the case when $\Omega = \mathbb{R}^d$, $d > 1$. Let $(\mathbf{a}, \mathbf{b}) = \times_{i=1}^d (a_i, b_i)$. Define

$$\mathcal{G}_n^+(\mathbb{R}^d) = \Big\{ \phi : \phi(\mathbf{x}) = \int_{(-\infty,\mathbf{x})} \prod_{i=1}^d (x_i - t_i)^{n-1}\ \mu(d\mathbf{t}) + c, \quad \text{where } c \in \mathbb{R},$$

$$\text{and } \mu \text{ is a positive measure on } \mathbb{R}^d \text{ such that } \int_{\mathbb{R}^d} \prod_{i=1}^d |x_i^n|\ \mu(d\mathbf{x}) < \infty. \Big\},$$

$$\mathcal{G}_n^-(\mathbb{R}^d) = \Big\{ \phi : \phi(\mathbf{x}) = -\int_{[\mathbf{x},\infty)} \prod_{i=1}^d (t_i - x_i)^{n-1}\ \mu(d\mathbf{t}) + c, \quad \text{where } c \in \mathbb{R},$$

$$\text{and } \mu \text{ is a positive measure on } \mathbb{R}^d \text{ such that } \int_{\mathbb{R}^d} \prod_{i=1}^d |x_i^n|\ \mu(d\mathbf{x}) < \infty. \Big\},$$

For $m > n$, we have $\mathcal{G}_n^+ \supset \mathcal{G}_m^+$ and $\mathcal{G}_n^- \supset \mathcal{G}_m^-$.

Let P be a probability measure on $(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d))$. Let F_P and $\overline{F}_P$ be the distribution function and the survival function associated with P, respectively:

$$F_P(\mathbf{x}) = P\{\times_{i=1}^d (-\infty, x_i]\}, \qquad \overline{F}_P(\mathbf{x}) = P\{\times_{i=1}^d (x_i, \infty)\}.$$

Define

$$F_P^1 = F_P, \qquad \overline{F}_P^1 = \overline{F}_P,$$

$$F_P^n(\mathbf{x}) = \int_{(-\infty,\mathbf{x}]} F_P^{n-1}(\mathbf{t})\ d\mathbf{t}, \qquad \overline{F}_P^n(\mathbf{x}) = \int_{(\mathbf{x},\infty)} \overline{F}_P^{n-1}(\mathbf{t})\ d\mathbf{t}.$$

The following theorem holds.

THEOREM 4.1. *Let P_1, P_2 be probability measures on* $(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d))$.

(a) $P_1 \precsim^{\mathcal{G}_n^+} P_2$ *if and only if* $\overline{F}_{P_1}^n(\mathbf{x}) \le \overline{F}_{P_2}^n(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^d$,

(b) $P_1 \precsim^{\mathcal{G}_n^-} P_2$ *if and only if* $F_{P_1}^n(\mathbf{x}) \ge F_{P_2}^n(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^d$.

PROOF. (a) Given any d-dimensional survival function $\overline{F}_P$, the following representation holds

$$\overline{F}_P^n(\mathbf{x}) = [(n-1)!]^{-d} \int_{(\mathbf{x},\infty)} \prod_{j=1}^d (t_j - x_j)^{n-1}\ dF_P(\mathbf{t}). \tag{4.1}$$

If $\phi(\mathbf{t}) = [(n-1)!]^{-d} \prod_{j=1}^d (t_j - x_j)_+^{n-1}$, then $\phi \in \mathcal{G}_n^+$, therefore $P_1 \precsim^{\mathcal{G}_n^+} P_2 \Longrightarrow \overline{F}_{P_1}^n(\mathbf{x}) \le \overline{F}_{P_2}^n(\mathbf{x})\ \forall \mathbf{x} \in \mathbb{R}^d$.

We have to prove the converse. If $\overline{F}^n_{P_1}(\mathbf{x}) \leq \overline{F}^n_{P_2}(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^d$, then

$$\int_{\mathbf{R}^d} \overline{F}^n_{P_1}(\mathbf{x})\ \mu(d\mathbf{x}) \leq \int_{\mathbf{R}^d} \overline{F}^n_{P_2}(\mathbf{x})\ \mu(d\mathbf{x}),$$

when μ is a positive measure. Then, by (4.1),

$$\int_{\mathbf{R}^d} [(n-1)!]^{-d} \int_{(\mathbf{x},\infty)} \prod_{j=1}^{d} (t_j - x_j)^{n-1}\ dF_{P_1}(\mathbf{t})\ \mu(d\mathbf{x}) \leq$$

$$\int_{\mathbf{R}^d} [(n-1)!]^{-d} \int_{(\mathbf{x},\infty)} \prod_{j=1}^{d} (t_j - x_j)^{n-1}\ dF_{P_2}(\mathbf{t})\ \mu(d\mathbf{x})$$

and, by Fubini's Theorem,

$$[(n-1)!]^{-d} \int_{\mathbf{R}^d} \int_{(-\infty,\mathbf{t})} \prod_{i=1}^{d} (t_i - x_i)^{n-1}\ \mu(d\mathbf{x})\ dF_{P_1}(\mathbf{t}) \leq$$

$$[(n-1)!]^{-d} \int_{\mathbf{R}^d} \int_{(-\infty,\mathbf{t})} \prod_{i=1}^{d} (t_i - x_i)^{n-1}\ \mu(d\mathbf{x})\ dF_{P_2}(\mathbf{t}).$$

The proof of part (b) is analogous. ∎

The univariate case has been studied by Rolski and Stoyan (1974) and Rolski (1976) in the case of $n \in \mathbb{N}$. Fishburn (1976), (1980a) employed fractional integrals to study the general case of $n \in [1, \infty)$. The multivariate case with $n = 2$ has been investigated by Bergmann (1978) and Mosler (1982), (1984). The bivariate case has been studied by Scarsini (1985).

The conditions used to define classes $\mathcal{G}_n^+$ and $\mathcal{G}_n^-$ have an economic meaning in terms of multivariate risk aversion. We start with the univariate case and show how the utility functions in $\mathcal{G}_n^-$ can be characterized in terms of preferences among lotteries. The lotteries characterizing the class $\mathcal{G}_n^-$ are defined recursively in terms of the lotteries characterizing $\mathcal{G}_{n-1}^-$.

Consider a decision maker whose utility functions ϕ is in the class $\mathcal{G}_1^-(\mathbb{R})$. For every $x \in \mathbb{R}$, for every $h > 0$ and for every pairs of lotteries

$\mathrm{L}_1(x) = x$ with probability 1 $\qquad$ $\mathrm{M}_1(x) = x + h$ with probability 1

she will prefer M_1 over L_1. Every decision maker whose utility function is in

$\mathcal{G}_n^-(\mathbb{R})$ prefers M_k over L_k, $\forall k \in \{1,\ldots,n\}$, where

$$L_k(x) = \begin{cases} L_{k-1}(x) & \text{w.p. } 1/2 \\ M_{k-1}(x+h) & \text{w.p. } 1/2 \end{cases} \qquad M_k(x) = \begin{cases} M_{k-1}(x) & \text{w.p. } 1/2 \\ L_{k-1}(x+h) & \text{w.p. } 1/2 \end{cases}$$

To be precise, we should make explicit the dependence of the lotteries upon h, but we want to avoid cumbersome notation. Preference of M_2 over L_2 corresponds to concavity of the utility function, i.e. to risk aversion (Pratt (1964)).

The construction for the multivariate case is similar in its structure to the univariate one. We describe the bivariate case more extensively. Consider $\phi \in \mathcal{G}_1^-(\mathbb{R}^2)$. It satisfies

$$\mathop{\Delta}_{s_1=x_1}^{x_1+h_1} \mathop{\Delta}_{s_2=x_2}^{x_2+h_2} \phi(s_1,s_2) \leq 0 \qquad \forall x_1,x_2 \in \mathbb{R},\ \forall h_1,h_2 > 0.$$

which is equivalent to the preference of lottery M_1 over lottery L_1, where

$$L_1(x_1,x_2) = \begin{cases} (x_1,x_2) & \text{w. p. } 1/2 \\ (x_1+h_1, x_2+h_2) & \text{w. p. } 1/2, \end{cases}$$

$$M_1(x_1,x_2) = \begin{cases} (x_1,x_2+h_2) & \text{w. p. } 1/2 \\ (x_1+h_1, x_2) & \text{w. p. } 1/2, \end{cases}$$

(where again the dependence on $\mathbf{h}$ has not been made explicit). This preference represents bivariate risk aversion (see Richard (1975)). Bivariate risk aversion is different and independent from risk aversion à la Arrow-Pratt. It has nothing to see with concavity of the utility function and is defined only in terms of different combinations of commodities.

A utility function in $\mathcal{G}_n^-(\mathbb{R}^2)$ implies preference of M_k over L_k, $\forall k \in \{1,\ldots,n\}$, where

$$L_k(x_1,x_2) = \begin{cases} L_{k-1}(x_1,x_2) & \text{w. p. } 1/4 \\ M_{k-1}(x_1+h_1,x_2) & \text{w. p. } 1/4 \\ M_{k-1}(x_1,x_2+h_2) & \text{w. p. } 1/4 \\ L_{k-1}(x_1+h_1,x_2+h_2) & \text{w. p. } 1/4 \end{cases}$$

$$\mathrm{M}_k(x_1, x_2) = \begin{cases} \mathrm{M}_{k-1}(x_1, x_2) & \text{w. p. } 1/4 \\ \mathrm{L}_{k-1}(x_1 + h_1, x_2) & \text{w. p. } 1/4 \\ \mathrm{L}_{k-1}(x_1, x_2 + h_2) & \text{w. p. } 1/4 \\ \mathrm{M}_{k-1}(x_1 + h_1, x_2 + h_2) & \text{w. p. } 1/4. \end{cases}$$

In general, for the d-dimensional case, we have $\phi \in \mathcal{G}_n^-(\mathbb{R}^d)$ if and only if $\mathrm{M}_k(\mathbf{x})$ is preferred to $\mathrm{L}_k(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^d$, $\forall \mathbf{h} \in \mathbb{R}_+^d \setminus \{0\}$, $\forall k \in \{1, \ldots, n\}$. The lottery $\mathrm{L}_1(\mathbf{x})$ has 2^{d-1} equally likely outcomes $\mathbf{z} \in Z$ (where $\mathbf{z} \in Z$ if and only if $z_i =$ either x_i or $x_i + h_i$ and the number of $z_i = x_i + h_i$ is even); the lottery $\mathrm{M}_1(\mathbf{x})$ has 2^{d-1} equally likely outcomes $\mathbf{w} \in W$ (where $\mathbf{w} \in W$ if and only if $w_i =$ either x_i or $x_i + h_i$ and the number of $w_i = x_i + h_i$ is odd). $\mathrm{L}_k(\mathbf{x})$ has 2^d equally likely outcomes each of which is a lottery: either $\mathrm{L}_{k-1}(\mathbf{z})$, $\mathbf{z} \in Z$ or $\mathrm{M}_{k-1}(\mathbf{w})$, $\mathbf{w} \in W$. $\mathrm{M}_k(\mathbf{x})$ has 2^d equally likely outcomes each of which is a lottery: either $\mathrm{M}_{k-1}(\mathbf{z})$, $\mathbf{z} \in Z$ or $\mathrm{L}_{k-1}(\mathbf{w})$, $\mathbf{w} \in W$.

It is worth noticing that, if two lotteries that we are comparing are marginalized (i.e. if one of the commodities is omitted in the lotteries), then they become equal. The reason for using the term risk aversion for the above preferences among lotteries is that the lottery L_k always contains the worst possible outcome and it is reasonable to assume that a risk averter wants to avoid it.

For an investigation of further multivariate risk postures and stochastic dominance with respect to them, see Mosler (1987).

Necessary conditions can be established for n-th degree stochastic dominance. These conditions, which involve the moments of the distributions, have been established by Fishburn (1980b), O'Brien (1984) in the univariate case, and by O'Brien and Scarsini (1991) in the multivariate case.

5. Preservation Under Transformations. In this section we will consider random variables with values in $(\Omega', \mathcal{S}')$, which are transformations or compositions of other random variables with values in $(\Omega, \mathcal{S})$. Given a random variable X, we will denote $\mathcal{L}(X)$ by P_X.

Given a measurable tranformation $h : (\Omega, \mathcal{S}) \to (\Omega', \mathcal{S}')$, we want to determine conditions on $\mathcal{F}$ and $\mathcal{F}'$ under which

$$P_X \precsim^{\mathcal{F}} P_Y \quad \text{implies} \quad P_{h(X)} \precsim^{\mathcal{F}'} P_{h(Y)}. \tag{5.1}$$

First assume that Ω is partially ordered, $\Omega' = \mathbb{R}$, $h : \Omega \to \mathbb{R}$ is increasing and $\mathcal{F}' = \mathcal{F}_1(\mathbb{R})$. If $\mathcal{F} = \mathcal{F}_1$, then the implication is well known. In the case

$\mathcal{F}' = \mathcal{F}_1(\mathbb{R})$, Marshall (1991) gave necessary and sufficient conditions on $\mathcal{F}$ for the validity of (5.1). For general $\Omega', h, \mathcal{F}, \mathcal{F}'$ we have the following result.

THEOREM 5.1. *Let $\psi \circ h \in \mathcal{F}$, whenever $\psi \in \mathcal{F}'$. Then* (5.1) *holds.*

The proof is obvious. We list several examples in which the condition of Theorem 5.1 is met (see also Mosler (1982, pp. 78 ff.)). Ω and Ω' will be endowed with a partial ordering and/or a linear structure as necessary.

EXAMPLE 5.1. Consider an increasing transformation $h : (\Omega, \leq) \to (\Omega', \leq')$, $\mathcal{F} = \mathcal{F}_1(\Omega)$, $\mathcal{F}' = \mathcal{F}_1(\Omega')$.

EXAMPLE 5.2. Let $\Omega' = \mathbb{R}^d$ with the usual ordering, let h be a convex transformation $\Omega \to \mathbb{R}$ and let $\mathcal{F} = \mathcal{F}_{10}$, $\mathcal{F}' = \mathcal{F}_{11}$.

EXAMPLE 5.3. Let $\Omega = \mathbb{R}^d$, $\Omega' = \mathbb{R}$,

$$\mathcal{F}_{13} = \{\phi \in C^1(\mathbb{R}^d) : \partial\phi/\partial x_j \text{ is nonnegative, decreasing and convex for all } j\}.$$

A differentiable function is in $\mathcal{F}_{13}$ if and only if it is in $\mathcal{G}_3^+(\mathbb{R})$ with respect to each argument x_j. $\mathcal{F}_{13}$ includes the utility functions which are risk averse in every attribute with increasing second derivative (cf. univariate third degree stochastic dominance, Whitmore (1970)). If $\mathcal{F} = \mathcal{F}_{13}$, $\mathcal{F}' = \mathcal{F}_{13}$, and $h \in \mathcal{F}_{13}$, then the conditions of Theorem 5.1 are satisfied.

EXAMPLE 5.4. Again, let $\Omega = \mathbb{R}^d$ and $\Omega' = \mathbb{R}$. A univariate utility function $\phi \in C^2(\mathbb{R}^d)$ has decreasing absolute risk aversion (DARA) with respect to x_j if and only if $-\partial/\partial x_j \log(\partial\phi(x)/\partial x_j)$ is decreasing in x_j. Let $\mathcal{F}_{14}$ denote the set of those functions which have DARA in every argument x_j. With $\mathcal{F} = \mathcal{F}_{14}$, $h \in \mathcal{F}_{14}$ and $\mathcal{F}'$ the set of univariate DARA utility functions, (5.1) holds. For $d = 1$, see Vickson (1977).

Let Ω be a linear space, $\mathcal{F}$ a set of functions $\Omega \to \mathbb{R}$. $\mathcal{F}$ is called scale invariant if $\forall\phi \in \mathcal{F}$, $\forall\alpha > 0$, $\phi_\alpha \in \mathcal{F}$, where $\phi_\alpha(x) = \phi(\alpha x)$. $\mathcal{F}$ is translation invariant if $\forall\phi \in \mathcal{F}$, $\forall y \in \Omega$, $\phi^y \in \mathcal{F}$, where $\phi^y(x) = \phi(x + y)$. For instance, it is easy to see that $\mathcal{F}_j$ is translation and scale invariant when $j \in \{1, 2, 3, 4, 5, 10, 11, 13, 14\}$, it is scale but not translation invariant when $j \in \{6, 7, 12\}$ and is neither translation nor scale invariant when $j \in \{8, 9\}$.

THEOREM 5.2. *Assume that $\mathcal{F}$ is scale invariant and $\alpha > 0$. Then*

$$P_X \precsim^{\mathcal{F}} P_Y \quad \textit{implies} \quad P_{\alpha X} \precsim^{\mathcal{F}} P_{\alpha Y}.$$

PROOF. The result follows from Theorem 5.1. ∎

THEOREM 5.3. *Assume that $\mathcal{F}$ is translation invariant, that X and Z are independent and that Y and V are independent. Then*

$$P_X \precsim^{\mathcal{F}} P_Y, \quad P_Z \precsim^{\mathcal{F}} P_V \quad \textit{implies} \quad P_{X+Z} \precsim^{\mathcal{F}} P_{Y+V}.$$

PROOF. Let $\phi \in \mathcal{F}$. For every z, x, we have $\phi^z, \phi^x \in \mathcal{F}$. Hence, by $P_X \precsim^{\mathcal{F}} P_Y$,

$$\int \phi^z(x)\, dP_X(x) \le \int \phi^z(x)\, dP_Y(x) = \int \phi^x(z)\, dP_Y(x),$$

$$\begin{aligned}
\int \phi\, dP_{X+Z} &= \int\int \phi^z(x)\, dP_X(x)\, dP_Z(z) \\
&\le \int\int \phi^x(z)\, dP_Y(x)\, dP_Z(z) \\
&= \int\int \phi^x(z)\, dP_Z(z)\, dP_Y(x) \\
&\le \int\int \phi^x(z)\, dP_V(z)\, dP_Y(x) \\
&= \int \phi\, dP_{Y+V},
\end{aligned}$$

where the last inequality stems from $P_Z \precsim^{\mathcal{F}} P_V$. ∎

COROLLARY 5.1. *Assume that $\mathcal{F}$ is scale and translation invariant, $X_1, \ldots, X_d$ are independent random variables and $Y_1, \ldots, Y_d$ are independent, too. Let $\alpha_1, \ldots, \alpha_d > 0$, $\beta \in \mathbb{R}$. Then*

$$P_{X_i} \precsim^{\mathcal{F}} P_{Y_i} \quad i = 1, \ldots, d \quad \textit{implies} \quad P_{\beta + \sum \alpha_i X_i} \precsim^{\mathcal{F}} P_{\beta + \sum \alpha_i Y_i}.$$

The proof of Corollary 5.1 is obvious. Analogous results may be obtained for dependent random variables when orderings of conditional distributions are employed.

THEOREM 5.4. *Assume that $\mathcal{F}$ is translation invariant. Let $P_{X|z}$ and $P_{Y|z}$ denote proper versions of the conditional distributions of X and Y given $Z = z$; then*

$$P_{X|z} \precsim^{\mathcal{F}} P_{Y|z}\ \forall z \quad \textit{implies} \quad P_{X+Z} \precsim^{\mathcal{F}} P_{Y+Z}.$$

PROOF. For all $\phi \in \mathcal{F}$,

$$\begin{aligned}\int \phi \, dP_{X+Z} &= \int\int \phi^z(x)\, dP_{X|z}\, dP_Z(z) \\ &\leq \int\int \phi^z(x)\, dP_{Y|z}\, dP_Z(z) \\ &= \int \phi\, dP_{Y+Z}. \qquad \blacksquare\end{aligned}$$

Similarly, Theorem 5.4 extends to the dependent case: If $P_{X|Z=z} \precsim^{\mathcal{F}} P_{Y|Z=z}$ $\forall z$ and $P_{Z|Y=y} \precsim^{\mathcal{F}} P_{V|Y=y}$ $\forall y$, then $P_{X+Z} \precsim^{\mathcal{F}} P_{Y+V}$. More preservation results are given in the papers by Arnold (1991), Eaton and Perlman (1991) and Marshall (1991).

6. Joint and Marginal Dominance. For $i = 1, \ldots, d$, let $(\Omega_i, \mathcal{S}_i)$ be a measurable space, endowed with some additional structure (see the above examples), and let $(\Omega, \mathcal{S}) = (\times_{i=1}^d \Omega_i, \bigotimes_{i=1}^d \mathcal{S}_i)$. If $P \in \mathcal{P}(\Omega, \mathcal{S})$ has marginals $P_1, \ldots, P_d$, we write $P \in \Gamma(P_1, \ldots, P_d)$. Let $P, Q \in \mathcal{P}(\Omega, \mathcal{S}), P \in \Gamma(P_1, \ldots, P_d)$, $Q \in \Gamma(Q_1, \ldots, Q_d)$. For some $\mathcal{F}$ consider

$$P \precsim^{\mathcal{F}(\Omega)} Q \tag{6.1}$$

and

$$P_i \precsim^{\mathcal{F}(\Omega_i)} Q_i, \quad i = 1, \ldots, d. \tag{6.2}$$

THEOREM 6.1. *For $\phi \in \mathcal{F}(\Omega_i)$ define $\overline{\phi}(x_1, \ldots, x_i, \ldots, x_d) = \phi(x_i)$, $(x_1, \ldots, x_i, \ldots, x_d) \in \Omega$. Assume that $\overline{\phi} \in \mathcal{F}(\Omega)$, whenever $\phi \in \mathcal{F}(\Omega_i)$, $i = 1, \ldots, d$. Then* (6.1) *implies* (6.2).

THEOREM 6.2. *Assume that $\precsim^{\mathcal{F}}$ is a set dominance ordering, $\precsim^{\mathcal{F}} = \precsim_{\mathcal{A}}$, and that $\mathrm{pro}_i^{-1}(A) \in \mathcal{A}(\Omega)$ whenever $A \in \mathcal{A}(\Omega_i)$, $i = 1, \ldots, d$ (where pro_i denotes the i-th projection). Then* (6.1) *implies* (6.2).

The proofs are obvious. It may be easily checked that $\mathcal{A}_j$ meets the assumptions of Theorem 6.2 when $j \in \{1, 2, 3, 4, 6\}$ and that the hypotheses of Theorem 6.1 are satisfied when $j \in \{10, 11, 13, 14\}$. Hence these orderings are preserved under marginalization.

When a suitable regularity condition is assumed, n-th degree stochastic dominance is preserved under marginalization, too. The condition that insures the preservation is called "margin-regularity". For details we refer to O'Brien and Scarsini (1991).

The reverse implication ((6.2) $\Longrightarrow$ (6.1)) is in general not true. However, there exist some results if P and Q are both product measures and some weaker results when P and Q have the same dependence structure. The latter case will be presented in detail.

THEOREM 6.3. *Let* $P = P_1 \otimes \ldots \otimes P_d$, $Q = Q_1 \otimes \ldots \otimes Q_d$. *Then* (6.2) *implies* (6.1), *when* $\mathcal{F} = \mathcal{F}_j$, $j \in \{1,2,3,4,10,11,13,14\}$.

PROOF. Proofs for all j can be found in Mosler (1982). If $j \in \{1,2,3,4\}$ and $P_i(\mathcal{B}_i) = Q_i(\mathcal{B}_i)$ for all i, the theorem follows from Theorem 6.4 below. ∎

A function $C : [0,1]^d \to [0,1]$ is called a copula if it satisfies the following properties

(i) $C(x_1, \ldots, x_d) = 0$, if at least one $x_i = 0$,

(ii) $C(1, \ldots, 1, x_k, 1 \ldots, 1) = x_k$,

(iii) $\triangle_{s_1=x_1}^{y_1} \cdots \triangle_{s_d=x_d}^{y_d} C(s_1, \ldots, s_d) \geq 0 \qquad \forall\, x_i \leq y_i \quad i = 1, \ldots, d.$

(see Sklar (1959), Schweizer and Sklar (1983)).

A relation on a space Ω is called a (strict) weak order if it is negatively transitive and asymmetric. Let $\prec$ be a strict weak order. Let $b \sim a$ if $\neg(a \prec b)$ and $\neg(b \prec a)$. Then $\sim$ is an equivalence relation. Let $a \precsim b$ if $a \prec b$ or $a \sim b$. Let $(\Omega, \mathrm{Bor}(\Omega), \prec)$ be a weakly ordered Polish space (WOPS), endowed with the Borel σ-field. For $x \in \Omega$, define $A_x = \{y : x \precsim y\}$ and let $\mathcal{B} = \{A_x : x \in \Omega\}$. Then $\mathcal{B}$ is an increasing class of measurable subsets of Ω, i. e., for $x, y \in \Omega$, either $A_x \subset A_y$, if $y \precsim x$, or $A_x \supset A_y$, if $x \precsim y$.

For $i = 1, \ldots, d$, let $(\Omega_i, \mathrm{Bor}(\Omega_i), \prec_i)$ be a WOPS, and let $\mathcal{B}_i$ be the increasing class of subsets induced by $\prec_i$. Since every weak order $\prec_i$ on a space Ω_i induces an increasing class $\mathcal{B}_i$ and vice versa, we can indicate a WOPS also by $(\Omega_i, \mathrm{Bor}(\Omega_i), \mathcal{B}_i)$. Let $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_d)$. Let P be a probability measure on $\left(\times_{i=1}^d \Omega_i, \bigotimes_{i=1}^d \mathrm{Bor}(\Omega_i)\right)$, $P \in \Gamma(P_1, \ldots, P_d)$. Then there exists a copula $C_P^{\mathcal{B}}$ such that, for $A_i \in \mathcal{B}_i$ $(i = 1, \ldots, d)$,

$$P(A_1 \times \ldots \times A_d) = C_P^{\mathcal{B}}(P_1(A_1), \ldots, P_d(A_d)).$$

Define $P_i(\mathcal{B}_i)$ the range of $A \mapsto P_i(A)$, $A \in \mathcal{B}_i$. Then $C_P^{\mathcal{B}}$ is unique on $\times_{i=1}^d P_i(\mathcal{B}_i)$. Therefore $C_P^{\mathcal{B}}$ is unique on $[0,1]^d$ if $P_i(\mathcal{B}_i) = [0,1]$, for $i = 1, \ldots, d$.

Given $P_1, \ldots, P_d$, $C_P^{\mathcal{B}}$ uniquely determines $P(B)$ for every $B \in \bigotimes_{i=1}^d \sigma(\mathcal{B}_i)$, where $\sigma(\mathcal{B}_i)$ is the σ-field generated by $\mathcal{B}_i$. Therefore, if $\sigma(\mathcal{B}_i) = \mathrm{Bor}(\Omega_i)$, $i = 1, \ldots, d$, then P is uniquely determined by $C_P^{\mathcal{B}}, P_1, \ldots, P_d$. Details can be found in Scarsini (1989).

THEOREM 6.4. *If* $\times_{i=1}^d P_i(\mathcal{B}_i) = \times_{i=1}^d Q_i(\mathcal{B}_i)$, $C_P^{\mathcal{B}} = C_Q^{\mathcal{B}}$, and $\mathcal{F} = \mathcal{F}_1$ *(the class of increasing functions), then* (6.2) *implies* (6.1).

PROOF. Call $\mathcal{U}(\Omega)$ the class of upper sets in Ω. We have to prove that $P(A) \le Q(A)\ \forall A \in \mathcal{U}(\times_{i=1}^d \Omega_i)$. We start considering the case of $P_i(\mathcal{B}_i) = Q_i(\mathcal{B}_i) = [0,1], i = 1,\ldots,d$. Define

$$G_i^P(x) = P_i\{y : y \precsim_i x\}, \quad i = 1,\ldots,d$$

and

$$\mathbf{G}^P(\mathbf{x}) = \left(G_1^P(x_1),\ldots,G_d^P(x_d)\right), \qquad x_i \in \Omega_i,\ i = 1,\ldots,d.$$

$\mathbf{G}^P \colon \times_{i=1}^d \Omega_i \to [0,1]^d$. If we write $\mathbf{G}^P(A) = \{\mathbf{y} : \mathbf{y} = \mathbf{G}^P(\mathbf{x}),\ \mathbf{x} \in A\}$, then

$$P(A) = P\{\mathbf{x} : \mathbf{x} \in A\} = \mu_P(\mathbf{G}^P(A)),$$

where μ_P is the measure induced by $C_P^{\mathcal{B}}$ on $\left([0,1]^d, \mathrm{Bor}([0,1]^d)\right)$. Since $C_P^{\mathcal{B}}$ is unique on $\left(\times_{i=1}^d P_i(\mathcal{B}_i)\right)$, then μ_P is unique on $\bigotimes_{i=1}^d \sigma(P_i(\mathcal{B}_i))$. If $A \in \mathcal{U}\left(\times_{i=1}^d \Omega_i\right)$, then $\mathbf{G}^P(A) \in \mathcal{U}([0,1]^d)$. From (6.2) it follows that

$$G_i^P(x) \ge G_i^Q(x) \qquad \forall x \in \Omega_i \quad i = 1,\ldots,d \tag{6.3}$$

and therefore

$$\mathbf{G}^P(A) \subset \mathbf{G}^Q(A).$$

Since $\mu_P = \mu_Q$, then

$$\mu_P(\mathbf{G}^P(A)) \le \mu_Q(\mathbf{G}^Q(A)),$$

that is $P(A) \le Q(A)$.

When $P_i(\mathcal{B}_i) \neq [0,1]$, even if A is upper, generally $\mathbf{G}^P(A)$ is not. The proof can be adjusted to encompass also this case. For $A \in \mathcal{U}(\times_{i=1}^d(\Omega_i))$, let

$$\overline{\mathbf{G}_P(A)} = \mathop{\times}_{i=1}^{d} P_i(\mathcal{B}_i) \setminus \mathbf{G}^P(A)$$

and let

$$\mathbf{G}_{\mathcal{U}}^P(A) = \bigcup_{\substack{B \cap \overline{\mathbf{G}^P(A)} = \emptyset \\ B \in \mathcal{U}([0,1]^d)}} B.$$

If $P_i(\mathcal{B}_i) = [0,1], i = 1,\ldots,d$, then $\mathbf{G}_{\mathcal{U}}^P(A) = \mathbf{G}^P(A)$.

Let $\nu = P \circ (\mathbf{G}_P)^{-1}$. If $P_i(\mathcal{B}_i) \neq [0,1]$, then $\exists s,t \in [0,1]$ such that $C_P^{\mathcal{B}}$ is not unique on $[0,1]^{d-1} \times (s,t)$ and ν concentrates positive mass on some subset of $[0,1]^{d-1} \times \{t\}$. Consider the measure $\widetilde{\nu}$ on $([0,1]^d, \mathrm{Bor}([0,1]^d))$

such that $\widetilde{\nu} = \nu$ on $[0,1]^{d-1} \times ([0,1] \setminus (s,t))$ and the mass concentrated by ν on $[0,1]^{d-1} \times \{t\}$ is uniformly spread by $\widetilde{\nu}$ on the right over the interval $[0,1]^{d-1} \times (s,t)$.

Repeat the procedure for all the points t where the phenomenon appears (the discontinuity points of G_i^P) and for all the coordinates $1,\ldots,d$. Eventually, (in a countable number of steps) a measure ν^* with continuous marginals on $[0,1]^d$ is obtained. The distribution function corresponding to ν^* is a copula of P with respect to $\mathcal{B}$ (see Schweizer and Sklar (1983), Scarsini (1989), for this construction), and therefore

$$\nu^*(\mathbf{G}_{\mathcal{U}}^P(A)) = \nu(\mathbf{G}^P(A)) = P(A).$$

Since $\times_{i=1}^d P_i(\mathcal{B}_i) = \times_{i=1}^d Q_i(\mathcal{B}_i)$, then the distribution function associated to ν^* is also a copula of Q with respect to $\mathcal{B}$. Therefore, if (6.2) holds, i.e. if (6.3) holds, then $\mathbf{G}_{\mathcal{U}}^P(A), \mathbf{G}_{\mathcal{U}}^Q(A) \in \mathcal{U}([0,1]^d)$ and

$$\mathbf{G}_{\mathcal{U}}^P(A) \subset \mathbf{G}_{\mathcal{U}}^Q(A),$$

hence

$$\nu^*(\mathbf{G}_{\mathcal{U}}^P(A)) \leq \nu^*(\mathbf{G}_{\mathcal{U}}^Q(A)),$$

i.e.,

$$P(A) \leq Q(A) \qquad \forall A \in \mathcal{U}(\mathop{\times}_{i=1}^{d} \Omega_i). \qquad \blacksquare$$

The result was proved by Scarsini (1988b) for the case $\Omega_i = \mathbb{R}$, with the natural order.

Of course, since orderings $\precsim_{\mathcal{A}_2}$, $\precsim_{\mathcal{A}_3}$, $\precsim_{\mathcal{A}_4}$ are finer than $\precsim_{\mathcal{A}_1}$ on $\mathbb{R}^d$, but they are equivalent on $\mathbb{R}$, it is clear that, if $C_P^{\mathcal{B}} = C_Q^{\mathcal{B}}$, then (for $j = 2,3,4$), $P_i \precsim_{\mathcal{A}_j} Q_i$ $\forall i = 1,\ldots,d$, implies (6.1).

7. Dominance for Stochastic Processes. Some dominance conditions for stochastic processes can be embedded in the framework previously described. We can consider a stochastic process as a random variable with values in a suitable functional space. If we endow it with a topological structure that makes it Polish, then the above results for Polish spaces apply. In this section we will be concerned with sufficient conditions for some of the orderings. In particular we will try to express these conditions in a way that is natural for stochastic processes, for instance through the marginal laws of finite dimensional vectors drawn from the processes. For instance, let $\mathbf{X} = \{X_n, n \in \mathbb{N}\}$, $\mathbf{Y} = \{Y_n, n \in \mathbb{N}\}$ be two discrete-time real valued random processes. Let

$P_1, P_2 \in \mathcal{P}(\mathbb{R}^{\mathbb{N}}, \mathrm{Bor}(\mathbb{R}^{\mathbb{N}}))$, $P_1 = \mathcal{L}(\mathbf{X})$, $P_2 = \mathcal{L}(\mathbf{Y})$. Consider the following two conditions

$$(a) \qquad P_1 \precsim^{\mathcal{F}} P_2$$

$$(b) \qquad P_1^{\mathbf{i}} \precsim^{\mathcal{G}_d} P_2^{\mathbf{i}} \quad \forall d \in \mathbb{N},\ \forall \mathbf{i} \in \mathbb{N}^d,$$

where $\mathbf{i} = \{i_1, \ldots, i_d\} \in \mathbb{N}^d$, and $P_1^{\mathbf{i}} = \mathcal{L}(X_{i_1}, \ldots, X_{i_d})$, $P_2^{\mathbf{i}} = \mathcal{L}(Y_{i_1}, \ldots, Y_{i_d})$.

For suitable choices of $\mathcal{F}$ and $\mathcal{G}_d$, condition (b) implies (a). For instance, when $\mathcal{F}$ and $\mathcal{G}_d$ are the classes of increasing functions, the implication has been proved by Kamae, Krengel and O'Brien (1977), where the more general case of processes with values in a POPS was studied.

For the case of real valued discrete time stochastic processes, we will give a different proof of Kamae, Krengel and O'Brien's result. Our proof has the advantage of working for other types of set dominance, and not only for the usual stochastic ordering. Kamae, Krengel and O'Brien resort to Theorem 3.2. Such result does not hold for all set dominances, so, in order to achieve a general result, we have to work directly on the classes $\mathcal{A}_j$.

THEOREM 7.1. *Let* $\mathbf{X}, \mathbf{Y}$ *be discrete time real valued stochastic processes, with* $\mathcal{L}(\mathbf{X}) = P_1$, $\mathcal{L}(\mathbf{Y}) = P_2$, $\mathcal{L}(X_1, \ldots, X_d) = P_1^d$, $\mathcal{L}(Y_1, \ldots, Y_d) = P_2^d$. *If, for* $\mathcal{A} = \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_6$,

$$P_1^d \precsim_{\mathcal{A}(\mathbb{R}^d)} P_2^d \qquad \forall d \in \mathbb{N}_+, \tag{8.1}$$

then

$$P_1 \precsim_{\mathcal{A}(\mathbb{R}^{\mathbb{N}})} P_2. \tag{8.2}$$

PROOF. Let $A \in \mathcal{A}(\mathbb{R}^{\mathbb{N}})$, let $\widetilde{A}^d$ be the projection of A on $\mathbb{R}^d$ and let A^d be the cylinder generated by $\widetilde{A}^d$ ($A^d = \widetilde{A}^d \times \mathbb{R} \times \mathbb{R} \cdots$). Then $A^d \in \mathcal{A}(\mathbb{R}^{\mathbb{N}})$. By (8.1), $P_1(A^d) \le P_2(A^d)$. We have $A^d \supset A$ and $I_{A^d} \searrow I_A$. Therefore by the Lebesgue monotone convergence theorem, $P_1(A^d) \searrow P_1(A)$ and $P_2(A^d) \searrow P_2(A)$, therefore $P_1(A) \le P_2(A)$. Since this is true for any $A \in \mathcal{A}$, we obtain (8.2). ∎

Bassan and Scarsini (1991) proved the implication for the classes of convex, concave, increasing convex and increasing concave functions, but for these results an assumption of continuity of the functions in $\mathcal{F}$ must be added. Since these orderings cannot be defined in terms of any set dominance, then a different argument must be used to prove the result, namely the approximation has to be carried out on the functions of the classes $\mathcal{F}$ and $\mathcal{G}_d$, rather that on the probabilities of some sets. We refer to Bassan and Scarsini (1991) for details.

REFERENCES

ALFSEN, E.M. (1971). *Compact convex sets and boundary integrals.* Springer, Berlin.

ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6**, 170–176.

ARNOLD, B.C. (1991) Preservation and attenuation of inequality as measured by the Lorenz order. This volume.

BASSAN, B. and SCARSINI, M. (1991). Convex orderings for stochastic processes, *Comment. Math. Univ. Carolinae* **32**, 115–118.

BERGMANN, R. (1978). Some classes of semi-ordering relations for random vectors and their use for comparing covariances, *Math. Nachr.* **82**, 103–114.

BERGMANN, R. (1991). Stochastic orders and their application to a unified approach to various concepts of dependence and association. This volume.

BRUMELLE, S. L. and VICKSON, R. G. (1975). A unified approach to stochastic dominance. In Ziemba, W. T. and Vickson, R. G. (eds.), *Stochastic Optimization Models in Finance*, Academic Press, New York.

CAMBANIS, S., SIMONS, G. and STOUT, W. (1976). Inequalities for $\mathcal{E}k(X,Y)$ when the marginals are fixed. *Z. Wahrsch. Verw. Gebiete* **36**, 285–294.

DHARMADHIKARI, S. and JOAG-DEV, K. (1988). *Unimodality, Convexity, and Applications.* Academic Press, San Diego.

EATON, M. L. and PERLMAN, M. D. (1991). Multivariate probability inequalities: Convolution theorems, composition theorems and concentration inequalities. This volume.

FISHBURN, P. C. (1976). Continua of stochastic dominance relations for bounded probability distributions. *J. Math. Econom.* **3**, 295–311.

FISHBURN, P. C. (1980a). Continua of stochastic dominance relations for unbounded probability distributions. *J. Math. Econom.* **7**, 271–285.

FISHBURN, P. C. (1980b). Stochastic dominance and moments of distributions. *Math. Oper. Res.* **5**, 94–100.

FISHBURN, P. C. and VICKSON, R. G. (1978). Theoretical foundations of stochastic dominance. In Whitmore, G. A. and Findlay, M. C. (eds.), *Stochastic Dominance: An Approach to Decision Making Under Risk.* Health, Lexington, MA.

HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1967). *Inequalities*, 2nd edition. Cambridge University Press, Cambridge.

KAMAE, T., KRENGEL, U. and O'BRIEN G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**, 899–912.

KEMPERMAN, J. H. B. (1983). Moment problems for sampling without replacement. *Indag. Math.* **35**, 149–188.

KIHLSTROM, R. E. and MIRMAN, L. J. (1974). Risk aversion with many commodities. *J. Econom. Theory* **8**, 361–388.

LEHMANN, E. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26**, 399–419.

LEVHARI, D., PAROUSH, J. and PELEG, B. (1975). Efficiency analysis for multivariate distributions. *Rev. Econom. Stud.* **42**, 87–91.

LEVY, H. and LEVY, A. (1984). Multivariate decision making. *J. Econom. Theory* **32**, 36–51.

MARSHALL, A. W. (1991). Multivariate stochastic orderings and generating cones of functions. This volume.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

MEYER, P. A. (1966). *Probability and Potentials.* Blaisdell, Boston, Massachusetts.

MOSLER, K. C. (1982). *Entscheidungsregeln bei Risiko: Multivariate stochastische Dominanz.* Springer, Berlin.

MOSLER, K. C. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Sci.* **30**, 1311–1322.

MOSLER, K. C. (1987). Multiattribute risk posture and decomposition of utilities. Universität der Bundeswehr Hamburg.

MULIERE, P. and SCARSINI, M. (1989). Multivariate decisions with unknown price vector. *Econom. Lett.* **29**, 13–19.

NEVIUS, S. E., PROSCHAN, F. and SETHURAMAN, J. (1977). Schur functions in statistics II: Stochastic majorization. *Ann. Statist.* **5**, 263–273.

O'BRIEN, G. L. (1984). Stochastic dominance and moment inequalities. *Math. Oper. Res.* **9**, 475–477.

O'BRIEN, G. L. and SCARSINI, M. (1991). Multivariate stochastic dominance and moments. *Math. Oper. Res.* **16**, 382–389.

PRATT, J. M. (1964). Risk aversion in the small and in the large. *Econometrica* **32**, 122–136.

RICHARD, S. F. (1975). Multivariate risk aversion, utility independence and separable utility functions. *Management Sci.* **22**, 12–21.

ROLSKI, T. (1976). Order relations in the set of probability distribution functions and their applications in queueing theory. *Dissertationes Math.* **132**. Warszawa.

ROLSKI, T. and STOYAN, D. (1974). Two classes of semi-orderings and their application in the queueing theory. *Z. Angew. Math. Mech.* **54**, 127–128.

RÜSCHENDORF, L. (1980). Inequalities for the expectation of Δ–monotone functions. *Z. Wahrsch. Verw. Gebiete* **54**, 341–349.

RÜSCHENDORF, L. (1981). Ordering of distributions and rearrangement of functions. *Ann. Probab.* **9**, 276–283.

SCARSINI, M. (1985). Stochastic dominance with pair-wise risk aversion. *J. Math. Econom.* **14**, 187–201.

SCARSINI, M. (1986). Comparison of random cash flows. *IMA J. Math. Management* **1**, 25–32.

SCARSINI, M. (1988a). Dominance conditions for multivariate utility functions. *Management Sci.* **34**, 454–460.

SCARSINI, M. (1988b). Multivariate stochastic dominance with fixed dependence structure. *Oper. Res. Lett.* **7**, 237–240.

SCARSINI, M. (1989). Copulae of probability measures on product spaces. *J. Multivariate Anal.* **31**, 201–219.

SCHWEIZER, B. and SKLAR, A. (1983). *Probabilistic Metric Spaces.* Elsevier, New York.

SHAKED, M. (1980). On mixtures from exponential families. *J. Roy. Statist. Soc., Ser. B* **42**, 192–198.

SHERMAN, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.* **26**, 763–766.

SKLAR, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

STOYAN, D. (1977). *Qualitative Eigenschaften und Abschätzungen stochastischer Modelle.* R. Oldenbourg, München.

STOYAN, D. (1983). *Comparison Methods for Queues and Other Stochastic Models.* Wiley, Chichester.

STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36**, 423–439.

Tchen, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab.* **8**, 814–827.

Vickson, R. G. (1977). Stochastic dominance tests for decreasing absolute risk-aversion II: General random variables. *Management Sci.* **23**, 478–489.

Whitmore, G. A. (1970). Third degree stochastic dominance. *Amer. Econom. Rev.* **60**, 457–459.

Fachbereich Wirtschafts- und
Organisationswissenschaften
Universität der Bundeswehr Hamburg
Postfach 70 08 22
D-2000 Hamburg 70, Germany

Dipartimento di Metodi Quantitativi
Università D'Annunzio
Viale Pindaro 42
I-65127 Pescara, Italy