Discussion Paper No. 905

STOCHASTIC DOMINANCE

by

Ricard Torres*

September 1990

---

* MEDS Department, Northwestern University, Evanston, Illinois, U.S.A., and Universitat Pompeu Fabra, Barcelona, Spain.

The author thanks Nancy Stokey for her guidance and support. Thanks are also due to Hugo Hopenhayn, Cassius Ionescu Tulcea and Larry Jones for their comments and encouragement. Financial support under NSF Grant SES-86-06755 and from the Ministerio de Educación y Ciencia of Spain are gratefully acknowledged, as well as the hospitality of the Universitat Autònoma de Barcelona in May-June 1989.

All errors are the author's sole responsibility.

ABSTRACT

We develop in this paper a systematic study of the stochastic dominance ordering in spaces of measures. We collect and present in an orderly fashion results that are spread out in the Applied Probability and Mathematical Economics literature, and extend most of them to a somewhat broader framework. Several original contributions are made on the way. We provide a sharp characterization of conditions that permit an equivalent definition of stochastic dominance by means of *continuous* and monotone functions. When the preorder of the original space is closed, we offer an extremely simple equivalent characterization of stochastic dominance, a result of which we have found no parallel in the literature. We develop original methods that shed light into the inheritance of the antisymmetric property by the stochastic dominance ordering. We study how the topological properties of the preorder translate to the stochastic dominance preorder. A class of spaces in which monotone and continuous functions are convergence-determining is described. Finally, conditions are given that guarantee that (order) bounded stochastically monotone nets have a limit point.

INDEX

Introduction.

There has been in the last two decades a steady growth in the use by economists of dynamic stochastic models. As a consequence, new tools have been developed that permit to analyze, characterize and show the existence of well-behaved equilibria in these models. The development of new tools allows the study of more complex problems and, at the same time, a systematization and simplification of the early contributions in the field (see Stokey–Lucas–Prescott, 1989, chapters 10, 13 and 16). Monotonicity has always played a crucial role in economic modelling, and our study of the properties of orderings in spaces of measures may prove helpful in using arguments based on monotonicity in stochastic models.

Stochastic dominance has been also used in the development of game theoretic models with non-expected utility preferences (Dekel–Safra–Segal, 1989), as well as in models with Bayesian learning (see Bikhchandani–Segal–Sharma, 1990, and references therein). The results we present here might be used to extend the work that has been done in these areas to non-finite models.

Our work here is, however, exclusively focused on the development of tools that allow a systematic study of the intrinsic properties of the stochastic dominance preorder, and we do not consider any applications. We collect and present in an integrated framework many results that are spread out in a variety of publications, extend most of them to a broader setting and make a number of original contributions.

One of our aims is to work in an initial setting as general as possible, so that it encompasses most of the cases one encounters in economic modelling, and yet it is amenable to definite conclusions. In particular, we want to develop results that can be used in dealing with spaces that are not necessarily finite. This is why we have to use topological tools. On the other hand, we also want the orderings considered to be as general as possible, so most of our work concentrates on general preorders (not necessarily complete). We do not consider strict (irreflexive) orderings, but one can adapt without much difficulty our results to this case.

Stochastic dominance is a natural way of extending an order (preorder) relation defined on a topological space, to the set of its probability (or finite nonnegative) measures. We proceed to a systematic study of how the topological properties of the original preorder translate to the stochastic dominance ordering. The main contributions of this paper can be summarized as follows. We define a new concept, which turns out to be a condition that is sufficient for stochastic dominance to be characterized by monotone *and continuous* functions; this condition is also necessary, except for abnormal cases. We extend to a much broader framework previous work on two aspects. The first consists of (very useful) alternative characterizations of stochastic dominance. The second is concerned with conditions under which stochastic dominance is an order (antisymmetric), if the original space is endowed with an order. Using a representation theorem due to Aumann (1964) and Gihman and Skorohod (1979), we present a novel characterization of stochastic dominance; it is in some sense the simplest possible characterization, and therefore it can be useful for applications. We next study the topological properties of the induced ordering in the space of measures, and, as a by-product, we obtain a result that has important applications in the field of weak convergence of stochastic processes; namely, we describe a class of spaces in which increasing and continuous functions are convergence-determining. This result can be viewed as a generalization of the characterization of weak convergence by means of distribution functions in finite-dimensional euclidean spaces. Finally, we give a condition that guarantees the convergence of (order) bounded stochastically monotone sequences (nets) of measures.

The paper is organized as follows. In section 1, we set up the notation, give definitions and present certain results that are going to be used later on. We devote section 2 to a further study of topological properties of orderings. Several new results that are developed there may have some interest in their own. We also illustrate how our regularity properties are satisfied in most applications. Section 3 is concerned with the characterization of stochastic dominance by means of semicontinuous and continuous functions. In section 4, three alternative characterizations of stochastic dominance are presented. Section 5 treats the antisymmetry of the stochastic ordering. Finally, in section 6 the topological properties of stochastic dominance are analyzed, and some of its consequences are derived.

1. Definitions and Miscellaneous Results.

ORDER

Let X be a set. A **relation** or **correspondence** R on X is a subset of $X \times X$. According to widespread usage, we will write sometimes xRy or $y \in R(x)$ instead of $(x, y) \in R$.

A relation R is **reflexive** if $\forall a$, aRa. It is **antisymmetric** if aRb and bRa implies $a = b$. It is **transitive** if aRb and $bRc \Rightarrow aRc$. The relation is a **preorder** if it is reflexive and transitive, and an **order** if, in addition, it is antisymmetric. The preorder is **total** if $\forall a$ and b, aRb or bRa. We denote preorders by the symbol $\leq$ or a similar one.

It is customary to distinguish between a preorder $\leq$ and its **graph** G, defined as $\{(x,y) \in X \times X : x \leq y\}$. In general, we follow the same terminological convention with any correspondence. A **correspondence** C from a set X to a set Y is a subset of $X \times Y$. If $A \subset X$, we define $C(A) \doteq \{y \in Y : \exists x \in A, (x,y) \in C\}$, if $x \in X$, $C(x) \doteq C(\{x\})$ and if $B \subset Y$, $C^{-1}(B) \doteq \{x \in X : C(x) \cap B \neq \emptyset\}$.

Let $(X, \leq)$ and $(Y, \leq')$ be preordered spaces. A function $f : X \to Y$ is called **increasing** (**decreasing**) if $y \leq x \Rightarrow f(y) \leq' f(x)$ $(f(x) \leq' f(y))$. A set $A \subset X$ is called **increasing** (**decreasing**) if $x \in A$ and $x \leq y$ $(y \leq x) \Rightarrow y \in A$. A set or function is said to be **monotone** if it is either increasing or decreasing. A set $A \subset X$ is called **$\leq$-connected** if $x, z \in A$ and $x \leq y \leq z \Rightarrow y \in A$. A set $A \subset X$ is increasing if and only if its indicator function I_A is increasing. A set $A \subset X$ is increasing if and only if its complement A^c is decreasing. A real-valued function f is increasing iff $-f$ is decreasing. If X, Y and Z are preordered spaces, and $f : X \to Y$ and $g : Y \to Z$ are increasing functions, so is its composition $g \circ f$. The duality between increasing and decreasing sets or functions derives from the fact that, if $(X, \leq)$ is a preordered (ordered) space, so is $(X, \leq')$, where $x \leq' x' \iff x' \leq x$. This permits us to consider in most cases only increasing sets or functions, with the understanding that the same conditions or propositions hold true for their decreasing counterparts.

The collection of increasing (decreasing, $\leq$-connected) sets always contains X and $\emptyset$, and is closed under arbitrary intersections. Therefore, given an arbitrary set A, we can define $i(A)$, the smallest increasing set containing A, and we have $i(A) = \{x \in X : \exists y \in A, y \leq x\}$. In an analogous manner are defined $d(A)$ and $c(A)$, the decreasing and $\leq$-connected sets generated by A. We have $c(A) = i(A) \cap d(A) = \{y \in X : \exists x, z \in A, x \leq y \leq z\}$. As usual, we use the notation $i(x)$ and $d(x)$ for $i(\{x\})$ and $d(\{x\})$. A **preorder interval** is a set of the form $[a, b] \doteq i(a) \cap d(b) = \{x \in X : a \leq x \leq b\}$. In particular, if we view a preorder $\leq$ as a correspondence C on X defined by its graph, we have that $i(A) = C(A)$ and $d(B) = C^{-1}(B)$. We will use these notations interchangeably in the paper, we hope without danger of confusion.

Let $(X, \leq)$ be a preordered space, and let $K \subset X$. We define the **induced preorder** in K by $x \leq_K y \iff x \leq y$ in X. It follows from the definition that the increasing sets of K (endowed with the induced preorder) are exactly the intersections of K and the increasing sets of X.

If I is an arbitrary index set[1], and $(X_i, \leq_i)_{i \in I}$ is a collection of preordered spaces, we define the **product preorder** $\leq$ on $X = \Pi_{i \in I} X_i$ by $(x_i)_{i \in I} \leq (y_i)_{i \in I} \iff \forall i, x_i \leq_i y_i$. The product preorder is the largest preorder for which all projections are increasing.

A preordered set I is said to be upward filtering (or directed) if for each i and j in I, there is k in I such that $i \leq k$ and $j \leq k$. A net in a set X is a function from a directed set I into X.

TOPOLOGY AND ORDER

If X is a topological space and $A \subset X$, then $\text{int}(A)$, $\overline{A}$ and $\partial(A)$ denote, respectively, the interior, closure and boundary of A. By a **neighborhood** (nbh) of a point x, we mean a set that contains an open set to which x belongs. A topological space is called **separated** (or Hausdorff) if any two distinct points have disjoint neighborhoods. A separated space X is **completely regular** (denoted cr in the subsequent) if, given a point $x \in X$ and a neighborhood V of x, there is a continuous $f : X \to [0,1]$ such that $f(x) = 1$ and $f(V^c) = \{0\}$. A separated space is called **normal** if any two disjoint closed sets have disjoint neighborhoods. A collection of open sets is a **basis** if every open set is the union of sets belonging to that collection. A topological space is **separable** if it has a countable dense subset. A metric space is separable iff it has a countable basis. A cr space with a countable basis is metrizable. When we refer to a topological space in the subsequent, we mean a *separated* (Hausdorff) topological space.

Let X and Y be topological spaces. A real-valued function f on X is called **upper semicontinuous** (usc) if for each real a, the set $f^{-1}(-\infty, a)$ is open, and **lower semicontinuous** (lsc) if $-f$ is usc. A (nonempty-valued) correspondence $C : X \to Y$ is **upper hemicontinuous** (uhc) if for each closed F in Y, $C^{-1}(F)$ is closed, and **lower hemicontinuous** (lhc) if for each open U in Y, $C^{-1}(U)$ is open. A (nonempty-valued) correspondence is **continuous** if it is both upper and lower hemicontinuous. If Y is compact, and C a nonempty-valued correspondence with a closed graph, then C is uhc.

A preorder in a topological space is called **closed** if its graph is closed (in the product topology).

Nachbin (1965) defines concepts relating order and topology in a set, and analyzes its properties. In our study of stochastic dominance, the topological properties of the preorder

[1] From now on, all indexing sets (and, in general, *any* sets) are assumed to be nonempty, unless we explicitly state the contrary.

play an essencial role, and Nachbin's work is fundamental for all of our results. We present now several results of Nachbin which are the basis of this work, and prove some of them for the sake of completeness.

Proposition 1.1 *Let X be a topological space with a closed preorder. If K is a compact part of X, then $i(K)$ is closed.*

Proof: Let (x_i) be a net in $i(K)$ converging to x. There are elements (y_i) in K s.t. $y_i \leq x_i$, $\forall i$. Since K is compact, there is a subnet[2] (y_j) converging to some y in K. Since (x_j) converges to x as well and the preorder is closed, it follows that $y \leq x$, and thus $x \in i(K)$. ■

The previous lemma implies that preorder intervals are closed whenever the preorder is closed. Taking into account the observations we made about induced preorders, this lemma implies also that, if X and K are as above, then a subset of K is closed (open) and increasing with respect to $\leq_K$ iff it is the intersection of K and a set closed (open) and increasing in X.

We say that two preordered topological spaces $(X, \leq)$ and $(Y, \leq')$ are **isomorphic** if there is a bijective function $f : X \to Y$ that establishes a homeomorphism between X and Y and, at the same time, it is an order morphism, i.e. $a \leq b \iff f(a) \leq' f(b)$.

If X is a preordered topological space, we denote by $\mathbf{C}(X)$ the collection of continuous and bounded real-valued functions defined on X; $\mathbf{I}(X)$ are the increasing functions in $\mathbf{C}(X)$; and $\mathbf{F}(X)$ the functions in $\mathbf{I}(X)$ that take values in $[0,1]$.

We say that a preordered topological space X is **normally preordered** if [A closed and decreasing, B closed and increasing, $A \cap B = \emptyset$] implies that [there are U open and decreasing and V open and increasing, s.t. $U \cap V = \emptyset$, $A \subset U$, $B \subset V$]. Nachbin (1965, Thm1, p30) establishes the equivalence between the previous definition and the following property: whenever A and B are as above, there is f in $\mathbf{F}(X)$ s.t. $f(A) = \{0\}$ and $f(B) = \{1\}$. If X is normally preordered and the preorder is an order, we say that it is normally ordered. The following is the analog of Nachbin (1965, Thm. 4, p. 48).

Proposition 1.2 *If X is a compact space with a closed preorder, then it is normally preordered.*

Proof: We first show that given a decreasing set F and an open set U containing F, there is an open and decreasing set V such that $F \subset V \subset U$. By lemma 1.1, it suffices to take $V = \left(i(U^c)\right)^c$.

[2]See Kelley (1955) for a definition of subnet, and a proof of the property we mention.

Now, let A be closed and decreasing, B closed and increasing, and suppose that $A \cap B = \emptyset$. Since X is normal, there are disjoint open sets U and U' such that $A \subset U$ and $B \subset U'$. Applying the result above (and its parallel for increasing sets) we obtain a decreasing and open set V and an increasing and open set V' such that $A \subset V \subset U$ and $B \subset V' \subset U'$. ∎

We will be later using extensively a condition that, for closed preorders, is weaker than normality. We give it therefore a name which we think is suggestive of its nature. We call a preorder G on a topological space **separating** if, whenever $x \not\leq y$, there is $f \in \mathbf{F}(X)$ such that $f(x) > f(y)$. Equivalently, $x \leq y \iff \forall f \in \mathbf{F}(X),\ f(x) \leq f(y)$. It is clear that a separating preorder is always closed, and that if G is normal and has closed sections ($i(x)$ and $d(x)$ are closed, for each x), then it is separating. Hence, a normal preorder is separating iff it is closed.

Remark. In any closed preorder, the functions that are usc and increasing always "separate" points in the above sense. We just need to consider the collection of indicator functions $\{I_{i(x)} :\ x \in X\}$.

A separating *order* (antisymmetric) in a topological space X is said to be **completely regular** if it satisfies the following condition: $\forall x \in X$ and for each neighborhood V of x, there are f and g in $\mathbf{F}(X)$ such that $f(x) = 1,\ g(x) = 0$, and $\forall y \in X \setminus V,\ f(y) = 0$ or $g(y) = 1$.

For instance, $\mathbf{R}$, $\mathbf{R}^n$ and $\mathbf{R}^\infty$ are completely regular ordered spaces. Notice that this property implies that monotone sets generate the topology of the space.

Proposition 1.3 *The property of being a completely regular ordered space is inherited by subspaces and arbitrary products.*

Proof: If $Y \subset X$, then $f \in \mathbf{F}(X) \Rightarrow f_Y \in \mathbf{F}(Y)$, where f_Y denotes the restriction of f to Y. It follows that Y is cro if X is.

Suppose that $(X_i)_{i \in I}$ are cro, and let $X = \Pi_{i \in I} X_i$. Denote the projections by P_i. The fact that the product order is separating in X is immediate. Let $x \in X$ and let V be a neighborhood of x. Then there is a finite set $J \subset I$ and nbh's U_j of $x_j, j \in J$, such that $\cap_{j \in J} P_j^{-1}(U_j) \subset V$. For each j, let f_j and g_j satisfy the conditions stated in the definition of a cro space for x_j and U_j. Define $f =\ \min\ \{f_j \circ P_j : j \in J\}$ and $g =\ \max\ \{g_j \circ P_j : j \in J\}$. Then f and g satisfy the condition required in the definition of a cro space. ∎

Call a compact space Y a **compact ordered space**, if it is endowed with a closed *order* (antisymmetric). Nachbin formulated a generalization of the Stone-Čech compactification theorem of general topology. We are going to use the following restricted version of it.

Theorem 1.4 *X is a completely regular ordered space iff it is isomorphic to a subspace of a compact ordered space. Moreover, if X is a cro space, there is a compact ordered space $\beta(X)$, and an isomorphism ψ between X and a dense subset of $\beta(X)$ such that, for any continuous and increasing function f defined on X and taking values in a compact ordered space Y, there is a continuous and increasing function F defined on $\beta(X)$ and taking values in Y such that $f = F \circ \psi$.*

The fact that a compact ordered space is cro is proved in Nachbin (1965), Theorem 7, p. 55. The rest of the proof is based on the use of the evaluation function $\psi : X \to [0,1]^{\mathbf{F}(X)}$, and follows closely the proof of the above mentioned result of general topology. $\beta(X)$ is the closure of $\psi(X)$ in $[0,1]^{\mathbf{F}(X)}$, which is endowed with the product order and topology inherited from $[0,1]$. With these indications the proof is routine, though tedious.

TOPOLOGY AND MEASURE

In what follows, all spaces we consider are assumed to be (topologically) completely regular. Measurability in X is to be understood with respect to the Borel σ-algebra $\mathbf{B}(X)$, the one generated by the open sets in X. If Y is an arbitrary subspace, then $\mathbf{B}(Y)$ equals the trace of $\mathbf{B}(X)$ on Y, i.e. $\{A \cap Y : A \in \mathbf{B}(X)\}$. Denote by $\mathbf{K}(X)$ the collection of compact subsets of X. A finite and nonnegative measure μ is said to be a **Radon measure** if for each measurable A, $\mu(A) = \sup\{\mu(K) : K \subset A, K \in \mathbf{K}(X)\}$. If X is a complete and separable metric space (Polish space), then every finite and nonnegative measure defined on $\mathbf{B}(X)$ is a Radon measure. More generally, this is true for any separated topological space that is the continuous image of a Polish space (Suslin space). We denote by $\mathbf{M}_+(X)$ the collection of all finite nonnegative Radon measures, and by $\mathbf{P}(X)$ its subset of probability measures; they are separated topological spaces when endowed with the topology of simple convergence (weak topology) induced by $\mathbf{C}(X)$: $\mu_i \to \mu \iff \forall f \in \mathbf{C}(X),\ \mu_i(f) \doteq \int f d\mu_i \to \mu(f)$.

Let X and Y be cr spaces, and let $f : X \to Y$ and $\mu \in \mathbf{M}_+(X)$. We say that f is μ-measurable if the inverse image under f of any Borel set in Y belongs to the completion of $\mathbf{B}(X)$ with respect to μ. We say that f is **universally measurable** if it is μ-measurable, for each μ. Given any μ-measurable function f, there is always a universally measurable function f' that is equal to f μ-almost everywhere.

Let Y be an arbitrary subset of X. We say that a measure μ is **supported by** (or **concentrated in**) Y, if there is $A \subset Y$ such that $A \in \mathbf{B}(X)$ and for all $B \in \mathbf{B}(X)$, $\mu(B) = \mu(B \cap A)$. The set of measures in $\mathbf{M}_+(X)$ that are supported by Y is homeomorphic to $\mathbf{M}_+(Y)$ (Dellacherie–Meyer (1975), Theorem 58, p. 115). The **support** of a measure is the smallest *closed* set that supports (in the above sense) this measure. To avoid confusions with the terminology, we will refer very rarely to the support of a measure.

If Y is a measurable subset of X, $\mu \in \mathbf{M}_+(X)$ and $f \in \mathbf{C}(X)$, we denote by μ_Y and by f_Y the restrictions of μ and f to Y. The definitional property of a Radon measure implies that, if f is nonnegative:

$$\mu(f) \;=\; \sup\,\{\mu_K(f_K) :\; K \in \mathbf{K}(X)\}$$

Let X and Y be cr spaces. Given $\lambda \in \mathbf{M}_+(X \times Y)$, we say that μ and ν are its (respective) marginals iff for each $A \in \mathbf{B}(X)$ and $B \in \mathbf{B}(Y)$, $\mu(A) = \lambda(A \times Y)$ and $\nu(B) = \lambda(X \times B)$. Denote by H the vector subspace of $\mathbf{C}(X \times Y)$ formed by the functions of the form $(f \oplus g)(x, y) = f(x) + g(y)$, where $f \in \mathbf{C}(X)$ and $g \in \mathbf{C}(Y)$. Then μ and ν are the marginals of λ iff for each $f \oplus g$ in H, $\lambda(f \oplus g) = \mu(f) + \nu(g)$. The following results are due to Hoffmann–Jørgensen (1977)[3]:

(a) Let (λ_i) be a net in $\mathbf{M}_+(X \times Y)$ and let μ_i and ν_i be the marginals of λ_i. Suppose that there are $\mu \in \mathbf{M}_+(X)$ and $\nu \in \mathbf{M}_+(Y)$ such that $\mu_i \to \mu$ and $\nu_i \to \nu$. Then (λ_i) has a cluster point λ with marginals μ and ν.
(In the Polish case the proof is a straightforward application of Prohorov's Theorem. In the general case, it follows along the same line by applying a more refined compactness criterion.)

(b) Let Λ be a closed and convex set in $\mathbf{P}(X \times Y)$. The upper support function of Λ is defined as

$$\sigma_\Lambda(\varphi) \;\doteq\; \sup\;\{\lambda(\varphi) :\; \lambda \in \Lambda\}\,, \quad \text{for } \varphi \in \mathbf{C}(X \times Y)$$

Suppose that $\mu \in \mathbf{P}(X)$ and $\nu \in \mathbf{P}(Y)$ satisfy:

$$\mu(f) + \nu(g) \;\leq\; \sigma_\Lambda(f \oplus g), \quad \text{for all} \;\; f \in \mathbf{C}(X) \text{ and } g \in \mathbf{C}(Y)$$

Then there is $\lambda \in \Lambda$ with marginals μ and ν.
(Note that a separation argument implies that $\lambda \in \Lambda$ iff $\lambda(\varphi) \leq \sigma_\Lambda(\varphi)$, for all φ.)

[3] I thank Rody Manuelli for pointing out this reference to me.

2. Further properties of preordered topological spaces.

Let f be a bounded real-valued function on a preordered space X. We define the upper and lower increasing/decreasing envelopes of f as:

Upper increasing envelope: $\overline{f}^i(x) = \sup \{f(y) : y \leq x\}$

Lower increasing envelope: $\underline{f}^i(x) = \inf \{f(y) : x \leq y\}$

Upper decreasing envelope: $\overline{f}^d(x) = \sup \{f(y) : x \leq y\}$

Lower decreasing envelope: $\underline{f}^d(x) = \inf \{f(y) : y \leq x\}$

Let G be the graph of the preorder, and define G^{-1} by $(x, y) \in G^{-1}$ iff $(y, x) \in G$. View G and G^{-1} as correspondences from X to X, that with no danger of confusion we can denote by the same letters. Notice that there is no ambiguity in the notation we defined previously: the inverse image of a set by G coincides with its direct image by G^{-1}.

Let X be a topological space. It follows immediately from the definitions that:

(i) G is uhc iff $d(F)$ is closed for all F closed.

(ii) G is lhc iff $d(U)$ is open for all U open.

(iii) G^{-1} is uhc iff $i(F)$ is closed for all F closed.

(iv) G^{-1} is lhc iff $i(U)$ is open for all U open.

In particular, we have shown before that if X is compact and the preorder is closed, then both G and G^{-1} are uhc. In noncompact cases, uhc is usually harder to be encountered than lhc is, as we will see shortly; for example, the order in $\mathbf{R}^n$ is not uhc, but it is lhc. We now offer an alternative characterization of lower hemicontinuity:

Proposition 2.1

(1) G is lhc iff [A decreasing $\Rightarrow$ $\mathrm{int}(A)$ decreasing] iff [A increasing $\Rightarrow$ $\overline{A}$ increasing].

(2) G^{-1} is lhc iff [A increasing $\Rightarrow$ $\mathrm{int}(A)$ increasing] iff [A decreasing $\Rightarrow$ $\overline{A}$ decreasing].

Proof: We prove the result for G only. Clearly the two right statements are equivalent (by complementation).

Necessity. Let A be increasing, then $\mathrm{int}(A) \subset i(\mathrm{int}(A)) \subset A$, and since the set in the middle is open by lhc, it must equal $\mathrm{int}(A)$, i.e. $\mathrm{int}(A)$ is increasing.

Sufficiency. Let A be open, $y \in A$ and $x \leq y$. Let V be an open nbh of y contained in A. By hypothesis $\mathrm{int}(d(V))$ is decreasing, and since V is open and contained in $d(V)$, it is contained in its interior. Therefore, x belongs to $\mathrm{int}(d(V))$, which constitutes an open nbh of x contained in $d(A)$. Hence, the latter set is open. ∎

The relationship between the continuity properties of G and G^{-1} and the upper and lower envelopes of continuous functions are given in the following proposition (which is actually a version of Berge's Theorem of the Maximum).

Proposition 2.2 *Let $f \in \mathbf{C}(X)$. Then we have:*

(1) If G is uhc, then $\overline{f}^d$ is usc and $\underline{f}^i$ is lsc.

(2) If G is lhc, then $\overline{f}^d$ is lsc and $\underline{f}^i$ is usc.

(3) If G^{-1} is uhc, then $\overline{f}^i$ is usc and $\underline{f}^d$ is lsc.

(4) If G^{-1} is lhc, then $\overline{f}^i$ is lsc and $\underline{f}^d$ is usc.

Proof: Again, we only prove the results for G.
Suppose it is uhc.
$\overline{f}^d(x) \geq a \iff \forall n,\ \exists y \in G(x),\ f(y) \geq a - \frac{1}{n}$.
Let $F_n = \{y : f(y) \geq a - \frac{1}{n}\}$. Then $\{x : \overline{f}^d(x) \geq a\} = \bigcap_n d(F_n)$, which is a closed set. This shows that $\overline{f}^d$ is usc. Exchange the $\geq$ by $\leq$ and a similar proof shows that $\underline{f}^i$ is lsc.

Suppose now G is lhc.
$\overline{f}^d(x) > a \iff \exists y \in G(x),\ f(y) > a$.
Let $V = \{y : f(y) > a\}$. Then $\{x : \overline{f}^d(x) > a\} = d(V)$, which is an open set. This shows that $\overline{f}^d$ is lsc. Exchange the $>$ by $<$ and a similar proof shows that $\underline{f}^i$ is usc. ∎

We get therefore continuity of the respective upper and lower envelopes of a continuous function when G (G^{-1}) is continuous (both uhc and lhc). The result that we will mainly use in the subsequent is the upper semicontinuity of $\overline{f}^d$ and $\overline{f}^i$ whenever X is compact and has a closed preorder. In general, if X is a Suslin metric space (in particular, if it is homeomorphic to a Borel subset of a compact metric space), we can show that the upper and lower envelopes of any measurable function are universally measurable, provided that G is measurable (or, more generally, analytic) in the product space.

To see that the order in $\mathbf{R}^n$ is not uhc, consider the closed set $F = \{(x,y) \in \mathbf{R}^2 : x < 0,\ y \geq -1/x\}$, for which $i(F) = \{(x,y) :\ y > 0\}$, which is not closed.

However, the order in $\mathbf{R}$ is continuous. Actually, one can show that if X has a *total* preorder and the topology has a basis of $\leq$-connected nbh's, then the preorder is closed and continuous, and it is an order iff X is separated (Hausdorff).

We are going to show now that normality and lower hemicontinuity are satisfied in a class of spaces most used in applications. In order to avoid more definitions and introduction of new notation, we are going to work with topological vector spaces, but it should be remarked that only the group structure (without commutativity) is used; therefore, everything we state next is valid for topological groups as well.

Let X be a preordered topological vector space, and suppose that the preorder satisfies the following compatibility condition:

$$(*) \qquad x \leq y, \ z \in X \ \Rightarrow \ x + z \leq y + z$$

If we define $P = \{x : \ 0 \leq x\}$, then $x \leq y$ iff there is $p \in P$ such that $y = x + p$. The preorder is closed iff P is closed. If A is open, then $i(A) = A + P$ and $d(A) = A - P$ are open, regardless of whether or not P is closed. That is,

Proposition 2.3 *Let X be a preordered topological vector space such that $(*)$ is satisfied. Then both G and G^{-1} are lhc.*

By taking into account the possibility of defining the "discrete" order ($x \leq y$ iff $x = y$), we can see that, except for trivial cases, only topologically normal spaces can be expected to have normal preorders. We consider now metrizable topological vector spaces with preorders that satisfy $(*)$. A metrizable topological vector space always has a translation-invariant metric compatible with the topology (this is also true for a topological group).

Proposition 2.4 *Let X be a metrizable topological vector space, with a preorder satisfying $(*)$. Then the preorder is normal.*

Proof: Let d be a translation-invariant metric. Let A be closed and decreasing, B closed and increasing, $A \cap B = \emptyset$. Let:

$$U = \{x : \ d(x, A) < 2d(x, B)\}; \quad V = \{x : \ d(x, B) < 2d(x, A)\}$$

$A \subset U$, $B \subset V$, U and V are open and disjoint. Suppose $x \in U$ and $y \leq x$. Let $p = x - y \in P$. Let $z \in A$. Then $d(x, z) = d(y + p, z) = d(y, z - p) \geq d(y, A)$, since $z - p \in A$. Hence, $d(x, A) \geq d(y, A)$. Let now $z \in B$. Then $d(y, z) = d(x - p, z) = d(x, z + p) \geq d(x, B)$, since $z + p \in B$. Hence, $d(y, B) \geq d(x, B)$. Thus:

$$d(y, A) \leq d(x, A) < 2d(x, B) \leq 2d(y, B)$$

So $y \in U$. A similar argument shows that V is increasing. ∎

Observe that the previous proposition implies in particular that a normal preorder need not be closed.

We next display another important class of spaces that have the property that every closed preorder defined on them is separating. This property is weaker than normality (plus closedness) but it will be sufficient to characterize the stochastic dominance ordering by means of monotone continuous functions.

If X has a closed normal preorder, any continuous and increasing function defined on a compact subset and taking values in $[0, 1]$, can be extended to a similar function defined on the entire space. For a proof, we refer to Nachbin (1965), Thm.2, p.36, and Thm.6, p.69.

Proposition 2.5 *Let X be a locally compact and σ-compact space, and let G be a closed preorder on X. Then G is separating.*

Proof: X can be represented as a countable union of open sets (S_m), such that the closure of each of them is compact and $\overline{S}_m \subset S_{m+1}$.

Suppose that $x \not\leq y$. Let $\overline{m}$ be s.t. both x and y are in $S_{\overline{m}}$.

Since $\overline{S}_{\overline{m}}$ has a separating preorder, we can define on this space $f_{\overline{m}}$ continuous, increasing and taking values in $[0, 1]$ so that $f_{\overline{m}}(x) = 1$ and $f_{\overline{m}}(y) = 0$. We define inductively increasing and continuous functions $f_n : \overline{S}_n \to [0, 1]$, for $n > \overline{m}$, as follows. Suppose f_n has been defined; then let f_{n+1} be a continuous and increasing extension of f_n to $\overline{S}_{n+1}$.

Finally, given $z \in X$, let $n \geq \overline{m}$ be s.t. $z \in \overline{S}_n$, then we can unambiguously define a function $f(z) = f_n(z)$.

If U is open in $[0, 1]$, then $f^{-1}(U) = \bigcup_{m \geq \overline{m}} f_m^{-1}(U) = \bigcup_{m \geq \overline{m}} \left(S_m \cap f_m^{-1}(U) \right)$ is an open set. This shows that f is continuous.

If $z \leq z'$, let m be s.t. S_m contains both. Then $f_m(z) \leq f_m(z')$, which implies $f(z) \leq f(z')$. ■

In particular, any locally compact space with a countable basis is σ-compact, and therefore satisfies the above hypotheses. The foremost examples in this class are the euclidean spaces $\mathbf{R}^n$.

Let X be a topological space and H an arbitrary subset of $X \times X$. We define $\mathbf{Q}(H)$ as the set of all continuous functions $f : X \to [0, 1]$ such that $f(x) \leq f(y)$ whenever $(x, y) \in H$. We define then a preorder on X, denoted $\Sigma(H)$, by setting: $(x, y) \in \Sigma(H) \iff \forall f \in \mathbf{Q}(H), f(x) \leq f(y)$.

Lemma 2.6 *$\Sigma(H)$ is the smallest separating preorder containing H. Moreover, $\mathbf{F}_\Sigma(X) = \mathbf{Q}(H)$*[4].

[4]If we relaxed the requirement of continuity we would get the smallest preorder containing H.

Proof: The fact that $\Sigma(H)$ is a preorder is obvious. Since it contains H, we have that $\mathbf{F}_{\Sigma}(X) \subset \mathbf{Q}(H)$. But the opposite inequality and the fact that $\Sigma(H)$ is separating are immediate consequences of its definition.

Let J be any separating preorder containing H. Then any $f : X \to [0,1]$ continuous and J-increasing is in $\mathbf{Q}(H)$, and hence in $\mathbf{F}_{\Sigma}(X)$. This implies that $\Sigma(H) \subset J$. ∎

Our next result is similar in spirit to Nachbin's compactification result for *ordered* spaces (Theorem 1.4 in this paper). It basically says that, if we restrict ourselves to *preorders* (and topologically cr spaces), then the property of being separating completely characterizes the subspaces of compact spaces endowed with closed preorders. Let us also remark that separation (unlike normality) is a property inherited by subspaces and products.

Theorem 2.7 *A cr space has a separating preorder iff it is isomorphic to a subspace of a compact space endowed with a closed preorder. Moreover, if X is a cr space and G a separating preorder on it, there is a compact space Y endowed with a closed preorder such that X is isomorphic to a dense subspace of Y and, further, each continuous, bounded and increasing real-valued function in X can be extended to a function in the entire space possessing the same characteristics.*

Proof: Let Y be the Stone-Čech compactification of X. To avoid cumbersome notation, view X as a subset of it.

Define in Y the smallest separating preorder $\Sigma(G)$ as above. We claim that $\Sigma(G)$ restricted to X coincides with G. Suppose that $(x, x') \in (X \times X) \setminus G$. Since G is separating, there is some $f : X \to [0,1]$, increasing and continuous, such that $f(x) > f(x')$. By the Stone-Čech Theorem, we can extend f to a continuous function $F : Y \to [0,1]$. It is clear that $F \in \mathbf{Q}(G)$, and therefore $(x, x') \notin \Sigma(G)$.

Taking into account Lemma 2.6, the last paragraph shows actually that the functions in $\mathbf{F}(X)$ are precisely the restrictions to X of the functions in $\mathbf{F}_{\Sigma}(Y)$. This implies immediately the same result for any function in $\mathbf{I}(X)$. ∎

Theorem 2.7 will be used in the sequel to characterize stochastic dominance by means of monotone continuous functions.

Let us finally stress an important point. Suppose X has a separating *order* G. Nothing in the construction of $\Sigma(G)$ guarantees that it is an *order* in Y. Actually, from Nachbin's theorem we know that if G is separating but not completely regular (in the order sense), then $\Sigma(G)$ *cannot* be an order in Y.

Our next result will not be explicitly used in the sequel, but it seems interesting enough to deserve its inclusion here. Suppose that $f : X \to \mathbf{R}$ is bounded, usc and increasing. Call U_f the set of all continuous and increasing functions g such that $f \leq g$. Note that U_f is downward filtering, since increasing and continuous functions are closed under lattice ($\vee$, $\wedge$) operations. We then have:

Theorem 2.8 *Let X be a cr space with a closed normal preorder. If $f : X \to \mathbf{R}$ is bounded, usc and increasing, then it is the lower envelope of the family U_f of increasing and continuous functions greater than f.*

Proof: We are claiming that, for each $x \in X$,

$$f(x) = \inf \{g(x) : g \text{ increasing and continuous, } g \geq f\}$$

Suppose that $f(x) < a$. We must find g increasing and continuous s.t. $g(x) \leq a$ and $f \leq g$. If $a \geq \|f\|$, then we take $g = \|f\|$, so we can assume $a < \|f\|$.

Let $V = \{y : f(y) < a\}$. Then V is an open and decreasing set, that contains the closed and decreasing set $d(x)$. Hence, $d(x)$ and V^c qualify as in the definition of normality, and therefore we may find a function $h \in \mathbf{F}(X)$ s.t. $h|_{d(x)} = 0$ and $h|_{V^c} = 1$. Define $g(y) = a + (\|f\| - a)h(y)$. Then $g(x) = a$, and we have $y \in V \Rightarrow f(y) < a \leq g(y)$, and $y \in V^c \Rightarrow a \leq f(y) \leq \|f\| = g(y)$. Therefore, g satisfies our requirements. ∎

Corollary 2.9 *In a cr space with a closed normal preorder, every bounded, lsc and increasing real-valued function f is the upper envelope of the family L_f of continuous and increasing real-valued functions that are $\leq f$.*

Proof: Consider the parallel of the preceding Theorem for decreasing (instead of increasing) functions. Then, apply that result to $-f$. ∎

We analyze now the relationship between the continuity properties of the correspondences G and G^{-1} and the topological properties of the preorders they define.

Theorem 2.10 *Suppose that either G or G^{-1} is continuous. If X is a (topologically) normal space, then the preorder is normal. In the general (topologically completely regular) case, the preorder is separating if, additionally, its graph is closed.*

Proof: Assume that G is continuous, and X a topologically normal space. Let A and B be disjoint closed sets such that A is decreasing and B increasing. Since X is normal, there is a continuous function $f : X \to [0,1]$ that equals 0 on A and 1 on B. By proposition 2.2, the increasing function $\underline{f}^i$ is continuous. If $x \in A$, the $\underline{f}^i(x) = 0$, since $f(x) = 0$. If $x \in B$, since $i(x) \subset B$, we have that $\underline{f}^i(x) = 1$. This shows that G is a normal preorder. If G^{-1} were continuous, instead of G, then we would have used the function $\overline{f}^i$ (which is continuous), and obtained the same result.

Suppose now that X is a (topologically) completely regular space, the preorder is closed and G is a continuous correspondence. Suppose that $x \not\lesssim y$. Since the preorder is closed, $i(x)$ is a closed set which does not contain y. By complete regularity, we may find a continuous function $f : X \to [0,1]$ that equals 1 on $i(x)$ and for which $f(y) = 0$. Then the continuous and increasing function $\underline{f}^i$ satisfies $\underline{f}^i(x) = 1$ and $\underline{f}^i(y) = 0$. If, instead, G^{-1} is continuous, the proof follows along the same lines suggested in the normality case. ∎

We finish by setting up notation to be used later on. Let X be a cr space endowed with a closed preorder G. Let K be a compact subset of X, and $f : K \to \mathbf{R}$ a bounded and measurable increasing function. We define the **canonical increasing extension** of f, denoted f^e, as follows. Let $\alpha = \sup\{f(x) : x \in K\}$, and $\beta = \inf\{f(x) : x \in K\}$. We then set:

$$f^e(x) = \begin{cases} \alpha, & \text{if } x \in i(K) \setminus K; \\ f(x), & \text{if } x \in K; \\ \beta, & \text{if } x \in i(K)^c. \end{cases}$$

By proposition 1.1 $i(K)$ is closed, it is therefore clear that f^e thus defined is a measurable and increasing extension of f.

3. Characterization by semicontinuous and continuous functions.

Assume in the sequel that X is a (topologically) completely regular space, with a preorder defined on it. Denote by $\mathcal{I}(X)$ the set of all bounded, increasing and $\mathbf{B}(X)$-measurable real-valued functions on X, and by $\mathcal{D}(X)$ its decreasing counterparts. We define the **stochastic dominance** preorder (denoted $\preceq$) on $\mathbf{M}_+(X)$ by setting

$$\mu \preceq \nu \iff \forall f \in \mathcal{I}(X),\ \mu(f) \leq \nu(f)$$

Equivalently, we can define $\preceq$ by means of functions in $\mathcal{D}(X)$, by reversing the inequality on the right.

By considering the constant functions equal to 1 and -1, we can see that two measures are $\preceq$-related only if they have the same mass. We could relax this requirement and define a richer preorder by requiring the functions in the definition to be, additionally, nonnegative. Actually, many of the results we present in the subsequent hold for this alternative ordering. However, there are compelling reasons to preserve the given definition. Three main orderings have been traditionally considered in $\mathbf{M}_+(X)$, and in the regular cases all of them are attached to a convex cone in $\mathbf{C}(X)$ closed under the $\wedge$ operation. The most natural order is the one generated by $\mathbf{C}_+(X)$, and referred to occasionally as the setwise ordering. Another is the one generated by the cone of convex functions in $\mathbf{C}(X)$, X being a compact and convex subset of some locally convex topological vector space (in the economics and finace literature, this ordering is usually termed "second-order stochastic dominace"). Finally, there is the ordering we study in this paper, and, as we shall see later, its similarities with the second one are best preserved if we stick to the definition given above.

We are mainly interested in this section in determining in what cases the preorder we have defined will coincide with the one generated by *continuous* and increasing functions. This will obviously require some topological regularity properties of the preorder in X. It turns out that the property of being *separating* is *precisely* (under mild measurability requirements) what is needed.

We start with some (most of them well-known) approximation results.

Lemma 3.1 *Let $f : X \to \mathbf{R}$ be $\mathbf{B}(X)$-measurable, nonnegative and increasing. Then there is a sequence (f_n) of nonnegative simple functions in $\mathcal{I}(X)$, such that $f_n \uparrow f$.*

Proof: It follows from the standard Lebesgue approximation.
For $1 \leq k \leq n2^n$, let $A_k = \{x :\ f(x) \geq \frac{k}{2^n}\}$. Then set

$$f_n(x) = \frac{1}{2^n} \sum_{k=1}^{n2^n} I_{A_k}(x)$$

f_n is increasing because each A_k is, and we have

$$f(x) \geq n \Rightarrow f_n(x) = n; \quad \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \Rightarrow f_n(x) = \frac{k-1}{2^n}$$

therefore, $f_n \leq f_{n+1}$. ∎

Even if f were not necessarily nonnegative, a similar approximation shows that $\mu \preceq \nu$ iff $\mu(f) \leq \nu(f)$ for each increasing and measurable f for which both integrals are defined.

Corollary 3.2 *Suppose that $\mu(1) = \nu(1)$. Then $\mu \preceq \nu$ iff for each increasing and measurable A, $\mu(A) \leq \nu(A)$.*

Proof: Necessity is true a fortiori, and sufficiency follows by extending the inequality to simple functions in $\mathcal{I}(X)$ and applying the previous lemma. The extension from nonnegative increasing functions to bounded increasing functions is trivial. ∎

This corollary gives rise to a well-known characterization of stochastic dominance in the real line, used frequently as alternative definition. In $\mathbf{R}$ a set is decreasing iff it is a left-unbounded interval. Since open intervals in this class can be approximated by increasing sequences of right-closed intervals, we have that two measures of equal mass satisfy $\mu \preceq \nu$ iff $\forall x,\ F_\nu(x) \leq F_\mu(x)$.

In $\mathbf{R}^n$ there are well known counterexamples of the distribution-function characterization, which stem from the fact that the sets of the form $d(x)$ form too narrow a class to characterize $\preceq$.

We next present the most important approximation result, which is crucially based on the inner regularity of Radon measures. It could be worded by saying that, whenever the preorder is closed, Radon measures are monotonically regular.

Lemma 3.3 *Suppose the preorder in X is closed. Then for any $\mu \in \mathbf{M}_+(X)$ and $A \in \mathbf{B}(X)$ increasing, we have*

$$\begin{aligned} \mu(A) &= \sup\ \{\mu F :\ F \subset A,\ F \text{ closed increasing}\} \\ \mu(A) &= \inf\ \{\mu U :\ A \subset U,\ U \text{ open increasing}\} \end{aligned}$$

Proof:

$$\begin{aligned} \mu(A) &= \sup\ \{\mu K :\ K \subset A,\ K \in \mathbf{K}(X)\} \\ &\leq \sup\ \{\mu(i(K)) :\ K \subset A,\ K \in \mathbf{K}(X)\} \\ &\leq \sup\ \{\mu F :\ F \subset A,\ F \text{ closed increasing}\} \leq \mu(A) \end{aligned}$$

The open approximation is obtained from the parallel result for decreasing sets by complementation. ■

Corollary 3.4 *If the preorder is closed and $\mu(1) = \nu(1)$, then $\mu \preceq \nu$ iff $\mu(F) \leq \nu(F)$ for each F closed and increasing.*

Therefore, whenever the preorder is closed, stochastic dominance is characterized by upper semicontinuous (or lower semicontinuous) and increasing functions.

We now pass to characterization by continuous functions. Our next proposition could be derived from theorem 2.8 and general properties of Radon measures (Dellacherie–Meyer (1975), Thm. 49, p. 108), but we can provide a very simple proof.

Recall that $\mathbf{I}(X)$ is defined as the set of all increasing functions in $\mathbf{C}(X)$.

Proposition 3.5 *Suppose the preorder in X is closed and normal. Then $\mu \preceq \nu \iff \forall f \in \mathbf{I}(X),\ \mu(f) \leq \nu(f)$.*

Proof: Let F be closed and increasing. Let $\varepsilon > 0$ be given. By lemma 3.3, there is some U open and increasing, $F \subset U$, satisfying $\nu(U) \leq \varepsilon + \nu(F)$.

By normality, there is $f \in \mathbf{F}(X)$ s.t. $f|_F = 1$ and $f|_{U^c} = 0$. Then

$$\mu(F) \leq \int f d\mu \leq \int f d\nu = \int_U f d\nu \leq \nu(U) \leq \varepsilon + \nu(F)$$

Since ε was arbitrary, $\mu(F) \leq \nu(F)$, so Corollary 3.4 applies. ■

In particular (Proposition 1.2), in compact spaces with closed preorders stochastic dominance is characterized by monotone continuous functions.

The following Lemma will be used in the general proof of characterization by monotone continuous functions, and also in the derivation of other results in section 6.

Lemma 3.6 *Let Y be a cr space with a closed preorder, and let X be an arbitrary subspace of it. Then $(\mathbf{M}_+(X), \preceq_X)$ is isomorphic (according to the definition given in section 1) to the subspace of $(\mathbf{M}_+(Y), \preceq_Y)$ of measures supported by X.*

Proof: Let μ and ν be measures in $\mathbf{M}_+(X)$. Call μ^e and ν^e the measures in $\mathbf{M}_+(Y)$, supported by X, that correspond to μ and ν.

Suppose first that $\mu \preceq_X \nu$. Then, if A is measurable and increasing in Y, so is $A \cap X$ in X, and therefore $\mu^e(A) = \mu(A \cap X) \leq \nu(A \cap X) = \nu^e(A)$. This implies that $\mu^e \preceq_Y \nu^e$.

Suppose now that $\mu^e \preceq_Y \nu^e$. Let F be closed and increasing in X, and let $\varepsilon > 0$. There is a compact $K \subset F$ satisfying: $\mu(F) \leq \varepsilon + \mu(K)$. Since the preorder in X is the restriction of the preorder in Y, we have that $i_X(K) = i_Y(K) \cap X$, and therefore $\nu^e(i_Y(K)) = \nu(i_X(K)) \leq \nu(F)$. Hence

$$\mu(F) \leq \varepsilon + \mu(K) = \varepsilon + \mu^e(K) \leq \varepsilon + \mu^e(i_Y(K)) \leq \varepsilon + \nu^e(i_Y(K)) \leq \varepsilon + \nu(F)$$

Since ε was arbitrary, it follows that $\mu(F) \leq \nu(F)$, and one can apply corollary 3.4. ■

Theorem 3.7 *If the preorder in X is separating, then stochastic dominance is characterized functions in* $\mathbf{I}(X)$.

Proof: By theorem 2.7, we can view X as a subspace of a compact space Y endowed with a closed preorder, and such that every function in $\mathbf{I}(X)$ is the restriction to X of a function in $\mathbf{I}(Y)$.

We know by propositions 1.2 and 3.5, that functions in $\mathbf{I}(Y)$ characterize $\preceq_Y$. Therefore, the conclusion of the Theorem is a consequence of the previous lemma and the identification of the functions in $\mathbf{I}(X)$ with restrictions to X of functions in $\mathbf{I}(Y)$. ■

In the next proposition, we show that the condition of being separating is almost (except perhaps for uninteresting cases) necessary in order to characterize stochastic dominance by continuous functions.

Observe that, given any preorder in X, we have that $x \leq y \iff f(x) \leq f(y)$, for *all* increasing functions $f : X \to [0,1]$. The following condition requires the preorder to satisfy some regularity from the measurable viewpoint.

We say that the preorder satisfies **condition M** (for measurability) if $x \leq y \iff f(x) \leq f(y)$, for all measurable increasing functions $f : X \to [0,1]$.

For instance, if X is a separable metric space and the graph of the preorder is measurable in $X \times X$, then condition M holds. More generally, condition M will be satisfied provided that either the upper sections ($i(x)$) or the lower sections ($d(x)$) of the preorder are measurable (not necessarily both).

Proposition 3.8 *Suppose that the preorder satisfies condition M. Then stochastic dominance is characterized by monotone continuous functions if and only if the preorder is separating.*

Proof: Only necessity remains to be shown, but its proof is a tautology. Suppose that the preorder is not separating, so that there is $x \not\leq y$ such that for all functions in $\mathbf{I}(X)$, $f(x) \leq f(y)$. Therefore, $\delta_x(f) \leq \delta_y(f)$ for all f in $\mathbf{I}(X)$. By condition M, there is g measurable and increasing such that $g(x) > g(y)$, which implies $\delta_x(g) > \delta_y(g)$. ■

4. Alternative characterizations.

In this section we prove a characterization of stochastic dominance (in the case of closed preorders), by means of measures in $X \times X$ that are concentrated in G. The origins of this characterization are traceable back to Nachbin. See also Hoffmann–Jørgensen (1977), Kamae–Krengel–O'Brien (1977) (that provide a brief historical note with further references) and Strassen (1965). In the case of spaces where the disintegration of measures (existence of regular conditional probabilities) is possible, this implies an alternative characterization, similar to the Blackwell–Sherman–Fell–Cartier Theorem in the case of "second-order stochastic dominance" (Meyer (1966), Theorem 36, p. 288). This result is finally used in conjunction with a representation theorem due to Aumann and Gihman–Skorohod, in order to provide a very simple characterization of stochastic dominance.

To motivate the product space representation, suppose that G is measurable (this is not even necessary, but the generality we gain otherwise is irrelevant) and λ is a measure on $X \times X$ concentrated in G. Let μ and ν be the respective marginals. If A is increasing and measurable, then $(A \times X) \cap G = (A \times A) \cap G \subset (X \times A) \cap G$. Therefore:

$$\mu(A) = \lambda(A \times X) = \lambda((A \times X) \cap G) \leq \lambda((X \times A) \cap G) = \lambda(X \times A) = \nu(A)$$

That is, $\mu \preceq \nu$.

We are going to show next that, if G is closed, then the opposite is also true, so that $\mu \preceq \nu$ iff there is λ concentrated in G with marginals μ and ν. The two first lemmas are merely technical, and are used to avoid some problems related to measurability.

Suppose from now on that G is closed. Let Λ be the set of probability measures concentrated in G. Then Λ is a closed and convex set in $\mathbf{P}(X)$. Recall from section 1 the definition of the upper support function of Λ,

$$\sigma_\Lambda(\varphi) \;:=\; \sup\ \{\lambda(\varphi):\ \lambda \in \Lambda\}, \quad \text{for } \varphi \in \mathbf{C}(X \times X)$$

Recall from section 2 the definition of the upper increasing envelope of a bounded function f:

$$\overline{f}^i(y) = \sup\{f(x):\ (x, y) \in G\}$$

Lemma 4.1 *For any f and g in $\mathbf{C}(X)$,*

$$\sigma_\Lambda(f \oplus g) = \sup\{f(x) + g(y):\ (x, y) \in G\} = \sup\{\overline{f}^i(y) + g(y):\ y \in X\}$$

Proof: Call α, β and γ, respectively, the three numbers above.
$\forall(x,y) \in G$, $\delta_{(x,y)} \in \Lambda$ implies $\alpha \geq \beta$. And $\forall\lambda \in \Lambda$, $\lambda(f \oplus g) \leq \beta$, implies $\alpha \leq \beta$. Thus, $\alpha = \beta$.

Now, $\forall(x,y) \in G$, $f(x) + g(y) \leq \gamma$, implies $\beta \leq \gamma$. Let $\varepsilon > 0$. Let y be so that $\bar{f}^i(y) + g(y) + \frac{\varepsilon}{2} \geq \gamma$, and given y, let $x \leq y$ be so that $f(x) + \frac{\varepsilon}{2} \geq \bar{f}^i(y)$. We have then

$$\beta + \varepsilon \geq f(x) + g(y) + \varepsilon \geq \bar{f}^i(y) + g(y) + \frac{\varepsilon}{2} \geq \gamma$$

Which shows that $\beta \geq \gamma$. ∎

If K is a compact subset of X, we denote f_K the restriction of f to K. $\bar{f}^i_K$ denotes the upper increasing envelope of f_K *with respect to* $G|_K$ (observe that this function will not coincide, in general, with the restriction to K of $\bar{f}^i$). The restriction of a measure μ to K is denoted μ_K.

Lemma 4.2 *Let μ and ν be in $\mathbf{P}(X)$, and suppose that $\mu \preceq \nu$. Let $f \in \mathbf{C}_+(X)$. Then*

$$\sup\left\{\mu_K\left(\bar{f}^i_K\right) : \ K \in \mathbf{K}(X)\right\} \leq \sup\left\{\nu_K\left(\bar{f}^i_K\right) : \ K \in \mathbf{K}(X)\right\}$$

Proof: Let α be the number on the left and β the one on the right. Let $\varepsilon > 0$ be given. First, note that for all K, $0 \leq \bar{f}^i_K \leq \|f\|$. Call g^K the canonical increasing extension of $\bar{f}^i_K$, as defined at the end of section 2. Then $0 \leq g^K \leq \|f\|$ as well. Suppose that $f \neq 0$, otherwise the result is trivial.

Since f is nonnegative, we may choose K so that the following two conditions are satisfied:

$$\nu(K^c) \leq \frac{\varepsilon}{2\|f\|}, \qquad \text{and} \qquad \alpha \leq \frac{\varepsilon}{2} + \mu_K\left(\bar{f}^i_K\right)$$

We have

$$\begin{aligned} \int_{K^c} g^K \, d\mu &= \mu(g^K) - \mu_K\left(\bar{f}^i_K\right) \geq 0 \\ \int_{K^c} g^K \, d\nu &= \nu(g^K) - \nu_K\left(\bar{f}^i_K\right) \geq 0 \end{aligned}$$

Hence,

$$\mu_K\left(\bar{f}^i_K\right) \leq \mu(g^K)$$

$$\nu(g^K) \leq \nu_K\left(\bar{f}^i_K\right) + \|f\|\nu\left(K^c\right) \leq \frac{\varepsilon}{2} + \nu_K\left(\bar{f}^i_K\right)$$

Therefore,

$$\alpha \leq \frac{\varepsilon}{2} + \mu_K\left(\bar{f}^i_K\right) \leq \frac{\varepsilon}{2} + \mu\left(g^K\right) \leq \frac{\varepsilon}{2} + \nu\left(g^K\right) \leq \varepsilon + \nu_K\left(\bar{f}^i_K\right) \leq \varepsilon + \beta$$

Since ε was arbitrary, we conclude $\alpha \leq \beta$. ■

Theorem 4.3 *Let G be a closed preorder on X, and let $\mu(1) = \nu(1)$. Then $\mu \preceq \nu$ iff there is λ concentrated in G with marginals μ and ν.*

Proof: If $\mu = \nu = 0$, the result is trivial. Otherwise, $\mu \preceq \nu$ iff the same is true for the respective normalized measures. Therefore we can assume that μ and ν are probability measures. Let f and g be in $\mathbf{C}_+(X)$. Then we have:

$$\begin{aligned}\mu(f) + \nu(g) &= \nu(g) + \sup\left\{\mu_K(f_K) : K \in \mathbf{K}(X)\right\} \\ &\leq \nu(g) + \sup\left\{\mu_K\left(\bar{f}^i_K\right) : K \in \mathbf{K}(X)\right\} \\ &\leq \nu(g) + \sup\left\{\nu_K\left(\bar{f}^i_K\right) : K \in \mathbf{K}(X)\right\}\end{aligned}$$

Call α the last term, and let $\beta = \sup\left\{\nu_K\left(g_K + \bar{f}^i_K\right) : K \in \mathbf{K}(X)\right\}$. It is clear that $\alpha \geq \beta$. We will prove the opposite inequality. Let $\varepsilon > 0$, and let $\hat{K}$ be such that

$$\sup\left\{\nu_K\left(\bar{f}^i_K\right) : K \in \mathbf{K}(X)\right\} \leq \frac{\varepsilon}{2} + \nu_{\hat{K}}\left(\bar{f}^i_{\hat{K}}\right) \qquad \text{and} \qquad \nu(g) \leq \frac{\varepsilon}{2} + \nu_{\hat{K}}(g_{\hat{K}})$$

Then $\alpha \leq \varepsilon + \nu_{\hat{K}}\left(g_{\hat{K}} + \bar{f}^i_{\hat{K}}\right) \leq \varepsilon + \beta$. Thus $\alpha = \beta$, so

$$\begin{aligned}\mu(f) + \nu(g) &\leq \sup\left\{\nu_K\left(g_K + \bar{f}^i_K\right) : K \in \mathbf{K}(X)\right\} \\ &\leq \sup\left\{\bar{f}^i(y) + g(y) : y \in X\right\} \qquad = \sigma_\Lambda(f \oplus g)\end{aligned}$$

We have proved so far that, if f and g are in $\mathbf{C}_+(X)$, then $\mu(f) + \nu(g) \leq \sigma_\Lambda(f \oplus g)$.

Suppose now that f and g are in $\mathbf{C}(X)$. Then $f + \|f\|$ and $g + \|g\|$ are in $\mathbf{C}_+(X)$, and therefore

$$\mu\left(f + \|f\|\right) + \nu\left(g + \|g\|\right) \leq \sigma_\Lambda\Big(\left(f + \|f\|\right) \oplus \left(g + \|g\|\right)\Big) = \left(\|f\| + \|g\|\right) + \sigma_\Lambda(f \oplus g)$$

Implying that $\mu(f) + \nu(g) \leq \sigma_\Lambda(f \oplus g)$. We can therefore apply Hoffmann–Jørgensen's Theorem (see section 1), and conclude that there is $\lambda \in \Lambda$ with marginals μ and ν.

The converse of this result we have seen previously that holds with full generality. ■

Closedness is essential for the product space representation implied by the previous Theorem (and therefore for its consequences we derive next)[5]. An example illustrating the fact is provided at the end of section 5.

A **stochastic kernel** (or Markov kernel) is a function $\rho : X \times \mathbf{B}(Y) \to [0,1]$ satisfying:
(i) $\forall x \in X$, $\rho(x, \cdot)$ is a probability measure on $\mathbf{B}(Y)$.
(ii) $\forall B \in \mathbf{B}(Y)$, $\rho(\cdot, B)$ is a $\mathbf{B}(X)$ measurable function on X.

Given a stochastic kernel ρ and $\mu \in \mathbf{M}_+(X)$, we can define a measure on Y, that we denote $\rho(\mu)$, by setting

$$\rho(\mu)(B) = \int_X \rho(x)(B)\ \mu(dx), \qquad \forall B \in \mathbf{B}(Y)$$

If Y is a separable metric space, then a stochastic kernel ρ can be always identified with a measurable function from X to $\mathbf{P}(Y)$, because the Borel field in $\mathbf{P}(Y)$ is countably generated. Let us briefly justify this identification. First of all, it is clear that a measurable function from X to $\mathbf{P}(Y)$ induces a stochastic kernel, since the map $\lambda \mapsto \lambda(A)$, for given $A \in \mathbf{B}(Y)$, is Borel measurable in $\mathbf{P}(Y)$. Suppose now that Y is a separable metric space. The Borel sets in $\mathbf{P}(Y)$ are in this case generated by the collection of functions of the form $\hat{f}(\mu) = \mu(f)$, where $f \in \mathbf{C}(Y)$. Let a stochastic kernel ρ as above be given, and define a function $\theta : X \to \mathbf{P}(Y)$ by associating to each x the probability measure $\theta(x) = \rho(x, \cdot)$. Approximation by simple functions combined with property (ii) imply that, for each $f \in \mathbf{C}(Y)$, the function $\hat{f} \circ \theta$ is measurable in X. Hence, a monotone class argument yields that θ is a measurable function.

This equivalent (for the separable metric case) representation of a stochastic kernel will be used in the sequel, since it is more convenient.

The following characterization is due to Meyer (1966) (Theorem 53, p. 303) for the compact metric case, and was generalized to the Polish case by Kamae–Krengel–O'Brien (1977).

Theorem 4.4 *If X is a separable metric space and G a closed preorder on it, then $\mu \preceq \nu$ iff there is a stochastic kernel $\rho : X \to \mathbf{P}(X)$ such that $\nu = \rho(\mu)$ and, for each x, $\rho(x)$ is concentrated in $i(x)$*[6].

[5] From Lemma 4.2 and the proof of Theorem 4.3, we can derive yet another characterization of stochastic dominance for the case of closed preorders. Given $f \in \mathbf{C}_+(X)$ and $\nu \in \mathbf{M}_+(X)$, define the support function

$$p_\nu(f) = \sup\left\{\nu_K\left(\bar{f}^i_K\right) :\ K \in \mathbf{K}(X)\right\}$$

If $\mu(1) = \nu(1)$, then $\mu \preceq \nu$ iff for all $f \in \mathbf{C}_+(X)$, $\mu(f) \leq p_\nu(f)$. In particular, if $\bar{f}^i$ is measurable (for instance, if X is compact), then the same characterization is true if we replace p_ν by the support function $q_\nu(f) = \nu\left(\bar{f}^i\right)$, defined for all (not necessarily nonnegative) functions in $\mathbf{C}(X)$.
The last equivalence could be used in conjunction with a Theorem of Strassen (1969) to obtain an alternative proof of the next result we present, but the proof that follows is more direct.
[6] Equivalently, $\forall x$, $\delta_x \preceq \rho(x)$.

Proof: Sufficiency can be verified directly or, alternatively, apply Theorem 4.3, because the measure induced by μ and ρ on $X \times X$ is concentrated in G.

We will prove necessity using Theorem 4.3. The trivial case where the measures have mass zero can be disregarded, so we assume that both μ and ν are probability measures. Suppose $\mu \preceq \nu$ and let λ be concentrated in G and have the former as marginals. Denote by $\pi : X \times X \to X$ the first projection, that is, $\pi(x, y) = x$, for all $(x, y) \in X \times X$; π is a continuous map and μ is the image of λ under π (this is an alternative definition of what a marginal is). Since we are considering Radon (hence tight) measures, propositions 71–73, pp. 126–129, of Dellacherie–Meyer (1975) imply the existence of a stochastic kernel $\theta : X \to \mathbf{P}(X \times X)$ such that $\lambda = \theta(\mu)$ and, for μ-almost all x, $\theta(x)$ is supported by $\{x\} \times X$. Since λ is concentrated in G, there is a measurable set A, with $\mu(A) = 0$, such that for $x \notin A$, $\theta(x)$ is concentrated in $\{x\} \times G(x)$. By setting $\theta'(x) = \theta(x)$ on A^c and $\theta'(x) = \delta_x$ on A, we obtain a (measurable) stochastic kernel such that $\theta'(x)$ is concentrated in $\{x\} \times G(x)$, for *all* x.

Call ψ the (continuous) function from $\mathbf{P}(X \times X)$ to $\mathbf{P}(X)$ that maps each probability measure on $X \times X$ into its second marginal. Define $\rho = \psi \circ \theta'$, the composition of θ' and ψ. It is a matter of verification to see that ρ thus defined satisfies the properties stated above. ■

It is important to stress that the generality of the previous theorem is a consequence of the fact that we are restricting our consideration to Radon measures. However, if X is homeomorphic to a (universally) measurable subset of a compact metric space (equivalently, a Polish space), then each measure is tight, so there is no restriction.

Let us now enter into the discussion of our last result. Let $\mu \in \mathrm{M}_+(X)$. The easiest and most natural way of generating measures that dominate μ in the stochastic sense, is to consider functions $\varphi : X \to X$ that satisfy the property that, for each x, $x \leq \varphi(x)$; the image of μ under any such function φ will stochastically dominate μ, as it is readily verified. Simple examples show that this method of generating dominating measures is not general enough (suppose $x \leq y$, $y \not\leq x$; let $\mu = \frac{2}{3}\delta_x + \frac{1}{3}\delta_y$ and $\nu = \frac{1}{3}\delta_x + \frac{2}{3}\delta_y$). The characterization we provide next shows that all possibilities of dominating measures are exhausted if one considers stochastic mixtures of collections of functions of the above form.

Let $(S, \mathcal{S}, \mathcal{P})$ be a probability space, and let $\xi : X \times S \to X$ be a jointly measurable function such that, for all (x, s), $\xi(x, s) \geq x$. Given $\mu \in \mathrm{M}_+(X)$, one can verify without much effort that the image measure of $\mu \otimes \mathcal{P}$ under ξ stochastically dominates μ. We are going to show that, in this way, *all* measures that stochastically dominate μ can be obtained. Moreover, the probability space $(S, \mathcal{S}, \mathcal{P})$ can always be taken to be the interval $[0, 1]$ with Lebesgue measure (uniform distribution).

We start by discussing the representation theorem of Aumann and Gihman–Skorohod for stochastic kernels[7]. The only applications of this representation theorem we have been able to find in the economics or game theoretic literature, are Aumann's original contribution and the discussion about it in Milgrom and Weber (1985). We think that the theorem can be fruitfully applied to other contexts as well. For instance, in the literature on Anonymous Games (see Mas Colell, 1984), it implies that one can actually start from a given representation of the game and end up with a well-defined equilibrium for that representation. A version of this representation result has recently been (independently) proved and used by Rustichini (1990), in order to show that mixed strategies can be consistently defined in (nice) strategy spaces with the cardinality of the continuum.

We say that a (separable metric) space is a **Borel space** if it is homeomorphic to a Borel subset of a compact metric space (equivalently, of a Polish space). If X and Y are Borel spaces, a bijective function $f : X \to Y$ is said to be a Borel isomorphism if it is measurable and so is its inverse. All uncountable Borel spaces are Borel isomorphic; particularly, they are all Borel isomorphic to the interval $[0,1]$ (Dellacherie–Meyer, 1975, Theorem 80, p. 249). The Aumann–Gihman–Skorohod representation Theorem states that, if $(S, \mathcal{S})$ is a measurable space, X a Borel space, and $\rho : S \times \mathbf{B}(X) \to [0,1]$ a stochastic kernel, then there is a jointly measurable function $\xi : S \times [0,1] \to X$ such that

$$\rho(s, A) \;=\; m\,\{t :\; \xi(s,t) \in A\}, \qquad \forall A \in \mathbf{B}(X)$$

where m is the Lebesgue measure in $[0,1]$. We shall sketch the proof of this result, because its understanding will be important for the proof of our characterization. First, assume that $X = [0,1]$. For fixed $s \in S$, let $\xi(s,\cdot)$ be defined as the "generalized inverse" of the distribution function of the measure $\rho(s,\cdot)$, that is,

$$\xi(s,t) \;=\; \inf\,\{x :\; F(s,x) \geq t\}$$

where $F(s,x) = \rho(s,[0,x])$. Then the property we want to prove for ξ follows from $\xi(s,t) \leq x \iff t \leq F(s,x)$, and a monotone class argument (Dynkin's π-λ theorem). Using property (ii) in the definition of a stochastic kernel, monotonicity and right–continuity of the distribution functions and rational approximations, one can derive the joint measurability of F, which by a similar argument implies the joint measurability of ξ. It is going to be important later on to observe at this point that, by construction, the image of $\xi(s,\cdot)$ is contained in the support of the measure $\rho(s,\cdot)$.

[7] We have not been able to trace back the origins of this theorem. This sort of representations are familiar in probability, and the basic idea of the one we discuss was also used by Blumenthal–Corson (1972). Skorohod is usually credited with the best known representation theorem (for measures weakly convergent), and Aumann does not mention any references, so we think fair to attach the names of these authors to the theorem. In our discussion, we follow most closely Aumann's version (Aumann, 1964, Lemma F, pp. 644–645). Two slightly different versions can be found in Gihman–Skorohod (1979), Lemmas 1.2 and 2.2.

In general, if X is any uncountable Borel space we use the fact that X and $[0,1]$ are Borel isomorphic, and if X is countable one can adapt without difficulties the proof sketched above.

Let us finally observe that, if f is a bounded and measurable real-valued function on X, then

$$\int_X f(x)\,\rho(s,dx) \;=\; \int_0^1 f\Big(\xi(s,t)\Big)\,m(dt)$$

and if $\mathcal{P}$ is a measure on $(S,\mathcal{S})$, then $\rho(\mathcal{P})$ coincides with the image measure of $\mathcal{P}\otimes m$ under ξ.

Lemma 4.5 *Suppose that X is a Borel space endowed with a closed preorder. Then $\mu \preceq \nu$ if and only if there is a jointly measurable function $\xi : X\times[0,1] \to X$, such that, for all (x,t), $\xi(x,t) \geq x$, and ν is the image of $\mu\otimes m$ under ξ, where m denotes Lebesgue measure.*

Proof: Sufficiency has been proved before. To prove necessity, apply first Theorem 4.4 to find a stochastic kernel $\rho : X \to \mathbf{P}(X)$ such that $\nu = \rho(\mu)$, and, for each x, $\rho(x)$ is concentrated in $i(x)$. Next use the representation theorem, and notice that our previous observation implies that, for all (x,t), $\xi(x,t) \geq x$, because $\xi(x,t)$ belongs to the support of the measure $\rho(x)$, which is contained in $i(x)$. ∎

Theorem 4.6 *The statement of the previous Lemma remains true if we assume that X is any separable metric space and μ and ν are finite and nonnegative Radon measures.*

Proof: Let B be a common σ-compact support for μ and ν. Then we can identify μ and ν with measures on the Borel space B (recall Lemma 3.6), and apply the previous Lemma. The function ξ whose existence is implied by that Lemma can be extended to a (measurable) function ξ' defined on X, by setting $\xi'(x,t) = x$, for $x \in B^c$. ∎

Let us finally remind that one can view (i.e. represent) a measure on X as the image of the Lebesgue measure on $[0,1]$ by a random variable, and in this case the last Theorem (or some of the previous alternative representations) implies immediately that $\mu \preceq \nu$ if and only if they are the respective images of two random variables ζ and χ such that, for all $t \in [0,1]$, $\zeta(t) \leq \chi(t)$.

5. Antisymmetric property.

We are going to show that if G is a closed *order* in X, then $\preceq$ is an order (antisymmetric) in $\mathbf{M}_+(X)$. This has been shown in the compact metric case by Hopenhayn–Prescott (1987) and in the Polish case by Kamae–Krengel (1978). Their proofs rely heavily on the fact that these spaces have countable bases, and are therefore not applicable in the context we consider. The approach followed here highlights better the essence of the problem, and helps in understanding why one cannot dispense with the closedness assumption. We start by motivating the result with a discussion of the compact case, in which antisymmetry is an immediate consequence of the Stone–Weierstrass Theorem.

Suppose that X is a compact space and has a closed order defined on it. The vector space generated by $\mathbf{I}(X)$ in $\mathbf{C}(X)$ can be written $\mathbf{I}(X) \ominus \mathbf{I}(X)$, because it is formed by functions of the form $f - g$, where f and g are in $\mathbf{I}(X)$. Every constant function is increasing, and therefore this vector space contains the constants. Since $\mathbf{I}(X)$ is closed under the lattice operations, so is $\mathbf{I}(X) \ominus \mathbf{I}(X)$. Since G is separating and it is an order, $\mathbf{I}(X)$ separates the points of X. It follows then from the Stone–Weierstrass Theorem that $\mathbf{I}(X) \ominus \mathbf{I}(X)$ is dense in $\mathbf{C}(X)$, and therefore $\mu = \nu$ iff $\mu(f) = \nu(f)$ for each $f \in \mathbf{I}(X)$.

The general case is not so straightforward. The key to our proof is Lemma 5.2, in which the role played by closedness of G is clear.

Lemma 5.1 *Let X be a cr space and G a closed preorder on it. Suppose that $\mu \preceq \nu$ and $\nu \preceq \mu$. Let A and B be disjoint measurable sets such that A is decreasing. Let $\lambda \in \mathbf{M}_+(X \times X)$ be concentrated in G and have μ and ν as marginals. Then $\lambda(A \times B) = 0$.*

Proof: We recall that A is decreasing iff $(X \times A) \bigcap G = (A \times A) \bigcap G$. Since the support of λ is contained in G and, by hypothesis, $\mu(A) = \nu(A)$, we have:

$$\nu(A) = \lambda(X \times A) = \lambda(A \times A) = \lambda(A \times X) = \mu A$$

Hence, $\lambda(A \times B) \leq \lambda(A \times A^c) = 0$. ∎

Recall that we defined $G^{-1} \doteq \{(x, y) : (y, x) \in G\}$.

Lemma 5.2 *Let the hypotheses of Lemma 5.1 hold. Suppose that K is a compact set contained in the complement of G^{-1}. Then $\lambda(K) = 0$.*

Proof: Since $\left(G^{-1}\right)^c$ is open and X is a regular topological space, for each $(x, y) \in K$, there are an open nbh of x, U_x, and an open nbh of y, V_y, such that $\overline{U}_x \times \overline{V}_y \subset \left(G^{-1}\right)^c$.

Then we have that $d(\overline{U}_x) \cap i(\overline{V}_y) = \emptyset$, because if z were in the intersection, then there would be $x' \in \overline{U}_x$ and $y' \in \overline{V}_y$ such that $y' \le z \le x'$, which is impossible, since in this case (x', y') would be in G^{-1}.

Since K is compact, we can find a finite set I such that (with simplified notation),

$$K \subset \bigcup_{i \in I} (U_i \times V_i)$$

Let K_1 and K_2 be the projections of K. For each i, we have $\overline{U}_i \supset \overline{U_i \cap K_1}$ and $\overline{V}_i \supset \overline{V_i \cap K_2}$. Let $A_i = d\left(\overline{U_i \cap K_1}\right)$ and $B_i = i\left(\overline{V_i \cap K_2}\right)$. The sets A_i and B_i are closed by proposition 1.1, and

$$K \subset \bigcup_{i \in I} ((U_i \cap K_1) \times (V_i \cap K_2)) \subset \bigcup_{i \in I} (A_i \times B_i)$$

By construction, A_i and B_i satisfy the assumptions of Lemma 5.1, for each i. Thus, $\lambda(A_i \times B_i) = 0$, $\forall i \in I$. Hence $\lambda(K) \le \sum_{i \in I} \lambda(A_i \times B_i) = 0$. ∎

Theorem 5.3 *Let the hypotheses of Lemma 5.1 hold, and suppose that, additionally, G is an order (antisymmetric). Then $\mu = \nu$.*

Proof: Since λ is a Radon measure, lemma 5.2 implies that $\lambda((G^{-1})^c) = 0$. Since $G \cap G^{-1} = \Delta$ (the diagonal) and the support of λ is contained in G, this means that the support of λ is contained in Δ.

Let $A \in \mathbf{B}(X)$, then $\mu(A) = \lambda(A \times X) = \lambda(A \times A) = \lambda(X \times A) = \nu(A)$. ∎

We remark that Lemma 5.2 shows that, in the *preorder* case, if μ and ν dominate each other in the stochastic sense, then λ is concentrated in $G \cap G^{-1}$.

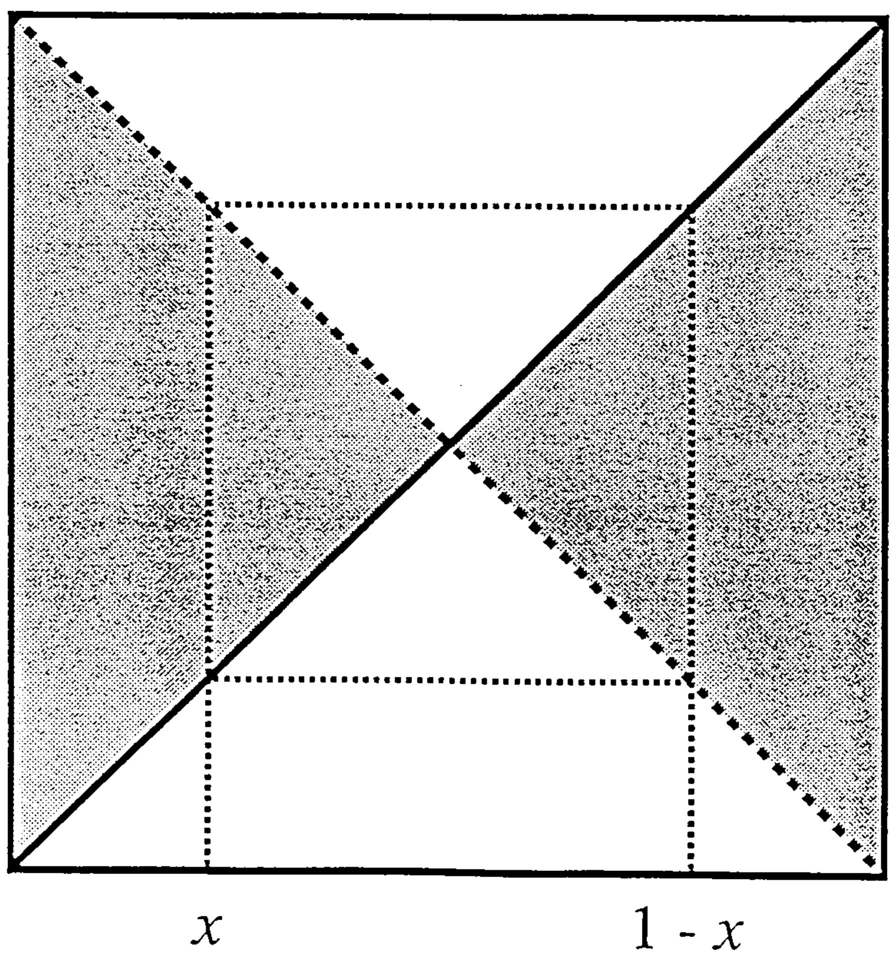
There is a special case in which antisymmetry is preserved, even though the graph of the order G is not closed. This happens whenever the smallest closed preorder containing G is actually an order (it is easy to see that one can always define a smallest closed preorder containing any set in $X \times X$, but not a smallest closed order). This observation can be used to extend our results here to the case of strict (irreflexive) orders.

We present now an example that shows that closedness of G is essential both for the product space representation (Theorem 4.3) and for the preservation of antisymmetry (with the proviso implied by the previous observation).

Let $X = [0,1]$, let $\le$ denote the usual order, and define a new order G by

$$G(x) = \begin{cases} [x, 1-x), & \text{if } x < \frac{1}{2}; \\ \{x\}, & \text{if } x = \frac{1}{2}; \\ (1-x, x], & \text{if } x > \frac{1}{2}. \end{cases}$$

It is readily seen that G thus defined is an order, and that it is not closed, because the pairs of different points that add up to one are in its closure but do not belong to it. Let now μ be the uniform distribution on $\left[0, \frac{1}{2}\right]$, and ν the uniform distribution on $\left[\frac{1}{2}, 1\right]$. Increasing sets with respect to G are precisely intervals symmetric around $\frac{1}{2}$, except perhaps for the endpoints. Therefore, μ and ν dominate each other in the stochastic dominance preorder induced by G. One can show (as the picture suggests) that no measure supported by G can have μ and ν as marginals, although the two measures can be obtained as marginals of the uniform distribution on the set $\left\{(x, y) : \ x \leq \frac{1}{2}, \text{ and } x + y = 1\right\}$, which is a measure supported by a set disjoint from G.



6. Topological properties.

In this section, we analyze the topological properties that the stochastic dominance ordering inherits from the preorder on X. The first one of them, and perhaps most important, is closedness. In the case in which stochastic dominance is generated by continuous and monotone functions, closedness is an immediate consequence. In the general case, it is derived without difficulty from theorem 4.3.

Next, we present a result that has interesting applications for weak convergence of stochastic processes. In the "order" language, it says that $\mathbf{M}_+(X)$ is a completely regular ordered space whenever so is X. As a consequence, we may characterize weak convergence by means of monotone continuous functions.

We finally formulate a condition that guarantees that stochastic preorder intervals are compact. This result is important in order to assure that increasing sequences (nets) of measures bounded above have limit points.

Theorem 6.1 *If the preorder on X is closed, then stochastic dominance is a closed preorder on $\mathbf{M}_+(X)$.*

Proof: Suppose $\mu_i \preceq \nu_i$, $\mu_i \to \mu$ and $\nu_i \to \nu$. Let λ_i be concentrated in G and have marginals μ_i and ν_i. By Hoffmann–Jørgensen's Theorem (see section 1), we know that (λ_i) has a cluster point λ with marginals μ and ν. This implies that $\mu \preceq \nu$, because λ is concentrated in G. ∎

Another easy consequence of our previous findings is,

Proposition 6.2 *Suppose the preorder in X is separating. Then stochastic dominance is a separating preorder. Moreover, if $\mu \not\preceq \nu$, then μ and ν can be separated by a linear and increasing continuous function.*

Proof: Since $\preceq$ is characterized by continuous and increasing functions in this case, if $\mu \not\preceq \nu$, there exists some f in $\mathbf{F}(X)$ such that $\mu(f) > \nu(f)$. ∎

We have not attempted to test for normality in $\mathbf{M}_+(X)$, since separation of closed and disjoint monotone sets can hardly be expected to be accomplished by any *linear* function, and therefore it seems of little interest for applications.

The following propositions involve the concept of a completely regular ordered space (in Nachbin's terminology) which, we remind, is *stronger* than a cr space endowed with an order. For a formal definition, refer to section 1.

Lemma 6.3 *Let $(X, \leq)$ be a compact ordered space. Then $(\mathbf{M}_+(X), \preceq)$ is a completely regular ordered space.*

Proof: The fact that $\preceq$ is separating in $\mathbf{M}_+(X)$ has been proved in the previous proposition. We also know from section 5 that $\preceq$ is an order. We just have to show that the remaining defining property of a cro space is satisfied.

Let $\mu \in \mathbf{M}_+(X)$ and let V be a nbh of μ. Since $\mathbf{I}(X) \ominus \mathbf{I}(X)$ is dense in $\mathbf{C}(X)$, there is a finite set I, a collection of strictly positive numbers (ε_i) and functions f_i in $\mathbf{I}(X)$, such that the set

$$U = \{\nu \in \mathbf{M}_+(X) : \ \forall i \in I, \ |\mu(f_i) - \nu(f_i)| \leq \varepsilon_i\}$$

is a neighborhood of μ containted in V. Now, we can define, for each $i \in I$, the following functions:

$$p_i(\nu) = \min\left\{ 1 \, , \ \max\left\{0, \left(\frac{1}{\varepsilon_i}\right) \left(\varepsilon_i + \nu(f_i) - \mu(f_i)\right)\right\}\right\}$$

$$q_i(\nu) = \min\left\{ 1 \, , \ \max\left\{0, \left(\frac{1}{\varepsilon_i}\right) \left(\nu(f_i) - \mu(f_i)\right)\right\}\right\}$$

Clearly, p_i and q_i are continuous, increasing and comprised between 0 and 1. For each i, $p_i(\mu) = 1$ and $q_i(\mu) = 0$. Suppose $\nu \notin U$; then for some i, $|\mu(f_i) - \nu(f_i)| > \varepsilon_i$, so either $\nu(f_i) - \mu(f_i) > \varepsilon_i$, of $\mu(f_i) - \nu(f_i) > \varepsilon_i$; in the former case, $q_i(\nu) = 1$, and in the latter $p_i(\nu) = 0$. Define

$$p(\nu) = \min\left\{p_i(\nu) : \ i \in I\right\}$$

$$q(\nu) = \max\left\{q_i(\nu) : \ i \in I\right\}$$

These functions satisfy the requirement in the definition of a completely regular ordered space for μ and V. ∎

Theorem 6.4 *$(X, \leq)$ is a completely regular ordered space iff $(\mathbf{M}_+(X), \preceq)$ is a completely regular ordered space.*

Proof: Sufficiency is a consequence of the isomorphism between X and the subset $\{\delta_x : x \in X\}$ of $\mathbf{M}_+(X)$.

Necessity. By Nachbin's compactification Theorem, X is isomorphic to a subset of a compact ordered space Y. Therefore, by lemma 3.6, $(\mathbf{M}_+(X), \preceq_X)$ is isomorphic to a subspace of $(\mathbf{M}_+(Y), \preceq_Y)$, and the statement above follows because the latter is a cro space by previous theorem. ∎

Corollary 6.5 *If X is a completely regular ordered space, then the weak topology in $\mathbf{M}_+(X)$ is generated by the increasing functions in $\mathbf{C}(X)$, that is, a net $\mu_i \to \mu$ if and only if $\mu_i(f) \to \mu(f)$ for each f in $\mathbf{I}(X)$.*

Proof: In the compact case, the result is a consequence of the Stone–Weierstrass Theorem, and in the general case it follows from Theorem 6.4, Nachbin's compactification Theorem and the homeomorphism between the measures in a cr space supported by a subset and the measures *defined* on that subset (see section 1). ■

Thus, in a completely regular ordered space, continuous and increasing functions are convergence-determining. This corollary can be used in applied work to show weak convergence to a candidate limiting measure, and it may also simplify in certain cases (separable metric spaces) measurability proofs. One way of viewing this result is as a generalization of the familiar distribution-function characterization of weak convergence in $\mathbf{R}^n$.

We now address the following problem: suppose that (μ_i) is a net in $\mathbf{M}_+(X)$ such that for each function f in $\mathbf{I}(X)$, $\mu_i(f)$ converges to some limit, say $l(f)$. Is there a measure μ that is the weak limit of (μ_i)? The answer is yes in the compact ordered case, since then (μ_i) has a cluster point μ, which by corollary 6.5 is actually its weak limit.

In the general case the answer is, however, negative. We just need to consider the sequence $(\delta_n)_{n\in\mathbf{N}}$, in $\mathbf{R}$.

Therefore, we look for a less ambitious objective. We are going to find conditions that guarantee that, if (μ_i) is increasing, and there is some ν such that, for all i, $\mu_i \preceq \nu$, then (μ_i) converges. The sufficient condition we are going to derive is compactness of the stochastic preorder intervals.

We say that a preordered cr space X satisfies **condition K** if, for each K compact in X, $c(K) = i(K) \cap d(K)$ is compact.

Condition K is stronger than requiring preorder intervals in X to be compact. However, if the preorder intervals of X are compact and each compact set is contained in some preorder interval, then it is obvious that condition K holds. On the other hand, Condition K is inherited by arbitrary products. For instance, it holds in $\mathbf{R}$, $\mathbf{R}^n$, $\mathbf{R}^\infty$, and, in general, in $\mathbf{R}^T$ (T any set). We can construct an example of a locally compact space with a countable base, endowed with a closed order for which order intervals are compact, but where condition K is not satisfied.

Theorem 6.6 *Suppose X has a closed preorder that satisfies condition K. Then $\mathbf{M}_+(X)$ has compact preorder intervals.*

Proof: Let $[\mu, \nu]$ be a preorder interval in $\mathbf{M}_+(X)$. Let $\varepsilon > 0$ be given. Let K be compact and such that

$$\mu(K^c) \le \frac{\varepsilon}{2} \qquad \text{and} \qquad \nu(K^c) \le \frac{\varepsilon}{2}$$

We can assume further that K is $\le$-connected, since otherwise we replace it by the compact set $c(K)$, which obviously satisfies also the property above. Notice that, if K is $\le$-connected, $i(K) \setminus K$ is an increasing set.

Let now λ be in $[\mu, \nu]$. Say that the mass of the measures in the interval is t, i.e. $t = \mu(X) = \nu(X) = \lambda(X)$. Then

$$\begin{aligned} \lambda(i(K)) &\ge \mu(i(K)) \ge \mu(K) \ge t - \frac{\varepsilon}{2} \\ \lambda(i(K) \setminus K) &\le \nu(i(K) \setminus K) \le \nu(K^c) \le \frac{\varepsilon}{2} \end{aligned}$$

Hence: $\lambda(K) = \lambda(i(K)) - \lambda(i(K) \setminus K) \ge t - \varepsilon$

Since λ was arbitrary, this shows that $[\mu, \nu]$ is relatively compact, and we already know it is closed. ∎

Let us remark that examples can be constructed showing that condition K is not necessary for stochastic preorder intervals to be compact.

Now we can derive our final result. We say that a net (μ_i) in $\mathbf{M}_+(X)$ is increasing, if $i \le i'$ implies $\mu_i \preceq \mu_{i'}$. The net is said to be bounded above if there is some ν such that, for all i, $\mu_i \preceq \nu$ (in other words, by boundedness we refer to *order* boundedness).

Proposition 6.7 *Let the hypotheses of Theorem 6.6 hold. Suppose that (μ_i) is an increasing and bounded above net in $\mathbf{M}_+(X)$. Then (μ_i) has a cluster point. Moreover, if X is an ordered space, then the net has a (unique) limit.*

Proof: Fix i_0 in I. Then $\mu_i \in [\mu_{i_0}, \nu]$, for all $i \ge i_0$, where ν is an upper bound of the net. By compactness, there is a cluster point μ of the net, which satisfies the property that, for each $f \in \mathbf{I}(X)$, $\mu(f) = \sup \{\mu_i(f) : i \in I\}$. If X is ordered, then μ is unique, by theorem 5.3. ∎

Let us finally observe that one can work out a theory of extremal measures (via Zorn's Lemma), paralleling the development of the analogous theory in the case of second-order stochastic dominance (Meyer, 1966), although one of the basic ingredients of that theory (the representation via distributions on the set of extremal points) would be missing.

Taking into account the last observation, Proposition 6.7 (combined with its counterpart for decreasing nets bounded below) can be used to furnish an alternative proof, as well as a generalization, of the fixed point theorem for monotone Markov processes stated in Hopenhayn–Prescott (1987).

7. References.

Aumann, R. (1964), "Mixed and Behavior Strategies in Infinite Extensive Games," *Advances in Game Theory*, Annals of Mathematical Studies no. 52, Princeton U.P., Princeton (N.J.), pp. 627–650.

Bikhchandani, S., U. Segal and S. Sharma (1990), "Stochastic Dominance under Bayesian Learning," Working Paper 9002, Dept. of Economics, University of Toronto.

Blumenthal, R. and H. Corson (1972), "On Continuous Collections of Measures," *Proc. 6th Berkeley Symp. on Math. Stat. and Prob.*, 2, pp. 33–40.

Dekel, E., Z. Safra and U. Segal (1989), "Existence of Nash Equilibrium with Non-Expected Utility Preferences," mimeo.

Dellacherie, C. and P. A. Meyer (1975), *Probabilités et Potentiel*, Chapters I–IV, Hermann, Paris.

Gihman, I. and A. Skorohod (1979), *Controlled Stochastic Processes*, Springer-Verlag, New York.

Hoffmann–Jørgensen, J. (1977), "Probability in Banach Spaces," École d'été de Probabilités de Saint–Flour VI–1976, *Lecture Notes in Mathematics*, 598, Springer–Verlag, Berlin.

Hopenhayn, H. and E. Prescott (1987), "Invariant Distributions for Monotone Markov Processes," Discussion Paper 242, Dept. Economics, University of Minnesota.

Kamae, T. and U. Krengel (1978), "Stochastic Partial Ordering," *Annals of Probability*, 6, pp. 1044–1049.

Kamae, T., U. Krengel and G. L. O'Brien (1977), "Stochastic Inequalities on Partially Ordered Sets," *Annals of Probability*, 5, pp. 899–912.

Kelley, J. L. (1955), *General Topology*, Van Nostrand, New York.

Mas Colell, A. (1984), "On a Theorem of Schmeidler," *Journal of Mathematical Economics*, 13, pp. 201–206.

Meyer, P.A. (1966), *Probabilités et Potentiel*, Hermann, Paris.

Milgrom, P. and R. Weber (1985), "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, 10, pp. 619–632.

Nachbin, L. (1965), *Topology and Order*, Van Nostrand, New York.

Rustichini, A. (1990), "Mixing on Function Spaces," Northwestern University (forthcoming in the CMSEMS series of discussion papers).

Stokey, N. and R. E. Lucas (with E. Prescott) (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge (Mass.).

Strassen, V. (1965), "The Existence of Probability Measures with given Marginals," *Annals of Mathematical Statistics*, 36, pp. 423–439.