Dependence Orderings for Multivariate Distributions: The Supermodular Ordering and Related Orderings

Margaret Meyer and Bruno Strulovici

Oxford + Northwestern

May 21, 2015

Introduction

In many contexts, valuable to have criteria for determining whether one set of random variables displays a **greater degree of dependence** than another. Here we adopt the "stochastic dominance approach":

- Relate partial orderings expressed directly in terms of joint prob. dists. to orderings expressed indirectly through properties of objective functions whose expectations are used to evaluate the dists.
- Since expectations of additively separable functions depend only on marginal dists., attitudes towards **dependence** must be represented through **non-separability** properties.
- Motivated by economic applications, we propose **"supermodularity"** (Topkis, 1978) to capture a preference for greater dependence.
  - The arguments of supermodular functions are "complementary": cross-partial derivatives are non-negative.
- **Our objectives:** To characterize the **"supermodular stochastic ordering"** and related orderings, identify sufficient conditions in specific environments, and explore applications.

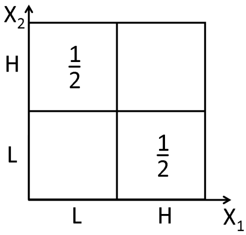
Overview

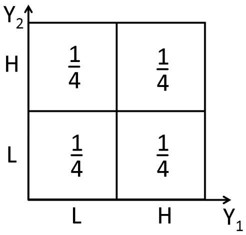
1. Motivations
   - Economic motivations/applications
   - Why aren't familiar tools for measuring "correlation" adequate?
2. Supermodular functions, elementary transformations, and our dual characterization of the supermodular (SPM) stochastic ordering
3. How to apply our characterization: constructive methods
4. Related orderings: increasing SPM ordering; symmetric SPM ordering; ordering based on dispersion of cdfs of order statistics
5. Sufficient conditions in specific environments
6. Application to systemic risk

Motivations/applications (1)

(Ex ante) comparisons of ex post inequality under uncertainty (Meyer and Mookherjee, 1987):

- In group settings where individual rewards are uncertain, groups may be concerned, ex ante, about ex post inequality of rewards. Then evaluating relative appeal of reward schemes $X \equiv (X_1, \ldots, X_n)$ and $Y \equiv (Y_1, \ldots, Y_n)$ requires comparing dependence in X and Y, e.g. by comparing $E[w(X_1, \ldots, X_n)]$ vs. $E[w(Y_1, \ldots, Y_n)]$ for any supermodular ex post welfare function w.

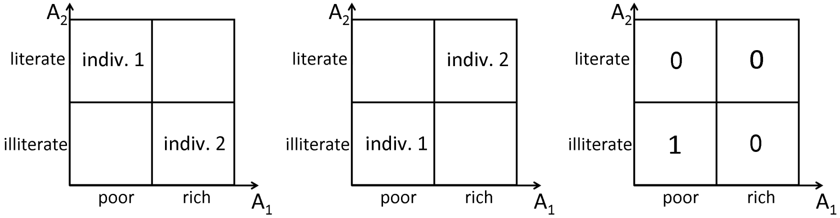




Motivations/applications (2)

Comparisons of multidimensional deprivation (Atkinson and Bourguignon, 1982):

- Individual-level data on n attributes $(A_1^i, \ldots, A_n^i)$ (e.g. income, health, education)
- Evaluate $\sum_i D(A_1^i, \ldots, A_n^i)$, where $D(\cdot)$ is the individual deprivation function. Supermodularity of D reflects a dislike of positive interdependence in population distribution of $(A_1, \ldots, A_n)$.



Motivations/applications (3)

Comparisons of systemic risk in financial systems (Beale et al, 2011):

- Under most definitions, "systemic risk" is greater, the more positively dependent are bank returns or bank failures.
- For ex., Beale et al propose a "systemic cost function"

$$c(Y_1, \ldots, Y_n) = C(\sum_{i=1}^{n} I_{\{Y_i \leq s_i\}}),$$

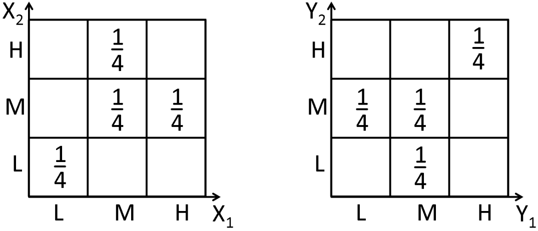
  where Y_i is bank i's return and s_i is bank i's failure threshold. $C(\cdot)$ is convex since "as more banks fail in the same time period, the economic disruption tends to increase disproportionately".

- For all convex C and for all $(s_1, \ldots, s_n)$, this systemic cost function $c(Y_1, \ldots, Y_n)$ is supermodular. Consequently, holding fixed the marginal distributions of the Y_i and "increasing their dependence" increases expected systemic cost.

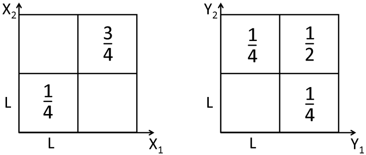
Why aren't familiar tools adequate? (1)

Many orderings (e.g. linear correlation coefficient) are not invariant to coordinate relabeling or coarsening of categories:

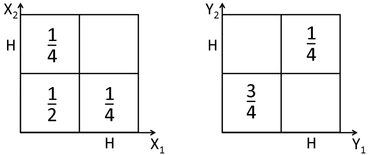
- Here, $corr(X_1, X_2) > (<)\, corr(Y_1, Y_2) \iff M > (<)\, \frac{L+H}{2}$



If $M = H$



If $M = L$



Why aren't familiar tools adequate? (2)

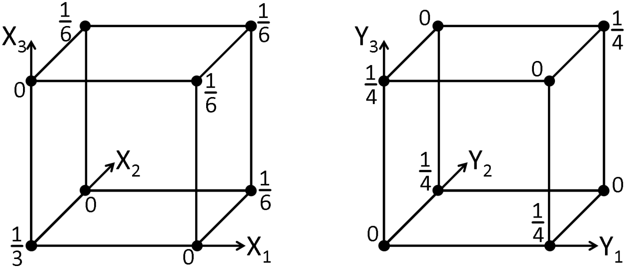
"Dependence" is a more subtle concept in ≥ 3 dimensions than in 2:

- In 2 dimensions, positive and negative dependence are "mirror images", but for $n > 2$ dimensions, this symmetry breaks down.
  - For $n = 2$ and for any plausible concept of positive dependence, (Y_1, Y_2) positively dependent $\Longrightarrow$ $(-Y_1, Y_2)$ negatively dependent.
  - For $n > 2$, no simple way to convert a positively dependent random vector $(Y_1, Y_2, \ldots, Y_n)$ into a negatively dependent one—e.g. $(-Y_1, Y_2, \ldots, Y_n)$ is not negatively dependent.

Why aren't familiar tools adequate? (3)

For $n > 2$ dimensions, pairwise measures are not always adequate:

- $(X_1, X_2, X_3) \in \{0,1\}^3$ and $(Y_1, Y_2, Y_3) \in \{0,1\}^3$ have identical marginal dists.
- For each pair (i,j) and for $k = 0, 1$, $Pr(X_i = X_j = k) > Pr(Y_i = Y_j = k)$, so clearly (X_i, X_j) are more positively dependent than (Y_i, Y_j) for each (i,j).
- But $Pr(X_1 = X_2 = X_3 = 1) < Pr(Y_1 = Y_2 = Y_3 = 1)$, so if objective function w equals 1 at (1,1,1) and 0 elsewhere, then $Ew(X) < Ew(Y)$.



General setting

- Distributions on support $L = \times_{i=1}^{n} L_i$, where L_i is a **finite** subset of $\mathbb{R}$ and $|L_i| = l_i$.
- L is a finite lattice, with usual order $z \leq v \Leftrightarrow z_i \leq v_i \ \forall i$.
- L has $d = \prod_{i=1}^{n} l_i$ elements.
- Wlog, as explained below, take $L_i = \{0, 1, \ldots, l_i - 1\}$.
- Objective functions w and distributions f on L can be viewed as vectors in $\mathbb{R}^d$. The expected value of w given f is the scalar product of w with f:

$$E[w|f] = \sum_{z \in L} w(z) f(z) = w \cdot f.$$

The supermodular ordering ($\succeq SPM$)

A function w is *supermodular (SPM)* on L, written $w \in \mathcal{S}$, if $w(z \wedge v) + w(z \vee v) \geq w(z) + w(v)$ for all $z, v \in L$. Define e_i to be the unit vector in the i^{th} dimension. Topkis (1978) has shown that

$$w \in \mathcal{S} \iff w(z + e_i + e_j) + w(z) \geq w(z + e_i) + w(z + e_j)$$

for all $i \neq j$ and z s.t. $z + e_i + e_j \in L$.

- For continuous v.'s and smooth w, w is SPM iff $\frac{\partial^2 w}{\partial z_i \partial z_j}(z) \geq 0 \ \forall z, \forall i \neq j$.

Definition

Let the random vectors Y and X have distributions g and f, respectively, on L. Y dominates X (g dominates f) acc. to the **supermodular ordering**, written $Y \succeq_{SPM} X$ ($g \succeq_{SPM} f$), if and only if $Ew(Y) \geq Ew(X)$ ($w \cdot g \geq w \cdot f$) for all supermodular functions w.

The supermodular ordering ($\succeq SPM$)

Invariance to monotonic relabelings of the coordinates:

- If w is SPM, then $\tilde{w}(z) \equiv w(r_1(z_1), \ldots, r_n(z_n))$ is SPM whenever $\{r_i\}_{i=1}^n$ are all nondecreasing or all nonincreasing. Hence the SPM ordering is preserved by monotonic relabelings of the coordinates from $(z_1, \ldots, z_n)$ to $(r_1(z_1), \ldots, r_n(z_n))$.
  - So it is wlog to take $L_i = \{0, 1, \ldots, l_i - 1\}$.
  - And the SPM ordering is preserved by coarsening of the support.

(Adjacent) elementary transformations (ET's)

Definition

For any $z \in L$ s.t. $z + e_i + e_j \in L$, an **(adjacent) elementary transformation (ET)** $t^z_{i,j}$ is the function on L s.t.

$$t^z_{i,j}(z) = t^z_{i,j}(z + e_i + e_j) = 1 \text{ and } t^z_{i,j}(z + e_i) = t^z_{i,j}(z + e_j) = -1,$$

and $t^z_{i,j}(v) = 0$ for all other nodes $v \in L$. Let $\mathcal{T}$ denote the class of all such $t^z_{i,j}$. If for some $z \in L$, some i, j, and some scalar $\alpha > 0$, $g - f = \alpha t^z_{i,j}$, then we say that g is obtained from f by an ET of size α.

- ET's increase dependence while leaving marginal distributions unchanged.
- ET's as defined above affect only **two** of the n dimensions and affect values only at four **adjacent** points in L.
- **w is SPM iff for every ET, the expectation of w is (weakly) increased:**

$$w \in \mathcal{S} \iff w \cdot t \geq 0 \qquad \forall t \in \mathcal{T}$$

Dual characterization of the supermodular ordering

Theorem

$g \succeq_{SPM} f$ if and only if there exist **nonnegative** *coefficients $\{\alpha_t\}_{t\in\mathcal{T}}$ such that*

$$g - f = \sum_{t\in\mathcal{T}} \alpha_t t.$$

Proof: Define $\mathcal{T}^C = \{\sum_{t\in\mathcal{T}} \alpha_t t : \alpha_t \geq 0 \quad \forall t \in \mathcal{T}\}$. Since $w \in \mathcal{S} \iff w \cdot t \geq 0 \quad \forall t \in \mathcal{T}$, $\mathcal{S}$ is the dual cone of $\mathcal{T}^C$. Given $\mathcal{T}^C$ is closed and convex, it follows from duality that $\mathcal{T}^C$ is the dual cone of $\mathcal{S}$, i.e.,

$$(g - f) \in \mathcal{T}^C \iff w \cdot (g - f) \geq 0 \quad \forall w \in \mathcal{S}. \qquad \blacksquare$$

Dual characterization of the supermodular ordering

Theorem

$g \succeq_{SPM} f$ if and only if there exist **nonnegative** *coefficients $\{\alpha_t\}_{t\in\mathcal{T}}$ such that*

$$g - f = \sum_{t\in\mathcal{T}} \alpha_t t.$$

Recall: We defined ET's to affect only **two** of the n dimensions and only four **adjacent** points in L. With this narrow definition, our set of ET's is **minimal**:

Proposition

All elements of $\mathcal{T}$ are extreme rays of $\mathcal{T}^C$, i.e. no ET can be expressed as a non-negative weighted sum of other ET's.

This result greatly simplifies the pratical application of the theorem above.

Implications of the decomposition theorem

Necessary conditions for $g \succeq_{SPM} f$ ($X \succeq_{SPM} Y$):

1. g and f have identical marginal dists.—since ET's are marginal-preserving
2. $Cov(r(Y_i), s(Y_j)) \geq Cov(r(X_i), s(X_j))$ for all increasing $r(\cdot)$, $s(\cdot)$—implied by identical marginals and supermodularity of $w(z) = r(z_i) \cdot s(z_j)$
3. g dominates f acc. to the "concordance ordering" (Joe, 1990):

$$g \succeq_{CONC} f \Leftrightarrow \forall v \in L, \sum_{z \geq v} g(z) \geq \sum_{z \geq v} f(z) \text{ and } \sum_{z \leq v} g(z) \geq \sum_{z \leq v} f(z)$$

—implied by supermodularity of $w(z) = I_{\{z \geq v\}}$ and $w(z) = I_{\{z \leq v\}}$

Two dimensions vs. more than two dimensions

Proposition

i) *Given f, g with identical 1-diml. marginals, if and only if $n = 2$, there is a* **unique** *set of coefficients $\{\alpha_t\}_{t\in\mathcal{T}}$, of* **arbitrary** *sign, s.t. $g - f = \sum_{t\in\mathcal{T}} \alpha_t t$.*

ii) *For $n = 2$, $g \succeq_{SPM} f \iff g \succeq_{CONC} f$.*

iii) *For $L = \{0,1\}^3$, $g \succeq_{SPM} f \iff g \succeq_{CONC} f$.*

iv) *For any support L strictly larger than $L = \{0,1\}^3$,*

$$g \succeq_{SPM} f \implies g \succeq_{CONC} f \quad \text{but} \quad g \succeq_{CONC} f \not\Longrightarrow g \succeq_{SPM} f.$$

ii) due to Tchen (1980); **iii)** to Hu, Yie, and Ruan (2005); **iv)** to Müller and Scarsini (2000) and Meyer and Strulovici (2012)

How to apply the characterization of $\succeq_{SPM}$ for $n > 2$?

We use our dual characterization to develop several constructive methods:

1. For a given pair g and f, we formulate a linear program whose optimal value is zero if and only if $g \succeq_{SPM} f$.

How to apply the characterization of $\succeq_{SPM}$ for $n > 2$?

We use our dual characterization to develop several constructive methods:

1. For a given pair g and f, we formulate a linear program whose optimal value is zero if and only if $g \succeq_{SPM} f$.

2. To compare many distributions, we develop an algorithm that, for any discrete support L, generates a minimal set of inequalities characterizing $\succeq_{SPM}$ on L.
    - Our algorithm is based on the "double description method" (Motzkin et al., 1953) used to switch btw. alternative representations of polyhedral cones.
    - For any L, it generates the extreme rays of the cone of supermodular functions on L.
    - Easy to code, but algorithmic complexity is high.

How to apply the characterization of $\succeq_{SPM}$ for $n > 2$?

We use our dual characterization to develop several constructive methods:

1. For a given pair g and f, we formulate a linear program whose optimal value is zero if and only if $g \succeq_{SPM} f$.
2. To compare many distributions, we develop an algorithm that, for any discrete support L, generates a minimal set of inequalities characterizing $\succeq_{SPM}$ on L.
   - Our algorithm is based on the "double description method" (Motzkin et al., 1953) used to switch btw. alternative representations of polyhedral cones.
   - For any L, it generates the extreme rays of the cone of supermodular functions on L.
   - Easy to code, but algorithmic complexity is high.
3. With small supports, or various forms of symmetry, we provide direct derivations of the inequalities on $g - f$ that are necessary and sufficient for $g \succeq_{SPM} f$—see Meyer and Strulovici (2012).

The increasing supermodular ordering ($\succeq_{ISPM}$)

In many contexts, we want the random variables not just to be **more dependent** (rather than less) but also to be **higher** (rather than lower). Say $g \succeq_{ISPM} f$ if and only if $w \cdot g \geq w \cdot f$ for all increasing supermodular functions w.

Theorem

$g \succeq_{ISPM} f$ if and only if the following two conditions hold:

1. *for each i, the i^{th} marginal distribution of g stochastically dominates the i^{th} marginal of f;*
2. *$g \succeq_{SPM} (f + \gamma)$, where γ vanishes everywhere except on the "bottom edges" of L and is such that $(f + \gamma)$ and g have identical marginal distributions.*

The symmetric supermodular ordering ($\succeq_{SSPM}$)

In many contexts, natural or convenient to assume that the SPM objective functions are **symmetric** with respect to the n dimensions.

- For ex., a symmetric systemic cost function would be unaffected by permutations of the banks' returns, and symmetry of an ex post welfare function reflects anonymity of individuals.

We define $Y \succeq_{SSPM} X$ if and only if $Ew(Y) \geq Ew(X)$ for all symmetric SPM w.

For any dist. f on $L = (L_1)^n$, its **symmetrized version** f^{symm} is defined by

$$f^{symm}(z) = \frac{1}{n!} \sum_{\sigma \in \Sigma(n)} f(\sigma(z)),$$

where $\Sigma(n)$ is the set of all permutations σ of n-diml. vectors.

Proposition

$$g \succeq_{SSPM} f \iff g^{symm} \succeq_{SPM} f^{symm}.$$

The symmetric supermodular ordering ($\succeq_{SSPM}$)

For $L = \{0,1\}^n$ (**binary** random vectors), the SSPM ordering has a simple form:

Proposition

On $L = \{0,1\}^n$, $Y \succeq_{SSPM} X$ if and only if $\sum_{i=1}^n I_{\{Y_i=1\}} \succeq_{CX} \sum_{i=1}^n I_{\{X_i=1\}}$, where $\succeq_{CX}$ denotes the univariate convex ordering.

- Proof: Any symmetric w defined on $\{0,1\}^n$ can be written as

$$w(z) = \phi(\sum_{i=1}^n z_i)$$

  for some ϕ defined on $\{0,1,\ldots,n\}$. And a symmetric w on $\{0,1\}^n$ is SPM if and only if ϕ is convex.

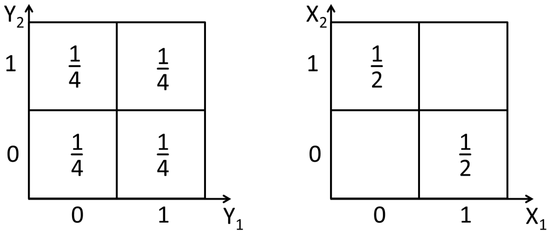
- **Implication:** For **two-point** supports, for any number of dimensions, comparison acc. to the **symmetric** SPM ordering reduces to a well-understood **one-dimensional** problem.

The dispersion ordering of dependence ($\succeq_{DISP}$)

Another notion of greater dependence in Y than in X: the cdfs of the order statistics of Y are closer together or less dispersed than those of X.

- Let $Y_{(j)}$ denote the j^{th} smallest value from $(Y_1, \ldots, Y_n)$ and $F_{Y_{(j)}}$ its cdf.
- Use the **majorization ($\prec$)** ordering of vectors (Hardy et al) to formalize lower dispersion.
- For (Y_1, Y_2) and (X_1, X_2) below,

$$(F_{Y_{(1)}}(0), F_{Y_{(2)}}(0)) = (\frac{3}{4}, \frac{1}{4}) \prec (1, 0) = (F_{X_{(1)}}(0), F_{X_{(2)}}(0)).$$



The dispersion ordering of dependence ($\succeq_{DISP}$)

- For random vector Y and $s \in L$, define **the binary coarsening of** Y **corresponding to** s, Y^s, by $Y_i^s = 0$ if $Y_i \leq s_i$ and $Y_i^s = 1$ if $Y_i > s_i$.
- $Y_{(j)}^s$ is the j^{th} order statistic of Y^s, and $F_{Y_{(j)}^s}$ is its cdf.

Definition

$Y \succeq_{DISP} X$ if and only if for all $s \in L$, the cdfs of the order statistics of Y^s are less dispersed than those of X^s, that is,

$$(F_{Y_{(1)}^s}(0), \ldots, F_{Y_{(n)}^s}(0)) \prec (F_{X_{(1)}^s}(0), \ldots, F_{X_{(n)}^s}(0)) \qquad \forall s \in L.$$

Remark: If $Y \succeq_{DISP} X$, then Y and X have identical 1-diml. marginals.

The dispersion ordering of dependence ($\succeq_{DISP}$)

Proposition

For all $n \geq 2$,

$$Y \succeq_{DISP} X \iff \forall s \in L,\ Y^s \succeq_{SSPM} X^s \iff \forall s \in L,\ \sum_{i=1}^{n} I_{\{Y_i > s_i\}} \succeq_{CX} \sum_{i=1}^{n} I_{\{X_i > s_i\}}.$$

Thus, the dispersion ordering is closely related to the SPM ordering and easily checkable pointwise.

Proposition

i) *For* $n = 2$, $Y \succeq_{SPM} X \iff Y \succeq_{DISP} X$.

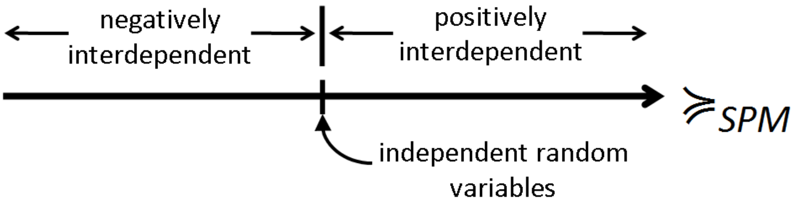
ii) *For* $n = 3$, $Y \succeq_{SPM} X \implies Y \succeq_{DISP} X \iff Y \succeq_{CONC} X$.

iii) *For* $n > 3$, $Y \succeq_{SPM} X \implies Y \succeq_{DISP} X \implies Y \succeq_{CONC} X$.

Sufficient conditions for supermodular ordering

We derive two **formally similar** theorems applicable to **different** environments:

1. Use $\succeq_{SPM}$ to compare **positive** dependence in mixture distributions (mixtures of conditionally independent random variables)
2. Use $\succeq_{SSPM}$ to compare asymmetric independent distributions acc. to degree of asymmetry across dimensions, which is **equivalent** to using $\succeq_{SPM}$ to compare **negative** dependence in symmetrized versions of independent distributions



Mixture distributions: aggregate vs. idiosyncratic shocks

"Mixture dists." (mixtures of conditionally independent random variables) incorporate both aggregate and idiosyncratic shocks. Intuitively, when aggregate shocks increase relative to idiosyncratic shocks, dependence increases.

Our objective: To identify an ordering, for mixture dists., of the relative size of aggregate vs. idiosyncratic shocks that implies $\succeq_{SPM}$-dominance.

Mixture distributions: aggregate vs. idiosyncratic shocks

"Mixture dists." (mixtures of conditionally independent random variables) incorporate both aggregate and idiosyncratic shocks. Intuitively, when aggregate shocks increase relative to idiosyncratic shocks, dependence increases.

Our objective: To identify an ordering, for mixture dists., of the relative size of aggregate vs. idiosyncratic shocks that implies $\succeq_{SPM}$-dominance.

Parametric example: $X_i = \theta + \epsilon_i$; $\quad \theta$, $\{\epsilon_i\}_{i=1}^n$ independent

- $\theta \sim B(\eta_\theta, p)$ is aggregate shock; $\{\epsilon_i\} \sim B(\eta - \eta_\theta, p)$ are idiosyncratic shocks
- Raising η_θ while holding η fixed leaves the marginal dist. of each X_i unchanged but increases the relative importance of the aggregate shock.
- Our mixture distribution theorem implies that raising η_θ, holding η fixed, makes $(X_1, \ldots, X_n)$ more supermodularly dependent.

Our theorem is non-parametric.

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Cautionary example: Even with the additive structure $X_i = \theta + \epsilon_i$, for *arbitrary* distributions an increase in $Var(\theta)$, holding marginal dists. fixed, does not generally make $(X_1, \ldots, X_n)$ more supermodularly dependent.

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Cautionary example: Even with the additive structure $X_i = \theta + \epsilon_i$, for *arbitrary* distributions an increase in $Var(\theta)$, holding marginal dists. fixed, does not generally make $(X_1, \ldots, X_n)$ more supermodularly dependent.

- $X_i = \theta_x + \epsilon_i$ and $Y_i = \theta_y + \delta_i$
- θ_x, $\{\epsilon_i\}_{i=1}^n$ independent and θ_y, $\{\delta_i\}_{i=1}^n$ independent
- θ_x and $\{\delta_i\}_{i=1}^n$ iden. dist. on $\{-2, 0, 2\}$, with probs. $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$, resp.
- θ_y and $\{\epsilon_i\}_{i=1}^n$ iden. dist. on $\{-1, 0, 1\}$, with probs. $\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$, resp.
- The random vectors X and Y have identical marginal distributions.
- θ_x is more variable than θ_y, and, $\forall i$, ϵ_i is less variable than δ_i, in the sense of the convex ordering.

But X and Y cannot be ranked according to $\succeq_{SPM}$:

- $P(X_1 \geq 3, X_2 \geq 3) = \frac{1}{4}\left(\frac{1}{8}\right)^2 < \frac{1}{8}\left(\frac{1}{4}\right)^2 = P(Y_1 \geq 3, Y_2 \geq 3)$
- $P(X_1 \geq 2, X_2 \geq 2) = \frac{1}{4}\left(\frac{7}{8}\right)^2 > \frac{7}{8}\left(\frac{1}{4}\right)^2 = P(Y_1 \geq 2, Y_2 \geq 2)$

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Example: **A**, **B**, **C** generate n-diml. symmetric mixture dists. on $L = \{0, 1, 2\}^n$:

- First, random selection of a row, with all rows equally likely: this represents realization of aggregate shock
- Second, n i.i.d. draws from the prob. dist. given by the selected row: this represents realizations of idiosyncratic shocks

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{8} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Example: **A**, **B**, **C** generate n-diml. symmetric mixture dists. on $L = \{0, 1, 2\}^n$:

- First, random selection of a row, with all rows equally likely: this represents realization of aggregate shock
- Second, n i.i.d. draws from the prob. dist. given by the selected row: this represents realizations of idiosyncratic shocks

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{8} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

- Intuitively, the **more different** are the rows, the **more important** is the aggregate shock and the **more interdependent** are $(X_1, \ldots, X_n)$.
- But we must hold fixed the "expected" dist. (i.e. the average of the rows), to ensure identical marginals. Here, **A**, **B**, **C** each generate a symmetric mixture dist. with marginals $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ on $\{0, 1, 2\}$.
- In each matrix, bottom row stochastically dominates top row $\Longrightarrow$ bottom row represents better realization of aggregate shock.

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Ex.: **A**, **B**, and **C** generate n-diml. symmetric mixture dists. on $L = \{0, 1, 2\}^n$.

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{8} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

- Intuitively, the **more different** are the rows, the **more important** is the aggregate shock and the **more interdependent** are $(X_1, \ldots, X_n)$.
- The rows of **C** are identical (i.e. no aggregate shock), the rows of **B** differ from each other, and the rows of **A** are differ from each other even more.
- **Result:** The mixture dist. derived from **A** $\succeq_{SPM}$-dominates that derived from **B**, which $\succeq_{SPM}$-dominates the independent dist. derived from **C**.

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Ex.: **A**, **B**, and **C** generate n-diml. symmetric mixture dists. on $L = \{0, 1, 2\}^n$.

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} \frac{3}{8} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

- Intuitively, the **more different** are the rows, the **more important** is the aggregate shock and the **more interdependent** are $(X_1, \ldots, X_n)$.
- The rows of **C** are identical (i.e. no aggregate shock), the rows of **B** differ from each other, and the rows of **A** are differ from each other even more.
- **Result:** The mixture dist. derived from **A** $\succeq_{SPM}$-dominates that derived from **B**, which $\succeq_{SPM}$-dominates the independent dist. derived from **C**.

Our theorem generalizes the example and formalizes the intuition above:

- We provide a partial ordering on $q \times l$ matrices representing "the rows are more different from one another, holding fixed the average of the rows".
- We prove that this ordering is sufficient for $\succeq_{SPM}$-dominance btw. the symmetric mixture dists. (of any dimension n) generated by the matrices.

Mixture dists.: sufficient conds. for $\succeq_{SPM}$-dominance

Theorem

Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ have symmetric mixture distributions generated from $q \times l$ row-stochastic matrices **A** *and* **B**, *respectively.*

Conditions **i)** *and* **ii)** *below are sufficient for $(X_1, \ldots, X_n) \succeq_{SPM} (Y_1, \ldots, Y_n)$:*

i) *The rows of* **A** *are stochastically ordered.*

ii) **A** *dominates* **B** *according to "cumulative column majorization", i.e.*

$$(\sum_{j=1}^{k} a_{1j}, \ldots, \sum_{j=1}^{k} a_{qj}) \succ (\sum_{j=1}^{k} b_{1j}, \ldots, \sum_{j=1}^{k} b_{qj}) \quad \forall k \in \{1, \ldots, l-1\}.$$

Remark: The theorem can be extended to **asymmetric** mixture distributions.

Heterogenous lotteries: sufficient conditions for $\succeq_{SSPM}$

Application: Optimal allocation of resources across projects (Hoeffding, 1956; Karlin and Novikoff, 1963; Bond and Gomes, 2009)

- Let $(X_1, \dots, X_n) \in \{0,1\}^n$ be outcomes (failure or success) on n projects
- $Pr(X_i = 1) = p_i$; conditional on $(p_1, \dots, p_n)$, outcomes **indep.** across tasks

Question: How to allocate a fixed $\sum_{i=1}^n p_i$ across projects to maximize $Ew(X_1, \dots, X_n)$ for w **symmetric** across projects?

Answer: If w is symmetric and SPM, optimal to choose **equal** success probability for all projects. If w is symmetric and submodular, optimal to choose success probabilities as **unequally** as possible (subj. to fixed $\sum_i p_i$).

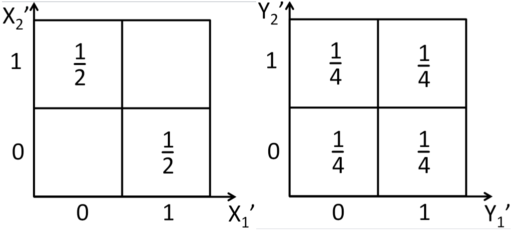
- For $X_i \in \{0,1\}$, any symmetric SPM $w = \phi(\sum_i X_i)$, where ϕ is convex. Dist. of $\sum_i X_i$ is "riskier", the less dispersed (in the sense of majorization) are the components of $(p_1, \dots, p_n)$.

$$\mathbf{A} = \begin{pmatrix} 1-p_1 & p_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1-p_n & p_n \end{pmatrix} \begin{matrix} \leftarrow \text{project } 1 \\ \\ \\ \\ \\ \leftarrow \text{project } n \end{matrix}$$

Heterogeneous lotteries: sufficient conds. for $\succeq_{SSPM}$

Key idea: Lower dispersion among heterogeneous independent lotteries, holding fixed the average of the lotteries $\implies \succeq_{SSPM}$-dominance of the indep. dists. $\iff \succeq_{SPM}$-dominance of the symmetrized versions of the distributions.

- *Example:* $P(X_1 = 1) = 0,\ P(X_2 = 1) = 1$ vs. $P(Y_1 = 1) = P(Y_2 = 1) = \frac{1}{2}$
- Symmetrized version of (X_1, X_2) is (X_1', X_2') and of (Y_1, Y_2) is (Y_1', Y_2')
- Since $(0, 1) \succ (\frac{1}{2}, \frac{1}{2})$, $(X_1, X_2) \preceq_{SSPM} (Y_1, Y_2)$ and $(X_1', X_2') \preceq_{SPM} (Y_1', Y_2')$



The X lotteries are more heterogeneous, and their symmetrized version X' displays more negative interdependence.

Sufficient conditions for $\succeq_{SSPM}$

Lower dispersion among heterogeneous independent lotteries, holding fixed the average of the lotteries $\implies \succeq_{SSPM}$-dominance of the independent dists.
$\iff \succeq_{SPM}$-dominance of the symmetrized versions of the distributions.

Theorem

Let $(X_1, \ldots, X_n)$ be **indep.** *with dist. f and $(Y_1, \ldots, Y_n)$ be* **indep.** *with dist. g. Let the dist. of X_i (resp. Y_i) be described by the i^{th} row of an $n \times l$ row-stochastic matrix* $\mathbf{A}$ *(resp.* $\mathbf{B}$*). Define $(X'_1, \ldots, X'_n) \sim f^{symm}$ and $(Y'_1, \ldots, Y'_n) \sim g^{symm}$.*

Conditions **i)** *and* **ii)** *below are sufficient for $(X_1, \ldots, X_n) \preceq_{SSPM} (Y_1, \ldots, Y_n)$ (equivalently, $X' \preceq_{SPM} Y'$):*

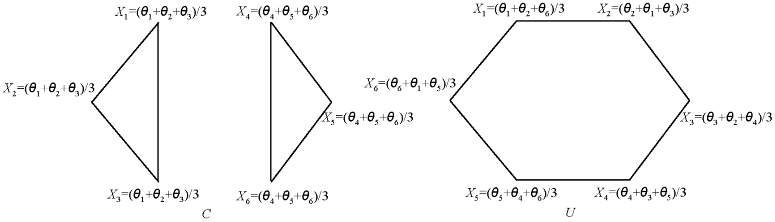
i) *The rows of* $\mathbf{A}$ *are stochastically ordered.*

ii) $\mathbf{A}$ *dominates* $\mathbf{B}$ *according to "cumulative column majorization", i.e.*

$$\left(\sum_{j=1}^{k} a_{1j}, \ldots, \sum_{j=1}^{k} a_{qj}\right) \succ \left(\sum_{j=1}^{k} b_{1j}, \ldots, \sum_{j=1}^{k} b_{qj}\right) \quad \forall k \in \{1, \ldots, l-1\}.$$

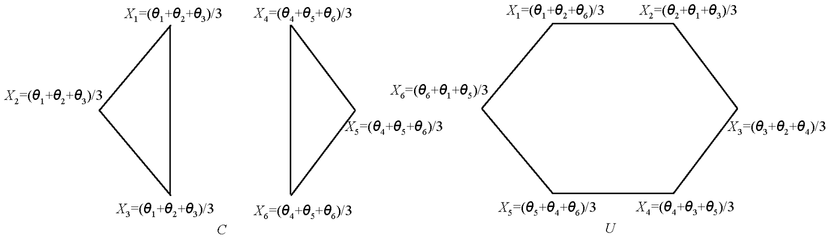
Asset-swapping among banks and systemic risk

Based on Allen, Babus, and Carletti's (2012) analysis of banking networks:



- Project returns $(\theta_1, \ldots, \theta_6) \in \{L, H\}^6$ are i.i.d., with $P(\theta_i = H) = p$.
- In network **C**, banks form 2 clusters; network **U** is unclustered.
- The marginal dist. of the return to each bank's portfolio is the same in the two networks, but interdependence differs.
- Suppose a bank fails (default status=1) if its reutrn is $\leq d \in [L, H)$, otherwise it is solvent (default status=0). Let banks' default statuses in clustered network be described by $(Y_1, \ldots, Y_6) \in \{0, 1\}^6$ and in unclustered network by $(Z_1, \ldots, Z_6) \in \{0, 1\}^6$.

Asset-swapping among banks and systemic risk



Proposition

For any prob. of project success p and for any common failure threshold d, $(Y_1, \ldots, Y_6) \succeq_{SSPM} (Z_1, \ldots, Z_6)$. Hence for any systemic cost function which is a symmetric and supermodular function of bank failures, expected systemic cost is higher under the clustered than under the unclustered network.

Conclusion

1. Provided a characterization of the supermodular ordering in terms of elementary, dependence-increasing transformations
2. Developed constructive methods for applying the characterization
3. Characterized related orderings (ISPM, SSPM, DISP) in terms of SPM ordering
4. Identified sufficient conditions for SPM or SSPM dominance in specific environments
5. Applications to welfare economics, committee decision-making, systemic risk, matching,...

Sufficient conditions for $\succeq_{SSPM}$

Application 2: Ex post inequality under uncertainty: tournaments vs. independent schemes

- Given n indivs. and n prizes $\{P_1, \ldots, P_n\}$, with $P_i \geq P_j \ \forall i < j$, a **tournament** assigns positive prob. only to outcome vectors that are permutations of $(1, \ldots, n)$. A **symmetric tournament** assigns equal prob. to each of these $n!$ vectors.
- Intuitively, tournaments generate **negative interdependence** in rewards.
- Compare tournament reward schemes with schemes that provide each indiv. with same marginal dist. over rewards but determine rewards independently.
- Does an arbitrary tournament generate "more negative interdependence" acc. to $\preceq_{SPM}$ than the corres. independent scheme? Or if not, acc. to $\preceq_{SSPM}$?

Proposition

(Meyer and Mookherjee, 1987) For any n, the dist. of prizes in a symmetric tournament is dominated acc. to $\preceq_{SPM}$ by the corresponding independent dist. with the same marginals.

Tournaments vs. independent schemes

Example: Tournament assigns prob. $\frac{1}{2}$ to outcome vectors (P_1, P_2, P_3) and (P_2, P_3, P_1).

- This tournament is **not** dominated acc. to $\preceq_{SPM}$ by the corresponding independent scheme.
- Reason: while rewards for indivs. 1 and 3 (or 2 and 3) are **negatively** dependent, rewards for indivs. 1 and 2 are **positively** dependent.
- This shows the subtlety of "negative interdependence" in $n > 2$ dimens.

But in this context, natural to use orderings which treat indivs. **symmetrically** (anonymously), so natural to use the weaker ordering $\preceq_{SSPM}$.

Corollary (of $\succeq_{SSPM}$ Theorem)

For any number of individuals (dimensions) n, given any (arbitrarily asymmetric) tournament, the prize distribution under the tournament is dominated acc. to $\preceq_{SSPM}$ *by the corresponding independent dist. with the same marginals. Equivalently, for any tournament, its symmetrized distribution displays more negative interdependence acc. to* $\preceq_{SPM}$ *than the symmetrized version of the corresponding independent scheme with the same marginals.*

The symmetric supermodular ordering ($\succeq SSPM$)

For $L = \{0,1\}^n$ (**binary** random vectors), the SSPM ordering has a simple form. Define $c(x) \equiv \sum_{i=1}^{n} I_{\{x_i=1\}}$. Any symmetric w defined on $L = \{0,1\}^n$ can be

written as $w(x) = \tilde{w}(c(x))$ for some $\tilde{w}$ defined on $\tilde{L}^1 \equiv \{0,1,\dots,n\}$. Any dist. f on L maps into a dist. $\tilde{f}$ on $\tilde{L}^1$ defined by $\tilde{f}(z) = \sum_{\{x:c(x)=z\}} f(x)$. ET's on L map into (adjacent) mean-preserving spreads on $\tilde{L}^1$, and a symmetric w defined on L is SPM iff ϕ defined on $\tilde{L}^1$ is convex.

Proposition

On $L = \{0,1\}^n$, $g \succeq_{SSPM} f$ if and only if $\tilde{g}$ dominates $\tilde{f}$ according to the univariate convex ordering on $\tilde{L}^1 = \{0,1,\dots,n\}$.

- For **two-point** supports, for any number n of dimensions, comparison acc. to the **symmetric** SPM ordering reduces to a well-understood **one-dimensional** problem.

The symmetric supermodular ordering ($\succeq SSPM$)

For **l-point** supports, for any number n of dimensions, comparison acc. to the **symmetric** SPM ordering reduces to an **(l-1)-dimensional** problem:

- For $x \in L = \{0, 1, \ldots, l-1\}^n$ and $k \in \{0, 1, \ldots, l-1\}$, define $\bar{c}^k(x) = \sum_{i=1}^n I_{\{x_i \geq k\}}$ and the "cumulative count vector" $\bar{c}(x) = (\bar{c}^1(x), \ldots, \bar{c}^{l-1}(x)) \in \tilde{L}^{l-1} \subset \{0, 1, \ldots, n\}^{l-1}$.
- For any symmetric w on L, $w(x) = \tilde{w}(\bar{c}(x))$ for some $\tilde{w}$ defined on $\tilde{L}^{l-1}$.
- Any dist. f on L maps into a dist. $\tilde{f}$ on $\tilde{L}^{l-1}$ with $\tilde{f}(z) \equiv \sum_{\{x:\bar{c}(x)=z\}} f(x)$.
- Each ET on L maps into **either** a univariate (adjacent) mean-preserving spread on $\tilde{L}^{l-1}$ **or** an ET on $\tilde{L}^{l-1}$.
- A symmetric w defined on L is SPM iff $\tilde{w}$ defined on $\tilde{L}^{l-1}$ is component-wise convex and SPM.

Proposition

On $L = \{0, 1 \ldots, l-1\}^n$, $g \succeq_{SSPM} f$ if and only if $\tilde{g}$ dominates $\tilde{f}$ according to the componentwise-convex and supermodular ordering on $\tilde{L}^{l-1}$.

Continuous Supports

- Let dists. F,G have continuous densities on a convex, compact lattice $L = \Pi_i L_i$ of $\mathbb{R}^n$.
- Say $G \succeq_{CSPM} F$ if for all integrable supermodular functions w on L, $E[w|G] \geq E[w|F]$.
- A finite coarsening $\tilde{L}$ of L is a finite partitioning $\tilde{L}_i$ of each L_i, s.t. $x \in \tilde{L}$ represents a hyper-rectangle $\Pi_i[l_i, u_i)$ of L. The coarsening of F on $\tilde{L}$ is the dist. $\tilde{F}$ s.t. $\forall x \in \tilde{L}$, $\tilde{F}(x) = F(\Pi_i[l_i, u_i))$. Similarly, the coarsened version $\tilde{w}$ of w on $\tilde{L}$ is

$$\tilde{w}(x) = \frac{\int_{\Pi_i[l_i,u_i)} w(t)dt}{\int_{\Pi_i[l_i,u_i)} dt}.$$

Proposition: $G \succeq_{CSPM} F$ if and only if for all finite coarsenings $\tilde{L}$ of L, $\tilde{G} \succeq_{SPM} \tilde{F}$ in the discrete sense.