

Increasing Interdependence of Multivariate Distributions

Margaret Meyer

Bruno Strulovici*

April 27, 2010

Abstract

In many economic contexts, it is of interest to know whether one set of random variables displays a greater degree of interdependence than another. Orderings of interdependence are useful in the assessment of ex post inequality under uncertainty; in comparisons of multidimensional inequality; in assessments of the degree of conformity of behavior in social learning situations; in comparisons of the efficiency of matching institutions; in the valuation of portfolios of assets or insurance policies; and in assessments of systemic risk. This paper explores five orderings of interdependence for multivariate distributions: greater weak association, the supermodular ordering, the convex-modular ordering, the dispersion ordering, and the concordance ordering. We show that for two dimensions, all five orderings are equivalent, whereas for an arbitrary number of dimensions $n > 2$, the five orderings are strictly ranked. For the special case of binary random variables, we establish some equivalence results among the orderings. We conclude by illustrating the application of our orderings to the comparison of interdependence in behavior in a model of learning in networks and to the assessment of ex post inequality under uncertainty.

*Email addresses: margaret.meyer@nuffield.ox.ac.uk b-strulovici@northwestern.edu

1 Introduction

1.1 Motivation

In many economic contexts, it is of interest to know whether one set of random variables displays a greater degree of interdependence than another. This paper explores several orderings of interdependence for multivariate distributions, establishes the relationships among the orderings, both in general as well as in important special cases, and illustrates their application to economic problems.

Orderings of interdependence are applicable in several welfare-economic contexts. In many group settings where individual outcomes (e.g. rewards) are uncertain, members of the group may be concerned, *ex ante*, about how unequal their *ex post* rewards will be (Meyer and Mookherjee, 1987; Ben-Porath et al, 1997; Gajdos and Maurin, 2004; Kroll and Davidovitz, 2003; Adler and Sanchirico, 2006). (This concern is distinct from concerns about the mean level of rewards and about their riskiness.) Comparisons of reward schemes then require comparisons of the degree of interdependence of the random rewards. Another welfare-economic example concerns comparisons of inequality when separate data are available on attributes such as income, health, and education (Atkinson and Bourguignon, 1982; Bourguignon and Chakravarty, 2002; Atkinson, 2003; and Decancq, 2007). As long as the function aggregating the different attributes into an overall measure of welfare or deprivation is not additively separable across attributes, comparisons of multidimensional inequality will necessitate comparisons of the degree of interdependence among the attributes.

In studies of social learning, where the goal is to understand the degree and determinants of conformity of behavior, interdependence orderings can be used to assess how the degree of conformity of (*ex ante* random) behavior varies with time and with changes in social structure (Choi, Gale, and Kariv, 2005). Interdependence orderings are also useful for comparing the efficiency of matching mechanisms. In many matching contexts, perfectly assortative matching would be efficient | such matching would correspond to a "perfectly positively dependent" joint distribution of the random variables representing quality in each category (dimension). When, however, matches are formed based only on noisy or coarse information (McAfee, 2002), or when search is costly (Shimer and Smith, 2000), or when signaling is constrained by market imperfections such as borrowing constraints (Fernandez and Gali, 1999), perfectly assortative matching will generally not arise. Fernandez and Gali (1999) and Meyer and Rothschild (2004) apply bivariate dependence orderings to compare the performance of different matching institutions.¹

¹In a related application, Prat (2002) explores how the composition of employee teams affects interdependence in (*ex ante* random) decisions of team members and shows how properties of the the production function translate into

In finance and insurance, valuing portfolios of assets or insurance policies requires assessing the degree of interdependence among asset returns or insurance claims (Müller and Stoyan, 2002, and Denuit et al, 2005). Financial economists and macroeconomists are, moreover, increasingly interested in measures and comparisons of "systemic risk" in financial and economic systems to capture the interdependence in the returns of different institutions, sectors, or regions (Hennessey and Lapan, 2003, Adrian and Brunnermeier, 2009, and Acharya, 2009).

For the special case of bivariate distributions, economists and statisticians have shown that two intuitive concepts of greater interdependence are in fact equivalent. Suppose we are comparing the degree of interdependence between (Y_1, Y_2) with that between (X_1, X_2) , where for each i , Y_i and X_i have the same marginal distribution. The first concept is "lower orthant dominance", which requires that for all points in the support of the random vectors Y and X , the cumulative distribution function of Y be at least as large as the c.d.f. of X : this captures the requirement that the components of Y are more likely than those of X to both be "low" together, for any thresholds determining the precise meaning of "low". Given the assumption of identical marginal distributions, lower orthant dominance is equivalent to "upper orthant dominance", which requires that the components of Y be more likely than those of X to both be "high" together. The second concept of greater interdependence is "supermodular dominance", which requires that $Ew(Y)$ be at least as large as $Ew(X)$ for all objective functions that are supermodular in their two arguments. Supermodularity (see Topkis, 1978, and Section 2.3 below) is a natural property of an objective function with which to capture a preference for interdependence, since it captures the idea that its arguments are complements, not substitutes: When an increasing function of two variables is supermodular and the variables are increased together, the resulting increase in the function is larger than the sum of the increases that would result from increasing each variable separately. It has been shown by Levy and Parousch (1974), Epstein and Tanny (1980), Tchen (1980), and Atkinson and Bourguignon (1982) that for two-dimensional random vectors with identical marginals, lower-orthant dominance of Y over X is equivalent to $Ew(Y) \geq Ew(X)$ for all supermodular objective functions.

Economists have, nevertheless, made very little progress in the development of orderings for comparing interdependence in multivariate, as opposed to bivariate, distributions. On the one hand, this is surprising, given the wide variety of applications for such orderings in both theoretical and empirical work. On the other hand, though, this lack of progress is less surprising given that, as we now argue, the n -dimensional case is substantially more difficult than the 2-dimensional case, for several reasons.

preferences over team composition.

First, whereas positive and negative interdependence are “mirror images” of each other in two dimensions, this symmetry breaks down for more than two dimensions. In two dimensions, for any plausible concept of positive dependence, if Y_1 and Y_2 are positively dependent, then $-Y_1$ and Y_2 are negatively dependent. Moreover, for Y_1 and Y_2 with identical uniform marginal distributions on $[0, 1]$, perfect positive dependence corresponds to $Y_2 = Y_1$, while perfect negative dependence corresponds to $Y_2 = 1 - Y_1$. For more than two dimensions, however, there is in general no simple way to convert a positively interdependent random vector (Y_1, Y_2, \dots, Y_n) into a negatively interdependent one. And even for $\{Y_i\}_{i=1}^n$ with identical uniform marginals on $[0, 1]$, while $Y_1 = Y_2 = \dots = Y_n$ represents perfect positive dependence, there is no obvious definition of perfect negative dependence.² Since multivariate concepts of greater and lesser interdependence should be applicable to distributions displaying either positive or negative dependence, the lack of symmetry between positive and negative dependence in $n > 2$ dimensions complicates the development of orderings.³

Second, for more than two dimensions, there are more distinct notions of greater interdependence than there are for two dimensions. Section 2 below presents five dependence orderings, three previously defined and two new to this paper. Whereas for two dimensions, all five orderings are equivalent, for an arbitrary number of dimensions $n > 2$, the five orderings are strictly ranked. Thus, the selection of orderings of interdependence is more complicated for multivariate distributions than for bivariate ones.

Finally, even for a given ordering of greater interdependence, determining whether two multivariate distributions can be ranked according to the ordering may be more difficult than for bivariate distributions. For two dimensions, it is straightforward to determine whether one distribution dominates another in the sense of “lower orthant dominance”: this requires comparing the cumulative distribution functions at every point in the common support. For more than two dimensions, some of the dependence orderings we study below can be implemented in an analogous fashion, by determining, for each point in the support, whether a given set of inequalities is satisfied. However, for other orderings, there exists no set of criteria that can be applied in a pointwise fashion; for

²For the set of all distribution functions with given marginals F_1, F_2, \dots, F_n , there exist upper and lower bounds for the distribution function, termed “Fréchet bounds”. The “upper Fréchet bound” is the natural candidate for the distribution exhibiting maximal positive dependence. However, while the “lower Fréchet bound” might seem like a natural candidate for the distribution displaying maximal negative dependence, the lower Fréchet bound is not in fact a proper distribution function except in very special cases (which do not include the uniform example described in the text). See Joe (1997) and Müller and Stoyan (2002) for more details.

³Unfortunately, this lack of symmetry between positive and negative dependence for more than two dimensions is not always recognized. See, for example, the discussion in Galeotti et al (2010, p.229, fn. 12).

implementing these orderings, more sophisticated techniques or algorithms need to be developed.

1.2 Outline

In the next section, we present five orderings of greater interdependence for multivariate distributions. We begin with an ordering inspired by the dependence concept of "association", a concept originally proposed by Esary, Proschan, and Walkup (1967) and widely used in the statistical literature.⁴ ⁵ Since the definition of (positive) association cannot be meaningfully reversed to define a concept of negative association, we turn to the concept of "weak association", originally defined by Burton, Dabrowski, and Dehling (1986). Based on this notion, we define a concept of greater interdependence: A random vector (Y_1, \dots, Y_n) displays greater weak association than a random vector (X_1, \dots, X_n) if the random vectors have identical marginals and, for all disjoint subsets A, B of $\{1, \dots, n\}$ and nondecreasing functions r, s , $Cov(r(Y_i, i \in A), s(Y_j, j \in B)) \geq Cov(r(X_i, i \in A), s(X_j, j \in B))$. This ordering has been defined in the actuarial literature (see Denuit et al (2005)), where it has been (somewhat confusingly) termed the "correlation order".

We then present the (multivariate) supermodular ordering, according to which (Y_1, \dots, Y_n) displays greater interdependence than (X_1, \dots, X_n) if for all supermodular objective functions, $Ew(Y_1, \dots, Y_n) \geq Ew(X_1, \dots, X_n)$. This is the natural multivariate generalization of the concept of supermodular dominance discussed above for bivariate distributions. Levy and Parousch (1974) and Epstein and Tanny (1980) proposed this ordering in the context of asset allocation, while Atkinson and Bourguignon (1982) proposed it for the assessment of two-dimensional inequality. Meyer and Mookherjee (1987) and Meyer (1990) contain early proposals to use it as a multivariate dependence ordering, focusing on the assessment of ex post inequality under uncertainty. In the statistical literature, the supermodular ordering was initially studied by Tchen (1980) in the bivariate case and then formalized for the multivariate case by Shaked and Shanthikumar (1997). Meyer and Strulovici (2010) use duality methods to characterize the supermodular ordering and develop several constructive methods to implement this characterization.

In Section 2.4, we introduce a new dependence ordering, which we term the "convex-modular or-

⁴Association, though a strong concept of dependence, is strictly weaker than "affiliation", which is familiar to economists from the work of Milgrom and Weber (1982), who used it to formalize positive interdependence among bidders' valuations in auctions. De Castro (2009) also notes that affiliation is strictly stronger than other concepts of positive dependence.

⁵Though association has not been widely used in economics, in a pair of recent papers, Calvo-Armengol and Jackson (2004, respectively 2007) show that the employment statuses (respectively, wages) of individuals connected by a social network are positively dependent in the sense of association.

dering". We say an objective function is convex-modular if it has the form $\phi(\sum_{i=1}^n r_i(z_i))$, where $\phi(\cdot)$ is convex and $\{r_i\}_{i=1}^n$ are nondecreasing. (Y_1, \dots, Y_n) is said to be more interdependent than (X_1, \dots, X_n) in the sense of the convex-modular ordering if for all convex-modular objective functions, $Ew(Y_1, \dots, Y_n) \geq Ew(X_1, \dots, X_n)$. Greater interdependence for Y than for X corresponds here to greater riskiness, in the sense of Rothschild and Stiglitz (1970), of *any* aggregate $\sum_{i=1}^n r_i(Y_i)$ compared to $\sum_{i=1}^n r_i(X_i)$, for any choice of $\{r_i\}_{i=1}^n$ nondecreasing. Convex-modular objective functions are a natural way of capturing insurance companies' preferences over the degree of dependence in the claims arising from portfolios of insurance policies.

Section 2.5 introduces another new dependence ordering, which we term the dispersion ordering. Consider a random vector (Y_1, \dots, Y_n) such that, with probability 1, $Y_1 = Y_2 = \dots = Y_n$. For this perfectly positively dependent distribution, all of the order statistics of (Y_1, \dots, Y_n) are, with probability 1, equal, so the cumulative distribution functions of the order statistics are equal everywhere. For any other random vector (X_1, \dots, X_n) with a distribution symmetric across dimensions and the same univariate marginal as (Y_1, \dots, Y_n) , the order statistics of (X_1, \dots, X_n) will in general not all be equal, so their cdf's will diverge from one another. This observation suggests comparing interdependence in random vectors by comparing the dispersion of the cdf's of the order statistics, with lower dispersion representing greater interdependence.

For random vectors with symmetric distributions, Shaked and Tong (1985) proposed a dependence ordering along these lines, using the majorization ordering of vectors to compare dispersion of the cdf's of the order statistics.⁶ We reformulate their ordering in a manner which suggests a new ordering that is both stronger and naturally applicable to asymmetric distributions. To each point (s_1, \dots, s_n) in the support of a random vector (Y_1, \dots, Y_n) , there corresponds a binary coarsening of Y , Y^s , defined by $Y_i^s = 0$ if $Y_i \leq s_i$ and $Y_i^s = 1$ if $Y_i > s_i$. We say that Y displays greater interdependence than X according to the dispersion ordering if, for each point s in the support, the distribution functions of the order statistics of Y^s are less dispersed than the distribution functions of the order statistics of X^s , where lower dispersion is assessed with the majorization ordering.

Finally, we present the concordance ordering, which was formalized for multivariate distributions by Joe (1990). A random vector Y is more concordant than X if, for any point in the support, the components of Y are more likely to be all higher than at that point, relative to those of X , and also more likely to be all lower than at that point, relative to those of X . The concordance ordering combines the requirement of upper-orthant dominance with that of lower-orthant dominance | with

⁶A vector a is said to be majorized by a vector b , written $a \prec b$, if i) the components of the vectors have the same total sum, and ii) for all k , the sum of the k largest entries of a is weakly smaller than the sum of the k largest entries of b (see Hardy, Littlewood, and Polya (1952)).

more than two dimensions, even if Y and X have identical marginal distributions, upper-orthant dominance and lower-orthant dominance are no longer equivalent.

Section 3 is devoted to the relationships among the five orderings just described. Whereas for two dimensions, all five orderings are equivalent (Theorem 1), Theorem 2 shows that for an arbitrary number of dimensions $n > 2$, the five orderings are strictly ranked: Greater weak association is strictly stronger than the supermodular ordering, which is strictly stronger than the convex-modular ordering. In turn, the convex-modular ordering is strictly stronger than the dispersion ordering, which is strictly stronger than the concordance ordering. (If the number of dimensions is exactly three, then the dispersion and concordance orderings are equivalent.)

Section 4 focuses on binary random variables. Binary random variables, besides being common in theoretical, experimental, and empirical applications, also help to illuminate the structure of and relationships among the interdependence orderings. We study a variety of special cases with binary random variables and highlight i) equivalences among the dependence orderings that arise in these special cases and ii) easily checkable and easily interpretable necessary and sufficient conditions for the orderings to hold.

Section 5 presents two applications of our orderings. In Section 5.1, we study the influence of network structure on the degree of conformity of behavior in social learning situations. We adapt the theoretical approach of Gale and Kariv (2003) and its experimental implementation, with binary actions, in Choi, Gale, and Kariv (2005). We illustrate how our orderings, and in particular our results in Section 4 for binary random variables, can be applied to compare interdependence in behavior as network structure varies.

In Section 5.2, we illustrate the use of our orderings in the assessment of ex post inequality under uncertainty. When groups dislike ex post inequality, tournament reward schemes would appear to be particularly unappealing, since they generate a form of negative correlation among rewards: if one individual receives a higher reward, another person must receive a lower one. We use our interdependence orderings to formally compare tournaments with schemes that provide each individual with the same marginal distribution over rewards but determine rewards *independently*. We show that a symmetric tournament generates a very strong form of negative interdependence in rewards but that arbitrarily asymmetric tournaments do not dominate their corresponding independent counterparts even according to the weakest of our dependence orderings. This is a reflection of the subtleties of negative interdependence in more than two dimensions. Nevertheless, we show that if we adopt an ordering which treats individuals symmetrically with respect to the assessment of interdependence, such as the symmetric supermodular or symmetric convex-modular ordering,

then in three dimensions, any tournament, no matter how asymmetric, does display more negative interdependence than its independent counterpart.

We offer a brief conclusion in Section 6.

2 Orderings of Greater Interdependence

2.1 Preliminaries

We consider multivariate distributions with the same number, n , of variables and identical, finite support. Focusing on finite supports simplifies notation, avoids some uninteresting technical issues, and clarifies the underlying structure of the orderings and the relationships between them.

Formally, let L_i denote the finite, totally ordered set of values taken by the i^{th} random variable, and let L denote the cartesian product of the L_i 's. L_i is a finite subset of \mathbb{R} , and L is a finite lattice of \mathbb{R}^n with the following partial order: $z \leq v$ if and only if $z_i \leq v_i$ for all $i \in N = \{1, \dots, n\}$. If l_i denotes the cardinality of L_i , then L has $d = \prod_{i=1}^n l_i$ elements.

All of the interdependence orderings we consider below will be invariant to monotonic relabelings of the elements of the support for each component. To state this invariance property formally: If the random vector (Y_1, Y_2, \dots, Y_n) dominates the random vector (X_1, X_2, \dots, X_n) according to some ordering \succeq_O , denoted $Y \succeq_O X$, then

$$(r_1(Y_1), r_2(Y_2), \dots, r_n(Y_n)) \succeq_O (r_1(X_1), r_2(X_2), \dots, r_n(X_n)) \quad (1)$$

whenever $r_i : \mathbb{R} \rightarrow \mathbb{R}$, $i \in \{1, \dots, n\}$, are all nondecreasing. Given that the orderings all satisfy this invariance, it is without loss of generality, and convenient notationally, to assume henceforth that the sets L_i of values taken by the i^{th} random variable each have the form $\{0, 1, \dots, l_i - 1\}$.

Such an invariance property is a natural desideratum for an ordering of interdependence.⁷

Note, though, that most orderings of variability, such as Rothschild and Stiglitz's (1970) ordering of greater riskiness, will fail to satisfy this invariance.⁸

⁷See, for example, the discussion in Joe (1997, p. 39).

⁸For the support $\{0, 1, 2, 3, 4, 5\}$, suppose that the difference between two distributions g and f is $\{\epsilon, -2\epsilon, \epsilon, \epsilon, -2\epsilon, \epsilon\}$, for some $\epsilon > 0$. Then g is derived from f by two mean-preserving spreads, one on $\{0, 1, 2\}$ and the other on $\{3, 4, 5\}$. If, however, the support is transformed to $\{0, 2, 3, 5\}$ by relabeling 1 as 0 and 4 as 5, keeping all other labels unchanged, then on this relabeled support, the difference between the distributions is $\{-\epsilon, \epsilon, \epsilon, -\epsilon\}$,

For any $z \in L$, let $z + e_i$ denote the element v of L , whenever it exists, such that $v_j = z_j$ for all $j \in N \setminus \{i\}$ and v_i is the smallest element of L_i greater than but not equal to z_i . For example, if $L = 0, 1^2$, $(0, 0) + e_1 = (1, 0)$ and $(1, 0) + e_2 = (0, 0) + e_1 + e_2 = (1, 1)$.

The lattice structure of the support L and its corresponding order are useful for comparing distributions. One may label the d elements (or "nodes") of L and view real functions on L as vectors of \mathbb{R}^d , where each coordinate of the vector corresponds to the value of the function at a specific node of L . Similarly, a multivariate distribution whose support is L can be represented as an element of the unit simplex Δ_d of \mathbb{R}^d . For any function $w : L \rightarrow \mathbb{R}$ and distribution $f \in \Delta_d$, the expected value of w given f is then the scalar product of w with f , seen as vectors of \mathbb{R}^d :

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

where \cdot denotes the scalar product of w and f in \mathbb{R}^d .

Suppose the random vectors Y and X have distributions g and f , respectively, and the former distribution dominates the latter according to some ordering " \succeq ". We will use the phrases " Y dominates X according to the ordering \succeq " and " g dominates f according to the ordering \succeq " interchangeably, and we denote the former by $Y \succeq X$ and the latter by $g \succeq f$.

Most (though not all) of the orderings we consider in this paper are what we term "difference-based orderings", in that whether or not two distributions g and f satisfy the ordering depends only on $\delta \equiv g - f$, the difference between their distributions. We will frequently exploit this convenient property of difference-based orderings.⁹

2.2 Greater Weak Association

We begin with an ordering of greater interdependence that is inspired by the dependence concept of association. Esary, Proschan, and Walkup (1967) defined association as follows:

DEFINITION 1 (Association) *A random vector Y with support L is **associated** if for all nondecreasing functions $r, s : L \rightarrow \mathbb{R}$, $\text{Cov}(r(Y), s(Y)) \geq 0$.*

so now f is a mean-preserving spread of g , and hence the ranking of riskiness is reversed.

⁹Some of the orderings of interdependence considered in this paper can be modified so they are responsive not only to the interdependence of the elements of a random vector but also to the levels of the elements. For clarity of focus, we will not explicitly consider these modifications of the orderings or the relations between them.

Note that if one tried to define a concept of negative association by reversing the inequality in the definition above, this concept would be uninteresting: Y could only be negatively associated in this strong sense if it were constant (to see this, consider functions $r = s$). This is one motivation for studying a less stringent version of association, as defined by Burton, Dabrowski, and Dehling (1986). In contrast to "association", which allows the functions r and s both to depend on the entire vector Y , "weak association" restricts them to depend on *disjoint* components of Y .

DEFINITION 2 (Weak Association) *A random vector Y with support $L = \times_{i=1}^n L_i$ is **weakly associated** if for any pair (A, B) of disjoint subsets of $\{1, \dots, n\}$ and nondecreasing functions $r : \times_{i \in A} L_i \rightarrow \mathbb{R}$ and $s : \times_{j \in B} L_j \rightarrow \mathbb{R}$, $\text{Cov}(r(Y_i, i \in A), s(Y_j, j \in B)) \geq 0$.*

It is clear from the definition that association is a stronger concept than weak association. Hu, Müller, and Scarsini (2004) have shown that it is strictly stronger, even for two dimensions. A meaningful concept of negative association (Joag-Dev and Proschan (1983)) is defined by reversing the inequality in the definition of weak association.

DEFINITION 3 (Negative Association) *A random vector Y with support $L = \times_{i=1}^n L_i$ is **negatively associated** if for any pair (A, B) of disjoint subsets of $\{1, \dots, n\}$ and nondecreasing functions $r : \times_{i \in A} L_i \rightarrow \mathbb{R}$ and $s : \times_{j \in B} L_j \rightarrow \mathbb{R}$, $\text{Cov}(r(Y_i, i \in A), s(Y_j, j \in B)) \leq 0$.*

We can now define an interdependence ordering corresponding to weak association and negative association as follows.

DEFINITION 4 (Greater Weak Association) *Y displays **greater weak association** than X , denoted $Y \succeq_{GWA} X$, if they have identical univariate marginal distributions and for all disjoint subsets A, B of $\{1, \dots, n\}$ and nondecreasing functions $r : \times_{i \in A} L_i \rightarrow \mathbb{R}$ and $s : \times_{j \in B} L_j \rightarrow \mathbb{R}$,*

$$\text{Cov}(r(Y_i, i \in A), s(Y_j, j \in B)) \geq \text{Cov}(r(X_i, i \in A), s(X_j, j \in B)).$$

In the insurance literature, this ordering has been (somewhat confusingly) termed the "correlation order" | see Denuit et al. (2005). Note that the definition of greater weak association assumes that the random vectors being compared have identical univariate marginals; this is not an implication of the condition of greater covariance per se.

In the next section, we show that for $n > 3$, greater weak association is not a difference-based ordering, though Theorem 1 in Section 3 implies that for $n = 2$, it is difference-based. To verify the claim for $n = 2$ directly, note that in this case, the only non-trivial partition of $\{1, 2\}$ into disjoint

subsets is $\{1\}, \{2\}$. Then the definition's requirement that Y and X have identical univariate marginals ensures that $Cov(r(Y_1), s(Y_2)) - Cov(r(X_1), s(X_2)) = E[r(Y_1)s(Y_2)] - E[r(X_1)s(X_2)]$, and the sign of the right-hand side depends only on the difference between the distributions of Y and X . That greater weak association is invariant to monotonic coordinate relabelings is apparent from the definition.

The greater weak association ordering has the desirable feature that (Y_1, \dots, Y_n) are weakly associated if and only if Y displays greater weak association than its "independent counterpart", defined as the random vector X such that (X_1, \dots, X_n) are independent and, for each i , X_i and Y_i have the same distribution.¹⁰ It might seem tempting to define Y as displaying "greater association" than X if Y and X have identical marginals and for all non-decreasing functions r and s defined on L , $Cov(r(Y), s(Y)) \geq Cov(r(X), s(X))$. However, such a definition would have the unappealing consequence that an associated random vector Y would not necessarily display greater association than its independent counterpart.¹¹

2.3 The Supermodular Ordering

For any $z, v \in L$, denote by $z \wedge v$ the component-wise minimum (or "meet") of z and v , i.e., the element of L such that $(z \wedge v)_i = \min\{z_i, v_i\} \in L_i$ for all $i \in N$. Let $z \vee v$ similarly denote the component-wise maximum (or "join") of z, v . A function w is said to be *supermodular* (on L) if $w(z \wedge v) + w(z \vee v) \geq w(z) + w(v)$ for all $z, v \in L$. Supermodular functions are characterized by the following property (see Topkis, 1978):

$$w \in \mathcal{S} \iff w(z + e_i + e_j) + w(z) \geq w(z + e_i) + w(z + e_j) \quad (2)$$

for all $i \neq j$ and z such that $z + e_i + e_j$ is well-defined (i.e., such that z_i is not the upper bound of L_i and z_j is not the upper bound of L_j).¹²

DEFINITION 5 (Supermodular Ordering) *Let the random vectors Y and X have distributions g and f , respectively. The distribution g dominates the distribution f according to the **supermodular ordering**, written $g \succeq_{SPM} f$, if and only if $E[w|g] \geq E[w|f]$ for all supermodular functions w .*

¹⁰Similarly, (Y_1, \dots, Y_n) are negatively associated if and only if the independent counterpart of Y displays greater weak association than Y .

¹¹For the support $\{0, 1, 2\}^2$, let $Pr(Y_1 = Y_2 = 0) = Pr(Y_1 = Y_2 = 1) = Pr(Y_1 = Y_2 = 2) = \frac{1}{4}$ and $Pr(Y_1 = 0, Y_2 = 2) = Pr(Y_1 = 2, Y_2 = 0) = \frac{1}{8}$. Then it can be checked that Y is associated, but for the increasing functions $r(Y_1, Y_2) = I_{\{Y_1 \geq 1, Y_2 \geq 1\}}$ and $s(Y_1, Y_2) = I_{\{Y_1 = 2 \text{ or } Y_2 = 2\}}$, $Cov(r, s)$ is strictly smaller for Y than for its independent counterpart.

¹²For functions w defined on \mathbb{R}^n and twice differentiable, an equivalent characterization is: w is supermodular if and only if $\frac{\partial^2 w}{\partial z_i \partial z_j} \geq 0$ and twice

It is clear from the definition that the supermodular ordering is a difference-based ordering.

To see most clearly the appeal of the supermodular ordering as an ordering of greater interdependence, consider two distributions g and f such that, for some $z \in L$ such that $z + e_i + e_j$ is well-defined, the difference $\delta \equiv g - f$ satisfies

$$\delta(z) = \delta(z + e_i + e_j) = -\delta(z + e_i) = -\delta(z + e_j) = \alpha \quad (3)$$

for some $\alpha > 0$, and such that $\delta(v) = 0$ for all other nodes v of L . In such a case, we say the distribution g is obtained from f by an *elementary transformation* (ET) of size α on L which leaves unchanged the probability of all nodes other than z , $z + e_i$, $z + e_j$, and $z + e_i + e_j$ and which raises the probability of nodes z and $z + e_i + e_j$ by the common amount α , while reducing the probability of nodes $z + e_i$ and $z + e_j$ by the same amount. Intuitively, such ET's increase the degree of interdependence of a multivariate distribution, as for some pair of components i and j , they make jointly high and jointly low realizations more likely, while making realizations where one component is high and the other low less likely. Furthermore, they raise interdependence without altering the marginal distribution of any component. From (2), a function w is supermodular if and only if $w \cdot \delta \geq 0$ for any δ of the form (3). Hence the class of supermodular functions is precisely the class for which the expectation is raised by any ET as defined in (3).

Meyer and Strulovici (2010) use duality methods to characterize the supermodular ordering and develop several constructive methods to implement this characterization. In their characterization, the elementary transformations defined above play a similar role to that of mean-preserving spreads in Rothschild and Stiglitz (1970) and Pigou-Dalton transfers in Atkinson (1970) and Dasgupta, Sen, and Starrett (1973).¹³

A necessary condition for $g \succeq_{SPM} f$ is that g and f have identical univariate marginal distributions. To see this, note that for any dimension $i \in \{1, \dots, n\}$ and any $k \in L_i$, the functions $\bar{w}(x) = I_{\{x_i \geq k\}}$ and $\underline{w}(x) = I_{\{x_i < k\}}$ are both supermodular. Therefore $g \succeq_{SPM} f$ implies that, for all $i \in \{1, \dots, n\}$

¹³Elementary transformations of bivariate distributions were also used by Epstein and Tanny (1980) and Tchen (1980) to prove the equivalence, in two dimensions, of the supermodular and lower-orthant orderings for distributions with identical marginals. Our definition of ET's in the text is more restrictive, as it requires that the four points affected by the ET be adjacent points in the lattice; this more restrictive definition allows a much simpler proof of the two-dimensional result and, more importantly, greatly facilitates the constructive methods for multivariate distributions developed in Meyer and Strulovici (2010).

and any $k \in L_i$,

$$\begin{aligned} 0 \leq E[\bar{w}|g] - E[\bar{w}|f] &= \sum_{x: x_i \geq k} g(x) - \sum_{x: x_i \geq k} f(x) \\ \text{and } 0 \leq E[\underline{w}|g] - E[\underline{w}|f] &= \sum_{x: x_i < k} g(x) - \sum_{x: x_i < k} f(x), \end{aligned} \quad (4)$$

and these inequalities together imply that g and f have identical univariate marginal distributions.

Since for any supermodular function $w(z_1, \dots, z_n)$, the function $w(r_1(z_1), \dots, r_n(z_n))$ is also supermodular whenever all the functions $\{r_i\}_{i=1}^n$ are nondecreasing, the supermodular ordering is invariant to monotonic relabelings of coordinates.

2.3.1 The Symmetric Supermodular Ordering

In many contexts, it is natural to assume that the supermodular objective functions being used to compare distributions are symmetric with respect to the components of the random vectors. For example, when the function w is an ex post welfare function defined on the realized utilities of n individuals, as in the assessment of ex post inequality under uncertainty (see Section 5), it is natural to assume that welfare is invariant to permutations of a given n -vector of utilities over the individuals. We now formally define the symmetric supermodular ordering.

Call a lattice $L = \times_{i=1}^n L_i$ symmetric if $L_i = L_j$ for all $i \neq j$. For a symmetric lattice, let the cardinality of L_i equal l , so the lattice has $d = l^n$ nodes. Let θ denote a real function on a symmetric lattice L , or equivalently a vector of \mathbb{R}^d . Depending on the context, θ can represent an objective function w or a probability distribution f . We will say that the function θ is *symmetric on L* if $\theta(z) = \theta(\sigma(z))$ for all $z \in L$ and for all permutations $\sigma(z)$ of z .

DEFINITION 6 (*Symmetric Supermodular Ordering*) *Let the random vectors Y and X have distributions g and f , respectively, on a symmetric lattice. The distribution g dominates the distribution f according to the **symmetric supermodular ordering**, written $g \succeq_{SSPM} f$, if and only if $E[w|g] \geq E[w|f]$ for all symmetric supermodular functions w .*

For an arbitrary (not necessarily symmetric) function θ , the *symmetrized version* of θ , θ^{symm} , is defined as follows: for any z ,

$$\theta^{symm}(z) = \frac{1}{n!} \sum_{\sigma \in \Sigma(n)} \theta(\sigma(z)), \quad (5)$$

where $\Sigma(n)$ is the set of all permutations of $\{1, \dots, n\}$. Importantly, if w is a supermodular function, then w^{symm} is supermodular. For a symmetric supermodular function w , let $\mathcal{W}^{symm}(w)$ denote the

set of supermodular functions \hat{w} on L such that the symmetrized version of \hat{w} is w , i.e., $\hat{w}^{symm} = w$. Note that $\{\mathcal{W}^{symm}(w)\}$ is a partition of the set of all supermodular functions on the symmetric lattice L .

We can now state the following useful result:

PROPOSITION 1 *Given a pair of distributions g, f defined on a symmetric lattice L , the following three statements are equivalent:*

- i) $g \succeq_{SSPM} f$;
- ii) $g^{symm} \succeq_{SSPM} f^{symm}$;
- iii) $g^{symm} \succeq_{SPM} f^{symm}$.

Proof. To show that i) \Rightarrow ii) \Rightarrow iii): If for all symmetric supermodular w , $w \cdot f \leq w \cdot g$, then for all symmetric supermodular w , $w \cdot f^{symm} \leq w \cdot g^{symm}$. This is ii). In turn, if for some symmetric supermodular w , $w \cdot f^{symm} \leq w \cdot g^{symm}$, then $\hat{w} \cdot f^{symm} \leq \hat{w} \cdot g^{symm}$ for all $\hat{w} \in \mathcal{W}^{symm}(w)$. Therefore, since $\{\mathcal{W}^{symm}(w)\}_w$ partitions the set of all supermodular functions on L , $w \cdot f^{symm} \leq w \cdot g^{symm}$ for all symmetric supermodular w implies that $\hat{w} \cdot f^{symm} \leq \hat{w} \cdot g^{symm}$ for all supermodular \hat{w} , which is iii).

To show that iii) \Rightarrow i): If for all supermodular w , $w \cdot f^{symm} \leq w \cdot g^{symm}$, then for all supermodular w , $w^{symm} \cdot f^{symm} \leq w^{symm} \cdot g^{symm}$. This is equivalent to $w^{symm} \cdot f^{symm} \leq w^{symm} \cdot g^{symm}$ for all symmetric supermodular w^{symm} . This in turn implies that for all symmetric supermodular w^{symm} , $w^{symm} \cdot f \leq w^{symm} \cdot g$. ■

Proposition 1 states that one can characterize the symmetric supermodular ordering in terms of the supermodular order applied to symmetric distributions. Furthermore, when attention is restricted to symmetric distributions, the supermodular order is equivalent to the symmetric supermodular one. Proposition 1 will be used frequently in what follows to simplify the analysis of the symmetric supermodular ordering by focusing on symmetric distributions.

For random vectors for which each component has a common binary support, say $\{0, 1\}$, so the lattice $L = \{0, 1\}^n$, the symmetric supermodular ordering has a very simple form. For a random vector Y with support $L = \{0, 1\}^n$, define $c(Y) = \sum_{i=1}^n I_{\{Y_i=1\}}$, the number of components of Y for which the realization takes the value 1. To state the result, we first recall the definition of the univariate convex ordering: For random variables Z and W with support $S \subseteq \mathbb{R}$, Z dominates V

according to the convex ordering, written $Z \succeq_{CX} V$, if $Ew(Z) \geq Ew(V)$ for all convex functions $w : S \rightarrow \mathbb{R}$. Since $w(z) = z$ and $w(z) = -z$ are both convex functions, $Z \succeq_{CX} V$ implies that $EZ = EV$. The convex ordering is equivalent to the ordering of greater riskiness studied by Rothschild and Stiglitz (1970).

PROPOSITION 2 *For random vectors Y and X distributed on $L = \{0, 1\}^n$, $Y \succeq_{SSPM} X$ if and only if $c(Y) \succeq_{CX} c(X)$.*

Proof. Any symmetric function w defined on $L = \{0, 1\}^n$ can be written as

$$w(X_1, \dots, X_n) = \phi(c(X_1, \dots, X_n)), \quad (6)$$

for some function $\phi : \{0, 1, \dots, n\} \rightarrow \mathbb{R}$. Furthermore, a symmetric function w on $\{0, 1\}^n$ is supermodular if and only if the function $\phi(\cdot)$ in (6) is convex. ■

In Sections 2.4 and 2.5, we propose two new orderings of greater interdependence.

2.4 The Convex-Modular Ordering

In many contexts, the objective functions that are used to evaluate the degree of interdependence of multivariate random variables have the form $w(z) = \phi(r_1(z_1) + \dots + r_n(z_n))$, where $\phi(\cdot)$ is convex and $\{r_i\}_{i=1}^n$ are nondecreasing. We will term such functions "convex-modular", as they take a convex transformation of a modular (i.e. additively separable) aggregate, $\sum_{i=1}^n r_i(z_i)$. It is easy to see that any convex-modular function is supermodular, and therefore the expectation of any convex-modular function is increased by any elementary transformation of the form defined in (3). The proof of Proposition 2 rested on the observation that any symmetric supermodular function w defined on $L = \{0, 1\}^n$ is convex-modular, with $\sum_{i=1}^n r_i(z_i) = \sum_{i=1}^n I_{\{z_i=1\}}$. Convex-modular functions arise naturally in an insurance context, where Z represents a vector of losses incurred by individuals $1, \dots, n$, all of whom are insured by a given insurer, and where the insurance contract of individual i obliges the insurer to pay compensation $r_i(Z_i)$, which would take the form $r_i(Z_i) = \min\{m_i, \max\{(1 - \beta_i)(Z_i - d_i), 0\}\}$ for a policy with a deductible d_i , a copayment rate β_i for the insured, and a compensation limit m_i . The total compensation paid out by the insurer is then $\sum_{i=1}^n r_i(Z_i)$, and the insurer is concerned with the riskiness of this total, so evaluates the cost of this payout using a convex objective function ϕ .¹⁴

This motivates us to define the following ordering:

¹⁴See Denuit et al (2005) for more details.

DEFINITION 7 (*Convex-Modular Ordering*) Let the random vectors Y and X have distributions g and f , respectively. The distribution g dominates the distribution f according to the **convex-modular ordering**, written $g \succeq_{CXMOD} f$, if and only if $E[w|g] \geq E[w|f]$ for all convex-modular functions w .

This definition is equivalent to the requirement that $E[w|g] \geq E[w|f]$ for all functions w that are nonnegative weighted sums of convex-modular functions. The convex-modular ordering is clearly difference-based and invariant to monotonic coordinate relabelings. It follows from the definition that $Y \succeq_{CXMOD} X$ if and only if, for all nondecreasing $\{r_i\}_{i=1}^n$, $\sum_{i=1}^n r_i(Y_i) \succeq_{CX} \sum_{i=1}^n r_i(X_i)$.

Since $\bar{w}(Z) = I_{\{Z_i \geq k\}}$ and $\underline{w}(Z) = -I_{\{Z_i \geq k\}}$ are both convex-modular functions, it follows, from the same logic as for the supermodular ordering, that $Y \succeq_{CXMOD} X$ implies that Y and X have identical marginals.

By analogy with what we did for the supermodular ordering, it is natural to define a symmetric counterpart of the convex-modular ordering. It is clear from our definition in (5) of the symmetrized version of a function, that if w is a convex-modular function, then w^{symm} is a nonnegative weighted sum of convex-modular functions. Let \mathcal{CM}^* denote the set of nonnegative weighted sums of convex-modular functions (where the dependence on a given L is implicit). It follows from the above observation that \mathcal{CM}^* is closed under symmetrization (although the set of convex-modular functions itself is not). As a consequence, we define the symmetric convex-modular ordering as follows:

DEFINITION 8 (*Symmetric Convex-Modular Ordering*) Let the random vectors Y and X have distributions g and f , respectively, on a symmetric lattice. The distribution g dominates the distribution f according to the **symmetric convex-modular ordering**, written $g \succeq_{SCXMOD} f$, if and only if $E[w|g] \geq E[w|f]$ for all symmetric functions $w \in \mathcal{CM}^*$.

As noted above, $Y \succeq_{CXMOD} X$ if and only if $E[w|g] \geq E[w|f]$ for all $w \in \mathcal{CM}^*$. Using this equivalence, it is then straightforward to adapt the proof of Proposition 1 to show:

PROPOSITION 3 Given a pair of distributions g, f defined on a symmetric lattice L , the following three statements are equivalent:

- i) $g \succeq_{SCXMOD} f$;
- ii) $g^{symm} \succeq_{SCXMOD} f^{symm}$;
- iii) $g^{symm} \succeq_{CXMOD} f^{symm}$.

2.5 The Dispersion Ordering

Another notion of greater interdependence in Y than in X reflects the idea that the distribution functions of the order statistics of Y should be "closer together" or less dispersed than the distribution functions of the order statistics of X . To understand the link between dispersion of order statistics and interdependence, consider the following distributions for (Y_1, Y_2) and (X_1, X_2) on support $\{0, 1\}^2$, where each of Y_1, X_1, Y_2 , and X_2 takes each of the two values 0 and 1 with probability $\frac{1}{2}$. Let Y_1 and Y_2 be perfectly positively dependent: each of the realizations $(0, 0)$ and $(1, 1)$ occurs with probability $\frac{1}{2}$. Let X_1 and X_2 be perfectly negatively dependent: each of the realizations $(0, 1)$ and $(1, 0)$ occurs with probability $\frac{1}{2}$. For the order statistics of Y , $\min\{Y_1, Y_2\}$ and $\max\{Y_1, Y_2\}$, $Pr(\min\{Y_1, Y_2\} = 0) = Pr(\max\{Y_1, Y_2\} = 0) = \frac{1}{2}$, so the two order statistics have the same distribution. On the other hand, for those of X , $Pr(\min\{X_1, X_2\} = 0) = 1$ while $Pr(\max\{X_1, X_2\} = 0) = 0$, so the two order statistics have distributions as different as possible in this context. The qualitative lesson of this example is that for the more positively dependent random vector Y , the distribution functions of the order statistics are more similar (less dispersed) than for X .

The majorization ordering of vectors can be used to formalize the notion of lower dispersion. A vector a is said to be majorized by a vector b , written $a \prec b$, if i) the components of the vectors have the same total sum, and ii) for all k , the sum of the k largest entries of a is weakly smaller than the sum of the k largest entries of b (see Hardy, Littlewood, and Polya (1952)).

Let $Y_{(j)}$ denote the j^{th} order statistic of Y , i.e. the j^{th} smallest value from (Y_1, \dots, Y_n) , and define $X_{(j)}$ similarly. Let $F_{Y_{(j)}}$ and $F_{X_{(j)}}$ denote the c.d.f.'s of these order statistics. For random vectors with *symmetric* distributions, Shaked and Tong (1985) suggested the following dependence ordering.

DEFINITION 9 (*Symmetric Dispersion Ordering*) For random vectors Y, X with symmetric distributions on a symmetric lattice L , the distribution of Y dominates that of X according to the **symmetric dispersion ordering**, written $Y \succeq_{SDISP} X$, if the distribution functions of the order statistics of Y are less dispersed than the distribution functions of the order statistics of X , that is,

$$(F_{Y_{(1)}}(b_0), \dots, F_{Y_{(n)}}(b_0)) \prec (F_{X_{(1)}}(b_0), \dots, F_{X_{(n)}}(b_0)) \quad \forall b = (b_0, \dots, b_0) \in L. \quad (7)$$

We now reformulate this dependence ordering in a manner which suggests a new ordering which is both stronger and naturally applicable to asymmetric distributions. Each vector $b = (b_0, \dots, b_0) \in L$ can be seen as generating a (componentwise) binary coarsening of the support of the random

vector Y from L to $\{0, 1\}^n$ and a corresponding coarsened version of Y , Y^b , such that $Y_i^b = 0$ if $Y_i \leq b_0$ and $Y_i^b = 1$ if $Y_i > b_0$. Let $Y_{(j)}^b$ denote the j^{th} order statistic of Y^b and $F_{Y_{(j)}^b}$ the distribution function of this order statistic. Then a condition equivalent to (7) is

$$(F_{Y_{(1)}^b}(0), \dots, F_{Y_{(n)}^b}(0)) \prec (F_{X_{(1)}^b}(0), \dots, F_{X_{(n)}^b}(0)) \quad \forall b = (b_0, \dots, b_0) \in L. \quad (8)$$

For asymmetric multivariate distributions, there is no particular reason to confine attention to binary coarsenings generated by *vectors with equal components* $b = (b_0, \dots, b_0)$. Given a lattice $L = \times_{i=1}^n L_i$ with $L_i = \{0, 1, \dots, l_i - 1\}$, define $L_i = \{-1, 0, 1, \dots, l_i - 1\}$ and $L = \times_{i=1}^n L_i$. We propose a new dependence ordering, which strengthens condition (8) by requiring that it hold for every vector $s = (s_1, s_2, \dots, s_n) \in L$. For a random vector Y , define its coarsening corresponding to the vector $s \in L$, Y^s , by $Y_i^s = 0$ if $Y_i \leq s_i$ and $Y_i^s = 1$ if $Y_i > s_i$. Let $Y_{(j)}^s$ denote the j^{th} order statistic of Y^s and $F_{Y_{(j)}^s}$ the distribution function of this order statistic.¹⁵

DEFINITION 10 (Dispersion Ordering) For random vectors Y, X distributed on L , consider the set of all binary coarsenings of Y and X , Y^s and X^s , respectively, corresponding to some $s \in L$. The distribution of Y dominates that of X according to the **dispersion ordering**, written $Y \succeq_{DISP} X$, if for all $s \in L$, the distribution functions of the order statistics of Y^s are less dispersed than the distribution functions of the order statistics of X^s , that is,

$$(F_{Y_{(1)}^s}(0), \dots, F_{Y_{(n)}^s}(0)) \prec (F_{X_{(1)}^s}(0), \dots, F_{X_{(n)}^s}(0)) \quad \forall s \in L. \quad (9)$$

The following proposition provides insight into the dispersion ordering by presenting some equivalent formulations. To state it, we define, for any $s \in L$,

$$c^s(Y^s) = \sum_{i=1}^n I_{\{Y_i^s=1\}} = \sum_{i=1}^n I_{\{Y_i > s_i\}},$$

which counts the number of components of Y^s that equal 1, or equivalently, the number of components of Y that strictly exceed the corresponding component of s .

PROPOSITION 4 *The following three conditions are equivalent:*

¹⁵Considering coarsenings corresponding to every s in the *extended* lattice \bar{L} allows us to include coarsened variables Y^s for which $Y_i^s = 1$ with probability 1 for some component i . This makes the treatment of such coarsened variables analogous to the treatment of coarsened variables Y^s for which $Y_i^s = 0$ with probability 1 for some i . Since the definition of the SDISP ordering requires (7), or equivalently (8), to hold only for vectors b with equal components, requiring it to hold for all $b = (b_0, \dots, b_0) \in \bar{L}$ would leave the definition effectively unchanged, since it would add only the trivial condition corresponding to $b = (-1, \dots, -1)$, for which both vectors of distribution functions in (7) are $(0, \dots, 0)$.

i) For Y, X with support L , $Y \succeq_{DISP} X$;

ii) For all $s \in L$, $Y^s \succeq_{SSPM} X^s$.

iii) For all $s \in L$, $c^s(Y^s) \succeq_{CX} c^s(X^s)$.

Proof. Part a): Since for any s , Y^s and X^s have support $\{0, 1\}^n$, the equivalence of ii) and iii) follows from Proposition 2 in Section 2.2.1. To show that ii) implies i), rewrite (9) as

$$E \left[\sum_{j=1}^k I_{\{Y_{(j)}^s=0\}} \right] \leq E \left[\sum_{j=1}^k I_{\{X_{(j)}^s=0\}} \right] \quad \forall k \in \{1, \dots, n\}, \forall s \in L, \quad (10)$$

with equality required for $k = n$. Now observe that

$$\begin{aligned} \sum_{j=1}^k I_{\{Y_{(j)}^s=0\}} &= \sum_{j=1}^k I_{\{c^s(Y^s) < n-(j-1)\}} = \sum_{j=1}^k (1 - I_{\{c^s(Y^s) \geq n-(j-1)\}}) \\ &= k - \max\{c^s(Y^s) - (n - k), 0\}. \end{aligned} \quad (11)$$

Since for any $s \in L$, $c^s(Y^s)$ is a symmetric function of Y^s , and $\max\{z - a, 0\}$ is convex in z for any $a \in \mathbb{R}$, $\max\{c^s(Y^s) - (n - k), 0\}$ is a symmetric supermodular function of Y^s for all $k \in \{1, \dots, n\}$.¹⁶ Therefore, ii) implies that (10) holds for all $k \in \{1, \dots, n\}$. Setting $k = n$ in () gives

$$\sum_{j=1}^n I_{\{Y_{(j)}^s=0\}} = n - c^s(Y^s),$$

and since both $c^s(Y^s)$ and $-c^s(Y^s)$ are symmetric supermodular functions of Y^s , ii) implies that for $k = n$, (10) holds with equality, as required.

To show that i) implies iii), first note that for random variables Z and V with support $\{0, 1, \dots, n\}$, Z dominates V according to the convex ordering if and only if $E(Z) = E(V)$ and $E[\max\{Z - a, 0\}] \geq E[\max\{V - a, 0\}]$ for all $a \in \{1, \dots, n - 1\}$.¹⁷ Given (2.5), the equality in (10) for $k = n$ implies that $E(c^s(Y^s)) = E(c^s(X^s))$, and the inequality in (10) for $k \in \{1, \dots, n - 1\}$ implies that $E[\max\{c^s(Y^s) - a, 0\}] \geq E[\max\{c^s(X^s) - a, 0\}]$ for all $a \in \{0, 1, \dots, n - 1\}$. ■

It is apparent from Proposition 4 that the dispersion ordering is a difference-based ordering and is invariant to monotonic relabelings of coordinates. If $Y \succeq_{DISP} X$, then Y and X have identical univariate marginals. To see this, first take $s \in L$ to be a vector with i^{th} component equal to $a \in L_i$ and all other components equal to -1. Then for $k = 1$, (10) can be rewritten as

$$E[I_{\{c^s(Y^s) < n\}}] \leq E[I_{\{c^s(X^s) < n\}}] \Leftrightarrow Pr(Y_i \leq a) \leq Pr(X_i \leq a).$$

¹⁶ $\max\{c^s(Y^s) - (n - k), 0\}$ is also a symmetric convex-modular function of Y^s .

¹⁷See, for example, Jewitt (1987) and Hardy, Littlewood, and Polya (1929).

Now take s to be a vector with i^{th} component equal to $a \in L_i$ and all other components equal to $l_i - 1$. Then it follows from the equality in (10) for $k = n$ and the inequality for $k = n - 1$ that

$$E[I_{\{Y_{(n)}^s=0\}}] \geq E[I_{\{X_{(n)}^s=0\}}] \Leftrightarrow Pr(Y_i \leq a) \geq Pr(X_i \leq a).$$

It follows, therefore, that for all i and for all $a \in L_i$, $Pr(Y_i \leq a) = Pr(X_i \leq a)$, and hence Y and X have identical marginals.

2.6 The Concordance Ordering

Another intuitively appealing notion of greater interdependence, the concordance ordering, has been formalized for multivariate distributions by Joe (1990).

DEFINITION 11 (*Concordance Ordering*) *Let the random vectors Y and X have distributions g and f , respectively. The distribution g dominates the distribution f according to the **concordance ordering**, written $g \succeq_{CONC} f$, if and only if*

$$Pr(Y \geq z) \geq Pr(X \geq z) \quad \text{and} \quad Pr(Y \leq z) \geq Pr(X \leq z) \quad \forall z \in L.$$

For any node in the support, the concordance ordering requires that the components of Y be more likely to be all higher than at that node, relative to those of X , and also more likely to be all lower than at that node, relative to those of X .

It is easy to see from the definition that the concordance ordering, like the supermodular ordering, is a difference-based ordering and that it satisfies invariance to monotonic relabelings of coordinates, as defined in (1). It is also straightforward to confirm, as is well known, that $Y \succeq_{CONC} X$ implies that Y and X have identical univariate marginals.¹⁸

3 Relationships among the Orderings

We are now in a position to establish the relationships among the orderings of interdependence defined in Section 2.

¹⁸First take $z \in L$ to be a vector with i^{th} component equal to a and all other components 0. Then $Pr(Y \geq z) \geq Pr(X \geq z)$ becomes $Pr(Y_i \geq a) \geq Pr(X_i \geq a)$. Similarly, taking z to be a vector with i^{th} component equal to $a - 1$ and all other components equal to $l_i - 1$, $Pr(Y \leq z) \geq Pr(X \leq z)$ becomes $Pr(Y_i \leq a - 1) \geq Pr(X_i \leq a - 1)$. Hence, for all i and $a \in L_i$, $Pr(Y_i \geq a) = Pr(X_i \geq a)$, so Y and X have identical marginals.

THEOREM 1 (ORDERINGS: TWO DIMENSIONS) *For two dimensions, the following orders are equivalent: greater weak association, supermodular ordering, convex-modular ordering, dispersion ordering, and concordance ordering.*

The equivalence for $n = 2$ between greater weak association, the supermodular ordering, and the concordance ordering is well known (see Meyer (1990, Prop. 2) and Müller and Stoyan (2002, Theorem 3.8.2) for references). That the convex-modular and dispersion orderings are also equivalent then follows from this, combined with part a) of Theorem 2 below.

THEOREM 2 (ORDERINGS: THREE OR MORE DIMENSIONS) a) *For $n \geq 3$, greater weak association is strictly stronger than the supermodular ordering, which is strictly stronger than the convex-modular ordering, which is strictly stronger than the dispersion ordering, which is at least as strong as the concordance ordering.*

b) *For $n = 3$, the dispersion ordering is equivalent to the concordance ordering, whereas for $n > 3$, the dispersion ordering is strictly stronger than the concordance ordering.*

Proof. The proof that $Y \succeq_{GWA} X$ implies $Y \succeq_{SPM} X$ is in Appendix A. To prove that this implication is strict, consider Example 1:

EXAMPLE 1: Let $L = \{0, 1\}^3$ and let X and Y have symmetric distributions f and g , respectively. Let $f(x) = \frac{1}{3}$ if $\sum_{i=1}^3 x_i = 1$ and $f(x) = 0$ otherwise. Let $g(y)$ take the values $\frac{17}{66}, \frac{11}{66}, \frac{5}{66}, \frac{1}{66}$ when $\sum_{i=1}^3 y_i$ takes the values 0, 1, 2, 3, respectively. For these symmetric distributions, Propositions 1 and 2 together imply that $g \succeq_{SPM} f$ on $L = \{0, 1\}^3$ if and only if $c(Y) \succeq_{CX} c(X)$ on $\{0, 1, 2, 3\}$. It is easily checked that $Ec(Y) = Ec(X)$, $Pr(c(Y) = 3) \geq Pr(c(X) = 3)$, and $Pr(c(Y) = 0) \geq Pr(c(X) = 0)$, establishing that $c(Y) \succeq_{CX} c(X)$ and hence that $g \succeq_{SPM} f$. However, for $r(z_1) = I_{\{z_1=1\}}$ and $s(z_2, z_3) = I_{\{z_2=z_3=1\}}$, $Cov(r(Y_1), s(Y_2, Y_3)) = \frac{1}{66} - \frac{1}{3} \cdot \frac{1}{11} < 0$, while $Cov(r(X_1), s(X_2, X_3)) = 0$, so $g \succeq_{GWA} f$ does not hold.

Since every convex-modular function is supermodular, the supermodular ordering implies the convex-modular ordering. We show that the implication is strict by providing, in Appendix B, an example of a supermodular function on $L = \{0, 1, 2\}^3$ which cannot be written as a nonnegative weighted sum of convex-modular functions.

The fact that the convex-modular ordering implies the dispersion ordering follows from the proof of Proposition 4 and the facts that the functions $\max\{c^s(Y) - (n - k), 0\}$, for $k \in \{1, \dots, n\}$, and $-c^s(Y)$ are both convex-modular.

To show that the dispersion ordering implies the concordance ordering, observe that, for $k = 1$, the left-hand side of (10) can be rewritten as

$$E[I_{\{Y_{(1)}^s=0\}}] = 1 - E[I_{\{c^s(Y)=n\}}] = 1 - Pr(Y_i > s_i \forall i).$$

Therefore, (10) implies that for all $s \in L$, $Pr(Y > s) \geq Pr(X > s)$. Similarly, it follows from the equality in (10) for $k = n$ and the inequality for $k = n - 1$ that, for all $s \in L$,

$$E[I_{\{Y_{(n)}^s=0\}}] \geq E[I_{\{X_{(n)}^s=0\}}],$$

which implies that for all $s \in L$, $Pr(Y \leq s) \geq Pr(X \leq s)$. Hence $Y \succeq_{CONC} X$.

Now we show that for $n = 3$, $Y \succeq_{CONC} X$ implies $Y \succeq_{DISP} X$. First, for any $s \in L$, showing that (10) holds with equality for $k = n = 3$ is equivalent to showing that

$$\sum_{i=1}^3 I_{\{Y_i \leq s_i\}} = \sum_{i=1}^3 I_{\{X_i \leq s_i\}},$$

and this is true, since $Y \succeq_{CONC} X$ implies that Y and X have identical marginal distributions. Given the equality in (10) for $k = 3$, it remains to show that for all $s \in L$,

$$E[I_{\{Y_{(1)}^s=0\}}] \leq E[I_{\{X_{(1)}^s=0\}}] \quad \text{and} \quad E[I_{\{Y_{(3)}^s=0\}}] \geq E[I_{\{X_{(3)}^s=0\}}].$$

The first of these inequalities is equivalent to $Pr(Y > s) \geq Pr(X > s)$ and the second to $Pr(Y \leq s) \geq Pr(X \leq s)$, as shown in the proof that $Y \succeq_{DISP} X$ implies $Y \succeq_{CONC} X$, and both inequalities therefore follow from $Y \succeq_{CONC} X$.

Example 2 shows that the convex-modular ordering is strictly stronger than the dispersion ordering:

EXAMPLE 2: Let $L = \{0, 1, 2\} \times \{0, 1\} \times \{0, 1\}$. Since both the convex-modular and the dispersion ordering are difference-based orderings, it is sufficient to specify $\delta \equiv g - f$, the difference between the distributions of Y and X . Let $\delta(z_1, z_2, z_3) = \epsilon > 0$ if $\sum_{i=1}^3 z_i$ is even and $\delta(z_1, z_2, z_3) = -\epsilon < 0$ if $\sum_{i=1}^3 z_i$ is odd. It is easily checked that $Y \succeq_{CONC} X$ and, since $n = 3$, it follows that $Y \succeq_{DISP} X$. However, for the convex-modular function $w(z) = \max\{(\sum_{i=1}^3 z_i) - 2, 0\}$, we have $w \cdot (g - f) = -\epsilon < 0$, so $Y \not\succeq_{XMOD} X$ does not hold.

To show that for $n > 3$, the dispersion ordering is strictly stronger than the concordance ordering, consider Example 3:

EXAMPLE 3: Let $L = \{0, 1\}^4$, and let $\delta(z_1, z_2, z_3, z_4) \equiv g - f = \epsilon > 0$ if $\sum_{i=1}^4 z_i$ is even and $\delta(z_1, z_2, z_3, z_4) = -\epsilon < 0$ if $\sum_{i=1}^4 z_i$ is odd. Again, it is easily checked that $Y \succeq_{CONC} X$. Let

$s = (0, 0, 0, 0)$ and consider the convex function of $c^s(z)$ defined by $w = \max\{c^s(z) - 2, 0\}$. We have $w \cdot (g - f) = -2\epsilon < 0$, so by Proposition 4, $Y \succeq_{DISP} X$ does not hold. ■

Example 1, which we used in the proof to show that, while $Y \succeq_{GWA} X$ implies $Y \succeq_{SPM} X$, the converse is not true, can be extended to show that greater weak association is not a difference-based order.

EXAMPLE 1': Replace f and g from Example 1 with f' and g' , respectively, where $f'(x) = \frac{1}{2}$ if $\sum_{i=1}^3 x_i = 0$, $f'(x) = \frac{1}{6}$ if $\sum_{i=1}^3 x_i = 1$, and $f'(x) = 0$ otherwise, and $g'(y)$ takes the values $\frac{50}{66}, 0, \frac{5}{66}, \frac{1}{66}$ when $\sum_{i=1}^3 y_i$ takes the values 0, 1, 2, 3, respectively. By construction, $g' - f' = g - f$, and we've shown above that $g \succeq_{GWA} f$ does not hold. Yet we can show (details are in Appendix C) that $g' \succeq_{GWA} f'$. Hence for $n \geq 3$, greater weak association is not a difference-based order.

4 Binary Random Variables

In many economic contexts, the random variables whose interdependence is to be assessed are binary. Theoretical models often focus on binary action spaces or binary outcome spaces for tractability. For example, Calvo-Armengol and Jackson's (2004) study of the effects of social networks on interdependence in individuals' employment outcomes suppressed wage variation among employed workers and focused only on whether workers were employed or unemployed. Experimental studies often focus on binary choice spaces to simplify the subjects' decision problems as well as to simplify the data analysis. For example, Choi, Gale, and Kariv (2005), in their experimental study of the effect of network structure on social learning and the resulting interdependence among agents' decisions, focused on a decision environment with only two states of the world, two signals, and two possible actions. In empirical work, for example on multidimensional inequality, binary classifications, such as whether or not income is below the poverty line or whether or not an individual is literate, are often inevitable features of the data.

Binary random variables, besides being common, also help to illuminate the structure of and relationships among the interdependence orderings. This section studies a variety of special cases with binary random variables. Our aims are to highlight i) equivalences among the dependence orderings that arise in these special cases and ii) easily checkable and easily interpretable necessary and sufficient conditions for the orderings to hold.

4.1 Symmetric Distributions or Symmetric Objective Functions

The following result is valid for any number n of dimensions.

PROPOSITION 5 *a) For random vectors Y and X with symmetric distributions on $L = \{0, 1\}^n$, the following conditions are equivalent:*

i) $Y \succeq_{SPM} X$;

ii) $Y \succeq_{CXMOD} X$;

iii) $Y \succeq_{DISP} X$;

iv) $Y \succeq_{SDISP} X$;

v) $c(Y) \succeq_{CX} c(X)$.

b) For random vectors Y and X distributed on $L = \{0, 1\}^n$, the following conditions are equivalent:

i) $Y \succeq_{SSPM} X$;

ii) $Y \succeq_{\$CXMOD}$

that Y and X have identical marginals, whereas as shown in Section 2.5, identical marginals is a necessary condition for $Y \succeq_{DISP} X$. Consequently, on $L = \{0, 1\}^n$, the dispersion ordering is strictly stronger than the SSPM ordering and the convex ordering of $c(Y)$ over $c(X)$.¹⁹

4.2 Three Dimensions

PROPOSITION 6 *For random vectors Y and X distributed on $L = \{0, 1\}^3$, the following conditions are equivalent:*

- i) $Y \succeq_{SPM} X$;
- ii) $Y \succeq_{CXMOD} X$;
- iii) $Y \succeq_{DISP} X$;
- iv) $Y \succeq_{CONC} X$;
- v) $c(Y) \succeq_{CX} c(X)$; Y and X have identical marginals; and for all $i \neq j$, $Pr(Y_i = Y_j) \geq Pr(X_i = X_j)$.

Proof. All of the orderings in the proposition are difference-based orderings, so it is sufficient to work with $\delta \equiv g - f$. For $L = \{0, 1\}^3$, δ is represented in Figure 1a. Each of conditions i)-v) implies that Y and X have identical marginals, and this in turn implies that once the values of $\delta(1, 1, 1) \equiv a$, $\delta(0, 1, 1) \equiv b_1$, $\delta(1, 0, 1) \equiv b_2$, and $\delta(1, 1, 0) \equiv b_3$ are specified, the remaining values are determined. Given identical marginals, $Y \succeq_{CONC} X$ if and only if

$$a \geq 0, \quad a + b_k \geq 0 \quad \forall k \in \{1, 2, 3\}, \quad \text{and} \quad 2a + \sum_{i=1}^3 b_i \geq 0. \quad (12)$$

In (12), the first and third inequalities are equivalent, respectively, to

$$Pr(c(Y) = 3) \geq Pr(c(X) = 3) \quad \text{and} \quad Pr(c(Y) = 0) \geq Pr(c(X) = 0).$$

The inequality $a + b_k \geq 0$ is equivalent to $Pr(Y_i = Y_j = 1) \geq Pr(X_i = X_j = 1)$, which, given identical marginals, is equivalent to $Pr(Y_i = Y_j) \geq Pr(X_i = X_j)$. Hence conditions iv) and v) are equivalent. We now provide a simple constructive proof that:

$$\text{For } L = \{0, 1\}^3, \quad Y \succeq_{CONC} X \quad \text{implies} \quad Y \succeq_{SPM} X. \quad (13)$$

¹⁹When Y and X have *symmetric* distributions on $\{0, 1\}^n$, though, then $c(Y) \succeq_{CX} c(X)$, which implies $Ec(Y) = Ec(X)$, furthermore implies, given the symmetric distributions, that Y and X have identical marginals, so under the hypotheses of part a), $Y \succeq_{DISP} X$ and $c(Y) \succeq_{CX} c(X)$ are equivalent.

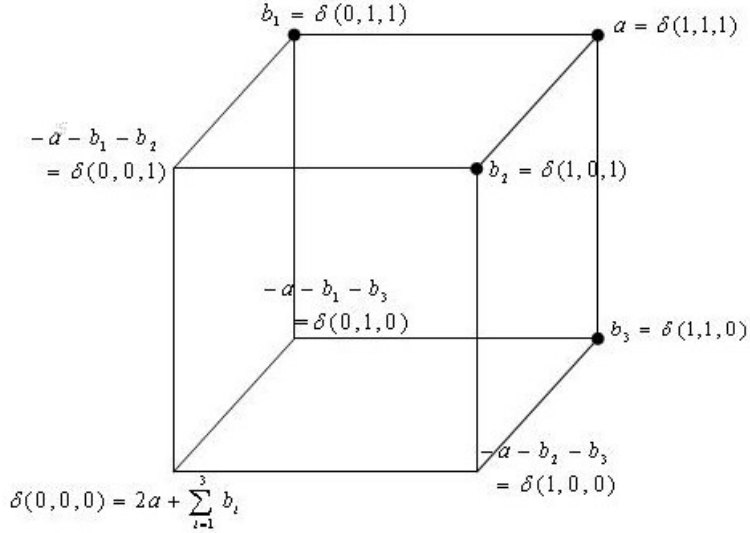


Figure 1a: $\delta = g - f$ for $L = \{0, 1\}^3$

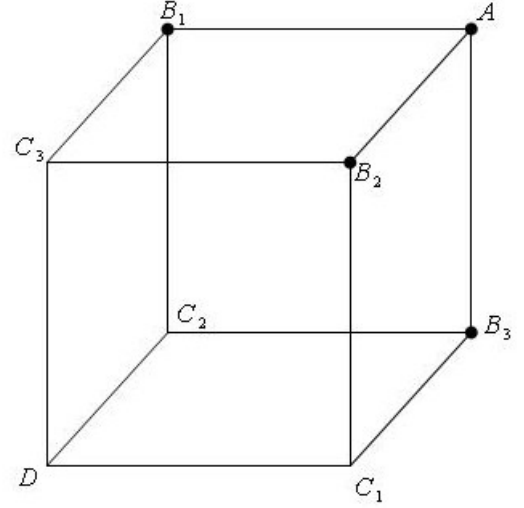


Figure 1b: Labels for the nodes of $L = \{0, 1\}^3$

This, in conjunction with Theorem 2, will complete the proof.

Our constructive proof of (13) decomposes δ into 6 elementary transformations (ET's) of the form defined in (3): for each of the 6 faces of the cube $L = \{0, 1\}^3$, there is one ET involving the 4 nodes on that face.²⁰ Using the labels for the nodes in Figure 1b, let the ET involving nodes A , B_i , B_j , and C_k have size $\bar{\alpha}_{ij}$, and the ET involving nodes B_k , C_i , C_j , and D have size $\underline{\alpha}_{ij}$. It is easily checked that these 6 ET's sum to δ if and only if

$$a = \bar{\alpha}_{12} + \bar{\alpha}_{13} + \bar{\alpha}_{23} \quad \text{and} \quad \forall i, j, k \in \{1, 2, 3\}, i \neq j \neq k, \quad a + b_k = \bar{\alpha}_{ij} + \underline{\alpha}_{ij}. \quad (14)$$

Now set

$$\bar{\alpha}_{ij} = a \left(\frac{a + b_k}{3a + \sum_{i=1}^3 b_i} \right) \quad \text{and} \quad \underline{\alpha}_{ij} = (2a + \sum_{i=1}^3 b_i) \left(\frac{a + b_k}{3a + \sum_{i=1}^3 b_i} \right). \quad (15)$$

It is apparent by inspection that the equations (14) are satisfied and that, if the inequalities (12) defining the concordance ordering here hold, then for all $i \neq j$, $\bar{\alpha}_{ij} \geq 0$ and $\underline{\alpha}_{ij} \geq 0$. Thus, $g \succeq_{CONC} f$ implies the existence of a sequence of nonnegative ET's that sum to $g - f$. Since each ET raises the expectation of any supermodular function, it follows that $g \succeq_{SPM} f$. ■

²⁰Hu, Xie, and Ruan (2005, pp. 188-9) proved (13) in a very indirect manner using the tool of “majorization with respect to weighted trees”.

4.3 Four Dimensions and “Top-to-Bottom” Symmetry

The equivalence demonstrated in Proposition 6 for the three-dimensional cube between the supermodular ordering and the concordance ordering breaks down if we increase either i) the number of dimensions, as shown by Example 3 in the proof of Theorem 2, or ii) the number of points in the support for any random variable, as shown by Example 2, used in the same proof. Nevertheless, some interesting equivalences do persist in higher dimensions, and can be demonstrated using a similar constructive method of proof.

Consider four-dimensional random vectors with support $L = \{0, 1\}^4$, and assume now that their distributions satisfy a symmetry condition we term “top-to-bottom symmetry”. We say that the distribution of a random vector Z satisfies *top-to-bottom symmetry* if for any $a \in \{0, 1\}^4$, $P(Z = a) = P(Z = 1 - a)$. Top-to-bottom symmetry arises naturally in a variety of settings. We give two examples: matching with frictions and social learning in networks.

In a matching context, suppose the four dimensions represent managers, supervisors, workers, and firms, and suppose that for each dimension, there is one representative (individual or firm) with high quality ($z_i = 1$) and one with low quality ($z_i = 0$). Production requires forming a “team” consisting of exactly one manager, one supervisor, one worker, and one firm, and the output of such a team is a supermodular function of the qualities of each of its four components. Supermodularity of the production function implies that it would be output-maximizing for the four high-quality individuals/firm to be matched and for the four low-quality individuals/firm to be matched. However, informational frictions may prevent such an outcome being reached and cause the matching process to be stochastic. Nevertheless, as long as the stochastic process is certain to generate two teams, each consisting of one representative from each dimension, the distribution over teams satisfies “top-to-bottom symmetry”.

In the next section, we illustrate how our orderings can be used to examine the influence of network structure on the degree of interdependence of behavior in social learning situations. In the scenario we analyze (based on Choi, Gale, and Kariv, 2005), the symmetry of the environment with respect to the two possible states of the world, signals, and actions generates distributions over agents’ actions which satisfy “top-to-bottom symmetry”.

PROPOSITION 7 *For random vectors Y and X with distributions on $L = \{0, 1\}^4$ that satisfy “top-to-bottom symmetry”, the following conditions are equivalent:*

- i) $Y \succeq_{SPM} X$;

ii) $Y \succeq_{CXMOD} X$;

iii) $Y \succeq_{DISP} X$;

iv) $c(Y) \succeq_{CX} c(X)$ and for all $i \neq j$, $Pr(Y_i = Y_j) \geq Pr(X_i = X_j)$.

The proof, which is in Appendix D, has a similar structure to that of Proposition 6. For $\delta \equiv g - f$, we apply the definition of the dispersion ordering, for each $s \in L$, and show that when $L = \{0, 1\}^4$ and the distributions satisfy top-to-bottom symmetry, $Y \succeq_{DISP} X$ if and only if the two conditions listed in iv) hold. (Top-to-bottom symmetry itself ensures that Y and X have identical marginal distributions.) We then adapt the construction used to prove Proposition 6 to show that if the conditions in iv) hold, then there exists a sequence of nonnegative ET's that sum to $g - f$. Since each ET raises the expectation of any supermodular function, it follows that $g \succeq_{SPM} f$ and hence, by Theorem 2, conditions i)-iv) in Proposition 7 are all equivalent.

Even though, in the environment of Proposition 7, the SPM, CXMOD, and DISP orderings are equivalent, the GWA ordering is strictly stronger and the concordance ordering strictly weaker.²¹

5 Applications

5.1 Interdependence of Behavior in Networks

Our first application illustrates how our orderings of interdependence can be used to study the influence of "social structure" on the degree of conformity of individual beliefs or choices in social learning situations. We adapt the theoretical approach of Gale and Kariv (2003), who model social structure by means of a social network and who assume that agents can observe only the actions of agents to whom they are directly connected by the network. A simplified version of Gale and Kariv's model has been studied experimentally by Choi, Gale, and Kariv (2005). Our orderings can in principle be applied to the theoretical predictions, the experimental data, or a combination

²¹To show that the GWA ordering is strictly stronger, let X have a uniform distribution on $\{0, 1\}^4$ and let Y be obtained from X by two ET's as defined in (3), each of size $\alpha \in (0, \frac{1}{16})$, one involving $(0, 0, 1, 1)$, $(1, 0, 1, 1)$, $(0, 1, 1, 1)$, and $(1, 1, 1, 1)$ and the other involving $(0, 0, 0, 0)$, $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, and $(1, 1, 0, 0)$. Both Y and X satisfy top-to-bottom symmetry, and by construction, $Y \succeq_{SPM} X$. However, for $r(z_1, z_2, z_3) = I_{\{z_1 + z_2 + z_3 \geq 2\}}$ and $s(z_4) = I_{\{z_4 = 1\}}$, $Cov(r(Y_1, Y_2, Y_3), s(Y_4)) < 0 = Cov(r(X_1, X_2, X_3), s(X_4))$, so $Y \not\succeq_{GWA} X$ does not hold. To show that the concordance ordering is strictly weaker, note that in Example 3, in which $Y \succeq_{CONC} X$ holds but $Y \not\succeq_{DISP} X$ does not, $\delta = g - f$ satisfies top-to-bottom symmetry.

of the two. Here, we use the orderings to compare the degree of interdependence in actions in the theoretical model, as we vary the structure of the social network.

Our economic environment matches that of Choi et al (2005), except that we study networks of four agents rather than three, in order to generate more variation in the predicted distributions. There are two equally likely states of the world, $\omega = 0$ and $\omega = 1$. Before period 1, each agent, independently and with probability q , observes a private signal which, with known probability, matches the state of the world. Conditional on the state, the signals are i.i.d. In each of periods $1, 2, \dots$, each agent $i \in \{A, B, C, D\}$ simultaneously chooses an action $i_t \in \{0, 1\}$ to maximize his current-period expected payoff, which is 1 if his action matches ω and 0 otherwise. After each period, each agent observes the actions of those agents to whom he is directly connected by the network and updates his beliefs using Bayesian reasoning. If an agent is indifferent between two actions, he chooses the action that matches his signal, if he received one; if he did not receive a signal and is ever indifferent, then randomizes symmetrically.

We use the orderings studied in Sections 2 through 4 to compare the degree of interdependence in agents' actions as the network structure varies. Because the environment is completely symmetric with respect to the two states of the world, it follows that for any value of q and for any precision of the private signals, the distributions of actions under any network structure always satisfy "top-to-bottom symmetry": letting Z denote the vector of action choices, $Pr(Z = a) = Pr(Z = 1 - a)$ for any $a \in \{0, 1\}^4$. The easily checkable conditions in iv) of Proposition 7 can therefore be used to compare interdependence in agents' actions according to the SPM, CXMOD, and DISP orderings.

As Choi et al (2005) demonstrate, when networks are incomplete, so agents do not observe the actions of all of their neighbors' neighbors, Bayesian reasoning can become very complex. Since our purpose here is to illustrate the application of our dependence orderings, not to conduct an exhaustive analysis, we make two assumptions that significantly simplify the computations. First, we focus on actions in period 2. Second, we focus on the limiting distributions of actions as the probability q of each agent receiving a signal goes to zero. For any level of precision of the private signals, letting q go to zero maximizes the impact of the network structure (relative to correlations in private signals) on interdependence.²² In this limit, each agent in period 1 randomizes symmetrically between the two actions and, in period 2, he copies the period-1 action chosen by a strict majority of his neighbors, if such a strict majority exists; if it does not, he randomizes

²²In this limit, the distributions of actions are independent of the precision of the signal.

symmetrically.²³²⁴

The four network structures we compare are illustrated in the left-hand column of Figure 2.²⁵ In the complete network, each agent is connected to all other agents. In the diamond network, agents A and C are both connected to both B and D but not to each other; similarly, B and D are both connected to both A and C but not to each other. The line network is similar to the diamond network except that there is no link between A and D. In the star network, agents B, C, and D are each linked only to A.

Figure 2 also presents, for each network structure, the ex ante distribution (computed before the start of the game) of the period-2 actions of the four players, shown as probabilities (expressed in units of $\frac{1}{64}$) for the 16 nodes of the four-dimensional cube $L = \{0,1\}^4$. The diagram at the bottom left shows which of the four axes measures each agent's action. The bottom row of the figure presents, as a benchmark, the distribution of actions if each agent independently chooses each of the two actions with probability $\frac{1}{2}$.

Before applying the dependence orderings formally, we briefly interpret the behavior summarized by the distributions in Figure 2. In the complete network, if three or four agents choose the same action in period 1 (which happens with probability $\frac{5}{8}$), then all agents will see that action chosen by a strict majority of their neighbors, so in period 2 all agents will choose that action. If two agents choose 0 in period 1 and the other two choose 1 (which happens with probability $\frac{3}{8}$), then those who chose 0 will see a majority of their neighbors choosing 1, so will choose 1 in period 2, while those who chose 1 will, by the same logic, choose 0 in period 2. The probabilities in Figure 2 then follow from the above observations, combined with the symmetry across agents and states of the world.

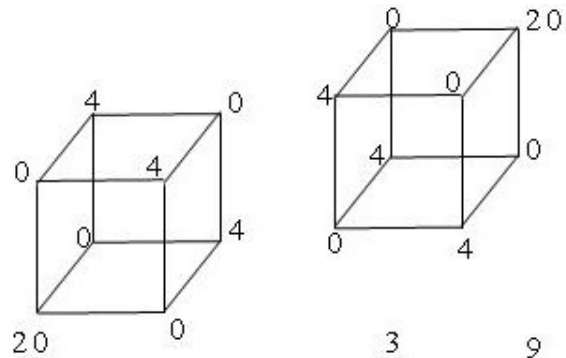
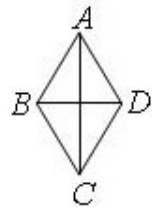
In the diamond network, if the period-1 actions of B and D match, then A and C will both copy this action, so the period-2 actions of A and C will match. A and C will also match in period 2 if B and D do not match in period 1 but A and C randomly choose the same period-2 action. Thus,

²³These limiting distributions are approximations to the distributions arising when q is very small; when q is very small, though any number of agents between 0 and 4 may end up receiving signals, the state in which no agent receives a signal is much more likely than the others and is thus the dominant influence on the outcome.

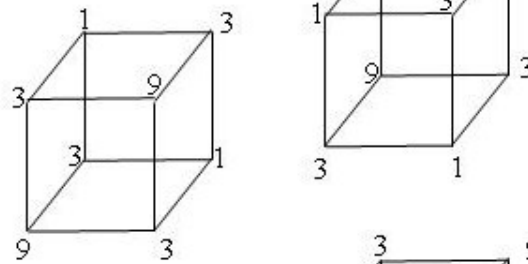
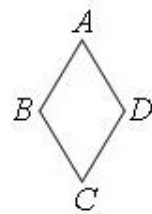
²⁴Another limiting case which also highlights the impact of network structure on interdependence is that in which $q = 1$ and the probability that the private signal matches the state goes to $\frac{1}{2}$. In this case, too, the random vector of period-1 actions has a uniform distribution over the 16 nodes of $\{0,1\}^4$; in period 2, each agent chooses the period-1 action chosen by a strict majority of those he observes, *including himself*, if such a strict majority exists; if it does not, he follows his signal.

²⁵In our networks, in contrast to Choi et al (2005), if agent i is linked to (and hence can observe the action of) agent j , then agent j is linked to i .

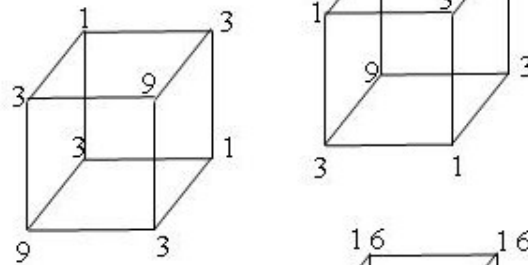
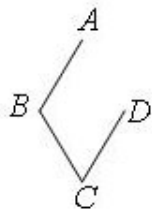
Complete



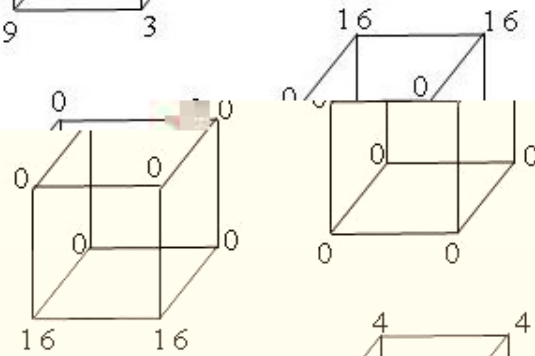
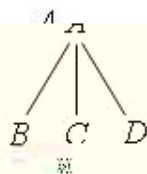
Diamond



Line



Star



Completely random choices

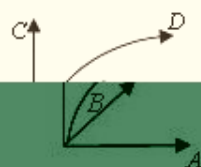
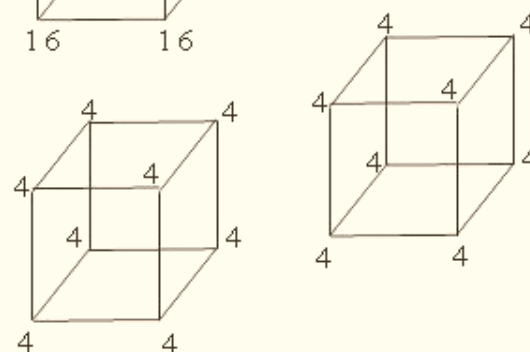


Figure 2: Distributions of choices in period 2, in the limit, as the probability of each individual receiving a signal in period 1 goes to 0 (probabilities expressed in units of $1/64$)

the overall probability of A and C matching in period 2 is $\frac{3}{4}$, and similarly for B and D matching. However, period-2 actions of A and C are independent of the period-2 actions of B and D.

In the limiting case we are considering, the distribution of period-2 actions is the same for the line network as for the diamond network, even though in the line network, agents A and D cannot observe each other's actions. In the line network, it remains true that if B and D match in period 1, then A and C will match in period 2. If in period 1, B and D do not match, C will randomize symmetrically (as in the diamond network), but A will copy B's period-1 action. The overall

The proof of greater weak association in parts i) and ii) is quite involved. In contrast, proving that the same rankings hold for the weaker supermodular ordering is, in this setting, extremely straightforward in light of Proposition 7, since using Figure 2 it is simple to confirm that condition *iv*) is satisfied for each pair of distributions compared in parts i) and ii) of the claim.

Part iii) of the claim is also proved using Proposition 7. The second part of condition *iv*) of Proposition 7 says that for any two agents, the probability of their period-2 actions matching must be at least as high under one distribution as under the other. Consider first the probability that the period-2 actions of A and C match. This probability is strictly higher in the complete, diamond, and line networks than in the star network, because in the complete, diamond, and line networks, A and C observe at least one individual in common in period 1, so their period-2 choices match with probability greater than $\frac{1}{2}$, whereas in the star network, A and C have no common neighbors, so their period-2 choices are independent. On the other hand, the probability that the period-2 actions of B and C match equals 1 in the star network, where B and C both copy the period-1 action of A; in the complete, diamond, and line networks, the probability that B and C match in period 2 is strictly less than 1.²⁶

If, in comparing interdependence, we are prepared to adopt an ordering which treats agents symmetrically (even if asymmetrically positioned in the network), then we can apply Proposition 5, part b) to show:

CLAIM 2 In the limit as the probability of each agent receiving a signal goes to zero, both the distribution of period-2 actions in the star network and that in the complete network are more interdependent according to the SSPM and SCXMOD orderings than the corresponding distribution under the diamond or line network.

By Proposition 5, checking the SSPM and SCXMOD orderings in this setting with binary actions merely requires checking whether the distributions of the number of agents choosing action 1 can be ranked according to the convex ordering. Since the number of agents here is four, and the distributions display “top-to-bottom symmetry”, $c(Y) \succeq_{CX} c(X)$ reduces to the following two inequalities:

$$Pr\left(\sum_{i=1}^4 Y_i = 4\right) \geq Pr\left(\sum_{i=1}^4 X_i = 4\right),$$

²⁶In fact, even the concordance ordering is too weak to rank the star network against the complete, diamond, or line networks, because in this environment with “top-to-bottom symmetry”, the second part of condition *iv*) in Proposition 7 is also a necessary condition for the concordance ordering.

$$2Pr(\sum_{i=1}^4 Y_i = 4) + Pr(\sum_{i=1}^4 Y_i = 3) \geq 2Pr(\sum_{i=1}^4 X_i = 4) + Pr(\sum_{i=1}^4 Y_i = 3).$$

It is easy to confirm, using Figure 2, that these inequalities are satisfied when Y has the distribution of actions under the star or complete networks and X the distribution of actions under the diamond or line networks. However, when Y has the distribution under the complete network and X the distribution under the star network, then the first inequality above is satisfied strictly, while the second is violated. Hence, the distributions of actions under the complete and star networks cannot be ranked according to their interdependence, even when we weaken our ordering to one such as SSPM or SCXMOD which treats agents symmetrically.

5.2 Ex Post Inequality under Uncertainty: Tournaments vs. Independent Schemes

This section illustrates how our orderings of interdependence can be applied to the assessment of ex post inequality under uncertainty. In many group settings where individual outcomes (e.g. rewards) are uncertain, members of the group may be concerned, ex ante, about how unequal their ex post rewards will be. Meyer and Mookherjee (1987) argued that a group or a planner's aversion to ex post inequality can be formalized by adopting an ex post welfare function that is supermodular in the realized utilities of the individuals, and thus that the supermodular ordering be used to rank multivariate distributions according to their degree of ex post inequality. More generally, we can use any of the orderings studied in Sections 2 through 4 to compare reward schemes according to the degree of interdependence of the uncertain utilities of the members of a group.

Tournament reward schemes, which distribute a prespecified set of rewards among individuals, one to each person, are a pervasive form of reward scheme in practice. There are many reasons for this, relating to incentives and the cost of implementation.²⁷

However, when groups dislike ex post inequality, tournament reward schemes should be particularly unappealing, since they generate a form of negative interdependence among rewards: if one individual receives a higher reward, this must be accompanied by another person's receiving a lower reward. This intuitive reasoning motivates us to use our interdependence orderings to formally

²⁷Tournaments can in some environments with agent moral hazard relax the incentive-insurance tradeoff facing the principal. Also, by committing the principal to distribute a prespecified total level of rewards, they deter the principal from underreporting agents' performances. The set of rewards is often determined by exogenous factors and hence difficult to alter, as with promotion contests in firms, and tournament reward schemes require only ordinal information about performances, so are cheaper to implement than schemes requiring cardinal information.

compare tournament reward schemes with schemes that provide each individual with the same marginal distribution over rewards but determine rewards *independently*. Besides shedding light on tournaments, these comparisons will also illuminate some of the subtleties associated with concepts of negative interdependence for distributions in more than two dimensions.

Since our interdependence orderings are invariant to monotonic relabelings of coordinates, we will focus henceforth on n -agent tournaments for which the prespecified set of prizes is the set $L_n \equiv \{0, 1, \dots, n-1\}$. With probability 1, a tournament distributes each of the n prizes, one to each individual. Thus, for any tournament, the n -dimensional ex ante distribution of prizes assigns strictly positive probability to at most $n!$ nodes in $L = L_n \times \dots \times L_n$, where these nodes are the permutations of L_n . Our discussion here suppresses consideration of incentives and, in order to focus purely on rankings of interdependence, compares only multivariate distributions with identical marginals.²⁸

A *symmetric tournament* assigns *equal* probability to each of the $n!$ possible prize allocations. Symmetric tournaments generate a very strong form of negative interdependence in rewards, as shown by the following proposition:

PROPOSITION 8 *For any number of individuals n , the distribution of prizes in a symmetric tournament is dominated according to the greater weak association ordering by the independent distribution with identical (uniform) marginal distributions.*

Proof. Joag-Dev and Proschan (1983) showed that the prize distribution under a symmetric tournament (termed by them a "permutation distribution") displays negative association, as defined in Section 2.2. The proposition then follows from this result and the definitions of negative association and greater weak association. ■

Symmetric tournaments are, however, non-generic in the class of tournaments, so it is natural to ask: For what orderings of interdependence is it the case that *any* tournament, no matter how asymmetric across individuals, displays less interdependence (i.e. more negative interdependence) than the corresponding independent distribution with identical marginals? The following simple example, for $n = 3$, shows that none of the five orderings compared in Theorem 2 will always rank a tournament as less interdependent than the corresponding independent distribution.

Consider a tournament which assigns probability $\frac{1}{2}$ to each of the outcomes $(0, 1, 2)$ and $(1, 2, 0)$.

²⁸Thus, tradeoffs between interdependence and risk, or between interdependence and ex ante fairness, are suppressed.

The independent distribution with identical marginals gives individual 1 each of the prizes 0 and 1 with probability $\frac{1}{2}$, individual 2 each of the prizes 1 and 2 with probability $\frac{1}{2}$, and individual 3 each of the prizes 2 and 0 with probability $\frac{1}{2}$. Let (X_1, X_2, X_3) have the distribution under the tournament and (Y_1, Y_2, Y_3) have the corresponding independent distribution. Then for $z = (1, 2, 0)$,

$$Pr(Y \geq z) = \frac{1}{4} < \frac{1}{2} = Pr(X \geq z),$$

so $Y \succeq_{CONC} X$ does not hold and hence, by Theorem 2, the tournament does not display less interdependence than the independent distribution according to the dispersion, convex-modular, supermodular, or greater weak association orderings either. In this example, the outcome $(1, 2, 0)$ is better than the outcome $(0, 1, 2)$ for both individuals 1 and 2, though it is worse for individual 3. Thus, under the tournament distribution, while the rewards of individuals 1 and 3 are negatively interdependent, as are the rewards of individuals 2 and 3, the rewards of individuals 1 and 2 are positively interdependent. The five interdependence orderings in Theorem 2 all fail to rank the independent distribution above the tournament in this example because for each ordering to hold, it is a necessary condition that each pairwise joint distribution be rankable according to the ordering.

Nevertheless, if we adopt an ordering which treats individuals symmetrically with respect to the assessment of interdependence, then we can show

PROPOSITION 9 *For three individuals, given any (arbitrarily asymmetric) tournament, the prize distribution under the tournament displays more negative interdependence according to the symmetric supermodular and symmetric convex-modular orderings than the corresponding independent distribution with identical marginals.*

A crucial role in the proof of Proposition 9 (which is provided in Appendix F) is played by the following explicit characterization of the SSPM and SCXMOD orderings on $L = \{0, 1, 2\}^3$. Since this characterization is of interest in its own right, we present it as a proposition.

PROPOSITION 10 *For random vectors Y and X with distributions g and f , respectively, on $L = \{0, 1, 2\}^3$, let δ^s denote the symmetrized version of $\delta = g - f$. The sets of conditions i), ii), and iii) below are equivalent:*

$$i) \ g \succeq_{SSPM} f;$$

$$ii) \ g \succeq_{SCXMOD} f;$$

$$iii) \ g^{symm} \text{ and } f^{symm} \text{ have identical marginals, and } \delta^s \text{ satisfies the following inequalities:}$$

- 1) $\delta^s(2, 2, 2) \geq 0$;
- 2) $\delta^s(0, 0, 0) \geq 0$;
- 3) $2\delta^s(2, 2, 2) + 3\delta^s(2, 2, 1) + 3\min\{0, \delta^s(2, 2, 0)\} \geq 0$;
- 4) $2\delta^s(0, 0, 0) + 3\delta^s(1, 0, 0) + 3\min\{0, \delta^s(2, 0, 0)\} \geq 0$;
- 5) $2\delta^s(2, 2, 2) + 4\delta^s(2, 2, 1) + 2\delta^s(2, 1, 1) + 2\delta^s(2, 2, 0) + 2\delta^2(2, 1, 0) + \min\{0, \delta^s(2, 2, 0)\} + \min\{0, \delta^s(2, 0, 0)\} \geq 0$.

Proposition 10 is proved in Appendix F. To prove that iii) implies i), we adopt a constructive approach analogous to the proofs of Propositions 6 and 7. We show that if the conditions in iii) hold, then there exists a sequence of nonnegative ET's that sum to δ^s and hence convert f^{symm} to g^{symm} . Therefore, $g^{symm} \succeq_{SPM} f^{symm}$, which by Proposition 1 is equivalent to $g \succeq_{SSPM} f$.

One might, at first glance, suppose that when we adopt an ordering, such as SSPM or SCXMOD, which treats individuals symmetrically with respect to the assessment of interdependence, then the comparison of an asymmetric tournament with its corresponding independent distribution should just reduce to the comparison of a symmetric tournament with its corresponding independent distribution. This reasoning, however, is flawed. It is true that, as Propositions 1 and 2 show, comparing two distributions according to the SSPM or SCXMOD orderings is equivalent to comparing the symmetrized versions of the distributions according to the supermodular or convex-modular orderings, respectively. And it is also true that, for any (arbitrarily asymmetric) tournament, the symmetrized version of the distribution of prizes it generates matches the distribution under a symmetric tournament. However, as we vary the degree of asymmetry across individuals in the tournament we are considering, we vary not only the corresponding independent distribution with identical marginals, *but also the symmetrized version of this distribution and the degree of interdependence it displays*. To see this, consider at one extreme a symmetric tournament; its corresponding independent distribution is clearly also symmetric, so the symmetrized version of this independent distribution is independent. Consider at the other extreme a degenerate tournament which with probability 1 generates the outcome $(2, 1, 0)$. The corresponding independent distribution is also degenerate and coincides with the tournament distribution; hence its symmetrized version coincides with the distribution under a symmetric tournament, which we know from Proposition 8 displays a strong form of negative interdependence. Thus, even when using orderings that treat individuals symmetrically, comparisons of arbitrary tournaments against their corresponding independent distributions do not reduce to the comparison made in Proposition 8: Proposition 9 does not follow from Proposition 8.²⁹

²⁹Using very different methods, Meyer and Strulovici (2010) show that the conclusion of Proposition 9 holds for any number of individuals n . However, they do not provide an analog of Proposition 10.

6 Conclusion

In this paper we examined five orderings of interdependence for multivariate distributions. While greater weak association, the supermodular ordering, and the concordance ordering have received some attention in the statistics and economics literatures, this paper introduces the dispersion ordering and the convex-modular ordering. While in two dimensions, all five orderings are equivalent (Theorem 1), for an arbitrary number of dimensions $n > 2$, Theorem 2 showed that the five orderings are strictly ranked. In general, the dispersion and concordance orderings are easier than the others to work with, since for any multivariate support, checking whether or not two distributions can be ranked reduces to checking, for each point in the support, whether or not a fixed set of inequalities holds. For a range of multivariate special cases, involving binary variables and/or various forms of symmetry, we demonstrated some equivalences among the orderings and provided easily checkable and easily interpretable necessary and sufficient conditions for the orderings to hold.

We briefly illustrated the application of our orderings to the assessment of interdependence of behavior in a model of learning in networks and to the assessment of ex post inequality under uncertainty. However, multivariate dependence orderings have many more potential applications in economics and finance, including to comparisons of multidimensional inequality; assessments of the efficiency of matching institutions; valuations of portfolios of assets or insurance policies; and assessments of systemic risk.

7 References

- Acharya, V. V., "A Theory of Systemic Risk and Design of Prudential Bank Regulation," *Journal of Financial Stability*, 2009, 5, 224-255.
- Adler, M. D., and Sanchirico, C. W., "Inequality and Uncertainty: Theory and Legal Applications", *University of Pennsylvania Law Review*, 2006, 155, 279.
- Adrian, T., and Brunnermeier, M.K., "CoVar", 2009, FRB of New York Staff Report No. 348.
- Atkinson, A. B., "On the Measurement of Inequality", *Journal of Economic Theory*, 1970, 2, 244-263.
- Atkinson, "Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches", *J. of Econ. Inequality*, 2003, 1, 51-65.

Atkinson, A. B., and Bourguignon, F., "The Comparison of Multi-Dimensioned Distributions of Economic Status", *Review of Economic Studies*, 1982, 49, 183-201.

Ben-Porath, E., Gilboa, I., and Schmeidler, D., "On the Measurement of Inequality under Uncertainty", *Journal of Economic Theory*, 1997, 75, 194-204.

Bourguignon, F., and Chakravarty, S.R., "Multi-dimensional poverty orderings", *DELTA W.P.* 2002-22.

Burton, R.M., Dabrowski, A.R., and Dehling, H., "An Invariance Principle for Weakly Associated Random Variables", *Stochastic Processes and their Applications*, 1986, 23, 301-306.

Calvo-Armengol, A., and Jackson, M.O., "The Effects of Social Networks on Employment and Inequality", *American Economic Review*, 2004, 94, 426-454.

Calvo-Armengol, A., and Jackson, M.O., "Networks in Labor Markets: Wage and Employment Dynamics and Inequality", *Journal of Economic Theory*, 2007, 132, 27-46.

de Castro, L.I., "Aliation and Dependence in Economic Models", Disc. Paper 1479, Center for Mathematical Studies in Economics and Management Science, Northwestern University, 2009.

Christodides, T.C., and Vaggelatos, E., "A Connection between Supermodular Ordering and Positive/Negative Association", *Journal of Multivariate Analysis*, 2004, 88, 138-151.

Choi, S., Gale, D., and Kariv, S., "Behavioral Aspects of Learning in Social Networks: An Experimental Study", *Advances in Applied Microeconomics*, 2005, 13.

Dasgupta, P., Sen, A.K., and Starrett, D., "Notes on the Measurement of Inequality", *Journal of Economic Theory*, 1973, 6, 180-187.

Decancq, K., "Multidimensional Inequality and the Copula", mimeo, K. U. Leuven, Nov. 2007.

Denuit, M., Dhaene, J., Goovaerts, M.J., and Kaas, R., *Actuarial Theory for Dependent Risks: Measures, Orders, and Models*, 2005, Wiley.

Epstein, L., and Tanny, S., "Increasing Generalized Correlation: A Definition and Some Economic Consequences", *Canadian Journal of Economics*, 1980, 13, 16-34.

Esary, J.D., Proschan, F., and Walkup, D.W., "Association of Random Variables, with Applications", *Annals of Mathematical Statistics*, 1967, 44, 1466-1474.

Fernandez, R., and Gali, J., "To Each According to?? Markets, Tournaments, and the Matching Problem with Borrowing Constraints", *Review of Economic Studies*, 1999, 66, 799-824.

Gajdos, T., and Maurin, E., "Unequal Uncertainties and Uncertain Inequalities: An Axiomatic Approach", *Journal of Economic Theory*, 2004, 116, 93-118.

Gale, D., and Kariv, S., "Bayesian Learning in Social Networks", *Journal of Economic Theory*, 2003, 45, 329-346.

Galeotti, A., Goyal, S., Jackson, M., Vega-Redondo, F., and Yariv, L., "Network Games", *Review of Economic Studies*, 2010, 77, 218-244.

Hardy, G.H., Littlewood, J.E., and Polya, G., "Some Simple Inequalities Satisfied by Convex Functions", *Messenger of Mathematics*, 1929, 58, 145-52.

Hardy, G.H., Littlewood, J.E., and Polya, G., *Inequalities*, Cambridge: Cambridge University Press, 2nd edition, 1952.

Hennessy, D. A., and Lapan, H. E., "A Definition of 'More Systematic Risk' with Some Welfare Implications", *Economica*, 2003, 70, 493-507.

Hu, T., Müller, A., and Scarsini, M., "Some Counterexamples in Positive Dependence", *Journal of Statistical Planning and Inference*, 2004, 153-158.

Hu, T., Xie, C., and Ruan, L., "Dependence Structures of Multivariate Bernoulli Random Vectors", *Journal of Multivariate Analysis*, 2005, 94, 172-195.

Joag-Dev, K., and Proschan, F., "Negative Association of Random Variables, with Applications", *Annals of Statistics*, 1983, 11, 286-95.

Joe, H., "Multivariate Concordance", *Journal of Multivariate Analysis*, 1990, 35, 12-30.

Joe, H., *Multivariate Models and Dependence Concepts*, London: Chapman and Hall/CRC, 1997.

Kroll, Y., and Davidovitz, L., "Inequality Aversion versus Risk Aversion", *Economica*, 2003, 70, 19-29.

Levy, H., and Paroush, J., "Toward Multivariate Efficiency Criteria", *Journal of Economic Theory*, 1974, 7, 129-42.

McAfee, R. P., "Coarse Matching", *Econometrica*, 2002, 70, 2025-34.

Meyer, M.A., "Interdependence in Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty", *Nuffield College Disc. Paper No. 49*, April 1990.

Meyer, M. A., and Mookherjee, D., "Incentives, Compensation and Social Welfare", *Review of Economic Studies*, 1987, 45, 209-226.

Meyer, M. A., and Rothschild, M., "(In)efficiency in Sorting and Matching", *mimeo*, Oxford, 2004.

Meyer, M.A., and Strulovici, B., "The Supermodular Stochastic Ordering", *draft*, Oxford, 2010.

Milgrom, P.R., and Weber, R.J., "A Theory of Auctions and Competitive Bidding", *Econometrica*, 1982, 50, 1089-1122.

Motzkin, T.S., Rai, H., Thompson, G.L., and Thrall, R.M., "The Double Description Method", in Kuhn, H.W., and Tucker, A.W. (eds.), *Contributions to the Theory of Games II*, Princeton: Princeton University Press, 1953.

Müller, A., and Stoyan, D., *Comparison Methods for Stochastic Models and Risks*, 2002, Wiley.

Prat, A., "Should a Team Be Homogenous?", *European Economic Review*, 2002, 46, 1187-1207.

Rothschild, M. and Stiglitz, J. E. "Increasing risk: I. A definition", *Journal of Economic Theory*, 1970, 2, 225-243.

Shaked, M., and Shanthikumar, J.G., "Supermodular Stochastic Orders and Positive Dependence of Random Vectors", *Journal of Multivariate Analysis*, 1997, 61, 86-101.

Shaked, M., and Tong, Y.L., "Some Partial Orderings of Exchangeable Random Variables by Positive Dependence", *Journal of Multivariate Analysis*, 1985, 17, 333-349.

Shimer, R., and Smith, L., "Assortative Matching and Search", *Econometrica*, 2000, 68, 343-369.

Tchen, A.H., "Inequalities for Distributions with Given Marginals", *Annals of Probability*, 1980, 8, 814-827.

Topkis, D.M., "Minimizing a Submodular function on a Lattice", *Operations Research*, 1978, 26, 305-331.

Appendices

A Proof that $Y \succeq_{GWA} X$ implies $Y \succeq_{SPM} X$

This proof builds on Cristo des and Vaggelatou's (2004) proof that if (Y_1, \dots, Y_n) is weakly associated, then (Y_1, \dots, Y_n) dominates its independent counterpart according to the supermodular ordering. First, we define a new random vector Z that has the same distribution as X and is independent of Y . We then show by induction that $Y \succeq_{SPM} Z$ and hence $Y \succeq_{SPM} X$.

For $n = 2$, the result is proved in the references cited after Theorem 2. Suppose that it is true for $n = m - 1$, so that for all supermodular w , $Ew(Y_2, \dots, Y_m) \geq Ew(Z_2, \dots, Z_m)$. Then

$$\begin{aligned} Ew(Z_1, Y_2, \dots, Y_m) &= \sum_{i=0}^{l_i-1} E[w(i, Y_2, \dots, Y_m) | Z_1 = i] Pr(Z_1 = i) \\ &= \sum_{i=0}^{l_i-1} E[w(i, Y_2, \dots, Y_m)] Pr(Z_1 = i) \end{aligned} \quad (16)$$

$$= \sum_{i=0}^{l_i-1} E[w(i, Y_2, \dots, Y_m)] Pr(Y_1 = i) \quad (17)$$

$$\begin{aligned} &\geq \sum_{i=0}^{l_i-1} E[w(i, Z_2, \dots, Z_m)] Pr(Y_1 = i) \\ &= \sum_{i=0}^{l_i-1} E[w(i, Z_2, \dots, Z_m) | Y_1 = i] Pr(Y_1 = i) \\ &= Ew(Y_1, Z_2, \dots, Z_m). \end{aligned} \quad (18)$$

The equality in (16) uses the independence of Z from Y , and that in (17) the assumption that Y and X , hence Y and Z , have identical marginal distributions (along with the fact that $E[w(i, Y_2, \dots, Y_m)]$ is a univariate function of i). The inequality follows by applying the induction hypothesis for each value of i . The equality in (18) uses the independence of Z from Y .

Now we show that for all supermodular w ,

$$Ew(Y_1, Y_2, \dots, Y_m) - Ew(Z_1, Z_2, \dots, Z_m) \geq Ew(Z_1, Y_2, \dots, Y_m) - Ew(Y_1, Z_2, \dots, Z_m),$$

which combined with the previous steps yields the desired result. First observe that we can write

$$w(Y_1, Y_2, \dots, Y_m) - w(Z_1, Y_2, \dots, Y_m) = \sum_{i=0}^{l_i-1} (I_{\{Y_1 > i\}} - I_{\{Z_1 > i\}}) (w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)).$$

Hence

$$\begin{aligned}
& Ew(Y_1, Y_2, \dots, Y_m) - Ew(Z_1, Y_2, \dots, Y_m) \\
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Z_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))]\} \\
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Z_1 > i\}}] E[w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)]\} \tag{19}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Y_1 > i\}}] E[w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)]\} \tag{20}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{l_i-1} Cov(I_{\{Y_1 > i\}}, w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)) \\
&\geq \sum_{i=0}^{l_i-1} Cov(I_{\{Z_1 > i\}}, w(i+1, Z_2, \dots, Z_m) - w(i, Z_2, \dots, Z_m)) \\
&= Ew(Z_1, Z_2, \dots, Z_m) - Ew(Y_1, Z_2, \dots, Z_m). \tag{21}
\end{aligned}$$

The equality in (19) uses the independence of Z from Y , and that in (20) the assumption that Y and X , hence Y and Z , have identical marginal distributions. The inequality holds since i) $I_{\{Y_1 > i\}}$ is an increasing function of Y_1 ; ii) $(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))$ is an increasing function of (Y_2, \dots, Y_m) since w is supermodular; and iii) by hypothesis, $Y \succeq_{GWA} X$ and hence $Y \succeq_{GWA} Z$. The equality in (21) follows from the logic of the first four equalities. \blacksquare

B Proof that for $n \geq 3$, the supermodular ordering is strictly stronger than the convex-modular ordering

Since every convex-modular function is supermodular, the supermodular ordering implies the convex-modular ordering. To show that the implication is strict, we provide here an example of a supermodular function on $L = \{0, 1, 2\}^3$ which cannot be written as a nonnegative weighted sum of convex-modular ones:

Define w as follows: $w(2, 2, 2) = 3$, $w(2, 2, 1) = w(2, 1, 2) = w(1, 2, 2) = 2$, $w(2, 2, 0) = w(2, 0, 2) = w(0, 2, 2) = 1$, $w(2, 1, 1) = w(1, 2, 1) = w(1, 1, 2) = 1$, $w(0, 2, 1) = 1$, and $w(z) = 0$ for all other nodes $z \in L$. We first show that this function $w(x)$ is not itself convex-modular. Suppose it were. Then clearly the function $\phi(\cdot)$ would have to take values in $\{0, 1, 2, 3\}$. If $\sum_{i=1}^3 r^i(x_i)$ were strictly

larger at $(0, 2, 2)$ than at $(0, 2, 1)$, then $\phi(\cdot)$ would not be convex, since $\phi(\cdot)$ would rise from 0 to 1 but then remain constant at 1 even though $\sum_{i=1}^3 r^i(x_i)$ increased. If, instead, $\sum_{i=1}^3 r^i(x_i)$ took on the same value at $(0, 2, 2)$ as at $(0, 2, 1)$, then since $\sum_{i=1}^3 r^i(x_i)$ is modular (additively separable) in the x_i 's, it would follow that $\sum_{i=1}^3 r^i(x_i)$ took on the same value at $(1, 2, 2)$ as at $(1, 2, 1)$. However, $w(1, 2, 2) = 2 > 1 = w(1, 2, 1)$. Thus, we reach a contradiction, so $w(z)$ as defined above is not convex-modular.

In Meyer and Strulovici (2010), we show how, for any finite support L , the "double description method", conceptualized by Motzkin et al (1953), can be used to determine the extreme rays of the cone of supermodular functions on L . For $L = \{0, 1, 2\}^3$, we show there that the function w defined above is an extreme ray of the cone of supermodular functions and hence cannot be non-trivially expressed as a nonnegative weighted sum of supermodular functions. This, combined with the fact that it is not itself convex-modular, shows that it cannot be expressed as a nonnegative weighted sum of convex-modular functions.

C Proof that in Example 1', $g' \succeq_{GWA} f'$

Since f' and g' are symmetric distributions, all partitions of $\{Z_1, Z_2, Z_3\}$ into disjoint sets $\{Z_i\}$ and $\{Z_j, Z_k\}$ will yield the same covariances. As is well known, any increasing functions on $L = \{0, 1\}^3$ can be written as nonnegative weighted combinations of indicator functions, so it is sufficient to focus on functions r, s that are increasing indicator functions. Take $r(z_1) = I_{\{z_1=1\}}$. If $s(z_2, z_3)$ is a function of only z_2 or only z_3 , then $Cov(r(Y_1), s(Y_2, Y_3)) \geq Cov(r(X_1), s(X_2, X_3))$ follows from the facts that i) $g' - f' = g - f$, where g, f are defined in Example 1; ii) $g \succeq_{SPM} f$ and the supermodular ordering is a difference-based order, so $g' \succeq_{SPM} f'$; and iii) by Theorem 1, for $n = 2$, the greater weak association and supermodular orderings are equivalent.

It remains only to consider $s(z_2, z_3) = I_{\{z_2=z_3=1\}}$ and $s(z_2, z_3) = I_{\{z_2+z_3 \geq 1\}}$. For $s(z_2, z_3) = I_{\{z_2=z_3=1\}}$,

$$Cov(r(Y_1), s(Y_2, Y_3)) - Cov(r(X_1), s(X_2, X_3)) = 0,$$

while for $s(z_2, z_3) = I_{\{z_2+z_3 \geq 1\}}$,

$$Cov(r(Y_1), s(Y_2, Y_3)) > 0 > Cov(r(X_1), s(X_2, X_3)).$$

Therefore, $g' \succeq_{GWA} f'$.

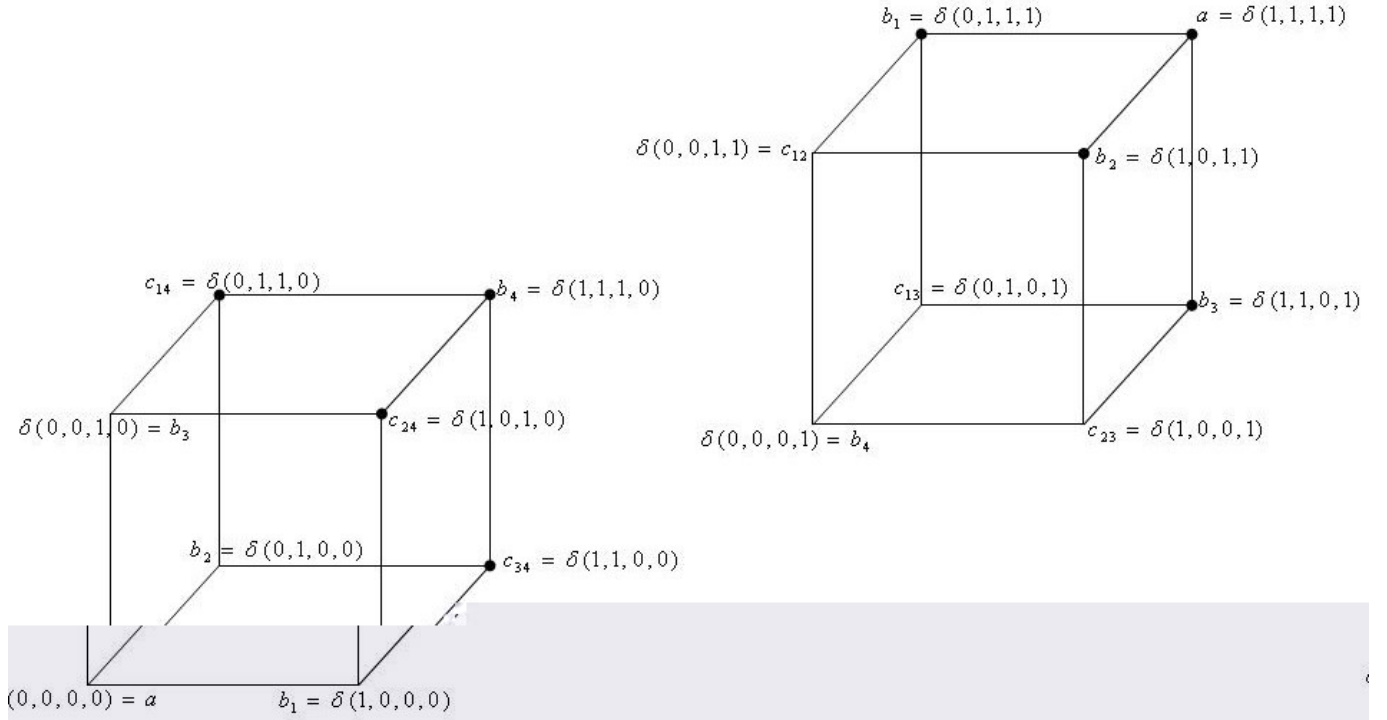


Figure 3: $\delta = g - f$ for $L = \{0,1\}^4$
and with "top-to-bottom symmetry"
(which implies $c_{14} = c_{23}$,
 $c_{24} = c_{13}$, $c_{34} = c_{12}$)

D Proof of Proposition 7

Since the SPM, CXMOD, and DISP orderings are all difference-based orderings, it is sufficient to work with $\delta \equiv g - f$, which for distributions on $L = \{0,1\}^4$ satisfying top-to-bottom symmetry is depicted in Figure 3. In terms of the components of δ , as defined in Figure 3, $c(Y) \succeq_{CX} c(X)$ corresponds to $a \geq 0$ and $2a + \sum_{i=1}^4 b_i \geq 0$, and $Pr(Y_i = Y_j) \geq Pr(X_i = X_j)$ for all $i \neq j$ corresponds to $a + b_i + b_j + c_{ij} \geq 0$ for all $i \neq j$. Note that top-to-bottom symmetry of δ implies that $c_{12} = c_{34}$, $c_{13} = c_{24}$, and $c_{23} = c_{14}$.

We first show that conditions iii) and iv) in the proposition are equivalent. For $s = (0,0,0,0)$, $Y \succeq_{DISP} X$ implies $a \geq 0$ and $2a + \sum_{i=1}^4 b_i \geq 0$; for $s = (1,0,0,0)$ and all permutations thereof, it implies $a + b_i \geq 0$ for all i ; for $s = (1,1,0,0)$ and all permutations thereof, it implies $a + b_i + b_j + c_{ij} \geq 0$ for all $i \neq j$; and for $s = (1,1,1,0)$ and all permutations thereof, it implies that X and Y have identical marginal distributions. That X and Y have identical marginals is already implied by top-

to-bottom symmetry. Moreover, top-to-bottom symmetry, in conjunction with $a + b_i + b_j + c_{ij} \geq 0$ for all $i \neq j$, implies $a + b_i \geq 0$ for all i . To verify this last claim, observe that

$$\begin{aligned} a + b_i + b_j + c_{ij} &\geq 0 \\ a + b_i + b_k + c_{ik} &\geq 0 \\ a + b_i + b_l + c_{il} &\geq 0 \end{aligned}$$

together imply

$$3a + 2b_i + \sum_{j=1}^4 b_j + c_{ij} + c_{ik} + c_{il} \geq 0. \quad (22)$$

Since identical marginals and top-to-bottom symmetry imply

$$a + \sum_{j=1}^4 b_j + c_{ij} + c_{ik} + c_{il} \geq 0,$$

(22) thus reduces to $a + b_i \geq 0$. Therefore, under the hypotheses of the proposition, $Y \succeq_{DISP} X$ is equivalent to $a \geq 0$, $2a + \sum_{i=1}^4 b_i \geq 0$, and, for all $i \neq j$, $a + b_i + b_j + c_{ij} \geq 0$. Since, as indicated above, these three inequalities are together equivalent to condition iv), we have shown that conditions iii) and iv) are equivalent.

We now decompose δ into 24 elementary transformations (ET's) of the form defined in (3): for each of the 24 faces of the hypercube $L = \{0, 1\}^4$, there is one ET involving the 4 nodes on that face. We will abuse notation slightly and let the value of δ at a given node also serve as the label for that node. Let the two ET's involving the nodes a, b_i, b_j , and c_{ij} (there are two such ET's because of the top-to-bottom symmetry of δ) have size $\beta_{ij} = \beta_{ji}$. Let the two ET's involving the nodes b_i, c_{ik}, c_{il}, b_j (once again, there are two such ET's because of the top-to-bottom symmetry) have size $\alpha_{ij} = \alpha_{ji}$. There are 6 distinct values of β_{ij} and 6 distinct values of α_{ij} . For the 24 ET's so defined to sum to δ , it is necessary and sufficient that β_{ij} and α_{ij} satisfy

$$\begin{aligned} \sum_{i < j} \beta_{ij} &= a \\ b_i + \beta_{ij} + \beta_{ik} + \beta_{il} &= \alpha_{ij} + \alpha_{ik} + \alpha_{il} \quad \forall i \neq j \neq k \neq l \\ -c_{ij} + \beta_{ij} + \beta_{kl} &= \alpha_{ik} + \alpha_{il} + \alpha_{jk} + \alpha_{jl} \quad \forall i \neq j \neq k \neq l. \end{aligned} \quad (23)$$

These three (sets of) equations ensure that each of the (sets of) nodes labeled a, b_i , and c_{ij} , respectively, in Figure 3 is transformed from its values under the distribution f to its values under

the distribution g by the sequence of ET's just defined. The equations (23) can be rearranged to

$$\begin{aligned}\sum_{i < j} \beta_{ij} &= a \\ b_i + \beta_{ij} + \beta_{ik} + \beta_{il} &= \alpha_{ij} + \alpha_{ik} + \alpha_{il} \quad \forall i \neq j \neq k \neq l \\ 2\alpha_{ij} + 2\beta_{kl} &= a + b_i + b_j + c_{ij} \quad \forall i \neq j \neq k \neq l.\end{aligned}\tag{24}$$

Noting the similarity between the equations (24) and the equations (14) for the three-dimensional cube, set

$$\begin{aligned}\beta_{kl} &= \frac{a(a + b_i + b_j + c_{ij})}{4a + \sum_{i=1}^4 b_i}, \quad \forall i \neq j \neq k \neq l \\ \alpha_{ij} &= \frac{(2a + \sum_{i=1}^4 b_i)(a + b_i + b_j + c_{ij})}{2(4a + \sum_{i=1}^4 b_i)}, \quad \forall i \neq j.\end{aligned}\tag{25}$$

Recalling that top-to-bottom symmetry of δ implies that $c_{12} = c_{34}$, $c_{13} = c_{24}$, and $c_{23} = c_{14}$, it is easily checked that with these choices for β_{kl} and α_{ij} , equations (24) are satisfied. Furthermore, if $a \geq 0$, $2a + \sum_{i=1}^4 b_i \geq 0$ and for all $i \neq j$, $a + b_i + b_j + c_{ij} \geq 0$, then β_{kl} and α_{ij} as defined in (25) are all nonnegative. Thus, condition (iv) in the proposition implies the existence of a sequence of nonnegative ET's that sum to $\delta = g - f$. Since each ET raises the expectation of any supermodular function, it follows that $g \succeq_{SPM} f$. By Theorem 2, it then follows that conditions i)-iv) are all equivalent.

E Proof of Claim 1

The four agents are denoted A, B, C, and D, and we also use these letters to denote the action choice, which is either 0 or 1, of the corresponding agent.

We first prove Part ii) of the claim, that the distribution of actions in the complete network displays greater weak association than the distribution of actions in the diamond (or line) network. First consider the case of $r(A, B)$ and $s(C, D)$. (The case of $r(A, D)$ and $s(B, C)$ is symmetric.) Each function r and s takes four possible values. By bilinearity of the covariance, we can scale these functions and translate them without affecting inequalities. Without loss of generality, thus, set $r_{00} = s_{00} = 0$ and $r_{11} = s_{11} = 1$.³⁰ Since r, s must be increasing, the remaining four values r_{01} , r_{10} , s_{10} and s_{01} may take any values between 0 and 1.

Using the above values, it is easy to show, given the probability distributions induced by the

³⁰If either $r_{11} = r_{00}$ or $s_{11} = s_{00}$, then the claim is trivially true.

complete and diamond networks, as shown in Figure 2, that the difference in covariances satisfies

$$\begin{aligned} & Cov^{comp}[r(A, B), s(C, D)] - Cov^{diam}[r(A, B), s(C, D)] \\ &= \frac{1}{64} \{6 - 2(r_{10} + r_{01} + s_{01} + s_{10}) - 2(r_{10}s_{10} + r_{01}s_{01}) + 6(r_{01}s_{10} + r_{10}s_{01})\}. \end{aligned}$$

This expression has the sign of

$$2 + 2(r_{01} - r_{10})(s_{10} - s_{01}) + 2(1 - r_{01})(1 - s_{10}) + 2(1 - r_{10})(1 - s_{01}),$$

which is nonnegative, since the second term is bounded by 2 in absolute value, and the third and fourth terms are nonnegative.

Now consider the case of $r(A, C)$ and $s(B, D)$. Since A, C and B, D are independent under the diamond distribution, the covariance for that distribution is zero, and we need only show that $r(A, C)$ and $s(B, D)$ have non-negative covariance for the distribution under the complete network. Using the same normalization and notation as above, we can compute, using Figure 2,

$$\begin{aligned} & Cov^{comp}[r(A, C), s(B, D)] = \\ & \frac{1}{64} \{20 + 4(r_{10} + r_{01})(s_{10} + s_{01}) - (3 + r_{10} + r_{01})(3 + s_{10} + s_{01})\}. \end{aligned}$$

This expression has the sign of

$$5 + 3(1 - r_{01})(1 - s_{10}) + 3(1 - r_{10})(1 - s_{01}) + 3r_{10}s_{10} + 3r_{01}s_{01},$$

which is clearly nonnegative.

Finally, consider the case of $r(D)$ and $s(A, B, C)$. (For partitions of the set of agents into a set of 3 agents and a set of 1, the problem is symmetric under all permutations of the agents.) Again, without loss of generality, suppose that $r(0) = 0$ and $r(1) = 1$. Then,

$$E^{comp}[r(D), s(A, B, C)] = Pr^{comp}(D = 1)E^{comp}[s(A, B, C)|D = 1].$$

Therefore,

$$Cov^{comp}(r(D), s(A, B, C)) = Pr^{comp}(D = 1) \{E^{comp}[s(A, B, C)|D = 1] - E^{comp}[s(A, B, C)]\},$$

with a similar expression for the diamond distribution. Therefore, since the complete and diamond distributions have equal marginals, $Pr^{comp}(D = 1) = Pr^{diam}(D = 1)$, and it suffices to show that for any increasing function s ,

$$E^{comp}[s(A, B, C)|D = 1] + E^{diam}[s(A, B, C)] \geq E^{comp}[s(A, B, C)] + E^{diam}[s(A, B, C)|D = 1]. \quad (26)$$

To show (26), we sum the two probability distributions on the left-hand side (the conditional distribution for the complete network and the unconditional one for the diamond) and, separately, sum those on the right-hand side (the unconditional one for the complete network and the conditional one for the diamond). Then we show that the difference, (A, B, C) , between these sums makes a nonnegative scalar product with any increasing function s . We compute (in units of $\frac{1}{64}$)

$$\begin{aligned}(1, 1, 1) &= 14, & (1, 1, 0) &= -6, & (1, 0, 1) &= 2, & (0, 1, 1) &= -6, & (1, 0, 0) &= 6, \\ & & (0, 1, 0) &= -2, & (0, 0, 1) &= 6, & (0, 0, 0) &= -14.\end{aligned}$$

Observe that i) $(1, 1, 1) > 0$; $(1, 1, 0) < 0$, $(0, 1, 1) < 0$, and $(0, 1, 0) < 0$; and $(1, 1, 1) + (1, 1, 0) + (0, 1, 1) + (0, 1, 0) = 0$. Similarly, observe that ii) $(1, 0, 1) > 0$, $(1, 0, 0) > 0$, and $(0, 0, 1) > 0$; $(0, 0, 0) < 0$; and $(1, 0, 1) + (1, 0, 0) + (0, 0, 1) + (0, 0, 0) = 0$. It follows from i) and ii) that for any increasing function s on $\{0, 1\}^3$, $\sum s \geq 0$, and therefore (26) holds. This completes the proof that the distribution under the complete network displays greater weak association than the distribution under the diamond (or line) network.

Part i): In the benchmark case when all agents randomize independently and symmetrically, the covariance of r and s is always zero since r and s are defined on disjoint subsets of agents. Thus we need only confirm that for all increasing functions r and s , $Cov(r, s) \geq 0$ for the complete, diamond (and line), and star networks. Moreover, given the proof of Part ii) (above), once we show that the distribution under the diamond network displays greater weak association than the independent benchmark, it will follow that the same is true for the complete network.

For the diamond network, we have already observed that A, C and B, D are independent, and hence $Cov(r(A, C), s(B, D)) = 0$. Consider therefore the case of $r(A, B)$ and $s(C, D)$. (The case of $r(A, D)$ and $s(B, C)$ is symmetric.) Using the same normalization and notation as above, we can compute, using Figure 2,

$$\begin{aligned}Cov^{diam}[r(A, B), s(C, D)] &= \\ \frac{1}{64} \{ &9 + 3(r_{10} + r_{01} + s_{01} + s_{10}) + 9(r_{10}s_{10} + r_{01}s_{01}) + (r_{01}s_{01} + r_{10}s_{10}) - 4(1 + r_{10} + r_{01})(1 + s_{10} + s_{01}) \}.\end{aligned}$$

This expression has the sign of

$$3 + 3(r_{10} - r_{01})(s_{10} - s_{01}) + r_{10}s_{10} + r_{01}s_{01} + (1 - r_{01})(1 - s_{01}) + (1 - r_{10})(1 - s_{10}),$$

which is nonnegative, since the second term is bounded by 3 in absolute value, and all subsequent terms are nonnegative.

Now consider the case of $r(D)$ and $s(A, B, C)$. (All permutations of the agents will leave $Cov(r, s)$ unchanged.) Normalizing $r(0) = 0$ and $r(1) = 1$, we have

$$Cov^{diam}(r(D), s(A, B, C)) = Pr^{diam}(D = 1) \left\{ E^{diam}[s(A, B, C)|D = 1] - E^{diam}[s(A, B, C)] \right\},$$

so it suffices to show that

$$E^{diam}[s(A, B, C)|D = 1] - E^{diam}[s(A, B, C)] \geq 0 \quad (27)$$

for all increasing functions s . The difference, (A, B, C) , between the conditional and unconditional distributions on the left-hand side of (27) satisfies (again using units of $\frac{1}{64}$)

$$\begin{aligned} (1, 1, 1) &= 6, & (1, 1, 0) &= 2, & (1, 0, 1) &= -6, & (0, 1, 1) &= 2, & (1, 0, 0) &= -2, \\ & & (0, 1, 0) &= 6, & (0, 0, 1) &= -2, & (0, 0, 0) &= -6. \end{aligned}$$

Since for each $(A, C) \in \{0, 1\}^2$, $(A, 1, C) > 0$ and $(A, 1, C) + (A, 0, C) = 0$, it follows that for any increasing function s on $\{0, 1\}^3$, $\sum s \geq 0$, and therefore (27) holds. This completes the proof that the distribution under the diamond network displays greater weak association than the independent benchmark.

We now turn to the star network and consider the case of $r(A, B)$ and $s(C, D)$. (All permutations of the agents will leave $Cov(r, s)$ unchanged.) Using the same normalization and notation as above, we can use Figure 2 to compute

$$Cov^{star}[r(A, B), s(C, D)] = \frac{1}{4}(1 + r_{01}) - \frac{1}{8}(1 + r_{01} + r_{10}) = \frac{1}{8}(1 + (r_{01} - r_{10})),$$

which is nonnegative.

For the star network, A is independent of B, C, D , so $Cov(r(A), s(B, C, D)) = 0$. Now consider $r(D)$ and $s(A, B, C)$. (All remaining permutations of the agents will leave $Cov(r, s)$ unchanged.) Following the logic used for the diamond network above, it suffices to show that

$$E^{star}[s(A, B, C)|D = 1] - E^{star}[s(A, B, C)] \geq 0 \quad (28)$$

for all increasing s . The difference, (A, B, C) , between the conditional and unconditional distributions on the left-hand side of (28) satisfies (again using units of $\frac{1}{64}$)

$$\begin{aligned} (1, 1, 1) &= (0, 1, 1) = 16, & (0, 0, 0) &= (1, 0, 0) = -16 \\ (1, 1, 0) &= (1, 0, 1) = (0, 1, 0) = (0, 0, 1) = 0. \end{aligned}$$

It is easy to see for this that for any increasing function s on $\{0, 1\}^3$, $\sum s \geq 0$, and therefore (28) holds. Therefore the distribution under the star network displays greater weak association than the independent benchmark.

Part iii) of Claim 1 is proved in the text following the statement of the claim.

F Proofs of Propositions 9 and 10

We first prove Proposition 10 and then use this proposition to prove Proposition 9.

Proof of Proposition 10: We proceed in three steps. Step 1 shows that the conditions in iii) imply that $g^{symm} \succeq_{SPM} f^{symm}$ and hence, by Proposition 1, that $g \succeq_{SSPM} f$. Step 2 is simply to note that it follows from the definitions of the symmetric supermodular and symmetric convex-modular orderings that $g \succeq_{SSPM} f$ implies $g \succeq_{SCXMOD} f$. Step 3 shows that $g \succeq_{SCXMOD} f$ implies the conditions in iii).

Step 1: For convenience, we let the values taken by δ^s , the symmetrized version of $\delta = g - f$, on $L = \{0, 1, 2\}^3$ be as represented in Figure 4. As in the proof of Proposition 7, we will abuse notation slightly and let the value of δ^s at a given node in Figure 4 also serve as the label for that node.

The condition that g^{symm} and f^{symm} have identical marginal distributions is equivalent to

$$a + 2b + c + 2d + 2e + f = b + 2c + g + 2e + 2h + i = d + 2e + h + 2f + 2i + j = 0. \quad (29)$$

We now decompose δ^s into 36 elementary transformations (ET's) of the form defined in (3), and we show that the conditions in iii) of Proposition 10 are sufficient for the existence of a decomposition of δ^s in which all ET's have nonnegative sign. We construct a sequence of ET's simultaneously from the "top" of L , i.e., the node $(2, 2, 2)$, and the "bottom", i.e., the node $(0, 0, 0)$, since the structure of L (though not the actual values of δ^s) is similar viewed from the top and from the bottom. We proceed in several stages, representing progression from the top and bottom of L towards its center.

Since there does not exist a unique decomposition of δ^s into a sequence of ET's, at some stages below we need to introduce undetermined variables to describe the sizes of the ET's. After describing the sequence of ET's, we show how to assign values to these undetermined variables to identify sufficient (and also necessary) conditions for the sizes of all ET's in the sequence to be nonnegative.

First stage: Assign to the 3 ET's involving nodes (a, b, b, c) size $a/3$ each. Similarly, assign to the

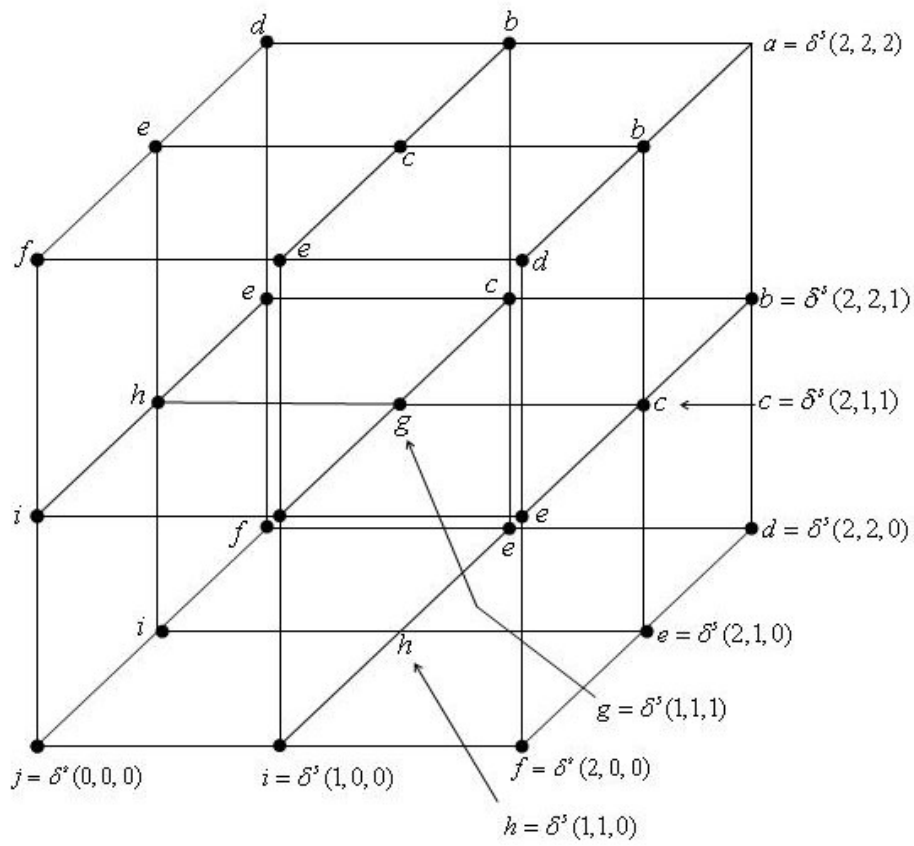


Figure 4: $\delta^s \equiv g^{\text{symm}} - f^{\text{symm}}$
for $L = \{0,1,2\}^3$

3 ET's involving nodes (j, i, i, h) size $j/3$ each. This guarantees that the sum of these ET's will match δ^s at $(2, 2, 2)$ and $(0, 0, 0)$.

Second stage:

- For each node b , assign to the 2 ET's involving nodes (b, c, d, e) size λ_b each, and assign to the 1 ET involving nodes (b, c, c, g) size $b + 2a/3 - 2\lambda_b$ each.
- For each node i , assign to the 2 ET's involving nodes (i, h, f, e) size λ_i each, and assign to the 1 ET involving nodes (i, h, h, g) size $i + 2j/3 - 2\lambda_i$ each.

λ_b and λ_i are the undetermined variables.

Third stage:

- For each node d , assign to the 1 ET involving nodes (d, e, e, h) size $d + 2\lambda_b$.
- For each node f , assign to the 1 ET involving nodes (f, e, e, c) size $f + 2\lambda_i$.

Fourth stage: For each node c , assign to the 2 ET's involving nodes (c, e, g, h) size $a + 2b + c + d + e - (\lambda_b + \lambda_i)$ each.

It is easily checked, using the equations (29) corresponding to identical marginal distributions for g^{symm} and f^{symm} , that the sequence of ET's defined above sums to δ^s .

Set

$$\lambda_b = \max\{0, -d/2\} \geq 0 \quad \text{and} \quad \lambda_i = \max\{0, -f/2\} \geq 0.$$

Then the following conditions are sufficient to ensure that every ET in the sequence defined above has nonnegative size:

1. $a \geq 0$;
2. $j \geq 0$;
3. $2a + 3b + 3 \min\{0, d\} \geq 0$;
4. $2j + 3i + 3 \min\{0, f\} \geq 0$;
5. $2a + 4b + 2c + 2d + 2e + \min\{0, d\} + \min\{0, f\} \geq 0$.

These 5 conditions correspond to the 5 conditions on δ^s listed in iii) of Proposition 10. Hence iii) implies $g^{symm} \succeq_{SPM} f^{symm}$ and therefore, by Proposition 1, $g \succeq_{SSPM} f$, which is i).

Step 3: Since for $k \in \{0, 1, 2\}$, the symmetric functions $\bar{w}(z_1, z_2, z_3) = \sum_{i=1}^3 I_{\{z_i \geq k\}}$ and $\underline{w}(z_1, z_2, z_3) = -\sum_{i=1}^3 I_{\{z_i \geq k\}}$ are both convex-modular, $g \succeq_{SCXMOD} f$ implies that for all $k \in \{0, 1, 2\}$, $\sum_{i=1}^3 Pr(Y_i = k) = \sum_{i=1}^3 Pr(X_i = k)$, which in turn implies that g^{symm} and f^{symm} have identical marginal distributions. Hence equations (29) hold.

The 5 conditions above can be expanded and simplified, using (29), to yield:

1. $a \geq 0$;
2. $j \geq 0$;
3. $2a + 3b \geq 0$;
4. $2a + 3b + 3d \geq 0$;
5. $2j + 3i \geq 0$;
6. $2j + 3i + 3f \geq 0$;
7. $3a + 6b + 3c \geq 0$;
8. $3j + 6i + 3h \geq 0$;
9. $3a + 6b + 3c + 3d \geq 0$;
10. $6a + 12b + 6c + 6d + 6e \geq 0$.

For each of these 10 inequalities, we list below a symmetric $w \in \mathcal{CM}^*$ such that the correspondingly numbered inequality is equivalent to $Ew(Y) \geq Ew(X)$. (Recall that \mathcal{CM}^* was defined in Section 2.4 as the set of nonnegative weighted sums of convex-modular functions.) This establishes that $g \succeq_{SCXMOD} f$ implies the conditions in iii). The 10 symmetric functions are:

1. $w^1(z) = \max\{0, (\sum_{i=1}^3 I_{\{z_i \geq 2\}}) - 2\}$;
2. $w^2(z) = \max\{0, 1 - (\sum_{i=1}^3 I_{\{z_i \geq 1\}})\}$;
3. $w^3(z) = 2 \max\{0, [\sum_{i=1}^3 (I_{\{z_i=2\}} + \frac{1}{2}I_{\{z_i=1\}})] - 2\}$;
4. $w^4(z) = \max\{0, (\sum_{i=1}^3 I_{\{z_i \geq 2\}}) - 1\}$;

$$5. \ w^5(z) = 2 \max\{0, 1 - [\sum_{i=1}^3 (I_{\{z_i=2\}} + \frac{1}{2} I_{\{z_i=1\}})]\};$$

$$6. \ w^6(z) = \max\{0, 2 - (\sum_{i=1}^3 I_{\{z_i \geq 1\}})\};$$

$$7. \ w^7(z) = 4 \max\{0, [\sum_{i=1}^3 (I_{\{z_i=2\}} + \frac{3}{4} I_{\{z_i=1\}})] - 2.25\};$$

$$8. \ w^8(z) = 4 \max\{0, .75 - [\sum_{i=1}^3 (I_{\{z_i=2\}} + \frac{1}{4} I_{\{z_i=1\}})]\};$$

$$9. \ w^9(z) = 2 \max\{0, [\sum_{i=1}^3 (I_{\{z_i=2\}} + 1$$

The proof of Proposition 10 showed that the 5 inequalities in iii) of that proposition can be expanded, and simplified using the identity of the marginals of f^{symm} and g^{symm} , to the 10 numbered inequalities above involving a, b, c , etc. Observe that only one of these inequalities, the tenth one, involves the variable e , which as Figure 4 shows represents $\delta^s(2, 1, 0) = g^{symm}(2, 1, 0) - f^{symm}(2, 1, 0)$.

Since $f^{symm}(z) = 0$ for any z which is not a permutation of $(2, 1, 0)$, it follows that $\delta^s \equiv g^{symm} - f^{symm} \geq 0$ for any z which is not a permutation of $(2, 1, 0)$. Therefore, for any tournament we consider, all of the variables in Figure 4 other than e are nonnegative, and hence the first nine inequalities above are automatically satisfied.

To verify that the tenth inequality is satisfied, for any tournament, we parameterize the 3×3 bistochastic matrix T summarizing the tournament as follows:

$$\begin{pmatrix} t & 1-t-r & r \\ 1-t-l & t+r+l+b-1 & 1-r-b \\ l & 1-l-b & b \end{pmatrix}.$$

Using this parameterization to compute g^{symm} , and thence $g^{symm} - f^{symm}$, we find that

$$6a + 12b + 6c + 6d + 6e \geq 0 \Leftrightarrow F(t, l, b, r) \geq 0,$$

$$\text{where } F(t, l, b, r) = 2rt + 2bl + lr + bt + 1 - r - t - b.$$

We need to show that $F(t, l, b, r) \geq 0$ for all feasible 3×3 bistochastic matrices T , where the feasibility constraints are

$$\begin{aligned} t + r &\leq 1; & t + l &\leq 1; & r + b &\leq 1; & l + b &\leq 1; \\ t &\geq 0; & r &\geq 0; & l &\geq 0; & b &\geq 0; & t + r + l + b &\geq 1. \end{aligned} \quad (30)$$

There is a unique stationary point of F , $t = l = r = b = \frac{1}{3}$, and at this point, $F = \frac{1}{3}$, but this point is not a local minimum of F , since $F(1, 0, b, 0) = 0$ for any b . Therefore, since F is continuous and differentiable, the minimum of F is attained on the boundary of the feasible set defined by (30). It is a straightforward exercise to confirm that everywhere along the boundary of the feasible set, $F(t, l, b, r) \geq 0$. This establishes that for any tournament, f and the corresponding g satisfy the conditions in iii) of Proposition 10, and therefore $g \succeq_{SSPM} f$ and $g \succeq_{SCXMOD} f$. ■