

Meyer, Margaret; Strulovici, Bruno

Working Paper

Increasing interdependence of multivariate distributions

Discussion Paper, No. 1523

Provided in Cooperation with:

Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Meyer, Margaret; Strulovici, Bruno (2011) : Increasing interdependence of multivariate distributions, Discussion Paper, No. 1523, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:

<https://hdl.handle.net/10419/59619>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Northwestern University

2001 Sheridan Road 580 Leverone Hall Evanston, IL 60208-2014 USA

Discussion Paper #1523
March 1, 2011

“Increasing Interdependence of Multivariate Distributions”

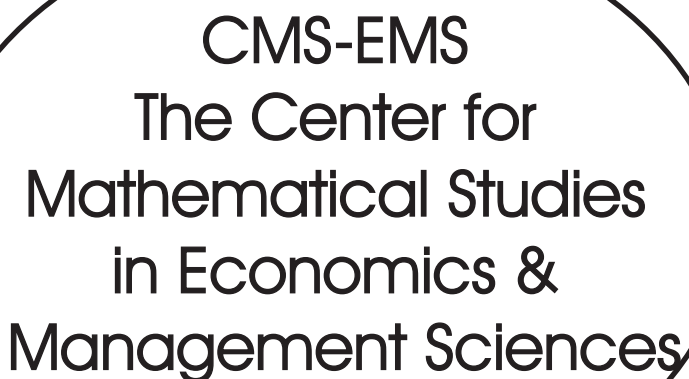
Key words: dependence ordering; stochastic orders;
supermodularity; weak association; concordance

JEL classification: D63, D81, G11, G22

Margaret Meyer
Nuffield College and Oxford University

Bruno Strulovici
Northwestern University

www.kellogg.northwestern.edu/research/math



CMS-EMS
The Center for
Mathematical Studies
in Economics &
Management Sciences

Increasing Interdependence of Multivariate Distributions

Margaret Meyer

Bruno Strulovici*

March 1, 2011

Abstract

Orderings of interdependence among random variables are useful in many economic contexts, for example, in assessing ex post inequality under uncertainty; in comparing multidimensional inequality; in valuing portfolios of assets or insurance policies; and in assessing systemic risk. We explore five orderings of interdependence for multivariate distributions: greater weak association, the supermodular ordering, the convex-modular ordering, the dispersion ordering, and the concordance ordering. 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 equivalences among the orderings.

Keywords: dependence ordering; stochastic orders; supermodularity; weak association; concordance

JEL classification codes: D63, D81, G11, G22

*Meyer: Nuffield College and Dept. of Economics, Oxford University, margaret.meyer@nuffield.ox.ac.uk; Strulovici: Dept. of Economics, Northwestern University, b-strulovici@northwestern.edu. We thank Larry Epstein, Ian Jewitt, Paul Klemperer, and Jeremy Large, as well as seminar audiences at Oxford, ESSET Gerzensee, and the Paris Tech-JET Conference on Inequality and Risk. We are very grateful to the referees and the Associate Editors for their very helpful suggestions.

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 and establishes the relationships among the orderings, both in general as well as in the important special case when all random variables are binary.

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 [40]; Ben-Porath et al [9]; Gajdos and Maurin [25]; Kroll and Davidovitz [36]; Adler and Sanchirico [2]; Chew and Sagi [17]). (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 application concerns comparisons of inequality when individual-level data are available on multiple attributes such as income, health, and education (Atkinson and Bourguignon [6]; Bourguignon and Chakravarty [11]; Atkinson [5]). As long as the function aggregating the different attributes into an overall measure of welfare or deprivation for individuals is not additively separable across attributes, comparisons of multidimensional inequality will necessitate comparisons of the degree of interdependence among the attributes.¹

Interdependence orderings are also useful for comparing the efficiency of matching mechanisms. In many matching contexts, perfectly assortative matching would be efficient, and it 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 [38]), or when search is costly (Shimer and Smith [50]), or when signaling is constrained by borrowing constraints (Fernandez and Gali [24]), perfectly assortative matching will generally not arise. Fernandez and Gali [24] and Meyer and Rothschild [41] apply bivariate dependence orderings to compare the performance of different matching institutions.²

¹There are important formal differences between these two welfare-economic applications. In the measurement of *ex post* inequality under uncertainty, the number of dimensions represents the number of individuals in the group, whereas in the measurement of multidimensional inequality, the number of dimensions represents the number of attributes.

²In a related application, Prat [46] explores how the composition of employee teams affects interdependence in

Orderings of interdependence can also be applied to make comparisons across decision-making groups of the degree of alignment in the preferences of group members. Boland and Proschan [10] and Baldiga and Green [7] focus on assessments of the similarity among different individuals' ordinal preferences over alternatives. In a strategic model of consensus building within a committee, Caillaud and Tirole [13] study how the degree of interdependence of members' ex ante uncertain payoffs from a proposal affects the proposer's optimal persuasion strategy.

In theoretical or experimental studies of information transmission in networks (Calvo-Armengol and Jackson [14], Choi, Gale, and Kariv [19]), interdependence orderings can be used to assess how the degree of concordance of individuals' ex ante uncertain actions or outcomes varies with time and with social structure, as represented by the network.

Finally, 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 [45]; Denuit et al [21]). Financial economists and macroeconomists are, moreover, increasingly interested in measures and comparisons of “systemic risk” in financial and economic systems that reflect the interdependence in the returns of different financial institutions, sectors, or regions (Hennessey and Lapan [29]; Adrian and Brunnermeier [3]; Acharya [1]; Beale et al [8]).

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 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 w that are supermodular. Supermodularity (see Topkis [52] 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: For a supermodular function (of two or more arguments), the effect of an increase in the value of any argument is larger, the larger are the values of the other arguments.

(ex ante random) decisions of team members and shows how properties of the the production function translate into preferences over team composition.

It has been shown, in the economics literature by Levy and Parousch [37], Epstein and Tanny [22], and Atkinson and Bourguignon [6] and in the statistics literature by Tchen [51], 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 w .

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.

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 marginals 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.³ 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.⁴

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

³For 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 [35] and Müller and Stoyan [45] for more details.

⁴Unfortunately, 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 [26, fn. 12].

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 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. Three of these have received some attention in the statistics and economics literatures, while two of them are new.

We begin with an ordering inspired by the dependence concept of “association”, a concept proposed by Esary, Proschan, and Walkup [23] and widely used in statistics.⁵ ⁶ Since the definition of association cannot be meaningfully reversed to yield a concept of negative association, we turn to the concept of “weak association” and use it to define the ordering we term “greater weak association”. In the actuarial literature (see Denuit et al [21]), this ordering has been (somewhat confusingly) named the “correlation order”.

We then present the (multivariate) supermodular ordering, which is the natural multivariate generalization of the concept of supermodular dominance discussed above for bivariate distributions. Meyer and Mookherjee [40] and Meyer [39] contain early proposals to use it as a multivariate dependence ordering, in the context of comparisons of ex post inequality under uncertainty. In the statistical literature, the multivariate supermodular ordering was formalized by Shaked and

⁵Association, though a strong concept of dependence, is strictly weaker than “affiliation”, which is familiar to economists from the work of Milgrom and Weber [43], who used it to formalize positive interdependence among bidders’ valuations in auctions. De Castro [16] also notes that affiliation is strictly stronger than many other concepts of positive dependence.

⁶Though association has not been widely used in economics, in a pair of recent papers, Calvo-Armengol and Jackson [14, respectively 15] show that the employment statuses (respectively, wages) of individuals connected by a social network are positively dependent in the sense of association.

Shanthikumar [48].

In Section 2.4, we introduce a new dependence ordering, which we term the “convex-modular ordering”. Greater interdependence for Y than for X according to the convex-modular ordering corresponds to greater riskiness, in the sense of Rothschild and Stiglitz [47], 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”. The dispersion ordering is motivated by the idea that interdependence in random vectors can be compared by comparing the dispersion of the cdf’s of their order statistics, with lower dispersion representing greater interdependence. For the special case of random vectors with symmetric distributions, Shaked and Tong [49] proposed a dependence ordering along these lines, using the majorization ordering of vectors (Hardy, Littlewood, and Polya [28]) to compare dispersion of the cdf’s of the order statistics. The dispersion ordering we propose uses majorization in a similar way and has the advantage of being applicable to asymmetric distributions. We show that greater interdependence for Y than for X according to the dispersion ordering corresponds to greater riskiness, in the sense of Rothschild and Stiglitz [47], of the summary statistic $\sum_{i=1}^n I_{\{Y_i > s_i\}}$ compared to $\sum_{i=1}^n I_{\{X_i > s_i\}}$, for *any* point (s_1, \dots, s_n) in the support. A natural application of the dispersion ordering is to the assessment of systemic risk in financial systems, where systemic risk is greater the more “positively correlated” are bank failures. The summary statistic $\sum_{i=1}^n I_{\{Y_i > s_i\}}$ (respectively, $\sum_{i=1}^n I_{\{Y_i \leq s_i\}}$) represents the number of banks that would be solvent (respectively, insolvent) if bank i ’s failure threshold (which may not be known by regulators) were s_i , and the systemic cost function is a convex function of the number of insolvent banks. Importantly, and in contrast to the greater weak association, supermodular, and convex-modular orderings, it is straightforward to determine whether or not any two n -dimensional random vectors Y and X can be ranked according to the dispersion ordering; it is sufficient to check, for each point in the support, whether a set of $n + 1$ inequalities is satisfied.

Finally, we present the concordance ordering, which was formalized for multivariate distributions by Joe [34] and which combines the requirement of upper-orthant dominance with that of lower-orthant dominance.⁷ The concordance ordering shares with our dispersion ordering the advantage of being easily checkable; for any number of dimensions, it is sufficient to check, for each point in the support, whether two inequalities are satisfied.

⁷With more than two dimensions, even if Y and X have identical marginals, upper-orthant and lower-orthant dominance are no longer equivalent.

Section 3 presents our main results, which establish the relationships among the five orderings just described. Theorem 1 shows that for two dimensions, all five orderings are equivalent. Theorem 2, on the other hand, shows that for an arbitrary number of dimensions $n > 2$, the five orderings are strictly ranked.

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 orderings that arise in these special cases and ii) easily checkable and easily interpretable necessary and sufficient conditions for the orderings to hold. We offer a brief conclusion in Section 5.

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.

As argued by Joe [35, p.39], a natural desideratum for an ordering of interdependence is invariance to monotonic relabelings of the elements of the support for each component. Formally, we say that the ordering \succeq_O is invariant to monotonic relabelings of coordinates if

$$(Y_1, \dots, Y_n) \succeq_O (X_1, \dots, X_n) \implies (r_1(Y_1), \dots, r_n(Y_n)) \succeq_O (r_1(X_1), \dots, r_n(X_n)) \quad (1)$$

whenever $r_i : \mathbb{R} \rightarrow \mathbb{R}$, $i \in \{1, \dots, n\}$, are all nondecreasing. This invariance property captures the idea that dependence orderings should be ordinal in that they should compare the extent to which higher (respectively, lower) realizations in one dimension are accompanied by higher (respectively, lower) realizations in other dimensions, regardless of the cardinal scales used for each dimension. An example of monotonic relabeling of coordinates is a componentwise coarsening of the support, in which some adjacent elements in the support of each component are combined. For example,

starting with the lattice $L = \{0, 1, 2\}^n$, the nondecreasing transformations r_i with, for all i , $r_i(0) = 0$ and $r_i(1) = r_i(2) = 1$ coarsen the lattice to $\{0, 1\}^n$ by combining the realizations $\{1, 2\}$ in each dimension.

All of the interdependence orderings we consider in this paper will be invariant to monotonic relabelings of coordinates, as defined in (1). This is straightforward to confirm from their formal definitions below. Consistent with our definition of invariance in (1), in the formal definitions of the orderings, when we consider monotonic functions, we allow these functions to be nondecreasing rather than requiring them to be increasing. None of our results is affected by this choice, but proofs are simplified by the ability to use nondecreasing indicator functions (denoted by I).

While all of our interdependence orderings satisfy the invariance property in (1), some widely used orderings, such as the (bivariate) linear correlation coefficient, fail to satisfy this invariance and, a fortiori, are not invariant to a coarsening of the support. A higher value of the linear correlation coefficient for one bivariate empirical distribution than another is therefore not a robust indicator of greater dependence in the former distribution, since the ranking might be reversed by aggregation of the data. To see this, for $L = \{l, m, h\}^2$, where $l < m < h$, let (Y_1, Y_2) have distribution g , where $g(l, m) = g(m, l) = g(m, m) = g(h, h) = \frac{1}{4}$, and let (X_1, X_2) have distribution f , where $f(l, l) = f(m, m) = f(m, h) = f(h, m) = \frac{1}{4}$. Then $\text{corr}(Y_1, Y_2) > (<) \text{corr}(X_1, X_2)$ if $\frac{(l+h)}{2} > (<) m$. This in turn implies that if the support L is coarsened by combining the realizations l and m in each dimension, then $\text{corr}(Y_1, Y_2) > \text{corr}(X_1, X_2)$, while if m and h are combined, then $\text{corr}(Y_1, Y_2) < \text{corr}(X_1, X_2)$.

Given that the orderings we consider below all satisfy the invariance property in (1), 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 $\{0, 1, \dots, l_i - 1\}$.

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 “O”. We will use the phrases “ Y dominates X according to the ordering O” and “ g dominates f according to the ordering O” interchangeably, and we denote the former by $Y \succeq_O X$ and the latter by $g \succeq_O f$.

Given two distributions g and f on L , define the difference between them by

$$\delta \equiv g - f. \tag{2}$$

Most (though not all) of the orderings we consider are “difference-based orderings”, in that whether or not two distributions g and f satisfy the ordering depends only on the corresponding δ . We will frequently exploit this convenient property of difference-based orderings.⁸

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 [23] defined association as follows:

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

A concept of negative association defined by reversing the inequality in the definition above 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 an important motivation for studying a less stringent concept, “weak association”, defined by Burton, Dabrowski, and Dehling [12], along with its negative counterpart, “negative association”, defined by Joag-Dev and Proschan [33]. In contrast to “association”, which allows both r and s to depend on the entire vector Y , weak association and negative association restrict them to depend on *disjoint* components of Y .⁹

⁸Some 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.

⁹It is clear from the definitions that association is a stronger concept than weak association. Hu, Müller, and Scarsini [30] have shown that it is strictly stronger, even for two dimensions.

DEFINITION 2 (*Weak Association and Negative Association*) A random vector Y with support $L = \times_{i=1}^n L_i$ is **weakly associated (respectively, 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 $s : \times_{j \in B} L_j \rightarrow \mathbb{R}$, $\text{Cov}(r(Y_i, i \in A), s(Y_j, j \in B)) \geq 0$ (respectively, $\text{Cov}(r(Y_i, i \in A), s(Y_j, j \in B)) \leq 0$).

We can now define an interdependence ordering corresponding to weak association and negative association as follows.¹⁰

DEFINITION 3 (*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 $s : \times_{j \in B} L_j \rightarrow \mathbb{R}$,

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

Note that Definition 3 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 $n \geq 3$, greater weak association is not a difference-based ordering, though Theorem 1 in Section 3 implies that for the special case of $n = 2$, it is.¹¹

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.¹² 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 , $\text{Cov}(r(Y), s(Y)) \geq \text{Cov}(r(X), s(X))$. 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.¹³

¹⁰In the insurance literature, this ordering has been termed the “correlation order”—see Denuit et al. [21].

¹¹To verify the claim for $n = 2$ directly, note that the only non-trivial partition of $\{1, 2\}$ into disjoint subsets is $\{1\}, \{2\}$. Then given Definition 3’s requirement that Y and X have identical marginals, $\text{Cov}(r(Y_1), s(Y_2)) - \text{Cov}(r(X_1), s(X_2)) = E[r(Y_1)s(Y_2)] - E[r(X_1)s(X_2)]$, and the sign of the right-hand side depends only on the difference between the distributions of Y and X .

¹²Similarly, (Y_1, \dots, Y_n) are negatively associated if and only if the independent counterpart of Y displays greater weak association than Y .

¹³For the support $\{0, 1, 2\}^2$, let $\text{Pr}(Y_1 = Y_2 = 0) = \text{Pr}(Y_1 = Y_2 = 1) = \text{Pr}(Y_1 = Y_2 = 2) = \frac{1}{4}$ and $\text{Pr}(Y_1 = 0, Y_2 = 2) = \text{Pr}(Y_1 = 2, Y_2 = 0) = \frac{1}{8}$. Then it can be checked that Y is associated. However, consider the nondecreasing functions $r(Y_1, Y_2) = I_{\{Y_1 \geq 1, Y_2 \geq 1\}}$ and $s(Y_1, Y_2) = I_{\{Y_1 = 2 \text{ or } Y_2 = 2\}}$. It is easy to confirm that $\text{Cov}(r(Y_1, Y_2), s(Y_1, Y_2))$ is strictly smaller for Y than for its independent counterpart.

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 [52]):

$$w(z + e_i + e_j) + w(z) \geq w(z + e_i) + w(z + e_j) \quad (3)$$

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).¹⁴

DEFINITION 4 (*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 .*

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

$$\delta(z) = \delta(z + e_i + e_j) = -\delta(z + e_i) = -\delta(z + e_j) = \alpha \quad (4)$$

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 (3), a function w is supermodular if and only if $w \cdot \delta \geq 0$ for any δ of the form (4). Hence the class of supermodular functions is precisely the class for which the expectation is raised by any ET as defined in (4).

¹⁴For functions w defined on \mathbb{R}^n and twice differentiable, an equivalent characterization is: w is supermodular if and only if $\frac{\partial^2 w}{\partial z_i \partial z_j} \geq 0$ for all $z \in \mathbb{R}^n$ and all $i \neq j$.

Meyer and Strulovici [42] 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 [47] and Pigou-Dalton transfers in Atkinson [4] and Dasgupta, Sen, and Starrett [20].¹⁵

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 $\bar{w}(z) = I_{\{z_i \geq k\}}$ and $\underline{w}(z) = I_{\{z_i < k\}}$ are both supermodular. Therefore $g \succeq_{SPM} f$ implies that, for all $i \in \{1, \dots, n\}$ and any $k \in L_i$,

$$\begin{aligned} 0 \leq E[\bar{w}|g] - E[\bar{w}|f] &= \sum_{z: z_i \geq k} g(z) - \sum_{z: z_i \geq k} f(z) \\ \text{and } 0 \leq E[\underline{w}|g] - E[\underline{w}|f] &= \sum_{z: z_i < k} g(z) - \sum_{z: z_i < k} f(z), \end{aligned} \quad (5)$$

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

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, 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 $i \neq j$. Let θ denote a real function on a symmetric lattice L . 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 5 (*Symmetric Supermodular Ordering*) *Let the random vectors Y and X have distri-*

¹⁵Elementary transformations of bivariate distributions were also used by Epstein and Tanny [22] and Tchen [51] 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 [42].

butions 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 ,

$$\theta^{symm}(z) = \frac{1}{n!} \sum_{\sigma \in \Sigma(n)} \theta(\sigma(z)), \quad (6)$$

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

Proof. \implies : If for all symmetric supermodular w , $w \cdot g \geq w \cdot f$, then for all symmetric supermodular w , $w \cdot g^{symm} \geq w \cdot f^{symm}$. In turn, if for some symmetric supermodular w , $w \cdot g^{symm} \geq w \cdot f^{symm}$, then $\hat{w} \cdot g^{symm} \geq \hat{w} \cdot f^{symm}$ 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 , $w \cdot g^{symm} \geq w \cdot f^{symm}$ for all symmetric supermodular w implies that $g^{symm} \succeq_{SPM} f^{symm}$.

\impliedby : If for all supermodular w , $w \cdot g^{symm} \geq w \cdot f^{symm}$, then for all supermodular w , $w^{symm} \cdot g^{symm} \geq w^{symm} \cdot f^{symm}$. This is equivalent to $w^{symm} \cdot g^{symm} \geq w^{symm} \cdot f^{symm}$ for all symmetric supermodular w^{symm} . This in turn implies that for all symmetric supermodular w^{symm} , $w^{symm} \cdot g \geq w^{symm} \cdot f$.

■

Proposition 1 states that one can characterize the symmetric supermodular ordering in terms of the supermodular order applied to symmetric distributions.

The symmetric supermodular ordering has a very simple form for random vectors for which each component has a binary support $\{0, 1\}$, so the lattice $L = \{0, 1\}^n$. To state the result, first define, for a random vector Y with support $L = \{0, 1\}^n$,

$$c(Y) \equiv \sum_{i=1}^n I_{\{Y_i=1\}}.$$

The “count function” $c(Y)$ gives the number of components of Y for which the realization takes the value 1. Now recall the definition of the univariate convex ordering:

DEFINITION 6 (*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 $EZ = EV$. The convex ordering is equivalent to the ordering of greater riskiness studied by Rothschild and Stiglitz [47].

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

$$w(Y_1, \dots, Y_n) = \phi(c(Y_1, \dots, Y_n)), \quad (7)$$

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 (7) 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 used to compare interdependence in multivariate distributions 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, $\sum_{i=1}^n r_i(z_i)$. 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 (4). 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 .¹⁶ These observations motivate us to define the following difference-based ordering:

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. 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.

It is natural to define a symmetric counterpart of the convex-modular ordering. It is clear from our definition in (6) 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 , $g \succeq_{SCXMOD} f \iff g^{symm} \succeq_{CXMOD} f^{symm}$.*

¹⁶See Denuit et al [21] for more details.

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, suppose (Y_1, Y_2) and (X_1, X_2) both have symmetric distributions on $\{0, 1\}^2$, and that $Pr(Y_i = 1) = Pr(X_i = 1) = \frac{1}{2}$. Let Y_1 and Y_2 be perfectly positively dependent: the realizations $(0, 0)$ and $(1, 1)$ both have probability $\frac{1}{2}$. Let X_1 and X_2 be perfectly negatively dependent: the realizations $(0, 1)$ and $(1, 0)$ both have 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. By contrast, 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 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 [28]). If $a \prec b$, then the components of a are less dispersed than the components of b .

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 [49] 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,*

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

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 , i.e. the j^{th} smallest value

from $(Y_1^b, \dots, Y_n^b) \in \{0, 1\}^n$, and let $F_{Y_{(j)}^b}$ denote the distribution function of this order statistic. Then a condition equivalent to (8) is

$$(F_{Y_{(1)}^b}(0), \dots, F_{Y_{(n)}^b}(0)) \prec (F_{X_{(1)}^b}(0), \dots, F_{X_{(n)}^b}(0)) \quad \forall b = (b_0, \dots, b_0) \in L. \quad (9)$$

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 the extended lattice $\bar{L} = \times_{i=1}^n \bar{L}_i$, where $\bar{L}_i = \{-1, 0, 1, \dots, l_i - 1\}$. We now propose a new dependence ordering, which strengthens condition (9) by requiring that it hold for every vector $s = (s_1, s_2, \dots, s_n) \in \bar{L}$.

For a random vector Y , define its (componentwise) binary coarsening corresponding to the vector $s \in \bar{L}$, Y^s , by

$$Y^s \equiv (Y_1^s, \dots, Y_n^s) \in \{0, 1\}^n \quad \text{where} \quad Y_i^s = \begin{cases} 0 & \text{if } Y_i \leq s_i \\ 1 & \text{if } Y_i > s_i. \end{cases}$$

Let $(Y_{(1)}^s, \dots, Y_{(n)}^s) \in \{0, 1\}^n$ denote the vector of order statistics of Y^s , that is, $Y_{(j)}^s$ equals the j^{th} smallest value from (Y_1^s, \dots, Y_n^s) . Thus, $Y_{(j)}^s = 0$ if there are at least j values of $i \in \{1, \dots, n\}$ such that $Y_i \leq s_i$. Let $F_{Y_{(j)}^s}$ denote the distribution function of $Y_{(j)}^s$.¹⁷

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 $s \in \bar{L}$. The distribution of Y dominates that of X according to the **dispersion ordering**, written $Y \succeq_{DISP} X$, if for all $s \in \bar{L}$, 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,

$$(F_{Y_{(1)}^s}(0), \dots, F_{Y_{(n)}^s}(0)) \prec (F_{X_{(1)}^s}(0), \dots, F_{X_{(n)}^s}(0)) \quad \forall s \in \bar{L}. \quad (10)$$

The following proposition illuminates the appeal and convenience of the dispersion ordering by presenting some equivalent formulations. To state it, we define, for any $s \in \bar{L}$,

$$c^s(Y^s) = \sum_{i=1}^n I_{\{Y_i^s=1\}} = \sum_{i=1}^n I_{\{Y_i > s_i\}}, \quad (11)$$

¹⁷Considering 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 (8), or equivalently (9), 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 (8) are $(0, \dots, 0)$.

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:*

- i) For Y, X with support L , $Y \succeq_{DISP} X$;
- ii) For all $s \in \bar{L}$, $Y^s \succeq_{SSPM} X^s$.
- iii) For all $s \in \bar{L}$, $c^s(Y^s) \succeq_{CX} c^s(X^s)$.

Proof. 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 2.3.1. To show that ii) implies i), first note that since $Y_{(j)}^s$ is the j^{th} smallest value from (Y_1^s, \dots, Y_n^s) , $F_{Y_{(j)}^s}(0) = Pr(Y_{(j)}^s = 0) = E[I_{\{Y_{(j)}^s=0\}}]$ is weakly decreasing in j . Hence the majorization condition (10) can be rewritten as

$$E \left[\sum_{j=1}^k I_{\{Y_{(j)}^s=0\}} \right] \leq E \left[\sum_{j=1}^k I_{\{X_{(j)}^s=0\}} \right] \quad \forall k \in \{1, \dots, n\}, \forall s \in \bar{L}, \quad (12)$$

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

$$\begin{aligned} \sum_{j=1}^k I_{\{Y_{(j)}^s=0\}} &= \sum_{j=1}^k I_{\{c^s(Y^s) < n-(j-1)\}} = \sum_{j=1}^k (1 - I_{\{c^s(Y^s) \geq n-(j-1)\}}) \\ &= k - \max\{c^s(Y^s) - (n - k), 0\}. \end{aligned} \quad (13)$$

Since for any $s \in \bar{L}$, $c^s(Y^s)$ 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 $k \in \{1, \dots, n\}$. Therefore, ii) implies that (12) holds for all $k \in \{1, \dots, n\}$. Setting $k = n$ in (13) gives

$$\sum_{j=1}^n I_{\{Y_{(j)}^s=0\}} = n - c^s(Y^s), \quad (14)$$

and since both $c^s(Y^s)$ and $-c^s(Y^s)$ are symmetric supermodular functions of Y^s , ii) implies that for $k = n$, (12) holds with equality, as required. Thus ii) implies i).

To show that i) implies iii), first note that for random variables Z and V with support $\{0, 1, \dots, n\}$, $Z \succeq_{CX} V$ if and only if $E(Z) = E(V)$ and, for all $a \in \{1, \dots, n-1\}$, $E[\max\{Z - a, 0\}] \geq E[\max\{V - a, 0\}]$.¹⁸ Given (14), the equality in (12) for $k = n$ implies that $E(c^s(Y^s)) = E(c^s(X^s))$.

¹⁸See, for example, Jewitt [32] and Hardy, Littlewood, and Polya [27].

Given (13), the weak inequality in (12) 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 $a \in \{1, \dots, n-1\}$. Hence for all $s \in \bar{L}$, $c^s(Y^s) \succeq_{CX} c^s(X^s)$. ■

Proposition 4 provides a simple interpretation of the dispersion ordering for random vectors Y and X . Consider any coarsening of the support $L = \times_{i=1}^n L_i$ into $\{0, 1\}^n$ generated by, for each dimension i , classifying all values $z_i \leq s_i$ as 0 and all values $z_i > s_i$ as 1, for some $s \in \bar{L}$. Then compare the distributions of the correspondingly coarsened random vectors Y^s and X^s according to the symmetric supermodular ordering, or equivalently, given Proposition 2, use the univariate convex ordering to compare the distributions of the random variables $c^s(Y^s)$ and $c^s(X^s)$ counting, respectively, the number of components of Y and of X such that $Y_i > s_i$ and $X_i > s_i$. $Y \succeq_{DISP} X$ if and only if, for any point s in the support, $\sum_{i=1}^n I_{\{Y_i > s_i\}}$ is riskier than $\sum_{i=1}^n I_{\{X_i > s_i\}}$ in the sense of Rothschild and Stiglitz [47].

A natural application of the dispersion ordering is to the assessment of systemic risk in financial systems, where the n dimensions represent banks and where a bank's financial health is often summarized by classifying it as either solvent or insolvent. The degree of interdependence in failures of financial institutions is of crucial importance because “as more banks fail in the same time period, the economic disruption tends to increase disproportionately” (Beale et al [8, p.1]). This disproportionate increase in cost, representing an aversion to positive interdependence, can be captured by a systemic cost function $\phi(k)$, where k is the number of banks that fail and $\phi : \{0, 1, \dots, n\} \rightarrow \mathbb{R}$ is convex. It is natural to want comparisons of expected systemic cost to be robust to uncertainty about the precise values of the failure thresholds for different banks. These considerations point towards using the dispersion ordering of the random vectors describing banks' returns to compare levels of systemic risk under different regulatory scenarios.

It is apparent from Proposition 4 that the dispersion ordering is a difference-based ordering. Furthermore, if $Y \succeq_{DISP} X$, then Y and X have identical univariate marginals. To see this, first take $s \in \bar{L}$ to be a vector with i^{th} component equal to $a \in L_i$ and all other components equal to -1. Then for any Y and X and for any $j \neq i$, $Y_j^s = X_j^s = 1$. Hence for $k = 1$, (12) can be rewritten as

$$E[I_{\{Y_{(1)}^s=0\}}] \leq E[I_{\{X_{(1)}^s=0\}}] \iff Pr(Y_i \leq a) \leq Pr(X_i \leq a).$$

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 for any Y and X and for any $j \neq i$, $Y_j^s = X_j^s = 0$. Hence it follows from the equality in (12) for $k = n$ and the inequality for $k = n - 1$ that

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

Therefore, for all i and for all $a \in L_i$, $Pr(Y_i \leq a) = Pr(X_i \leq a)$.

An important advantage of the dispersion ordering, relative to the greater weak association, super-modular, and convex-modular orderings, is that it is straightforward to check whether or not any two n -dimensional random vectors Y and X can be ranked according to the dispersion ordering. As shown in the proof of Proposition 4, checking whether or not, for a given $s \in \bar{L}$, $c^s(Y^s) \succeq_{CX} c^s(X^s)$ is equivalent to checking whether or not a set of $n - 1$ inequalities and 1 equality hold. Hence, by Proposition 4, to check whether or not $Y \succeq_{DISP} X$, it is sufficient to check a set of $n+1$ inequalities for each $s \in \bar{L}$.¹⁹

2.6 The Concordance Ordering

Another intuitively appealing notion of greater interdependence, the concordance ordering, has been formalized for multivariate distributions by Joe [34].

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 that the concordance ordering is a difference-based ordering. As is well known, $Y \succeq_{CONC} X$ implies that Y and X have identical univariate marginals.²⁰

The concordance ordering shares with our dispersion ordering the advantage of being easily checkable; it is clear from the definition that for any number of dimensions, it is sufficient to check, for each point in the support, whether two inequalities are satisfied.²¹

¹⁹In practice, several of the inequalities generated as s varies will hold trivially or be redundant. For example, for $L = \{0, 1\}^2$, so $\bar{L} = \{-1, 0, 1\}^2$, the $3 \times 3 \times 3$ inequalities that in principle need to be checked reduce to only 5, corresponding to the requirements that i) (Y_1, Y_2) and (X_1, X_2) have identical marginals (4 inequalities) and ii) $Pr(Y_1 = 0, Y_2 = 0) \geq Pr(X_1 = 0, X_2 = 0)$.

²⁰To see this, first 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.

²¹As with the dispersion ordering, in practice some of the inequalities will hold trivially or be redundant. For $L = \{0, 1\}^2$, the $2 \times 2 \times 2$ inequalities that in principle need to be checked reduce to the same 5 as for the dispersion

3 Relationships among the Orderings

We are now in a position to present our main results, which establish the relationships among the orderings of interdependence defined in Section 2.

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

Proof. The equivalence for $n = 2$ between greater weak association, the supermodular ordering, and the concordance ordering is well known (see Meyer [39, Prop. 2] and Müller and Stoyan [45, Theorem 3.8.2] for references). We now prove that for $n = 2$ the convex-modular and dispersion orderings are equivalent to the other three orderings, by showing that $Y \succeq_{SPM} X \implies Y \succeq_{CXMOD} X \implies Y \succeq_{DISP} X \implies Y \succeq_{CONC} X$.

Since every convex-modular function is supermodular, the supermodular ordering implies the convex-modular ordering. The fact that the convex-modular ordering implies the dispersion ordering follows from the proof of Proposition 4 and the facts that the functions $\max\{c^s(Y) - (2 - k), 0\}$, for $k \in \{1, 2\}$, and $-c^s(Y)$ are both convex-modular. Finally, observe that by Proposition 4, if $(Y_1, Y_2) \succeq_{DISP} (X_1, X_2)$, then for all $s \in \bar{L}$, $Pr(c^s(Y^s) = 0) \geq Pr(c^s(X^s) = 0)$ and $Pr(c^s(Y^s) = 2) \geq Pr(c^s(X^s) = 2)$. The first of these inequalities is equivalent to $Pr(Y_1 \leq s_1, Y_2 \leq s_2) \geq Pr(X_1 \leq s_1, X_2 \leq s_2)$, while the second is equivalent to $Pr(Y_1 > s_1, Y_2 > s_2) \geq Pr(X_1 > s_1, X_2 > s_2)$. Hence $(Y_1, Y_2) \succeq_{DISP} (X_1, X_2)$ implies $(Y_1, Y_2) \succeq_{CONC} (X_1, X_2)$. ■

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. i) The proof that $Y \succeq_{GWA} X$ implies $Y \succeq_{SPM} X$ is in Appendix A. The following example (which for future reference we label Example 1) proves that this implication is strict: Let $L = \{0, 1\}^3$ and let X and Y have 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

ordering.

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 $Pr(c(Y) = 0) \geq Pr(c(X) = 0)$, hence $c(Y) \succeq_{CX} c(X)$ and $g \succeq_{SPM} f$. 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 $g \not\succeq_{GWA} f$ does not hold.

ii) Since every convex-modular function is supermodular, $Y \succeq_{SPM} X$ implies $Y \succeq_{CXMOD} X$. 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.

iii) That $Y \succeq_{CXMOD} X$ implies $Y \succeq_{DISP} X$ 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. The following example (labeled Example 2 for future reference) proves that this implication is strict: 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 $Y \succeq_{CONC} X$. As proved in part iv) of the proof below, for $n = 3$, $Y \succeq_{CONC} X \Leftrightarrow Y \succeq_{DISP} X$. 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 $Y \not\succeq_{CXMOD} X$ does not hold.

iv) To show that $Y \succeq_{DISP} X$ implies $Y \succeq_{CONC} X$, observe that, for $k = 1$, the left-hand side of (12) can be rewritten as $E[I_{\{Y_{(1)}^s=0\}}] = 1 - Pr(Y_i > s_i \forall i)$. Therefore, (12) implies that for all $s \in \bar{L}$, $Pr(Y > s) \geq Pr(X > s)$. Similarly, it follows from the equality in (12) for $k = n$ and the inequality for $k = n - 1$ that, for all $s \in \bar{L}$, $E[I_{\{Y_{(n)}^s=0\}}] \geq E[I_{\{X_{(n)}^s=0\}}]$, 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 $s \in \bar{L}$, showing that (12) 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 (12) for $k = 3$, it remains to show that for all $s \in \bar{L}$,

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

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 above in the proof that $Y \succeq_{DISP} X$ implies $Y \succeq_{CONC} X$. Therefore, both inequalities follow from $Y \succeq_{CONC} X$.

The following example (labeled for future reference Example 3) proves that for $n > 3$, $Y \succeq_{CONC} X$ does not imply $Y \succeq_{DISP} X$: Let $L = \{0, 1\}^4$, let g, f represent the distributions of Y, X , 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 $Y \succeq_{CONC} X$. For $s = (0, 0, 0, 0)$, 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. \blacksquare

Remark 1: Given the close relationship shown by Proposition 4 between the dispersion ordering and the symmetric supermodular ordering, and given that both are strictly weaker than the supermodular ordering, it is natural to ask how these two orderings relate to each other. For arbitrary supports, neither implies the other; however, as we note in Section 4, in the special case of binary random variables, the dispersion ordering is strictly stronger than the symmetric supermodular ordering.

To confirm that for arbitrary supports, $Y \succeq_{DISP} X$ does not imply $Y \succeq_{SSPM} X$, consider Example 2, used in part iii) of the proof of Theorem 2. Expand the support from $L = \{0, 1, 2\} \times \{0, 1\} \times \{0, 1\}$ to the symmetric $L' = \{0, 1, 2\}^3$. For $z \in L' \setminus L$, let $\delta(z) \equiv g - f = 0$ and for $z \in L$, define $\delta(z)$ as in Example 2. Then $Y \succeq_{DISP} X$ but, for the symmetric supermodular function w defined in Example 2, $w \cdot \delta = w \cdot (g - f) = -\epsilon < 0$, so $Y \succeq_{SSPM} X$ does not hold.

To confirm that $Y \succeq_{SSPM} X$ does not imply $Y \succeq_{DISP} X$, consider Y, X with distributions g, f , respectively, on $L = \{0, 1\}^3$ such that $\delta = g - f$ is given by $\delta(0, 0, 0) = 0$, $\delta(z_1, z_2, z_3) = \epsilon > 0$ if $\sum_{i=1}^3 z_i = 1$, $\delta(1, 1, 0) = -4\epsilon$, $\delta(0, 1, 1) = \delta(1, 0, 1) = -\epsilon$, and $\delta(1, 1, 1) = 3\epsilon$. It is easy to check that for $k = 0, 1, 2, 3$, $Pr(c(Y) = k) - Pr(c(X) = k) = 0, 3\epsilon, -6\epsilon, 3\epsilon$, respectively, so $c(Y) \succeq_{CX} c(X)$ and hence $Y \succeq_{SSPM} X$. However, $P(Y_1 = 1) - P(X_1 = 1) < 0$, hence Y and X do not have identical marginals, and therefore $Y \succeq_{DISP} X$ does not hold.

Remark 2: For $n \geq 3$, the strongest of our interdependence orderings, greater weak association, is not a difference-based ordering. To show this, we extend Example 1, used in part i) of the proof of Theorem 2. For $L = \{0, 1\}^3$, $f(x) = \frac{1}{3}$ if $\sum_{i=1}^3 x_i = 1$ and $f(x) = 0$ otherwise. $g(y)$ takes 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. Now define $\tilde{f}(x) = \frac{1}{2}$ if $\sum_{i=1}^3 x_i = 0$, $\tilde{f}(x) = \frac{1}{6}$ if $\sum_{i=1}^3 x_i = 1$, and $\tilde{f}(x) = 0$ otherwise, and let $\tilde{g}(y)$ take 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, $\tilde{g} - \tilde{f} = g - f$,

and we showed in part i) of the proof of Theorem 2 that $g \succeq_{GWA} f$ does not hold. Yet (see Appendix C) $\tilde{g} \succeq_{GWA} \tilde{f}$. Hence for $n \geq 3$, greater weak association is not difference-based. Since the supermodular ordering is difference-based and the distributions in this example are symmetric, it follows that even for *symmetric* distributions, greater weak association is strictly stronger than the supermodular ordering whenever $n \geq 3$.

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 [14] 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 [19], 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 both i) equivalences among the 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) $Y \succeq_{SCXMOD} X$;

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

Proof. a) It is clear from Definitions 9 and 10 that $Y \succeq_{DISP} X$ implies $Y \succeq_{SDISP} X$. It follows from Definition 9 and Proposition 4 that for symmetric distributions on $L = \{0, 1\}^n$, $Y \succeq_{SDISP} X$ if and only if $c(Y) \succeq_{CX} c(X)$ —the only non-trivial choice of $b = (b_0, \dots, b_0)$ in Definition 9 is $(0, \dots, 0)$, in which case (8) coincides with $c(Y) \succeq_{CX} c(X)$. But Propositions 1 and 2 together imply that for symmetric distributions on $L = \{0, 1\}^n$, $Y \succeq_{SPM} X$ if and only if $c(Y) \succeq_{CX} c(X)$. The remaining equivalences then follow from Theorem 2.

b) The equivalence of $Y \succeq_{SSPM} X$ and $c(Y) \succeq_{CX} c(X)$ on $L = \{0, 1\}^n$ is shown in Proposition 2. Since every symmetric $w \in \mathcal{CM}^*$ is supermodular, i) implies ii). Since $c(Y)$ is symmetric and convex-modular, ii) implies iii). ■

For random vectors on $L = \{0, 1\}^n$ with asymmetric distributions, the dispersion ordering is strictly stronger than the three orderings shown to be equivalent in part b) of Proposition 5. With $s = (0, 0, \dots, 0)$, the majorization condition (10) defining $Y \succeq_{DISP} X$ is equivalent to $c(Y) \succeq_{CX} c(X)$, but other choices of s generate additional conditions. In particular, as shown in Section 2.5, $Y \succeq_{DISP} X$ implies that Y and X have identical marginals, whereas this is not an implication of the three orderings in part b) of Proposition 5.²²

4.2 Three Dimensions

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

²²The second example in Remark 1 following Theorem 2 is one for which $L = \{0, 1\}^3$ and $Y \succeq_{SSPM} X$, but Y and X do not have identical marginals, so $Y \succeq_{DISP} X$ does not hold.

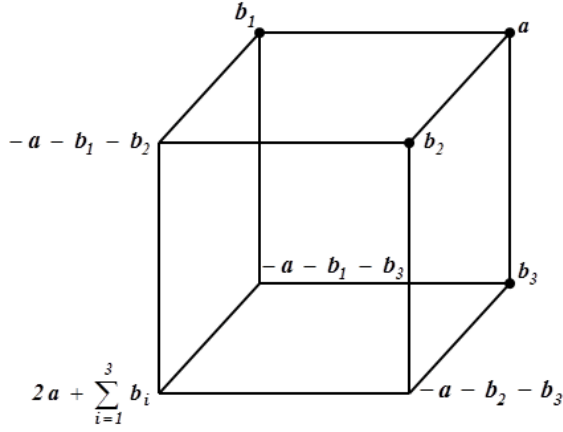


Figure 1a: $\delta = g - f$ for $L = \{0,1\}^3$

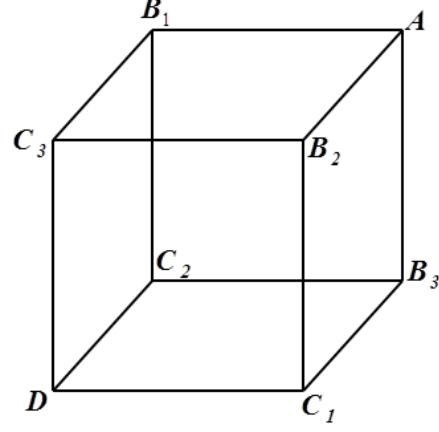


Figure 1b: Labels for the nodes of $L = \{0,1\}^3$

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 $L = \{0,1\}^3$, the values of δ at each of the 8 nodes are displayed in Figure 1a. Each of conditions i)-v) implies that Y and X have identical marginals. This in turn implies that once the 4 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 (corresponding to the 4 nodes marked with black dots in Figure 1a), the remaining 4 values are determined. Given identical marginals, $Y \succeq_{CONC} X$ if and only if

$$a \geq 0, \quad a + b_k \geq 0 \quad \forall k \in \{1, 2, 3\}, \quad \text{and} \quad 2a + \sum_{i=1}^3 b_i \geq 0. \quad (15)$$

In (15), 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:

$$\text{For } L = \{0, 1\}^3, \quad Y \succeq_{CONC} X \quad \text{implies} \quad Y \succeq_{SPM} X. \quad (16)$$

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

Our constructive proof of (16) decomposes δ into 6 elementary transformations (ET's) of the form defined in (4): for each of the 6 faces of the cube $L = \{0, 1\}^3$, there is one ET involving the 4 nodes on that face.²³ 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

$$a = \bar{\alpha}_{12} + \bar{\alpha}_{13} + \bar{\alpha}_{23} \quad \text{and} \quad \forall i, j, k \in \{1, 2, 3\}, i \neq j \neq k, \quad a + b_k = \bar{\alpha}_{ij} + \underline{\alpha}_{ij}. \quad (17)$$

Now set

$$\bar{\alpha}_{ij} = a \left(\frac{a + b_k}{3a + \sum_{i=1}^3 b_i} \right) \quad \text{and} \quad \underline{\alpha}_{ij} = (2a + \sum_{i=1}^3 b_i) \left(\frac{a + b_k}{3a + \sum_{i=1}^3 b_i} \right). \quad (18)$$

It is apparent by inspection that the equations (17) are satisfied and that, if the inequalities (15) 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 $g \succeq_{SPM} f$. ■

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 (used in part iv) of 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 part iii) of 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 $a \in \{0, 1\}^4$, $P(Z = a) = P(Z = (1, 1, 1, 1) - a)$. Top-to-bottom symmetry arises naturally in a variety of settings. We give two examples: matching with frictions and social learning in networks.

²³Hu, Xie, and Ruan [31, pp. 188-9] proved (16) in a very indirect manner using the tool of “majorization with respect to weighted trees”.

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”.

Choi, Gale, and Kariv [19] experimentally investigate the effect of network structure on the degree of interdependence of behavior in social learning situations. To simplify the computation of equilibrium behavior, they work with an environment that is symmetric with respect to the two possible states of the world, signals, and actions. As a result, in each period and for any network, the predicted distributions over agents’ actions 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 $s \in \bar{L}$, 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, we can show by example that the GWA ordering is strictly stronger and the concordance ordering strictly weaker.²⁴

5 Conclusion

In this paper we examined five orderings of interdependence for multivariate distributions and described a range of economic contexts in which these orderings are applicable, including multidimensional inequality comparisons, assessments of ex post inequality under uncertainty, valuations of portfolios of assets or insurance policies, and assessments of systemic risk in financial systems. 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. Our dispersion ordering is motivated by the link between greater interdependence in random vectors and lower dispersion of the cdf's of their order statistics, and we proved that greater interdependence for Y than for X according to the dispersion ordering corresponds to greater riskiness, in the sense of Rothschild and Stiglitz [47], of the summary statistic $\sum_{i=1}^n I_{\{Y_i > s_i\}}$ compared to $\sum_{i=1}^n I_{\{X_i > s_i\}}$, for *any* point (s_1, \dots, s_n) in the support. Similarly, Y is more interdependent than X according to the convex-modular ordering if for *any* choice of $\{r_i\}_{i=1}^n$ nondecreasing, the aggregate $\sum_{i=1}^n r_i(Y_i)$ is riskier than $\sum_{i=1}^n r_i(X_i)$. 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. For multivariate random vectors each of whose components are binary, we demonstrated some equivalences among the orderings and provided easily checkable and easily interpretable necessary and sufficient conditions for the orderings to hold. For arbitrary multivariate distributions, we emphasized that the dispersion and concordance orderings have the advantage, relative to the other three, of being checkable pointwise.

²⁴To 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 (4), 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 $Y \succeq_{GWA} X$ does not hold. To show that the concordance ordering is strictly weaker, recall Example 3, which was used in part iv) of the proof of Theorem 2. In that example, $\delta = g - f$ satisfies top-to-bottom symmetry, and as we showed, $Y \succeq_{CONC} X$ holds but $Y \succeq_{DISP} X$ does not.

6 References

- [1] V.V. Acharya, A theory of systemic risk and design of prudential bank regulation, *J. of Financial Stability* 5 (2009), 224-255.
- [2] M.D. Adler, C.W. Sanchirico, Inequality and uncertainty: theory and legal applications, *Univ. of Pennsylvania Law Review* 155 (2006), 279-377.
- [3] T. Adrian, M.K. Brunnermeier, CoVar, *Fed. Res. Bank of N.Y. Staff Rep.* 348 (2009).
- [4] A.B. Atkinson, On the measurement of inequality, *J. Econ. Theory* 2 (1970), 244-263.
- [5] A.B. Atkinson, Multidimensional deprivation: contrasting social welfare and counting approaches, *J. Econ. Inequality* 1 (2003), 51-65.
- [6] A.B. Atkinson, F. Bourguignon, The comparison of multi-dimensioned distributions of economic status, *Rev. Econ.Stud.* 49 (1982), 183-201.
- [7] K. Baldiga, J. Green, Choice-based measures of conflict in preferences, W.P., Harvard Univ., 2010.
- [8] N. Beale, D. Rand, H. Battey, K. Croxson, R. May, M. Nowak, Individual and systemic risk: the regulator's dilemma, working paper, Harvard Univ. and Oxford Univ., 2010.
- [9] E. Ben-Porath, I. Gilboa, D. Schmeidler, On the measurement of inequality under uncertainty, *J. Econ. Theory* 75 (1997), 194-204.
- [10] P.J. Boland, F. Proschan, Multivariate arrangement increasing functions with applications in probability and statistics, *J. Multivariate Anal.* 25 (1988), 286-298.
- [11] F. Bourguignon, S.R. Chakravarty, The measurement of multidimensional poverty, *J. Econ. Inequality* 1 (2003), 25-49.
- [12] R.M. Burton, A.R. Dabrowski, H. Dehling, An invariance principle for weakly associated random variables", *Stoch. Processes and their Applics.* 23 (1986), 301-306.
- [13] B. Caillaud, J. Tirole, Consensus building: how to persuade a group, *Amer. Econ. Rev.* 97 (2007), 1877-1900.
- [14] A. Calvo-Armengol, M.O. Jackson, The effects of social networks on employment and inequality, *Amer. Econ. Rev.* 94 (2004), 426-454.
- [15] A. Calvo-Armengol, M.O. Jackson, Networks in labor markets: wage and employment dynamics and inequality, *J. Econ. Theory* 132 (2007), 27-46.
- [16] L.I. de Castro, Affiliation and dependence in economic models, D.P. 1479, *Ctr. for Math. Stud. in Econ. and Mgt. Sci.*, Northwestern Univ., 2009.
- [17] S.H. Chew, J.S. Sagi, An inequality measure for uncertain allocations, *J. Econ. Theory*, this issue, 000 (0000) 000-000.
- [18] T.C. Christofides, E. Vaggelatou, A connection between supermodular ordering and positive/negative association", *J. Multivariate Anal.* 88 (2004), 138-151.
- [19] S. Choi, D. Gale, S. Kariv, Behavioral aspects of learning in social networks: an experimental study, in J. Morgan (Ed.), *Advances in Appl. Microeconomics*, 13 (2005).
- [20] P. Dasgupta, A.K. Sen, D. Starrett, Notes on the measurement of inequality, *J. Econ. Theory* 6 (1973), 180-187.
- [21] M. Denuit, J. Dhaene, M.J. Goovaerts, R. Kaas, *Actuarial Theory for Dependent Risks: Measures, Orders, and Models*, Wiley, Chichester, West Sussex, 2005.
- [22] L. Epstein, S. Tanny, Increasing generalized correlation: a definition and some economic consequences, *Can. J. Econ.* 13, (1980), 16-34.
- [23] J.D. Esary, F. Proschan, D.W. Walkup, Association of random variables, with applications, *Annals of Math. Stats.* 44 (1967), 1466-1474.
- [24] R. Fernandez, J. Gali, To each according to...? markets, tournaments, and the matching problem with borrowing constraints, *Rev. Econ. Stud.* 66 (1999), 799-824.
- [25] T. Gajdos, E. Maurin, Unequal uncertainties and uncertain inequalities: an axiomatic approach", *J. Econ. Theory* 116 (2004), 93-118.
- [26] A. Galeotti, S. Goyal, M.O. Jackson, F. Vega-Redondo, L. Yariv, *Network Games*, *Rev. Econ. Stud.* 77 (2010), 218-244.
- [27] G.H. Hardy, J.E. Littlewood, G. Polya, Some simple inequalities satisfied by convex functions, *Messenger of Maths.*, 58 (1929), 145-52.
- [28] G.H. Hardy, J.E. Littlewood, G. Polya, *Inequalities*, Cambridge University Press, Cambridge,

2nd ed., 1952.

- [29] D.A. Hennessy, H.E. Lapan, A definition of “more systematic risk” with some welfare implications, *Economica* 70 (2003), 493-507.
- [30] T. Hu, A. Müller, M. Scarsini, Some counterexamples in positive dependence, *J. Stat. Planning and Inference* 124 (2004), 153-158.
- [31] T. Hu, C. Xie, L. Ruan, Dependence structures of multivariate bernoulli random vectors, *J. Multivariate Anal.* 94 (2005), 172-195.
- [32] I. Jewitt, Risk aversion and the choice between risky prospects: the preservation of comparative statics results, *Rev. Econ. Stud.* 54 (1987), 73-86.
- [33] K. Joag-Dev, F. Proschan, Negative association of random variables, with applications, *Annals of Stats.* 11 (1983), 286-95.
- [34] H. Joe, Multivariate Concordance, *J. Multivariate Anal.* 35 (1990), 12-30.
- [35] H. Joe, *Multivariate Models and Dependence Concepts*, Chapman and Hall/CRC, London, 1997.
- [36] Y. Kroll, L. Davidovitz, Inequality aversion versus risk aversion, *Economica* 70 (2003), 19-29.
- [37] H. Levy, J. Paroush, Toward multivariate efficiency criteria, *J. Econ. Theory* 7 (1974), 129-42.
- [38] R.P. McAfee, Coarse matching, *Econometrica* 70 (2002), 2025-34.
- [39] M.A. Meyer, Interdependence in multivariate distributions: stochastic dominance theorems and an application to the measurement of ex post inequality under uncertainty”, *Nuffield Coll. D.P.* 49, 1990.
- [40] M.A. Meyer, D. Mookherjee, Incentives, compensation and social welfare, *Rev. Econ. Stud.* 45 (1987), 209-226.
- [41] M.A. Meyer, M. Rothschild, (In)efficiency in sorting and matching, mimeo, Oxford, 2004.
- [42] M.A. Meyer, B. Strulovici, The supermodular stochastic ordering, mimeo, Oxford and Northwestern, 2011.
- [43] P.R. Milgrom, R.J. Weber, A theory of auctions and competitive bidding, *Econometrica* 50 (1982), 1089-1122.
- [44] T.S. Motzkin, H. Raiffa, G.L. Thompson, R.M. Thrall, The double description method, in H.W. Kuhn, A.W. Tucker (Eds.), *Contributions to the Theory of Games II*, Princeton University Press, Princeton, 1953.
- [45] A. Müller, D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester, West Sussex, 2002..
- [46] A. Prat, Should a team be homogenous?, *Eur. Econ. Rev.* 46 (2002), 1187-1207.
- [47] M. Rothschild, J.E. Stiglitz, Increasing risk: I. a definition, *J. Econ. Theory* 2 (1970), 225-243.
- [48] M. Shaked, J.G. Shanthikumar, Supermodular stochastic orders and positive dependence of random vectors, *J. Multivariate Anal.* 61 (1997), 86-101.
- [49] M. Shaked, Y.L. Tong, Some partial orderings of exchangeable random variables by positive dependence, *J. Multivariate Anal.* 17 (1985), 333-349.
- [50] R. Shimer, L. Smith, Assortative matching and search, *Econometrica* 68 (2000), 343-369.
- [51] A.H. Tchen, Inequalities for distributions with given marginals, *Annals of Prob.* 8 (1980), 814-827.
- [52] D. M. Topkis, Minimizing a submodular function on a lattice, *Oper. Res.* 26,(1978), 305-331.

Appendices

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

This proof builds on Cristofides and Vaggelatou's [18] 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 in the proof of Theorem 1. 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

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

$$= \sum_{i=0}^{l_i-1} E[w(i, Y_2, \dots, Y_m)] Pr(Y_1 = i) \quad (20)$$

$$\begin{aligned} &\geq \sum_{i=0}^{l_i-1} E[w(i, Z_2, \dots, Z_m)] Pr(Y_1 = i) \\ &= \sum_{i=0}^{l_i-1} E[w(i, Z_2, \dots, Z_m | Y_1 = i)] Pr(Y_1 = i) \\ &= Ew(Y_1, Z_2, \dots, Z_m). \end{aligned} \quad (21)$$

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 (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 (21) 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

$$\begin{aligned}
& Ew(Y_1, Y_2, \dots, Y_m) - Ew(Z_1, Y_2, \dots, Y_m) \\
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Z_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))]\} \\
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Z_1 > i\}}] E[w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)]\} \tag{22}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{l_i-1} \{E[I_{\{Y_1 > i\}}(w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m))] \\
&\quad - E[I_{\{Y_1 > i\}}] E[w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)]\} \tag{23}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{l_i-1} Cov(I_{\{Y_1 > i\}}, w(i+1, Y_2, \dots, Y_m) - w(i, Y_2, \dots, Y_m)) \\
&\geq \sum_{i=0}^{l_i-1} Cov(I_{\{Z_1 > i\}}, w(i+1, Z_2, \dots, Z_m) - w(i, Z_2, \dots, Z_m)) \\
&= Ew(Z_1, Z_2, \dots, Z_m) - Ew(Y_1, Z_2, \dots, Z_m). \tag{24}
\end{aligned}$$

The equality in (22) uses the independence of Z from Y , and that in (23) 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 (24) 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, $Y \succeq_{SPM} X$ implies $Y \succeq_{CXMOD} X$. We show that the implication is strict by providing 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 [42], we show how, for any finite support L , the “double description method”, conceptualized by Motzkin et al [44], 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 in Remark 2 (Section 3), $\tilde{g} \succeq_{GWA} \tilde{f}$

Since \tilde{f} and \tilde{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 nondecreasing 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 nondecreasing 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) $\tilde{g} - \tilde{f} = g - f$, where g, f are defined in Example 1; ii) $g \succeq_{SPM} f$ and the supermodular ordering is a difference-based order, so $\tilde{g} \succeq_{SPM} \tilde{f}$; 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, $\tilde{g} \succeq_{GWA} \tilde{f}$.

D Proof of Proposition 7

Since the SPM, CXMOD, and DISP orderings are all difference-based orderings, it is sufficient to work with $\delta \equiv g - f$. Figure 2 displays the values of δ at each of the 16 nodes of $L = \{0, 1\}^4$, given top-to-bottom symmetry of g and f , and hence of δ . Top-to-bottom symmetry of δ implies that once the 8 values of $\delta(1, 1, 1, 1) \equiv a$, $\delta(0, 1, 1, 1) \equiv b_1$, $\delta(1, 0, 1, 1) \equiv b_2$, $\delta(1, 1, 0, 1) \equiv b_3$, $\delta(1, 1, 1, 0) \equiv b_4$, $\delta(0, 1, 1, 0) \equiv c_{14}$, $\delta(1, 0, 1, 0) \equiv c_{24}$, and $\delta(1, 1, 0, 0) \equiv c_{34}$ are specified (corresponding to the 8 nodes marked with black dots in Figure 2), the remaining 8 values are determined. Now $c(Y) \succeq_{CX} c(X)$ corresponds to $a \geq 0$ and $2a + \sum_{i=1}^4 b_i \geq 0$, while $Pr(Y_i = Y_j) \geq Pr(X_i = X_j)$ 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

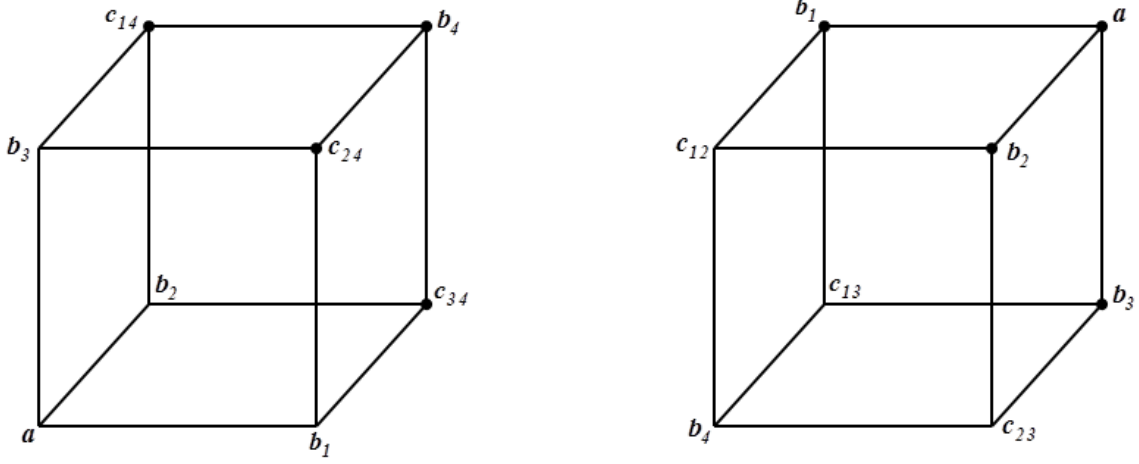


Figure 2: $\delta(z) = g(z) - f(z)$ for $L = \{0,1\}^4$ and with "top-to-bottom symmetry" (which implies $c_{14} = c_{23}$, $c_{24} = c_{13}$, $c_{34} = c_{12}$). In left cube, $z_4 = 0$; in right cube, $z_4 = 1$.

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

$$3a + 2b_i + \sum_{j=1}^4 b_j + c_{ij} + c_{ik} + c_{il} \geq 0. \quad (25)$$

Since identical marginals and top-to-bottom symmetry imply $a + \sum_{j=1}^4 b_j + c_{ij} + c_{ik} + c_{il} \geq 0$, (25) 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 (4): 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

$$\begin{aligned} \sum_{i < j} \beta_{ij} &= a \\ b_i + \beta_{ij} + \beta_{ik} + \beta_{il} &= \alpha_{ij} + \alpha_{ik} + \alpha_{il} \quad \forall i \neq j \neq k \neq l \\ -c_{ij} + \beta_{ij} + \beta_{kl} &= \alpha_{ik} + \alpha_{il} + \alpha_{jk} + \alpha_{jl} \quad \forall i \neq j \neq k \neq l. \end{aligned} \quad (26)$$

These three (sets of) equations ensure that each of the (sets of) nodes labeled a , b_i , and c_{ij} , respectively, in Figure 2 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 (26) can be rearranged to

$$\begin{aligned} \sum_{i < j} \beta_{ij} &= a \\ b_i + \beta_{ij} + \beta_{ik} + \beta_{il} &= \alpha_{ij} + \alpha_{ik} + \alpha_{il} \quad \forall i \neq j \neq k \neq l \\ 2\alpha_{ij} + 2\beta_{kl} &= a + b_i + b_j + c_{ij} \quad \forall i \neq j \neq k \neq l. \end{aligned} \quad (27)$$

Noting the similarity between the equations (27) and (17) for the 3-dimensional cube, set

$$\begin{aligned} \beta_{kl} &= \frac{a(a + b_i + b_j + c_{ij})}{4a + \sum_{i=1}^4 b_i}, \quad \forall i \neq j \neq k \neq l \\ \alpha_{ij} &= \frac{(2a + \sum_{i=1}^4 b_i)(a + b_i + b_j + c_{ij})}{2(4a + \sum_{i=1}^4 b_i)}, \quad \forall i \neq j. \end{aligned} \quad (28)$$

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 (27) 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 (28) 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, $g \succeq_{SPM} f$. By Theorem 2, it then follows that conditions i)-iv) are all equivalent.