Interdependence in Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty

Margaret A. Meyer
Nuffield College
Oxford, OX1 1NF
ENGLAND

April, 1990

Acknowledgment: I am very grateful to Tim Besley, Sushil Bikhchandani, Adam Brandenburger, Andrew Caplin, David Cox, Bob Gibbons, Terence Gorman, Ian Jewitt, Mervyn King, Paul Klemperer, Paul Milgrom, James Mirrlees, Dilip Mookherjee, and Barry Nalebuff for helpful comments and discussions.

J. E. L. Classification: 024, 025, 026

Keywords: stochastic dominance, correlation, inequality

Abstract

This paper analyzes the measurement of interdependence among n random variables. We adopt the stochastic dominance approach, relating concepts of interdependence expressed directly in terms of joint probability distributions to concepts expressed indirectly through properties of objective functions whose expectations are used to evaluate distributions. Since the expected values of additively separable objective functions depend only on the marginal distributions, attitudes towards correlation must be represented through non-separability properties. For the bivariate case, we present two stochastic dominance theorems, characterizing rankings of different strengths. For the multivariate case, we propose and chacterize three different rankings, each a natural extension of one of the bivariate ones.

This analysis of interdependence is applicable to a wide range of problems in choice theory and welfare economics. Our results are presented in the context of one such application: the measurement of inequality in an uncertain environment. This context motivates the "tournament axiom", which we view as a requirement for an objective function to represent a suitably strong aversion to negative interdependence. We apply our three multivariate stochastic dominance conditions to verify that the corresponding classes of objective functions satisfy the tournament axiom. We also analyze the relationships between our dominance conditions and the concepts of affiliation and association.

1. Introduction

The problem of the measurement of multidimensional correlation arises in several areas of choice theory and welfare economics. Its most familiar form involves an investor deciding how to allocate his endowment among several different risky assets within a given period; this decision requires assessing joint distributions of random asset returns (Hadar and Russell (1974), Levy and Paroush (1974), Epstein and Tanny (1980)). Similarly, an individual making investment decisions in a multiperiod context must evaluate joint distributions of random consumption levels in several different periods (Litzenberger and Ronn (1981)). In welfare economics, the analysis of inequality with respect to several different indicators of economic status, e.g. income and life expectancy, requires an assessment of the joint distribution over the population of these indicators (Atkinson and Bourguignon (1982)). The measurement of economic mobility focuses on the correlation between individual incomes at different points in time (Atkinson (1981)). A related example is the measurement of horizontal equity in the tax-benefit system: some proposals for assessing horizontal equity focus on the correlation between individuals' ranks in the pre-tax and post-tax income distributions (Feldstein (1976), Plotnick (1982)).[1]

In all of these settings, the decision maker's or social planner's evaluation of the multivariate distributions will reflect his preferences for positive or negative interdependence among the variables, as well as his preferences over their marginal distributions. Some work in these areas has concentrated on the development of indices for measuring correlation: for mobility, see Shorrocks (1978) and Cowell (1985), and for horizontal equity, see Plotnick (1982), King (1983), and Jenkins (1989). An index allows any two multivariate distributions to be ranked. An alternative approach to measuring interdependence, the stochastic dominance approach, seeks to characterize a partial ordering of multivariate distributions such that one distribution is ranked above another if and only if, for all objective functions W in a specified class, the expectation or sum of W is higher under the former distribution than under the latter.

[1] The problem of assessing the degree of concordance of several rankings of a given set of alternatives also involves measuring multidimensional correlation. The welfare-economic examples given above can be formulated in this way; in other settings, the rankings may represent individual preferences or beliefs (Reid (1990)).

Stochastic dominance theorems relate concepts of interdependence expressed directly in terms of probability distributions to concepts represented indirectly through properties of the objective function used to evaluate distributions. Theorems of this form have been presented for portfolio allocation by Levy and Paroush (1974), Hadar and Russell (1974), and Epstein and Tanny (1980), for multidimensional inequality by Atkinson and Bourguignon (1982), and for mobility by Atkinson (1981) and Kanbur and Stiglitz (1986). Despite providing only a partial ordering of distributions, the stochastic dominance approach is appealing (see especially Atkinson (1970)) because the assumptions on the underlying objective function are both explicit and fewer in number than in the development of indices. Furthermore, this approach highlights the inverse relationship between the strength of these assumptions and the stringency of the dominance conditions for ranking distributions.

For the most part, the stochastic dominance results for interdependence in multivariate distributions that appear in the economics literature are confined to the case of two dimensions.[2] While in some applications, such as the assessment of horizontal equity, the restriction to two dimensions is natural, in other contexts n-dimensional distributions will frequently need to be compared. The statistics literature contains numerous concepts of dependence for the bivariate case (see the survey by Jogdeo (1982)), but for more than two dimensions, attention has largely focused on the concepts of affiliation (Karlin and Rinott (1980)) and association (Esary, Proschan, and Walkup (1967)). The properties of affiliation and association arise naturally in many statistical settings, and affiliation in particular has many useful implications. However, affiliation and association are concepts of positive interdependence of random variables and do not represent notions of greater or less interdependence. Furthermore, these properties are both quite strong (see Section 4.6), and for economic applications weaker notions can be useful (for example, when considerable structure can be put on the preferences of the decision-maker or planner). There is thus considerable scope for the development, for economic applications, of stochastic dominance

[2] An exception is the paper by Hadar and Russell (1974), but they prove only a sufficiency result for preference between multivariate distributions and provide little interpretation of the assumptions on the objective function.

theorems characterizing concepts of greater interdependence for more than two dimensions.

This paper develops such stochastic dominance theorems in the context of a specific welfare-economic problem: the assessment of inequality in an uncertain environment. Such an assessment requires concepts of interdependence if one adopts the ex post approach to welfare evaluation under uncertainty, taking as a welfare objective the expected value of a social welfare function, W, defined on individual ex post utilities in each state.[3] Specifically, if $(\tilde{U}_1, \ldots, \tilde{U}_n)$ represent the random utility levels of n individuals, the welfare objective is $EW(\tilde{U}_1, \ldots, \tilde{U}_n)$. This welfare objective represents preferences over probability distributions of utility profiles. It seems reasonable to argue that a social planner or a society which, ex post, will value equality among realized utilities should, ex ante, prefer joint distributions of utilities which, for given marginal distributions, display greater positive interdependence.

This argument implies that, under uncertainty, an additively separable ex post welfare function W cannot represent an aversion to ex post inequality, since if W is additively separable its expectation depends only on the marginal distributions of individual welfare and is completely insensitive to the correlation properties of the joint distribution. Non-separability assumptions on W are therefore needed. The stochastic dominance theorems we derive characterize concepts of greater interdependence among random utilities (random variables) in terms of a preference for one joint distribution over another by all welfare functions (objective functions) possessing specific non-separability properties.

Section 2 motivates in greater detail the ex post approach to inequality measurement under uncertainty and discusses two contexts, organizations and taxation, in which considerable significance is attached to ex post equality (or ex post horizontal equity).

Sections 3 and 4 present the formal results on the measurement of interdependence. In Section 3, which treats the bivariate case, we propose two types of non-

[3] Hammond (1983) provides an axiomatic foundation for such a welfare objective. Meyer and Mookherjee (1987) discuss the relationship between such an objective and attitudes towards correlation of utilities.

separability property for the welfare function, which we term "weak complementarity" and "strong complementarity". Each of these properties represents a preference for a different type of "elementary transformation" of the joint distribution of utilities which, while leaving the marginal distributions unchanged, can be viewed as increasing the correlation. We present two stochastic dominance theorems, the first characterizing in different ways partial orderings equivalent to preference by all weakly complementary welfare functions and the second doing the same for strongly complementary functions. The first theorem is new, while the second extends existing economic and statistical results by providing a new representation of the partial ordering (which possesses a natural extension to n dimensions). These theorems are useful as both intuitive and formal stepping stones to the n-dimensional case.

Section 4 begins with a heuristic discussion of the issues involved in developing useful extensions of these results to more than two dimensions. If the only assumption on the welfare function is that it prefers joint distributions producing *no* ex post inequality at all to joint distributions, with the same marginals, producing *some* ex post inequality, then although we can characterize the associated partial ordering, it is so stringent as to be of little use. Developing weaker partial orderings requires specifying how the welfare function will rank distributions neither of which generates complete ex post equality. Any such specification, however, implicitly involves assessing the level of inequality in realized utility profiles for the n individuals, and there is considerably more scope for choice in this assessment than in the two-individual case. Consequently, the scope for choice among non-separability properties of the welfare function is considerably broader in the n-dimensional than in the two-dimensional case.

To gauge whether any particular set of assumptions on the welfare function captures a sufficiently strong aversion to ex post inequality (negative interdependence), we impose an axiom, referred to as the "tournament axiom". This axiom is motivated by the negative correlation of utilities generated by tournament reward schemes and requires that for every member of a class of welfare functions, any tournament generate lower expected ex post welfare than the corresponding reward scheme which gives each individual the same marginal distribution of rewards but determines rewards independently.

In Sections 4.2 through 4.4, we propose and characterize three stochastic dominance rankings for interdependence among n random variables. The appeal of these rankings has three sources. First, each ranking is a very natural extension of one of the dominance conditions derived for the bivariate case. Second, the classes of non-separable objective functions which correspond to the rankings each contain specific forms adopted in the choice theory and welfare economics literatures. Third, each ranking can be used to show that the corresponding class of welfare functions satisfies the tournament axiom and so represents a sufficiently strong aversion to negative interdependence. Section 4.5 analyzes the relationships among the three stochastic dominance conditions and among the three classes of welfare functions, and Section 4.6 treats the relationships of the stochastic dominance conditions with the concepts of affiliation and association.

Section 5 discusses several directions for future research. All proofs are in the Appendix.

2. The Ex Post Approach to Measuring Inequality under Uncertainty, and the Tournament Axiom

The ex post approach to welfare evaluation under uncertainty takes as a welfare objective the expected value of a welfare function, W, defined on individual ex post utilities in each state: $EW(\tilde{U}_1, \ldots, \tilde{U}_n)$. In contrast, the ex ante approach takes individuals' ex ante expected utilities as the arguments of the welfare function: $W(E\tilde{U}_1, \ldots, E\tilde{U}_n)$.[4] In the latter case, welfare depends only on the marginal distributions of individual utilities, whereas in the former case, the entire joint distribution of utilities is in general relevant. A social planner or a society which, ex post, will value equality among realized utilities should, ex ante, prefer joint distributions of utilities which display greater positive interdependence, for any given marginal distributions and hence for any given degree of risk faced by individuals. Such preferences can

[4] The contrast between the ex ante and the ex post approaches has been examined by, among many others, Diamond (1967), Mirrlees (1974), Myerson (1981), Hammond (1981), (1982), (1983), and Broome (1984).

be expressed with the ex post approach, but not with the ex ante approach. However, such preferences cannot be represented by an ex post welfare function that is additively separable, since the expectation of such a function depends only on the marginal distributions. Thus, whereas concavity assumptions on the welfare function may be sufficient for expressing aversion to inequality in ex ante utility levels or in deterministic utility levels, non-separability assumptions are needed to express aversion to ex post inequality when probability distributions of utility profiles are being compared.

In what contexts involving uncertainty might a group possess strong preferences for ex post equality? Organizations are obvious settings. Sociologists have extensively studied how social comparisons within peer groups shape workers' assessments of fairness within organizations and thereby influence worker morale, unity, and cooperation (Baron (1987)). Baron has hypothesized that social comparisons are more influential, and hence that ex post inequality is less, when co-workers are demographically homogeneous and have frequent contact with one another and when information about wage differentials is public.[5] In the economics literature, Lazear (1989) argues that reducing ex post inequality in rewards when compensation is based on relative performance can benefit organizations by reducing incentives for employees to "sabotage" the work of others. In a similar vein, Milgrom (1988) suggests that reducing ex post inequality can deter employees from expending resources on activities to influence those who wield authority. Ex post inequality within organizations can be increased not only by narrowing wage (or utility) differentials, but also by enhancing the correlation of rewards across individuals.[6] [7]

---

[5] Pfeffer and Davis-Blake's (1985) findings on patterns of wage dispersion across administrative positions in colleges and universities are consistent with these hypotheses.

[6] Meyer and Mookherjee (1987) illustrate this point in their analysis of compensation schemes under moral hazard in an organization which values ex post equality.

[7] In some settings, increasing the correlation of rewards for the group as a whole may be difficult, as when the total number of promotions is fixed in advance. However, the concern for ex post equality may be significantly stronger within particular subgroups than between these subgroups, for the reasons suggested by Baron (above). In such cases, the organization can largely satisfy preferences for ex post equality by enhancing the correlation of rewards of individuals within subgroups. (The constraint

Preferences for ex post equality may also be important in the evaluation of tax systems. While random taxes can sometimes increase the ex ante expected utility of all individuals (Stiglitz (1982)), they are horizontally inequitable ex post. Yet ex ante welfare functions and ex post utilitarian ones are both completely insensitive to ex post horizontal inequity. A non-additively-separable ex post welfare function would capture the tradeoff between the efficiency benefits of random taxation and its equity costs.

The Tournament Axiom

The formal analysis that follows, when interpreted in the context of inequality measurement under uncertainty, applies to a group within which the concern for ex post equality is strong. The welfare objective is the expected value of an ex post welfare function W defined on the n-dimensional space of utility profiles for the n individuals.[8] We will impose non-separability conditions on W to incorporate an aversion to negative interdependence of utilities. Since as we will see, there is considerable scope for choice among non-separability properties in n dimensions, we develop an axiom, the "tournament axiom", to gauge whether any class of non-separable welfare functions represents a sufficiently strong aversion to negative interdependence.

The tournament axiom is motivated by the correlation properties of rank-order tournament reward schemes. In a tournament, a set of prizes is pre-specified, and they are awarded, one to each individual, on the basis only of the rank order of individual performance indicators. Focusing solely on the distribution of rewards and ignoring effort decisions, we see that a tournament generates extreme negative correlation of rewards. If the prizes are all distinct, no two agents can receive equal utility. Furthermore, a rise in one agent's rank and reward is necessarily accompanied by a reduction in the rank and reward of at least one other agent. The tournament

---

on the overall distribution of rewards will then generate negative correlation between rewards of different subgroups.)

[8] Since we will always compare joint distributions with identical marginals, all of our results would continue to hold if the welfare objective also contained an ex ante component, $W(E\tilde{U}_1, \ldots, E\tilde{U}_n)$, capturing a concern for ex ante fairness.

axiom we develop requires that for every member of a class of welfare functions, the prize distribution under any given tournament yield lower expected welfare than the prize distribution under the scheme which imposes the same risks on individuals as the tournament but determines their rewards independently.

Formally, the probability distribution over utility profiles generated by an arbitrary tournament can be described by a set of prizes in utility terms $(t_1, \ldots, t_n)$ and a set of $n!$ probabilities, one for each possible allocation of prizes to individuals. If the welfare function is symmetric over individuals, ex post welfare is the same for all allocations. We can summarize the $n!$ probabilities by an $n \times n$ matrix $P = \{p_{ij}\}$, where p_{ij} is the probability that agent i receives the jth prize. The matrix P is bistochastic, since each agent is certain to win exactly one of the n prizes and each prize is certain to be awarded to exactly one person. Note that this representation allows for the marginal distributions over the n prizes to differ across individuals. (In the special case of a symmetric tournament, $p_{ij} = \frac{1}{n}$ for all i, j.)

For any tournament, we can use the associated P matrix to construct a reward scheme that, for each individual, generates the same marginal distribution of utility as the tournament but determines individuals' rewards independently. This *randomized independent scheme* (RIS) conducts n independent lotteries, with the lottery for individual i assigning probability p_{ij} that he receives the jth prize.

A tournament and its associated RIS impose exactly the same degree of risk on each individual and differ only in the correlation among utilities that they induce. An additively separable welfare function therefore has the same expected value under the two schemes. Since a tournament produces negative correlation and the RIS independence, an ex post welfare function which is averse to negative correlation should always prefer a RIS to the tournament from which it was derived. Formally, we impose the

Tournament Axiom: *For all symmetric welfare functions in a specified class and for all tournaments, expected ex post social welfare is at least as high under the RIS associated with the tournament as under the tournament itself, i.e. for all symmetric*

W in the class,

$$\sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \prod_{k=1}^{n} p_{ki_k} W(t_{i_1}, \ldots, t_{i_n}) \geq W(t_1, \ldots, t_n) \tag{1}$$

for all vectors $(t_1, \ldots, t_n)$ *and all* $n \times n$ *bistochastic matrices* $P = \{p_{ij}\}$.

The right-hand side of (1) represents the deterministic social welfare under a tournament (when W is symmetric). The left-hand side is the expected value of W over the n^n possible prize allocations under the RIS.[9]

For any $n \times n$ bistochastic matrix P, there exists a tournament, described by $n!$ probabilities, that gives rise to the marginal distributions represented by P.[10] Thus, in checking (1) for all bistochastic matrices, we are not examining a set larger than the set of all tournaments. On the other hand, for $n > 2$, many tournaments map into the same P matrix, so some information is lost in summarizing the $n!$ probabilities by the matrix. Nevertheless, for symmetric welfare functions, all information relevant for comparing the tournament and the RIS is contained in the P matrix.

For each class of welfare functions we examine, we verify that the tournament axiom is satisfied by using the stochastic dominance condition we show to be equivalent to preference by all welfare functions in the class. If every tournament and

---

[9] We have stated the tournament axiom in "weak form", requiring only a weak inequality in the expected welfare comparison. Similarly, throughout the paper we present the dominance theorems with only weak inequalities. All of our results, however, are easily carried over to "strong forms". The strong form of the tournament axiom entails the additional requirement that, for each tournament, there exist a W in the specified class for which the inequality in (1) is strict (as long as the tournament and the RIS do not produce identical distributions of utilities). This version of the axiom can be verified by the strong forms of the dominance theorems, which involve at least one strict inequality in the stochastic dominance conditions and at least one W in the specified class for which the inequality in expected welfare is strict. We have chosen to present the weak forms purely for expositional convenience, but note that it is the strong form of the tournament axiom that rules out additively separable welfare functions.

[10] This follows from a result in Mirsky (1963), which ensures that we can solve $(n-1)^2 + 1$ linearly independent equations (corresponding to the number of independent entries in the bistochastic matrix, plus the constraint that the $n!$ probabilities sum to 1) for $n!$ variables, despite the restriction that these variables be non-negative.

its associated RIS can be ranked according to the appropriate stochastic dominance condition, then (1) follows.

In order to focus purely on measures of, and attitudes to, correlation, we compare only joint distributions with identical marginals. Thus, tradeoffs between correlation and risk, or between correlation and ex ante expected utility, are suppressed.[11] Since it is natural that the welfare function be non-decreasing in the utilities of individuals (though not intrinsically relevant to representing a preference for correlation), we will show that, for our stochastic dominance theorems, the partial ordering of distributions is the same whether or not the requirement that W be non-decreasing is imposed.

The analysis that follows treats distributions with discrete supports. In a discrete setting, the intuition behind the theorems and their proofs emerges much more clearly. This setting is also the natural one when considering tournaments, where the set of prizes is discrete.

3. Measures of Interdependence for Bivariate Distributions

Given a bivariate distribution of utilities $(\widetilde{U}_1, \widetilde{U}_2)$, call a transformation of the following form an *elementary transformation on identical intervals* (ETI). Given any $U^L < U^H$, the probabilities of (U^H, U^H) and (U^L, U^L) are increased by d ($d > 0$), while the probabilities of (U^H, U^L) and (U^L, U^H) are decreased by d. ETI's leave the marginal distributions of individual utilities unchanged, but they increase the likelihood of ex post equal interpersonal utility distributions and reduce the likelihood of unequal ones. The expected value of an additively separable ex post welfare function is unchanged by the increased correlation in utilities resulting from an ETI. We will call a welfare function *weakly complementary* if its expected value is never decreased by an ETI:

[11] Our focus on attitudes to correlation, and on the non-separability properties of objective functions which represent these attitudes, distinguishes our multivariate stochastic dominance theorems from those of Kolm (1977), developed in the context of multidimensional inequality. Kolm focuses on preferences between distributions whose marginals display different degrees of riskiness; these preferences can be represented by concavity assumptions on the objective function.

Definition: $W(U_1, U_2)$ is *weakly complementary* if and only if, for all $U^L < U^H$,

$$W(U^H, U^H) + W(U^L, U^L) - W(U^H, U^L) - W(U^L, U^H) \geq 0. \tag{2}$$

Observe that in an ETI, the same values of U^H and U^L must apply for both individuals, so that when they both do well or both do badly there is no ex post inequality at all. A more general type of transformation, without this requirement, is an *elementary transformation on non-identical intervals* (ETN). Given $U_1^L < U_1^H$ and $U_2^L < U_2^H$, the probabilities of (U_1^H, U_2^H) and (U_1^L, U_2^L) are increased by d ($d > 0$), while the probabilities of (U_1^H, U_2^L) and (U_1^L, U_2^H) are decreased by d.[12] A welfare function that weakly prefers all ETN's is *strongly complementary*:

Definition: $W(U_1, U_2)$ is *strongly complementary* if and only if, for all $U_1^L < U_1^H$, $U_2^L < U_2^H$,

$$W(U_1^H, U_2^H) + W(U_1^L, U_2^L) - W(U_1^H, U_2^L) - W(U_1^L, U_2^H) \geq 0.$$

To understand what is implied by adopting strong, rather than weak, complementarity, consider how the consequences of an ETN for the distribution of ex post equality, as measured by $- \mid U_1 - U_2 \mid$, vary with the relative positions of the intervals (U_1^L, U_1^H) and (U_2^L, U_2^H). If one interval is a subset of the other, then the distribution of $- \mid U_1 - U_2 \mid$ after the ETN first-order stochastically dominates the original distribution. On the other hand, if the intervals are disjoint, the ETN generates a mean-preserving reduction in the variability of $- \mid U_1 - U_2 \mid$, i.e. a second-order stochastic improvement in ex post equality. In the intermediate case where the intervals partially overlap, the ETN always raises the mean of $- \mid U_1 - U_2 \mid$ and, depending on the degree of overlap, the distribution of $- \mid U_1 - U_2 \mid$ after the ETN either first-order or second-order stochastically dominates the original distribution. Therefore, if welfare were a function only of $- \mid U_1 - U_2 \mid$, the assumption that all ETN's weakly increased expected welfare would imply that welfare was increasing and concave in $- \mid U_1 - U_2 \mid$; in contrast, the weaker assumption that all ETI's weakly increased

---

[12] Transformations of this form were initially discussed (under a variety of names) by Hamada (1974), Tchen (1976), and Epstein and Tanny (1980).

expected welfare would imply only that welfare was smaller when $- \mid U_1 - U_2 \mid$ was negative than when it was 0.[13]

ETI's are useful for characterizing the stochastic dominance conditions which are equivalent to preference by all weakly complementary welfare functions.

Proposition 1: *Let the random vectors (X_1, X_2) and (Y_1, Y_2) each have symmetric joint distributions, with identical discrete support $\{a_1, \ldots, a_M\} \times \{a_1, \ldots, a_M\}$ and identical marginals. Then the following statements are equivalent:*

(i) $P(X_1 = a_k, X_2 = a_l) \leq P(Y_1 = a_k, Y_2 = a_l) \quad \forall k, l \in \{1, \ldots, M\}, \quad k \neq l$

(ii) The distribution of (X_1, X_2) can be derived from that of (Y_1, Y_2) by a finite sequence of ETI's.

(iii) For all weakly complementary W, $EW(X_1, X_2) \geq EW(Y_1, Y_2)$.

(iv) For all non-decreasing and weakly complementary W, $EW(X_1, X_2) \geq EW(Y_1, Y_2)$.

Before presenting several stochastic dominance conditions equivalent to preference by all strongly complementary welfare functions, we state the definition of second-order stochastic dominance for discrete distributions with identical means.

Definition: Let the random variables λ and θ have identical discrete support $\{b_1, \ldots, b_N\}$. Assume that the expectations of λ and θ are equal, which implies

$$\sum_{r=1}^{N-1} (b_{r+1} - b_r) P(\lambda \leq b_r) = \sum_{r=1}^{N-1} (b_{r+1} - b_r) P(\theta \leq b_r).$$

The distribution of λ *second-order stochastically dominates* the distribution of θ (equivalently, is less risky in the sense of Rothschild and Stiglitz (1970)) if and only if

$$\sum_{r=1}^{s} (b_{r+1} - b_r) P(\lambda \leq b_r) \leq \sum_{r=1}^{s} (b_{r+1} - b_r) P(\theta \leq b_r) \qquad \forall s \in \{1, \ldots, N-2\}. \quad (3)$$

Given the assumption of equal means, an equivalent set of conditions is

$$\sum_{r=s}^{N} (b_r - b_{r-1}) P(\lambda \geq b_r) \leq \sum_{r=s}^{N} (b_r - b_{r-1}) P(\theta \geq b_r) \qquad \forall s \in \{3, \ldots, N\}. \quad (4)$$

[13] This interpretation reveals the existence of categories of transformations intermediate between ETI's and ETN's. The "inequality reducing reversals" defined by Plotnick (1982) in developing a ranking for horizontal inequity are an example of an intermediate category.

Proposition 2: *Let the random vectors* (X_1, X_2) *and* (Y_1, Y_2) *have identical discrete support* $\{a_1, \ldots, a_M\} \times \{a_1, \ldots, a_M\}$ *and identical marginal distributions. Then the following statements are equivalent:*

(i) $P(X_1 \geq a_k, X_2 \geq a_l) \geq P(Y_1 \geq a_k, Y_2 \geq a_l) \quad \forall k, l \in \{1, \ldots, M\}$

(ii) $P(X_1 \leq a_k, X_2 \leq a_l) \geq P(Y_1 \leq a_k, Y_2 \leq a_l) \quad \forall k, l \in \{1, \ldots, M\}$

(iii) The distribution of (X_1, X_2) *can be derived from that of* (Y_1, Y_2) *by a finite sequence of ETN's.*

(iv) For all non-decreasing functions f^1 *and* f^2 *defined on* $\Re$, *the distribution of* $f^1(Y_1) + f^2(Y_2)$ *second-order stochastically dominates the distribution of* $f^1(X_1) + f^2(X_2)$.

(v) For all non-decreasing functions f^1 *and* f^2 *defined on* $\Re$, $cov[f^1(X_1), f^2(X_2)] \geq cov[f^1(Y_1), f^2(Y_2)]$.

(vi) For all strongly complementary W, $EW(X_1, X_2) \geq EW(Y_1, Y_2)$.

(vii) For all non-decreasing and strongly complementary W, $EW(X_1, X_2) \geq EW(Y_1, Y_2)$.

The equivalence of (i), (ii), (iii), (vi), and (vii) was proved in the statistics literature by Tchen (1976) and in the economics literature by Epstein and Tanny (1980). The equivalence of (v) and (i) is noted in the statistics literature (see the survey by Jogdeo (1982)). The equivalence of (iv) to the other conditions is a new result, and is particularly useful in suggesting an extension to n dimensions. The Appendix includes proofs that (iv) and (v) are equivalent to the other conditions, making use of the existing results.

Propositions 1 and 2 characterize two distinct concepts of greater interdependence in bivariate distributions. Both concepts are partial orderings of distributions. The stochastic dominance conditions in Proposition 1 imply those in Proposition 2 (since every ETI is an ETN) but, as is easily checked, the reverse implication does not hold.

The condition $\text{cov}(X_1, X_2) \geq \text{cov}(Y_1, Y_2)$ defines a complete ordering of distributions and clearly is implied by, but does not imply, the stochastic dominance conditions in Proposition 2 (see (v)). Epstein and Tanny (1980, Theorem 7) showed why the covariance is of limited usefulness as a measure of interdependence when distributions are not restricted to be normal: the condition $\text{cov}(X_1, X_2) \geq \text{cov}(Y_1, Y_2)$ is equivalent to the condition that, for all welfare functions of the form $W(U_1, U_2) = f^1(U_1) + f^2(U_2) + \alpha U_1 U_2$ (where $\alpha > 0$), $EW(X_1, X_2) \geq EW(Y_1, Y_2)$, and furthermore this is the *largest* class of welfare functions for which the equivalence holds. Thus, if one is not willing to restrict attention to welfare functions of this special form, a larger covariance is not sufficient to guarantee higher expected welfare.

Proposition 3: *The tournament axiom is satisfied for the set of symmetric, weakly complementary welfare functions (and a fortiori for the set of symmetric, strongly complementary welfare functions).*

With two individuals, there are only two possible utility levels for each of them in the tournament and the RIS. The tournament axiom is satisfied because, for $\alpha \in (0, 1)$, there is a positive probability under the RIS that the individuals will receive equal rewards, compared to zero probability under the tournament, and any weakly complementary welfare function therefore prefers the former scheme.

4. Measures of Interdependence for Multivariate Distributions

4.1 The First Step

A very natural extension of an ETI to n dimensions is a transformation of the following form, to be called a NETI (where the "N" stands for "n-dimensional"): given any set of utility levels $\{U^1, \ldots, U^n\}$, not all equal, the probability of each of the $n!$ outcome vectors which are permutations of $(U^1, \ldots, U^n)$ is decreased by d, while the probability of each of the n outcome vectors $(U^1, \ldots, U^1), \ldots, (U^n, \ldots, U^n)$ is increased by $(n-1)!d$. A NETI leaves the marginal distributions unchanged but transfers probability mass from outcomes with some ex post inequality to outcomes with no ex post inequality at all. Call a welfare function *perfect correlation loving* if its expected value is never decreased by a NETI:

Definition: $W(U_1, \ldots, U_n)$ is *perfect correlation loving* if and only if, for all $\{U^1, \ldots, U^n\}$

for which it is not the case that $U^1 = U^2 = \ldots = U^n$,

$$(n-1)! \sum_{i=1}^{n} W(U^i, \ldots, U^i) \geq \sum_{\substack{(i_1,\ldots,i_n), \\ \text{permutations of} \\ (1,\ldots,n)}} W(U^{i_1}, \ldots, U^{i_n}). \tag{5}$$

This definition captures an appealing, and very basic, notion of ex post inequality aversion in n dimensions: the expected value of any perfect correlation loving welfare function is (weakly) increased by replacing an arbitrary symmetric multivariate distribution by one which has the same marginals but in which individual utilities are perfectly correlated (so all the probability mass lies along the grand diagonal). While appealing, this notion is, however, extremely weak, as is shown by the stringency of the corresponding stochastic dominance conditions:

Proposition 4: *Let the random vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ each have symmetric joint distributions, with identical discrete support $\{a_1, \ldots, a_M\} \times \cdots \times \{a_1, \ldots, a_M\}$ and identical marginals. Then the following statements are equivalent:*

(i) $P(X_1 = a_{i_1}, \ldots, X_n = a_{i_n}) \leq P(Y_1 = a_{i_1}, \ldots, Y_n = a_{i_n})$ for all $(i_1, \ldots, i_n)$ for which it is not the case that $i_1 = i_2 = \ldots = i_n$.

(ii) The distribution of $(X_1, \ldots, X_n)$ can be derived from that of $(Y_1, \ldots, Y_n)$ by a finite sequence of NETI's.

(iii) For all perfect correlation loving W, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

(iv) For all symmetric and perfect correlation loving W, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

The stochastic dominance condition (i) for $(X_1, \ldots, X_n)$ to be more interdependent than $(Y_1, \ldots, Y_n)$ requires that *every* outcome vector in which there is *not* complete ex post equality have lower probability under X than under Y. This is a strong condition even for two dimensions, but it becomes increasingly stringent as the number of dimensions increases: while the number of outcomes with *no* ex post inequality remains unchanged, the number of outcomes with *some* ex post inequality rises rapidly. The stringency of the condition is highlighted by the fact that it is violated if Y represents the distribution of utilities under a symmetric tournament ($p_{ij} = \frac{1}{n}$ for all i, j) and X the distribution under the corresponding RIS. (Given any outcome vector in which the utilities are neither all equal nor all distinct, the tournament assigns it zero probability, while the RIS assigns it strictly positive probability.) From Proposition 4, it then follows that the class of symmetric, perfect correlation loving welfare functions fails to satisfy the tournament axiom. This failure results because the tournament axiom requires a preference for independent distributions over negatively correlated ones, whereas the weak condition defining perfect correlation loving functions requires only a preference for distributions with perfect positive correlation over all others.

The implication is that developing useful rankings of interdependence in n dimensions requires imposing stronger conditions on W than that it be perfect correlation loving. Equivalently, we need to specify preferences over a larger set of transformations than the set of NETI's — over transformations which do not add probability mass only to outcomes with no inequality and thus do not completely equalize individual utilities. However, two complications arise.

First, there are many types of marginal-preserving transformations which could be considered. How many individuals (dimensions) should a transformation affect? When $n = 2$, the only possible answer is 2, but when $n > 2$, we can consider probabilistically equalizing the utilities of any subset of individuals. Also, should the utility levels whose probabilities are altered by a transformation be the same for all individuals affected, as with ETI's, or potentially different, as with ETN's?

Second, specifying preferences over any particular type of transformation implicitly involves reckoning the level of inequality in each n-dimensional outcome vector whose probability is altered by the transformation and then assigning weights to the different inequality levels. When $n = 2$, there is scope for choice in the assignment of weights, but the measure of inequality itself is straightforward: $|U_1 - U_2|$. When $n > 2$, however, there is no unambiguously best measure of inequality for realized profiles of utilities $(U_1, \ldots, U_n)$.[14]

As a consequence of these complications, there is considerably more scope in the

[14] Because NETI's add probability mass only to outcomes with all utilities equal, a preference for NETI's does *not* rest on the adoption of any particular cardinal measure of inequality for realized profiles of utilities.

n-dimensional than in the two-dimensional case for a diversity of views as to which types of marginal-preserving transformations increase interdependence. Correspondingly, there is considerably broader scope for judgments about which non-separability properties of the welfare function incorporate a preference for greater interdependence in multivariate distributions.

In Sections 4.2, 4.3, and 4.4, we propose three classes of non-additively-separable welfare functions, each one a subset of the set of perfect correlation loving welfare functions. In each case, we characterize the stochastic dominance conditions for multivariate distributions which are equivalent to preference by every member of the specified class. The appeal of these classes as representations of preferences for interdependence has three sources: 1) members of each class have been used in the choice theory and welfare economics literatures; 2) each class is narrow enough to satisfy the tournament axiom; and 3) each of the corresponding stochastic dominance conditions extends in a natural way one of the concepts of greater interdependence analyzed in Section 3 in the bivariate case.

4.2 Stochastic Dominance Condition A

One appealing approach to the multidimensional problem is to assume that the interdependence among n random utilities can be assessed by considering each pair of utilities in turn and assessing the interdependence of that pair, independently of the correlation of those utilities with the others. This assumption is captured by an ex post welfare function whose expectation depends only on the pairwise joint distributions of utilities , or equivalently, which is pairwise separable across individuals :

$$W = \sum_{i=1}^{n} \sum_{j=1}^{n} V^{ij}(U_i, U_j) \qquad (6)$$

With this assumption, the n-dimensional problem is reduced to a set of two-dimensional problems. Let us define Class A as the set of ex post welfare functions of the form (6), where for each $i \neq j$, V^{ij} is weakly complementary.

To interpret Class A in terms of preferences over marginal-preserving transformations, define a GETI (where the "G" stands for "generalized") as follows: given any configuration of utilities for individuals other than i and j, denoted $\bar{U}_{-i-j}$, and given any $U^L < U^H$, the probabilities of $(U^H, U^H; \bar{U}_{-i-j})$ and $(U^L, U^L; \bar{U}_{-i-j})$ are increased by d, while the probabilities of $(U^H, U^L; \bar{U}_{-i-j})$ and $(U^L, U^H; \bar{U}_{-i-j})$ are decreased by d. GETI's reduce ex post inequality between a pair of utilities, for any fixed realizations of the other $n-2$ utilities. The expected value of any welfare function in Class A is (weakly) increased by any GETI, and furthermore the increase is independent of $\bar{U}_{-i-j}$.

Further perspective on Class A is provided by the fact that any perfect correlation loving W that is symmetric and pairwise separable is a symmetric member of Class A ($W = \sum_{i=1}^{n} Z(U_i) + \sum_{i=1}^{n} \sum_{j=i+1}^{n} V(U_i, U_j)$, where V is symmetric and weakly complementary).

The stochastic dominance condition equivalent to preference by all welfare functions in Class A is simply the stochastic dominance condition (i) in Proposition 1, applied to each of the pairwise joint distributions.

Proposition 5: *Let the random vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ each have symmetric joint distributions, with identical discrete support $\{a_1, \ldots, a_M\} \times \cdots \times \{a_1, \ldots, a_M\}$ and identical marginals. Then the following statements are equivalent:*

(i) (SDA) $P(X_i = a_k, X_j = a_l) \leq P(Y_i = a_k, Y_j = a_l) \quad \forall i, j \in \{1, \ldots, n\}, i \neq j$

$$\text{and} \quad \forall k, l \in \{1, \ldots, M\}, k \neq l$$

(ii) For all W in Class A, i.e. representable as $\sum_{i=1}^{n} \sum_{j=1}^{n} V^{ij}(U_i, U_j)$, where for all $i \neq j$ V^{ij} is weakly complementary, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

(iii) For all non-decreasing W in Class A, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

The stochastic dominance condition (SDA) in Proposition 5 is applied to prove

Proposition 6: *The tournament axiom is satisfied for the set of symmetric welfare functions in Class A.*

The class of symmetric, perfect correlation loving welfare functions fails to satisfy the tournament axiom (Section 4.1), but when pairwise separability is imposed (reducing the class to the symmetric functions in Class A), the tournament axiom is satisfied (Proposition 6). With pairwise separability imposed, the stochastic dominance condition, (SDA), requires only that every unequal utility *pair* have lower probability under one distribution than under the other; this is significantly weaker than the requirement in the absence of pairwise separability ((i) in Proposition 4), that the probability ranking hold for every utility *vector* displaying some degree of inequality.[15]

We could consider the subset of Class A in which each V^{ij} is strongly complementary: the corresponding stochastic dominance conditions would be those in Proposition 2, applied to each of the pairwise joint distributions. Our reason for focusing on Class A and (SDA) is to stress that, when pairwise separability is adopted, *only* weak complementarity is needed for the tournament axiom to be satisfied.

The assumption of pairwise separability derives additional appeal from the fact that the Gini coefficient, expressed in terms of utilities, is based on a welfare function of this form (see Sen (1973, p. 31) and Sheshinski (1972)):

$$W = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \min(U_i, U_j) = \overline{U}(1 - G)$$

where

$$G = \frac{1}{2n^2\overline{U}} \sum_{i=1}^{n} \sum_{j=1}^{n} |U_i - U_j|$$

In this formula, $\overline{U}$ represents average utility and G the Gini coefficient. The function $\min(U_i, U_j)$ is in fact strongly complementary. As Sen (1973), Yitzhaki (1979), and Hey and Lambert (1980) have noted, $n^2\overline{U}G$ can be viewed as the aggregate level (over all pairs of individuals) of "relative deprivation", where the relative deprivation of individual i with respect to j is $\max(0, U_j - U_i)$. By generalizing the definition of relative deprivation, and summing over all pairs, one can derive other measures of total welfare that belong to Class A.

4.3 Stochastic Dominance Condition B

[15] It can be shown by example that the tournament axiom is not satisfied by the class of symmetric welfare functions whose expectation is (weakly) increased by any GETI, but which are not necessarily pairwise separable.

An alternative approach to the multidimensional problem is to confine attention to welfare functions that can be represented as

$$W = \widetilde{W}(\sum_{i=1}^{n} f^i(U_i)),$$

for some set $\{f^1, \ldots, f^n\}$ of non-decreasing functions and some convex function $\widetilde{W}$. The expression $\sum_{i=1}^{n} f^i(U_i)$ is a one-dimensional sufficient statistic for the vector of utilities. Call the set of such welfare functions Class B.

To interpret Class B in terms of preferences over marginal-preserving transformations, call a GETN a transformation which, given any $\bar{U}_{-i-j}$, $U_i^L < U_i^H$, and $U_j^L < U_j^H$, increases by d the probabilities of $(U_i^H, U_j^H; \bar{U}_{-i-j})$ and $(U_i^L, U_j^L; \bar{U}_{-i-j})$ and decreases by d the probabilities of $(U_i^H, U_j^L; \bar{U}_{-i-j})$ and $(U_i^L, U_j^H; \bar{U}_{-i-j})$. For any $\bar{U}_{-i-j}$, GETN's increase the interdependence between U_i and U_j in the same sense that, in the bivariate case, ETN's increase interdependence. The expected value of any welfare function in Class B is (weakly) increased by any GETN: by the monotonicity of the f^i and the additive separability of $\sum_i f^i(U_i)$, any GETN produces a mean-preserving spread in the distribution of $\sum_i f^i(U_i)$, and convexity of $\widetilde{W}$ guarantees that this spread (weakly) increases expected welfare. Furthermore, the increase in expected welfare varies with $\bar{U}_{-i-j}$ only if $\sum_{k \neq i,j} f^k(U_k)$ varies.

In a one-period portfolio problem in which the returns U_i on different assets are additive, a risk averse decision-maker will have an objective function of this form, except that $\widetilde{W}$ will be concave instead of convex. Such an objective function incorporates a preference for negative, rather than positive, interdependence of the random variables U_i; any inequalities derived for Class B will be reversed if $\widetilde{W}$ is concave. A welfare function in Class B could provide a natural representation of the preferences of a union facing uncertainty and concerned about the ex post equality of its members. Most models of union behavior (see the survey by Oswald (1985)) assume that the objective function is the sum of members' utilities or the expected utility of a representative member; being additively separable, these functions are sensitive only to the riskiness and not to the correlation of workers' utilities. By contrast, a welfare function that is a convex transform of the sum of members' utilities, and hence in Class B, favors joint distributions that display positive interdependence. Furthermore, this form of non-separable function is convenient for analyzing the effects on union behavior of a preference for ex post equality, since its indifference curves are identical to those of the utilitarian objective function: only under uncertainty do the two functions generate different behavioral predictions or prescriptions.

A welfare function in Class B arises naturally if we confine attention to distributions in which each random utility has an identical two-point support $\{U^L, U^H\}$. Then if the welfare function is symmetric, it can be expressed as

$$W = \widetilde{W}(\sum_{i=1}^{n} I_{\{U_i = U^H\}}),$$

and convexity of $\widetilde{W}$ follows as long as GETI's are assumed to increase expected welfare. This case would arise in the union setting if wages and unemployment benefit were uniform across workers and if all workers not laid off worked a common number of hours. It would arise for a group of employees currently receiving the same salary and all facing the same possible promotion. It would arise with random taxation when it was optimal to use no more than two tax rates for identical individuals.[16] Finally, and more generally, a symmetric objective function would take this form in any setting where a classification of each component into two categories was significantly easier than a finer classification, and where the categories naturally took the same form for each component.

The stochastic dominance condition equivalent to preference by all welfare functions in Class B is the natural n-dimensional extension of condition (iv) in Proposition 2:

Proposition 7: *Let the random vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ have identical discrete support and identical marginal distributions. Then the following statements are equivalent:*

(i) (SDB) *The distribution of $\sum_{i=1}^{n} f^i(Y_i)$ second-order stochastically dominates the distribution of $\sum_{i=1}^{n} f^i(X_i)$, for all non-decreasing functions $\{f^1, \ldots, f^n\}$.*

---

16 In Stiglitz's (1982) model, an optimal random tax takes only two levels when the tax is determined before the labor supply decision.

(ii) For all W in Class B, i.e. representable as $\widetilde{W}(\sum_{i=1}^n f^i(U_i))$*, where* $\widetilde{W}$ *is convex and* f^i *is non-decreasing for all i,* $EW(X_1,\ldots,X_n) \geq EW(Y_1,\ldots,Y_n)$*.*

(iii) For all non-decreasing W in Class B, $EW(X_1,\ldots,X_n) \geq EW(Y_1,\ldots,Y_n)$*.*

Furthermore, statements (i), (ii), and (iii) remain equivalent if in each case we require $f^i(\cdot) \equiv f^j(\cdot)$ *for all* $i \neq j$*.*

For welfare functions in Class B, the multivariate problem reduces to the comparison of the riskiness of the univariate distributions of $\sum_i f^i(X_i)$ and $\sum_i f^i(Y_i)$, for all non-decreasing functions $\{f^1,\ldots,f^n\}$. When the sufficient statistic for the vector of utilities naturally assumes a particular form, $\sum_{i=1}^n \hat{f}^i(U_i)$, as in the examples discussed above, so the welfare function is known up to the form of $\widetilde{W}$, the equivalent stochastic dominance condition is the requirement (weaker than (SDB)) of second-order stochastic dominance of the distribution of $\sum_{i=1}^n \hat{f}^i(U_i)$. We work with the stronger condition (SDB) in what follows; this makes our results more powerful and applicable to a wider class of welfare functions (all those in Class B).

Proposition 8: *The tournament axiom is satisfied for the set of symmetric welfare functions in Class B.*

Proposition 8 is easily proved, because *symmetric* functions in Class B take the form $\widetilde{W}(\sum_i f(U_i))$, for $\widetilde{W}$ convex and f non-decreasing, and the corresponding stochastic dominance condition is (by Proposition 7) second-order stochastic dominance of $\sum_i f(Y_i)$ over $\sum_i f(X_i)$ for all non-decreasing f. With f the same for all i, $\sum_i f(U_i)$ is deterministic under any tournament but variable under the associated RIS, so the dominance condition is clearly satisfied for any tournament/RIS pair.

4.4 Stochastic Dominance Condition C

Proposition 2 shows that, when $n = 2$ and the distributions of X and Y have identical marginals, the conditions

$$P(X_i \geq z_i \quad \forall i \in \{1,\ldots,n\}) \geq P(Y_i \geq z_i \quad \forall i \in \{1,\ldots,n\}) \quad \forall z = (z_1,\ldots,z_n) \tag{7a}$$

$$P(X_i \leq z_i \quad \forall i \in \{1,\ldots,n\}) \geq P(Y_i \leq z_i \quad \forall i \in \{1,\ldots,n\}) \quad \forall z = (z_1,\ldots,z_n) \tag{7b}$$

are equivalent to each other and to the condition (SDB) that $\sum_i f^i(Y_i)$ second-order stochastically dominates $\sum_i f^i(X_i)$ for all non-decreasing $\{f^1, \ldots, f^n\}$. Given $n = 2$ and identical marginals, the following two conditions are also equivalent to (SDB) and to each other:[17]

$$E\left[\prod_{i=1}^{n} f^i(X_i)\right] \geq E\left[\prod_{i=1}^{n} f^i(Y_i)\right], \text{ for all non-negative-valued, non-decreasing } \{f^1, \ldots, f^n\} \tag{8a}$$

$$E\left[\prod_{i=1}^{n} f^i(X_i)\right] \geq E\left[\prod_{i=1}^{n} f^i(Y_i)\right], \text{ for all non-negative-valued, non-increasing } \{f^1, \ldots, f^n\} \tag{8b}$$

When $n > 2$, however, (7a) and (7b) are no longer equivalent, nor are (8a) and (8b) equivalent, even for identical marginals (as is easily checked). Moreover, as is proved in Section 4.5, when $n > 2$, (SDB) is strictly stronger than (7a) and (7b) taken together, and also strictly stronger than (8a) and (8b) together. Still, (7a) and (7b) capture an intuitively appealing notion of greater interdependence: for every outcome vector, the probability of all individuals doing at least as well and the probability of all doing at least as badly are higher under one distribution than under the other. We now demonstrate that in $n > 2$ dimensions, (7a) and (7b) together are equivalent to (8a) and (8b) together, and we construct a class of welfare functions for which these are the corresponding stochastic dominance conditions.

Assume the random vectors X and Y have identical discrete support. Define C^+ as the set of welfare functions of the form

$$W^z(U_1, \ldots, U_n) = \begin{cases} 1 & \text{if } U_i \geq z_i \quad \forall i \\ 0 & \text{otherwise} \end{cases}$$

for some $z = (z_1, \ldots, z_n)$ in the support, and define C^- as the set of welfare functions of the form

$$W^z(U_1, \ldots, U_n) = \begin{cases} 1 & \text{if } U_i \leq z_i \quad \forall i \\ 0 & \text{otherwise} \end{cases}$$

for some z. Given any z, observe that for the corresponding W^z in C^+,

$$EW^z(X_1, \ldots, X_n) - EW^z(Y_1, \ldots, Y_n) = P(X_i \geq z_i \quad \forall i) - P(Y_i \geq z_i \quad \forall i),$$

---

[17] The proof of this result is an easy extension of the proof of Proposition 2. The restriction in (8a) and (8b) to non-negative-valued functions f^i is inessential for the result for $n = 2$ but essential for Proposition 9 below.

while for the corresponding W^z in C^-,

$$EW^z(X_1,\ldots,X_n) - EW^z(Y_1,\ldots,Y_n) = P(X_i \leq z_i \quad \forall i) - P(Y_i \leq z_i \quad \forall i).$$

Now define $\bar{C}^-$ as the set of welfare functions of the form

$$W^z(U_1,\ldots,U_n) = \begin{cases} n-k & \text{if } U_i > z_i \text{ for exactly } n-k+1 \text{ distinct } i \text{ values} \\ & \text{and } U_i \leq z_i \text{ otherwise, } k = 1,\ldots,n \\ 0 & \text{if } U_i \leq z_i \quad \forall i \end{cases}$$

for some z, and define $\bar{C}^+$ as the set of welfare functions of the form

$$W^z(U_1,\ldots,U_n) = \begin{cases} n-k & \text{if } U_i < z_i \text{ for exactly } n-k+1 \text{ distinct } i \text{ values} \\ & \text{and } U_i \geq z_i \text{ otherwise, } k = 1,\ldots,n \\ 0 & \text{if } U_i \geq z_i \quad \forall i \end{cases}$$

for some z.

Lemma 1: *Let the random vectors $X = (X_1,\ldots,X_n)$ and $Y = (Y_1,\ldots,Y_n)$ have identical discrete support and identical marginal distributions. Given any $z = (z_1,\ldots,z_n)$ in the support, for the corresponding W^z in $\bar{C}^-$,*

$$EW^z(X_1,\ldots,X_n) - EW^z(Y_1,\ldots,Y_n) = P(X_i \leq z_i \quad \forall i) - P(Y_i \leq z_i \quad \forall i),$$

and for the corresponding W^z in $\bar{C}^+$,

$$EW^z(X_1,\ldots,X_n) - EW^z(Y_1,\ldots,Y_n) = P(X_i \geq z_i \quad \forall i) - P(Y_i \geq z_i \quad \forall i).$$

Classes C^+, C^-, $\bar{C}^-$, and $\bar{C}^+$ each incorporate a preference for greater interdependence in the sense that the expected value of a welfare function belonging to any one of them is (weakly) increased by any GETN. Elements of C^+ and $\bar{C}^-$ are non-decreasing, while elements of C^- and $\bar{C}^+$ are non-increasing. The classes also differ in terms of how the increase in expected value from a GETN varies with the utility levels of the individuals not affected by the transformation.[18]

[18] More formally, consider the discrete analogue of the k^{th}-order derivative of W with respect to k distinct arguments. If for $k = 2$ and for every pair of arguments, this expression is non-negative everywhere, W prefers every GETN. This is the case for all W in C^+, C^-, $\bar{C}^-$, or $\bar{C}^+$. The sign of this expression for $k > 2$ determines how the effect of a GETN varies with the utilities of the individuals not affected. For W in C^+ or $\bar{C}^+$, this expression is non-negative everywhere, for all $k > 2$. On the other hand, for W in C^- or $\bar{C}^-$, this expression is non-positive everywhere if $k > 2$ and odd, and is non-negative everywhere if $k > 2$ and even.

Define Class C as the set of linear combinations, with non-negative weights, of elements of C^+, C^-, $\bar{C}^-$, and $\bar{C}^+$. The expected value of any welfare function in Class C is (weakly) increased by any GETN.

Proposition 9: *Let the random vectors $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ have identical discrete support and identical marginal distributions. Then the following statements are equivalent:*

(i) (SDC) $P(X_i \geq z_i \quad \forall i) \geq P(Y_i \geq z_i \quad \forall i)$ and $P(X_i \leq z_i \quad \forall i) \geq P(Y_i \leq z_i \quad \forall i) \qquad \forall (z_1, \ldots, z_n)$ in the support of X and Y.

(ii) $E[\prod_{i=1}^n f^i(X_i)] \geq E[\prod_{i=1}^n f^i(Y_i)]$, for all sets of non-negative-valued functions $\{f^1, \ldots, f^n\}$ such that either all f^i are non-decreasing or all f^i are non-increasing.

(iii) For all W in Class C, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

(iv) For all non-decreasing W in Class C, $EW(X_1, \ldots, X_n) \geq EW(Y_1, \ldots, Y_n)$.

The proof of Proposition 9 shows that any multiplicatively separable function $W(U_1, \ldots, U_n) = \prod_{i=1}^n f^i(U_i)$ belongs to Class C, if $\{f^1, \ldots, f^n\}$ are non-negative-valued and either all non-decreasing or all non-increasing.[19] The multiplicatively separable function $\frac{1}{\gamma} \prod_{i=1}^n (U_i)^{\frac{\gamma}{n}}$, where $\gamma \in (0, 1]$ and where U_i represents consumption in period i, was used as an intertemporal utility function by Litzenberger and Ronn (1981) to study hedging behavior in a multi-period investment setting.

The Rawlsian social welfare function

$$R(U_1, \ldots, U_n) \equiv \min(U_1, \ldots, U_n)$$

also belongs to Class C. If each U_i assumes values in $\{a_1, \ldots, a_M\}$, where $a_M > \ldots > a_1 \geq a_0 \equiv 0$, R can be expressed as

$$R = \sum_{m=1}^{M} (a_m - a_{m-1}) I_{\{U_i \geq a_m \quad \forall i\}},$$

[19] Since $\prod_i f^i(U_i) = \exp\{\sum_i \ln f^i(U_i)\} = \exp\{-\sum_i \ln(\frac{1}{f^i(U_i)})\}$, and $\exp(x)$ and $\exp(-x)$ are both convex functions, any multiplicatively separable function satisfying the above assumptions on $\{f^i\}$ also belongs to Class B.

a linear combination with non-negative weights of elements of C^+. Any non-negative-valued, monotonically increasing transformation of R has a similar representation, differing only in the weights, and so also belongs to Class C.

Proposition 10: *The tournament axiom is satisfied for the set of symmetric welfare functions in Class C.*

The proof of Proposition 10, unlike that of Proposition 8, is quite difficult. With Class B, restricting attention to *symmetric* functions *weakened* the corresponding dominance condition from (SDB) to second-order stochastic dominance of $\sum_i f(Y_i)$ over $\sum_i f(X_i)$ for all non-decreasing f, and the latter condition was easily checked for every tournament/RIS pair. With Class C, by contrast, imposition of symmetry on the welfare function does *not* produce an analogous simplification. Given a symmetric welfare function, we can confine attention to distributions which have been "symmetrized" by the averaging of probability mass over all outcome vectors which are permutations of one another. However, for symmetric distributions, the dominance condition equivalent to preference by all symmetric elements of Class C can be shown to be (SDC) itself, not a weaker condition. Verifying (SDC) for the symmetrized distributions corresponding to every tournament/RIS pair requires a lengthy argument.

4.5 Relationships among the Stochastic Dominance Conditions and among the Classes of Welfare Functions

In the preceding subsections, we have proposed three stochastic dominance rankings, (SDA), (SDB), and (SDC), each representing a concept of greater interdependence in multivariate distributions. Each ranking generalizes one of the partial orderings which we characterized in Section 3 for the bivariate case. When applied only to the bivariate case, the three conditions are related according to the following diagram:

$$\text{For } n=2: \qquad \text{(SDA)} \implies \text{(SDB)} \iff \text{(SDC)}$$

(For $n = 2$, (SDA) is (i) of Proposition 1, (SDB) is (iv) of Proposition 2, and (SDC) is (i) (or (ii)) of Proposition 2.)

Proposition 11: *When n can assume any value greater than or equal to 2, (SDB) $\Longrightarrow$ (SDC), but no other implication holds between (SDA), (SDB), and (SDC).*

Remark 1: We noted in Section 4.2 that for the subset of Class A in which $W = \sum_{i=1}^{n}\sum_{j=1}^{n} V^{ij}(U_i, U_j)$ with V^{ij} *strongly* complementary for $i \neq j$, the corresponding stochastic dominance conditions are those in Proposition 2, applied to each of the pairwise joint distributions. It is easy to show that these conditions are implied by each of (SDA), (SDB), and (SDC), but do not imply any of them.

Remark 2: Despite the fact that (SDB) $\Longrightarrow$ (SDC), satisfaction of the tournament axiom by the symmetric welfare functions in Class C does *not* follow trivially from satisfaction of the axiom by the symmetric functions in Class B. The reason is that, as Proposition 7 shows, the stochastic dominance condition equivalent to preference by all *symmetric* elements of Class B is second-order stochastic dominance of $\sum_i f(Y_i)$ over $\sum_i f(X_i)$ for all non-decreasing functions f, and this condition is weaker than (SDB). Proposition 11 does *not* show that (SDC) is implied by this weaker condition.

Proposition 11 shows that, comparing (SDA) on the one hand with (SDB) and (SDC) on the other, these conditions are strong and weak in different respects. (SDA) is weak in that it restricts attention to the pairwise joint distributions but strong in the conditions it imposes on these distributions, that they be related by a sequence of transformations of a very narrow form (ETI's). (SDB) and (SDC) are weak in that they are satisfied by distributions related by a sequence of transformations of a relatively broad form (GETN's, not just GETI's), but they are strong in not allowing attention to be restricted to the pairwise joint distributions.

Proposition 11 also makes clear that there is not a unique extension to n dimensions of the bivariate stochastic dominance conditions in Proposition 2. Both (SDB) and (SDC) extend these conditions in extremely natural ways, but the former condition is strictly stronger than the latter.

Proposition 12: *None of the three classes of welfare functions A, B, or C is a subset of either of the others.*

4.6 Relationships of the Stochastic Dominance Conditions with Affiliation and Association

The concepts of affiliation and association of random variables have been extensively studied in the statistics literature and fruitfully applied in economic analyses of uncertainty and information (see, for example, Milgrom (1981), Milgrom and Weber (1982), and Jewitt (1987)). These are both concepts of *positive* interdependence, in contrast to our conditions (SDA), (SDB), and (SDC), which are partial orderings representing *greater* interdependence in one distribution than another. We can examine the relationships among these different notions by letting $(Y_1, \ldots, Y_n)$ be independent, letting $(X_1, \ldots, X_n)$ have identical marginals to $(Y_1, \ldots, Y_n)$, and letting the joint distribution of $(X_1, \ldots, X_n)$ vary.

We take the following definitions from Milgrom and Weber (1982). Given $x = (x_1, \ldots, x_n)$ and $x' = (x'_1, \ldots, x'_n)$, let $x \vee x'$ denote $(\max(x_1, x'_1), \ldots, \max(x_n, x'_n))$ and $x \wedge x'$ denote $(\min(x_1, x'_1), \ldots, \min(x_n, x'_n))$.

Definition: A subset A of R^n is *increasing* if its indicator function 1_A is non-decreasing.

Definition: A subset S of R^n is a *sublattice* if $x \vee x'$ and $x \wedge x'$ are in S whenever x and x' are.

Definition: The random variables $(X_1, \ldots, X_n)$ are *associated* if for all increasing sets A and B, $P(X \in (A \cap B)) \geq P(X \in A)P(X \in B)$, or equivalently, $P(X \in (A^C \cap B^C)) \geq P(X \in A^C)P(X \in B^C)$, where A^C denotes the complement of A in the support of the random variables.

Definition: The random variables $(X_1, \ldots, X_n)$ are *affiliated* if for all increasing sets A and B and every sublattice S, $P(X \in (A \cap B) \mid X \in S) \geq P(X \in A \mid X \in S)P(X \in B \mid X \in S)$, i.e. if the variables are associated conditional on any sublattice.

Result (Milgrom and Weber (1982, Theorem 24)): Let $p(x_1, \ldots, x_n)$ be the probability mass function of the discrete random variables $(X_1, \ldots, X_n)$. $(X_1, \ldots, X_n)$ are affiliated if and only if

$$p(x \vee x')p(x \wedge x') \geq p(x)p(x') \qquad \forall (x, x') \quad \text{in the support.} \tag{9}$$

Proposition 13: *Assume that $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ have identical discrete support and identical marginal distributions, and that $(Y_1, \ldots, Y_n)$ are independent.*

(i) If $(X_1, \ldots, X_n)$ are associated (or affiliated), then X and Y satisfy (SDC).

(ii) For $n = 2$, if (X_1, X_2) are associated (or affiliated), then X and Y satisfy (SDB).

(iii) Affiliation (or association) of $(X_1, \ldots, X_n)$ does not imply that X and Y satisfy (SDA).

(iv) Association (or affiliation) of $(X_1, \ldots, X_n)$ is not implied by X and Y satisfy (SDA), nor by X and Y satisfy (SDB), nor by X and Y satisfy (SDC).

Proposition 13 shows that, for multivariate distributions, affiliation and association capture different, but neither stronger nor weaker, notions of interdependence than (SDA) does (taking the benchmark distribution as an independent one). On the other hand, affiliation and association are strictly stronger criteria than (SDC) and, for $n = 2$, than (SDB) (with an independent distribution as benchmark). Whether association or affiliation of X implies that X and Y satisfy (SDB) for $n > 2$ is an open question.

5. Directions for Future Research

i) The tournament axiom is used in this paper as a standard with which to test whether a given class of objective functions incorporates a sufficiently strong aversion to negative interdependence. We have shown that the axiom is satisfied by the three classes of ex post welfare functions we propose. One could also seek to characterize the largest set of objective functions for which the axiom is satisfied.

ii) Our stochastic dominance theorems all take the following form: the distribution of X dominates that of Y according to a specified stochastic dominance condition if and only if $EW(X) \geq EW(Y)$ for all objective functions W in a specified class.

These theorems can be interpreted as axiomatizations of the stochastic dominance conditions in terms of properties of objective functions. One could also ask, for each theorem, whether the specified class of objective functions is the *maximal* class for which the theorem holds. If it is, then a theorem of the following form is true: W is in a specified class if and only if $EW(X) \geq EW(Y)$ for all X and Y such that X dominates Y according to a specified stochastic dominance condition. Theorems of the latter type can be viewed as axiomatizations of the classes of objective functions in terms of the stochastic dominance conditions.

For the bivariate case, it is easy to show that the classes of weakly complementary and strongly complementary welfare functions are maximal for the dominance conditions in Propositions 1 and 2, respectively. It is similarly easy to show in n dimensions that the class of perfect correlation loving welfare functions is maximal for the dominance condition in Proposition 4. Moreover, Class A can be shown to be maximal for the condition (SDA).[20]

Class B is not maximal for the condition (SDB). To see this, define the larger class, $\hat{B}$, as the set of non-negative linear combinations of elements of B, and observe that it follows easily from Proposition 7 that X and Y satisfy (SDB) if and only if $EW(X) \geq EW(Y)$ for all W in Class $\hat{B}$.

Class C is not maximal for the condition (SDC). Define Class $\hat{C}$ to consist of all welfare functions expressible as an element of C plus an arbitrary constant. Since all welfare functions in C^+, C^-, $\bar{C}^-$, and $\bar{C}^+$ (from whch Class C is constructed) are non-negative-valued, an element of $\hat{C}$ corresponding to a negative constant is not necessarily in C. Yet subtraction of a constant from the welfare function does not change the ranking of distributions, so (SDC) is equivalent to preference by all welfare functions in Class $\hat{C}$.

Are Classes $\hat{B}$ and $\hat{C}$ maximal for the conditions (SDB) and (SDC), respectively? For $n = 2$, (SDB) and (SDC) reduce to conditions (iv) and (i) in Proposition 2, for

[20] The proof of this result closely parallels the demonstration in Brandenburger (1985) that W is additively separable and monotonically increasing if and only if $EW(X) \geq EW(Y)$ for all X and Y such that each marginal distribution of X first-order stochastically dominates the corresponding marginal distribution of Y.

which the class of strongly complementary welfare functions is maximal. It can be shown that for $n = 2$, Classes $\hat{B}$ and $\hat{C}$ are both equivalent to the class of strongly complementary welfare functions, so in the bivariate case $\hat{B}$ and $\hat{C}$ are maximal. In the multivariate case, we know from Proposition 11 that (SDB) is strictly stronger than (SDC), and it can be shown that $\hat{B}$ contains $\hat{C}$. (By contrast, Proposition 12 showed that B does not contain C.) However, whether in the multivariate case $\hat{B}$ and $\hat{C}$ are maximal for (SDB) and (SDC), respectively, remains an open question.

iii) Our stochastic dominance theorems provide guidance in the development of indices of interdependence in multivariate distributions. As Cowell (1985) notes, an axiomatic development of summary indices for distributions requires a key economic assumption, in addition to mathematical and structural assumptions. With indices of inequality, for example, the key economic assumption is the Pigou-Dalton principle of transfers. With indices of interdependence, the key economic assumption will concern the class of marginal-preserving elementary transformations which increase interdependence. As we have seen, any such assumption is equivalent to an assumption about the particular type of non-separability displayed by the objective function in the stochastic dominance approach.

One might seek to develop an index of interdependence by finding a scalar measure of the "difference" or "distance" between an arbitrary symmetric distribution, F, and a distribution, G, which has the same marginals but in which the random variables are perfectly correlated. Such an approach would be analogous to that of Atkinson (1970), who developed his index of inequality by comparing an arbitrary distribution with one with the same mean but income equally distributed. Given a particular concave welfare function, Atkinson based his index on the "certainty equivalents" of the two distributions. In the current setting, given a particular perfect correlation loving, non-decreasing objective function W, one could determine how far the support of every component of G must be translated downwards (retaining perfect correlation and the shape of the marginals) to reduce the expectation of W to its value under F; this translation, $\Delta_{F,G}$, could be used as a scalar measure of interdependence in F. Or one could construct an index based on the "certainty equivalents" of F and G, that is, the values x_F and x_G such that $W(x_F, \ldots, x_F)$ equals expected welfare under F and

$W(x_G, \ldots, x_G)$ equals expected welfare under G. Such indices would clearly depend on the particular perfect correlation loving objective function chosen. The stochastic dominance theorems in this paper could be useful in guiding the selection, because they reveal to what features of distributions different subsets of perfect correlation loving objective functions are sensitive.

References

A. B. Atkinson, "On the Measurement of Inequality", *Journal of Economic Theory* 2 (1970), 244-263.

A. B. Atkinson, "The Measurement of Economic Mobility", in P. J. Eggelshoven and L. J. Gemerden, eds., *Inkommens Verdeling en Openbare Financien.* Leiden: Het Spectrum, 1981. Reprinted in A. B. Atkinson, *Social Justice and Public Policy.* Brighton: Wheatsheaf Books, 1983.

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

J. N. Baron, "The Employment Relation as a Social Relation", Working Paper, Stanford Graduate School of Business, Sept. 1987.

A. Brandenburger, "Stochastically Monotone Functions", Mimeo, Stanford Graduate School of Business, Jan. 1985.

J. Broome, "Uncertainty and Fairness", *Economic Journal* , 94 (1984), 624-632.

F. Cowell, "Measures of Distributional Change: An Axiomatic Approach", *Review of Economic Studies* , 52 (1985), 135-151.

P. Diamond, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: A Comment", *Journal of Political Economy* , 75 (1967), 765-766.

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

J. D. Esary, F. Proschan, and D. Walkup, "Association of Random Variables, with Applications", *Annals of Mathematical Statistics*, 38 (1967), 1466-1474.

M. Feldstein, "On the Theory of Tax Reform", *Journal of Public Economics*, 6 (1976), 77-104.

J. Hadar and W. R. Russell, "Rules for Ordering Uncertain Prospects", *American Economic Review,* 59 (1969), 25-34.

J. Hadar and W. R. Russell, "Stochastic Dominance in Choice under Uncertainty", in M. Balch, D. McFadden, and S. Wu, eds., *Essays on Economic Behavior under Uncertainty.* Amsterdam: North Holland, 1974.

K. Hamada, "Comment on 'Stochastic Dominance in Choice under Uncertainty' ", in M. Balch, D. McFadden, and S. Wu, eds., *Essays on Economic Behavior under Uncertainty.* Amsterdam: North Holland, 1974.

P. Hammond, "Ex Ante and Ex Post Welfare Optimality uncer Uncertainty", *Economica*, 48 (1981), 235-250.

P. Hammond, "Utilitarianism, Uncertainty, and Information", in A. Sen and B. Williams, eds., *Utilitarianism and Beyond.* Cambridge: Cambridge University Press, 1982.

P. Hammond, "Ex Post Optimality as a Dynamically Consistent Objective for Collective Choice under Uncertainty", in P. Pattanaik and M. Salles, eds., *Social Choice and Welfare.* Amsterdam: North Holland, 1983.

J. D. Hey and P. J. Lambert, "Relative Deprivation and the Gini Coefficient: A Comment", *Quarterly Journal of Economics* , 95 (1980), 567-573.

S. P. Jenkins, "Social Welfare Function Measures of Horizontal Inequity", Working Paper, University of Bath, Aug. 1989.

I. Jewitt, "Risk Aversion and the Choice Between Risky Prospects: The Preservation of Comparative Statics Results", *Review of Economic Studies*, 54 (1987), 73-85.

K. Jogdeo, "Concepts of Dependence", in S. Kotz and N. L. Johnson, eds., *Encyclopedia of Statistical Sciences.* New York: John Wiley and Sons, 1982.

S. M. R. Kanbur and J. E. Stiglitz, "Intergenerational Mobility and Dynastic Inequality", Mimeo, Princeton University, April 1986.

S. Karlin and Y. Rinott, "Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions", *Journal of Multivariate Analysis*, 10 (1980), 467-498.

M. A. King, "An Index of Inequality, with Applications to Horizontal Equity and Social Mobility", *Econometrica*, 51 (1983), 99-115.

S.-C. Kolm, "Multidimensional Egalitarianisms", *Quarterly Journal of Economics*, 91 (1977), 1-13.

E. P. Lazear, "Pay Equality and Industrial Politics", *Journal of Political Economy*, 97 (1989), 561-580.

H. Levy and J. Paroush, "Toward Multivariate Efficiency Criteria", *Journal of Economic Theory* , 7 (1974), 129-142.

R. H. Litzenberger and E. I. Ronn, "The Effect of Time-Complementarity on Optimal Consumption-Investment Decisions", Research Paper No. 619, Stanford Graduate School of Business, 1981.

M. A. Meyer and D. Mookherjee, "Incentives, Compensation, and Social Welfare", *Review of Economic Studies* , 54 (1987), 209-226.

P. R. Milgrom, "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, 12 (1981), 380-391.

P. R. Milgrom, "Employment Contracts, Influence Activities, and Efficient Organization Design", *Journal of Political Economy*, 96 (1988), 42-60.

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

J. Mirrlees, "Notes on Welfare Economics, Information, and Uncertainty", in M. Balch, D. McFadden, and S. Wu, eds., *Essays on Economic Behavior under Uncertainty.* Amsterdam: North Holland, 1974.

L. Mirsky, "Results and Problems in the Theory of Doubly Stochastic Matrices", *Z. Wahrscheilichkeitstheorie und Verw. Gebeite*, 1 (1963), 319-334.

R. Myerson, "Utilitarianism, Egalitarianism, and the Timing Effect in Social Choice Problems", *Econometrica* , 49 (1981), 883-897.

A. J. Oswald, "The Economic Theory of Trade Unions: An Introductory Survey", *Scandinavian Journal of Economics*, 87 (1985), 160-193.

J. Pfeffer and A. Davis-Blake, "Determinants of Salary Inequality in Organizations", Mimeo, Stanford Graduate School of Business, 1985.

R. Plotnick, "The Concept and Measurement of Horizontal Inequity", *Journal of Public Economics*, 17 (1982), 373-391.

G. C. Reid, "Analysing Rankings, with an Application to the Financing of Small Entrepreneurial Firms", *Economic Journal Supplement*, 100 (1990), 200-205.

M. Rothschild and J. E. Stiglitz, "Increasing Risk: I. A Definition", *Journal of Economic Theory* , 2 (1970), 225-243.

A. K. Sen, *On Economic Inequality.* Oxford: Clarendon Press, 1973.

E. Sheshinski, "Relation between a Social Welfare Function and the Gini Index of Inequality", *Journal of Economic Theory* , 4 (1972), 98-100.

A. F. Shorrocks, "The Measurement of Mobility", *Econometrica* , 46 (1978), 1013-1024.

J. E. Stiglitz, "Utilitarianism and Horizontal Equity", *Journal of Public Economics*, 18 (1982), 1-33.

A. H. T. Tchen, "Inequalities for Distributions with Given Marginals", Ph.D. Thesis, Department of Statistics, Stanford University, June 1976.

S. Yitzhaki, "Relative Deprivation and the Gini Coefficient", *Quarterly Journal of Economics* , 93 (1979), 321-324.

Appendix

Proof of Proposition 1: (i) $\Rightarrow$ (ii): We will show that (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i), and that (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (ii).

(i) $\Rightarrow$ (ii): We identify a sequence of $\binom{n}{2}$ ETI's, each corresponding to a different (k,l) pair (with $k < l$), that converts the distribution of (Y_1, Y_2) into that of (X_1, X_2). For each (k,l), the ETI increases the probabilities of (a_k, a_k) and (a_l, a_l) by $P(Y_1 = a_k, Y_2 = a_l) - P(X_1 = a_k, X_2 = a_l) \geq 0$ and reduces the probabilities of (a_k, a_l) and (a_l, a_k) by the same amount. Given the symmetry of the distributions, it is clear that this sequence converts the probabilities of all of the unequal outcome pairs under the Y distribution into those under the X distribution. That the probabilities of all of the equal outcome pairs are so converted follows from the assumption of identical marginals for (X_1, X_2) and (Y_1, Y_2) and the fact that ETI's leave the marginals unchanged.

(ii) $\Rightarrow$ (i) follows immediately from the definition of an ETI.

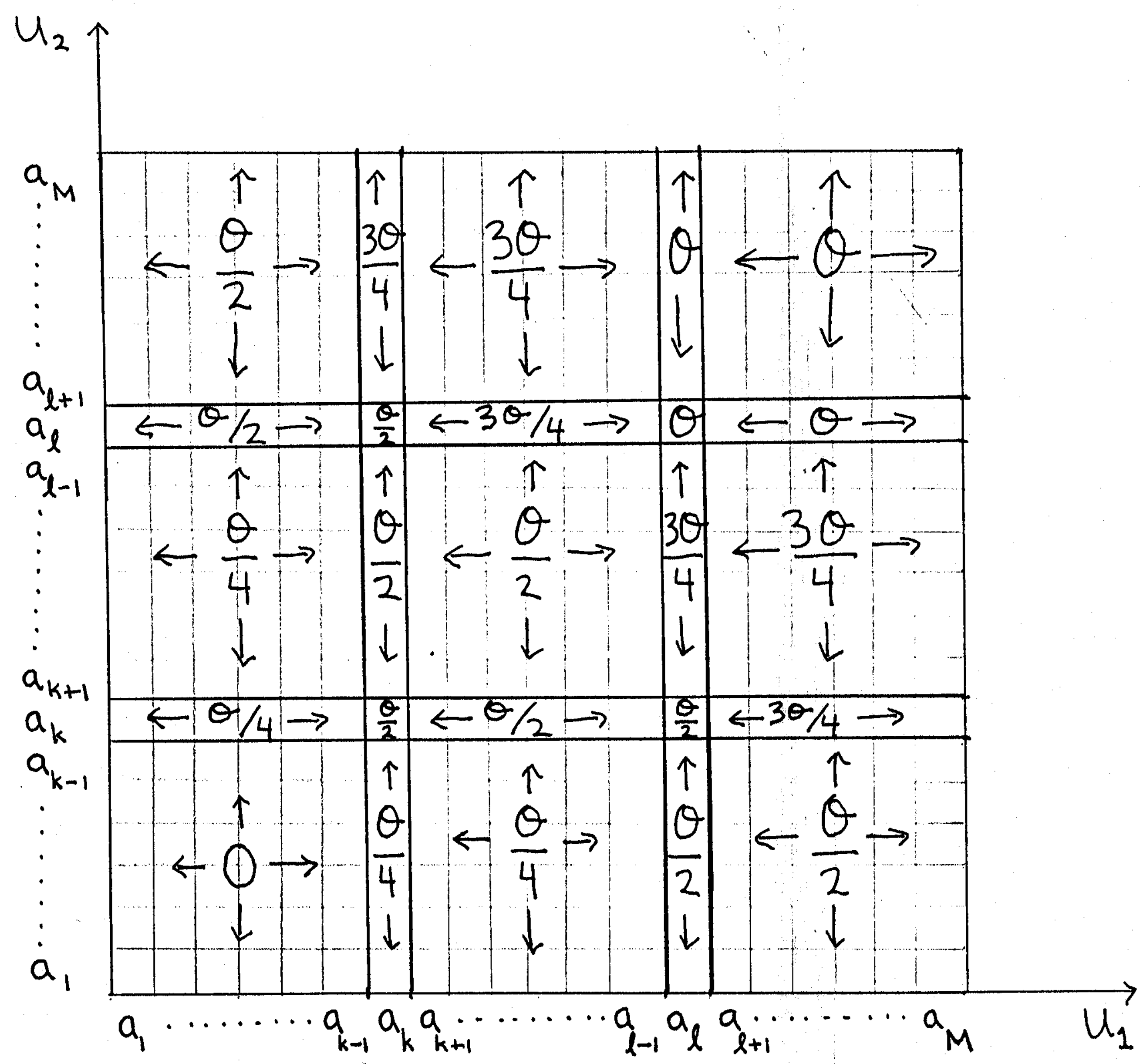
(ii) $\Rightarrow$ (iii) follows from the fact that ETI's weakly increase the expected value of all weakly complementary W, by definition.

(iii) $\Rightarrow$ (iv) is obvious.

(iv) $\Rightarrow$ (ii): Define a "negative ETI" as an ETI for which $d < 0$. Given symmetry and identity of the marginals, *any* distribution of (Y_1, Y_2) can be converted into that of (X_1, X_2) by a sequence of $\binom{n}{2}$ transformations, each an ETI or a negative ETI, and each corresponding to a different (k,l) pair (with $k < l$): this is a simple extension of the argument proving (i) $\Rightarrow$ (ii). Now suppose (ii) does not hold. Then in the sequence just described, there must be at least one negative ETI. Suppose that a negative ETI corresponds to the particular pair (k,l) (with $k < l$). Now consider the non-decreasing and weakly complementary function W^{kl} defined in Figure 1. (If $a_k = a_1$ or $a_l = a_M$, W^{kl} is simply the function depicted, truncated in the obvious fashion.) W^{kl} has been chosen so that the weak inequality (2) in the definition of weak complementarity is an equality for all (U^L, U^H) pairs except $U^L = a_k$, $U^H = a_l$: since the negative ETI corresponding to (k,l) reduces the expectation of W^{kl}, and since all other transformations in the sequence leave the expectation unchanged,

FIGURE 1: The function $W^{kl}(u_1, u_2)$,

where $u_1, u_2 \in \{a_1, \ldots, a_M\}$ and $a_k < a_l$



$EW^{kl}(X_1, X_2) < EW^{kl}(Y_1, Y_2)$. Thus (iv) does not hold. Q.E.D.

Remark A1: Dropping the reqirement that W be non-decreasing makes the proof much more transparent. To show (iii) $\Rightarrow$ (i), for any pair (k, l) with $k < l$ let $\hat{W}^{kl}$ be the weakly complementary function

$$\hat{W}^{kl} = \begin{cases} 0 & \text{if } (U_1, U_2) \in \{(a_k, a_l), (a_l, a_k)\} \\ 1 & \text{otherwise} \end{cases}$$

Then
$$\begin{aligned} &E\hat{W}^{kl}(X_1, X_2) - E\hat{W}^{kl}(Y_1, Y_2) \\ &= P(Y_1 = a_k, Y_2 = a_l) + P(Y_1 = a_l, Y_2 = a_k) \\ &\quad - P(X_1 = a_k, X_2 = a_l) - P(X_1 = a_l, X_2 = a_k) \\ &\geq 0 \qquad \text{by (iii).} \end{aligned}$$

(i) follows, by the symmetry of the joint distributions. It is trivial to verify that for $\hat{W}^{kl}$, the weak inequality (2) is an equality for all (U^L, U^H) pairs except (a_k, a_l).

Proof of Proposition 2: The text cites references in which the equivalence of (i), (ii), (iii), (vi), and (vii) is proved. We will prove that (iv) and (v) are equivalent to the other conditions by showing that (iii) $\Rightarrow$ (iv) and (iv) $\Rightarrow$ (i), and that (vi) $\Rightarrow$ (v) and (v) $\Rightarrow$ (i).

(iii) $\Rightarrow$ (iv): For all non-decreasing functions f^1 and f^2, each ETN produces a mean-preserving spread of the distribution of $f^1(U_1) + f^2(U_2)$. Therefore (iv) follows from the results of Rothschild and Stiglitz (1970).

(iv) $\Rightarrow$ (i): For any (k, l), let $f^1(U_1) = I_{\{U_1 \geq a_k\}}$ and $f^2(U_2) = I_{\{U_2 \geq a_l\}}$, so $f^1(U_1) + f^2(U_2)$ can assume the three values $\{0, 1, 2\}$. Using the definition (4) of second-order stochastic dominance, (iv) implies

$$P(I_{\{Y_1 \geq a_k\}} + I_{\{Y_2 \geq a_l\}} = 2) \leq P(I_{\{X_1 \geq a_k\}} + I_{\{X_2 \geq a_l\}} = 2),$$

which implies

$$P(Y_1 \geq a_k, Y_2 \geq a_l) \leq P(X_1 \geq a_k, X_2 \geq a_l).$$

(vi) $\Rightarrow$ (v): Let $W = f^1(U_1)f^2(U_2)$. If f^1 and f^2 are non-decreasing functions, W is strongly complementary. (vi) implies $E[f^1(X_1)f^2(X_2)] \geq E[f^1(Y_1)f^2(Y_2)]$, and since the distributions of (X_1, X_2) and (Y_1, Y_2) have identical marginals, we have

$$E[f^1(X_1)f^2(X_2)] - [Ef^1(X_1)][Ef^2(X_2)] \geq E[f^1(Y_1)f^2(Y_2)] - [Ef^1(Y_1)][Ef^2(Y_2)].$$

(v) $\Rightarrow$ (i): For any (k, l), let $f^1(U_1) = I_{\{U_1 \geq a_k\}}$ and $f^2(U_2) = I_{\{U_2 \geq a_l\}}$. Given the identity of the marginals of X and Y, (v) implies $P(X_1 \geq a_k, X_2 \geq a_l) \geq P(Y_1 \geq a_k, Y_2 \geq a_l)$. Q.E.D.

Proof of Proposition 3: The set of 2×2 bistochastic matrices is a one-parameter family:

$$\begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}$$

for $\alpha \in [0, 1]$. Welfare under any tournament with utility prizes t_1 and t_2 is deterministic and equals $W(t_1, t_2)$. Expected welfare under the RIS corresponding to a particular α is

$$\begin{aligned} &\alpha(1-\alpha)[W(t_1, t_1) + W(t_2, t_2)] + (\alpha^2 + (1-\alpha)^2)W(t_1, t_2) \\ &\geq \alpha(1-\alpha)[2W(t_1, t_2)] + (\alpha^2 + (1-\alpha)^2)W(t_1, t_2) \\ &= W(t_1, t_2) \end{aligned}$$

using symmetry and weak complementarity of W. Q.E.D.

Proof of Proposition 4: The proof that (i), (ii), and (iii) are equivalent is a straightforward generalization of the argument for equivalence of (i), (ii), and (iii) in Proposition 1 (using the proof that (iii) $\Rightarrow$ (i) given in Remark A1). (iii) $\Rightarrow$ (iv) is obvious, and (iv) $\Rightarrow$ (i) holds because the proof that (iii) $\Rightarrow$ (i) uses a symmetric W. Q.E.D.

Proof of Proposition 5: (i) $\Rightarrow$ (ii): For W in Class A,

$$EW(U_1, \ldots, U_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} EV^{ij}(U_i, U_j),$$

where the expectation in each term in the sum is taken only with respect to the pairwise joint distribution of U_i and U_j. For terms with $i = j$, $EV^{ii}(X_i, X_i) = EV^{ii}(Y_i, Y_i)$, since by assumption X and Y have identical marginals. For $i \neq j$,

(i) implies, given weak complementarity of V^{ij} and Proposition 1, $EV^{ij}(X_i, X_j) \geq EV^{ij}(Y_i, Y_j)$. Summing over all (i, j) pairs:

$$EW(X_1, \ldots, X_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} EV^{ij}(X_i, X_j) \geq \sum_{i=1}^{n}\sum_{j=1}^{n} EV^{ij}(Y_i, Y_j) = EW(Y_1, \ldots, Y_n).$$

(ii) $\Rightarrow$ (iii) is obvious.

(iii) $\Rightarrow$ (i): Suppose that (i) does not hold, so there exist i, j with $i \neq j$ and k, l with $k \neq l$ such that $P(X_i = a_k, X_j = a_l) > P(Y_i = a_k, Y_j = a_l)$. By Proposition 1, it follows that there exists a non-decreasing and weakly complementary V^{ij} such that $EV^{ij}(X_i, X_j) < EV^{ij}(Y_i, Y_j)$. Let $W(U_1, \ldots, U_n) = V^{ij}(U_i, U_j)$. Then W is non-decreasing and in Class A, and $EW(X_1, \ldots, X_n) < EW(Y_1, \ldots, Y_n)$, so (iii) does not hold.

Q.E.D.

Proof of Proposition 6: A tournament and its associated RIS do not in general yield joint distributions of utilities that are symmetric with respect to individuals. Nevertheless, since W is symmetric, expected social welfare is not affected by redistributing probability among outcome vectors which are permutations of one another, so as to make the joint distributions symmetric. This is accomplished by assigning to an outcome the average of the probabilities, under the original distribution, of all permutations of that outcome. Given a tournament and its associated RIS, let the distribution of Y be the "symmetrized" joint distribution of utilities under the tournament and the distribution of X the "symmetrized" joint distribution under the RIS (where the RIS is constructed from the tournament before either distribution is symmetrized). For both X and Y, symmetry implies that the probability that agent i wins prize k must equal $\frac{1}{n}$ for all agents and for all prizes. Thus X and Y have identical marginals. We will show that, for all tournaments, the distribution of X dominates that of Y in the sense of condition (SDA) ((i) of Proposition 5). It will then follow from Proposition 5 that the tournament axiom is satisfied for the set of symmetric welfare functions in Class A.

Let the prizes awarded by the given tournament be $(t_1, \ldots, t_n)$. For both X and Y, symmetry implies that the pairwise joint distributions are the same for all pairs (i, j) of individuals. The pairwise joint distributions derived from the symmetrized tournament assign zero probability to agents' i and j receiving the same prize and make each of the $n(n-1)$ unequal prize pairs equally likely. Thus

$$P(Y_i = t_k, Y_j = t_l) = \frac{1}{n(n-1)} \quad \forall i, j \in \{1, \ldots, n\}, i \neq j, \quad \forall k, l \in \{1, \ldots, n\}, k \neq l$$

In the symmetrized distribution under the RIS, the probability that agent i receives t_k and agent j receives t_l, $k \neq l$, equals the sum, over all ordered pairs (r, s) of agents, of the probability that r gets t_k and s gets t_l in the original RIS, divided by $n(n-1)$, the number of ordered pairs (r, s). Thus for all $i \neq j$ and $k \neq l$,

$$\begin{aligned} P(X_i = t_k, X_j = t_l) &= \frac{1}{n(n-1)} \sum_{r=1}^{n} \sum_{\substack{s=1 \\ s \neq r}}^{n} p_{rk} p_{sl} \\ &= \frac{1}{n(n-1)} \sum_{r=1}^{n} p_{rk}(1 - p_{rl}) \quad \text{since} \sum_{s=1}^{n} p_{sl} = 1 \\ &= \frac{1}{n(n-1)} (1 - \sum_{r=1}^{n} p_{rk} p_{rl}) \quad \text{since} \sum_{r=1}^{n} p_{rk} = 1 \\ &\leq \frac{1}{n(n-1)} = P(Y_i = t_k, Y_j = t_l) \end{aligned},$$

so (SDA) is satisfied. Q.E.D.

Proof of Proposition 7: Since $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ have identical marginals, $E(\sum_{i=1}^n f^i(X_i)) = E(\sum_{i=1}^n f^i(Y_i))$, for all $\{f^1, \ldots, f^n\}$. The equivalence of (i), (ii), and (iii), both with and without the restriction to identical functions f^i, follows from the results of Hadar and Russell (1969) and Rothschild and Stiglitz (1970) on second-order stochastic dominance in one dimension, adapted for convex, rather than concave, objective functions and applied successively for each allowable set of functions $\{f^1, \ldots, f^n\}$. Q.E.D.

Proof of Proposition 8: All symmetric welfare functions in Class B can be written as $\widetilde{W}(\sum_i f(U_i))$, for some non-decreasing function f (the same for all i) and some convex function $\widetilde{W}$. Let the distribution of Y represent the joint distribution of utilities under a given tournament and that of X the joint distribution under the corresponding RIS. By construction, Y and X have identical marginals, implying

$E(\sum_i f(Y_i)) = E(\sum_i f(X_i))$. Under the tournament, $\sum_i f(Y_i)$ is deterministic for all f; under the corresponding RIS, $\sum_i f(X_i)$ is variable. Therefore (SDB) holds with the restriction that f is the same for all i, so by Proposition 7, (ii) holds with this restriction, and the tournament axiom is satisfied for the set of symmetric welfare functions in Class B. Q.E.D.

Proof of Lemma 1: We will prove the lemma for the special case in which the distributions of X and Y are symmetric over individuals and $(z_1, \ldots, z_n) = (\bar{z}, \ldots, \bar{z})$. For the general case, the proof proceeds along exactly the same lines, but the absence of symmetry makes the notation significantly more complicated. We will also prove only the claim for W^z in $\bar{C}^-$; for W^z in $\bar{C}^+$, the steps are the same, with inequality signs reversed in the appropriate places.

Given $z = (\bar{z}, \ldots, \bar{z})$ and an arbitrary permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$, define

$$\Delta p(k) \equiv P(X_{i_1} > \bar{z}, \ldots, X_{i_k} > \bar{z}, X_{i_{k+1}} \leq \bar{z}, \ldots, X_{i_n} \leq \bar{z}) \\ - P(Y_{i_1} > \bar{z}, \ldots, Y_{i_k} > \bar{z}, Y_{i_{k+1}} \leq \bar{z}, \ldots, Y_{i_n} \leq \bar{z}).$$

(Since the distributions of X and Y are assumed symmetric, $\Delta p(k)$ is the same for all permutations $(i_1, \ldots, i_n)$.) In this notation, we want to show that, for $z = (\bar{z}, \ldots, \bar{z})$ and the corresponding W^z in $\bar{C}^-$,

$$EW^z(X) - EW^z(Y) = \Delta p(0). \tag{A1}$$

Note that W^z assumes the value $n-k$ in $\binom{n}{k-1}$ regions, since there are $\binom{n}{k-1}$ regions where $U_i > \bar{z}$ for exactly $n-k+1$ values of i. Then

$$EW^z(X) - EW^z(Y) = \sum_{k=1}^{n-1} (n-k) \binom{n}{k-1} \Delta p(n-k+1). \tag{A2}$$

Showing that the right-hand side of (A2) equals $\Delta p(0)$ requires three steps.

Step 1: Use the equality of the marginals of X and Y to derive an expression for $\Delta p(1)$ in terms of $\Delta p(2), \ldots, \Delta p(n)$:

$$\Delta p(1) = -\sum_{k=1}^{n-1} \binom{n-1}{k-1} \Delta p(n-k+1) \tag{A3}$$

Step 2: Use the equality of the marginals to derive an expression for $\Delta p(0)$ in terms of $\Delta p(1), \ldots, \Delta p(n-1)$:

$$\Delta p(0) = -\sum_{k=2}^{n} \binom{n-1}{k-2} \Delta p(n-k+1) \tag{A4}$$

Step 3: Use (A3) to substitute for $\Delta p(1)$ in (A4). This yields a formula for $\Delta p(0)$ in terms of $\Delta p(2), \ldots, \Delta p(n)$ which we show to be equivalent to the right-hand side of (A2):

$$\begin{aligned}
\Delta p(0) &= -\sum_{k=2}^{n-1} \binom{n-1}{k-2} \Delta p(n-k+1) \\
&\quad + (n-1)\sum_{k=1}^{n-1} \binom{n-1}{k-1} \Delta p(n-k+1) \\
&= (n-1)\Delta p(n) + \sum_{k=2}^{n-1} \left[(n-1)\binom{n-1}{k-1} - \binom{n-1}{k-2} \right] \Delta p(n-k+1) \\
&= (n-1)\Delta p(n) + \sum_{k=2}^{n-1} (n-k) \binom{n}{k-1} \Delta p(n-k+1) \\
&= \sum_{k=1}^{n-1} (n-k) \binom{n}{k-1} \Delta p(n-k+1) \qquad \text{Q.E.D.}
\end{aligned}$$

Proof of Proposition 9: We will show (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i).

(i) $\Rightarrow$ (iii) Let j index the points $z = (z_1, \ldots, z_n)$ in the support of X and Y, and let the corresponding elements of C^+, C^-, $\bar{C}^-$, and $\bar{C}^+$ be denoted W_+^j, W_-^j, $\bar{W}_-^j$, and $\bar{W}_+^j$. Then any W in Class C can be expressed as

$$W = \sum_j \alpha_j W_+^j + \sum_j \beta_j W_-^j + \sum_j \gamma_j \bar{W}_-^j + \sum_j \delta_j \bar{W}_+^j, \qquad \text{where} \quad \alpha_j, \beta_j, \gamma_j, \delta_j \geq 0 \quad \forall j.$$

$$\begin{aligned}
EW(X) - EW(Y) &= \sum_j \alpha_j [EW_+^j(X) - EW_+^j(Y)] + \sum_j \beta_j [EW_-^j(X) - EW_-^j(Y)] \\
&\quad + \sum_j \gamma_j [E\bar{W}_-^j(X) - E\bar{W}_-^j(Y)] + \sum_j \delta_j [E\bar{W}_+^j(X) - E\bar{W}_+^j(Y)]
\end{aligned}$$

For each j, the first and fourth terms in square brackets take the form $P(X_i \geq z_i \forall i) - P(Y_i \geq z_i \forall i)$, and the second and third terms in square brackets take the form

$P(X_i \le z_i \forall i) - P(Y_i \le z_i \forall i)$, for the $(z_1, \ldots, z_n)$ corresponding to the index j (using Lemma 1), and therefore (i) implies (iii).

(iii) $\Rightarrow$ (ii): We will show that for any set of non-negative-valued functions $\{f^1, \ldots, f^n\}$ such that either all f^i are non-decreasing or all f^i are non-increasing, $W = \prod_{i=1}^n f^i(U_i)$ belongs to Class C, so $E[\prod_{i=1}^n f^i(X_i)] \ge E[\prod_{i=1}^n f^i(Y_i)]$ follows from (iii).

Write the support of X and Y as $\{a_1, \ldots, a_M\} \times \cdots \times \{a_1, \ldots, a_M\}$. Suppose first that each f^i is non-negative-valued and non-decreasing. Then each f^i can be represented as

$$f^i(U_i) = \sum_{k_i=1}^{M} \theta^i_{k_i} I_{\{U_i \ge a_{k_i}\}}, \qquad \text{where} \quad \theta^i_{k_i} \ge 0 \qquad \forall i,\ \forall k_i.$$

The product $\prod_{i=1}^n f^i(U_i)$ therefore consists of a sum of terms of the form

$$\prod_{i=1}^{n} \theta^i_{k_i} I_{\{U_i \ge a_{k_i}\}} = \left(\prod_{i=1}^{n} \theta^i_{k_i}\right) I_{\{U_i \ge a_{k_i} \forall i\}}, \qquad \text{for some} \quad (k_1, \ldots, k_n).$$

Since $I_{\{U_i \ge a_{k_i} \forall i\}}$ belongs to C^+ for every $(k_1, \ldots, k_n)$, $\prod_{i=1}^n f^i(U_i)$ belongs to C.

Similarly, suppose that each f^i is non-negative-valued and non-increasing. Then each f^i can be represented as

$$f^i(U_i) = \sum_{k_i=1}^{M} \theta^i_{k_i} I_{\{U_i \le a_{k_i}\}}, \qquad \text{where} \quad \theta^i_{k_i} \ge 0 \qquad \forall i,\ \forall k_i.$$

Thus $\prod_{i=1}^n f^i(U_i)$ consists of a sum of terms of the form

$$\prod_{i=1}^{n} \theta^i_{k_i} I_{\{U_i \le a_{k_i}\}} = \left(\prod_{i=1}^{n} \theta^i_{k_i}\right) I_{\{U_i \le a_{k_i} \forall i\}}, \qquad \text{for some} \quad (k_1, \ldots, k_n).$$

Since $I_{\{U_i \le a_{k_i} \forall i\}}$ belongs to C^- for every $(k_1, \ldots, k_n)$, $\prod_{i=1}^n f^i(U_i)$ belongs to C.

(ii) $\Rightarrow$ (i): For any $(z_1, \ldots, z_n)$, let $f^i(U_i) = I_{\{U_i \ge z_i\}}$ for all i. Then $\prod_{i=1}^n f^i(U_i) = I_{\{U_i \ge z_i \forall i\}}$, so $E[\prod_{i=1}^n f^i(X_i)] \ge E[\prod_{i=1}^n f^i(Y_i)]$ implies $P(X_i \ge z_i \forall i) \ge P(Y_i \ge z_i \forall i)$. Similarly, if we let $f^i(U_i) = I_{\{U_i \le z_i\}}$ for all i, then $E[\prod_{i=1}^n f^i(X_i)] \ge E[\prod_{i=1}^n f^i(Y_i)]$ implies $P(X_i \le z_i \forall i) \ge P(Y_i \le z_i \forall i)$.

(iii) $\Rightarrow$ (iv) is obvious.

(iv) $\Rightarrow$ (i): Every W in C^+ or $\bar{C}^-$ is non-decreasing. With W successively equal to every member of C^+ and then to every member of $\bar{C}^-$, the inequality $EW(X) - EW(Y) \geq 0$ implies (i) (using Lemma 1). Q.E.D.

Proof of Proposition 10: Suppose a given tournament awards prizes $(t_1, \ldots, t_n)$, with $t_1 \geq \ldots \geq t_n$, so t_j is the prize with rank j (from the top). For ease of notation, describe a given outcome vector $(U_1, \ldots, U_n)$, where each $U_i \in \{t_1, \ldots, t_n\}$, by the vector of ranks of the prizes, $(r_1, \ldots, r_n)$, where each $r_i \in \{1, \ldots, n\}$. Thus $U_i = t_j$ implies $r_i = j$.

For symmetric W, expected welfare is not affected by redistributing probability mass in joint distributions to make them symmetric (as in the proof of Proposition 6). Given a tournament and its associated RIS, let the distribution of Y be the "symmetrized" distribution of ranks under the tournament and the distribution of X the "symmetrized" distribution of ranks under the RIS. By symmetry, X and Y have identical marginals. We will show that, for all tournaments, the distribution of X dominates that of Y in the sense of condition (SDC). Then (SDC) will also be satisfied if X and Y represent symmetrized distributions of utilities instead of ranks, and satisfaction of the tournament axiom for the set of symmetric functions in Class C will follow from Proposition 9.

Furthermore, if we can establish that for all bistochastic matrices P and for all vectors of ranks $(r_1, \ldots, r_n)$, the distributions of X and Y satisfy $P(X_i \leq r_i \forall i) \geq P(Y_i \leq r_i \forall i)$, then it will follow that for all $(r_1, \ldots, r_n)$, $P(X_i \geq r_i \forall i) \geq P(Y_i \geq r_i \forall i)$. To see this, observe first that if we rearrange the columns of a bistochastic matrix so that for $j = 1, \ldots, n$ the columns corresponding to ranks j and $n+1-j$ exchange places, the new matrix is still bistochastic. Then any inequality that we establish for the event "each individual i does at least as well as rank r_i" for all bistochastic matrices will imply the validity of the analogous inequality for the event "each individual i does at least as badly as rank $n+1-r_i$". Since we will establish the former type of inequality for *all* vectors of ranks, the latter type will be valid for all such vectors as well.

The symmetrized joint distribution of ranks under a tournament assigns to each permutation of $(1, \ldots, n)$ a probability $\frac{1}{n!}$. By symmetry, $P(Y_i \leq r_i \forall i)$ is the same for all $(r_1, \ldots, r_n)$ that are permutations of one another.

Lemma 2: *Let $(r_1, \ldots, r_n)$ be ordered so that $r_1 \leq \cdots \leq r_n$.*

(i) If there exists a j such that $r_j < j$, then $P(Y_i \leq r_i \forall i) = 0$.

(ii) If for all j, $r_j \geq j$, then

$$P(Y_i \leq r_i \forall i) = \frac{1}{n!} \prod_{j=1}^{n} \left[\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n-j) \right].$$

Proof of Lemma 2: $P(Y_i \leq r_i \ \forall\ i)$ equals $\frac{1}{n!}$ times the number of permutations $(j_1, \ldots, j_n)$ of $(1, \ldots, n)$ such that $j_i \leq r_i$ for all i. Let us call such permutations admissible. To prove (i), observe that when $r_1 \leq \cdots \leq r_n$ and for some j, $r_j < j$, there are no admissible permutations. To prove (ii), we count admissible permutations given that $r_1 \leq \cdots \leq r_n$ and $r_j \geq j$ for all j. Observe that the number of allowable positions for n is $\sum_{k=1}^{n} I_{\{r_k = n\}}$. The number of allowable positions for $n-1$ is $\sum_{k=1}^{n} I_{\{r_k \geq n-1\}} - 1$ (we must subtract 1 to take account of the position already occupied by n). Similarly, the number of allowable positions for j is $\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n-j)$ (subtracting $n-j$ takes account of the positions already occupied by n, $n-1, \ldots, j+1$). Note that $r_1 \leq \cdots \leq r_n$ and $r_j \geq j$ for all j imply that $\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n-j) \geq 1$ for all j. The total number of admissible permutations for this case is

$$\prod_{j=1}^{n} \left[\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n-j) \right].$$

Q.E.D.

Now consider the symmetrized joint distribution of ranks under the RIS. By symmetry, $P(X_i \leq r_i \forall i)$ is the same for all $(r_1, \ldots, r_n)$ that are permutations of one another.

Lemma 3: *Let $(r_1, \ldots, r_n)$ be ordered so that $r_1 \leq \cdots \leq r_n$ and suppose that $r_j \geq j$ for all j. For all bistochastic matrices P, the derived distribution of $(X_1, \ldots, X_n)$ satisfies*

$$P(X_i \leq r_i \forall i) \geq \frac{1}{n!} \prod_{k=1}^{n} (r_k - k + 1).$$

Proof of Lemma 3: Given $(r_1, \ldots, r_n)$, to calculate $P(X_i \leq r_i \forall i)$, we first take a particular permutation $(i_1, \ldots, i_n)$ of individuals $(1, \ldots, n)$ and, from the matrix P, calculate the probability that each individual i_k does at least as well as rank r_k under the original, asymmetric RIS. We then sum this probability over all $n!$ permutations $(i_1, \ldots, i_n)$ and divide by $n!$.

Formally, given any bistochastic matrix P and given a particular permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$, the probability, under the original, asymmetric RIS, that each agent i_k does at least as well as rank r_k is

$$\prod_{k=1}^{n} \sum_{j=1}^{r_k} p_{i_k j}.$$

Then

$$n!P(X_i \leq r_i \forall i) = \sum_{\substack{(i_1,\ldots,i_n), \\ \text{permutations of} \\ (1,\ldots,n)}} \prod_{k=1}^{n} \sum_{j=1}^{r_k} p_{i_k j}. \tag{A5}$$

Define, for $m \leq n$,

$$G(r_1, \ldots, r_m; n) \equiv \sum_{\substack{(i_1,\ldots,i_m), \\ \text{permutations of} \\ m \text{ from } \{1,\ldots,n\}}} \prod_{k=1}^{m} \sum_{j=1}^{r_k} p_{i_k j},$$

the outer sum being taken over all orderings $(i_1, \ldots, i_m)$ of m distinct elements from $\{1, \ldots, n\}$. For notational convenience, we will write

$$G(r_1, \ldots, r_m; n) \equiv \sum_{i_1 \neq \ldots \neq i_m}^{n} \prod_{k=1}^{m} \sum_{j=1}^{r_k} p_{i_k j}.$$

Observe that $G(r_1, \ldots, r_n; n)$ equals the right-hand side of (A5). We will show by induction that

$$G(r_1, \ldots, r_m; n) \geq \prod_{k=1}^{m} (r_k - k + 1) \quad \forall\, m \leq n$$

and hence that

$$G(r_1, \ldots, r_n; n) \geq \prod_{k=1}^{n} (r_k - k + 1),$$

which is the inequality we seek.

For $m = 1$,

$$\begin{aligned}
G(r_1; n) &\equiv \sum_{i_1=1}^{n} \sum_{j=1}^{r_1} p_{i_1 j} \\
&= \sum_{i_1=1}^{n} (1 - \sum_{j=r_1+1}^{n} p_{i_1 j}) \qquad \text{since } \sum_{j=1}^{n} p_{i_1 j} = 1 \\
&= n - \sum_{j=r_1+1}^{n} \sum_{i_1=1}^{n} p_{i_1 j} \\
&= n - \sum_{j=r_1+1}^{n} 1 \qquad \text{since } \sum_{i_1=1}^{n} p_{i_1 j} = 1 \\
&= n - (n - (r_1 + 1) + 1) \\
&= r_1 - 1 + 1.
\end{aligned}$$

Now assume that

$$G(r_1, \ldots, r_{m-1}; n) \equiv \sum_{i_1 \neq \ldots \neq i_{m-1}}^{n} \prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \geq \prod_{k=1}^{m-1} (r_k - k + 1). \qquad (A6)$$

Then

$$\begin{aligned}
G(r_1, \ldots, r_m; n) \equiv& \sum_{i_1 \neq \ldots \neq i_m}^{n} \prod_{k=1}^{m} \sum_{j=1}^{r_k} p_{i_k j} \\
=& \sum_{i_1 \neq \ldots \neq i_m}^{n} \left[\prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \right] - \sum_{i_1 \neq \ldots \neq i_m}^{n} \left[\prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \right] \left[\sum_{j=r_m+1}^{n} p_{i_m j} \right] \\
=& (n - m + 1) \sum_{i_1 \neq \ldots \neq i_{m-1}}^{n} \left[\prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \right] \\
&- \sum_{i_1 \neq \ldots \neq i_{m-1}}^{n} \left[\prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \right] \left[\sum_{j=r_m+1}^{n} \left(1 - \sum_{S=1}^{m-1} p_{i_S j} \right) \right] \\
=& [n - m + 1 - (n - (r_m + 1) + 1)] \sum_{i_1 \neq \ldots \neq i_{m-1}}^{n} \prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \\
&+ \sum_{i_1 \neq \ldots \neq i_{m-1}}^{n} \left[\prod_{k=1}^{m-1} \sum_{j=1}^{r_k} p_{i_k j} \right] \left[\sum_{j=r_m+1}^{n} \sum_{S=1}^{m-1} p_{i_S j} \right] \\
\geq& (r_m - m + 1) G(r_1, \ldots, r_{m-1}; n) \\
\geq& (r_m - m + 1) \prod_{k=1}^{m-1} (r_k - k + 1) \quad \text{using (A6) and } r_m \geq m \quad \forall m
\end{aligned}$$

$$= \prod_{k=1}^{m} (r_k - k + 1).$$

Q.E.D.

When $r_1 \leq \cdots \leq r_n$ and for some j, $r_j < j$, then by Lemma 2(i), $P(Y_i \leq r_i \forall i) = 0$, so $P(X_i \leq r_i \forall i) \geq P(Y_i \leq r_i \forall i)$. For the case when $r_j \geq j$ for all j, the following lemma, combined with Lemmas 2 and 3, implies $P(X_i \leq r_i \forall i) \geq P(Y_i \leq r_i \forall i)$.

Lemma 4: *Let $(r_1, \ldots, r_n)$ be ordered so that $r_1 \leq \cdots \leq r_n$ and suppose that $r_j \geq j$ for all j. Then*

$$\prod_{k=1}^{n} (r_k - k + 1) = \prod_{j=1}^{n} \left[\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n - j) \right].$$

Proof of Lemma 4: Given $r_1 \leq \cdots \leq r_n$ and $r_j \geq j$ for all j, define

$$R(r_1, \ldots, r_n) \equiv \prod_{k=1}^{n} (r_k - k + 1)$$

and

$$T(r_1, \ldots, r_n) \equiv \prod_{j=1}^{n} \left[\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n - j) \right].$$

We can simplify T by writing

$$\sum_{k=1}^{n} I_{\{r_k \geq j\}} - (n - j) = \sum_{k=1}^{j} I_{\{r_k \geq j\}} + \sum_{k=j+1}^{n} I_{\{r_k \geq j\}} - (n - j)$$

$$= \sum_{k=1}^{j} I_{\{r_k \geq j\}} \quad \text{since} \quad r_k \geq j \quad \forall\, k \geq j + 1$$

Thus

$$T(r_1, \ldots, r_n) = \prod_{j=1}^{n} \left(\sum_{k=1}^{j} I_{\{r_k \geq j\}} \right).$$

We now use a graphical representation to show that for each of the n factors in T, there is a corresponding, distinct factor in R with the same value. This clearly implies that T equals R. On a grid, plot the points (k, r_k), for $k = 1, \ldots, n$. Since $r_1 \leq \cdots \leq r_n$, the function mapping k into r_k is non-decreasing, and since $r_k \geq k$ for all k, all of the points lie on or above the 45° line. Draw a path connecting the points according to the following sequence of rules (see Figure 2):

(i) Use horizontal segments to connect any two points (k, r_k) and $(k+1, r_{k+1})$ for which $r_k = r_{k+1}$.

(ii) Use diagonal segments parallel to the 45° line to connect any two points (k, r_k) and $(k+1, r_{k+1})$ for which $r_k + 1 = r_{k+1}$.

(iii) For points (k, r_k) and $(k+1, r_{k+1})$ such that $r_k+1 < r_{k+1}$, draw vertical segments downwards from $(k+1, r_{k+1})$ until the point $(k+1, r_k+1)$ is reached. Then use a diagonal segment parallel to the 45° line to connect (k, r_k) with $(k+1, r_k+1)$.

(iv) If $r_1 > 1$, draw vertical segments downwards to connect $(1, r_1)$ with $(1, 1)$.

Let S denote the set of points (i, z_i), where i and z_i are both integers, on the path generated by these rules. The path never goes below the 45° line. Think of the path as having direction, from (n, n) to $(1, 1)$. Note that the set of points $\{k, r_k\}_{k=1,\ldots,n}$ is in general a strict subset of S.

The points in S fall into three mutually exclusive and exhaustive classes:

$$\begin{aligned}
&\text{Class I: } \{(i, z_i) \in S \mid \exists (l, z_l) \in S \text{ with } l < i, z_l = z_i\} \\
&\text{Class II: } \{(i, z_i) \in S \mid \exists (l, z_l) \in S \text{ with } l = i, z_l > z_i\} \\
&\text{Class III:} \{(i, z_i) \in S \mid \nexists (l, z_l) \in S \text{ with } l < i, z_l = z_i \\
&\qquad\qquad \text{and} \nexists (l, z_l) \in S \text{ with } l = i, z_l > z_i\}
\end{aligned}$$

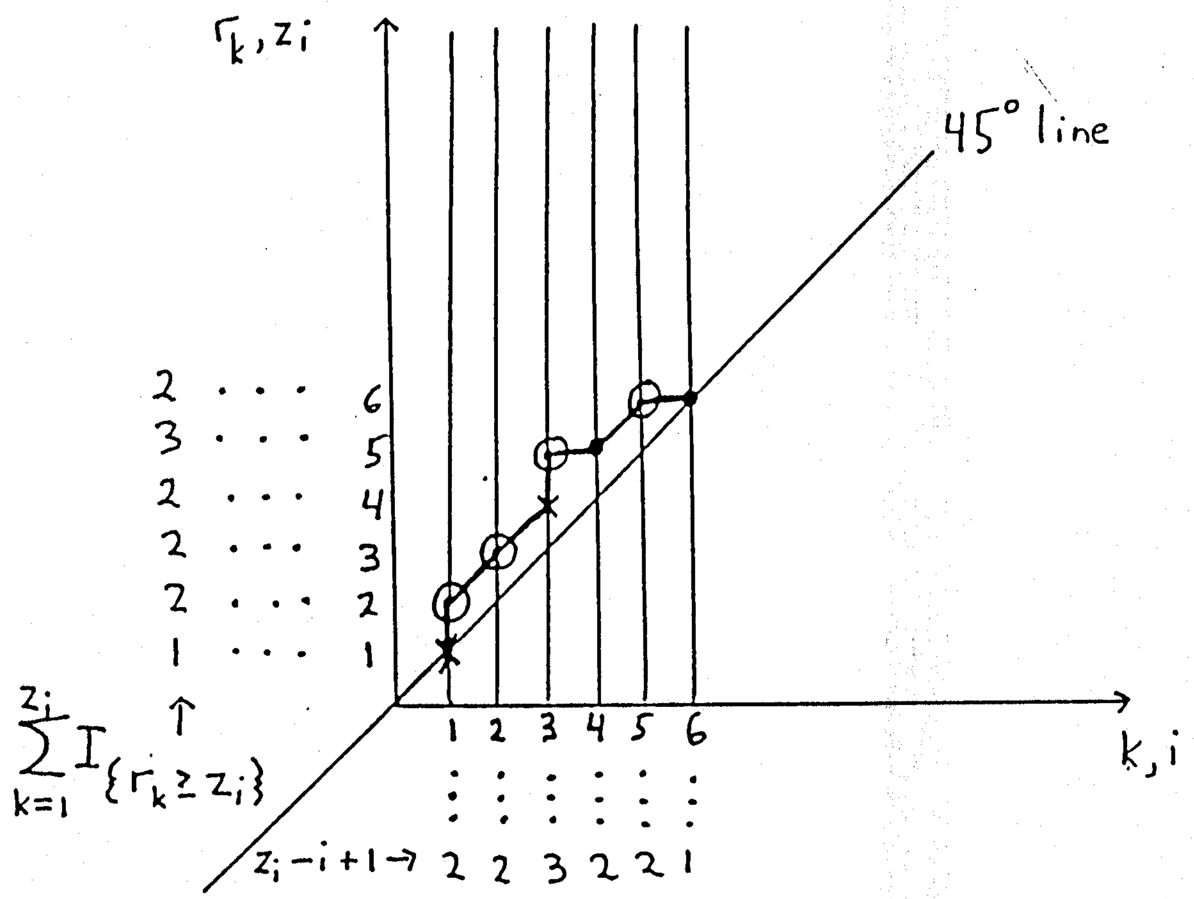
The fact that Classes I and II are disjoint follows from the rules (i)–(iv). The union of Classes I and III is the set of points $\{k, r_k\}_{k=1,\ldots,n}$. From this it follows immediately that the sets of i coordinates of points in I and III are disjoint, with union $\{1, \ldots, n\}$. The sets of z_i coordinates of points in II and III are also disjoint, with union $\{1, \ldots, n\}$ and with no z_i value appearing more than once in either set. Therefore, there are equal numbers of points in I and II. Figure 2 shows S and its three subsets for a particular vector $(r_1, \ldots, r_n)$ when $n = 6$.

Lemma 4.1: *For points (i, z_i) in Class III,*

$$z_i - i + 1 = \sum_{k=1}^{z_i} I_{\{r_k \geq z_i\}}.$$

FIGURE 2: Example of method of proof of Lemma 4:

$n=6$ and $(r_1, r_2, r_3, r_4, r_5, r_6) = (2, 3, 5, 5, 6, 6)$



- • Class I: $\{(6,6), (4,5)\}$
- ✗ Class II: $\{(3,4), (1,1)\}$
- ⊙ Class III: $\{(5,6), (13,5), (2,3), (1,2)\}$

Proof of Lemma 4.1: The right-hand side above equals the number of points in the set $\{k, r_k\}_{k=1,\dots,n}$ for which $k \le z_i$ and $r_k \ge z_i$. Since by assumption, $r_k \ge k$ for all k, the point (z_i, r_{z_i}) necessarily satisfies these conditions. Then it can be checked that the right-hand side equals 1 plus the horizontal distance of (i, z_i) from the 45° line, which equals 1 plus the vertical distance of (i, z_i) from the 45° line, which equals $z_i - i + 1$. Q.E.D.

Lemma 4.2: *The points in Classes I and II can be put into one-to-one correspondence $\{I_1, I_2, \dots, I_L\}$, $\{II_1, II_2, \dots, II_L\}$ so that $z_i - i + 1$ evaluated for I_l equals $\sum_{k=1}^{z_i} I_{\{r_k \ge z_i\}}$ evaluated for II_l, for all $l = 1, \dots, L$.*

Proof of Lemma 4.2: Each point in Class I is at the right-hand endpoint (the beginning) of a horizontal segment of length 1, and each point in Class II is at the bottom endpoint (the end) of a vertical segment of length 1. Since the path begins and ends on the 45° line and since the only diagonal segments used are parallel to the 45° line, the number of horizontal segments must equal the number of vertical segments, thus verifying that Classes I and II contain equal numbers of points.

Moreover, for each horizontal segment beginning m units to the left of the 45° line, there is a vertical segment ending m units above the 45° line. This is the one-to-one correspondence we seek:

(i) A point (i, z_i) in Class I m units to the left of the 45° line is m units above the 45° line, and therefore $z_i - i + 1 = m + 1$.

(ii) A point (i, z_i) in Class II m units above the 45° line is m units to the left of the 45° line, and it is easily verified that $\sum_{k=1}^{z_i} I_{\{r_k \ge z_i\}}$, the number of points (k, r_k) with $k \le z_i$ and $r_k \ge z_i$, is $m + 1$. Q.E.D.

Combining Lemmas 4.1 and 4.2, we conclude that the values of $z_i - i + 1$ associated with points in Classes I or III can be put into one-to-one correspondence with the values of $\sum_{k=1}^{z_i} I_{\{r_k \ge z_i\}}$ associated with points in Classes II or III. Since the union of I and III is $\{k, r_k\}_{k=1,\dots n}$ and since the set of z_i values for the union of II and III is

$\{1, \dots, n\}$, it follows that

$$\prod_{k=1}^{n}(r_k - k + 1) = \prod_{j=1}^{n}\left(\sum_{k=1}^{j} I_{\{r_k \geq j\}}\right). \qquad \text{Q.E.D.}$$

Lemma 4 completes the proof of Proposition 10. Q.E.D.

Proof of Proposition 11:

(SDB) $\Rightarrow$ (SDC): For any $(z_1, \dots, z_n)$, let $f^i(U_i) = I_{\{U_i \geq z_i\}}$ for all i. Then $\sum_{i=1}^{n} f^i(U_i)$ takes values in the set $\{0, \dots, n\}$, and second-order stochastic dominance of the distribution of $\sum_{i=1}^{n} f^i(Y_i)$ over that of $\sum_{i=1}^{n} f^i(X_i)$ implies, using the definition (4),

$$P(\sum_{i=1}^{n} I_{\{Y_i \geq z_i\}} = n) \leq P(\sum_{i=1}^{n} I_{\{X_i \geq z_i\}} = n),$$

from which it follows that $P(Y_i \geq z_i \forall i) \leq P(X_i \geq z_i \forall i)$. Similarly, given $(z_1, \dots, z_n)$, now define $f^i(U_i) = I_{\{U_i > z_i\}}$ for all i. (SDB) implies, using the definition (3),

$$P(\sum_{i=1}^{n} I_{\{Y_i > z_i\}} = 0) \leq P(\sum_{i=1}^{n} I_{\{X_i > z_i\}} = 0),$$

which implies $P(Y_i \leq z_i \forall i) \leq P(X_i \leq z_i \forall i)$. Thus (SDC) is satisfied.

(SDC) $\not\Rightarrow$ (SDB): Figures 3a and 3b show symmetric, four-dimensional distributions with identical marginals, for X and Y respectively, in which each component has a two-point support $\{U^L, U^H\}$. (SDC) is satisfied since

$$P(X_1 = U^H, X_2 = U^H, X_3 = U^H, X_4 = U^H) = \frac{1}{6} \geq 0 = P(Y_1 = U^H, Y_2 = U^H, Y_3 = U^H, Y_4 = U^H)$$

$$P(X_1 \geq U^L, X_2 = U^H, X_3 = U^H, X_4 = U^H) = \frac{1}{6} \geq \frac{1}{8} = P(Y_1 \geq U^L, Y_2 = U^H, Y_3 = U^H, Y_4 = U^H)$$

$$P(X_1 \geq U^L, X_2 \geq U^L, X_3 = U^H, X_4 = U^H) = \frac{1}{6} + \frac{1}{9} \geq 2(\frac{1}{8}) = P(Y_1 \geq U^L, Y_2 \geq U^L, Y_3 = U^H, Y_4 = U^H)$$

$$P(X_1 = U^L, X_2 = U^L, X_3 = U^L, X_4 = U^L) = \frac{1}{6} \geq 0 = P(Y_1 = U^L, Y_2 = U^L, Y_3 = U^L, Y_4 = U^L)$$

$$P(X_1 \leq U^H, X_2 = U^L, X_3 = U^L, X_4 = U^L) = \frac{1}{6} \geq \frac{1}{8} = P(Y_1 \leq U^H, Y_2 = U^L, Y_3 = U^L, Y_4 = U^L)$$

and since all of the other inequalities comprising (SDC) follow either from the symmetry of the distributions or from the equality of the marginals.

FIGURE 3

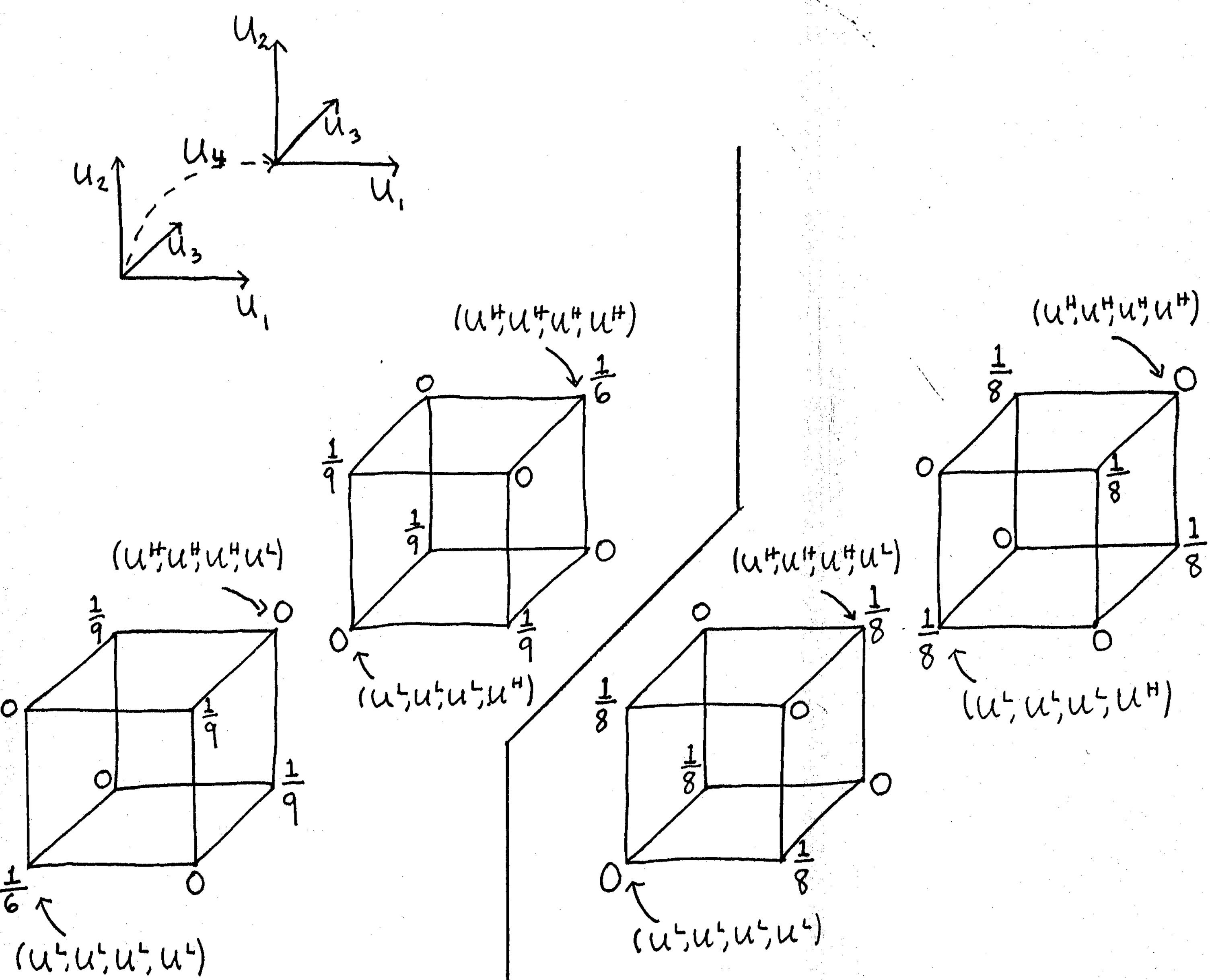


FIGURE 3a: (X_1, X_2, X_3, X_4)

FIGURE 3b: (Y_1, Y_2, Y_3, Y_4)

Example showing that (SDC) $\not\Rightarrow$ (SDB)

To show that (SDB) is violated, let $f^i(U^H) = 1$ for all i and $f^i(U^L) = 0$ for all i. Using the definition (4), (SDB) requires

$$2P(\sum_i f^i(Y_i) = 4) + P(\sum_i f^i(Y_i) = 3) \le 2P(\sum_i f^i(X_i) = 4) + P(\sum_i f^i(X_i) = 3).$$

However, the left-hand side equals $2 \cdot 0 + 4 \cdot \frac{1}{8} = \frac{1}{2}$, while the right-hand side equals $2 \cdot \frac{1}{6} + 4 \cdot 0 = \frac{1}{3}$, so (SDB) is violated.

(SDA) $\not\Rightarrow$ (SDC) and (SDA) $\not\Rightarrow$ (SDB): Figures 4a and 4b show symmetric, three-dimensional distributions with identical marginals, for X and Y respectively, in which each component has a two-point support $\{U^L, U^H\}$. (SDA) is satisfied since

$$P(X_1 = U^H, X_2 = U^L) = \frac{1}{6} \le \frac{1}{4} = P(Y_1 = U^H, Y_2 = U^L)$$

and since all of the other inequalities comprising (SDA) follow from the symmetry of the distributions. (SDC) is violated since

$$P(X_1 = U^H, X_2 = U^H, X_3 = U^H) = \frac{1}{6} < \frac{1}{4} = P(Y_1 = U^H, Y_2 = U^H, Y_3 = U^H).$$

Since (SDB) $\Rightarrow$ (SDC), it follows from (SDA) $\not\Rightarrow$ (SDC) that (SDA) $\not\Rightarrow$ (SDB).

(SDB) $\not\Rightarrow$ (SDA) and (SDC) $\not\Rightarrow$ (SDA): By Proposition 2, (SDB) and (SDC) are equivalent for $n = 2$, and we have already observed in the text that for $n = 2$ neither condition implies (SDA). (Figures 5a and 5b provide a two-dimensional example in which (SDC) is satisfied but (SDA) is violated.) Q.E.D.

Proof of Proposition 12: The expected value of any W in Class B or Class C is (weakly) increased by any GETN. However, for W in Class A, V^{ij} is not required to be strongly complementary, so the expectation of W can be strictly decreased by some GETN. Therefore, $W \in A$ does not imply either $W \in B$ or $W \in C$.

The remaining claims are proved by two examples. For each example, $n = 3$, each U_i can assume three possible values ($U^L < U^M < U^H$), and W is symmetric in its three arguments. Given this symmetry, the definitions below omit outcome vectors which are permutations of those listed.

FIGURE 4

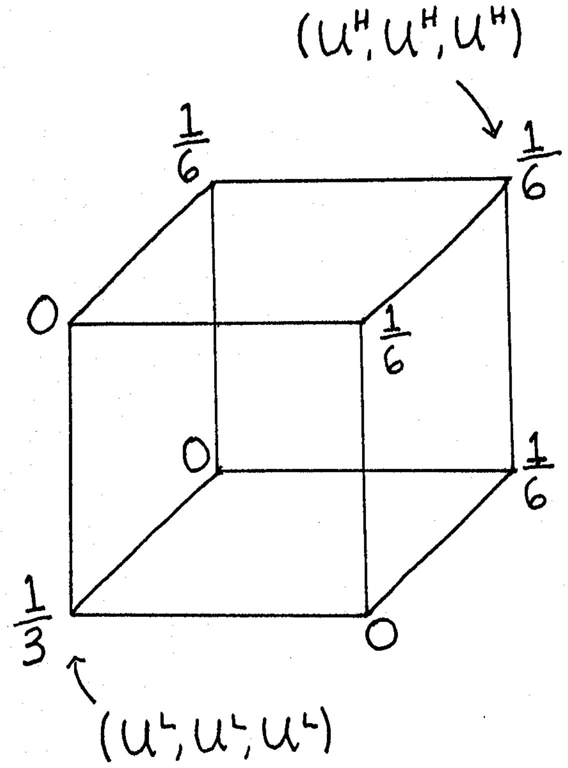


FIGURE 4a: (X_1, X_2, X_3)

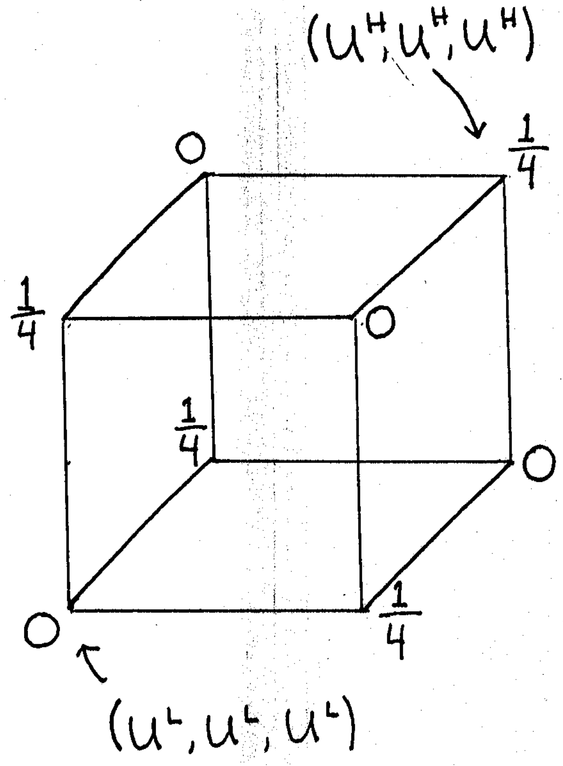


FIGURE 4b: (Y_1, Y_2, Y_3)

Example showing that (SDA) $\not\Longrightarrow$ (SDC)

Example 1: Let W^1 be symmetric and

$$W^1(U^H,U^H,U^H)=3, \quad W^1(U^H,U^H,U^M)=2,$$
$$W^1(U^H,U^H,U^L)=1, \quad W^1(U^H,U^M,U^M)=1,$$
$$W^1(U_1,U_2,U_3)=0 \qquad \text{for } (U_1,U_2,U_3) \text{ which are not permutations of the vectors above}$$

Let $f(U^L)=0$, $f(U^M)=1$, and $f(U^H)=2$. Let $\tilde{W}(0)=\tilde{W}(1)=\tilde{W}(2)=\tilde{W}(3)=0$, $\tilde{W}(4)=1$, $\tilde{W}(5)=2$, and $\tilde{W}(6)=3$. Then W^1 can be represented as

$$W^1(U_1,U_2,U_3)=\tilde{W}(f(U_1)+f(U_2)+f(U_3)),$$

with f non-decreasing and $\tilde{W}$ convex. Thus W^1 is in Class B. However,

$$W^1(U_1,U^H,U^H)+W^1(U_1,U^M,U^M)-W^1(U_1,U^H,U^M)-W^1(U_1,U^M,U^H)$$

is not invariant with respect to U_1, so W^1 is not pairwise separable and thus not in Class A. Moreover, it is straightforward though tedious to verify that any representation of W^1 as a linear combination of elements of C^+, C^-, $\bar{C}^-$, and $\bar{C}^+$ *must* have some of the weights negative, so W^1 is not in Class C. Therefore $W \in B$ does not imply either $W \in A$ or $W \in C$.

Example 2: Let W^2 differ from W^1 only in that $W^2(U^M,U^M,U^M)=1$ (instead of 0). W^2 is a linear combination (with both weights equal to 1) of the following two symmetric functions, the first an element of C^+ and the second an element of $\bar{C}^-$:

$$W^3(U_1,U_2,U_3)=\begin{cases}1 & \text{if } U_i \geq U^M \quad \forall\, i=1,2,3\\ 0 & \text{otherwise}\end{cases}$$

and

$$W^4(U^H,U^H,U^H)=2$$
$$W^4(U^H,U^H,U^M)=W^4(U^H,U^H,U^L)=1$$
$$W^4(U_1,U_2,U_3)=0 \qquad \text{for } (U_1,U_2,U_3) \text{ which are not permutations of the vectors above}$$

Thus W^2 is in Class C. However, since

$$W^2(U^L, U^M, U^M) < W^2(U^M, U^M, U^M) = W^2(U^M, U^M, U^H) < W^2(U^M, U^H, U^H),$$

if we try to put W^2 in the form $\tilde{W}(\sum_{i=1}^{3} f^i(U_i))$, we find that $\tilde{W}$ is not globally convex. Thus W^2 is not in Class B. Furthermore, as in Example 1,

$$W^2(U_1, U^H, U^H) + W^2(U_1, U^M, U^M) - W^2(U_1, U^H, U^M) - W^2(U_1, U^M, U^H)$$

is not invariant with respect to U_1, so W^2 is not in Class A. Therefore $W \in C$ does not imply either $W \in A$ or $W \in B$. Q.E.D.

Proof of Proposition 13: (i) For all points $z = (z_1, \ldots, z_n)$ in the support and for all $i = 1, \ldots, n$, define the increasing sets $A_i^z = \{(U_1, \ldots, U_n) \mid U_i \geq z_i\}$. Repeated application of the definition of association yields

$$P(X_i \geq z_i \quad \forall i) = P(X \in \bigcap_{i=1}^{n} A_i^z) \geq \prod_{i=1}^{n} P(X \in A_i^z) = \prod_{i=1}^{n} P(X_i \geq z_i).$$

The equality of the marginals of X and Y and the independence of $(Y_1, \ldots, Y_n)$ give

$$\prod_{i=1}^{n} P(X_i \geq z_i) = \prod_{i=1}^{n} P(Y_i \geq z_i) = P(Y_i \geq z_i \quad \forall i).$$

Therefore $P(X_i \geq z_i \forall i) \geq P(Y_i \geq z_i \forall i)$. Now define for all z and for all i the increasing sets $B_i^z = \{(U_1, \ldots, U_n) \mid U_i > z_i\}$. Then $(B_i^z)^C = \{(U_1, \ldots, U_n) \mid U_i \leq z_i\}$. Using association of X, equality of the marginals of X and Y, and independence of Y in the same manner as above yields $P(X_i \leq z_i \forall i) \geq P(Y_i \leq z_i \forall i)$. Thus (SDC) is satisfied if X is associated, and a fortiori if X is affiliated.

(ii) For $n = 2$, (SDB) and (SDC) are equivalent.

(iii) Figure 5a shows a two-dimensional distribution for X which satisfies the affiliation inequality (9) (and is therefore associated as well). Figure 5b shows the independent distribution for Y which has identical marginals to those of X. Since

$$P(X_1 = U^M, X_2 = U^H) = \frac{1}{6} > \frac{1}{9} = P(Y_1 = U^M, Y_2 = U^H),$$

X and Y do not satisfy (SDA).

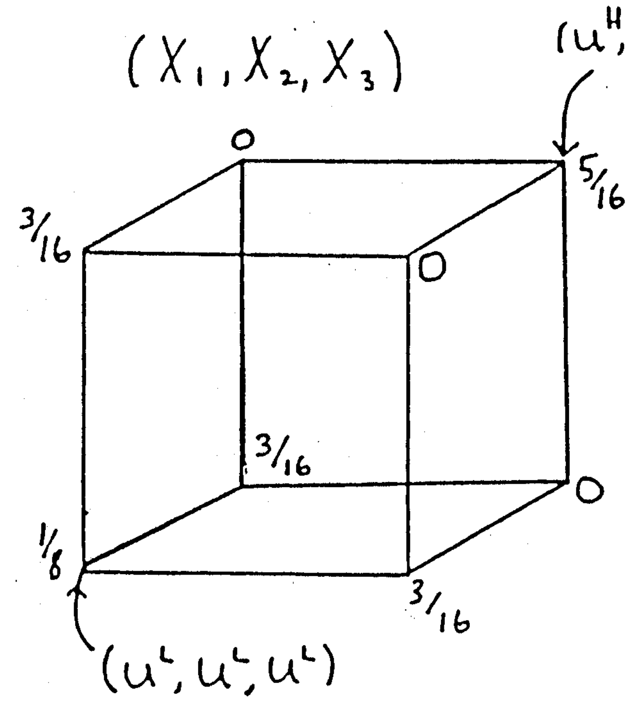
FIGURE 5: Example in which (X_1, X_2) are affiliated and (Y_1, Y_2) are independent, but X and Y do not satisfy (SDA)

FIGURE 5a

X_2 \ X_1	u^L	u^M	u^H
u^H	0	$\frac{1}{6}$	$\frac{1}{6}$
u^M	0	$\frac{1}{6}$	$\frac{1}{6}$
u^L	$\frac{1}{3}$	0	0

FIGURE 5b

Y_2 \ Y_1	u^L	u^M	u^H
u^H	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
u^M	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
u^L	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$



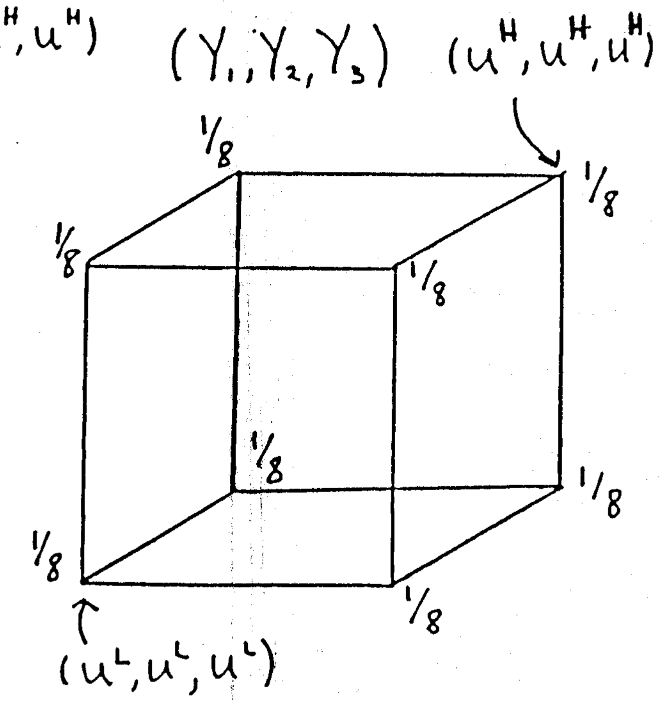


FIGURE 6a

FIGURE 6b

FIGURE 6: Example in which X and Y satisfy (SDA), (SDB), and (SDC), and (Y_1, Y_2, Y_3) are independent, but (X_1, X_2, X_3) are not associated

(iv) Figures 6a and 6b show symmetric, three-dimensional distributions with identical marginals, for X and Y, respectively, in which each component has a two-point support $\{U^L, U^H\}$. (Y_1, Y_2, Y_3) are independent. It is straightforward to check that X and Y satisfy (SDA) and (SDC). X and Y can be shown to satisfy (SDB) by converting the distribution of Y into that of X by a sequence of three GETI's: since any GETI produces a mean-preserving spread of the distribution of $\sum_{i=1}^{n} f^i(U_i)$, for any set of non-decreasing functions $\{f^1, \ldots, f^n\}$, it follows from Rothschild and Stiglitz (1970) that X and Y satisfy (SDB).[1]

However, (X_1, X_2, X_3) are not associated. To see this, define for $i = 1, 2, 3$ the increasing sets

$$A_i = \{(U_1, U_2, U_3) \mid U_j = U^L, U_k = U^L\}^C, \quad \text{where} \quad j \neq i, k \neq i, j \neq k$$

$$\text{Then} \qquad A_1 \cap A_2 \cap A_3 = (A_1^C \cup A_2^C \cup A_3^C)^C$$

$$= \{(U_1, U_2, U_3) \mid U_r = U^L, U_s = U^L, \quad \text{for some} \quad r, s \in \{1, 2, 3\}, r \neq s\}^C$$

$$= \{(U_1, U_2, U_3) \mid U_r = U^H, U_s = U^H, \quad \text{for some} \quad r, s \in \{1, 2, 3\}, r \neq s\}$$

Now $P(X \in A_i) = \frac{11}{16}$ for $i = 1, 2, 3$ and $P(X \in (A_1 \cap A_2 \cap A_3)) = \frac{5}{16}$. Since

$$P(X \in A_1)P(X \in A_2)P(X \in A_3) = \left(\frac{11}{16}\right)^3 > \frac{5}{16} = P(X \in (A_1 \cap A_2 \cap A_3)),$$

(X_1, X_2, X_3) are not associated, and therefore not affiliated. Q.E.D.

---

[1] Each GETI in the appropriate sequence affects the utilities of a different pair (i, j) of individuals, leaving the utility of the remaining individual at U^H. Using the notation $(U_i, U_j; \bar{U}_{-i-j})$, for each (i, j) the GETI increases the probabilities of $(U^H, U^H; U^H)$ and $(U^L, U^L; U^H)$ by $\frac{1}{16}$ and decreases the probabilities of $(U^H, U^L; U^H)$ and $(U^L, U^H; U^H)$ by $\frac{1}{16}$.