© The Journal of Risk and Insurance, 1995 Vol. 62, No. 1, 30-49

Second-Degree Stochastic Dominance Decisions and Random Initial Wealth with Applications to the Economics of Insurance

Ronny Aboudi
Dominique Thon

ABSTRACT

It is known that if random prospect Y stochastically dominates prospect X by the second degree, then all risk-averse agents prefer to hold Y and some random initial wealth Z rather than X and the same initial wealth Z, if Z is statistically independent of both X and Y. We give more general conditions for the invariance of the stochastic dominance ordering to initial wealth to hold: that X be more positively dependent on Z than Y is, and that either X and Z or Y and Z be positively dependent. Applications to the choice of an insurance deductible and of insurance coverage are given.

Introduction

A weakness of many economic analyses relying on a second-degree stochastic dominance ordering is that the variable which appears as an argument of the utility function is typically not total wealth, as should be, but rather some concept of partial wealth. The proposition, for example, that a risk-averse agent, at constant mean, prefers to have a small rather than a large insurance deductible on the basis of the distribution of his or her insurable wealth is always vulnerable to the attack that it is the distribution of total wealth and not only insurable wealth which matters to the agent. It is not difficult to concoct an example where the result no longer holds if it is the distribution of total wealth which is assessed. Many other economic applications could be given. Incorporating random initial wealth in decision making under uncertainty is a major concern of a number of recent papers, and this article contributes to this growing literature.

Ronny Aboudi is Associate Professor of Management Science at the University of Miami. Dominique Thon is Professor of Economics at the Bodö Graduate School of Business, Norway.
The authors thank A. Holtmann, H. Schlesinger, L. Thorlund-Petersen, and two anonymous referees for useful comments. Thon thanks the Department of Economics of the University of Miami for research support and the Norwegian Research Council for Science and the Humanities for financial support.

In this context, it is of interest to identify as undemanding as possible sufficient conditions for decisions taken under the second-degree dominance rule applied to partial wealth to be invariant to the remaining wealth of the agent, called here *initial wealth.* The joint distribution of the partial wealth under consideration and the initial wealth is, of course, the key. It is known that if random prospect Y stochastically dominates random prospect X by the second degree, then all risk-averse agents will prefer to hold Y and some initial wealth Z rather than X and the same Z as long as the latter is statistically independent of both X and Y (Hadar and Russell, 1971, and Levy and Sarnat, 1971; see also Levy and Kroll, 1978, for extensions). This article provides more general conditions which include independence as a particular case for the above invariance of the stochastic dominance ordering to initial wealth to hold. Some restrictions are needed, which are particularly natural in an insurance context: essentially that there be no moral hazard and that the insurance contracts preclude overinsurance.

The next section provides the main result: the sufficient conditions are that X be more positively dependent on Z than Y is, in a well defined sense, and that *either* (X,Z) or (Y,Z) be positively dependent in a well defined sense. Then, we discuss the concept of dependence used here and show that, of the many possible such concepts, it is the most suitable one. Then, we apply the main results to the choice of an insurance deductible and to the choice of insurance coverage by a risk-averse agent in a situation where fair insurance is available. The conclusion relates our results to the recent economic literature on this subject and to the germane subject of characterizing the risk premia when the insurance is partial and/or initial wealth is random.

Main Results

Let in general X and Y be discrete random variables taking values $x = (x_1,\dots, x_n)$ and $y = (y_1,\dots, y_n)$ with a common set of probabilities $\pi = (\pi_1,\dots, \pi_n) \in \mathbb{R}^n_+$; $\sum\pi_i = 1$, pertaining to a set of relevant states of nature $1,\dots, n$. Assume furthermore that x and y are similarly ordered, that is, $(x_i - x_{i-1})(y_i - y_{i-1}) \geq 0$; $i = 2,\dots, n$. We assume without loss of generality x and y to be increasingly ordered; formally,
$x, y \in \boldsymbol{D}^n$ with $\boldsymbol{D}^n = \{w \in \mathbb{R}^n \mid w_1 \leq \dots \leq w_n\}$. Let $S_{n\pi}$ be the set of discrete random variables taking values $w_1 \leq \dots \leq w_n$ with prescribed probabilities π. We thus assume that $X,Y \in S_{n\pi}$. The economic interpretation of those restrictions is discussed below.

Let Z be a discrete random variable taking values $z = (z_1,\dots, z_m)$ with probabilities $p = (p_1,\dots, p_m)$. We assume without loss of generality that $z \in \boldsymbol{D}^m$. Let $A = \{\alpha_{ij}\}$ represent the joint probabilities of (X,Z), that is,
$\alpha_{ij} = \Pr(X = x_i, Z = z_j)$ and $B = \{\beta_{ij}\}$ the joint probabilities of (Y,Z) defined analogously. Let $\hat{\alpha}_{ij} = \sum_{k=1}^{i} \sum_{r=1}^{j} \alpha_{kr}$ be defined as $\Pr(X \leq x_i, Z \leq z_j)$ and $\hat{\beta}_{ij}$ be defined likewise.

$$\begin{array}{cc} \downarrow X \; Z\rightarrow & \begin{array}{ccc} z_1 & z_2 & z_3 \end{array} & \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} & \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} \end{bmatrix} & \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{array} \\ & \begin{array}{ccc} p_1 & p_2 & p_3 \end{array} & \end{array} \qquad \begin{array}{cc} \downarrow Y \; Z\rightarrow & \begin{array}{ccc} z_1 & z_2 & z_3 \end{array} & \\ \begin{array}{c} y_1 \\ y_2 \\ y_3 \\ y_4 \end{array} & \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \\ \beta_{41} & \beta_{42} & \beta_{43} \end{bmatrix} & \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{array} \\ & \begin{array}{ccc} p_1 & p_2 & p_3 \end{array} & \end{array}$$

Consider now the following concept of positive statistical dependence.

Definition 1

Z is *positively regression dependent* on X, denoted $Z_{prd}X$, if

$$\sum_{j=1}^{k} \alpha_{1j}/\pi_1 \geq \sum_{j=1}^{k} \alpha_{2j}/\pi_2 \geq \ldots \geq \sum_{j=1}^{k} \alpha_{nj}/\pi_n; \quad k = 1,\ldots, m.$$

This definition is a specialization of the more general condition that, with the obvious definitions for x and z, $\Pr(Z \leq z \mid X = x)$ be decreasing in x for all z; that is, the smaller the value of X, the larger the probability that Z is small. This concept of dependence is due to Lehmann (1966) and is also known as "Z is stochastically increasing in X," as the conditional distribution of Z given $X = x_1$ is stochastically dominated by the first degree by the conditional distribution of Z given $X = x_2$ for $x_1 < x_2$, as can easily be checked.

The condition $Z_{prd}X$ has the intuition that Z and X tend to move together. A number of alternative concepts of dependence have the same intuition but do not apply here; see the next section. Although the paraphrase that Z and X tend to move together casts Z and X in a symmetrical role, Definition 1 is not symmetrical: $Z_{prd}X$ does not imply $X_{prd}Z$. The following joint distribution is a counterexample; the $\bar{A}$ matrix is defined as $\bar{A} = \{\alpha_{ij}/\pi_i\}$:

$$\begin{array}{cc} \downarrow X \; Z\rightarrow & \begin{array}{ccc} z_1 & z_2 & z_3 \end{array} & \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \begin{bmatrix} 1.8/24 & 0.6/24 & 0 \\ 4.5/24 & 1.8/24 & 0.9/24 \\ 1.8/24 & 0 & 12.6/24 \end{bmatrix} & \begin{array}{c} 0.1 \\ 0.3 \\ 0.6 \end{array} \\ & \begin{array}{ccc} 8.1/24 & 2.4/24 & 13.5/24 \end{array} & \end{array} \qquad \bar{A} = \begin{bmatrix} 18/24 & 6/24 & 0 \\ 15/24 & 6/24 & 3/24 \\ 3/24 & 0 & 21/24 \end{bmatrix}$$

The asymmetry is also apparent from the fact that, if matrix A satisfies the conditions of Definition 1, then $\bar{A}z \in \boldsymbol{D}^n$ for $z \in \boldsymbol{D}^m$ (Marshall and Olkin, 1979, p. 32), the interpretation of which is that $E(Z|x_i)$ is increasing in i. On the other hand, $E(X|z_j)$ need not be increasing in j.

By definition and with $X,Y \in S_{n\pi}$, Y stochastically dominates X by the second degree if

$$\sum_{i=1}^{n} \pi_i U(x_i) \leq \sum_{i=1}^{n} \pi_i U(y_i) \text{ for all increasing concave functions } U; \quad (1)$$

that is, all risk-averse agents prefer Y to X. This we write as $X \preceq^2 Y$. If, in addition, $\Sigma_{i=1}^{n} \pi_i x_i = \Sigma_{i=1}^{n} \pi_i y_i$, then Y is said to dominate X at constant mean, and X is said to be more risky than Y; this we write as $X \lhd^2 Y$. Obviously, $X \lhd^2 Y$ implies that $X \preceq^2 Y$.

Because of our assumption that $X,Y \in S_{n\pi}$, the above definition of $X \preceq^2 Y$ is equivalent to a convenient partial sums condition (see Cheng, 1977, and Pečarić, Proschan, and Tong, 1992, Th. 12.12, for the corresponding result for convex functions; Thon and Thorlund-Petersen, 1988).

Lemma 1

Let $X, Y \in S_{n\pi}$. Then expression (1) is equivalent to

$$\sum_{i=1}^{k} \pi_i x_i \leq \sum_{i=1}^{k} \pi_i y_i; \qquad k = 1,\ldots,n.$$

Let $X \preceq^2 Y$ hold, and $(X,Z) \sim A$; $(Y,Z) \sim B$, with A, B the joint probability matrices defined above and ~ indicating "jointly distributed according to." We now show that, if $Z_{prd}X$ and (X,Z) and (Y,Z) have the same joint probability matrices, then $X \preceq^2 Y$ implies that $X + Z \preceq^2 Y + Z$. The proof is given in the Appendix.

Theorem 1

Let $X,Y \in S_{n\pi}$ and Z be random variables as defined above. Let $X \preceq^2 Y$. Then, if $Z_{prd}X$ and (X,Z) and (Y,Z) have the same joint probabilities, that is, if

$$\sum_{j=1}^{k} \alpha_{1j}/\pi_1 \geq \sum_{j=1}^{k} \alpha_{2j}/\pi_2 \geq \ldots \geq \sum_{j=1}^{k} \alpha_{nj}/\pi_n; \quad k = 1,\ldots,n.$$

and

$$\beta_{ij} = \alpha_{ij}; \text{ all } i,j,$$

then $\quad X + Z \preceq^2 Y + Z.$

Obviously, in the theorem, $Z_{prd}X$ and $Z_{prd}Y$ are equivalent. The nature of the result dictates that X and Y have the same marginal probabilities and that x and y be similarly ordered, that is, that $X,Y \in S_{n\pi}$. In an insurance context, making the assumption that X and Y have the same marginal probabilities is to assume away moral hazard: if x_i is insurable wealth before insurance under state i and y_i is insurable wealth after partial insurance, the probabilities have not changed.

The further assumption that x and y be similarly ordered corresponds to the fundamental incentive-preserving property of a partial insurance contract, which is that, if, under state i, pre-insurance income is larger than under state j, then the same must also hold after insurance. In other words, it is never profitable for an insured agent to incur a larger loss than the one he or she is subjected to by nature. This is essentially a "no overinsurance" assumption. Insurance contracts typically preclude overinsurance because it causes obvious moral hazard problems; contracts with deductible or less than full coverage, such as considered below satisfy this assumption by construction.

To give some intuitive content to the theorem, assume that originally the agent had income Z + X, of which X is insurable. Let Y be reachable from X by taking a fair partial insurance on X alone, that is, $X \lhd^2 Y$, for example a fair insurance with deductible such as discussed below. Then the condition that $Z_{prd}X$ can be interpreted as saying that X is not a good hedge against Z, and, therefore, the theorem says that, barring moral hazard and assuming the insurance contract not to exhibit overinsurance, to insure the insurable wealth is indeed to insure total wealth. In case of "negative dependence" between X and Z, this does not follow; in case of "strong negative dependence" to insure X alone may even deinsure total wealth because the risk present in X reduces the risk in X + Z; that is, X is a good hedge against Z to start with. To take an extreme example, let x = z = (1, 2, 3), y = (1.5, 2, 2.5), π = p = (1/3, 1/3, 1/3). One has $X \lhd^2 Y$. But if the pairs (X,Z), (Y,Z) both have joint probabilities

$$\begin{bmatrix} 0 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1/3 & 0 & 0 \end{bmatrix},$$

it is easy to check that $Y + Z \lhd^2 X + Z$; that is, to insure X alone makes total wealth more risky. Consider now the following conditions on the joint cumulative probabilities $\hat{\alpha}_{ij}$ and $\hat{\beta}_{ij}$.

Definition 2

(X,Z) are more positively dependent than (Y,Z), denoted $(Y,Z) \preceq^d (X,Z)$ if $\beta_{ij} \leq \hat{\alpha}_{ij}$, i = 1,..., n; j = 1,..., m.

This partial ordering on joint distributions is related to but distinct from, the "more concordant than" partial ordering. In general, let V_1, V_2, W_1, W_2 be random variables, and let (V_1,W_1) and (V_2,W_2) have joint distributions H_1 and H_2, respectively. Then, following Tchen (1980), we say that (V_1, W_1) are *more concordant than* (V_2,W_2) if V_1 and V_2 have the same marginal distributions; likewise for W_1 and W_2, and if $H_1 \leq H_2$. In other words, (V_1,W_1) dominates (V_2,W_2) by joint first-degree dominance. We write this as $(V_2,W_2) \preceq^c (V_1,W_1)$. This ordering was introduced in economics by Epstein and Tanny (1980) as "more correlated than." For economic applications, see, for example, Epstein and Tanny (1980), Pratt and Zeckhauser (1987), and Aboudi and Thon (1993).

Definition 3

Specialized to our case, with $X,Y \in S_{n\pi}$ and $W_1 = W_2 = Z$, Definition 2 becomes $(Y,Z) \preceq^c (X,Z)$ if $\hat{\beta}_{ij} \leq \hat{\alpha}_{ij}$; i = 1,..., n; j = 1,..., m, and $x_i = y_i$, i = 1,..., n.

The following Lemma is a specialization of a result of Epstein and Tanny (1980, Corollary 5, p. 21). Note that, by Definition 3, X and Y have the same marginals and thus X+Z and Y+Z have the same mean.

Lemma 2

If $(Y,Z) \preceq^c (X,Z)$, then $X+Z \trianglelefteq^2 Y+Z$.

Theorem 1, which assumes that $A = B$ but allows the marginals to differ, and Lemma 2, which imposes common marginals but allows for $A \neq B$, can be combined into our first main result: If Z is positively regression dependent on X and (X,Z) are more positively dependent than (Y,Z), then $X \preceq^2 Y$ implies $X+Z \preceq^2 Y+Z$. Thus, if all risk-averse agents prefer Y to X in isolation, so do they with common initial wealth Z. Equality of the means of X and Y is not assumed. The case of statistical independence discussed in the literature falls off as a particular case.

Theorem 2

Let $X,Y \in S_{n\pi}$ and Z be random variables as defined above. Let $X \preceq^2 Y$. Then, if $Z_{prd}X$ and if $(Y,Z) \preceq^d (X,Z)$, that is, if

$$\sum_{j=1}^{k} \alpha_{1j} /\pi_1 \geq \sum_{j=1}^{k} \alpha_{2j}/\pi_2 \geq \dots \geq \sum_{j=1}^{k} \alpha_{nj}/\pi_n; \quad k = 1,\dots,m$$

and

$$\beta_{ij} \leq \hat{\alpha}_{ij}; \text{ all } i,j,$$

then $\quad X + Z \preceq^2 Y + Z.$

Proof. Introduce the random variable Y', where Y' and Y are identically distributed but $\Pr(Y' \leq y_i , Z \leq z_j) = \alpha_{ij}$. It is clear that $X \preceq^2 Y'$. By assumption, $Z_{prd}X$ and by construction of (Y',Z), (X,Z) and (Y',Z) have the same joint probabilities. Thus, by Theorem 1, $X+Z \preceq^2 Y'+Z$. By Lemma 2, since Y' and Y are identically distributed, $Z = Z$, and $\hat{\beta}_{ij} \leq \hat{\alpha}_{ij}$, one gets $Y'+Z \trianglelefteq^2 Y+Z$. By transitivity of $\preceq^2$, one has $X+Z \preceq^2 Y+Z$.■

Our second main result is that, if it is now on Y that Z is positively regression dependent and as in Theorem 2, $(Y,Z) \preceq^d (X,Z)$, then $X \preceq^2 Y$ again implies that $X+Z \preceq^2 Y+Z$.

Theorem 3

Let $X,Y \in S_{n\pi}$ and Z be random variables as defined above. Let $X \preceq^2 Y$. Then, if $Z_{prd}Y$ and if $(Y,Z) \preceq^d (X,Z)$, that is, if

$$\sum_{j=1}^{k} \beta_{1j}/\pi_1 \geq \sum_{j=1}^{k} \beta_{2j}/\pi_2 \geq \dots \geq \sum_{j=1}^{k} \beta_{nj}/\pi_n; \quad k = 1,\dots,m$$

and

$$\beta_{ij} \leq \hat{\alpha}_{ij}; \text{ all } i,j,$$

then $\quad X + Z \preceq^2 Y + Z.$

Proof. Introduce the random variable X', where X' and X are identically distributed but $\Pr(X' \leq x_i, Z \leq z_j) = \beta_{ij}$. It is clear that $X' \preceq^2 Y$. By construction, $(X',Z) \preceq^c (X, Z)$; so, by Lemma 1, $X+Z \trianglelefteq^2 X'+Z$. Since $Z_{prd}Y$, by definition of X', we have $Z_{prd}X'$. Also, (X',Z) and (Y,Z) have the same joint probabilities; thus, by Theorem 1, $X'+Z \preceq^2 Y+Z$. By transitivity of $\preceq^2$, one has $X+Z \preceq^2 Y+Z$.■

Note that $Z_{prd}Y$ and $(Y,Z) \preceq^d (X,Z)$ together do not imply $Z_{prd}X$. For example,

$$A = \begin{bmatrix} 6 & 1 & 1 \\ 2 & 6 & 0 \\ 0 & 1 & 7 \end{bmatrix} \text{x } 1/24 \qquad B = \begin{bmatrix} 8/3 & 8/3 & 8/3 \\ 8/3 & 8/3 & 8/3 \\ 8/3 & 8/3 & 8/3 \end{bmatrix} \text{x } 1/24$$

One has $Z_{prd}Y$ and $(Y,Z) \preceq^d (X,Z)$ but not $Z_{prd}X$. Thus, Theorem 2 does not follow from Theorem 3. That Theorem 2 does not imply Theorem 3 is obvious.

Concepts of Bivariate Dependence

In this section, we discuss the concept of positive regression dependence (PRD) and its relationship to other bivariate dependence concepts. We show that those other concepts of dependence are either unnecessarily restrictive or insufficient for Theorems 2 and 3 to hold.

To fix ideas, let us discuss the dependence of X and Z and Theorem 2. The implications of this discussion for Theorem 3 are immediate. Let X, Z have densities f^x, f^z, distribution functions F^x, F^z, respectively, joint density g^{xz}, and joint distribution function G^{xz}. The natural dependence concept corresponding to $\preceq^d$ is the symmetrical concept of positive quadrant dependence (PQD), defined as follows (Epstein and Tanny, 1980).

Definition 4

Random variables X and Z with range [a,b] are positively quadrant dependent, denoted (X,Z)PQD, if $G^{xz} \geq F^x F^z$, for all values in [a,b].

This is, of course, saying that X and Z are more concordant (Definition 3) than if they were independent. It is easy to show that, if (Y,Z)PQD and $(Y,Z) \preceq^d (X,Z)$, then (X,Z)PQD. Thus, with positive dependence defined as PQD, the intuition holds that, if Y and Z are positively dependent and (X,Z) are more positively dependent than (Y,Z), then X and Z are positively dependent. Other concepts of bivariate dependence have appeared in the mathematical literature (see Lehmann, 1966, and Tong, 1980, for a review) and the economic literature on portfolio theory (Brumelle, 1974, and Hildreth and Tesfatsion, 1977).

Definition 5

Z is left tail decreasing in X, denoted $Z_{ltd}X$ if $\Pr(Z \leq z|X \leq x)$ is decreasing in x for all z.

Definition 6

Z is right tail increasing in X, denoted $Z_{rti}X$ if $\Pr(Z > z|X > x)$ is increasing in x for all z.

Neither of those two concepts is symmetrical, and neither implies the other one (Tong, 1980). Another nonsymmetrical concept of positive dependence is the following one, corresponding to Wright's (1987) concept of negative expectation dependence.

Definition 7

Z is positively expectation dependent on X, (Z|X)PED if $E(Z|X \leq x) \leq E(Z)$ for all x.

The following concept of dependence is symmetrical; so is statistical independence, denoted (X,Z)I, and positive covariance ($cov_{xz} \geq 0$).

Definition 8

X and Z are positively likelihood ratio dependent, denoted (X,Z)PLRD if $G^{xz}(z',x)G^{xz}(z,x') \leq G^{xz}(z,x)G^{xz}(z',x')$ for all $z' > z$, $x' > x$.

The implications shown in Figure 1 are known to exist and to be strict, that is, their converses do not hold (Tong, 1980, and Wright, 1987). For completeness, we have included the alternative version of the nonsymmetrical concepts.

Figure 1

$$\begin{tikzcd}
 & & & Z_{ltd}X \arrow[ddr] & & & \\
 & & Z_{prd}X \arrow[ur] \arrow[r] & Z_{rti}X \arrow[dr] & & & \\
(X,Z)I \arrow[r] & (X,Z)PLRD \arrow[ur] \arrow[dr] & & & (X,Z)PQD \arrow[r] & (Z|X)PED \arrow[r] & Cov_{xz}\geq 0 \\
 & & X_{prd}Z \arrow[r] \arrow[dr] & X_{ltd}Z \arrow[ur] & & & \\
 & & & X_{rti}Z \arrow[uur] & & &
\end{tikzcd}$$

We now show that, in Theorem 2, $Z_{prd}X$ cannot be replaced by a dependence concept which is not more restrictive. From Figure 1, this requires a counterexample to each one of the following statements:

$$X \preceq^2 Y \text{ and } Z_{ltd}X \rightarrow X + Z \preceq^2 Y + Z \quad (2)$$
$$X \preceq^2 Y \text{ and } Z_{rti}X \rightarrow X + Z \preceq^2 Y + Z \quad (3)$$
$$X \preceq^2 Y \text{ and } X_{prd}Z \rightarrow X + Z \preceq^2 Y + Z \quad (4)$$

To make the counterexamples easy to check, we assume that $n = m = 3$; $\pi_i = p_i = 1/3$, $i = 1, 2, 3$. Let v_i, $i=1,...$ be the increasingly ordered union of the values taken by random variables X+Z and Y+Z. Producing suitable counterexamples to the above implications requires showing that for some v_i, $G_2(v_i) > F_2(v_i)$, with G_2, F_2 defined as in expression (A3), and F the distribution function of X+Z, G the distribution function of Y+Z. In all three cases, we assume the joint probability matrices A and B to be identical.

Counterexample to expression (2).

$$\text{Let } A = B = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 2 \end{bmatrix} 1/18$$

This matrix satisfies $Z_{ltd}X$ but not $Z_{prd}X$. Let X = (0, 1, 20), Y = (0, 10, 11), Z = (0, 100, 200). Simple calculations show that $X \preceq^2 Y$ but not $X+Z \preceq^2 Y+Z$, as $G_2(20) > F_2(20)$.

Counterexample to expression (3).

$$\text{Let } A = B = \begin{bmatrix} 3 & 2.5 & 2.5 \\ 5 & 1.5 & 1.5 \\ 0 & 4 & 4 \end{bmatrix} 1/24$$

This matrix satisfies $Z_{rti}X$ but not $Z_{prd}X$. With X = (1, 20, 21), Y = (10, 11, 50), and Z as in the previous counterexample, we find $G_2(20) > F_2(20)$. Incidently, this and the previous counterexample also show that $Z_{ltd}X$ and $Z_{rti}X$ do not imply each other.

Counterexample to expression (4).

$$\text{Let } A = B = \begin{bmatrix} 3.5 & 3 & 1.5 \\ 4.5 & 3.5 & 0 \\ 0 & 1.5 & 6.5 \end{bmatrix} 1/24$$

This matrix satisfies $X_{prd}Z$ but not $Z_{prd}X$. With X,Y,Z as in the previous counterexample, we find again $G_2(20) > F_2(20)$.

An implication of the above counterexamples is that positive quadrant dependence, the most natural concept of dependence to be used in conjunction with $\preceq^d$, will not do in place of positive regression dependence in Theorem 2. If, now, in Figure 1, every (X,Z) is replaced by (Y,Z), the counterexamples show, *mutatis mutandis,* that $Z_{prd}Y$ in Theorem 3 cannot either be replaced by a weaker dependence concept.

For future reference, we say, as is natural, that Z is negatively regression dependent on X, $Z_{nrd}X$, if $\Pr(Z \leq z \mid X = x)$ is increasing in x (Lehmann, 1966), which can easily be checked to be equivalent to $Z_{prd}(-X)$. With $X,Y \in S_{n\pi}$, this gives the natural counterpart of Definition 1.

Definition 9

Z is negatively regression dependent on X, denoted by $Z_{nrd}X$, if

$$\sum_{j=1}^{k} \alpha_{1j}/\pi_1 \leq \sum_{j=1}^{k} \alpha_{2j}/\pi_2 \leq \ldots \leq \sum_{j=1}^{k} \alpha_{nj}/\pi_n; \qquad k = 1,\ldots,m.$$

The above example might suggest the conjecture that, in general, "if $Z_{nrd}X$ and $\beta_{ij} = \alpha_{ij}$, all i,j, then $X \preceq^2 Y$ implies that $Y + Z \preceq^2 X + Z$." That this is not the case is clear from the fact that the latter cannot hold if, for example, the lowest value of Y is larger than the lowest value of X.

Applications to Insurance Deductible and Insurance Coverage

Suppose a risk-averse agent has initial random wealth Z taking values $z_1 \leq z_2 \leq \ldots \leq z_m$, with probabilities p_i, i = 1,..., m. The agent is subject to a random loss, L, taking values $\ell = (\ell_i)_1^n$, $\ell_1 > \ell_2 > \ldots > \ell_n > 0$, with probabilities π_i, i = 1,..., n; n > 1. The agent can take a fair insurance with deductible δ, a nonnegative scalar, $0 \leq \delta \leq \ell_1$. The fair premium, which of course depends on δ, is calculated according to

$$\Pi(\delta) = \Sigma_{i=1}^{n} \pi_i \max(0, \ell_i - \delta).$$

This is the standard setup; see, for example, Mossin (1968) and Doherty and Schlesinger (1983a). Define the vector $b(\ell, \delta)$ corresponding to the discrete random variable $B(L,\delta)$ taking values $b_i(\ell_i,\delta)$ with probabilities π_i:

$$b(\ell,\delta) = \{b_i(\ell_i,\delta)\}_1^n = \{-\min(\ell_i, \delta) - \Pi(\delta)\}_1^n.$$

For any i, $-\min(\ell_i, \delta)$ is the negative income to which the agent is directly subjected, with probability π_i, while $-\Pi(\delta)$ is the additional negative income corresponding to the premium payment, independent of i. Then, $B(L,\delta)$ is the agent's partial income in addition to initial income Z when he or she chooses deductible δ, and $Z + B(L,\delta)$ is the agent's total income. Note that $b(\ell,\delta)$ contains only negative values and, as shown below, $b_i(\ell_i,\delta)$ is increasing in i. When deciding whether to take deductible δ or δ', the agent evaluates random income $Z+B(L,\delta)$ versus $Z+B(L,\delta')$.

Mossin (1968) shows that, with deterministic Z, a risk-averse agent always prefers a zero deductible. Doherty and Schlesinger (1983a) show that this remains true with a stochastic Z if Z and L are statistically independent. They also show that, if (Z,L) are distributed according to a bivariate normal distribution, then the results hold under the condition that $\mathrm{Cov}(Z, L) \leq 0$. We generalize those results using a stochastic dominance argument and our main results. We first prove that partial income $B(L,\delta)$ can be ordered by $\triangleleft^2$ according to the size of the deductible δ.

Lemma 3

If $\delta < \delta'$, then $B(L,\delta') \triangleleft^2 B(L,\delta)$.

Proof. To simplify, we write b_i for $b_i(\ell_i,\delta) = -\min(\ell_i,\delta) - \Pi(\delta)$ and b_i' for $b_i(\ell_i,\delta') = -\min(\ell_i, \delta') - \Pi(\delta')$; i = 1,..., n. Step 1 establishes that b_i, b_i', and

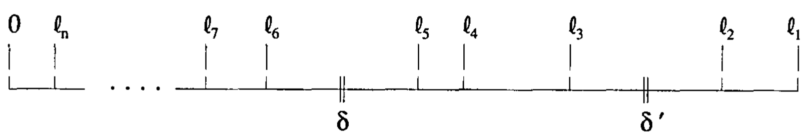
$b_i' - b_i$ are increasing (in the sense of nondecreasing) in $i = 1,\dots, n$; Step 2 completes the proof.

Step 1. We have

for any i such that $\ell_i \geq \delta'$: $b_i = -\delta - \Pi(\delta)$; $b_i' = -\delta' - \Pi(\delta')$; b_i, b_i', and $b_i' - b_i$ are constant in i.

for any i such that $\delta < \ell_i < \delta'$: $b_i = -\delta - \Pi(\delta)$; $b_i' = -\ell_i - \Pi(\delta')$; b_i is constant and b_i' increasing in i; $b_i' - b_i$ is increasing in i.

for any i such that $\ell_i \leq \delta$: $b_i = -\ell_i - \Pi(\delta)$; $b_i' = -\ell_i - \Pi(\delta')$; b_i and b_i' are increasing in i; $b_i' - b_i$ is constant in i.



For the largest i such that $\delta < \ell_i < \delta'$(if any): $b_{i+1} \geq b_i$, $b_{i+1}' \geq b_i'$, and $b_{i+1}' - b_{i+1} \geq b_i' - b_i$. For the largest i such that $\ell_i \geq \delta'$, whether or not there is some i such that $\delta < \ell_i < \delta'$, we have that $b_{i+1} \geq b_i$, $b_{i+1}' \geq b_i'$, and $b_{i+1}' - b_{i+1} \geq b_i' - b_i$. Thus, b_i, b_i', and $b_i' - b_i$ are increasing in $i = 1,\dots, n$.

Step 2. We have $\Sigma_{i=1}^n \pi_i \max(\delta,\ell_i) \leq \Sigma_{i=1}^n \pi_i \max(\delta', \ell_i)$, which is equivalent to $\Pi(\delta) + \delta \leq \Pi(\delta') + \delta'$, equivalent again to $b_1' \leq b_1$, as $b_1 = -\delta - \Pi(\delta)$, $b_1' = -\delta' - \Pi(\delta')$.

We also have $\Pi(\delta') \leq \Pi(\delta)$.

If $\ell_n \leq \delta$, then $b_n = -\ell_n - \Pi(\delta)$, $b_n' = -\ell_n - \Pi(\delta')$; thus, $b_n \leq b_n'$.

If $\delta < \ell_n < \delta'$, then $\Pi(\delta) = \Sigma_{i=1}^n \pi_i(\ell_i - \delta) = \Pi(\ell_n) + \ell_n - \delta$ and $\Pi(\delta') \leq \Pi(\ell_n) = \Pi(\delta) + \delta - \ell_n$, or $-\delta - \Pi(\delta) \leq -\ell_n - \Pi(\delta')$, which is $b_n \leq b_n'$.

If $\ell_n \geq \delta'$, then

$\Pi(\delta) = \Sigma_{i=1}^n \pi_i (\ell_i - \delta) = \Sigma_{i=1}^n \pi_i(\ell_i - \delta') + \delta' - \delta = \Pi(\delta') + \delta' - \delta$.

Thus, $\Pi(\delta) + \delta = \Pi(\delta') + \delta'$. Then $b_n = -\delta - \Pi(\delta)$, $b_n' = -\delta' - \Pi(\delta')$ and, thus, $b_n \leq b_n'$.

Thus, b_i, b_i', and $b_i' - b_i$ are increasing in $i = 1,\dots, n$, $b_n \leq b_n'$, and $b_1' \leq b_1$. Then there exists some $\kappa < n$ such that $b_i' \leq b_i$; $i = 1,\dots, \kappa$ and $b_i \leq b_i'$; $k = \kappa+1,\dots, n$. It follows that $\pi_i b_i' \leq \pi_i b_i$, $i = 1,\dots, \kappa$ and $\pi_i b_i \leq \pi_i b_i'$, $i = \kappa+1,\dots, n$. The former gives

$$\Sigma_{i=1}^k \pi_i b_i' \leq \Sigma_{i=1}^k \pi_i b_i \,;\; k = 1,\dots, \kappa, \tag{5}$$

while the latter gives $\Sigma_{i=k}^n \pi_i b_i \leq \Sigma_{i=k}^n \pi_i b_i'$; $k = \kappa+1,\dots, n$, which, together with $\Sigma_{i=1}^n \pi_i b_i = \Sigma_{i=1}^n \pi_i b_i'$, which holds by construction, implies that

$$\Sigma_{i=1}^k \pi_i b_i' \leq \Sigma_{i=1}^k \pi_i b_i \,;\; k = \kappa+1,\dots, n. \tag{6}$$

Together, expressions (5) and (6) give $\Sigma_{i=1}^k \pi_i b_i' \leq \Sigma_{i=1}^k \pi_i b_i$; $k = 1,\dots, n$. As b_i and b_i' are increasing in $i = 1,\dots, n$, by Lemma 1, $B(L,\delta') \triangleleft^2 B(L,\delta)$.■

Note that, if $Z_{prd}B(L,\delta)$ for some δ, then it is also the case for any other δ, as the ordering of $b(\ell,\delta)$ is the same for all δ. As in the literature cited above, we abstract from any moral hazard problem associated with the deductible; that is, we assume the joint probabilities of Z and B(L,δ) to be the same for all δ, $0 \leq \delta$. Lemma 3 and Theorem 1 (specialized in the present argument to $\triangleleft^2$) with $X = B(L,\delta')$, $Y = B(L,\delta)$, $\delta < \delta'$ then imply that if $Z_{prd}B(L,\delta')$ (or, equivalently, $Z_{prd}B(L,\delta)$), then a risk-averse agent always prefers a lower to a higher deductible, as the former gives a distribution of total wealth which dominates the one given by the latter.

It is convenient to restate the result in terms of conditions on the joint distribution of (L,Z), the loss and initial wealth. As $[b_i(\ell_i,\delta)]_1^n$, $[-\min(\ell_i,\delta)]_1^n$ and thus $[-\ell_i]_1^n$ are similarly ordered, then $Z_{prd}B(L,\delta)$ is equivalent to $Z_{nrd}L$, Definition 9. We have thus a generalization of the earlier results: If initial wealth and insurable losses are negatively regression dependent, then a risk-averse agent always prefers a lower to a higher deductible; in particular, a zero deductible is optimal.

Suppose now that the agent wants to insure a fraction of a possible loss instead of, as before, insure the losses with a deductible. Let λ be the value of the agent's insurable income when no loss occurs; it is subject to a random loss L, taking as before values $\lambda \geq \ell_1 > \ell_2 > \ldots > \ell_n > 0$ with probabilities π_i, $i = 1,\ldots, n$; $n > 1$. Following Mossin (1968), let the *coverage* be expressed in absolute value by η, $0 \leq \eta \leq \lambda$; that is, in case of a loss ℓ_i, the agent receives a compensation $(\eta/\lambda)\ell_i$. Assume the insurance to be fair for all levels of coverage; then the premium is calculated as $\Pi(\eta) = (\eta/\lambda)\Sigma_{i=1}^n \pi_i \ell_i$.

The agent's partial wealth, H(η), that he or she holds in addition to initial wealth Z can thus be represented by

$$h(\eta) = [h_i]_1^n = [\lambda - \ell_i(1 - \eta/\lambda) - \Pi(\eta)]_1^n.$$

We now show that h(η) can be ordered by $\triangleleft^2$ according to η. For any i, with $\eta' < \eta$:

$$\ell_i(1 - \eta/\lambda) - \ell_i(1 - \eta'/\lambda) = \ell_i(\eta' - \eta)/\lambda$$

is increasing in i. Then,

$$h_i(\eta') - h_i(\eta) = [\lambda - \ell_i(1 - \eta'/\lambda) - \Pi(\eta')] - [\lambda - \ell_i (1 - \eta/\lambda) - \Pi(\eta)]$$

is increasing in i. Furthermore, $[h_i(\eta)]_1^n$ and $[h_i(\eta')]_1^n$ are also increasingly ordered. It is easy to show that $h_n(\eta) < h_n(\eta')$; $h_1(\eta') < h_1(\eta)$ and $\Sigma_{i=1}^n\pi_i h_i(\eta) = \Sigma_{i=1}^n\pi_i h_i(\eta') = \lambda - \Sigma_{i=1}^n\pi_i\ell_i$. Then, by an argument similar to the one of Step 2 of the proof of Lemma 3, one obtains $H(\eta') \triangleleft^2 H(\eta)$. If now Z is, as before, initial wealth taking values $z_1 \leq z_2 \leq \ldots \leq z_m$, with probabilities p_i, $i = 1,\ldots, m$, we have by an argument similar to the one of the deductible case that if $Z_{prd}H(\eta)$ (or, equivalently, $Z_{prd}H(\eta')$), then $Z+H(\eta') \triangleleft^2 Z+H(\eta)$; that is, all risk-averse agents prefer a higher to a lower coverage. It is easy to show along the same lines that $Z_{prd}H(\eta)$ is equivalent to $Z_{nrd}L$. Thus, if initial wealth and losses are negatively regression dependent, then a risk-averse agent always prefers a higher to a lower coverage; in particular, total coverage ($\eta = \lambda$) is optimal. We summarize our results in the following proposition.

Proposition 1

If a risk-averse agent whose wealth is negatively regression dependent on insurable losses has access to fair insurance, then he or she will always prefer a low to a high deductible and a high to a low coverage.

Conclusion

Let us for the sake of concreteness limit our concluding remarks to choices at constant mean, illustrated by choices between alternative partial fair insurance schemes. Note that Theorems 2 and 3 are in effect more general, as they allow the mean of the preferred distribution to be larger than the one of its alternative, as follows from the definition of stochastic dominance. Better-than-fair insurance is normally not available. On the other hand, a second-degree stochastic argument can never explain the taking of insurance, or of more insurance, when loading is present; our results, like many expected utility results appearing in the literature would thus not survive the introduction of any degree of loading. Obtaining results concerning risk-averse agents' insurance decisions with positive loading requires characterizations of their utility function which extend beyond mere concavity (see Mossin, 1968; Schlesinger, 1981; Turnbull, 1983; Eeckoudt and Kimball, 1992; and Kimball, 1993). Economic decisions in the nature of partial fair insurance arrangements constitute the natural field of application of our results.

Expected utility arguments expressing the desirability of a complete fair insurance in the case of risk aversion and in the absence of randomness in initial wealth, as well as the characterization of the corresponding risk premium of an agent have been known for some time; such results follow from concavity alone (Pratt, 1964). The more recent literature has been concerned with the same two classes of problems—the decisions of risk-averse agents and the characterization of the risk premium—but has extended the analysis in two directions. The first one is the consideration of partial rather than complete insurance. Stochastic dominance theory is perfectly suited for obtaining results about such choices by risk-averse agents (Rothschild and Stiglitz, 1970) with the obvious limitation that the insurance needs to be fair, as noted above. That the fair insurance be partial rather than complete raises nontrivial issues regarding the definition of the risk premium of an agent and the derivation of results about it (e.g., Ross, 1981). The second direction taken by recent research is to recognize the ubiquity of randomness in initial wealth (also known as background risk or, in an insurance context, uninsurable risk) and derive results either on the decisions of the agent as done here (Doherty and Schlesinger, 1983a; Schlesinger and Doherty, 1985) or on the risk premium (Doherty and Schlesinger, 1986; Doherty, Loubergé, and Schlesinger, 1987).

This article deals with partial fair insurance in the presence of background risk and it offers what seem to be the least demanding among the known necessary conditions for the partial insurance decisions of all risk-averse agents to be invariant to background risk. It has, on the other hand, nothing to offer regarding the characterization of the risk premium. Kihlstrom, Romer, and

Williams (1981), using more restrictive assumptions on the joint distribution (independence) and a smaller class of utility functions (nonincreasing absolute risk aversion), are able to prove, essentially, that the risk premium increases with risk aversion. Another result along this line is provided by Pratt and Zeckhauser (1987), who have characterized the risk premium for a class of risk-averse agents—those exhibiting so-called proper risk aversion—with a background risk germane, but not identical, to the one used in this article. As the pieces fall into place, an interesting direction for future research would be to obtain characterizations of the risk premium, hopefully for all risk-averse agents, with assumptions on the background risk such as the ones utilized here.

Appendix

Lemma A1

Let $x_1 \le ... \le x_n$, $y_1 \le ... \le y_n$. If $\sum_{i=1}^{k} \pi_i x_i \le \sum_{i=1}^{k} \pi_i y_i$, $k = 1,..., n$, and $a_1 \ge a_2 ... \ge a_n \ge 0$, then $\sum_{i=1}^{k} a_i \pi_i x_i \le \sum_{i=1}^{k} a_i \pi_i y_i$, $k = 1,..., n$.

Proof. $\sum_{i=1}^{k} a_i \pi_i x_i = (a_1-a_2)\pi_1 x_1 + (a_2-a_3)(\pi_1 x_1+\pi_2 x_2) + (a_3-a_4)(\pi_1 x_1 + \pi_2 x_2 + \pi_3 x_3) + ... + (a_{k-1}-a_k)(\pi_1 x_1 + \pi_2 x_2 + ... + \pi_{k-1} x_{k-1}) + a_k(\pi_1 x_1 + \pi_2 x_2 + ... + \pi_k x_k)$. The result then follows from

$a_i-a_{i+1} \ge 0$, $a_k \ge 0$ and $\pi_1 x_1 + \pi_2 x_2 + ... + \pi_i x_i \le \pi_1 y_1 + \pi_2 y_2 + ... + \pi_i y_i$.■

Let Q^{yz} be the matrix with element $q_{ij}^{yz} = y_i + z_j$, $i = 1,..., n$; $j = 1,..., m$. Let T_r^{yz} be the subset of pairs of indices (i,j) such that $y_i + z_j$, with $(i,j) \in T_r^{yz}$ belongs to the set of the r smallest elements of Q^{yz}. Define Q^{xz} and T_r^{yz} similarly. Let $N = n \times m$.

Lemma A2

For all $r = 1,..., N$, if $(p,q) \in T_r^{yz}$, then $(i,j) \in T_r^{yz}$ for all (i,j) with $i \le p$ and $j \le q$.

Proof. As $y_i \le y_p$ for all $i \le p$, and $z_j \le z_q$ for all $j \le q$, then $y_i + z_j \le y_p + z_q$. Thus, the matrix Q^{yz} is partitioned between the q_{ij}^{yz}s such that $(i,j) \in T_r^{yz}$ and the other ones by an "increasing line." For example, with $y = (1.75, 1.8, 4)$, $z = (5, 7, 10)$, and $r = 5$:

$$\begin{bmatrix} 6.75 & 8.75 & 11.75 & .. \\ 6.8 & 8.8 & 11.8 & .. \\ 9 & 11 & 14 & .. \\ . & . & . & .. \\ . & . & . & .. \end{bmatrix}. \blacksquare$$

Lemma A3

If $X \preceq^2 Y$, and $Z_{prd}X$, then for a given r,

$$\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(x_i + z_j) \le \sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(y_i + z_j). \tag{A1}$$

Proof. Expression (A1) is equivalent to

$$\sum_{(i,j)\in T_r^{yz}} \alpha_{ij} x_i \le \sum_{(i,j)\in T_r^{yz}} \alpha_{ij} y_i \ . \tag{A2}$$

Let $\zeta = \max_{(i,j)\in T_r^{yz}} \{i\}$, and $\delta_i = \max_{(i,j)\in T_r^{yz}} \{j\}$, $i=1,\ldots,\zeta$. Then,

$$\sum_{(i,j)\in T_r^{yz}} [\alpha_{ij}/\pi_i]\pi_i x_i = \sum_{i=1}^{\zeta} a_i \pi_i x_i, \quad \text{where } a_i = \sum_{j=1}^{\delta_i} \alpha_{ij}/\pi_i \ ,$$

δ_i is the length of row i to the left of the partition; by Lemma A2, it follows that $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_\zeta$. Thus, $Z_{prd}X$ and $\alpha_{ij} \geq 0$ imply $a_1 \geq a_2 \geq \ldots a_\zeta \geq 0$. Lemma A1 then implies expression (A2).■

Proof of Theorem 1

Let F and G be the distribution functions of discrete random variables V and W, respectively; F(u) is defined as Pr(V ≤ u). For $V \preceq^2 W$ to hold, we need (see, e.g., Hanoch and Levy, 1969)

$$\int_a^t G(u)\,du \leq \int_a^t F(u)\,du, \text{ all } t\in[a,b],$$

where [a,b] is the range of the random variables. This we shall write as

$$G_2(t) \leq F_2(t), \text{ all } t\in[a,b]. \tag{A3}$$

If the random variable V takes values $v_1 \leq \ldots \leq v_n$ with probabilities $\alpha_1,\ldots, \alpha_n$, and $v_\tau \leq t \leq v_{\tau+1}$, then

$$\begin{aligned} F_2(t) &= \alpha_1(v_2-v_1) + (\alpha_1+\alpha_2)(v_3-v_2) + \ldots + (\alpha_1+\ldots+\alpha_{\tau-1})(v_\tau-v_{\tau-1}) \\ &\quad + (\alpha_1+\ldots+\alpha_\tau)(t-v_\tau) \\ &= -\alpha_1 v_1 - \alpha_2 v_2 - \ldots -\alpha_\tau v_\tau + \alpha_1 t + \alpha_2 t + \ldots + \alpha_\tau t \\ &= \sum_{i=1}^{\tau} \alpha_i (t-v_i). \end{aligned} \tag{A4}$$

and

$$\text{for any } S \subseteq \{1,2,\ldots,\tau\},\ F_2'(t) = \sum_{j\in S} \alpha_j (t-v_j) \leq F_2(t). \tag{A5}$$

Now, $G_2(t)$ and $F_2(t)$ are piecewise linear and expression (A3) need be checked only at those values t taken by the random variable (see, e.g., Fishburn and Vickson, 1978, p. 79). If now we take V = X + Z and W = Y + Z, the values in question are $x_i + z_j$ and $y_i + z_j$.

We first look at the points $x_i + z_j$. Define $(x+z)_{(0)} = -\infty$. Let $(x+z)_{(k)}$ denote the kth smallest element of $\{x_i + z_j\}$, i = 1,..., n: j = 1,..., m. Let now F_2 and G_2 (see expression (A3)) correspond to X+Z and Y+Z, respectively. Let r be the largest index such that $(y+z)_{(r)} \le (x+z)_{(s)}$. Then, by expression (A4),

$$G_2((x+z)_{(s)}) = -\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(y_i+z_j) + [\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}](x+z)_{(s)}.$$

Let $S_r^{xz} = \{(i,j)\in T_r^{yz} \mid x_i + z_j \le (x+z)_{(s)}\}$. Clearly, $S_r^{xz} \subseteq T_r^{xz}$. It is also clear that $S_r^{xz} \subseteq T_s^{xz}$; thus, by equation (A5), $F_2((x+z)_{(s)}) \ge F_2'((x+z)_{(s)})$, where

$$F_2'((x+z)_{(s)}) = -\sum_{(i,j)\in S_r^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in S_r^{xz}} \alpha_{ij}](x+z)_{(s)}.$$

Now, $F_2'((x+z)_{(s)}) \ge G_2((x+z)_{(s)})$ if and only if

$$-\sum_{(i,j)\in S_r^{xz}} \alpha_{ij}(x_i+z_i) + [\sum_{(i,j)\in S_r^{xz}} \alpha_{ij}](x+z)_{(s)} \ge$$

$$-\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(y_i+z_j) + [\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}](x+z)_{(s)},$$

which is equivalent to

$$\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(y_i+z_j) \ge \sum_{(i,j)\in S_r^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in T_r^{yz}\setminus S_r^{xz}} \alpha_{ij}](x+z)_{(s)}, \tag{A6}$$

where $T_r^{yz}\setminus S_r^{xz} = \{(i,j) \mid (i,j) \in T_r^{yr} \text{ and } (i,j)\notin S_r^{xz}\}$. By Lemma A3,

$$\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(y_i+z_j) \ge \sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(x_i+z_j). \tag{A7}$$

By construction of S_r^{xz}, $(x+z)_{(s)} \le x_i + z_j$ for all $(i,j) \in T_r^{yz}\setminus S_r^{xz}$. Hence, since $S_r^{xz} \subseteq T_r^{yz}$, $T_r^{yz} = S_r^{xz} \cup (T_r^{yz}\setminus S_r^{xz})$; therefore,

$$\sum_{(i,j)\in T_r^{yz}} \alpha_{ij}(x_i+z_j) \ge \sum_{(i,j)\in S_r^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in T_r^{yz}\setminus S_r^{xz}} \alpha_{ij}](x+z)_{(s)}. \tag{A8}$$

By expressions (A7) and (A8), expression (A6) holds. Thus, since $F_2((x+z)_{(s)}) \ge F_2'((x+z)_{(s)})$,

$$F_2((x+z)_{(s)}) \ge G_2((x+z)_{(s)}) \text{ for all } s = 1,\dots, N. \tag{A9}$$

We now prove the same result at all points $y_i + z_j$. We have

$$G_2((y+z)_{(r)}) = -\sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}(y_i+z_j) + [\sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}](y+z)_{(r)}.$$

Define

$$S_{r-1}^{xz} = \{(i,j) \mid (i,j)\in T_{r-1}^{yz}, x_i+z_j \le (y+z)_{(r)}\}$$

and

$$F_2'((y+z)_{(r)}) = -\sum_{(i,j)\in S_{r-1}^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in S_{r-1}^{xz}} \alpha_{ij}](y+z)_{(r)}.$$

Let q be the largest index such that $(x+z)_{(q)} \leq (y+z)_{(r)}$. It is clear that $S_{r-1}^{xz} \subseteq T_q^{xz}$. Thus, by expression (A5), we have

$$F_2((y+z)_{(r)}) \geq F_2'((y+z)_{(r)}).$$

Now, $F_2'((y+z)_{(r)}) \geq G_2((y+z)_{(r)})$ is equivalent to

$$\sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}(y_i+z_j) \geq \sum_{(i,j)\in S_{r-1}^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in T_{r-1}^{yz}\setminus S_{r-1}^{xz}} \alpha_{ij}](y+z)_{(r)}. \quad \text{(A10)}$$

By Lemma A3,

$$\sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}(y_i+z_j) \geq \sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}(x_i+z_j). \quad \text{(A11)}$$

Since $S_{r-1}^{xz} \subseteq T_{r-1}^{yz}$, $T_{r-1}^{xz} = S_{r-1}^{xz} \cup (T_{r-1}^{yz}\setminus S_{r-1}^{xz})$. Also, since $x_i + z_j \geq (y+z)_{(r)}$ for all $(i,j) \in T_{r-1}^{yz}\setminus S_{r-1}^{xz}$, we have

$$\sum_{(i,j)\in T_{r-1}^{yz}} \alpha_{ij}(x_i+z_j) \geq \sum_{(i,j)\in S_{r-1}^{xz}} \alpha_{ij}(x_i+z_j) + [\sum_{(i,j)\in T_{r-1}^{yz}\setminus S_{r-1}^{xz}} \alpha_{ij}](y+z)_{(r)}. \quad \text{(A12)}$$

Expressions (A11) and (A12) provide expression (A10), which is

$$F_2((y+z)_{(r)}) \geq G_2((y+z)_{(r)}) \text{ for } r = 1, 2, \ldots, N. \quad \text{(A13)}$$

We have proven that expressions (A9) and (A13), which together are equivalent to $X+Z \preceq^2 Y+Z$, hold.■

References

Aboudi, R. and D. Thon, 1993, Expected Utility and the Siegel Paradox: A Generalization, *Journal of Economics,* 57(1): 69-93.

Brumelle, S., 1974, When Does Diversification Between Two Investments Pay? *Journal of Finance and Quantitative Analysis,* 9: 473-483.

Cheng, K., 1977, Majorization: Its Extensions and Preservation Theorems, Technical Report 121, Department of Statistics, Stanford University, Stanford, California.

Doherty, N., H. Loubergé, and H. Schlesinger, 1987, Additive and Multiplicative Risk Premium with Multiple Sources of Risk, *Scandinavian Actuarial Journal,* 41-49.

Doherty, N. and H. Schlesinger, 1983a, The Optimal Deductible for an Insurance When Initial Wealth Is Random, *Journal of Business,* 56(4), 555-565.

Doherty, N. and H. Schlesinger, 1983b, Optimal Insurance in Incomplete Markets, *Journal of Political Economy,* 91(6): 1045-1054.

Doherty, N. and H. Schlesinger, 1986, A Note on Risk Premiums with Random Initial Wealth, *Insurance: Mathematics and Economics,* 5: 183-185.

Eeckoudt, L. and M. Kimball, 1992, Background Risk, Prudence and Insurance Demand, in: G. Dionne, ed., *Contributions to Insurance Economics* (Boston: Kluwer Academic Publishers).

Epstein, L. and S. Tanny, 1980, Increasing Generalized Correlation: A Definition and Some Economic Consequences, *Canadian Journal of Economics,* 8(1): 16-34.

Fishburn P. and R. Vickson, 1978, Theoretical Foundations of Stochastic Dominance, in: G. Whitmore and M. Findlay, eds., *Stochastic Dominance: An Approach to Decision Making Under Risk* (Lexington, Mass.: D. C. Heath).

Hadar J. and W. Russell, 1971, Stochastic Dominance and Diversification, *Journal of Economic Theory,* 3: 288-305.

Hanoch, G. and H. Levy, 1969, The Efficiency Analysis of Choices Involving Risk, *Review of Economic Studies,* 36: 335-346.

Hildreth, C. and L. Tesfatsion, 1977, A Note on Dependence Between a Venture and a Current Prospect, *Journal of Economic Theory,* 15: 381-391.

Kihlstrom, R., D. Romer, and S. Williams, 1981, Risk Aversion with Random Initial Wealth, *Econometrica,* 49(4): 911-920.

Kimball, M., 1993, Standard Risk Aversion, *Econometrica,* 61(3): 589-611.

Lehmann, E., 1966, Some Concepts of Dependence, *Annals of Mathematical Statistics,* 37: 1137-1153.

Levy, H. and Y. Kroll, 1978, Investment Decision Rules, Diversification, and the Investor's Initial Wealth, *Econometrica,* 46(5): 1231-1237.

Levy, H. and M. Sarnat, 1971, A Note on Portfolio Selection and Investor Wealth, *Journal of Financial and Quantitative Analysis,* 6: 639-642.

Marshall, A. and I. Olkin, 1979, *Inequalities: Theory of Majorization and Its Applications* (New York: Academic Press).

Mossin, J., 1968, Aspects of Rational Insurance Purchasing, *Journal of Political Economy,* 76: 553-568.

Pečarić, J., F. Proschan, and Y. Tong, 1992, *Convex Functions, Partial Orderings and Statistical Applications* (New York: Academic Press).

Pratt, J., 1964, Risk Aversion in the Small and in the Large, *Econometrica,* 32: 122-136.

Pratt, J. and R. Zeckhauser, 1987, Proper Risk aversion, *Econometrica,* 55(1): 143-154.

Ross, S., 1981, Some Stronger Measures of Risk Aversion in the Small and the Large with Applications, *Econometrica,* 49(3): 621-638.

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

Schlesinger, H., 1981, The Optimal Level of Deductible in Insurance Contracts, *Journal of Risk and Insurance,* 48: 465-481.

Schlesinger, H. and N. Doherty, 52, Incomplete Markets for Insurance: An Overview, *Journal of Risk and Insurance,* 52: 402-423.

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

Thon, D. and L. Thorlund-Petersen, 1988, Stochastic Dominance and Friedman-Savage Utility Functions, *Mathematical Social Sciences,* 16: 305-317.

Tong, Y., 1980, *Probability Inequalities in Multivariate Distributions* (New York: Academic Press).

Turnbull, S., 1983, Additional Aspects of Rational Insurance Purchasing, *Journal of Business,* 56(2): 217-229.

Wright, R., 1987, Expectation Dependence of Random Variables, with an Application in Portfolio Theory, *Theory and Decisions,* 22: 111-124.