Increasing generalized correlation: a definition and some economic consequences

LARRY G. EPSTEIN and STEPHEN M. TANNY / Institute for Policy Analysis, University of Toronto

Abstract. The question 'when is a random variable Y riskier or more variable than another random variable X ?' has recently been answered in the literature in a manner that is consistent with expected utility theory. This paper provides a similarly natural and theoretically sound definition for the statement that 'the random variables Y_1 and Y_2 are more correlated or positively interdependent than the random variables X_1 and X_2.' The usefulness of the definition is demonstrated by applying it to the determination of the effects of increased correlation on behaviour in some standard economic models.

Corrélation généralisée croissante : une définition et quelques implications économiques. Récemment on a pu répondre à la question 'quand une variable aléatoire Y est-elle plus aléatoire ou plus variable qu'une variable aléatoire X ?' d'une façon qui est consistente avec la théorie de l'utilité basée sur l'espérance mathématique. Ce mémoire veut construire une définition tout aussi naturelle et théoriquement robuste pour la proposition 'les variables aléatoires Y_1 et Y_2 sont davantage co-reliées ou davantage positivement interdépendantes que les variables aléatoires X_1 et X_2.' Les auteurs montrent l'utilité de cette définition en l'appliquant à la calibration des effets d'une corrélation plus grande sur le comportement dans des modèles économiques conventionnels.

The question 'when is a random variable Y riskier or more variable than another random variable X?' has been answered in the literature in a manner consistent with expected utility theory (see for example Hanoch and Levy, 1969; Hadar and Russell, 1969; and especially Rothschild and Stiglitz (RS), 1970). These papers consider only scalar random variables (see, however, Brumelle and Vickson, 1975, and the references therein). Once the analysis is extended to a multivariate, and in particular a bivariate, framework it seems reasonable to ask whether a similarly natural and theoretically sound definition may be provided for the statement that 'the random variables Y_1 and Y_2 are more correlated (positively interdependent or interrelated) than the random variables X_1 and X_2.'

The research described in this paper was supported in part by grants from the Canada Council and the National Research Council. This paper was presented at the Canadian Economic Theory Conference, Montreal, May 1978.

Canadian Journal of Economics / Revue canadienne d'Economique, XIII, no. 1
February / février 1980. Printed in Canada / Imprimé au Canada.

0008–4085 / 80 / 0000–0016 $01.50 /

The formulation and theoretical justification of such a definition is the objective of the first part of this paper. In the second part we demonstrate the usefulness of the definition by applying it to the determination of the effects of increased correlation on behaviour in some standard economic models.

The structure and approach of the first part are similar to those of RS. We proceed as follows: the key notion of an elementary correlation increasing transformation (CIT) of a given bivariate probability distribution is defined and is used to motivate the first definition of greater correlation. The CIT is then used to define correlation-averse and correlation-affine utility functions. They in turn are used to formulate the following plausible alternative definition of greater correlation: Y_1 and Y_2 are more correlated than X_1 and X_2 if all expected utility-maximizers who are correlation averters (lovers) prefer (disprefer) (X_1, X_2) to (Y_1, Y_2). The third section proves that the above two definitions of greater correlation are equivalent, and the fourth section compares them with others used in the literature. In particular, we point out the limited theoretical validity of the linear (Pearsonian) correlation coefficient or the covariance as measures of the positive interdependence of two random variables.

Many of the notions and results described may be found in scattered references in the literature. One contribution of this paper is to bring them into focus as the essential components of a natural and theoretically sound definition of greater correlation. In addition, we feel that our approach to Theorem 6, via elementary correlation-increasing transformations, provides further insight and a new perspective regarding the definition of greater correlation.

The second part contains a more extensive analysis of the effects of increased correlation in an expected-utility framework than may be found in existing literature. Portfolio diversification is discussed first, and then the analysis of portfolio diversification is extended to the case where future consumption good prices, as well as asset returns, are uncertain. The use of an asset as a hedge against uncertain inflation is considered. Finally, we analyse the effects of correlated price expectations in a two-period model of the behaviour of a competitive firm.

Proofs of the principal theorems in the text are collected in an appendix. Proofs of most of the remaining theorems may be found in Epstein and Tanny (1978).

A DEFINITION

Certain notational conventions are adopted. Derivatives are denoted, as is customary, by primes or by subscripted variables. Upper case letters generally refer to random variables (rv's) and lower case letters to deterministic variables. $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ are rv's with cumulative distribution functions (cdf's) F and G respectively. The corresponding marginal cdf's are denoted by $F^{(i)}$ and $G^{(i)}$, $i = 1, 2$, and the density functions by f and g. In general, we adhere to the convention that Y_1 and Y_2 are more correlated than X_1 and X_2. $F \leqslant G$ will mean $F(t_1, t_2) \leqslant G(t_1, t_2)$ for all t_1 and t_2.

Our analysis is limited to probability distributions that have compact support. For convenience, in the first part it will be assumed that rv's take on values in the unit square $[0, 1] \times [0, 1]$ with certainty. In addition we initially consider principally discrete rv's. Some of the results that we establish are then extended by standard limiting arguments to arbitrary rv's. The set $\{(a_i, b_j) : i = 1, 2, \ldots, p; j = 1, 2, \ldots, q\}$, abbreviated by $\{(a_i, b_j)\}_{p,q}$, denotes the set of possible realizations of a pair of discrete rv's and f_{ij}, g_{ij} the corresponding probabilities $\Pr(X_1 = a_i, X_2 = b_j)$, $\Pr(Y_1 = a_i, Y_2 = b_j)$ respectively. The realizations are numbered so that $a_1 < a_2 < \ldots < a_p$ and $b_1 < b_2 < \ldots < b_q$.

$u(x_1, x_2)$ denotes a von Neumann–Morgenstern utility index defined and continuous for $x_1 \geqslant 0$, $x_2 \geqslant 0$. The consequences of differentiability will be considered occasionally, but differentiability is not a maintained hypothesis.

Elementary correlation-increasing transformations

The following geometrically motivated definition of a correlation-increasing transformation (suggested by the comments of Hamada, 1974) seems intuitively correct.

DEFINITION 1: Let X and Y be discrete rv's. Then G(g or Y) is said to differ from F(f or X) by an elementary correlation-increasing transformation (CIT) if there exist $i_1 < i_2$ and $j_1 < j_2$ such that

$$g_{ij} - f_{ij} = \begin{cases} \epsilon, (i,j) = (i_1, j_1) \text{ or } (i_2, j_2) \\ -\epsilon, (i,j) = (i_1, j_2) \text{ or } (i_2, j_1) \\ 0, \text{otherwise,} \end{cases}$$

where $\epsilon > 0$.

The definition is illustrated in Figure 1, where all the non-zero values of $g - f$ are indicated. A CIT shifts weight towards realizations where both underlying rv's are 'small' or 'large' and away from realizations where one rv is 'large' and the other 'small.' Note that a CIT leaves the marginal distributions unchanged, as we would expect of a change in a bivariate distribution which is to affect only the interdependence of two rv's.

The concept of a CIT is the beginning of a definition of greater correlation. To satisfy transitivity, such a definition requires a criterion for deciding whether G could have been obtained from F by a finite sequence of CIT's. (This is the analogue of the procedure followed by RS in basing a definition of greater variability on the concept of a mean preserving spread.) This criterion is described in the following basic result.

THEOREM 1: Let F and G correspond to discrete rv's. There exists a sequence of cdf's $F = F_0, F_1, \ldots, F_n = G$ of discrete rv's such that F_k differs from F_{k-1} by a CIT, $k = 1, 2, \ldots, n$, if and only if $F \leqslant G$ and $F^{(i)} = G^{(i)}$, $i = 1, 2$.

Theorem 1 can be used to motivate the following definition.

DEFINITION 2: If F and G are arbitrary cdf's, we define the partial ordering $\leqslant_D$ as follows: $F \leqslant_D G$ if and only if $F \leqslant G$ and $F^{(i)} = G^{(i)}$, $i = 1, 2$.

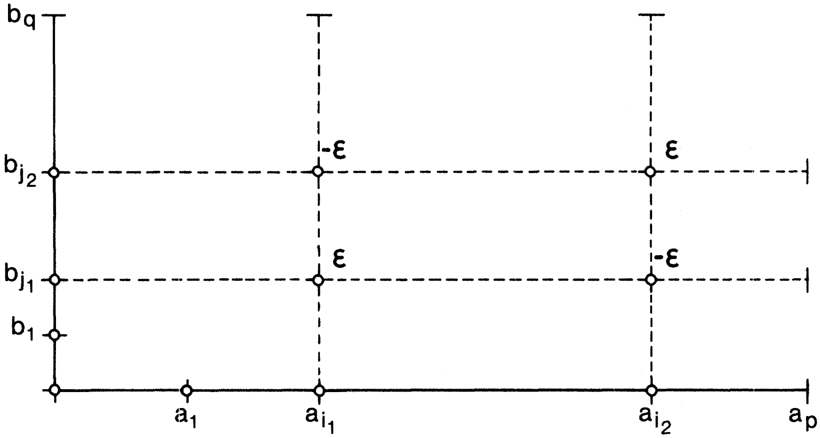

FIGURE 1

By Theorem 1, $F \leqslant_D G$ can be interpreted as stating that G exhibits greater correlation than F, or that Y_1 and Y_2 (X_1 and X_2) are more positively (negatively) correlated than X_1 and X_2 (Y_1 and Y_2), at least in the case of discrete rv's. This interpretation is extended to arbitrary rv's by observing that, if $F \leqslant_D G$, the transition from F to G may be approximated arbitrarily closely by a sequence of CIT's. More precisely:

THEOREM 2: Suppose $F \leqslant_D G$. Then there exist sequences $\{F_n\}$and $\{G_n\}$, cdf's of discrete rv's, such that $F_n \to F$ and $G_n \to G$ pointwise, and further, $F_n \leqslant_D G_n$ for all n.

The proof is similar to that of Lemma 2 of RS (232–3).

$F \leqslant_D G$ has been defined in terms of the probabilities of events of the form $(X_1 \leqslant t_1, X_2 \leqslant t_2)$.

$F \leqslant G$ ($\Pr(X_1 \leqslant t_1, X_2 \leqslant t_2) \leqslant \Pr(Y_1 \leqslant t_1, Y_2 \leqslant t_2)$) asserts roughly that the probability that X_1 and X_2 both realize 'small' values is no greater than the probability that Y_1 and Y_2 both realize 'equally small' values, suggesting that Y_1 and Y_2 are more positively interdependent than X_1 and X_2. But clearly there are other events that seem no less basic a priori and could also be used to define greater correlation. The following theorem, therefore, is essential in justifying the specific definition of $\leqslant_D$ we have adopted.

THEOREM 3: Let F and G have equal marginals. Then the following statements, each understood to be valid for all t_1 and t_2, are equivalent:

(a) $\Pr(X_1 \leqslant t_1, X_2 \leqslant t_2) \leqslant \Pr(Y_1 \leqslant t_1, Y_2 \leqslant t_2)$,

(b) $\Pr(X_1 \leqslant t_1, X_2 \geqslant t_2) \geqslant \Pr(Y_1 \leqslant t_1, Y_2 \geqslant t_2)$,

(c) $\Pr(X_1 \geqslant t_1, X_2 \leqslant t_2) \geqslant \Pr(Y_1 \geqslant t_1, Y_2 \leqslant t_2)$,

(d) $\Pr(X_1 \geqslant t_1, X_2 \geqslant t_2) \leqslant \Pr(Y_1 \geqslant t_1, Y_2 \geqslant t_2)$.

Of course inequalities (b), (c), and (d) are readily verified for rv's that differ by a CIT.

Attitudes towards correlation

This section formulates another plausible definition of greater correlation based on individual preferences over bivariate distributions.

Rothschild and Stiglitz define the scalar random variable Y' to be riskier than the scalar random variable X' if all risk averters prefer X' to Y'. Risk aversion, of course, corresponds to concavity of the utility index $v(x)$. But it can be demonstrated that v is concave if and only if expected utility unambiguously falls when the underlying probability distribution undergoes a mean preserving spread. Thus the following definition, where the CIT is used to assess attitudes towards correlation, and Definition 4 below, constitute the exact analogue of the RS procedure.

DEFINITION 3: Let $u(x_1, x_2)$ be a utility function. $u(x_1, x_2)$ is said to be correlation-averse (CAV), correlation-affine (CAF), or correlation-neutral (CN) according as expected utility is reduced, increased, or unaffected by a CIT. More precisely, assume that (Y_1, Y_2) differs from (X_1, X_2) by a CIT, so that

$$Eu(Y_1, Y_2) - Eu(X_1, X_2) = \epsilon\left[u(a_{i_2}, b_{j_2}) + u(a_{i_1}, b_{j_1}) - u(a_{i_1}, b_{j_2}) - u(a_{i_2}, b_{j_1})\right],$$

where the notation is consistent with Definition 1. Therefore, u is CAV, CAF, or CN according as

$$u(x_1, x_2) + u(y_1, y_2) \lesseqgtr u(x_1, y_2) + u(y_1, x_2), \tag{1}$$

whenever $(x_1 - y_1)(x_2 - y_2) > 0$.

It follows that

$$u(x_1, x_2) - u(x_1, y_2) \lesseqgtr u(y_1, x_2) - u(y_1, y_2),$$

or

$$u(x_1, x_2) - u(y_1, x_2) \lesseqgtr u(x_1, y_2) - u(y_1, y_2), \tag{2}$$

whenever $(x_1 - y_1)(x_2 - y_2) > 0$, so the increment in utility induced by a given increment of one attribute depends upon the level of the other attribute. This observation suggests immediately that when u is differentiable, u is CAV, CAF, or CN according as the cross partial derivative $u_{x_1 x_2} \lesseqgtr 0$, a result proved by Richard (1975).[1]

The following examples of utility functions demonstrate the link between attitudes towards correlation and attitudes towards risk.

THEOREM 4 (a) If $u(x_1, x_2) = \phi(a_1 x_1 + a_2 x_2)$, $a_1 > 0$, $a_2 > 0$ $(a_2 < 0)$, then u is CAV, CAF, or CN according as ϕ is concave (convex), convex (concave), or linear.

(b) Let u be CAV (CAF) and non-decreasing in both arguments. If v is more (less) risk-averse than u (Kihlstrom and Mirman, 1974), that is, $v(x_1, x_2) = \phi\left[u(x_1, x_2)\right]$, where ϕ is increasing and concave (convex), then v is CAV (CAF).

1 Richard (1975) used (1) to define what he called multivariate risk-averting, -seeking, and -neutral utility functions. He also proved Theorem 4(b) and (c) below under the assumption that $u_{x_1 x_2}$ exists.

(c) $u(x_1, x_2)$ is CN if and only if $u(x_1, x_2) = a(x_1) + b(x_2)$.

(a) and (b) show in what sense risk aversion implies or contributes to correlation aversion. The two attitudes are equivalent when the two attributes are perfect substitutes. In general, however, we can say only that the more risk-averse the individual, the greater 'likelihood' that the individual is also correlation-averse.

The following definition of greater correlation is surely reasonable:

DEFINITION 4: G is said to exhibit greater correlation than F (denoted $F \leqslant_u G$) if $Eu(X_1, X_2) \geqslant (\leqslant)\, Eu(Y_1, Y_2)$ for all utility functions u that are CAV (CAF).

We note that we could ostensibly weaken the definition by specifying that u be non-decreasing or non-increasing in one or both arguments. But in fact the weakening is only apparent as the two definitions are equivalent; e.g. it may be shown that $F \leqslant_u G$ is *equivalent* to the statement that $Eu(X_1, X_2) \geqslant (\leqslant)\, Eu(Y_1, Y_2)$ for all *increasing* utility functions u that are CAV(CAF).

As an immediate corollary of Theorem 4(a), we can relate increases in correlation to increases in risk in the following intuitively consistent manner:

COROLLARY 5: If (Y_1, Y_2) is more correlated than (X_1, X_2) in the sense of $\leqslant_u$, then for all $a_1 > 0$, $a_2 > (<)\, 0$, $a_1 Y_1 + a_2 Y_2$ is more (less) variable than $a_1 X_1 + a_2 X_2$ in the sense of Rothschild and Stiglitz.

A definition of increased correlation

We have formulated two plausible definitions of greater correlation. In fact they are equivalent and henceforth denoted $F \leqslant_c G$.

THEOREM 6: $F \leqslant_D G$ if and only if $F \leqslant_u G$.

Hadar and Russell (1974, Theorem 3) proved that $F \leqslant_D G \Rightarrow F \leqslant_u G$; the complete theorem was proved by Levy and Paroush (1974, Theorem 1, Corollary 3). In both cases, however, it is assumed that $u_{x_1 x_2}$ exists almost everywhere and that the distributions are absolutely continuous (assumptions with which we have dispensed).[2] Neither pair of authors seems aware, or at least makes explicit, that their analysis is relevant to a natural definition of increased correlation. This relevance is made clear by Theorem 1 and the notion of the CIT.

It seems natural to call X_1 and X_2 positively (negatively) correlated or interdependent rv's if $F^{(1)}F^{(2)} \leqslant_c ({}_c\geqslant)\, F$.

Other notions of correlation

The most frequently used measure of the interdependence between two rv's is the linear correlation coefficient or equivalently, given fixed marginal dis-

2 Their proofs apply integration by parts. Levy and Paroush consider questions of convergence for distributions with non-compact support. Our proof of Theorem 6 may be found in Epstein and Tanny (1978). It makes use of the argument underlying common stochastic dominance tests, as clearly described by Brumelle and Vickson (1975); namely, that the characteristic function corresponding to the event $(X_1 \leqslant t_1, X_2 \leqslant t_2)$, say, may be approximated arbitrarily closely by a continuous (increasing) and CAF utility function. (Also applied are the Helly-Bray Theorem, Rao (1965, 97), and the compact supports of all probability distributions.)

tributions, the covariance. The restrictive nature of mean-variance analysis has been frequently and thoroughly discussed in the literature, at least in models where only a single uncertain attribute (e.g. total wealth) is of concern. Thus, for example, variance has been shown to be a theoretically satisfactory measure of variability, given unrestricted probability distributions, if and only if utility is quadratic. The use of covariance is similarly of limited theoretical validity.

THEOREM 7: (a) $F \leqslant_c G \Rightarrow \text{cov}(X_1, X_2) \leqslant \text{cov}(Y_1, Y_2)$.

(b) $[\text{cov}(X_1, X_2) \leqslant \text{cov}(Y_1, Y_2)$ and u CAV (CAF) $\Rightarrow Eu(X_1, X_2) \geqslant (\leqslant) Eu(Y_1, Y_2)] \Leftrightarrow u$ is of the form

$$u(x_1, x_2) = a(x_1) + b(x_2) + \alpha x_1 x_2,$$

where u is CAV (CAF) if $\alpha \leqslant (\geqslant) 0$.

Part (a) is proved by Hadar and Russell (1974), while (b) may be proved by arguments analogous to those used in this paper. Of course (b) proves that the converse of (a) is false.

Covariance is also an adequate measure of interdependence if we restrict attention to bivariate normal distributions F and G. In that case Levy and Paroush (1974, 140) have observed that $F \leqslant_D G$ if and only if G differs from F precisely by having a larger covariance, where we extend the definition of $\leqslant_D$ in the obvious manner to distributions with non-compact supports.

Spurred principally by the search for sufficient conditions for the optimality of portfolio diversification, several authors have considered other notions of interdependence (see for example Samuelson, 1967, 7–8; Scheffman, 1975; Hildreth and Tesfatsion, 1977). These papers, especially the latter (Theorems 1 and 2), as well as Hadar and Russell (1974, 238–9), have extensively examined the relationships between the various notions some of which are mentioned below. We point out, in relating those discussions to ours, that (i) though several of the notions in the above papers seem plausible at first glance, none stands out as superior to the others as the most appropriate measure of interdependence – they are all lacking a strong intuitive justification comparable to that which we have provided above by means of the CIT; (ii) the former are concerned with notions of positive or negative interdependence, rather than with a definition of more or less interdependence. Moreover, it is not clear in most cases whether a plausible extension to the more general framework is possible.

Even if we restrict attention to notions of positive or negative interdependence, we can present a strong argument in favour of the definition described in this paper. Consider the following alternative definitions of the statement that 'X_1 and X_2 are negatively interdependent':

C.1 $\quad \text{cov}(X_1, X_2) \leqslant 0$;

C.2 $\quad F(x_1, x_2) \leqslant F^{(1)}(x_1)F^{(2)}(x_2)$, $\quad$ for all x_1, x_2 in $[0, 1]$.

We would argue that a proper notion of negative interdependence should be 'ordinal,' or invariant to increasing transformations of the rv's, i.e. X_1 and X_2 should be said to be negatively interdependent if and only if $\theta_1(X_1)$ and $\theta_2(X_2)$ can be said to be negatively interdependent for all increasing functions θ_1 and θ_2. C.2 is invariant with respect to increasing transformations, while C.1 is invariant only to linear transformations. Moreover, if C.1 is strengthened to make it invariant with respect to all increasing transformations, the resulting definition is equivalent to C.2. This is the content of the following theorem.

THEOREM 8: The following statements are equivalent:

(a) $\text{cov}(\theta_1(X_1), \theta_2(X_2)) \leqslant 0$ for all increasing θ_1 and θ_2.

(b) $F(x_1, x_2) \leqslant F^{(1)}(x_1)F^{(2)}(x_2)$ for all x_1, x_2 in $[0, 1]$.

Once it is accepted that an ordinal notion is desirable, therefore, C.2 emerges as the natural definition. Samuelson's (1967, 8) definition, which requires that $\partial \Pr(X_i \leqslant x_i / X_j = x_j)/\partial x_j \leqslant 0$ for $i \neq j$, is also ordinal. However, it can be shown to be stronger than C.2 and thus unnecessarily restrictive since it does not appear to be useful in generating any behavioural propositions not derivable using C.2. (See below for examples of the latter propositions.) Finally, it may be shown that the notions investigated by Hildreth and Tesfatsion (1977) and Scheffman (1975) are both equivalent to C.2, when they are strengthened to be invariant to increasing transformations.

SOME ECONOMIC CONSEQUENCES

This part of the paper demonstrates the usefulness of our definition for deriving comparative static effects of increased correlation. Following Rothschild and Stiglitz (1971), we consider decision problems of the form

$$\max_{\alpha \geqslant 0} EU(\alpha; X_1, X_2), \tag{3}$$

where $U(\alpha; x_1, x_2)$ represents the individual's utility function, which depends on a decision variable α and on the exogenous variable $x = (x_1, x_2)$. The individual entertains expectations concerning the possible future values of x described by the vector rv $X = (X_1, X_2)$, and chooses α to maximize ex ante expected utility.

Let $\alpha^* > 0$ be the unique optimal decision which must satisfy the first-order condition

$$EU_\alpha(\alpha^*; X) = 0.$$

Assume that U is strictly concave in α. A modification of the now familiar RS argument shows that an increase in the correlation between X_1 and X_2 will increase (reduce) α^* if $U_\alpha(\alpha^*; x_1, x_2)$ is CAF (CAV), or, assuming differentiability, if $U_{\alpha x_1 x_2}(\alpha^*; x_1, x_2) \geqslant (\leqslant)\ 0$. Moreover, if U_α is neither CAF nor CAV, the effect of an increase in correlation is ambiguous.

We proceed to analyse some specific instances of the general decision problem (3).

Portfolio diversification

We begin with an analysis of the effect of the correlation of asset returns on the optimality and degree of portfolio diversification in a two-asset model.[3]

Consider an individual that solves

$$\max_{0 \leq \alpha \leq 1} Eu(\alpha X_1 + (1 - \alpha)X_2), \tag{4}$$

where u is a strictly concave utility index of wealth, X_1 and X_2 are the stochastic gross returns of the two assets, and α is the decision variable. We call the optimal portfolio diversified if $0 < \alpha^* < 1$. Note that by Theorem 4(a), a risk-averse individual is averse to greater correlation between asset returns.

THEOREM 9: Let $Y = (Y_1, Y_2)$ be such that diversification is optimal, and let (X_1, X_2) be more negatively correlated than (Y_1, Y_2). Then for (X_1, X_2) diversification is also optimal.[4]

Samuelson (1967, 6) has shown that diversification is optimal for all risk averters if the two assets are independent and have identical means. Thus the theorem implies that diversification is also optimal for risk averters if the assets have equal means and are negatively correlated in the sense described above, a result that has been proven by Hadar and Russell (1974). In general, the theorem demonstrates that more negative interdependence between assets can only strengthen the case for diversification (see Samuelson, 1967, 7).

Several authors (cited in the previous section) have formulated notions of negative correlation which they show to be sufficient for diversification to be optimal. We now show that there is a sense in which negative correlation, defined as above, is necessary as well as sufficient for portfolio diversification.

It is easier to consider the weak inequalities $0 \leq \alpha^* \leq 1$ where it is not optimal to go short in either asset. Therefore we describe the relationship between negative interdependence and the non-optimality of short holdings. Brumelle (1974, 479) and Hadar and Russell (1971, 299) have noted that if $0 \leq \alpha^* \leq 1$ for all risk averters, then $EX_1 = EX_2$ necessarily. Henceforth, we maintain the assumption of equal means for all assets in a given portfolio.

THEOREM 10: Consider the general portfolio problem (4), where $EX_1 = EX_2$. The following two statements are equivalent:

(a) $F \leq F^{(1)}F^{(2)}$,

(b) In (4) let asset returns be described by $(\theta_1(X_1), \theta_2(X_2))$ instead of (X_1, X_2), where θ_1 and θ_2 are increasing functions such that $E\theta_1(X_1) = E\theta_2(X_2)$. The solution α^* satisfies $0 \leq \alpha^* \leq 1$ for any such functions θ_1 and θ_2.

Some perspective on the theorem is provided by restricting attention to a mean-variance world. The theorem remains valid if (a) is modified to $\text{cov}(X_1, X_2) \leq 0$ and if the transformations θ_1 and θ_2 in (b) are restricted to be linear.

3 The analysis may be extended, largely unsatisfactorily, to n assets by forcing the n-asset model into a two-asset framework as in Scheffman (1975, 282–4).

4 The proof is straightforward and is omitted. It makes use of the fact that the functions $(x_1 - x_2)u'(x_2)$ and $(x_1 - x_2)u'(x_1)$ are correlation-averse and -affine respectively.

Thus the theorem provides further justification for the view that the condition $F \leqslant F^{(1)}F^{(2)}$ is an extension to the case of an inverse, non-linear interdependence of the standard notion of an inverse, linear interdependence between two rv's represented by a negative covariance.

By paying due attention to strict versus weak inequalities we may prove an analogous theorem which summarizes the relationship between negative interdependence and portfolio diversification.

To conclude this section, we turn from the question of the optimality of diversification to an examination of the degree of diversification. Consider the following question: if an investor diversifies when asset returns are (Y_1, Y_2), will a change to the more negatively correlated returns (X_1, X_2) induce him to 'diversify more,' in the sense that he will divide his total investment more equally between the two assets?

Denote by α^* $(\alpha^{**}) > 0$ the optimal decisions, given (X_1, X_2) and (Y_1, Y_2) respectively, and define $U(\alpha;\, X_1,\, X_2) \equiv u(\alpha X_1 + (1 - \alpha)X_2)$. Then α^* is determined by

$$EU_\alpha(\alpha^*; X_1X_2) = E[u'(\alpha^*X_1 + (1 - \alpha^*)X_2)(X_1 - X_2)] = 0. \tag{5}$$

From above, $\alpha^{**} \geqslant (\leqslant)\, \alpha^*$ if $U_{\alpha X_1X_2}(\alpha^*; X_1, X_2) \geqslant (\leqslant) = 0$. But

$$U_{\alpha X_1 X_2} = (1 - 2\alpha^*)u'' + \alpha^*(1 - \alpha^*)(X_1 - X_2)u''', \tag{6}$$

which can be uniformly signed only in special cases.

For example, if u is quadratic the second term on the right side of (6) vanishes and a reduction in correlation reduces (increases) α^* if $\alpha^* > (<) 1/2$. In fact we can show that $\alpha^* < 1/2 \Rightarrow \alpha^* < \alpha^{**} < 1/2$, and that $\alpha^* > 1/2 \Rightarrow \alpha^* > \alpha^{**} > 1/2$. These statements remain valid if instead of postulating a quadratic utility function we assume that asset returns are bivariate normal, in which case the investor may be viewed as maximizing a utility function of the expected value and variance of total wealth. Thus in a mean-variance world more negative correlation induces great diversification.

In general, however, the second term in (6) cannot be ignored. The sign of the latter can be interpreted in the following way: consider the first-order equation (5). When asset returns become more positively correlated two partial influences may be isolated: (i) there is an increase in correlation between aggregate future wealth and X_1 (X_2) weighted by the factor $\alpha^*(1 - \alpha^*)$; (ii) for any given holding of each asset there is an increase in the variability of future wealth, e.g. $\alpha^* Y_1 + (1 - \alpha^*) Y_2$ is riskier, in the sense of RS (1970), than $\alpha^*X_1 + (1 - \alpha^*)X_2$ (see Corollary 5). Moreover, it is readily demonstrated that the two terms on the right side of (6) correspond precisely to the qualitative impacts of (i) and (ii) respectively on the optimal portfolio. In general, an investor will respond unambiguously to greater correlation by increasing the 'degree of specialization' in his portfolio only if he is perfectly compensated for the induced increase in the variability of total future wealth.

Such compensation is unnecessary in a mean-variance framework because the optimal portfolio is unaffected by the change in variability ($u''' = 0$).[5]

A portfolio problem with uncertain prices

Generally, future consumption rather than future wealth is the objective of portfolio decisions. When future prices are certain, this distinction is inconsequential for the analysis of portfolio problems. However, in the more realistic situation where there is some uncertainty about future prices or about the future rate of inflation, the conventional portfolio model must be modified.

We consider the problem

$$\max_{0 \leq \alpha \leq 1} Eu\left(\frac{\alpha X + (1 - \alpha)}{Q}\right), \tag{7}$$

where X denotes the gross return to the single nominally risky asset, called bonds. The only other asset, called money, has the nominally certain net return of zero. Q represents expectations concerning the future price of a single composite good, and $u(c)$ is a twice differentiable, strictly concave, von-Neumann Morgenstern utility index of future consumption.

We are interested in the effect on the demand for bonds of the correlation between X and Q, i.e. of the extent to which bonds serve as a hedge against uncertain inflation. This question has been investigated by Boonekamp (1978) under the assumption of small risks, where, in the spirit of Samuelson (1970), the use of covariance may be justified as a measure of interdependence.[6] Some of his results can be generalized using the notion of correlation developed above.

The following terminology and notation will be adopted. Bonds are a positive (negative) hedge against inflation if X and Q are positively (negatively) correlated. Y is a better hedge against inflation if (Y, Q) is more correlated than (X, Q). RRA denotes the measure of relative risk aversion $-cu''(c)/u'(c)$. Finally, $W \equiv \alpha X + (1 - \alpha)$, $U(\alpha; X, Q) \equiv u(\alpha X + (1 - \alpha)/Q)$ and $h(W; Q) \equiv u(W/Q)$. Note that positive correlation between X and Q is desirable if and only if $U_{XQ} = \alpha u'(\text{RRA} - 1) > 0$, i.e. RRA > 1.

Conditions under which the hedging power of bonds encourages the positive holding of bonds are described in the following theorem.

THEOREM 11: (a) Suppose that $\alpha^* > 0$ in (7). Then there is a positive demand for bonds also if the return to bonds is described by Y and if (i) RRA > 1 and Y is a better hedge or (ii) RRA < 1 and X is a better hedge.

(b) Suppose that $EX = 1$. Then the demand for bonds is zero if Q is certain. There is a positive (zero) demand for bonds if Q is uncertain and if (iii) RRA $>$ ($<$) 1 and bonds are a positive hedge or (iv) RRA $<$ ($>$) 1 and bonds are a negative hedge.

5 The change in variability acts in an additve fashion, i.e. as though expected utility changed from $Eu(\alpha X_1 + (1 - \alpha)X_2)$ to $Eu(\alpha X_1 + (1 - \alpha) X_2 + S)$ for any α, where S is a stochastic lump-sum wealth, $E[S/X_1, X_2] = 0$. In a mean variance world α^* is the same in both problems.

6 Other related papers include Roll (1973) and Fischer (1975).

If RRA $= 1$ identically, the demand for bonds vanishes.[7]

That risky bonds may be demanded even if they yield an expected zero net return in nominal terms is of course not surprising when we realize that neither a bond nor money is safe in real terms if inflation is uncertain. The dependence on the RRA of an investor's attitude towards correlation between X and Q and of the influence of the latter on the attractiveness of investment in bonds is also not surprising in view of the following observations. An increase in correlation between bond returns and Q leaves the real return to money $1/Q$ unaffected, but it affects the real return to bonds in two ways. First, the function x/q is correlation-averse, so that $E[X/Q] > E[Y/Q]$ if Y is a better hedge. Second, a moment's thought suggests that Y/Q should be less variable than X/Q. In fact we can prove that the change in real returns from Y/Q to $X/Q + E[Y/Q] - E[X/Q]$ constitutes an increase in riskiness in the sense of Rothschild and Stiglitz.[8] Therefore, the reduction in the expected real return is at least partially offset by a reduction in the variability of the real return, when the hedging power of bonds is increased. The tradeoff between the two effects depends on the degree of risk aversion as measured by the RRA.

We note Arrow's (1965) observation that RRA $\geqslant 1$ for large consumption values is necessary if utility is bounded from above. To the extent that the latter is a reasonable hypothesis, a positive demand for bonds is made more likely by greater positive hedging power, at least when rv's are such that 'sufficiently' large consumption is assured.

Suppose now that the demand for bonds is positive and we investigate the way in which the magnitude of the demand is affected by the hedging power of bonds. Let α^* (α^{**}) > 0 be optimal, given (X, Q) $((Y, Q))$. From the first-order condition

$$EU_\alpha(\alpha^*; X, Q) = E[h_w(\alpha^* X + 1 - \alpha^*; Q)X] = 0, \tag{8}$$

we see that $\alpha^{**} - \alpha^*$ has the sign of

$$h_{WQ} + \alpha^* X h_{WWQ}, \tag{9}$$

when the latter is uniformly signed.

Focusing on (8) helps to isolate the following partial effects induced by an increase in the hedging power of bonds: (i) the correlation between bond returns and Q is increased, and (ii) for the given bond holdings α^* the correlation between total future wealth W and Q is increased in proportion to α^*. It is easy to see that the terms h_{WQ} and $\alpha^* X h_{WWQ}$ represent the qualitative effects on bond holdings of (i) and (ii) respectively. $h_{WQ} = u'(\text{RRA} - 1)/Q^2$, so that the partial change (i) increases (reduces) the demand for bonds if RRA $>$ ($<$) 1, consistent with the expectations promoted by Theorem 11. However,

7 The theorem is a simple consequence of the observation that the function $v(x, q) \equiv (x - 1)u'(1/q)/q$ is correlation-averse, -affine or -neutral according as RRA $>$, $<$, or $=$ 1. See also Hadar and Russell (1971, 303).

8 Prove first that when all rv's are discrete and when (Y, Q) differs from (X, Q) by a CIT, then $X/Q + E[Y/Q] - E[X/Q]$ differs from Y/Q by a mean preserving spread. Then use the limiting arguments of RS (Lemma 2) and Theorem 2 above.

the over-all impact on α^* may differ because of the 'wealth effect,' unless the investor is perfectly compensated for the increase in correlation between W and Q. (For example, if RRA is constant then RRA $< 1 \Rightarrow \alpha^{**} < \alpha^*$; but if RRA $>$ 1 the sign of $\alpha^{**} - \alpha^*$ is ambiguous.) We note therefore that Boonekamp's findings for the case of small risks do not generalize directly, as they correspond to the effects of (i) only.

A firm's production problem

Consider a firm that produces a single output y using the two-factor strictly concave production function $y = f(L, \alpha)$, where L and α denote labour and capital respectively. Capital must be ordered immediately, but production and sales do not take place until next period. Corresponding to y, L, and α are the prices P, W, and q, where P and W are rv's that describe expectations about next period's (discounted) prices for output and labour respectively. The firm wishes to maximize expected profits, i.e. it solves

$$\max_{\alpha \geqslant 0} Eg(P, W; \alpha) - q\alpha, \tag{10}$$

where

$$g(p, w; \alpha) \equiv \max_{y,\, L \geqslant 0} \{py - wL \,|\, y = f(L, \alpha)\} \tag{11}$$

is the variable profit function corresponding to f (see Diewert, 1974). It gives the maximum variable profits attainable given the capital stock α and output and labour prices p and w.

(The following important property of the variable profit function, known as Hotelling's Lemma, will be useful: denote by $\hat{L}(p, w; \alpha)$ and $\hat{y}(p, w; \alpha)$ the solutions to the optimization problem (10). Then these short-run demand and supply functions are given by $\hat{L} = -g_w$ and $\hat{y} = g_p$. We also note that g is strictly concave in α and convex and linearly homogeneous in prices. The latter two properties imply that $g_{pw} = -pg_{pp}/w < 0$ and that $g_{\alpha pw} = -pg_{\alpha pp}/w = -wg_{\alpha ww}/p$.)

The effects on the demand for capital of increased variability of price expectations in such a model have been analysed by Hartman (1976) and Epstein (1978). Here we wish to investigate the effects of correlation between output price and wage rate expectations. Therefore, we identify $g(P, W; \alpha) - q\alpha$ with the function $U(\alpha; P, W)$ of (3).

The producer is necessarily averse to correlation between P and W because $U_{pw} = g_{pw} < 0$. It is clear from Hotelling's Lemma that the attitude towards correlation induced by the technology depends on the short-run substitution-complementarity relations among future decision variables. The unambiguous attitude in this model is a consequence of the assumption that there are only two variables in the short run, output and labour, so that a regressive relationship (Hicks, 1946, chap 7) between the product and factor is ruled out. (It is also a consequence of the assumption of profit risk neutrality on the part of the producer. If he were averse to profit uncertainty and maximized the expected value of a concave utility function of profits, the aversion to profit risk would induce a degree of affinity for price and wage rate correlation that could partially or completely offset the technologically induced version.)

A more interesting question is what happens to the optimal level of α when price and wage rate expectations become more correlated. The first-order condition that defines the unique solution α^* to (10) is

$$Eg_{\alpha}(P, W; \alpha^*) - q = 0. \tag{12}$$

Therefore, the effect of increased correlation depends on the sign of $g_{\alpha pw}$. $g_{\alpha pp}$ and $g_{\alpha ww}$ have the same sign, which is the opposite of the sign of $g_{\alpha pw}$. The former signs determine (by the Hartman and Epstein papers) the qualitative impacts on α^* of increased variability of P and W respectively. Therefore, in this simple model increases in correlation and increases in variability have opposite qualitative effects.

We can say more in the special case of a CES production function that is homogeneous of degree $\mu < 1$ and has elasticity of substitution σ. From equations (19) and (27) of Hartman (1976), we may conclude that greater price and wage rate correlation increases (reduces) the demand for capital if $\sigma > 1/(1 - \mu)$ $(\sigma < 1/(1 - \mu))$.

APPENDIX: PROOFS OF THEOREMS

THEOREM 1: Necessity follows immediately from the definitions of $F \leqslant G$ and a CIT, together with the earlier remark that a CIT leaves the marginal distributions unchanged. For sufficiency we extend and subsequently generalize some results of Hardy, Littlewood, and Polya (HLP) (1934, 45) on the majorization of finite sequences.

LEMMA 1: Let f_i, g_i, $i = 1, 2, \ldots, n$ be two sequences with $0 \leqslant f_i, g_i \leqslant 1$ for all i. Suppose that

$$\text{(a)} \sum_{i=1}^{k} f_i \leqslant \sum_{i=1}^{k} g_i, \qquad k = 1, 2, \ldots, n-1,$$

$$\text{(b)} \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} g_i.$$

We write $(f_i) \propto (g_i)$. Further, for $k < l$ define a transformation $T = T(k, l; \delta)$ on sequences by $T(k, l; \delta)(f_i) = (f_i')$, where

$$f_i' = \begin{cases} f_i + \delta & i = k \\ f_i - \delta & i = l \\ f_i & i \neq k, l. \end{cases}$$

Then (g_i) can be obtained from (f_i) by the successive application of a finite number of transformations $T_1, T_2, \ldots, T_s$, and

$$(f_i) \propto T_1(f_i) \propto T_1T_2(f_i) \propto \ldots \propto T_1T_2 \ldots T_s(f_i) = (g_i). \tag{13}$$

Note that $T_1T_2 \ldots T_k(f_i)$ means that T_1 is applied first, then T_2 is applied to $T_1(f_i)$, and so on.

A special case of the Lemma is proved by HLP in which it is assumed that $f_1 \geqslant f_2 \geqslant \ldots \geqslant f_n$ and $g_1 \geqslant g_2 \geqslant \ldots \geqslant g_n$.

PROOF: It is obvious that $(f_i) \propto T(f_i)$ for any sequence (f_i) and transformation T, so we need only show the existence of $T_1, T_2, \ldots, T_s$ such that $T_1T_2 \ldots T_s(f_i) = (g_i)$. We proceed by induction on the number m of non-zero differences among the elements $(g_i - f_i)$. Clearly $m = 1$ is impossible while $m = 2$ is obvious. Notice that (a) implies that the first non-zero difference is positive. Let k be the first index such that $g_k - f_k = \delta_k > 0$; then by (b) there exists $l > k$ where l is the first index such that $f_l - g_l = \delta_l > 0$. Set $\delta = \min(\delta_k, \delta_l)$ and $(f_i') = T(k, l; \delta)$. Then $(f_i) \propto (f_i') \propto (g_i)$, and the set of elements $(g_i - f_i')$ has at most $m - 1$ non-zero values; hence we can apply the induction hypothesis to conclude the proof.

Return now to the proof of the theorem. Note that $F^{(i)} = G^{(i)}$, $i = 1, 2$, is equivalent to

$$\text{(i)} \sum_{i=1}^{p} f_{ij} = \sum_{i=1}^{p} g_{ij}, \qquad \text{for all } j,$$

$$\text{(ii)} \sum_{j=1}^{q} f_{ij} = \sum_{j=1}^{q} g_{ij}, \qquad \text{for all } i,$$

while the restriction $F \leqslant G$ means that

$$\text{(iii)} \sum_{(i_0, j_0)} f_{ij} \leqslant \sum_{(i_0, j_0)} g_{ij}$$

for all (i_0, j_0), where the sum is taken over all (i, j) such that $(i, j) \leqslant (i_0, j_0)$. Finally, we have

$$\text{(iv)} \sum_{i,j} f_{ij} = \sum_{i,j} g_{ij} = 1,$$

$$0 \leqslant f_{ij}, g_{ij} \leqslant 1 \text{ for all } i \text{ and } j.$$

We abbreviate (i) to (iv) by writing $(f_{ij}) \propto^* (g_{ij})$.

Let $\tau = \tau(k, l; r, s; \epsilon)$ be a correlation increasing transformation (CIT) defined by $\tau(k, l; r, s; \epsilon)(f_{ij}) = (f_{ij}')$, where $k < r$, $l < s$, and

$$f_{ij}' = \begin{cases} f_{ij} + \epsilon, & (i,j) = (k, l) \text{ or } (r, s), \\ f_{ij} - \epsilon, & (i,j) = (k, s) \text{ or } (r, l), \\ f_{ij}, & \text{otherwise.} \end{cases}$$

We always assume that the choice of ϵ is restricted so that (f_{ij}') is a probability matrix. It is obvious that $(f_{ij}) \propto^* \tau(f_{ij})$ for any τ. We show that it is possible to find $\tau_1, \tau_2, \ldots, \tau_s$ such that

$$(f_{ij}) \propto^* \tau_1 (f_{ij}) \propto^* \ldots \propto^* \tau_1\tau_2 \ldots \tau_s(f_{ij}) = (g_{ij}),$$

where $\tau_1\tau_2 \ldots \tau_s(f_{ij})$ means τ_1 is applied to (f_{ij}), τ_2 to $\tau_1(f_{ij})$, and so on.

We proceed by induction on p. For $p = 1$ the result is trivial since $(f_{ij}) = (g_{ij})$; for $p = 2$ note that the sequences (f_{ij}) and (g_{ij}) satisfy the conditions of the Lemma, and hence there is a sequence of T-transformations transforming (f_{ij}) into (g_{ij}). If $T(k, l; \delta)$ is any such transformation acting on (f_{ij}), let $\tau(1, k; 2, l; \delta)$ act on (f_{ij}). By (i) to (iv) it follows that the sequence of τ transformations so defined transform (f_{ij}) into (g_{ij}).

Assume the result up to $p - 1 \geqslant 2$. We prove it for p by generalizing the preceding argument for $p = 2$. Consider the sequences (f_{1j}), (g_{1j}). If $(f_{1j}) = (g_{1j})$ we can apply the induction assumption to the remaining $p - 1$ rows in an obvious manner to obtain the desired result, so assume $(f_{1j}) \neq (g_{1j})$. It follows from (ii) and (iii) that the sequences satisfy $(f_{1j}) \propto^* (g_{1j})$ so there is a sequence of T-transformations $T_1, T_2, \ldots, T_r$ such that

$$(f_{1j}) \propto^* T_1(f_{1j}) \propto^* T_1T_2(f_{1j}) \propto^* T_1T_2 \ldots T_r(f_{1j}) = (g_{1j}).$$

Consider first $T_1 = T(k, l; \delta)$, $k < l$. It follows that

$$f_{1k} + \delta \leqslant g_{1k}, \tag{14}$$

where we may assume that $f_{1j} = g_{1j}$ for $j < k$. We also have

$$f_{1l} - \delta \geqslant g_{1l}. \tag{15}$$

We now operate on the 'columns' k, l of the array of probabilities to show that we can reduce elements in column k by amounts which sum to δ while we add equal amounts to *corresponding* elements in the same row in column l. The only constraint on the sizes of these amounts is that as a result of these changes no element becomes negative or exceeds 1 (i.e. we are left with a probability matrix). In this way we can define a collection of transformations $\{\tau(1, k; j, l; \delta_j)\colon \Sigma_j\delta_j = \delta, \delta_j > 0\}$ which act upon (f_{ij}) in a way that extends the action of T_1 on the sequence (f_{ij}). In the same way we find for each $T_1, T_2, \ldots, T_r$ the corresponding set of τ-transformations. After all such τ-transformations have been applied to (f_{ij}) we obtain (f_{ij}'), where $(f_{1j}) = (f_{1j}')$, so we can apply induction to the remaining $p - 1$ rows in (f_{ij}') and (g_{ij}) to obtain the desired result.

To establish the existence of the τ-transformations $\tau(1, k; j, l; \delta_j)$ for T_1, we proceed constructively. The only constraint on the δ_j is that the resulting matrix is a probability matrix. We choose δ_j for the entry (j, k), $j \geqslant 2$, as the maximum possible so that $f_{jk} - \delta_j \geqslant 0$ and $f_{jl} + \delta_j \leqslant 1$, and $\Sigma\delta_j < \delta$. We stop when we reach the equality $\Sigma_j\delta_j = \delta$, which must occur because of the following argument. Since

$$\sum_{j=2}^{p} f_{jk} = \sum_{j=1}^{p} g_{jk} - f_{1k} = (g_{1k} - f_{1k}) + \sum_{j=2}^{p} g_{jk} \geqslant \delta,$$

a total of at least δ can be removed from the entries in column k. Also, $f_{1l} - \delta \geqslant g_{1l} \geqslant 0$ so

$$\sum_{j=2}^{p} f_{jl} \leqslant \sum_{j=1}^{p} g_{jl} - \delta,$$

it follows that any entry f_{jl} in column l in rows 2 to p can be increased by any amount not exceeding δ without violation of the upper bound constraint. Thus, we can set $\delta_2 = \min(f_{2k}, \delta)$, $\delta_3 = \min(f_{3k}, \delta - \delta_2)$, and in general

$$\delta_j = \min\left(f_{jk}, \delta - \sum_{r=2}^{j-1} \delta_r\right).$$

In this way we generate the values of δ_j for the τ-transformations for T_1, which concludes the proof.

THEOREM 4: (a) Take $a_1 = a_2 = 1$ for simplicity. u is CAV if and only if

$$\phi(x_1 + x_2) + \phi(y_1 + y_2) \leqslant \phi(x_1 + y_2) + \phi(y_1 + x_2) \tag{16}$$

whenever $(x_1 - y_1)(x_2 - y_2) > 0$. The left side of (16) may be viewed as twice the expected value of a lottery L paying $(x_1 + x_2)$ or $(y_1 + y_2)$ each with probability $\frac{1}{2}$. Similarly the right side of (16) corresponds in an obvious manner to a lottery L'. Note that

$$\min\{y_1 + y_2, x_1 + x_2\} < x_1 + y_2, x_2 + y_1 < \max\{y_1 + y_2, x_1 + x_2\},$$

and that the lotteries have the same expected values. Therefore, L is a mean preserving spread of L' and from RS all risk averters prefer L' to L. Conversely, if all such L' are preferred to the corresponding L, take $x_1 + y_2 = y_1 + x_2 = t$. Then $\frac{1}{2}\phi(x_1 + x_2) + \frac{1}{2}\phi(y_1 + y_2) \leqslant \phi(t)$ and $\frac{1}{2}(x_1 + x_2) + \frac{1}{2}(y_1 + y_2) = t$, proving that ϕ is concave. The remainder of (a) is proved similarly.

(b) Let u be non-decreasing and CAV and $\phi(t)$ increasing and concave. Given $(x_1 - y_1)(x_2 - y_2) > 0$, we must show that $\phi(u(x_1, x_2)) + \phi(u(y_1, y_2)) \leqslant \phi(u(x_1, y_2)) + \phi(u(y_1, x_2))$. As in (a), define the lottery L that pays $u(x_1, x_2)$ or $u(y_1, y_2)$, each with probability $\frac{1}{2}$, and define the lottery L' that pays $u(x_1, y_2)$ or $u(y_1, x_2)$, each with probability $\frac{1}{2}$. Then

$$\min\{u(x_1, x_2), u(y_1, y_2)\} \leqslant u(x_1, y_2), u(y_1, x_2) \leqslant \max\{u(x_1, x_2), u(y_1, y_2)\},$$

so that L has more weight in the tails than does L'. L also has a smaller expected payoff because u is CAV. Therefore L' is preferred to L according to the increasing, concave utility function ϕ. The remainder of (b) is proved similarly.

(c) Sufficiency of additivity is clear. For necessity, take $y_1 = y_2 = 0$ in (1). Then correlation neutrality implies that for all $x_1, x_2 > 0$, $u(x_1, x_2) = u(0, 0) + u(0, x_2) + u(x_1, 0)$. By continuity, the equality may be extended to $x_1, x_2 \geqslant 0$.

THEOREM 8: Since C.2 is ordinal and C.2 => C.1, it is enough to show that (a) => (b). (a) states that $E[\theta_1(X_1)\theta_2(X_2)] \leqslant E\theta_1(X_1)\cdot E\theta_2(X_2)$ for all increasing functions θ_1 and θ_2. Fix points t_1, t_2 in [0, 1]. We may find a sequence of differentiable and (strictly) increasing functions that converges pointwise to the function $h(x_1)$ and a similar sequence that converges to the function $g(x_2)$, where

$$h(x_1) = \begin{cases} 0 & x_1 < t_1 \\ 1 & x_1 \geqslant t_1 \end{cases} \qquad g(x_2) = \begin{cases} 0 & x_2 < t_2 \\ 1 & x_2 \geqslant t_2 \end{cases} .$$

Therefore, by the bounded convergence theorem, $E[h(X_1)g(X_2)] \leqslant Eh(X_1) \cdot Eg(X_2)$, or $\Pr(X_1 \geqslant t_1, X_2 \geqslant t_2) \leqslant \Pr(X_1 \geqslant t_1) \cdot \Pr(X_2 \geqslant t_2)$. C.2 follows by Theorem 3. Thus the theorem is valid even if the transformations θ_1 and θ_2 in the statement of the theorem are required to be differentiable and strictly increasing.

THEOREM 10: (a) = >(b): $F \leqslant F^{(1)}F^{(2)}$ implies the corresponding inequality for the cdf and marginal cdf's associated with $(\theta_1(X_1), \theta_2(X_2))$. (b) follows from the discussion immediately following Theorem 9.

(b) = >(a): By a result in Brumelle (1974, 479), (b) = $>E[\theta_1(X_1)/\theta_2(X_2) \leqslant b] \geqslant E[\theta_2(X_2)/\theta_2(X_2) \leqslant b]$ for all b in $\theta_2([0, A])$ = $>E[\theta_1(X_1)/X_2 \leqslant a] \geqslant E[\theta_2(X_2)/X_2 \leqslant a]$ for all a in $[0, A]$. (By limiting arguments, these inequalities may be extended from continuous and strictly increasing functions θ_1 and θ_2 to weakly increasing (or constant) functions having a jump discontinuity.) Pick t_1 in $[0, A]$ such that $\Pr(X_1 \geqslant t_1) > 0$. Define

$$\theta_1(x_1) = \begin{cases} 0, & x_1 < t_1, \\ \dfrac{1}{\Pr(X_1 \geqslant t_1)}, & x_1 \geqslant t_1 \end{cases} \qquad \theta_2(x_2) = 1 \text{ for all } x_2.$$

Then $E\theta_1(X_1) = E\theta_2(X_2) = 1$, and the above inequalities imply

$$\Pr(X_1 \geqslant t_1, X_2 \leqslant a)/\Pr(X_1 \geqslant t_1) \geqslant \Pr(X_2 \geqslant a),$$

or

$$\Pr(X_1 \geqslant t_1, X_2 \leqslant a) \geqslant \Pr(X_1 \geqslant t_1)\Pr(X_2 \geqslant a), \tag{17}$$

for all a and all t_1 for which $\Pr(X_1 \geqslant t_1) > 0$. If $\Pr(X_1 \geqslant t_1) = 0$, (17) holds trivially. Therefore, by Theorem 3, $F(x_1, x_2) \leqslant F^{(1)}(x_1)F^{(2)}(x_2)$ for all x_1, x_2, proving (a).

REFERENCES

Arrow, K.J. (1965) *Aspects of the Theory of Risk Bearing* (Helsinki: Yrjö Jahnsson Lectures)

Boonekamp, C.F.J. (1978) 'Inflation, hedging and the demand for money.' *American Economic Review* 68, 821–33

Brumelle, S.L. (1974) 'When does diversification pay?' *Journal of Financial and Quantitative Analysis* 9, 473–83

Brumelle, S.L. and R.G. Vickson (1975) 'A unified approach to stochastic dominance.' In W.T. Ziemba and R.G. Vickson, eds, *Stochastic Optimization Models in Finance* (New York: Academic Press)

Diewert, W.E. (1974) 'Applications of duality theory.' In M.D. Intriligator and D.A. Kendrick, eds, *Frontiers of Quantitative Economics*, 2 (New York: North-Holland)
Epstein, L.G. (1978) 'Production flexibility and the behaviour of the competitive firm under price uncertainty.' *Review of Economic Studies* 45, 251–61
Epstein, L.G. and S.M. Tanny (1978) 'Increasing generalized correlation: a definition and some economic consequences.' Working Paper 7803, Institute for Policy Analysis, University of Toronto
Fischer, S. (1975) 'The demand for index bonds.' *Journal of Political Economy* 83, 509–32
Hadar, J. and W.R. Russell (1969) 'Rules for ordering uncertain prospects.' *American Economic Review* 59, 25–34
Hadar, J. and W.R. Russell (1971) 'Stochastic dominance and diversification.' *Journal of Economic Theory* 3, 288–305
Hadar, J. and W.R. Russell (1974) 'Diversification of interdependent prospects.' *Journal of Economic Theory* 7, 231–40
Halmos, P.R. (1961) *Measure Theory* (New York: Van Nostrand)
Hamada, K. (1974) 'Comments on "Stochastic dominance in choice under uncertainty."' In M.S. Balch, D.L. McFadden, and S.Y. Wu, eds, *Essays on Economic Behaviour Under Uncertainty* (New York: North-Holland)
Hanoch, G. and H. Levy (1969) 'The efficiency analysis of choices involving risk.' *Review of Economic Studies* 69, 335–46
Hardy, G.H., J.E. Littlewood, and G. Polya (1934) *Inequalities* (London: Cambridge University Press)
Hartman, R. (1976) 'Factor demand with output price uncertainty.' *American Economic Review* 66, 675–81
Hildreth, C. and L. Tesfatsion (1977) 'A note on dependence between a venture and a current prospect.' *Journal of Economic Theory* 15, 381–91
Hicks, J. (1946) *Value and Capital* (London: Oxford University Press)
Kihlstrom, R. and L. Mirman (1974) 'Risk aversion with many commodities.' *Journal of Economic Theory* 8, 361–88
Levy, H. and J. Paroush (1974) 'Toward multivariate efficiency criteria.' *Journal of Economic Theory* 7, 129–42
Rao, C.R. (1965) *Linear Statistical Inference and Its Applications* (New York: John Wiley)
Richard, S.F. (1975) 'Multivariate risk aversion, utility independence and separable utility functions.' *Management Science* 22, 12–21
Roll, R. (1973) 'Assets, money and commodity price inflation under uncertainty.' *Journal of Money, Credit, and Banking* 5, 903–23
Rothschild, M. and J. Stiglitz (1970) 'Increasing risk: I A definition.' *Journal of Economic Theory* 2, 225–43
Rothschild, M. and J. Stiglitz (1971) 'Increasing risk: II Its economic consequences.' *Journal of Economic Theory* 3, 66–84
Samuelson, P. (1967) 'General proof that diversification pays.' *Journal of Financial and Quantitative Analysis* 2, 1–13
Samuelson, P. (1970) 'The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments.' *Review of Economic Studies* 37, 537–42
Scheffman, D.T. (1975) 'A definition of generalized correlation and its application for portfolio analysis.' *Economic Inquiry* 13, 277–86