1973

The Diversification Problem in Portfolio Models

David T. Scheffman

20273

Research Report 7318

THE DIVERSIFICATION PROBLEM IN PORTFOLIO MODELS

by

D. T. Scheffman

The Diversification Problem in Portfolio Models

D. T. Scheffman*
University of Western Ontario

I. Introduction

There is an old adage which states that "the prudent investor should diversify his holdings". In this paper we attempt to analyze this proposition within the framework of a one-period optimal asset choice model. The introduction of mean-variance analysis to the economic theory of choice under uncertainty by Markowitz and Tobin provided the first formal economic model which was consistent with this adage. This model provided the framework needed for the important breakthrough in modern monetary and financial market theory, with the main impetus coming perhaps from Tobin's seminal article, "Liquidity Preference as Behavior Towards Risk". The two main contributions of mean-variance analysis for these breakthroughs were:

a) Mean-variance analysis provided a formal but relatively simple construct for analyzing choice under uncertainty.

b) The most important implication of the model was that an investor with a mean-variance criterion would generally diversify rather than specialize. This implication provided a congruence between theory and actual investors' behavior which only existed on an ad hoc basis prior to the introduction of mean-variance analysis. (It should perhaps be recalled by the reader that the most popular pre-Markowitz theory of choice under uncertainty postulated that investors were expected-value maximizers, and since this had very unrealistic implications for investor behavior, ad hoc explanatory variables such as risk and liquidity premiums were appended to the model.)

An alternative to the mean-variance criterion theory of choice under uncertainty was provided by the Von-Neumann-Morgenstern expected utility hypothesis. As we now know, the expected utility hypothesis is probably a more adequate model of choice under uncertainty than mean-variance analysis--i.e., it can be shown that except under very special assumptions, an investor with a mean-variance criterion can be shown to make totally implausible choices, and this type of implausible behavior would not occur for any expected-utility criterion investor. The expected-utility hypothesis model also provides a formal construct for analyzing choice under uncertainty, but it is considerably more complicated than the mean-variance model. Because of the difficulty in deriving propositions from the expected utility model, there has been much more research concerned with extending mean-variance analysis than with discovering meaningful implications of the expected-utility model. This would seem to be an unfortunate imbalance because of the inherent superiority of the expected utility hypothesis.

Many of the important questions which a theory of choice under uncertainty should answer have been thoroughly considered within the mean-variance model. In particular, the conditions under which a mean-variance criterion investor will diversify are fairly well known. However, very little analysis has been considered with discovering the appropriate diversification conditions for the expected utility model. Besides discovering and analyzing such conditions, our analysis will provide three contributions. First, we will analyze and extend a definition of correlation first behavior, ad hoc explanatory variables such as risk and liquidity premiums were appended to the model.)

An alternative to the mean-variance criterion theory of choice under uncertainty was provided by the Von-Neumann-Morgenstern expected utility hypothesis. As we now know, the expected utility hypothesis is probably a more adequate model of choice under uncertainty than mean-variance analysis--i.e., it can be shown that except under very special assumptions, an investor with a mean-variance criterion can be shown to make totally implausible choices, and this type of implausible behavior would not occur for any expected-utility criterion investor. The expected-utility hypothesis model also provides a formal construct for analyzing choice under uncertainty, but it is considerably more complicated than the mean-variance model. Because of the difficulty in deriving propositions from the expected utility model, there has been much more research concerned with extending mean-variance analysis than with discovering meaningful implications of the expected-utility model. This would seem to be an unfortunate imbalance because of the inherent superiority of the expected utility hypothesis.

Many of the important questions which a theory of choice under uncertainty should answer have been thoroughly considered within the mean-variance model. In particular, the conditions under which a mean-variance criterion investor will diversify are fairly well known. However, very little analysis has been considered with discovering the appropriate diversification conditions for the expected utility model. Besides discovering and analyzing such conditions, our analysis will provide three contributions. First, we will analyze and extend a definition of correlation first proposed by Samuelson,[1] which we will show provides a more useful definition of statistical dependence than linear (Pearsonian) correlation for the expected utility model. Our second contribution is derived from the apparent philosophy behind the continued research on the mean-variance model. Since, as we have indicated, the expected utility hypothesis is inherently superior as a theory of rational choice, seemingly the only defense for continued use of the mean-variance model is the argument that this model is in some sense a reasonable approximation to the expected utility model. One very indirect means of considering the validity of this argument is to compare the conditions necessary for establishing equivalent results in the two models. Our analysis will provide one such indirect test by allowing comparison of the diversification theorems of the two models. Of final interest in our exposition is our method of analysis, which is indicative of a general way of approaching problems of the type we consider.

II. The Setting of the Problem

The question: "under what conditions will an investor diversify?" is not meaningfully posed unless we consider specific preferences and probability distributions. Restricting ourselves to such specifics would yield results which would only be of very limited interest. Because one of the purposes of our analysis is to provide a comparison of results in the mean-variance and expected utility models, we will focus on the main diversification theorem of mean-variance analysis. This theorem gives conditions under which any risk-averse investor will hold positive amounts of each of the available assets.

III. The Mean-Variance Model

The mean-variance model is defined by the following assumptions. An investor must choose from among n securities, $i = 1,\ldots,n$ in planning his portfolio for one period. The one period yields of at least some of the assets are uncertain. p_i^t is the price he must pay for one unit of the i^{th} security at time t. p_i^{t+1} is the (possibly) unknown price of the security at time t + 1. It is assumed that the investor knows (or has a fixed estimate of) $E(p_i^{t+1}/p_i^t) = \mu_i$ and $E[(p_i^{t+1}/p_i^t - \mu_i)(p_j^{t+1}/p_j^t - \mu_j)] = \sigma_{ij}$. He has an initial wealth of W_0. His portfolio optimization problem then becomes:

a) $$\max_{\{X_i\}} U[E\{\sum_i p_i^{t+1} X_i\}, \text{var } \{\Sigma_i p_i^{t+1} X_i\}]$$

subject to $\sum_i p_i^t X_i = W_0$, where X_i is the number of units bought of the i^{th} security and U is a mean-variance utility function.

Let $U_1 = \partial U/\partial[E\{\sum_{i=1}^{n} P_i^{t+1} X_i\}]$ and $U_2 = \partial U/\partial[\text{var}\{\sum_{i=1}^{n} P_i^{t+1} X_i\}]$. The usual assumptions about U require $U_1 > 0$, $U_2 < 0$. (If $U_2 > 0$ we call the investor a risk lover.) Of course, $U_2 < 0$ seems consistent with the notion of a "prudent" investor.

The solution of problem (a) can be broken down into two parts: i) Derive the set of feasible efficient mean-variance combinations--i.e., the efficiency locus in mean-variance space. (A point on the efficiency locus represents the maximum possible mean for a given variance.) ii) Choose the preferred mean-variance combination from the efficiency locus.

Every point on the efficiency locus will be chosen by some risk-averse mean-variance utility function. Therefore a necessary condition for <u>all</u>

risk-averse investors to hold positive amounts of each asset is that all points on the efficiency locus must correspond to portfolios with positive amounts of each asset. It is well known that if there is no riskless asset and all mean yields are not equal then the efficiency locus (in mean-standard deviation space) is a half-parabola and some points on this locus correspond to non-diversified portfolios.[2] It is also well known that if there is a riskless asset then the efficiency locus is a straight line, where again some points correspond to non-diversified portfolios.[3] (We assume here that the mean of the riskless asset is smaller than at least one risky asset.) <u>Therefore a necessary condition for all efficient portfolios to be diversified is that all mean yields are equal and there is no riskless asset</u>.

We can now state the well-known diversification theorems of mean-variance analysis. Let Σ be the matrix of variances and covariances of the asset yields (i.e., $\Sigma = [\sigma_{ij}]$).

<u>Theorem I</u>. The necessary and sufficient conditions for any risk-averse mean-variance investor to hold positive amounts of each asset are: i) no asset is riskless; ii) all assets have the same mean yield; iii) Σ^{-1} has positive row sums.

The usual statement of the diversification theorem is the following Corollary:

<u>Corollary</u>. If: i) no asset is riskless; ii) all assets have the same mean yield; iii) $\sigma_{ij} \leqq 0$ for $i \neq j$
then any risk-averse mean-variance investor will hold positive amounts of each asset.

Since these results are well known, we will not prove them here (the author will furnish proofs on request). The assumptions of the Corollary are strongly over-sufficient in fulfilling the condition that Σ^{-1} have positive row sums. Consideration of the formula for Σ^{-1} for n=2 or 3 makes it clear that a considerable amount of positive correlation is consistent with Σ^{-1} having positive row sums. This is also indicated by various empirical studies of stock yields (e.g., Evans and Archer)[4] which show that while stock yields are positively correlated there are still significant reductions in variance for large portfolios possible through diversification.

IV. The Expected Utility Model

The expected utility model choice problem can be written:

a′) $$\max_{\{X_i\}} E\{U(\sum_i p_i^{t+1} X_i)\} \text{ subject to } \sum_i p_i^t X_i = W_0 .$$

By defining $Z_i = p_i^{t+1}/p_i^t$, $\lambda_i = \frac{p_i^t X_i}{W_0}$, this can be rewritten:

b′) $$\max_{\{\lambda_i\}} E\{U((\Sigma\lambda_i Z_i)W_0)\} \text{ subject to } \Sigma \lambda_i = 1 .$$

The usual assumptions about U are $U' > 0$, $U'' < 0$. (If $U'' > 0$, we call the investor a risk lover.) $U'' < 0$ is consistent with the notion of a "prudent" investor.

a) Comparison of the Expected Utility Model and the Mean-Variance Model

In the expected utility model, Theorem 1 is no longer valid as the following example shows. There are three states of nature possible, A, B and C, and each has a probability of 1/3 of occurrence. There are two assets, the random yields of which are denoted Z_1 and Z_2. The value of the asset yields in each of the states is given in the following table.

	A	B	C
Z_1	.5	1	1.5
Z_2	0	3	0

$$E\{Z_1\} = 1/3 \cdot .5 + 1/3 \cdot 1 + 1/3 \cdot 1.5 = 1$$

$$E\{Z_2\} = 1/3 \cdot 0 + 1/3 \cdot 3 + 1/3 \cdot 0 = 1$$

$$\sigma_{12} = 1/3(-1/2)(-1) + 0 + 1/3(1/2)(-1) = 0$$

Thus the asset yields have equal means and zero covariance. For a mean-variance criterion utility function, by Theorem 1 a positive amount of each would be held. Assume $W_0 = 1$. Let U be some utility function with $U' > 0$, $U'' < 0$.

$$E\{U(\lambda Z_1 + (1-\lambda) Z_2)\} = 1/3\ U(.5\lambda) + 1/3\ U(\lambda + 3(1-\lambda)) + 1/3\ U(1.5\lambda)$$

$$\frac{dE\{U\}}{d\lambda} = 1/6\ U'(.5\lambda) - 2/3\ U'(-2\lambda + 3) + 1/2\ U'(1.5\lambda) = F(\lambda).$$

The first order conditions for maximizing $E\{U\}$ require that $F(\lambda^*) = 0$, for λ^* the maximizer of $E\{U\}$.

$$F(0) = 1/6\ U'(0) - 2/3\ U'(3) + 1/2\ U'(0) = 2/3\ U'(0) - 2/3\ U'(3)$$

$$F(1) = 1/6\ U'(.5) - 2/3\ U'(1) + 1/2\ U'(1.5)$$

Since $U'' < 0$, we must have $F(0) > 0$, and $F'(\lambda) < 0$, for all λ. Therefore $\lambda^* > 0$. Furthermore it is clearly possible to find a concave U such that $F(1) > 0$ (e.g., for $U(x) = \log x$, $F(1) = 0$). For U such that $F(1) > 0$ the solution requires $\lambda^* > 1$--i.e., the investor "shorts" the second asset. Of course given the nature of the assets this is an entirely plausible solution. Thus Theorem 1 is not valid in the expected utility model.

It is not even true that if one prospect has a bigger mean and smaller variance than another prospect that the first is necessarily preferred to the second.[5]

As we mentioned in the introduction to this essay, the mean-variance criterion can be shown to lead to implausible choices. Consider the following example.

Let E = expected value of any prospect

V = variance of any prospect.

Consider the utility function

$U(E,V) = E/(1+V)$, for $E \geqq 0$

This utility function has the properties $U_1 > 0$, $U_2 < 0$.

Now consider the following two prospects:

P1: 10 with probability .99

101 with probability .01

P2: 10 with probability .98

101 with probability .02

$E_1 = 11$, $V_1 = 81.99$; $E_2 = 12$, $V_2 = 162.34$.

For this utility function, P1 is preferred to P2, but this ranking is completely implausible.

Of course what is going wrong in this example is that variance is not a good indication of risk. Clearly, in any sense meaningful for portfolio selection P2 is <u>less</u> risky than P1. Mean-variance disciples would argue that if, as in this case, the probability distributions considered are highly skewed, then semi-variance instead of variance should be used. This however does not circumvent the main difficulty, which is that any simple measure of risk is going to be inadequate. The reason for this is that only in special circumstances is the mean-variance criterion (or its common modifications) consistent with the expected utility axioms.[6] As is well known, these special circumstances are: either the investor's preferences can be represented by the expected value of a quadratic utility function or that the probability distributions be confined to Gaussian.

Although much of the recent literature makes one of these assumptions,[7] neither is plausible. The quadratic utility function: $U(x) = x - bx^2$ exhibits wealth satiation and increasing absolute risk aversion.[8] The assumption of Gaussian probability is not logically consistent with the model since because prices cannot be negative, the Z_i's cannot be negative. Thus the question arises as to how we are to interpret the results of these various mean-variance models if we reject the mean-variance criterion. Some of the literature seems to argue that these mean-variance models are in some sense approximations to an expected utility model. Most comparisons of the results derivable from the two models would seem to indicate that the approximation is not very good. A further indication of this will be a comparison of the diversification theorems for the expected utility model with Theorem I of the first section.

There has been a small amount of literature concerned with the diversification problem in the expected utility model. Hadar and Russell[9] enlarged the scope of the "stochastic dominance" literature by considering stochastic dominance for multivariate distributions. However, they concentrated on combinations of <u>independent</u> random variables, which is only of very limited interest for the diversification problem. In an earlier paper Samuelson[10] attempted to go beyond independence in his search for diversification theorems, but met with only partial success. Our analysis will be, in large part, a development of ideas Samuelson advanced in that paper.

b) A Definition of Nonlinear Correlation and Its Properties

As we saw in the earlier example, in our search for a theorem analogous to Theorem I in the expected utility model, non-positive (Pearsonian) correlation between the asset yields will not be a strong enough assumption. Samuelson[11] showed that if the yields were distributed independently with equal means then a risk-averting expected utility maximizer would hold a positive amount of each asset. He then argued that since going from zero to negative correlation in the mean-variance model makes everything even better, going from independence to some stronger type of negative correlation would have the same result in the expected utility model. He proposed a stronger form of negative correlation which we will discuss next. Although he stated this stronger form of negative correlation for the general n-asset problem, this is not a strong enough condition except in the two asset case. Thus the following development will concentrate on the two asset case. We will then see that based on the two asset theorems we can derive, a stronger condition than Samuelson proposed will give us some n-asset results. Let Z_1 and Z_2 be random variables with joint density function denoted $dP(z_1, z_2)$ and joint cumulative distribution function denoted $P(z_1, z_2)$. Let $P(z_1|z_2) = \text{Prob}\,\{Z_1 \leqq z_1 | Z_2 = z_2\}$.

Definition: Z_1 will be said to be negatively S-correlated with Z_2 if $\partial P(z_1|z_2)/\partial z_2 \geqq 0$, for all (z_1, z_2). Z_1 will be said to be positively S-correlated with Z_2 if $\partial P(z_1|z_2)/\partial z_1 \leqq 0$, for all (z_1, z_2).

Note: The differentiability here is not necessary,--the definition applies also for finite differences between z_2 points with positive density.

For example, consider the following density functions for the random variables Z_1 and Z_2:

i) There are three states of nature, A, B and C, each with a probability of occurrence of 1/3.

	A	B	C
Z_1	1	3	5
Z_2	7	4	0

$P(z_1|z_2)$: $P(1|0) = 0$, $P(1|4) = 0$, $P(1|7) = 1$

$P(3|0) = 0$, $P(3|4) = 1$, $P(3|7) = 1$

$P(5|0) = P(5|4) = P(5|7) = 1$

Therefore Z_1 is negatively S-correlated with Z_2.

ii) $dP(z_1, z_2) = z_1 + z_2$, $0 \leqq z_i \leqq 1$.

$$P(z_1|z_2) = \int_0^{z_1} dP(z_1|z_2) = \left[\int_0^{z_1} (Z_1 + z_2)dZ_1\right] \div \left[\int_0^{1} (Z_1 + z_2)dZ_1\right]$$

$$= (z_1^2/2 + z_1z_2)/(1/2 + z_2) .$$

$$\frac{\partial P(z_1|z_2)}{\partial Z_2} = \frac{(1/2+z_2) \cdot z_1 - (z_1^2/2+z_1z_2)}{(1/2 + z_2)^2} = \frac{1/2 - z_1^2/2}{(1/2 + z_2)^2}$$

which is $\geqq 0$ because $0 \leqq z_1 \leqq 1$.

Therefore Z_1 is negatively S-correlated with Z_2.

iii) $dP(z_1, z_2) = z_1 + 1/z_2$, $0 \leqq z_1 \leqq 1$, $1 \leqq z_2 \leqq b$

where b is such that $(b - 1)/2 + \log b = 1$.

$$P(z_1|z_2) = [{}_0\int^{z_1} (Z_1 + 1/z_2)dZ_1] \div [{}_0\int^{1} (Z_1 + 1/z_2)dZ_1]$$

$$= (z_1^2/2 + z_1/z_2)/(1/2 + 1/z_2) .$$

$$\frac{\partial P(z_1|z_2)}{\partial Z_2} = \frac{(1/2+1/z_2) \cdot (-z_1/z_2^2) - (z_1^2/2+z_1/z_2)(-1/z_2^2)}{(1/2 + 1/z_2)^2}$$

$$= \frac{1/2(-z_1+z_1^2)}{z_2^2(1/2+1/z_2)^2}$$

which is $\leqq 0$ because $0 \leqq z_1 \leqq 1$.

Therefore Z_1 is positively S-correlated with Z_2 .

iv) If Z_1 and Z_2 are distributed as bivariate normal then Z_1 is negatively S-correlated with Z_2 if $\rho \leqq 0$ and Z_1 is positively S-correlated with Z_2 if $\rho \geqq 0$. To see this, recall for the bivariate normal

$$P(z_1|z_2) = {}_{-\infty}\int^{z_1} k \exp [\frac{-(Z_1 - b)^2}{2\sigma_1^2(1-\rho^2)}]dZ_1$$

where $b = \mu_1 + \rho(\sigma_1/\sigma_2)[z_2 - \mu_2]$.

Therefore $\partial P(z_1|z_2)/\partial z_2 = \frac{2k(\rho\sigma_1/\sigma_2)}{2\sigma_1^2(1-\rho^2)} {}_{-\infty}\int^{z_1} (Z_1 - b) \exp [\quad]dZ_1$

and the sign of this expression = - sign ρ because

${}_{-\infty}\int^{z_1} (Z_1 - b) \exp [\quad]dZ_1 < 0$ for all $z_1 < \infty$ because $E\{Z_1|z_2\} = b$.

The definition of S-correlation is not a symmetric one--i.e., Z_1 may be negatively S-correlated with Z_2 but Z_2 may not be negatively S-correlated with Z_1. (As we shall see in Lemma 5 it is not possible for Z_1 to be strictly negatively S-correlated with Z_1.) The following example shows that S-correlation is not a symmetric property.

There are two states of nature, A and B, with prob $\{A\} = 9/10$, prob $\{B\} = 1/10$. The density function for random variables Z_1 and Z_2 is:

State A	State B
$Z_1 = 5$, with probability = 1	$Z_1 = 10$, with probability = 1
$Z_2 = \begin{cases} 3, \text{ with probability} = 1/2 \\ 4, \text{ with probability} = 1/2 \end{cases}$	$Z_2 = \begin{cases} 1, \text{ with probability} = 1/2 \\ 3, \text{ with probability} = 1/6 \\ 4, \text{ with probability} = 1/3 \end{cases}$

From this density function we get the cumulative conditional distribution functions:

$P(Z_2 \leqq 1 | Z_1 = 5) = 0$

$P(Z_2 \leqq 1 | Z_1 = 10) = 1/2$

$P(Z_2 \leqq 3 | Z_1 = 5) = 1/2$

$P(Z_2 \leqq 3 | Z_1 = 10) = 2/3$

$P(Z_2 \leqq 4 | \quad\quad) = 1$

$P(Z_1 \leqq 5 | z_2 = 1) = 0$

$P(Z_1 \leqq 5 | z_2 = 3) = 27/28$

$P(Z_1 \leqq 5 | z_2 = 4) = 27/29$

We see from the definition (using first differences between points of positive density) that Z_2 is negatively S-correlated with Z_1, but since $P(Z_1 \leqq 5 | Z_2 = 4) - P(Z_1 \leqq 5 | Z_2 = 3) < 0$, Z_1 is not negatively S-correlated with Z_2.

Z_2 negatively S-correlated with Z_1 means that for $\epsilon > 0$, $P(Z_2 \leqq a | Z_1 = b + \epsilon) - P(Z_2 \leqq a | Z_1 = b) > 0$ (for b, ϵ such that b and b + ϵ are points of positive density for Z_1). This means that an increase in the given value of Z_2 shifts the conditional cumulative distribution $P(z_2|z_1)$ to the left. Thus negative S-correlation is plausible as a stronger version of negative correlation. The shifting of the conditional cumulative distribution function is depicted in the following diagram. To determine the properties of negative S-correlation we need the following mathematical results.

The following result is obvious.

Lemma 1: Let $g(y)$ and $h(y)$ be continuous functions which are Riemann-Stieltjes integrable with respect to P where $dP(y)$ is a probability density function.

a) If $g(y)$ and $h(y)$ are both monotone non-increasing or are both monotone non-decreasing, $\text{cov}(g,h) \geqq 0$.

b) If $g(y)$ is monotone non-increasing and $h(y)$ is monotone non-decreasing then $\text{cov}(g,h) \leqq 0$.

Of course if g and h are not constant at all points of positive density the inequalities in a) and b) are strict inequalities.

The following result is from the paper by Hanoch and Levy.[12]

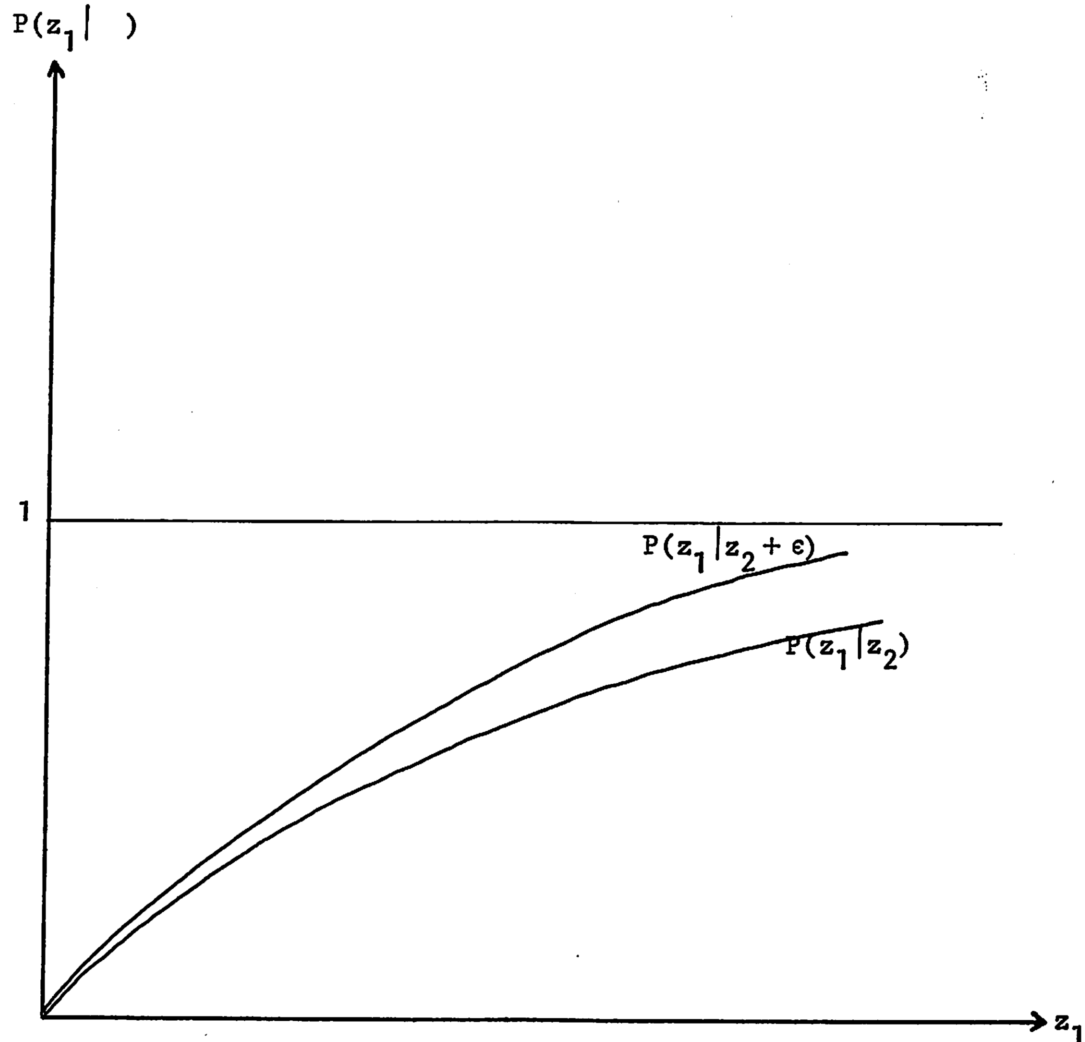
Lemma 2: Let $F(x)$ and $G(x)$ be cumulative distribution functions and let $h(x)$ be a function which is Riemann-Stieltjes integrable with respect to F and G. Let $E_F\{h\} = \int h(x)dF(x)$, etc.

If $F(x) \leqq G(x)$, for all x and (<) holds for some x_0 then:

$E_F\{h(x)\} > (<) E_G\{h(x)\}$ if $h' > 0 (h' < 0)$.

(Again--the differentiability is not necessary here. The basic strategy of the proof is to integrate $H(x)[dF(x) - dG(x)]$ by parts.)

* For the rest of this essay we will assume $\underline{Z_i \geq 0}$, which is a reasonable assumption since $Z_i = p_i^{t+1}/p_i^t$.


P(z1|)
1
P(z1|z2+ε)
P(z1|z2)
z1

We can now prove the following results.

Let $E\{Z_1|z_2\} = \int z_1 dP(z_1|z_2)$ -- i.e., the conditional expectation of Z_1 given $Z_2 = z_2$.

Lemma 3: If $\partial P(z_1|z_2)/\partial z_2 \geqq 0$, for all (z_1, z_2), then $\partial E\{Z_1|z_2\}/\partial z_2 \leqq 0$. (The differentiability is not necessary here.)

Proof: $E\{Z_1|z_2\} = \int z_1 dP(z_1|z_2)$.
Let $\epsilon > 0$.
$E\{Z_1|z_2 + \epsilon\} - E\{Z_1|z_2\} = \int z_1 dP(z_1|z_2 + \epsilon) - \int z_1 dP(z_1|z_2)$.

By the assumptions of the Lemma,
$P(z_1|z_2) \leqq P(z_1|z_2 + \epsilon)$ for $\epsilon > 0$.
Furthermore, the function $U(z_1) = z_1$ is monotone-increasing.

Therefore by Lemma 2
$\int z_1 dP(z_1|z_2) \geqq \int z_1 dP(z_1|z_2 + \epsilon)$,
so that $E\{Z_1|z_2 + \epsilon\} - E\{Z_1|z_2\} \leqq 0$ for $\epsilon > 0$
and the result is proved. (Notice if $\partial P(z_1|z_2)/\partial z_2 > 0$, then $\partial E\{Z_1|z_2\}/\partial z_2 < 0$.)

Corollary: If $\partial P(z_1|z_2)/\partial z_2 \leqq 0$ for all (z_1, z_2), then $\partial E\{Z_1|z_2\}/\partial z_2 \geqq 0$.

Lemma 5: If Z_1 is negatively S-correlated with Z_2 (or vice versa) then $\sigma_{12} \leqq 0$.

Proof: $\sigma_{12} = \int\int z_1 z_2 dP(z_1, z_2) - E\{Z_1\} E\{Z_2\}$

$$\int\int z_1 z_2 dP(z_1, z_2) = \int z_1 E\{Z_2|z_1\} dP(z_1)$$

By Lemma 4, $E\{Z_2|z_1\}$ is a monotone non-increasing function of z_1. Therefore, by Lemma 1, $\text{cov}(Z_1, E\{Z_2|z_1\}) \leqq 0$. But of course $\text{cov}(Z_1, E\{Z_2|z_1\}) = \text{cov}(Z_1, Z_2)$, and the result is proved. (Notice if $\partial P(z_1|z_2) > 0$, then $\sigma_{12} < 0$.)

<u>Corollary</u>: If Z_1 is positively S-correlated with Z_2 (or vice versa) then $\sigma_{12} \geqq 0$.

Consider the following example:

There are three states of nature, A, B, and C, each with probability = 1/3. The values of random variables Z_1 and Z_2 in each state are given in the following table:

	A	B	C
Z_1	1	5	0
Z_2	1	2	3

$\sigma_{12} = -1/3$

$P(z_1|z_2)$: $P(1|1) = 1$

$P(1|2) = 0$

$P(1|3) = 1$

Therefore Z_1 is not negatively S-correlated with Z_2, and so negative S-correlation is a <u>stronger</u> condition than negative (Pearsonian) correlation.

We state without proof the following obvious result:

<u>Lemma 5</u>: If $\partial P(z_1|z_2)/\partial z_2 = 0$, $\forall(z_1, z_2)$ then Z_1 and Z_2 are independent random variables.

c) Some Diversification Results for the Two-Asset Case

The previous examples, lemmas and comments give us a fairly clear picture of the properties of S-correlation. We will now show that this definition of correlation is of interest because of the diversification theorems we can prove using it. In the following theorem we will assume that there are only two assets available.

Theorem II: If $E\{Z_1\} = E\{Z_2\}$, neither asset is riskless, and Z_2 is negatively S-correlated with Z_1, then any risk-averse ($U'' < 0$) expected utility criterion investor will hold positive amounts of each asset.

Proof: The investor's maximization problem becomes:

$\max_{\{\lambda\}} E\{U[(\lambda Z_1 + (1-\lambda)Z_2)W_o]\}$. Without loss of generality we will let $W_o = 1$.

Therefore we must show that the solution to the problem:

$\max_{\{\lambda\}} E\{U(\lambda Z_1 + (1-\lambda)Z_1)\}$ is such that $0 < \lambda^* < 1$.

Suppose $\lambda^* \geqq 1$. The First Order Conditions (F.O.C.) for the problem require:

$$1) \quad \iint z_1U'(\lambda^*z_1 + (1-\lambda^*)z_2)dP(z_1,z_2) = \int\int z_2U'(\lambda^*z_1 + (1-\lambda^*)\ z_2)\ dP(z_1,\ z_2)$$

Iterating integrals we have:

$$2) \quad \int\int z_2U'(\lambda^*z_1 + (1-\lambda^*)\ z_2)\ dP(z_1,\ z_2)$$

$$= \int [\int z_2U'(\lambda^*z_1 + (1-\lambda^*)\ z_2)\ dP(z_2|z_1)]\ dP(z_1)$$

where $dP(z_1)$ is the marginal density of Z_1.

$\partial U'/\partial z_2 = (1-\lambda^*)\ U'' \geqq 0$ since $\lambda^* \geqq 1$. Thus U' is a monotone non-decreasing function of z_2.

Let $\text{cov}_{z_1}(Z_2, U')$ be the covariance of Z_2 and U' relative to the density function $dP(z_2|z_1)$. Then by a) of Lemma 1 we have $\text{cov}_{z_1}(Z_2, U') \geqq 0$ (z_2 is a monotone increasing and U' is a monotone non-decreasing function of z_2) so,

3) $\int z_2 U' dP(z_2|z_1) \geqq E\{Z_2|z_1\} \cdot E\{U'|z_1\}$ for <u>all</u> z_1.

Then from 2):

4) $\int\int z_2 U' dP(z_1, z_2) \geqq \int E\{Z_2|z_1\}\ E\{U'|z_1\}\ dP(z_1)$.

5) By Lemma 3, $\partial E\{Z_2|z_1\}/\partial z_1 \leqq 0$.

6) $\partial/\partial z_1 [E\{U'|z_1\}] = E\{\lambda^* U''|z_1\} + \int U'\ \partial/\partial z_1 [dP(z_2|z_1)]$

$(E\{U'|z_1\} = \int U' dP(z_2|z_1))$.

Since $\lambda^* \geqq 1$ and $U'' < 0$ by assumption, we have $E\{\lambda^* U''|z_1\} < 0$.

Now we must determine the sign of $\int U'\ \partial/\partial Z_1 [dP(z_2|z_1)]$.

For $\epsilon > 0$, consider

$\int U'(\lambda^* z_1 + (1-\lambda^*)\ z_2)\ dP(z_2|z_1 + \epsilon) - \int U'(\lambda^* z_1 + (1-\lambda^*)\ z_2)\ dP(z_2|z_1)$

$\partial/\partial z_2 [U'(\lambda^* z_1 + (1-\lambda^*)\ z_2)] = (1-\lambda^*) U'' \geqq 0$.

Therefore by Lemma 2

$\int U' dP(z_2|z_1 + \epsilon) - \int U' dP(z_2|z_1) \leqq 0$,

since $\partial P(z_2|z_1)/\partial z_1 \geqq 0$. (In Lemma 2, let $F = P(z_2|z_1 + \epsilon)$, $G = P(z_2|z_1)$, $h = U'$, and $h' \geqq 0$.)

7) Therefore $\partial/\partial z_1\ E\{U'|z_1\} < 0$.

From 5), 7) and a) of Lemma 1 we have

8) $\operatorname{cov}(E\{Z_2|z_1\}, E\{U'|z_1\}) \geqq 0.$

Therefore we can rewrite 4):

9) $\int\int z_2 U' dP(z_1, z_2) \geqq E\{Z_2\} E\{U'\}.$

Going back to 1), by iterating integrals we have

10) $\int\int z_1 U' dP(z_1, z_2) = \int z_1 E\{U'|z_1\} dP(z_1),$

giving us from 9) and 2):

11) $\int z_1 E\{U'|z_1\} dP(z_1) \geqq E\{Z_2\} E\{U'\}$

But as we have already shown, $\partial E\{U'|z_1\}/\partial z_1 < 0$ so by b) of Lemma 1

12) $\operatorname{cov}(Z_1, E\{U'|z_1\}) < 0$ which means:

13) $\int z_1 E\{U'|z_1\} dP(z_1) < E\{Z_1\} E\{U'\}.$

But 13) is a contradiction of 11) since by assumption, $E\{Z_1\} = E\{Z_2\}$. Thus we have shown $\lambda^* < 1$. If we repeat the proof by assuming that $\lambda^* < 0$ we can again arrive at a contradiction in an analogous manner. Therefore $0 < \lambda^* < 1$. Notice that although we have demonstrated that the definition of negative S-correlation is not necessarily symmetric, we need only assume it for one direction.

Theorem II is of course of limited interest because of the two-asset framework. However, Theorem II does allow comparison with the Corollary to Theorem I, and we see the assumptions of Theorem II (negative S-correlation) are much stronger than the assumptions of the Corollary (non-positive linear correlation).

Theorem II generates our first n-asset result in the form of the following Corollary:

Corollary: Let $(\lambda_1^*, \ldots, \lambda_{n-1}^*)$ be the maximizers of the problem:

a) $\max\limits_{\{\lambda_i\}} E\{U(\sum_{i=1}^{n-1} \lambda_i Z_i)\}$ subject to $\sum_{i=1}^{n-1} \lambda_i = 1$, where $U' > 0$, $U'' < 0$.

Suppose $\sum_{i=1}^{n-1} \lambda_i^* Z_i$ is negatively S-correlated with Z_n and $E\{\sum_{i=1}^{n-1} \lambda_i^* Z_i\} \leq E\{Z_n\}$.

Then the solution of the problem:

b) $\max\limits_{\{\lambda_i\}} E\{U(\sum_{i=1}^{n} \lambda_i Z_i)\}$ subject to $\sum_{i=1}^{n} \lambda_i = 1$ has the property that

$\lambda_n \neq 0$. [It can be seen that this Corollary is a stronger version of Samuelson's Theorem III.[13]]

Proof: Let $\sum_{i=1}^{n-1} \lambda_i^* Z_i = X$, and $Z_n = Y$.

Then by the assumptions of the Corollary and by Theorem II we have that the solution of the problem $\max\limits_{\{\lambda\}} E\{U(\lambda X + (1-\lambda)Y)\}$ has the property that $\lambda < 1$.

Therefore it cannot be optimal to have $\lambda_n = 0$ since the solution of 1) is admissible as a solution of b).

This Corollary is of some practical interest since it gives conditions under which an investor can tell that his portfolio is inefficient. Thus if an investor's portfolio is negatively S-correlated with an asset which he does not hold, and if this asset has a mean yield at least as large as the portfolio, the portfolio is inefficient.

d) Some Non-Diversification Results

As in the case of the Corollary to Theorem I, the assumption of negative S-correlation is strongly over-sufficient. This can be seen in the next material, where we derive a non-diversification theorem. This theorem has already been found to be useful in the theory of warrant pricing.[14] We will see that we need a much stronger condition than positive S-correlation to guarantee that an investor will "short" one of the assets in a given situation. This analysis requires the following strengthening of Lemma 1, which the author has proved elsewhere.[15]

Lemma 6: Let $\Psi(y)$, $\phi(y)$ and $v(y)$ be functions which are Riemann-Stieltjes integrable with respect to P, where $dP(y)$ is a probability density function defined on $[0, \infty]$, and $v(y) \geqq 0$ for $y \geqq 0$.

Suppose:

a) $\exists\ \bar{y} > 0$ such that $\Psi(y) \leqq \phi(y)$ for all $0 \leqq y \leqq \bar{y}$ and $\Psi(y) \geqq \phi(y)$ for all $y \geqq \bar{y}$, and these inequalities are strict inequalities at some points of positive density.

b) $v(y)$ is a monotone non-increasing (non-decreasing) function.

c) $\int \phi(y)\, dP(y) \geqq (\leqq) \int \Psi(y)\, v(y)\, dP(y)$

Then $\int \phi(y)\, v(y)\, dP(y) \geqq (\leqq) \int \Psi(y)\, v(y)\, dP(y)$

and if $v(y)$ is a strictly decreasing (strictly increasing) function this is a strict inequality.

By interpreting Ψ, ϕ and v correctly, the corollary gives us a proof of Lemma 1 (e.g., let $\phi = g$, $\Psi = \bar{g}$, and $v = f$, and we have a proof for a) of Lemma 1 for the case f and g non-increasing). We will have to make use of Lemma 6 in our proof of Theorem III.

<u>Definition</u>: Z_1 will be said to be <u>strongly</u> positively S-correlated with Z_2 if Z_1 is positively S-correlated with Z_2 and $E\{Z_1|z_2\}$ is a strictly convex function of z_2. (By Lemma 3, $E\{Z_1|z_2\}$ is a monotone-increasing function of z_2.)

<u>Example</u>: Let $Z_1 \sim U[0, 1]$ and let $Z_2 = 3/2\ Z_1^2$.

Then $P\{z_2|z_1\} = \begin{cases} 1, \text{ if } z_2 \leqq (3/2)z_1^2, \ 0 \leqq z_1 \leqq 1 \\ 0, \text{ otherwise} \end{cases}$

and $E\{Z_2|z_1\} = (3/2)z_1^2$. Clearly, Z_2 is positively S-correlated with Z_1 and $E\{Z_2|z_1\}$ is a strictly convex function of z_1. Therefore Z_2 is strongly positively S-correlated with Z_1.

The theory of warrant pricing presented in the paper by Samuelson and Merton[16] has as one result that the yield on a warrant is strongly positively S-correlated with its associated stock.

Theorem III: If $U' > 0$, $U'' < 0$, $E\{Z_1\} = E\{Z_2\}$ and if Z_1 is <u>strongly</u> positively S-correlated with Z_2, then the solution of the problem

$$\max_{\{\lambda\}} E\{U(\lambda Z_1 + (1 - \lambda)Z_2)\}$$

has the property that $\lambda^* \notin [0, 1]$.

<u>Proof</u>: Suppose $0 \leqq \lambda^* \leqq 1$.

The First Order Conditions for the problem require:

1) $\iint z_1 U'(\lambda^* z_1 + (1 - \lambda^*)z_2)\, dP(z_1, z_2) = \iint z_2 U'(\lambda^* z_1 + (1 - \lambda^*)z_2)\, dP(z_1, z_2).$

Iterating integrals we have:

2) $\iint z_1 U'(\lambda^* z_1 + (1 - \lambda^*)z_2)\, dP(z_1, z_2)$

$$= \int [\int z_1 U'(\lambda^* z_1 + (1 - \lambda^*)z_2)\, dP(z_1|z_2)]\, dP(z_2).$$

$\partial U'/\partial z_1 = \lambda^* U'' \leqq 0$ since $\lambda^* \geqq 0$.

Let $\text{cov}_{z_2}(Z_1, U')$ be the covariance of Z_1 and U' relative to the density function $dP(z_1|z_2)$. Then by b) of Lemma 1 we have $\text{cov}_{z_2}(Z_1, U') \leqq 0$, so

3) $\int z_1 U'(\lambda^* z_1 + (1 - \lambda^*)z_2)\, dP(z_1|z_2) \leqq E\{Z_1|z_2\}E\{U'|z_2\}$, for <u>all</u> z_2.

Iterating the integrals in 1) and substituting 3) into 1) we have:

4) $\int E\{Z_1|z_2\}E\{U'|z_2\}\, dP(z_2) \geqq \int z_2 E\{U'|z_2\}\, dP(z_2).$

Applying the notation of Lemma 6, let $v(z_2) = E\{U'|z_2\}$,

$\Psi(z_2) = E\{Z_1|z_2\}$, $\Phi(z_2) = z_2$, and $dP = dP(z_2)$. By assumption, $\Psi(z_2)$ is strictly convex and since $\int \Psi(z_2)\, dP(z_2) = E\{Z_1\} = \int \Phi(z_2) dP(z_2) = E\{Z_2\}$, $\exists\ \bar{z}_2$ such that $\Psi(z_2) < \Phi(z_2)$ for $z_2 < \bar{z}_2$ and $\Psi(z_2) > \Phi(z_2)$ for $z_2 > z_2$.

5) $\partial/\partial z_2[E\{U'|z_2\}] = E\{(1 - \lambda^*)U''\ |z_2\} + \int U'\ \partial/\partial z_2[dP(z_1|z_2)].$

Since $(1 - \lambda^*) \geqq 0$ and $U'' < 0$, $E\{(1 - \lambda^*)U''\} \leqq 0$. Furthermore, by Lemma 2, since Z_1 is positively S-correlated with Z_2 and $\lambda^* \geqq 0$, $\int U'\ \partial/\partial z_2[dP(z_1|z_2)] \leqq 0$. To see this, consider

$\int U'\ dP(z_1|z_2 + \epsilon) - \int u''\ dP(z_1|z_2)$, for $\epsilon > 0$.

Since Z_1 is positively S-correlated with Z_2, $P(z_1|z_2 + \epsilon) \leqq P(z_1|z_2)$, for all (z_1, z_2). Also, $\partial/\partial z_1(U') = \lambda^* U'' \leqq 0$. Therefore, by Lemma 2

$\int U' dP(z_1|z_2 + \epsilon) - \int U' dP(z_1|z_2) \leqq 0.$

6) Thus, $\partial/\partial z_2[E\{U'|z_2\}] \leqq 0$, so $v(z_2)$ is non-increasing.

Since $\int \Psi(z_2)dP(z_2) = E\{Z_1\} = E\{Z_2\} = \int \Phi(z_2)dP(z_2)$ by Lemma 6.

7) $\int E\{Z_1|z_2\}E\{U'|z_2\}dP(z_2) \leqq \int z_2 E\{U'|z_2\}dP(z_2).$

Since it must be the case that either $\lambda^* > 0$ or $(1 - \lambda^*) > 0$, if $\lambda^* > 0$ then 4) holds with strict inequality which is a contradiction of 7), or if $(1 - \lambda^*) > 0$, then 7) holds with strict inequality which is a contradiction of 4).

Therefore $\lambda^* \notin [0, 1]$.

As mentioned earlier, the theory of warrant pricing developed by Samuelson and Merton[17] has as a result that a warrant must be strongly positively S-correlated with its associated stock. Thus, in a world where investors have "similar" subjective probability distributions, the warrant must have a higher expected yield than its associated stock; otherwise both could not exist.

<u>Corollary</u>: Let $(\lambda_1^*, \ldots, \lambda_{n-1}^*)$ be the maximizers of the problem:

a) $\max_{\{\lambda\}} E\{U(\sum_{i=1}^{n-1} \lambda_i Z_i)\}$ subject to $\sum_{i=1}^{n-1} \lambda_i = 1$, where $U' > 0$, $U'' < 0$.

Suppose $\sum_{i=1}^{n-1} \lambda_i^* Z_i$ is <u>strongly</u> positively S-correlated with Z_n and $E\{Z_n\} \geqq E\{\sum_{i=1}^{n-1} \lambda_i^* Z_i^*\}$. Then the solution of the problem:

b) $\max_{\{\lambda\}} E\{U(\sum_{i=1}^{n} \lambda_i Z_i)\}$ subject to $\sum_{i=1}^{n} \lambda_i = 1$ has the property that $\lambda_n \neq 0$.

Proof: Let $\sum_{i=1}^{n-1} \lambda_i^* Z_i = X$ and $Z_n = Y$.

Then by the assumptions of the Corollary and by Theorem 3 we have that the solution of the problem

1) $\max_{\{\lambda\}} E\{U(\lambda X + (1 - \lambda) Y)\}$ has the property that $\lambda \notin [0, 1]$. Therefore it cannot be optimal to have $\lambda_n = 0$ since the solution of 1) is an admissible solution of b).

e) Some n-Asset Results

Of course Theorems II and III are of very limited interest since they apply only to the two-asset case. The Corollaries to these theorems are much more interesting since they give conditions under which in some circumstances a risk averter can tell that his current portfolio is non-optimal (i.e., if his portfolio doesn't contain an asset which satisfies the assumptions of one of the Corollaries relative to his current portfolio).

We are now ready to develop an n-asset theorem. The proofs of the two-asset theorems are suggestive of the type of assumptions we will need for the n-asset case in that the obvious type of assumptions to make are those which will get us back to a two-asset framework.

Definition: Consider a family of random variables $(Z_1, \ldots, Z_n)$ with joint density function $dP(z_1,\ldots,z_n)$. Let

$\Lambda = \{\lambda | \lambda = (\lambda_1,\ldots,\lambda_n)$ where $\underline{\lambda_i \geq 0}$ and $\sum_{i=1}^{n} \lambda_i = 1\}$.

Let $Y(\lambda)$ be the random variable generated by taking a linear combination of the Z_i's according to the vector λ, i.e., $Y(\lambda) = \sum_{i=1}^{n} \lambda_i Z_i$.

We will say that the <u>family</u> of random variables $(Z_1,\ldots,Z_n)$ has the <u>negative S-correlation property</u> if for <u>any</u> j, and for any $\lambda \epsilon \Lambda$ with $\lambda_j = 0$, $Y(\lambda)$ is negatively S-correlated with Z_j. (In the case of n=2, this definition is equivalent to our definition of negative S-correlation in the two variable case).

This is obviously a very strong definition, but there are some density functions which satisfy the definition as the following example shows.

<u>Example</u>: $dP(z_1, z_2, z_3) = (z_1 + z_2 + z_3)dz_1\, dz_2\, dz_3$ for $0 \leqq z_1 \leqq (2/3)^{\frac{1}{4}}$. Let $(2/3)^{\frac{1}{4}} = b$. One can easily show that

$$P\{\lambda Z_1 + (1-\lambda)Z_2 \leqq k | z_3\} = \frac{k^3\left[\frac{1}{6\lambda^2(1-\lambda)^2}\right] + z_3 k^2\left[\frac{1}{2\lambda(1-\lambda)}\right]}{b^3\left[\frac{1}{6\lambda^2(1-\lambda)^2}\right] + z_3 b^2\left[\frac{1}{2\lambda(1-\lambda)}\right]}$$

Using this formula it is easily shown that $\partial P\{\lambda Z_1 + (1-\lambda)Z_2 \leqq k | z_3\}/\partial z_3 \geqq 0$ (using the fact that $z_3 \leqq b$). It is also clear that if $(Z_1,\ldots,Z_n)$ are independent or are joint Gaussian distributed with $\sigma_{ij} \leqq 0$ for $i \neq j$, they have this property.

From the previous definition we can now prove an n-asset theorem. However, this theorem is different in character from the previous theorems in that the assumptions of the theorem <u>do not allow shorting</u>. This is not a totally objectionable assumption, however, since institutional arrangements make going short considerably more difficult than going "long" in an asset. We again assume that no assets are riskless.

Theorem IV: If $U' > 0$, $U'' < 0$, $E\{Z_i\} = E\{Z_j\}$ for all i, j=1,..., n and the family of random variables $(Z_1, \ldots, Z_n)$ has the negative S-correlation property then the problem:

$\max_{\{\lambda_i\}} E\{U(\sum_{i=1}^{n} \lambda_i Z_i)\}$ subject to $\sum_{i=1}^{n} \lambda_i = 1$ and $\lambda_i \geqq 0$ has the property that $\lambda_i^* > 0$ for all i.

Proof: Suppose that the solution vector $(\lambda_1^*, \ldots, \lambda_n^*)$ is such that $\lambda_1^* = \ldots = \lambda_{k-1}^* = 0$; $\lambda_k^*, \ldots, \lambda_n^* > 0$ and of course $\sum_{i=k}^{n} \lambda_i^* = 1$.

Let $Z^* = \sum_{i=k}^{n} \lambda_i^* Z_i$. Since, by assumption, $\lambda_2^*, \ldots, \lambda_{k-1}^* = 0$, the solution

of the problem $\max_{\{\lambda\}} E\{U(\lambda Z_1 + (1 - \lambda)Z^*)\}$ subject to $\lambda \geqq 0$ will have the property that $\lambda^* = 0$. But Z^* is, by the assumptions of the theorem, negatively S-correlated with Z_1 and since $\sum_{i=k}^{n} \lambda_i^* = 1$ and $E\{Z_i\} = E\{Z_j\}$ for all i, j, $E\{Z^*\} = E\{Z_1\}$.

Therefore we are back in the situation of Theorem II, slightly modified by the constraint: $\lambda \geqq 0$. However, this only modifies step 1) of the proof of Theorem II, where now we must invoke the Kuhn-Tucker F.O.C.'s:[18]

$$1') \quad \iint z_1 U'(\lambda z_1 + (1 - \lambda)z^*)dP(z^*, z_1) \leqq \iint z^* U'(\lambda z_1 + (1 - \lambda)z^*)dP(z^*, z_1).$$

It is easily seen that the rest of the steps in the proof of Theorem II follow exactly. Therefore, $\lambda_1^* > 0$, and the theorem is proved.

By using an even stronger definition than the previous one, we can prove a stronger theorem. Using the notation of the previous definition:

<u>Definition</u>: We will say that the <u>family</u> of random variables $(Z_1, \ldots, Z_n)$ has the <u>strong negative S-correlation property</u> if for any $\lambda, \hat{\lambda} \in \Lambda$ such that $\lambda' \hat{\lambda} = 0$ (i.e., λ and $\hat{\lambda}$ are orthogonal) $Y(\lambda)$ is negatively S-correlated with $Y(\hat{\lambda})$.

Clearly, if $(Z_1, \ldots, Z_n)$ are independent or if they have a joint Gaussian distribution with $\sigma_{ij} \leqq 0$ for $i \neq j$, they satisfy the strong negative S-correlation property.

We now have the apparatus necessary to derive the analogue of Theorem I for the Expected-Utility model.

<u>Theorem V</u>: If $U' > 0$, $U'' < 0$, $E\{Z_i\} = E\{Z_j\}$ for all $i, j = 1, \ldots, n$ and the family of random variables $(Z_1, \ldots, Z_n)$ has the <u>strong</u> negative S-correlation property, then the problem:

$\max\limits_{\{\lambda_i\}} E\{U(\sum_{i=1}^{n} \lambda_i Z_i)\}$ subject to $\Sigma\lambda_i = 1$ has the property that $\lambda_i^* > 0$ for all i.

[It can be seen that this theorem is a stronger version of Samuelson's Corollary II.][19]

<u>Proof</u>: Suppose $\lambda_1^*, \ldots, \lambda_{k-1}^* \leqq 0$, $\lambda_k^*, \ldots, \lambda_n^* > 0$ and $\lambda_j^* < 0$ for some j such that $1 \leqq j \leqq k-1$. (If no $\lambda_j^* < 0$, then we are in the case of Theorem 4, where the F.O.C.'s hold with equality).

Since $\sum_{i=1}^{n} \lambda_i^* = 1$, $\sum_{i=k}^{n} \lambda_i^* \geqq 1$. Also, by our previous assumption, $\sum_{i=1}^{k-1} \lambda_i^* < 0$. Let $X_1 = - \sum_{i=1}^{k-1} \lambda_i^* Z_i / \sum_{i=1}^{k-1} \lambda_i^*$ and $X_2 = \sum_{i=k}^{n} \lambda_i^* Z_i / \sum_{i=k}^{n} \lambda_i^*$. Using this notation, we can rewrite the previous problem:

$$\max_{\{\lambda\}} E\{U(\lambda X_1 + (1 - \lambda)X_2)\}.$$

It is easily seen from our construction that $X_1 = Y(\lambda)$ for some $\lambda \in \Lambda$, $X_2 = Y(\hat{\lambda})$ for some $\hat{\lambda} \in \Lambda$ and $\lambda' \hat{\lambda} = 0$.

Therefore, X_1 and X_2 are negatively S-correlated by the assumptions of the theorem. Furthermore, since $E\{Z_i\} = E\{Z_j\}$, for all i, j it is easily seen that $E\{X_1\} = E\{X_2\}$. Thus we are back in the case of Theorem II, and the result is proved.

V. Summary

We have now derived the diversification theorems for the expected utility model. In our derivation we have seen that, with suitable modifications, Samuelson's conjectural definition of non-linear correlations is a very useful definition of statistical dependence for the theory of decision making under uncertainty. The results which we have obtained are very weak, even in comparison with analogous results in the mean-variance model, but this is yet another strong indication that the mean-variance model probably cannot be considered an adequate approximation of the expected utility model. Thus, it is time that more research be directed towards the expected utility model in a multivariate context. Our results are a small advance in that direction.

FOOTNOTES

*The author expresses gratitude to Paul Samuelson and Harold Freeman for many helpful discussions.

[1]Samuelson, P. A., "General Proof that Diversification Pays," Journal of Financial and Quantitative Analysis, 3, 1-13, 1967.

[2]See, for example: Merton, R. D., "An Analytic Derivation of the Efficient Portfolio Frontier," Alfred P. Sloan School of Management (M.I.T.) Working Paper 493-70, October 1970.

[3]See Merton, op. cit.

[4]Evans, J. L., and Archer, S. H., "Diversification and the Reduction of Dispersion: An Empirical Analysis," Journal of Finance, 761-769, December 1968.

[5]For examples of this see: Hanoch, G., and Levy, H., "The Efficiency Analysis of Choices Involving Risk," The Review of Economic Studies, 36, 335-346, 1969.

[6]For an examination of most of the common "non-expected utility" criteria see: Borch, K. H., The Economics of Uncertainty, Princeton University Press.

[7]See, for example: Mossin, J., "Equilibrium in a Capital Asset Market," Econometrica, 34, 768-783, 1966.

[8]If an investor's utility function displays increasing absolute risk aversion he will buy more insurance against a possible loss of fixed size as his income increases. See Pratt, J. W., "Risk Aversion in the Small and in the Large," Econometrica, 32, Jan.-April 1964.

[9]Hadar, J., and Russell, W., "Stochastic Dominance and Diversification," Journal of Economic Theory, 1971, 288-305.

[10]Samuelson, op. cit.

[11]Samuelson, op. cit.

[12]Hanoch and Levy, op. cit.

[13]Samuelson, op. cit.

[14]Samuelson, P. A. and Merton, R. K., "A Complete Model of Warrant Pricing that Maximizes Utility," Industrial Management Review, 10, 17-46, 1969.

[15]Samuelson and Merton, *op. cit.*, Appendix A.

[16]Samuelson and Merton, *op. cit.*

[17]Samuelson, *op. cit.*

[18]See, for example: Baumol, W. J., Economic Theory and Operations Analysis, Third Edition, Prentice-Hall, Chapter 7.