Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 587

August 1997

ON THE INVERSE OF THE COVARIANCE MATRIX IN PORTFOLIO ANALYSIS

Guy V.G. Stevens

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author. Recent IFDPs are available on the Web at www.bog.frb.fed.us.

ON THE INVERSE OF THE COVARIANCE MATRIX IN PORTFOLIO ANALYSIS

Guy V.G. Stevens*

Abstract: The goal of this study is the derivation and application of a direct characterization of the *inverse* of the covariance matrix central to portfolio analysis. As argued below, such a specification, in terms of a few primitive constructs, provides new and illuminating expressions for such key concepts as the optimal holding of a given risky asset and the slope of the risk-return efficiency locus faced by the individual investor. The building blocks of the inverse turn out to be the regression coefficients and residual variance obtained by regressing the asset's excess return on the set of excess returns for all other risky assets.

Keywords: portfolio analysis, covariance matrix, inverse matrix.

* The author is a Senior Economist in the Division of International Finance, Board of Governors of the Federal Reserve System. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other persons associated with the Federal Reserve System. My thanks go to Michael Adler, John Ammer, Neil Ericsson, William Helkie and, especially, to Michael McCarthy and Charles Thomas for very helpful discussions and comments.

ON THE INVERSE OF THE COVARIANCE MATRIX IN PORTFOLIO ANALYSIS

Guy V.G. Stevens

I. Introduction

The goal of this study is the derivation and application of a direct characterization of the *inverse* of the covariance matrix $[\sigma_{ij}]$ central to portfolio analysis. As argued below, such a specification of the inverse, in terms of a few primitive constructs, provides new and illuminating expressions for such key concepts as (1) the optimal holding of a given risky asset, (2) the slope of the risk-return efficiency locus faced by the individual investor, and (3) the pricing of risky assets in the Capital Asset Pricing Model. The building blocks of the inverse matrix turn out to be the non-diversifiable part of each asset's variance of return and coefficients obtained by regressing the asset's excess return on the set of excess returns for all other risky assets.

II. Preliminaries

It is well known that every optimizing mean-variance investor will choose a portfolio falling on his or her risk-return efficiency frontier -- the locus of portfolios of minimum variance conditional on a given expected return. As shown by Mossin (1973) and others, with the existence of a riskless asset, this frontier is a straight line in mean-standard deviation space, with a slope, *dE/dS,* equal to $(\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}})^{1/2}$ -- where $\boldsymbol{C}^{-1}$, the inverse of the matrix of asset variances and covariances, is the subject of this paper, and $\bar{\boldsymbol{m}}$ (and its transpose, $\bar{\boldsymbol{m}}'$) is the vector of the excess expected return of each asset over the riskfree rate: $\bar{r}_i - r_f$. To illustrate this derivation, assume markets for N risky assets, each with stochastic return $\tilde{r}_i$ and expected return $\bar{r}_i$, along with the opportunity for unlimited lending and borrowing of a riskless asset with return r_f. For an investor with initial wealth *W*, the expected return on any portfolio, *E(Y)*, can be defined as:

$$E(Y) = \bar{Y} = r_f W + \sum_{i=1}^{N} z_i (\bar{r}_i - r_f) = r_f W + \boldsymbol{z}'\bar{\boldsymbol{m}}. \tag{1}$$

In addition to previously defined terms, $\boldsymbol{z}$ is the $N \times 1$ column vector of nominal security holdings, with elements z_i. The variance of Y is:

$$V(Y) = \sum_{i=1}^{N} \sum_{j=1}^{N} z_i z_j \sigma_{ij} = \boldsymbol{z}'\boldsymbol{C}\boldsymbol{z}. \tag{2}$$

Minimizing *V(Y)* subject to a predetermined level of *E(Y)* yields the following set of first order conditions for points on the risk-return efficiency frontier:

$$2\boldsymbol{C}\boldsymbol{z} - \lambda\bar{\boldsymbol{m}} = 0, \tag{3}$$

where λ is the Lagrange multiplier for the constraint.

Solving this system of equations for $\boldsymbol{z}$ leads to the expression for the vector of optimal holdings of risky assets along the efficiency frontier:

$$\boldsymbol{z}* = \lambda\boldsymbol{C}^{-1}\bar{\boldsymbol{m}}/2. \tag{4}$$

Although the level of vector $\boldsymbol{z}*$ depends on the unknown λ -- and, therefore, generally on the investor's utility function and the required expected return -- equation (4) does fix the *ratios* of the various risky assets along the efficiency frontier in any optimal portfolio: the famous portfolio separation theorem discovered by Tobin (1958). These optimal ratios will be preference-free, depending only on the investor's estimates of expected excess returns, and, once again, the elements, c_{ij}^{-1}, of the inverse of the covariance matrix:

$$\frac{z^*_i}{z^*_j} = \frac{\sum_{k=1}^{N} \bar{m}_k \; c_{ik}^{-1}}{\sum_{k=1}^{N} \bar{m}_k \; c_{jk}^{-1}}. \tag{5}$$

By combining equations (4), (1) and (2), one derives the investor's risk-return frontier -- which happens to be linear in expected return and standard deviation:

$$E(Y) - r_f W = \left(\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}}\right)^{1/2}\sqrt{V(Y)}\,. \tag{6}$$

Equations (4) and (5) and (6) emphasize the importance of the elements of the *inverse* of the covariance matrix $\boldsymbol{C}$. Typically in portfolio analysis, however, we do not characterize these elements directly, but only indirectly, as those elements that map the original covariance matrix, $\boldsymbol{C}$, into the identity matrix. The purpose of this note is to derive a *direct characterization* of the elements of $\boldsymbol{C}^{-1}$, one that relies on a few key constructs and that leads to straightforward explanations of the optimal ratios in (5) and the slope of the risk-return locus in (6).

II. Derivation

The derivation of $\boldsymbol{C}^{-1}$ below adapts a useful partitioning technique developed by Anderson and Danthine (1981) for their study of hedging in futures markets. Partition the set of the *N* first order conditions (3), above, between the first equation and a *N-1* equation block; in matrix notation, the partitioned system appears as follows:

$$\begin{bmatrix} \sigma_{11} & \boldsymbol{\sigma_{1j}} \\ \boldsymbol{\sigma_{j1}} & \boldsymbol{C_{N-1}} \end{bmatrix} \begin{bmatrix} z_1 \\ \boldsymbol{z_{N-1}} \end{bmatrix} = \frac{\lambda}{2} \begin{bmatrix} \bar{m}_1 \\ \bar{\boldsymbol{m}}_{\boldsymbol{N-1}} \end{bmatrix}, \tag{7}$$

The scalars $\sigma_{11}, z_1, \bar{m}_1,$ are the variance, asset level, and expected excess return for asset *1*; $\boldsymbol{\sigma_{1j}}$ is the $1{\times}N\text{-}1$ row vector of covariances between the first asset and the *N-1* other assets, and $\boldsymbol{\sigma_{j1}}$ is its transpose. The matrix $\boldsymbol{C_{N-1}}$ in the bottom block is the *N-1* square submatrix of the covariance matrix

$\boldsymbol{C}$ formed by eliminating its first row and column; finally, $\boldsymbol{z}_{N-1}$ and $\bar{\boldsymbol{m}}_{N-1}$ are the *N-1* column vectors made up of all but the first elements of the original $\boldsymbol{z}$ and $\bar{\boldsymbol{m}}$ vectors, respectively.

To facilitate the derivation, I shall use familiar econometric notation for the four submatrices in (7). Let $\mathbf{Y}$ be a $T \times 1$ column vector of observations on variable 1 (taken around the mean), and $\mathbf{X}$ the $T \times N\text{-}1$ matrix of observations on the remaining *N-1* variables. The variance of variable 1, σ_{11}, the scalar in the upper left-hand corner of the covariance matrix in (7), can be expressed as $\mathbf{Y'Y}/\mathrm{T}$; $\boldsymbol{\sigma}_{1j}$ equals $\mathbf{Y'X}/\mathrm{T}$ and $\boldsymbol{\sigma}_{j1}$ is $\mathbf{X'Y}/\mathrm{T}$; finally, the bottom *N-1* square block, $\boldsymbol{C}_{N-1}$, is $\mathbf{X'X}/\mathrm{T}$.

Standard results on partitioned matrix inversion indicate that the inverse can be partitioned similarly to $\boldsymbol{C}$ in (7).[1] If the matrix with submatrices $\boldsymbol{A}_{ij}$ is indeed $\boldsymbol{C}^{-1}$ we have the following:

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11} & \boldsymbol{\sigma}_{1j} \\ \boldsymbol{\sigma}_{j1} & \boldsymbol{C}_{N-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{Y'Y}/T & \boldsymbol{Y'X}/T \\ \boldsymbol{X'Y}/T & \boldsymbol{X'X}/T \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \tag{8}$$

System (8) leads to the following four equations sufficient to determine each of the submatrices:[2]

$$\begin{aligned} \boldsymbol{A}_{11}(\boldsymbol{Y'Y}/T) + \boldsymbol{A}_{12}(\boldsymbol{X'Y}/T) &= \boldsymbol{I} \\ \boldsymbol{A}_{11}(\boldsymbol{Y'X}/T) + \boldsymbol{A}_{12}(\boldsymbol{X'X}/T) &= \mathbf{0} \\ \boldsymbol{A}_{21}(\boldsymbol{Y'Y}/T) + \boldsymbol{A}_{22}(\boldsymbol{X'Y}/T) &= \mathbf{0} \\ \boldsymbol{A}_{21}(\boldsymbol{Y'X}/T) + \boldsymbol{A}_{22}(\boldsymbol{X'X}/T) &= \boldsymbol{I} \end{aligned} \tag{9}$$

Let us initially take the second equation in the set (9) and solve it for $\boldsymbol{A}_{12}$:

---

[1] See, for example, Goldberger (l964), pp. 27-28, and Theil (1971), pp. 16-19.

[2] Recall that $\boldsymbol{A}_{11}$ and the identity matrix in the first equation in set (9) are both scalars. It should also be noted that $\boldsymbol{A}_{21}$ is the transpose of $\boldsymbol{A}_{12}$, because of the symmetry of $\boldsymbol{C}^{-1}$.

$$A_{12} = -A_{11}(Y'X)(X'X)^{-1} \tag{10}$$

Note that the $1 \times N\text{-}1$ matrix A_{12} equals the (scalar) inverse element $-A_{11}$ times a term that is the row vector of *regression* coefficients, β_1', the result of regressing the returns from asset *1* on those of all the other *N-1* risky assets. Next substitute (10) into the first equation of set (9) and solve for A_{11}. After factoring out the number of observations, *T*, and inverting the resulting matrix, we get:

$$A_{11}/T = \left[Y'\,[I - X(X'X)^{-1}X']\,Y\right]^{-1} \tag{11}$$

Because the product $[I - X(X'X)^{-1}X']Y$ can be shown to equal the vector of residuals, u, from the above regression, and because $I - X(X'X)^{-1}X'$ is idempotent (its square is equal to itself), equation (11) can be simplified as follows:

$$A_{11}/T = \left[Y'[I - X(X'X)^{-1}X']\,[I - X(X'X)^{-1}X']\,Y\right]^{-1} = \left[Y - X\hat{\beta}]'[Y - X\hat{\beta}]\right]^{-1} = [u'u]^{-1} \tag{12}$$

The multiple regression coefficient, R_1^2, for the regression of returns from the first asset on those for all the other assets is defined as 1- $u'u/Y'Y$; therefore, $u'u$ can be expressed as $Y'Y(1\text{-}R_1^2)$. This leads directly to the final expressions for the first row and, by symmetry, the first column of C^{-1} :[3]

$$A_{11} = \frac{T}{u'u} = \frac{T}{(1-R_1^2)Y'Y} = \frac{1}{\sigma_{11}(1-R_1^2)} \tag{13}$$

$$A_{12} = -A_{11}(Y'X)(X'X)^{-1} = \frac{-\beta_1'}{\sigma_{11}(1-R_1^2)} \tag{14}$$

The common factor in these elements of the inverse is $\sigma_{11}(1-R_1^2)$, that part of the variance of the first return that *cannot* be explained by a regression on the other risky returns; this is shown in equation (13) to be equivalent to the estimate of the variance of the residual of that regression and will play an

[3] It might be noted that the original version of this paper, Stevens (1995), contains an alternative derivation of equations (13) and (14) that does not rely on partitioned inversion or the standard econometric notation.

important role in the applications below.

Although equations (13) and (14) are only part of $\boldsymbol{C}^{-1}$, they are sufficient, in conjunction with equation (4), to obtain the final expression for the optimal level of the first asset:

$$z*_1 = \left[\frac{1}{\sigma_{11}(1 - R_1^2)} \right] \frac{\lambda \bar{m}_1}{2} + \sum_{j=2}^{N} \left(\left[- \frac{\beta_{1j}}{\sigma_{11}(1 - R_1^2)} \right] \frac{\lambda \bar{m}_j}{2} \right) \tag{15}$$

A. *Alternative 1 for Completing* $\underline{\boldsymbol{C}}^{-1}$

In determining the remaining elements of $\boldsymbol{C}^{-1}$, rather than focusing on solving the last two equations in set (9) for $\boldsymbol{A}_{22}$, we shall exploit equation (15) and the fact that the choice of a particular asset as the first or "***Y***" variable is clearly an arbitrary one. Let us permute the rows and columns of $\boldsymbol{C}$ to move the moments of some other asset i to the first row and column, thus forming a new covariance matrix, $\boldsymbol{C*}$. By repeating the partitioned matrix inversion steps in equations (9)-(14), above, we can derive the elements of the first row and column of $\boldsymbol{C*}^{-1}$, the matrices $\boldsymbol{A*}_{11}$ and $\boldsymbol{A*}_{12}$. We now have $\boldsymbol{A}*_{11} = 1/\sigma_{ii}(1-R_i^2)$ and $\boldsymbol{A*}_{12}$ a $1 \times N\text{-}1$ vector with elements $-\beta_{ij}/\sigma_{ii}(1-R_i^2)$; the subscript i refers to the regression where the ith variable of $\boldsymbol{C}$ is now taken as the dependent variable.

The major remaining question is to determine the relationship between the elements of the first row of $\boldsymbol{C*}^{-1}$ and the ith row of $\boldsymbol{C}^{-1}$. Substituting the above elements from the first row of $\boldsymbol{C*}^{-1}$ into equation (15), we get one expression for $z*_i$; further, the ith row of the original optimal solution, equation (4), provides a second expression for $z*_i$, this time in terms of the elements of the ith row of $\boldsymbol{C}^{-1}$. The difference between these two alternative expressions for $z*_i$ implies the following:

$$0 = \left[\frac{1}{\sigma_{ii}(1 - R_i^2)} - \boldsymbol{c}_{\boldsymbol{ii}}^{\boldsymbol{-1}} \right] \frac{\lambda \bar{m}_i}{2} + \sum_{j \neq i}^{N} \left(\left[- \frac{\beta_{ij}}{\sigma_{ii}(1 - R_i^2)} - \boldsymbol{c}_{\boldsymbol{ij}}^{\boldsymbol{-1}} \right] \frac{\lambda \bar{m}_j}{2} \right) \tag{16}$$

Since the excess expected returns, $\bar{m}_i$ and $\bar{m}_j$, may assume any value, the only way for (16) to hold in general is for each term in square brackets to be identically equal to zero. Thus, for any row i in $\boldsymbol{C}^{-1}$, c_{ii}^{-1} must equal the reciprocal of $\sigma_{ii}(1 - R_i^2)$ and c_{ij}^{-1} must equal $-\beta_{ij} / \sigma_{ii}(1 - R_i^2)$.

The upshot of all of the above is the following direct characterization of $\boldsymbol{C}^{-1}$:[4]

$$\begin{bmatrix} \dfrac{1}{\sigma_{11}(1 - R_1^2)} & -\dfrac{\beta_{12}}{\sigma_{11}(1 - R_1^2)} & \cdots & -\dfrac{\beta_{1N}}{\sigma_{11}(1 - R_1^2)} \\ -\dfrac{\beta_{21}}{\sigma_{22}(1 - R_2^2)} & \dfrac{1}{\sigma_{22}(1 - R_2^2)} & \cdots & -\dfrac{\beta_{2N}}{\sigma_{22}(1 - R_2^2)} \\ \vdots & \vdots & \ddots & \vdots \\ -\dfrac{\beta_{N1}}{\sigma_{NN}(1 - R_N^2)} & -\dfrac{\beta_{N2}}{\sigma_{NN}(1 - R_N^2)} & \cdots & \dfrac{1}{\sigma_{NN}(1 - R_N^2)} \end{bmatrix} \tag{17}$$

[4] There are at least two questions on which one might want further verification -- namely, (1) proof that the inverse (17) is indeed symmetric, and (2) further evidence that $\boldsymbol{CC}^{-1} = \boldsymbol{I}$. Neither is immediately obvious by multiplying the various rows and columns of the two matrices, because we are forced to multiply estimated coefficients that relate to different regressions. However, by adapting some results of Johnston (1972), p. 132 ff., one can relate the elements in matrix (17) to the determinant, $|\boldsymbol{C}|$, and the various cofactors of the elements in $\boldsymbol{C}$, $\boldsymbol{COF}_{ij}$. Johnston develops his relationships in terms of the correlation matrix, but that can easily be rewritten in terms of our covariance matrix. Following Johnston's derivation of his equation (5-34) one can show that a diagonal element in (17) equals $\boldsymbol{COF}_{ii} / |\boldsymbol{C}|$, where the first term is the cofactor of the diagonal element c_{ii}. Moreover, the regression coefficient, β_{ij}, equals $-\boldsymbol{COF}_{ij} / \boldsymbol{COF}_{ii}$. Putting these two results together, the element c_{ij}^{-1} equals $\boldsymbol{COF}_{ij} / |\boldsymbol{C}|$; similarly, we have $c_{ji}^{-1} = \boldsymbol{COF}_{ji} / |\boldsymbol{C}|$. However, since $\boldsymbol{C}$ is symmetric $\boldsymbol{COF}_{ij} = \boldsymbol{COF}_{ji}$, and the two elements are equal -- verifying the symmetry of the inverse. (For what it may be worth, the above equalities posit some possibly interesting relationships between regression coefficients from different regressions).

These expressions for the elements of $\boldsymbol{C}^{-1}$ facilitate showing that $\boldsymbol{CC}^{-1} = \boldsymbol{I}$. Recalling that by symmetry the columns are identical to the corresponding rows of the inverse, when multiplying the ith row of $\boldsymbol{C}$ with the ith column of the inverse, one gets the sum of products of the elements in the ith row of $\boldsymbol{C}$ each multiplied by its corresponding cofactor, all divided by $|\boldsymbol{C}|$ -- the net result being *1*. All off-diagonal elements of $\boldsymbol{CC}^{-1}$ must be equal to zero, since they involve an expansion by alien cofactors.

B. *A Second Alternative for Completing* $\boldsymbol{C}^{-1}$

Consider the fundamental expression for the element, c^{-1}_{ij}, in $\boldsymbol{C}^{-1}$ as the ratio of two determinants. Where $|\boldsymbol{C}|$ is the determinant of the matrix, $\boldsymbol{COF_{ij}}$ is the cofactor of the *ij*th element of $\boldsymbol{C}$, and $\boldsymbol{M_{ij}}$ is the minor of that same element, we have: [5]

$$\boldsymbol{c}^{-1}_{ij} = \frac{\boldsymbol{COF_{ji}}}{|\boldsymbol{C}|} = \frac{(-1)^{i+j}\boldsymbol{M_{ji}}}{|\boldsymbol{C}|} = \frac{(-1)^{i+j}\boldsymbol{M_{ij}}}{|\boldsymbol{C}|} \tag{18}$$

The last part of the equality holds because of the symmetry of $\boldsymbol{C}$.

Using (18), consider the expressions for an element, c^{-1}_{ii}, on the diagonal of $\boldsymbol{C}^{-1}$ and for the element $c*^{-1}_{1j}$, where, as above, $\boldsymbol{C}$* is formed by moving the ith row and column of $\boldsymbol{C}$ to the first row and column of $\boldsymbol{C}$ *****. Note first that $|\boldsymbol{C}|$ equals $|\boldsymbol{C}*|$, since the number of permutations of rows and columns to move from $\boldsymbol{C}$ to $\boldsymbol{C}$* is even, $(-1)^{2(i-1)}$. Note also that the minor, M_{ii}, in the numerator of c^{-1}_{ii} is identical to $M*_{11}$: in forming the matrix $\boldsymbol{C}$* only the ith row and column have changed position; but since this row and column is eliminated in forming *both* of the above minors, all of the other elements of these two determinants are unchanged, and they must be equal. Finally, in forming the respective cofactors, the sign attaching to both of these minors must be positive: $(-1)^2 = (-1)^{2i}$. Applying these results to equation (18), we find that the leading element of $\boldsymbol{C*}^{-1}$, $\boldsymbol{c*}^{-1}_{11}$, the value of which has been determined above, equals the diagonal element, $\boldsymbol{c}^{-1}_{ii}$, of $\boldsymbol{C}^{-1}$**:**

[5] See Goldberger (1964), p.24, for the theorem. For completeness, it might be noted that the (first) minor of of c_{ij} is the determinant formed by eliminating the *ith* row and *j*th column of the determinant of $\boldsymbol{C}$; the cofactor of c_{ij} is the signed minor: $(-1)^{i+j} M_{ij}$.

$$c^{-1}_{ii} = \frac{\boldsymbol{COF_{ii}}}{|\boldsymbol{C}|} = \frac{(-1)^{2i}\boldsymbol{M_{ij}}}{|\boldsymbol{C}|} = \frac{(-1)^{2}\boldsymbol{M}*_{ij}}{|\boldsymbol{C}*|} = \boldsymbol{c}*^{-1}_{11} = 1/\sigma_{ii}(1-R_i^2) \tag{19}$$

Continuing with the same ***C**** matrix, for the off-diagonal elements in the first row $\boldsymbol{c}*^{-1}_{1j}$, the story is slightly more complicated, but similar. The complication arises because the minors of the off-diagonal elements in ***C****, although containing the same elements, are not identical to those for the corresponding off-diagonal element in ***C***. Since the *i*th row of ***C*** (the first row of ***C****) is still eliminated in forming all of the relevant minors, the permutation of this row continues to pose no problem; however, the *i*th *column* of ***C*** does appear in all off-diagonal minors and appears in a different position in the relevant minors of ***C****. However, the changes in sign caused by the permutations of this column are offset by changes in sign in the cofactor of the relevant inverse elements. We have two cases:

(1) For a given element of the ***C*** matrix, c_{ij} , if $j>i$, we have $(-1)^{i-1}\boldsymbol{M}_{ij} = M*_{1j}$ leading to:

$$c^{-1}_{ij} = \frac{(-1)^{i+j}\boldsymbol{M_{ij}}}{|\boldsymbol{C}|} = \frac{(-1)^{i+j}(-1)^{1-i}\boldsymbol{M^*_{1j}}}{|\boldsymbol{C}|} = \frac{(-1)^{1+j}\boldsymbol{M^*_{1j}}}{|\boldsymbol{C}^*|} = \boldsymbol{c}*^{-1}_{1j} = -\beta_{ij}/\sigma_{ii}(1-R_i^2) \tag{20}$$

(2) For a given element of the ***C*** matrix, c_{ij} , if $j<i$, we have $(-1)^{i-2}\boldsymbol{M}_{ij} = \boldsymbol{M*}_{1,j+1}$ leading to:[6]

$$c^{-1}_{ij} = \frac{(-1)^{i+j}\boldsymbol{M_{ij}}}{|\boldsymbol{C}|} = \frac{(-1)^{i+j}(-1)^{2-i}\boldsymbol{M^*_{1,j+1}}}{|\boldsymbol{C}|} = \frac{(-1)^{1+j+1}\boldsymbol{M^*_{1,j+1}}}{|\boldsymbol{C}^*|} = \boldsymbol{c}*^{-1}_{1,j+1} = -\beta_{ij}/\sigma_{ii}(1-R_i^2) \tag{21}$$

To summarize, we have shown that all the elements in rows *2* through *N* of $\boldsymbol{C}^{-1}$ are equal to elements in the first row of an appropriately defined $\boldsymbol{C}^{*-1}$; these latter elements of course can

[6] In equations (20) and (21) the row and column subscripts of the regression coefficient, β_{ij}, refer to the rows and columns of the original covariance matrix, ***C***.

be evaluated by applying the partitioned matrix inversion procedure, equations (9) through (14), leading again to the inverse matrix (17).

IV. Implications

A. _Asset Holdings_

On first glance the inverse (17), that is so central in the expressions for the optimal level (4) and ratios (5) of dollar holdings of risky assets, seems to contain a welter of intriguing but not particularly illuminating terms. It turns out, however, that these terms combine to yield understandable and intuitively attractive expressions for the holding of a given asset, both in special cases and in the general case with arbitrary, non-zero covariances.

Consider first the special case of independent returns. With all off-diagonal elements zero in the original variance-covariance matrix, the inverse matrix (17) is also diagonal, with each element equal to the reciprocal of a given asset's variance, $1/\sigma_{ii}$. Using the implication of equation (4) that asset holdings are proportional to $\boldsymbol{C}^{-1}\bar{\boldsymbol{m}}$, in the independence case the holding for any asset *i* is therefore proportional to the ratio of its excess expected return to its variance, $\bar{m}_i/\sigma_{ii}$.

Although considerably more complicated, the general case can be interpreted as a natural generalization of the independence case. Let us consider in turn the denominator and numerator of the expression implied by equation (4) and the general form of the inverse matrix (17). Since (4) shows that the holding of risky asset *i* is proportional to the matrix product of the *i*th row of $\boldsymbol{C}^{-1}$ with the column vector of excess expected returns, the denominator of the expression becomes $\sigma_{ii}(1 - R_i^2)$ -- instead of σ_{ii} in the independence case. The squared multiple regression coefficient appearing in the denominator, R_i^2, equals that maximum percentage of the variance of the return of asset *i* that can be explained by a linear combination of the returns of all other available risky assets; because the optimal linear combination minimizes the residual variance, it is easily shown that the denominator is the minimum *non-diversifiable* part of asset *i*'s variance.[7] The coefficients of this optimal linear combination are calculated via a least-squares regression; in other contexts, this optimal combination has also been called the *pure hedge* or a regression hedge[8]

The numerator in the general expression, instead of $\bar{m}_i$ in the independence case, becomes $\bar{m}_i - \sum_{k \neq i} \beta_{ik} \bar{m}_k$. This, too, has intuitive appeal, as soon as one recalls that the regression line to which the β_{ik}s apply passes through the point of the means. Letting α_i equal the intercept of the regression for the *i*th asset, we have therefore that $\bar{m}_i - \alpha_i - \sum_{k \neq i} \beta_{ik} \bar{m}_k = 0$; hence the numerator of the general expression is equal to the *intercept* of this regression equation. As such, the numerator equals that part of the expected excess return of asset *i* that *cannot* be accounted for by the excess expected return of the same linear combination of assets that minimizes the residual variance of asset *i*'s return -- i.e., the numerator equals the difference between $\bar{r}_i - r_f$ and the expected costs of the optimal hedge. Thus, as contrasted with the raw or unadjusted expected returns and variances that determine asset holdings in the independence case, the expression for holdings in the general case uses the same concepts, but in *adjusted* form -- adjusted for that part of the asset's expected excess return and variance that can be explained by the optimal linear combination of other risky assets:

---

[7] Consider the "portfolio" formed by a dollar in asset *i* and the amount $-\beta_{ik}$ in each of the other assets *k*, where β_{ik} is the appropriate coefficient, appearing in the inverse matrix (17), from the multiple regression of the excess return for the ith asset on the excess returns of all the other assets. By definition, for any sample period, the value or observed return of this "portfolio" will be the residual from the least-squares multiple regression defined above. Since a property of the regression is the minimization of the variance of this residual over the sample period, or the maximization of the explanation of the variance of the return of asset *i*, no other linear combination of these asset returns can reduce this residual variance further. The variance of this "portfolio" will be $\sigma_{ii}(1 - R_i^2)$.

[8] In Anderson and Danthine's 1981 study of hedging in future's markets, the optimal linear combination balancing their "cash" position was denoted as the *pure* hedge (p.1187). In an international setting, Adler and Dumas (1980) identify an asset's currency risk exposure as a coefficient in a particular linear regression.

$$z*_i = (\lambda/2)\,\frac{\bar{m}_i - \sum_{k\neq i}\beta_{ik}\bar{m}_k}{\sigma_i^2(1-R_i^2)} \tag{22}$$

Despite the welter of extra terms, the general expression for the holdings of a given asset is, therefore, a natural generalization of that for the independence case.

B. *The Investor's Risk-Return Frontier*

Equation (6), above, shows the investor's optimal tradeoff between risk and return to be linear with slope $(\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}})^{1/2}$. The direct characterization of $\boldsymbol{C}^{-1}$ in (17) can show us both how to interpret this expression intuitively and how it changes in response to changes in the underlying structural elements -- the expected returns, variances and covariances.

In the simplest case, a world with a single risky asset, $\boldsymbol{C}^{-1}$ collapses to $1/\sigma_{11}$ and the slope, dE/dS, reduces to $\bar{m}_1/\sqrt{\sigma_{11}}$, the only possible tradeoff between risk and return in such a world. The second simplest case involves adding a second risky asset, but with a return independent from the first. This changes $(\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}})^{1/2}$ and dE/dS to the square root of $\bar{m}_1^2/\sigma_{11} + \bar{m}_2^2/\sigma_{22}$. If the investor optimizes -- in this case by taking a diversified portfolio -- the optimal tradeoff becomes a function of the expected excess returns of the two assets: the square of dE/dS being a weighted average of the squared excess returns, the weights being the reciprocals of the variances of the respective assets. Because of the power of diversification, the investor's tradeoff between risk and return can easily be shown to have *improved* (the slope increased) over the single asset case, irrespective of the second asset's expected return or variance; even a negative excess expected return improves the investor's opportunity set.[9] Because of the diagonality of (17) in the case of independent asset returns, the above

[9]For the independence case, the numerator of equation (22) shows that a negative expected excess return implies an optimal *short* position for the asset in question.

results are easily generalized to any number of assets. Thus, for N assets:

$$\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}} = \sum_{i=1}^{N} \left[\bar{m}_i^2/\sigma_{ii}\right]. \tag{23}$$

The most realistic cases of course are those where the asset returns are correlated -- where the off-diagonal elements and the multiple correlation coefficients in (17) are non-zero. How the knowledge of $\boldsymbol{C}^{-1}$ helps in the analysis can be illustrated by a consideration of the general 2-asset case. This case is more important than it might seem, because one of the two assets could be the overall market portfolio.

After some algebra, the general form of the risk-return tradeoff can be related to that for the independence case as follows:

$$\bar{\boldsymbol{m}}'\boldsymbol{C}^{-1}\bar{\boldsymbol{m}} = \frac{\bar{m}_1^2}{\sigma_1^2} + \frac{\bar{m}_2^2}{\sigma_2^2} + \left[\frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\right]\left[\rho^2(\bar{m}_1\sigma_2 - \bar{m}_2\sigma_1)^2 + 2\rho\bar{m}_1\bar{m}_2\sigma_1\sigma_2(\rho-1)\right]. \tag{24}$$

The only new symbol introduced in (24) is the correlation coefficient ρ, the square of which in this two-asset case equals both R_1^2 and R_2^2 (both appearing in (17)); for clarity, we also denote variances in (24) by the square of the standard deviation, σ_i^2 rather than σ_{ii}.

With (24) one can address two significant questions about the effect of a non-zero covariance: whether the non-zero correlation between the two risky returns improves or worsens the investor's risk-return tradeoff relative to the two-asset independence case; and whether the addition of the second asset improves the tradeoff over that offered by the first asset alone. Improvement relative to the independence case depends on the sign of the sum inside the right brackets; improvement relative to the trade-off offered by the first asset alone depends on whether the product of the terms within brackets is greater (less in absolute value) than $-\bar{m}_2^2/\sigma_2^2$.

One immediate implication of (24) is that when both expected excess returns are positive and the two returns are *negatively* correlated, the risk-return tradeoff must improve -- in both senses defined above. In this case, the last term in (24) becomes positive and, since all other terms are always positive, the difference between the tradeoff in this and the independence case must be positive.

For a *positive* ρ, the analysis of (24) leads to the answer "it depends" for both questions. For example, consider the case where the expected excess returns and variances of both assets are equal, causing the squared term in the right hand expression to equal zero. We can then simplify the whole right hand expression (the difference between the risk-return tradeoff in the general and the independence cases) into: $-2\rho\bar{m}_1\bar{m}_2 / \left[\sigma_1\sigma_2(1+\rho)\right] = -2\rho\bar{m}^2 / \left[\sigma^2(1+\rho)\right]$. In this special case, it is first apparent that any positive correlation *worsens* the risk-return tradeoff relative to the independence case. Moreover, it can be proved that for $\rho < 1$, the difference above is greater (less in absolute value) than $-\bar{m}^2/\sigma^2$; hence, the addition of a second asset does in this case improve the risk-return tradeoff over what was available for the first asset alone. However, as ρ approaches its limit of 1, the contribution of the second asset to the risk-return tradeoff can be shown to approach zero.

All of the above examples with a positive correlation were for the special case where both assets had identical means and variances (a more general, but equivalent restriction is that the *ratio* of each asset's mean and variance is equal). That the above result -- that a positive correlation worsens the tradeoff relative to the independence case -- is not general can be seen by examining a specific counter-example. Consider the following case where the excess returns of the two assets are allowed to diverge: $\rho = .95,\ m_1 = 10,\ m_2 = 2,\ \mathit{with}\ \sigma_{11} = \sigma_{22} = 4.$ Of the two terms in the rightmost bracket in equation (24), the positive squared difference between the means (231.04) far outweighs the last, negative term (-7.6). The explanation for this improvement in the risk-return tradeoff over the independence case, despite the positive correlation, can be understood by noting that equation (22) tells us that, for this given set of values, the optimal holding of asset 2 turns negative. With the high positive covariance between the two returns, the optimizing investor can control the risk buildup of a heavy investment in the high-return asset 1 by going *short* in asset 2.

C. <u>*Linkages to the Capital Asset Pricing Model*</u>

Assuming that the above decision framework applies to all investors and, further, that all share common beliefs with respect to the variance-covariance matrix and expected returns, one can derive the various results of the Capital Asset Pricing Model (CAPM). Of particular interest for this paper is the CAPM *security market line*. Derivable from the summation of equation (3) over all investors, the excess expected return of any asset, in equilibrium, can be expressed, in a number of alternative ways, as follows:[10]

$$\bar{r}_{i\$} - r_f = \left[\frac{(\bar{r}_{M\$} - r_f)}{\sigma^2_{M\$}}\right]\sigma_{iM\$} = \left[\frac{(\bar{r}_M - r_f V_M)}{\sigma^2_M}\right]\sigma_{iM} \ , \tag{25}$$

where the subscript *M* refers to the market portfolio, and a subscripted *$* indicates an expected return or variance *per dollar* of investment -- the absence of an *$* referring to an expected value or other moment for the asset or the market.[11] Thus $\bar{r}_M$, the expected return on the market portfolio, equals $\Sigma[V_i \bar{r}_{i\$}]$, where V_i is the total market value of the *i*th asset; correspondingly, the expected return per dollar invested in the market portfolio, $\bar{r}_{M\$}$, equals $\bar{r}_M / V_M$, where $V_M = \Sigma V_i$, the total value of the market. Similarly, $\sigma^2_{M\$} = \sigma^2_M / V^2_M$. Finally, the systematic risk of a marginal dollar in asset *i*, its covariance with the market, is: $\sigma_{iM} = E[\tilde{r}_i, r_M] = E[\tilde{r}_i, \Sigma V_j \tilde{r}_j] = \Sigma V_j \sigma_{ij}$,, which is V_M times the covariance per dollar of the market portfolio.

---

[10] See, e.g., Copeland and Weston (1983), pp. 187-189 or Elton and Gruber (1987), chapter 11, for good expositions and derivations of the security market line and its components. For future reference, one can also derive from the summation of equations (3) that $2(r_M - r_f)/\sigma_M^2 = 1/\Sigma\lambda_j$, where the last term is the summation of each investor's equilibrium value of λ.

[11] As noted a number of times above, prior to equation (25) all rates of return were returns *per dollar* of investment, despite the omission of the subscript *$*.

Equation (25) tells the now-standard CAPM story, linking the excess expected return on a asset to the product of its systematic risk and the market price of risk. On the other hand, a very different equilibrium expression for $\bar{r}_i - r_f$ can be derived from equation (22). Given the assumptions of the CAPM, we can add the expression for $\boldsymbol{z_i}$ over each investor j, to derive the following:

$$2V_i[\sigma_i^2(1-R_i^2)] = (\sum_{j=1}^{N}\lambda_j)[\bar{m}_i - \sum_{k\neq i}\beta_{ik}\bar{m}_k]\,. \tag{26}$$

It was established in the previous section that the term $\sigma_i^2(1-R_i^2)$ is the variance of a dollar invested in asset i when optimally hedged by going short β_{ik} in each other asset k; thus, if all of asset i is hedged in this way, its variance would be V_i^2 times the term in brackets. The left hand side of (26), therefore, is the derivative of this variance, the marginal change in the variance caused by an added dollar's investment in asset *i*.

Since $\bar{m}_i = \bar{r}_i - r_f$, and the sum of the Lagrange multipliers equals the reciprocal of the market price of risk,[12] (26) can be rewritten like (25), with the expected excess return on the left hand side:

$$\bar{r}_i - r_f = [(\bar{r}_M - r_f V_M)/\sigma_M^2]\,V_i[\sigma_i^2(1-R_i^2)] + \sum_{k\neq i}\beta_{ik}\bar{m}_k\,. \tag{27}$$

Unlike (25), however, the excess return on asset *i* is now (also) shown to equal the sum of expected costs due to the marginal increase in risk and to the use of the optimal hedge. Equating the right-hand sides of (27) and (25) shows the relationship between these two notions of marginal cost and risk:

$$\sigma_{iM} = V_i[\sigma_i^2(1-R_i^2)] + [\sigma_M^2/(\bar{r}_M - r_f V_M)]\sum_{k\neq i}\beta_{ik}\bar{m}_k \tag{28}$$

On the left hand side, from (25) and expressed in units of risk, is the traditional measure of systematic

[12] Summing equations (3) over all investors, and premultiplying the resulting equation by the row vector of market values of all assets, V_i, we get $2\sigma_M^2 = \Sigma\lambda_j(\bar{r}_M - r_f V_M)$.

risk. Because all alternative ways of increasing the holding of asset *i* must, in equilibrium, bear the same cost, this systematic risk must equal the sum of the marginal risk costs and hedging costs of the combination investment of a dollar's increase in *i* balanced by the various changes in the other holdings that constitute the optimal hedge.

V. Summary and Conclusions

This paper derives and applies the inverse of the covariance matrix central to portfolio analysis. As shown in equation (17) the inverse is composed of two key elements: (1) the non-diversifiable part of each asset's variance of return $[\sigma_i^2(1 - R_i^2)]$ and (2) the set of coefficients obtained by regressing the expected excess return for a given asset on the expected excess returns of all other available assets. It is of some interest to note that *everything* in $\boldsymbol{C}^{-1}$ relates to the characteristics of the N regressions that minimize each asset's residual variance -- which, for good reason, may be termed the optimal hedge regressions.

Knowledge of the inverse matrix leads to equation (22), an illuminating expression for the optimal holding of any given asset *i*. The numerator is proportional to the difference between asset *i*'s expected excess return and the expected excess return of its optimal hedging combination (the intercept of its optimal hedge regression). The denominator is that part of asset *i*'s variance that cannot be diversified away (the residual variance of the optimal hedge regression).

The inverse of the covariance matrix was also shown to be a central element in the expression for an investor's risk-return frontier and instrumental in providing an alternative expression for the CAPM's security-market line. Knowledge of $\boldsymbol{C}^{-1}$ was shown to be useful for analyzing shifts in the former, either because of changes in the underlying covariances or because of the introduction of new assets. The derivation of equation (22) led both to alternatives to traditional CAPM equations and to a clarification of the relationship between an asset's non-diversifiable risk and the traditional measure of its systematic risk.

References

Adler, Michael and Bernard Dumas, "The Exposure of Long-Term Foreign Currency Bonds," *Journal of Financial and Quantitative Analysis,* November 1980, *15*, 973-994.

Anderson, Ronald W. and Jean-Pierre Danthine, "Cross Hedging," *Journal of Political Economy,* December 1981, *89,* 1182-1196.

Copeland, Thomas E. and J. Fred Weston, *Financial Theory and Corporate Policy,* 2nd. ed., Reading Mass.: Addison Wesley, 1983.

Elton, Edwin J. and Martin J. Gruber, *Modern Portfolio Theory and Investment Analysis,* 3rd ed., New York: Wiley, 1987.

Goldberger, Arthur S., *Econometric Theory*, New York: Wiley, 1964

Johnston, J., *Econometric Methods, 2nd. ed.,* New York: McGraw-Hill, 1972.

Mossin, Jan, *Theory of Financial Markets*, Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Stevens, Guy V.G., "On the Inverse of the Covariance Matrix in Portfolio Analysis," *International Finance Discussion Paper #528,* Board of Governors of the Federal Reserve System, Oct. 1995.

Theil, Henri, *Principles of Econometrics,* New York: Wiley, 1971.

Tobin, James, "Liquidity Preference as Behavior Towards Risk," *The Review of Economic Studies,* February 1958, *25,* 65-86.