Orthogonal Frontiers and Alternative Mean-Variance Efficiency Tests
Author(s): Bruce N. Lehmann
Source: *The Journal of Finance*, Vol. 42, No. 3, Papers and Proceedings of the Forty-Fifth Annual Meeting of the American Finance Association, New Orleans, Louisiana, December 28-30, 1986 (Jul., 1987), pp. 601-619
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328373
Accessed: 26/11/2009 08:10

THE JOURNAL OF FINANCE • VOL. XLII, NO. 3 • JULY 1987

Orthogonal Frontiers and Alternative Mean-Variance Efficiency Tests

BRUCE N. LEHMANN*

ABSTRACT

This paper catalogues properties of minimum norm orthogonal portfolios: portfolios which minimize a quadratic objective function and have returns uncorrelated with those of a candidate portfolio that is not mean-variance efficient. The analysis shows that the dollar versions of these portfolios correspond to estimators of zero beta rates based on alternative statistical criteria and grouping procedures while costless orthogonal portfolios represent candidate mean-variance efficiency tests. It also develops inference procedures for zero and unit net investment portfolios of individual securities (instead of grouped portfolios) that have zero expected betas. The resulting mean-variance efficiency tests are reasonably insensitive to the underlying statistical assumptions.

I. Introduction

MOST MEAN-VARIANCE efficiency tests follow from the following simple intuition. If a given portfolio p is mean-variance efficient, all portfolios that cost a dollar and have returns uncorrelated with those of p (i.e., zero beta portfolios) have identical expected returns. Hence, all costless zero beta portfolios have expected returns of zero when p is mean-variance efficient. However, the situation is altered considerably when p is not a mean-variance efficient portfolio. In this case, Roll (1980) has shown that there are portfolios that cost a dollar and have returns uncorrelated with those of p at all levels of mean return. As a consequence, there are costless zero beta portfolios at all levels of expected return as well when p is inefficient. Mean-variance efficiency tests typically involve asking whether particular costless zero beta portfolios have zero expected returns.

The translation of this intuition into operational test statistics is a problem in sampling theory. For example, when tests are based on costless portfolios with zero sample betas, appropriate allowance must be made for the impact of sampling error (in the sample betas) on the distribution of its mean return. Such analysis is extremely complicated when the number of assets included in such portfolios exceeds the number of time series observations on asset returns. Hence, previous investigations have restricted their attention to the special case where the number of assets is taken to be smaller than the number of time series observations.

An enormous amount has been learned about both the statistical properties and economic interpretation of two such tests in these circumstances: Shanken's (1985) CSR test and the related likelihood ratio test of Kandel (1984). A partial list of relevant papers includes Roll (1979, 1985), Gibbons (1982), Stambaugh

* Department of Economics and Graduate School of Business, Columbia University and National Bureau of Economic Research.

(1982), Jobson and Korkie (1982), Ross (1983), Kandel (1984), Shanken (1985, 1986), Amsler and Schmidt (1985), Gibbons, Ross, and Shanken (1986), and MacKinlay (1986). These papers provide both Monte Carlo evidence and analytical bounds on the finite sample distributions of these test statistics as well as economic interpretations of their magnitudes in terms of the geometry of the mean-variance efficient set. For a given set of assets, these methods have the desirable property that they provide uniformly most powerful tests (at least, asymptotically) against the unspecified alternative hypothesis of inefficiency.

The practical implementation of these asymptotically equivalent tests requires the selection of an asset menu that is short relative to the available number of time series observations. The most commonly employed solution is to group the available assets into a smaller number of portfolios. For example, since Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973), it has been common practice to form between ten and forty portfolios from the universe available on the CRSP tapes, often using the sample beta from a previous period as the sorting criterion to improve the precision of the estimated zero beta rate.

As Roll (1979) has emphasized, this strategy is likely to drastically reduce the power of the resulting test unless the grouping variable happens to be related to a relevant alternative hypothesis. "Measurement errors are *the* problem in testing a portfolio's efficiency via the linearity relation" and the usual "portfolio grouping procedure not only tends to remove sampling error but it also tends to remove α, the only measure of portfolio p's inefficiency" (pp. 141, 143). Securities can be grouped by beta to enhance the efficiency of estimated risk premia and mitigate measurement error in sample betas but the resulting intercepts would tend to be the same across groups unless they happened to be related to the magnitudes of the betas. Of course, if securities could be grouped both to retain estimation efficiency under the null hypothesis and power under the alternative, grouping need not be pernicious but such an outcome would require either remarkable prescience or good fortune.

This paper provides some analytical tools which are useful for the development of alternative mean-variance efficiency tests and procedures for estimating risk premia which do not require grouping the available universe of securities into a much smaller number of portfolios. It is well-known that the mean-variance efficient frontier and the locus of minimum variance portfolios with returns uncorrelated with those of an inefficient portfolio are both parabolas in mean-variance space. It is not as widely appreciated that the solution of arbitrary quadratic programming portfolio problems typically yield solutions that are parabolas in mean-variance space as well. The tools developed below involve the characterization of different parabolas in mean-variance space. In addition, inference procedures are provided for sample versions of some of these parabolas.

The layout of the paper is as follows. The next section discusses a number of population orthogonal (i.e., zero beta) frontiers in mean-variance space and describes their relations to commonly employed statistical procedures. The emphasis is on characterization so that more questions are raised than are resolved. The subsequent section provides the relevant sampling theory for a class of asymptotically orthogonal frontiers. The final section contains concluding remarks.

II. Alternative Orthogonal Frontiers

Roll (1980) provided a detailed characterization of the locus of minimum variance portfolios with returns uncorrelated with those of a given inefficient portfolio p. He showed that this frontier is a parabola in mean-variance space which inherits many of the mathematical properties of the mean-variance efficient set. The relations among these orthogonal portfolios simplify several aspects of mean-variance analysis: the interpretation of alternative estimators of zero beta rates, the geometry of mean-variance analysis, the relations between the minimum variance orthogonal and efficient frontiers, and the characterization of deviations from the security market line.

This section provides a mathematical characterization of the location in mean-variance space of particular parabolas that, in general, lie inside the minimum variance orthogonal frontier. In particular, these parabolas correspond to portfolios whose returns are uncorrelated with those of a given inefficient portfolio p that solve arbitrary quadratic programming problems and so these are termed minimum norm orthogonal portfolios. The relations among these portfolios are not nearly so nice and useful as those of the minimum variance orthogonal frontier. In fact, minimum norm orthogonal portfolios appear to be useful only for characterizing alternative estimators of zero beta portfolios and for suggesting corresponding mean-variance efficiency tests. Hence, the discussion will involve two kinds of minimum norm orthogonal portfolios: those that cost a dollar and those that are costless.

To make matters concrete, let $\mathbf{E}$ denote the vector of expected returns on N given assets, let V be their associated covariance matrix (assumed to be nonsingular), and let $\boldsymbol{\beta}_p$ be their vector of (population) regression coefficients on the returns of the given inefficient portfolio p (i.e., $\boldsymbol{\beta}_p = \text{cov}(\mathbf{R}_t, R_{pt})/\text{var}(R_{pt})$). Similarly, let E_p denote the expected return of p and let σ_p^2 denote the corresponding variance of its returns. In addition, notation is required to take account of the common practice of grouping individual security returns into a smaller number of portfolios.[1] Hence, W_M is a given $N \times M$ matrix of portfolio weights (i.e. $W_M'\boldsymbol{\iota} = \boldsymbol{\iota}$) of full column rank used to group individual securities into a smaller number of portfolios. When grouping is not employed, the number of portfolios M is set equal to the number of securities N and the matrix of 'portfolio weights' W_M is simply an identity matrix. Finally, the vectors $\boldsymbol{\iota}$ and $\mathbf{0}$ will always denote (suitably conformable) vectors of ones and zeros, respectively, while the vector $\mathbf{e}_i$ will always denote a (suitably conformable) vector of zeros except for a one in the ith position.

[1] Another common practice in applied work is to include measured security characteristics such as dividend yield or the logarithm of market capitalization in the model for expected returns. The inclusion of a given $N \times K$ matrix Δ ($\Delta = (\boldsymbol{\delta}_1 \boldsymbol{\delta}_2 \cdots \boldsymbol{\delta}_K)$) of individual firm characteristics (that might help explain expected security returns when p is not efficient) that is of full column rank (that is linearly independent of the vectors of expected returns, betas, and ones) involves a straightforward extension of the analysis presented below. The two amendments are that all costless and dollar zero beta portfolios need to be interpreted as having weights orthogonal to Δ and that alternative mean-variance efficiency tests can be based on the costless zero beta portfolios that have weights orthogonal to all but the ith characteristic. The relevant details are omitted to conserve space.

The following proposition characterizes the linear zero beta portfolios that have typically been employed in applied work.

PROPOSITION 1. *Let A be a given $N \times N$ symmetric matrix chosen so that the $W_M' A W_M$ is nonsingular. Then the means and variances of the dollar zero beta portfolios which are solutions to the quadratic programming problem:*

$$\min \mathbf{w}_{zA}' A \mathbf{w}_{zA} \quad \text{s.t.} \quad \mathbf{w}_{zA}' \boldsymbol{\iota} = 1; \qquad \mathbf{w}_{zA}' \mathbf{E} = \mathrm{E}_{zA};$$
$$\mathbf{w}_{zA} = W_M \mathbf{w}_{zM}; \quad \text{and} \quad \mathbf{w}_{zA}' \boldsymbol{\beta}_p = 0 \quad (1)$$

lie on a parabola in mean-variance space. Any property of mean-variance efficient portfolios which depends only upon their means, variances, and covariances is also true for these orthogonal portfolios.

Proof. The dollar minimum norm frontier is the solution to:

$$\min \mathbf{w}_{zM}' A_M \mathbf{w}_{zM} \quad \text{s.t.} \quad \mathbf{w}_{zM}' Y_M = \mathbf{y}_{zM}$$

where $A_M = W_M' A W_M$, $Y_M = W_M'(\mathbf{E}\boldsymbol{\iota}\boldsymbol{\beta}_p)$, and $\mathbf{y}_{zM}' = (E_{zM}10)$. The corresponding Lagrangean is:

$$\mathscr{L} = (1/2)\mathbf{w}_{zM}' A_M \mathbf{w}_{zM} + (\mathbf{w}_{zM}' Y_M - \mathbf{y}_{zM}')\boldsymbol{\mu}$$

with solution:

$$\mathbf{w}_{zM} = A_M^{-1} Y_M \boldsymbol{\mu}$$

Since the solution must satisfy the constraints:

$$\mathbf{y}_{zM} = Y_M' \mathbf{w}_{zM} = Y_M A_M^{-1} Y_M \boldsymbol{\mu} \rightarrow \boldsymbol{\mu} = (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{y}_{zM}$$

so that the solution is:

$$\mathbf{w}_{zM} = A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{y}_{zM}$$

Consequently, letting V_M denote $W_M' V W_M$, the variance of portfolio zM is simply:

$$\sigma_{zM}^2 = \mathbf{y}_{zM}' (Y_M' A_M^{-1} Y_M)^{-1} Y_M' A_M^{-1} V_M A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{y}_{zM}$$

which is clearly quadratic in its mean (i.e., $\mathbf{y}_{zM}' = (E_{zM}10)$) so that the locus of minimum norm portfolios is a parabola in mean-variance space. Let G_M be the upper left-hand 2×2 submatrix of

$$(Y_M' A_M^{-1} Y_M)^{-1} Y_M' A_M^{-1} V_M A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1}$$

Then the covariance of any two minimum norm orthogonal portfolios is:

$$\sigma_{zz'} = (E_{zM}1) G_M (E_{z'M}1)'$$

This is the same structure as the relations among efficient portfolios with G_M replacing the inverse of the efficient set 'information matrix' $(\mathbf{E}\boldsymbol{\iota})' V^{-1} (\mathbf{E}\boldsymbol{\iota})$ so that the means, variances, and covariances of the minimum norm orthogonal frontier are related by identical functions (but with different parameters). Q.E.D.

It is worth emphasizing that the mathematical relations among minimum norm orthogonal portfolios are considerably less useful than the corresponding relations among mean-variance efficient or minimum variance orthogonal portfolios. The reason is that the portfolios along the minimum norm orthogonal frontier cannot be readily identified with particular solutions to the quadratic programming problem (1). For example, the minimum norm orthogonal portfolio is not the minimum variance portfolio on the minimum norm orthogonal frontier in general. Only when A bears some obvious relation to V can the link between the programming problem and the minimum norm orthogonal frontier be determined explicitly.

When sample moments replace the population moments $\boldsymbol{\beta}_p$, $\mathbf{E}$, and V, different choices of A and W_M correspond to alternative linear estimators of zero beta portfolios that have been employed in the literature. When individual securities are employed (i.e., W_M is an $N \times N$ identity matrix), various choices of the matrix A correspond to the common weighted least squares estimators. For example, when A is taken to be the identity matrix, the dollar minimum norm orthogonal portfolio provides the ordinary least squares estimator of the zero beta rate. Similarly, A is often taken to be the diagonal matrix D of market model residual variances ($D = \text{Diag}(\Sigma)$; $V = \boldsymbol{\beta}_p \boldsymbol{\beta}_p' \sigma_p^2 + \Sigma$) yielding the weighted least squares zero beta rate. Finally, when the sample covariance matrix of security returns is nonsingular, the full generalized least squares estimator is the minimum variance zero beta portfolio. In this case, the dollar minimum norm orthogonal frontier is the minimum variance orthogonal frontier studied by Roll (1980).

This does not exhaust the list of dollar orthogonal frontiers that have been employed in the literature, primarily because of the omission of alternative grouping procedures. As noted above, the excess of securities over time series observations renders commonplace the practice of grouping the N securities into a much smaller number of portfolios after sorting on the basis of some characteristic such as previous period beta, market capitalization, dividend yield, or own variance and employ different choices for A. For example, Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) selected W_M on the basis of previous period beta and employed the ordinary least squares estimator (i.e., $A = I$) while Shanken (1985) grouped on the basis of firm size and employed the generalized least squares procedure (i.e., $A = V$).[2]

What can be said about the relative locations of the minimum norm orthogonal frontiers corresponding to alternative choices of A and W_M? Clearly, the minimum variance orthogonal frontier of Roll (1980) lies outside all of the other frontiers in general. Similarly, for a given portfolio grouping matrix W_M, the optimal (i.e., minimum variance) choice of the weighting matrix A is V, the covariance matrix of security returns, so long as $W_M' V W_M$ is nonsingular. Unfortunately, the relative positions associated with different weighting matrices

[2] Similarly, some investigators have included other characteristics in their expected return models. Examples include Litzenberger and Ramaswamy (1979) who employed a weighted least squares estimator (with $A = D$) and set Δ (defined in footnote 1) to be the vector of associated dividend yields and Shanken (1985) employed the generalized least squares estimator on size portfolios and set Δ to be the vector of size portfolio ranks.

and grouping procedures cannot be determined analytically in the absence of *a priori* knowledge of the structure of $\mathbf{E}$, V, and $\boldsymbol{\beta}_p$ and their relations to W_M and A. In particular, nothing can be said about the comparative merits of grouping securities into portfolios and the employment of alternative weighting matrices A. This would appear to be an important empirical question.

These frontiers are useful for characterizing alternative minimum norm frontiers of costless zero beta portfolios. As noted in the introduction, if p is a mean-variance efficient portfolio, all dollar zero beta portfolios have the same expected return and so all costless zero beta portfolios have zero expected returns. This need not be the case when p is inefficient so that these costless orthogonal portfolios represent candidate mean-variance efficiency tests.

PROPOSITION 2. *Let A be a given $N \times N$ symmetric matrix chosen so that $W_M' A W_M$ is nonsingular. Then the costless zero beta portfolios which are solutions to the quadratic programming problem:*

$$\min \mathbf{w}_{xA}' A \mathbf{w}_{xA} \quad \text{s.t.} \quad \mathbf{w}_{xA}' \boldsymbol{\iota} = 0; \qquad \mathbf{w}_{xA}' \mathbf{E} = E_{xA};$$

$$\mathbf{w}_{xA} = W_M \mathbf{w}_{xM}; \quad \text{and} \quad \mathbf{w}_{xA}' \boldsymbol{\beta}_p = 0 \qquad (2)$$

lie on a parabola in mean-variance space that passes through the origin and on a line in the positive quadrant of mean-standard deviation space that also passes through the origin. The weights of these costless portfolios are proportional to $A^{-1}\boldsymbol{\alpha}_{z0A}$ where $\boldsymbol{\alpha}_{z0A}$ ($\boldsymbol{\alpha}_{z0A} = \mathbf{E} - \boldsymbol{\iota} E_{z0A} - \boldsymbol{\beta}_p(E_p - E_{z0A})$) is the vector of deviations from the security market line computed with respect to the minimum norm orthogonal portfolio. The weights are also proportional to the difference in the weights of any two portfolios on the minimum norm orthogonal frontier which solves (1).

Proof. The costless minimum norm frontier is the solution to:

$$\min \mathbf{w}_{xM}' A_M \mathbf{w}_{xM} \quad \text{s.t.} \quad \mathbf{w}_{xM}' Y_M = \mathbf{e}_1 E_{xM}$$

The solution can be computed directly from that given for the dollar minimum norm orthogonal frontier with $\mathbf{e}_1 E_{xM}$ replacing $\mathbf{y}_{zM}$ so that:

$$\mathbf{w}_{xM} = A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1 E_{xM}$$

Consequently, the variance of portfolio xM is simply:

$$\sigma_{xM}^2 = E_{xA}^2 \mathbf{e}_1' (Y_M' A_M^{-1} Y_M)^{-1} Y_M' A_M^{-1} V_M A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1$$

which is proportional to its squared mean so that the locus of costless minimum norm portfolios is a parabola passing through the origin. Similarly, its standard deviation is proportional by the scale factor:

$$[\mathbf{e}_1' (Y_M' A_M^{-1} Y_M)^{-1} Y_M' A_M^{-1} V_M A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1]^{1/2}$$

to its mean, establishing its linearity in mean-standard deviation space. The relation:

$$\boldsymbol{\alpha}_{z0A} = Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1 [\mathbf{e}_1' (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1]^{-1}$$

follows directly from the definition of deviations from the security market line and the minimum norm orthogonal portfolio. Similarly, the difference of the portfolio weights of any two portfolios on the orthogonal frontier is:

$$\mathbf{w}_{z1M} - \mathbf{w}_{z2M} = A_M^{-1} Y_M (Y_M' A_M^{-1} Y_M)^{-1} \mathbf{e}_1 (E_{z1M} - E_{z2M})$$

from the weights given in the proof of Proposition 1 and are identical to the weights of portfolio xM for $E_{xM} = (E_{z1M} - E_{z2M})$. Q.E.D.

One costless minimum norm orthogonal portfolio has been studied extensively in the literature on mean variance efficiency tests. For a given choice of portfolio grouping matrix W_M, the costless minimum variance orthogonal frontier is a straight line in mean-standard deviation space with slope $(\boldsymbol{\alpha}_{z0M}' V_M^{-1} \boldsymbol{\alpha}_{z0M})^{1/2}$ where $\boldsymbol{\alpha}_{z0M}$ is the vector of deviations from the security market line computed with respect to the minimum variance zero beta portfolio.[3] This establishes that the costless zero beta portfolio with weights $V_M^{-1} \boldsymbol{\alpha}_{z0M}$ has the largest squared Sharpe ratio of all costless zero beta portfolios constructed from this set of portfolios. Hence, its mean return will have the largest expected t statistic of any such portfolio. Shanken's (1985) CSR test involves estimation of the squared Sharpe ratio of the sample version of this portfolio and employing finite sample bounds or asymptotic distribution theory to ascertain whether this estimate is likely under the null hypothesis of mean-variance efficiency.

Once again, little can be said about the relative locations of alternative costless minimum norm frontiers. Of course, alternative minimum norm frontiers will, in general, lie inside the costless minimum variance frontier and, as noted above, the minimum norm frontier involving the weighting matrix V will yield smaller variance portfolios than other choices of the weighting matrix A for a given portfolio grouping matrix W_M so long as $W_M' V_M W_M$ is nonsingular. Unfortunately, the relative positions of frontiers based on alternative grouping strategies and weighting matrices cannot be determined analytically and, hence, neither can the associated power functions of the corresponding test statistics even when their sampling distributions are tractable. This, too, would appear to be an important empirical question.

III. Inference for a Class of Sample Frontiers

The preceding discussion pointed to strategies for constructing dollar and costless zero beta portfolios based on alternative programming problems and grouping strategies. However, it did not suggest procedures for performing inference based on sample versions of these portfolios. Shanken (1983, 1985) provided a detailed characterization of the relevant asymptotic theory (and argued convincingly for

[3] Its portfolio weights $\mathbf{w}_{z0M}$, mean return E_{z0M}, and return variance σ_{z0M}^2 are given by:

$$\mathbf{w}_{z0M} = [\sigma_{oM}^2/(\sigma_{oM}^2 - \sigma_{pM^*}^2)]\mathbf{w}_{pM^*} - [\sigma_{pM^*}^2/(\sigma_{oM}^2 - \sigma_{pM^*}^2)]\mathbf{w}_{oM}$$

$$E_{z0M} = (E_{pM^*}\sigma_{oM}^2 - E_{oM}\sigma_{pM^*}^2)/(\sigma_{oM}^2 - \sigma_{pM^*}^2)$$

$$\sigma_{z0M}^2 = \sigma_{oM}^2\sigma_{pM^*}^2/(\sigma_{pM^*}^2 - \sigma_{oM}^2)$$

where $\mathbf{w}_{pM^*}(\mathbf{w}_{pM^*} = V_M^{-1}\boldsymbol{\beta}_{pM}/\boldsymbol{\iota}' V_M^{-1}\boldsymbol{\beta}_{pM})$, E_{pM^*}, and $\sigma_{pM^*}^2$ are the portfolio weights, mean return, and return variance of the portfolio of these N assets with maximum correlation with portfolio p and $\mathbf{w}_{oM}$, E_{oM}, and σ_{oM}^2 are the corresponding numbers for the minimum variance portfolio of the given grouping of these N assets.

its relevance in finite samples) for independently and identically distributed returns when the number of assets was less than the number of time series obervations.[4] His analysis covered both the sample dollar and costless minimum variance zero beta portfolios as well as the sample dollar minimum norm orthogonal portfolio and can be extended in a straightforward fashion to accommodate sample costless minimum norm portfolios.

There is no analysis in the literature of the empirically relevant case where the number of securities exceeds the number of available time series observations in the absence of *a priori* restrictions on the structure of V.[5] It appears to be unlikely that the brute force approach of replacing the relevant population moments with the corresponding sample estimates will yield tractable and close approximations to the sampling distributions of the sample orthogonal frontiers. The basic reason is that the sampling errors of the portfolio weight vectors do not vanish as the number of assets grows and their sample moments do not have tractable finite sample distributions.[6]

This section considers a second strategy: the employment of sample frontiers which are, to some extent, purged of the effects of sampling error in the estimates of the betas. The portfolios associated with these frontiers solve programming problems which are different from those in the previous section in order to facilitate statistical inference. It proves natural to compute the portfolio weights of such portfolios in one period and their returns in a subsequent period to avoid complications in the relevant sampling distributions. This procedure is shown to yield simple inference procedures which are robust with respect to the statistical assumptions underlying the construction of the portfolio weights.[7] Throughout the discussion, security returns are assumed to be independently (but not nec-

[4] The basic idea can be easily illustrated for the costless minimum variance portfolio. The sample version of this portfolio has weights $\hat{\Sigma}^{-1}\hat{\boldsymbol{\alpha}}_{z0}$ (where $\hat{\boldsymbol{\alpha}}_{z0} = \hat{\mathbf{E}} - \iota\hat{E}_{z0} - \hat{\boldsymbol{\beta}}_p(\hat{E}_p - \hat{E}_{z0})$) where $\hat{\Sigma}$ is the sample market model residual covariance matrix, $\hat{\mathbf{E}}$ is the sample mean return vector, $\hat{\boldsymbol{\beta}}_p$ is the vector of ordinary least squares betas, and $\hat{E}_{z0}$ is the generalized least squares estimate of the minimum variance zero beta rate derived from these estimates. Since $\hat{E}_{z0}$ is a consistent estimate, $\hat{\boldsymbol{\alpha}}_{z0}$ 'behaves like' a vector of intercepts from simple regressions in large samples with mean vector zero (under the null hypothesis of efficiency) and covariance matrix $[1 + (E_p - E_{z0})^2/\sigma_p^2]\Sigma$. Hence, Shanken's CSR test statistic $T\hat{\boldsymbol{\alpha}}_{z0}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\alpha}}_{z0}[1 + (\hat{E}_p - \hat{E}_{z0})^2/\hat{\sigma}_p^2]^{-1}$ is asymptotically distributed as $T^2(N-1, T-2)$.

[5] See Shanken (1983) for an analysis involving restrictions on the structure of V.

[6] Such an analysis can be constructed in principle. For example, let $\hat{\boldsymbol{\alpha}}_{z0A}$ be any consistent estimator of the minimum norm vector of deviations from the security market line associated with a given nonstochastic choice of A such that the quantity $\sqrt{T}\hat{\boldsymbol{\alpha}}_{z0A}$ is asymptotically normally distributed with mean vector zero (under the null hypothesis of efficiency) and covariance matrix Ω. Let $\hat{S}$ be a $\sqrt{T}$ consistent estimator for Ω. Then the quantity $T[\hat{\boldsymbol{\alpha}}_{z0A}A^{-1}\hat{\boldsymbol{\alpha}}_{z0A} - trA^{-1}\hat{S}]/\sqrt{2trA^{-2}\hat{S}^2}$ is asymptotically (as T grows indefinitely) normally distributed with mean zero and variance one. It may be possible to improve on this (probably poor) asymptotic approximation when S is a sample covariance matrix with a Wishart distribution (such as a market model residual covariance matrix under normality) because the first two moments of $\hat{S}$ do not depend on the number of assets N. However, I do not think that effective improvements along these lines will be easy to obtain. Moreover, I know of no strategy for doing inference for the more relevant case of fixed T and N growing without bound without restrictions on the structure of V.

[7] Empirical examples in this vein include Litzenberger and Ramaswamy (1979) and Lehmann (1986a, 1986b). These papers do not provide the kind of rigorous statistical analysis presented below.

essarily normally or identically) distributed over time with finite second moments and constant (unconditional) means and betas.[8]

The analysis presumes a given set of estimators for the relevant population moments. By analogy with the preceding section, let $\hat{\mathbf{E}}_t$ denote the vector of estimated mean returns on the N given assets, let $\hat{V}_t$ be the associated covariance matrix estimate (which is not assumed to be nonsingular), and let $\hat{\boldsymbol{\beta}}_{pt}$ be an estimate of the vector of regression coefficients on the returns of the given inefficient portfolio p. Similarly, let $\hat{E}_{pt}$ denote the estimated expected return of p and let $\hat{\sigma}^2_{pt}$ denote the corresponding estimate of the variance of its returns. Finally, the sample weighting matrices are denoted as $\hat{A}_t$ to indicate their possible dependence on parameter estimates. The dependence of all of these parameters on t indicates that these estimates are assumed to be based on returns up to time t (i.e., on $\mathbf{R}_s$ and R_{ps}, $s \leq t$). The remaining notation is identical to that of Section 2.

It is worth emphasizing that no particular estimators for the parameters are being postulated *a priori*. They may be sample moments, the restricted maximum likelihood estimates arising from some postulated asset pricing model, or adjusted estimates yielded by some measurement error correction or empirical Bayes scheme. Similarly, the weighting matrices can depend on parameter estimates and sample moments as well. Hence, the results presented below cover a wide range of estimation strategies.[9]

The following proposition characterizes the class of costless and dollar portfolios for which inference procedures are developed. These are, on average, orthogonal portfolios in that they have zero expected betas (i.e., averaging over repeated, sufficiently independent computations of these weights) and have weights that are distributed independently of subsequent returns under the independence assumption.

PROPOSITION 3. *Consider the dollar portfolios which are solutions to the quadratic programming problem:*

$$\min \mathbf{w}'_{zAt}\hat{A}_{zt}\mathbf{w}_{zAt} \quad \text{s.t.} \quad \mathbf{w}'_{zAt}\boldsymbol{\iota} = 1;$$
$$\mathbf{w}'_{zAt}\hat{\mathbf{E}}_t = E_{zA}; \quad \text{and} \quad \mathbf{w}'_{zAt}\hat{\boldsymbol{\beta}}_{pt} = \hat{\delta}_{zAt} \tag{3}$$

for given nonsingular symmetric $\hat{A}_{zt}$ and the corresponding costless portfolios which are solutions to the quadratic programming problem:

$$\min \mathbf{w}'_{xAt}\hat{A}_{xt}\mathbf{w}_{xAt} \quad \text{s.t.} \quad \mathbf{w}'_{xAt}\iota = 0; \qquad \mathbf{w}'_{xAt}\hat{\mathbf{E}}_t = E_{xA}; \qquad \mathbf{w}'_{xAt}\hat{\beta}_{pt} = \hat{\delta}_{xt}E_{xA} \tag{4}$$

[8] All of the asymptotic results obtained below can be obtained (under the independence assumption) with nonconstant unconditional means and betas and weaker second moment conditions so long as the time series averages of these means and betas and normalized sums of squares and cross-products of returns converge to finite limits. See White (1984) for alternative combinations of restrictions that will yield these results at the expense of unnecessarily complicating the proofs. However, weakening the independence assumption (beyond an alternative martingale difference model would substantially alter the analysis.

[9] Some of the results obtained in Propositions 3 and 5 do explicitly require that means, variances, and covariances among the parameter estimates exist and are finite. These moment restrictions can be replaced by weaker assumptions or by a corresponding analysis of their probability limits.

for given nonsingular symmetric $\hat{A}_{xt}$ (where the weighting matrices $\hat{A}_{zt}$ and $\hat{A}_{xt}$ need not be the same). Then the returns of these portfolios for $s > t$ are distributed independently of the estimated portfolio weights so long as the values of $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ are independent of returns for $s > t$. When $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ are set equal to zero, these are simply zero sample beta portfolios. Moreover, when $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ reflect the sampling error moments:

$$\hat{\delta}_{zAt} = \frac{E\{(\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)'\hat{A}_{zt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{zt}^{-1}\hat{Y}_t)^{-1}(\mathbf{e}_1 E_{zA} + \mathbf{e}_2)\}}{[1 - E\{(\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)'\hat{A}_{zt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{zt}^{-1}\hat{Y}_t)^{-1}\mathbf{e}_3\}]}$$

$$\hat{\delta}_{xt} = \frac{E\{(\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)'\hat{A}_{xt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{xt}^{-1}\hat{Y}_t)^{-1}\mathbf{e}_1\}}{[1 - E\{(\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)'\hat{A}_{xt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{xt}^{-1}\hat{Y}_t)^{-1}\mathbf{e}_3\}]}$$

$$\hat{Y}_t = (\hat{\mathbf{E}}_t \iota \hat{\boldsymbol{\beta}}_{pt}) \tag{5}$$

(assumed to exist and be finite) then the portfolios have zero unconditional (i.e., expected) betas.[10]

Proof. Let $\hat{\mathbf{y}}_{zAt}' = (E_{zA} 1 \hat{\delta}_{zAt})$ and let $\hat{\mathbf{y}}_{xAt}' = (E_{xA} 0 \hat{\delta}_{xt} E_{xA})$. From the proof of Proposition 1, the corresponding portfolio weights are given by:

$$\mathbf{w}_{zAt} = \hat{A}_{zt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{zt}^{-1}\hat{Y}_t)^{-1}\hat{\mathbf{y}}_{zAt}$$

$$\mathbf{w}_{xAt} = \hat{A}_{xt}^{-1}\hat{Y}_t(\hat{Y}_t'\hat{A}_{xt}^{-1}\hat{Y}_t)^{-1}\hat{\mathbf{y}}_{xAt}$$

which are clearly distributed independently of subsequent security returns due to their assumed independence. Their unconditional betas are given by:

$$\beta_{zAt} = \text{cov}\{\mathbf{w}_{zAt}'\mathbf{R}_s, R_{ps}\}/V(R_{ps}) = E\{\mathbf{w}_{zAt}'\boldsymbol{\beta}_p\} = E\{\mathbf{w}_{zAt}'[\hat{\boldsymbol{\beta}}_{pt} - (\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)]\}$$
$$= \hat{\delta}_{zAt} - \hat{\delta}_{zAt} = 0$$

$$\beta_{xAt} = \text{cov}\{\mathbf{w}_{xAt}'\mathbf{R}_s, R_{ps}\}/V(R_{ps}) = E\{\mathbf{w}_{xAt}'\boldsymbol{\beta}_p\} = E\{\mathbf{w}_{xAt}'[\hat{\boldsymbol{\beta}}_{pt} - (\hat{\boldsymbol{\beta}}_{pt} - \boldsymbol{\beta}_p)]\}$$
$$= \hat{\delta}_{xt}E_{xA} - \hat{\delta}_{xt}E_{xA} = 0$$

due to the definition of the measurement error corrections $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$. Q.E.D.

There are two aspects of Proposition 3 that are useful for the development of sampling theory for this class of portfolios. The first is that their weights are independent of subsequent returns which suggests the possibility of constructing inference procedures utilizing subsequent returns on such portfolios. The second is that their expected betas are zero (given appropriate sampling error corrections) so that portfolio returns based on sequences of computed portfolio weights reflect (on average) the returns of orthogonal portfolios over sufficiently long time periods given sufficient independence in the sequence of portfolio weights. Proposition 4 utilizes the first feature and Proposition 5 exploits both insights.

Suppose that security returns are normally, independently, and identically distributed. In addition, suppose that portfolio weights $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$ are con-

[10] Both the population and sample frontiers of the costless portfolios lie on parabolas in mean-variance space passing through the origin and on lines in the positive quadrant of mean-standard deviation space that also pass through the origin, as is clear from the proof. The corresponding dollar frontiers do not have this convenient representation.

structed along the lines of Proposition 3 and the returns on these portfolios are computed over a subsequent sample of T observations. These portfolio returns will then be normally, identically, and independently distributed as well with means E_{zAt} and E_{xAt} and betas β_{zAt} and β_{xAt}, respectively (i.e., $E_{zAt} = \mathbf{w}'_{zAt}\mathbf{E}$). As is well-known, exact finite sample inference procedures are available for the sample moments of these portfolio returns in this circumstance. A number of these are collected in Proposition 4 for the pair of portfolios zA and xA.

PROPOSITION 4. *Let the values of $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ used to construct the portfolio weights $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$ be distributed independently of $\mathbf{R}_s$ $(s > t)$. Let $\{\mathbf{R}_s; t < s \leq T + t + 1\}$ be a sample of T observations on the security returns which are assumed to be normally, independently, and identically distributed. Let $\hat{\mathbf{E}}_T$, $\hat{V}_T$, $\hat{\Sigma}_T$, $\hat{\boldsymbol{\beta}}_{pT}$, and $\hat{\sigma}^2_{pT}$ denote the associated sample mean vector, covariance matrix, market model residual covariance matrix, ordinary least squares betas, and sample variance of p, respectively. Similarly, let $\hat{E}_{zAT}$, $\hat{E}_{xAT}$, $\hat{\sigma}^2_{zAT}$, and $\hat{\sigma}^2_{xAT}$ denote the sample means and variance estimates of the returns on $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$. Finally, let $\hat{\beta}_{zAT}$ and $\hat{\beta}_{xAT}$ denote the ordinary least squares betas of these two portfolios and let $\hat{\sigma}^2_{\beta zAT}$ and $\hat{\sigma}^2_{\beta xAT}$ be their corresponding sample variances (i.e., $\hat{\sigma}^2_{\beta zAT} = \mathbf{w}'_{zAt}\hat{\Sigma}_T\mathbf{w}_{zAt}/T\hat{\sigma}^2_{pT}$). Then all of the following quantities have t distributions*:

$$\sqrt{T}(\hat{E}_{zAT} - E_{zAt})/\sqrt{\hat{\sigma}^2_{zAT}} \sim t(T-1)$$

$$\sqrt{T}(\hat{E}_{xAT} - E_{xAt})/\sqrt{\hat{\sigma}^2_{xAT}} \sim t(T-1)$$

$$(\hat{\beta}_{zAT} - \beta_{zAt})/\sqrt{\hat{\sigma}^2_{\beta zAT}} \sim t(T-2)$$

$$(\hat{\beta}_{xAT} - \beta_{xAt})/\sqrt{\hat{\sigma}^2_{\beta xAT}} \sim t(T-2)$$

and are distributed independently of $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$. If, in addition, β_{xAt} is zero (i.e., conditional on $\mathbf{w}_{xAt}$), E_{xAt} is equal to zero under the null hypothesis of mean-variance efficiency and can be tested with the usual t statistic for the significance of $\hat{E}_{xAT}$. The joint hypothesis that both E_{xAt} and β_{xAt} are zero can be tested with the usual F test:

$$(T-2)[\hat{E}^2_{xAT} + \hat{\beta}^2_{xAT}\hat{\sigma}^2_{pT}]/2\mathbf{w}'_{xAt}\hat{\Sigma}_T\mathbf{w}_{xAt} \sim F(2, T-2)$$

so that inferences need not be conditional on the assumption that β_{xAt} is zero. Finally, if β_{xAt} is nonzero, an asymptotic Shanken (1985) T^2 test based on portfolios zA and xA can be constructed. Hence, the test statistic:

$$\frac{\sqrt{T}[\hat{\boldsymbol{\beta}}_{xAT}(\hat{E}_{zAT} - \hat{E}_{pT}) + (1 - \hat{\boldsymbol{\beta}}_{zAT})\hat{E}_{xAT}]}{\sqrt{\begin{array}{l}\hat{\beta}^2_{xAT}(\hat{\sigma}^2_{zAT} - \hat{\sigma}^2_{pT}) + 2(1 - \hat{\beta}_{zAT})\hat{\beta}_{xAT}\hat{\sigma}_{xzAT} \\ \quad + (1 - \hat{\beta}_{zAT})^2\hat{\sigma}^2_{xAT}\left[1 + \dfrac{(\hat{E}_{pt} - \hat{E}^*_{zT})^2}{\hat{\sigma}^2_{pT}}\right]\end{array}}} \xrightarrow{D} t(T-2)$$

under the null hypothesis of efficiency where $\xrightarrow{D}$ denotes convergence in distribution and $\hat{E}^*_{zT}$ is the GLS estimate of the 0 beta rate.

Proof. The proof is given only for portfolio zA since the corresponding proof for xA is identical. Since $\hat{\mathbf{E}}_T \sim N(\mathbf{E}, V/T)$ independently of $\mathbf{w}_{zAt}$ (by Proposition

3) the quantity:

$$\mathbf{w}_{zAt}'\hat{\mathbf{E}}_T \sim N(\mathbf{w}_{zAt}'\mathbf{E},\ \mathbf{w}_{zAt}'V\mathbf{w}_{zAt}/T)$$

independently of $\mathbf{w}_{zAt}$ as well. Similarly, since $\mathbf{w}_{zAt}'V\mathbf{w}_{zAt}$ is nonzero with probability one, the quantity:

$$\mathbf{w}_{zAt}'\hat{V}_T\mathbf{w}_{zAt}/\mathbf{w}_{zAt}'V\mathbf{w}_{zAt} \sim \chi^2(\mathrm{T}-1)$$

independently of $\mathbf{w}_{zAt}$. Hence:

$$\sqrt{T}(\hat{E}_{zAT} - E_{zAt})/\sqrt{\hat{\sigma}_{zAT}^2} \sim t(T-1)$$

has an exact t distribution independent of $\mathbf{w}_{zAt}$. Similarly, $\hat{\boldsymbol{\beta}}_{pT} \sim N(\boldsymbol{\beta}_p,\ \Sigma/T\sigma_p^2)$ independently of $\mathbf{w}_{zAt}$, so that:

$$\mathbf{w}_{zAt}'\hat{\boldsymbol{\beta}}_{pT} \sim N(\mathbf{w}_{zAt}'\boldsymbol{\beta}_p,\ \mathbf{w}_{zAt}'\Sigma\mathbf{w}_{zAt}/T\sigma_p^2)$$

independently of $\mathbf{w}_{zAt}$ as well. Similarly, since $\mathbf{w}_{zAt}'\Sigma\mathbf{w}_{zAt}$ is nonzero with probability one, the market model residual variance of portfolio zA satisfies:

$$\mathbf{w}_{zAt}'\hat{\Sigma}_T\mathbf{w}_{zAt}/\mathbf{w}_{zAt}'\Sigma\mathbf{w}_{zAt} \sim \chi^2(T-2)$$

independently of $\mathbf{w}_{zAt}$. Hence:

$$(\hat{\beta}_{zAT} - \beta_{zAt})/\sqrt{\hat{\sigma}_{\beta zAT}^2} \sim t(T-2)$$

has an exact t distribution independent of $\mathbf{w}_{zAt}$.

When β_{xAt} is zero, the expected return E_{zAt} is zero under the null hypothesis of efficiency, so that the associated t statistic for $\hat{E}_{zAT}$ constitutes an appropriate mean variance efficiency test. The F statistic for the joint hypothesis that both E_{zAt} and β_{xAt} are zero given in the proposition follow from simple algebraic manipulation of the usual F statistic that the intercept and slope are both zero in the market model regression of the returns on xA on those of p. In this simple two asset case, the square root of Shanken's (1985) CSR T^2 test statistic is proportional to the t statistic for the hypothesis that the mean return on the costless portfolio that is long one dollar of zA, short one dollar of p, and levered up by the ratio $\hat{\lambda}_{\beta T} = (1 - \hat{\beta}_{zAT})/\hat{\beta}_{xAT}$ of xA is zero. Hence, by Shanken's analysis:

$$\frac{\sqrt{T}[(\hat{E}_{zAT} - \hat{E}_{pT}) + \hat{\lambda}_{\beta T}\hat{E}_{xAT}]}{\sqrt{[\hat{\sigma}_{zAT}^2 - \hat{\sigma}_{pT}^2 + 2\hat{\lambda}_{\beta T}\hat{\sigma}_{xzAT} + \hat{\lambda}_{\beta T}^2\hat{\sigma}_{xAT}^2]\left[1 + \dfrac{(\hat{E}_{pt} - \hat{E}_{zT}^*)^2}{\hat{\sigma}_{pT}^2}\right]}} \xrightarrow{D} t(T-2)$$

asymptotically, after some algebraic manipulation. The statistic given in the proposition is obtained by multiplying the numerator and denominator of this statistic by $\hat{\beta}_{xAT}$ (which is asymptotically valid due to the consistency of $\hat{\beta}_{xAT}$) to include the relevant case where β_{xAt} is equal to zero. Q.E.D.

The most noteworthy feature of Proposition 4 is the robustness it lends to inference procedures regarding the efficiency of portfolio p. If an investigator is confident that the procedure used to construct the portfolio weights $\mathbf{w}_{xAt}$ yielded a zero beta portfolio, inferences can be based on a simple t test. Uncertainty about the validity of this presumption can be accounted for in an F test for the joint hypothesis that both the mean and beta of this portfolio are equal to zero. Failure to reject this joint hypothesis would constitute evidence in favor of the efficiency of p while rejection could follow either from the inefficiency of p or when portfolio xA has a nonzero beta. The last two possibilities can be distinguished using a Shanken (1985) CSR T^2 test for the efficiency of p where the assets underlying the test are portfolios zA and xA. Hence, inferences regarding the efficiency of p need not rest on assumptions regarding the statistical properties of the estimated portfolio weights.

Note also that these results allow considerable flexibility in the implementation of this sort of mean-variance efficiency test. As emphasized above, a wide variety of choices for the expected return and beta estimators as well as sample dependent weighting matrices and measurement error corrections $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ can be accommodated in this framework. Similarly, there was no requirement that portfolio zA and xA be derived from similar quadratic programming problems. In fact, these results can be adapted in a straightforward fashion to the case of many imperfectly correlated dollar and costless portfolios, each associated with possibly different quadratic programming problems.

The portfolio strategy underlying Proposition 4 has both advantages and disadvantages. The principal advantage is the availability of simple finite sample inference procedures (under the maintained assumption of the normality of returns).[11] The main disadvantage is that it does not exploit the other implication of Proposition 3: the expected betas of portfolios constructed along the lines of zA and xA are zero (given adequate measurement error corrections). This observation can be utilized by altering the portfolio strategy implicit in Proposition 4.

The basic idea is to revise the portfolio weights $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$ as new returns arrive. In particular, suppose that, at each time t, the portfolio weights $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$ are constructed on the basis of returns from the preceding L periods (i.e., the estimates $\hat{\mathbf{E}}_t$, $\hat{V}_t$, $\hat{\boldsymbol{\beta}}_{pt}$, $\hat{E}_{pt}$, and $\hat{\sigma}^2_{pt}$ as well as the weighting matrices $\hat{A}_{zt}$ and $\hat{A}_{xt}$ and measurement error corrections $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ depend on $\mathbf{R}_s$ and R_{ps} for $t - L \leq s \leq t$). In addition, suppose that the constructed portfolio weights have finite, but possibly time varying, first and second moments. These assumptions (coupled with the independence of returns) are sufficient to insure that both the average portfolio weight vector and normalized sums of squared portfolio weights converge to well-defined limits.[12]

Inferences can be based on the returns of this more active portfolio strategy:

$$
\begin{aligned}
R_{zAt+1} &= \mathbf{w}'_{zAt}\mathbf{R}_{t+1} \\
R_{xAt+1} &= \mathbf{w}'_{xAt}\mathbf{R}_{t+1} \qquad (6)
\end{aligned}
$$

[11] The proof of Proposition 5 can be adapted in a straightforward fashion to prove Proposition 4 for the case of independently but not normally or identically distributed returns with all of the t distributions obtained in Proposition 4 replaced by asymptotic unit normal distributions while two times the F statistic has an asymptotic $\chi^2(2)$ distribution.

[12] Once again, all that is required for the analysis is that normalized sums and sums of squares of portfolio weights converge to appropriate finite limits. See White (1984) for alternative assumptions that yield these results.

because their unconditional means and betas are given by:

$$E_{zA} = E\{\mathbf{w}'_{zAt}\mathbf{R}_{t+1}\} = \mathbf{w}'_{zA}\mathbf{E}; \quad \beta_{zA} = E\{\mathbf{w}'_{zAt}\boldsymbol{\beta}_p\} = \mathbf{w}'_{zA}\boldsymbol{\beta}_p$$

$$E_{xA} = E\{\mathbf{w}'_{xAt}\mathbf{R}_{t+1}\} = \mathbf{w}'_{xA}\mathbf{E}; \quad \beta_{xA} = E\{\mathbf{w}'_{xAt}\boldsymbol{\beta}_p\} = \mathbf{w}'_{xA}\boldsymbol{\beta}_p \tag{7}$$

where $\mathbf{w}_{zA}$ and $\mathbf{w}_{xA}$ are the limits of the sums $\sum_{t=1}^{T} E\{\mathbf{w}_{zAt}\}/T$ and $\sum_{t=1}^{T} E\{\mathbf{w}_{xAt}\}/T$, respectively. Given appropriate estimators of the measurement error corrections δ_{zAt} and δ_{xt}, the betas of these portfolios should converge to their unconditional means of zero in large samples. Hence, the mean return on portfolio xA should converge to its unconditional mean of zero if portfolio p is mean-variance efficient.

Inference procedures are more complicated here since the returns on these portfolios will be both serially correlated and heteroskedastic even if returns are independently and identically distributed (an assumption not made below). This occurs in this setting for the same reason that it occurs in the analysis of value-weighted portfolios. Returns are serially correlated because the evolution of weights is, to some extent, predictable and, hence, there is predictable variation in mean returns. Fortunately, this serial correlation does not affect inferences regarding sample moments such as means and betas in this setting. This occurs because deviations from means (and mean betas) are unpredictable given the independence of returns and this is what is relevant for inference in large samples. The variances of the portfolio returns will not be (conditonally) constant due to the variation in portfolio weights so that the problem of heteroskedasticity remains.

Fortunately, it is possible to take account of conditional heteroskedasticity with the methods of Hansen (1982) and White (1984). Let $\hat{\gamma}_{zAT}$ and $\hat{\gamma}_{xAT}$ and let $\hat{\beta}_{zAT}$ and $\hat{\beta}_{xAT}$ denote the ordinary least squares market model intercepts and betas, respectively, of portfolios zA and xA, respectively, computed over a sample of T observations. Let X_T denote the market model regressor matrix whose sth row $\mathbf{x}'_s$ equals $(1 R_{ps})$. Finally, let $\hat{\epsilon}_{zAs}$ and $\hat{\epsilon}_{xAs}$ denote the sth market model residual for portfolios zA and xA, respectively, and let $\hat{\boldsymbol{\epsilon}}'_s$ be the vector $(\hat{\epsilon}_{zAs}\ \hat{\epsilon}_{xAs})$. Hansen (1982) and White (1984) have shown that the vector $(\hat{\gamma}_{zAT}\,\hat{\beta}_{zAT}\,\hat{\gamma}_{xAT}\,\hat{\beta}_{xAT})$ is then asymptotically normally distributed with covariance matrix $\hat{\Omega}$ of the form HSH with H equal to $[I_2 \otimes (X'_T X_T)^{-1}]$ and S equal to $[\sum_{s=1}^{T} (\hat{\boldsymbol{\epsilon}}_s\hat{\boldsymbol{\epsilon}}'_s)\otimes(\mathbf{x}_s\mathbf{x}'_s)]$. For later reference, let $\hat{\Omega}_{zAT}$ denote the upper 2×2 diagonal block (i.e., the asymptotic covariance matrix of $(\hat{\gamma}_{zAT}\hat{\beta}_{zAT})$) and let $\hat{\Omega}_{xAT}$ denote the corresponding lower 2×2 diagonal block (i.e., the asymptotic covariance matrix of $(\hat{\gamma}_{xAT}\hat{\beta}_{xAT})$) of $\hat{\Omega}$. These observations make it possible to construct the asymptotic inference procedures given in Proposition 5.

PROPOSITION 5. *Let the choices of $\hat{\delta}_{zAT}$ and $\hat{\delta}_{xt}$ used to construct the portfolio weights $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$ both be distributed independently of $\mathbf{R}_s$ $(s > t)$. Suppose the sequence of portfolio weight vectors satisfy the assumptions given in the text (i.e., finite memory and finite, but possibly time varying, first and second moments). Let $\{\mathbf{R}_t;\ 1 \le t \le T\}$ be a sample of T observations on the security returns which are assumed to be independently distributed over time with constant (unconditional) means and betas. Let $\hat{E}_{zAT}$, $\hat{E}_{xAT}$, $\hat{\sigma}^2_{zAT}$, and $\hat{\sigma}^2_{xAT}$ denote the corresponding sample means and variances of the returns on portfolio strategies zA and xA over* *these T observations. Finally, let $\hat{\gamma}_{zAT}$ and $\hat{\gamma}_{xAT}$ and let $\hat{\beta}_{zAT}$ and $\hat{\beta}_{xAT}$ denote the ordinary least squares market model intercepts and betas, respectively, of these two portfolios and let $\hat{\sigma}\beta^2_{zAT}$ and $\hat{\sigma}\beta^2_{xAT}$ denote their heteroskedastic consistent sample variances (i.e., the lower right-hand elements of $\hat{\Omega}_{zAT}$ and $\hat{\Omega}_{xAT}$, respectively) computed over these T observations. Then, each of the following random variables converges in distribution to unit normal random variables:*

$$\sqrt{T}(\hat{E}_{zAT} - E_{zA})/\sqrt{\hat{\sigma}^2_{zAT}} \xrightarrow{D} N(0, 1)$$

$$\sqrt{T}(\hat{E}_{xAT} - E_{xA})/\sqrt{\hat{\sigma}^2_{xAT}} \xrightarrow{D} N(0, 1)$$

$$(\hat{\beta}_{zAT} - \beta_{zA})/\sqrt{\hat{\sigma}^2_{\beta zAT}} \xrightarrow{D} N(0, 1)$$

$$(\hat{\beta}_{xAT} - \beta_{xA})/\sqrt{\hat{\sigma}^2_{\beta xAT}} \xrightarrow{D} N(0, 1)$$

where $\xrightarrow{D}$ denotes convergence in distribution. If, in addition, the choices of $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ used to construct the portfolio weights are unbiased estimates of the corresponding population measurement error corrections that are distributed independently of the parameter estimates underlying $\mathbf{w}_{zAt}$ and $\mathbf{w}_{xAt}$, β_{xA} (as well as β_{zA}) is zero and, hence, E_{xA} is equal to zero under the null hypothesis of mean-variance efficiency and can be tested with the usual asymptotic t statistic for the significance of $\hat{E}_{xAT}$. The joint hypothesis that both E_{xA} and β_{xA} are zero can be tested with the asymptotic chi-squared test:

$$T(\hat{\gamma}_{xAT}\hat{\beta}_{xAT})\hat{\Omega}^{-1}_{xAT}(\hat{\gamma}_{xAT}\hat{\beta}_{xAT})' \xrightarrow{D} \chi^2(2)$$

so that inferences need not be conditional on the assumption that β_{xA} is zero. Finally, if β_{xA} is nonzero, an asymptotic version of Shanken's (1985) *T^2 test based on portfolios zA and xA can be constructed. Hence, the test statistic:*

$$\sqrt{T}[\hat{\beta}_{xAT}\hat{\gamma}_{zAT} + (1 - \hat{\beta}_{zAT})\hat{\gamma}_{xAT}]/\sqrt{\hat{\boldsymbol{\theta}}'\hat{\Omega}\hat{\boldsymbol{\theta}}} \xrightarrow{D} N(0, 1)$$

under the null hypothesis of efficiency where the vector $\hat{\boldsymbol{\theta}}'$ equals

$$(\hat{\beta}_{xAT} - \hat{\gamma}_{xAT}\ (1 - \hat{\beta}_{zAT})\ \hat{\gamma}_{zAT}).$$

Proof. Once again, the proofs will be presented only for portfolio zA since the corresponding proofs for portfolio xA are identical. Let $r_{zt} = \mathbf{w}'_{zAt-1}(\mathbf{R}_t - \mathbf{E})$ be the deviations from means of the increments to $\hat{E}_{zAT}$ and let $q_{zt} = [\mathbf{w}'_{zAt-1}(\mathbf{R}_t - \hat{\mathbf{E}}_T)(R_{pt} - \hat{E}_p)/\hat{\sigma}^2_{pT}] - \beta_{zA}$ where $\hat{E}_{pT}$ and $\hat{\sigma}^2_{pT}$ denote the sample mean and variance of portfolio p over these T observations. Since security returns are independently distributed over time and the portfolio weights are distributed independently of subsequent returns, the quantity $S_T = \sum_{t=1}^T r_{zt}$ is a zero mean martingale. Similarly, the quantity $U_T = \sum_{t=1}^T q_{zt}$ converges uniformly to a zero mean martingale by Kronecker's lemma (cf. Hall and Heyde (1980)) as $\hat{\sigma}^2_{pT}$ converges to its probability limit (intuitively because $q^*_{zt} = [\mathbf{w}'_{zAt-1}(\mathbf{R}_t - \mathbf{E})(R_{pt} - E_p)/\sigma^2_p] - \beta_{zA}$ is a zero mean martingale increment). Since the portfolio weights and security returns were assumed to have finite second moments:

$$E\{S^2_T\}^{-1}\sup r^2_{zs} \rightarrow 0$$

$$E\{U^2_T\}^{-1}\sup q^2_{zs} \rightarrow 0$$

in probability since any squared observation is bounded with probability one. Similarly, the independence of returns and the finite memory and moments of portfolio weights implies that:

$$E\{S_T^2\}^{-1} \sum_{s=1}^{T} r_{zs}^2 \rightarrow 1$$

$$E\{U_T^2\}^{-1} \sum_{s=1}^{T} q_{zs}^2 \rightarrow 1$$

in probability as well. Hence, by the central limit theorem for martingales (cf. Scott (1973) and Hall and Heyde (1980)):

$$E\{S_T^2\}^{-1/2} S_T \xrightarrow{D} N(0, 1)$$

$$E\{U_T^2\}^{-1/2} U_T \xrightarrow{D} N(0, 1)$$

Finally, the convergence of $\hat{\sigma}_{zAT}^2$ to $E\{S_T^2\}/T$ and of $\hat{\sigma}_{\beta zAT}^2$ to $E\{U_T^2\}/T$ establishes the unit normal distributions given in the proposition.

The finite memory and moments of the sequence of portfolio weight vectors ensures that β_{zA} and β_{xA} will converge to their unconditional means of zero (given in Proposition 3) when $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ are unbiased estimates of the corresponding population measurement error corrections that are distributed independently of $\hat{Y}_T$ and subsequent returns. In this case, E_{xA} is equal to zero under the null hypothesis of mean-variance efficiency. Finally, the convergence of the two test statistics follow from the usual asymptotic properties of quadratic forms in asymptotically normal random variables. Q.E.D.

The econometric setting of Proposition 5 corresponds closely to the environments of empirical studies such as Litzenberger and Ramaswamy (1979) and Lehmann (1986a, 1986b).[13] Like Proposition 4, it suggests that inferences about mean-variance efficiency obtained from this econometric practice can be made quite robust (at least asymptotically) with respect to the assumptions underlying the computation of the weights of the portfolios. The use of unbiased measurement error corrections $\hat{\delta}_{zAt}$ and $\hat{\delta}_{xt}$ that are distributed independently of the corresponding parameter estimates insures that, on average, portfolios zA and xA are zero beta portfolios. In these circumstances, inferences about mean-variance efficiency can be obtained from a simple asymptotic t test. Of course, the adequacy of the measurement error corrections can be tested with an asymptotic t test for the beta of portfolio xA. Similarly, the joint hypothesis that both the mean and beta of xA are zero can be tested with an asymptotic chi-squared test and an asymptotic test which does not require these portfolios to have zero betas based on portfolios p, zA, and xA (analogous to Shanken's (1985) CSR test) is available as well. Once again, these procedures would appear to constitute a rich set of asymptotic diagnostic tests and safeguards.

The main disadvantage of these results is their reliance on asymptotic distribution theory. While appropriate Monte Carlo and sample evidence is required

[13] The main difference is that these investigators did not study N given securities but instead permitted securities to enter and exit the sample. The results of Proposition 5 remain valid in these circumstances so long as the sample selection rules underlying this practice are well-approximated by a random sampling scheme for which the required moments exist and which do not lead to correlations between the portfolio weights and subsequent returns.

to determine the conformity of finite sample distributions to their asymptotic approximations, there is reason for optimism in this application, at least relative to the rate of convergence of the multivariate CSR test to its asymptotic chi-squared distribution. As Shanken (1985) has emphasized, the main reason for the slow convergence of the CSR test (as well as the likelihood ratio test) to its asymptotic distribution is that the number of time series observations must be appreciably larger than the number of assets for the sampling variation in the relevant residual covariance matrix to be negligible. That is not an issue in this application since the number of assets in the analysis never exceeded three: portfolios p, zA, and xA.

The primary determinant (besides the number of time series observations) of the rate of convergence of the random variables discussed in Proposition 5 to their asymptotic distributions (as well as the distance of their finite sample distributions from their large sample limits) is the rate at which the sample estimate of the asymptotic variance converges to the population variance of the random variable.[14] In this application, this depends on the rate at which the average squared portfolio weight vector implicit in the portfolio returns as well as the average population covariance matrix of returns converge to their probability limits. It seems reasonable to suppose that the number of time series observations need not be appreciably larger than the number of lagged returns underlying the computation of the portfolio weights to insure sufficiently rapid convergence of these quantities. In the applications cited above, the number of time series observations ran from 500 to 700 while the number of lagged returns used to compute the portfolio weights was 60, a ratio which suggests confidence in the relevance of the asymptotic approximations. Of course, this discussion is merely suggestive and more concrete statements require both sample and simulation evidence.

IV. Conclusion

It is now widely appreciated that the locus of minimum variance zero beta portfolios associated with a given inefficient portfolio is a parabola in mean-variance space as shown by Roll (1980). This observation has provided the basis for interpretations of the relations among alternative estimators of zero beta rates as well as geometric interpretations of mean-variance efficiency tests. The analysis presented above demonstrates that there are numerous parabolas in mean-variance space which correspond to orthogonal portfolios that cost a dollar and solve different quadratic programming problems. The population versions of these minimum norm orthogonal frontiers are not as interesting as the minimum variance orthogonal frontier studied extensively by Roll (1980). However, the sample versions are more interesting in that they characterize the relations among alternative estimators of zero beta rates corresponding to different norms

[14] See Hall and Heyde (1980) for a discussion of the analogues of the Berry-Esseen theorem (their Theorems 3.9 and 3.10) and of the law of the iterated logarithm (their Theorem 4.7) for martingales of the sort underlying Proposition 5. They verify that these bounds depend in a nonlinear fashion on the rate of convergence of the sample asymptotic variance as well as a number of complicated constants.

and grouping schemes. While it is not possible to determine the relative locations of these estimators in mean-variance space, the analysis serves as a reminder that minimum variance orthogonal portfolios associated with a given grouping strategy do not possess obvious optimality properties.

All costless zero beta portfolios associated with a given portfolio have zero expected returns when that portfolio is mean-variance efficient but this will not be the case in general when the candidate portfolio is inefficient. This provides the motivation for studying another class of orthogonal frontiers: the locus of costless zero beta portfolios which minimize a quadratic objective function for given mean return. These orthogonal frontiers are parabolas in mean-variance space that pass through the origin and, hence, are lines in mean-standard deviation space passing through the origin. The costless minimum variance orthogonal frontier is the most interesting population frontier since it yields to the costless zero beta portfolio with the largest squared Sharpe ratio. The feasible sample versions of alternative minimum norm orthogonal frontiers correspond to different possible mean-variance efficiency tests because the population slopes of the corresponding sample lines in mean-standard deviation space are zero under the null hypothesis of efficiency. The inability to determine the relative magnitudes of either the sample or population values of these slopes associated with different grouping strategies and norms serves as a reminder that the mean-variance efficiency tests that examine the costless minimum variance orthogonal portfolios associated with a given grouping procedure do not possess obvious optimal power properties. The quote from Roll (1979) in the introduction suggests that the converse is equally likely.

In fact, the main advantage of the now ubiquitous multivariate tests based on grouped assets is the remarkably detailed analysis of the finite sample properties of these procedures that has appeared in the recent literature. Similarly, the principal barrier to the implementation of alternative mean-variance efficiency tests based on individual assets is the seemingly intractable nature of the sampling theory of sample costless minimum norm orthogonal portfolios because the number of individual assets available greatly exceeds the number of time series observations on their returns. This suggests that the development of alternative tests with good finite sample properties based on individual assets is not likely to proceed from the analysis of such sample portfolios.

The analysis in the preceding section pursues a strategy that was introduced in Litzenberger and Ramaswamy (1979): constructing the portfolio weights associated with a given programming problem in one period and drawing inferences from their subsequent returns. The analysis provides both finite sample and asymptotic theory for the inferences associated with two such strategies. The most striking feature of the results is that the diagnostic tests and safeguards that follow from the practice of separating portfolio weight construction from return measurement greatly reduces the dependence of inference on the statistical properties of both returns and the procedures for constructing portfolio weights. Hence, the mean-variance efficiency tests developed here constitute a promising alternative to the usual multivariate tests which require portfolio grouping, although considerable sample and simulation evidence on the finite sample properties of the new procedures as well as on the relative power of the procedures is required. This investigation is left for future research.

REFERENCES

Amsler, Christine E. and Peter Schmidt, 1985, "A Monte Carlo Investigation of the Accuracy of Multivariate CAPM Tests." *Journal of Financial Economics* 14, pp. 359–376.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, "The Capital Asset Pricing Model: Some Empirical Tests." In Michael C. Jensen, ed., *Studies in the Theory of Capital Markets.* New York: Praeger.

Fama, Eugene F. and James D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81, pp. 607–636.

Gibbons, Michael R., 1982, "Multivariate Tests of Financial Models: A New Approach." *Journal of Financial Economics* 10, pp. 3–28.

———, Stephen A. Ross, and Jay Shanken, 1986, "A Test of the Efficiency of a Given Portfolio." Research Paper No. 853, Graduate School of Business, Stanford University.

Hall, P. and C. C. Heyde, 1980, *Martingale Limit Theory and Its Application.* New York: Academic Press.

Hansen, Lars P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50, pp. 1029–54.

Jobson, J. D. and Bob M. Korkie, 1982, "Potential Performance and Tests of Portfolio Efficiency." *Journal of Financial Economics* 10, pp. 433–466.

Kandel, Shmuel, 1984, "The Likelihood Ratio Test Statistic of Mean Variance Efficiency without a Riskless Asset." *Journal of Financial Economics* 13, pp. 575–592.

Lehmann, Bruce N., 1986a, "Residual Risk Revisited." National Bureau of Economic Research Working Paper No. 1908.

———, 1986b, "Mean Variance Efficiency Tests in Large Cross-sections." Graduate School of Business, Columbia University.

Litzenberger, Robert and Krishna Ramaswamy, 1979, "The Effects of Personal Taxes and Dividends on Capital Asset Prices." *Journal of Financial Economics* 7, pp. 163–195.

MacKinlay, Craig, 1986, "On Multivariate Tests of the CAPM." The Wharton School, University of Pennsylvania.

Roll, Richard W., 1977, "A Critique of the Asset Pricing Theory's Tests—Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics* 4, pp. 129–176.

———, 1979, "Testing a Portfolio for Ex Ante Mean Variance Efficiency." In Edwin E. Elton and Martin J. Gruber, *Portfolio Theory, Twenty Five Years After.* New York: North-Holland.

———, 1980, "Orthogonal Portfolios." *Journal of Financial and Quantitative Analysis* 15, pp. 1005–1023.

———, 1985, "A Note on the Geometry of Shanken's CSR T^2 Test for Mean Variance Efficiency." *Journal of Financial Economics* 14, pp. 349–358.

Ross, Stephen A., 1983, "A Test of the Efficiency of a Given Portfolio." School of Organization and Management, Yale University.

Scott, D. J., 1973, "Central Limit Theorems for Martingales and for Processes with Stationary Increments Using a Skorokhod Representation Approach." *Advances in Applied Probability* 5, pp. 119–137.

Shanken, Jay, 1983, "An Asymptotic Analysis of the Traditional Risk Return Model." Graduate School of Business, University of California, Berkeley.

———, 1985, "Multivariate Tests of the Zero Beta CAPM." *Journal of Financial Economics* 14, pp. 327–348.

———, 1986, "Testing Portfolio Efficiency when the Zero Beta Rate is Unknown." *Journal of Finance* 41, pp. 269–76.

Stambaugh, Robert F., 1982, "On the Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis." *Journal of Financial Economics* 10, pp. 237–268.

White, Halbert J., 1984, *Asymptotic Theory for Econometricians.* New York: Academic Press.