Journal of Financial Economics 30 (1991) 165-191. North-Holland

Small sample tests of portfolio efficiency

Guofu Zhou*
Washington University, St. Louis, MO 63130, USA

Received January 1991, final version received July 1991

This paper presents an eigenvalue test of the efficiency of a portfolio when there is no riskless asset, complementing the test of Gibbons, Ross, and Shanken (1989). Besides optimal upper and lower bounds, an easily-implented numerical method is provided for computing the exact P-value. Our approach makes it possible to draw statistical inferences on the efficiency of a given portfolio both in the context of the zero-beta CAPM and with respect to other linear pricing models. As an application, using monthly data for every consecutive five-year period from 1926 to 1986, we reject the efficiency of the CRSP value-weighted index for most periods.

1. Introduction

A fundamental problem in finance is whether or not a particular portfolio is efficient. In a framework in which all uncertainties about returns on risky assets are determined only by the means and variances of the returns, an efficient portfolio must have the maximum expected reward (mean return) for a given level of risk (variance). Since Markowitz's (1952) seminal contribution to modern portfolio theory, the question of efficiency problem has been of concern not only to individual investors and financial managers with respect to portfolio choice, but also to researchers in finance with respect to the validity of various equilibrium pricing models. It is well known that the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) and the zero-beta CAPM of Black (1972) rely on the efficiency of the market portfolio. Because of its intrinsic importance, the efficiency of a portfolio has been studied and tested extensively in the finance literature.

*I have benefited from the comments of Wayne Ferson, Campbell Harvey, Christopher Lamoureux, Francis Longstaff, Craig MacKinlay, William Schwert (the editor), and especially Philip Dybvig. I am also especially indebted to Jay Shanken (the referee), who provided many suggestions that substantially improved the paper. Financial support from the Fossett Foundation is gratefully acknowledged.

Studies of portfolio efficiency have traditionally been cast in the form of testing the Sharpe–Lintner CAPM and the zero-beta CAPM. However, these tests, which are based on cross-sectional regressions, suffer from an errors-in-variables problem. Gibbons (1982) was the first to develop a portfolio efficiency test within a multivariate statistical framework, although a similar multivariate formulation can be traced back to MacBeth (1975). Gibbons suggests the likelihood ratio test (LRT), which relies on an asymptotic (chi-squared) distribution. However, Stambaugh's (1982) simulations results show that, in fact, the asymptotic χ^2 is not a good approximation of the exact distribution. The efficiency hypothesis is rejected too often, and the test results become less reliable as the number of assets increases. This problem was further demonstrated by Shanken's (1985) theoretical study in which, with 40 assets and 60 periods, the asymptotic P-value is 0.01 whereas the true P-value is actually 0.92. Clearly, knowing the exact distribution or small sample properties of the test statistics is very important for making correct inferences.

When there exists a riskless asset, the exact test of a portfolio's efficiency is provided by Gibbons, Ross, and Shanken (1989). If a riskless asset does not exist, the problem is more complex because the zero-beta rate, which is the return on the zero-beta portfolio, is unknown and has to be estimated from the data. Moreover, the zero-beta rate enters the efficiency constraints by multiplication with other parameters, so that the hypothesis to be tested is nonlinear and hence it is more difficult to develop an exact test. Gibbons (1982) provides a Gauss–Newton numerical procedure for estimating the zero-beta rate, but his method is, in general, computationally intensive. Fortunately, Kandel (1984, 1986) obtains an explicit solution and provides some geometric interpretations for the maximum likelihood estimators. Shanken (1986) extends this result to market model parameterization and to the multibeta case. Although the estimation problem is elegantly solved, the testing problem remains. Due to problems in using the asymptotic tests, Shanken (1985) derives a small-sample lower bound on the P-value of the LRT test, in addition to his approximate P-value based on Hotelling's T^2-distribution. Furthermore, Shanken (1986) provides a small-sample upper bound on the P-value which is useful for rejecting efficiency without resorting to asymptotic results.

In this paper, we present an eigenvalue test for the efficiency of a given portfolio when there does not exist a riskless asset. In particular, if the given portfolio is the market portfolio, our procedure represents an exact test of the zero-beta CAPM. In contrast with the exact test of Gibbons, Ross, and Shanken (1989) in the riskless asset case, our test is unfortunately more complex. In our approach, the LRT is shown to be a transformation of an eigenvalue that allows us to apply known eigenvalue distributions of multivariate analysis to obtain the exact distribution and optimal upper and lower bounds. The exact distribution is shown to be dependent on a nonnegative nuisance parameter. As this nuisance parameter increases, the exact P-value also increases. Nevertheless, a maximum likelihood estimator of the nuisance parameter is easily obtained and a computational method is provided to give the exact P-value at any given level of the nuisance parameter. In addition, we provide optimal upper and lower bounds on the exact P-value so that inferences can be made without knowing the nuisance parameter. The upper bound follows an F-distribution and the lower bound is nonstandard but can still easily be computed. The upper bound on the P-value, which is tighter than that of Shanken (1986), is particularly useful when the hypothesis is close to being rejected. For example, if we find an upper bound of 5%, the true significance level must be less than or equal to 5%, so we can assert that the efficiency is rejected at the usual 5% level. Our methodology also applies to a variety of other linear pricing models.

In section 2, we provide a detailed account of the approach. First, the model and the efficiency implications are described. Second, to understand our new approach, the maximum likelihood estimation of the parameters is discussed. Third, a test statistic which is a monotonic function of the likelihood ratio is proposed. While the primary focus is to derive the exact distribution and the optimal upper and lower bounds, some of the statistical and economic interpretations are also explored. In section 3, we examine the Gibbons, Ross, and Shanken (1989) test from an eigenvalue perspective. In section 4, using monthly data for consecutive five-year periods from 1926–1986, we test the efficiency of the CRSP value-weighted index in the market model with twelve industry groups. At the 10% significance level, the efficiency hypothesis is rejected for all but two periods. At the 5% level, efficiency is rejected in eight of the twelve periods. To see whether the rejection is due to the 'January effect', we repeat our tests with the returns on January deleted. Conclusions and some remarks about future research are offered in the final section.

2. The model and the test of efficiency

2.1. *Asset pricing restrictions*

Our objective is to examine the efficiency of a given portfolio, p, whose return is r_{pt} in period t. Notice that efficiency cannot be tested without making assumptions about the distributions or law of motions of the asset returns. Different assumptions yield different empirical models, and so the efficiency hypothesis yields different constraints on the parameters of the models to be tested.

In our analysis, following Gibbons (1982), we assume that the returns are well-described by the 'market-model' regression:

$$r_{it} = \alpha_i + \beta_i r_{pt} + \varepsilon_{it}, \qquad i = 1, \ldots, N, \tag{1}$$

where

r_{it} = return on asset i in period t,
r_{pt} = return on the given portfolio in period t,
ε_{it} = disturbances or random errors, and
N = number of assets.

In consideration of the statistical tools available, we follow the tradition of most studies by assuming that the disturbances are independent over time, and jointly normally distributed each period with a mean of zero and a nonsingular covariance matrix $\boldsymbol{\Sigma}$. This implies that the returns are correlated contemporaneously but not across time and that the disturbances have the following properties:

$$\mathrm{E}\varepsilon_{it}\varepsilon_{js} = \begin{cases} \sigma_{ij} & \text{if} \quad s = t, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that we examine the efficiency of the portfolio without a riskless asset, so that if the portfolio is efficient, we obtain the following nonlinear constraints on the parameters of the model:

$$\mathrm{H}_0: \quad \alpha_i = \gamma_0(1 - \beta_i), \qquad i = 1, \ldots, N. \tag{2}$$

This is the null hypothesis to be tested. The hypothesis is a nonlinear one, because both the 'zero-beta' rate, γ_0, and β_i are unknown and have to be estimated from the data. Further, both γ_0 and β_i enter the constraints by multiplication with one another. There are many standard approaches for dealing with linear constraints, but few available for studying nonlinear ones. In fact, both the estimation of the parameters under the constraints (2) and the testing problem present a nonstandard problem in multivariate analysis. The focus of the paper is to derive a test of the validity of these nonlinear constraints and to give optimal upper and lower bounds for its complex exact distribution.

2.2. Maximum likelihood estimators

There are two maximum likelihood (ML) estimators of the parameters, unconstrained and constrained, that are needed later to construct the test statistics. For the unconstrained ML estimators, a standard formula is available [e.g., Anderson (1984, pp. 287–289)]. However, it is more difficult to obtain the constrained ML estimators, i.e., those values of parameters that maximize the likelihood function under the constraints (2). Gibbons (1982) provides a Gauss–Newton numerical procedure, but his method, as shown by Shanken (1989), is only asymptotically equivalent to the ML estimator. Fortunately, an exact closed-form solution was found by Shanken (1986) for the multibeta case, which covers our market-model parameterization. The constrained estimators thus become straightforward to evaluate. Nevertheless, Shanken's expression does not seem to allow for easy study of the statistical properties of either the estimators or the likelihood ratio test.

We develop a new approach to obtain a determinant equation that governs the constrained ML estimator. The striking feature of our approach is that it links the distributional properties of the ML estimators and the likelihood ratio to known results of multivariate analysis. In particular, our simple expression for the likelihood ratio allows us to find a test statistic whose exact distribution can be obtained. Since our approach appears to be relevant to a variety of linear pricing models, we present it in some detail.

To simplify the presentation that follows, we rewrite the return-generating process (1) in matrix form:

$$\boldsymbol{R} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{E}, \tag{3}$$

where $\boldsymbol{R}$ is a T (observations) $\times N$ (asset) matrix of returns, $\boldsymbol{X}$ is a $T \times 2$ matrix with the first column a vector of ones and the second the portfolio returns, $\boldsymbol{\Theta}$ is a $2 \times N$ coefficient matrix with the N alphas, $\boldsymbol{\alpha}$, in the first row and the N betas, $\boldsymbol{\beta}$, in the second, and $\boldsymbol{E}$ is the $T \times N$ disturbance matrix.

Notice the fact that the efficiency constraints (2) do not contain any elements of the parameter matrix $\boldsymbol{\Sigma}$, so we can obtain the ML estimators in two steps. First, conditional on $\boldsymbol{\Theta}$, we maximize the likelihood function with respect to $\boldsymbol{\Sigma}$. Then, in the second step, we maximize the so-called 'concentrated' likelihood function (the likelihood function in which $\boldsymbol{\Sigma}$ is replaced by its conditional estimator) with respect to $\boldsymbol{\Theta}$ alone.

By initially not imposing any constraints but then imposing the constraints (2) on this maximization, we obtain the unconstrained and constrained ML estimators, respectively. Although the unconstrained ML estimator is known, we derive its formula in a slightly different way in order to highlight our approach to the derivation of the constrained ML estimator.

With the assumption of normality on the disturbances, the log-likelihood function is

$$\log L(\boldsymbol{\Theta}, \boldsymbol{\Sigma}) = -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma}| - \tfrac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Omega}, \tag{4}$$

where 'tr' denotes the trace of a matrix, and $\boldsymbol{\Omega}$ is an $N \times N$ matrix:

$$\boldsymbol{\Omega} \equiv (\boldsymbol{R} - \boldsymbol{X\Theta})'(\boldsymbol{R} - \boldsymbol{X\Theta}). \tag{5}$$

Conditional on $\boldsymbol{\Theta}$, it is known [(Anderson, 1984, p. 62, lemma 3.2.2)] that the conditional estimator of $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma} = \frac{1}{T}(\boldsymbol{R} - \boldsymbol{X\Theta})'(\boldsymbol{R} - \boldsymbol{X\Theta}).$$

Replacing $\boldsymbol{\Sigma}$ in the log-likelihood function (4) by this conditional estimator, we have the concentrated likelihood function:

$$\log L(\boldsymbol{\Theta}) = -\frac{T}{2}\log|\boldsymbol{\Omega}| + C, \tag{6}$$

where $C = -NT[\log(2\pi) - \log T + 1]/2$. Thus, in order to maximize the concentrated likelihood function, we need only to minimize $|\boldsymbol{\Omega}|$, the determinant of the $\boldsymbol{\Omega}$ matrix (5). It is easy to verify that

$$\boldsymbol{\Omega} = \left(\boldsymbol{R} - \boldsymbol{X}\hat{\boldsymbol{\Theta}}_u\right)'\left(\boldsymbol{R} - \boldsymbol{X}\hat{\boldsymbol{\Theta}}_u\right) + \left(\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}_u\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}_u\right),$$

where $\hat{\boldsymbol{\Theta}}_u$ is defined as

$$\hat{\boldsymbol{\Theta}}_u = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}.$$

Now, the semi-positive definite matrix $\boldsymbol{\Omega}$ is a sum of two other positive semi-definite matrices of which the first one does not depend on $\boldsymbol{\Theta}$, so its determinant must be minimized if the determinant of the second is. Therefore, the unconstrained ML estimator of $\boldsymbol{\Theta}$ is equal to $\hat{\boldsymbol{\Theta}}_u$, a well-known result. It is worth mentioning that, as far as the computation of $\hat{\boldsymbol{\Theta}}_u$ is concerned, its ith column can be obtained by running an ordinary least-squares (OLS) regression of the ith asset on the market return, i.e., regressing the ith equation of the market model (1). The unconstrained ML estimator of $\boldsymbol{\Sigma}$ is then given by the cross-products of the residuals divided by T.

To obtain the constrained estimator, we make the following transformation of parameters:

$$\boldsymbol{\Phi} \equiv \begin{pmatrix} \alpha_1 & \cdots & \alpha_N \\ \beta_1 - 1 & \cdots & \beta_N - 1 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}(\eta_1, \ldots, \eta_N). \tag{7}$$

Under the null hypothesis that the constraints (2) are true, the column rank of $\boldsymbol{\Phi}$ must be less than or equal to one. It can thus be expressed as the right-hand side. On the other hand, for any parameter values of $\boldsymbol{\xi} = (\xi_1, \xi_2)'$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)'$ that minimize $|\boldsymbol{\Omega}|$, the corresponding $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$'s must also minimize $|\boldsymbol{\Omega}|$ and satisfy the constraints. Further, as shown later, given a solution of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ to the minimization problem, an arbitrary scale such as $k\boldsymbol{\xi}$ and $(1/k)\boldsymbol{\eta}$ is still a solution. Although the estimators of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are not unique, there exists one and only one ML estimator of the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

With the above transformation, we are ready to find the $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ that minimize $|\boldsymbol{\Omega}|$. This is an unconstrained minimization problem. We have first:

$$\boldsymbol{\Omega} = \left(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\boldsymbol{\eta}\right)'\left(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\boldsymbol{\eta}\right), \qquad \tilde{\boldsymbol{X}} \equiv \boldsymbol{X}\boldsymbol{\xi}, \qquad \boldsymbol{Y} \equiv \boldsymbol{R} - \mathbf{1}\boldsymbol{R}_p',$$

where $\mathbf{1}$ is a $N \times 1$ vector of 1's and $\boldsymbol{R}_p$ is the vector of the returns on the given portfolio whose efficiency is to be examined. In comparison with (5), the estimator of $\boldsymbol{\eta}$ conditional on $\boldsymbol{\xi}$ must have the same algebraic form as the unconstrained ML estimator, and hence it is given by

$$\hat{\boldsymbol{\eta}} = \left(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}'\boldsymbol{Y}.$$

Next, we choose some suitable $\boldsymbol{\xi}$ to minimize the 'concentrated' determinant:

$$|\boldsymbol{\Omega}| = \left|\left(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\eta}}\right)'\left(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\eta}}\right)\right|.$$

After a few simplifications, it has the following tractable form:

$$|\boldsymbol{\Omega}| = |\boldsymbol{Y}'\boldsymbol{Y}|\left(1 - \frac{\boldsymbol{\xi}'\boldsymbol{X}'\boldsymbol{Y}(\boldsymbol{Y}'\boldsymbol{Y})^{-1}\boldsymbol{Y}'\boldsymbol{X}\boldsymbol{\xi}}{\boldsymbol{\xi}'(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\xi}}\right).$$

This expression shows that only the second term in the larger parentheses needs to be maximized. Consider the maximum of $(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x})/(\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x})$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices and $\boldsymbol{x}$ is any vector with the same column number. By the result of matrix theory [Anderson (1984, p. 590, theorem A.2.4)] we know that the maximum is equal to the largest eigenvalue of $\boldsymbol{A}$ with respect to $\boldsymbol{B}$:

$$(\boldsymbol{A} - \lambda\boldsymbol{B})\boldsymbol{x} = 0,$$

or the largest root of $|\boldsymbol{A} - \boldsymbol{B}\lambda| = 0$. The maximum is obtained when $\boldsymbol{x}$ is the corresponding eigenvector. Therefore, if λ_1 and λ_2 where $\lambda_1 \geq \lambda_2$ are the roots of

$$|X'Y(Y'Y)^{-1}Y'X - \lambda X'X| = 0, \tag{8}$$

the 'concentrated' determinant $|\boldsymbol{\Omega}|$ is minimized if $\boldsymbol{\xi}$ is the eigenvector corresponding to λ_1, and the minimum is given by

$$|\boldsymbol{\Omega}_r| = |Y'Y|(1-\lambda_1). \tag{9}$$

By using the transformation (7), the constrained ML estimator of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be obtained. Although this provides an alternative way for obtaining the constrained ML estimator, Shanken (1986) is computationally more efficient, because our method requires evaluation of eigenvalues and eigenvectors, whereas Shanken's approach requires only the solution of a quadratic equation.[1] However, as we pointed out before, the value of our method is that it allows us to apply the classical distributional results of multivariate statistics to derive an exact test of the null hypothesis (2).

2.3. *The eigenvalue test and its optimal bounds*

Consider now an important formula for the likelihood ratio (LR). By definition, the LR is the ratio of the likelihood function at the constrained ML estimator to that at the unconstrained estimator. Thus,

$$LR \equiv \frac{\max_{\gamma_0, \boldsymbol{\beta}, \boldsymbol{\Sigma}} L(\gamma_0, \boldsymbol{\beta}, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\Theta}, \boldsymbol{\Sigma}} L(\boldsymbol{\Theta}, \boldsymbol{\Sigma})} = \left(\frac{|\boldsymbol{\Omega}_u|}{|\boldsymbol{\Omega}_r|}\right)^{T/2}, \tag{10}$$

where the second equality follows from (6), and $\boldsymbol{\Omega}_u$ is the $\boldsymbol{\Omega}$ matrix evaluated at the unconstrained ML estimator. Now it can be shown that

$$|\boldsymbol{\Omega}_u| = |Y'Y|(1-\lambda_1)(1-\lambda_2), \tag{11}$$

so, combining (9) with (11), we have the interesting relationship between the likelihood ratio and the eigenvalue λ_2,

$$LR = (1-\lambda_2)^{T/2}. \tag{12}$$

Since it is always true that $0 \leq LR \leq 1$, the range of λ_2 must be between 0 and 1, that is, $0 \leq \lambda_2 \leq 1$. This fact can also be proved directly from (8).

[1]Actually, the computational advantage is not clear in the market model, because the eigenvalues and eigenvectors are also analytically determined by a quadratic equation. But the advantage will be important in a multi-beta model.

To test the efficiency constraints (2), we use λ_2 as the test statistic. Since λ_2 is a monotonic function of the likelihood ratio, this test is statistically equivalent to the likelihood ratio test The exact distribution of λ_2 has a simpler form than the LRT and is thus easier to implement.

Why is λ_2 a measure of deviation from the null? There are at least two statistical interpretations. The first is in terms of the likelihoods: the bigger the λ_2, the smaller the LR, i.e., the lower the ratio of the maximum likelihood under the null to the maximum likelihood under the alternative. Thus as λ_2 increases we can plausibly believe that the null is not valid. The second interpretation is in terms of model fitting. It is shown [see (10)] that LR is also a positive power function of the ratio $(|\boldsymbol{\Omega}_u|/|\boldsymbol{\Omega}_r|)$. Recall that $|\boldsymbol{\Omega}_r|$ and $|\boldsymbol{\Omega}_u|$, the so-called generalized variances, measure how the model fits the data under both the null and the alternative [Anderson (1984, pp. 259–263)]. Their ratio thus indicates the relative goodness of fit. As a result, the smaller the λ_2, the higher the ratio $(|\boldsymbol{\Omega}_u|/|\boldsymbol{\Omega}_r|)$, which implies the fit becomes better under the null, making it more difficult to reject the null.

What economic departures from the null have been measured by λ_2? If there is a riskless asset, Gibbons, Ross, and Shanken (1989) show that

$$LR = \left(\frac{1}{1+Q}\right)^{T/2}, \qquad Q \equiv \frac{S(p^*)^2 - S(p)^2}{1+S(p)^2}, \tag{13}$$

where $S(p^*)$ and $S(p)$ are the Sharpe measures (i.e., the maximum sample mean excess return per unit of sample standard deviation) of the ex post efficient portfolio, p^*, and the (market) portfolio, p, whose efficiency is under study. If p is indeed efficient, we would expect the difference between $S(p^*)$ and $S(p)$ to be small, and hence Q should be small. If $S(p^*)$ is sufficiently greater than $S(p)$, the return per unit of risk (Sharpe measure) for portfolio p is much lower than that for the ex post efficient portfolio, making it natural to reject the null hypothesis that the portfolio p is ex ante efficient. In the present zero-beta case, even though there does not exist a riskless asset, the same relation is still valid [see Shanken (1985, 1986)] if we interpret the maximum likelihood estimator of γ_0 as the risk-free borrowing and lending rate. By using (12) and (13), it immediately follows that

$$\lambda_2 = \frac{Q}{1+Q}. \tag{14}$$

Thus, λ_2 is an increasing function of Q. As a result, it measures the ex post departure of the null that the efficiency constraints (2) are valid. If we find ex post that the Sharpe measure of the market portfolio $S(p)$ is very close to that of the efficient portfolio $S(p^*)$, Q and hence λ_2 will be close to zero, and we have little evidence against the null. Otherwise, we may reject the null because of a big difference, or equivalently, a high λ_2 value.

The exact density (and distribution) of λ_2 is unfortunately much more complex than the exact test of Gibbon, Ross, and Shanken (1989) in the riskless asset case. For one thing, it depends on unknown parameters even under the null hypothesis. More specifically, let $\omega_1 \geq \omega_2 \geq 0$ be the roots of

$$|S\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'S' - \omega I| = 0, \tag{15}$$

where S is a 2×2 nonsingular matrix such that $X'X = S'S$. It was observed earlier that the null hypothesis is true if and only if the rank of $\boldsymbol{\Phi}$ is less than or equal to one [or equivalently, the rank of $(S\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'S')$ is not greater than one]. Therefore, under the null we must have $\omega_2 = 0$, although we do not know the value of ω_1. The exact density is shown to be dependent upon the nuisance parameter ω_1 (see Theorem 4 in appendix A). To emphasize the dependence, we write the probability $\mathrm{Prob}(\lambda_2 < x)$ as $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ wherever it is necessary. (This should not be confused with conditional probabilities.) In practice, the exact P-value is of interest for drawing inferences. Although we have to know ω_1 to compute the P-value, we can use an estimated value of ω_1 (in particular, the maximum likelihood estimator) and the method described in appendix B to arrive at the P-value. Numerical results show that a small perturbation of ω_1 does not change the P-value very much.

Theoretically, however, it is of interest to know how the probability $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ changes as the true but unknown parameter ω_1 varies. This is useful because it indicates the bias when we evaluate $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ at an estimated value of ω_1. The following theorem asserts the monotonicity of this probability in ω_1:

Theorem 1. Under the null hypothesis that the constraints (2) *are true, if* $N \geq 2$ *and* $T \geq N + 2$, *then the probability* $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ *is a decreasing function in* ω_1,

$$\mathrm{Prob}(\lambda_2 < x|\omega_1') < \mathrm{Prob}(\lambda_2 < x|\omega_1''),$$

whenever $\omega_1' > \omega_1''$, $0 \leq \omega_1', \omega_1'' \leq 1$.

Proof. See appendix A.

Theorem 1 states that, if we compute $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ at a value of ω_1 higher than the true value, we will underestimate the true probability. In other words, we will have a lower bound for the true probability. Alternatively, the result will be an upper bound if we use a value of ω_1 that is lower than the true value.

The monotonicity implies that if $\lim_{\omega_1 \to \infty} \mathrm{Prob}(\lambda_2 < x|\omega_1)$ exists, it must be a lower bound on $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ for all possible values of ω_1. Formally, based on Schott (1984), who in turn extended Saw (1974), we have:[2]

Theorem 2. Under the null hypothesis that the constraints (2) are true, if $N \geq 2$ and $T \geq N+2$, then we have the following lower bound on the distribution of λ_2:

$$\mathrm{P}(\lambda_2 \leq x|\omega_1) \geq \mathrm{P}\left[F_{N-1,T-N} \leq \frac{x(T-N)}{(1-x)(N-1)}\right], \qquad \forall\, \omega_1 \geq 0.$$

The bound is optimal in the sense that the limit of the left-hand side, $\lim_{\omega_1 \to \infty} \mathrm{Prob}(\lambda_2 < x|\omega_1)$, obtains the bound as the parameter ω_1 goes to infinity.

Proof. See appendix A.

By the same token, $\mathrm{Prob}(\lambda_2 < x|0)$ is the optimal upper bound on the probability $\mathrm{Prob}(\lambda_2 < x|\omega_1)$. Theorem 1 implies that the latter is less than or equal to the former; the equality occurs if and only if $\omega_1 = 0$. Based upon Nanda (1947) or Pillai (1956), a closed-form expression for the distribution of this optimal upper bound can be derived:

Theorem 3. Under the null hypothesis that the constraints (2) are true, if $N \geq 2$ and $T \geq N+2$, then we have the following upper bound[3] on the distribution of λ_2:

$$\mathrm{P}(\lambda_2 \leq x|\omega_1) \leq 1 - k_2\left(2\int_0^{1-x} t^{2n_2+1}(1-t)^{2n_1+1}\,\mathrm{d}t \right.$$

$$\left. - x^{n_1+1}(1-x)^{n_2+1}\int_0^{1-x} t^{n_2}(1-t)^{n_1}\,\mathrm{d}t\right),$$

[2] In terms of our framework, Shanken (1986) also derived a lower bound on $\mathrm{Prob}(\lambda_2 < x|\omega_1)$, but our bound is tighter and enjoys the optimality pointed out in Theorem 2. However, his bound is more general in that it is also a bound under the alternative. An additional lower bound can also be derived based on Muirhead and Chikuse (1975) which we will not pursue since we do not consider alternative hypotheses here.

[3] Shanken (1985) also derived an upper bound on $\mathrm{Prob}(\lambda_2 < x|\omega_1)$ in addition to his approximate P-value based on Hotelling's T^2-distribution.

where $n_1 = (N-3)/2$, $n_2 = (T-N-3)/2$, $k_1 = \Gamma(T-1)/4\Gamma(N-1)\Gamma(T-N-1)$, $k_2 = k_1/(n_1+n_2+2)$, *and* $\Gamma(a)$ *is the gamma function (evaluated at* a*). The bound is optimal in the sense that the left-hand side* $\mathrm{P}(\lambda_2 \leq x | \omega_1)$ *obtains the bound at* $\omega_1 = 0$.

Proof. See appendix A.

Both Theorem 2 and Theorem 3 provide lower and upper bounds that are easily computed on the probability Prob($\lambda_2 < x | \omega_1$). They translate immediately into upper and lower bounds on the *P*-value. An upper bound on the *P*-value, if it is small, is useful for rejecting the null hypothesis. On the other hand, a lower bound on the *P*-value, if it is large, is useful for accepting the hypothesis. Other than being optimal and easy to compute, the bounds also have the attractive feature that they are independent of any unknown parameters. This is in contrast with the exact distribution which depends on omega, a nuisance parameter that is a complex function of the unknown true alpha's and beta's as well as the covariance matrix.

3. The GRS test: An eigenvalue perspective

We have shown that the likelihood ratio test of the efficiency of a portfolio with no riskless asset is equivalent to an eigenvalue test. An unusual feature about the eigenvalue test is that its exact distribution not only has a complex form, but it also depends on a nuisance parameter even under the null. This is in contrast with the Gibbons, Ross, and Shanken (1989) test (in which a riskless asset exists) which has a simple *F*-distribution under the null; no nuisance parameters are involved. To shed light on this difference, it is of interest to derive their test from an eigenvalue perspective.

If there is a riskless asset, we can interpret the returns, r_{it}, in the market model (1) as excess returns; the efficiency of the given portfolio r_p thus implies the following constraints on the parameters:

$$\mathrm{H}_0: \quad \alpha_i = 0, \qquad i = 1, \ldots, N. \tag{16}$$

As in the previous section, the likelihood ratio for testing this hypothesis can be shown to have the same form as (10) except that the constrained determinant $|\boldsymbol{\Omega}_r|$ is replaced by

$$|\boldsymbol{\Omega}_r| = \left|\left(\boldsymbol{R} - \boldsymbol{R}_p\tilde{\boldsymbol{\beta}}\right)'\left(\boldsymbol{R} - \boldsymbol{R}_p\tilde{\boldsymbol{\beta}}\right)\right|,$$

where $\tilde{\boldsymbol{\beta}} = (\boldsymbol{R}_p\boldsymbol{R}_p')^{-1}\boldsymbol{R}_p'\boldsymbol{R}$ is the constrained estimator of $\boldsymbol{\beta}$ under (16). Presumably, we can follow (9) or (11) in decomposing $|\boldsymbol{\Omega}_r|$ into a product of eigenvalues and possibly some other factors. However, there will be no simplification in the likelihood ratio expression because, due to the nature of the hypothesis, the eigenvalues in the decomposition do not have the same form as those in (11). As a result, a different method of decomposition has to be used.

In fact, there is a general and well-known method of decomposition in multivariate statistical analysis which is applicable to generic linear hypotheses, of which (16) is a special case. For both linear hypotheses and many others, the idea is to transform the likelihood ratio into a function of eigenvalues. Then the likelihood ratio test is obtained and results on eigenvalue distributions are used to obtain its distribution. Very often, functions of some or all of the eigenvalue themselves are chosen as the test statistics. Since these tests do not seem to be widely known in the econometrics literature, we discuss them briefly in what follows. Based on this general discussion, we obtain, in particular, the Gibbon, Ross, and Shanken test from an eigenvalue perspective.

Consider testing a general linear hypothesis in a multivariate regression model. The matrix form of the model, (3), can clearly be generalized into a multivariate regression by interpreting $\boldsymbol{X}$ as a $T \times K$ matrix and $\boldsymbol{\Theta}$ as a $K \times N$ matrix. Letting $\boldsymbol{A}$ and $\boldsymbol{C}$ be known $q \times K$ and $q \times N$ matrices, we want to examine if there are linear relationships among the parameters $\boldsymbol{\Theta}$,

$$\mathrm{H}_0\colon \quad \boldsymbol{A}\boldsymbol{\Theta} = \boldsymbol{C}. \tag{17}$$

In the univariate case, this is the familiar linear hypothesis testing problem in a regression model and we have the commonly used F-test. In the multivariate case, the testing problem becomes much more complex, although there are many well-known tests available whose exact distribution can be computed (at least in principle); most of these tests are derived from an eigenvalue perspective.

To test (17), it can be shown [see, e.g., Muirhead (1982, ch. 10) for details] that the likelihood ratio test has the eigenvalue decomposition:

$$LRT = \left(\prod_{i=1}^{s} (1 + f_i) \right)^{-T/2},$$

where $s = \min(q, N)$, $f_1 > \cdots > f_s > 0$ are positive eigenvalues of $\boldsymbol{\Omega}_u^{-1/2} \boldsymbol{H} \boldsymbol{\Omega}_u^{-1/2}$, and $\boldsymbol{H} = \boldsymbol{\Omega}_r - \boldsymbol{\Omega}_u$ and $\boldsymbol{\Omega}_r$ and $\boldsymbol{\Omega}_u$ are the constrained and unconstrained cross-products of the residuals. If hypothesis (17) is true, we expect the LRT to be close to one. Other than the LRT itself, we may use functions of the eigenvalues, $f_1 > \cdots > f_s > 0$, to examine how close the LRT is to one. For example, if the maximum root, f_1, is close to zero, we expect the *LRT* to be close to one. This line of reasoning leads to the following well-known tests of (17).

Wilk's Λ-test:

$$\Lambda = \prod_{i=1}^{s} (1 + f_i).$$

Hotelling–Lawley's T_0^2-test:

$$T_0^2 = \sum_{i=1}^{s} f_i.$$

Pillai's test:

$$V = \sum_{i=1}^{s} \frac{f_i}{1 + f_i}.$$

Roy's largest root test:

$$W = f_1.$$

All tests reject the null for large observed statistic values. To understand why no nuisance parameters are present in the null case, notice that the distribution of $\boldsymbol{\Omega}_u$ is central Wishart, independent of the distribution of $\boldsymbol{H}$ which is noncentral Wishart with the noncentrality parameter:

$$\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{A\Theta} - \boldsymbol{C})'\left[\boldsymbol{A}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{A}'\right]^{-1}(\boldsymbol{A\Theta} - \boldsymbol{C})\boldsymbol{\Sigma}^{-1/2}.$$

Under the null, $\boldsymbol{\Delta} = 0$, and the eigenvalues are thus independent of any unknown parameters of the model. For example, the exact distribution of Wilk's Λ-test is a product of independent beta random variables with degrees of freedom dependent only on T, K, N, and q.

Returning to the efficiency hypothesis (16), we have $s = 1$. The above four tests collapse to be equivalent to a single test: the likelihood ratio test. In this case, using the relationship between beta and F-distributions, the LRT has an F-distribution [see, e.g., Muirhead (1982, p. 458)]. This is the same result obtained by Gibbon, Ross, and Shanken.

It should be noted that the above tests may have usefulness beyond the scope of this paper. As alternative exact small-sample tests, they may be used to test the mean–variance spanning hypothesis of Huberman and Kandel (1987). Moreover, because financial theories do not specify which benchmark portfolios should be included in the right-hand side of the multi-beta model, it may be interesting to examine the zero hypothesis of the coefficient matrix of those benchmark portfolios.[4] This hypothesis can be tested jointly with the efficiency hypothesis. Economic intuition together with statistical tools are helpful in identifying the benchmark portfolios and analyzing their contributions to the efficiency tests.

4. Empirical results

As an application, we use our test to examine the efficiency of the CRSP value-weighted index. Twelve value-weighted industry portfolios are used, following Sharpe (1982), Breeden, Gibbons, and Litzenberger (1989), Gibbons, Ross, and Shanken (1989), and Ferson and Harvey (1990). The market return is the value-weighted New York Stock Exchange return available from the Center for Research in Security Prices (CRSP) at the University of Chicago. All returns used in the market model are raw returns (in contrast, returns in excess of the riskless rate must be used when there is a riskless asset).

Given that the twelve industry portfolios are adequately described by the market model, the empiricist must choose an appropriate sample size. If the sample size T is too small, the maximum likelihood estimation of the true parameters may not be very accurate. In fact, T must be greater than N both for the invertibility of the covariance matrix and the requirements of the theorems. On the other hand, concerns of parameter stability suggest that T not be too large. Following most practices, we choose five-year subperiods with $T = 60$ as the sample size.

The results are provided in table 1. The first two columns are the maximum likelihood estimators of ω_1 and γ_0 under the null hypothesis that the restrictions (2) are valid.[5] There are no standard errors reported for the estimations because no such results have become available yet. The next column reports the value of the statistic λ_2, followed by its exact P-value, P_0, evaluated at the maximum likelihood estimator of ω_1 (see appendix B for the details of computation). By comparison, we also report the approximate P-values, P_S and P_{JK}, of Shanken (1985) and Jobson and Korkie (1982). An upper bound from Shanken (1986), U_1, is included to compare with our optimal upper bound U_2 (Theorem 2). The final column reports the optimal lower bound,[6] L, a result of Theorem 3. The overall P-value is reported at

[4]Shanken (1987a) and Kandel and Stambaugh (1987) discuss how to test portfolio efficiency if only proxies on the true benchmark portfolios are available.

[5]Somewhat surprisingly, the constrained ML estimator of ω_1 is identical numerically to its unconstrained ML estimator, a fact to be established theoretically.

[6]Shanken (1985) also derives a lower bound.

Table 1

Tests of the efficiency of the CRSP value-weighted index.

The efficiency is examined by using the market model:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it}, \quad i = 1, \ldots, N; \qquad \mathrm{E}\varepsilon_t \varepsilon_s' = \boldsymbol{\Sigma} \quad \text{if } s = t, \quad 0 \text{ otherwise},$$

where r_{it} is the return on the ith industry-sorted portfolio, r_{mt} the return on the CRSP value-weighted index, and $N = 12$ is the total number of industries. The data are monthly returns and the sample size is $T = 60$ in each of the five-year testing periods. The following hypothesis is to be tested:

$$\mathrm{H}_0\colon \quad \alpha_i = \gamma_0(1 - \beta_i), \qquad i = 1, \ldots, N.$$

In the table $\hat{\omega}_1$ and $\hat{\gamma}_0$ are the maximum likelihood estimators of the nuisance parameter[a] ω_1 and the zero-beta rate γ_0. The test statistic is λ_2 which is equivalent to the likelihood ratio test: $LR = (1 - \lambda_2)^{T/2}$. P_0 is the P-value, $\mathrm{Prob}(\lambda_2 > x)$, evaluated at $\hat{\omega}_1$. P_S and P_{JK} are approximate P-values based on Shanken (1985) and Jobson and Korkie (1982). Based on Shanken (1986) and Theorem 2, U_1 and U_2 are the upper bounds of the P-value which, based on Theorem 3, is also bounded below by L, a sum of two integrations.[b] These bounds are valid for all possible values of ω_1. The last row reports the overall P-value for the whole 1926/2–1986/1 period computed by aggregating the subperiod P-values (P_0's) using Shanken's (1985) normal density approach.

Period	$\hat{\omega}_1$	$\hat{\gamma}_0$	λ_2	P_0	P_S	P_{JK}	U_1	U_2	L
1926/2–1931/1	156.62	0.0043	0.3541	0.017	0.011	0.017	0.027	0.019	< 0.001
1931/2–1936/1	397.49	−0.0181	0.1953	0.410	0.321	0.360	0.478	0.413	0.124
1936/2–1941/1	235.91	< −0.0001	0.2908	0.079	0.053	0.060	0.107	0.082	0.004
1941/2–1946/1	46.28	−0.0151	0.2817	0.085	0.065	0.074	0.128	0.099	0.006
1946/2–1951/1	192.38	−0.0041	0.4062	0.004	0.002	0.002	0.006	0.004	< 0.001
1951/2–1956/1	380.50	0.0049	0.3396	0.026	0.016	0.018	0.037	0.027	< 0.001
1956/2–1961/1	119.83	0.0182	0.3517	0.018	0.011	0.013	0.026	0.020	< 0.001
1961/2–1966/1	52.71	0.0124	0.4203	0.002	0.001	0.001	0.004	0.003	< 0.001
1966/2–1971/1	85.96	0.0157	0.3641	0.012	0.008	0.009	0.020	0.014	< 0.001
1971/2–1976/1	66.59	0.0058	0.3478	0.018	0.012	0.014	0.030	0.022	< 0.001
1976/2–1981/1	125.43	−0.0066	0.3843	0.007	0.004	0.005	0.012	0.008	< 0.001
1981/2–1986/1	111.94	0.0256	0.1461	0.680	0.595	0.644	0.750	0.690	0.388
1926/2–1986/1			overall P-value: < 0.001						

[a]The parameter $\omega_1 \geq 0$ is the largest root of the determinate equation: $|\boldsymbol{S\Phi\Sigma}^{-1}\boldsymbol{\Phi}'\boldsymbol{S}' - \omega \boldsymbol{I}| = 0$, where $\boldsymbol{S}$ is a 2×2 nonsingular matrix such that $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{S}'\boldsymbol{S}$, while $\boldsymbol{X}$ is a $T \times 2$ matrix with the first column being a vector of ones and the second the returns of the index; $\boldsymbol{\Phi}$ is a $2 \times N$ parameter matrix with the N alphas in the first row and the N betas minus 1 in the second and $\boldsymbol{I}$ is the identity matrix of order 2.

[b]To be more precise, the variable L is defined as

$$L = k_2\left(2\int_0^{1-x} t^{2n_2+1}(1-t)^{2n_1+1}\,\mathrm{d}t - x^{n_1+1}(1-x)^{n_2+1}\int_0^{1-x} t^{n_2}(1-t)^{n_1}\,\mathrm{d}t\right),$$

where $n_1 = 4.5$, $n_2 = 22.5$, $k_2 = 885/29\Gamma(11)\Gamma(47)$, and $\Gamma(\cdot)$ is the gamma function.

the bottom row following Shanken (1985), who proposes a general method of aggregating P-values over subperiods. The novelty of this aggregation is that it allows one to make inferences over the whole period without assuming the stationarity of the parameters during that period. However, it is still necessary to assume the stationarity in each of the subperiods as well as the independence of the return residuals across subperiods.

We observe from table 1 that the nonnegative nuisance parameter ω_1 is likely to be large for our return data. Of all the subperiods, the lowest estimator is 46.28 and the highest is 397.49. Since they are increasing functions of ω_1, the exact P-values evaluated at the estimated ω_1's should not be far from the upper bounds in the same subperiod. This is verified by comparing the values of P_0 with U_2. The various P-values, P_0, P_S, and P_{JK}, are remarkably close to one another; the differences are within 1% in most subperiods. Since the true P-value is the exact distribution at the true but unknown parameter value ω_1, we do not know the accuracy of P_0, P_S, and P_{JK}. Notice that P_{JK}, which is the P-value of Bartlett's corrected large sample test [Bartlett (1938, p. 38)], is justified only asymptotically. Although P_S is a small-sample approximation, it is difficult to assess the error. The disadvantage of using P_0 is that it can only be computed by specifying a value of ω_1, due to the unfortunate fact that the exact distribution depends on the nuisance parameter [as is true of the exact distributions of many other likelihood ratio tests in canonical correlation analysis, e.g., Muirhead (1982, ch. 11)]. However, there are important advantages to using our exact distribution. First, the dependence on the nuisance parameter becomes less and less important as the sample size increases. Second, given a finite sample, if the estimator of ω_1 is close to the true ω_1, P_0 is also close to the true P-value; if the estimator of ω_1 is the true ω_1, P_0 must be the true P-value. Furthermore, our results allow computation of the exact P-values at different assumed values of ω_1 around the ML estimator, so that the sensitivity to inference can be examined.

The statistical significance of the efficiency of the CRSP value-weighted index can be seen immediately from the optimal upper and lower bounds. At the usual 5% significance level, we clearly reject efficiency for all the subperiods except four. If we raise the significance level to 10%, there are only two subperiods, February 1931 to January 1936 and February 1981 to January 1986, for which we cannot reject efficiency, possibly due to either the inability to bound the true P-value at the given significance level or the high significance level the index entertains. For example, for the periods February 1936 to January 1941 and February 1941 to January 1946 at the 5% significance level, the values of U_2 for these two periods are greater than 5%, but the values of L are not. However, the P_0's, the exact P-values evaluated at the ML estimators of ω_1, suggest that we do not reject efficiency in these two periods at the 5% level. For the periods February 1931 to January 1936 and

February 1981 to January 1986, since the P-values are greater than 12% and 38%, respectively, we are certain that we cannot reject the efficiency of the CRSP value-weighted index at both the 5% and 10% significance levels. In the above analysis, it is observed that Shanken's upper bound, U_1, derived from a different motivation, performs reasonably well. However, there are cases where the exact P-values are needed to make inference.

To understand why efficiency is rejected in so many subperiods, we want to compare the sample Sharpe measure of the index with that of the sample tangent portfolio. If there is a big difference between them, we would expect to reject the efficiency hypothesis from the familiar mean–variance analysis. Given the symmetric treatment of positive and negative Sharpe measures, we can consider only the positive ones without loss of generality. For example, let us consider why we have rejected the efficiency in the period of February 1926 to January 1931. For this period, $U_2 = 0.019$. As discussed in the second section, if we regard $\gamma_0 = 0.0043\%$ as the riskless rate of return (monthly), then the sample Sharpe measure of the index is $S(p) = 0.0394$. However, a value of $\lambda_2 = 0.3541$ implies that the ex post efficient portfolio has a Sharpe measure of $S(p^*) = 0.7420$, which is ten times as high as $S(p)$. As a result, the rejection of index efficiency comes as no surprise.

One may argue that the rejection of the efficiency of the index is caused by either the 'size effect' or the 'January effect' [see, e.g., Keim (1983) and Lamoureux and Sanger (1989)]. Since the portfolios used here are obtained by the industry grouping procedure of Sharpe (1982), it seems plausible to assume the absence of the 'size effect' and focus our attention instead on the 'January effect'. For example, if the market model is reasonably well-specified for February–December and the index is efficient, an inclusion of the abnormal returns in January may result in an inaccurately higher value of the statistic λ_2, which leads to rejections of the efficiency hypothesis. Therefore, we rerun the above test with the returns in January deleted. Although the results are not reported here, they are very similar to those with the returns on January included. In almost all the periods, there are no significant differences in either the statistic λ_2 or the P-values. No matter how weak or strong the January effect in industry-grouped portfolio, it does not appear to have much impact on the multivariate efficiency tests, which is consistent with a similar conclusion reached by Shanken (1985). In summary, it seems that, given the assumption that the market model is well-specified in those periods, the rejection of efficiency is most likely caused by the fundamentally inefficient behavior of the index.

5. Conclusions and topics for future research

This paper presents an eigenvalue test for the efficiency of a given portfolio when there does not exist a riskless asset. To facilitate its application, optimal upper and lower bounds are provided which can be easily computed. The eigenvalue test makes it possible to draw statistical inferences on the efficiency of a given portfolio not only in the context of the zero-beta CAPM but also with respect to other linear pricing models. The statistical and economic interpretations of our test are explored, and then as an application of the eigenvalue test, we study the efficiency of the CRSP value-weighted index in a single zero-beta CAPM model with twelve industry groups. Using monthly data for consecutive five-year periods from 1926–1986, both with and without January returns, efficiency is rejected for most of the periods at the usual significance levels.

Several topics for future research can be suggested:

(1) An accuracy study of the maximum likelihood estimation through further application of our approach. A practitioner wishing to use the zero-beta CAPM in portfolio analysis may be interested not only in getting good estimates of the parameters but also in assessing the possible noise in the estimation. While both Shanken (1986) and this paper have presented the maximum likelihood estimators of the parameters, such as the zero-beta rate γ_0, the standard errors or the finite moments properties are not known.

(2) Further empirical studies of indices other than the CRSP value-weighted index studied here, or even an international version of the zero-beta CAPM. It would also be interesting to explore the causes of the rejections found in this paper. For example, rejection may be due to misspecification of the market model in which case Brennan's (1970) after-tax CAPM which incorporates the effects of both taxes and dividends may better manifest the efficiency of the indices. Alternatively, it would be interesting to extend the study to Litzenberger and Ramaswamy's (1979) model that explicitly considers taxes, dividends, and margin constraints.

(3) A generalization of the test to the multi-beta CAPM. Recall the market model of section 2:

$$r_{it} = \alpha_i + \beta_i r_{pt} + \varepsilon_{it}, \qquad i = 1, \ldots, N. \tag{1}$$

If we regard r_{pt} as a $1 \times K$ vector containing the returns on the K reference portfolios, then β_i becomes $\boldsymbol{\beta}_i$, a $K \times 1$ vector consisting of the K betas of asset i with respect to the reference portfolios. The economic interpretations of the model are well known and efficiency implies the following constraints on the parameters:

$$\mathrm{H}_0\colon \quad \alpha_i = \gamma_0(1 - \boldsymbol{\beta}_i'\mathbf{1}_K), \qquad i = 1, \ldots, N, \tag{2'}$$

where $\mathbf{1}_K$ is a $K \times 1$ vector of 1's. The test statistic and its optimal and lower bounds may be generalized and its exact distribution may be obtained explicitly in principle. However, instead of a simple infinite series expression as Theorem 4 of appendix A, zonal polynomials such as in James (1964) [alternatively see Muirhead (1982, ch. 7)] have to be used. Nevertheless, the computational method of appendix B may be extended to evaluate the exact P-values. Furthermore, it appears that the test is also relevant to Shanken's (1989) more general version of the multi-beta CAPM.

(4) A Bayesian posterior analysis and an odds ratio test of the zero-beta CAPM. Prior information and prior densities may be explored, based on which the posterior densities are easily obtained. The omegas are the functions of interest for posterior analysis. Although the integrations that have to be evaluated cannot be done analytically, the Monte Carlo integration with importance sampling [Geweke (1988, 1989)] can be utilized to obtain reliable numerical results, as demonstrated in Harvey and Zhou (1990) and Zhou (1990). Notice that the null is valid if and only if the parameter $\omega_2 = 0$. Thus it is possible to obtain odds ratios, which lead to the posterior probabilities for the validity of the efficiency hypothesis.

(5) Applying the approach to alternative linear pricing models. For example, the exact distribution clearly offers a lower bound on the test of the CCAPM of Breeden, Gibbons, and Litzenberger (1989) and has the potential to give the exact distribution. In addition, it is of great interest to apply the method into the testing of the arbitrage pricing theory and the conditional CAPM models [e.g., Gibbons and Ferson (1985)]. Moreover, the method can be extended to test the rank of the unknown regression parameter matrix, making it widely applicable in economics, especially in simultaneous models. One such application is in the study of the term structure of interest rates along the lines of Hansen and Hodrick (1983), Campbell (1987), and Stambaugh (1988).

Appendix A: Proofs of the theorems

All of the proofs are based upon the following result about the joint density of λ_1 and λ_2:

Lemma. *If* $N \geq 2$ *and* $T \geq N+2$, *the joint density of* λ_1 *and* λ_2 *is given by*

$$d(\lambda_1, \lambda_2) = k_1(\lambda_1 - \lambda_2)(\lambda_1\lambda_2)^{n_1}[(1-\lambda_1)(1-\lambda_2)]^{n_2}$$

$$\times \operatorname{etr}\left(-\tfrac{1}{2}\tilde{\boldsymbol{\Omega}}\right) {}_1F_1\left(\frac{T}{2}; \frac{N}{2}; \frac{\tilde{\boldsymbol{\Omega}}}{2}, \boldsymbol{\Lambda}\right),$$

where $\tilde{\boldsymbol{\Omega}} = \boldsymbol{M}\boldsymbol{\Sigma}^{-1}\boldsymbol{M}'$, $\boldsymbol{M} = \boldsymbol{L}\boldsymbol{\Phi}$, $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2)$, etr *is the exponential function of the trace operator, and* ${}_1F_1(a; b; \boldsymbol{M}_1; \boldsymbol{M}_2)$ *is the hypergeometric function of matrix arguments* [*see James* (*1964*) *or Muirhead* (*1982, ch. 7*)].

Proof. Notice first that both $X'X$ and $Y'Y$ are assumed to be nonsingular matrices throughout the paper. (These assumptions are harmless.) Since X is fixed in the empirical model, if the return on the market portfolio is not constant, $X'X$ should be invertible. Although Y is a random matrix, the nonsingularity of Σ implies that $Y'Y$ must be nonsingular with probability one [Dykstra (1970)].

To obtain the density, it is worth noting that λ_1 and λ_2 [see (8)] are also nonzero roots of

$$|Y'X(X'X)^{-1}X'Y - \lambda Y'Y| = 0. \tag{A.1}$$

Let H be a $T \times T$ orthogonal matrix such that

$$HX = \begin{pmatrix} L \\ 0 \end{pmatrix} \quad \text{and} \quad Z \equiv HY = \begin{pmatrix} U \\ V \end{pmatrix},$$

where L, U, and V are 2×2, $2 \times N$, and $(T-2) \times N$ matrices, respectively. Then it follows that $Y'X(X'X)^{-1}X'Y = U'U$, and thus we can write (A.1) as

$$|U'U - \lambda(U'U + V'V)| = 0. \tag{A.2}$$

Now notice that

$$\begin{aligned} \operatorname{tr} \Sigma^{-1}(Y - X\Phi)'(Y - X\Phi) &= \operatorname{tr} \Sigma^{-1}(HY - HX\Phi)'(HY - HX\Phi) \\ &= \operatorname{tr} \Sigma^{-1}\left[Z - \begin{pmatrix} L\Phi \\ 0 \end{pmatrix}\right]'\left[Z - \begin{pmatrix} L\Phi \\ 0 \end{pmatrix}\right]. \end{aligned}$$

From the density of Y we obtain the joint density of U and V (up to a constant):

$$|\Sigma|^{-T/2} \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}(U - L\Phi)'(U - L\Phi)\right) \times \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}V'V\right). \tag{A.3}$$

This implies that U and V must be independent and each of the rows has a normal distribution. Consider the joint density of eigenvalues, $\delta_1 \geq \delta_2$, of

$$|U(V'V)^{-1}U' - \delta I| = 0. \tag{A.4}$$

For the distribution of eigenvalues of commonly-used matrices in multivariate analysis, James (1964) provides a detailed discussion and analytically derives the joint densities. As one of the results, the joint density of δ_1 and δ_2 is shown to be [see also Muirhead (1982, p. 454, theorem 10.4.5)]

$$k_1(\delta_1-\delta_2)\frac{(\delta_1\delta_2)^{(N-3)/2}}{[(1+\delta_1)(1+\delta_2)]^{T/2}}$$

$$\times\operatorname{etr}\left(-\tfrac{1}{2}\tilde{\boldsymbol{\Omega}}\right){}_1F_1\left(\frac{T}{2};\frac{N}{2};\frac{\tilde{\boldsymbol{\Omega}}}{2},\boldsymbol{\Delta}(\boldsymbol{I}+\boldsymbol{\Delta})^{-1}\right),$$

where $\boldsymbol{\Delta}=\operatorname{diag}(\delta_1,\delta_2)$. Transforming the variables back to λ, $\lambda_1=\delta_1/(1+\delta_1)$, and $\lambda_2=\delta_2/(1+\delta_2)$, we immediately obtain the lemma. Q.E.D.

Denote by $W_2(N,\boldsymbol{I}_2,\tilde{\boldsymbol{\Omega}})$ and $W_2(T-N,\boldsymbol{I}_2)$ the noncentral and central Wishart densities, respectively. By using James (1964) [or Muirhead (1982, p. 450, theorem 10.4.2)] and the lemma, we have:

Corollary. If $N\geq 2$ and $T\geq N+2$, the joint density of λ_1 and λ_2 has exactly the same form as the joint density of the eigenvalues of $\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}$, where $\boldsymbol{A}\sim W_2(N,\boldsymbol{I}_2,\tilde{\boldsymbol{\Omega}})$ and $\boldsymbol{B}\sim W_2(T-N,\boldsymbol{I}_2)$.

The corollary suggests that we can apply any result about the latter density function to the density function of the former, which is useful in that it is often easier to study the second density function (for which there are many results available).

Proof of Theorem 1

The proof follows from our lemma and Perlman and Oklin (1980). In a different context, the result is the well-known monotonicity of power functions.

Proof of Theorem 2

The proof follows from the corollary and Schott (1984).

Proof of Theorem 3

The proof follows from the corollary and Nanda (1947) or Pillai (1956).

An Explicit Expression for the Exact Density

Based on the lemma and Muirhead (1975), an explicit expression for the exact density of λ_2 under the null is given by:

Theorem 4. Under the null hypothesis that the constraints (2) are true, if $N \geq 2$ and $T \geq N+2$, then the exact density of λ_2 is

$$f(\lambda_2) = k_1 \, \mathrm{e}^{-\omega_1/2} \sum_{j=0}^{\infty} \sum_{k=0}^{j} C_{j,k} \lambda_2^{n_1} (1-\lambda_2)^{2n_2+k+2}$$

$$\times F(-n_1 - j + k, n_2 + 1, n_2 + k + 3, 1 - \lambda_2),$$

where $n_1 = (N-3)/2$, $n_2 = (T-N-3)/2$, $\Gamma(a)$ is the gamma function (evaluated at a), $k_1 = \Gamma(T-1)/4\Gamma(N-1)\Gamma(T-N-1)$, and

$$C_{j,k} = \frac{(T/2)_j(-j)_k\left(\frac{1}{2}\omega_1\right)^j\left(\frac{1}{2}\right)_k \Gamma(n_2+1)\Gamma(k+2)}{(N/2)_j j!(k!)^2 \Gamma(n_2+k+3)},$$

with factorial function $(a)_j \equiv a(a+1)\cdots(a+(j-1))$, $\forall a$, and $F(a,b,c;z)$ is the hypergeometric function (or Gaussian function):

$$F(a,b,c;z) \equiv 1 + \sum_{j=1}^{\infty} \frac{(a)_j(b)_j}{(c)_j j!} z^j.$$

Proof. Under the null, ω_1 is the single nonzero eigenvalue of $\tilde{\boldsymbol{\Omega}}$:

$$|\boldsymbol{L\Phi\Sigma}^{-1}\boldsymbol{\Phi}'\boldsymbol{L}' - \omega \boldsymbol{I}| = 0,$$

or (15). By the lemma, we need to integrate out λ_1 in the joint density. By using Muirhead (1975, corollary 1), we have

$$\begin{aligned}
\Upsilon &\equiv {}_1F_1\left(T/2; N/2; \tilde{\boldsymbol{\Omega}}/2, \boldsymbol{\Lambda}\right) \\
&= \sum_{j=0}^{\infty} \frac{(T/2)_j}{(N/2)_j} \frac{\left(\frac{1}{4}(\lambda_1+\lambda_2)\omega_1\right)^j}{j!} \\
&\quad \times {}_2F_1\left(-\tfrac{1}{2}j; -\tfrac{1}{2}j + \tfrac{1}{2}; 1; \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}\right)^2\right).
\end{aligned}$$

Notice it remains difficult to get a closed-form expression for Υ which contains terms of $(\lambda_1+\lambda_2)$. To eliminate these terms, we apply Rainville (1960, p. 65, theorem 24) with $y/(1-y)=(\lambda_1-\lambda_2)/(\lambda_1+\lambda_2)$ or $y=(\lambda_1-\lambda_2)/(2\lambda_1)$, to obtain the result that

$$\Upsilon=\sum_{j=0}^{\infty}\frac{(T/2)_j}{(N/2)_j}\frac{\left(\frac{1}{2}\omega_1\lambda_1\right)^j}{j!}\,{}_2F_1\left(-j;\tfrac{1}{2};1;\frac{\lambda_1-\lambda_2}{\lambda_2}\right). \tag{A.5}$$

To obtain the marginal density of λ_2, let $(1-\lambda_2)t=1-\lambda_1$ or $t=(1-\lambda_1)/(1-\lambda_2)$ and $\lambda_2=\lambda_2$. Expanding ${}_2F_1$ into a series and integrating it term by term with respect to t and then using the identity [Rainville (1960, p. 47, theorem 16)]:

$$\begin{aligned}&\int_0^1 t^{n_2}(1-t)^{k+1}\left[1-(1-\lambda_2)t\right]^{n_1+j-k}\,\mathrm{d}t\\&=\frac{\Gamma(n_2+1)\Gamma(k+2)}{\Gamma(n_2+k+3)}\\&\quad\times{}_2F_1(-n_1-j+k,n_2+1,n_2+k+3,1-\lambda_2),\end{aligned}$$

we immediately obtain Theorem 4. Q.E.D.

Appendix B: The computation of the exact distribution

We evaluate the exact distribution only under the null that the restrictions (2) are valid, for our method extended to the general case becomes complex. The null case implies that $\omega_2=0$, but ω_1 remains unknown. Nevertheless, for any given x between 0 and 1, the exact value of $\mathrm{Prob}(\lambda_2<x\,|\,\omega_1)$ can be evaluated at the maximum likelihood estimator of ω_1. Therefore we assume below that ω_1 is given.

The corollary offers a straightforward method for obtaining reliable computational results. By the corollary, $\lambda_2=\delta_2/(1+\delta_2)$, where δ_2 is the smallest eigenvalue of the matrix $\boldsymbol{A}\boldsymbol{B}^{-1}$ with $\boldsymbol{A}\sim W_2(N,\boldsymbol{I}_2,\tilde{\boldsymbol{\Omega}})$ and $\boldsymbol{B}\sim W_2(T-N,\boldsymbol{I}_2)$. Now, by the familiar property of the hypergeometric function of matrix arguments, we know $\tilde{\boldsymbol{\Omega}}$ matters in the distribution only through its eigenvalues. This allows us to replace $\tilde{\boldsymbol{\Omega}}$ by $\mathrm{diag}(\omega_1,\omega_2)=\mathrm{diag}(\omega_1,0)$. So, by Muirhead (1982, p. 448, theorem 10.3.8), we have $\boldsymbol{A}=\boldsymbol{U}'\boldsymbol{U}$ and $\boldsymbol{B}=\boldsymbol{V}'\boldsymbol{V}$ whose elements are random variables or zeros as follows:

$$\boldsymbol{U}=\begin{pmatrix}\sqrt{\chi_N^2(\omega_1)} & N(0,1)\\ 0 & \sqrt{\chi_{N-1}^2}\end{pmatrix}\quad\text{and}\quad\boldsymbol{V}=\begin{pmatrix}\sqrt{\chi_{T-N}^2} & N(0,1)\\ 0 & \sqrt{\chi_{T-N-1}^2}\end{pmatrix},$$

where $\chi^2_N(\omega_1)$ denotes a noncentral χ^2 random variate with noncentrality parameter ω_1, χ^2_{N-1} denotes a central χ^2, $N(0,1)$ a standard normal, and all the random variates are independent from one another. It is thus straightforward to generate U and V. For each such generation, δ_2, as the smallest eigenvalue of

$$|(UV^{-1})'(UV^{-1}) - \delta_2 I_2| = 0,$$

can be analytically solved. Thus $\mathrm{Prob}(\delta_2 < x/(1-x)|\omega_1)$, which is equal to $\mathrm{Prob}(\lambda_2 < x|\omega_1)$, can be easily determined by generating a number of U and V and then computing the percentage of the δ_2's which satisfy $\delta_2 < x/(1-x)$. The P-value is then given by one minus this percentage. Furthermore, tables can be made as a function of T, N, and ω_1.

An alternative method is to evaluate the integration of $f(\lambda_2)$ (see Theorem 4) over the interval $[0, x)$ as a series of incomplete beta functions, or else to integrate $f(\lambda_2)$ numerically. Still another method is to evaluate the density by computing a six-dimensional integral suggested by the corollary. However, we advocate the previous approach, which is essentially a Monte Carlo integration approach [see, e.g., Geweke (1988, 1989)], because it is easy to implement and generally applicable. In our present applications, the computational load is very light and so we choose the number of repetitions to be 100,000, which should deliver very satisfactory results. Indeed, for $\omega_1 = 0$, we obtain the same value (0.124), which is the last entry in table 1 for the period 1931/2–1936/1, as the analytical result from Theorem 3. On a SPARCstation 1 +, all computations reported in the paper were finished in 40 minutes. For readers who are interested in more details, a Fortran program is available upon request.

References

Anderson, Theodore W., 1984, An introduction of multivariate statistical analysis, 2nd ed. (Wiley, New York, NY).

Bartlett, M.S., 1938, Further aspects of the theory of multiple regression, Proceedings of the Cambridge Philosophical Society 34, 33–47.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, Journal of Business 45, 444–454.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265–296.

Breeden, Douglas T., Michael Gibbons, and Robert H. Litzenberger, 1989, Empirical tests of the consumption based CAPM, Journal of Finance 44, 231–262.

Brennan, Michael J., 1970, Taxes, market valuation and corporate financial policy, National Tax Journal 23, 417–427.

Campbell, John Y., 1987, Stock returns and the term structure, Journal of Financial Economics 18, 373–399.

Constantine, A.G., 1963, Some noncentral distribution problems in multivariate analysis, Annals of Mathematical Statistics 34, 1270–1285.

Dybvig, Philip H., 1983, An explicit bound on individual assets' deviations from APT pricing in a finite economy, Journal of Financial Economics 12, 483–496.

Dykstra, Richard L., 1970, Establishing the positive definiteness of the sample covariance matrix, Annals of Mathematical Statistics 41, 2153–2154.

Erdelyi, A., W. Magnus, F. Oberhettinger, and F. Tricomi, 1953, Higher transcendental functions, Vol. I (McGraw-Hill, New York, NY).

Ferson, Wayne E. and Campbell R. Harvey, 1990, The variation of economic risk premiums, Journal of Political Economy 99, 385–415.

Geweke, John F., 1988, Antithetic acceleration of Monte Carlo integration in Bayesian inference, Journal of Econometrics 38, 73–90.

Geweke, John F., 1989, Bayesian inference in econometric models using Monte Carlo integration, Econometrica 57, 1317–1339.

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, Journal of Financial Economics 10, 3–27.

Gibbons, Michael R. and Wayne Ferson, 1985, Testing asset pricing models with changing expectations and an unobservable market portfolio, Journal of Financial Economics 14, 217–236.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, Econometrica 57, 1121–1152.

Grinblatt, Mark and Sheridan Titman, 1983, Factor pricing in a finite economy, Journal of Financial Economics 12, 497–507.

Hansen, Lars P. and Robert J. Hodrick, 1983, Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models, in: J.A. Frenkel, ed., Exchange rates and international macroeconomics (University of Chicago Press, Chicago, IL).

Harvey, Campbell R. and Guofu Zhou, 1990, Bayesian inference in asset pricing tests, Journal of Financial Economics 26, 221–254.

Huberman, Gur and Shmuel Kandel, 1987, Mean–variance spanning, Journal of Finance 32, 339–346.

James, Alan T., 1964, Distributions of matrix variates and latent roots derived from normal samples, Annals of Mathematical Statistics 35, 475–501.

Jobson, J.D. and Bob Korkie, 1982, Potential performance and tests of portfolio efficiency, Journal of Financial Economics 10, 433–466.

Kandel, Shmuel, 1984, The likelihood ratio tests statistic of mean–variance efficiency without a riskless asset, Journal of Financial Economics 13, 575–592.

Kandel, Shmuel, 1986, The geometry of the likelihood estimator of the zero-beta return, Journal of Finance 31, 339–346.

Kandel, Shmuel and Robert F. Stambaugh, 1987, On correlations and inferences about mean–variance efficiency, Journal of Financial Economics 18, 61–90.

Kandel, Shmuel and Robert F. Stambaugh, 1989, A mean–variance framework for tests of asset pricing models, Review of Financial Studies 2, 125–156.

Keim, Donald B., 1983, Size related anormalies and stock return seasonality: Further empirical evidence, Journal of Financial Economics 12, 13–32.

Kloek, Teun and Herman K. van Dijk, 1978, Bayesian estimates of equation system parameters: An application of integration by Monte Carlo, Econometrica 46, 1–20.

Lamoureux, Christopher G. and Gary C. Sanger, 1989, Firm size and turn-of-the year effects in the OTC/NASDAQ market, Journal of Finance 34, 1219–1245.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13–37.

Litzenberger, Robert H. and Krishna Ramaswamy, 1979, The effects of personal taxes and dividends on capital asset prices: Theory and empirical evidence, Journal of Financial Economics 7, 163–196.

MacBeth, James D., 1975, Tests of two parameter models of capital market equilibrium, Ph.D. dissertation (University of Chicago, Chicago, IL).

MacKinlay, A. Craig, 1987, On multivariate tests of the capital asset pricing model, Journal of Financial Economics 18, 341–372.

Markowitz, Harry M., 1952, Portfolio selection, Journal of Finance 7, 77–91.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, Econometrica 41, 867–887.
Muirhead, Robb J., 1975, Expressions for hypergeometric functions of matrix argument with applications, Journal of Multivariate Analysis 5, 283–293.
Muirhead, Robb J., 1982, Aspects of multivariate statistical theory (Wiley, New York, NY).
Muirhead, Robb J. and Yasuko Chikuse, 1975, Approximations for the distribution of the extreme latent roots of three matrices, Annals of the Institute of Statistical Mathematics 27, 473–478.
Nanda, D.N., 1947, Distribution of a root of a determinantal equation, Annals of Mathematical Statistics 18, 47–57.
Perlman, Michael D., 1980, Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations, Annals of Statistics 8, 247–263.
Perlman, Michael D. and Ingram Oklin, 1980, Unbiasedness of invariant tests for MANOVA and other multivariate problems, Annals of Statistics 8, 1326–1341.
Pillai, K.C.S., 1956, On the distribution of the largest or the smallest root of a matrix in multivariate analysis, Biometrika 43, 122–127.
Rainville, Earl D., 1960, Special functions (Macmillan, New York, NY).
Roll, Richard, A critique of the asset pricing theory's tests, Part I: On past and potential testability of the theory, Journal of Financial Economics 4, 129–176.
Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341–360.
Ross, Stephen A., 1980, A test of the efficiency of a given portfolio, Prepared for the World Econometrics Meeting, Aix-en-Provence, France.
Saw, John G., 1974, A lower bound for the distribution of a partial product of latent roots, Communications in Statistics 3, 665–669.
Schott, James R., 1984, Optimal bounds for the distributions of some test criteria for tests of dimensionality, Biometrika 71, 561–567.
Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, Journal of Financial Economics 14, 327–348.
Shanken, Jay, 1986, Testing portfolio efficiency when the zero-beta rate is unknown: A note, Journal of Finance 41, 269–276.
Shanken, Jay, 1987a, Multivariate proxies and asset pricing relations: Living with Roll's critique, Journal of Financial Economics 18, 91–110.
Shanken, Jay, 1987b, A Bayesian approach to testing portfolio efficiency, Journal of Financial Economics 19, 195–216.
Shanken, Jay, 1989, On the estimation of beta-pricing models, A paper presented at the 1989 American Finance Meeting in Atlanta, GA.
Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.
Sharpe, William F., 1982, Factors in New York Stock Exchange security returns, 1931–1979, Journal of Portfolio Management 8, 5–19.
Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, Journal of Financial Economics 10, 237–268.
Stambaugh, Robert F., 1988, The information in forward rates: Implications for models of the term structure, Journal of Financial Economics 21, 41–70.
Sugiyama, T., 1972, Percentile points of the largest latent root of a matrix and power calculations for testing $\boldsymbol{\Sigma} = \boldsymbol{I}$, Journal of the Japanese Statistical Society 3, 1–8.
Zhou, Guofu, 1990, A Bayesian analysis of time series with applications to stationarity and causality, Ph.D. dissertation (Duke University, Durham, NC).