23 Statistical methods in tests of portfolio efficiency: A synthesis

Article *in* Handbook of Statistics · December 1996

DOI: 10.1016/S0169-7161(96)14025-6

Jay Shanken
Emory University

G. S. Maddala and C. R. Rao, eds., *Handbook of Statistics, Vol. 14*

23

Statistical Methods in Tests of Portfolio Efficiency: A Synthesis*

Jay Shanken

This paper provides a review of statistical methods that have been used in testing the mean-variance efficiency of a portfolio, with or without a riskless asset. Topics considered include asymptotic properties of the two-pass methodology for estimating coefficients in the linear relation between expected returns and betas; the errors-in-variables problem in two-pass estimation; small-sample properties and economic interpretation of multivariate tests of expected return linearity in beta.

1. Introduction

The tradeoff between risk and expected return in the formation of an investment portfolio is a central focus of modern financial theory. In this review, we explore the ways in which statistical methods have been used to evaluate this tradeoff and test the "efficiency" of a portfolio. The emphasis is on methodology rather than empirical findings.

Formally, a portfolio is characterized by a set of security or asset weights that sum to one. The return on the portfolio is the corresponding weighted average of security returns. Here, return refers to the change in price over the period plus any cash flow received (interest or dividends) at the end of the period, all divided by the beginning-of-period price. In a single-period context, if the rates of return on the available investments are jointly normally distributed, then a risk-averse (strictly concave utility function) investor will exhibit a preference for expected return and an aversion to variance of return.[1] In order to maximize expected utility, such an investor will combine securities in what is termed an *efficient portfolio*, i.e., a portfolio that (i) has the smallest possible variance of return given its expected return and (ii) the largest possible expected return given its variance.

* Thanks to Dave Chapman, Aditya Kaul, Jonathan Lewellen, John Long, Ane Tamayo, and Guofu Zhou for helpful comments on earlier drafts.

[1] See Chamberlain (1983) for more general conditions.

More generally, any portfolio that satisfies condition (i) is said to be a *minimum-variance portfolio*.[2] We now consider statistical methods for testing whether a given portfolio satisfies these conditions.

Assume that a set of N risky securities and a portfolio p are given. The return on security i over period t is denoted R_{it} and the return on the portfolio is R_{pt}. The $N+1$ returns are taken to be linearly independent. It is well known [Fama (1976), Roll (1977) and Ross (1977)] that p is a minimum-variance portfolio if and only if there is a constant, γ_{0p}, such that the vector of expected security returns, $r_1, \dots r_N$, is an exact linear function of the vector of security betas on R_p; i.e.,

$$r_i = \gamma_{0p} + \beta_i(r_p - \gamma_{0p}), \; i = 1, 2, \dots, N, \tag{1.1}$$

where r_p is the expected return on portfolio p and the betas are slope coefficients in the time-series regressions of (realized) security returns on the returns of p:

$$R_{it} = \alpha_i + \beta_i R_{pt} + \varepsilon_{it} \quad \textit{and} \quad \mathrm{E}(\varepsilon_{it}) = \mathrm{E}(\varepsilon_{it} R_{pt}) = 0. \tag{1.2}$$

Moreover, a minimum-variance portfolio p is efficient if and only if the additional restriction, $r_p > \gamma_{0p}$, is satisfied, where the "zero-beta rate," γ_{0p}, is the expected return on any security (or portfolio) that has a beta of zero relative to p. Thus, in the efficient portfolio case, expected return is an increasing linear function of beta.

The equivalence between the minimum-variance property and the expected return-beta relation arises from the fact that the beta coefficient determines the *marginal* contribution that a security makes to the total risk (variance) of portfolio p. This equivalence is of great import for the testing of portfolio efficiency since the hypothesis can be viewed as a restriction on the parameters in the multivariate linear regression system (1.2).

Combining (1.1) and (1.2), we have the hypothesis

$$H_{01}: \quad \alpha_i = \gamma_{0p}(1 - \beta_i), \quad i = 1, \dots, N, \tag{1.3}$$

a joint restriction on the intercepts and slopes in the time-series regressions. This condition asserts the existence of a single number, γ_{0p}, for which the intercept-slope relation holds across the given N securities. If investors can borrow or lend at a known riskfree rate, r_f, and p is presumed efficient with respect to the set of all portfolios of both the risky securities and the riskless asset, then $\gamma_{0p} = r_f$.[3] Otherwise, γ_{0p} is unknown and must be estimated.

According to H_{01}, the ratio of alpha to one minus beta for any N–1 securities is equal to the ratio for the remaining security. Thus, $2N$ parameters (the alphas and betas) are reduced to a set of just $N+1$ parameters (the betas and γ_{0p}) under the

[2] It is convenient to exclude from this definition the global minimum variance portfolio, i.e., the portfolio with the lowest variance of return, regardless of expected return. Also, we assume below that at least two portfolios have distinct expected returns.

[3] A negative position in the riskless asset amounts to borrowing, and the riskless rate is assumed to be the same for both borrowing and lending.

N–1 restrictions implicit in (1.3) [Gibbons (1982)]. The restriction is nonlinear in a statistical sense when γ_{0p} is unknown, since γ_{0p} and β_{ip} enter multiplicatively and both must be estimated.

2. Testing efficiency with a riskless asset

2.1. Univariate tests

Before going on to the general case, we focus on the much simpler scenario in which γ_{0p} is known and equal to the return on a riskless security. In this case, it is convenient to consider the *excess-return* version of the system (1.2); i.e., we now view R_{it} as the return on security i in excess of the riskless rate and r_i is the corresponding expected excess return.[4] The excess zero-beta rate in (1.1) is then zero, and hence, by (1.3), so are the time-series regression intercepts in (1.2). Thus, the main hypothesis of interest is now

$$H_{02}: \quad \alpha_i = 0, \quad i = 1, \ldots, N. \tag{2.1}$$

A test of this restriction on the excess-return regression model is a test that the given portfolio satisfies the minimum-variance property in the presence of a riskless asset.

An early study by Black, Jensen, and Scholes (1972) examines the efficiency of an equal-weighted stock market index using monthly excess returns over the period 1931–65. The equal-weighted index is used as a proxy for the value-weighted market portfolio of all financial assets. The latter portfolio is predicted to be an efficient portfolio under the assumptions of the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), a theory of financial market equilibrium. Black, Jensen, and Scholes report t-tests on the intercepts for a set of ten stock portfolios, with two of the ten significant at the 0.05 level (two-sided). The estimated intercepts are negative for the portfolios with relatively high estimated betas and positive for those with lower betas.

2.2. Multivariate tests

2.2.1. F-test on the intercepts

More recently, Gibbons, Ross, and Shanken (1989) apply a multivariate F-test of H_{02} to the Black, Jensen, and Scholes data and fail to reject the joint hypothesis that the intercepts are all zero [see related work by Jobson and Korkie (1982, 1985) and MacKinlay (1987)]. Use of the F-test presumes that the disturbances in (1.2) are independent over time and jointly normally distributed, each period,

[4] In this context, all probability statements can be viewed as conditional on the riskless rate series. In general, the total return and excess return time-series specifications need not be strictly consistent when the riskless rate varies over time.

with mean zero and nonsingular cross-sectional covariance matrix Σ, conditional on the vector of returns, R_p. Let T equal the length of the given time-series of returns for the N assets and portfolio p. The F-statistic, with degrees of freedom N and T–N–1, equals $(T\text{–N–}1)N^{-1}(T\text{–}2)^{-1}$ times the Hotelling T^2 statistic

$$Q \equiv T\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}/[1+\bar{R}_p^2/s_p^2], \tag{2.2}$$

where $\bar{R}_p$ and s_p are the sample mean and standard deviation of excess return for p; $\hat{\alpha}$ is the N-vector of OLS intercept estimates and $\hat{\Sigma}$ is the unbiased estimate of Σ, computed from crossproducts of OLS residuals divided by T–2.

The conditional covariance matrix of the alpha estimates, given R_p, equals the product of the denominator in (2.2), a function of R_p, and the residual covariance matrix, Σ, divided by T. Thus, the T^2 statistic is a quadratic form in the alphas, weighted by the inverse of the estimated covariance matrix of the alphas. When $N = 1$, Q is just the square of the usual univariate t-statistic on the intercept. More generally, it can be shown that Q is the maximum squared (univariate) t-statistic for alpha, where the maximum is taken over all portfolios of the N assets.[5]

Since Q has the same distribution unconditionally and conditional on R_p, the F-test does not require that R_p itself be normally distributed; the disturbances are assumed to be jointly normally distributed, however. Affleck-Graves and McDonald (1989) present simulation evidence indicating that the multivariate tests are robust to deviations from normality of the residuals, although MacKinlay and Richardson (1991) report a sensitivity to conditional heteroskedasticity. Zhou (1993) reaches similar conclusions.

Given our assumptions, the zero intercept restriction implies that expected excess returns for the N assets are proportional to the betas, both unconditionally and *conditional* on R_p. Extremely high or low returns for p, in a given sample period, tell us nothing about whether the intercepts are zero. Accordingly, the test statistic in (2.2) depends on the mean return of portfolio p only through its squared value, not its level. Portfolio efficiency entails the additional restriction that the *ex ante* mean excess return, r_p, exceeds zero, however, and this hypothesis can and should be evaluated separately through a simple t-test on the sample mean, $\bar{R}_p$.[6]

2.2.2. Power and economic interpretation of the F-test

Gibbons, Ross, and Shanken (1989) provide an interesting economic interpretation of the F-statistic that requires some additional notation. Let SH(p) equal the ratio, r_p/σ_p, of expected excess return to standard deviation of return for portfolio

[5] See Gibbons, Ross, and Shanken, section 6, for a proof and an economic interpretation of this relation.

[6] Since Q is independent of $\bar{R}_p$ under the null hypothesis of efficiency, the p-value for the joint hypothesis that the intercepts are zero and $r_p > 0$ (probability that at least one of the two statistics is in the relevant tail areas) equals the sum of the two p-values minus their product.

p and let $sh(p)$ be the corresponding sample quantity. These reward/risk measures are referred to as *Sharpe ratios.* Using this terminology, an efficient portfolio can be characterized as one with the maximum possible Sharpe ratio, while a minimum-variance portfolio maximizes the squared (absolute) Sharpe ratio.[7] If portfolios are plotted as points in a graph with expected excess return on the vertical axis and standard deviation of return on the horizontal axis, then the Sharpe ratio for p equals the slope of a ray through p emanating from the origin; in the case of a minimum-variance portfolio, the ray is tangent to the graph.

Gibbons, Ross, and Shanken show that Q in (2.2) equals

$$T[sh(*)^2 - sh(p)^2]/[(1 + sh(p)^2], \tag{2.3}$$

where $sh(*)$ is the sample Sharpe ratio with maximum squared value over all portfolios. Examining the numerator of (2.3), we see that, other things equal, the F-statistic is larger the lower is the squared Sharpe ratio for portfolio p in relation to the maximum squared sample ratio. Thus, the F-statistic is large when p is "far" from the *ex post* minimum-variance frontier.

Of course, in any sample, there will be portfolios whose *sample* Sharpe ratios dominate p's, even if p is truly an *ex ante* minimum-variance portfolio. The F-test provides a basis for inferring whether the difference, $sh(*)^2 - sh(p)$, is within the range of random outcomes that would reasonably be anticipated under the null hypothesis. This assessment naturally depends on the precision of the alpha estimates.

Given the assumptions above, Gibbons, Ross, and Shanken show, further, that the F-statistic is distributed, under the alternative, as noncentral F with noncentrality parameter

$$\lambda = T[SH(*)^2 - SH(p)^2]/[1 + sh(p)^2]. \tag{2.4}$$

Again, the distribution is conditioned on R_p, the independent variable in the time-series regressions, and depends on R_p through the *ex post* Sharpe ratio. In this context, $sh(p)$ may be viewed as a constant, and hence the noncentrality parameter in (2.4) is just the (conditional) population counterpart of the sample statistic, Q, in (2.3). Under the null hypothesis that p is a minimum-variance portfolio, p attains the maximum squared *ex ante* ratio. In this case, λ equals zero and we have a central F distribution as earlier.

The power of the F-test is known to be an increasing function of the noncentrality parameter. Therefore, given $sh(p)$, power is greater the further is the square of $SH(p)$ from the maximum squared (population) ratio; i.e., the greater is the deviation from ex ante efficiency in this metric. Holding the ex ante deviation constant, power decreases as the square of $sh(p)$ increases, reflecting the lower

[7] See Merton (1973a) and Litzenberger and Huang (1988).

(conditional) precision with which the intercepts are estimated when this sample quantity is high.

In order to implement the F-test, the residual covariance matrix, $\hat{\Sigma}$, must be invertible, which requires that N be at most equal to T–2. Analysis in Gibbons, Ross, and Shanken (1989) suggests that much smaller values of N should be used in order to maximize power, however. This is related to the fact that the number of covariances that must be estimated increases rapidly with the number of assets. Although increasing N can increase the noncentrality parameter in (2.4), by increasing the maximum Sharpe measure, apparently this benefit is eventually offset by the additional noise in estimating Σ and its inverse.

Given the thousands of stocks available for analysis and the requirement that N be (much) less than T, some procedure is needed to reduce the number of assets. Although subsets of stocks could be used, the test is more commonly applied to portfolios of stocks. This has the advantage, for a given N, of reducing the residual variances, thereby increasing the precision with which the alphas are estimated.[8] On the other hand, as Roll (1979) has noted, individual stock expected return deviations can cancel out in portfolios, which would reduce power. The expected power of the test thus depends on the researcher's prior beliefs as to the likely sources of portfolio inefficiency.[9]

2.3. Other tests

The likelihood ratio test (LRT) and the Lagrange multiplier (Rao's score) test statistics are both monotonic transformations of the T^2 statistic (modified Wald test) in (2.2) and thus need not be considered separately from the F-test.[10] In particular,

$$LRT = T\ln[1 + Q/(T-2)]. \tag{2.5}$$

Lo and MacKinlay (1990) have emphasized that the use of portfolio grouping in multivariate tests, together with the exploration of a wide variety of potentially relevant firm ranking variables, can lead to substantial "data-snooping" biases; i.e., the appearance of statistical significance even when the null hypothesis of efficiency is true. An alternative diagonal version of the multivariate test, suggested by Affleck-Graves and McDonald (1990), is interesting in this regard since it does not require grouping. As such, it also avoids Roll's concerns about the use of portfolio-based tests. The diagonal test appears to have desirable power characteristics in simulations, but the distribution of the test statistic is unknown.

[8] There are additional motivations for the use of portfolios. Some stocks come and go over time and using portfolios allows one to use longer time series than would otherwise be possible. Also, portfolios formed by periodically ranking on some economic characteristic *may* have fairly constant betas even though individual security betas change over time. Note that the composition of each portfolio changes over time in this context.

[9] See the related analysis of power issues in MacKinlay (1995).

[10] See related work by Evans and Savin (1982).

It would be helpful to have some sort of approximate distribution theory for this approach.

In the remainder of this section, we consider several different variations on the multivariate framework – joint confidence intervals, tests of approximate efficiency, Bayesian approaches to testing efficiency, and tests of conditional efficiency.

2.3.1. Joint confidence intervals

In some contexts, one is interested in the mean-variance efficiency of an index primarily for the purpose of obtaining (statistically) efficient estimates of asset expected returns, via the linear relation (1.1). For example, in capital budgeting applications, the required discount rates for a set of projects might equal the expected returns (adjusted for financial leverage) of some industry portfolios. Here, the magnitude of deviations from the expected return relation is important. Shanken (1990, p.110) suggests examining joint confidence intervals for the alphas, in such a case, since the p-value for the F-test is not very informative in this regard.

The simultaneous confidence interval approach exploits the fact, noted earlier, that the T^2 statistic in (2.2) equals the maximum squared univariate t-statistic for the alphas, where the maximum is taken over all possible portfolios of the given assets.[11] The intervals consist of alphas within k sample standard errors of the OLS estimates, where the constant k is the relevant fractile of the T^2 distribution or, equivalently, $N(T-2)(T-N-1)^{-1}$ times the fractile of an F distribution with degrees of freedom N and $T-N-1$. Alternatively, the Bonferroni approach may be used to obtain (conservative) joint confidence intervals for the N alphas. In this case, one divides the designated error probability by N and then computes conventional confidence intervals based on a t distribution with $T-2$ degrees of freedom.

2.3.2. Tests of approximate efficiency

In a portfolio investment context, one may not be interested in the expected returns, alone, but rather in the extent to which a given portfolio deviates from efficiency. This, recall, is reflected in the noncentrality parameter λ, in (2.4), which depends on both the alphas and the residual covariance matrix, Σ. Kandel and Stambaugh (1987) and Shanken (1987b) utilize the multivariate framework to formulate tests of *approximate* efficiency. This enables the researcher to test for "economically significant" departures from mean-variance efficiency. It is also of interest in testing positive theories like the CAPM, mentioned earlier.

Roll (1977) emphasizes that inferences about the efficiency of a stock index proxy do not tell us whether the true market portfolio is efficient, as required by

[11] See Morrison (1976), Chapter 4, for a discussion of joint confidence intervals. Asymptotic versions of these methods [e.g., Shanken (1990)] based on chi-square or normal distributions follow the same logic.

the asset pricing theory. Kandel and Stambaugh and Shanken show, however, that efficiency of the true market portfolio, along with an a priori belief about the correlation between the proxy and the market, can be used to bound the extent to which the proxy is inefficient. If the bound is violated, efficiency of the true market portfolio is rejected.

For example, Shanken rejects efficiency of the true market portfolio, over the period 1953–83, assuming the correlation with an equal-weighted stock index proxy exceeds 0.7. This tempers the concerns about testability raised by Roll somewhat, as he also conjectured that most reasonable proxies would be fairly highly correlated with the true market portfolio, whether the latter is efficient or not.

2.3.3. Bayesian tests of efficiency

Making use of the fact that the distribution of the test statistic for the minimum-variance property is known under both the null and the alternative, given normality, Shanken (1987a) explores a Bayesian approach to testing portfolio efficiency. Harvey and Zhou (1990) and Kandel, McCulloch, and Stambaugh (1995) extend this analysis by considering prior distributions formulated over the entire parameter space of the multivariate regression model.[12] The relation (2.4) is important in this context, as it facilitates an assessment of the economic significance of deviations from the null hypothesis and the related formulation of meaningful priors on the unknown parameters.

2.3.4. Tests of conditional efficiency

We have assumed, thus far, that asset betas are constant over time. However, if we condition on variables characterizing different states of the economy, betas may well vary. The regression framework is easily extended to accommodate changes in the betas if one is willing to specify the relevant state variables, say interest rates, and postulate some functional relation to beta.

For example, suppose there is a single, stationary, mean-zero state variable, z_{t-1}, known at the beginning of period t, and the conditional beta is

$$\beta_{it-1} = \overline{\beta}_i + c_i z_{t-1}. \tag{2.6}$$

Here, $\overline{\beta}_i$ is the long-run average beta for security i and c_i indicates the sensitivity of i's conditional beta to variation in the state variable. Substituting β_{it-1} for β_i in (1.2) and assuming ε_{it} has zero mean conditional on both z_{t-1} and R_{pt},

$$R_{it} = \alpha_i + \overline{\beta}_i R_{pt} + c_i(z_{t-1}R_{pt}) + \varepsilon_{it} \tag{2.7}$$

is an expanded regression equation from which the parameters of interest may be estimated and the zero-intercept restriction tested. This approach to efficiency

[12] Also see related work by McCulloch and Rossi (1990).

tests is developed in Campbell (1985) and Shanken (1990) in the context of an intertemporal CAPM [Merton (1973b)].[13]

In addition to time-varying betas, the expected return or risk of portfolio p may change over time. This does not pose a problem, though, since the regression analysis is conditioned on the returns for p, as noted earlier. An F-test of the joint zero-intercept restriction is still appropriate if the disturbances in (2.7) have constant variance (over time) conditional on both R_{pt} and z_{t-1}. Shanken (1990) finds strong evidence of conditional residual heteroskedasticity, however, and employs an asymptotic chi-square test based on the heteroskedasticity-consistent covariance matrix of the intercept estimates [White (1984)]. This approach is also adopted by MacKinlay and Richardson (1991), in exploring the impact of residual heteroskedasticity conditional on the contemporaneous realization of R_p.

3. Testing efficiency without a riskless asset

Since U.S. Treasury bills are only nominally riskless, the assumption that there is a riskless asset may not be appropriate if one is concerned with the efficiency of a portfolio in real (inflation-adjusted) terms. Even in the nominal case, if there are restrictions on borrowing [Black (1972)], or an investor's riskless borrowing rate exceeds the T-bill rate [Brennan (1971)], then the zero-beta rate for an efficient portfolio can be greater than the T-bill rate and must be estimated. In this section, therefore, we treat γ_{0p} as an unknown parameter and consider tests of the nonlinear restriction (1.3). The regression variables in (1.2) can now be viewed as either total returns or excess returns; in the latter case, γ_{0p} is the excess zero-beta rate.

3.1. Traditional two-pass estimation techniques

Given the "bilinear" nature [Brown and Weinstein (1983)] of the relation (1.3), an intuitively appealing approach to estimation entails first, estimating the alphas and betas from time-series regressions (1.2), for each security, and then running a cross-sectional regression of the N alpha estimates on one minus the N beta estimates (no constant) in order to estimate γ_{0p}. This is effectively the approach adopted by Black, Jensen, and Scholes (1972) [see related discussion in Blume and Friend (1973)].

Another approach, essentially that of Fama and MacBeth (1973), is to regress the cross-section of mean security returns on the betas and a constant.[14] The intercept in this cross-sectional regression (CSR) is taken as the estimate of γ_{0p}

[13] Also see related work by Ferson, Kandel, and Stambaugh (1987) and Harvey (1989).

[14] There are many variations on this approach. Here, we assume that each asset beta is estimated from a single time-series regression over the entire period. See Jensen (1972) for a review of the early development of the literature.

and the slope coefficient on beta is an estimate of $\gamma_{1p} \equiv r_p - \gamma_{0p}$.[15] We focus primarily on the Fama-MacBeth version of the "two-pass" methodology in the remainder of this review, as it is the approach used most often in the literature.[16]

It is well known that security returns are cross-sectionally correlated, due to common market and industry factors, and also heteroskedastic. For example, small-firm returns tend to be more volatile than large-firm returns. As a result, the usual formulas for standard errors, based on a scalar covariance matrix assumption, are not appropriate for the OLS CSR's run by Black, Jensen, and Scholes and Fama and MacBeth.

Recognizing this problem, Fama and MacBeth run CSR's each month, generating time-series of estimates for both γ_{0p} and γ_{1p}. Means, standard errors, and "t-statistics" are then computed from these time series and inference proceeds in the usual manner, as if the time series are independently and identically distributed. Since the true variance of each monthly estimator depends on the covariance matrix of returns, cross-sectional correlation and heteroskedasticity are reflected in the time series of monthly estimates. However, given the fact that the same beta estimates are used in each monthly cross-sectional regression, the monthly gamma estimates are not serially independent. This dependence is ignored by the traditional two-pass procedure.

The fact that there is an error component common to each of the monthly cross-sectional regressions, due to beta estimation error, makes the small-sample distribution of the mean gamma estimator difficult to evaluate. This is a form of the "generated regressor" problem [Pagan (1983)], as it is sometimes called in the econometrics literature. While consistency (as $T \to \infty$) of the beta estimates implies consistency of the gamma estimates, the "Fama-MacBeth standard errors" computed from the time series of CSR estimates are generally inconsistent estimates of the asymptotic standard errors [Shanken (1983, 1992)].

Let X be the $N \times 2$ matrix $[1_N, \beta]$ of ones and betas and $\hat{X}$ the corresponding matrix, $[1_N, \hat{\beta}]$, with estimated betas. Let R_t be the N-vector of security returns for period t and $\bar{R}$ the N-vector of sample mean returns. In this notation, equation (1.1) implies

$$R_t = X\Gamma + \text{error} = \hat{X}\Gamma + [\text{error} - \gamma_{1p}(\hat{\beta} - \beta)], \tag{3.1}$$

where $\Gamma \equiv (\gamma_{0p}, \gamma_{1p})'$ and "error" is the unexpected component of return. If $A \equiv (X'X)^{-1} X'$ and $\hat{A}$ is the corresponding estimator, then the second-pass estimator of the gammas is $\hat{\Gamma} \equiv (\hat{\gamma}_0, \hat{\gamma}_1)' \equiv \hat{A}\bar{R}$, the mean of the monthly estimators, $\hat{\Gamma}_t \equiv \hat{A}R_t$.

[15] Although γ_{0p} and γ_{1p} are treated as separate parameters, the constraint that $\gamma_{1p} = r_p - \gamma_{0p}$ is implicitly imposed if p is an equal-weighted portfolio of the N assets used in an OLS CSR. The Fama-MacBeth approach can also be used in asset pricing tests where the "factor" is, say, a macroeconomic variable, rather than a portfolio return [e.g., Chen, Roll, and Ross (1986) and Shanken and Weinstein (1990)], and the constraint on the gammas is no longer appropriate.

[16] The various results summarized here all have straightforward extensions to the Black, Jensen, and Scholes specification. See Shanken (1992).

Since the gamma estimates are linear combinations of asset returns, they have an intuitively appealing portfolio interpretation [Fama (1976, Chapter 9)]. Note that AX is a 2×2 identity matrix. Focusing on the first row of A, we see that the estimate of γ_{0p} is the sample mean return on a standard (weights sum to one) portfolio with a beta (weighted-average asset beta) of zero. Similarly, the estimate of the risk premium γ_{1p} is the mean return on a zero-investment portfolio (weights sum to zero) with a beta of one - properties shared by the mean *excess* return for p in the riskless asset case.

Using (3.1), Shanken (1992) shows that the sample covariance matrix of the $\hat{\Gamma}_t'$s, used in computing Fama-MacBeth standard errors, converges to $A\Sigma A' + M$, where M is a 2×2 matrix with $\sigma_p{}^2$ in the lower right corner and zeroes elsewhere.[17] The first term, $A\Sigma A'$, arises from the return residuals in (1.2); the diagonal elements capture the residual variation in the portfolio estimators. The second term, M, accounts for "systematic" variation related to R_p and reflects the fact that the estimates of γ_{0p} and γ_{1p} are returns on portfolios with betas of zero and one, respectively. It follows that the variance of the mean excess return for p is a lower bound on the variance of $\hat{\gamma}_1$.

As noted earlier, the traditional method of computing standard errors for the gamma estimates ignores beta estimation error. When this measurement error is recognized, the asymptotic covariance matrix of $\hat{\Gamma}$, i.e., the covariance matrix of the limiting multivariate normal distribution of $\sqrt{T}(\hat{\Gamma} - \Gamma)$, is:[18]

$$(1 + \gamma_{1p}^2/\sigma_p^2)A\Sigma A' + M, \tag{3.2}$$

The additional term in (3.2) arises from the fact that i) the asymptotic covariance matrix for $\hat{\beta}$ is Σ/σ_p^2 and, ii) the impact of measurement error in $\hat{\beta}$ on the CSR disturbance is, by (3.1), proportional to γ_{1p}. Thus, the traditional standard errors are too low, except for the case in which measurement error in beta is irrelevant, i.e., under the null hypothesis that γ_{1p} equals zero.[19] Asymptotic confidence intervals for the gammas always require the use of adjusted standard errors.

Asymptotically valid standard errors are easily obtained from (3.2) by substituting consistent estimates for the various parameters. For γ_{0p}, this amounts to multiplying the Fama-MacBeth variance by the errors-in-variables adjustment term, $(1 + \hat{\gamma}_1^2/s_p^2)$. For γ_{1p}, s_p^2 is subtracted from the Fama-MacBeth variance before multiplying by the adjustment term and is then added back.

[17] This follows from the fact that the covariance matrix of R_t is $\Sigma + \beta\beta'\sigma_p{}^2$, and that $A\beta$ is the second column of M.

[18] Gibbons (1980) independently derives the asymptotic distribution for the Black, Jensen, and Scholes estimator, a special case of Shanken (1992).

[19] In the "multifactor" context, the adjustment term is a quadratic form in the vector of factor risk-premia with weighting matrix equal to the inverse of the factor covariance matrix. Now, an asymptotic "*t*-statistic" for the null hypothesis that a given factor's risk premium is zero always requires that the adjustment term be incorporated since the other factor premia need not be zero under the null.

3.2. Tests of linearity against a specific alternative

The estimation results above are relevant for testing whether $\gamma_{1p} > 0$, a necessary condition for p to be an efficient portfolio. The analysis *assumes* linearity of the expected return relation, however, and this must be tested separately. The simplest approach is to include other independent variables along with beta in the CSR and test whether the coefficients on the additional variables differ from zero. If so, then beta is not the sole determinant of cross-sectional variation in expected returns and efficiency is rejected. This is the approach taken by Fama and MacBeth (1973), who use beta-squared and residual variance as additional variables. Their evidence supports linearity in beta with a positive risk premium. Consistent with the results of Black, Jensen, and Scholes (1972), they also find that $\hat{\gamma}_0$ is significantly greater than the T-bill rate while $\hat{\gamma}_1$ is less than the mean excess market index return.

Supposing, for simplicity, that the additional cross-sectional variables are constant over time and measured without error, the asymptotic analysis above is easily modified. The additional variables are included in the X matrix and a row and column of zeroes are added to the matrix M, for each extra variable. The asymptotic covariance matrix of the expanded gamma estimator is then given by (3.2). Note that measurement error in the betas affects the standard errors of the additional coefficients, even though the associated independent variables are measured without error. Moreover, the adjustment term, $1 + \gamma_{1p}^2/\sigma_p^2$, must always be included in testing linearity, as γ_{1p} need not be zero under the linearity hypothesis.

In contrast to the multivariate approach, the coefficient-based test of this section requires that the researcher formulate a specific alternative hypothesis to linearity. This can be an advantage if the null hypothesis is rejected, as the test provides concrete information concerning the deviations from linearity. The downside is that the test will have limited power, or none at all, against other potentially relevant alternatives. In addition, there is the inherent invitation to data mining, i.e., the tendency of researchers to explore various alternatives and to publish the results of experiments which, nominally, indicate statistical significance, while discarding the "unsuccessful" experiments.

The multivariate approach to testing has the potential to reject any deviation from expected return linearity with power converging to one as $T \to \infty$. The general nature of this "goodness-of-fit" approach is not without its downside, however, as it is likely to be less powerful against some alternatives than a more focused test. As discussed earlier, it also has its own data-mining problems.

3.3. Maximum likelihood and modified regression estimation

Gibbons (1982) proposes that classical maximum likelihood estimation (MLE) be used to estimate the betas and gammas in (1.3) simultaneously. Since MLE is asymptotically efficient (as $T \to \infty$), it is of interest to compare the efficiency of two-pass estimation to that of MLE. The asymptotic analysis of the OLS second-pass estimator, considered above, easily generalizes to weighted-least-squares (WLS) or generalized-least squares (GLS) versions of the estimator based on sample estimates of the variances and covariances.[20] One merely redefines the matrix A.

It turns out that the asymptotic covariance matrix of the second-pass GLS estimator is the same as that for MLE and hence GLS is asymptotically efficient.[21] In fact, the second-pass GLS estimator of Γ is identical to a one-step Gauss-Newton (linearization) procedure that Gibbons uses to simplify the computations. A straightforward computational procedure for exact MLE was subsequently developed in Kandel (1984) and extended in Shanken (1992).

Although two-pass estimation is consistent, as $T \to \infty$, it suffers from an errors-in-variables problem since β, the independent variable in the cross-sectional relation, is measured with error. Thus, the slope (risk premium) estimator is biased toward zero and the bias is not eliminated asymptotically by increasing the number of securities; i.e., the estimator is not *N-consistent.*[22] Recognizing this, the early studies group securities into portfolios in order to reduce the variance of the error in estimating betas. Concerned about possible reductions in efficiency, elaborate techniques are used to ensure that a substantial spread in portfolio betas is maintained. Assuming the residual covariance matrix is (approximately) diagonal, Black, Jensen, and Scholes (1972) show that the resulting estimator is N-consistent.

In proposing MLE, Gibbons (1982) conjectures that simultaneous estimation of betas and gammas should provide a solution to the errors-in-variables problem. However, simulation evidence in Amsler and Schmidt (1985) indicates that the GLS CSR (they call it "Newton-Raphson") estimator outperforms MLE in terms of mean-square error; GLS is biased upward while MLE is biased downward. Some support for Gibbons' conjecture is provided in Shanken (1992), however, in that a version of MLE with the residual covariance matrix constrained to be diagonal is shown to be N-consistent. Thus, the benefits of MLE may only be realized with a large number of assets.

Although simultaneous estimation of betas and gammas is one path to N-consistency, a modified version of the second-pass estimator is also N-consistent [Litzenberger and Ramaswamy (1979) and Shanken (1992)]. The modified estimator is based on the observation that inconsistency of the second-pass estimator is driven by systematic bias in the lower right element of the $\hat{X}'\hat{X}$ matrix. Conditioning on the time series of returns for portfolio p, we have:

[20] In fact, the same estimator is obtained whether the residual covariance matrix or the (total) covariance matrix of returns is used. This was first noted by Litzenberger and Ramaswamy (1979) for WLS.

[21] This is true despite the fact that the OLS estimator of β, used in the CSR, is inefficient. Also, we assume that the constraint, $\gamma_{1p} = r_p - \gamma_{0p}$, is imposed when appropriate.

[22] More formally, it does not converge to the sample mean return on p minus the zero-beta rate, the "ex post price of risk."

$$\mathrm{E}(\hat{\beta}'\hat{\beta}) = \beta'\beta + tr(\Sigma)/(Ts_p^2), \tag{3.3}$$

where $tr(.)$ is the sum of the diagonal elements of a matrix. Subtracting off $tr(\hat{\Sigma})/(Ts_p^2)$ from the lower right element of $\hat{X}'\hat{X}$, therefore, yields an N-consistent estimator of Γ, provided the residual covariance matrix, Σ, is (approximately) diagonal.[23] The asymptotic distribution of the estimator, as $T \to \infty$, is unaffected by this modification.[24]

Recall, from classical errors-in-variables analysis, that the slope estimator ($\hat{\gamma}_1$) is attenuated toward zero by a factor equal to the variance of the true independent variable (β), divided by the variance of the proxy variable ($\hat{\beta}$). This attenuation factor is less than one, since the latter variance equals the sum of the true variance and the measurement error variance. It is easily verified that the slope component of the modified estimator, described above, equals the regression slope estimator divided by an estimate of the attenuation factor.[25]

The results for MLE and modified CSR estimation suggest that the traditional use of portfolio grouping techniques to address the errors-in-variables problem may be unnecessary. An interesting issue that has not been adequately explored, however, concerns the relative efficiency of (modified) OLS or WLS estimation with a very large set of securities and MLE or GLS estimation with a more modest number of portfolios and a full covariance matrix.

3.4. Multivariate tests

3.4.1. Likelihood ratio and CSR T^2 tests

The first step toward a multivariate test of linearity is taken by MacBeth (1975), who uses a variation on Hotelling's T^2 test to evaluate whether the residuals from Fama-MacBeth CSR's systematically deviate from zero. The test does not fully take into account all of the existing parameter uncertainty, however. Gibbons (1982) formulates a likelihood ratio test (LRT) of the nonlinear restriction (1.3) under the assumption of temporally independent and identically jointly normally distributed returns. Inference is then based on the usual asymptotic chi-square distribution. Unlike MacBeth's approach, the LRT accounts, at least asymptotically, for all relevant parameter uncertainty. As we shall see, though, the asymptotic test suffers from serious small-sample problems.

[23] Unfortunately, this can result in a negative diagonal element in finite samples.

[24] WLS and GLS versions of the modified CSR estimator have also been derived, and additional variables measured without error can be included as in section 3.2. See the references cited earlier. Kim (1995) develops an MLE procedure that accomodates the use of betas estimated from prior data. The modified regression approach can also be applied using prior betas. In this case, T, s_p, and the residual variance estimates substituted in (3.3) come from the time-series regressions used to estimate the betas.

[25] Banz (1981) considers errors-in-variables biases in the gammas when additional variables like firm size are considered along with beta in cross-sectional regressions. The coefficient on beta is still biased toward zero, while the "size effect" is overstated.

The connection between the LRT and the multivariate T^2 test is explored in Shanken (1985). He shows that the relation (2.5) continues to hold for this model with the following expression substituted for Q:

$$Q_{\mathrm{MLE}} \equiv Te'\hat{\Sigma}^{-1}e/(1+\gamma_{1\mathrm{MLE}}^2/s_p^2), \tag{3.4}$$

where

$$e \equiv \bar{R} - \hat{X}\Gamma_{\mathrm{MLE}},$$

$\hat{\Sigma}$ is the unbiased estimate of the residual covariance matrix, s_p^2 is the sample variance of return for portfolio p, and $\Gamma_{\mathrm{MLE}} \equiv (\gamma_{0\mathrm{MLE}}, \gamma_{1\mathrm{MLE}})'$ is the MLE for Γ. Shanken refers to the corresponding test based on the GLS CSR estimate of Γ as the CSR test (CSRT).[26]

3.4.2. Small-sample inference

The test statistic in (3.4) is a direct generalization of Q in (2.2), for the riskless asset case, as $\hat{\alpha}$ is obtained from the residual vector e by substituting the riskless rate and portfolio p's mean excess return for $\gamma_{0\mathrm{MLE}}$ and $\gamma_{1\mathrm{MLE}}$, respectively.[27] In other words, Q in (2.2) is just a constrained version of Q_{MLE} in (3.4). This parallel suggests that the T^2 distribution might be useful in approximating the small-sample distributions of the LRT and the CSRT.[28] By this logic, $(T-N+1)(N-2)^{-1}(T-2)^{-1}Q_{\mathrm{MLE}}$ (and the corresponding CSRT statistic) should be approximately distributed as F with degrees of freedom $N-2$ and $T-N+1$. Here, $N-2$ replaces N from the riskless asset case, since two additional cross-sectional parameters, γ_{0p} and γ_{1p}, are now estimated.

Shanken (1985) shows, further, that ignoring estimation error in the betas and omitting the errors-in-variables adjustment term (denominator of (3.4)) in computing the CSRT "F-statistic" yields a lower bound on the exact p-value for the test. On the other hand, ignoring estimation error in Γ_{MLE} and treating the gammas as if they were known yields an upper bound on the true p-value. In this case, the "F-statistic" is computed as in Section 2.2.1 with degrees of freedom N and $T-N-1$ [Shanken (1986)]. Zhou (1991) derives the exact distribution of the LRT and finds that it depends on a nuisance parameter that must be estimated. Optimal bounds that do not depend on the unknown parameter are also provided.

Inferences based on small-sample analysis of the multivariate test differ dramatically from those based on the asymptotic chi-square distribution. For ex-

[26] See Kandel (1984) and Roll (1985) for geometric perspectives on the LRT and CSRT, respectively.

[27] This follows from the usual relation between the (time-series) regression estimates and the means of the regression variables.

[28] This observation is made with the benefit of hindsight. In fact, most of the work on the multivariate statistical model with γ_{0p} unknown was done *before* the riskless asset case was analyzed in depth.

ample, whereas Gibbons (1982) obtains an asymptotic p-value less than 0.001 in testing the efficiency of a stock index, Shanken (1985) reports that a small-sample *lower* bound on the true p-value is 0.75. This difference is driven by the fact that error in estimating the residual covariance matrix is not reflected in the limiting chi-square distribution. The estimate of the *inverse* of the residual covariance matrix is quite noisy in small samples and severely biased upward when the number of assets, N, is large relative to the time-series length, T.[29] In Gibbons' case, the test was applied over subperiods with $N = 40$ and $T \doteq 60$. Jobson and Korkie (1982) reach a similar conclusion about Gibbons' test using a Bartlett correction factor [also see Stambaugh (1982)]. Amsler and Schmidt (1985) find that this correction and Shanken's CSRT both perform quite well in simulations under joint normality.

4. Related work

Given a subset of a larger set of assets, it is natural to ask whether some portfolio of the assets in the subset is a minimum-variance portfolio with respect to the larger set. The minimum-variance problem considered in this review is a special case in which the subset consists of a single portfolio. Most of the results discussed here have straightforward generalizations to the multiple-portfolio or "multifactor" case.

A related question is whether a given subset of risky assets actually *spans* the entire minimum-variance frontier of the larger set. This is a stronger restriction than that considered above, which Huberman and Kandel (1987) refer to as "*intersection.*" They show that the spanning condition amounts to a joint restriction that the intercepts equal zero and the betas for each asset sum to one in the multifactor version of (1.2). This is tested using a small-sample F-statistic.

There is also a literature that treats the efficient portfolio as an unobserved "latent variable." A time-series model of conditional expectations is postulated and used to derive testable cross-sectional restrictions on the joint distribution of observed security returns. See Gibbons and Ferson (1985) and Hansen and Hodrick (1983) for early examples of latent variable models. A recent paper by Zhou (1994) provides analytical generalized method of moment tests for latent variable models, permitting applications with many more assets than was previously computationally feasible.

[29] The first and second moments of the distribution of the sample covariance matrix do not depend on N, whereas the moments of the distribution of the inverse involve expressions with T–N in the denominator. See Press (1982), pp. 107–120, for the basic properties of Wishart and inverted Wishart distributions.

References

Amsler, C. and P. Schmidt (1985). A Monte Carlo investigation of the accuracy of CAPM tests. *Economics* **14**, 359–375.

Affleck-Graves, J. and B. McDonald (1989). Nonnormalities and tests of asset pricing theories. *J. Finance* **44**, 889–908.

Affleck-Graves, J. and B. McDonald (1990). Multivariate tests of asset pricing: The comparative power of alternative statistics. *J. Financ. Quant. Anal.* **25**, 163–183.

Banz, R. (1981). The relationship between return and market value of common stocks. *J. Financ. Econom.* **9**, 3–18.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *J. Business* **45**, 444–455.

Brennan, M. (1971). Capital market equilibrium with divergent borrowing and lending rates. *J. Financ. Quant. Anal.* **6**, 1197–1205.

Black, F., M. C. Jensen and M. Scholes (1972). The capital asset pricing model: Some empirical tests. In: M. C. Jensen, ed., Studies in the theory of capital markets, Praeger, New York, NY.

Blume, M. and I. Friend (1973). A new look at the capital asset pricing model. *J. Finance* **28**, 19–34.

Brown, S. and M. Weinstein (1983). A new approach to testing asset pricing models: The bilinear paradigm. *J. Finance* **38**, 711–743.

Campbell, J. (1985). Stock returns and the term structure. NBER Working Paper.

Chamberlain G. (1983). A characterization of the distributions that imply mean-variance utility functions. *J. Econom. Theory* **29**, 185-2–1.

Chen, N. F., R. Roll and S. Ross (1986). Economic forces and the stock market. *J. Business* **59**, 383–403.

Evans, G. and N. Savin (1982). Conflict among the criteria revisited: The W, LR and LM tests. *Econometrica* **50**, 737–748.

Fama, E. F. (1976). *Foundations of Finance*. Basic Books, New York, NY.

Fama E. F. and J. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *J. Politic. Econom.* **81**, 607–636.

Ferson, W., S. Kandel and R. Stambaugh (1987). Tests of asset pricing with time-varying expected risk premiums and market betas. *J. Finance* **42**, 201–220.

Gibbons, M. (1980). Estimating the parameters of the capital asset pricing model: A minimum expected loss approach. Unpublished manuscript, Graduate School of Business, Stanford, University.

Gibbons, M. (1982). Multivariate tests of financial models: A new approach. *J. Financ. Econom.* **10**, 3–27.

Gibbons, M. and W. Ferson (1985). Testing asset pricing models with changing expectations and an unobservable market portfolio. *J. Financ. Econom.* **14**, 217–236.

Gibbons, M., S. Ross and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica* **57**, 1121–1152.

Hansen, L. and R. Hodrick (1983). Risk-averse speculation in the forward foreign exchange market: An econometric analysis of linear models. In: J. J. Frenkel ed., *Exchange rates and international macroeconomics*. Cambridge, MA: National Bureau of Economic Research, 113–146.

Harvey, C. (1989). Time-varying conditional covariances in tests of asset-pricing models. *J. Financ. Econom.* **24**, 289–317.

Harvey, C. and G. Zhou (1990). Bayesian inference in asset pricing tests. *J. Financ. Econom.* **26**, 221–254.

Huberman, G. and S. Kandel (1987). Mean-variance spanning. *J. Finance* **42**, 873–888.

Jensen, M. 1972, Capital markets: Theory and evidence. *Bell J. Econom. Mgmt. Sci.* **3**, 357–398.

Jobson, J. D. and B. Korkie (1982) Potential performance and tests of portfolio efficiency. *J. Financ. Econom.* **10**, 433–466.

Jobson, J. D. and B. Korkie (1985). Some test of linear asset pricing with multivariate normality. *Canad. J. Administ. Sci.* **2**, 114–138.

Kandel, S. (1984). The likelihood ratio test statistic of mean-variance efficiency without a riskless asset. *J. Financ. Econom.* **13**, 575–592.

Kandel, S. and R. F. Stambaugh (1987). On correlations and inferences about mean-variance efficiency. *J. Financ. Econom.* **18**, 61–90.

Kandel, S. R. McCulloch and R. F. Stambaugh (1995). Bayesian inference and portfolio efficiency. *Rev. Financ. Stud.* **8**, 1–53.

Kim, D. (1995). The errors in the variables problem in the cross-section of expected stock returns. *J. Finance* **50**, 1605–1634.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econom. Statist.* **47**, 13–37.

Litzenberger, R. and K. Ramaswamy (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *J. Financ. Econom.* **7**, 163–195.

Litzerberger, R. and C-f Huang (1988). *Foundations for Financial Economics*. Elsevier Science Publishing Company, Inc., North Holland.

Lo, A. W. and A. C. MacKinlay (1990). Data-snooping biases in tests of financial asset pricing models. *Rev. Financ. Stud.* 3, 431–467.

MacBeth, J. (1975). Tests of the two parameter model of capital market equilibrium. Ph.D. Dissertation, University of Chicago, Chicago, IL.

MacKinlay, A. C. (1987). On multivariate tests of the CAPM. *J. Financ. Econom.* **18**, 341–371.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *J. Financ. Econom.* **38**, 3–28.

MacKinlay A. C. and M. Richardson (1991). Using generalized method of moments to test mean-variance efficiency. *J. Finance* **46**, 511–527.

McCulloch, R. and P. E. Rossi (1990). Posterior, predictive, and utility based approaches to testing the arbitrage pricing theory. *J. Financ. Econom.* **28**, 7–38.

Merton, R. (1973a). An analytic derivation of the efficient portfolio frontier. *J. Financ. Quant. Anal.*, 1851–1872.

Merton, R. (1973b). An intertemporal capital asset pricing model. *Econometrica* **41**, 867–887.

Morrison, D. (1976). Multivariate statistical methods. McGraw-Hill, New York.

Pagan, A. (1983). Econometric issues in the analysis of regressions with generated regressors. *Internat. Econom. Rev.* **25**, 221–247.

Press, S. J. (1982). *Applied Multivariate Analysis*. Robert E. Krieger Publishing Company, Malabar, Florida.

Roll, R. (1977). A critique of the asset pricing theory's test - Part 1: On past and potential testability of the theory. *J. Financ. Econom.* **4**, 129–176.

Roll, R. (1979). A reply to Mayers and Rice. *J. Financ. Econom.* **7**, 391–399.

Roll, R. (1985). A note on the geometry of Shanken's CSR T^2 test for mean/variance efficiency. *J. Financ. Econom.* **14**, 349–357.

Ross, S. (1977). The capital asset pricing model, short sales restrictions and related issues. *J. Finance* **32**, 177–183.

Shanken, J. (1983). An asymptotic analysis of the traditional risk-return model. Ph.D. Dissertation, Carnegie Mellon University, Chapter 2.

Shanken, J. (1985). Multivariate tests of the zero-beta CAPM. *J. Financ. Econom.* **14**, 327–348.

Shanken, J. (1986). Testing portfolio efficiency when the zero-beta rate is unknown: A note. *J. Finance* **41**, 269–276.

Shanken, J. (1987a). A Bayesian approach to testing portfolio efficiency. *J. Financ. Econom.* **19**, 195–215.

Shanken, J. (1987b). Multivariate proxies and asset pricing relations: Living with the Roll critique. *J. Financ. Econom.* **18**, 91–110.

Shanken, J. (1990). Intertemporal asset pricing: An empirical investigation. *J. Econometrics* **45**, 99–120.

Shanken, J. and M. Weinstein (1990). Macroeconomic variables and asset pricing: Further results. Working Paper, University of Rochester.

Shanken, J. (1992). On the estimation of beta-pricing models. *Rev. Financ. Stud.* **5**, 1–33.

Stambaugh, R. F. (1982). On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis. *J. Financ. Econom.* **10**, 237–268.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance* **19**, 425–442.

White H. (1984). *Asympototic theory for econometricians*. Academic Press, Orlando, Florida.

Zhou, G. (1991). Small sample tests of portfolio efficiency. *J. Financ. Econom.* **30**, 165–191.

Zhou, G. (1993). Asset pricing tests under alternative distributions. *J. Finance* **48**, 1927–1942.

Zhou, G. (1994). Analytical GMM tests: Asset pricing with time-varying risk premiums. *Rev. Financ. Stud.* **7**, 687–709.