Journal of Financial Economics 18 (1987) 341-371. North-Holland

ON MULTIVARIATE TESTS OF THE CAPM

A. Craig MacKINLAY*

University of Pennsylvania, Philadelphia, PA 19104-6367, USA

Received November 1985, final version received November 1986

This paper evaluates the power of multivariate tests of the Capital Asset Pricing Model. The results indicate that when employing an unspecified alternative hypothesis, the ability of the tests to distinguish between the CAPM and other pricing models is poor. An upper bound is derived for the distance the alternative distribution of the test statistic can be from the null distribution when the deviations from the CAPM are due to missing factors. This upper bound explains the low power of the tests.

1. Introduction

The Capital Asset Pricing Model (CAPM) is an important asset pricing model in financial economics. It has been the subject of considerable research. Recent research has focused on multivariate statistical tests of the CAPM. This paper analyzes whether such multivariate tests can distinguish between the CAPM and other pricing models.

The first multivariate test of the CAPM in the literature is by MacBeth (1975); however, Gibbons (1980, 1982) presented the first extensive treatment. Further work containing both empirical and theoretical results includes Stambaugh (1981, 1982), Jobson and Korkie (1982), Shanken (1983, 1985, 1986) and Amsler and Schmidt (1985). Work providing some theoretical results includes Kandel (1984a, 1984b), Roll (1985), and Gibbons, Ross and Shanken (1986).[1] The thrust of these papers has been the development and study of testable implications of the model. Relatively little attention has been given to power considerations.[2]

*This paper is based on part of my dissertation. I am grateful to my dissertation committee: Eugene Fama, Arnold Zellner, Richard Leftwich, Robert Hamada, Shmuel Kandel, Robert Kohn, and Robert Stambaugh for their assistance. I would like to thank Allan Kleidon, Andrew Lo, Krishna Ramaswamy, Jay Shanken (the referee), and G. William Schwert (the editor) for helpful comments. I remain responsible for any errors.

[1] In addition, Marsh (1985) presents multivariate tests in the context of the term structure of interest rates.

[2] An exception is Gibbons, Ross and Shanken (1986). They present some power results for the test of the efficiency of a given portfolio.

In the literature, there are indications that these tests (with an unspecified alternative hypothesis) may have low power. The fact that the Sharpe–Lintner model[3] can be rejected when tested as a restriction on the Black model,[4] and cannot be rejected when tested as restrictions on the excess return market model, suggests that the multivariate tests with an unspecified alternative may be weak. Further evidence of low power is the apparent insensitivity of the tests to the number of assets considered or the index used as a market proxy.[5] It is unclear whether these tests are capable of detecting economically important deviations from the model.

This paper focuses on the multivariate tests of the Sharpe–Lintner model. These tests are chosen because exact distributional results for the test statistics are available. The Black model is not included, but where applicable, the power results are essentially the same. An analysis of the Black model is included in MacKinlay (1985).

A major constraint on the tests is the stationarity assumption for asset returns.[6] This requirement usually limits the test periods to range between five and seven years. Throughout this paper the test period is taken to be five years (60 monthly observations). To reduce the impact of the constraint, some tests using 240 observations are also considered. The parameters are adjusted to correspond to four observations per month. The objective is to see if substantial power gains can be made by using weekly data rather than monthly data. An increase in power will result from more precise estimation of the covariance matrices allowing sharper tests. The weekly observation interval mitigates non-trading problems and within week seasonality which affects studies using daily data.

Section 2 presents the basic statistical framework. Section 3 describes the data used for the analysis. In section 4, the power of the tests is investigated under two plausible alternative hypotheses. The results indicate the tests have low power against these alternatives.

Section 5 presents a detailed analysis of the power characteristics of the tests. The analysis shows that the type of deviation from the model is an important determinant of the power. If the deviations are cross-sectionally random, the tests can have reasonable power, but if the deviations are due to omitted factors, the tests have low power. For the case of omitted factors, an

[3] The Sharpe–Lintner version is derived by Sharpe (1964) and Lintner (1965).

[4] The Black version is derived by Black (1976).

[5] See Stambaugh (1982).

[6] Three alternative assumptions regarding the distribution of security returns through time are common. They are: (1) nominal returns are independently and identically distributed through time [see Gibbons (1980)]; (2) real returns are IID through time [see Stambaugh (1981)]; (3) excess returns are IID through time [see Jobson and Korkie (1982)]. Using five years of stock data it is empirically difficult to determine the most reasonable assumption because of the high variability of stock returns.

upper bound for the non-centrality parameter of the distribution of the test statistic exists. Some implications of this upper bound are presented. Section 6 reports the empirical evidence, and section 7 contains a summary.

2. The statistical framework for testing the CAPM

Assume asset returns follow a multivariate normal distribution, and that excess asset returns are independently and identically distributed through time. Excess asset returns are defined as the return in excess of the treasury bill rate. With these assumptions asset returns can be described by the excess return market model,

$$\begin{aligned} z_t &= \alpha + \beta z_{mt} + e_t, \qquad t = 1, \dots, T, \\ Ee_t &= 0, \\ Ee_s e_t' &= \Sigma \quad \text{if} \quad s = t, \\ &= 0 \quad \text{if} \quad s \neq t, \end{aligned} \tag{1}$$

where

z_t $= (N \times 1)$ vector of excess asset returns for time period t,
z_{mt} = excess market return for time period t,
e_t $= (N \times 1)$ disturbance vector,
$\alpha, \beta = (N \times 1)$ parameter vectors,
Σ $= (N \times N)$ disturbance covariance matrix.

Throughout the paper, N will refer to the number of left-hand side assets (or portfolios of assets) and T will refer to the number of time observations.[7]

In the presence of a riskless asset the Sharpe–Lintner model, in a one-period world, posits a restricted relation between the excess returns on assets and the excess return on the market portfolio,

$$Ez_t = \beta_t Ez_{mt}. \tag{2}$$

From (1) and (2) we can see the N restrictions imposed on the excess return market model by the Sharpe–Lintner model are $\alpha = 0$.

[7]Strictly speaking, given the assumption that excess asset returns are independently and identically distributed, the proper specification of the excess return market model relies on the asset weights in the market portfolio not changing. Although in reality the weights do change, the fixed weight assumption is likely to be a good working approximation since there are a large number of assets in the market portfolio. See Ferson, Kandel and Stambaugh (1987) and references therein for further discussions of this issue.

The test of these N restrictions against an unspecified alternative is the test for a zero intercept in a multivariate regression model.[8] Specifically, let

$$Z' = [z_1 \quad z_2 \quad \cdots \quad z_T],$$

$$B' = [\alpha \quad \beta],$$

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_{m1} & z_{m2} & \cdots & z_{mT} \end{bmatrix},$$

$$E' = [e_1 \quad e_2 \quad \cdots \quad e_T].$$

Using the above notation we can express the excess return market model in a multivariate regression framework as

$$Z = XB + E. \tag{3}$$

The unbiased estimators for B and Σ are

$$\hat{B} = (X'X)^{-1}X'Z, \tag{4}$$

$$\hat{\Sigma} = (T-2)^{-1}(Z - X\hat{B})'(Z - X\hat{B}). \tag{5}$$

Conditional on X, these estimators are independent and their distributions are[9]

$$\mathrm{vec}(\hat{B}) \sim \mathrm{N}\big(\mathrm{vec}(B), \Sigma \otimes (X'X)^{-1}\big),$$

$$(T-2)\hat{\Sigma} \sim \mathrm{Wishart}(T-2, \Sigma).$$

By recognizing that

$$\alpha = [I_N \otimes C]\mathrm{vec}(B), \tag{6}$$

where $C = [1 \; 0]$, we can isolate the distribution of $\hat{\alpha}$ conditional on the excess market return,

$$\hat{\alpha} \sim \mathrm{N}\left(\alpha, T^{-1}\left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)\Sigma\right),$$

[8] It is assumed that the market portfolio cannot be formed as a linear combination of the left-hand side assets. With this assumption the residual covariance matrix is full rank.

[9] The analysis requires N to be less than $T-1$ in order for the sample estimator of Σ to be full rank.

where $[T^{-1}(1+\hat{\mu}_m^2/\hat{\sigma}_M^2)]$ is the (1,1) element of $(X'X)^{-1}$ with

$$\hat{\mu}_m = \frac{1}{T}\sum_t z_{mt}, \qquad \hat{\sigma}_m^2 = \frac{1}{T}\sum_t (z_{mt} - \hat{\mu}_m)^2.$$

The test statistic for testing $\alpha = 0$ is

$$\theta_1 = \frac{(T-N-1)T}{(T-2)N}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}. \tag{7}$$

From the distributional results for $\hat{\alpha}$ and $\hat{\Sigma}$ it follows that the distribution of θ_1, conditional on the market return, is F with N degrees of freedom in the numerator and $T-N-1$ degrees of freedom in the denominator.[10] The value of the non-centrality parameter of the F-distribution is

$$\lambda = T\left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)^{-1} \alpha'\Sigma^{-1}\alpha. \tag{8}$$

Under the null hypothesis the non-centrality parameter equals zero and the distribution of θ_1 is a central F. Because under the null hypothesis the non-centrality parameter is zero independent of the market return, the central F is also the unconditional distribution of θ_1. The test of the null hypothesis using θ_1 is the uniformly most powerful invariant test and is the likelihood ratio test. [See Muirhead (1982, pp. 212–215).]

Gibbons, Ross and Shanken (1986) present a derivation of the above test using a geometric approach. An important contribution of their paper is the economic interpretation they present. They show that θ_1 can be expressed in terms of the squared Sharpe measures of the tangency portfolio and the market portfolio (which under the alternative hypothesis is not the tangency portfolio), that is

$$\theta_1 = \frac{(T-N-1)}{N}\left[\left(\frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2} - \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)\Big/\left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)\right], \tag{9}$$

where $\hat{\mu}_p$ and $\hat{\sigma}_p^2$ are the sample mean and variance of the tangency portfolio excess return. θ_1 is an increasing function of the difference between the squared Sharpe measures of the tangency portfolio and the market portfolio.

[10] The distribution of θ_1 conditional on the market return follows from the Hotelling T^2 literature and is a direct application of theorem 6.3.1 in Muirhead (1982, p. 211).

With the basic statistical framework in hand we can now proceed with an analysis of the power of the tests.

3. Specification of parameters for analysis

To conduct the analysis, it is necessary to specify the expected excess return of the market, the standard deviation of the market return, the excess return market model residual covariance matrix, and the betas of the portfolios. We use sample estimates from actual monthly returns for this purpose. The 30-year period from January 1954 to December 1983 inclusive is divided into six five-year periods. For each period, we compute the mean and the standard deviation of the excess return on the CRSP equal weighted index. Table 1 reports these values. They are used in the analysis as the expected excess return and the standard deviation of the excess return for the market portfolio.

To obtain some diversity in the parameters we use two portfolio formation methods to assign values to the betas and residual covariance matrices. We form portfolios for each of the six periods. One method uses out-of-period betas as the sorting variable. For this method, the portfolios include all stocks with a complete set of returns on the CRSP monthly return file for the five-year test period and for five years either prior to or after the test period. We compute the beta of each stock using a market model regression for the five years out-of-period. If the stock has returns for both the five years preceding the test period and for the five years succeeding the test period, the average of the prior and post period beta is used for the out-of-period beta. The eligible stocks are assigned to portfolios based on their out-of-period betas, with portfolio 1 assigned the stocks with the highest out-of-period betas and portfolio 20 (or 40) the stocks with the lowest out-of-period betas. An equal number of stocks are assigned to each portfolio except extra stocks (the remainder of the number of eligible stocks divided by the number of portfolios) are assigned sequentially, one per portfolio, beginning with portfolio 1.

The second portfolio formation method uses the market value of the equity at the beginning of the period as the sorting variable. All stocks with complete returns for the five-year test period are assigned to portfolios based on their beginning of period market value. The largest firms are assigned to portfolio 1 and the smallest firms are assigned to portfolio 20 (or 40). An equal number of stocks are assigned to each portfolio except for the extra stocks which are handled in the same manner as they are for the beta sorted portfolios.

The returns for the portfolios are computed by using an equal weighted average of the returns of the included stocks. The number of stocks eligible for inclusion in the portfolios ranged from 910 for the beta sorted portfolios for the January 1959 to December 1963 period to 1275 for the size sorted portfolios for the January 1974 to December 1978 period. The excess portfolio returns are regressed on the excess return of the CRSP equal weighted market

Table 1

Sample estimates of selected parameters of the distribution of excess returns on the CRSP equal weighted index in the time period January 1954 to December 1983.

Time period	Date	Market mean excess return (% per month) μ_m	Standard deviation of market excess return (% per month) σ_m	$\frac{\mu_m^2}{\sigma_m^2}$ [a]	Sample estimates of $(\iota-\beta)'\Sigma^{-1}(\iota-\beta)$ [b]: 20 portfolios, Beta-sorted[c]	20 portfolios, Size-sorted[d]	40 portfolios, Beta sorted	40 portfolios, Size-sorted
1	1/54–12/58	1.6	3.5	0.21	7,319	1,608	17,004	13,262
2	1/59–12/63	0.6	4.1	0.021	3,189	1,193	8,839	4,510
3	1/64–12/68	1.5	4.0	0.14	2,536	1,151	3,915	6,340
4	1/69–12/73	−1.0	6.2	0.026	2,052	1,770	3,096	4,504
5	1/74–12/78	0.6	6.7	0.013	1,305	737	3,222	1,372
6	1/79–12/83	1.3	5.2	0.063	1,702	816	3,247	2,684

[a] The mean of market excess return squared divided by the variance of the market excess return calculated using monthly observations of the CRSP equal weighted index.

[b] β and Σ are computed using the excess return market model with monthly observations. The index used is the CRSP equal weighted index. The value of $(\iota-\beta)'\Sigma^{-1}(\iota-\beta)$ relates to the power of the tests when the risk-free rate is measured with error.

[c] These portfolios are formed using the out-of-period market model beta as the sorting variable.

[d] These portfolios are formed using the beginning-of-period value of equity as the sorting variable.

index to obtain sample estimates of the betas and of the residual covariance matrices.

4. Evaluation of the power characteristics

This section evaluates the power of the multivariate tests of the Sharpe–Lintner model for two cases.[11] In the first case, we introduce the violations by assuming the risk-free return is the treasury bill return plus a constant. By letting the constant deviate from zero, the power of the tests is documented. This setup amounts to having the market portfolio on the efficient frontier of risky assets but not being the tangency portfolio. It approximates a situation where the Black model is valid yet the Sharpe–Lintner model is not. The example is useful for illustrative purposes. Jobson and Korkie (1982) test the Sharpe–Lintner model nested in the excess return market model and do not reject the model. Yet most studies, beginning with Black, Jensen and Scholes (1972) and Fama and MacBeth (1973), reject the Sharpe–Lintner model by testing it as a restriction on the Black model.

The second case involves tests of the Sharpe–Lintner model when the true model is a two-factor pricing model. The market is chosen to be the first factor, and a normally distributed variable that has a positive mean and is independent of the market, is chosen to be the second factor. The coefficients on the second factor are chosen independent of the market betas. The objective is to document the ability of the tests to distinguish the CAPM from alternative pricing models.

We consider one five-year time period in the analysis. With more than one time period of data available, the power of an aggregated test will be higher than the power for a single period. However, the ability to aggregate will not influence relative comparisons across alternatives.

4.1. Case I – Risk-free rate measured with error

Assume that the treasury bill rate is equal to the true risk-free rate minus a constant. Let r_{Ft}^* be the treasury bill rate. Then

$$r_{Ft}^* = r_{Ft} - \gamma, \tag{10}$$

where r_{Ft} is the true risk-free rate and γ is a constant.

When the excess return market model is estimated using the treasury bill rate as a risk-free measure, we have

$$\alpha = \gamma(\iota - \beta), \tag{11}$$

where ι is a $(N \times 1)$ vector of ones. The null hypothesis is true when γ equals

[11] Gibbons, Ross and Shanken (1986) also consider the power of the test. They consider a case where the excess return market model residual covariance matrix has equal off diagonal elements.

zero. As γ deviates from zero the violation of the null hypothesis becomes more severe. Increasing γ shifts the opportunity set of risky assets upward without altering its shape. The market portfolio remains on the efficient frontier of risky assets. However, the market is no longer the tangency portfolio with respect to the treasury bill rate.

The setup under the alternative hypothesis can be approximately related to a situation where the Black model is appropriate but the Sharpe–Lintner model is not. Here, the treasury bill is not an appropriate measure of the riskfree return. The market portfolio need only be on the efficient frontier and not be the tangency portfolio. The relationship to the Black model is only approximate because in the Black model framework the market portfolio need not be on the minimum variance boundary when the opportunity set is expressed in excess returns rather than in real returns. However, when considering common stocks in the tests, the approximation should be adequate.

Recall from section 2 that the test statistic of the Sharpe–Lintner model under the null hypothesis has a central F-distribution. Under the alternative hypothesis the distribution, conditional on the market return, is a non-central F with non-centrality parameter λ. Using the value for α from (11) and the expression for λ from (8), we obtain an expression for the non-centrality parameter:

$$\lambda = T\left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)^{-1} (\iota - \beta)'\Sigma^{-1}(\iota - \beta)\gamma^2. \tag{12}$$

Given values of $(\iota - \beta)'\Sigma^{-1}(\iota - \beta)$ and $\hat{\mu}_m^2/\hat{\sigma}_m^2$, we can investigate the power of the test by varying γ, computing the corresponding value of λ, and using the non-central F-distribution. For a given significance level, we find the critical value of the appropriate central F and then find the proportion of the non-central F (for the given non-centrality parameter) above that value. We tabulate the power for sixty observations over a five-year interval as well as for 240 observations over this interval. These cases are selected to roughly correspond to monthly and weekly observations.

Table 1 reports the values for $(\iota - \beta)'\Sigma^{-1}(\iota - \beta)$ and $\hat{\mu}_m^2/\hat{\sigma}_m^2$ which are calculated from the sample parameters for each of the six time periods. These values are calculated using sixty observations per five-year period. In time periods when the excess return market model residuals have lower variability the value of $(\iota - \beta)'\Sigma^{-1}(\iota - \beta)$ is highest. In these periods, holding the deviation from the null constant, the tests will be more powerful. The subsequent analysis employs the values from the beta sorted portfolios for time periods 1, 3 and 5. Given the deviation considered, portfolios sorted to maximize the dispersion in betas are desirable.

Table 2 reports the power of the test for various values of γ. It is clear that the multivariate test of the Sharpe–Lintner model is not useful if the error is in

Table 2

Power table for Sharpe–Lintner model test using the excess return market model as the alternative hypothesis.[a]

γ_0 (monthly value)[b]	20 portfolios – 60 observations: Non-centrality parameter[c]	20 portfolios – 60 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.05	20 portfolios – 60 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.01	20 portfolios – 240 observations: Non-centrality parameter	20 portfolios – 240 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.05	20 portfolios – 240 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.01	40 portfolios – 60 observations: Non-centrality parameter	40 portfolios – 60 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.5	40 portfolios – 60 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.01	40 portfolios – 240 observations: Non-centrality parameter	40 portfolios – 240 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.05	40 portfolios – 240 observations: $p(\text{reject } \alpha=0 \mid \alpha=\alpha_0)$ 0.01
Panel A: Based on time period 1 (1/54–12/58) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01
0.002	1.45	0.07	0.02	1.67	0.08	0.02	3.37	0.07	0.02	3.88	0.11	0.03
0.004	5.81	0.16	0.05	6.68	0.24	0.09	13.5	0.16	0.04	15.5	0.40	0.18
0.006	13.1	0.37	0.15	15.0	0.57	0.32	30.4	0.38	0.14	34.9	0.86	0.66
0.008	23.2	0.67	0.38	26.7	0.88	0.71	54.0	0.68	0.36	62.0	0.99	0.97
0.010	36.3	0.89	0.69	41.7	0.98	0.95	84.3	0.90	0.66	96.9	1.00	1.00
Panel B: Based on time period 3 (1/64–12/68) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01
0.002	0.53	0.06	0.01	0.59	0.06	0.01	0.82	0.06	0.01	0.91	0.06	0.01
0.004	2.14	0.08	0.02	2.35	0.10	0.03	3.30	0.07	0.02	3.63	0.10	0.03
0.006	4.81	0.14	0.04	5.29	0.19	0.06	7.42	0.10	0.03	8.17	0.20	0.06
0.008	8.54	0.24	0.08	9.41	0.35	0.15	13.2	0.16	0.04	14.5	0.37	0.16
0.010	13.4	0.38	0.16	14.7	0.56	0.31	20.6	0.25	0.08	22.7	0.61	0.35
Panel C: Based on time period 5 (1/74–12/78) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01
0.002	0.31	0.05	0.01	0.31	0.06	0.01	0.76	0.05	0.01	0.77	0.06	0.01
0.004	1.24	0.07	0.02	1.25	0.07	0.02	3.05	0.07	0.02	3.08	0.09	0.02
0.006	2.78	0.10	0.02	2.81	0.11	0.03	6.87	0.10	0.02	6.94	0.17	0.05
0.008	4.95	0.14	0.04	5.00	0.18	0.06	12.2	0.15	0.04	12.3	0.31	0.12
0.010	7.73	0.21	0.07	7.81	0.28	0.11	19.1	0.23	0.07	19.3	0.51	0.27

[a] Deviations from the Sharpe–Lintner model are introduced by assuming the risk-free rate is measured with error. γ_0 is the (constant) measurement error. The value of the excess return market model intercept vector (α_0) is specified using $\alpha_0 = \gamma_0(1-\beta)$. Sample estimates are used to assign values to the excess return market model parameter β. When γ_0 equals zero, α_0 equals zero and the Sharpe–Lintner model is true. The null hypothesis is $\alpha = 0$ and the alternative hypothesis is $\alpha \neq 0$.

[b] The monthly value of γ_0 specified corresponds to the cases of 60 observations. For the cases of 240 observations the value of γ_0 used is the monthly value divided by 4.

[c] The non-centrality parameter (λ) is calculated by substituting the sample estimates of the excess return market model residual covariance matrix (Σ), the mean excess return of the market (μ_m), and the variance of the excess return of the market (σ_m^2) into $\lambda = T\alpha_0'\Sigma^{-1}\alpha_0/(1+\mu_m^2/\sigma_m^2)$.

the measurement of the risk-free return. Suppose the measurement error is 0.4 percent per month. From table 2 the power of the tests, at the five percent significance level, ranges from 0.07 (for twenty portfolios and sixty observations using time period 5 parameters) to 0.40 (for forty portfolios and 240 observations using time period 1 parameters). Most of the values are less than 0.10. Given the importance of 0.4 percent per month (or about 5 percent annually), the power seems unsatisfactory. Using the parameter values from time period 5 (panel C), the power is very low even for a value of 1.0 percent per month for γ. The power ranges from 0.21 to 0.51 at the five percent significance level. Recognizing that the expected excess return of the market is generally assumed to be about 6 to 8 percent annually (or about 0.50 to 0.75 percent per month), from an economic perspective, it appears the tests are unlikely to detect large deviations.

Despite the low power, we can obtain some insight of the gains from more frequent observations. There appear to be substantial gains in power in both the case of twenty portfolios and the case of forty portfolios. For example, in time period 3, for forty portfolios, and γ equal to 0.6 percent per month, the power increases from 0.10 for sixty observations to 0.20 for 240 observations at the five percent significance level. The source of the gain is the more precise estimation of the residual covariance matrix.

We can illustrate the importance of a specific alternative hypothesis using tests of the Sharpe–Lintner model. The alternative model is the observed risk-free rate is the true risk-free rate plus a constant. For the alternative model we have

$$z_t = (\iota - \beta)\gamma + \beta z_{mt} + v_t, \tag{13}$$

where v_t is the disturbance vector. The model is in the form of the model Gibbons (1982) and Shanken (1985) consider. For this model, a one-step estimator of γ is asymptotically efficient. [See Shanken (1983).] The estimator is

$$\hat{\gamma} = \left[(\iota - \hat{\beta})'\hat{\Sigma}^{-1}(\iota - \hat{\beta})\right]^{-1}(\iota - \hat{\beta})'\hat{\Sigma}^{-1}\hat{\alpha}. \tag{14}$$

The restriction the Sharpe–Lintner model imposes on the alternative model is γ equals zero. Instead of testing N restrictions, only one restriction is being tested. The test closely relates to the tests of the Sharpe–Lintner model that Black, Jensen and Scholes (1972) and Fama and MacBeth (1973) present. To compute the power, the standard deviation of the estimator of γ is necessary. One can use the asymptotic standard deviation. But given evidence[12] that in finite samples the asymptotic standard deviation understates the true standard

[12] See Gibbons (1980) or MacKinlay (1985).

deviation, using the asymptotic standard deviation for computing the power will overstate the power.

To alleviate this overstatement, we use standard deviations computed from simulations. 500 samples of excess returns are simulated with γ equal to zero for each set of parameters. For each sample $\hat{\gamma}$ is calculated using (14). From these 500 simulated estimates the standard deviation is calculated.[13] Table 3 reports the power results using this standard deviation.

Comparing table 2 and table 3, one can see a substantial increase in the power of the tests when using a specific alternative.[14] For example, for time period 3 with twenty portfolios and sixty observations, when γ equals 0.006, the power of the test using a specific alternative is 0.49 at the five percent significance level and the power using an unspecified alternative is 0.14 at the same significance level. Large power gains are present for all cases considered.

4.2. Case II – A two-factor model

In the second case, we introduce violations of the null hypothesis by assuming that excess returns are generated by a two factor model. The model is

$$
\begin{aligned}
& z_t = \beta z_{mt} + \delta z_{ht} + u_t, \\
& E u_t = 0, \\
& E u_s u_t' = \sigma^2 I \quad \text{if} \quad s = t, \\
& \qquad\quad = 0 \quad\ \ \text{if} \quad s \neq t,
\end{aligned}
\tag{15}
$$

where

$z_{ht} \sim \mathrm{N}(\mu_h, \sigma_h^2)$, IID through time, and independent of z_{mt} and u_t,
$\delta \ = (N \times 1)$ parameter vector.

The model is designed with three primary objectives: (1) the model should be consistent with the excess return market model and its parameters; (2) the model must not be consistent with the one-period CAPM; and (3) the model should be consistent with one of the competing asset pricing models. The alternative model's parameters are specified to attain these objectives. The model is consistent with the excess return market model and a two-factor

[13] These results are in MacKinlay (1985).

[14] Naturally, the power gains documented depend on the researchers ability to specify a reasonable specific alternative hypothesis. Given the difficulty of such a task, it is unlikely that the gains in practice would be as large.

Table 3

Power table for the Sharpe–Lintner model test using a specific alternative hypothesis.[a]

γ_0 (monthly value)[b]	20 portfolios – 60 observations			20 portfolios – 240 observations			40 portfolios – 60 observations			40 portfolios – 240 observations		
	$\gamma_0/\sigma_{\gamma_0}$[c]	$p(\text{reject } \gamma=0 \mid \gamma=\gamma_0)$ 0.05	0.01	$\gamma_0/\sigma_{\gamma_0}$	$p(\text{reject } \gamma=0 \mid \gamma=\gamma_0)$ 0.05	0.01	$\gamma_0/\sigma_{\gamma_0}$	$p(\text{reject } \gamma=0 \mid \gamma=\gamma_0)$ 0.05	0.01	$\gamma_0/\sigma_{\gamma_0}$	$p(\text{reject } \gamma=0 \mid \gamma=\gamma_0)$ 0.05	0.01
Panel A: Based on time period 1 (1/54–12/58) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01
0.002	1.00	0.17	0.06	1.19	0.22	0.08	1.05	0.18	0.06	1.72	0.41	0.20
0.004	2.00	0.52	0.28	2.38	0.66	0.42	2.11	0.56	0.32	3.45	0.93	0.81
0.006	3.00	0.85	0.66	3.57	0.95	0.84	3.16	0.89	0.72	5.17	1.00	1.00
0.008	4.00	0.98	0.92	4.76	1.00	0.99	4.21	0.99	0.95	6.89	1.00	1.00
0.010	5.00	1.00	0.99	5.95	1.00	1.00	5.27	1.00	1.00	8.62	1.00	1.00
Panel B: Based on time period 3 (1/64–12/68) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01
0.002	0.65	0.10	0.03	0.69	0.11	0.03	0.49	0.08	0.02	0.85	0.14	0.04
0.004	1.29	0.25	0.10	1.39	0.29	0.12	0.98	0.17	0.06	1.69	0.39	0.19
0.006	1.94	0.49	0.26	2.08	0.55	0.31	1.46	0.31	0.13	2.54	0.72	0.48
0.008	2.58	0.73	0.50	2.78	0.79	0.58	1.95	0.50	0.26	3.38	0.92	0.79
0.010	3.23	0.90	0.74	3.47	0.93	0.81	2.44	0.68	0.44	4.23	0.99	0.95
Panel C: Based on time period 5 (1/74–12/78) sample estimates												
0.000	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.05	0.01	0.00	0.01	0.05
0.002	0.47	0.08	0.02	0.53	0.08	0.02	0.49	0.08	0.02	0.78	0.12	0.04
0.004	0.93	0.15	0.05	1.05	0.18	0.06	0.98	0.17	0.06	1.56	0.35	0.15
0.006	1.40	0.29	0.12	1.58	0.35	0.16	1.46	0.31	0.13	2.34	0.65	0.41
0.008	1.86	0.46	0.24	2.11	0.56	0.32	1.95	0.50	0.26	3.12	0.88	0.71
0.010	2.33	0.64	0.40	2.63	0.75	0.52	2.44	0.68	0.44	3.91	0.97	0.91

[a] Deviations from the Sharpe–Lintner model are introduced by assuming the risk-free rate is measured with error. γ_0 is the (constant) measurement error. When γ_0 equals zero the Sharpe–Lintner model is true. The null hypothesis is $\gamma = 0$ and the alternative hypothesis is $\gamma \neq 0$.

[b] The monthly value of γ_0 specified corresponds to the cases of 60 observations. For the cases of 240 observations the value of γ_0 used is the monthly value divided by 4.

[c] σ_0 is the standard deviation of the estimator of γ. γ_0/σ_0 is calculated using values of σ_0 derived from simulations. The power results in this table assume that the estimator of γ divided by its standard deviation has a standard normal distribution.

arbitrage pricing model,[15] and inconsistent with the CAPM. This setup facilitates an investigation of the ability of the test to distinguish the CAPM from a plausible alternative model.

For the analysis the mean and variance of the excess return on the market are set to 0.01 and 0.0016, respectively, for sixty monthly observations, and adjusted appropriately for 240 observations. The mean of the second factor is set to 1.0. For the variance of the second factor we consider five values - 2.0, 4.0, 6.0, 9.0, and 16.0. The residual variance of the two-factor model (σ^2) is set to 0.0001 for twenty portfolios and 0.0002 for forty portfolios. These values for the variance of the second factor and the variance of the two-factor model residuals will make the covariance matrix of the excess return market model residuals from this two-factor model roughly consistent with possible sample estimates.

Several values for the variance of the second factor are considered in order to vary the factor's importance. Generally, as we shall see later, the importance of the factor can be quantified by its mean squared divided by its variance. The higher the value of this quantity, the greater the importance of the factor.[16] In the cases considered the factor's mean squared divided by its variance ranges from 0.0625 to 0.5. These values are all higher than one would typically propose as being the population value of this quantity for the excess market return.[17]

For the model to be well specified, it is necessary for the weighted sum of the delta coefficients to be zero.[18] To satisfy this requirement the delta coefficients are assigned equally spaced values from $-r/2$ to $r/2$ where r is a prespecified range. For example, with 40 portfolios and a range of 0.01, the deltas extend from -0.005 to 0.005 with incremental changes of 0.0002564. After specifying the values of the delta coefficients, the coefficients are randomly assigned to one of the portfolios.

With knowledge of the range of the second factor coefficients and the mean of the second factor, statements concerning the difference in the expected returns of two assets in this two-factor world are possible.[19] With a range of 0.005, the expected returns of two portfolios with the same market beta can differ by one half percent per month. With a range of 0.01, the possible difference is one percent per month. The implied difference for two individual

[15] The arbitrage pricing model is due to Ross (1976).

[16] Of course, the coefficients associated with the factors are also relevant. See section 5.

[17] Using 30 years of monthly data from January 1954 to December 1983, the sample values for the mean excess return of the market squared divided by the variance of the excess return are 0.026 for the equal weighted CRSP index and 0.016 for the value weighted CRSP index.

[18] This condition is a result of the fact that the market portfolio return is a weighted sum of individual asset returns. When not all assets are included the condition need not hold exactly. However, it is likely that the condition holds approximately.

[19] These implied differences assume the deltas of the assets are independent of the betas.

securities with the same market beta is even larger, with its magnitude depending on the cross-sectional distribution of the deltas and the ability to form portfolios based on the true deltas. Clearly, these differences are economically important, and it is of interest to see if the tests can detect the presence of such a second factor.

For the results presented, two values of r (on a monthly basis) are considered. For forty portfolios, the values of r considered are 0.01 and 0.005. For twenty portfolios the values of r considered are 0.00974 and 0.00487. The ranges for twenty portfolios are chosen to be equal to the ranges that will result from ordering the forty portfolios by their delta value and then forming a new portfolio from every two portfolios. This procedure assumes one can sort the forty portfolios based on their true delta coefficient and consequently conclusions from power comparisons of the twenty-portfolio case versus the forty-portfolio case are of limited usefulness.

With the given specification the non-centrality parameter of the distribution of θ_1 can be calculated. Using

$$\alpha = \delta\mu_h, \tag{16}$$

$$\Sigma = \delta\delta'\sigma_h^2 + \sigma^2 I, \tag{17}$$

we have

$$\lambda = T\mu_h^2\delta'\left[\delta\delta'\sigma_h^2 + \sigma^2 I\right]^{-1}\delta\left[1 + \frac{\mu_m^2}{\sigma_m^2}\right]^{-1}, \tag{18}$$

conditional on the sample mean and variance of the market being equal to their population values. The value of λ can be calculated for each portfolio–observation–second-factor variance–coefficient range combination. Given λ, we compute the power analytically using the non-central F-distribution. Table 4 reports the power of the test with this two-factor alternative.

Table 4 is divided into five panels based on the variance of the second factor. Panel A contains the results when the second factor variance is 2.0. The power of the test ranges from 0.16 to 0.80 at the five percent level of significance. The power is 0.16 with forty portfolios, sixty observations and a second factor coefficient range of 0.005. With twenty portfolios, 240 observations and a second-factor coefficient range of 0.00974 the power is 0.80. Although this may seem to be a reasonable level of power, recall we are considering a factor substantially more important than the market and coefficient values that could lead to expected returns on two securities with the same market betas differing by over twelve percent on an annual basis. As we proceed from panel A through the table, the power situation degenerates. In

Table 4

Power of tests of Sharpe–Lintner model with the excess return market model as the alternative hypothesis. The true model is a two-factor model with the excess market return as the first factor and a second factor that is orthogonal to the first factor. The power is derived analytically using the F-distribution.[a]

Number of portfolios	Number of observations (T)	Range of second-factor coefficients[b]	Residual variance (σ^2)[c]	Non-centrality parameter (λ)	Power of test 0.05	Power of test 0.01
Panel A: Second-factor monthly variance 2.0						
20	60	0.487	1.0	13.18	0.38	0.15
20	240	0.487	0.25	13.70	0.52	0.28
20	60	0.974	1.0	22.06	0.64	0.36
20	240	0.974	0.25	22.82	0.80	0.59
40	60	0.5	2.0	13.19	0.16	0.04
40	240	0.5	0.5	13.68	0.35	0.15
40	60	1.0	2.0	21.97	0.27	0.09
40	240	1.0	0.5	22.79	0.61	0.35
Panel B: Second-factor monthly variance 4.0						
20	60	0.487	1.0	8.97	0.25	0.09
20	240	0.487	0.25	9.38	0.35	0.15
20	60	0.974	1.0	12.35	0.35	0.14
20	240	0.974	0.25	12.92	0.49	0.25
40	60	0.5	2.0	8.99	0.12	0.03
40	240	0.5	0.5	9.37	0.23	0.08
40	60	1.0	2.0	12.36	0.15	0.04
40	240	1.0	0.5	12.91	0.33	0.13
Panel C: Second-factor monthly variance 6.0						
20	60	0.487	1.0	6.80	0.19	0.06
20	240	0.487	0.25	7.12	0.26	0.10
20	60	0.974	1.0	8.59	0.24	0.08
20	240	0.974	0.25	8.99	0.33	0.14
40	60	0.5	2.0	6.82	0.10	0.02
40	240	0.5	0.5	7.11	0.17	0.05
40	60	1.0	2.0	8.59	0.12	0.03
40	240	1.0	0.5	8.98	0.22	0.07
Panel D: Second-factor monthly variance 9.0						
20	60	0.487	1.0	5.00	0.14	0.04
20	240	0.487	0.25	5.23	0.19	0.06
20	60	0.974	1.0	5.90	0.17	0.04
20	240	0.974	0.25	6.17	0.22	0.08
40	60	0.5	2.0	5.01	0.08	0.02
40	240	0.5	0.5	5.23	0.13	0.04
40	60	1.0	2.0	5.90	0.09	0.02
40	240	1.0	0.5	6.17	0.15	0.04
Panel E: Second-factor monthly variance 16.0						
20	60	0.487	1.0	3.09	0.10	0.03
20	240	0.487	0.25	3.23	0.13	0.02
20	60	0.974	1.0	3.41	0.11	0.03
20	240	0.974	0.25	3.56	0.14	0.04
40	60	0.5	2.0	3.09	0.07	0.02
40	240	0.5	0.5	3.23	0.10	0.02
40	60	1.0	2.0	3.41	0.07	0.02
40	240	1.0	0.5	3.56	0.10	0.02

[a] Under the alternative hypothesis the non-centrality parameter of the F-distribution is $\lambda = T\mu_h^2\delta[\delta\delta'\sigma_h^2 + \sigma^2 I]^{-1}\delta[1 + \mu_m^2/\sigma_m^2]^{-1}$, where T = number of observations; μ_h = mean of the second factor; σ_h^2 = variance of the second factor; δ = vector of second-factor coefficients; σ^2 = residual variance of the two-factor model; μ_m = mean excess return of the market; σ_m^2 = variance of excess return of the market.

For 60 observations, the values of the parameters (not specified in the table) are $\mu_h = 1.0$; $\mu_m = 0.01$; and $\sigma_m^2 = 0.0016$. For 240 observations, the values are $\mu_h = 0.25$; $\mu_m = 0.0025$; and $\sigma_m^2 = 0.0004$.

[b] The second-factor coefficients are centered about zero. The range is scaled by 10^2.

[c] The residual variance of the two-factor model is scaled by 10^4.

panel E, which considers the case where the second-factor variance is 16.0, the power is very low with a maximum of 0.14 at the five percent significance level for all the cases. In this situation the second-factor mean squared divided by its variance is 0.0625, a value similar to sample estimates for the excess return of the market. This implies that if the true model is a two-factor model, with the second factor and the market of about equal importance, the tests are very unlikely to distinguish between the single-factor CAPM and the two-factor model.

From table 4 we can draw some conclusions concerning the test design. Increasing the frequency of observation from sixty observations per period to 240 observations per period results in considerable power increases for the cases with a low second-factor variance. The increased power results from more precise estimation of the excess return market model residual covariance matrix.[20] As the variance of the second factor increases, the gains diminish because the deviation from the CAPM is difficult to detect independent of the precision of the residual covariance matrix estimator. When the alternative model is a multi-factor model, the tests, using an unspecified alternative hypothesis, sixty observations, and forty portfolios, are virtually useless. These tests have low power even when the second factor is important. The tests using twenty portfolios are consistently more powerful than the tests using forty portfolios. However, this result is not general but depends on the ability to group the assets into portfolios in a manner that does not wash out the deviation from the CAPM.

5. Analysis of the power characteristics

The results of section 4 indicate that the multivariate tests lack the power to detect plausible deviations from the CAPM. Yet, in contrast to these results, Gibbons (1980, 1982) and Stambaugh (1981) present simulation results indicating the tests have reasonable power. We solve this discrepancy by examining the link between economically plausible deviations and the non-centrality parameter of the test statistic distribution.

In the Sharpe–Lintner model framework, deviations from the model exist when any of the elements of the vector α have a non-zero value (see section 2). To link the deviations to a non-centrality parameter one needs to specify this vector, and then given appropriate values for T, Σ and $[1+\hat{\mu}_m^2/\hat{\sigma}_m^2]$, compute the non-centrality parameter using eq. (8). For the initial analysis in this

[20] MacKinlay (1985) takes this analysis one step further by considering the power for the extreme case where the residual covariance matrix is known. The results indicate that for twenty and forty portfolios much of the possible power gains are realized by going from 60 to 240 observations.

section $[1 + \hat{\mu}_m^2/\hat{\sigma}_m^2]$ will be approximated by one, and then we have

$$\lambda = T\alpha'\Sigma^{-1}\alpha. \tag{19}$$

First, consider the specification where the elements of α do not obey any particular relation across assets but are zero on average. This specification is similar to that considered by Gibbons (1980, 1982) and by Stambaugh (1981) in the evaluation of the tests of the Black model. The elements of α are chosen in the same manner as the second-factor coefficient vector elements are chosen in section 4. This method randomly locates equally spaced values of α coefficients in the α vector. The values of the α coefficients are specified by dividing the given range centered about zero into N equally spaced points. For example, with forty portfolios and a range of 0.01, the α coefficients take on the values 0.00500, 0.00474, 0.00449, ..., -0.00474, and -0.00500. Using the excess return market model residual covariance matrices previously employed (see section 3), the value of the non-centrality parameter can be calculated using

$$\lambda = T\bar{\bar{\alpha}}'\Sigma^{-1}\bar{\bar{\alpha}}, \tag{20}$$

where $\bar{\bar{\alpha}}$ is a $(N \times 1)$ randomly assigned parameter vector. For the twelve sample estimates of the residual covariance matrix, 200 values of λ are randomly generated for the α coefficients having a range of 0.00974 for twenty portfolios and a range of 0.01 for forty portfolios. To obtain non-centrality parameters for other ranges these values are appropriately scaled. The other ranges considered are 0.00195 and 0.00487 for twenty portfolios, and 0.002 and 0.005 for forty portfolios. For each λ the power of the test is calculated assuming that λ is the non-centrality parameter of the alternative distribution. The average power for each covariance matrix and range combination is then calculated using the mean of the power across the 200 values. Table 5 reports the results for the case of sixty observations. Consistent with the results of other studies, the tests, using this alternative hypothesis, have considerable power. At the five percent significance level, with the range of the alpha coefficients set to 0.00487, the null hypothesis will be rejected about 90 percent of the time for twenty portfolios. With a range of 0.005 for the alpha coefficients, the null hypothesis will be rejected about the same amount of the time for forty portfolios. This 90 percent rejection rate is substantially higher than the rejection rates of 19 and 10 percent, for twenty and forty portfolios respectively, we find using the same specification for alpha in a two-factor model framework (in section 4 and table 4, panel C).

The dramatic difference in the two situations illustrates the importance of the covariance structure of the residuals in the power analysis. When the deviations from the model are randomly introduced without regard to the

Table 5

Power summary for Sharpe–Lintner tests with the excess return market model intercept vector (α) randomly assigned values.[a]

Number of portfolios	Sorting variables[b]	Test period[c]	Panel A $\bar{\hat{\lambda}}$[d]	Panel A $\sigma(\hat{\lambda})$[e]	Panel A Power[f]	Panel B $\bar{\hat{\lambda}}$	Panel B $\sigma(\hat{\lambda})$	Panel B Power	Panel C $\bar{\hat{\lambda}}$	Panel C $\sigma(\hat{\lambda})$	Panel C Power
20	Beta	1	8.8	2.1	0.25	55	13	0.97	219	52	1.00
		3	6.1	1.5	0.17	38	9.5	0.88	152	38	1.00
		5	6.4	2.0	0.18	40	13	0.88	160	51	1.00
	Size	1	9.3	2.1	0.26	58	14	0.98	232	54	1.00
		3	6.9	1.9	0.19	43	12	0.91	172	48	1.00
		5	7.4	2.0	0.21	46	13	0.93	185	51	1.00
40	Beta	1	18	5.7	0.22	115	36	0.94	459	142	1.00
		3	14	4.6	0.17	87	28	0.87	349	114	1.00
		5	21	10	0.26	132	64	0.94	527	255	1.00
	Size	1	22	7.2	0.27	136	45	0.96	542	179	1.00
		3	16	5.6	0.20	101	35	0.91	403	139	1.00
		5	17	5.8	0.21	106	36	0.92	425	145	1.00

[a] Under the null hypothesis $\alpha = 0$. For each panel the range of the elements of α is changed. The range associated with each panel is:

Panel	Range of alphas: 20 portfolios	Range of alphas: 40 portfolios
A	0.00195	0.002
B	0.00487	0.005
C	0.00974	0.01

[b] Beta and size identify the portfolio formation method used to compute the sample residual covariance matrix. The beta sorted portfolios are formed using the out-of-period market model beta as the sorting variable. The size sorted portfolios are formed using the beginning-of-period market value of equity as the sorting variable.

[c] Test period identifies the five-year time period used to compute the sample residual covariance matrix. Test period 1 is 1/54 to 12/58. Test period 3 is 1/64 to 1/68. Test period 5 is 1/74 to 1/78.

[d] $\hat{\lambda}$ is the non-centrality parameter estimated using $\hat{\lambda} = T\bar{\bar{\alpha}}'\Sigma^{-1}\bar{\bar{\alpha}}$. $\bar{\bar{\alpha}}$ is randomly assigned equally spaced fixed values without replacement about zero in the specified range. The mean of the sample of non-centrality parameters (defined as $\bar{\hat{\lambda}}$) is based on 200 replications for each residual covariance matrix (Σ) considered.

[e] $\sigma(\hat{\lambda})$ is the standard deviation of the sample of $\hat{\lambda}$'s.

[f] The power presented is the average of the power associated with each $\hat{\lambda}$, at the five percent level of significance. For twenty portfolios the appropriate distribution for the test statistic is $F_{20,39}(\lambda)$. For forty portfolios the appropriate distribution is $F_{40,19}(\lambda)$. The same random sample is used for each panel.

covariance structure of the residuals, the tests have reasonable power. However, when the same sort of deviations are introduced using a factor model, the tests are very weak. When the alternative hypothesis is a two-factor model, the deviations are reflected in the residual covariance matrix, as well as the alpha vector. When the magnitude of the deviation is larger, the residual variance is also larger, making the deviation more difficult to detect. The covariances are also important. With deviations introduced by a factor model, the residuals of assets with deviations with the same sign will be positively correlated and residuals with deviations with different signs will be negatively correlated (neglecting other influences on the covariance structure). This phenomena results in weaker evidence against the null hypothesis than if, for example, the residuals are uncorrelated.

Statements concerning the power of the test against alternatives as the arbitrage pricing model [Ross (1976)] or the intertemporal CAPM [Merton (1973)] are not possible without consideration of the residual covariance matrix structure. We can establish an upper bound on the value of the non-centrality parameter if the true model is a factor model. Consider the two-factor model introduced in section 4,

$$\begin{aligned} z_t &= \beta z_{mt} + \delta z_{ht} + u_t, \\ Eu_t &= 0, \\ Eu_s u_t' &= \Phi \quad \text{if} \quad s = t, \\ &= 0 \quad \text{if} \quad s \neq t, \\ z_{ht} &\sim \mathrm{N}(\mu_h, \sigma_h^2) \quad \text{independent of } z_{mt} \text{ and } u_t. \end{aligned} \tag{21}$$

One factor is the market portfolio and the other factor is a normally distributed variable orthogonal to the market. Cross-sectional independence of the errors is not imposed. From this model, the parameters of the excess return market model are

$$\alpha = \delta\mu_h, \tag{22}$$

$$\Sigma = \delta\delta'\sigma_h^2 + \Phi. \tag{23}$$

For the F-test of the Sharpe–Lintner model, the non-centrality parameter of the distribution of the test statistic is

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \alpha'\Sigma^{-1}\alpha. \tag{24}$$

α and Σ from eqs. (22) and (23) can be substituted into eq. (24) giving

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \mu_h^2 \delta' \left[\delta\delta'\sigma_h^2 + \Phi\right]^{-1} \delta. \tag{25}$$

Analytically inverting the residual covariance matrix gives

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \mu_h^2 \delta' \left[\Phi^{-1} - \frac{\sigma_h^2}{1 + \sigma_h^2 \delta' \Phi^{-1} \delta} \Phi^{-1} \delta\delta' \Phi^{-1}\right] \delta.$$

Simplifying we have

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \frac{\mu_h^2}{\sigma_h^2} \left[\frac{\sigma_h^2 \delta' \Phi^{-1} \delta}{1 + \sigma_h^2 \delta' \Phi^{-1} \delta}\right]. \tag{26}$$

To establish the upperbound of λ, we use the fact that

$$0 < \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \leq 1, \tag{27}$$

since $\hat{\mu}_m^2 / \hat{\sigma}_m^2$ is non-negative, and the fact that

$$0 < \frac{\sigma_h^2 \delta' \Phi^{-1} \delta}{1 + \sigma_h^2 \delta' \Phi^{-1} \delta} \leq 1, \tag{28}$$

since $\sigma_h^2 \delta' \Phi^{-1} \delta$ is non-negative. Using (27) and (28) the upperbound for λ is established. From (26) we have

$$\lambda \leq T \frac{\mu_h^2}{\sigma_h^2}. \tag{29}$$

The mean of the factor squared divided by the variance of the factor, and the length of time period determine the upper bound. In the appendix, this upperbound for the non-centrality parameter is generalized for an alternative model with the market and multiple factors.

This upper bound has implications concerning the ability of the tests to distinguish between the one-period CAPM and other alternative factor pricing models. Suppose the second factor has a mean and variance equal to the mean and variance of the excess market return. What are the chances of detecting this second factor with an unspecified alternative hypothesis? Using 0.006 as the monthly expected excess return on the market and 0.04 as the monthly standard deviation of the excess return, and using five years (sixty months) as the length of time period, one can calculate the maximum power of the tests for a given number of portfolios. From eq. (29), the upper bound of the non-centrality parameter is 1.35. This implies that for any number of portfolios the tests have little power to reject the null hypothesis.

This low upper bound using plausible excess market return parameters also implies that inferences may not be overly sensitive to the exact identification of the market portfolio.[21] Stambaugh (1981, 1982) presents results consistent with this implication. Also, if the market is not identified as a factor, the same analysis leads to an upper bound on the power of a test for the equality of all expected returns. Thus, it is not surprising that in many time periods Shanken (1985) is unable to reject the hypothesis that the expected returns on all assets are equal.

It is well known that the problem of testing the CAPM is directly related to the problem of testing the mean–variance efficiency of a given portfolio. Kandel and Stambaugh (1987) and Shanken (1987) consider the problem of testing the mean–variance efficiency of the market portfolio. They address the question of the sensitivity of inferences to the portfolio selected as the proxy for the market portfolio. They explore the question of how small the correlation between the market proxy portfolio return and the true tangency portfolio return must be to reverse inferences about mean–variance efficiency. In a mean–variance framework, the upper bound from (29) allows an ex ante statement about the sensitivity of inferences to the proxy chosen for the market portfolio. To do this using the two-factor model in this paper, it is necessary to interpret the second factor as the excess return on a portfolio and to interpret the market portfolio as a proxy for the market portfolio. Then, using the assumption that the second-factor portfolio is orthogonal to the proxy for the market portfolio and the condition that the second-factor portfolio and the proxy for the market portfolio can be combined to form the tangency portfolio,[22] we can express the expected excess return of the second-factor portfolio squared divided by the variance of the second factor in terms of the means and variances of the market proxy and tangency portfolio excess returns and the correlation between the returns of these portfolios. This allows the upper bound on the non-centrality parameter of the distribution of the test statistic to be expressed in terms of the proxy portfolio and tangency portfolio parameters and hence, a statement about the sensitivity of inferences to these parameters.

[21] Roll (1977) emphasizes this potential problem.

[22] This condition follows from the fact that the intercept in the two-factor model is equal to zero.

6. The empirical evidence

We present tests of the Sharpe–Lintner model for completeness. All tests are conducted using excess returns where one-month treasury bill returns are used as the risk-free asset return.

Table 6 reports the results of the Sharpe–Lintner model tests using monthly data and the excess return market model as the alternative hypothesis. The number of restrictions tested is equal to the number of portfolios (either twenty or forty). The CRSP equal weighted index is used as a proxy for the market portfolio return. Six five-year time periods are considered, beginning with January 1954 and ending with December 1983. Although the model can be rejected at the five percent significance level in some subperiods, it cannot be rejected for the overall thirty-year period for either the beta sorted portfolios or the size sorted portfolios at the five percent level. The lowest overall p-value is 0.082 for the twenty size-sorted portfolios.[23]

These results are consistent with previous results that have employed the market model as the alternative hypothesis. Using an unspecified alternative hypothesis, violations of the CAPM are difficult to detect.

Table 7 reports tests of the Sharpe–Lintner model using the alternative that the observed risk-free rate is the true risk-free rate minus a constant. These tests can also be interpreted as tests of the Sharpe–Lintner model with the Black model as the alternative. These tests are similar to tests Black, Jensen and Scholes (1972) present. The null hypothesis is the expected zero beta portfolio excess return is equal to zero. As in previous studies, the estimates of the expected excess return on the zero beta portfolio are generally greater than zero. The only exception is the fifth time period which includes the years when the market had a large negative return. For the test with twenty beta-sorted portfolios the overall p-value for the null hypothesis less than 0.001. This value differs markedly from the overall p-value of 0.25 for the test of the same model using the same data but an unspecified alternative hypothesis. For forty portfolios the p-value with a specific hypothesis is 0.056 versus a p-value of 0.96 with the vague alternative. These results illustrate the potential for increased power using a specific alternative hypothesis.

The final empirical results are tests of the CAPM using weekly data. In previous sections, it is shown that power gains are possible using more

[23] For the test and some of the following tests, it is necessary to aggregate independent F-statistics to obtain an overall test statistic. The F-statistics are summed together to form an overall test statistic. The null distribution of the aggregate test statistic is approximated by a chi-square distribution. To get the chi-square approximation, the F-distribution for the individual period is approximated by a chi-square distribution and the individual period chi-square distributions are added together. For example, in table 6, the F-test of the Sharpe–Lintner model with twenty portfolios has a null F-distribution with 20 and 39 degrees of freedom. The $F_{20,39}$ can be approximated by (0.086) χ^2 with 12.28 degrees of freedom by matching the first two moments of the distributions. Then, the chi-square distribution for the individual periods can be aggregated giving a null distribution for the six time periods of (0.086) χ^2 with 49.12 degrees of freedom.

Table 6

Tests of Sharpe–Lintner model using five-year time periods of monthly observations from 1/54 to 12/83. The alternative model is the excess return market model.[a]

		Beta-sorted portfolios				Size-sorted portfolios			
		20 portfolios		40 portfolios		20 portfolios		40 portfolios	
Time period	Date	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value
1	1/54–12/58	2.11	0.023	0.94	0.58	1.26	0.26	1.46	0.19
2	1/59–12/63	1.65	0.089	0.99	0.53	1.63	0.094	0.89	0.63
3	1/64–12/68	0.75	0.75	0.48	0.97	1.53	0.13	0.70	0.83
4	1/69–12/73	1.14	0.35	1.18	0.36	1.00	0.48	1.11	0.42
5	1/74–12/78	0.68	0.82	0.61	0.91	1.82	0.054	1.22	0.33
6	1/79–12/83	0.66	0.84	0.52	0.96	0.60	0.89	0.70	0.83
Overall[b]		6.99	0.25	4.72	0.96	7.84	0.082	6.08	0.68

[a] The test presented is the Sharpe–Lintner F-test. The appropriate null distributions are (from left to right) $F_{20,39}$, $F_{40,19}$, $F_{20,39}$, and $F_{40,19}$.

[b] The overall p-values are calculated by approximating the F-distribution with a chi-square distribution and then using the sum of the chi-square distributions for inferences.

Table 7

Tests of Sharpe–Lintner model using five-year time periods of monthly observations from 1/54 to 12/83. Securities are assigned to portfolios using the out-of-period beta. A specific alternative hypothesis is used.[a]

		20 portfolios		40 portfolios	
Time period	Date	$\hat{\gamma}$ (% per month)	$t(\hat{\gamma})^{b,c}$	$\hat{\gamma}$ (% per month)	$t(\hat{\gamma})^{b,c}$
1	1/54–12/58	0.71	3.55	0.44	2.32
2	1/59–12/63	0.74	2.74	0.41	1.71
3	1/64–12/68	0.22	0.71	0.051	0.12
4	1/69–12/73	0.37	1.06	0.53	1.29
5	1/74–12/78	−0.21	−0.49	−0.61	−1.49
6	1/79–12/83	0.58	1.57	0.10	0.26

[a]For the results of the table the restriction $\alpha = \gamma(\iota - \beta)$ is imposed on the intercept vector of the excess return market model. α is the intercept vector and β is the coefficient vector associated with the excess market return. The specific hypothesis tested is $\gamma = 0$.

[b]The t-statistics are calculated using the simulated standard errors of γ. Since the simulated standard errors generally exceed the asymptotic standard errors, this procedure provides a more conservative test.

[c]The overall p-values are 0.0004 for twenty portfolios and 0.056 for forty portfolios. The p-values are calculated using the assumption that the sum of the t-statistics squared has a chi-square distribution with six degrees freedom.

frequent observations. For these tests, we construct weekly returns from the CRSP daily stock return tape. The time period considered is the 1120 weeks from July 4, 1962 to December 20, 1983 inclusive. The 1120 week period is divided into four periods of 280 weeks. Sets of twenty and forty portfolios are formed based on the out of period betas in the same manner as for monthly data. To be eligible for inclusion a stock must have complete returns for the 280-week period under consideration and at least one adjacent 280-week period. The number of stocks eligible for inclusion range from 1235 for the first time period to 1883 in the third time period. Weekly treasury bill returns are constructed from monthly returns by assuming the returns are equal for each week in the month. Although this method of approximation will smooth the weekly returns, the effect on the tests should be minimal. The test results reported in table 8 differ from the tests with monthly data.[24] The Sharpe–Lintner model is rejected in all cases. However, these results should only be interpreted as being suggestive. Unlike for monthly returns, extensive diagnostics assessing the appropriateness of the assumption that returns are independently and identically distributed have not been undertaken for weekly returns.

[24]The results using weekly data and monthly data are not directly comparable. For the weekly results both NYSE and AMEX stocks are used. For the monthly results only NYSE stocks are used.

Table 8

Tests of Sharpe–Lintner model using 280 week time periods of weekly observations from 7/6/62 to 12/28/83. Securities are assigned to portfolios using the out-of-period beta. The excess return market model is the alternative hypothesis.

		20 portfolios		40 portfolios	
Time period	Date	Test statistic[a]	p-value	Test statistic[a]	p-value
1	7/6/62–11/22/67	2.43	0.0008	1.63	0.014
2	11/23/67–4/4/73	0.90	0.59	0.69	0.92
3	4/5/73–8/16/78	2.88	0.0001	1.81	0.0037
4	8/17/78–12/28/83	1.61	0.050	1.08	0.35
Overall[b]	7.82	—[c]	5.21	0.02	

[a]The null distribution for twenty portfolios is $F_{20,259}$, and for forty portfolios the null distribution is $F_{40,239}$.

[b]The overall p-values are calculated by approximating the F-distribution with a chi-square distribution and then using the sum of the chi-square distributions for inferences.

[c]Less than 0.0001.

The empirical results are consistent with the analysis of the first five sections. Using the market model as the alternative hypothesis, the monthly data are consistent with the CAPM. However, the Sharpe–Lintner version of the CAPM can be rejected at low significance levels with a specific alternative hypothesis. Tests conducted with weekly data are not consistent with the CAPM, but further empirical analysis of the appropriateness of the distributional assumptions adopted is necessary before relying on these tests.

7. Summary

This paper addresses the ability of multivariate tests to detect economically important deviations from the Capital Asset Pricing Model. The results indicate that, with an unspecified alternative hypothesis, an important determinant of the power is the type of deviation present. The tests can have reasonable power if the deviation is random across assets. But if the deviation is the result of missing factors (as is the case in many competing models), the tests are quite weak. There exists an upper bound (depending on the missing factor parameters) on the distance the distribution of the test statistic under the alternative can be from the distribution under the null hypothesis. This distance will be relatively small for reasonable missing factor parameters.

Power gains are possible by introducing a specific alternative hypothesis. Using a specific alternative hypothesis we reject the Sharpe–Lintner version of the CAPM. These findings are consistent with earlier tests of the model and with other work which has rejected the CAPM by using a specific hypothesis. For example, see Banz (1981), who rejects the CAPM by specifying an alternative hypothesis with the deviation related to the market value of the equity.

The dependence of the power on the number of portfolios included and the observation interval is investigated. We consider systems of both twenty and forty portfolios. The findings generally favor the use of twenty portfolios, although the results are dependent on the ability to form portfolios without eliminating the violation of the model. The power of tests with forty portfolios and sixty monthly observations is very low when using an unspecified alternative. Under ideal conditions significant increases in power are possible by measuring returns more frequently. In practice the gains may not be as large because decreasing the observation frequency below a monthly interval strains the normality and independence assumptions.

The results suggest that one should be cautious in interpreting the rejection of one model against an unspecified alternative hypothesis as evidence in favor of an alternative model. If an alternative model is available, the relevant comparison is between the current model and the alternative model. A rejection of the current model against an unspecified alternative is often interpreted as evidence in favor of the alternative model. This phenomena has happened somewhat with tests of the CAPM against an unspecified alternative. Initially some researchers interpreted Gibbons' rejection and more recently Shanken's (1985, 1987) rejection of the CAPM as evidence in favor of the Arbitrage Pricing Theory. However, this paper illustrates that the distribution of the test statistic in an APT world is likely not to be very different from the distribution in a CAPM world making such an interpretation, without further investigation, inappropriate.

Appendix: Derivation of non-centrality parameter upper bound

True model specification:

$$\underset{(N\times 1)}{z_t} = \underset{(N\times 1)}{\beta}\, z_{mt} + \underset{(N\times k)}{\Lambda}\; \underset{(k\times 1)}{f_t} + \underset{(N\times 1)}{u_t}\,,$$

$$Ef_t = \mu, \qquad \operatorname{var}(f_t) = V, \tag{A.1}$$

$$Eu_t = 0, \qquad \operatorname{var}(u_t) = \Psi.$$

z_{mt}, f_t and u_t are independent of each other.

Excess return market model specification:

$$z_t = \alpha + \beta z_{mt} + e_t,$$
$$Ee_t = 0, \qquad \mathrm{var}(e_t) = \Sigma. \tag{A.2}$$

For the Sharpe–Lintner model F-test, the non-centrality parameter of the distribution of the test statistic (conditional on z_{mt}) is

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \alpha'\Sigma^{-1}\alpha. \tag{A.3}$$

Taking the expectation of (A.1) and (A.2) gives

$$\alpha = \Lambda\mu, \tag{A.4}$$

$$\Sigma = \Lambda V \Lambda' + \Psi. \tag{A.5}$$

Decompose V such that

$$V = LL', \tag{A.6}$$

where L is of dimension $(k \times k)$ and of full rank. Then define

$$\Gamma = \Lambda L, \tag{A.7}$$

$$\theta = L^{-1}\mu. \tag{A.8}$$

Then

$$\alpha = \Gamma\theta, \tag{A.9}$$

$$\Sigma = \Gamma\Gamma' + \Psi. \tag{A.10}$$

Substitution of (A.9) and (A.10) into (A.3) gives

$$\lambda = T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \theta'\Gamma'[\Gamma\Gamma' + \Psi]^{-1}\Gamma\theta. \tag{A.11}$$

$[\Gamma\Gamma' + \Psi]$ can be inverted analytically [see Morrison (1976, p. 69)],

$$[\Gamma\Gamma' + \Psi]^{-1} = \Psi^{-1} - \Psi^{-1}\Gamma(I + \Gamma'\Psi^{-1}\Gamma)^{-1}\Gamma'\Psi^{-1}. \tag{A.12}$$

Substitution of (A.12) into (A.11) gives

$$\lambda = T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\Gamma'\left[\Psi^{-1}-\Psi^{-1}\Gamma(I+\Gamma'\Psi^{-1}\Gamma)^{-1}\Gamma'\Psi^{-1}\right]\Gamma\theta.$$

Simplifying

$$\begin{aligned}\lambda &= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\left[\Gamma'\Psi^{-1}\Gamma-\Gamma'\Psi^{-1}\Gamma(I+\Gamma'\Psi^{-1}\Gamma)^{-1}\Gamma'\Psi^{-1}\Gamma\right]\theta\\
&= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\left[\Gamma'\Psi^{-1}\Gamma\left(I-(I+\Gamma'\Psi^{-1}\Gamma)^{-1}\Gamma'\Psi^{-1}\Gamma\right)\right]\theta\\
&= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\left[\Gamma'\Psi^{-1}\Gamma(I+\Gamma'\Psi^{-1}\Gamma)^{-1}\right]\theta\\
&= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\left[(I+\Gamma'\Psi^{-1}\Gamma)(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}\theta\\
&= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}\theta. \qquad \text{(A.13)}\end{aligned}$$

To establish the upper bound consider the following identity:

$$\begin{aligned}\theta'\theta &= \theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]\theta\\
&= \theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}\theta\\
&\quad +\theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}(\Gamma'\Psi^{-1}\Gamma)^{-1}\theta. \qquad \text{(A.14)}\end{aligned}$$

From (A.14)

$$\theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}\theta \leq \theta'\theta, \qquad \text{(A.15)}$$

since

$$\theta'\left[I+(\Gamma'\Psi^{-1}\Gamma)^{-1}\right]^{-1}(\Gamma'\Psi^{-1}\Gamma)^{-1}\theta \geq 0.$$

Since $\hat{\sigma}_m^2/\hat{\sigma}_m^2$ is non-negative we have

$$0 < \left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \leq 1. \qquad \text{(A.16)}$$

Since $\theta'[I + (\Gamma'\Psi^{-1}\Gamma)^{-1}]^{-1}\theta$ and $T[1 + \hat{\mu}_m^2/\hat{\sigma}_m^2]^{-1}$ are non-negative it follows from (A.13), (A.15), and (A.16) that

$$\lambda \leq T\theta'\theta, \tag{A.17}$$

which established an upper bound on the non-centrality parameter. Using (A.6) and (A.8), (A.17) can be expressed as

$$\lambda \leq T\mu'(L^{-1})'L^{-1}\mu \quad \text{or} \quad \lambda \leq T\mu'V^{-1}\mu. \tag{A.18}$$

Two examples that can be helpful to interpret (A.18) follow:

(1) There is the market plus one factor ($k = 1$). Then

$$\lambda \leq T\frac{\mu_1^2}{\sigma_1^2}.$$

(2) The K factors are orthogonal to each other (i.e., V is diagonal). Then

$$\lambda \leq T\sum_{k=1}^{K}\frac{\mu_k^2}{\sigma_k^2},$$

where

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_K \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_K^2 \end{bmatrix}$$

References

Amsler, C. and P. Schmidt, 1985, A Monte Carlo investigation of the accuracy of multivariate CAPM tests, Journal of Financial Economics 14, 359–376.

Banz, R., 1981, The relationship between return and market value of common stocks, Journal of Financial Economics 9, 3–18.

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

Black, F., M. Jensen and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, in: M. Jensen, ed., Studies in the theory of capital markets (Praeger, New York) 79–121.

Fama, E. and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, Journal of Political Economy 81, 607–636.

Ferson, W., S. Kandel and R. Stambaugh, 1987, Tests of asset pricing with time-varying risk premiums and market betas, Journal of Finance, forthcoming.

Gibbons, M., 1980, Econometric methods for testing a class of financial models: An application of the nonlinear multivariate regression model, Ph.D. dissertation (University of Chicago, Chicago, IL).

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

Gibbons, M., S. Ross and J. Shanken, 1986, A test of the efficiency of a given portfolio, Research paper no. 853 (Graduate School of Business, Stanford University, Stanford, CA).

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

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

Kandel, S., 1984b, On the exclusion of assets from tests of the mean variance efficiency of the market-portfolio, Journal of Finance 39, 63–75.

Kandel, S. and R. Stambaugh, 1987, On correlations and the sensitivity of inferences about mean–variance efficiency, Journal of Financial Economics 18, forthcoming.

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

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

MacKinlay, A.C., 1985, An analysis of multivariate financial tests, Ph.D. dissertation (University of Chicago, Chicago, IL).

Marsh, T., 1985, Asset pricing model specification and the term structure evidence, Working paper no. 1420-83 (Sloan School of Management, MIT, Cambridge, MA).

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

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

Muirhead, R., 1982, Aspects of multivariate statistical theory (Wiley, New York).

Roll, R., 1977, A critique of the asset pricing theory's tests – Part 1: On past and potential testability of the theory, Journal of Financial Economics 4, 129–176.

Roll, R., 1985, A note on the geometry of Shanken's CSR T^2 test for mean/variance efficiency, Journal of Financial Economics 14, 349–357.

Ross, S., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341–360.

Shanken, J., 1983, An asymptotic analysis of the traditional risk–return model, Unpublished manuscript (University of California at Berkeley, CA).

Shanken, J., 1985, Multivariate tests of the zero–beta CAPM, Journal of Financial Economics 14, 327–348.

Shanken, J., 1986a, Testing portfolio efficiency when the zero–beta rate is unknown: A note, Journal of Finance 41, 269–276.

Shanken, J., 1987, Proxies and asset pricing relations: Living with the Roll critique, Journal of Financial Economics 18, forthcoming.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.

Stambaugh, R., 1981, Missing assets, measuring the market, and testing the capital asset pricing model, Ph.D. dissertation (University of Chicago, Chicago, IL).

Stambaugh, R., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, Journal of Financial Economics 10, 237–268.