Journal of Financial Economics 14 (1985) 327–348. North-Holland

MULTIVARIATE TESTS OF THE ZERO-BETA CAPM

Jay SHANKEN*

University of California at Berkeley, Berkeley, CA 94720, USA

Received October 1982, final version received January 1985

A 'cross-sectional regression test' (CSRT) of the CAPM is developed and its connection to the Hotelling T^2 test of multivariate statistical analysis is explored. Algebraic relations between the CSRT, the likelihood ratio test and the Lagrange multiplier test are derived and a useful small-sample bound on the distribution function of the CSRT is obtained. An application of the CSRT suggests that the CRSP equally-weighted index is inefficient, but that the inefficiency is not explained by a firm size-effect from February to December.

1. Introduction

The Capital Asset Pricing Model (CAPM) has been the focus of empirical testing for over a decade. The primary implication of the model is the mean-variance efficiency of the market portfolio. This is equivalent to the existence of a positive linear relation between an asset's expected return and its covariance with the market return.[1] Thus

$$E_i = \gamma_0 + \gamma_1 \beta_i, \qquad i = 1, \dots, N, \tag{1}$$

where

E_i ≡ expected return on asset i,
β_i ≡ $\mathrm{cov}(R_i, R_m)/\mathrm{var}(R_m)$ is the beta of asset i,
γ_0 ≡ expected return on a zero-beta portfolio,
γ_1 ≡ positive market risk premium,
N ≡ number of left-hand-side assets.

*Thanks to Nai Fu Chen, Craig MacKinlay, Dick Roll, Steve Ross, Gonzalo Rubio, Bill Schwert, G. Sharathchandra, Ken Singleton, Alex Tajirian, Jerry Warner and especially Mike Gibbons (the referee) and Rob Stambaugh for helpful comments on earlier versions of this paper. Particular thanks to Rex Thompson, my thesis advisor at Carnegie Mellon University, for valuable feedback in the early stages of this research. Workshops at Yale University, the University of Southern California, the University of Rochester, Stanford University and the University of California, Berkeley, were of great value in formulating these ideas.

[1]See Fama (1976), Roll (1977), or Ross (1977).

An important test of (1) is conducted by Fama and MacBeth (1973) using an equally-weighted stock index as a proxy for the market portfolio. Their approach is to specify an alternative hypothesis of the form

$$E_i = \gamma_0 + \gamma_1 \beta_i + \gamma_2 Z_i, \tag{2}$$

where Z_i is some attribute of security i. They use residual variance and beta-squared. Under the null hypothesis (1), $\gamma_2 = 0$. Fama and MacBeth test this restriction using, now familiar, two-pass cross-sectional-regression (CSR) methods. They fail to reject efficiency of the market index. Using a similar experimental design, several subsequent studies have investigated the cross-sectional explanatory power of dividend yield and firm size. Much of the evidence appears to be inconsistent with the efficiency of commonly used market proxies.[2]

MacBeth (1975) tests (1) using a multivariate Hotelling T^2 statistic. In contrast to the Fama–MacBeth test, the multivariate approach does not require the specification of a particular alternative hypothesis. The T^2 test directly assesses the statistical significance of deviations from (1), given observed returns on a set of assets. Inference based on MacBeth's test is complicated, however (as he recognizes), by the fact that the test requires observability of market betas as well as a time series of returns on the minimum variance zero-beta portfolio. In practice, estimates of these quantities are substituted in the test statistic, rendering its distribution unknown, even in large samples. Nonetheless, MacBeth's work constitutes an important early exploration of the multivariate framework.

Gibbons (1982) employs maximum likelihood techniques in a multivariate test of (1). Inference is based on a standard likelihood ratio test (LRT) statistic, in conjunction with its limiting chi-squared distribution. Using a one-step Gauss–Newton computational method, a strong statistical rejection of the efficiency of the equally-weighted index is obtained. The LRT procedure incorporates (at least asymptotically) the fact that all parameters in (1) are unknown and estimated jointly. In this respect, it appears to dominate the approximate T^2 procedure discussed above, which does not fully account for the existing parameter uncertainty. The question remains, however, whether asymptotic statistical analysis adequately captures the features of the distribution of the LRT that are important to the inference process for finite sample sizes.

Stambaugh (1982) reports some simulation evidence indicating that, in fact, the LRT does *not* conform closely to its limiting distribution. The null hypothesis is rejected too often, the problem becoming more serious as the

[2] Banz (1981) and Litzenberger and Ramaswamy (1979) are good examples.

number of left-hand-side assets increases.[3] A related Lagrange multiplier test (LMT) appears to conform more closely to the chi-squared distribution. Using the LMT, Stambaugh fails to reject (1) at conventional significance levels with a variety of market proxies.

Jobson and Korkie (1982) modify the LRT statistic using 'Bartlett's correction factor'. Transformed in this manner, the statistics reported by Gibbons no longer reject the null hypothesis, consistent with Stambaugh's results. Jobson and Korkie also report tests of the zero-beta model using statistical methodology developed for the Sharpe–Lintner specification, in which there is an observable riskless rate of return r. An estimate of γ_0 is substituted for r. Kandel (1984) shows that substitution of the MLE of γ_0 for r in the LRT statistic of the Sharpe–Lintner model produces the LRT statistic of the zero-beta model.[4] Thus, this class of tests of the zero-beta model need not be considered separately.[5]

Still another multivariate test is developed by Shanken (1980). The test statistic is a simple quadratic form in the residuals from a generalized least squares (GLS) version of the traditional CSR procedure. Like the LRT and LMT, this 'CSR test' is asymptotically distributed as chi-squared. Its small-sample behavior is the main focus of this paper.

The existing literature leaves several important questions unanswered. We know very little, from an analytic perspective, about the small-sample properties of the tests that have been proposed or of the relations between the various tests. Simulation evidence favors the LMT over the LRT, but we are necessarily uncertain as to the reliability of the LMT outside the range of parameter values investigated thus far. Ideally, we would like to know the exact small-sample distribution of our test statistic. While the research presented below does not take us quite that far, it does provide useful insights into the nature of this distribution. Since many asset pricing models share essentially the same statistical structure as the zero-beta model, the relevance of this research extends beyond the CAPM context.

Section 2 introduces the cross-sectional-regression test (CSRT) and develops some relations between it and other multivariate tests. The small-sample behavior of the CSRT is analyzed in section 3. Empirical results are reported in section 4, and a summary of the paper is presented in section 5. Some

[3] Gibbons also reports simulation evidence on the LRT for $N = 5$. Significant departure from the chi-squared distribution was observed only when the length of the time series was equal to 30. With the benefit of hindsight we know this was due to the small number of assets employed for computational purposes.

[4] See Gibbons, Ross and Shanken (1984) for a complete analysis of the specification with a riskless asset.

[5] The degrees of freedom associated with Sharpe–Lintner tests are inappropriate for zero-beta tests, however, since γ_0 is an additional unknown parameter that must be estimated.

alternative specifications are considered in appendix A. Technical proofs are deferred to appendices B and C.

2. A cross-sectional-regression T^2 test of linearity

2.1. Testing for equality of expected returns

In this section, a test of the hypothesis that all securities in a given set have the same expected return is formulated. Let R_t be an N-vector of returns at time t, distributed as multivariate normal with mean vector E and covariance matrix V. Returns are assumed to be serially independent. The null hypothesis is that there exists a scalar γ such that $E = \gamma 1_N$, where 1_N is a vector of ones.

To simplify the problem, let R_t^* be an N^*-vector ($N^* \equiv N - 1$) obtained by subtracting R_{Nt} from $R_{1t}, \ldots, R_{N^*t}$. The assumptions above imply R_t^* is independently and identically normally distributed over time. The null hypothesis is equivalent to the condition $E(R_t^*) = 0$. The standard multivariate procedure for testing this condition is based on Hotelling's T^2 statistic

$$T\bar{R}^{*\prime} S^{-1} \bar{R}^*,$$

where $\bar{R}^*$ is the time series mean and S the usual unbiased sample covariance matrix of the R_t^*. T is the length of the time-series.

Inference is facilitated by the fact that

$$F \equiv T^2 (m - n + 1)/mn \tag{3}$$

is distributed as an F variate with degrees of freedom n and $m - n + 1$. The corresponding distributions will be denoted $F(n, m - n + 1)$ and $T^2(n, m)$. In the application above, $n = N^*$ and $m = T - 1$. When $n = 1$, we have the square of a Student t variate with m degrees of freedom, distributed as $F(1, m)$. Thus, the T^2 test is a multivariate generalization of the standard two-sided t-test.

A CSR interpretation of the T^2 test for equality of expected returns will prove useful in developing and understanding tests of the CAPM hypothesis. The T^2 test statistic above is identical to[6]

$$Te'\hat{V}^{-1}e,$$

where

$$e \equiv \bar{R} - \hat{\gamma} 1_N \quad \text{and} \quad \hat{\gamma} \equiv \left(1_N' \hat{V}^{-1} 1_N\right)^{-1} 1_N' \hat{V}^{-1} \bar{R}.$$

[6] The proofs of this and other assertions made in section 2 are similar to the proofs of results in section 3 and are omitted.

$\bar{R}$ is the time series mean of the R_t and $\hat{\gamma}$ is the MLE of γ for this model. $\hat{V}$ is the usual unbiased sample covariance matrix of the R_t. Thus, the test statistic is a GLS-weighted residual sum of squares from a CSR of $\bar{R}$ on 1_N.[7]

2.2. *Testing the CAPM with beta known*

The CAPM security market relation may be expressed as

$$E = X\Gamma, \tag{4}$$

where

$$X \equiv [1_N : \beta] \quad \text{and} \quad \Gamma \equiv (\gamma_0, \gamma_1)'.$$

β is an N-vector of security betas. In this section assume that beta is known. Analysis of this case enhances our understanding of the more realistic specification considered later.

To test (4) we consider a natural generalization of the T^2 statistic of section 2.1:

$$Q \equiv Te'\hat{V}^{-1}e, \tag{5}$$

where

$$e = \bar{R} - X\hat{\Gamma} \quad \text{and} \quad \hat{\Gamma} \equiv (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}\bar{R}.$$

$\hat{\Gamma}$ is the MLE of Γ for this model in which Γ, E and V are the unknown parameters.[8] Q is distributed as $T^2(N-2, T-1)$.[9] The degrees of freedom $N-2$ reflect the fact that there are now two independent variables in the GLS CSR. As before, $T-1$ is the divisor which makes $\hat{V}$ an unbiased estimator of V.

The intuition underlying the CSRT is quite simple. Q is essentially a 'goodness of fit' measure, with respect to the relation (4). When the null hypothesis is true, the residual vector e is randomly distributed about zero. The randomness arises from the fact that E and Γ are unknown and must be estimated. Since $\hat{V}^{-1}$ is positive definite, the test statistic is always positive.

[7] The term GLS is used loosely throughout the paper in that we always work with an estimated covariance matrix, not the true parameters.

[8] This specification ignores the information which knowledge of beta provides about the covariance matrix of returns. While this information could certainly be incorporated, pursuit of this topic would be inconsistent with our pedagogical aims in this section. In section 3, the artificial assumptions of this section are dropped.

[9] In contrast to MacBeth (1975), we do not require that a time series of returns on a zero-beta portfolio be observed.

Table 1

Comparison of p-values obtained under alternative distributional assumptions.[a]

Test statistic	p-values	
	$\chi^2(38)$	$T^2(38,59)$
38.2	0.46	1.00
49.5	0.10	0.97
60.7	0.01	0.92
103.2	0.00	0.50
170.9	0.00	0.10

[a] If the test statistic is distributed as $\chi^2(38)$, the probability of observing a value greater than or equal to 49.5 is 0.10; if the distribution of the statistic is Hotelling T^2 with 38 and 59 degrees of freedom, the probability of observing a value greater than or equal to 49.5 is 0.97.

When the null is false, however, e is distributed about some non-zero vector which reflects the deviation from (4). This imparts an upward bias to the test statistic. Unusually large values of Q, therefore, suggest rejection. The notion 'unusual' is evaluated relative to the distribution of Q under the null hypothesis.

In the current context this distribution is known, but under more realistic assumptions the problem of identifying the distribution of the test statistic is more complex. It is common, in such situations, to resort to the use of asymptotic statistical approximations. The adequacy of such approximations varies greatly from one problem to another, however. Some valuable insight into the present problem may be obtained by comparing asymptotic inferences with those based on the T^2 distribution, the exact distribution of Q under our simplifying assumption that beta is known.

As $T \to \infty$, $T^2(N-2, T-1)$ converges in distribution to $\chi^2(N-2)$. Suppose one were to use a chi-squared table to assess the statistical significance of the CSRT statistic Q. How would the inferences be affected? Table 1 provides some comparisons for $N=40$ and $T=60$, which reflect the sample sizes in Gibbons (1982). A test statistic as large as 103 would be observed about half the time with Q distributed as $T^2(38,59)$. The corresponding p-value[10] for a $\chi^2(38)$ variate, however, is less than 10^{-5}. $Q=62$ is sufficient to reject the null at the 0.01 level based on the chi-squared distribution. Yet this would occur more than ninety percent of the time when (4) is, in fact, true.

To summarize, reliance on asymptotic theory would result in excessive rejection of the null hypothesis when it is true, i.e., a large type I error. The discrepancy between the small sample and asymptotic distributions can be attributed entirely to the stochastic behavior of $\hat{V}^{-1}$.[11] This follows from the

[10] The p-value is the probability of exceeding a given level of the test statistic.

[11] The expected value of $\hat{V}^{-1}$ is $[(T-1)/(T-N-2)]V^{-1}$. The bias is considerable when N is not small compared to T. See Press (1970, pp. 107, 112).

observation that if $\hat{V}^{-1}$ is replaced by V^{-1} in (5), the exact distribution of Q is $\chi^2(N-2)$, the same as the limiting distribution.[12]

Recall that when $n=1$, the T^2 statistic is the square of a Student t variate. With the variance known, the corresponding statistic is the square of a standard normal variate, i.e., $\chi^2(1)$. When m is small, the t density has much more mass in its tails than the normal density. There is a greater probability of obtaining extreme observations with the former, due to the stochastic behavior of the sample variance. Outliers, positive or negative, map into large positive values of the squared variable. This provides some intuition for the excessive rejection of the null when the chi-squared table is used to assess statistical significance. The t distribution converges rapidly to standard normal as $m \to \infty$. The multivariate case is more complicated, however, as the *relative* values of m and n (equivalently T and N) become relevant.

2.3. *Alternative tests of linearity*

Except for a degrees of freedom adjustment, the CSRT of the previous section is a Wald test. The actual Wald statistic is $Q^* \equiv [T/(T-1)]Q$. Two other tests, the LRT and the LMT, were mentioned in the introduction. Under mild regularity conditions, the three test statistics have the same limiting chi-squared distribution. For the model considered here, the statistics are actually increasing transformations of each other:[13]

$$LRT = T\ln(1+Q^*/T) \quad \text{and} \quad LMT = TQ^*/(T+Q^*). \tag{6}$$

It follows that the ordering

$$LMT < LRT < Q^* \tag{7}$$

must, with probability one, hold in every sample.[14]

In our example, with $N=40$ and $T=60$, the median of Q is 103.2. The corresponding values of LRT and LMT, computed from (6), are 60.7 and 38.2, respectively. Each of these statistics therefore, has a true p-value of 0.5. Evaluated with respect to $\chi^2(38)$, the p-values are 0.00, 0.01, and 0.46 for Q, LRT and LMT, respectively. This example demonstrates that when one relies on asymptotic inference, the conclusions reached may depend on the particular test statistic adopted.[15] In particular, p-values for the CSRT and the LRT may

[12] See Theil (1971, p. 240).

[13] Evans and Savin (1982) derive the same relations in the case of linear restrictions on the classical linear regression model.

[14] Berndt and Savin (1977) derive these inequalities in the case of linear restrictions on a multivariate linear regression model.

[15] Of course, if significance is assessed with respect to the true distribution of each test statistic in (7), then the tests must produce identical inferences.

seriously overstate the statistical significance of the results. These observations are consistent with the simulation evidence discussed in section 1. More will be said about these issues later.

3. Testing linearity when beta is unknown

3.1. An errors-in-variables adjustment

In this section, beta is added to the space of unknown parameters and the more realistic model specification is analyzed. It is convenient, in this case, to work with the 'market model' parameterization of the joint distribution of returns:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}, \qquad i = 1, \ldots, N.$$

Let Σ be the covariance matrix of the ε_{it}. Σ is the conditional covariance matrix of returns, given R_m. The sample covariance matrix of the N time series of OLS market model residuals (with $T-2$ in the denominator) is an unbiased estimator of Σ and is denoted $\hat{\Sigma}$. $\hat{\beta}$ is the N-vector of OLS estimators and $\hat{X} \equiv [1_N \colon \hat{\beta}]$.

Although the CAPM constrains expected return to be linear in beta, this constraint is nonlinear in a statistical sense, as the unknown parameters γ_1 and β_i enter multiplicatively. The MLE $\hat{\Gamma}_M$ for this model is distinct from the GLS CSR estimator which we define as[16]

$$\hat{\Gamma}_C \equiv (\hat{X}'\hat{\Sigma}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Sigma}^{-1}\bar{R}.$$

A quadratic form, similar to that of section 2, plays a central role in the analysis below. Use of the same symbol Q should not be a source of confusion. Since two estimators of Γ are of interest, the definition is given in terms of an arbitrary estimator $\hat{\Gamma}$:

$$Q \equiv Te'\hat{\Sigma}^{-1}e \quad \text{where} \quad e \equiv \bar{R} - \hat{X}\hat{\Gamma}. \tag{8}$$

A subscript, M for MLE or C for CSR, indicates the particular estimator employed.

The simple formulas for LRT and LMT, assuming beta known, continue to hold for the current specification with a few modifications. We define an

[16] The GLS CSR estimator is actually identical to a one-step Gauss–Newton estimator which employs OLS $\hat{\beta}$ as the initial consistent estimator of β. Furthermore, the same estimator is obtained with $\hat{V}$ in place of $\hat{\Sigma}$ [see Shanken (1982, ch. 2)]. When beta is unknown, working with $\hat{\Sigma}$ facilitates the small-sample statistical analysis. For example, assuming joint normality, $\hat{\Sigma}$ is conditionally independent of $\hat{\beta}$, given R_m, while $\hat{V}$ is not.

adjusted version of Q as

$$Q^A \equiv Q/(1+\hat{\gamma}_1^2/s_m^2) \tag{9}$$

where s_m^2 is the MLE of $\mathrm{var}(R_m)$, and let $Q^* \equiv [T/(T-2)]Q^A$, reflecting the use of the unbiased estimator of Σ in Q. LRT and LMT are now given by (6) with $\hat{\Gamma}_M$ employed throughout.[17] The corresponding one-step test statistics are obtained by using $\hat{\Gamma}_C$ in Q^A. Q_C^A is a natural generalization of the CSRT statistic of the previous section to the case where beta is unknown. It will be referred to as the CSRT throughout the remainder of this paper.[18]

Introduction of uncertainty with respect to beta has resulted in an adjustment to the quadratic form of cross-sectional residuals. The particular form of the adjustment is a consequence of two facts. First, the covariance matrix of the beta estimation errors is proportional to the covariance matrix of the market model disturbances; specifically

$$\mathrm{var}[\sqrt{T}(\hat{\beta}-\beta)] = (1/s_m^2)\Sigma.$$

Second, the impact of error in estimating beta depends on the value of γ_1 in the linear expected return relation. The variance of this effect is thus proportional to γ_1^2. Additional details are contained in appendix B.

Since the adjustment term $1+\hat{\gamma}_1^2/s_m^2$ exceeds one, Q^A is always less than Q. As $T \to \infty$, Q^A, LRT and LMT converge in distribution to $\chi^2(N-2)$. Without the 'errors-in-variables adjustment', all three tests would reject the null hypothesis too often, asymptotically, even though $\hat{\beta}$ converges to β, as $T \to \infty$. In contrast, the excessive rejection discussed in section 2, due to the stochastic behavior of the estimated covariance matrix, is only a problem in small samples.

As earlier, the various tests may produce very different inferences in small samples when significance is evaluated with respect to the chi-squared distribution. Indeed, it follows from (6) that LMT is bounded above by T. Suppose $N=40$, $T=60$. Since the 99th percentile of $\chi^2(38)$ is 61.2, it is impossible to reject the null hypothesis at this level with the LMT using its asymptotic

[17] These results extend those cited in footnotes 13 and 14 since the constraints are nonlinear when beta is unknown. The general strategy employed in the proof is as follows. The first-order condition for maximization of the restricted likelihood function with respect to beta is used to obtain the MLE of beta as a function of $\hat{\Gamma}_M$ and $\hat{\beta}$. Similarly, the MLE of Σ is expressed in terms of $\hat{\Gamma}_M$, $\hat{\beta}$ and $\hat{\Sigma}$. Recall that $\hat{\beta}$ and $\hat{\Sigma}$ are the unrestricted estimators based on ordinary least squares 'market model' regressions. Algebraic manipulation then leads to (6). The details are available on request.

[18] Q^* is not a Wald statistic for the current nonlinear statistical specification. This can be inferred indirectly from the fact that the mean of the Wald test statistic was much lower than the mean of the LRT statistic in Stambaugh's simulations, while (7) implies Q^* must have a higher mean than LRT in every sample.

distribution! More generally, the inequalities (7) explain why LMT rejects less often than LRT when using asymptotic chi-squared tests.

3.2. Characteristics of the small-sample distribution of the CSRT

Introduction of uncertainty with respect to beta significantly complicates analysis of the exact distributions of our test statistics. We focus on the statistic Q_C. In appendix B, the small-sample distribution of Q_C is shown to be a mixture of non-central T^2 distributions. The random non-centrality parameter is unobservable, as it is a function of the difference between the true betas and the OLS estimates. A useful implication, however, is that the exact distribution function of Q_C is bounded above by the central $T^2(N-2, T-2)$ distribution function.[19] Equivalently, p-values computed with respect to this distribution understate the true p-values. Since the GLS estimator minimizes Q (viewed as a function of $\hat{\Gamma}$), $Q_C \leq Q_M$. Hence the distribution function of Q_M is also bounded above by the central $T^2(N-2, T-2)$ distribution function.

Inverting the relation (6), Q_C can be computed from the one-step LRT statistics and estimates reported in Gibbons (1982). Applying the results above to his ten subperiod statistics, an aggregate p-value of 0.75 is obtained.[20] This is a lower bound on the true p-value and contrasts dramatically with the p-value (less than 0.001) obtained from the asymptotic chi-squared distribution. In particular, no rejection of the null hypothesis is suggested. The LMT statistics in Stambaugh (1982) have been transformed in a similar manner. Once again, no rejection is indicated, thus supporting the general empirical conclusions of that paper. It should be emphasized that, in these cases, our inferences are not dependent on asymptotic approximations. In this respect, they differ from all previous work on the zero-beta model.

Suppose the p-value for Q, relative to the central T^2, is smaller than the specified significance level. Rejection of the null hypothesis might not be appropriate, as we know the true p-value is larger than that computed. Thus, an approximation to the actual distribution function of Q is needed.[21] We have seen that standard asymptotic results, involving the chi-squared distribution, are not very helpful here. A somewhat unconventional asymptotic analysis, presented in appendix C, would appear to be more promising. The main implications of that analysis are summarized below.

[19] The fact that the non-central F distribution function is decreasing in the non-centrality parameter is used. See Johnson and Kotz (1970, p. 193).

[20] The statistics are aggregated by a method discussed in section 4 below. For completeness, it should be noted that Gibbons' one-step estimator takes the OLS CSR estimator as the initial consistent estimator, whereas our results assume the GLS CSR estimator is used. This simplifies some of the formulas. Computationally, the difference is negligible. Also note that Gibbons imposes the constraint $\gamma_0 + \gamma_1 = E_m$. this is discussed in our appendix A.

[21] A lower bound on the distribution function might also be useful and is currently under investigation.

The CAPM, together with our assumptions of temporal independence and joint normality, yields the following constraint on expected return, conditional on market return:

$$E(\bar{R}_i | R_m) = \gamma_0 + \bar{\gamma}_1 \beta_i,$$

where $\bar{\gamma}_1 \equiv \gamma_1 + (\bar{R}_m - E_m)$ is the ex post market price of risk. (E_m is the expected return on the market portfolio.) The CSRT and the equivalent Lagrange multiplier and likelihood ratio tests essentially evaluate this conditional linearity hypothesis. An interesting implication is that such tests are meaningful, even if the empiricist knows ex post that realized market returns deviated greatly from expectations in a given subperiod. Furthermore, it is the (approximate) joint normality of the *conditional* distribution of returns, given R_m, that is of relevance from the statistical perspective.[22] The ex ante distribution of R_m plays no particular role in the linearity tests, although the realized sample quantities $\bar{R}_m$ and s_m^2 do.

Use of asymptotic principles to obtain an approximate distribution for the random non-centrality parameter, referred to at the beginning of this section, suggests that Q_C is approximately distributed as $(1 + \bar{\gamma}_1^2/s_m^2)T^2(N-2, T-2)$, given R_m; i.e., as a multiple of a *central* T^2. Note that the degrees of freedom $T-2$ is the divisor which makes $\hat{\Sigma}$ an unbiased estimator of Σ. Substituting the estimator $\hat{\gamma}_{1C}$ for $\bar{\gamma}_1$ yields $T^2(N-2, T-2)$ as an approximate distribution for Q_C^A.[23] Thus, our analysis simply boils down to replacing Q by Q_C^A and modifying the degrees of freedom in the CSRT of section 2.2.

The adequacy of our approximation depends, in part, on the accuracy of the estimator of $\bar{\gamma}_1$. When the ratio $\bar{\gamma}_1^2/s_m^2$ is small, precise estimation may not be crucial. Precision, in this context, involves the variability of the estimator around $\bar{\gamma}_1$, conditional on R_m. Variation in $\bar{\gamma}_1$, due to the difference between realized and expected market return, is irrelevant. This potentially large source of variation becomes relevant, however, when testing for the existence of a positive ex ante risk premium. A test of this hypothesis is of interest provided that the linearity relation (1) is not rejected by the CSRT.

Recent simulation evidence reported by Amsler and Schmidt (1985) reflects favorably on the use of the T^2 approximation to the CSRT statistic, in the usual CAPM contexts. The evidence indicates that Q_M^A may be satisfactorily

[22] The unconditional joint distribution of returns is certainly relevant from an economic equilibrium perspective.

[23] The analysis in appendix C is based on the properties of Q_C. It is not necessary that the CSR estimator be used in the denominator of the adjusted statistic Q^A, however. Given the results of Shanken (1982, ch. 2), it might be preferable to use the MLE of γ_1, or a CSR estimator, modified for errors-in-variables induced bias, in the denominator. Such modifications had little effect on the results in section 4, but may be more relevant in other applications.

approximated by a T^2 distribution as well. MacKinlay (1984) also presents simulation results for several multivariate tests. One must, of course, be careful in extrapolating from any of these results to new situations.

In comparing the T^2 CSRT to the various asymptotic chi-squared tests, it is important to recognize the following. All of these tests implicitly deal with the 'errors-in-variables problem' in the same manner.[24] Each test statistic is a transformation of the quadratic form Q which is adjusted downward to reflect the noise introduced through estimation of beta. What distinguishes the T^2 test from the other approaches is the way in which it incorporates the considerable (small-sample) variability in the estimator of the covariance matrix. As was illustrated earlier, this variability has a substantial impact on the properties of the resulting test statistic.

4. Empirical applications

Multivariate tests hold out the promise, in principle, of permitting the researcher to detect departures from an asset pricing relation without having to venture a guess as to the potential source of misspecification. In practice, however, there are a number of limiting factors. Since the tests involve inversion of an $N \times N$ covariance matrix, computational considerations necessitate some form of data reduction. Furthermore, invertibility of the covariance estimator requires that T be greater than N, while stationarity considerations dictate that T not be too large.

One alternative is to use a subset of the available securities in the test. Given the considerable variability of individual security returns, however, such a test might not be very powerful. More commonly, securities are aggregated into portfolios, thereby reducing variability. There is the danger, however, as Roll (1979) has noted, that individual deviations from the asset pricing relation may cancel out in the portfolios and escape detection. Hence the importance of using a 'suitable' grouping variable.[25]

If the empiricist is willing to model the nature of the deviation from a given expected return relation, then a traditional experimental design as in (2) might be preferred to the multivariate approach. For example, it might be hypothesized that the deviation is linearly related to dividend yield. In other contexts, a

[24] Note that the concern here is with the influence of estimation error in $\hat{\beta}$ on the properties of the statistic for testing the validity of the expected return relation. The expression 'errors-in-variables problem' usually refers to the bias in the CSR estimator of gamma, induced by error in estimating beta.

[25] The common procedure of grouping on the basis of estimated beta is generally motivated by a desire to obtain efficient estimates of the gamma parameters under the null hypothesis that the CAPM expected return relation is valid. Roll's concern, on the other hand, is with the power of the test, i.e., the ability to reject the null hypothesis when it is, in fact, false. Grouping on beta need not be 'optimal' from this perspective.

grouping variable might be specified without a strong prior as to the functional form of its relation to expected return. In particular, monotonicity of the relation might be questionable. The generality of the multivariate approach can be valuable in such cases, as no particular model of the deviations is required. Even if the traditional approach is adopted and a relation between expected return and some variable Z is established, one may wish to further assess the significance of deviations from the expanded relation (2), which includes Z as an independent variable. Here again, the multivariate test can be a useful tool of analysis. An application of this sort is presented below.

Previous multivariate tests of the CAPM, employing beta-sorted or industry portfolios, all fail to reject the model when the small-sample characteristics of the test statistics are accounted for. Since many studies have indicated that deviations from the CAPM are related to firm size, it is of interest to determine whether the CSRT rejects the null hypothesis when size is used as the grouping variable. This is the starting point of our empirical analysis.

For each of three subperiods, the following steps are taken: (i) all securities on the CRSP monthly return tape with complete data for the subperiod are ranked on the basis of total value of all shares outstanding at the end of the month preceding the subperiod, and (ii) the securities are grouped into twenty equally-weighted portfolios. Each portfolio contains approximately the same number of securities. The portfolios are ranked from one to twenty, portfolio one containing the smallest firms and portfolio twenty the largest.[26] The portfolio rank will serve as the variable Z in our analysis.

The three subperiods, each of length $T = 74$ (months), are February 1953 to March 1959, April 1959 to May 1965, and June 1965 to July 1971.[27] Real returns are computed using the consumer price index. CSR tests of the efficiency of the CRSP equally-weighted index are reported in table 2. The constraint $\gamma_0 + \gamma_1 = E_m$ is imposed.[28] Also reported, for comparison, are tests of the hypothesis that the twenty portfolios have the same expected return. F statistics are obtained by the transformation in (3).

In attempting to aggregate subperiod F statistics, we are confronted with the fact that sums of F variates do not conform to a tabulated distribution. The following procedure has therefore been adopted. For each subperiod statistic,

[26] Initially, forty-one portfolios were formed and the twenty-first portfolio was deleted for reasons unrelated to the present study. Twenty portfolios were obtained from the forty by equally weighting the first and second, third and fourth, etc. The use of twenty portfolios is motivated by the desire to be comparable with some of the previous multivariate studies. Analysis, not reported here, employing ten portfolios fully supports the conclusions obtained with twenty portfolios. Similar results were also obtained with a value-weighted index.

[27] A comprehensive revision of the CPI was completed in January 1953. Wage and price controls were imposed in August 1971. Note that our subperiods coincide with the first three subperiods in Stambaugh (1982).

[28] See appendix A.

Table 2

Cross-sectional-regression F tests for identical expected real returns (*EQUAL*), efficiency of the CRSP equally-weighted index (*CAPM*),[a] and linearity of expected return in beta and portfolio rank (*SIZE*). All tests use 20 value-sorted portfolios. The lower the rank of a portfolio the smaller are the firms in that portfolio. January returns have been deleted for tests in the second panel.

Distribution specification	Subperiod test statistics (*p*-value) 2/53–3/59	4/59–5/65	6/65–7/71	(overall *p*-value)
	(A) *Including January returns*			
$F(19,55)$	2.05	0.82	1.46	
EQUAL	(0.02)	(0.67)	(0.14)	(0.061)
$F(19,54)$	1.96	1.45	1.47	
CAPM	(0.03)	(0.14)	(0.13)	(0.009)
$F(17,56)$	1.57	0.88	0.92	
SIZE	(0.11)	(0.59)	(0.56)	(0.308)
	(B) *Excluding January returns*			
$F(19,49)$	2.35	0.90	1.69	
EQUAL	(0.01)	(0.58)	(0.07)	(0.018)
$F(19,48)$	2.22	1.44	1.69	
CAPM	(0.01)	(0.15)	(0.07)	(0.003)
$F(17,50)$	2.44	0.94	1.71	
SIZE	(0.01)	(0.54)	(0.07)	(0.014)

[a] The constraint $\gamma_0 + \gamma_1 = E_m$ is imposed as described in appendix A.

the standard normal z corresponding to the given p-value is determined. The subperiod z's are then added and divided by $\sqrt{3}$, to obtain an aggregate N(0,1) statistic from which an aggregate p-value may be determined.

While equality of expected returns cannot be rejected at the 0.05 level, efficiency of the equally-weighted index *is* rejected at the 0.01 level. Note that rejection of the efficiency relation (1) using value-sorted portfolios does not necessarily imply a 'size-effect'; i.e., the deviations from (1) need not be monotonically related to firm size. To determine whether the rejection is indeed driven by size, the mean return vector $\bar{R}$ was regressed on a constant, $\hat{\beta}$, and a proxy for size – the portfolio rank. The results, not reported here, are consistent with the evidence in Banz (1981), who uses a traditional CSR approach to document a size effect.[29]

The finding of a 'significant' coefficient on the size variable is sufficient to reject efficiency of the CRSP index. It is of further interest, however, to determine whether size completely 'accounts for' the misspecification of (1). This sort of question is not typically addressed in CSR studies, but is easily

[29] The results are reported in Shanken (1982, ch. 3).

handled within the multivariate framework. We wish to test the constraint (2), where Z_i is our size rank variable. The CSRT statistic is similar to that in (9); e is now the N-vector of residuals from the GLS CSR which includes Z_i as an additional independent variable. This Q_C^A is approximately distributed as $T^2(N-3, T-2)$; $N-3$, since there are three independent variables in the CSR, and $T-2$ since $\hat{\Sigma}^{-1}$ is still used in the quadratic form (8).[30] Based on the evidence in table 2, the hypothesis that size completely accounts for the deviations from (1) cannot be rejected.

Unadjusted for errors-in-variables, the statistics for testing the size specification are 1.57, 0.90 and 0.92. Comparing these with the adjusted statistics in table 1, we find very little difference. This is due to the fact that the largest value of $\hat{\gamma}_1^2/s_m^2$, which occurred in the second subperiod, was only 0.02. The overall p-value of 0.29 for the unadjusted statistics underestimates the true p-value. Thus, we are assured that our failure to reject the null hypothesis is not due to problems with asymptotic inference.[31]

Keim (1983) has recently reported evidence that the size effect is more pronounced in January than in the other months. He attributes nearly half of the effect, over the period 1963–1979, to January abnormal returns, and suggests that two separate phenomena may be at work. To assess the extent to which the rejection of efficiency observed here is due to a 'January effect', the tests described above were rerun with January returns deleted. Consistent with Keim's observations, both the magnitude and significance of the size effect (not reported here) were substantially reduced. In light of this, one might expect a weaker rejection of (1) by the CSRT with the January returns deleted. This need not be the case, however.

Suppose the market model is reasonably well specified for February–December, but the return process is fundamentally different in January. If the CRSP index is inefficient with respect to the February–December joint distribution then the probability of detecting this departure from the null hypothesis could increase with the January source of variation removed. In other words, elimination of data might actually increase the power of the test. The results reported in the second panel of table 2 are consistent with this scenario.

Equality of expected returns and efficiency of the equally-weighted index are strongly rejected by the data. In each case, the p-values are lower than the corresponding entries in the first panel. In striking contrast to the earlier results, inclusion of the size rank variable does not greatly improve the fit of the expected return relation with the January returns deleted. Thus, the CRSP index appears to be inefficient even apart from the size and January effects

[30] The proof is similar to that in appendix C and is omitted. Note that Z_i is a known variable. Estimation error in the additional independent variable would further complicate the statistical analysis.

[31] The small-sample bound on the distribution function does rely on the assumption of normality, however. The proof is similar to that in appendix B and is omitted.

found in historical returns. Further exploration of this phenomenon is left to future research.[32]

5. Summary and conclusions

Several multivariate tests of the zero-beta CAPM have been proposed in the literature. Since the three basic test statistics are exact transformations of one another, there is really just one test. The three alternative test statistics conform to a given ordering in every sample, yet all three have the same asymptotic chi-squared distribution. If this distribution is taken as the reference point for drawing inferences, different conclusions may be reached, depending on which test statistic is employed – a problem previously encountered elsewhere in the econometrics literature.

Simulation evidence reported by Stambaugh (1982) indicates that the likelihood ratio test does not conform well to the chi-squared distribution. The null hypothesis is rejected too often when the number of market model equations N is even moderately large in relation to the time series length T. While the Lagrange multiplier test performs better in simulations, we have observed analytically that it suffers from the reverse problem – it accepts the null hypothesis too often when N is large in relation to T.

An analysis of the cross-sectional regression test (CRST) has revealed its close relation to the Hotelling T^2 test and highlighted the central statistical role played by the estimator of the covariance matrix. The T^2 distribution has been proposed as a useful approximation to the exact distribution of the CSRT, after a simple adjustment for error in the estimation of beta. It was noted that this proposal appears to be consistent with the existing simulation evidence.

It has been proven that under certain circumstances, frequently encountered in practice, inferences may be made without appeal to asymptotic approximations. This occurs when the value of the CSRT statistic, unadjusted for 'errors-in-variables', is sufficiently small, indicating that the null hypothesis cannot be rejected. This was observed to be the case for the data in Gibbons (1982) and Stambaugh (1982) and constitutes the only small-sample result obtained thus far, in the zero-beta literature.

An empirical application of the CSRT suggests that the CRSP equally-weighted index is inefficient, but that the inefficiency is not explained by a firm size-effect from February to December. This application illustrates the value of

[32]As do several previous studies, we require that firms have complete data for a given subperiod. A 'survivorship bias' related to firm size would not appear to explain the differences between the two panels of table 1, however. Of course, when viewed as tests of the CAPM, our results are subject to the usual ambiguities associated with the use of market proxies. See Roll (1977) for a discussion of these issues.

the multivariate test as a tool to be used in conjunction with more traditional methods and not necessarily as an alternative to those methods.

Appendix A: Alternative specifications

In this appendix, extensions to a multi-factor asset pricing specification are stated. The proofs involve minor modifications of the one-factor case and are omitted. Let δ be a $K \times 1$ random vector and consider the following generalization of the usual market model regressions:

$$R_{it} = \alpha_i + \beta_i \delta_t + \varepsilon_{it}, \qquad i = 1, \ldots, N, \tag{A.1}$$

where the ε_{it} are assumed to be independent and identically distributed over time, independent of δ, with mean zero and covariance matrix Σ. Furthermore, the joint distribution of the ε_{it} is assumed to be multivariate normal. β_i is a row vector of regression coeffcieints. α and β are the corresponding $N \times 1$ and $N \times K$ matrices of parameters, assumed to be constant over time.

The relation to be tested is $E = X\Gamma$, where X is now $N \times (K+1)$ and Γ is $(K+1) \times 1$ with γ_1 $K \times 1$. Sufficient conditions for this asset pricing relation to hold are given in the Connor (1983) equilibrium extension of Ross's Arbitrage Pricing Theory. $\hat{\Gamma}$, $\hat{X}$, $\hat{\Sigma}$, e and Q are defined as in section 3, with the obvious modifications. In addition,

$$Q^A \equiv Q/\left(1 + \hat{\gamma}_1' \Delta^{-1} \hat{\gamma}_1\right),$$

where Δ is the sample covariance matrix of the δ_t, with T in the denominator.

The conditional distribution of Q_C, given δ and $\hat{\beta}$, is $T^2(N-K-1, T-K-1; \lambda)$ with λ random. Q_C^A is approximately distributed as $T^2(N-K-1, T-K-1)$, which converges to $\chi^2(N-K-1)$, as $T \to \infty$. $N-K-1$ reflects the fact that there are $K+1$ columns in X, while $T-K-1$ is the divisor which makes $\hat{\Sigma}$ an unbiased estimator of Σ.

If the components of δ are portfolio returns, for which the multi-factor asset pricing relation is assumed to hold, then

$$\gamma_0 1_K + \gamma_1 = E(\delta). \tag{A.2}$$

Combining (A.2) with $E = X\Gamma$ and noting from (A.1) that $\alpha = E - \beta E(\delta)$ yields

$$\alpha = \gamma_0 Z \quad \text{where} \quad Z \equiv 1_N - \beta 1_K. \tag{A.3}$$

The economic specification based on the constraint (A.3) no longer includes γ_1 as a separate vector of parameters. This entails a few simple changes in the test procedure.

The cross-sectional residual vector e is now defined as $\hat{\alpha} - \hat{Z}\hat{\gamma}_0$, where $\hat{\alpha}$ is the vector of OLS time series estimates of the α_i. $\hat{Z}$ is obtained by substituting $\hat{\beta}$ for β in Z. $\hat{\gamma}_0$ is a given estimator of γ_0. In particular, the estimator $\hat{\gamma}_{0C}$ is now obtained from the GLS CSR of $\hat{\alpha}$ on $\hat{Z}$ with covariance matrix $\hat{\Sigma}$. Q and Q^A are defined as before, with $\hat{\gamma}_1$ equal to $\bar{\delta} - \hat{\gamma}_0 1_K$.

The conditional distribution of Q_C, given δ and $\hat{\beta}$, is now $T^2(N-1, T-K-1; \lambda)$ and Q_C^A is approximately distributed as $T^2(N-1, T-K-1)$ which converges to $\chi^2(N-1)$, as $T \to \infty$. $N-1$ reflects the fact that Z is the sole independent variable in the CSR. The same covariance estimator $\hat{\Sigma}$ is employed whether (A.2) is incorporated or not. Hence the same degrees of freedom $T-K-1$ in each case. We assume that the components of δ and the N left-hand-side assets constitute a set of $N+K$ linearly independent assets. This ensures that $\hat{\Sigma}$ and Δ are invertible.

Eqs. (6) and (7), which relate LMT and LRT to Q^*, continue to hold for the specifications considered in this appendix. Q^* is now defined as $[T/(T-K-1)]Q$, since the estimator $\hat{\Sigma}$ in Q is the unbiased estimator of Σ. In all cases, the indicated central T^2 approximation for Q_C^A provides an upper bound on the distribution function of the unadjusted statistic Q_C.

Appendix B: Conditional distribution of Q_C given R_m and $\hat{\beta}$

In this appendix, the distribution of Q_C is shown to be a mixture of non-central T^2's, under the null hypothesis that $E = X\Gamma$ for some Γ. Recall that

$$Q_C = Te'\hat{\Sigma}^{-1}e, \tag{B.1}$$

where

$$e \equiv \bar{R} - \hat{X}\hat{\Gamma}, \qquad \hat{\Gamma} \equiv \hat{A}\bar{R},$$

and

$$\hat{A} \equiv (\hat{X}'\hat{\Sigma}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Sigma}^{-1}.$$

Averaging the market model regression equations over time, we have

$$\bar{R} = \alpha + \beta\bar{R}_m + \bar{\varepsilon},$$

where α is the N-vector of intercepts. Noting that $\alpha = E - \beta E_m$,

$$\bar{R} = E + \beta(\bar{R}_m - E_m) + \bar{\varepsilon}.$$

Imposing the null hypothesis $E = X\Gamma$,

$$\bar{R} = X\bar{\Gamma} + \bar{\varepsilon},$$

where

$$\bar{\Gamma} \equiv (\gamma_0, \bar{\gamma}_1) \quad \text{and} \quad \bar{\gamma}_1 \equiv \gamma_1 + \bar{R}_m - E_m.$$

It follows that

$$\bar{R} = \hat{X}\bar{\Gamma} + [\bar{\varepsilon} - \bar{\gamma}_1 U] \quad \text{where} \quad U \equiv \hat{\beta} - \beta. \tag{B.2}$$

Let $\hat{B} \equiv I_N - \hat{X}\hat{A}$. As $\hat{A}\hat{X} = I_2$, $\hat{B}\hat{X} = 0$, hence

$$e = \hat{B}\bar{R} = \hat{B}Y \quad \text{where} \quad Y \equiv [\bar{\varepsilon} - \bar{\gamma}_1 U]. \tag{B.3}$$

Since $\hat{X}$ has rank 2, there exists an invertible $N \times N$ matrix C, which depends on $\hat{\beta}$, such that

$$C\hat{X} = [0 : I_2]',$$

where 0 is a $2 \times (N-2)$ matrix of zeroes. It is straightforward to verify that (i) $\hat{\Gamma}$ in (B.1) equals the estimator obtained from the GLS regression of $C\bar{R}$ on $C\hat{X}$ with covariance matrix $C\hat{\Sigma}C'$, and (ii) Q_C in (B.1) is unaltered when e is replaced by Ce and $\hat{\Sigma}$ by $C\hat{\Sigma}C'$. We may, therefore, assume without loss of generality that

$$\hat{X} = [0 : I_2]'. \tag{B.4}$$

Let $\hat{\Sigma}$ and $\hat{\Sigma}^{-1}$ be partitioned as follows:

$$\hat{\Sigma} = \begin{bmatrix} P & L \\ L' & K \end{bmatrix} \quad \text{and} \quad \hat{\Sigma}^{-1} = \begin{bmatrix} H & F \\ F' & G \end{bmatrix}, \tag{B.5}$$

where P and H are $(N-2)\times(N-2)$, K and G are 2×2, and L and F are $(N-2)\times 2$. Let Y_1 consist of the first $N-2$ components of Y. Using (B.1) and (B.3)–(B.5),

$$\hat{A} = G^{-1}[F' : G] = [G^{-1}F' : I_2],$$

and

$$e' = Y'\hat{B}' = Y_1'[I_{N-2} : -FG^{-1}]. \tag{B.6}$$

By the formula for a partitioned inverse,

$$\hat{\Sigma}^{-1} = \begin{bmatrix} P^{-1} + P^{-1}LGL'P^{-1} & -P^{-1}LG \\ -GL'P^{-1} & G \end{bmatrix}.$$

Using this, it is straightforward (although tedious) to verify that

$$[I_{N-2}: -FG^{-1}]\hat{\Sigma}^{-1}[I_{N-2}: -FG^{-1}]' = P^{-1}. \tag{B.7}$$

It follows from (B.6) and (B.7) that

$$e'\hat{\Sigma}^{-1}e = Y'\hat{B}'\hat{\Sigma}^{-1}\hat{B}Y = Y_1'P^{-1}Y_1. \tag{B.8}$$

All probabilistic statements below are conditional on R_m. $\sqrt{T}\bar{\varepsilon}$ is distributed as $N(0, \Sigma)$ and $(T-2)\hat{\Sigma}$ has a Wishart distribution with parameters $T-2$ and Σ [see Anderson (1958, p. 183)]. Furthermore, $\bar{\varepsilon}$, $\hat{\beta}$ and $\hat{\Sigma}$ are mutually independent [see Shanken (1982, ch. 2, app. B)]. The remaining statements are conditional on $\hat{\beta}$ as well as R_m. $\sqrt{T}Y$ is distributed as $N(-\sqrt{T}\bar{\gamma}_1 U, \Sigma)$. The subvector $\sqrt{T}Y_1$ is distributed as $N(E_1, \Sigma_{11})$, where E_1 consists of the first $N-2$ components of $-\sqrt{T}\bar{\gamma}_1 U$ and Σ_{11} is the principal submatrix of Σ of order $N-2$. The random submatrix P in (B.5), multiplied by $T-2$, has a Wishart distribution with parameters $T-2$ and Σ_{11}, independent of Y_1. Therefore, $TY_1'P^{-1}Y_1$ has a non-central T^2 distribution with non-centrality parameter $E_1'\Sigma_{11}^{-1}E_1$ and degrees of freedom $N-2$ and $T-2$ [see Morrison (1976, p. 131)]. By (B.8), this is also the distribution of Q_C in (B.1). Finally, the algebra used to establish (B.8) implies that

$$\begin{aligned} \lambda = E_1'\Sigma_{11}^{-1}E_1 &= (-\sqrt{T}\bar{\gamma}_1 U)'B'\Sigma^{-1}B(-\sqrt{T}\bar{\gamma}_1 U) \\ &= T\bar{\gamma}_1^2 U'B'\Sigma^{-1}BU, \end{aligned} \tag{B.9}$$

where B is obtained by replacing $\hat{\Sigma}$ by Σ in $\hat{B}$. In this context, $-\sqrt{T}\bar{\gamma}_1 U$ plays the role of Y and E_1 the role of Y_1 in (B.8). Note that the non-centrality parameter λ is a function of the random variable $\hat{\beta}$, but not of $\hat{\Sigma}$.

Appendix C: An approximate distribution for Q_C

In this appendix, both asymptotic and small-sample statistical principles are applied to obtain an approximate distribution for Q_C. The analysis builds upon the results of appendix B. Again, all probabilistic statements are conditional on R_m. Thus $\bar{R}_m$, $\bar{\gamma}_1$ and s_m^2 may be viewed as constants.

The random non-centrality parameter in (B.9) can be written as

$$\lambda = (\bar{\gamma}_1^2/s_m^2)Z'B'\Sigma^{-1}BZ \quad \text{where} \quad Z \equiv \sqrt{T}s_m U.$$

Since $\hat{\beta}$ is normally distributed with mean β and covariance matrix $(Ts_m^2)^{-1}\Sigma$ [see Anderson (1958, p. 183)], Z is distributed as $N(0, \Sigma)$. As $T \to \infty$, B converges in probability to the corresponding matrix with $\hat{\beta}$ replaced by β. Since $B'\Sigma^{-1}B\Sigma$ is idempotent of rank $N-2$, it follows that $Z'B'\Sigma^{-1}BZ$ converges in distribution to $\chi^2(N-2)$ [see Graybill (1961, p. 83)]. A proof of the following lemma may be obtained from the author.

Lemma. If λ is distributed as $c\chi^2(n)$ and Q is conditionally (on λ) distributed as non-central $T^2(n, m; \lambda)$ then Q is unconditionally distributed as $(1+c)T^2(n, m)$; i.e., as a constant multiple of the corresponding central T^2 distribution.

In appendix B, it was demonstrated that the conditional distribution of Q_C, given $\hat{\beta}$ and R_m, is non-central $T^2(N-2, T-2; \lambda)$. The analysis above suggests $(\bar{\gamma}_1^2/s_m^2)\chi^2(N-2)$ as an approximation to the distribution of λ. Utilizing this approximation and applying the lemma above, we obtain $(1+\bar{\gamma}_1^2/s_m^2)T^2(N-2, T-2)$ as an approximate distribution for Q_C. It should be emphasized that the small-sample behavior of $\hat{\Sigma}^{-1}$ is fully incorporated in this approximation since λ is not a function of $\hat{\Sigma}^{-1}$ and our results are exact, conditional on λ. A more standard asymptotic analysis would derive $(1+\gamma_1^2/\sigma_m^2)\chi^2(N-2)$ as the limiting distribution of Q_C, thereby failing to reflect the stochastic behavior of $\hat{\Sigma}^{-1}$. Of course, the suggested approximation for Q_C converges in distribution to $(1+\gamma_1^2/\sigma_m^2)\chi^2(N-2)$, as $T \to \infty$.

References

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

Anderson, T., 1958, An introduction to multivariate statistical analysis (Wiley, New York).

Banz, R., 1981, The relations between returns and market values of common stocks, Journal of Financial Economics 9, 3–18.

Berndt, E. and N. Savin, 1977, Conflict among criteria for testing hypotheses in the multivariate linear regression model, Econometrica 45, 1263–1278.

Connor, G., 1983, A unified beta pricing theory, Journal of Economic Theory, forthcoming.

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

Fama, E., 1976, Foundations of finance (Basic Books, New York).

Fama, E. and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, Journal of Political Economy 91, 607–636.

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

Gibbons, M., S. Ross and J. Shanken, 1984, Testing portfolio efficiency in the presence of a riskless asset, Working paper (Graduate School of Business, University of California, Berkeley, CA).

Graybill, F., 1961, An introduction to linear statistical models, Vol. 1 (McGraw-Hill, New York).

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

Johnson, N. and S. Kotz, 1970, Continuous univariate distributions – 2 (Houghton-Mifflin, Boston, MA).

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

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

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

MacKinlay, C., 1984, An analysis of multivariate financial tests, Working paper (Graduate School of Business, University of Chicago, Chicago, IL).

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

Press, J., 1972, Applied multivariate analysis (Holt, Rinehart and Winston, New York).

Roll, R., 1977, A critique of the asset pricing theory's tests, Part I, Journal of Financial Economics 4, 129–176.

Roll, R., 1979, A reply to Mayers and Rice, Journal of Financial Economics 7, 391–399.

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

Ross, S., 1977, The capital asset pricing model, short sales restrictions and related issues, Journal of Finance 32, 177–183.

Shanken, J., 1980, Some asymptotic properties of statistical methods in finance, Unpublished manuscript.

Shanken, J., 1982, An analysis of the traditional risk-return model, Ph.D. dissertation (Carnegie-Mellon University, Pittsburgh, PA).

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.

Theil, H., 1971, Principles of econometrics (Wiley, New York).