THE JOURNAL OF FINANCE • VOL. XLVI, NO. 2 • JUNE 1991

Using Generalized Method of Moments to Test Mean-Variance Efficiency

A. CRAIG MACKINLAY and
MATTHEW P. RICHARDSON*

ABSTRACT

This paper develops tests of unconditional mean-variance efficiency under weak distributional assumptions using a Generalized Method of Moments framework. These tests are potentially more robust than commonly employed tests which rely on the assumption that asset returns are normally distributed and temporarily i.i.d. Using returns for size-based portfolios from 1926 to 1988 we show that the conclusion concerning the mean-variance efficiency of market indexes can be sensitive to the test considered.

THE APPLICATION OF MULTIVARIATE statistical techniques in financial economics has become common practice in recent years. Much of the development has been in the area of testing asset pricing models. The assumption that asset returns follow a time invariant multivariate normal distribution permits tests of restrictions on model parameters in a pooled time series–cross-section framework. These tests can be interpreted as tests of the mean-variance efficiency of a portfolio or combination of portfolios. Gibbons (1982), Jobson and Korkie (1982), Stambaugh (1982), Shanken (1985, 1987), Kandel and Stambaugh (1987), MacKinlay (1987), and Gibbons, Ross, and Shanken (1989) are examples of work presenting multivariate tests of the capital asset pricing model.[1] Connor and Korajczyk (1988) and Lehmann and Modest (1988) present multivariate tests of the arbitrage pricing model; Breeden, Gibbons, and Litzenberger (1989) present multivariate tests of the consumption based CAPM. Although these papers assume the normality of asset returns, little analysis has been conducted to understand the sensitivity of the inferences to violations of this assumption.[2,3]

* The Wharton School of The University of Pennsylvania, Philadelphia, Pennsylvania 19104-6367. We thank Michael Gibbons, Andrew Lo, Krishna Ramaswamy, Tom Smith, René Stulz, and a referee for helpful comments. Research support received from the Geewax-Terker Research Fund and the University of Pennsylvania Research Foundation is gratefully acknowledged. The first author also acknowledges support from the Batterymarch Fellowship and the National Science Foundation (SES-8821583).

[1] Gibbons, Ross, and Shanken (1989) present their results in the context of testing the efficiency of a given portfolio.

[2] Actually, from a statistical perspective, the tests require that asset returns conditional on the factor portfolios be i.i.d. and multivariate normal.

[3] MacKinlay (1985) and Affleck-Graves and McDonald (1989) do present simulation evidence for the robustness of these tests to deviations from normality. The results indicate the tests are quite robust to deviations. However, these studies focus on 5-year test periods with twenty and more portfolios.

This paper proposes multivariate tests which are robust to departures from normality. Such tests are of interest for two reasons. First, the normality assumption is, in general, not necessary from a theoretical perspective to derive the models. Rather, the normality assumption is adopted for statistical convenience. Without this assumption finite sample properties of asset pricing model tests are difficult to derive. Second, nonnormality of security returns on a monthly basis (more so on a weekly basis) has been documented. Fama (1965, 1976), Blattberg and Gonedes (1974), and Hsu (1982) contain evidence concerning the nonnormality of returns. In addition, numerous studies have presented evidence of heteroskedasticity in returns.[4]

We employ a Generalized Method of Moments (GMM) framework to develop tests of mean-variance efficiency. The GMM based tests are valid under much weaker distributional assumptions than most previous tests. Using a GMM based test, we find that misspecified distributional assumptions can have a quantifiably adverse effect on statistical inference.

The paper is organized as follows. Section I presents a statement of the mean-variance efficiency of a portfolio and reviews the commonly used test statistic. In Section II, we provide GMM based tests of whether a portfolio p is mean-variance efficient. Most importantly, we employ only minimal restrictions beyond those implied by mean-variance efficiency. In order to compare the various approaches, Section III investigates the effect on the test statistics of contemporaneous conditional heteroskedasticity. In Section IV, we apply the analysis of the previous sections to actual data. It is shown that this analysis is economically meaningful. Section V concludes the paper.

I. Tests of Mean-Variance Efficiency

In this section, we review mean-variance efficiency and the commonly employed test statistic. If a given portfolio p is mean-variance efficient (when a risk-free asset exists) then:

$$E[\tilde{r}_{it}] = \beta_i E[\tilde{r}_{pt}] \quad \forall i \tag{1}$$

where

$$\tilde{r}_{it} = \text{excess return on asset } i \text{ time period } t;$$

$$\beta_i = \frac{\text{cov}(\tilde{r}_{it}, \tilde{r}_{pt})}{\text{var}(\tilde{r}_{pt})}; \tag{2}$$

$$\tilde{r}_{pt} = \text{excess return on portfolio } p \text{ time period } t.$$

[4] Schwert and Seguin (1989) is a recent example of work providing evidence of heteroskedasticity in stock returns, and Barone-Adesi and Talwar (1983) and Diebold, Im, and Lee (1989) provide evidence of heteroskedasticity in market model residuals.

For N assets, these conditions can be expressed as restrictions on a system of excess return market model regression equations:

$$\begin{aligned} \tilde{r}_{it} &= \alpha_i + \beta_i \tilde{r}_{pt} + \tilde{\epsilon}_{it}, \quad i = 1, \ldots, N, \\ E[\tilde{\epsilon}_{it}] &= 0 \\ E[\tilde{\epsilon}_{it}\tilde{r}_{pt}] &= 0 \\ \alpha_i &= 0 \end{aligned} \tag{3}$$

where $\tilde{\epsilon}_{it}$ is the disturbance term for asset i in period t.

For the sample estimators the first two restrictions, $E[\tilde{\epsilon}_{it}] = E[\tilde{\epsilon}_{it}\tilde{r}_{pt}] = 0$, will always be true for all i from least squares projection theory. That is, ordinary least squares (OLS) implies that the β_i's given by equation (2) always satisfy these restrictions in the system of equations in (3). The third condition, $\alpha_i = 0$ for all i, however, imposes testable restrictions on the data and is true if portfolio p is mean-variance efficient.

For the subsequent analysis, it is convenient to introduce further notation. Define α as the $(N \times 1)$ vector $(\alpha_1, \alpha_2, \ldots, \alpha_N)'$, β as the $(N \times 1)$ vector $(\beta_1, \beta_2, \ldots, \beta_N)'$, and δ as the $(2N \times 1)$ vector $(\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_N, \beta_N)'$. The $(N \times N)$ unconditional disturbance covariance matrix will be denoted by Σ and is assumed to be positive definite.[5]

Given the i.i.d. assumption for excess returns and multivariate normality, most papers have used the Wald statistic to test the mean-variance efficiency of a portfolio. The Wald test statistic for the efficiency of portfolio p is:

$$\phi_0 = T\hat{\alpha}'\left[\left(1 + \frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}\right)\hat{\Sigma}\right]^{-1}\hat{\alpha} \tag{4}$$

where $\hat{\mu}_p$ is the sample mean excess return of portfolio p, $\hat{\sigma}_p^2$ is the maximum likelihood estimator of the variance of portfolio p's excess return, $\hat{\Sigma}$ is the maximum likelihood estimator of Σ, and T is the number of time series observations. Asymptotically, under the null hypothesis, ϕ_0 will have a chi-square distribution with N degrees of freedom. In finite samples, $[(T - N - 1)/(NT)]\phi_0$ will have a F distribution with N degrees of freedom in the numerator and $T - N - 1$ degrees of freedom in the denominator.[6]

II. GMM Test

In this section, we present a GMM-based test of whether a portfolio is unconditionally mean-variance efficient. For the most part, we employ only restrictions implied by mean-variance portfolio theory. Unlike many other authors, we do not make strong assumptions concerning the disturbance term

[5] This assumption requires that r_{pt} not be a linear combination of the left-hand side portfolio returns.

[6] See, for example, Gibbons, Ross, and Shanken (1989) for derivations of the Wald and F statistics.

in a market model regression.[7] With the GMM framework, the disturbance term can be both serially dependent and conditionally heteroskedastic. We assume that there exists a riskless rate of interest R_{ft} for each t and that excess asset returns (over R_{ft}) are stationary and ergodic with finite fourth moments.[8,9]

Consider a sample of T time series observations, $t = 1, \ldots, T$ and of N assets, $i = 1, \ldots, N$. Define $(2N \times 1)$ vectors $f_t(\delta)$ and $g_T(\delta)$ as follows:

$$f_t(\delta) = \begin{bmatrix} \tilde{\epsilon}_{1t}(\alpha_1, \beta_1) \\ \tilde{\epsilon}_{1t}(\alpha_1, \beta_1)\tilde{r}_{pt} \\ \vdots \\ \tilde{\epsilon}_{it}(\alpha_i, \beta_i) \\ \tilde{\epsilon}_{it}(\alpha_i, \beta_i)\tilde{r}_{pt} \\ \vdots \\ \tilde{\epsilon}_{Nt}(\alpha_N, \beta_N) \\ \tilde{\epsilon}_{Nt}(\alpha_N, \beta_N)\tilde{r}_{pt} \end{bmatrix} \tag{5}$$

$$g_T(\delta) = \frac{1}{T}\sum_{t=1}^{T} f_t(\delta). \tag{6}$$

From the excess return market model we have the moment condition $E[f_t(\delta)] = 0$. We exploit this moment condition for estimation and testing using a GMM approach. The GMM approach involves selecting an estimator to set linear combinations of the moment condition to zero. Specifically, for some matrix A with row dimension equal to the number of parameters and column dimension equal to the length of $g_T(\cdot)$, $\hat{\delta}$ is chosen so that $Ag_T(\hat{\delta}) = 0$. Hansen (1982) develops the sampling theory for such an estimator. Within the class of estimators that set linear combinations of $g_T(\cdot)$ equal to zero, Hansen (1982) shows that the optimal GMM weighting matrix is given by A^* where:[10]

$$A^* = D_0' S_0^{-1}, \tag{7}$$

$$D_0 = E\left[\frac{\partial g_T(\delta)}{\partial \delta'}\right], \tag{8}$$

$$S_0 = \sum_{l=-\infty}^{+\infty} E[f_t(\delta) f_{t-l}(\delta)']. \tag{9}$$

[7] An exception in Shanken (1990) who employs the White (1980) covariance matrix estimator which does not require the stronger distributional assumptions. In the absence of serial correlation, this covariance matrix estimator will be identical to the estimator from the exactly identified GMM case in Section II.A.

[8] If a time series is ergodic, then the time average over a period of T observations converges in mean square to the corresponding ensemble average as T increases to ∞.

[9] See Hansen (1982) for technical details concerning the required assumptions.

[10] The weighting matrix is optimal in the sense that the difference between the covariance matrix of the parameter estimators using any other weighting matrix and the covariance matrix using A^* will be a positive semidefinite matrix. The estimator will not depend on the weighing matrix in the case of exact identification since all elements of $g_T(\cdot)$ can be set to 0.

Using a GMM framework, there are two ways to test whether portfolio p is mean-variance efficient; that is, whether $\alpha_i = 0$, $i = 1, \ldots, N$. One approach is to estimate the unrestricted system and then test the hypothesis $\alpha = 0$ using the unrestricted estimates. The second approach is to substitute the restrictions $\alpha = 0$ into equation (5), estimate the restricted system, and then test the overidentifying restrictions. Next we briefly outline these two approaches.

A. Unrestricted Case

Consider the sample moment vector given in equation (6) for N assets, $i = 1, \ldots, N$. For each asset, there are two sample moments, $[1/T\Sigma_{t=1}^{T}\tilde{\epsilon}_{it}(\alpha_i, \beta_i)]$ and $[1/T\Sigma_{t=1}^{T}\tilde{\epsilon}_{it}(\alpha_i, \beta_i)\tilde{r}_{pt}]$ and two parameters, (α_i, β_i). There are, therefore, $2N$ equations and $2N$ unknown parameters, the system is exactly identified, and the parameters can be chosen to set the sample moments equal to zero. Since setting the sample moments equal to zero is equivalent to deriving the normal equations from OLS, this GMM procedure is equivalent to OLS regression for each i.

Hansen (1982) shows that the GMM parameter estimator $\hat{\delta}$ will have an asymptotic normal distribution with mean δ and asymptotic variance-covariance matrix $[D_0' S_0^{-1} D_0]^{-1}$. In practice, D_0 and S_0 will be unknown; however, the asymptotic results are still valid for consistent estimators of D_0 and S_0, which we denote D_T and S_T. An assumption with respect to S_0 is necessary to reduce the summation to a finite number of terms and permit construction of a consistent estimator.[11] The test statistic can be constructed employing the usual framework for testing linear restrictions. Let ϕ_1 be the test statistic. Then under the null hypothesis we have:

$$\phi_1 = T\hat{\alpha}'\Big[R\big[D_T' S_T^{-1} D_T\big]^{-1} R'\Big]^{-1}\hat{\alpha} \overset{a}{\sim} \chi_N^2 \tag{10}$$

where $R = I_N \otimes (1\ 0)$ and $R\hat{\delta} = \hat{\alpha}$.

B. Restricted Case

Consider substituting the mean-variance portfolio restrictions $\alpha_i = 0$ for all i into $f_t(\delta)$ in equation (5). Then, for this restricted case we have:

$$g_T(\alpha = 0, \beta) = \frac{1}{T}\sum_{t=1}^{T} f_t(\alpha = 0, \beta). \tag{11}$$

For each asset, there are two sample moments but only one parameter β_i to be estimated. There are, therefore, $2N$ equations and N unknown parameters, the system is overidentified, and all the sample moments cannot be set equal to zero. It is possible, however, to set the optimal linear combination of

[11] See Hansen and Singleton (1982), Eichenbaum, Hansen, and Singleton (1988), Newey and West (1987), and Richardson and Smith (1990) for a discussion of various estimation procedures for D_T and S_T.

the moments $A^*g_T(\beta)$ equal to zero. Under the null hypothesis, the parameter estimators from this procedure $\hat{\beta}$ will asymptotically have a normal distribution with mean β and variance-covariance matrix $[D_0' S_0^{-1} D_0]^{-1}$ (where $D_0 = E[\frac{\partial g_T(\alpha = 0, \beta)}{\partial \beta'}]$). Using Hansen's (1982) results, a test of the N overidentifying restrictions of the model ($\alpha = 0$) can then be performed using:

$$\phi_2 = T g_T(\hat{\beta})' S_T^{-1} g_T(\hat{\beta}) \overset{a}{\sim} \chi_N^2. \tag{12}$$

The required assumptions are the stationarity and ergodicity of $\tilde{r}_{it}$, $i = 1, \ldots, N$, and $\tilde{r}_{pt}$ and the existence of fourth moments for excess asset returns. The mean-variance efficient model imposes restrictions which can be tested directly. This is in contrast to previous tests of mean-variance efficiency which have placed strong distributional assumptions of the $\tilde{\epsilon}_{it}$ in regression equation (3).

In the next section, we look at one particular test of mean-variance efficiency in some detail. Because many existing tests are special cases of this test, the conditions under which these existing tests of mean-variance efficiency are appropriate can be determined.

III. Conditional Heteroskedasticity

We consider the case of contemporaneous conditional heteroskedasticity to illustrate the difference between some previous tests of mean-variance efficiency and a GMM based test. With contemporaneous conditional heteroskedasticity, the variances of the market model residuals of equation (3) will be dependent upon the contemporaneous portfolio return. We assume that $E[f_t(\delta) f_{t-l}'(\delta)] = 0$ for all $l \neq 0$. This assumption implies that there is no serial dependence in either $\tilde{\epsilon}_{it}$ or $\tilde{\epsilon}_{it}\tilde{r}_{pt}$.[12] Consider the unrestricted model of Section II.A. Applying the GMM procedure to the moment condition in equation (4) gives us the OLS estimators of (α_i, β_i) for each i. Then, given the covariance matrix of $\hat{\alpha}$, we can construct a test similar to ones that appear in Gibbons, Ross, and Shanken (1989), Breeden, Gibbons, and Litzenberger (1989), MacKinlay (1987), and Jobson and Korkie (1982). The difference with the GMM based test is the calculation of the asymptotic variance of $\hat{\alpha}$, $\text{var}(\hat{\alpha})$.

In Appendix A, for this case, the expression for the asymptotic covariance matrix of the GMM $\hat{\alpha}$ estimator is derived:

$$\text{var}(\hat{\alpha}) \equiv \Omega = \left(1 + \frac{\mu_p^2}{\sigma_p^2}\right)\Sigma + \frac{1}{\sigma_p^2}\Psi \tag{13}$$

[12] The results can be generalized to allow for serial correlation using, for example, the technique of Newey and West (1987).

where

$\Sigma = N \times N$ covariance matrix of the disturbance vector $\tilde{\epsilon}_t$ with (i, j)th element $[\sigma_{ij}]$,

$\Psi = N \times N$ "correction factor" matrix for contemporaneous conditional heteroskedasticity, with (i, j)th element $\left[\frac{\mu_p^2}{\sigma_p^2}\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)} - 2\mu_p\left(1 + \frac{\mu_p^2}{\sigma_p^2}\right)\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt})}\right]$ where $\sigma(\cdot, \cdot)$ denotes the covariance operator.

Given Ω, we can test the hypothesis $\alpha = 0$ with the test statistic $T\hat{\alpha}'\Omega^{-1}\hat{\alpha}$. Since asymptotically $\sqrt{T}(\hat{\alpha} - \alpha)$ is distributed $N(0, \Omega)$, this test statistic will have a chi-square distribution with N degrees of freedom under the null hypothesis. In applications, Ω can be replaced with a consistent estimator $\hat{\Omega}$ without altering the asymptotic distribution of the test statistic.

Since the estimator of α is the same for the GMM procedure and the procedure based on the normality assumption, the differences between the tests relate to the appropriate asymptotic covariance matrix for $\hat{\alpha}$. The first term in equation (13), $(1 + \frac{\mu_p^2}{\sigma_p^2})\Sigma$, has been given as the asymptotic covariance matrix of the intercepts $\hat{\alpha}$ in most empirical applications of mean-variance efficiency and is used in constructing ϕ_0.[13] The multivariate normality assumption implies the $\tilde{\epsilon}_{it}$'s are contemporaneously conditionally homoskedastistic, which is a sufficient condition to make the second term of (13) zero. It is in this sense that previous tests are special cases. That is, by including additional restrictions such as $\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt})} = 0$ and $\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)} = 0$ for all i and j in the null hypothesis of mean-variance efficiency, GMM test becomes asymptotically equivalent to previous tests.[14] These additional assumptions, however, are not implied by mean-variance theory and may not be consistent with actual equilibrium returns.

A. A Specific Example: Multivariate Student t

The use of the multivariate Student t as a return distribution can be motivated both empirically and theoretically. One empirical stylized fact from the distribution of returns literature is that returns have fatter tails and are more peaked than one would expect from a normal distribution. This is consistent with returns coming from a multivariate Student t.[15] Further,

[13] See Gibbons, Ross, and Shanken (1989) or MacKinlay (1987).

[14] However, the cost of the more general framework is the loss of small sample results for the test statistic distribution. See Gibbons, Ross, and Shanken (1989) and MacKinlay (1987) for discussion of the small sample results.

[15] Blattberg and Gonedes (1974) suggest the Student t as a distribution for asset returns. The kurtosis of a multivariate Student t distribution with ν degrees of freedom is $3\frac{(\nu - 2)}{(\nu - 4)}$ which exists for $\nu > 4$ and will be greater than the kurtosis of a normal distribution.

the multivariate Student t is a return distribution for which mean-variance analysis is consistent with expected utility maximization, making the choice theoretically appealling.[16]

In the multivariate Student t case, the regression equations are the same as for the multivariate normal case except that the conditional variance of the error term $\tilde{\epsilon}_{it}$ is no longer independent of $\tilde{r}_{pt}$ (see Zellner (1971) p. 388). With some manipulation to the formula given in Zellner, it can be shown that

$$\text{var}\left(\tilde{\epsilon}_t \mid \tilde{r}_{pt}\right) = \left[\frac{\nu - 2}{\nu - 1}\right]\left[1 + \frac{\left(\tilde{r}_{pt} - \mu_p\right)^2}{(\nu - 2)\sigma_p^2}\right]\Sigma \tag{14}$$

where $\epsilon_t = (\epsilon_{1t}\,\epsilon_{2t} \dots \epsilon_{Nt})'$ and ν is the degrees of freedom of the Student t. Using this formula, it is possible to calculate $\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt})}$ and $\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)}$ explicitly. These calculations result in the following expressions:

$$\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt})} = \mu_p\left[\frac{\nu - 2}{\nu - 1} + \frac{1}{\nu - 1} - 1\right]\sigma_{ij} = 0, \tag{15}$$

$$\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)} = \frac{1}{\nu - 1}\left[\frac{E\left(\tilde{r}_{pt} - \mu_p\right)^4}{\sigma_p^2} - \sigma_p^2\right]\sigma_{ij}. \tag{16}$$

By Student t distribution properties, $E(\tilde{r}_{pt} - \mu_p)^4 = 3\sigma_p^4(\frac{\nu - 2}{\nu - 4})$, and we can rewrite the second covariance as

$$\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)} = \frac{2\sigma_p^2}{(\nu - 4)}\sigma_{ij}. \tag{17}$$

Substituting these values into equation (13) for Ω gives:

$$\Omega = \left[1 + \frac{(\nu - 2)}{(\nu - 4)}\frac{\mu_p^2}{\sigma_p^2}\right]\Sigma. \tag{18}$$

Using (18) and the sample estimates of $\frac{\mu_p^2}{\sigma_p^2}$ and Σ for the test statistic we have:

$$\phi_3 = T\hat{\alpha}\hat{\Omega}^{-1}\hat{\alpha} = T\hat{\alpha}\left[\left(1 + \frac{(\nu - 2)}{(\nu - 4)}\frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}\right)\hat{\Sigma}\right]^{-1}\hat{\alpha}. \tag{19}$$

[16] See Ingersoll (1987) p. 104 for a discussion of this result.

Comparing (19) with equation (4) we have:

$$\left[\frac{1+\dfrac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}}{1+\dfrac{(\nu-2)}{(\nu-4)}\dfrac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}}\right]\phi_0. \tag{20}$$

Hence, with the multivariate Student t distribution, we can examine a case where the bias of the common test statistic can be quantified. An appropriate test statistic in an environment where returns are multivariate Student t can be expressed as the usual statistic scaled by a factor which accounts for the conditional heteroskedasticity. The magnitude of the impact of the contemporaneous conditional heteroskedasticity will depend on the degrees of freedom of the Student t and the sample Sharpe ratio of the portfolio being tested. As the degrees of freedom become large, the Student t is well approximated by the normal distribution, and the conditional heteroskedasticity will be reduced. Asymptotically, the conditional heteroskedasticity will vanish.

As can be seen from equation (20), the bias calculation involves the population parameters μ_p and σ_p. Since these parameters are unknown, we assume that the sample estimates represent the population values in order to get reasonable measures of the magnitude of the bias. Specifically, we take sample estimates of the mean and variance on the CRSP equally weighted index for 5-year intervals between January 1954 and December 1983 (see MacKinlay (1987) p. 347). Since there is variation in the sample estimates of the Sharpe measure across subperiods, the minimum and maximum of these will tend to be downward and upward biased estimates of the population Sharpe measure, respectively. The magnitude of the bias in the tests presented, therefore, should be interpreted cautiously.

For the analysis, we consider two values for the degrees of freedom of the Student t, $\nu = 5$ and $\nu = 10$. These choices are motivated by the empirical estimates of Blattberg and Gonedes (1974). Three values for the number of assets are considered, $N = 1$, $N = 5$, and $N = 10$. Table I reports the results. The misspecification can be substantial. For example, when ν is 5 for the January 1954 to December 1958 period, ϕ_0 overstates ϕ_3 by as much 35%. A test using ϕ_0 that would have a size of 5% if returns are i.i.d. multivariate normal will have a size of 19.2% for Student t returns with $N = 10$. With one lefthand side portfolio ($N = 1$) the actual size would be 9.1%. The error in the size of the test will increase as N increases and will decrease as the degrees of freedom of the Student t increase. The dependence of the bias explicitly upon μ_p^2/σ_p^2 is evident in this Table. In contrast to the 1954 to 1958 time period, the January 1974 to December 1978 sample value of μ_p^2/σ_p^2 is small and the bias is minimal.

The above example provides an analytical illustration of the potential for incorrect inferences when returns are assumed to be i.i.d. and multivariate

Table I

Misspecification of Tests of Mean-Variance Efficiency

Misspecification for tests of mean-variance efficiency which assume the distribution of excess asset returns is i.i.d. multivariate normal and the distribution is actually i.i.d. Student t with ν degrees of freedom. Values of the Sharpe measure squared $(\hat{\mu}_p^2/\hat{\sigma}_p^2)$ are reported for the CRSP equally weighted index. The bias is the ratio of the Wald test statistic to the Student t test statistic minus 1. The size of the test is reported for 1, 5, and 10 portfolios. This size is the actual size of a test conducted using the 5% critical value for the Wald test with multivariate normality.

Time period	ν	$\frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}$	Bias	size-0.05 $N = 1$	size-0.05 $N = 5$	size-0.05 N = 10
01/54–12/58	5	0.21	34.7%	0.091	0.145	0.192
	10		5.8%	0.057	0.063	0.068
01/59–12/63	5	0.021	4.1%	0.055	0.059	0.062
	10		0.7%	0.051	0.052	0.052
01/64–12/68	5	0.14	24.6%	0.079	0.114	0.144
	10		4.1%	0.055	0.059	0.062
01/69–12/73	5	0.026	5.1%	0.056	0.061	0.066
	10		0.8%	0.051	0.052	0.052
01/74–12/78	5	0.013	2.6%	0.053	0.056	0.058
	10		0.4%	0.050	0.051	0.051
01/79–12/83	5	0.063	11.9%	0.064	0.078	0.090
	10		2.0%	0.052	0.054	0.056

normal, yet are actually multivariate Student t. A solution for this misspecification is to use the GMM statistic which asymptotically has a chi-square distribution under either assumption. In the next section, we explore the differences between these statistics in the actual data.

IV. Empirical Results

In this section, we compare different test statistics of whether a given portfolio is mean-variance efficient. Specifically, we compare inferences using the Wald and F statistics' which are based on ϕ_0 with the GMM statistic ϕ_1 for the exactly identified case. The tests are applied to the CRSP value-weighted and equally weighted portfolios. The F test is included to quantify the extent to which the use of large sample theory is an issue. Similar inferences using the F statistic and the Wald statistic indicate the use of asymptotic theory is appropriate since the distribution for the F statistic is a finite sample result. Throughout Tables II and III the Wald statistic and the GMM statistic are scaled by $(T - N - 1)/T$. This adjustment will not affect the asymptotic properties and is intended to improve the finite sample behavior of the Wald test. Without this adjustment, the Wald test will reject too often in finite samples.[17]

[17] See Gibbons, Ross, and Shanken (1989).

Table II

Empirical Tests of Mean-Variance Efficiency of CRSP Indices

Empirical tests of mean-variance efficiency of CRSP value-weighted and equally weighted indices using 10 portfolios constructed and weighted on the basis of market value of equity. Results are presented for monthly data from 01/26 to 12/88 and for two subperiods of equal length. Three test statistics are reported: the F-test statistic, the Wald statistic, and the GMM statistic. The p-values are below the statistics in parentheses.

Portfolio p	Time Period	$\frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2}$	F^a	Wald[b]	GMM[b]	$\frac{\text{Wald}}{\text{GMM}} - 1$
Value-Weighted	01/26–12/88	0.0127	1.899 (0.042)	18.99 (0.040)	21.16 (0.020)	−10.2%
	01/26–06/57	0.0165	1.750 (0.068)	17.50 (0.064)	20.45 (0.025)	−14.4%
	07/57–12/88	0.0092	1.729 (0.073)	17.29 (0.068)	17.71 (0.060)	−2.4%
Equally Weighted	01/26–12/88	0.0169	1.620 (0.096)	16.20 (0.094)	19.66 (0.033)	−17.6%
	01/26–06/57	0.0192	2.839 (0.002)	28.39 (0.002)	39.21 (0.00002)	−27.6%
	07/57–12/88	0.0167	1.325 (0.215)	13.25 (0.210)	14.01 (0.173)	−5.4%

[a] Under the null hypothesis distributed $F_{10,745}$.
[b] Under the null hypothesis asymptotically distributed χ^2_{10}.

We use the ϕ_1 statistic defined in equation (10) for the GMM test. The statistic is calculated using

$$D_T = \frac{1}{T}\sum_{t=1}^{T}\left[I_N \otimes x_t x_t'\right], \tag{21}$$

$$S_T = \frac{1}{T}\sum_{t=1}^{T}\left[\hat{\epsilon}_t\hat{\epsilon}_t' \otimes x_t x_t'\right], \tag{22}$$

where $x_t = (1\ r_{pt})'$ and the $(N \times 1)$ disturbance vectors $(\hat{\epsilon}_t)$ are calculated using the OLS parameter estimates.

The tests are performed using monthly observations from January 1926 to December 1988 on ten value-weighted portfolios. NYSE and AMEX firms are allocated to portfolios based on their beginning of year market value of equity. The results are presented in Table II. For the overall time period, the Wald statistic understates the GMM statistic by 10.2% and 17.6% for the value-weighted and equally weighted indices, respectively. This can have important consequences for statistical inference at conventional levels of significance. For example, the Wald and F statistics cannot reject the mean-variance efficiency of the equal-weighted index at the 5% level, whereas the GMM statistic with a p-value of 0.033 does reject. A similar pattern holds for the results for the value-weighted index and for subperiods. The

Table III

Simulation Evidence for Distributions of the Test Statistics under the Null Hypothesis that the CRSP Indices are Mean-Variance Efficient

Panel A contains results when excess returns are simulated to be i.i.d. multivariate normal. The number of time series observations is 756. The parameters required for the simulation are specified using the parameter estimates from excess return market model with the specified CRSP index and the ten size sorted portfolios. Panel B contains results generated by bootstrapping the excess return data for the 756 months from January 1926 to December 1988. For each time series observation, the vector used in the resampling procedure contains the specified index excess return and the 10 size sorted portfolio fitted residuals from the excess return market model. For each replication excess returns are constructed using the resampled market index excess return and the residuals with α equal to 0 with β equal to the sample estimate. The results for each panel are based on 10,000 replications. Columns two through six summarize the test statistics' distributions, and columns seven through ten present the size of test using the χ^2_{10} critical values

Panel A. Returns Distributed Multivariate Normal									
Statistic	Mean	Variance	Median	90th%-tile	95th%-tile	99th%-tile	size = 0.10	size = 0.05	size = 0.01
χ^2_{10}	10.00	20.00	9.34	15.99	18.31	23.21	0.100	0.050	0.010
Value-Weighted:									
Wald	10.01	20.23	9.36	15.89	18.16	23.04	0.098	0.048	0.009
GMM	10.02	20.26	9.36	15.90	18.18	23.07	0.097	0.048	0.010
Equally Weighted:									
Wald	10.02	20.43	9.30	15.96	18.46	23.57	0.100	0.052	0.012
GMM	10.03	20.47	9.31	15.97	18.45	23.58	0.100	0.053	0.011

Panel B. Returns Drawn from Multivariate Empirical Distribution									
Statistic	Mean	Variance	Median	90th%-tile	95th%-tile	99th%-tile	size = 0.10	size = 0.05	size = 0.01
Value-Weighted:									
Wald	9.40	18.41	8.73	15.17	17.46	22.02	0.078	0.039	0.007
GMM	10.17	20.89	9.45	16.33	18.74	23.65	0.111	0.058	0.012
Equally Weighted:									
Wald	8.89	16.92	8.24	14.42	16.70	21.13	0.061	0.029	0.005
GMM	10.02	20.66	9.31	16.14	18.48	23.71	0.104	0.053	0.012

GMM statistic is consistently higher than the Wald statistic, providing stronger evidence against the mean-variance efficiency of the indices. The difference between the Wald test and the GMM test is more pronounced for the January 1926 to June 1957 time period than for the July 1957 to December 1988 period. This observation is consistent with the i.i.d. multivariate normality assumption for excess asset returns being a better approximation in the later time period.

Having established that inferences can be sensitive to the test statistic considered, we now address the source of the difference. Is this difference due to the inappropriateness of the Wald statistic or due to perverse small sample behavior of the GMM statistic? Table III reports small sample behavior of the Wald and GMM statistics under two different simulation scenarios. In all the simulations the data is created so that portfolio p is mean-variance efficient. Additionally, the parameters of the return generating process are selected to match those of the actual data, that is, the 10 size portfolios and the specified index. Each simulation consists of 10,000 replications. For a given index, the Wald statistic and the GMM statistic are computed using the same simulated returns.

For the first experiment the 10 portfolios returns and mean-variance efficient candidate portfolio return are simulated from a multivariate normal distribution using a random number generator. This distributional assumption matches the most common assumption made in the literature for statistical inference. The results for this distributional assumption are given in Table III Panel A. Both the Wald and GMM statistics are well-behaved for a sample of 756 time series observations and 10 portfolios. For example, their means are 10.01 and 10.02 using the value-weighted index parameters versus a theoretical mean of 10.00; and at the 95th percentile the statistics' simulated values are 18.16 and 18.18 versus 18.31 theoretically. More important though, the small sample behavior of the Wald and GMM statistics is almost identical. Therefore, the difference in the actual results does not seem due to "perverse" small sample behavior on the part of the GMM statistic.

In order to empirically assess the possible misspecification of the Wald statistic, a bootstrapping experiment is performed. In this simulation, the returns are constructed by drawing with replacement the disturbance vector $\tilde{\epsilon}_t$ and the efficient portfolio return $\tilde{r}_{pt}$ from the empirical distribution of the 10 size portfolios' market model residuals and the specified index.[18] The simulation relies on the temporal i.i.d. assumption but maintains any contemporaneous conditional heteroskedasticity and cross-correlation present in the actual data. The results are reported in Table III Panel B. The Wald statistic's empirical critical values understate the theoretical critical values, and the empirical sizes of the test are too low. For example, using the equally weighted index, the Wald statistic's mean drops from 10.02 to 8.89 and its

[18] See equation (3) for the form of the market model. In addition, note that the β_i were chosen to match those of the actual data while the α_i were all set equal to zero.

95th percentile falls from 18.46 to 16.70. In contrast, the GMM statistic's values increase only slightly.

The difference in the actual results in Table II seems, therefore, due to the stronger distributional assumptions required for the Wald statistic. For example, using the bootstrap distributions from the simulations used to create Table III Panel B, the p-value of the Wald statistic for the efficiency of the equally weighted index decreases from 0.094 to 0.057, whereas the GMM statistic p-value changes from 0.033 to 0.035. Much of the difference between the Wald and GMM statistic can be explained from the simulation evidence. However, the lower p-values for the GMM statistic in Table II may also reflect differences in the properties of the statistics under alternative hypotheses.

It is interesting to note that the biases introduced by violations of the i.i.d. and multivariate normality assumptions can be in either direction. In the example of Section III where excess returns are distributed multivariate Student t, the conventional test will reject the null hypothesis too often. In contrast, in the empirical example of this section, the conventional test appears to accept the null hypothesis too often.

V. Conclusion

The finance literature contains many different test statistics for whether a given portfolio is mean-variance efficient. Most of the multivariate tests depend on the market model disturbance vector being conditionally homoskedastic and i.i.d.. Under theoretically consistent distributional assumptions, this need not be the case. For example, in the presence of contemporaneous conditional heteroskedasticity, there can be large enough divergences from the true asymptotic variance to have substantial effects on the specification of these multivariate statistics.

We illustrate that, using theoretical large sample critical values, conclusions regarding the mean-variance efficiency of a portfolio can be materially affected. In contrast, a more robust test statistic largely eliminates the above problems. This statistic captures as special cases some of the more widely used tests in the literature. Furthermore, to the extent that mean-variance efficient testing is simply an application of linear regression techniques, this paper has applications elsewhere in the literature.

Appendix A

Using Hansen (1982) and the serial independence assumption, it is possible to derive the expression for the asymptotic variance-covariance matrix of the estimators:

$$\begin{aligned}\mathrm{Var}(\hat{\delta}) &= \left[D_0' S_0^{-1} D_0\right]^{-1} \\ &= D_0^{-1}\left[S_0^i + S_0^{ii}\right] D_0^{-1}\end{aligned}$$

where

$$D_0^{-1} = \begin{pmatrix} 1+\frac{\mu_p^2}{\sigma_p^2} & -\frac{\mu_p}{\sigma_p^2} & 0 & 0 & \cdots & 0 & 0 \\ -\frac{\mu_p}{\sigma_p^2} & \frac{1}{\sigma_p^2} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1+\frac{\mu_p^2}{\sigma_p^2} & -\frac{\mu_p}{\sigma_p^2} & \cdots & \vdots & \vdots \\ 0 & 0 & -\frac{\mu_p}{\sigma_p^2} & \frac{1}{\sigma_p^2} & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 1+\frac{\mu_p^2}{\sigma_p^2} & -\frac{\mu_p}{\sigma_p^2} \\ 0 & 0 & \cdots & \cdots & \cdots & -\frac{\mu_p}{\sigma_p^2} & \frac{1}{\sigma_p^2} \end{pmatrix}$$

$$S_0^i = \begin{pmatrix} \sigma_1^2 & \sigma_1^2\mu_p & \sigma_{12} & \sigma_{12}\mu_p & \cdots & \sigma_{1N} & \sigma_{1N}\mu_p \\ \sigma_1^2\mu_p & \sigma_1^2(\mu_p^2+\sigma_p^2) & \sigma_{12}\mu_p & \sigma_{12}(\mu_p^2+\sigma_p^2) & \cdots & \sigma_{1N}\mu_p & \sigma_{1N}(\mu_p^2+\sigma_p^2) \\ \sigma_{21} & \sigma_{21}\mu_p & \sigma_2^2 & \sigma_2^2\mu_p & \cdots & \sigma_{2N} & \sigma_{2N}\mu_p \\ \sigma_{21}\mu_p & \sigma_{21}(\mu_p^2+\sigma_p^2) & \sigma_2^2\mu_p & \sigma_2^2(\mu_p^2+\sigma_p^2) & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{N1} & \sigma_{N1}\mu_p & \cdots & \cdots & \cdots & \sigma_N^2 & \sigma_N^2\mu_p \\ \sigma_{N1} & \sigma_{N1}(\mu_p^2+\sigma_p^2) & \cdots & \cdots & \cdots & \sigma_N^2\mu_p & \sigma_N^2(\mu_p^2+\sigma_p^2) \end{pmatrix}$$

$$S_0^{ii} = \begin{pmatrix} 0 & \sigma_{(\tilde\epsilon_{1t}^2,\, \tilde r_{pt})} & 0 & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt})} & \cdots & 0 & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt})} \\ \sigma_{(\tilde\epsilon_{1t}^2,\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{1t}^2,\, \tilde r_{pt}^2)} & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt}^2)} & \cdots & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt}^2)} \\ 0 & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt})} & 0 & \sigma_{(\tilde\epsilon_{2t}^2,\, \tilde r_{pt})} & \cdots & \vdots & \vdots \\ \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{2t},\, \tilde r_{pt}^2)} & \sigma_{(\tilde\epsilon_{2t}^2,\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{2t}^2,\, \tilde r_{pt}^2)} & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt})} & \cdots & \cdots & \cdots & 0 & \sigma_{(\tilde\epsilon_{Nt}^2,\, \tilde r_{pt})} \\ \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{1t}\tilde\epsilon_{Nt},\, \tilde r_{pt}^2)} & \cdots & \cdots & \cdots & \sigma_{(\tilde\epsilon_{Nt}^2,\, \tilde r_{pt})} & \sigma_{(\tilde\epsilon_{Nt}^2,\, \tilde r_{pt}^2)} \end{pmatrix}$$

Since $\hat\alpha = [I_N \otimes (1 \;\; 0)]\hat\delta$, we can calculate the asymptotic variance-covariance matrix of the $\hat\alpha$ estimator explicitly. Performing the matrix operations above, the asymptotic variance-covariance matrix of $\hat\alpha$ is

$$\operatorname{var}(\hat\alpha) \equiv \Omega = \left(1 + \frac{\mu_p^2}{\sigma_p^2}\right)\Sigma + \frac{1}{\sigma_p^2}\Psi$$

where $\Sigma = N \times N$ variance-covariance matrix of the $\tilde{\epsilon}_t$, with (i, j)th element $[\sigma_{ij}]$.

$\Psi = N \times N$ "correction factor" matrix for contemporaneous conditional heteroskedasticity, with (i, j)th element

$$\left[\frac{\mu_p^2}{\sigma_p^2}\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt}^2)} - 2\mu_p\left(1 + \frac{\mu_p^2}{\sigma_p^2}\right)\sigma_{(\tilde{\epsilon}_{it}\tilde{\epsilon}_{jt},\, \tilde{r}_{pt})}\right].$$

REFERENCES

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

Barone-Adesi, G. and P. Talwar, 1983, Market models and heteroskedasticity of security returns, *Journal of Business and Economic Statistics* 4, 163-168.

Blattberg, R. and N. Gonedes, 1974, A comparison of the stable and student distributions as statistical models of stock prices, *Journal of Business* 47, 244-280.

Breeden, D., M. Gibbons, and R. Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, *Journal of Finance* 44, 231-262.

Connor, G. and R. Korajczyk, 1988, Risk and return in an equilibrium APT: Application of a new test methodology, *Journal of Financial Economics* 21, 255-290.

Diebold, F., J. Im, and J. Lee, 1989, Conditional heteroskedasticity in the market model, Working paper, University of Pennsylvania, *Journal of Accounting, Auditing, and Finance*, Forthcoming.

Eichenbaum, M., L. Hansen, and K. Singleton, 1988, A time series analysis of representative agent models of consumption and leisure choice under uncertainty, *Quarterly Journal of Economics* CIII, 51-78.

Fama, E., 1965, The behavior of stock market prices, *Journal of Business* 38, 34-105.

——, 1976, *Foundations of Finance*, (Basic Books, New York).

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

——, S. Ross, and J. Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Hansen, L., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1269-1286.

—— and K. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269-1286.

Hsu, D., 1982, A Bayesian robust detection of shifts in the risk structure of stock market returns, *Journal of the American Statistical Society* 77, 29-39.

Ingersoll, J., 1987, *Theory of Financial Decision Making*, (Rowman & Littlefield, Totowa, NJ).

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

Kandel, S. and R. Stambaugh, 1987, On correlations and inferences about mean-variance efficiency, *Journal of Financial Economics* 18, 61-90.

Lehmann, B. and D. Modest, 1988, The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics* 21, 213-254.

MacKinlay, A., 1985, An analysis of multivariate financial tests, Unpublished dissertation, University of Chicago.

——, 1987, On multivariate tests of the CAPM, *Journal of Financial Economics* 18, 341-372.

Newey, W. and K. West, 1987, A simple positive definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-705.

Richardson, M. and T. Smith, 1990, Tests of financial models in the presence of overlapping observations, *Review of Financial Studies*, Forthcoming.

Schwert, G. and P. Seguin, 1990, Heteroskedasticity in stock returns, *Journal of Finance*, Forthcoming.

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

——, 1987, Multivariate proxies and asset pricing relations: Living with the Roll critique, *Journal of Financial Economics* 18, 91-110.

——, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99-120.

Stambaugh, R., 1982, On the exclusion of assets from tests of the two parameter model, *Journal of Financial Economics* 10, 235-268.

White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817-838.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley & Sons, New York).