

# International asset pricing with alternative distributional specifications\*

Campbell R. Harvey<sup>a</sup> and Guofu Zhou<sup>b</sup>

<sup>a</sup>Duke University, Durham, NC 27706, USA

<sup>b</sup>Washington University, St. Louis, MO 63130, USA

Received January 1992. Final version accepted November 1992

## Abstract

The unconditional mean-variance efficiency of the Morgan Stanley Capital International world equity index is investigated. Using data from 16 OECD countries and Hong Kong and maintaining the assumption of multivariate normality, we cannot reject the efficiency of the benchmark. However, residual diagnostics reveal significant departures from normality. We test the sensitivity of the results by specifying error structures that are *t*-distributed and mixtures of normal distributions. Even after relaxing the i.i.d. assumption, we cannot reject the mean-variance efficiency of the world portfolio. Our results suggest that differences in country risk exposure, measured against the MSCI world portfolio, will lead to differences in expected returns.

## 1. Introduction

As world capital markets have become more integrated, there is increased interest in international asset allocation. A number of mutual funds offer country index portfolios. Pension funds are beginning to realize the benefit of international participation. As such, it is important to be able to reliably assess the risk of investing in different national markets and to understand how that risk is rewarded.

To investigate the relation between risk and expected returns, an asset pricing model must be specified. However, the usefulness of the standard asset pricing paradigms has been challenged in recent cross-sectional work. But, with growing international integration, the case for the NYSE-AMEX portfolio being the 'market portfolio' is strained. Indeed, Harvey (1991) finds that the U.S. equity portfolio is unconditionally dominated by a more general 'world' portfolio. As

*Correspondence to:* C.R. Harvey, Fuqua School of Business, Duke University, Durham, NC 27706, USA.

\*Part of paper this paper was written while the first author was visiting the University of Chicago. We have benefitted from the many suggestions of two anonymous referees, and the Editor, Franz Palm.

such, it is of considerable interest to investigate the viability of a global asset pricing model.

The goal of this paper is to test a single factor world capital asset pricing model. This model implies that the cross-section of expected asset returns is related to the assets' sensitivities to movements in a world market portfolio. Testing this model is equivalent to testing the unconditional mean-variance efficiency of the world market portfolio.

Most previous work starts with the assumption of multivariate normality of the joint returns distribution.<sup>1</sup> However, it is also well known that equity returns, in many cases, depart from normality. So an important part of our analysis is to assess the sensitivity of the inference to the distributional assumptions.

We begin our investigation by assuming, as others have, that returns are drawn from a multivariate normal distribution. The multivariate tests fail to reject the unconditional efficiency of this benchmark portfolio.

The next part of the paper examines the distributional properties of the market model residuals. Skewness and kurtosis statistics are presented and exact probability values are presented for the test of the null hypothesis that the data are generated from a normal distribution. Tests for conditional heteroskedasticity are also presented here. The evidence suggests that there are departures from normality in many cases.

Given the evidence against normality, we modify the multivariate asset pricing tests for alternative distributional specifications. Tests are presented for two cases of the multivariate- $t$  distribution and two cases of mixtures of normals. The distributions are chosen to be, in one case, close to normal and in the other representing a substantial departure from the normal. The multivariate efficiency tests are remarkably robust to changes in the distributional specifications.

Finally, we present tests that drop all distributional assumptions except for stationarity and ergodicity. Multivariate tests based on the generalized method of moments estimation are robust to departures from normality and to conditional heteroskedasticity. Even with this general specification, the mean-variance efficiency of the world portfolio cannot be rejected. However, when the tests are specialized to the returns of the group of seven countries, there is some evidence against the efficiency of this benchmark.

The paper is organized as follows. In the second section, the approaches to testing mean-variance efficiency under the assumption of multivariate normality are detailed. Next, the tests are modified to accommodate alternative distributional specifications: the multivariate  $t$ -distribution and the mixture of normals. Finally, tests based on the generalized method of moments are presented. The

<sup>1</sup>For a review and a list of references on tests of international asset pricing models, see Solnik (1977), Adler and Dumas (1983) and Stulz (1983, 1992).

data are described and some summary statistics are presented in the third section. The fourth section contains the empirical results. Some concluding remarks are offered in the final section.

## 2. Methodology

### 2.1. An international CAPM

Tests of the single factor capital asset pricing model of Sharpe (1964) and Lintner (1965) begin with a statistical representation of the returns generating process. The multivariate linear factor model or 'market model' regression is:

$$r_t = \alpha + \beta r_{pt} + \varepsilon_t, \quad t = 1, \dots, T, \quad (1)$$

where  $r$  is a  $N \times 1$  (number of assets) vector representing returns in excess of a reference asset,  $\alpha$  is the intercept vector,  $\beta$  is the slope vector. The common factor is  $r_p$  which, if a traded security, represents the return on the asset. The slope vector,  $\beta$ , is the sensitivity of the portfolio return to changes in this benchmark return. Finally,  $\varepsilon$  is the disturbance vector with mean zero and a variance-covariance matrix  $\Sigma$ . The disturbance also has the following properties:

$$E\varepsilon_{it}\varepsilon_{js} = \begin{cases} \sigma_{ij}, & s = t; \\ 0, & \text{otherwise;} \end{cases} \quad (2)$$

for all assets  $i$  and  $j$ . Multivariate normality of the asset returns is sufficient for the statistical specification in (1).

Assuming that agents minimize the variance of their portfolio for a given level of expected return and assuming that the reference asset which is used to calculate the excess returns is riskless, Sharpe (1964) and Lintner (1965) show that expected excess returns will be proportional to an expected benchmark return:

$$E[r_t] = \beta E[r_{pt}], \quad t = 1, \dots, T, \quad (3)$$

In the model of Sharpe and Lintner, this benchmark portfolio,  $r_{pt}$ , is a value-weighted portfolio of all invested wealth. Investors choose a portfolio strategy by leveraging up or down this market portfolio.

In terms of (1), the asset pricing model imposes the following restriction:

$$H_0: \alpha = \mathbf{0}. \quad (4)$$

Of course, market portfolio is assumed to be efficient in terms of expected return and variance. In this sense, a test of the Sharpe–Lintner model is also a test of the mean-variance efficiency of the given portfolio which is specified as the

benchmark. Roll (1977) makes a powerful case for caution in interpreting tests of the CAPM. If the benchmark is incorrectly specified, then we may falsely reject the asset pricing model. Indeed, it may never be possible to test the CAPM because of the difficulty in identifying the market portfolio, but the efficiency of the a given benchmark can be tested.

Roll's (1977) critique of tests of the CAPM was directed at empirical work which used U.S. equity returns and some version of the NYSE equity stock portfolio as the market portfolio. Stambaugh (1982) investigated the sensitivity of inference to different specifications of this market portfolio. However, Stambaugh's different benchmarks were exclusively composed of U.S. assets.

While it might have been reasonable in early empirical work to assume that some U.S. portfolio is a plausible proxy for all invested wealth, it is more difficult to make this case today. World capital market have become much more integrated. Chan, Karolyi and Stulz (1992, figure 1) show that U.S. share of world equity markets has dropped below 40% in recent years. Furthermore, Chan, Karolyi and Stulz, as well as Harvey (1991), show that the risk premium on the U.S. 'market' portfolio is driven not by the variance of the market return (as one would expect with a domestic version of the CAPM) but by the covariance with a world market return benchmark.

However, for (1) to hold in an international setting, it is necessary to make additional assumptions. Capital markets must be integrated. In addition, investors are mean-variance optimizers in a common numeraire currency which is assumed to be the U.S. dollar. The sufficient conditions for the international CAPM to hold are detailed in Solnik (1974), Stulz (1981, 1983) and Adler and Dumas (1983).

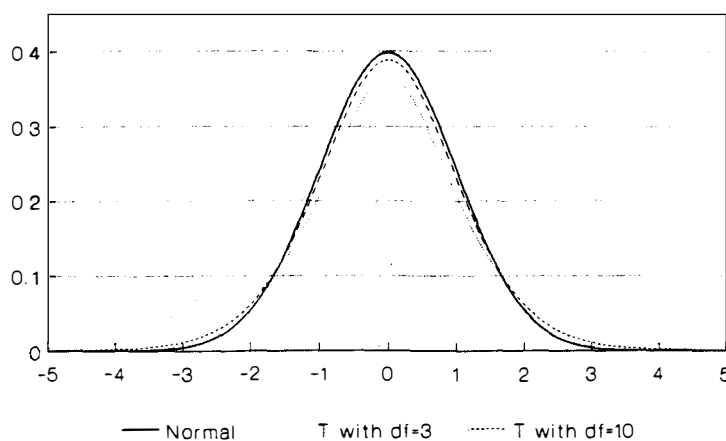


Fig. 1. A comparison of a standardized normal distribution (solid curve) with two  $t$  distributions. The  $t$  distribution is  $f(x) = C/\sigma [1 + x^2/v\sigma^2]^{-(1+v)/2}$  where  $C = \Gamma((1+v)/2)/\Gamma(v/2)(v\pi)^{1/2}$  is a constant,  $v$  is the degrees of freedom,  $\sigma^2$  is the variance and  $\Gamma(-)$  is the gamma function. The distributions are computed for  $v=3$  (dotted curve) and for  $v=10$  (dashed curve).

Tests of the international CAPM are now joint tests of the mean-variance efficiency of the world market portfolio and the integration of world capital markets. Interestingly, there has been little work assessing the unconditional mean-variance efficiency of a world market portfolio. Cumby and Glen (1990) are unable to reject the efficiency of the Morgan Stanley Capital International (MSCI) world portfolio using a sample of mutual funds and six years of data. Harvey (1991, figure 1) shows that the U.S. equity portfolio is unconditionally dominated (lower expected return and higher variance) by the MSCI world portfolio.

The world market portfolio also plays a role in the conditional asset pricing tests of Harvey (1991), Chan, Karolyi and Stulz (1992) and Ferson and Harvey (1992). With some approaches, such as the latent variables methodology, the benchmark need not be prespecified. However, the evidence presented in Harvey, Solnik and Zhou (1992) suggests that the first latent factor is 95% correlated with the MSCI world index.<sup>2</sup> Of course, the unconditional efficiency is a special case of conditional efficiency – when sufficient restrictions are placed on the conditioning information.

Given the specification of the MSCI world equity index as the market portfolio, we will explore alternative ways to test the restriction in (4)  $\alpha=0$ . Special attention will be paid to the sensitivity of inferences to the distribution specifications.

## 2.2. *The classical F-test*

Gibbons, Ross and Shanken (1989) propose a multivariate test of the intercept restriction (4). They propose the statistic:

$$W_u = \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} / (1 + \hat{\theta}_p^2), \quad (5)$$

where  $\hat{\theta}_p$  is the ratio of the mean to the standard deviation (unadjusted for degrees of freedom) of the given portfolio return and  $\hat{\Sigma}$  is the unbiased OLS estimator (cross-product of the residuals divided by  $T-2$ ) of the variance-covariance of the disturbances in (1).

The test statistic is:

$$[T(T-N-1)/N(T-2)] W_u, \quad (6)$$

where  $N$  is the number of assets and  $T$  represents the number of time-series observations. This statistic follows a non-central  $F$ -distribution with numerator

<sup>2</sup>The latent variables approach was pioneered by Hansen and Hodrick (1983) and Gibbons and Ferson (1985). Recent empirical tests using international data are also found in Bekaert and Hodrick (1992), Campbell and Hamao (1992), Chang, Pinegar and Ravichandran (1991) and Harvey (1991).

and denominator degrees of freedom,  $N$  and  $(T-N-1)$ , respectively. The noncentrality parameter is:

$$\lambda = [T/(1 + \theta_p^2)] \mathbf{a}' \Sigma^{-1} \mathbf{a}. \quad (7)$$

However, under the null hypothesis  $\mathbf{a} = \mathbf{0}$ , which implies that  $\lambda = 0$ . Hence, the statistic is distributed as a central  $F$ .

Gibbons, Ross and Shanken (1989) also provide the geometric intuition behind the  $W_u$  statistic.<sup>3</sup> The condition  $\mathbf{a} = \mathbf{0}$  implies the benchmark portfolio is tangent to a ray originating from the zero point on the expected excess return axis and touching the unconditional mean-variance frontier. Of course, given some data, we can construct this ray (to an ex post efficient portfolio) and measure its slope. Similarly, we could measure the slope of the ray with the same origin point to the given portfolio (say the MSCI world market portfolio). The  $W_u$  statistic is just the ratio of one plus the squared slope to the tangency portfolio to one plus the squared slope to the given portfolio. If the given portfolio is efficient or close to efficient, this ratio is small and the test statistic is small. However, if the given portfolio is well inside the frontier, the ratio will be large and the statistic will yield evidence against the mean-variance efficiency.<sup>4</sup>

The approach of Gibbons, Ross and Shanken (1989) relies upon the multivariate normality of the disturbance distribution and on conditional homoskedasticity. With any deviation from these assumptions, the statistic has unknown properties. Our goal is to investigate what happens to the inference when the distributional specification changes.

### 2.3. *Exact tests for normality*

Both the Gibbons, Ross and Shanken (1989) test assumes that the distribution of the residuals is multivariate normal. We conduct diagnostics of the residuals and present exact tests for normality. In other words, the p-value is exactly the p-value with the given sample size in contrast to the asymptotic p-value where the number of observations must be very large.<sup>5</sup> We do this by calculating the sample third and fourth central moments, the skewness and kurtosis, and comparing the sample moments with those expected if the data are drawn from a normal distribution.

<sup>3</sup>Also see the discussion in Shanken (1985, 1986) and Harvey and Zhou (1990).

<sup>4</sup>See Harvey and Zhou (1992) for a Bayesian analysis of the mean-variance efficiency tests in an international setting.

<sup>5</sup>Asymptotic tests are presented in, for example, Richardson and Smith (1992).

Let  $X_1, \dots, X_T$  be the observations on a random variable  $X$ . The (sample) skewness and kurtosis for each of residuals is defined as:

$$b_1 = \frac{1}{T} \sum_{i=1}^T c_i^3 / s^3 \quad \text{and} \quad b_2 = \frac{1}{T} \sum_{i=1}^T c_i^4 / s^4, \quad (8)$$

where  $c_i \equiv X_i - \bar{X}$ , and  $\bar{X}$  and  $s^2$  are the sample mean and sample variance respectively.

When  $X$  is from a symmetric distribution, we expect, on average, to observe an equal number of positive  $c_i$ 's and negative  $c_i$ 's, implying  $b_1$  should be close to zero. Indeed, if the data are from a normal distribution, it is well known that [see, e.g., Pearson (1930)]

$$E(b_1) = 0. \quad (9)$$

If the observed  $b_1$  is large, the data are skewed and we reject the normality. To get some intuition about the kurtosis, we notice that  $c_i^2$  is the (squared) distance between  $X_i$  and  $\bar{X}$ . When  $X_i$  and  $\bar{X}$  are close,  $c_i^2$  is close to zero. Therefore,  $b_2$  measures the overall departure of the data points from the 'central' point  $\bar{X}$ . If the data points are heavily concentrated near  $\bar{X}$ ,  $b_2$  will be close to zero. On the other hand, if the data points depart greatly from  $\bar{X}$ ,  $b_2$  will be large. Both of these cases are violations of normality. In the normality case, it can be shown that

$$E(b_2) = 3(T-1)/(T+1). \quad (10)$$

So, if the data actually comes from a normal distribution, we would expect to get on the average a value of about 3 for  $b_2$ .

The exact p-value of both the skewness and kurtosis test is computed by a method suggested in Zhou (1992). To motivate the method, notice that if we can draw samples from the residuals, the distribution of both the skewness and kurtosis can then be determined by the samples, because they are functions of the samples, and hence the p-value. The null hypothesis is that the residuals are normal with mean zero. Under the null, the residuals follow  $N(0, \sigma^2)$ , where  $\sigma$  is the standard deviation. Because  $\sigma$  is an unknown parameter, samples of the residuals cannot be drawn from the normal distribution. However, there is an important property of both the skewness and kurtosis – they are invariant to scaling. That is, samples from  $N(0, \sigma^2)$  or samples from  $N(0, 1)$  will have the same values of skewness and kurtosis. Therefore, as far as the distribution of both the skewness and kurtosis is concerned, we can simply draw samples from  $N(0, 1)$ . With say 10,000 samples, the means of the skewness and kurtosis parameters are given by the sample means. The p-values are determined from the sample distributions.

The distributions of both the skewness and kurtosis are given by an integration over a  $T$ -dimensional space. For example,  $\text{Prob}(b_1 < x)$  is the integration

over the region  $b_1 < x$  of the standard normal densities multiplied by  $dx_i$ ,  $i = 1, \dots, T$ , where  $x = (x_1, \dots, x_T)$  is in the  $T$ -dimensional space. Therefore, the above method for the computation of the p-values is also a Monte Carlo integration approach. As shown in Zhou (1992), with 10,000 samples, the results are 2 to 3 digits accurate, which for our purposes are satisfactory.

#### 2.4. Tests assuming alternative distributions

Although we don't know the true distributions of the residuals, it is of interest to consider two alternative multivariate distributions which contain the multivariate normal as special cases. These distributions are chosen to provide a closer match to the data than the usual assumption of a multivariate normal distribution.

The first is the multivariate  $t$  with density:

$$f(\mathbf{E}) = C |\Sigma|^{-1/2} [1 + \mathbf{E}' \Sigma^{-1} \mathbf{E} / \nu]^{-(N+\nu)/2}, \quad (11)$$

where  $C = \Gamma((N+\nu)/2) / \Gamma(\nu/2) (\nu\pi)^{N/2}$  is a constant,  $\mathbf{E}$  is an  $N \times 1$  vector,  $\nu$  represents the degrees of freedom and  $\Gamma(-)$  is the gamma function. With a multivariate  $t$  assumption on the residuals, the first two moments are:

$$E(\mathbf{E}) = \mathbf{0}, \quad \text{and} \quad \text{Cov}(\mathbf{E}) = \frac{\nu}{\nu-2} \Sigma, \quad \nu > 2. \quad (12)$$

The multivariate  $t$  has fatter tails than the normal. As the degrees of freedom get large, this distribution becomes identical to the normal. However, to have finite second moments, the degrees of freedom must be greater than two.

The idea of obtaining the exact p-values of the Gibbons, Ross and Shanken (1989) test is similar to the way we calculated the exact p-values for the normality tests using the sample skewness and kurtosis. It is easy to write the  $W_u$  statistic as a function of the residuals. Thus, if the residuals can be drawn from the multivariate  $t$  distribution, it is straightforward to obtain the exact p-values of the multivariate test. However, the parameter  $\Sigma$  is unknown. Nevertheless, the theoretical results in Zhou (1992) show that the multivariate test is invariant to the specification of  $\Sigma$ . Hence, to obtain p-value, we can assume  $\Sigma = \mathbf{I}$ , and the method follows precisely the same steps outlined for the normality tests.

Of course, our goal is to assess the sensitivity of the multivariate test proposed by Gibbons, Ross and Shanken (1989) to the choice of distributional assumptions. We choose two different  $t$  distributions. The first has 3 degrees of freedom and represents a large deviation from normality. The second distribution has 10 degrees of freedom and resembles more closely a normal distribution. Fig. 1 provides a comparison of a univariate normal distribution to these two



$t$  distributions. It is evident that the  $t$  distributions have skewness and kurtosis that may better approximate the data.

The second set of distributions are referred to as mixture normal distributions. This distribution has a density function:

$$f(\mathbf{E}) = (1 - \varepsilon) \frac{|\Sigma|^{-1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2} \mathbf{E}' \Sigma^{-1} \mathbf{E}\right) + \varepsilon \frac{|\Sigma|^{-1/2}}{(2\pi\gamma)^{N/2}} \exp\left(-\frac{1}{2\gamma} \mathbf{E}' \Sigma^{-1} \mathbf{E}\right), \quad (13)$$

where  $\varepsilon$  and  $\gamma$  are the degrees of freedom. When  $\varepsilon = 0$ , the distribution is the usual multivariate normal form. When  $0 < \varepsilon < 1$ , the distribution is a mixture of two normal distributions.

The mean and covariance of the mixture normal distribution are:

$$E(\mathbf{E}) = \mathbf{0}, \quad \text{and} \quad \text{Var}(\mathbf{E}) = (1 - \varepsilon + \varepsilon\gamma)\Sigma. \quad (14)$$

In our efficiency test, we fix the parameter  $\gamma = 10$  so that the second normal distribution has covariance matrix 10 times as large as the first.<sup>6</sup> The p-values are then computed at  $\varepsilon = 3\%$  and  $\varepsilon = 50\%$ , representing small and large deviations from the normality.

Fig. 2 compares these two mixture normal distributions with a normal distribution. The mixture normals have fatter tails than the normal distribution.

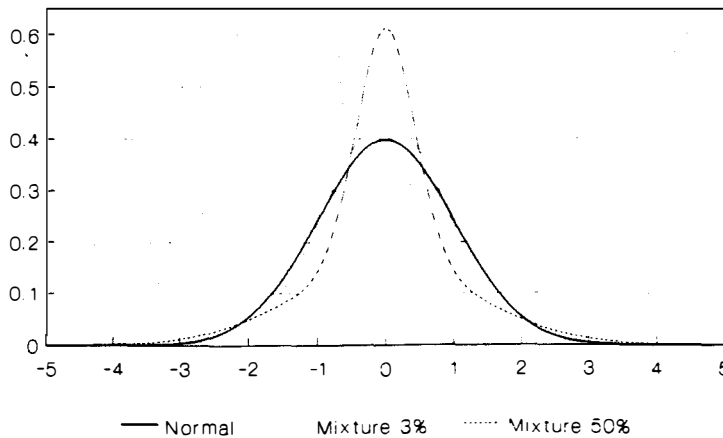


Fig. 2. A comparison of a standardized normal distribution (solid curve) and two mixture normal distributions. The mixture normal is  $f(x) = (1 - \varepsilon) 1/(2\pi)^{1/2} \sigma \exp(-x^2/2\sigma^2) + \varepsilon (1/(2\pi\lambda))^{1/2} \sigma \exp(-x^2/2\lambda\sigma^2)$ , where  $\varepsilon$  and  $\lambda$  are the degrees of freedom and  $\sigma^2$  is the variance. When  $\varepsilon = 0$ , the distribution is normal. When  $0 < \varepsilon < 1$ , the distribution is a mixture of two normal distributions. We set the parameter  $\lambda = 10$  so that the second normal distribution has covariance matrix 10 times as large as the first. The distributions are then computed at  $\varepsilon = 3\%$  (dotted curve) and  $\varepsilon = 50\%$  (dashed curve), representing small and large deviations from the normality.

<sup>6</sup>Our goal in choosing  $\gamma = 10$  is to obtain a distribution that substantially deviates from the normal.

As with the  $t$  distribution, the mixture normal is centered around zero and does not account for the skewness in the disturbance distribution. However, these two distributions may more closely approximate the true distribution of the data and provide a check of the robustness of the inference to departures from normality.

Similar to our approach with the multivariate  $t$ , we recalculate the multivariate test of whether the intercepts are zero assuming that the residual distribution is multivariate mixture normal. The procedure for computing the exact p-values is the same for the multivariate mixture normal except that samples are now drawn from the mixture normal distribution with  $\Sigma = \mathbf{I}$ .

### 2.5. Multivariate GMM tests

The tests presented so far assume that the data are drawn from particular probability distributions. Initial test statistics have been calculated assuming the data are normal. We then considered four alternative distributions: two versions of the multivariate  $t$  and two mixture normal distributions. In all four of these cases, the residuals in (1) are assumed to be i.i.d. We now present tests which use Hansen's (1982) generalized method of moments. No particular distributional assumptions are needed other than the data are stationary and ergodic. The residual distribution need not be i.i.d.

The GMM begins with the disturbance definition in (1):

$$\varepsilon_t = r_t - \alpha - \beta r_{pt} . \quad (15)$$

The idea of the GMM approach<sup>7</sup> is to use sample moments conditions to replace the population moment conditions of the model restrictions. The parameter estimates are obtained as the solution of minimizing a sum of the weighted sample moments conditions.

The relevant population moment conditions are:

$$E(\varepsilon_t) = \mathbf{0} , \quad (16)$$

and

$$E(r_{pt}\varepsilon_t) = \mathbf{0} . \quad (17)$$

The corresponding sample moment conditions are:

$$g_{ur}(a, \beta) = \sum_{t=1}^T \begin{pmatrix} \varepsilon_t \\ r_{pt}\varepsilon_t \end{pmatrix} , \quad (18)$$

<sup>7</sup>See Harvey (1989), Richardson and Smith (1991) and MacKinlay and Richardson (1991) for a description of how to use the GMM in asset pricing tests.

divided by the number of observations. The subscript  $ur$  denotes ‘unrestricted’. The dimension of  $g_{ur}$  is  $2N$  by 1 where  $N$  is the number of assets.

At the true parameters, (16) and (17) are satisfied exactly and so their sample analog should be small. This implies that the weighted residuals and the cross products should be small,

$$J = g_{ur}(\alpha, \beta)' w_{ur} g_{ur}(\alpha, \beta), \quad (19)$$

where  $w_{ur}$ ,  $2N \times 2N$ , is the weighting matrix. However, when (19) is minimized, it will only equal zero if there is a parameter for each moment condition (model is exactly identified). In general, we cannot require that all the moment conditions be satisfied because there may be more moment conditions than parameters available (model is overidentified).

The parameter estimator that minimizes (19) is Hansen’s GMM. The weighting matrix  $w_{ur}$  can be optimally chosen to minimize the asymptotic covariance matrix of the GMM estimator. The covariance matrix is

$$(\mathbf{D}'_{ur} w_{ur}^{-1} \mathbf{D}_{ur})^{-1}, \quad (20)$$

where  $\mathbf{D}_{ur}$ ,  $2N \times m$ , are the derivatives of  $g_{ur}$  with respect to the parameters evaluated at the GMM estimator, and  $m$  is the number of parameters. The procedure of Hansen (1982), Newey and West (1987a), or Andrews (1991) can be used to construct the optimal weighting matrix which is a consistent estimator of the underlying distribution of the residuals.

There are three ways to use the GMM approach to test efficiency. The first, suggested by MacKinlay and Richardson (1991), uses the GMM procedure without imposing the efficiency restrictions that  $\alpha = \mathbf{0}$ . In this case, there are  $m = 2N$  parameters and  $2N$  sample moments conditions. The system is exactly identified and we can choose the estimator of the alphas and betas so that all the sample moments conditions are zero. Since the model is linear, the parameters will be precisely the ones estimated with ordinary least squares. However, the standard errors will be different in the presence of heteroskedasticity.

Let the asymptotic covariance matrix for the alpha estimators alone be  $\mathbf{D}_\alpha$ , which is obtained from (20) and is an  $N \times N$  matrix. Then, under the null, the covariance matrix can be factored out to give a  $\chi^2$  test for the efficiency:

$$H_1 = \hat{\alpha}' \mathbf{D}_\alpha^{-1} \hat{\alpha} \sim \chi^2_N. \quad (21)$$

Of course, this test is just the Gibbons, Ross and Shanken (1989) test using the variance-covariance matrix obtained from the GMM estimation.

The second way to test  $\alpha = \mathbf{0}$  is to estimate the restricted model:

$$\varepsilon_t = r_t - \beta r_{pt}. \quad (22)$$

Like the unrestricted model in (15), there are  $N$  (number of countries) equations and two moment conditions for each equation. Hence, the number of orthogonality conditions is  $2N$ . But with the restricted model (22), there are only

$N$  parameters (one  $\beta$  for each country, the  $\alpha$  are restricted to equal zero). Hence, the system is overidentified and the  $\chi^2$  statistic has  $N$  degrees of freedom. This statistic,

$$H_2 = \mathbf{g}'_r \mathbf{w}_r \mathbf{g}_r \simeq \chi^2_N, \quad (23)$$

provides a test of the model's restrictions. In this case, the subscript ' $r$ ' denotes restricted.

The third test<sup>8</sup> follows Gallant and Jorgenson (1979).<sup>9</sup> The idea is to estimate the restricted model (22) using the unrestricted weighting matrix that results from the estimation of the parameters in (15).

$$H_3 = \mathbf{g}'_r \mathbf{w}_{ur} \mathbf{g}_r - \mathbf{g}'_{ur} \mathbf{w}_{ur} \mathbf{g}_{ur} \sim \chi^2_N. \quad (24)$$

The statistic is  $\chi^2$  distributed with degrees of freedom equal to the number of restrictions, which in our case, is  $N$ .

Of course, all of these tests are asymptotically equivalent. However, in small samples, their performance could be different. Hence, it is of interest to examine many different ways to test for efficiency. In our empirical work, we focus on the last two GMM-based efficiency tests.

### 3. Data and preliminary analysis

#### 3.1. Sources

The data in this study are from Morgan Stanley Capital International. Monthly data on equity indices for 16 OECD countries and Hong Kong<sup>10</sup> are available from December 1969 to May 1989. These indices are value weighted and are calculated with dividend reinvestment. Morgan Stanley also calculates a value weighted world equity index which serves as the market portfolio.

The MSCI international indices are composed of stocks that broadly represent stock composition in the different countries. For example, there is a 99.1% correlation between the MSCI U.S. excess return and the New York Stock Exchange value-weighted return calculated by the Center for Research in

<sup>8</sup>A fourth test based on an analytic solution is also presented in MacKinlay and Richardson (1991).

<sup>9</sup>For implementations of this statistic in asset pricing tests see Newey and West (1987b), Eichenbaum, Hansen, and Singleton (1988), Harvey (1989), and Eichenbaum and Hansen (1990).

<sup>10</sup>The 16 OECD countries are Australia, Austria, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom, and the United States. Morgan Stanley also has data on Finland and New Zealand but only from December 1987. Data are available for Singapore/Malaysia but dividend data are not available for the full period. As a result, these countries are omitted from the empirical analysis.

Security Prices (CRSP) at the University of Chicago.<sup>11</sup> There is a 95% correlation between the MSCI Japanese excess return and the Nikkei 225 return. An important difference between the MSCI indices and other national indices such as CRSP is the exclusion of investment companies and foreign domiciled companies. These stocks are excluded to avoid double counting. The weight that each country commands in the MSCI world index is also reported.<sup>12</sup>

All returns are calculated in excess of the U.S. Treasury bill that is closest to 30 days to maturity on the last trading day of the month. Data from 1970–1988 are drawn from the CRSP Government Bond File. The data for 1989 are from the Wall Street Journal. Holding period returns are calculated in the same way as Fama (1984).

### 3.2. *Summary statistics*

Means, standard deviations, as well as tests for autocorrelation, heteroskedasticity, skewness and kurtosis of the monthly returns from January 1970 to May 1989 are provided in Table 1. The highest mean excess return over the sample is from the Hong Kong market. Hong Kong also has the highest volatility. The United States has one of the lowest average returns. However, the volatility of the U.S. stock returns is lower than any other country.

The world market portfolio is the value-weighted average of the country returns. The world portfolio has a lower standard deviation than any individual country. Comparing the country portfolios to the world portfolio, there are seven countries (including the U.S.) that are unconditionally dominated by the world market portfolio. That is, given a choice between investing in one of these countries and the world portfolio, the world portfolio is a better investment for the risk averse investor because it delivers a lower unconditional standard deviation and a higher unconditionally expected return.

Of the 17 countries and the world portfolio, the chi-square tests for the presence of autocorrelation reveal that only five countries have significant autocorrelation. These countries are Austria, Denmark, Italy, Japan and Spain. The test results are driven by high first-order serial correlation in Austria, Italy, Japan and Spain. In addition, both Austria and Denmark have significant seasonal autocorrelations.

<sup>11</sup>However, MSCI index is more weighted towards large firms than the CRSP value weighted index. See Harvey (1991).

<sup>12</sup>MacDonald (1989), French and Poterba (1991) and Fedenia, Hodder and Triantis (1991) show that the MSCI world index gives too much weight to the Japanese stocks because of the large amount of cross corporate ownership. However, other indices such as the FT-Actuaries World Index which is compiled by The Financial Times, Goldman, Sachs and Co., and Country NatWest/Wood Mackenzie suffer from the same problem. In March 1989, Japan composed 42.9% of the MSCI index and as of June 1989, 40.7% of the FT-Actuaries index.

Table 1

Summary statistics for monthly excess returns (calculated in U.S. dollars) for 17 country indices and the world market portfolio based on data from January 1970 to May 1989 (233 observations).

Portfolio	Mean	Std. Dev.	Returns		Squared return		Returns	
			$\chi^2$ $\rho_{1:6} = 0$	$\chi^2$ $\rho_{1:12} = 0$	$\chi^2$ $\rho_{1:6} = 0$	$\chi^2$ $\rho_{1:12} = 0$	Skewness	Kurtosis
Australia	0.423	8.227	1.054 (0.983)	6.487 (0.890)	2.662 (0.850)	4.943 (0.960)	-0.724 (0.000)	7.347 (0.000)
Austria	0.550	5.371	24.198 (0.000)	44.120 (0.000)	21.049 (0.002)	47.899 (0.000)	1.217 (0.000)	7.668 (0.000)
Belgium	0.865	6.005	7.052 (0.316)	14.793 (0.253)	12.109 (0.060)	14.950 (0.244)	0.463 (0.004)	5.945 (0.000)
Canada	0.447	5.886	7.486 (0.278)	12.959 (0.372)	13.769 (0.032)	18.267 (0.108)	-0.380 (0.019)	4.787 (0.000)
Denmark	0.715	5.530	19.594 (0.003)	30.305 (0.003)	0.670 (0.995)	5.702 (0.930)	0.423 (0.008)	4.908 (0.000)
France	0.658	7.348	7.218 (0.301)	11.744 (0.466)	7.255 (0.298)	10.168 (0.601)	0.009 (0.955)	4.312 (0.001)
Germany	0.477	5.983	5.891 (0.436)	13.256 (0.351)	9.935 (0.127)	12.232 (0.427)	-0.015 (0.921)	3.657 (0.032)
Hong Kong	1.699	12.752	4.140 (0.658)	7.605 (0.815)	18.971 (0.004)	19.184 (0.084)	0.867 (0.000)	12.033 (0.000)
Italy	0.235	7.760	19.155 (0.004)	26.438 (0.009)	13.823 (0.032)	17.046 (0.148)	0.319 (0.046)	4.035 (0.006)
Japan	1.326	6.085	11.249 (0.081)	23.069 (0.027)	3.627 (0.727)	5.574 (0.936)	0.209 (0.185)	3.507 (0.087)

The Netherlands	0.739	5.567	6.053 (0.417)	16.491 (0.170)	3.944 (0.684)	7.854 (0.796)	- 0.047 (0.768)	5.120 (0.000)
Norway	0.940	8.277	15.776 (0.015)	19.702 (0.073)	5.238 (0.514)	10.099 (0.607)	- 0.102 (0.517)	3.831 (0.014)
Spain	0.368	6.491	22.565 (0.001)	36.946 (0.000)	9.214 (0.162)	11.233 (0.509)	0.027 (0.865)	5.364 (0.000)
Sweden	0.915	6.213	4.439 (0.617)	5.545 (0.937)	14.265 (0.027)	15.304 (0.225)	0.009 (0.959)	3.135 (0.664)
Switzerland	0.447	5.668	4.038 (0.672)	6.287 (0.901)	4.531 (0.605)	11.312 (0.502)	0.003 (0.988)	4.485 (0.001)
United Kingdom	0.728	7.909	14.679 (0.023)	16.834 (0.156)	11.624 (0.071)	12.629 (0.397)	1.378 (0.000)	12.780 (0.000)
United States	0.339	4.715	6.878 (0.332)	10.911 (0.537)	7.374 (0.288)	15.094 (0.236)	- 0.245 (0.123)	5.224 (0.000)
World	0.505	4.188	11.790 (0.067)	15.538 (0.213)	2.273 (0.893)	8.519 (0.743)	- 0.459 (0.005)	4.981 (0.000)

The country returns are calculated in U.S. dollars in excess of the holding period return on the Treasury bill that is closest to 30 days to maturity. The data are from Morgan Stanley Capital International.  $\rho_j$  represents the autocorrelation coefficient at lag  $j$ . The p-values associated with the skewness and kurtosis statistics are tests of the null hypothesis that the data is drawn from a normal distribution.

Table 1 also presents tests for autoregressive conditional heteroskedasticity (ARCH). Interestingly, there are only four countries that exhibit significant ARCH patterns. However, this test does not rule out more general forms of conditional heteroskedasticity.

The final columns in table 1 present skewness and kurtosis statistics as well as the exact p-value testing the null hypothesis that the data are drawn from a normal distribution. Significant skewness is found in nine of the 18 portfolio returns. Interestingly, the U.S. and Japan, the two largest equity portfolios, do not exhibit significant skewness.<sup>13</sup>

The analysis of kurtosis shows much more dramatic departures from normality. All but two of the portfolios exhibit significant kurtosis. For the U.S., we can reject that the data are drawn from a normal distribution. With Japan, the p-value is 0.086, so we cannot reject that the data are drawn from a normal distribution at the standard 5% level of significance.<sup>14</sup>

The significant departures from normality detected in the returns foreshadow the conclusions from diagnostic tests on the disturbances in (15). The non-normality in the returns will probably carry over into non-normality of the disturbances.

## 4. Empirical results

### 4.1. Multivariate tests under normality

Table 2 reports the results of estimating (15) with ordinary least squares. The MSCI world returns account for between 9 and 75% of the variance of the country returns. In all but one of the regressions, the intercept is indistinguishable from zero using the classical p-value. The exception is Japan where the intercept is close to three standard errors from zero.

The Gibbons, Ross and Shanken (1989) test of whether the intercepts are zero is presented at the bottom of table 2. For all 17 countries under consideration, the test does not reject the null hypothesis that the world portfolio is mean-variance efficient. The classical p-value is 26%. The test was recalculated using the Group of Seven (G-7) countries.<sup>15</sup>

<sup>13</sup>We also computed the asymptotic p-values which, in most cases, were consistent with the exact p-values. This is due to the reasonably large sample (233 observations).

<sup>14</sup>It is possible to obtain an exact p-value for the test of multivariate normality. However, the estimated CPU time on a Sun Workstation is 2000 hours compared to only two hours for the univariate tests. Given that the univariate results find departures from normality, it is likely that the multivariate tests would also find departures.

<sup>15</sup>The G-7 countries are Canada, France, Germany, Italy, Japan, United States and the United Kingdom.



Table 2

Ordinary least squares regressions of the country index excess returns (calculated in U.S. dollars) on the excess return on the world market portfolio based on the assumption of i.i.d., multivariate normal errors,  $\varepsilon_{jt}$ . The regressions are based on monthly data from January 1970 to May 1989 (233 observations). The equation estimated is:

$$r_{j,t} = \alpha_j + \beta_j r_{wt} + \varepsilon_{jt}$$

portfolio	$\alpha$	$\beta$	$\bar{R}^2$
Australia	-0.0015 (0.0045)	1.1254 (0.1059)	0.325
Austria	0.0036 (0.0034)	0.3741 (0.0807)	0.081
Belgium	0.0042 (0.0031)	0.8884 (0.0740)	0.381
Canada	-0.0008 (0.0026)	1.0428 (0.0620)	0.549
Denmark	0.0040 (0.0032)	0.6150 (0.0769)	0.214
France	0.0012 (0.0039)	1.0667 (0.0916)	0.367
Germany	0.0008 (0.0033)	0.7799 (0.0787)	0.295
Hong Kong	0.0108 (0.0077)	1.2247 (0.1834)	0.158
Italy	-0.0015 (0.0047)	0.7639 (0.1111)	0.166
Japan	0.0086 (0.0031)	0.9184 (0.0741)	0.397
The Netherlands	0.0024 (0.0025)	0.9837 (0.0588)	0.546
Norway	0.0044 (0.0047)	0.9882 (0.1126)	0.247
Spain	0.0005 (0.0039)	0.6344 (0.0930)	0.164
Sweden	0.0054 (0.0036)	0.7407 (0.0846)	0.246
Switzerland	-0.0002 (0.0028)	0.9162 (0.0655)	0.456
United Kingdom	0.0009 (0.0039)	1.2627 (0.0924)	0.445
United States	-0.0015 (0.0016)	0.9733 (0.0372)	0.746

Multivariate tests of the restriction  $\alpha_j = 0$

Number of assets	F-statistic	Degrees of freedom	P-value
17	1.212	17,215	0.2565
7	1.842	7,225	0.0803

Standard errors are in parentheses. The country returns are calculated in U.S. dollars in excess of the holding period return on the Treasury bill that is closest to 30 days to maturity. The equity data are from Morgan Stanley Capital International.

These countries account for 90% of the world index. While the classical p-values are smaller, the null hypothesis that the world portfolio is mean-variance efficient is not rejected at standard significance levels.<sup>16</sup> The p-values are smaller because Japan is given a larger relative weight when there are fewer portfolios examined.

The tests suggest that the efficiency of the MSCI world market portfolio cannot be rejected. An alternative view is the portfolio is inefficient and test is misleading because of the distributional assumptions. As such, it is important to conduct diagnostics on the regression residuals.

#### *4.2. Are the regression residuals normal?*

Diagnostics of the regression residuals is the next step in our analysis. This is particularly important given the evidence in table 1 that many of the returns deviated from the normality assumption. Tests for autocorrelation and heteroskedasticity of the residuals are presented in table 3. The tests use both six and twelve lags of the returns and squared returns in order to isolate the influence of any seasonality. Exact tests for skewness and kurtosis are also detailed in table 3.

Ljung and Box (1978) tests of the first six autocorrelations reveal significant test statistics in 6 of the 17 countries. When the order of the test is increased to 12, we can reject the hypothesis of zero autocorrelation in two additional countries: Canada and the United States.

Tests of ARCH are also presented in table 3. While significant ARCH effects are found in only 5 countries, these countries include the United States, Germany and the United Kingdom which are some of the largest portfolios. As is true with the results in table 1, the ARCH tests focus on one particular type of heteroskedasticity, the autoregressive. It is possible that the residuals exhibit a more general type of heteroskedasticity not detected with this test.

Skewness and kurtosis statistics are presented for the regression residuals in the final columns of table 3. The exact p-value for the test of the null hypothesis that the residuals are normal is presented in parentheses. On the basis of the skewness test, we can reject normality for seven of the 17 countries. With the kurtosis test, there is evidence against the null hypothesis for 14 countries at the 5% level and 15 countries at the 10% level of confidence.

The message from table 3 is that the residual distributions are not normal. Next we re-estimate the multivariate tests assuming alternative distributional assumptions.

<sup>16</sup>Using a sample of 13 country indices and data from 1982: 1 to 1988: 6, Cumby and Glen (1990) also cannot reject the null hypothesis of the mean-variance efficiency of the MSCI world index.

Table 3

An analysis of the residuals from the linear regression of the country excess returns on the world excess return based on data from January 1970 to May 1989 (233 observations)

Portfolio	Residuals		Squared residuals		Residuals	
	$\chi^2$ $\rho_{1:6}=0$	$\chi^2$ $\rho_{1:12}=0$	$\chi^2$ $\rho_{1:6}=0$	$\chi^2$ $\rho_{1:12}=0$	Skewness	Kurtosis
Australia	4.765 (0.574)	13.094 (0.362)	8.524 (0.202)	15.629 (0.209)	-0.138 (0.378)	3.865 (0.013)
Austria	26.432 (0.000)	45.056 (0.000)	19.186 (0.004)	49.317 (0.000)	1.376 (0.000)	8.286 (0.000)
Belgium	5.619 (0.467)	10.996 (0.529)	2.264 (0.894)	10.403 (0.581)	1.059 (0.000)	5.849 (0.000)
Canada	11.772 (0.067)	24.680 (0.016)	11.262 (0.081)	12.913 (0.375)	-0.012 (0.940)	3.601 (0.047)
Denmark	15.991 (0.014)	23.439 (0.024)	3.249 (0.777)	6.557 (0.885)	0.693 (0.000)	5.138 (0.000)
France	3.265 (0.775)	5.770 (0.927)	6.244 (0.396)	10.556 (0.567)	0.159 (0.311)	4.001 (0.006)
Germany	16.123 (0.013)	26.878 (0.008)	9.260 (0.159)	24.015 (0.020)	0.012 (0.937)	3.406 (0.170)
Hong Kong	4.131 (0.659)	7.074 (0.853)	15.846 (0.015)	16.166 (0.184)	1.514 (0.000)	15.322 (0.000)
Italy	16.160 (0.013)	25.737 (0.012)	7.654 (0.265)	13.915 (0.306)	0.403 (0.012)	4.313 (0.001)
Japan	6.559 (0.364)	9.227 (0.683)	8.642 (0.195)	12.327 (0.420)	0.031 (0.849)	3.652 (0.034)
The Netherlands	8.348 (0.214)	13.190 (0.355)	3.921 (0.687)	12.305 (0.422)	0.249 (0.117)	3.637 (0.038)
Norway	29.862 (0.000)	34.208 0.001	11.079 0.086	14.658 0.261	0.610 (0.000)	3.857 (0.013)
Spain	16.564 (0.011)	24.305 (0.018)	7.251 (0.298)	9.669 (0.645)	-0.048 (0.764)	4.809 (0.000)
Sweden	3.801 (0.704)	5.660 (0.932)	17.792 (0.007)	21.788 (0.040)	0.077 (0.623)	3.572 (0.056)
Switzerland	4.178 (0.653)	7.138 (0.848)	2.988 (0.810)	9.564 (0.654)	0.129 (0.410)	2.973 (0.933)
United Kingdom	10.132 (0.119)	16.827 (0.156)	16.977 (0.009)	23.218 (0.026)	1.224 (0.000)	9.820 (0.000)
United States	9.705 (0.138)	22.849 (0.029)	20.118 (0.003)	35.992 (0.000)	0.243 (0.124)	4.258 (0.002)

The country returns are calculated in U.S. dollars in excess of the holding period return on the Treasury bill that is closest to 30 days to maturity. The data are from Morgan Stanley Capital International.  $\rho_j$  is the autocorrelation coefficient for lag  $j$ . The p-values associated with the skewness and kurtosis statistics are tests of the null hypothesis that the data is drawn from a normal distribution.

#### 4.3. Multivariate tests assuming alternative distributions

Table 4 presents the versions of the Gibbons, Ross and Shanken (1989) test when the residuals are assumed to follow two types of multivariate  $t$  distributions and mixture normal distributions. Both the mixture normal and the  $t$  distribution are chosen because these exhibit skewness and kurtosis which better approximate the empirical distribution of the data detailed in table 4.

The first panel of table 4 presents the tests with the  $t$  distributions. Two types are chosen: degrees of freedom equaling 10 and degrees of freedom equaling 3. This represents small and large departures from normality. However, the test results show that the p-value is virtually unaffected from the case of multivariate normality. In the case of a large departure from normality (degrees of freedom = 3), the p-value when 17 countries are used in the test is 0.248 compared to 0.257 when normality is assumed. With the group of seven countries, the p-value is 0.076 compared to the value of 0.080 reported in table 2.<sup>17</sup>

Table 4

Multivariate tests of the efficiency of the Morgan Stanley Capital International world portfolio using alternative distributional assumptions. Based on monthly data from January 1970 to May 1989 (233 observations).

Distribution of residuals	Number of countries	Exact p-value
Multivariate $t$ (Degrees of freedom = 10)	17	0.252
Multivariate $t$ (Degrees of freedom = 10)	7	0.077
Multivariate $t$ (Degrees of freedom = 3)	17	0.248
Multivariate $t$ (Degrees of freedom = 3)	7	0.076
Mixture normal (3% mixture)	17	0.252
Mixture normal (3% mixture)	7	0.079
Mixture normal (50% mixture)	17	0.259
Mixture normal (50% mixture)	7	0.082

The country returns are calculated in U.S. dollars in excess of the holding period return on the Treasury bill that is closest to 30 days to maturity. The equity data are from Morgan Stanley Capital International.

<sup>17</sup>Our empirical results showing the insensitivity of the test to departures from normality are consistent with the simulations in Affleck-Graves and MacDonald (1989).

The second panel of table 4 presents tests assuming that the residual distribution is multivariate mixture normal. Similar to the results with the multivariate  $t$ , the test is little changed from the results presented in table 2. With a 50% mixture, the p-value using 17 country portfolios is 0.252. With seven countries, the value is 0.082.

In summary, when the alternative distributional assumptions are made, the p-values of the multivariate test of the intercept restriction are generally lower. However, they are only slightly lower. The basic inference, that the MSCI portfolio is mean variance efficient, is unaffected by the assumption of an alternative parametric distributional specification.

#### *4.4. Multivariate tests without parametric distributional specifications*

The final table presents tests that make no parametric distributional assumptions. Hansen's (1982) generalized method of moments technique only requires that the data are stationary and ergodic – and even these are only sufficient assumptions. The technique is also robust to conditional heteroskedasticity of a general form.

Table 5 presents estimates of (15) in its unrestricted form. The parameter estimates are identical to the OLS estimates presented in table 2. However, the standard errors are different. Note that the standard errors on the slope coefficients are generally larger. However, the standard errors on the intercepts are, in most cases, smaller than the ones reported in table 2. This would lead one to expect that the multivariate tests would provide more evidence against the efficiency of the world index.

In estimating the restricted system (22), the p-values of the overidentifying conditions (test  $H_2$ ) are lower than the ones reported under the assumption of normality and conditional homoskedasticity. For the 17 country system, the p-value is 0.238 which does not change the inference. However, for the seven country system, the p-value drops to 0.063 which is close to the rejection region given the standard level of significance at 5%.

The second test estimates the restricted system (22) using the unrestricted weighting matrix (test  $H_3$ ). Harvey (1989, 1991) presents evidence that this test may be more powerful in detecting deviations from conditional mean-variance efficiency. Indeed, the test results in the last panel of table 5 indicate that the p-values drop substantially. With the 17 asset system, the p-value is 0.157 (compared to 0.257 under normality). With the 7 asset system, the p-value is cut in half to 0.047 compared to the one under normality.

Although the p-values drop, we still are unable to reject the mean-variance efficiency of the MSCI with 17 assets. When the number of assets is specialized to the group of seven countries, we now can reject efficiency at the 5% level.

Table 5

Generalized method of moments estimation of the linear relation between country index excess returns (calculated in U.S. dollars) and the excess return on the world market portfolio. The regressions are based on monthly data from January 1970 to May 1989 (233 observations). The equation estimated is:

$$\varepsilon_t = r_t - \alpha - \beta r_{wt}$$

Portfolio	$\alpha$	$\beta$	
Australia	-0.0015 (0.0046)	1.1254 (0.1600)	
Austria	0.0036 (0.0033)	0.3741 (0.0753)	
Belgium	0.0042 (0.0031)	0.8884 (0.0821)	
Canada	-0.0008 (0.0026)	1.0428 (0.0719)	
Denmark	0.0040 (0.0032)	0.6150 (0.0813)	
France	0.0012 (0.0039)	1.0667 (0.0937)	
Germany	0.0008 (0.0032)	0.7799 (0.0868)	
Hong Kong	0.0108 (0.0078)	1.2247 (0.1931)	
Italy	-0.0015 (0.0046)	0.7639 (0.1022)	
Japan	0.0086 (0.0030)	0.9184 (0.0935)	
The Netherlands	0.0024 (0.0024)	0.9837 (0.0687)	
Norway	0.0044 (0.0049)	0.9882 (0.1307)	
Spain	0.0005 (0.0038)	0.6344 (0.1060)	
Sweden	0.0054 (0.0035)	0.7407 (0.0875)	
Switzerland	-0.0002 (0.0027)	0.9162 (0.0750)	
United Kingdom	0.0009 (0.0036)	1.2627 (0.1478)	
United States	-0.0015 (0.0015)	0.9733 (0.0428)	
Test $\alpha = 0$	Number of countries	$\chi^2$	p-value
$g_r W_r g_r$	17	20.747	0.238
$g_r W_r g_r$	7	13.398	0.063
$g_r W_{ur} g_r - g_{ur} W_{ur} g_{ur}$	17	22.777	0.157
$g_r W_{ur} g_r - g_{ur} W_{ur} g_{ur}$	7	14.212	0.047

Heteroskedasticity-consistent standard errors are in parentheses. The country returns are calculated in U.S. dollars in excess of the holding period return on the Treasury bill that is closest to 30 days to maturity. The equity data are from Morgan Stanley Capital International.

## 5. Conclusions

With increasingly integrated world capital markets, it is a mistake to ignore the affect of global factors in risk management. This paper explores the use of a simple version of the Sharpe (1964) and Lintner (1965) capital asset pricing model in an international setting. This model posits the existence of a single factor – the value weighted world market portfolio. Each country has a risk exposure to the world portfolio and differential expected returns across countries are determined by the relative risk exposures.

The primary implication of this model is the mean-variance efficiency of this world market portfolio. Our proxy for the world market portfolio is the Morgan Stanley Capital International world equity portfolio. Our goal is to test its efficiency and to assess the sensitivity of these tests to the distributional specifications.

Our empirical work shows that the efficiency of the MSCI world index cannot be rejected at standard significance levels when all 17 country portfolios are used in the tests. When the tests are specialized to the group of seven countries, the p-values for the tests drop. However, this is not unexpected. The strongest evidence against the model comes from the Japanese equity portfolio. The expected return in Japan (January 1970–May 1989) is too high compared to the measured risk. In the test with seven countries, more relative weight is given to the Japanese portfolio and which leads to lower p-values. However, it should be noted that if recent data are used, the size of Japan's intercept is greatly reduced.

Our investigation of the model residuals finds evidence of heteroskedasticity and significant departures from normality in the form of skewness and kurtosis. We adjust the multivariate tests of efficiency to account for alternative distributional specifications: multivariate  $t$  and multivariate mixture normal. Exact tests are presented. Although the p-values are generally lower, the basic inference is unchanged.

Finally, we drop all parametric distributional assumptions. We present tests based on the generalized method of moments that allow the data to be non-i.i.d and allows for general forms of conditional heteroskedasticity. The p-values of the multivariate tests are substantially reduced. However, the main conclusion is unaffected.

Our paper suggests that it is difficult to reject the efficiency of the MSCI world market portfolio. Even after extensive sensitivity analysis, involving alternative distributional specifications (multivariate  $t$ , mixture normal, nonparametric versus multivariate normal), we cannot reject the efficiency of this portfolio. As such, our paper suggests that measuring the risk of securities by the covariance with the MSCI portfolio is a meaningful first step in global risk management.

## References

- Adler, Michael and Bernard Dumas, 1983, International portfolio selection and corporation finance: A synthesis, *Journal of Finance* 38, 925–984.
- Affleck-Graves, John and Bill McDonald, 1989, Nonnormalities and tests of asset pricing theories, *Journal of Finance* 44, 889–908.
- Andrews, Donald W.K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–858.
- Bekaert, Geert and Robert Hodrick, 1992, Characterizing predictable components in excess returns on equity and foreign exchange markets, *Journal of Finance* 47, 467–510.
- Campbell, John Y. and Yasushi Hamao, 1992, Predictable bond and stock returns in the United States and Japan: A study of long-term capital market integration, *Journal of Finance* 47, 43–70.
- Chan, K.C., G. Andrew Karolyi and René M. Stulz, 1992, Global financial markets and the risk premium on U.S. equity, *Journal of Financial Economics*, forthcoming.
- Chang, Eric C., J. Michael Pinegar and R. Ravichandran, 1991, Latent variables tests of the integration of European equity markets, Working paper, University of Maryland.
- Cumby, Robert E. and Jack D. Glen, 1990, Evaluating the performance of international mutual funds, *Journal of Finance* 45, 497–521.
- Eichenbaum, Martin S. and Lars P. Hansen, 1990, Estimating models with intertemporal substitution using aggregate time series data, *Journal of Business and Economic Statistics* 8, 53–69.
- Eichenbaum, Martin S., Lars P. Hansen, and Kenneth J. Singleton, 1988, A time series analysis of representative agent models of consumption and leisure choice under uncertainty, *Quarterly Journal of Economics* 103, 51–78.
- Fama, Eugene F. 1984, The information in the term structure, *Journal of Financial Economics* 13, 509–528.
- Fedenia, Mark, James E. Hodder and Alexander J. Triantis, 1991, Cross-holding and market return measures, Graduate School of Business, Stanford University.
- Ferson, Wayne E. and Campbell R. Harvey, 1992, The risk and predictability of international equity returns, Working paper, University of Chicago.
- French, Kenneth R. and James M. Poterba, 1991, Were Japanese stock prices too high? *Journal of Financial Economics* 29, 337–363.
- Gallant, A. Ronald and Dale W. Jorgenson, 1979, Statistical inference for a system of simultaneous, nonlinear, implicit equations in the context of instrumental variables estimation, *Journal of Econometrics* 11, 275–302.
- Gibbons, Michael R. and Wayne E. Ferson, 1985, Tests of asset pricing models with changing expectations and an unobservable market portfolio, *Journal of Financial Economics* 14, 217–236.
- Gibbons, Michael R., Stephen A. Ross and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.
- Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.
- Hansen, Lars P. and Robert J. Hodrick, 1983, Risk averse speculation in forward foreign exchange markets: An econometric analysis of linear models, in: Jacob A. Frenkel, ed., *Exchange rates and international macroeconomics* (University of Chicago Press, Chicago, IL).
- Harvey, Campbell R., 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289–317.
- Harvey, Campbell R., 1991, The world price of covariance risk, *Journal of Finance* 46, 111–157.
- Harvey, Campbell R. and Guofu Zhou, 1990, Bayesian inference in asset pricing tests, *Journal of Financial Economics* 26, 221–254.
- Harvey, Campbell R. and Guofu Zhou, 1992, A Bayesian approach to international asset pricing, Working paper, Duke University.
- Harvey, Campbell R., Bruno Solnik and Guofu Zhou, 1992, What determines expected international asset returns, Working paper, Duke University.
- Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.



- Ljung, G.M. and George E.P. Box, 1978, On a measure of lack of fit in time series models, *Biometrika* 65, 297–303.
- MacKinlay, A. Craig and Matthew P. Richardson, 1991, Using generalized method of moments to test mean-variance efficiency, *Journal of Finance* 46, 511–527.
- McDonald, Jack, 1989, The Mochiai Effect: Japanese Corporate Cross Holdings, *Journal of Portfolio Management*, 90–94.
- Newey, Whitney K. and Kenneth D. West, 1987a, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Newey, Whitney K. and Kenneth D. West, 1987b, Hypothesis testing with efficient method of moments estimation, *International Economic Review* 28, 777–787.
- Pearson, E.S., 1930, A further development of tests for normality, *Biometrika* 22, 239.
- Richardson, Matthew and Tom Smith, 1991, Tests of Financial Models in the presence of overlapping observations, *Review of Financial Studies* 4, 227–254.
- Richardson, Matthew and Tom Smith, 1992, A test for multivariate normality in stock returns, Working paper, Duke University, Durham, NC.
- Roll, Richard W., 1977, A critique of the asset pricing theory's tests; Part 1: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129–176.
- Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics* 14, 327–348.
- Shanken, Jay, 1986, Testing portfolio efficiency when the zero-beta rate is unknown: A note, *Journal of Finance* 41, 269–276.
- Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.
- Solnik, Bruno, 1974, An equilibrium model of the international capital market, *Journal of Economic Theory* 8, 500–524.
- Solnik, Bruno, 1977, Testing international asset pricing: Some pessimistic views, *Journal of Finance* 32, 503–511.
- Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.
- Stulz, René, 1981, A model of international asset pricing, *Journal of Financial Economics* 9, 383–406.
- Stulz, René, 1983, Pricing capital assets in an international setting: An introduction, *Journal of International Business Studies*, Winter, 55–73.
- Stulz, René, 1992, International portfolio choice and asset pricing: An integrative survey, Working paper, Ohio State University.
- Zhou, Guofu, 1992, Asset pricing tests under alternative distributions, Working paper, Washington University, St. Louis, MO.