Multivariate Asset Pricing Tests: Insights from Spherical Distribution Theory[1]

Wentworth Boynton[a,2], Jiajuan Liang[b], Man-Lai Tang[c]

[a,b] *College of Business, University of New Haven*

300 Boston Post Road, West Haven, CT 06516

[c] *Department of Mathematics, Hong Kong Baptist University*

Kowloon Tong, Hong Kong, China

[1]This research was partially supported by a research grant from Hong Kong Baptist University, and a University of New Haven 2009 Summer Research Fund.

[2]Corresponding author: E-mail: cboynton@newhaven.edu

Monte Carlo Study

The model for the Monte Carlo study in this section is from (MacKinlay, 1987):

$$\boldsymbol{z}_t = \boldsymbol{\alpha} + z_{mt}\boldsymbol{\beta} + \boldsymbol{e}_t, \qquad t = 1, \ldots, T, \tag{1}$$

where

1) $\boldsymbol{z}_t = (z_{1t}, ..., z_{Nt})'$: the $N \times 1$ vector of excess asset returns for time period t for $t = 1, \ldots, T$, T stands for the number of time periods, N stands for the number of assets;
2) z_{mt}: the excess portfolio return for the portfolio m at time period t;
3) $\boldsymbol{e}_t$: the $N \times 1$ vector of disturbance, $\boldsymbol{e}_t$ is assumed to be contemporaneously cross-correlated $\text{cov}(\boldsymbol{e}_t, \boldsymbol{e}_s) = \delta_{ts}\boldsymbol{\Sigma}$ with $\delta_{ts} = 1$ if $t = s$ and $\delta_{ts} = 0$ if $t \neq s$, and $\boldsymbol{e}_t$ has an N-dimensional normal distribution $N_N(\mathbf{0}, \boldsymbol{\Sigma})$;
4) $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$: $N \times 1$ vector of parameters;
5) $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)'$: $N \times 1$ vector of regression slopes

The hypothesis for testing the efficiency for portfolio m is:

$$H_0: \ \boldsymbol{\alpha} = \mathbf{0} \quad \text{versus} \quad H_1: \ \boldsymbol{\alpha} \neq \mathbf{0}. \tag{2}$$

In order to simulate the excess asset return data $\{\boldsymbol{z}_t = (z_{1t}, ..., z_{Nt})' : \ t = 1, \ldots, T\}$ from model (1), we specify the model parameters as follows.

1) Take the excess portfolio return data $\{z_{mt} : \ t = 1, \ldots, T\}$ as excess portfolio returns from January 1954 to December 1958 (column vw mar in the data set JLMZ2.xls from CRSP). These are 5-year-period data with $T = 60$ months, $N = 20$ and $N = 40$ assets, respectively;
2) The population excess asset return data $\{\boldsymbol{z}_t = (z_{1t}, ..., z_{Nt})' : \ t = 1, \ldots, T = 60\}$ are taken from the CRSP data set JLMZ2.xls with $N = 20$ and $N = 40$ assets, respectively;

3) The population regression slopes $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$ and the covariance matrix $\boldsymbol{\Sigma}$ for the disturbance vector are estimated from the maximum likelihood estimation: let

$$\begin{aligned}
\boldsymbol{X}' &= \begin{pmatrix} 1 & 1 & \ldots & 1 \\ z_{m1} & \boldsymbol{z}_{m2} & \ldots & z_{mT} \end{pmatrix} : \ 2 \times T, \ T = 60, \\
\boldsymbol{B}' &= [\boldsymbol{\alpha} \ \boldsymbol{\beta}] : \ N \times 2, \\
\boldsymbol{Z}' &= [\boldsymbol{z}_1 \ \boldsymbol{z}_2 \ \ldots \ \boldsymbol{z}_T] : \ N \times T.
\end{aligned} \tag{3}$$

Then the maximum likelihood estimates for $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ can be obtained from

$$\begin{aligned}
\hat{\boldsymbol{B}} &= [\hat{\boldsymbol{\alpha}} \ \hat{\boldsymbol{\beta}}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}, \\
\hat{\boldsymbol{\Sigma}} &= (T-2)^{-1}(\boldsymbol{Z} - \boldsymbol{X}\hat{\boldsymbol{B}})'(\boldsymbol{Z} - \boldsymbol{X}\hat{\boldsymbol{B}}).
\end{aligned} \tag{4}$$

4) $\boldsymbol{\alpha} = c\mathbf{1}_N$ is chosen as along the direction $\mathbf{1}_N = (1, \ldots, 1)'$: $N \times 1$ vector of ones, where the constant c controls how far $\boldsymbol{\alpha}$ differs from the origin $\alpha = \mathbf{0}$. $c = 0$ implies that the null hypothesis in (2) is true or the portfolio is efficient.

With the model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ given by (4), and the excess portfolio returns $\{z_{mt} : \ t = 1, \ldots, T\}$ as specified above, we can generate the empirical excess asset returns $\{\boldsymbol{z}_t = (z_{1t}, ..., z_{Nt})' : \ t = 1, \ldots, T\}$ from model (1).

The existing multivariate Hotelling T^2-type test was given by Shanken (1989):

$$\theta_1 = \frac{(T-N-1)T}{(T-2)N} \left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)^{-1} \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}, \tag{5}$$

where $\hat{\mu}_m$ and $\hat{\sigma}_m$ stand for the sample mean and the sample standard deviation from the data $\{z_{mt} : \ t = 1, \ldots, T\}$. Under the null hypothesis in (2), θ_1 has an F-distribution $F(N, T-N-1)$. Under the alternative hypothesis in (2), θ_1 has a noncentral F-distribution $F(N, T-N-1; \lambda)$ with the noncentral parameter

$$\lambda = T \left(1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right)^{-1} \boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}. \tag{6}$$

Note that the Hotelling T^2-type test θ_1 is applicable only if the observation time periods is greater than the number of assets N, i.e., $T > N$.

The purpose of our new test is to avoid the condition of $T > N$. Define the risk-adjusted returns

$$\boldsymbol{m}_t = \boldsymbol{z}_t - z_{mt}\boldsymbol{\beta}, \qquad t = 1, \ldots, T, \tag{7}$$

and consider the regression model

$$\boldsymbol{m}_t = \boldsymbol{\alpha} + \boldsymbol{e}_t, \qquad t = 1, \ldots, T, \tag{8}$$

where the random disturbance vectors $\{\boldsymbol{e}_t : \ t = 1, \ldots, T\}$ are assumed to have the same property as in model (1). We assume that the risk-adjusted return sampled data $\{\boldsymbol{m}_t : \ t = 1, \ldots, T\}$ can be observed from the market without being estimated by $\hat{\boldsymbol{m}}_t = \boldsymbol{z}_t - z_{mt}\hat{\boldsymbol{\beta}}$. That is, we assume that $\{\boldsymbol{m}_t : \ t = 1, \ldots, T\}$ is an i.i.d. sample available from the market. Let $\boldsymbol{Y}' = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_T]$: $N \times T$, and define the eigenvalue-eigenvector problem

$$\frac{1}{T}(\boldsymbol{Y}'\boldsymbol{Y})\boldsymbol{d}_i = \lambda_i \boldsymbol{d}_i, \quad \boldsymbol{u}_i = \boldsymbol{Y}\boldsymbol{d}_i, \ i = 1, \ldots, r, \tag{9}$$

where $\lambda_1 \geq \ldots \geq \lambda_r$ are the r positive eigenvalues of $\boldsymbol{Y}'\boldsymbol{Y}/T$ and $\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_r\}$ are the associated eigenvectors. For any $1 \leq q \leq r$, define the matrices

$$\begin{aligned} &\boldsymbol{U}_q = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_q) : \quad T \times q, \\ &\boldsymbol{H}_q = \boldsymbol{U}_q'\left(\frac{1}{T}\mathbf{1}_T\mathbf{1}_T'\right)\boldsymbol{U}_q, \quad \boldsymbol{G}_q = \boldsymbol{U}_q'\left(\boldsymbol{I}_T - \frac{1}{T}\mathbf{1}_T\mathbf{1}_T'\right)\boldsymbol{U}_q, \end{aligned} \tag{10}$$

where $\boldsymbol{I}_T$ stands for the $T \times T$ identity matrix.

Our new test for hypothesis (2) is given by

$$LF_q = \frac{T-q}{Tq}\mathrm{tr}(\boldsymbol{H}_q\boldsymbol{G}_q^{-1}) = \frac{T-q}{Tq}\left(\mathbf{1}_T'\boldsymbol{U}_q\boldsymbol{G}_q^{-1}\boldsymbol{U}_q'\mathbf{1}_T\right), \quad q = 1, \ldots, r. \tag{11}$$

It can be proved that LF_q has a null distribution (H_0 in (2) is true):

$$LF_q \sim F(q, T-q), \quad q = 1, \ldots, r. \tag{12}$$

The non-null distribution of LF_q is still an open problem. The null hypothesis in (2) is rejected for a large value of LF_q. Theoretically, any value of q such that $1 \leq q \leq r$ can give a test for hypothesis (2), where r is the number of positive eigenvalues in the eigenvalue-eigenvector problem (9). We will choose for values of q in the Monte Carlo study: 1) $q = 1$; 2) $q = [r/3]$, where $[\cdot]$ stands for the integer part of a real number; 2) $q = [r/2]$; and 4) $q = \min(N, T) - 1$.

Table 1 shows the empirical type error rates for the Hotelling T^2-type statistic θ_1 given by (5) and the LF_q-test given by (12), where the sampled data are generated from model (1) with $\boldsymbol{\alpha} = c\mathbf{1}_N$ ($c = 0$) and the model parameters as previously specified. The number of simulations=2,000. It is shown that the type I error rates for our new tests LF_q with different choices of q are well controlled for both $T > N$ and $T \leq N$.

Figures 1-2 show the power plots for both the Hotelling T^2-type statistic θ_1 given by (5) and the LF_q-test given by (12). Both Figure 1 and Figure 2 show that if $T >> N$ (T is much greater than N), the LF_q-test can still improve a little bit over the θ_1-test by choosing q as the maximum possible value. When $T > N$ and is close to N, the θ_1-test is going to lose power as expected, but the LF_q-test can still perform quite well by choosing $[r/3] \leq q \leq [r/2]$ with $r = \min(T, N) - 1$. When $T \leq N$, the θ_1-test is no longer applicable but LF_q-test can still perform quite well with $[r/3] \leq q \leq [r/2]$. Therefore, when the risk-adjusted returns from the assets are available from the market returns, the new test based on LF_q for portfolio efficiency is generally superior to the existing multivariate test based on the Hotelling T^2-type statistic θ_1.

Table 1. Empirical type I error rates

Test	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
	$T = 60,\ N = 20$			$T = 60,\ N = 40$		
θ_1	.0130	.0480	.0980	.0140	.0510	.0995
$LF_q,\ q = 1$	.0120	.0515	.1025	.0115	.0565	.1130
$LF_q,\ q = [r/3]$	.0090	.0475	.1060	.0095	.0545	.1070
$LF_q,\ q = [r/2]$	.0100	.0510	.0980	.0100	.0535	.1015
$LF_q,\ q = r$	.0080	.0510	.1025	.0125	.0485	.1050
	$T = 24,\ N = 20$			$T = 42,\ N = 40$		
θ_1	.0125	.0470	.0980	.0080	.0520	.0930
$LF_q,\ q = 1$	.0125	.0605	.1095	.0110	.0535	.0950
$LF_q,\ q = [r/3]$	.0100	.0570	.1065	.0100	.0515	.0900
$LF_q,\ q = [r/2]$	.0115	.0530	.1010	.0105	.0505	.0940
$LF_q,\ q = r$	.0110	.0485	.1000	.0095	.0495	.0940
	$T = 20,\ N = 20$			$T = 40,\ N = 40$		
$LF_q,\ q = 1$	.0070	.0535	.1125	.0095	.0440	.0905
$LF_q,\ q = [r/3]$	.0140	.0550	.1075	.0100	.0535	.1055
$LF_q,\ q = [r/2]$	.0130	.0520	.1015	.0120	.0525	.0990
$LF_q,\ q = r$	.0115	.0480	.0935	.0155	.0550	.0945
	$T = 15,\ N = 20$			$T = 20,\ N = 40$		
$LF_q,\ q = 1$	.0100	.0505	.1020	.0110	.0455	.0915
$LF_q,\ q = [r/3]$	.0085	.0550	.0975	.0075	.0520	.1105
$LF_q,\ q = [r/2]$	.0085	.0450	.0950	.0125	.0430	.0920
$LF_q,\ q = r$	.0060	.0400	.0945	.0135	.0595	.1090

Note: In Table 1, $r = \min(T, N) - 1$ for each case.

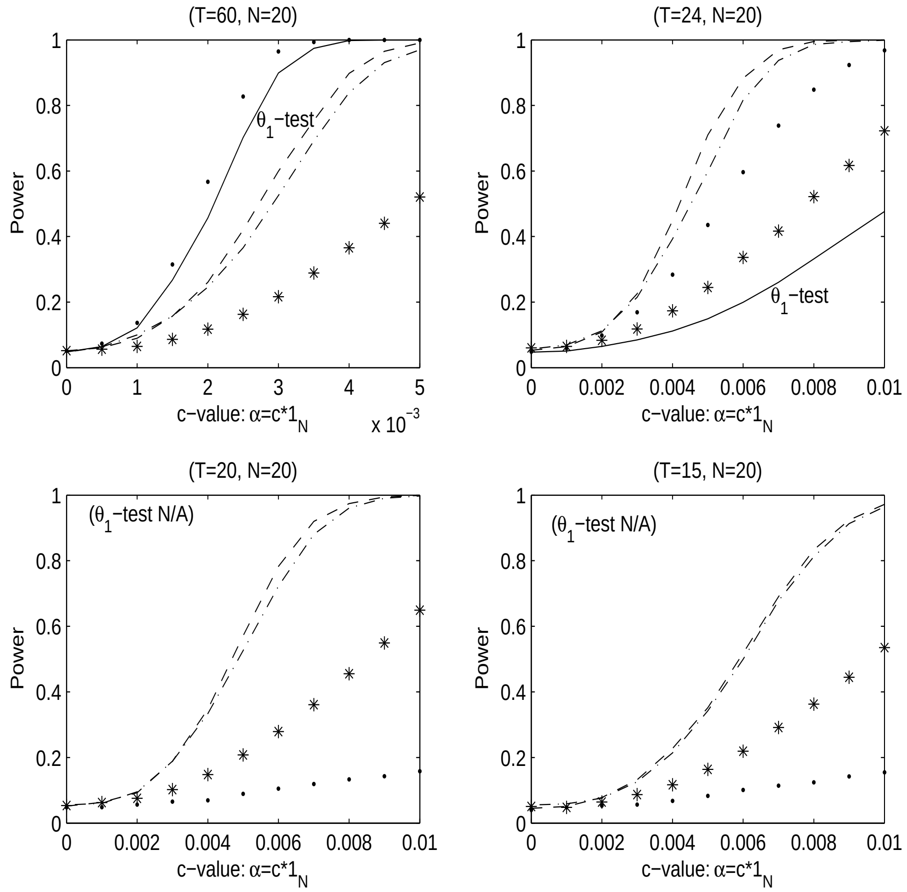

Fig. 1. Power comparison (level $\alpha = 5\%$) between the θ_1-test and the LF_q-test. In each plot, the real line stands for θ_1's power, the "*"-line for LF_1-test, the dashed line for $LF_{[r/2]}$-test, the dashdotted line for $LF_{[r/3]}$-test, the dotted line for LF_r-test, $r = \min(T, N) - 1$.

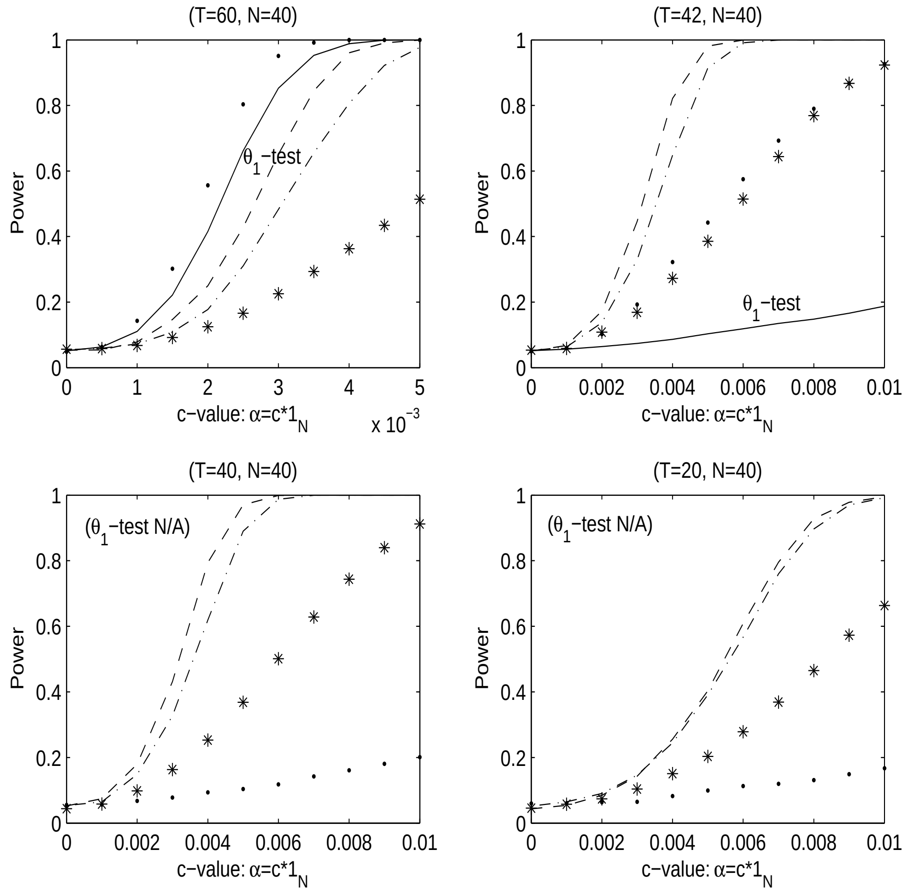

Fig. 2. Power comparison (level $\alpha = 5\%$) between the θ_1-test and the LF_q-test. In each plot, the real line stands for θ_1's power, the "*"-line for LF_1-test, the dashed line for $LF_{[r/2]}$-test, the dashdotted line for $LF_{[r/3]}$-test, the dotted line for LF_r-test, $r = \min(T, N) - 1$.