THE EFFECTS OF NONNORMALITY ON THE MARKET MODEL IN THE CLASS OF ELLIPTICAL DISTRIBUTIONS *

Jiro Hodoshima[†]
Nagoya City University, Nagoya 467-8501, Japan

First Version: October 2000 This Version: March, 2001

Abstract

This paper considers the effects of nonnormality on the statistical inference of the market model when the joint distribution of asset returns and a given portfolio return is elliptically distributed. The standard covariance matrices derived under normality for the least squares estimators of α and β are shown to overstate the accuracy of the least squares estimators whenever the joint distribution is in fact nonnormal in the class of elliptical distributions satisfying the representation of scale mixtures of multivariate normal distributions. Therefore, non-robustified tests of mean-variance efficiency, constructed with the standard covariance matrix derived under normality, result in rejecting the mean-variance efficiency hypothesis too often whenever the joint distribution is in fact nonnormal in the same class.

---

[*] I would like to thank Goufu Zhou for valuable comments and suggestions and Takeaki Kariya for helpful discussion and encouragement. An earlier version of the paper was presented at the 4th Columbia-JAFEE conference held at Tokyo in December 2000. This research started while I was visiting CORE at Universite Catholique de Louvain in 1999. I would like to thank the Japan Society for the Promotion of Science for making possible my visit to CORE. Financial support from Japan Securities Research Foundation is gratefully acknowledged for this research.

[†] Faculty of Economics, Nagoya City University, 1 Yamanohata, Mizuho-cho, Mizuho-ku, Nagoya 467-8501, Japan. Tel: +81-52-872-5715 Fax: +81-52-871-9429 Email:hodoshi@econ.nagoya-cu.ac.jp

1 INTRODUCTION

Testing mean-variance efficiency of a given portfolio in the market model is an important topic in financial econometrics. Many tests have been proposed; Gibbons (1982), Gibbons, Ross, and Shanken (GRS) (1989), Jobson and Korkie (1982), Kandel and Stambaugh (1987), MacKinlay (1987), and Shanken (1985, 1987) (see also Section 5.3 of Campbell, Lo, and MacKinlay (1997)) among others. The exact test by GRS (1989) constructed based on the least squares estimator (LSE) appears to be most popular among them. Those tests are derived assuming normality of the error term. The sensitivity of the tests to departures from the normality assumption has been studied by Affleck-Graves and McDonald (1989), MacKinlay and Richardson (MR) (1991), and Zhou (1993) among others. Affleck-Graves and McDonald (1989) investigate the effects of nonnormality in the market model error term on the exact test by GRS (1989) based on the simulation techniques and find the GRS test not sensitive to minor deviations from normality while sizably affected by higher levels of nonnormality. MR (1991) consider the case when a given portfolio return is also random and show that the tests of mean-variance efficiency derived under normality such as the exact test by GRS (1989) may over-reject the efficiency hypothesis too often when the joint distribution of asset returns to be explained and a given portfolio return is in fact nonnormal, by considering the asymptotic distribution of the GRS test statistic when the joint distribution is multivariate Student t distribution. It is shown in MR (1991) that the tests derived under the normality assumption such as the exact test by GRS (1989) may not possess the correct size since the standard covariance matrix of the LSE may not be valid when the joint distribution of the asset returns to be explained and a given portfolio return is not normally distributed, or equivalently as shown below, when the joint distribution of the two random vectors, the market model error term and a given portfolio return, is not normally distributed. The assumption of a given portfolio return being random seems appropriate since a portfolio return is usually treated as a random variable in its univariate models. Later, Zhou (1993) investigates the GRS test and the *exact*[1] test derived under alternative elliptical distributions for both the case when only the market model error term is elliptically distributed and the case when the joint distribution of the asset returns to be explained and a given portfolio return is elliptically distributed. The GRS test is found robust against departures from the normality assumption in the marginal distribution of the market model error term while not robust against departures from the normality assumption in the joint distribution of the asset returns to be explained and a given portfolio return. In Zhou (1993), based on the U.S. stock market data, the mean-variance efficiency is rejected by the GRS test derived under the normality assumption but no longer rejected by the *exact* test derived under elliptical distributions of multivariate Student t distribution and contaminated multivariate normal distribution. However, the findings on the efficiency test by Zhou (1993) are empirical, and the over-rejection property of the GRS test derived under normality when the joint distribution of the asset returns to be explained and a given portfolio return is in fact other than normal remains to be formally proved. It is still not clear when the nonrobustness of the efficiency tests derived assuming normality arises or whether the non-robustified tests such as the GRS test result in rejecting the efficiency hypothesis too often or too less when the underlying joint distribution of the asset returns to be explained and a given portfolio return is other than multivariate Student t distribution. In this paper, we intend to answer these questions when the underlying joint distribution is elliptically distributed.

Elliptical distributions are known in the financial economics literature to be the class of distributions under which mean-variance analysis is consistent with expected utility maximization (see, e.g., Chamberlain (1983), Ingersoll (1987), and Owen and Rabinovitch (1983)). The class of elliptical distributions contains multivariate normal distribution as well as many heavy tail distributions such as multivariate Cauchy distribution, multivariate Student t distribution, and scale mixture of multivariate normal distributions. It is now well known that

---

[1] The exact test under alternative elliptical distributions in Zhou (1993) is not, strictly speaking, *exact* since parameters of the marginal distribution of a given portfolio return need to be estimated and the numerical approximation is required.

heavy tail distributions such as multivariate Student distribution and contaminated multivariate normal distribution can describe the stock market data better than multivariate normal distribution. Elliptical distributions are thus an important class of distributions both in financial economics and in financial econometrics. Some important properties of the class of elliptical distributions have already been introduced and utilized by Zhou (1993) in the context of testing the mean-variance efficiency. However, we consider that the properties of the class of elliptical distributions appear not yet to have been fully introduced in the financial econometrics literature. In this paper, we make use of a result due to Chu (1973) on the conditional covariance matrix in the class of elliptical distributions in order to see the effects of nonnormality on the inference of the market model.

In this paper, we consider the market model when a given portfolio return is random and the joint distribution of the asset returns to be explained and a given portfolio return is in the class of elliptical distributions, under which the inference of the market model is known to be sensitive to departures from the normality assumption. We first derive the asymptotic distribution of the LSE for the market model in the class of elliptical distributions. We show that the asymptotic covariance matrix of the LSE can be simplified, making use of the property of the class of elliptical distributions. Then we focus on the asymptotic covariance matrix of the LSE, using the general formula of the conditional covariance matrix in the class of elliptical distributions. In a restricted class of elliptical distributions, i.e., the class of scale mixtures of multivariate normal distributions, we show that the asymptotic covariance matrix of the LSE under the nonnormal distributions is larger in the sense of positive semidefiniteness than that under the normal distribution and that the non-robustified tests of mean-variance efficiency, constructed with the covariance matrix derived under the normal distribution, reject the efficiency hypothesis too often asymptotically under the nonnormal distributions. Then, we illustrate the asymptotic bias of the non-robustified tests when the underlying distribution is multivariate Student t distribution, contaminated multivariate normal distribution, and Laplace distribution. The bias obtained by MR (1991) when the underlying distribution is multivariate Student t distribution is verified by our approach.

The paper is organized as follows. Section 2 presents the market model and introduces pertinent properties of the class of elliptical distributions. Section 3 derives the asymptotic distribution of the LSE, proves the properties of the asymptotic covariance matrix of the LSE, and illustrates the asymptotic bias of the non-robustified tests with examples. Section 4 concludes the paper.

2 MODEL

In this section, we first define the market model and then introduce the class of elliptical distributions. Let us assume that investors can borrow and lend at a riskfree rate of return, which is known for investors and hence not a random variable. Then, if a given portfolio p is mean-variance efficient, we have

$$E[R_{it}] = \beta_i E[R_{pt}] \qquad i = 1, \ldots, N; t = 1, \ldots, T \tag{1}$$

where E denotes expectation and

$$\begin{aligned} R_{it} &= time\ period\ t\ excess\ return\ on\ asset\ i \\ \beta_i &= \frac{cov[R_{it}, R_{pt}]}{V[R_{pt}]} \\ R_{pt} &= time\ period\ t\ excess\ return\ on\ portfolio\ p. \end{aligned}$$

where cov and V denote covariance and variance respectively and excess returns are defined as returns minus the risk free rate of return. Then we have the following market model regression equations;

$$R_{it} = \alpha_i + \beta_i R_{pt} + \epsilon_{it} \qquad i = 1, \ldots, N; t = 1, \ldots, T \tag{2}$$

where ϵ_{it} is an error term for asset i in period t and $\alpha_i = 0$. $E[\epsilon_{it}] = E[\epsilon_{it} R_{pt}] = 0$ holds. Hence we have

$$cov[\epsilon_{it}, R_{pt}] = 0 (i = 1, \cdots, N, t = 1, \cdots, T). \tag{3}$$

We consider, as the testing hypothesis of mean-variance efficiency of the portfolio p, the following testing hypothesis given by

$$H_0 : \alpha_i = 0 (i = 1, \ldots, N) \qquad versus \qquad H_1 : \alpha_i \neq 0 (i = 1, \ldots, N). \tag{4}$$

This testing problem has been studied by many authors, assuming normality as the underlying distribution of the error term (see, for example, Section 5.3 of Campbell, Lo, and MacKinlay (1997)). Although we only consider the market model in this paper, the regression model (2) is nothing but a simple regression model, and the results obtained in this paper remain to be valid and relevant in the general regression context if the corresponding stochastic assumptions we make in this paper are satisfied.

In MR (1991), the market model error term $\boldsymbol{\epsilon_t} = (\epsilon_{1t}, \ldots, \epsilon_{Nt})'$ and the portfolio excess return R_{pt} are assumed to be identically and independently distributed (i.i.d.) and to follow joint multivariate Student t distribution in order to find the effects of nonnormality on the asymptotic covariance matrix of the LSE $\hat{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$. In this paper, we intend to study the robustness problem of the tests derived under normality more systematically in the class of elliptical distributions and assume the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} to be i.i.d. and to follow some elliptical distribution, which is, as shown below, equivalent to the same assumptions for the joint distribution of $\boldsymbol{R_t} = (R_{1t}, \ldots, R_{Nt})'$ and R_{pt}.

Now let us define the class of elliptical distributions (cf., e.g., Devlin, Gnanadesikan, and Kettenring (1976), Fang and Zhang (1990) and Kariya and Sinha (1988)). Let $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote the class of elliptical distributions on R^p with $p \times 1$ location vector $\boldsymbol{\mu}$ and $p \times p$ positive definite scale matrix $\boldsymbol{\Sigma}$. We shall write $X \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ when X is elliptically distributed. A $p \times 1$ random vector X is defined to be elliptically distributed if and only if $A\boldsymbol{\Sigma}^{-1/2}(X - \boldsymbol{\mu})$ has the same distribution as $\boldsymbol{\Sigma}^{-1/2}(X - \boldsymbol{\mu})$ for all orthogonal $p \times p$ matrices A where $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}$. An equivalent definition of $X \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is if and only if the characteristic function is of the form $exp(it'\boldsymbol{\mu})\Lambda(t'\boldsymbol{\Sigma}t)$ for some function Λ. If X has a probability density function (pdf), then its pdf is given by

$$f(x) = c_p|\boldsymbol{\Sigma}|g((x - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(x - \boldsymbol{\mu})) \tag{5}$$

for some function g and some positive constant c_p. If the distribution has mean, then mean is $\boldsymbol{\mu}$. If the distribution has second moments, then covariance matrix is $\gamma\boldsymbol{\Sigma}$ where γ is a positive scalar independent of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and coincides with the variance of the univariate marginal distribution when $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = I_p$ where I_p denotes the $p \times p$ identity matrix.

Let us decompose $X, \boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ as follows; $X = (X_1', X_2')'$, $\boldsymbol{\mu} = \boldsymbol{(\mu_1', \mu_2')'}$, and $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix}$ where X_i and $\boldsymbol{\mu_i}$ denote $p_i \times 1$ subvectors of X and $\boldsymbol{\mu}$ respectively and $\boldsymbol{\Sigma_{ij}}$ denotes $p_i \times p_j$ submatrix of $\boldsymbol{\Sigma}$ $(i, j = 1, 2)$ with $p_1 + p_2 = p$. Then we have the following properties on marginal distributions and conditional distributions in the class of elliptical distributions. The marginal distribution of X_2 as well as the conditional distribution of X_1 given X_2 follow elliptical distributions, i.e., $X_2 \sim \mathcal{P}_E(\boldsymbol{\mu_2, \Sigma_{22}})$ and the conditional distribution of X_1 given $X_2 \sim \mathcal{P}_E(\boldsymbol{\mu_1} + \boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}(X_2 - \boldsymbol{\mu_2}), \boldsymbol{\Sigma_{11.2}})$ where $\boldsymbol{\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}$. Therefore, if the conditional expectation of X_1 given X_2, $E[X_1|X_2]$, exists, it is given by

$$E[X_1|X_2] = \boldsymbol{\mu_1} + \boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}(X_2 - \boldsymbol{\mu_2}). \tag{6}$$

If the conditional covariance matrix of X_1 given X_2, $V[X_1|X_2]$, exists, it is given by

$$V[X_1|X_2] = \delta(X_2)\boldsymbol{\Sigma_{11.2}}, \tag{7}$$

where $\delta(X_2)$ denotes some positive scalar and in general depends on the conditioning variable X_2.

The conditional expectation formula (6) shows that the coefficient of X_2 is given by $\boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}$, the same for any member of the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu, \Sigma})$ even though second moments of variances and covariances differ by a scalar in each member of $\mathcal{P}_E(\boldsymbol{\mu, \Sigma})$. For example, let us consider p-dimensional multivariate normal distribution and multivariate Student t distribution with ν degrees of freedom in the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu, \Sigma})$. Then their pdfs are given respectively by

$$\begin{aligned} f(x|\boldsymbol{\mu, \Sigma}) &= (2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2} exp\{-\frac{1}{2}(x - \boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(x - \boldsymbol{\mu})\} \qquad (8) \\ f(x|\boldsymbol{\mu, \Sigma}) &= \frac{\nu^{\nu/2}\Gamma((\nu + p)/2)|\boldsymbol{\Sigma}|^{-1/2}}{\pi^{p/2}\Gamma(\nu/2)}\{\nu + (x - \boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(x - \boldsymbol{\mu})\}^{-(p+\nu)/2}. \qquad (9) \end{aligned}$$

The coefficient of X_2 in the conditional expectation of X_1 given X_2 is, in both cases, given by $\boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}$ while covariance matrices of multivariate normal distribution and multivariate Student t distribution with ν degrees of freedom are different and given respectively by $\boldsymbol{\Sigma}$ and $\frac{\nu}{\nu-2}\boldsymbol{\Sigma}$ $(\nu > 2)$.

Since (6) implies

$$E[X_1] = (\boldsymbol{\mu_1} - \boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}\mu_2}) + \boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}E[X_2], \tag{10}$$

the conditional expectation formula (6) implies that beta is the same in the market model (1) for any member of the class of elliptical distributions with the same scale matrix $\boldsymbol{\Sigma}$. Similarly, the correlation matrix of X is identical for all members of the class of elliptical distributions with the same scale matrix $\boldsymbol{\Sigma}$ because the form of g in (5) corresponds to the form of γ in the covariance matrix, i.e., $\gamma\boldsymbol{\Sigma}(= V[X])$, and $\gamma's$ cancel in the correlation matrix.

In the market model regression equations (2), the joint distribution of the market model error term $\boldsymbol{\epsilon}$ and the excess portfolio return R_{pt} being elliptically distributed is equivalent to the joint distribution of the N excess asset returns $\boldsymbol{R_t}$ and the excess portfolio return R_{pt} being elliptically distributed, as shown below. The following property of the class of elliptical distributions holds; if $X \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\boldsymbol{a} + \boldsymbol{B}X \sim \mathcal{P}_E(\boldsymbol{a} + \boldsymbol{B\mu}, \boldsymbol{B\Sigma B'})$ where $\boldsymbol{a}$ is a $p \times 1$ vector of constants and $\boldsymbol{B}$ is a $p \times p$ nonsingular matrix of constants. Since

$$\begin{pmatrix} -\boldsymbol{\alpha} \\ 0 \end{pmatrix} + \begin{pmatrix} I_N & -\boldsymbol{\beta} \\ \boldsymbol{0} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{R_t} \\ R_{pt} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon_t} \\ R_{pt} \end{pmatrix},$$

$\begin{pmatrix} \boldsymbol{R_t} \\ R_{pt} \end{pmatrix}$ being in the class of elliptical distributions implies that $\begin{pmatrix} \boldsymbol{\epsilon_t} \\ R_{pt} \end{pmatrix}$ is also in the class of elliptical distributions. Conversely, since the matrix $\begin{pmatrix} I_N & -\boldsymbol{\beta} \\ \boldsymbol{0} & 1 \end{pmatrix}$ is full rank and since

$$\begin{pmatrix} I_N & -\boldsymbol{\beta} \\ \boldsymbol{0} & 1 \end{pmatrix}^{-1} \left\{ \begin{pmatrix} \boldsymbol{\epsilon_t} \\ R_{pt} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\alpha} \\ 0 \end{pmatrix} \right\} = \begin{pmatrix} \boldsymbol{R_t} \\ R_{pt} \end{pmatrix},$$

$\begin{pmatrix} \boldsymbol{\epsilon_t} \\ R_{pt} \end{pmatrix}$ being in the class of elliptical distributions implies that $\begin{pmatrix} \boldsymbol{R_t} \\ R_{pt} \end{pmatrix}$ is also in the class of elliptical distributions. Therefore, the assumption of marginal elliptical distribution for $\boldsymbol{\epsilon_t}$ is justified if the joint distribution of $\boldsymbol{R_t}$ and R_{pt} is elliptically distributed.

The general form of dependence on X_2 in the conditional covariance matrix of X_1 given X_2 in the class of elliptical distributions is provided by Chu (1973). We refer this property as follows;

Lemma 1.

When $X \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and the pdf of X is written as an integral of a set of multivariate normal pdfs given by

$$f(x) = \int_0^\infty w(t)(2\pi)^{-p/2}|t^{-1}\boldsymbol{\Sigma}|^{-1/2}exp\{-\frac{1}{2}(x-\boldsymbol{\mu})'(t^{-1}\boldsymbol{\Sigma})^{-1}(x-\boldsymbol{\mu})\}dt \tag{11}$$

where $w(t)$ denotes a weighting function defined by

$$w(t) = (2\pi)^{p/2}|t^{-1}\boldsymbol{\Sigma}|^{1/2}\mathcal{L}^{-1}[f(u)], \qquad 0 < t < \infty \tag{12}$$

where $\mathcal{L}[\cdot]$ is the Laplace transform operator, $\mathcal{L}^{-1}[\cdot]$ is the inverse operator, and $u = \frac{1}{2}(x-\boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(x-\boldsymbol{\mu})$, then the conditional covariance matrix of X_1 given X_2 is given by

$$V[X_1|X_2] = \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11.2}} \tag{13}$$

where $s = \frac{1}{2}(x_2-\boldsymbol{\mu_2})'\boldsymbol{\Sigma_{22}^{-1}}(x_2-\boldsymbol{\mu_2})$ and $f_2(s)$ is the marginal pdf of X_2.

This property of the conditional covariance matrix in the class of elliptical distributions has an important role to obtain our results, and appears not to have been made use of in the financial econometrics literature to our knowledge. This corresponds to Theorem 5[2] of

---

[2] Actually, Theorem 5 of Chu (1973) was stated as

$$E[(X_1-\mu_1)(X_1-\mu_1)'|X_2] = \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11.2}}$$

under our notation. As can be seen from the proof of this theorem in Chu (1973), however, the correct equation is equation (13), i.e.,

$$V[X_1|X_2] \equiv E[X_1X_1'|X_2] - E[X_1|X_2]E[X_1'|X_2] = \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11.2}},$$

not

$$E[(X_1-\mu_1)(X_1-\mu_1)'|X_2] = \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11.2}}.$$

Chu (1973). $\int_s^\infty f_2(s_2)ds_2/f_2(s)$ in Lemma 1 coincides with the scaling $\delta(X_2)$ in (7), and its general dependence on X_2 can be easily seen from its definition.

A weak sufficient condition for the existence of $\mathcal{L}^{-1}[f(u)]$ is provided by Chu (1973) and given as follows;

Condition 1 *$f(u)$ is differentiable when u is sufficiently large and $f(u) = o(u^{-k})$ for $k > 1$.*

where $f(u) = o(u^{-k})$ denotes $\lim_{u\to\infty} \frac{f(u)}{u^{-k}} = 0$. Therefore, the representation (11) of the pdf as an integral of a set of multivariate normal pdfs seems mostly satisfied in the class of elliptical distributions. We remark that $w(t)$ may take negative values and hence may not be a pdf[3] although $\int_0^\infty w(t)dt = 1$ (cf. Chu (1973)). If $w(t)$ is nonnegative, $w(t)$ is a pdf and the right hand side of (11) becomes a scale mixture of multivariate normal distributions. The necessary and sufficient condition for $w(t)$ to be a pdf in (11) is given by Andrews and Mallows (1974) as follows;

Condition 2 *The pdf f possesses the derivatives of all orders and satisfies*

$$\left(-\frac{d}{dx}\right)^k f\left(x^{\frac{1}{2}}\right) \geq 0 \qquad for \qquad x > 0.$$

The representation of the pdf of the elliptical distribution as scale mixtures of multivariate normal distributions is more restrictive than Chu's representation (11) where $w(t)$ is not necessarily nonnegative and also plays an important role in this paper as shown below. We remark that the restrictive class of elliptical distributions represented as scale mixtures of multivariate normal distributions still contains infinite numbers of elliptical distributions.

The representation (11) when $w(t)$ is not necessarily a pdf and also a pdf is fundamental to derive the results in this paper.

In addition, the following property of the conditional covariance matrix in the class of elliptical distributions due to Kelker (1970) has an important role in this paper. This property is also introduced in Zhou (1993). We refer this property as follows;

---

[3] An example of $w(t)$ which takes negative values is provided by Chu (1973) and given by a p-dimensional pdf $f(u) = c[1+(2u)^2]^{-1}$ with the corresponding $w(t)$ proportional to $t^{-p/2}\sin\frac{t}{2}$ where c denotes a constant.

Lemma 2.
When $X \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *the conditional covariance matrix of* X_1 *given* X_2 *is independent of* X_2 *only if* X *has normal distribution.*

This is Theorem 7 of Kelker (1970). Since the conditional covariance matrix of X_1 given X_2 is, as is well known, independent of X_2 if X is normal, Lemma 2 implies that independence of X_2 in the conditional covariance matrix of X_1 given X_2 is equivalent to normality of X in the class of elliptical distributions. We remark that this is so only when the joint distribution of X is in the class of elliptical distributions. When X is normal, $\delta(X_2) = 1$ follows from direct calculation of $\int_s^\infty f_2(s_2)ds_2/f_2(s)$. MR (1991) define the contemporaneous conditional homoskedasticity and heteroskedasticity when the conditional covariance matrix of X_1 given X_2 does not and does depend on X_2 respectively, but do not investigate when the contemporaneous conditional heteroskedasticity obtains, besides giving multivariate Student t distribution as an example of the contemporaneous conditional heteroskedasticity. Lemma 2 is relevant to see how restrictive the assumption of the contemporaneous conditional homoskedasticity is for the conditional covariance matrix of the market model error term given a portfolio return, which is the problem taken up by MR (1991). This equivalence property due to Kelker (1970) implies that the contemporaneous conditional homoskedasticity of the market model error term given a portfolio return is satisfied only under the joint normality of the market model error term and the given portfolio return in the market model (2). Since the class of elliptical distributions makes mean-variance analysis consistent with the expected utility maximization and contains many heavy tail distributions, this implies that the assumption of contemporaneous conditional homoskedasticity is very strong and restrictive.

3 THE ASYMPTOTIC PROPERTIES OF THE LSE AND THE TESTS DERIVED UNDER NORMLAITY IN THE CLASS OF ELLIPTICAL DISTRIBUTIONS

In this section, we first investigate the covariance matrix of the asymptotic distribution of the LSE $\hat{\boldsymbol{\alpha}}$ and also the LSE $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)'$ when the joint distribution of the asset returns to be explained and a given portfolio return is in the class of elliptical distributions. We make use of the two lemmas given in Section 2 to see the effects of nonnormality on the asymptotic covariance matrix of the LSE in the class of elliptical distributions. Then we study the effects of nonnormality on the efficiency tests constructed from the LSE such as the GRS test in the class of elliptical distributions.

We assume that the two random vectors, the market model error term $\boldsymbol{\epsilon_t}$ and the excess return on portfolio p given by R_{pt} (equivalently the N excess returns $\boldsymbol{R_t}$ and the excess portfolio return R_{pt}) are jointly i.i.d., follow some elliptical distribution $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and have the joint pdf f which possesses finite fourth moments and satisfies Condition 1 where a $(N+1) \times 1$ location vector $\boldsymbol{\mu} = (\boldsymbol{0'}, \mu_p)'$ and a positive definite $(N+1) \times (N+1)$ scale matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_{22} \end{pmatrix}$ where the $N \times 1$ subvector $\boldsymbol{0}$ of $\boldsymbol{\mu}$ denotes expectation of the market model error term, μ_p denotes expectation of the given excess portfolio return, and the scale matrix $\boldsymbol{\Sigma}$ is similarly decomposed as in $\boldsymbol{\mu}$. $\boldsymbol{\Sigma}$ is block-diagonal because $cov[\epsilon_{it}, R_{pt}] = 0 (i = 1, \ldots, N)$ as given in (3). As shown in Section 2, $(\boldsymbol{\epsilon_t'}, R_{pt})' \sim \mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ does not necessarily imply $V[\boldsymbol{\epsilon_t}] = \boldsymbol{\Sigma_{11}}$ or $V[R_{pt}] = \sigma_{22}$ when the second moments of $(\boldsymbol{\epsilon_t'}, R_{pt})'$ exist. Therefore, we shall denote $V[\boldsymbol{\epsilon_t}]$ and $V[R_{pt}]$ as the covariance matrix of $\boldsymbol{\epsilon_t}$ and the variance of R_{pt} respectively.

The LSE of $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')'$ is obtained by solving the moment conditions $\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\epsilon_t} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\epsilon_t} R_{pt} = \boldsymbol{0}$. Thus the LSE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is also the GMME and corresponds to the quasi maximum likelihood estimator (QMLE) of $\boldsymbol{\theta}$ under the normality of $\boldsymbol{\epsilon_t}$. Under our assumptions, the asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ can be shown by the standard method to be normal with mean $\boldsymbol{0}$ and covariance matrix given by

$$\boldsymbol{H^{-1} B H^{-1}} \tag{14}$$

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix} \bigotimes \boldsymbol{I_N}$$

and

$$\boldsymbol{B} = \begin{pmatrix} E[\boldsymbol{\epsilon_t \epsilon_t'}] & E[R_{pt}\boldsymbol{\epsilon_t \epsilon_t'}] \\ E[R_{pt}\boldsymbol{\epsilon_t \epsilon_t'}] & E[R_{pt}^2\boldsymbol{\epsilon_t \epsilon_t'}]@ \end{pmatrix}$$

where $\mu_p = E[R_{pt}], \boldsymbol{I_N}$ is an N dimensional identity matrix, and $\bigotimes$ denotes Kronecker product. When the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is normal, it follows that

$$\boldsymbol{B} = \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix} \bigotimes V[\boldsymbol{\epsilon_t}] \tag{15}$$

and

$$\boldsymbol{H^{-1}BH^{-1}} = \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix}^{-1} \bigotimes V[\boldsymbol{\epsilon_t}] \qquad (\equiv \boldsymbol{D}) \tag{16}$$

since $\boldsymbol{\epsilon_t}$ and R_{pt} become independent under the assumption of block-diagonal $\boldsymbol{\Sigma}$. $\boldsymbol{D}$ is the probability limit of the standard covariance matrix of the LSE and also that of the GMME and QMLE when the conditional distribution of $\boldsymbol{\epsilon_t}$ (or $\boldsymbol{R_t}$) given R_{pt} is normal (cf., for example, Section 5.3 of Campbell, Lo, and MacKinlay (1997)). On the other hand, when the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is nonnormal in the class of elliptical distributions, the equalities (15) and (16) no longer hold. Since $\boldsymbol{H}$ and $\boldsymbol{D}$ remain the same in the class of elliptical distributions, we need to evaluate $\boldsymbol{B}$ under nonnormality to see the effects of nonnormality in the class of elliptical distributions. When people use the LSE, they often use the standard covariance matrix whose probability limit is given by $\boldsymbol{D}$. Therefore, we compare $\boldsymbol{D}$ and $\boldsymbol{H^{-1}BH^{-1}}$ under nonnormality of $(\boldsymbol{\epsilon_t'}, R_{pt})'$ in the class of elliptical distributions to see whether we overstate or understate the accuracy of the LSE when the covariance matrix asymptotically equivalent to $\boldsymbol{D}$ is used for the inference of the LSE. By the property of the conditional expectation given by $E[Y] = E_Z[E[Y|Z]]$ where E_Z denotes expectation with respect to the marginal distribution of Z, we have

$$\begin{aligned} E[R_{pt}\boldsymbol{\epsilon_t \epsilon_t'}] &= E[(R_{pt} - \mu_p)\boldsymbol{\epsilon_t \epsilon_t'}] + \mu_p E[\boldsymbol{\epsilon_t \epsilon_t'}] \\ &= E[(R_{pt} - \mu_p)\boldsymbol{\epsilon_t \epsilon_t'}] + \mu_p V[\boldsymbol{\epsilon_t}] \end{aligned}$$

and

$$\begin{aligned} E[R_{pt}^2\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}] &= E[(R_{pt}^2-\mu_p^2-V[R_{pt}])\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}]+(\mu_p^2+V[R_{pt}])E[\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}] \\ &= E\left[\left\{(R_{pt}-\mu_p)^2+2\mu_p(R_{pt}-\mu_p)-V[R_{pt}]\right\}\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}\right]+(\mu_p^2+V[R_{pt}])V[\boldsymbol{\epsilon_t}] \\ &= E_{R_{pt}}\left[\left\{(R_{pt}-\mu_p)^2-V[R_{pt}]\right\}E[\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}|R_{pt}]\right]+2\mu_p E[(R_{pt}-\mu_p)\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}] \\ &\quad +(\mu_p^2+V[R_{pt}])V[\boldsymbol{\epsilon_t}]. \end{aligned}$$

Since $\boldsymbol{\Sigma}$ is block-diagonal and $E[\boldsymbol{\epsilon_t}]=\mathbf{0}$, $E[\boldsymbol{\epsilon_t}|R_{pt}]=\mathbf{0}$ by the conditional expectation formula (6). Hence $E[\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}|R_{pt}]$ is the conditional covariance matrix of $\boldsymbol{\epsilon_t}$ given R_{pt}. Therefore, it follows from Lemma 1 that

$$V[\boldsymbol{\epsilon_t}|R_{pt}]=E[\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}|R_{pt}]=\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11}}$$

where $s=\frac{1}{2}(R_{pt}-\mu_p)^2/\sigma_{22}$ and $f_2(s)$ is the marginal pdf of R_{pt}. The scalar $\int_s^\infty f_2(s_2)ds_2/f_2(s)$ indicates the conditional heteroskedasticity in the conditional covariance matrix $V[\boldsymbol{\epsilon_t}|R_{pt}]$, i.e., the conditional covariance matrix of $\boldsymbol{\epsilon_t}$ given R_{pt} being dependent on the conditioning variable R_{pt}. Since $(\boldsymbol{\epsilon_t'},R_{pt})'\sim\mathcal{P}_E(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with $\boldsymbol{\mu}=(\mathbf{0'},\mu_p)'$, $\boldsymbol{\epsilon_t}$ and R_{pt} are symmetrically distributed about $\mathbf{0}$ and μ_p respectively and hence $(R_{pt}-\mu_p)\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}$ is also symmetrically distributed about $\mathbf{0}$. Hence, $E[(R_{pt}-\mu_p)\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}]=\mathbf{0}$[4] , which implies $E[R_{pt}\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}]=\mu_p V[\boldsymbol{\epsilon_t}]$. Therefore we have the following result as the asymptotic distribution of the LSE.

Theorem 1.

When $(\boldsymbol{\epsilon_t'},R_{pt})'$ is i.i.d., follows the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with finite fourth moments, and has the pdf f satisfying Condition 1, then the asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is normal distribution with mean $\mathbf{0}$ and covariance matrix given by

$$\boldsymbol{H^{-1}BH^{-1}}$$

---

[4] The scalar $\int_s^\infty f_2(s_2)ds_2/f_2(s)$ is an even function of $R_{pt}-\mu_p$ since it depends on $R_{pt}-\mu_p$ through $s=\frac{1}{2}(R_{pt}-\mu_p)^2/\sigma_{22}$. Therefore, expectation of $(R_{pt}-\mu_p)\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}$ with respect to R_{pt} is zero when $(\boldsymbol{\epsilon_t'},R_{pt})'$ is in the class of elliptical distributions. The general property that $E_{R_{pt}}\left[(R_{pt}-\mu_p)\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right]$ is zero when $(\boldsymbol{\epsilon_t'},R_{pt})'\sim\mathcal{P}_E(\boldsymbol{\mu},\boldsymbol{\Sigma})$ is not recognized in MR (1991) although they derive it in the case of multivariate Student t distribution. But we do not need this property since $E[(R_{pt}-\mu_p)\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t'}]$ is zero as shown above by the symmetry argument.

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix} \bigotimes \boldsymbol{I_N}$$

and

$$\boldsymbol{B} = \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix} \bigotimes V[\boldsymbol{\epsilon_t}] + \boldsymbol{C}$$

with $\boldsymbol{C} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & E_{R_{pt}}[\{(R_{pt} - \mu_p)^2 - V[R_{pt}]\} \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}]\boldsymbol{\Sigma_{11}} \end{pmatrix}$.

Therefore, we need to evaluate $E_{R_{pt}}\left[\{(R_{pt} - \mu_p)^2 - V[R_{pt}]\} \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right]$ to obtain the asymptotic covariance matrix of the LSE $\hat{\boldsymbol{\theta}}$. This term is a one-dimensional scalar. If this term is positive (negative), then the LSE with its covariance matrix asymptotically equivalent to $\boldsymbol{D}$ in general overstates (understates) the accuracy of the LSE since the matrix $\boldsymbol{H^{-1}BH^{-1}} - \boldsymbol{D}$ is positive (negative) semidefinite. Therefore, evaluation of this term is critically important to assess the effects of nonnormality on the statistical inference of the market model in the class of elliptical distributions. MR (1991) show by direct calculation that this term is positive when the underlying joint distribution of the market model error term and the excess portfolio return is multivariate Student t distribution. In the following, we show that this term is always positive when the underlying joint distribution of the market model error term and the given portfolio return is nonnormal in the class of elliptical distributions, provided that the underlying joint distribution satisfies Condition 2, i.e., satisfies the representation of the elliptical distribution as scale mixtures of multivariate normal distributions.

Theorem 2.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ is i.i.d., follows the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with finite fourth moments, and has the pdf f satisfying Condition 2, then the pdf f of $\boldsymbol{Z} \equiv (\boldsymbol{\epsilon_t'}, R_{pt})'$ is written as a scale mixture of multivariate normal distributions given by

$$f(z) = \int_0^\infty w(t)(2\pi)^{-p/2}|t^{-1}\boldsymbol{\Sigma}|^{-1/2} exp\{-\frac{1}{2}(z - \boldsymbol{\mu})'(t^{-1}\boldsymbol{\Sigma})^{-1}(z - \boldsymbol{\mu})\}dt$$

where $w(t)$ is a pdf and

$$E_{R_{pt}}\left[\left\{(R_{pt} - \mu_p)^2 - V[R_{pt}]\right\} \frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right]$$

$$= \left\{ \int_0^\infty w(t) t^{-2} dt - \left(\int_0^\infty w(t) t^{-1} dt \right)^2 \right\} \sigma_{22}$$

which is positive if Z is nonnormal and zero if Z is normal.

The proof of Theorem 2 is given at Appendix. The asymptotic covariance matrix of the LSE $\hat{\boldsymbol{\theta}}$ is

$$\begin{aligned} \boldsymbol{H^{-1}BH^{-1}} &= \begin{pmatrix} 1 & \mu_p \\ \mu_p & \mu_p^2 + V[R_{pt}] \end{pmatrix}^{-1} \bigotimes V[\boldsymbol{\epsilon_t}] + \boldsymbol{H^{-1}CH^{-1}} \\ &= \boldsymbol{D} + \frac{1}{V[R_{pt}]^2} \begin{pmatrix} \mu_p^2 & -\mu_p \\ -\mu_p & 1 \end{pmatrix} \bigotimes \boldsymbol{F} \end{aligned}$$

where $\boldsymbol{F}$ is given by $E_{R_{pt}} \left[\{ (R_{pt} - \mu_p)^2 - V[R_{pt}] \} \frac{\int_s^\infty f_2(s_2) ds_2}{f_2(s)} \right] \boldsymbol{\Sigma_{11}}$. Since the second matrix on the farthest right hand side of the above equation of the asymptotic covariance matrix of the LSE $\hat{\boldsymbol{\theta}}$ is positive semidefinite by Theorem 2 when the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is nonnormal in the class of elliptical distributions satisfying Condition 2, the LSE $\hat{\boldsymbol{\theta}}$ in general overstates the accuracy of the LSE when the covariance matrix asymptotically equivalent to $\boldsymbol{D}$ is used for the inference of the LSE $\hat{\boldsymbol{\theta}}$. In particular, the submatrices $\frac{\mu_p^2}{V[R_{pt}]^2} \boldsymbol{F}$ and $\frac{1}{V[R_{pt}]^2} \boldsymbol{F}$ corresponding to the LSE $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are both positive definite. Therefore, both the LSE $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ overstate the accuracy asymptotically if the covariance matrices asymptotically equivalent to corresponding submatrices of $\boldsymbol{D}$ are used, when the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is nonnormal in the class of elliptical distributions satisfying Condition 2. This is a generalization of the result of MR (1991) in the case of multivariate Student t distribution. We summarize this as follows;

Corollary 1.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ is i.i.d., follows the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with finite fourth moments, and has the pdf f satisfying Condition 2 (the representation as a scale mixture of multivariate normal distributions), then both the LSE $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ overstate the accuracy asymptotically whenever the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is in fact nonnormal if the covariance matrices asymptotically equivalent to corresponding submatrices of $\boldsymbol{D}$ are used.

Corollary 1 implies the overrejection property of the LSE based tests derived under normality such as the GRS test when the joint distribution of $\boldsymbol{\epsilon_t}$ and R_{pt} is in fact nonnormal, which is also a generalization of MR (1991). We summarize this as follows.

Corollary 2.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ is i.i.d., follows the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with finite fourth moments, and has the pdf f satisfying Condition 2 (the representation as a scale mixture of multivariate normal distributions), the tests which use, as the asymptotic covariance matrix of $\hat{\boldsymbol{\alpha}}$, $\left(1 + \frac{\mu_p^2}{V[R_{pt}]^2}\right) V[\boldsymbol{\epsilon_t}]$, reject the mean-variance efficiency hypothesis too often asymptotically whenever the joint distribution of the $\boldsymbol{\epsilon_t}$ and R_{pt} is in fact nonnormal.

This corollary has an important implication since normality is, in the class of elliptical distributions satisfying the representation as scale mixtures of multivariate normal distributions, the only distribution under which the non-robustified tests are valid. Therefore, the normality assumption is very restrictive, and it is important to conduct the robust tests such as the GMME based test proposed by MR (1991).

Remark.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ is i.i.d., follows the class of elliptical distributions $\mathcal{P}_E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with finite fourth moments, and has the pdf f satisfying Condition 1 but not Condition 2, then the pdf f of $\boldsymbol{Z} \equiv (\boldsymbol{\epsilon_t'}, R_{pt})'$ is given by

$$f(z) = \int_0^\infty w(t)(2\pi)^{-p/2}|t^{-1}\boldsymbol{\Sigma}|^{-1/2} exp\{-\frac{1}{2}(z-\boldsymbol{\mu})'(t^{-1}\boldsymbol{\Sigma})^{-1}(z-\boldsymbol{\mu})\}dt,$$

where $w(t)$ is not a pdf and takes negative values, and

$$\begin{aligned} &E_{R_{pt}}\left[\left\{(R_{pt}-\mu_p)^2 - V[R_{pt}]\right\}\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] \\ &= \left\{\int_0^\infty w(t)t^{-2}dt - \left(\int_0^\infty w(t)t^{-1}dt\right)^2\right\}\sigma_{22}, \end{aligned}$$

which is not necessarily nonnegative. Therefore, in this case, the LSE $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ may or may not overstate the accuracy depending on the sign of $\left\{\int_0^\infty w(t)t^{-2}dt - (\int_0^\infty w(t)t^{-1}dt)^2\right\}$ if the respective standard covariance matrices derived under normality are used. Similarly, the tests, which use $\left(1 + \frac{\mu_p^2}{V[R_{pt}]^2}\right) V[\boldsymbol{\epsilon_t}]$ as the asymptotic covariance matrix $\hat{\boldsymbol{\alpha}}$, reject the *mean-variance efficiency hypothesis too often or too less depending on the sign of* $\left\{\int_0^\infty w(t)t^{-2}dt - (\int_0^\infty w(t)t^{-1}dt)^2\right\}$. *The class of elliptical distributions contains distributions with heavy tails as well as light tails.*

Example 1.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ follows multivariate normal distribution, $\int_s^\infty f_2(s_2)ds_2 = f_2(s)$. Then $E_{R_{pt}}\left[((R_{pt}-\mu_p)^2 - V[R_{pt}])\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] = E_{R_{pt}}[(R_{pt}-\mu_p)^2 - V[R_{pt}]] = 0$. In this case. $w(t)$ is a degenerate probability measure with $w(t) = 1$ when $t = 1$ and 0 otherwise and $\int_0^\infty w(t)t^{-2}dt = (\int_0^\infty w(t)t^{-1}dt)^2$. Hence, eitherway the asymptotic bias of the non-robustified tests is verified to be zero.

Example 2.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ follows multivariate Student t distribution with ν degrees of freedom, $\int_s^\infty f_2(s_2)ds_2/f_2(s) = \frac{1}{\nu-1}\{\nu + (R_{pt}-\mu_p)^2/\sigma_{22}\}$. Since $V[R_{pt}] = \frac{\nu}{\nu-2}\sigma_{22}(\nu > 2)$ and $V[\boldsymbol{\epsilon_t}] = \frac{\nu}{\nu-2}\boldsymbol{\Sigma_{11}}$, $\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\boldsymbol{\Sigma_{11}} = \frac{\nu-2}{\nu}\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\frac{\nu}{\nu-2}\boldsymbol{\Sigma_{11}} = \frac{\nu-2}{\nu}\frac{\nu}{\nu-1}\left\{1 + \frac{1}{(\nu-2)V[R_{pt}]}(R_{pt}-\mu_p)^2\right\}V[\boldsymbol{\epsilon_t}] = \frac{\nu-2}{\nu-1}\left\{1 + \frac{(R_{pt}-\mu_p)^2}{(\nu-2)V[R_{pt}]}\right\}V[\boldsymbol{\epsilon_t}]$, which is the same formula as equation (14) of MR (1991). Hence, in this case we have

$$
\begin{aligned}
&E_{R_{pt}}\left[((R_{pt}-\mu_p)^2 - V[R_{pt}])\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right]\boldsymbol{\Sigma_{11}} \\
&= \frac{\nu-2}{\nu-1}\left\{V[R_{pt}] - V[R_{pt}] + \frac{1}{(\nu-2)V[R_{pt}]}E[R_{pt}-\mu_p]^4 - \frac{1}{\nu-2}V[R_{pt}]\right\}V[\boldsymbol{\epsilon_t}] \\
&= \frac{1}{\nu-1}\left\{\frac{E[R_{pt}-\mu_p]^4}{V[R_{pt}]} - V[R_{pt}]\right\}V[\boldsymbol{\epsilon_t}],
\end{aligned}
$$

which coincides with the right hand side of equation (16) of MR (1991) and is positive definite under our assumptions; $\frac{E[R_{pt}-\mu_p]^4}{V[R_{pt}]} - V[R_{pt}]$ is positive by the Cauchy-Schwarz inequality if R_{pt} is a nondegenerate probability measure. Therefore, the asymptotic covariance matrix of the LSE $\hat{\alpha}$ becomes larger in the sense of positive definiteness than $(1 + \frac{\mu_p^2}{V[R_{pt}]})V[\boldsymbol{\epsilon_t}]$.

The above derivation can be also verified, making use of the representation of multivariate Student t distribution with ν degrees of freedom as a scale mixture of multivariate normal distributions in Theorem 2. The weighting function $w(t)$ for multivariate Student t distribution with ν degrees of freedom is given by

$$
w(t) = \frac{(\frac{\nu}{2})^{(\nu/2)}}{\Gamma(\frac{\nu}{2})}t^{\nu/2-1}exp\left\{-\frac{\nu t}{2}\right\}
$$

where $\Gamma(\cdot)$ denotes a gamma function. With this weighting function, computation of $\int_0^\infty w(t)t^{-2}dt - \left(\int_0^\infty w(t)t^{-1}dt\right)^2$ will lead to the same result obtained above.

Example 3.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ follows a contaminated multivariate normal distribution with the weighting function $w(t)$ given by

$$w(t) = \begin{cases} p & \text{for } t = 1 \\ 1-p & \text{for } t = \frac{1}{k} \end{cases}$$

where $0 < p < 1, k > 0$, and $k \neq 1$, we have

$$\begin{aligned} & E_{R_{pt}}\left[((R_{pt}-\mu_p)^2 - V[R_{pt}])\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] \\ &= \left\{\int_0^\infty w(t)t^{-2}dt - \left(\int_0^\infty w(t)t^{-1}dt\right)^2\right\}\sigma_{22} \\ &= \left\{(p + k^2(1-p)) - (p + k(1-p))^2\right\}\sigma_{22}. \end{aligned}$$

$p + k^2(1-p) > (p + k(1-p))^2$ when $0 < p < 1, k > 0$, and $k \neq 1$. This case was considered by Zhou (1993).

Example 4.

When $(\boldsymbol{\epsilon_t'}, R_{pt})'$ follows Laplace distribution with the marginal pdf of R_{pt} given by

$$f_2(r_{pt}) = (2\sigma_{22})^{-1/2} exp\left\{-\sqrt{2}|r_{pt} - \mu_p|/\sigma_{22}^{1/2}\right\},$$

$\int_s^\infty f_2(s_2)ds_2/f_2(s) = \sqrt{2}|R_{pt} - \mu_p|/\sigma_{22} + 1$. Then

$$\begin{aligned} E_{R_{pt}}\left[(R_{pt}-\mu_p)^2\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] &= \frac{\sqrt{2}}{\sigma_{22}^{1/2}}E\left[|R_{pt}-\mu_p|^{\frac{5}{2}}\right] + V\left[R_{pt}\right] \\ &= 2^{-3/4}\Gamma(\frac{7}{2})V[R_{pt}] + V[R_{pt}] \end{aligned}$$

and

$$E_{R_{pt}}\left[\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] = \frac{\sqrt{2}}{\sigma_{22}^{1/2}}E[|R_{pt}-\mu_p|] + 1 = 2.$$

Therefore, we have

$$E_{R_{pt}}\left[((R_{pt}-\mu_p)^2 - V[R_{pt}])\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] = \left(2^{-3/4}\Gamma(\frac{7}{2}) - 1\right)V[R_{pt}].$$

4 CONCLUSION

We have investigated the effects of nonnormality on the statistical inference of the market model in the class of elliptical distributions for the joint distribution of the asset returns and a given portfolio return or equivalently for the joint distribution of the market model error term and a given portfolio return. We have analyzed the asymptotic covariance matrix of the LSE of the market model in the class of elliptical distributions, comparing with the standard covariance matrix of the LSE derived under normality. We have found that the asymptotic covariance matrices of the LSE for α and β are larger in the sense of positive definiteness than the standard covariance matrices when the joint distribution is nonnormal in the class of elliptical distributions, provided that the joint distribution satisfies the representation of scale mixtures of multivariate normal distributions. Thus we have generalized the result of MR (1991) in the case of multivariate Student t distribution to a larger class of scale mixtures of normal distributions. Equivalently, the traditional non-robustified tests such as the exact test by GRS (1989) reject too often the null hypothesis of mean-variance efficiency when the joint distribution is scale mixtures of normal distributions. This implies that the assumption of the contemporaneous conditional homoskedasticity in the conditional covariance matrix is very strong and restrictive since it holds only under normality in the class of elliptical distributions and encourages the use of the robust tests, such as the GMME based tests proposed by MR (1991), which are robust against departures from the normality assumption or the contemporaneous conditional homoskedasticity assumption.

We have also quantified the difference of the non-robustified and robust asymptotic covariance matrices in multivariate Student t distribution, contaminated multivariate normal distribution, and Laplace distribution, making use of the properties of the class of elliptical distributions.

Considering nonnormal distributions is equivalent to considering heteoskedastic error term in the class of elliptical distributions as far as the contemporaneous conditional heteroskedasticity is concerned. Therefore, nonnormality and contemporaneous conditional heteroskedasticity can be treated simultaneously by introducing nonnormal distributions in the class of elliptical distributions. On the other hand, nonnormality and heteroskedasticity appear to have been mostly treated as different phenomena in the literature. But this is not true for the case of contemporaneous conditional heteroskedasticity in the class of elliptical distributions. Since financial data are well known to be better described by heavy tail distributions, making inferences explicitly incorporating contemporaneous conditional heteroskedasticity may be worthwhile for the market model.

APPENDIX: PROOF OF THEOREM 2

When the pdf of $\boldsymbol{Z} = (\boldsymbol{\epsilon}_t', R_{pt})'$ is given by

$$f(z) = \int_0^\infty w(t)(2\pi)^{-p/2}|t^{-1}\boldsymbol{\Sigma}|^{-1/2} exp\left\{-\frac{1}{2}(z-\mu)'(t^{-1}\boldsymbol{\Sigma})^{-1}(z-\mu)\right\} dt$$

where $w(t)$ is a weighting function, the marginal pdf of $R_{pt}, p_2(r_{pt}) = f_2(\frac{1}{2}(r_{pt}-\mu_p)^2/\sigma_{22}) = f_2(s)$ where $s = \frac{1}{2}(r_{pt}-\mu_p)^2/\sigma_{22}$, is given by

$$f_2(s) = \int_0^\infty w(t)(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}e^{-ts}dt$$

(cf. Chu (1973)). Then we have

$$\begin{aligned}
\int_s^\infty f_2(s_2)ds_2 &= \int_s^\infty \int_0^\infty w(t)(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}e^{-ts_2}dtds_2 \\
&= \int_0^\infty w(t)(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}\int_s^\infty e^{-ts_2}ds_2dt \\
&= \int_0^\infty w(t)t^{-1}(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}e^{-ts}dt \\
&= \int_0^\infty w(t)t^{-1}(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}exp\left\{-\frac{1}{2}t(r_{pt}-\mu_p)^2/\sigma_{22}\right\} dt
\end{aligned}$$

where $(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2}exp\left\{-\frac{1}{2}t(r_{pt}-\mu_p)^2/\sigma_{22}\right\}$ is the pdf of normal distribution with mean μ_p and variance $t^{-1}\sigma_{22}$. Hence, we have

$$\begin{aligned}
&E_{R_{pt}}\left[(R_{pt}-\mu_p)^2\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}\right] \\
&= \int_{-\infty}^\infty (r_{pt}-\mu_p)^2\frac{\int_s^\infty f_2(s_2)ds_2}{f_2(s)}f_2(s)dr_{pt} \\
&= \int_{-\infty}^\infty (r_{pt}-\mu_p)^2\int_s^\infty f_2(s_2)ds_2dr_{pt}
\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} (r_{pt}-\mu_p)^2 \int_0^{\infty} w(t) t^{-1} (2\pi)^{-1/2} |t^{-1}\sigma_{22}|^{-1/2} exp\{-\frac{1}{2} t (r_{pt}-\mu_p)^2/\sigma_{22}\} dt dr_{pt} \\
&= \int_0^{\infty} w(t) t^{-1} \int_{-\infty}^{\infty} (r_{pt}-\mu_p)^2 (2\pi)^{-1/2} |t^{-1}\sigma_{22}|^{-1/2} exp\{-\frac{1}{2} t (r_{pt}-\mu_p)^2/\sigma_{22}\} dr_{pt} dt \\
&= \int_0^{\infty} w(t) t^{-1} t^{-1} \sigma_{22} dt
\end{aligned}
$$

where the last equality holds since $(2\pi)^{-1/2}|t^{-1}\sigma_{22}|^{-1/2} exp\left\{-\frac{1}{2}t(r_{pt}-\mu_p)^2/\sigma_{22}\right\}$ is the pdf of normal distribution with mean μ_p and variance $t^{-1}\sigma_{22}$. Therefore, we have

$$
E_{R_{pt}}\left[(R_{pt}-\mu_p)^2 \frac{\int_s^{\infty} f_2(s_2) ds_2}{f_2(s)}\right] = \int_0^{\infty} w(t) t^{-2} dt \sigma_{22}.
$$

We also have

$$
\begin{aligned}
E_{R_{pt}}\left[\frac{\int_s^{\infty} f_2(s_2) ds_2}{f_2(s)}\right] &= \int_{-\infty}^{\infty} \frac{\int_s^{\infty} f_2(s_2) ds_2}{f_2(s)} f_2(s) dr_{pt} \\
&= \int_{-\infty}^{\infty} \int_s^{\infty} f_2(s_2) ds_2 dr_{pt} \\
&= \int_{-\infty}^{\infty} \int_0^{\infty} w(t) t^{-1} (2\pi)^{-1/2} |t^{-1}\sigma_{22}|^{-1/2} exp\{-\frac{1}{2} t (r_{pt}-\mu_p)^2/\sigma_{22}\} dt dr_{pt} \\
&= \int_0^{\infty} w(t) t^{-1} \int_{-\infty}^{\infty} (2\pi)^{-1/2} |t^{-1}\sigma_{22}|^{-1/2} exp\{-\frac{1}{2} t (r_{pt}-\mu_p)^2/\sigma_{22}\} dr_{pt} dt \\
&= \int_0^{\infty} w(t) t^{-1} dt
\end{aligned}
$$

where the last equality holds because

$$
\int_{-\infty}^{\infty} (2\pi)^{-1/2} |t^{-1}\sigma_{22}|^{-1/2} exp\{-\frac{1}{2} t (r_{pt}-\mu_p)^2/\sigma_{22}\} dr_{pt} = 1.
$$

Therefore,

$$
\begin{aligned}
&E_{R_{pt}}\left[((R_{pt}-\mu_p)^2 - V[R_{pt}]) \frac{\int_s^{\infty} f_2(s_2) ds_2}{f_2(s)}\right] \\
&= \left\{\int_0^{\infty} w(t) t^{-2} dt - (\int_0^{\infty} w(t) t^{-1} dt)^2\right\} \sigma_{22}
\end{aligned}
$$

since $V[R_{pt}] = \{\int_0^{\infty} w(t) t^{-1} dt\} \sigma_{22}$ (cf. Corollary 3.2 of Chu (1973)). Since $\int_0^{\infty} w(t)t^{-1}dt$ and $\int_0^{\infty} w(t)t^{-2}dt$ denote expectation of the inverse of the random variable associated with $w(t)$ and expectation of the inverse of its square respectively when $w(t)$ is a probability measure, the following relationship given by

$$
\int_0^{\infty} w(t) t^{-2} dt \geq \left(\int_0^{\infty} w(t) t^{-1} dt\right)^2
$$

holds by the Cauchy-Schwarz inequality where the equality holds only when $w(t)$ is a degenerate probability measure. Since $w(t)$ being a degenerate pdf is equivalent to Z being normal, this completes the proof.

REFERENCES

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

Andrews, D.F., and C.L. Mallows, 1974, Scale mixtures of normal distributions, *Journal of Royal Statistical Society B* 36, 99-102.

Campbell, J.Y., A. W. Lo, and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets* Princeton University Press, Princeton, NJ.

Chamberlain, G., 1983, A characterization of the distributions that imply mean-variance utility functions, *Journal of Economic Theory* 29, 185-201.

Chu, K.C., 1973, Estimation and decision for linear systems with elliptical random processes, *IEEE Trasactions on Automatic Control* AC-18, 499-505.

Devlin, S.J., R. Gnanadesikan, and J.R. Kettenring, 1976, Some multivariate applications of elliptical distributions, *Essays in Probability and Statistics,* S. Ikeda(ed.) Shinko Tsusho Co. Ltd. Tokyo, Japan.

Fang, K.T., and Y.T. Zhang, 1990, *Generalized Multivariate Analysis* Springer-Verlag, New-York.

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

Gibbons, M., S. Ross, and J. Shanken, 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, 1029-1054.

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

Jobson, H., 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.

Kariya, T., and B. K. Sinha, 1988, *Robustness of Statistical Tests* Academic Press, San Diego.

Kelker, D., 1970, Distribution theory of spherical distributions and a location-scale parameter generalization, *Sankhya A* 32, 419-430.

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

MacKinlay, A. C., and M. Richardson, 1991, Using generalized methods of moments to test mean-variance efficiency, *Journal of Finance* 46, 511-527.

Owen, J., and R. Rabinovitch, 1983, On the class of elliptical distributions and their applications to the theory of portfolio choice, *Journal of Finance* 38,745-752.

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

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

White, H., 1982, Maximum likelihood estimation of misspecified models, *Econometrica* 50, 1-25.
Zhou, G., 1993, Asset-pricing tests under alternative distributions,*Journal of Finance* 48, 1927-1942.