GCOE Discussion Paper Series

Global COE Program

Human Behavior and Socioeconomic Dynamics

Discussion Paper No. 301

Asset allocation under higher moments with the GARCH filter

Ryo Kinoshita

March 2013

GCOE Secretariat
Graduate School of Economics
OSAKA UNIVERSITY
1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

Asset allocation under higher moments with the GARCH filter

Abstract

Except for several pairs of utility functions and distribution functions, expected utility maximization problems do not have closed form solutions so that these problems often require complex numerical optimizations. The paper proposes an approximated solution to this problem using higher moments of returns. Utility functions are approximated by polynomials by the Taylor expansion, and thus expected utility functions are approximated by linear combinations of moments. With the GARCH effects, a simple approach to estimate conditional higher moments is given. In an empirical study, the strategy is compared to alternative strategies such as mean variance optimization and static optimization.

JEL Classification Code: G11, C58, C61

Keywords : Portfolio allocation, GARCH model, non-normality

1 Introduction

Markowitz (1952) proposed the now well known mean variance portfolio selection. When asset returns follow the multivariate normal distribution, this approach is a good approximation of the expected utility maximization for risk averse investors. However, when returns are non-normally distributed higher moments may affect the optimal portfolios since the mean variance criteria cannot capture the investor's preference about the asymmetric property or the heavy tailedness of distributions of asset returns. Recently, many papers have reported that asset returns are asymmetric and heavy tailed distribution. In addition, many papers reported that asset returns have conditional heteroskedasticities. Asset allocation strategies using the conditional heteroskedasticities are known as volatility timing. In this paper, I consider a simple asset allocation problem with non-normalities and volatility timing.

Previous studies have used two model specifications where asset allocation problems are studied under non-normalities. The first is the regime switching model. Ang and Bekaert (2002) and Guidolin and Timmerman (2008) consider asset allocation problems where returns are regime switching in international stock markets. In this specification, the conditional distributions that investors face are the mixtures of normal distributions. The other specification is the Generalized conditional heteroskedasticity model (GARCH) with non-normal innovations. Harvey et al (2010) and Jondeau and Rockinger (2012) focused on the skew normal distribution proposed by Sahu, Dey and Branco (2003), of which the co-skewness matrix and co-kurtosis matrix can be computed explicitly from the parameters of the distribution function to demonstrate the optimizations. Patton (2004) and Jondeau and Rockinger (2006) used the copula-GARCH model. With an assumption of a location-scale mixture of the multivariate normal distributions, Mencia and Sentana (2009) derived a mean-variance-skewness optimization problem. Under the GARCH model with non-normal innovations, conditional distribution is derived by scaling innovations with GARCH effects. Other related studies include the following, Adcock (2010) demonstrated an asset allocation problem under multivariate skew normal and skew-t distribution which were proposed by Azzalini and Capitanio (2003). Konno et al. (1991) derived a mean-absolute deviation-skewness optimization problem that can handle a very large number of assets. Theoretical properties of the mean-variance-skewness frontier were derived by Athayde and Flores (2004). Non-normalities of asset returns have been reported in many studies not directly related to asset allocation problems. Longin and Solnik (2001) reported non-normalities of international asset returns using the exceedance correlation measure. Bauwens and Laurent (2005) adapted a multivariate skew-t distribution for the multivariate GARCH model. For univariate series, Hansen (1994) considered skewed distributions for innovations of the GARCH model. Jondeau and Rockinger (2003) studied conditional shapes of distributions of asset returns using distributions proposed by Hansen (1994). Harvey (1999) proposed the GARCHS model, which models the conditional skewnesses in the manner of the GARCH formula.

Except for several pairs of distributions and utility functions, the expected utility functions cannot be computed analytically. When non-normalities are introduced, this is true for most cases. One of approaches for calculating expected utility function are numerical integrations called quadrature methods proposed by Tauchen and Hussey (1991). Ang and Chen (2002) employed this approach under the regime switching model. The second approach is Monte Carlo integration, which is used when random generations are possible. Patton (2004) used this approach under the copula-GARCH model. These two approaches require heavy computation to obtain the expected utility function, and the optimization is often unstable. The third approach is an approximation of the utility function by the Taylor expansion. Expansion up to the second moments results in the quadratic utility function, so it only focuses on the first two moments (mean variance approach). In the cases of non-normal distributions of returns, an approximation using higher moments may be valuable. Jondeau and Rockinger (2006) derived an optimal asset allocation problem with the expansion up to the first four moments using sample moments. Jondeau and Rokinger (2012) extended this to incorporate the dynamic conditional correlation GARCH (DCC-GARCH) model. They assume specific distribution for innovations, which restricts the structures of the moments. In this paper, I propose a simple expected utility maximization problem using an approximation by the Taylor expansion under the GARCH model without assuming any specific distributions. As in Jondeau and Rockinger (2006), Jondeau and Rockinger (2012) and Guidolin and Timmerman (2008), I expand an investor's utility function up to the first four moments. This paper can be considerd an extension of Jondeau and Rockinger (2006). For simplicity, it is assumed that asset returns follow the multivariate GARCH model with the correlation matrix, co-skewness matrix co-kurtosis matrix of innovations constant over time; however, the DCC GARCH is also applicable. The main idea is a two step estimation. I first estimate the multivariate GARCH model by variance targeting Gaussian Quasi-maximum likelihood estimation (QMLE). Given the estimated GARCH parameters, I estimate the standardized innovation series and the co-skewness matrix of innovations by those sample moments. Changing the location and scale of the correlation matrix, co-skewness matrix and co-kurtosis matrix of innovations by the GARCH effects, conditional covariance matrices, co-skewness matrices and co-kurtosis matrices are estimated. This two step estimation is easy to compute relative to the previous studies and is adaptive for many studies of the GARCH model. With the estimated conditional moments, optimal asset allocations are estimated by maximizing the approximated expected utility function in each period.

The effects of the degree of approximation for the expected utility functions are measured by the certainty equivalent returns (performance fee) as in Fleming, Kirby and Ostdiek (2001), Patton (2004) and Jondeau and Rockinger (2012). To evaluate these differences statistically, the bootstrap method is employed. The remainder of this paper is organized as follows. In the section 2, a model of the asset allocation problem and underlying asset returns processes are described. Derivation of optimal portfolio weights with approximated expected utility functions is explained. In the section 3, alternative portfolios and the methods to compare their performances are given. In section 4, an example of an empirical application is shown. and the section 5, concludes.

2 Model

In this section, an approach to the expected utility maximization problem for a risk averse investor is explained. In the first subsection, assumptions for return processes are given. Returns are assumed to be the multivariate GARCH model and innovations are identically and independently distributed and have finite third and fourth moments. Several previous studies assumed specific distributions for innovations (Jondeau and Rockinger (2012), Mencia and Sentana (2009) and Harvey et al. (2010)). In contrast to these studies, this paper does not assume any specific distributions for innovations. Under this setting, this paper demonstrates how to estimate time varying conditional covariance matrices, co-skewness matrices and co-kurtosis matrices which are induced by the GARCH effect. In the second subsection, an expected utility maximization problem and approximations are given. Except for several cases, the expected utility cannot be calculated analytically and numerical integrations are required. In this paper, utility functions are approximated by the Taylor expansion to evaluate expected utility. This makes it possible to approximate the expected utility function by a weighted sum of the elements of the moment matrices. The approach is easy to compute relative to numerical integration methods such as the quadrature method and Monte Carlo integration. By maximizing the approximated utility functions under the estimated conditional moments, conditional optimal portfolio weights are estimated for each period.

2.1 Return process

2.1.1 Estimations of the GARCH and central moment matrices

In this subsection, a model of asset returns is given, and the dynamics of conditional covariance matrices and co-skewness matrices and estimations are explained. There are a risk free asset and n risky assets with return vector $r_t = (r_{1t}, r_{2t}, \ldots, r_{nt})'$ and those processes are multivariate GARCH with constant conditional correlation. That is, the individual return series follow univariate asymmetric-GARCH (1,1) which is proposed by Glosten, Jagannathan and

Runkle (1993) and the correlation matrix of innovations $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{nt})'$ is constant over time. In addition, it is assumed that the co-skewness matrix and co-kurtosis matrix of the innovations are constant over time:

$$r_{it} = \mu_i + \sigma_{it}\epsilon_{it}, \tag{1}$$

$$\sigma_{it}^2 = w_i + \alpha_i \sigma_{it-1}^2 + \beta_i \epsilon_{it-1}^2 + \gamma \epsilon_{it-1}^2 I_{\{\epsilon_{it-1}<0\}}, \tag{2}$$

$$\epsilon_t \sim \text{iid}(\mathbf{0}, \Omega, \mathbf{S}, \mathbf{K}), \tag{3}$$

where $\mathbf{S}$ and $\mathbf{K}$ are the co-skewess matrix ($n \times n^2$) and co-kurtosis matrix ($n \times n^3$) respectively, the same as in Jondeau and Rockinger (2006) such that:

$$\mathbf{S} = \mathrm{E}[(\epsilon_t - \mu) \otimes (\epsilon_t - \mu)' \otimes (\epsilon_t - \mu)'], \tag{4}$$

$$\mathbf{K} = \mathrm{E}[(\epsilon_t - \mu)(\epsilon_t - \mu)' \otimes (\epsilon_t - \mu)' \otimes (\epsilon_t - \mu)']. \tag{5}$$

The co-skewness matrix consists of n submatrices such that

$$\mathbf{S} = \begin{bmatrix} S_1 & S_2 & \cdots & S_n \end{bmatrix},$$

where

$$\mathbf{S_i} = \begin{bmatrix} s_{i11} & s_{i12} & \cdots & s_{i1n} \\ s_{i21} & s_{i22} & \cdots & s_{i2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{in1} & s_{in2} & \cdots & s_{inn} \end{bmatrix}.$$

The co-kurtosis matrix consists of n submatrices such that

$$\mathbf{K} = \begin{bmatrix} \mathbf{K_1} & \mathbf{K_2} & \cdots & \mathbf{K_n} \end{bmatrix},$$

where K_i consists of submatrices such that

$$\mathbf{K_i} = \begin{bmatrix} \mathbf{K_{i1}} & \mathbf{K_{i2}} & \cdots & \mathbf{K_{in}} \end{bmatrix},$$

where

$$\mathbf{K_{ij}} = \begin{bmatrix} k_{ij11} & k_{ij21} & \cdots & k_{ij1n} \\ k_{ij21} & k_{ij22} & \cdots & k_{ij2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{ijn1} & k_{ijn1} & \cdots & k_{ijnn} \end{bmatrix},$$

where $s_{ijk} = \mathrm{E}[\epsilon_{it}\epsilon_{jt}\epsilon_{kt}]$ and $k_{ijkl} = \mathrm{E}[\epsilon_{it}\epsilon_{jt}\epsilon_{kt}\epsilon_{kt}]$. The co-skewness matrix and co-kurtosis matrix have $n(n+1)(n+2)/6$ and $n(n+1)(n+2)(n+3)/24$ non-redundant components respectively. Using these matrices, Jondeau and Rockinger (2006) proposed a framework to use higher moments for asset allocation problems with sample moments. I extend their approach to incorporate the GARCH effects. Mencia and Sentana (2009) also applied the GARCH model assuming that the distributions of ϵ_t are location-scale mixture of the multivariate normal distributions. Jondeau and Rockinger (2012) and Harvey et al. (2010) also used the GARCH model and higher moments assuming the innovations to be the skewed t-distribution proposed by Sahu, Dey and Branco (2003). In contrast to these previous studies, this paper does not assume any specific distributions for innovations so it has flexibility for the shape of ϵ_t. In this paper, a two step semiparametric estimation is employed. The first step is the estimation of GARCH parameters. The second step is the estimation of the shape of innovations, that is, an estimation of higher moment matrices.

For the GARCH estimation, variance targeting estimation is employed. Unconditional means and variances are estimated by the sample moments for each asset:

$$\hat{\mu}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it}, \tag{6}$$

$$\bar{\sigma}_i = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - \hat{\mu}_i)^2. \tag{7}$$

Given these unconditional moments, the GARCH parameters are estimated by the Gaussian

QMLE which uses the likelihood function of the multivariate normal distribution as follows:

$$\max \quad l(\alpha_i, \beta_i, \gamma_i, \Omega; \mathbf{r}) = -\frac{nT}{2}\log 2\pi - \frac{T}{2}\log|\Omega| - \sum_{t=1}^{T}\sum_{i=1}^{n}\log|\sigma_{it}| - \frac{1}{2}\sum_{i=1}^{T}\epsilon_t'\Omega^{-1}\epsilon_t, \quad (8)$$

$$\text{s.t } \sigma_{it}^2 = \omega_i + \alpha\sigma_{it-1}^2 + \beta_i\epsilon_{it-1}^2 + \gamma_i\epsilon_{it-1}^2 I_{\{\epsilon_{it-1}<0\}},$$

$$\frac{\omega_i}{1-\alpha_i-\beta_i-\frac{\gamma_i}{2}} = \bar{\sigma}_i^2,$$

where $\epsilon_{it} = (r_{it} - \hat{\mu}_i)/\sigma_{it}$. When the dynamics of scales are correctly specified, the Gaussian QMLE gives consistent estimations for $\{w_i, \alpha_i, \beta_i, \gamma_i, \Omega\}$. Given the estimated parameters, innovations for the GARCH are estimated:

$$\hat{\sigma}_{it}^2 = \hat{w}_i + \hat{\alpha}_i\hat{\sigma}_{it-1}^2 + \hat{\beta}_i\hat{\epsilon}_{it-1}^2 + \hat{\gamma}_i\epsilon_{it-1}^2 I_{\{\epsilon_{it-1}<0\}}, \quad (9)$$

$$\hat{\epsilon}_{it} = \frac{r_{it} - \hat{\mu}_i}{\hat{\sigma}_{it}}. \quad (10)$$

Using these standardized innovations, the co-skewness matrix and co-kurtosis matrix of innovations are estimated as follows:

$$\hat{\mathbf{S}} = \frac{1}{T}\hat{\epsilon}_t\hat{\epsilon}_t' \otimes \epsilon_t', \quad (11)$$

$$\hat{\mathbf{K}} = \frac{1}{T}\hat{\epsilon}_t\hat{\epsilon}_t' \otimes (\epsilon_t' \otimes \epsilon_t'). \quad (12)$$

These are $(n\times n^2)$ and $(n\times n^3)$ matrcies, and their components are simply $\hat{s}_{ijk} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{it}\hat{\epsilon}_{jt}\hat{\epsilon}_{kt}$, $\hat{k}_{ijkl} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{it}\hat{\epsilon}_{jt}\hat{\epsilon}_{kt}\hat{\epsilon}_{lt}$, Using the estimated GARCH parameters, volatility series and standarized innovation series are estimated. Since $\mathbf{S}$ and $\mathbf{K}$ are the third and fourth moment matrices of standardized innovations respectively, conditional moment matricies also vary because of the change in scale induced by the GARCH effects. Taking into account the GARCH effects, the conditional covariance matrix, co-skewness matrix and co-kurtosis matrix are as follows:

$$\mathbf{V}_t = \mathbf{G_t\Omega G_t}, \quad (13)$$

$$\mathbf{S}_t = \left[\sigma_{1t}\mathbf{G_tS_1G_t} \quad \sigma_{2t}\mathbf{G_tS_2G_t} \quad \cdots \quad \sigma_{nt}\mathbf{G_tS_nG_t}\right], \quad (14)$$

$$\mathbf{K}_t = \left[\sigma_{1t}\sigma_{1t}\mathbf{G_tK_{11}G_t} \quad \sigma_{1t}\sigma_{2t}\mathbf{G_tK_{12}G_t} \quad \cdots \quad \sigma_{nt}\sigma_{nt}\mathbf{G_tK_{nn}G_t}\right], \quad (15)$$

where $\mathbf{G_t} = \text{diag}\{\sigma_{1t} \ \cdots \ \sigma_{nt}\}$. The elements of the conditional co-skewness matrix are $\hat{s}_{ijkt} = \hat{\sigma}_{it}\hat{\sigma}_{jt}\hat{\sigma}_{kt}\hat{s}_{ijk}$ and the elements of the conditional co-kurtosis matrix are $\hat{k}_{ijklt} = \hat{\sigma}_{it}\hat{\sigma}_{jt}\hat{\sigma}_{kt}\hat{\sigma}_{lt}k_{ijkl}$. The conditional covariance matrix is the same as the constant conditional correlation model (CCC). Since the focus is not on the dynamics of the second moments, the constant correlation GARCH is assumed for simplicity. However the dynamic conditional correlation model (DCC) proposed by Engle (2002) can be applied. The conditional co-skewness matrix and co-kurtosis matrix are also scaled by the GARCH effects as well as the covariance matrix. In this way, conditional moments with the GARCH effects are estimated.

2.1.2 Transformation into noncentral moments from central moments

As showen later, when the utility function is approximated by a polynomial, the expected utility function is the weighted sum of the noncentral moments of a portfolio return. Since GARCH models are those of central moments, moment matrices estimated above are also central moments. Preliminary to an asset allocation problem, I transform central moments into noncentral moments using binomial expansion. Let $\mathbf{M_{2,t}}, \mathbf{M_{3,t}}$,and $\mathbf{M_{4,t}}$ be the first,second ,third, and fourth conditional noncentral moment matrices, respectively. Elements of the noncentral moments are computed from central moments as follows:

$$m_{2,ijt} = \text{E}_t[r_{it+1}r_{jt+1}] = m_{2,ijt} + \mu_i\mu_j, \tag{16}$$

$$m_{3,ijkt} = \text{E}_t[r_{it+1}r_{jt+1}r_{kt+1}] = s^3_{ijkt} + m_{2,ijt}\mu_k + m_{2,ikt}\mu_j + m_{2,jkt}\mu_i - 2\mu_i\mu_j\mu_k, \tag{17}$$

$$\begin{aligned} m_{4,ijklt} = \text{E}_t[r_{it+1}r_{jt+1}r_{kt+1}r_{lt+1}] = k^4_{ijklt} + m_{3,ijkt}\mu_l + m_{3,ijlt}\mu_k + m_{3,iklt}\mu_j + m_{3,jklt}\mu_i \\ - m_{2,ijt}\mu_k\mu_l - m_{2,ikt}\mu_j\mu_l - m_{2,ilt}\mu_j\mu_k - m_{2,jkt}\mu_i\mu_l - m_{2,jlt}\mu_i\mu_k - m_{2,klt}\mu_i\mu_j \\ + 3\mu_i\mu_j\mu_k. \end{aligned} \tag{18}$$

2.2 Asset allocation

In this subsection, a conditional expected utility maximization problem for a risk averse investor is explained under the assumptions of the return process described in the previous subsection. There are n risky assets with return vector $r_t = (r_{1t}, \ldots, r_{nt})'$ and a corresponding portfolio vector is denoted by $\mathbf{w_t} = (w_{1t}, \ldots, w_{nt})'$ and the weight for the risk free asset is denoted by $w_{ft} = 1 - \mathbf{w_t'1} = 1 - \sum_{i=1}^{n} w_{it} \geq 0$. Given the initial wealth is W_t at t, the myopic optimization problem is as follows:

$$\max_{\mathbf{w_t}} \quad \mathrm{E}_t[U(W_{t+1})] = \mathrm{E}_t[U(W_t(1 + \mathbf{w_t}'\mathbf{r_{t+1}} + w_{ft}r_{ft}))]. \tag{19}$$

Except for some pairs of a utility function and a distribution function, this problem does not have a closed form solution. For instance, a pair of the exponential utility function and the multivariate normal distribution has a closed form solution so that a numerical optimization is required to evaluate the integration. When specific distributions are assumed, the integration can be calculated with the quadrature methods proposed by Tauchen and Hussey (1991). Ang and Bekaert (2002) employed this approach to calculate the expected utility function where returns are regime switching. Another approach is Monte Carlo integration which was employed by Patton (2004) when random generations were possible. However, both approaches involve heavy computation. In this paper, another approach is employed to evaluate the expected utility function. The utility function is approximated by the Taylor expansion up to a finite number of moments as follows:

$$E_t[U(W_{t+1})] \approx \sum_{k=1}^{K} \frac{\theta_k}{k!} \mathrm{E}_t[(W_{t+1} - W_t)^k]. \tag{20}$$

When the expansion is up to the second moment, the problem results in mean variance optimization. When returns follow a multivariate normal distribution, the first two moments suffice for the expected utility maximization problem. However, when non-normalities are introduced, higher moments affect the investor's optimization problem. A great deal of empirical research suggests that asset returns are sometimes asymmetric and heavy tailed (Longin and Solnik (2001), Bauwen and Laurent (2003)). In this paper, the utility function is expanded up to the first four moments. Jondeau and Rockinger (2006) and Harvey et al (2010) also employed this approach. A case of Taylor expansion up to the fourth moments can be written as follows:

$$E_t[U(W_{t+1})] \approx e(\mathbf{w_t}) = U^{(1)}(W_t)m_{1,Pt+1}(\mathbf{w_t}) + \frac{1}{2}U^{(2)}(W_t) \cdot m_{2,Pt+1}(\mathbf{w_t}) + \frac{1}{6}U^{(3)}(W_t)m_{3,Pt+1}(\mathbf{w_t}) + \frac{1}{24}U^{(4)}(W_t)m_{4,Pt+1}(\mathbf{w_t}) \tag{21}$$

where $m_{1,Pt+1}(\mathbf{w_t}), m_{2,Pt+1}(\mathbf{w_t}), m_{3,Pt+1}(\mathbf{w_t})$ and $m_{4,Pt+1}(\mathbf{w_t})$ are the first, second, third and fourth noncentral moments of the portfolio, respectively. The noncentral moments of the portfolio are computed from the weighted sum of noncentral moment matrices as follows:

$$m_{1,Pt+1}(\mathbf{w_t}) = \mathbf{w_t'}\mu_{t+1} = \sum_{i=1}^{n} w_{it}\mu_i + w_{ft}r_{ft+1},$$

$$m_{2,Pt+1}(\mathbf{w_t}) = \mathbf{w_t'}\mathbf{M_{2,t+1}}\mathbf{w_t} = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{it}w_{jt}m_{2,ijt+1},$$

$$m_{3,Pt+1}(\mathbf{w_t}) = \mathbf{w_t}'\mathbf{M_{3,t+1}}(\mathbf{w_t} \otimes \mathbf{w_t}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} w_{it}w_{jt}w_{kt}m_{3,ijkt+1},$$

$$m_{4,Pt+1}(\mathbf{w_t}) = \mathbf{w_t}'\mathbf{M_{4,t+1}}(\mathbf{w_t} \otimes \mathbf{w_t} \otimes \mathbf{w_t}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} w_{it}w_{jt}w_{kt}w_{lt}m_{4,ijklt+1}.$$

To maximize this approximated utility function by standard techniques, the derivative vector and the Hessian matrix are required. The elements of the derivative vector and the Hessian matrix can be written using submatrices of the higher moment matrices. The elements of the derivative vector are as follows:

$$\frac{\partial m_{1,Pt+1}}{\partial w_{it}} = \mu_i - r_{ft+1},$$

$$\frac{\partial m_{2,Pt+1}}{\partial w_{it}} = 2\sum_{j=1}^{n} w_{it}m_{2,ijt+1},$$

$$\frac{\partial m_{3,Pt+1}}{\partial w_{it}} = 3\mathbf{w_t}'\mathbf{M_{3,t+1}}\mathbf{w_t} = 3\sum_{j=1}^{n}\sum_{k=1}^{n} w_{jt}w_{kt}m_{3,ijkt+1},$$

$$\frac{\partial m_{4,Pt+1}}{\partial w_{it}} = 4\mathbf{w_t}'\mathbf{M_{4,t+1}}(\mathbf{w_t} \otimes \mathbf{w_t}) = 4\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} w_{jt}w_{kt}w_{lt}m_{4,ijklt+1}.$$

The elements of the Hessian matrix are as follows:

$$\frac{\partial^2 m_{2,Pt+1}}{\partial w_{it}\partial w_{jt}} = 2m_{2,ijt+1}$$
$$\frac{\partial^2 m_{3,Pt+1}}{\partial w_{it}\partial w_{jt}} = 6\sum_{k=1}^{n} m_{3,ijkt+1}w_{kt}$$
$$\frac{\partial^2 m_{4,Pt+1}}{\partial w_{it}\partial w_{jt}} = 12\mathbf{w_t}'\mathbf{M_{4,t+1}}\mathbf{w_t} = 12\sum_{k=1}^{n}\sum_{l=1}^{n} m_{4,ijklt+1}w_{kt}w_{lt}.$$

In this paper, I consider an investor whose preference is represented by a constant relative risk aversion (CRRA) utility function (i.e. $U(W) = \frac{W^{1-\lambda}}{1-\lambda}$) and normalize the initial wealth $W = 1$. The approximated utility function is as follows:

$$\begin{aligned} \mathrm{E}_t[U(1 + \mathbf{w_t'}\mathbf{r_{t+1}} + w_{ft}r_{ft+1})] \approx\; & m_{1,Pt+1} - \frac{\lambda}{2}m_{2,Pt+1}(\mathbf{w_t}) \\ & + \frac{\lambda(\lambda+1)}{6}m_{3,Pt+1}(\mathbf{w_t}) - \frac{\lambda(\lambda+1)(\lambda+2)}{24}m_{4,Pt+1}(\mathbf{w_t}). \end{aligned} \tag{22}$$

For risk averse investors, the effects of the third and fourth moments are positive and negative respectively. This is consistent with the intuition that an investor tries to avoid a large negative return and a fat-tailed distribution. It is logical to assume that the CRRA utility function is useful since optimizations are not affected by the level of initial wealth. If the initial wealth is W, then the optimal weight is $W\omega_t^*$ where ω_t^* is the optimal portfolio when the initial wealth is unity. To sum up, replacing the moments matrices with estimated series, the conditional optimal portfolio weights given information up to $t-1$ are selected by solving following problem:

$$\begin{aligned} \max_{\mathbf{w_t}} \quad & \hat{e}(\mathbf{w_t}) = \hat{m}_{1,Pt+1} - \frac{\lambda}{2}\hat{m}_{2,Pt+1}(\mathbf{w_t}) + \frac{\lambda(\lambda+1)}{6}\hat{m}_{3,Pt+1}(\mathbf{w_t}) - \frac{\lambda(\lambda+1)(\lambda+2)}{24}\hat{m}_{4,Pt+1}(\mathbf{w_t}) \\ \text{s.t} \quad & \hat{m}_{1,Pt+1}(\mathbf{w_t}) = \mathbf{w_t'}\hat{\mu}_i + w_{ft}r_{ft+1}, \\ & \hat{m}_{2,Pt+1}(\mathbf{w}_t) = \mathbf{w_t}'\hat{\mathbf{M}}_{2,t+1}\mathbf{w_t}', \\ & \hat{m}_{3,Pt+1}(\mathbf{w}_t) = \mathbf{w_t}'\hat{\mathbf{M}}_{3,t+1}(\mathbf{w_t}\otimes\mathbf{w_t}), \\ & \hat{m}_{4,Pt+1}(\mathbf{w}_t) = \mathbf{w_t}'\hat{\mathbf{M}}_{4,t+1}(\mathbf{w_t}\otimes\mathbf{w_t}\otimes\mathbf{w_t}), \end{aligned}$$

The optimization problem is solved by the Newton-Rapthon method for each t. Since the problem is an approximation of the expected utility maximization, the performance of an approximation is of interest. By casting estimated portfolio weights in the argument of utility function, realized utilities can be calculated. In the following section, I measure the difference between alternative strategies.

3 Comparison of performance

In section 2, an approximated expected utility maximization problem is introduced. When returns are non-normally distributed, including higher moments will better approximate the expected utility function. In this section I determine whether including the higher moments and volatility timing is valuable or not.

3.1 Alternative portfolios

In this subsection, alternative asset allocation strategies are defined. To study the effect of higher moments and volatility timings, I compare static and dynamic portfolios changing the degree of approximations. The first strategy is the static portfolio with mean and variance which maximizes the approximated expected utility function up to first two moments. The approximated expected utility function is as follows:

$$\mathrm{E}_t[U(1 + \mathbf{w_t}'\mathbf{r_{t+1}} + w_{ft} r_{ft+1})] \approx m_{Pt+1} - \frac{\lambda^2}{2} m^2_{Pt+1}(\mathbf{w_t}). \tag{23}$$

The second strategy is the static portfolio with first four moments. This portfolio captures the non-normalities of returns but the GARCH effects are not considered. The third strategy is the dynamic mean-variance portfolio which changes portfolio weights depending on the GARCH effects. This portfolio captures the GARCH effects but non-normalities are not considered.

3.2 Method of the comparison

To measure the difference of alternative strategies, I use the concept of certainty equivalent return. Let $\mathbf{w_t}^*$ and $\mathbf{w_t}^{**}$ be allocations under different strategies. Then an unconditional certainty equivalent return is m such that

$$\mathrm{E}[U(1+\mathbf{w_t}^{*'}r_{t+1}+m)] = \mathrm{E}[U(1+\mathbf{w_t}^{**'}r_{t+1})]. \tag{24}$$

Certainty equivalennt return m is a required extra return where investors become indifferent to both strategies. This can be interpreted as the investor being willing to pay m to switch his strategy from $\mathbf{w_t}^*$ to $\mathbf{w_t}^{**}$. This enables an investor to compare the cost of calculating the higher moments and its benefit. The certainty equivalent return is estimated by solving sample analogs of (24):

$$\frac{1}{T}\sum_{t=1}^{T} U(1+\hat{\mathbf{w}}_{\mathbf{t}}' r_{t+1} + w_{ft} r_{ft+1} + \hat{m}) = \frac{1}{T}\sum_{t=1}^{T} U(1+\tilde{\mathbf{w}}_{\mathbf{t}}' r_{t+1} + w_{ft} r_{ft+1}) \tag{25}$$

by the grid search method.

3.3 Bootstrap

To compare the expected utilities of two alternative strategies statistically, I employ a bootstrap method. To derive a bootstrap distributions of certainty equivalent returns, the following procedures are iterated.

1. Estimate the GARCH model (1), (2) and (3) using the return series $\{\mathbf{r}_t^{(k)}\}$.
2. Calculate the certainty equivalent return $m^{(k)}$ using the original return series $\mathbf{r_t}$ and the estimated parameters.
3. Using the original return series $\mathbf{r_t}$ and the estimated parameters, the mean, variance

and innovation series $\{\mu_i^{(k)}, \sigma_{it}^{2(k)}, \epsilon_t^{(k)}\}$ are estimated as in section 2

$$\begin{aligned}
\hat{\mu}_i^{(k)} &= \frac{1}{T}\sum_{t=1}^{T} r_{it}^{(k)}, \\
\hat{\sigma}_{i0}^{2(k)} &= \frac{1}{T}\sum_{t=1}^{T}(r_{it}^{(k)} - \hat{\mu}_i^{(k)})^2, \\
\hat{\sigma}_{it}^{2(k)} &= \hat{w}_i^{(k)} + \hat{\alpha}_i^{(k)}\hat{\sigma}_{it-1}^{2(k)} + \hat{\beta}_i^{(k)}\hat{\epsilon}_{it-1}^{2(k)} + \hat{\gamma}_i^{(k)}\hat{\epsilon}_{it-1}^{2(k)} I_{\{\epsilon_{it-1}<0\}}, \\
\hat{\epsilon}_{it}^{(k)} &= \frac{r_{it} - \hat{\mu}_i^{(k)}}{\hat{\sigma}_{it}^{(k)}}.
\end{aligned}$$

4. Resampling $\epsilon_t^{(k)}$, the bootstrap series of innovations $\{\epsilon_t^{(k*)}\}$ is computed.

5. A bootstrap series of returns $\{\mathbf{r_t}^{(k+1)}\}$ are computed as follows:

$$r_{it}^{(k+1)} = \hat{\mu}_i^{(k)} + \hat{\sigma}_{it}^{(k)}\epsilon_{it}^{(k*)}.$$

The null hypothesis is $m = 0$ and the alternative is $m \neq 0$. The null hypothesis implies that the two strategies are indifferent, at least unconditionally. Demeaning the bootstrap sample of m, the distribution of the certainty equivalent return under the null is simulated. The p-value for $\hat{m}$ is computed as

$$p(\hat{m}) = 2\min\left\{\frac{1}{B}\sum_{k=1}^{B} I(m^{(k)} - \bar{m} \leq \hat{m}), \frac{1}{B}\sum_{k=1}^{B} I(m^{(k)} - \bar{m} > \hat{m})\right\} \tag{26}$$

where $\bar{m} = \frac{1}{B}\sum_{k=1}^{B} m^{(k)}$. In the following section, an illustrative empirical application is demonstrated.

4 Empirical Application

In this section, an empirical application for the Tokyo Stock Exchange (TSE) is demonstrated. The most popular index of the TSE is TOPIX which consists of all the stocks in the first section of the TSE. In the TSE, TOPIX includes 33 categorizes industries. I construct portfolios using several indexes and a risk-free asset. The selected categories are Mining, Iron

& Steel, Transportation Equipments and Securities & Commodity Futures and the risk free rate is proxied by the commercial paper rate. The data are weekly returns from 1999/12/28 to December 2011/4/29.

4.1 Estimation of the GARCH and moment matrices

In Table 1, the estimated parameters of the GARCH model are reported. The co-skewness matrix is 4×4^2 and the co-kurtosis matrix is 4×4^3 and only non-redundant elements are displayed. As the correlation matrix shows, the indexes are positively correlated so that there is little gain from diversification that reducing the variance of a portfolio. However, there is an another type of gain from diversification that increases a skewness or reduces a kurtosis. Under the multivariate normal distribution, $\mathrm{E}[\epsilon_{it}^3] = 0, \mathrm{E}[\epsilon_{it}\epsilon_{jt}^2] = 0, \mathrm{E}[\epsilon_{it}^4] = 3, \mathrm{E}[\epsilon_{it}^2\epsilon_{jt}^2] = 1, \mathrm{E}[\epsilon_{it}\epsilon_{jt}^3] = 0$. Co-skewnesses $\mathrm{E}[\epsilon_{1t}, \epsilon_{2t}^2]$ can be interpreted as correlation of ϵ_{1t} and ϵ_{2t}^2. The negative co-skewness $\mathrm{E}[\epsilon_{1t}\epsilon_{2t}^2] = -0.11$ implies that when the Mining index experience large return in absolute value, the Iron & Steel index tends to experience negative return. Since rare event of Mining and negative return of Iron & Steel are correlated, this combination raises the possibility of large negative return. This has negative effect on skewness of a portfolio return, and thus this combination can be interpreted as risky. Co-kurtosis $\mathrm{E}[\epsilon_1^2\epsilon_2^2]$ can be interpreted as the covariance of ϵ_{1t}^2 and ϵ_{2t}^2. The positive co-kurtosis $\mathrm{E}[\epsilon_{1t}^2\epsilon_{2t}^2] = 1.69$ implies that when the Mining index experiences a rare event, Iron & Steel index also tends to experience rare event. This combination raises the possibility of very rare event, and so can be interpreted as risky. These bootstrap resamplings are repeated 100 times with the procedure shown in section 3. Bootstrap standard errors are reported in parentheses.

Table 1: GARCH parameters and moment matrices of innovations

With the estimated GARCH parameters and higher moment matrices of innovations, conditional moment matrices are estimated every week. Given the conditional moments, optimal asset allocations are estimated for each strategy. As in Jondeau and Rockinger (2012), I use the parameters of risk aversion $\lambda = 5, 10, 15$. When this value is large the investor is more risk averse. In the following subsections the performance of each strategy is reported.

4.2 Descriptive statistics of the realized return

Table 2 represents the descriptive statistics of the realized return for each strategy. Each row reports the summary statistics of the portfolio return of the dynamic higher moment portfolio, dynamic mean-variance portfolio, static higher moment portfolio and static mean-variance portfolio. Two static portfolios shows very similar performance. This implies that higher moment does not has value unconditionally. The dynamic higher moment portfolio have large mean, large variance, large skewness and large kurtosis than that of dynamic mean variance portfolio. In terms of the mean variance criterion, the portfolio may be interpreted as a high risk, high return portfolio; however, in terms of the expected utility, the risk is not so large since relatively large skewness makes the probability of negative return small.

Table 2: Summary statistics of the realized returns

4.3 Expected utility and certainty equivalent returns

For several values of risk aversion, certainty equivalent returns are estimated. Certainty equivalent returns are measured between the dynamic higher moment portfolio and other strategies respectively. Table 3 presents the computation results. The first column shows the certainty equivalent return for each strategy and the dynamic higher moment portfolio. The second column shows the bootstrap p-values explained in section 3. Figure 1 presents the bootstrap distribution of the certainty equivalent returns with $\lambda = 15$. The result shows that the dynamic higher moment portfolio significantlly improve the performance when it is compared to dynamic mean-variance portfolio. This impies that considering higher moments is valuable in term of the expected utility when it is combined to the volatility timing. The value 0.0000056 of the certainty equivalent return between the dynamic higher moment portfolio and the dynamic mean variance portfolio implies that investors are willing to pay

0.0000056 to switch the strategy from dynamic mean variance portfolio to dynamic higher moment portfolio. Since average risk free rate during this period is 0.0000385, it can be interpreted that considering higher moments has about seventh part of value of the risk free rate when $\lambda = 15$. Also figure 1 shows, the most of simulated certainty equivalent returns are positive. Comparing to static mean variance, or static higher moment portfolio, certainty equivalent returns are 0.000020 but is not statistically significant. As figure 2 and 3 show, the value of the certainty equivalent returns are volatile. This imply that the value of volality timing is ambiguous.

Table 3: Certainty equivalent returns and bootstrap p-values

Figure 1: Bootstrap distribution of the certainty equivalent returns (a) Dynamic mean variance portfolio

Figure 2: Bootstrap distribution of the certainty equivalent returns (b) Static higher moment portfolio

Figure 3: Bootstrap distribution of the certainty equivalent returns (c) Static mean variance portfolio

5 Conclusion and future work

In this paper a simple asset allocation problem with higher moments and GARCH effects is introduced. The approach is easy to estimate since the time series estimations are based on the standard estimation of the GARCH and portfolio optimizations are solved by the Newton-Rapthon method with explicitly computed derivative vectors and Hessian matrices. The performance of the strategies is measured by the bootstrap method. The empirical results demonstrated that conditional higher moments with GARCH effects improve an investor's utility. For most risk aversion parameters, the certainty equivalent return that investors are willing to pay to switch their strategy to the portfolio with higher moments and GARCH effects is positive. This implies that information of the higher moments of the return has value at some situations where non-normalities are observed. The paper considers only a simple model with constant moment matrices of the GARCH innovations. However the dynamic conditional correlation model is applicable. Additionally, conditional skewness and kurtosis should be topics for future research.

References

Adcock, C.J. (2010), 'Asset pricing and portfolio selection based on the multivariate extended-skew Student-t distribution" *Annual of Operations Research*, 176, 221-234.

Ang, C.J. and Bekaert (2002), 'International asset allocation with regime shifts" *Review of financial studies*, 15, 1137-1187.

Athayde, G.M.d. and Flores, R.G. (2004), 'Finding a maximum skewness portfolio-a general solution to three-moments portfolio choice," *Journal of Economic Dynamics and Control*, 28, 1335-1352.

Azzalini, A. and Capitanio, A. (2003), 'Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," *Journal of the Royal Statistical Society: Series B* , 65, 367-389.

Bauwens, L. and Laurent, S. (2003), 'A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models," *Journal of Economic Dynamics and Control*, 28, 1335-1352.

Engle, R. (2002), 'Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20, 339-350.

Fleming, J., Kirby, C. and Ostdiek, B. (2001), 'The economic value of volatility timing," *Journal of Finance*, 56, 329-352.

Glosten, L., Jagannathan, R. and Runkle, D. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779-1801.

Guidolin, M. and Timmerman, A. (2008), 'International asset allocation under regime switching, skew and kurtosis preference," *Review of Financial Studies*, 21, 889-935.

Hansen, B.E. (1994), 'Autoregressive conditional density estimation," *International Economic Review*, 35, 705-730.

Harvey, C.R. and Siddique, A. (1999), 'Autoregressive conditional skewness," *Journal of Financial and Quantitative Analysis*, 34, 465-487.

Harvey, C.R., Liechty, J.C. and Liechty, M.W (2010), 'Portfolio selection with higher moments," *Quantitative Finance*, 10, 469-485.

Jondeau, E. and Rockinger, M. (2003), "Conditional volatility, skewness and kurtosis: Existence, persistence, and comovements," *Journal of Economic Dynamic and Control*, 27, 1699-1737.

Jondeau, E. and Rockinger, M. (2006), "Optimal asset allocation under higher moments," *European Financial Management*, 12, 29-35.

Jondeau, E. and Rockinger, M. (2007), "The copula-garch model of conditional dependencies: An international stock market application," *Journal of International Money and Finance*, 25, 827-853.

Jondeau, E. and Rockinger, M. (2012), "On the importance of time variability in higher moments for asset allocation," *Journal of Financial Econometrics*, 10, 84-123.

Longin, F. and Solnik, B. (2001), 'Extreme correlation of international equity markets," *Journal of Finance*, 56, 649-676.

Konno, H., Shirakawa, H. and Yamazaki, H. (1993), ‘A mean-absolute deviation-skewness portfolio optimization model,” *Annual of Operations Research*, 45, 205-220.

Markowitz, H. (2009), ‘Portfolio selection,” *Journal of Finance*, 7, 77-91.

Mencia, J. and Sentana, E. (2009), ‘Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation,” *Journal of Econometrics*, 153, 105-121.

Patton, A.J. (2004), ‘On the out of sample importance of skewness ans asymmetric dependence for asset allocation,” *Journal of Financial Econometrics*, 2, 130-168.

Sahu, S.K., Dey, D.K. and Branco, M.D. (2003), ‘A new class of multivariate skew distributions with applications to bayesian regression models,” *Canadian Journal of Statistics*, 31, 129-150.

Tauchen, G. and Hussey, R. (1991), ‘Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models,” *Econometrica*, 59, 371-396.

Table 1: GARCH parameters and moment matrices of innovations

	μ	$\bar{\sigma}^2$	α	β	γ			
Mining	0.00098	0.0021	0.88	0.052	0.12			
	(0.0000)	(0.0001)	(0.14)	(0.0005)	(0.0015)			
Iron & Steel	0.0011	0.0019	0.91	0.0738	0.18			
	(0.0000)	(0.0002)	(0.16)	(0.0008)	(0.0010)			
Transportation	0.0010	0.0013	0.81	0.02	0.22			
Equipments	(0.0000)	(0.0002)	(0.15)	(0.0010)	(0.0017)			
Securities &	0.00038	0.0030	0.91	0.021	0.017			
Commodity Futures	(0.0002)	(0.0001)	(0.14)	(0.0005)	(0.0023)			
Correlation matrix								
	1.00							
	0.47	(0.00)	1.00					
	0.34	(0.00)	0.55	(0.00)	1.00			
	0.31	(0.00)	0.59	(0.02)	0.59	(0.01)	1.0000	
Co-skewness matrix								
S_1	0.00	(0.07)						
	-0.11	(0.02)	-0.04	(0.01)				
	-0.20	(0.01)	-0.17	(0.01)	-0.20	(0.02)		
	-0.06	(0.01)	-0.03	(0.00)	-0.06	(0.00)	0.04	(0.00)
S_2								
			0.21	(0.03)				
			-0.14	(0.02)	-0.24	(0.01)		
			0.07	(0.01)	-0.07	(0.01)	0.12	(0.01)
S_3								
					-0.18	(0.07)		
					-0.14	(0.02)	0.06	(0.01)
S_4								
							0.51	(0.03)

Co-kurtosis matrix								
K_{11}	3.78	(3.66)						
	1.69	(0.46)	1.80	(0.16)				
	1.22	(0.28)	1.16	(0.19)	1.46	(0.24)		
	1.12	(0.11)	1.14	(0.03)	0.93	(0.03)	1.38	(0.01)
K_{12}								
			1.66	(0.18)				
			1.08	(0.21)	1.10	(0.27)		
			1.02	(0.03)	0.71	(0.04)	0.97	(0.01)
K_{13}								
					1.17	(0.48)		
					0.74	(0.06)	0.70	(0.01)
K_{14}								
							1.02	(0.03)
K_{21}								
			3.70	(0.22)				
			1.93	(0.33)	1.93	(0.45)		
			1.96	(0.07)	1.36	(0.08)	1.86	(0.02)
K_{22}								
					2.00	(0.76)		
					1.38	(0.12)	1.29	(0.03)
K_{23}								
							1.89	(0.07)
K_{33}								
					3.57	(1.51)		
					1.87	(0.24)	1.72	(0.05)
K_{34}								
							1.85	(0.03)
K_{44}								
							3.69	(0.30)

Table 2: Summary statistics of the realized returns

	Mean	Sd	Skew	Kurt	Max	Min
$\lambda = 5$						
Dynamic higher moment	0.00047	0.0010	-0.14	1.06	0.034	-0.038
Dynamic mean-variance	0.00036	0.00080	-0.19	0.91	0.028	-0.033
Static higher moment	0.00028	0.068	-0.41	5.55	0.043	-0.43
Static mean-variance	0.00027	0.068	-0.42	5.56	0.043	-0.43
$\lambda = 10$						
Dynamic higher moment	0.00026	0.0050	-0.13	1.06	0.017	-0.019
Dynamic mean-variance	0.00020	0.0040	-0.19	0.91	0.014	-0.016
Static higher moment	0.00016	0.0034	-0.41	5.55	0.043	-0.022
Static mean-variance	0.00016	0.0034	-0.42	5.55	0.043	-0.022
$\lambda = 15$						
Dynamic higher moment	0.00018	0.0033	-0.13	1.06	0.012	-0.013
Dynamic mean-variance	0.00014	0.0026	-0.19	0.91	0.0092	-0.010
Static higher moment	0.00012	0.0023	-0.42	5.55	0.14	-0.014
Static mean-variance	0.00012	0.0023	-0.42	5.55	0.14	-0.014

Table 3: Certainty equivalent returns and bootstrap p-values

	C.E	P-value
$\lambda = 5$		
Dynamic mean-variance	0.000017	0.02
Static higher moment	0.000061	0.14
Static mean-variance	0.000061	0.15
$\lambda = 10$		
Dynamic mean-variance	0.0000082	0.01
Static higher moment	0.000031	0.14
Static mean-variance	0.000031	0.14
$\lambda = 15$		
Dynamic mean-variance	0.0000056	0.01
Static higher moment	0.000020	0.14
Static mean-variance	0.000020	0.14

C.E means the certainty equivalent returns between Dynamic higher moment portfolio and other portfolios.

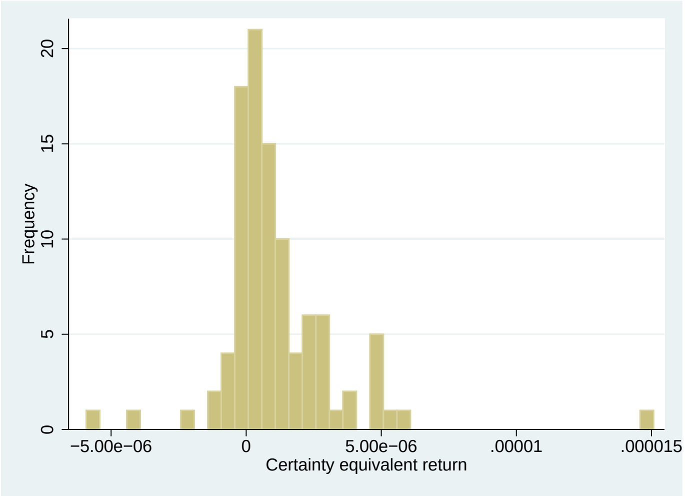

Figure 1: Bootstrap distribution of the certainty equivalent returns

(a) Dynamic mean variance portfolio

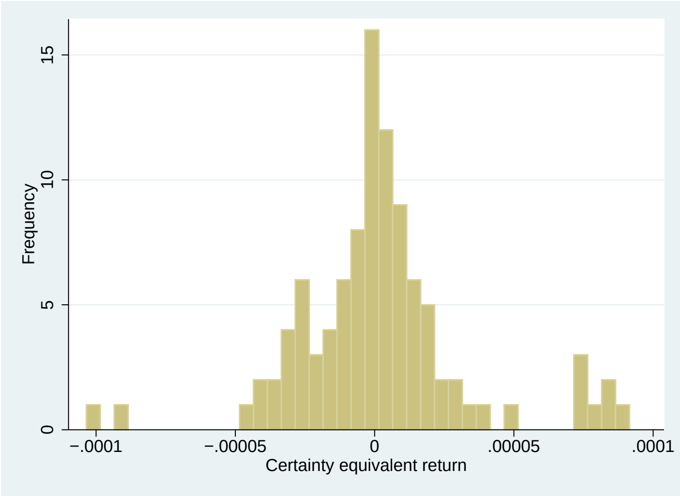

Figure 2: Bootstrap distribution of the certainty equivalent returns

(b) Static higher moment portfolio

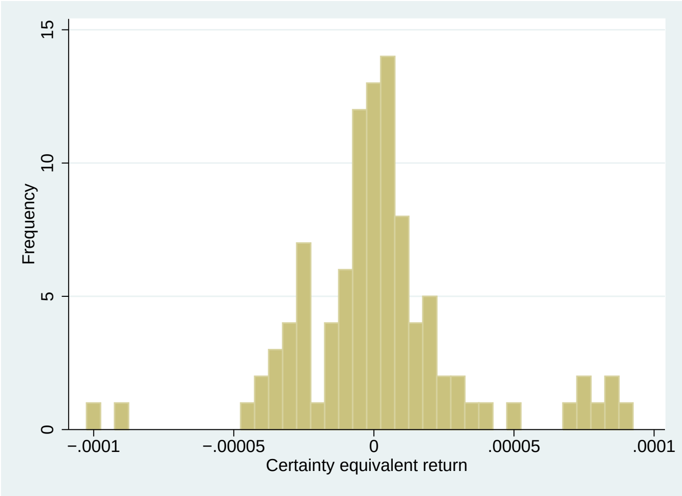

Figure 3: Bootstrap distribution of the certainty equivalent returns
(c) Static mean variance portfolio