Future hedge: Static case

Student : Jang-Yi Li Advisor: Dr. Yuan-Chung Sheu

Department of Applied Mathematics
National Chiao Tung University
Hsinchu, 300, Taiwan, R.O.C.

June 2005

Abstract

This paper presents a review of different theoretical approaches to the optimal futures hedge ratios. Under current futures prices are unbiased, different hedge ratios are the same as the minimum variance hedge ratio. We introduce the class of hyperbolic distributions which can be fitted to the empirical log-returns with high accuracy, and simulate it to estimate GSV hedge ratio. Last, we compare these futures hedge ratios in several markets. In particular, we propose conditional correlations with a bivariate hyperbolic distributions method to dominate GSV hedge ratios.

Contents

List of Tables

List of Figures

1 Introduction

One of the best uses of derivative securities such as futures [1] contracts is in hedging. The basic concept of hedging is to combine investments in the spot market and futures market to form a portfolio that will eliminate fluctuations in its value. Hence, There is a considerable body of literature on the determination of optimal hedge ratios based on different criteria. For example, one of the most widely-used hedging strategies is based on the minimization of the variance of the hedged portfolio. The minimum-variance hedge ratios (abbreviated MVHs) completely include risk of the hedged portfolio. However, MVHs are popular because they are easy to estimate with econometric methods.

After the axiomatization of the expected utility hypothesis by John von Neumann and Oskar Morgenstern (1944), economists began immediately seeing the potential applications of expected utility to economic issues like portfolio choice, insurance, etc. These simple applications tended to use simple models where outcomes were expressed as a single commodity, "wealth", thus the set of outcomes X, became merely the real line, R. However, utility maximizing hedging strategies require the specification of a utility function, and the most commonly proposed utility functions have been quadratic, logarithmic, and exponential functions. Furthermore, Benninga, Eloder, and Zilcha [8] have shown that, under certain conditions, MVHs are consistent with expected-utility maximization regardless of the utility function chosen. No matter what type of function is chosen, there is always the concern that the proposed function will not accurately describe the preferences of hedgers.

In recent years using generalized semivariance for determining the hedge ratio has been suggested (see, Chen, Lee and Shrestha[11], Lien and Tse [15][16]). The relationship between the generalized semivariance (abbreviated GSV) and expected utility as discussed by Bawa [5]. The GSV, due to its emphasis on the returns below the target return, is relation with the risk perceived by managers(see Lien and Tse [17]). It can be shown that $n < 1$ represents a risk-seeking investor and $n > 1$ represents a risk-averse investor by Fishburn [13].
Lien and Tse [15] have shown that if the and spot returns are jointly and normally

[1] In generally, a hedger may be choose options to hedge spot on the basis of lower partial moments. However, it is proved that futures outperformed options in generalized semivariance criterion by Lien[17].

distributed and if the future price is unbiased (i.e., expected futures price change is zero) the GSV hedge ratio will be equal to the minimum-variance hedge ratio.

Distributional assumptions for the returns on the underlying assets play a key role in valuation theories for derivative securities. A class of distributions that is very often able to fit the distributions of financial data is the class of generalized hyperbolic distributions. The class of generalized hyperbolic distributions includes the hyperbolic distributions, the normal inverse Gaussian distributions, and variance-gamma distributions in the financial literature. The hyperbolic distribution was introduced by Barndorff-Nielsen [3]. Almost twenty years later the generalized hyperbolic law was found to provide a very good model for the distributions of daily returns of stocks from a number of leading German enterprisers (see, Eberlein [12]), such that generalized hyperbolic family of probability laws are investigated and discussed by Bibby and Sørensen [9], Eberlein [12], Prause[18], Barndorff-Nielsen [3], and so on. Computer algorithms are described for simulation of the generalized inverse Gaussian, generalized hyperbolic and hyperbolic distributions. The efficiencies of the algorithms are found by Atkinson [2].

Understanding the relation between volatility and correlation is imperative for financial analysts who measure and manage market risks. Recent research on domestic and international stock markets suggests the notion of correlation asymmetry. Generally speaking, it is found that correlations between stocks and the aggregate market are much greater on the downside that on the upside (see, Ang [1]).

It is quite clear that there are several different criteria, which have different hedge ratios. This paper is laid out in main aims. First, we review these different techniques and approaches in order to examine their relations. Second, we find an extension to generalized hyperbolic distribution fit data better than generalized hyperbolic distribution. Third, we adopt the class of hyperbolic distributions to numerically estimate hedge ratio in GSV criteria, and investigate GSV criterion about different target return. Fourth, correlation asymmetry is also an important issue for risk managers who are concerned with downside risk particularly when the market is bearish or unusually volatile. Finally, we compare these futures ratios, and provide a summary and conclusions.

2 The Generalized Hyperbolic and Inverse Gaussian Distributions

Almost twenty years later the hyperbolic law was found to provide a very good model for the distributions of daily log-returns of derivative securities. Hyperbolic distributions are characterized by their log-density being a hyperbola. Recall that for normal distributions the log-density is a parabola, Hence the hyperbolic distribution provides the possibility of modelling heavier tails.

The class of generalized hyperbolic distributions (abbreviated **GH**) can be obtained by mean-variance mixtures of normal distributions where the mixing distribution is a generalized inverse Gaussian distribution (abbreviated **GIG**). More precisely, a random variable Z has the generalized hyperbolic distribution if:

$$Z|Y = y \sim N(\mu + \beta y, y),$$

where Y is a **GIG** random variable and $N(\mu + \beta y, y)$ denotes a normal distribution with mean $\mu + \beta y$ and variance y. The **GIG** with parameters $\delta, \psi > 0$ has density

$$d_{GIG(\lambda,\delta,\psi)}(x) = \frac{(\psi/\delta)^{\lambda}}{2K_{\lambda}(\delta\psi)} x^{\lambda-1} e^{-\frac{1}{2}(\delta^2 x^{-1} + \psi^2 x)} \mathbb{I}_{\{x>0\}}, \tag{1}$$

where the normalizing constant:

$$K_{\lambda}(t) = \frac{1}{2}\int_0^{\infty} x^{\lambda-1} e^{-\frac{1}{2}t(x^{-1}+x)}\, dx, \qquad t > 0, \tag{2}$$

is the modified Bessel function of the third kind with index λ. If the **GIG**(λ, δ, ψ) law with parameter $\lambda = -\frac{1}{2}$, it is called the inverse Gaussian distribution (abbreviated **IG**). Its mean and variance are

$$\delta/\psi, \qquad \text{and} \qquad \delta/\psi^3$$

respectively. For other values of λ numerical approximations of the integral in equation (2). The hyperbolic distribution is defined as a normal variance-mean mixture where the mixing distribution is the **GIG**(λ, δ, ψ) law with parameter $\lambda=1$.

In the context of generalized hyperbolic distributions, the Bessel functions are thoroughly discussed in Barndorff-Nielsen [3]. Here we make summary of properties that will be used later.

Lemma 2.1.

$$
\begin{aligned}
K_{\pm\frac{1}{2}}(t) &= \sqrt{\frac{\pi}{2t}}e^{-t} \\
K_{\lambda}(t) &= K_{-\lambda}(t) \\
K_{\lambda+1}(t) &= \frac{2\lambda}{t}K_{\lambda}(t) + K_{\lambda-1}(t) \quad \textit{(by parts)} \\
\frac{dK_{\lambda+1}(t)}{dt} &= \frac{-1}{2}\left(K_{\lambda+1}(t) + K_{\lambda-1}(t)\right) \\
&= \frac{-\lambda}{t}K_{\lambda}(t) - K_{\lambda-1}(x) \\
K_{\lambda}(t) &\approx \sqrt{\frac{\pi}{2t}}e^{-t} \quad \textit{for} \quad t \to \infty \\
\textit{and} \quad K_{\lambda}(t) &\approx \frac{1}{2}\Gamma(\lambda)\left(\frac{t}{2}\right)^{-\lambda}, \quad \forall\lambda > 0,\ \textit{as} \quad t \to 0,
\end{aligned}
$$

where $\Gamma(\lambda)$ *is gamma function.*

Then the probability density function of **GH** can be calculated by

$$
\begin{aligned}
& d_{GH(\lambda,\beta,\delta,\psi,\mu)}(x) = \int_0^{\infty} d_{N(\mu+\beta y,y)}(x)d_{GIG(\lambda,\delta,\psi)}(y)\, dy \\
=& \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \frac{(\psi/\delta)^{\lambda}}{2K_{\lambda}(\delta\psi)} y^{\lambda-1-\frac{1}{2}} e^{\frac{-1}{2}\left[\frac{1}{y}\left(\delta^2+(x-\mu)^2\right)+y\left(\beta^2+\psi^2\right)\right]+\beta(x-\mu)}\, dy \\
=& \left(\frac{\psi}{\delta}\right)^{\lambda} \frac{e^{(x-\mu)\beta}}{\sqrt{2\pi}K_{\lambda}(\delta\psi)} \left[\frac{\delta^2+(x-\mu)^2}{\beta^2+\psi^2}\right]^{\frac{\lambda-\frac{1}{2}}{2}} K_{\lambda-\frac{1}{2}}\left(\left[\left(\beta^2+\psi^2\right)\left(\delta^2+(x-\mu)^2\right)\right]^{\frac{1}{2}}\right).
\end{aligned}
$$

Hence, the probability density function of the hyperbolic d_H can be written as:

$$
d_{H(\beta,\delta,\psi,\mu)}(x) = \left(\frac{\psi}{\delta}\right) \frac{e^{-\sqrt{(\beta^2+\psi^2)(\delta^2+(x-\mu)^2)}+(x-\mu)\beta}}{2\sqrt{\beta^2+\psi^2}K_1(\delta\psi)} \tag{3}
$$

Sometimes another parametrization of the hyperbolic distribution with $\alpha = \sqrt{\beta^2+\psi^2}$ is used in order to represent the moment generating function clearly.

Then the probability density function of the hyperbolic $\mathbf{H}(\alpha, \beta, \delta, \mu)$ law can be written as:

$$
d_{H(\alpha,\beta,\delta,\mu)}(x) = \frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2-\beta^2})} e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)}
$$

The normal inverse Gaussian (abbreviated **NIG**) distributions were introduced by Barndorff-Nielsen (1995) as a subclass of the generalized hyperbolic distribution obtained for $\lambda =$

$-1/2$. The density of the **NIG** distribution d_{NIG} is given by:

$$d_{NIG(\beta,\delta,\psi,\mu)}(x) = \frac{\delta}{\pi}\left[\frac{\beta^2+\psi^2}{\delta^2+(x-\mu)^2}\right]^{\frac{1}{2}} e^{\delta\psi+(x-\mu)\beta} K_1(\left[\left(\beta^2+\psi^2\right)\left(\delta^2+(x-\mu)^2\right)\right]^{\frac{1}{2}}).$$

Because of the exponential form of the density (1), the moment generating function of **GIG**(λ, δ, ψ) distributions is simply the ratio of the forming constants corresponding to the parameters $(\lambda, \delta, \sqrt{\psi^2 - 2u})$ and (λ, δ, ψ):

$$M_{GIG(\lambda,\delta,\psi)}(u) = \int_0^\infty e^{ux} d_{GIG(\lambda,\delta,\psi)}(x)dx = \left(\frac{\psi}{\sqrt{\psi^2-2u}}\right)^\lambda \frac{K_\lambda(\delta\sqrt{\psi^2-2u})}{K_\lambda(\delta\psi)} \tag{4}$$

with the restriction $2u < \psi^2$.

We now consider the case $\delta \to 0^+$ to obtain the limit moment-generating function. From Lemma 2.1 we conclude

$$\lim_{\delta\to 0^+} M_{GIG(\lambda,\delta,\psi)}(u) = \left(1 - \frac{2}{\psi^2}u\right)^{-\lambda}$$

which is the moment-generating function of the $\Gamma_{(\lambda,\psi^2/2)}$[2] -distribution. Further, the **GIG** distribution satisfies the following property. For completeness we provide below the well-know formula for the moments of the **GIG** and properties of the **GIG** distribution. We summarize the results here:

Lemma 2.2. *Let $X \sim GIG(\lambda, \delta, \psi)$ and $\Gamma_{(\lambda,\psi^2/2)}$ random variables be independent Then*

$$\mathbb{E}[X^r] = \left(\frac{\delta}{\psi}\right)^r \frac{K_{\lambda+r}(\delta\psi)}{K_\lambda(\delta\psi)}, \quad \gamma, \psi > 0 \tag{5}$$

$$\begin{aligned} \boldsymbol{GIG}(\lambda, 0, \psi) &\stackrel{\mathcal{D}}{=} \Gamma_{(\lambda,\psi^2/2)} \\ c\boldsymbol{GIG}(\lambda, \delta, \psi) &\stackrel{\mathcal{D}}{=} \boldsymbol{GIG}(\lambda, \delta\sqrt{c}, \psi/\sqrt{c}), \ \forall\, c > 0 \\ \boldsymbol{GIG}(\lambda, \delta, \psi) &\stackrel{\mathcal{D}}{=} \boldsymbol{GIG}(-\lambda, \delta, \psi) + \Gamma_{(\lambda,\psi^2/2)} \\ \boldsymbol{GIG}(\lambda, \delta, \psi) &\stackrel{\mathcal{D}}{=} \frac{1}{\boldsymbol{GIG}(-\lambda, \psi, \delta)} \end{aligned} \tag{6}$$

,where $\stackrel{\mathcal{D}}{=}$ denote distributional equivalences.

Proof It is provided in Appendices.

[2]The density of a $\Gamma_{\lambda,\beta}$-distribution is given $f(x) = \frac{\beta^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-x\beta}$.

From Lemma 2.2, **GIG** easily follows that it's mean and variance are

$$\left(\frac{\delta}{\psi}\right)\frac{K_{\lambda+1}(\delta\psi)}{K_\lambda(\delta\psi)}, \quad \text{and} \quad \left(\frac{\delta}{\psi}\right)^2\left[\frac{K_{\lambda+2}(\delta\psi)}{K_\lambda(\delta\psi)} - \frac{K^2_{\lambda+1}(\delta\psi)}{K^2_\lambda(\delta\psi)}\right].$$

Using (4), we immediately get for $|\beta+u|<\alpha$

$$\begin{aligned} M_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u) &= e^{u\mu}M_{GIG(\lambda,\delta,\sqrt{\alpha^2-\beta^2})}\left(\frac{u^2}{2}+u\beta\right) \\ &= e^{u\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}}\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})} \end{aligned} \tag{7}$$

GH as well as the via (4) associated **GIG** mixing distributions therefore possess moments of arbitrary order. In particular, we taking derivatives of M_{GH} to get the results as follows:

$$\begin{aligned} \mu_{GH(\lambda,\alpha,\beta,\delta,\mu)} &= \mu + \beta\mathbb{E}[GIG(\lambda,\delta,\psi)], \\ Var[GH(\lambda,\alpha,\beta,\delta,\mu)] &= \mathbb{E}[GIG(\lambda,\delta,\psi)] + \beta^2 Var[GIG(\lambda,\delta,\psi)] \\ \mathbb{E}[\{GH(\lambda,\alpha,\beta,\delta,\mu)-\mu_{GH}\}^3] &= 3\beta Var[GIG(\lambda,\delta,\psi)] + \beta^3\mathbb{E}[\{GIG(\lambda,\delta,\psi)-\mu_{GIG}\}^3] \\ \mathbb{E}[\{GH(\lambda,\alpha,\beta,\delta,\mu)-\mu_{GH}\}^4] &= 3\mathbb{E}[GIG^2(\lambda,\delta,\psi)] + \beta^4\mathbb{E}[\{GIG(\lambda,\delta,\psi)-\mu_{GIG}\}^4] \\ &+ 6\beta^2\mathbb{E}[GIG(\lambda,\delta,\psi)\{GIG(\lambda,\delta,\psi)-\mu_{GIG}\}^2], \end{aligned}$$

where μ_{GIG} is mean of $GIG(\lambda,\delta,\psi)$.
Thus, we know that the normal inverse Gaussian distribution **NIG**$(\alpha,\beta,\delta,\mu)$ with the mean and variance of the distribution are

$$\mu + \frac{\delta}{\alpha\sqrt{1-(\beta/\alpha)^2}} \quad \text{and} \quad \frac{\delta}{\alpha(1-(\beta/\alpha)^2)^{3/2}}$$

using by (7). It is obtain from moment M_{GH}. It means sums of independent **GH** distributed random variables are no longer **GH** distributed. The same holds for **GIG** distribution. The only exception is the subclass of normal inverse Gaussian distributions belonging to the parameter $\lambda=-1/2$. Its moment generating function is given by

$$M_{NIG}(\alpha,\beta,\delta,\mu)(u) = e^{u\mu}\frac{e^{\delta\sqrt{\alpha^2-\beta^2}}}{e^{\delta\sqrt{\alpha^2-(\beta+u)^2}}}.$$

Using the moment-generating-function technique, we obtain that independent normal inverse Gaussian random variables add in the following way:

$$NIG(\alpha,\beta,\delta_1,\mu_1) + NIG(\alpha,\beta,\delta_2,\mu_2) = NIG(\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2).$$

Further, we propose an extension to the **GH** family of probability laws. In particular we add one more scaling parameter. Consider a random variable Z' has distribution if:

$$Z'|Y = y \sim N(\mu + \beta y, \xi y),$$

where Y is a **GIG** random variable, $N(\mu + \beta y, \xi y)$ denotes a normal distribution with mean $\mu + \beta y$, variance ξy, and ξ is a positive constant. Then the probability density function of **GH** can be calculated by

$$d_{GH(\lambda,\beta,\delta,\psi,\mu,\xi)}(x) = \\ (\frac{\psi}{\delta})^{\lambda} \frac{e^{(x-\mu)\beta/\xi}}{\sqrt{2\pi\xi}K_{\lambda}(\delta\psi)} \left[\frac{\delta^2 + \frac{(x-\mu)^2}{\xi}}{\frac{\beta^2}{\xi} + \psi^2}\right]^{\frac{\lambda-\frac{1}{2}}{2}} K_{\lambda-\frac{1}{2}}(\left[\left(\frac{\beta^2}{\xi} + \psi^2\right)\left(\delta^2 + \frac{(x-\mu)^2}{\xi}\right)\right]^{\frac{1}{2}}). \quad (8)$$

In particular, the probability density function of the hyperbolic $\mathbf{H}(\alpha, \beta, \delta, \mu, \xi)$ law can be written as:

$$d_{H(\beta,\delta,\psi,\mu,\xi)}(x) = \left(\frac{\psi}{\delta\sqrt{\xi}}\right) \frac{e^{-\sqrt{(\frac{\beta^2}{\xi}+\psi^2)(\delta^2+\frac{(x-\mu)^2}{\xi})}+\frac{(x-\mu)\beta}{\xi}}}{2\sqrt{\frac{\beta^2}{\xi}+\psi^2}K_1(\delta\psi)}$$

when $\xi = 1$ the above formula it has become (3). In similarly, we obtain that mean and variance are given by

$$\begin{aligned}
\mathbb{E}[GH(\lambda,\beta,\delta,\psi,\mu,\xi)] &= \mu + \beta\mathbb{E}[GIG(\lambda,\delta,\psi)], \\
Var[GH(\lambda,\beta,\delta,\psi,\mu,\xi)] &= \xi\mathbb{E}[GIG(\lambda,\delta,\psi)] + \beta^2 Var[GIG(\lambda,\delta,\psi)] \\
\mathbb{E}[\{GH(\lambda,\alpha,\beta,\delta,\mu) - \mu_{GH}\}^3] &= 3\beta\xi Var[GIG(\lambda,\delta,\psi)] + \beta^3\mathbb{E}[\{GIG(\lambda,\delta,\psi) - \mu_{GIG}\}^3] \\
\mathbb{E}[\{GH(\lambda,\alpha,\beta,\delta,\mu) - \mu_{GH}\}^4] &= 3\xi^2\mathbb{E}[GIG^2(\lambda,\delta,\psi)] + \beta^4\mathbb{E}[\{GIG(\lambda,\delta,\psi) - \mu_{GIG}\}^4] \\
&+ 6\beta^2\xi\mathbb{E}[GIG(\lambda,\delta,\psi)\,\{GIG(\lambda,\delta,\psi) - \mu_{GIG}\}^2].
\end{aligned}$$

where μ_{GIG} is mean of $GIG(\lambda, \delta, \psi)$.
There are three reasons for our ξ. First, we extend $\mathbf{GH}(\lambda, \beta, \delta, \psi, \mu)$ to $\mathbf{GH}(\lambda, \beta, \delta, \psi, \mu, \xi)$ $\forall\ \xi > 0$. Second, $\sqrt{Y}$ statements volatility which is important quality in market. So it is given a parameter, such that it fit our data more precisely. Final, except spot of S&P500, we find that we adopt maximum-Likelihood estimation, [3] which is a fast and has closed

[3]Blæsid and Sørensen [10] had used to estimate parameters of hyperbolic distribution by the method , we will be appropriate in our section 5.3.1.

formula , to estimate our parameters, and theoretical $\mathbb{E}[GH(\lambda,\beta,\delta,\psi,\mu)]$ match with sample mean, but sample variance does not in our data (see Table 7). Thus, we can adjust theoretical $Var[GH(\lambda,\beta,\delta,\psi,\mu)]$ to match sample variance.

2.1 Multivariate distribution models

In econometrical finance the latter family of distributions has been used for multidimensional asset-return modelling. In this context, the generalized hyperbolic distribution replaces the Gaussian distribution that is not able to describe the fat tails and the skewness of most financial asset-return data. Reference is Eberlein, Keller [12], and Schmidt[19]. In fact that the multivariate **GH** distributions (abbreviated **MGH**) is also driven by a univariate mixing distribution is sufficient for many modelling purposes in finance, since the variance of all assets in a particular market usually change together. Multivariate generalized hyperbolic distributions were introduced and investigated by Barndorff-Nielsen [4]. The rough concept are stated as following:
Let the random variable **Z** be **MGH** distributions via the following variance-mean mixtures of multivariate normal distributions, i.e.,

$$\mathbf{Z}|Y=y \sim N_d(\mu+y\Delta\beta, y\Delta),$$

where $\Delta\in\mathbb{R}^{d\times d}$ is a symmetric positive-definite matrix with determinant $|\Delta|=1$ and $\mu,\beta\in\mathbb{R}^d$. The following density representation, which is given by Barndorff-Nielsen [4], is appropriate in our context.

Definition 2.3. *The d-dimensional generalized hyperbolic distribution is defined for x* $\in\mathbb{R}^d$

$$\begin{aligned} d_{MGH}(x) &= c_d\frac{K_{\lambda-d/2}\left(\alpha\sqrt{\delta^2+(x-\mu)'\Delta^{-1}(x-\mu)}\right)}{\left(\alpha^{-1}\sqrt{\delta^2+(x-\mu)'\Delta^{-1}(x-\mu)}\right)^{d/2-\lambda}}e^{(\beta'(x-\mu))} \\ c_d &= \frac{\left(\sqrt{\alpha^2-\beta'\Delta\beta}/\delta\right)^{\lambda}}{(2\pi)^{d/2}K_\lambda(\delta\sqrt{\alpha^2-\beta'\Delta\beta})} \end{aligned}$$

These parameters have the following domain of variation: $\lambda\in\mathbb{R},\beta,\mu\in\mathbb{R}^d$*,* $\gamma>0,\alpha^2>\beta'\Delta\beta$*. The positive matrix* $\Delta\in\mathbb{R}^{d\times d}$ *has a determinant* $|\Delta|=1$*. For* $\lambda=(d+1)/2$ *we obtain the multivariate hyperbolic and for*$\lambda=-\frac{1}{2}$ *the multivariate normal inverse Gaussian distribution.*

Lemma 2.4. *The moment generating function of the generalized hyperbolic distribution is given by*

$$M_{MGH(\lambda,\alpha,\beta,\delta,\mu)}(u) = e^{\mu u}\left(\frac{\alpha^2-\beta'\Delta\beta}{\alpha^2-(\beta+u)'\Delta(\beta+u)}\right)^{\lambda/2}\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u)'\Delta(\beta+u)})}{K_\lambda(\delta\sqrt{\alpha^2-\beta'\Delta\beta})}$$

where $\alpha^2 > (\beta+u)'\Delta(\beta+u)$.

Among the large number of dependence measures for multivariate random vectors the covariance and the correlation matrix are still the most favorite ones in most practical applications. In similarly, we obtain that mean and variance are given by

$$\mathbb{E}[MGH(\lambda,\alpha,\beta,\delta,\mu,\Delta)] = \mu + \Delta\beta\mathbb{E}[GIG(\lambda,\delta,\psi)], \tag{9}$$

$$Var[MGH(\lambda,\alpha,\beta,\delta,\mu,\Delta)] = \Delta\mathbb{E}[GIG(\lambda,\delta,\psi)] + \Delta\beta\beta'\Delta Var[GIG(\lambda,\delta,\psi)]. \tag{10}$$

Furthermore, first, it is a result of the mixing distribution which is a univariate **GIG** law and, hence, similar in all dimensions. Second, the one-dimensional margins of **MGH** distribution are not that flexible sine the parameters α and λ relate to the enter multivariate distribution and determine the some structural shape of distribution. Thus, Schmidt [19] introduces multivariate affine generalized hyperbolic (abbreviated **MAGH**) model.

Definition 2.5. *An d-dimensional random vector X is said to be **MAGH** distributed with location vector $\mu \in \mathbb{R}^d$ and scaling matrix $\sum \in \mathbb{R}^{d\times d}$ if it has following stochastic representation:*

$$X \stackrel{\mathcal{D}}{=} A'Y + \mu$$

for some lower triangular matrix $A \in \mathbb{R}$ such that $A'A = \sum$ is positive-definite and the random vector $Y = (Y_1, ..., Y_d)'$ consists of mutually independent random variables $Y_i \in GH(\lambda_i, \beta_i, \delta_i, \psi_i, 0)$, i=1,...,d. In particular the one-dimensional margins of Y are generalized hyperbolic distributed.

An outstanding property of **MAGH** distributions is that, after an affine-linear transformation, all one-dimensional margins can be fitted separately via different generalized hyperbolic distributions.

It takes advantage of using an affine-linear transformation in our model and improve it.

3 Existing Hedge Ratios

In this section, we briefly review different hedge ratios that have been proposed. First we consider a decision maker. At the decision date ($t = 0$), the agent engages in the production of Q ($Q > 0$) commodity units for sale at the terminal date ($t = 1$) at the random cash price P_1. In addition, at the decision date he can sell X commodity units in the futures market at the price F_0, but must repurchase them at the terminal date at the random futures price F_1. Then let the initial wealth be $V_0 = P_0 Q$ and the end-of-period wealth be $V_1 = P_1 Q + (F_0 - F_1)X$, so we consider the wealth return that is

$$\begin{aligned} \widetilde{r_\theta} &= \frac{V_1 - V_0}{V_0} = \frac{P_1 Q + F_0 X - F_1 X - P_0 Q}{P_0 Q} \\ &= \frac{P_1 - P_0}{P_0} - \frac{F_1 - F_0}{F_0}\left(\frac{F_0}{P_0}\frac{X}{Q}\right) = \widetilde{r_p} - \theta \widetilde{r_f} \end{aligned} \tag{11}$$

where $\widetilde{r_p} = \frac{P_1 - P_0}{P_0}$ and $\widetilde{r_f} = \frac{F_1 - F_0}{F_0}$ are called one-period returns on the spot and futures positions, respectively, and $h = \frac{X}{Q}$ is hedge ratio. The $\theta = h\frac{F_0}{P_0}$ is so-called the adjusted hedge ratio.

Sometimes, the hedge ratio is discussed in terms of price changes (profits) instead of returns. In this case the profit on the hedged portfolio ΔV is given by

$$\begin{aligned} \Delta V &= V_1 - V_0 = P_1 Q + F_0 X - F_1 X - P_0 Q \\ &= \triangle P Q - \triangle F X = (\triangle P - \triangle F h)Q, \end{aligned} \tag{12}$$

where $\triangle P = P_1 - P_0$, and $\triangle F = F_1 - F_0$.

3.1 Minimum variance hedge ratio

The most widely-used hedge ratio is minimum variance hedge ratio which is known as the MV hedge ratio, is obtained by minimization of the variance of θ which is given by

$$Var[\widetilde{r_\theta}] = \sigma_{r_p}^2 + \theta^2 \sigma_{r_f}^2 - 2\theta\rho\sigma_{r_p}\sigma_{r_f},$$

where $\sigma_{r_p}$and σ_{r_f} are standard deviations of $\widetilde{r_p}$ and $\widetilde{r_f}$, respectively; ρ is the correlation coefficient between $\widetilde{r_p}$ and $\widetilde{r_f}$.

Consider that

$$r_{MV}(\theta) = Var[\widetilde{r_\theta}].$$

Then we have

$$r_{MV}(\theta)' = \ 2\theta\sigma_{r_f}^2 - 2\rho\sigma_{r_p}\sigma_{r_f}.$$

This shows that

$$\begin{cases} r_{MV}(\theta)' > 0, & \text{when } \theta > \rho\frac{\sigma_{r_p}}{\sigma_{r_f}} \\ r_{MV}(\theta)' < 0, & \text{when } \theta < \rho\frac{\sigma_{r_p}}{\sigma_{r_f}}. \end{cases}$$

So, in this case, the MV hedge ratio is obtained by minimizing $Var[r_\theta]$ which is given by

$$\theta_{MV}^* = \rho\frac{\sigma_{r_p}}{\sigma_{r_f}}. \tag{13}$$

If we consider the risk is given by the variance of changes in the value of the hedged portfolio as follows:

$$Var[\Delta V] = Q^2\sigma_p^2 + X^2\sigma_f^2 - 2QX\sigma_{pf},$$

where σ_p and σ_f are standard deviations of $\triangle P$ and $\triangle F$ respectively.
Then we derives this hedge ratio by chosen h such that minimize the portfolio risk. In this case, the MV hedge ratio is obtained by minimizing $Var[\Delta V]$ which is given by

$$h_{MV}^* = \frac{\sigma_{pf}}{\sigma_f^2}.$$

For a price process $S_t \in \mathbb{R}^2$ we define returns $x_t \in \mathbb{R}^2$ by

$$x_t^{(i)} = \frac{S_t^{(i)} - S_{t-1}^{(i)}}{S_t^{(i)}} \approx \log S_t^{(i)} - \log S_{t-1}^{(i)}, \qquad 1 \le i \le 2,$$

which are approximated by the log-returns. Next, for the structure of **MGH**, suppose that (r_f, r_p) is a order pair and

$$(r_f, r_p) \in \mathbf{MGH}_2(\lambda, \alpha, \beta_i, \delta, \mu_i, \Delta), \ i = 1, 2, \tag{14}$$

where $\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix}$ is symmetry.
Let

$$\begin{aligned} \delta_{ff} &= \left[\beta_1^2\Delta_{11}^2 + 2\beta_1\beta_2\Delta_{11}\Delta_{12} + \beta_2^2\Delta_{12}^2\right] Var[GIG] \\ \delta_{fp} &= \left[\beta_1^2\Delta_{11}\Delta_{12} + \beta_1\beta_2(\Delta_{11}\Delta_{22} + \Delta_{12}^2) + \beta_2^2\Delta_{12}\Delta_{22}\right] Var[GIG] \\ \delta_{pp} &= \left[\beta_1^2\Delta_{21}^2 + 2\beta_1\beta_2\Delta_{21}\Delta_{22} + \beta_2^2\Delta_{22}^2\right] Var[GIG]. \end{aligned}$$

Then, we can re-form the equation (13) given by

$$\theta_{MV}^* = \frac{\Delta_{12}\mathbb{E}[GIG] + \delta_{fp}}{\Delta_{11}\mathbb{E}[GIG] + \delta_{ff}}, \tag{15}$$

3.2 Sharpe hedge ratio

Next, we consider the class of hedge ratio that incorporate both risk and expected return in the derivation of the hedge ratios. Since Howard and D'Antonio [14] that consider the agent is seeking the greatest expected portfolio return for a given level of portfolio risk as measured by the standard deviation of portfolio return (i.e., mean and standard deviation are the only important considerations in choosing a portfolio). They consider the optimal level of futures contracts by maximizing the ratio of the portfolio's excess return to its volatility, that is

$$\max_{\theta} \frac{\mu_{r_p} - \theta\mu_{r_f} - r_L}{\sigma_{r_\theta}}, \tag{16}$$

where $\sigma_{r_\theta} = [Var(\widetilde{r_\theta})]^{\frac{1}{2}}$, r_L is the risk-free interest rate.
Consider that

$$r_s(\theta) = \frac{\mu_{r_p} - \theta\mu_{r_f} - r_L}{\sigma_{r_\theta}},$$

Then we have

$$r_s^{'}(\theta) = \frac{\theta\left[-\sigma_{r_f}^2(\mu_{r_p} - r_L) + \mu_{r_f}\sigma_{r_f r_p}\right] + (\mu_{r_p} - r_L)\sigma_{r_p r_f} - \sigma_{r_p}^2\mu_{r_f}}{\sigma_{r_\theta}^3}, \tag{17}$$

Thus, the critical point for $r_s(\theta)$ is given by

$$\theta_s^* = \frac{(\frac{\sigma_{r_p}}{\sigma_{r_f}})^2\mu_{r_f} - \rho\frac{\sigma_{r_p}}{\sigma_{r_f}}(\mu_{r_p} - r_L)}{\rho\frac{\sigma_{r_p}}{\sigma_{r_f}}\mu_{r_f} - (\mu_{r_p} - r_L)}. \tag{18}$$

To use the Second Derivative Test we obtain

$$r_s^{''}(\theta_s^*) < 0, \text{ whenever } (\mu_{r_p} - r_L) > \rho\frac{\sigma_{r_p}}{\sigma_{r_f}}\mu_{r_f}. \tag{19}$$

Hence, it is obtained from the equation (17), there is a only critical point in equation (16). Next, we turn to consider that r_f, and r_p have **MGH** distributions. In great details, when we have the same assumption, that is equation (14) , we can re-form the equation (18); that is given by

$$\theta_s^* = \frac{\zeta_{pp}\left[(\mu_1 + \beta_1\Delta_{11} + \beta_2\Delta_{12})\mathbb{E}[GIG]\right] - \zeta_{fp}\left[(\mu_2 + \beta_1\Delta_{21} + \beta_2\Delta_{22})\mathbb{E}[GIG] - r_L\right]}{\zeta_{fp}\left[(\mu_1 + \beta_1\Delta_{11} + \beta_2\Delta_{12})\mathbb{E}[GIG]\right] - \zeta_{ff}\left[(\mu_2 + \beta_1\Delta_{21} + \beta_2\Delta_{22})\mathbb{E}[GIG] - r_L\right]}, \tag{20}$$

where

$$\zeta_{ff} = \Delta_{11}\mathbb{E}[GIG] + \delta_{ff}$$
$$\zeta_{fp} = \Delta_{12}\mathbb{E}[GIG] + \delta_{fp}$$
$$\zeta_{pp} = \Delta_{22}\mathbb{E}[GIG] + \delta_{pp}.$$

In Howard and D'Antonio [14], they discussed how to decrease risk of portfolios, when know the optimal θ_s^* . Thus, the measure of hedging effectiveness (abbreviated HE) is given by

$$\text{HE} = \theta_s^* \big/ (\frac{\mu_{r_p} - r_L}{\sigma_{r_p}}), \tag{21}$$

and consider that

$$\zeta = \frac{\mu_{r_f}/\sigma_{r_f}}{(\mu_{r_p} - r_L)/\sigma_{r_p}}, \tag{22}$$

where ζ is also-called the risk-return relative and it is restrict to $\mu_{r_p} > r_L$. [4] Thus, under the second-order conditions for this to be a maximum are

$$\mu_{r_p} > r_L \quad \text{and} \quad 1 > \zeta\rho.$$

And it is discussed to the situation for which the second-order conditions. Substituting the appropriate expression in equation (22) into equation (18), we obtain

$$\theta_s^* = \frac{\sigma_{r_p}}{\sigma_{r_f}}(\frac{\rho - \zeta}{1 - \zeta\rho}). \tag{23}$$

Using equation (21), we can derive an expression for the measure of hedging effectiveness. Hence we have

$$\text{HE} = \sqrt{\frac{(\rho - \zeta)^2}{1 - \rho^2} + 1}\ .$$

We observe the agent will short futures, hence $\rho > \zeta$. Since $\rho = \zeta$, then $\theta_s^* = 0$. This is a significant result and clearly demonstrates that the holding of futures as a hedge against price risk is not simply related to ρ but depends upon both ρ and the risk-return relative ζ. This shows that

$$\begin{cases} \text{HE} > 1 & \text{when } \rho \neq \zeta \\ \text{HE} = 1 & \text{when } \rho = \zeta. \end{cases}$$

[4] It makes the hedging effectiveness positive.

Furthermore, the relationship of ζ to ρ is important in determining the relationship among $\theta_s^*, \theta_{MV}^*$. Hence, we lead to the following proposition:

Proposition 3.1. *Suppose that the agent will short futures. If θ_s^* is exitance*

$$\begin{cases} \theta_s^* > \theta_{MV}^* & \text{when } 0 > \zeta \\ \theta_s^* = \theta_{MV}^* & \text{when } \zeta = 0 \\ \theta_s^* < \theta_{MV}^* & \text{when } \frac{1}{\rho} > \zeta > 0. \end{cases}$$

Proof Consider $\zeta > 0$. Hence, $\zeta(1-\rho^2) > 0$. It is implied that $\rho - \rho + \zeta(1-\rho^2) > 0$. Then $\frac{\rho-\zeta}{1-\zeta\rho} < \rho$. Hence, we observe the equation (23), then $\frac{\theta_s^*}{\theta_{MV}^*} < 1$. On the other hand, $\frac{\theta_s^*}{\theta_{MV}^*} > 1$. In particular, if $\zeta = 0$ then we can get $\theta_s^* = \theta_{MV}^*$. Since θ_s^* is exitance, $1 > \zeta\rho$.

Therefore, if the expected return on the futures contract is zero, then we can know the Sharpe hedge ratio reduces to minimum variance hedge ratio.

3.3 Maximum expected utility hedge ratio

Benninga, Edor and Zilcha [8] have show that minimum variance are consistent with expected-utility maximization regardless of the utility function chosen under unbiasedness and separability. Benninga, Edor and Zilcha [8] consider a decision maker whose utility is a function of terminal wealth $U(V_1)$, such that $U' > 0$ and $U'' < 0$. Thus the decision maker's problem may be written:

$$\max_X \mathbb{E}U(V_1). \tag{24}$$

Benninga, Edor and Zilcha [8] make two assumptions which are essential for our results:
(1) Unbiasedess. That is $F_0 = \mathbb{E}(F_1) = \mathbb{E}(P_2)$, where $\mathbb{E}$ denotes the expectation operator at time 0.
(2) Separability. Let $P_1 = \alpha_1 + \beta_1 F_1 + b_1$, where the joint probability density function of F_1 and b_1 are separable. First, Benninga, Edor ,and Zilcha [8] prove the following Lemma. Let $\tilde{X}(\omega)$ be a random variable, defined on Ω, which satisfies: $\mathbb{E}\tilde{X} < \infty$. Then

Lemma 3.2. *Let A and B be constants and*

$$\mathbb{E}\tilde{X}U'(A\tilde{X}+B) = (\mathbb{E}\tilde{X})\mathbb{E}U'(A\tilde{X}+B). \tag{25}$$

Then we must have $A = 0$.

Proof It is provided in Appendices.

Finally, Benninga, Edor and Zilcha [8] have that proposition :

Proposition 3.3. *Assume*

$$\max_X \mathbb{E}U(V_1) \tag{26}$$

is under assumptions unbiasedness and separability. Then the optimal hedge ratio is

$$X^* = \beta_1 Q.$$

Proof It is provided in Appendices.

Next, we turn to consider that

$$\max_h \mathbb{E}U(\Delta V)$$

in Benninga, Edor and Zilcha's assumptions. Then, the optimal hedge ratio is β_1, which is equal h^*_{MV}.

Hence, we find that Benninga, Edor, and Zilcha [8] have show that minimum variance are consistent with expected-utility maximization regardless of the utility function chosen.

3.4 Minimum generalized semivariance hedge ratio

In this case the optimal hedge ratio is obtained by minimizing the GSV given below:

$$L_n(c, X) = \int_{-\infty}^{c} (c-x)^n dF(x), \quad n > 0, \tag{27}$$

where $F(\cdot)$ denote the probability distribution function of the return on the hedged portfolio X. The GSV is specified by two parameters: the target return (c) and the power of the shortfall (n). The GSV, due to its emphasis on the returns below the target return, is relation with the risk perceived by managers(see [17]). It can be shown that $n < 1$ represents a risk-seeking investor and $n > 1$ represents a risk-averse investor by Fishburn [13]. Some properties of GSV are summarized as following:

1. For a given pair of c and n, GSV is uniquely determined by the probability distribution function. That is, if X_1 and X_2 have the same distribution, then

$$L_n(c, X_1) = L_n(c, X_2).$$

2. As c increases, GSV also increases. That is, when $c_1 > c_2$,

$$L_n(c_1, X) > L_n(c_2, X).$$

3. Suppose that the density function of X is symmetric at c. Let var(X) denote the variance of X, then

$$L_2(c, X) = \frac{\text{var}(X)}{2}.$$

If $c = 0$, GSV is identical to the semivariance which is a half of the variance.

It would be interesting to compare the optimal hedge ratio under the criteria of minimum variance and minimum GSV. To reduce the scope of comparison we consider the semivariance. Thus, we assume $n > 1$, and $f(r_p, r_f)$ is a joint p.d.f. of $\widetilde{r_p}$ and $\widetilde{r_f}$. From equation (27), the first-order condition [5] for θ^* is:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta r_f} r_f(c - r_p + \theta r_f)^{n-1} f(r_p, r_f) dr_p dr_f \bigg|_{\theta^*} = 0,$$

and second derivative condition is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta^* r_f} r_f^2(c - r_p + \theta r_f)^{n-2} f(r_p, r_f) dr_p dr_f > 0.$$

It is always satisfied for any θ^*, then we can find unique and existence optimal solution under $0 \leq \theta^* \leq a$ where $a \in \mathbb{R}$ and finite .

Lien and Tse [15] have shown that if the future and spot returns are jointly normally distributed and if the future price is unbiased (i.e., expected futures price change is zero) the GSV hedge ratio will be equal to the minimum-variance hedge ratio. We adopt the formulated including the symmetric bivariate generalized hyperbolic variables assumption. Suppose that $\widetilde{r_p}$ and $\widetilde{r_f}$ are bivariate normal variables. The joint distribution is characterized by the two means $\mu_i = E(r_i)$, $i = p,\ f$, and by the covariance terms $\sigma_{ij} = cov(r_p, r_f)$. Let $f(\cdot,\cdot)$ denote the joint density function, then equation (27) can be rewritten as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta r_f} (c + \theta r_f - r_p)^n f(r_p, r_f) dr_p dr_f. \tag{28}$$

Upon taking the partial derivative of $L_n(c, X)$ with respect to θ, we conclude that the GSV hedge ratio, θ^*_{nor}, must satisfy the following condition

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta r_f} n(c + \theta r_f - r_p)^{n-1} r_f f(r_p, r_f) dr_p dr_f \bigg|_{\theta^*_{nor}} = 0. \tag{29}$$

[5]The proof is provided in Appendices

We apply the variable transformation method by changing the variables the variables from r_p and r_f to x and y ,where $x = c + \theta r_f - r_p$ and $y = r_f$. As a result, equation (29) can be rewritten as follows:

$$\int_0^\infty \int_{-\infty}^\infty n x^{n-1} y f(c - x + \theta y, y) dy dx \Big|_{\theta^*_{nor}} = 0. \tag{30}$$

The joint density $f(c - x + \theta y, y)$ can be decomposed as $\frac{1}{2\pi\Lambda} e^{\frac{-1}{2}AB}$, where $\Lambda^2 = \sigma^2_{r_p}\sigma^2_{r_f} - (\sigma_{r_p r_f})^2$, and

$$\begin{aligned} A &= \left[y + (c - x - \mu_{r_p})(\theta\sigma^2_{r_f} - \sigma_{r_p r_f})\sigma^{-2}_{r_\theta} - \mu_{r_f}(\sigma^2_{r_p} - \theta\sigma_{r_p r_f})\sigma^{-2}_{r_\theta} \right]^2 \Lambda^{-2}\sigma^2_{r_\theta} \\ B &= (c - x - \mu_{r_p} + \theta\mu_{r_f})^2 \sigma^{-2}_{r_\theta}, \end{aligned}$$

Upon substituting the above decomposition into equation (30) and the integration with respect to y, we obtain

$$\int_0^\infty \frac{1}{\sqrt{2\pi}} n x^{n-1} \sigma^{-3}_{r_\theta} \begin{bmatrix} -(c - x - \mu_{r_p})(\theta\sigma^2_{r_f} - \sigma_{r_p r_f}) \\ +\mu_{r_f}(\sigma^2_{r_p} - \theta\sigma_{r_p r_f}) \end{bmatrix} e^{\frac{-1}{2}\left(\frac{c - x - \mu_{r_p} + \theta\mu_{r_f}}{\sigma_{r_\theta}}\right)^2} dx \Big|_{\theta^*_{nor}} = 0. \tag{31}$$

Next, we define $\omega = \frac{x - c + \mu_{r_p} - \theta\mu_{r_f}}{\sigma_{r_\theta}}$. After a change of variable from x to ω, and $k = \frac{-c + \mu_{r_p} - \theta\mu_{r_f}}{\sigma_{r_\theta}}$.

In the case of an unbiased futures market, $\mu_{r_f} = 0$, then the above equation can be reformulated as following:

$$(\theta\sigma^2_{r_f} - \sigma_{r_p r_f}) \int_k^\infty (n-1)(\omega - k)^{n-2} e^{-\omega/2} d\omega = 0,$$

where evaluated at $\theta = \theta^*_{nor}$.

Hence, the GSV hedge ratio is established at $\theta^*_{nor} = \theta^*_{MV} = \frac{\sigma_{r_p r_f}}{\sigma^2_{r_f}}$, $\forall\ n > 1$.

Next, we turn to consider that $\widetilde{r_f}$ and $\widetilde{r_p}$ are symmetric bivariate generalized hyperbolic variables. Hence, the equation (31) can be calculated as follows:

$$\int_0^\infty \int_0^\infty \frac{1}{\sqrt{2\pi}} n x^{n-1} \sigma^{-3}_\theta \begin{bmatrix} -(c - x - \mu_2) y (\theta\Delta_{11} - \Delta_{12}) \\ +\mu_1 y (\Delta_{22} - \theta\Delta_{12}) \end{bmatrix} e^{\frac{-1}{2}\left(\frac{c - x - \mu_2 + \theta\mu_1}{\sigma_\theta}\right)^2} d_{GIG}(y) dy dx \Big|_{\theta^*_{GH}} = 0$$

where $\sigma^2_\theta = y(\Delta_{22} + \theta^2\Delta_{11} - \theta\Delta_{12})$.

Upon the equation can be reformulated following as

$$\int_0^\infty n x^{n-1} \begin{bmatrix} -(c - x - \mu_2)(\theta\Delta_{11} - \Delta_{12}) \\ +\mu_1(\Delta_{22} - \theta\Delta_{12}) \end{bmatrix} d_{GH(\lambda, 0, \delta, \psi, \bar{\mu}, \bar{\xi})}(x) dx \Big|_{\theta^*_{GH}} = 0$$

where

$$\bar{\mu} = c + \theta\mu_1 - \mu_2$$
$$\bar{\xi} = \Delta_{22} + \theta^2\Delta_{11} - \theta\Delta_{12}.$$

Similarity, in the case of an unbiased futures market, $\mu_1 = 0$, then the above equation can be reformulated as following:

$$(\theta\Delta_{11} - \Delta_{12}) \int_{\mu_2 - c}^{\infty} (n-1)(\omega - \mu_2 + c)^{n-2} d_{GH(\lambda,0,\delta,\psi,0,\bar{\xi})}(\omega) d\omega = 0$$

where evaluated at $\theta = \theta^*_{GH}$.
Hence, in symmetric bivariate generalized hyperbolic variables, the GSV hedge ratio is established at $\theta^*_{GH} = \theta^*_{MV} = \frac{\Delta_{12}}{\Delta_{11}}, \quad \forall \ n > 1$.
Hence, we make a summary of the case follows as:
If returns have a symmetric bivariate generalized hyperbolic density or normal bivariate density in GSV criterion

$$\begin{cases} \theta^*_{GH} > \theta^*_{MV} & \text{as } (\mu_{r_f})\mu_1 < 0 \\ \theta^*_{GH} = \theta^*_{MV} & \text{as } (\mu_{r_f})\mu_1 = 0 \\ \theta^*_{GH} < \theta^*_{MV} & \text{as } (\mu_{r_f})\mu_1 > 0 \end{cases} \forall \, n > 1 \tag{32}$$

3.4.1 Estimation of GSV hedge ratios method

Numerical methods can be used to search for the optimum hedge ratio. In empirical studies, the true distribution is unknown but the GSV can be estimated from the sample by using the following sample counterpart. That is, for a given θ we can construct the data series for N from the data series of r_p, and r_f,

$$L_n(c, r_\theta) = \frac{1}{N} \sum_{i=1}^{N} (c - r_{i,\theta})^n I_{(c - r_{i,\theta})}, \tag{33}$$

where

$$I_{(c - r_{i,\theta})} = \begin{cases} 1, & \text{for } c \geq r_{i,\theta}, \\ 0, & \text{for } c < r_{i,\theta}. \end{cases}$$

So, the GSV can be estimated.

One can instead use the kernel density estimation method suggested by Lien and Tse [16]. The method is abstract described as follows: From the definition of a probability, if the random variable $\tilde{X}$ has density f, then

$$f(x) = \lim_{h \to 0} \frac{1}{2h} \Pr(x - h < \tilde{X} < x + h).$$

For any given h, we can of cause estimate $\Pr(x - h < \tilde{X} < x + h)$ by the proportion of the sample falling in the interval $(x - h < \tilde{X} < x + h)$. Thus a natural estimator $\hat{f}$ of the density is given by choosing a small number h and setting

$$\hat{f}\ (x) = \frac{1}{2hn}[No. \text{ of } X_1, ..., \ X_n \text{ falling in } (x - h, \ x + h)];$$

we shall call this the naive estimator.

To express the estimator more transparently, define the weight function w by

$$w(y) = \begin{cases} \frac{1}{2}, & \text{if } |y| < 1 \\ 0, & \text{otherwise.} \end{cases} \tag{34}$$

Then it is easy to see that the naive estimator can be written

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} w(\frac{x - X_i}{h}).$$

It follows from (34) that the estimate is constructed by placing a box of width $2h$ and height $(2nh)^{-1}$ on each observation and then summing to obtain the estimate.

Replace the weight function w by a kernel function K which satisfies the condition

$$\int_{-\infty}^{\infty} K(x)dx = 1.$$

In empirical studies, the true distribution is unknown and must be estimated. We estimate the probability distribution of the portfolio return for θ. That is, for a given θ we can construct the data series for m from the data series of r_p and r_f and estimate the probability distribution function of m. Let $K(\cdot)$ be a smooth probability density function. Suppose N random samples of m are given, m_1, m_2, ..., m_N, calculated from given θ. The probability density function of m at a given point x, denoted by $\hat{f}(x)$, can be estimated by:

$$\hat{f}_h\ (x) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{x - m_i}{h}), \tag{35}$$

where h is the window width [6] and $K(\cdot)$ is called the kernel function. Consequently, upon substituting the estimated density of m for the true unknown density from equation (27) the semivariance, that is $n = 2$, of m is estimated by

$$l_2(c, X) = \int_{-\infty}^{c} (c-x)^2 \hat{f}_h(x)dx = \frac{1}{Nh}\sum_{i=1}^{N}\int_{-\infty}^{c}(c-x)^2 K(\frac{x-m_i}{h})dx$$

After a change-in-variable from x to $y = \frac{x-m_i}{h}$ can be written as:

$$l_2(c, X) = \frac{1}{N}\sum_{i=1}^{N} q(c, m_i),$$

where

$$q(c, m_i) = \int_{-\infty}^{\frac{c-m_i}{h}} (c - yh - m_i)^2 K(y)dy.$$

4 Stochastic Dominance

Stochastic Dominance is based on the Von Neumann and Morgenstern paradigm. Von Neumann and Morgenstern, in their study, *introduced the expected utility hypothesis*: for every reasonable decision maker there exists a utility function $u(\cdot)$ such that individual prefers return X over return Y if and only if $\mathbb{E}[u(X)] > \mathbb{E}[u(Y)]$. However, in most situations, such a selection is not possible since complete information about an individual‘s preference set, and hence his utility function, is not available. So, we are interested in ordering rules for certain restricted classes of utility functions which defined as follows:

Definition 4.1. $U_1 = \left\{u(x)|\ u(x) \text{ is finite for every finite } x, u^{'}(x) > 0\ \forall x \in \mathbb{R}\right\};$

Definition 4.2. $U_2 = \left\{u(x)|\ u(x) \in U_1, -\infty < u^{''}(x) < 0\ \forall x \in \mathbb{R}\right\};$

Definition 4.3. $U_3 = \left\{u(x)|\ u(x) \in U_2, u^{'''}(x) > 0\ \forall x \in \mathbb{R}\right\};$

Definition 4.4. $U_4 = \left\{u(x)|\ u(x) \in U_2, r^{'}(x) \equiv (-u^{''}(x)/u^{'}(x))^{'} < 0\ \forall x \in \mathbb{R}\right\}.$

U_1 is the set of all increasing and continuously differentiable utility functions $u(x)$ assumed to have finite values for finite values of x. U_2 is the subset of U_1 with risk averse utility

[6] It is also called the smoothing parameter or bandwidth by some authors.

functions; U_3 is the subset of U_2 with utility functions having positive third derivative, while U_4 is the subset of U_2 with decreasing absolute risk averse utility functions.
Bawa ([5][6]) proposed the mean-lower partial moment criterion, which contains the mean-semivariance as a special case and is consistent with first, second, and third degree stochastic dominance relations. We apply his results to ours setting furthermore.

Theorem 4.5. *Assume $F(x)$ and $G(x)$ denote the probability distribution function Then*

(1) $\mathbb{E}[U(r_{\theta_F})] > \mathbb{E}[U(r_{\theta_G})]$ for any strictly increasing utility function U_1 iff $L_0(c, r_{\theta_F}) \leq L_0(c, r_{\theta_G})$ for every c with strict inequality for some c.

(2) $\mathbb{E}[U(r_{\theta_F})] > \mathbb{E}[U(r_{\theta_G})]$ for any strictly increasing, concave utility function U_2 iff $L_1(c, r_{\theta_F}) \leq L_1(c, r_{\theta_G})$ for every c with strict inequality for some c.

(3) $\mathbb{E}[U(r_{\theta_F})] > \mathbb{E}[U(r_{\theta_G})]$ for any strictly increasing, concave utility function U_3 with positive third derivatives iff $L_2(c, r_{\theta_F}) \leq L_2(c, r_{\theta_G})$ for every c with strict inequality for some c, and $\mathbb{E}(r_{\theta_F}) > \mathbb{E}(r_{\theta_G})$.

(4) Assume that $\mathbb{E}(r_{\theta_F}) = \mathbb{E}(r_{\theta_G})$. $\mathbb{E}[U(r_{\theta_F})] > \mathbb{E}[U(r_{\theta_G})]$ for any strictly increasing, concave utility function U_4 with decreasing risk aversion coefficients iff $L_2(c, r_{\theta_F}) \leq L_2(c, r_{\theta_G})$ for every c with strict inequality for some c. Where $\mathbb{E}$ denotes the expectation operator.

Proof The proof is shown in Appendices.

4.1 Stochastic Dominance under the exponential function assumption

We choose the $U(x) = 1 - e^{-x\eta}, \eta > 0$. Let $x = \widetilde{r_p} - \theta\widetilde{r_f}$. Then suppose $\widetilde{r_p}$ and $\widetilde{r_f}$, which probability density function are $f(r_f, r_p)$, are bivariate variables.
The decision maker is wished to choose θ to maximize the expected utility of wealth.

$$\max_{\theta} \mathbb{E}[1 - e^{-(\widetilde{r_p} - \theta\widetilde{r_f})\eta}] \tag{36}$$

Hence, consider

$$\min_{\theta} \int_{-\infty}^{\infty} e^{(\theta r_f - r_p)\eta} f(r_f, r_p) dr_f dr_p = \min_{\theta} e^{(\theta\mu_{r_f} - \mu_{r_p} + \frac{1}{2}(\theta^2\sigma_{r_f}^2 - 2\theta\rho\sigma_{r_f}\sigma_{r_p} + \sigma_{r_p}^2))\eta}.$$

Then we have

$$\theta_{nor}^* = \frac{\rho\sigma_{r_f}\sigma_{r_p} - \mu_{r_f}}{\sigma_{r_f}^2}. \tag{37}$$

In the same condition, suppose that a pair $\widetilde{r_f}$ and $\widetilde{r_p}$ are symmetric joint hyperbolic distributions $\mathbf{H}_2$.

Hence, we can obtain that when $\mu_f = 0$, then

$$\theta^*_{H_2} = \frac{\Delta_{12}}{\Delta_{11}},$$

which is equal the equation (15) as $\beta_1 = \beta_2 = 0$.

Further, form Theorem 4.5 we consider that $\mathbb{E}[U(r_p - \theta r_f)|r_f]$ with $U(x) = 1 - e^{-x\eta}, \eta > 0$. Hence, we can reformulate the formula (36) and consider

$$e^{\eta} \int_{-\infty}^{\infty} e^{\theta r_f} \left\{ \mathbb{E}[e^{-r_p}|r_f] \right\} f(r_f) dr_f,$$

where $f(r_f)$[7] is the probability density function of the **GH**.

Suppose that $r_p = \mu_{r_p} + \rho \frac{\sigma_{r_p}}{\sigma_{r_f}}(r_f - \mu_{r_f})$, [8] and The formula will be reformulated as follows:

$$\xi e^{\mu_f t} \left(\frac{\alpha_f^2 - \bar{\beta}_f^2}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right)^{\lambda/2} \frac{K_\lambda(\gamma_f \sqrt{\alpha_f^2 - (\bar{\beta}_f + t)^2})}{K_\lambda(\gamma_f \sqrt{\alpha_f^2 - \bar{\beta}_f^2})},$$

where $t = (\theta - \rho \frac{\sigma_{r_p}}{\sigma_{r_f}})\eta$ and, $\xi = e^{(-\mu_{r_p} + \mu_{r_f} \rho \frac{\sigma_{r_p}}{\sigma_{r_f}})\eta} > 0$.

In the case $\mu_f = 0$, we consider minimum above formulas when $\gamma_f \to 0$, [9] there is extreme value given by

$$\theta^*_{GH} = \rho \frac{\sigma_{r_p}}{\sigma_{r_f}} - \bar{\beta}_f \frac{1}{\eta}. \tag{38}$$

We obtain that if the density function of future log-return is symmetric at 0, then its hedge ratios becomes minimum variance ratio.

If we do not have any assumption under $U(x)$ then we can using Taylor series in place of $U(x)$. Hence,

$$\begin{aligned} \mathbb{E}[U(r_p - \theta r_f)] &\approx U(\mathbb{E}[r_p - \theta r_f]) + \frac{1}{2} var(r_p - \theta r_f) U''(\mathbb{E}[r_p - \theta r_f]) \\ &\approx U(0) + \mathbb{E}[r_p - \theta r_f] U'(0) + \frac{1}{2} \mathbb{E}[r_p - \theta r_f]^2 U''(0) + \frac{1}{2} var(r_p - \theta r_f) U''(0) \end{aligned}$$

We can get $\theta = \frac{\rho \sigma_{r_f} \sigma_{r_p} + \mu_{r_f} \mu_{r_p}}{\mu_{r_f}^2 + \sigma_{r_f}^2} + \mu_{r_f} \frac{U'(0)}{U''(0)}$, moreover, when $\mu_{r_f} = 0$, it reduces to minimum variance hedge ratio .

[7] see Appendices.
[8] It is called the best linear predictor. Since it has least mean square error.
[9] It is also called variance Gamma class, which is special class of **GH** distributions .

5 Numerical Method

5.1 Kernel estimate distribution

In empirical studies, the true distribution is unknown and must be estimated. It will be assumed that we have a sample $M_1, ..., M_N$ of independent, identically distributed observations from a continuous univariate distribution with probability density function f, which we are trying to estimate. For simplicity, we shall assume throughout this discussion that:
$K(\cdot)$ is a symmetric probability density function satisfying

$$\int K(t)dt = 1, \quad \int tK(t)dt = 0, \quad \text{and} \quad \int t^2K(t)dt = k_2 \neq 0$$

and the unknown density f has continuous derivatives of all orders required.

5.1.1 Mean integrated square error

To determine the optimal size of h, the minimizing the approximate mean integrated square error method is adopted. The first (Rosenblatt, 1956) and the most widely used way of placing a measure on the global accuracy of $\hat{f}_h$ as an estimator of f is the mean integrated square error (abbreviated MISE) defined by

$$MISE(\hat{f}_h) = \mathbb{E}\int \{\hat{f}_h(x) - f(x)\}^2 dx. \tag{39}$$

It is useful to note that, since the integrand is non-negative, the order of integration and expectation in (39) can be reversed to give the alternative forms

$$MISE(\hat{f}_h) = \int \{\mathbb{E}\hat{f}_h(x) - f(x)\}^2 dx + \int var\hat{f}_h(x)dx, \tag{40}$$

which gives the MISE as the sum of the integrated square bias and the integrated variance.

5.1.2 The bias and variance

Suppose $\hat{f}_h$ is defined in (35). We shall write

$$\begin{aligned} bias_h &= \mathbb{E}\hat{f}_h(x) - f(x) \\ &= \int h^{-1}K[(x-m)/h]f(m)dm - f(x). \end{aligned} \tag{41}$$

We shall now use (41) to obtain an approximate expression for the bias.
Make the change of variable $m = x - ht$ and use the assumption that K integrates to unity, to write

$$\begin{aligned} bias_h(x) &= \int K(t)f(x-ht)dt - f(x) \\ &= \int K(t)[f(x-ht) - f(x)]dt. \end{aligned}$$

A Taylor series expansion gives

$$f(x-ht) = f(x) - htf'(x) + \frac{1}{2}h^2t^2f''(x) + \cdots$$

so that, by the assumptions made about K,

$$\begin{aligned} bias_h(x) &= -hf'(x)\int tK(t)dt + \frac{1}{2}h^2f''(x)\int t^2K(t)dt + \cdots \\ &= \frac{1}{2}h^2f''(x)k_2 + R(h), \end{aligned} \tag{42}$$

where $R(h)$ is higher order terms in h.
The integrated square bias, required in formula (42) for the mean integrated square error, is then given by

$$\int bias_h(x)^2dx \approx \frac{1}{4}h^4k_2^2\int f''(x)^2dx. \tag{43}$$

We now turn to the variance, since the M_i are independent.

$$\begin{aligned} var\hat{f}_h(x) &= n^{-1}\int h^{-2}K[(x-m)h^{-1}]^2f(m)dm - n^{-1}\left\{\int h^{-1}K[(x-m)h^{-1}]f(m)dm\right\}^2 \\ &= n^{-1}\int h^{-2}K[(x-m)h^{-1}]^2f(m)dm - n^{-1}[f(x) + bias_h(x)]^2 \\ &\approx n^{-1}h^{-1}\int f(x-ht)K(t)^2dt - n^{-1}[f(x) + O(h^2)]^2, \end{aligned}$$

using the substitution $m = x - ht$ in the integral, and the approximation (42) for the bias. Assume that h is small and n is large, and expand $f(x-ht)$ as a Taylor series to obtain

$$\begin{aligned} var\hat{f}_h(x) &\approx n^{-1}h^{-1}\int[f(x) - htf'(x) + ...]K(t)^2dt + O(n^{-1}) \\ &= n^{-1}h^{-1}f(x)\int K(t)^2dt + O(n^{-1}) \\ &\approx n^{-1}h^{-1}f(x)\int K(t)^2dt. \end{aligned} \tag{44}$$

Since f is a probability density function, integrating (44) over x gives the simple approximation

$$\int var\hat{f}_h(x)dx \approx n^{-1}h^{-1}\int K(t)^2dt. \tag{45}$$

Using the approximations (43) and (45) in the (40), and we get the approximate mean integrated square error following as :

$$MISE(\hat{f}_h) \approx \frac{1}{4}h^4k_2^2\int f''(x)dx + n^{-1}h^{-1}\int K(t)^2dt. \tag{46}$$

The ideal value of h which is from the point of view of minimizing the approximate mean integrated square error (46). Hence, it can be show by simple calculus to be equal to h_{opt}, where

$$h_{opt} = k_2^{-2/5}\left\{\int K(t)^2dt\right\}^{1/5}\left\{\int f''(t)^2dt\right\}^{-1/5}n^{-1/5}. \tag{47}$$

The formula (47) for the optimal window width is somewhat disappointing since it shows that h_{opt} itself depends on the unknown density being estimated.

5.1.3 Reference to a standard distribution

A very easy and natural approach is to use a standard family of distributions to assign a value to the term $\int f''(t)^2dt$ in the expression (47) for the ideal window width. For example, the normal distribution with variance σ^2. We substituting the $\int f''(t)^2dt$ with the normal distribution with variance σ^2 has, setting ϕ to be the standard normal density i.e.,

$$\int f''(x)^2dx = \sigma^{-5}\int \phi''(x)^2dx = \frac{3}{8}\pi^{-1/2}\sigma^{-5}. \tag{48}$$

If a Gaussian kernel is being used, then the window width obtained from (47) would be, substituting the vale (48), we get

$$h_{opt} = (\frac{4}{3})^{1/5}\sigma n^{\frac{-1}{5}}. \tag{49}$$

It has advanced situation, we can adjust it to became $h = (\frac{4}{3})^{1/5}An^{\frac{-1}{5}}$, which $A =$ min(standard deviation, interquartile/1.34) is instead of σ in the formula (49). The method was introduced and investigated further in [20].

The Gaussian kernel density estimates based on the daily log returns of the Taiwan Stock Exchange Capitalization Weighted Stock Index (abbreviated TAIEX), S&P500 Index, and Nasaq100 Index over the period from 2000 January until the end of 2004 December are shown in Fgure1, and compared with normal distributions. And, all data sets of statistics information are listed in Table7. In numerical calculation, the density functions are estimated by the Matlab 7.0 of Toolbox.

In the Figure2, we can obtain that the empirical p.d.f have higher right and left tails in TAIEX Futures and TAIEX then theoretical distributions.

5.2 Simulation of random variables

5.2.1 Sampling from GIG distribution

Atkinson [2] proposed a method for **GIG**, which can be simulated via the rejection algorithm. We make summary following text. The generation utilizes a rejection method with a two part rejection-envelop. We define the $g\ : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ by the envelop

$$d_i(x) = \begin{cases} e^{sx} & (0 \leq x \leq t) \\ e^{-px} & (x > t) \end{cases}$$

, where $t = m(\lambda, \delta, \psi) = \frac{\lambda - 1 + \sqrt{(1-\lambda)^2 + \delta\psi}}{\psi}\ \forall\ \psi > 0.$
We find values of s and p to minimize

$$S_1 \int_0^t d_1(z)dz + S_2 \int_t^\infty d_2(z)dz = S_1 \frac{(e^{st} - 1)}{s} + S_2 \frac{e^{-pt}}{p} \tag{50}$$

,where

$$\begin{aligned} &S_1 = e_{(\lambda,\delta,\psi+2s)}(x_L) \quad \text{where} \quad x_L = m(\lambda, \delta, \psi + 2s) \\ &S_2 = e_{(\lambda,\delta,\psi-2p)}(x_R) \quad \text{where} \quad x_R = m(\lambda, \delta, \psi - 2p) \\ &\text{and} \quad e_{(\lambda,\delta,\psi)}(x) = x^{\lambda-1} e^{-0.5(\delta x^{-1} + \psi x)}, \end{aligned}$$

Hence, the algorithm GIG.
Set

$$r = k_1 \frac{(e^{st} - 1)}{s}$$

1. Generate U and U^*. If $U > r$, go to 2.
   $x = \frac{1}{s}\log\left(1 + \frac{sU}{k_1}\right)$.
   If $\log U^* > \log\{e_{(\lambda,\gamma,\psi+2s)}(x)/S_1\}$ go to 1. Otherwise return x.

2. $x = \frac{-1}{p}\log\left\{\frac{p}{k_2}(1-U)\right\}$.
   If $\log U^* > \log\{e_{(\lambda,\gamma,\psi-2p)}(x)/S_2\}$ go to 1. Otherwise return x.

where U^* is uniformly distributed on $(0,1)$ independently of U,

$$k_1 = \frac{(e^{st}-1)/s}{S_2(\frac{(e^{st}-1)/s}{S_2} + \frac{e^{-pt}/p}{S_1})} \qquad \text{and} \qquad k_2 = \frac{(e^{st}-1)/s}{S_1(\frac{(e^{st}-1)/s}{S_2} + \frac{e^{-pt}/p}{S_1})}.$$

The algorithm with highest efficiency is found by choosing the s, p to minimize the equation (50).

5.2.2 Sampling from GH distribution

Hence, the most natural way of simulating generalized hyperbolic variables stems from the fact that they can be represented as normal variance-mean mixtures. Since the mixing distribution is the generalized inverse Gaussian law, the resulting algorithm read as follows:

1. Simulate a random variable $Y \sim \mathbf{GIG}(\lambda, \delta, \psi)$;

2. Simulate a standard normal random variable ε, e.g. using the Box-Muller algorithm;

3. Return $X = \mu + \beta Y + \sqrt{Y}\varepsilon$.

An d-dimensional random vector X being conditionally normal distributed with mean vector $Y\Delta\beta$ and covariance matrix $Y\Delta$ is generated. More precisely,

1. Set $\Delta = L^T L$ via Cholesky decomposition;

2. Simulate a random variable $Y \sim \mathbf{GIG}(\lambda, \delta, \psi)$;

3. Simulate a standard normal random variable $\varepsilon \sim N(0, I_d)$, e.g. using the Box-Muller algorithm;

4. Return $X = \mu + Y\Delta\beta + \sqrt{Y}L'\varepsilon$.

We turn now to use the algorithm to simulation dependent of multivariate $\mathbf{H}$ as follows:

1. Simulate Y_f and Y_p of $\mathbf{GIG}(-1, \psi, \gamma)$ random variables, respectively,
2. Sample U form uniform distribution to generate random variables of exponential distribution;
3. Simulate Y_f and Y_p via lemma 2.2.

The algorithm is fast and efficient if we have a handy way of simulating generalized inverse Gaussian variables.

5.3 Estimation of parameters

5.3.1 Maximum-Likelihood Estimation

Blæsid and Sørensen [10] use maximum log-likelihood estimation in the situation where independent and identically (possibly multi-dimensional) hyperbolic distribution observations are considered. They adopted multinomial likelihood function obtained by only observing the number of observations in given intervals. More precisely, if $I_1, ..., I_k$ are disjoint intervals with union the entire real line and y_j denotes the number of observations in I_j, j=1,...,k, then the multinomial log-likelihood function is given by

$$\ell(\Theta) = \sum_{j}^{k} y_j \log p_j,$$

where p_j is the probability that joint hyperbolic distributed random variable takes a value in I_j, and $\Theta = (\theta_1, ..., \theta_m)$ is parameter vetoer.
In this paper, we also estimate the parameters of hyperbolic distributions via maximum log-likelihood.
Since there exist closed-form formulas for the densities of these laws. Given observed i.i.d data $x_1, x_2, ..., x_n$. We define the likelihood function as

$$like(\Theta) = f(x_1, x_2, ..., x_n | \Theta),$$

where f is the frequency function. Note that if the distribution is discrete the likelihood function gives the probability of observing the given data as a function of the parameter vetoer $\Theta = (\theta_1, ..., \theta_m)$. With maximum likelihood estimator we maximize the probability. Since $x_1, x_2, ..., x_n$ are assumed i.i.d and the natural logarithm is a monotonic function we may instead maximize the log-likelihood function

$$\begin{aligned}
\ell_{H(\alpha,\beta,\delta,\mu)}(\Theta) &= \sum_{1}^{n} \log\left[f(x_i|\Theta)\right] \\
&= n\left(\log\sqrt{\alpha^2-\beta^2} - \log 2 - \log\alpha - \log\delta - \log K_1(\delta\sqrt{\alpha^2-\beta^2})\right) \\
&\quad + \sum_{i=1}^{n}\left[-\alpha\sqrt{\delta^2+(x_i-\mu)^2} + \beta(x_i-\mu)\right].
\end{aligned}$$

Then we get the first derivatives of hyperbolic log-likelihood function with respect to the parameters:

$$\begin{aligned}
\frac{\partial \ell_H}{\partial\alpha} &= n\left[\frac{2\alpha}{\alpha^2-\beta^2} - \frac{1}{\alpha} + \frac{\alpha\delta}{\sqrt{\alpha^2-\beta^2}}\frac{K_0(\delta\sqrt{\alpha^2-\beta^2})}{K_1(\delta\sqrt{\alpha^2-\beta^2})}\right] - \sum_{i=1}^{n}\sqrt{\delta^2+(x_i-\mu)^2} \\
\frac{\partial \ell_H}{\partial\beta} &= n\left[\frac{-2\beta}{\alpha^2-\beta^2} - \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}\frac{K_0(\delta\sqrt{\alpha^2-\beta^2})}{K_1(\delta\sqrt{\alpha^2-\beta^2})} - \mu\right] + \sum_{i=1}^{n} x_i \\
\frac{\partial \ell_H}{\partial\delta} &= n\left[\sqrt{\alpha^2-\beta^2}\frac{K_0(\delta\sqrt{\alpha^2-\beta^2})}{K_1(\delta\sqrt{\alpha^2-\beta^2})}\right] - \alpha\delta\sum_{i=1}^{n}\frac{1}{\sqrt{\delta^2+(x_i-\mu)^2}} \\
\frac{\partial \ell_H}{\partial\mu} &= \alpha\sum_{i=1}^{n}\frac{(x_i-\mu)}{\sqrt{\delta^2+(x_i-\mu)^2}} - n\beta.
\end{aligned}$$

In principle, it is no problem to estimate high-dimensional skewed **GH** distributions efficiently. Moreover, numerical problems with the Bessel functions do occur frequently, thus Atkinson [2] had discussed with λ and $\hat{\Delta}$ must be positive definite. Consequently, Prause [18] propose to restrict the estimation to symmetric (i.e., $\beta = 0$) **GH** distributions. We decided to focus our attention to symmetric case **MGH** distributions have the following density

$$\frac{(\alpha/\delta)^{\lambda}}{2\pi^{d/2}K_{\lambda}(\alpha\delta)}\frac{K_{\lambda-\frac{d}{2}}(\alpha\sqrt{\delta^2+(x-\mu)'\Delta^{-1}(x-\mu)})}{(\alpha^{-1}\sqrt{\delta^2+(x-\mu)'\Delta^{-1}(x-\mu)})^{\frac{d}{2}-\lambda}}.$$

In the symmetric case $\mathbf{H}_2$ hyperbolic distributions, that is

$$\mathbf{H}_2 = \frac{(\alpha/\delta)^{3/2}}{2^{3/2}\sqrt{\pi}\alpha K_{\frac{3}{2}}(\delta\alpha)}e^{-\alpha\sqrt{\delta^2+(x-\mu')\Delta^{-1}(x-\mu)}}.$$

Thus, we obtain the log-likelihood function for $\mathbf{H}_2$ distributions is

$$\begin{aligned}\ell_{H_2} &= n\left[\frac{3}{2}\log\frac{\alpha}{\delta}-\frac{3}{2}\log 2-\frac{1}{2}\log\pi-\log\alpha-\log K_{\frac{3}{2}}(\alpha\delta)\right] \\ &\quad -\alpha\sum_{i=1}^{n}\sqrt{\delta^2+(x_i-\mu)'\Delta^{-1}(x_i-\mu)},\end{aligned}$$

where μ and Δ are estimated rapidly via equations (9) and (10) yield that $\hat{\mu}$ and $\hat{\Delta}$ in symmetric case.

We fit hyperbolic distributions, bivariate hyperbolic distribution, and joint normal distributions for the log returns of the TAIEX Futures, TAIEX, S&P500 Index, Nasaq100 Index data. Hence, a maximum likelihood estimation is performed by the program of Matlab 7.0 and the results for our data set are given by

Estimated parameters for the hyperbolic distribution

Parameters	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$	$^*P$-value
TAIEX Futures	69.995	1.5937	0.0027	-9.5635×10^{-4}	0.9277
TAIEX	89.647	3.3658	0.0095	-0.0014	0.9334
S&P500 Future	122.474	-5.8050	0.0059	8.0431×10^{-4}	0.9450
S&P500 Spot	125.312	-3.4285	0.0069	4.1696×10^{-4}	0.9628
Nasdaq100 Future	53.510	-3.8168	0.0028	0.0021	0.5959
Nasdaq100 Spot	54.506	-2.2591	0.0071	9.9398×10^{-4}	0.4994

P-value is used a way of testing the goodness of fit with the Kolmogorov-Smirnov test. The parameter β is called the risk premium parameter. We observe future and spot in Taiwan differ form others.

The generalized hyperbolic distributions have semi-heavy tails, that is

$$d_{GH(\lambda,\alpha,\beta,\delta,\mu)(x)}\sim|x|^{\lambda-1}e^{(\mp\alpha+\beta)x}\quad as\quad x\to\pm\infty\quad\forall\,\lambda>0.$$

(*Barndorff-Nielsen* and *Blæsild* 1981) [10]

Then, in the Nasdaq100 market, it is decreasing slower then others.

[10] It be obtained using Lemma 2.1 through **GH** density.

Estimated parameters for the symmetric bivariate hyperbolic distributions

Parameters	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\mu}\times 10^{-4}$	$\hat{\Delta}_{11}$	$\hat{\Delta}_{12}$	$\hat{\Delta}_{21}$	$^*P$-value
TAIEX			-2.8313				
Future/Spot	161.1448	0.0027	-2.8154	3.3632	2.7491	2.5445	0.1024
S&P500			-1.5092				
Future/Spot	298.1549	0.0017	-1.4577	4.3914	4.2464	4.3339	0.0935
Nasdaq100			-6.8444				
Future/Spot	157.1192	0.0000	-6.7680	4.5243	4.4855	4.6681	3.06×10^{-7}

P-value depends on n, that is number of simple, is used a way of testing the goodness of fit with the χ^2-test on (-0.1, 0.1). We plot those densities in Figure3,also. It is reject $\mathbf{H}_2$ and **GH** in the Nasdaq100, we can see following Table.

Estimated parameters for the symmetric **GH** distribution in Nasdaq100 market.

Parameters	λ	$\hat{\alpha}$	$\hat{\delta}$	P-value
Nasdaq100 Future/Spot				
	1	157.1192	0.0000	3.06×10^{-7}
	-0.7336	61.4088	0.017	2.1284×10^{-6}

Estimated parameters for the normal distribution

Parameters	$\hat{\mu}\times 10^{-4}$	$\hat{\sigma}$	P-value	H_0
TAIEX Future	-2.8313	0.0204	3.7457×10^{-5}	reject
TAIEX	-2.8154	0.0177	0.0168	reject
S&P500 Future	-1.5092	0.0128	0.0049	reject
S&P500 spot	-1.4577	0.0127	0.0133	reject
Nasdaq100 Future	-6.8444	0.0265	1.2275×10^{-4}	reject
Nasdaq100 spot	-6.7680	0.0269	9.3381×10^{-5}	reject

We observe that the several markets are rejected normal distributions, when we take significance of 0.005.

5.3.2 Minimum-Distance Estimation

A frequent problem in statistic and finance is to measure the goodness-of-fit of a theoretical distribution to real world data. To measure how close or how far, is an empirical distribution from a theatrical distribution, several distances have been proposed.

Among then, we introduction three: the Kolmogorov distance; Kuiper distance and the Anderson-Darling distance.

The Kolmogorov distance, which is called the Kolmogorov-Smirnov test statistic, defined as the greatest distance between empirical distribution and theoretical distribution, for all possible values:

$$D_{Kol} = \max_{x \in \mathbb{R}} |F_e(x) - F_{th}(x)|$$

where F_e is the empirical cumulative density function and F_{th} is the continuous theoretical cumulative density function.

F_e can be defined by:

$$F_e(x) = \frac{number\, of\, Y_i \leq x}{n},$$

where Y_i are the sample's elements and n is the sample number of elements.

The Kuiper distance is similar to the Kolmogorov distance, it is defined

$$D_K = \max_{x \in \mathbb{R}} \{F_e(x) - F_{th}(x)\} + \max_{x \in \mathbb{R}} \{F_{th}(x) - F_e(x)\}$$

The Anderson and Darling (1952) paper proposes a distance that would be viewed as Kolmogorov distance with weight. The formula of this distance with tail emphasis is:

$$D_{AD} = \max_{x \in \mathbb{R}} \frac{|F_e(x) - F_{th}(x)|}{\sqrt{F_{th}(x)(1 - F_{th}(x))}}.$$

The **AD** distance is especially relevant to VaR calculations, since it is more sensitive in tails than in distribution's middle range.

In this paper, we will choose ξ such that **AD** is minimal. The Kolmogorov and **AD** distance and the results for our data set are given by Table 6.

5.4 Statistical Testing

5.4.1 Kolmogorov-Smirnov test

Theorem 5.1. *Let F be a continuous c.d.f.,and let $X_1, ..., X_n$ be a sequence of i.i.d. r.v. with the c.d.f. F. Then*

If $n \to \infty$ the distribution of $\sqrt{n}D_n$ is asymptotically Kolmogorov's distribution with c.d.f.

$$Q(t) = 1 - 2\sum_{j}^{\infty}(-1)^{j-1}e^{-2k^2t^2},$$

that is

$$\lim_{n\to\infty} P\left(\sqrt{n}D_n \le t\right) = Q(t),$$

where D is called the Kolmogorov-Smirnov test statistic.

Since a large value of D_n would appear to be inconsistent with the null hypothesis, it follows that the p-value for this data set is given by

$$P - value = P\left\{D_n \ge d\right\}, \tag{51}$$

where d is a constant.
It follows from the preceding proposition that after the value of D_n is determined from the data, say, $D_n = d$, the p-value can be obtained by (51). We use the one-sample test to help us check empiric data and how good-fit in hyperbolic distributions.

5.4.2 χ^2 test

A way of testing the goodness of fit is with the χ^2-test. The χ^2-test counts the number of sample points falling into certain intervals and compares them with the expected number under the null hypothesis.
More precisely, suppose we have n independent observation $M_1, ..., M_N$ from the random variable $\mathbf{Z}$ and we want to test whether theses observations follow a law with distribution $\mathbf{Z}$, depending on h parameters, which we estimate by MLE method. First, make a partition $\mathcal{P} = \{A_1, ..., A_m\}$ of the support of $\mathbf{Z}$. The classes A_k can be chosen arbitrarily. Let N_k, k=1,...,m, be the number of observations M_i falling into the set A_k. We will compare these numbers with the theoretical frequency distribution ϖ_k, defined by

$$\varpi_k = P(\mathbf{Z} \in A_k), \quad k = 1, ..., m,$$

through the Pearson statistic

$$\hat{\chi}^2 = \sum_{k=1}^{m} \frac{(N_k - n\varpi_k)^2}{n\varpi_k}.$$

If necessary, we collapse outer cells, so that expected value ϖ_k of the observations becomes always greater than 5. General theory says that the Pearson statistic $\hat{\chi}^2$ follows a $\hat{\chi^2}$ distribution with $m - 1 - h$ degrees of freedom.

The P-value of the χ_2 statistic is defined as

$$P = P(\chi^2_{m-1-h} > \hat{\chi}^2).$$

It is clear that very mall P-values lead to a rejection of the null hypothesis. To be precise, we reject the hypothesis if the P-value is less than our level of significance, which we take to be to 0.05.

6 Data description and model specification

We propose an alternative class of multivariate distributions belonging to affine-linear transformed in order to improve **MAGH**.
Suppose that ε_1, and ε_2 are mutually independent of standard normal variables

$$\begin{aligned} r_f &\stackrel{\mathcal{D}}{=} \mu_f + \beta_f Y_f + \sqrt{\xi_f Y_f}\varepsilon_1 \\ r_p &\stackrel{\mathcal{D}}{=} \mu_p + \beta_p Y_p + \sqrt{\xi_p Y_p}(\rho\,\varepsilon_1 + \sqrt{1-\rho^2}\varepsilon_2) \end{aligned}$$

where μ_f, β_f, μ_p, and β_p are constants. If we let $\xi_f, \xi_p = 1$, and ignore β_f, β_p the model will reduce multivariate normal distribution. Further, we can obtain that

$$Cov[r_f, r_p] = \beta_f \beta_p Cov[Y_f, Y_p] + \sqrt{\xi_f \xi_p}\,\rho\,\mathbb{E}[\sqrt{Y_f Y_p}].$$

A disadvantage of the model is that the margins r_i $\forall\ i = f, p$ are mutually independent for any choice of the parameters of Y. In the other hands , a disadvantage of **MGH** distributions is that the margins r_i of **MGH** are not mutually independent for any choice of the parameters of $Y, \forall\ i = f, p$.

For the futures markets, the settlement prices are used. For the spot markets, the closing index values are used. We can get theoretical minimum variance hedge ratios and Sharpe hedge ratios in the symmetric MGH for our chosen markets. Unfortunately, we can not get theoretical hedge ratios in the semivariance for our chosen markets. Nevertheless, they can be got via numerical calculation. Therefor, in Table 1 we state theoretical hedge ratios of minimum variance and Sharpe. Then, theoretical hedge ratios are following as:

For risk-free is given 0.1%, we find that Sharpe hedge ratios are not exitance in S&P500 and Nasaq100 markets, but to in Taiwan markets. Sine it is satisfied the equation (19).

Table 1: Theoretical Hedge Ratios under $\mathbf{H}_2$ distribution

	θ^*_{MV}	θ^*_s
Tw	0.8175	1.2
S&P500	0.9670	1.2
Nasaq100	0.9914	1.2

All data sets are over the period 2000-2004. And basic statistic are given by Table 7

In the symmetrical $\mathbf{H}_2$, we obtain optimal θ^* form the equation (15) and (20).

For the Table 3, in the second column, thy are different target returns (TR). The number in parenthesis are the error of variance. Sine θ^*_{nor}, $\theta^*_{H_2}$ are almost the same in the Table 3 and we observe the fact from the equations (37) and (38). We are consequent that correlations are not constant between future and spot markets, and they made a great impact on semivariance hedge ratios. In particular, the target returns are below zeros in the semivariance.

Hence, we want to discussed the state and to improve semivariance hedge ratios by conditional correlations.

6.1 Conditional correlations

Recent research on domestic and international stock markets suggests the notion of correlation asymmetry. Correlation asymmetry has implications for several applications including portfolio diversification, risk management ([1][7]). Hence, we will relax correlation be constant by conditional correlation. Consider a pair of bivariate **MGH** random variables $\bar{r}_f$ and $\bar{r}_p$, respectively. Hence, there is a Lemma following as:

Lemma 6.1. *Let the log-returns on the spot $\bar{r}_p$ and futures $\bar{r}_f$ be symmetric* $\boldsymbol{MGH}_2$ *and unconditional correlation coefficient ρ, which is the correlation between spot and futures log-return. Consider Y is a* ***GIG*** *distribution. The conditional correlation is given*

$$\rho_E = \frac{\rho}{\sqrt{\rho^2 + (1-\rho^2)\frac{\mathbb{E}[Y]}{\mathbb{E}[Y|E]}}}.$$

where $\tau_1\, \tau_2 \in \mathbb{R}$, and E is even on interval $[\tau_1\, \tau_2]$ of $Pr\{Y\}$.

Hence, we can improve the semivariance hedge ratios in the structure of $\mathbf{H}_2$. The results are given by Table 4. There are one implications of the this result. When target return will be low, optimal hedge ratios chosen by the risk manager will be large lower then minimum variance hedge ratios. In generally speaking, when the target return is special low,like -0.005, optimal hedge ratios can be dominance in after the second or third standard deviation of distributions of **GIG**. We also see hedge ratios of GSV as $n = 1.5$ and $n = 4$. The hedge ratios are only estimated by the equation (33). The result are shown in Table 2 and Table 5. For a given target return, we observe that the optimal hedge ratios decrease as the order of the GSV increases.

7 Conclusions

In this paper, first, we have reviewed various approaches to deriving the optimal hedge ratio. Especially, we can derive the optimal hedge ratio by maximizing the expected exponential utility, which consistence with semivariance of hedge ratios under our assumptions. From the equation (38), we find that risk premium is impact on hedge ratios, too. Second, we showed that under some circumstances the assumption of symmetrical distribution is sufficient to have the semivariance hedge ratio converge to the minimum variance hedge ratio. All these approaches will lead to the same hedge ratio as the minimum variance hedge ratio if the futures price follows a pure martingale process and if the futures and spot prices are the best linear predictor assumption. Third, we add one more scaling parameter to the **GH** family in order to good fit sample variance, and given its density by the equation (8). Fourth, we find that Nasdq100 Future and spot log-return are hyperbolic distributions but not bivariate hyperbolic distributions during from 2000 to 2004 years. Finally, we propose an alternative class of bivariate hyperbolic distributions (using numerical method) belonging to affine transformed, and adopt hyperbolic distributions estimate semivariance of hedge ratios via numerical method. The results are shown in Hedge Ratios of Table, and relaxed correlation coefficients being constant. In particular, when the target return is below zero, the analysis of conational correlations are even more important for risk managers who are concerned with semivariance or higher order.

References

[1] Ang, A., J. Chen and Y. Xing (2001). Downside risk and the momentum effect, Working paper. Columbia University.

[2] Atkinson, A.C.(1982). The simulation of generalized inverse Gaussian and hyperbolic random variables, SIAM Journal of Scientific & Statistical Computing 3, 502-515

[3] Barndorff-Nielsen, O.E. (1977). Expoenentially decreasing distributions for the logarothm of partical size. Proceedings of the Royal Society London A 353, 401-419.

[4] Barndorff-Nielsen, O.E. (1978). Hyperbolic distributions and distributions on hyperbolae, Scand. J. Statist. 151-157.

[5] Bawa, V.S. (1975). Optimal Rules For Orderign Uncertain Prospects. Journa of Financial Economics 2, 95-121.

[6] Bawa, V.S. (1978). Safety-first, stochastic dominance, and optimal portfolio choice. Journal of Financial and Quantutative Analysis 13, 255-271.

[7] Bekaert, G. and C. R. Harvey. (1997). Emerging equity market volatility. Journal of Financial Economics 43, 29-77.

[8] Benninga, S., Elodor, R.,and Zilcha,I. (1983). Optimal hedging in the futures market under price uncertainty. Economics Letters 13, 141-145.

[9] Bibby, B.M. and Sørensen, M. (2003). Hyperbolic Processes in Finance. In: Rachev, S. T. (Ed.) Handbook of Heavy Tailed Distributions in Finance. Elsevier Science, 211–248.

[10] Blæsid, P. and Sørensen, M. (1992) 'hyp' a computer program for analyzing data by means of the hyperbolic distribution. Research Reort 248, Department of Theoretical Statistics, Aarhus University.

[11] Chen, S.S.,Lee, C.F., and Shrestha, K.(2001). On a mean-generalized semivariance approach to determining the hedge ratio. Journal of Futures Markets 21, 581-598.

[12] Eberlein, E. and keller, U. (1995). Hyperbolic distributions in finance. Bernoulli 1, 281-299.

[13] Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. American Economic Review, 67, 116-126.

[14] Howard, C.T.,and D'Antonio,L.J. (1984). A risk-return measure of hedging effectiveness. Journal of Financial and Quantitative Analysis, 19, 101-112.

[15] Lien, D.,and Tse.,Y.K. (1998). Hedging time-varying downside risk. Journal of Futures Markets, 18, 705-722.

[16] Lien, D.,and Tse.,Y.K. (2000). Hedging downside risk with futures contracts. Applied Financial Economics, 10, 163-170.

[17] Lien, D.,and Tse.,Y.K. (2001). Hedging downside risk:futures vs. options. International Review of Economics and Finance, 10, 159-169.

[18] Prause, K.(1999). The Generalized Hperbolic Model: Estimation, Financial Derivatitivate, and Risk Meansues. Mathematische Fakultät der Albert-Ludwigs-Universität Feriburg i. Br.

[19] Schmidt, R., Stützle, E.,and Hrycej, T.(2003). Multivariate distribution models with generalized hyperbolic margins. Postdoc-Programme of the German Academic Exchange Service.

[20] Scott, W. (1979). On optimal and data-based histograms. Biometrika, 33, 605-610.

[21] Silverman, B.W. (1986) Density Estimation for Statistics and Data Analyses Chapman and Hall, New York.

Appendices

A Lemma2.2

Proof Lemma 2.2 of (5)

$$\begin{aligned}
\mathbb{E}[X^r] &= \int_0^\infty \frac{(\psi/\delta)}{2K_\lambda(\delta\psi)} x^{\lambda+r-1} e^{\frac{-1}{2}(\delta^2 x^{-1}+\psi^2 x)} dx \\
&\overset{y=(\psi/\delta)x}{=} \frac{(\psi/\delta)}{2K_\lambda(\delta\psi)} \int_0^\infty \left(\frac{\delta}{\psi} y\right)^{\lambda+r-1} e^{\frac{-1}{2}\delta\psi(y^{-1}+y)} \left(\frac{\delta}{\psi}\right) dy \\
&= \left(\frac{\delta}{\psi}\right)^r \frac{K_{\lambda+r}(\delta\psi)}{K_\lambda(\delta\psi)}.
\end{aligned}$$

Lemma 2.2 of (6)

$$\begin{aligned}
& d_{GIG(\lambda,0,\psi)} * d_{GIG(-\lambda,\delta,\psi)}(x) = \\
= & \int_0^\infty \left(\frac{\psi^2}{2}\right)^\lambda \frac{(x-y)^{\lambda-1}}{\Gamma(\lambda)} e^{\frac{-\psi^2(x-y)}{2}} \mathbb{I}_{\{x-y\}}(y) \left(\frac{\psi}{\delta}\right)^{-\lambda} \frac{1}{2K_{-\lambda}(\delta\psi)} y^{-\lambda-1} \\
\times & e^{-\frac{1}{2}(\delta^2 y^{-1}+\psi^2 y)} dy \\
= & \left(\frac{\delta\psi}{2}\right)^\lambda \frac{e^{\frac{-\psi^2}{2}x}}{2K_\lambda(\delta\psi)\Gamma(\lambda)} \int_0^x \frac{1}{y^2}\left(\frac{x}{y}-1\right)^{\lambda-1} e^{-\frac{1}{2}\delta^2 y^{-1}} dy \\
\overset{z=x/y}{=} & \left(\frac{\delta\psi}{2}\right)^\lambda \frac{e^{\frac{-\psi^2}{2}x}}{2K_\lambda(\delta\psi)\Gamma(\lambda)} \frac{1}{x} \int_1^\infty (z-1)^{\lambda-1} e^{-\frac{1}{2}\delta^2 z x^{-1}} dz \\
= & \left(\frac{\delta\psi}{2}\right)^\lambda \frac{e^{-\frac{1}{2}(\delta^2 x^{-1}+\psi^2 x)}}{2K_\lambda(\delta\psi)\Gamma(\lambda)} \frac{1}{x} \int_1^\infty (z-1)^{\lambda-1} e^{-\frac{1}{2}\delta^2 (z-1) x^{-1}} dz \\
= & \left(\frac{\delta\psi}{2}\right)^\lambda \frac{e^{-\frac{1}{2}(\delta^2 x^{-1}+\psi^2 x)}}{2K_\lambda(\delta\psi)\Gamma(\lambda)} \frac{1}{x} \frac{\Gamma(\lambda)(2x)^\lambda}{\delta^{2\lambda}}. \\
= & \left(\frac{\psi}{\delta}\right)^\lambda \frac{1}{2K_\lambda(\delta\psi)} x^{\lambda-1} e^{-\frac{1}{2}(\delta^2 x^{-1}+\psi^2 x)} = d_{GIG(\lambda,\delta,\psi)}
\end{aligned}$$

B Lemma3.2

Proof Lemma 3.2

Define the following two sets on the sample space:

$$K_1 = \left\{\omega | \tilde{X}(\omega) > \mu\right\}, \qquad K_2 = \left\{\omega | \tilde{X}(\omega) \leq \mu\right\}.$$

It is clear that the following inequality holds if $A > 0$

$$U'\left(A\tilde{X}(\omega_1) + B\right) < U'\left(A\tilde{X}(\omega_2) + B\right), \text{ for all } \omega_1 \in K_1, \omega_2 \in K_2.$$

But since

$$\int_{K_1} \left(\tilde{X}(\omega_1) - \mu\right) dp(\omega_1) = \int_{K_2} \left(\tilde{X}(\omega_2) - \mu\right) dp(\omega_2).$$

We obtain that,

$$\int_{K_1} \left(\tilde{X}(\omega_1) - \mu\right) U'\left(A\tilde{X}(\omega_1) + B\right) dp(\omega_1) < \int_{K_2} \left(\tilde{X}(\omega_2) - \mu\right) U'\left(A\tilde{X}(\omega_2) + B\right) dp(\omega_2),$$

which is a contradiction to the equation (25) since $K_1 \cup K_2 = \Omega$. On the other hand, if $A < 0$ then we can apply the same discussion. Therefore the only case will be held, when $A = 0$.

Proof Proposition 3.3

Let us rewrite (24) as follows:

$$\max_X \mathbb{E}_{\tilde{F}_1, \tilde{b}_1} \left\{U\left[(\beta_1 Q - X)\tilde{F}_1 + XF_0 + \alpha_1 Q + Q\tilde{b}_1\right]\right\},$$

where the subscripts denote expectation with respect to $\tilde{F}_1$ and $\tilde{b}_1$. The first order condition is

$$\mathbb{E}_{\tilde{F}_1, \tilde{b}_1} \left\{(F_0 - \tilde{F}_1)U'\left[(\beta_1 Q - X)\tilde{F}_1 + XF_0 + \alpha_1 Q + Q\tilde{b}_1\right]\right\}\Big|_{X^*} = 0. \tag{52}$$

By separability assumption, the equation (52) can be written as

$$\mathbb{E}_{\tilde{b}_1} \left\{\mathbb{E}_{\tilde{F}_1} \left\{(F_0 - \tilde{F}_1)U'\left[(\beta_1 Q - X^*)\tilde{F}_1 + X^* F_0 + \alpha_1 Q + Q\tilde{b}_1\right]\right\}\right\} = 0.$$

The equation (52) certainly holds if for almost all ω,

$$\mathbb{E}_{\tilde{F}_1} \left\{(F_0 - \tilde{F}_1)U'\left[(\beta_1 Q - X^*)\tilde{F}_1 + X^* F_0 + \alpha_1 Q + Q\tilde{b}_1(\omega)\right]\right\} = 0.$$

We reformulate the equations (53),

$$\begin{aligned}
& \mathbb{E}_{\tilde{F}_1}\left\{\tilde{F}_1 U'\left[(\beta_1 Q - X^*)\tilde{F}_1 + X^* F_0 + \alpha_1 Q + Q\tilde{b}_1(\omega)\right]\right\} \\
= \; & F_0 \mathbb{E}_{\tilde{F}_1}\left\{U'\left[(\beta_1 Q - X^*)\tilde{F}_1 + X^* F_0 + \alpha_1 Q + Q\tilde{b}_1(\omega)\right]\right\} \\
= \; & \mathbb{E}_{\tilde{F}_1}(F_1)\, \mathbb{E}_{\tilde{F}_1}\left\{F_0 U'\left[(\beta_1 Q - X^*)\tilde{F}_1 + X^* F_0 + \alpha_1 Q + Q\tilde{b}_1(\omega)\right]\right\},
\end{aligned}$$

for almost all ω. Hence, holds at $\beta_1 Q - X^* = 0$.

Proof Let $r_\theta = r_p - \theta r_f$

$$\int_{-\infty}^{c} (c - r_\theta)^2 dG(r_\theta) = \int_{-\infty}^{c} (c - r_\theta)^2 g(r_\theta) dr_\theta,$$

where $G(\cdot)$ denote the probability distribution function of r_θ.

We have

$$G(a) = Pr\{\widetilde{r_\theta} \leq a\} = Pr\{\widetilde{r_p} - \theta\widetilde{r_f} \leq a\} = \int_{-\infty}^{\infty} Pr\{\widetilde{r_p} \leq a + \theta r_f | \widetilde{r_f} = r_f\} f_{r_f}(r_f) dr_f$$

$$= \int_{-\infty}^{\infty} F(a + \theta r_f | r_f) f_{r_f}(r_f) dr_f = \int_{-\infty}^{\infty} \int_{-\infty}^{a+\theta r_f} f(r_p, r_f) dr_p dr_f$$

and, $$\frac{G(a)}{da} = g(a) = \int_{-\infty}^{\infty} f(a + \theta r_f, r_f) dr_f.$$

Hence,

$$\int_{-\infty}^{c} (c - r_\theta)^2 g(r_\theta) dr_\theta = \int_{-\infty}^{c} \int_{-\infty}^{\infty} (c - r_\theta)^2 f(r_\theta + \theta r_f, r_f) dr_f dr_\theta$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{c} (c - r_\theta)^2 f(r_\theta + \theta r_f, r_f) dr_\theta dr_f$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{c} (c - r_\theta)^2 f(r_\theta + \theta r_f | r_f) dr_\theta f_{r_f}(r_f) dr_f$$

Since $r_\theta = r_p - \theta r_f$, we apply the variable transformation from r_h to r_p.

The result is

$$\int_{-\infty}^{\infty} \int_{-\infty}^{c+\theta r_f} (c - r_p + \theta r_f)^2 f(r_p | r_f) dr_p f_{r_f}(r_f) dr_f \tag{53}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{c+\theta r_f} (c - r_p + \theta r_f)^2 f(r_p, r_f) dr_p dr_f. \tag{54}$$

We turn to use the below equations to calculate first derivative condition of the equation (54)

$$H(y) = \int_{a(y)}^{b(y)} h(x, y) dx, \text{ then}$$

$$H'(y) = \int_{a(y)}^{b(y)} h_y(x, y) dx + h[b(y), y]b'(y) - h[a(y), y]a'(y)$$

We can obtain its first derivative condition is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta r_f} r_f(c-r_p+\theta r_f)f(r_p,r_f)dr_p dr_f,$$

and second derivative condition is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{c+\theta r_f} r_f^2 f(r_p,r_f)dr_p dr_f > 0.$$

C Theorem 4.5

Proof Theorem 4.5.

Without loss the generality. Assumes throughout that the interval $\mathbb{R}$=$(a, b]$ with $a > 0$ for U_i, $i = 1...4$.

We note that

$$\begin{aligned} \Delta\mathbb{E}u &\equiv \mathbb{E}U(r_{\theta_F}) - EU(r_{\theta_G}) \\ &= \int_a^b u(x)dF - \int_a^b u(x)dG \\ &= \int_a^b u(x)d(F(x) - G(x)). \end{aligned}$$

Carrying out integration by parts several times with $H_1(x) = F(x) - G(x)$,

$$H_n(x) = \int_a^x H_{n-1}(y)dy,$$

for $n \geq 2$.

We obtain the following equivalent expressions for $\Delta\mathbb{E}u$

$$\Delta\mathbb{E}u = -\int_a^b u'(x)H_1(x)dx, \tag{55}$$

$$\Delta\mathbb{E}u = -u'(b)H_2(b) + \int_a^b u''(x)H_2(x)dx, \tag{56}$$

$$\Delta\mathbb{E}u = -u'(b)H_2(b) + u''(b)H_3(b) - \int_a^b u'''(x)H_3(x)dx. \tag{57}$$

To prove (1), We obtain

$$L_0(c, r_{\theta_F}) \leq L_0(c, r_{\theta_G})$$

is implying $H_1(c) \leq 0$, $\forall$ c $\in\mathbb{R}$ and using the equation (55) proves the sufficiency.

Conversely, we want to find utility $u_1 \in U_1$ such that $\Delta\mathbb{E}u < 0$.

Consider following the function, with $\varepsilon = \gamma\delta$

$$\xi(x) = \begin{cases} \varepsilon, & a < x \leq x_0, \\ \varepsilon - \gamma(x - x_0), & x_0 \leq x \leq x_0 + \delta, \\ 0, & x_0 + \delta \leq x, \end{cases}$$

For some arbitrary x_0, $-\infty < 0 \leq a < x_0 < x_0 + \delta \leq b$. Hence, we obtain that utility function $u_1(x)$ defined by $u_1'(x) \equiv k - \xi'(x)$, where $k > 0$; then $u_1'(x) > 0, u_1(x) \in U_1$.

Hence, we obtain substituting in (55),

$$\Delta \mathbb{E} u_1 = \int_{x_0}^{x_0+\delta} -(\gamma + k) H_1(x) dx + M_1.$$

Thus, by choosing $(k+\gamma)$ large enough, we can make $\Delta \mathbb{E} u_1 < 0$. This completes proof of (1).

To prove (2), We obtain

$$L_1(c, r_{\theta_F}) \leq L_1(c, r_{\theta_G})$$

is implying $H_2(c) \leq 0$, $\forall$ c $\in \mathbb{R}$ and using the equation (56) proves the sufficiency. Conversely, we want to find utility $u_2 \in U_2$ such that $\Delta \mathbb{E} u < 0$.
We consider utility function $u_2(x)$ defined by $u_2''(x) \equiv -k + \xi'(x)$, where $k > 0$.
Then $u_2''(x) < 0$, $u_2'(x) \equiv k_1 - k(x-a) + \xi(x) > 0$ by appropriate choice of $k_1, u_2 \in U_2$.
We again obtain substituting in (56),

$$\Delta \mathbb{E} u_2 = M_2 + \int_{x_0}^{x_0+\delta} (-k-\gamma) H_2(x) dx$$

We can make $\Delta \mathbb{E} u_2 < 0$ as $(\gamma + k)$ large enough. This completes the proof of (2).

To prove (3), we consider utility function u_3 defined by $u_3''' \equiv \varepsilon_1 - \xi'(x)$, where $\varepsilon_1 > 0$.
We note that $u_3'''(x) > 0$, $u_3''(x) = -(k_1 - \varepsilon_1(x-a)) - \xi(x) < 0$ by appropriate choice of k_1. Similarly,

$$u_3'(x) \equiv k_2 - (x-a)\left(k_1 - \frac{\varepsilon_1}{2}(x-a)\right) - \int_a^x \xi(y) dy,$$

where

$$\int_a^x \xi(y) dy = \begin{cases} \varepsilon(x-a), & a \leq x \leq x_0, \\ \varepsilon(x-a) - \frac{\gamma}{2}(x-x_0)^2, & x_0 \leq x \leq x_0+\delta, \\ \varepsilon(x_0 - a + \delta) - \frac{\gamma}{2}\delta^2, & x_0 + \delta \leq x, \end{cases}$$

and k_2 is an arbitrary constant chosen such that $u_3'(x) > 0$ $\forall$ $x \in \mathbb{R}$; hence $u_3(x) \in U_3$.
We let

$$k_1 = \varepsilon(b-a)(1+k_1'), \tag{58}$$

$$k_2 = \varepsilon(b-a)^2(k_1' + k_2' + \frac{1}{2}) + \varepsilon(x_0 - a + \delta) - \frac{\gamma}{2}\delta^2 + k_3\gamma\delta^2; \tag{59}$$

$$M_3 = k_2'(b-a)^2 \Delta\mu - k_1'(b-a)H_3(b) - \int_a^b H_3(x) dx, \tag{60}$$

where $-H_2(b) \equiv \Delta\mu = \mathbb{E}(r_{\theta_F}) - \mathbb{E}(r_{\theta_G})$.
We use the equations (58-60) to substitute in the equation (57) and obtain

$$\begin{aligned}\Delta\mathbb{E}u_3 &= k_3\gamma\delta^2\Delta\mu + \varepsilon_1 M_3 - \gamma\int_{x_0}^{x_0+\delta} H_3(x)dx \\ &\leq \gamma\delta[k_3\delta\Delta\mu - M_4] + \varepsilon_1 M_3,\end{aligned}$$

where for $x \in [x_0, x_0+\delta]$, $H_3(x) \geq M_4 > 0$. Thus, if we let $\varepsilon_1 \to 0, \delta \to 0^+, \gamma \to \infty$ such that $\varepsilon = \gamma\delta \to d > 0$, then $\Delta\mathbb{E}u_3 < 0$ for sufficient small value of δ.
Next, we consider utility function $u_4(x)$ defined by

$$u_4'(x) = exp(e^{-px}),$$

where $p > 0$. Then we note that $u_4(x) \in U_3$ and obtain substituting in the equation (57)

$$\Delta\mathbb{E}u_4 = -exp(e^{-pb})[H_2(b) + pe^{-pb}H_3(b)] + \int_a^b exp(e^{-bx}\left(-p^2e^{-px}(1+e^{-px})\right)H_3(x)dx$$

It thus follows that since $(-p^2e^{-px}(1+e^{-px})) \to 0^-$ as $p \to 0$, if $H_2(b) > 0$, $\Delta E u_4 < 0$ for sufficiently small value of p. This completes proof of (3).

To prove (4). We obtain from (57), remembering that $r(x) \equiv -u''(x)/u'(x)$ and $H_2(b) = 0$,

$$\Delta\mathbb{E}u = u''(b)H_3(b) + \int_a^b u'(x)[r'(x) - (r(x))^2]H_3(x)dx.$$

For $u(x) \in U_4$, $r'(x) < 0$, $r(x) > 0$ and thus $H_3(x) \leq 0\ \forall\ x$ and < 0 for some x is a sufficient condition for dominance.
To prove necessity, we consider a utility function $u_5(x)$ defined by $r_5'(x) \equiv \xi'(x)$, and note that $u_5(x) \in U_4$, and we obtain

$$\begin{aligned}\mathbb{E}u_5 = &- \varepsilon^2\int_a^{x_0} u_5'(x)H_3(x)dx - \gamma\int_{x_0}^{x_0+\delta} u_5'H_3(x)dx \\ &- \int_{x_0}^{x_0+\delta} u_5'(x)\left(\varepsilon - \gamma(x-x_0)^2\right)H_3(x)dx.\end{aligned}$$

Suppose that $H_3(x) \geq M_1 > 0$, for $x \in [x_0, x_0+\delta]$ and $|H_3(x)| < M_2\ \forall\ x \in \mathbb{R}$, then $\Delta\mathbb{E}u_5 < 0$.
If

$$\gamma M_1\int_{x_0}^{x_0+\delta} u_5'(x)dx > \varepsilon^2 M_2\int_a^{x_0} u_5'(x)dx \tag{61}$$

By Mean value theorem, hence we get

$$\gamma M_1 \int_{x_0}^{x_0+\delta} u_5'(x)dx = \gamma\delta M_1 u_5'(x_c),$$

where $x_c \in [x_0, x_0+\delta]$.

$$M_1 u_5'(x_0) > \varepsilon M_2 \int_a^x u_5'(x)dx$$

We note that with $\delta \to 0$ such that $\varepsilon \to 0$, inequality (61) will hold for sufficiently small δ. This completes the proof of (4).

Assume we are given a strictly positive and stationary process $(\sigma_t)_{t\geq 0}$ independent of B_t with $\sigma_t \stackrel{\mathcal{D}}{=} GIG(\lambda, \gamma, \psi)$ for all $t \geq 0$

$$dX_t = (\mu + \beta\sigma_t^2)dt + \sigma_t dB_t,$$

where $X_t = \log \frac{F_1}{F_0}$.
Using the Itô formula we can describe the stochastic differential equation of F by

$$dF_t = F_t(\mu + \bar{\beta}\sigma_t^2)dt + F_t\sigma_t dB_t$$

where $\bar{\beta} = \beta + \frac{1}{2}$.
Take $dt = 1$, we approximately

$$r_f = \frac{F_1 - F_0}{F_0} \stackrel{\mathcal{D}}{=} N(\mu + \bar{\beta}\sigma_1^2, \sigma_1^2).$$

The p.d.f of $f(r_f)$ is $d_{GH(\lambda,\alpha,\bar{\beta},\gamma,\mu)}$

GH of hedge ratio

Proof the equation (38)

Consider $U(x) = 1 - e^{-x\eta}$

Let $x = \widetilde{r_p} - \theta \widetilde{r_f}$ and $\widetilde{r_p} = \mu_{r_p} + \rho \frac{\sigma_{r_p}}{\sigma_{r_f}} (\widetilde{r_f} - \mu_{r_f})$

We consider that

$$\begin{aligned}
& \int e^{\theta r_f \eta} \left\{ \mathbb{E}[e^{-r_p \eta} | r_f] \right\} f(r_f) dr_f \\
= & \int e^{(-\mu_{r_p} + \mu_{r_f} \rho \frac{\sigma_{r_p}}{\sigma_{r_f}} + r_f(\theta - \rho \frac{\sigma_{r_p}}{\sigma_{r_f}}))\eta} f(r_f) dr_f \\
= & \ \xi \int e^{r_f(\theta - \rho \frac{\sigma_{r_p}}{\sigma_{r_f}})\eta} f(r_f) dr_f \\
= & \ \xi e^{\mu_f t} \left(\frac{\alpha_f^2 - \bar{\beta}_f^2}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right)^{\lambda/2} \frac{K_\lambda(\gamma_f \sqrt{\alpha_f^2 - (\bar{\beta}_f + t)^2})}{K_\lambda(\gamma_f \sqrt{\alpha_f^2 - \bar{\beta}_f^2})} \\
\approx & \ e^{\mu_f t} \left(\frac{\alpha_f^2 - \bar{\beta}_f^2}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right)^{\lambda}, \quad as \quad \gamma_f \to 0
\end{aligned}$$

where $t = (\theta - \rho \frac{\sigma_{r_p}}{\sigma_{r_f}})\eta$ and, $\xi = e^{(-\mu_{r_p} + \mu_{r_f} \rho \frac{\sigma_{r_p}}{\sigma_{r_f}})\eta} > 0$

Hence,

$$\min_{\theta} e^{\mu_f t} \left(\frac{\alpha_f^2 - \bar{\beta}_f^2}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right)^{\lambda}. \tag{62}$$

Suppose that $\mu_f = 0$. To take the first deriver for θ

$$\left[2\lambda \frac{\bar{\beta}_f + t}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right] \left(\frac{\alpha_f^2 - \bar{\beta}_f^2}{\alpha_f^2 - (\bar{\beta}_f + t)^2} \right)^{\lambda} \eta$$

Let above formula be equal $h(\theta)$. Then

$$h'(\theta)|_{\theta^*} = 2\lambda \left. \frac{\alpha_f^2 + (\bar{\beta}_f + t)^2}{(\alpha_f^2 - (\bar{\beta}_f + t)^2)^2} \right|_{\theta^*} > 0.$$

where

$$\theta^*_{GH} = \rho \frac{\sigma_{r_p}}{\sigma_{r_f}} - \bar{\beta}_f \frac{1}{\eta}.$$

Figure 1: Densities of compounded return

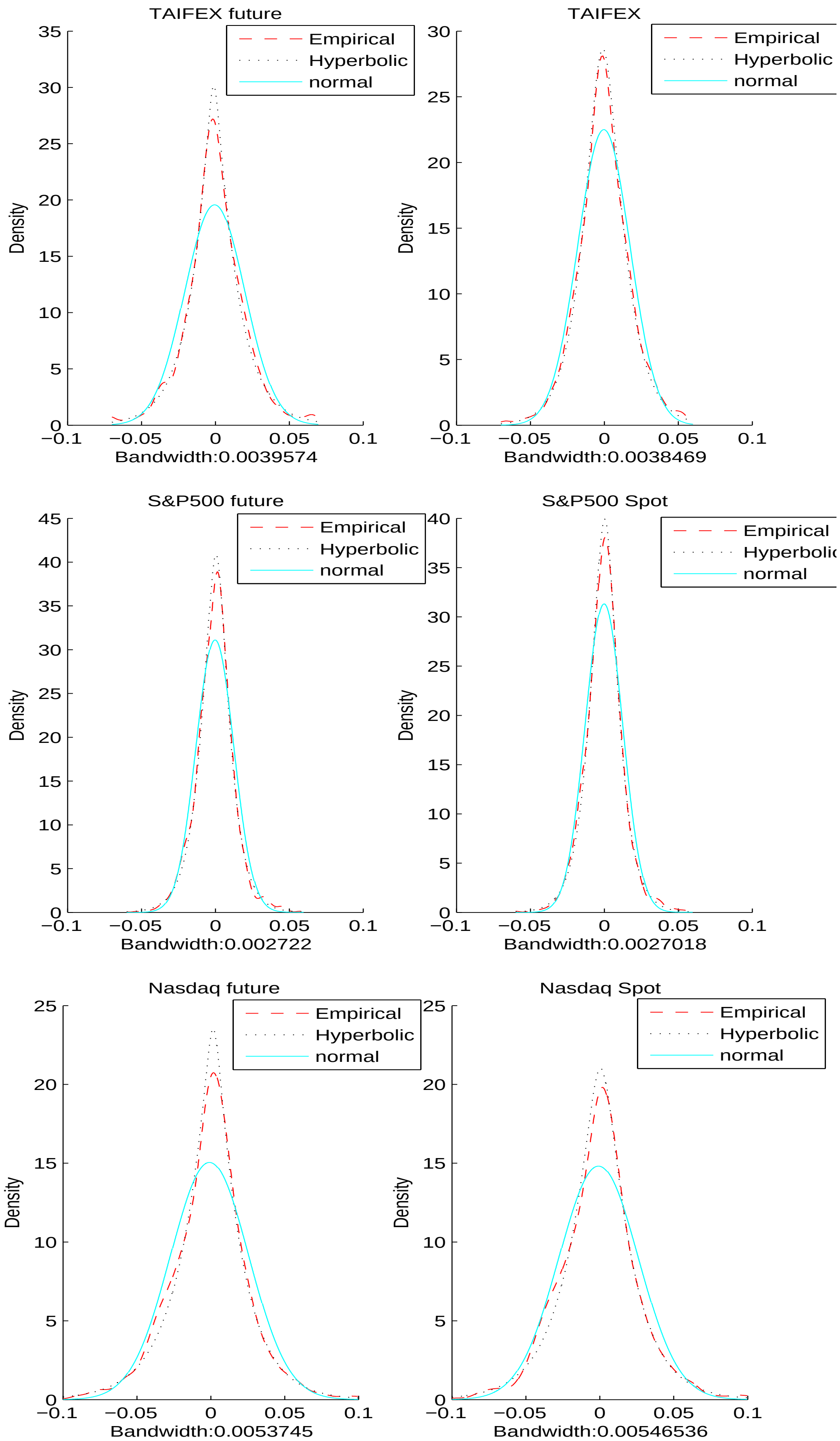

Figure 2: Log-densities of compounded return

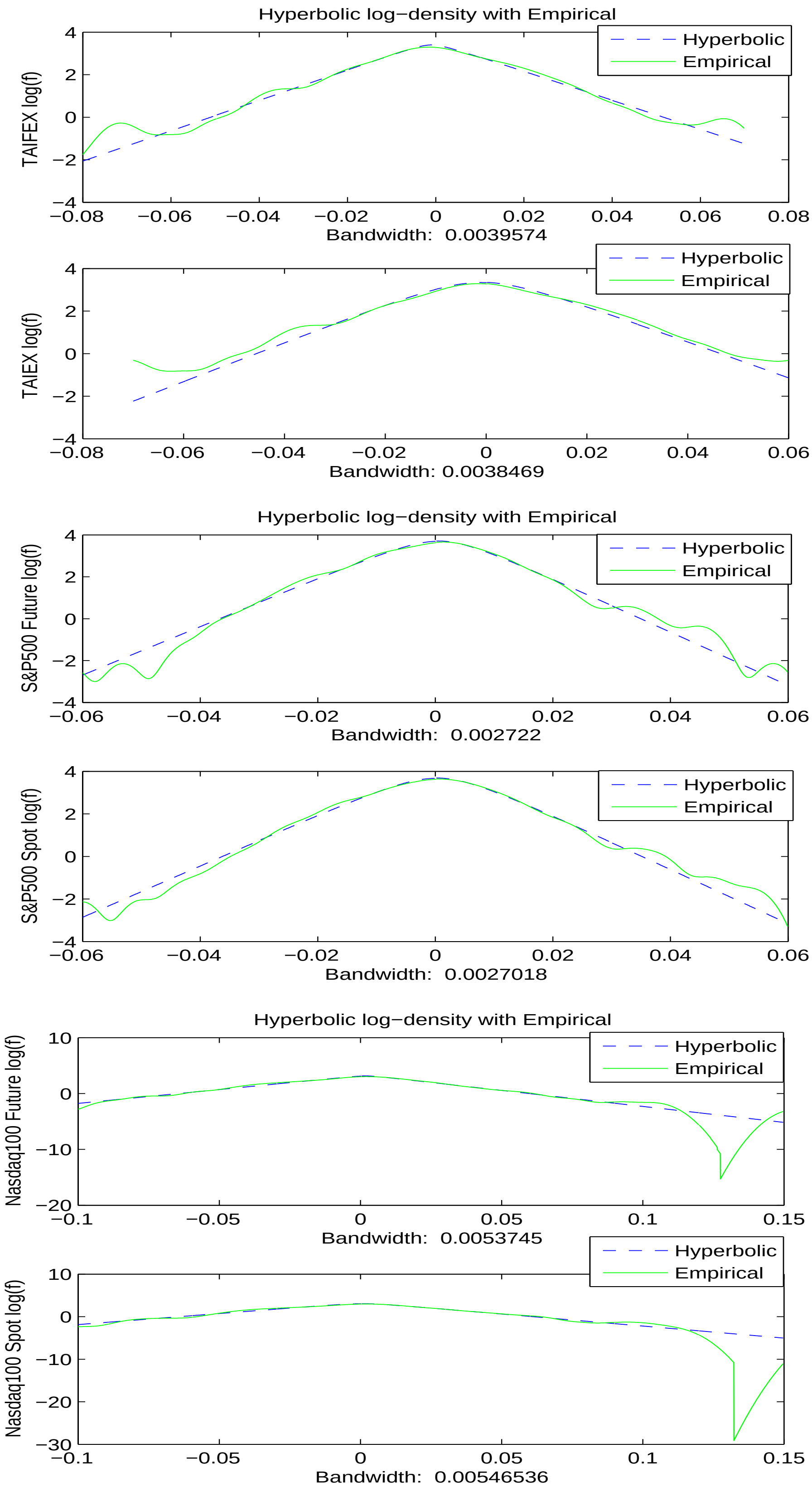

Figure 3: Symmetric $\mathbf{H}_2$ distributions

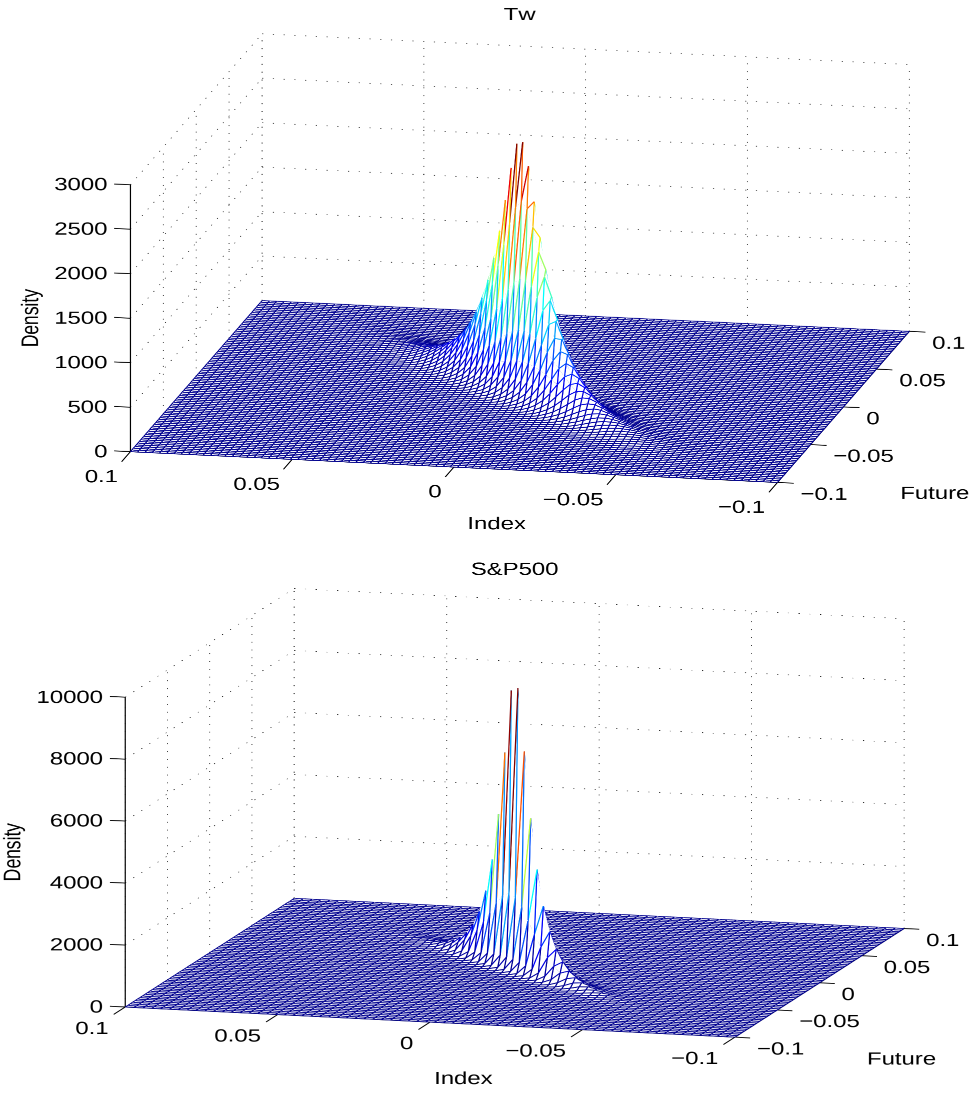

Table 2: GSV of Hedge Ratios (n=1.5)

	T R	θ^*_{nor}	$\theta^*_{H_2}$	$\theta^*_{H_\xi}$	$\theta^*_{H_\xi}$	$\theta^*_{H_\xi}$	θ_{ture}
				$\xi_f = 0.9824$	$\xi = 1$	$\xi_f = 1.015$	
				$\xi_p = 0.9789$	$\xi_p = 1$	$\xi_p = 0.934$	
	-0.005	0.8165	0.8182	0.8018	0.8025	0.7705	0.8266
		(2.54)	(2.99)	(4.44)	(4.35)	(3.58) $\times 10^{-4}$	
	0	0.8163	0.8183	0.8037	0.8050	0.7723	0.8224
		(1.42)	(1.41)	(2.31)	(2.97)	(2.13) $\times 10^{-4}$	
Tw	0.005	0.8164	0.8185	0.8034	0.8047	0.7724	0.8320
		(1.89)	(1.56)	(2.43)	(2.57)	(2.27) $\times 10^{-4}$	
				$\xi_f = 0.9960$	$\xi = 1$	$\xi_f = 1.072$	
				$\xi_p = 1$	$\xi_p = 1$	$\xi_p = 1.044$	
	-0.005	0.9650	0.9696	0.9649	0.9631	0.9506	0.9743
		(3.80)	(4.70)	(5.94)	(6.07)	(5.40) $\times 10^{-4}$	
	0	0.9670	0.9701	0.9605	0.9586	0.9457	0.9749
		(0.80)	(0.71)	(2.14)	(2.12)	(2.08) $\times 10^{-4}$	
S&P500	0.005	0.9659	0.9672	0.9605	0.9585	0.9458	0.9863
		(2.43)	(1.99)	(3.65)	(3.77)	(3.43) $\times 10^{-4}$	
				$\xi_f = 0.9670$	$\xi = 1$	$\xi_f = 0.992$	
				$\xi_p = 0.9765$	$\xi_p = 1$	$\xi_p = 1.147$	
	-0.005	0.9899		0.9949	0.9902	1.0638	1.0031
		(1.75)		(2.50)	(2.25)	(2.51) $\times 10^{-4}$	
	0	0.9918		0.9336	0.9886	1.0627	0.9971
		(0.83)		(1.23)	(1.22)	(1.39) $\times 10^{-4}$	
Nasaq100	0.005	0.9914		0.9935	0.9887	1.0630	0.9911
		(1.09)		(1.33)	(1.32)	(1.55) $\times 10^{-4}$	

Table 3: Semivariance of Hedge Ratios (GSV n=2)

	T R	θ^*_{nor}	$\theta^*_{H_2}$	$\theta^*_{H_\xi}$	$\theta^*_{H_\xi}$	$\theta^*_{H_\xi}$	θ_{ture}
				$\xi_f = 0.9824$	$\xi = 1$	$\xi_f = 1.015$	
				$\xi_p = 0.9789$	$\xi_p = 1$	$\xi_p = 0.934$	
	-0.005	0.8121	0.8123	0.7916	0.7903	0.7600	0.7884
		(1.65)	(3.65)	(4.14)	(4.33)	(3.57) $\times 10^{-4}$	
	0	0.8143	0.8176	0.7978	0.8011	0.7694	0.8131
		(1.24)	(2.12)	(2.82)	(2.97)	(2.75) $\times 10^{-4}$	
Tw	0.005	0.8161	0.8208	0.8039	0.8047	0.7733	0.8271
		(1.51)	(1.90)	(3.12)	(3.17)	(3.00) $\times 10^{-4}$	
				$\xi_f = 0.9960$	$\xi = 1$	$\xi_f = 1.072$	
				$\xi_p = 1$	$\xi_p = 1$	$\xi_p = 1.044$	
	-0.005	0.9608	0.9648	0.9699	0.9673	0.9531	0.8764
		(1.85)	(6.86)	(9.30)	(7.27)	(8.92) $\times 10^{-4}$	
	0	0.9656	0.9688	0.9611	0.9628	0.9470	0.9643
		(0.74)	(1.09)	(2.52)	(2.77)	(2.90) $\times 10^{-4}$	
S&P500	0.005	0.9687	0.9750	0.9612	0.9584	0.9447	0.9678
		(1.70)	(1.36)	(2.80)	(3.15)	(3.13) $\times 10^{-4}$	
				$\xi_f = 0.9670$	$\xi = 1$	$\xi_f = 0.992$	
				$\xi_p = 0.9765$	$\xi_p = 1$	$\xi_p = 1.147$	
	-0.005	0.9858		0.9933	0.9876	1.0606	0.9610
		(0.87)		(2.41)	(2.68)	(2.93) $\times 10^{-4}$	
	0	0.9899		0.9940	0.9882	1.0622	0.9854
		(0.80)		(1.50)	(1.44)	(1.99) $\times 10^{-4}$	
Nasaq100	0.005	0.9933		0.9933	0.9871	1.0625	0.9995
		(0.87)		(1.57)	(1.80)	(1.58) $\times 10^{-4}$	

In the second column, thy are different target returns (TR).

Table 4: Semivariance of Hedge Ratios under Conditional Correlations

	TR		-0.005	0	0.005
θ_{H_2}	θ_{ture}		0.7884	0.8131	0.8271
Tw	unconditional		0.8123(3.65)	0.8176(2.12)	0.8208(1.90)
	$\tau_1 = 0.5\sigma_1$	$\tau_2 = 7\sigma_1$	0.8092(4.11)	0.8137(2.04)	0.8168(1.92)
	$\tau_1 = \sigma_1$	$\tau_2 = 7\sigma_1$	0.7976(4.35)	0.8022(2.55)	0.8061(2.41)
	$\tau_1 = 2\sigma_1$	$\tau_2 = 7\sigma_1$	0.7535(6.29)	0.7574(4.27)	0.7604(4.10)
	$\tau_1 = 3\sigma_1$	$\tau_2 = 7\sigma_1$	0.6700(8.39)	0.6733(6.40)	0.6770(6.29)
	θ_{ture}		0.8764	0.9643	0.9678
S&P500	unconditional		0.9648(6.86)	0.9688(1.09)	0.9750(1.36)
	$\tau_1 = 0.5\sigma_2$	$\tau_2 = 7\sigma_2$	0.9596(6.51)	0.9660(1.249)	0.9722(1.390)
	$\tau_1 = \sigma_2$	$\tau_2 = 7\sigma_2$	0.9576(6.87)	0.9611(1.43)	0.9667(1.25)
	$\tau_1 = 2\sigma_2$	$\tau_2 = 7\sigma_2$	0.9278(7.42)	0.9349(2.63)	0.9413(2.51)
	$\tau_1 = 3\sigma_2$	$\tau_2 = 7\sigma_2$	0.8751(9.21)	0.8800(4.29)	0.8856(3.74)

The number in parenthesis are error of variance 10^{-4} times. σ_1 is standard deviation of Y of TAIEX market. σ_2 is standard deviation of Y of S&P500 market

	TR		-0.005	0	0.005	
θ_{nor}	θ_{ture}		0.7884	0.8131	0.8271	θ_{MV}
Tw	unconditional		0.8121(1.65)	0.8143(1.24)	0.8161(1.51)	0.8175
	$\tau_1 = 0.5\sigma_1$	$\tau_2 = 7\sigma_1$	0.8068(1.94)	0.8088(1.36)	0.8087(3.06)	0.8130
	$\tau_1 = \sigma_1$	$\tau_2 = 7\sigma_1$	0.7998(2.36)	0.7981(1.60)	0.7988(2.01)	0.8012
	$\tau_1 = 2\sigma_1$	$\tau_2 = 7\sigma_1$	0.7484(5.59)	0.7481(3.47)	0.7459(5.17)	0.7547
	$\tau_1 = 3\sigma_1$	$\tau_2 = 7\sigma_1$	0.6671(6.76)	0.6668(5.80)	0.6685(6.90)	0.6691
	θ_{ture}		0.8764	0.9643	0.9678	
S&P500	unconditional		0.9608 (1.85)	0.9656 (0.74)	0.9687(1.70)	0.9646
	$\tau_1 = 0.5\sigma_2$	$\tau_2 = 7\sigma_2$	0.9613 (3.22)	0.9637(9.24)	0.9625(2.62)	0.9646
	$\tau_1 = \sigma_2$	$\tau_2 = 7\sigma_2$	0.9538(3.21)	0.9567(1.16)	0.9555(3.24)	0.9582
	$\tau_1 = 2\sigma_2$	$\tau_2 = 7\sigma_2$	0.9270(3.80)	0.9304(1.55)	0.9301(5.33)	0.9314
	$\tau_1 = 3\sigma_2$	$\tau_2 = 7\sigma_2$	0.8739(5.97)	0.8737(3.40)	0.8732(6.54)	0.8747

Table 5: GSV of Hedge Ratios (n=4)
under Conditional Correlations

	TR	-0.005	0	0.005
θ_{H_2}	θ_{ture}	0.7702	0.7678	0.7714
Tw	unconditional	0.8134(7.82)	0.8138(7.80)	0.8138(7.86)
	$\tau_1 = 0.5\sigma_1$ $\tau_2 = 7\sigma_1$	0.8082(7.10)	0.8083(7.12)	0.8080(7.14)
	$\tau_1 = \sigma_1$ $\tau_2 = 7\sigma_1$	0.7986(11)	0.7987(11)	0.7986(11)
	$\tau_1 = 2\sigma_1$ $\tau_2 = 7\sigma_1$	0.7448(19)	0.7449(19)	0.7449(19)
	$\tau_1 = 3\sigma_1$ $\tau_2 = 7\sigma_1$	0.6646(22)	0.6647(22)	0.6646(23)
	θ_{ture}	0.8332	0.8452	0.8542
S&P500	unconditional	0.9648(6.09)	0.9652(4.79)	0.9641(5.14)
	$\tau_1 = 0.5\sigma_2$ $\tau_2 = 7\sigma_2$	0.9626(6.50)	0.9628(5.20)	0.9624(5.25)
	$\tau_1 = \sigma_2$ $\tau_2 = 7\sigma_2$	0.9558(7.23)	0.9562(6.313)	0.9565(6.60)
	$\tau_1 = 2\sigma_2$ $\tau_2 = 7\sigma_2$	0.9284(12)	0.9285(11)	0.9278(11)
	$\tau_1 = 3\sigma_2$ $\tau_2 = 7\sigma_2$	0.8712(21)	0.8713(20)	0.8709(20)
	$\tau_1 = 4\sigma_2$ $\tau_2 = 7\sigma_2$	0.7618(35)	0.7618(35)	0.7619(35)

The number in parenthesis are error of variance 10^{-4} times. σ_1 is standard deviation of Y of TAIEX market. σ_2 is standard deviation of Y of S&P500 market

	TR	-0.005	0	0.005	
θ_{nor}	θ_{ture}	0.7702	0.7678	0.7714	θ_{MV}
Tw	unconditional	0.8116(2.46)	0.8114(2.45)	0.8113(2.45)	0.8175
	$\tau_1 = 0.5\sigma_1$ $\tau_2 = 7\sigma_1$	0.8077(3.07)	0.8072(3.04)	0.8074(3.09)	0.8130
	$\tau_1 = \sigma_1$ $\tau_2 = 7\sigma_1$	0.7918(4.27)	0.7918(4.32)	0.7916(4.39)	0.8012
	$\tau_1 = 2\sigma_1$ $\tau_2 = 7\sigma_1$	0.7427(5.36)	0.7427(5.41)	0.7424(5.49)	0.7547
	$\tau_1 = 3\sigma_1$ $\tau_2 = 7\sigma_1$	0.6602(6.81)	0.6602(6.79)	0.6603(6.83)	0.6691
	θ_{ture}	0.8332	0.8452	0.8542	
S&P500	unconditional	0.9620(2.55)	0.9635 (1.42)	0.9637(1.92)	0.9646
	$\tau_1 = 0.5\sigma_2$ $\tau_2 = 7\sigma_2$	0.9597(2.21)	0.9605(1.22)	0.9592(1.88)	0.9646
	$\tau_1 = \sigma_2$ $\tau_2 = 7\sigma_2$	0.9556(2.59)	0.9555(1.84)	0.9551(2.34)	0.9582
	$\tau_1 = 2\sigma_2$ $\tau_2 = 7\sigma_2$	0.9249(3.11)	0.9257(2.92)	0.9261(3.36)	0.9314
	$\tau_1 = 3\sigma_2$ $\tau_2 = 7\sigma_2$	0.8695(4.48)	0.8701(4.30)	0.8697(4.62)	0.8747
	$\tau_1 = 4\sigma_2$ $\tau_2 = 7\sigma_2$	0.7631(10)	0.7633(10)	0.7630(10)	0.7687

Table 6: About distance

	TAIEX Future			TAIEX		
	Hyperbolic		Normal	Hyperbolic		Normal
$\xi_{f,p}$	1	1.015		1	0.934	
D_k	0.0153		0.0655	0.0152		0.0434
D_{AD}		0.0711			0.0543	
	S&P Future			Spot		
	Hyperbolic		Normal	Hyperbolic		Normal
$\xi_{f,p}$	1	1.072		1	1.044	
D_k	0.0148		0.0488	0.0142		0.0445
D_{AD}		0.0574			0.0530	
	Nasdaq 100 Future			Spot		
	Hyperbolic		Normal	Hyperbolic		Normal
$\xi_{f,p}$	1	0.992		1	1.147	
D_k	0.0217		0.0619	0.0234		0.0628
D_{AD}		0.0616			0.0915	

Table 7: Basic statistics

Products	number	mean $\times 10^{-4}$	variance$\times 10^{-4}$	skewness	kurtosis	corrlation
TAIEX Future	1261	-2.8313	4.1541	-0.0807	4.8848	
$H_f(\xi = 0.9824)$		-2.8313	4.1541	0.0933	5.8330	
$H_f(\xi = 1)$		-2.8313	4.2284	0.0925	5.8329	
TAIEX	1261	-2.8154	3.1429	0.0103	4.1432	
$H_s(\xi = 0.9789)$		-2.8154	3.1429	0.1211	5.0114	
$H_s(\xi = 1)$		-2.8154	3.2104	0.1198	5.0112	0.9399
S&P500 Future	1256	-1.5092	1.6454	0.0562	4.8801	
$H_f(\xi = 0.9960)$		-1.5092	1.6454	-0.1587	5.1544	
$H_f(\xi = 1)$		-1.5092	1.6520	-0.1584	5.1544	
S&P500 Spot	1256	-1.4577	1.6238	0.1221	4.7822	
$H_s(\xi = 1)$		-1.4691	1.6467	-0.0873	4.9928	0.9734
Nasdaq100 Future	1256	-6.8444	7.0224	0.1478	5.3631	
$H_f(\xi = 0.9670)$		-6.8444	7.0224	-0.2968	5.9377	
$H_f(\xi = 1)$		-6.8444	7.2596	-0.2920	5.9357	
Nasdaq100 Spot	1256	-6.7680	7.2455	0.3078	5.5348	
$H_s(\xi = 0.9765)$		-6.7680	7.2455	-0.1590	5.5822	
$H_s(\xi = 1)$		-6.7680	7.4195	-0.1571	5.5817	0.9760