Dai Jun (China)

A dynamic hedging model based on conditional higher moments

Abstract

By utilizing a Taylor series expansion of a utility function, this study investigates how the introduction of higher moments affects the investor's objective utility function. Using the bivariate GARCH-SK model and Independent Component Analysis (ICA), the authors build a dynamic model to describe the conditional higher moment risk of returns. Then, the empirical application of this new model is performed on the CSI 300 (China Securities Index 300) index futures and spot markets. The empirical results show that after the introduction of transaction costs, the optimal adjustment frequency of optimal hedge ratios will rise within a narrow range, as the investor's risk aversion coefficients increase. Additionally, although minimum-variance (MV) hedging strategy can effectively minimize hedged portfolio risks of variance and leptokurtosis, the dynamic utility maximized hedging strategy, which considers the conditional higher moments, can better balance revenues and risks, and generate the higher investors' utility.

Keywords: higher moments, stock index futures, dynamic hedging, independent component analysis.
JEL Classification: C10, G13.

Introduction

The determination of optimal hedge ratios is always the core of futures hedging theories. Following the seminal study of Johnson (1960), a large body of literature in futures hedging has centered on the MV strategy, which has the merits of computational and understanding simplicity. Nevertheless, the MV hedging strategy implicitly assumes that investors have infinitely great risk aversion coefficients, which is obviously unrealistic. Later, some scholars suggested the optimal hedge ratio should be determined under the expected utility maximization paradigm. With the utility function solely approximated by the mean and variance of asset returns, the random variables of asset returns are assumed to be normally distributed, which has been widely found to be a restrictive assumption in practice. A large amount of empirical evidence suggests that many financial asset returns display the significant features of peaks, fat tail, biased, etc, so that they are not normal (see Engle, 1982; Bollerslev, 1986; Hansen, 1994; Theodossiou, 1998; Harvey and Siddique, 1999; Wang and Fawson, 2001). Recent evidence from Chen et al. (2008) proposed a formal test on the joint normality of futures and their underlying spot returns. They documented that the null of normality was rejected for all twenty-five contacts considered in their article. Therefore, it is reasonable to incorporate higher moments like the skewness and leptokurtosis to describe the abnormal distributions of asset returns.

In recent years, the research of the impacts of higher moments on the optimal hedge ratio has come to the attention of many scholars. Gilbert et al. (2006) derived and applied a partial equilibrium hedge model to allowing for skewness (but not kurtosis) in the hedger's utility function. Their research showed that skewness could be a important factor to the undiversified agents, and the overall extent of speculation could either rise or fall, depending upon whether there was a price bias in the forward market. Another relevant contribution in this area was from Harris and Shen (2006), who considered cross-hedging with currencies rather than with futures. They proved that although MV hedging was likely to reduce the out-of-sample variance of hedge portfolios, the skewness and excess kurtosis of hedging returns were likely to fall and rise respectively. This result indicates that the benefit of hedging may be overstated, because the higher moments move exactly in the opposite direction to the utility maximum. Similarly, Brooks and Kat (2003) stated that hedge funds, which showed impressive performance on mean-variance grounds, often got less desirable higher-moment values than the traditional asset classes. Brooks et al. (2007) suggested that incorporating skewness into hedging decisions would generate lower but higher-variable optimal hedge ratio than the MV method did.

Due to the limited development of estimation methods for GARCH-SK model, the dynamic hedging model with higher moments of skewness and kurtosis has only achieved slow progress in recent years. Fairly recently, Zhang et al. (2009) firstly applied the GARCH-SK model to estimating dynamic hedge ratios based on conditional higher moments. They reported that compared with static strategy, their new dynamic hedging strategy could reduce the higher-moment risk and increase investor's utility. However, their empirical research is limited due to choosing only Hong Kong's Hang Seng Index as sample, and not considering the impacts of transaction costs and investor's risk aversions on the hedging strategy either. The study of Jing (2012) is one of the very few exceptions. She reported that adding a preference for positively skewed returns to traditional

---

Sponsored by: (1) China Scholarship Council (No. 201208420726); (2) Humanities and Social Sciences Project by Ministry of Education of China (No. 12YJC790024); (3) Humanities and Social Science by the Ministry of Education of Hubei Province, (No. 2012D026).

mean-variance models might not lead to more speculative hedging. Additionally, she proved that considering the aversion to excess kurtosis would cause investors to hedge more, and the overall empirical results did not support the view that the dynamic hedging model with higher moments was superior to traditional minimum variance strategies.

Although the previous studies have successfully introduced higher-moment risks into the determination of optimal hedge ratios, there still exist three drawbacks as shown below: firstly, the existing research literature of dynamic hedging model with higher moments seldom takes full consideration of the effects of transaction costs on the optimal hedge ratio. As is known to all, the dynamic adjustments of hedge ratios will generate huge transaction costs. Thus, neglecting transaction costs will inevitably lead to a huge mistake in the hedging strategy. Secondly, under the expected utility maximization paradigm, it is quite common to research the connections between investor's risk aversion coefficients and optimal hedge ratios. But this kind of research has not been reported in the previous studies of hedging models with higher moments. Thirdly, so far most literature chose the developed financial futures market as the sample. Because the degree of skewness and kurtosis is generally less in developed security markets than in developing security markets, the existing empirical results of the effectiveness of dynamic hedging models with higher moments is far from clear. In light of this, we choose the emerging CSI 300 index futures and its underlying spot index as samples, which are famous for being skewed and fat-tailed, and probe the effectiveness of the dynamic hedging model with higher moments in these new markets. Additionally, how the transaction costs and investors' risk aversions affect optimal hedge ratios is also the focus of this research.

The article is organized as follows. The next Section describes the impacts of the introduction of higher moments on the investor's objective utility function. Section 2 outlines the hedging model with higher moments. Section 3 will briefly discuss the measurement of higher moments. Section 4 introduces how higher moments are estimated by ICA. Section 5 describes the data, which is followed by our empirical findings. Final Section summarizes and concludes our findings.

1. Effects of skewness and kurtosis on investor's utility function

Let R_s and R_f denote the logarithmic changes of the spot and futures prices. Then, the returns on the hedged portfolio with the hedge ratio of h can be expressed as $R_p = R_s - hR_f$. To introduce the moments higher than second into the expected utility function, a straightforward technique is to take a Taylor series expansion of the utility function evaluated. Ignoring terms associated with moments higher than the fourth moment, the utility function can be rewritten by the Taylor series formula as:

$$E_{t-1}\left[U\left(R_p\right)\right]=\sum_{n=0}^{4}\frac{U^{(n)}\left(E_{t-1}\left(R_{p,t}\right)\right)}{n!}\times \\ \times E_{t-1}\left[\left(R_{p,t}-E_{t-1}\left(R_{p,t}\right)\right)^n\right], \tag{1}$$

where $U^{(n)}$ is the nth derivative of the utility function with respect to $E_{t-1}(R_{p,t})$, which denotes the expected return on the hedge portfolio at time $t-1$. Plugging the σ_p^2, s_p^3, and k_p^4, which represent the variance, third and fourth central moments of the random variable $R_{p,t}$ respectively, into Equation (1), the expected utility function of hedge portfolio can be transformed to:

$$E_{t-1}\left[U\left(R_{p,t}\right)\right]=U\left(\mu_{p,t}\right)+\frac{U^{(2)}\left(\mu_{p,t}\right)}{2!}\sigma_{p,t}^2+ \\ +\frac{U^{(3)}\left(\mu_{p,t}\right)}{3!}s_{p,t}^3+\frac{U^{(4)}\left(\mu_{p,t}\right)}{4!}k_{p,t}^4. \tag{2}$$

2. Hedging model with higher moments

Therefore, to estimate the utility maximizing optimal hedge ratio with higher moments can be transferred into an optimization problem as below:

$$\max_h U\left(u_{p,t},\sigma_{p,t}^2,s_{p,t}^3,k_{p,t}^4\right) \\ \begin{cases} u_{p,t}=W_t^T R_t \\ \sigma_{p,t}^2=W_t^T H_t W_t \\ s_{p,t}^3=W_t^T S_t W_t\otimes W_t \\ k_{p,t}^4=W_t^T K_t W_t\otimes W_t\otimes W_t, \end{cases} \tag{3}$$

where $R_t=\left[R_{s,t},R_{f,t}\right]^T$ denotes the vector of the spot and futures returns, $W_t=\left[1,-h\right]^T$ is the vector of spot and futures positions, and $u_t=E_{t-1}\left(R_t\right)$ represents the expectation of R_t. The time-varying matrices of co-variance H_t, co-skewness S_t, and co-kurtosis K_t of R_t can be expressed respectively as:

$$H_t=E_{t-1}\left[\left(R_t-u_t\right)\left(R_t-u_t\right)^T\right]=\begin{pmatrix} h_{ss} & h_{sf} \\ h_{sf} & h_{ff}\end{pmatrix} \\ S_t=E_{t-1}\left[\left(R_t-u_t\right)\left(R_t-u_t\right)^T\otimes\left(R_t-u_t\right)^T\right] \\ K_t=E_{t-1}\left[\left(R_t-u_t\right)\left(R_t-u_t\right)^T\otimes\left(R_t-u_t\right)^T \\ \otimes\left(R_t-u_t\right)^T\right], \tag{4}$$

where $\otimes$ denotes the Kronecker-product. According to Equation (4), the investor's objective utility function is a univariate function of independent variable h. Therefore, the optimal hedge ratio h^* is given by the first derivative of the utility function with respect to the hedge ratio equal to zero. In order to verify h^* is the unique maximizer of an objective utility function with higher moments, the first step is to consider the approximation of a utility function with a second order Taylor series expansion:

$$E_{t-1}\left[U\left(R_{P,t}\right)\right]=U\left(u_{p,t}\right)+\frac{U^{(2)}\left(u_{p,t}\right)}{2!}\sigma_{p,t}^{2}. \tag{5}$$

The first and second derivatives of the utility function with respect to h are collected as:

$$\begin{aligned}\frac{dU\left(\mu_{p,t},\sigma_{p,t}^{2}\right)}{\partial h}&=U_{1,t}\left(0,-1\right)u_{p,t}+2U_{2,t}\left(0,-1\right)H_tW_t\\&=-U_{1,t}u_f+2U_{2,t}\left(-h_{sf,t}+h_{ff}h\right)\end{aligned} \tag{6}$$

$$\frac{dU^{2}\left(\mu_{p,t},\sigma_{p,t}^{2}\right)}{\partial^{2}h}=2U_{2,t}h_{ff}.$$

Let the first derivative of the utility function in Equation (6) equals to zero, and we can obtain the optimal hedge ratio as:

$$\begin{aligned}&h^*=U_{1,t}\mu_f/U_{2,t}h_{ff}+h_{sf}/h_{ff}\\&U_{1,t}=\partial U\left(u_{p,t},\sigma_{p,t}^{2}\right)/\partial u_{p,t}\\&U_{2,t}=\partial U\left(u_{p,t},\sigma_{p,t}^{2}\right)/\partial\sigma_{p,t}^{2}.\end{aligned} \tag{7}$$

To judge the second derivative positive or negative, it is necessary to specify the exact form of a utility function. Following Gilbert et al. (2006), a constant absolute risk aversion (CARA) utility function, also known as the exponential utility function, is adopted.

$$U\left(R_p\right)=-\exp\left(-\gamma\mu_p\right), \tag{8}$$

where γ is the coefficient of risk aversion, and the higher γ, the more risk-averse the decision maker is. Because the second derivative of utility is $U_{2,t}=-\gamma^{2}/2\cdot\exp\left(-\gamma\mu\right)<0$, and variance of futures returns is $h_{ff}>0$, the second derivative of investor's objective utility $dU^{2}\left(\mu_{p,t},\sigma_{p,t}^{2}\right)/\partial^{2}h$ is always less than zero. Thus, the utility function can be graphed in the shape of a concave down parabola and reach the maximum value at the point of $h^*=U^{1}\mu_f/U^{2}h_{ff}+h_{sf}/h_{ff}$. Because the third and fourth terms in the Taylor series expansion, as shown in Equation (2), have much smaller impacts on investor's utility compared with the first and second terms, we can reasonably infer that the shape of utility function is mostly decided by the first and second terms in the Taylor series expansion. Therefore, the expected utility function consisting of higher moments is still a concave down parabola. That means by setting h^* to be the starting point and using gradient descent algorithm, we can seek out the utility maximizing hedge ratio. Furthermore, according to Equation (3), the estimations of H_t, S_t, and K_t are the preconditions for deriving the optimal hedge ratio. Therefore, the next section will give a brief introduction on how to estimate the above higher-moment matrices by the bivariate GARCH-SK model.

3. Measurement of higher-moment risks

Leon et al. (2005) proposed a univariate GARCH-SK type model to specify the time-variation in volatility, skewness, and kurtosis. In order to figure out the optimal hedge ratio based on conditional higher moments, the bivariate GARCH-SK model must be adopted to measure the impacts of conditional co-variance, co-skewness, and co-kurtosis on the optimal futures hedge ratio.

In order to figure out the optimal hedge ratio based on conditional higher moments, the bivariate GARCH-SK model must be adopted to measure the impacts of conditional co-variance, co-skewness, and co-kurtosis on the optimal futures hedge ratio, which can be defined as follows:

$$\begin{cases}Y_t=M_t+\varepsilon_t \quad \varepsilon_t\left|I_{t-1}\sim D\left(0,H_t,S_t,K_t\right)\right.\\ \eta_t=H_t^{-1/2}\varepsilon_t,\ \eta_t\left|I_{t-1}\sim D\left(0,I,S_t^*,K_t^*\right)\right.\\ vech\left(H_t\right)=B_0+\sum_{i=1}^{q_1}B_{1,i}vech\left(\eta_{t-i}\eta_{t-i}'\right)+\sum_{j=1}^{p_1}B_{2,j}vech\left(H_{t-j}\right)\\ vech\left(S_t\right)=\Gamma_0+\sum_{i=1}^{q_2}\Gamma_{1,i}vech\left(\left(\eta_{t-i}\eta_{t-i}'\right)\otimes\eta_{t-i}'\right)+\sum_{j=1}^{p_2}\Gamma_{2,j}vech\left(S_{t-j}\right)\\ vech\left(K_t\right)=\delta_0+\sum_{i=1}^{q_3}\delta_{1,i}vech\left(\left(\eta_{t-i}\eta_{t-i}'\right)\otimes\eta_{t-i}'\otimes\eta_{t-i}'\right)+\sum_{j=1}^{p_3}\delta_{2,j}vech\left(K_{t-j}\right).\end{cases} \tag{9}$$

I_{t-1} denotes all the information available in $t - 1$, $Y_t = (\gamma_{1t}, \cdots \gamma_{Nt})'$ is a $N \times 1$ vector with a mean vector of M_t. $\{\varepsilon_t\}$ is a random vector process with dimension $N \times 1$. $D(0, H_t, S_t, K_t)$ denotes an arbitrary distribution with conditional skewness and kurtosis. H_t is a conditional co-variance metric with dimension $N \times N$, S_t is a $N \times N^2$ co-skewness matrix of variables, K_t is a $N \times N^3$ co-kurtosis matrix of variables. η_t is a random vector standardized by the ε_t and $H_t^{-1/2}$, so there exist $E_{t-1}(\eta_t) = 0$ and $\mathrm{var}_{t-1}(\eta_t) = 1$. I is a unit matrix. S_t^* and K_t^* are the co-skewness and co-kurtosis matrices of η_t. B_0, Γ_0, and δ_0 are the vectors with the dimensions of $\tilde{N}^{(H)} \times 1$, $\tilde{N}^{(S)} \times 1$, and $\tilde{N}^{(K)} \times 1$ respectively. $B_{1,i}$ and $B_{2,j}$ are the square matrices with the same dimension of $\tilde{N}^{(H)} = N(N+1)/2$. $\Gamma_{1,i}$ and $\Gamma_{2,j}$ are the square matrices with the dimension of $\tilde{N}^{(S)} = N(N+1)(N+2)/6$. $\delta_{1,i}$ and $\delta_{2,j}$ are the square matrices with the dimension of $\tilde{N}^{(K)} = N(N+1)(N+2)(N+3)/24$. vech$(\cdot)$ denotes arithmetic operator, which can convert the lower triangular portion of $N \times N$ matrix into a $[N(N+1)/2] \times 1$ column vector.

In fact, the bivariate GARCH-SK model will encounter the much more serious problem of "curse of dimensionality" compared with the bivariate GARCH model. For overcoming the above difficulty, this paper will describe how to estimate first to fourth moments in the bivariate GARCH-SK model by a much easier way through ICA.

4. Simplified estimation of higher moments in the bivariate GARCH-SK

If financial time series can be regarded as a series of signals emitted by the financial system, some original source signals are assumed to have been mixed in some prescribed manner to form the observed asset returns series. Thus, ICA provides a mechanism to decompose the given signals of spot and futures returns series into statistically independent components.

Since only H_t, S_t, and K_t of the futures and spots returns need to be figured out, we can directly estimate the higher moments using ICA rather than evaluate the coefficients of the bivariate GARCH-SK model. Given $IC_t = (IC_{1t}, IC_{2t})'$ are the two independent components of returns vector Y_t, IC_{1t} and IC_{2t} are statistically independent, and there exists an invertible matrix, which can make:

$$IC_t = \psi Y_t. \tag{10}$$

Evaluating the conditional variance, skewness, and kurtosis of the both sides of Equation (10), we can get:

$$\begin{cases} \mathrm{var}(IC_t/I_{t-1}) = \mathrm{var}(\psi Y_t | I_{t-1}) = \psi\, \mathrm{var}(Y_t/I_{t-1})\psi' = \psi H \psi' \\ skew(IC_t/I_{t-1}) = skew(\psi Y_t | I_{t-1}) = \psi\, skew(Y_t/I_{t-1})(\psi' \otimes \psi') = \psi S_t (\psi' \otimes \psi') \\ kurt(IC_t/I_{t-1}) = kurt(\psi Y_t | I_{t-1}) = \psi\, kurt(Y_t/I_{t-1})(\psi' \otimes \psi' \otimes \psi') = \psi K_t (\psi' \otimes \psi' \otimes \psi'). \end{cases} \tag{11}$$

As seen in Equation (11), this formula has built up the great connection between variables of H_t, S_t, K_t and variables of $\mathrm{var}(IC_t/I_{t-1})$, $skew(IC_t/I_{t-1})$, and $kurt(IC_t/I_{t-1})$. Since the each component in ICA is independent, $\mathrm{var}(IC_t/I_{t-1})$, $skew(IC_t/I_{t-1})$, and $kurt(IC_t/I_{t-1})$ are all diagonal matrices, which means that the diagonal entries are the conditional variance, skewness, and kurtosis, and the entries outside the main diagonal are all zeros.

Therefore, as long as $IC_t = (IC_{1t}, IC_{2t})'$ are reached by Equation (10), using the univariate GARCH-SK model, we can calculate the conditional variance, skewness, and kurtosis of each IC_t and put them on the main diagonal to get diagonal matrices of $\mathrm{var}(IC_t/I_{t-1})$, $skew(IC_t/I_{t-1})$ and $kurt(IC_t/I_{t-1})$. Then, H_t, S_t and K_t can all be calculated from the ICA transfer matrix ψ and Equation (12) as shown below:

$$\begin{cases} H_t = \psi^{-1}\, \mathrm{var}(IC_t/I_{t-1})\psi'^{-1} \\ S_t = \psi^{-1} skew(IC_t/I_{t-1})(\psi' \otimes \psi')^{-1} \\ K_t = \psi^{-1} kurt(IC_t/I_{t-1})(\psi' \otimes \psi' \otimes)\psi'^{-1}. \end{cases} \tag{12}$$

5. Empirical results

5.1. Data and preliminary analysis. The price data employed in this article pertains to the China Shanghai Shenzhen 300 stock index futures, which is often abbreviated to CSI 300 index futures. The CSI 300 index futures contract started trading in the China Financial Futures Exchange (CFFEX) on April 16, 2010. Our data set consists of daily observations of the spot index and the futures prices from April 16, 2010 through January 2, 2014 (900 observations). Furthermore, we are using the closing price of the nearest contract month (which usually represents the most liquid contract) for the returns on the futures. To avoid thin markets and expiration effects, we roll over to the next nearest contract at least one week prior to the expiration of the current contract. For convenience of analysis, the daily changes of both index futures and spot are calculated by $\log(P_t) - \log(P_{t-1})$. Table 1 gives some standard Summary statistics along with the *Jarque-Bera* test normality and unit root test for the two assets.

Table 1. Descriptive statistics for log returns of CSI 300 index futures and underling spot

Series	Mean	Variance	Skewness	Kurtosis	J-B	Unit Root Test
Δf_t	-0.00042	0.00021	-0.0564	5.7269	279.014***	-40.333***
Δs_t	-0.00041	0.00020	-0.1828	4.8417	132.056***	-40.683***

Notes: *** Denotes that the null is rejected at the 1% significance level. J-B is the Jarque-Bera (1980) test for normality, and is chi-squares asymptotic with two degrees of freedom. The unit root test is the augmented Dickey-Fuller test.

As is shown in Table 1, the returns series of CSI 300 index futures and its spot have mean values less than zero, and display the classic non-normal features of peaks, fat tail, biased, etc. The *Jacque-Bera* statistic further testifies that the above two returns series do not obey normal distribution. Thus, the impacts of higher moments on the hedge ratio should be considered, whenever returns distribution with higher moments is employed. The augmented Dick-Fuller unit root test confirms that the return series of the futures and spot are both stationary.

5.2. Estimation of higher moments. Firstly, using fixed-point algorithm (see Hyvarinen, 1997), the return series of the futures and spot are discomposed into two independent components. The whole mathematic calculation process is realized by the software MATLAB, and the transfer matrix is expressed as:

$$\hat{\psi} = \begin{bmatrix} 212.3906 & -229.2530 \\ 101.7197 & -32.6115 \end{bmatrix}. \tag{13}$$

By the transfer matrix $\hat{\psi}$ and Equation (10), the two independent components ($IC1$, $IC2$) can be easily estimated as Figure 1 graphs.

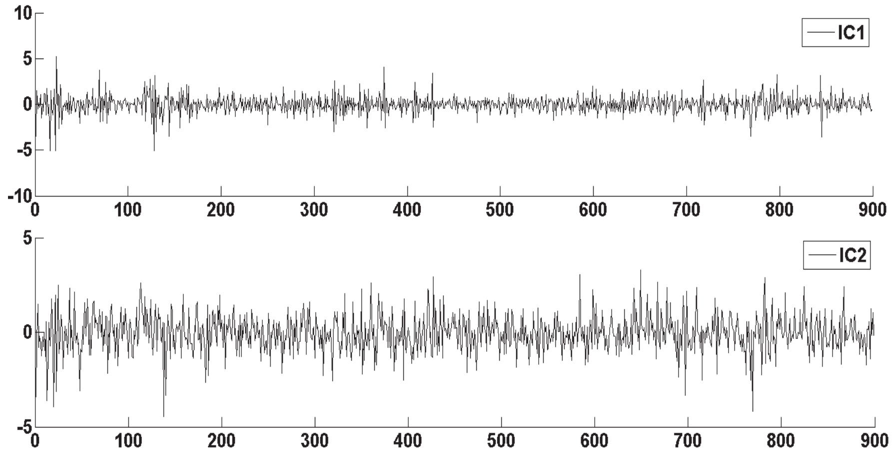

Fig. 1. Estimated independent components of the CSI 300 index futures and spot

Then, following Leon (2005), the GARCH-SK models for each independent component are formally defined as below:

$$\begin{cases} r_{i,t} = \varepsilon_{i,t}; \quad \varepsilon_{i,t} \left| I_{t-1} \sim D\left(0, \sigma_{i,t}, s_{i,t}, k_{it}\right)\right. \\ \eta_{i,t} = \sigma_{i,t}^{\frac{1}{2}} \varepsilon_{i,t}; \quad \eta_{i,t} \left| I_{t-1} \sim D\left(0, \sigma_{i,t}, s_{i,t}, k_{i,t}\right)\right. \\ \sigma_{i,t} = a_0 + a_1 \eta_{i,t-1}^2 + a_2 \sigma_{i,t-1} \\ s_{i,t} = b_0 + b_1 \eta_{i,t-1}^3 + b_2 s_{i,t-1} \\ k_{i,t} = c_0 + c_1 \eta_{i,t-1}^4 + c_2 k_{i,t-1} \\ i = 1,2 \end{cases} \tag{14}$$

Where $\sigma_{i,t}$ is the conditional variance of $r_{i,t}$, i represents the ith component, $s_{i,t}$ is the conditional skewness of $\eta_{i,t}$, and $k_{i,t}$ is the conditional kurtosis of $\eta_{i,t}$. For the GARCH-SK model, the constrains on the parameters are required to ensure that the conditional variance and kurtosis are both positive and stationary. In particular, the constrains include $a_0 \geq 0$, $0 < a_1 < 1$, $0 < a_2 < 1$, $-1 < b_1 < 1$, $-1 < b_2 < 1$, $a_1 + a_2 < 1$, $-1 < b_1 + b_2 < 1$, $c_0 \geq 0$, $0 < c_1 < 1$, $0 < c_2 < 1$, and $c_1 + c_2 < 1$.

Furthermore, the estimation of a GARCH-SK model need specify the distribution of $D(\cdot)$. Here, using the Gram-Charlier series expansion of normal density function, the density function for the standardized residuals $\eta_{i,t}$ conditioned on the information available in $t-1$ can be expressed as:

$$\begin{aligned} & f\left(\eta_{i,t} \middle| I_{t-1}\right) = \varphi(\eta_t)\psi^2(\eta_t)/\Gamma_t \\ & \varphi(\eta_t) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{\eta_t^2}{2}} \\ & \psi(\eta_t) = \left[1 + \frac{s_{i,t}}{6}\left(\eta_{i,t}^{\ 3} - 3\eta_{i,t}\right) + \frac{k_{i,t}-3}{24}\left(\eta_{i,t}^{\ 4} - 6\eta_{i,t}^{\ 2} + 3\right)\right] \\ & \Gamma_t = 1 + \frac{s_{i,t}^2}{6} + \frac{\left(k_{i,t}-3\right)^2}{24} \end{aligned} \tag{15}$$

$\varphi(\cdot)$ denotes probability density function of the standard normal distribution, and $\Psi(\cdot)$ is the fourth order of the Gram-Charlier polynomial. Then, the coefficients of the GARCH-SK model for the two independent components are computed by the method of maximum likelihood. Table 2 presented below gives the estimation results. From Table 2, the conditional higher central moments of the either independent component can be easily obtained by Equation (14).

Table 2. Estimation of the GARCH-SK model

	Parameter	IC_1	t	IC_2	t
Variance Equation	a_0	0.0745	135.9742	0.8895	42.4533
	a_1	0.1761	`78.4827	0.1454	11.6601
	a_2	0.1735	173.7716	0.3602	14.8961
Skewness Equation	b_0	-7.8281	-30.2602	-1.2741	-25.9439
	b_1	0.7966	440.5170	-0.9917	-0.0024
	b_2	-08066	-31.8981	0.0387	26.2949
Kurtosis Equation	c_0	0.3160	10.1498	0.7383	2.9624
	c_1	0.9665	301.1315	0.8074	12.5973
	c_2	0.0025	8.4385	0.0124	3.0538

Notes: Table 2 reports the estimation results of the GARCH-SK model for the continuous returns of CSI 300 futures and spot. The column t contains the t-statistics. The sample period is from April 16, 2010 to January 2, 2014, with total of 900 observations. The estimation is performed by the method of the maximum likelihood using Matlab software.

The conditional daily series of variance, skewness, and kurtosis for the two independent components are graphed respectively by Figures 2 to 4.

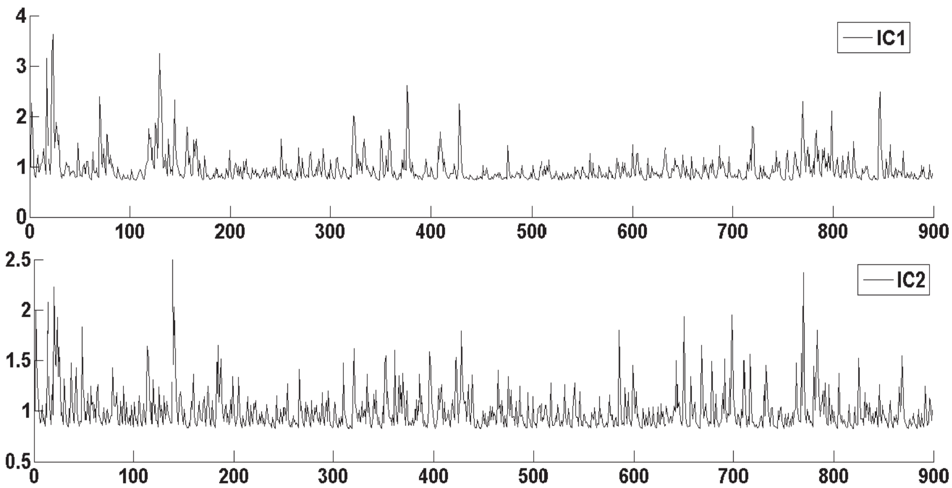

Fig. 2. Daily variance of independent components

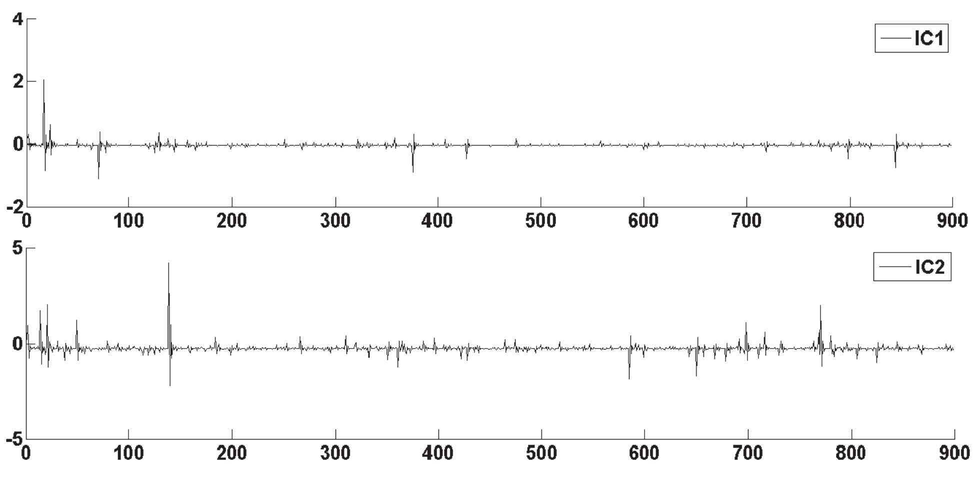

Fig. 3. Daily skewness of independent components

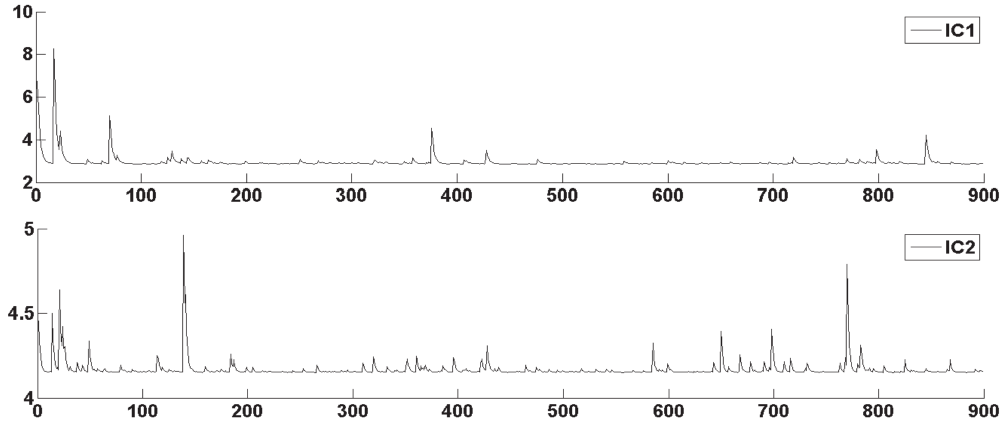

Fig. 4. Daily kurtosis of independent components

As shown in Figures 2 to 4, the conditional daily series of variance, skewness, and kurtosis for the two independent components display obvious characters of time-varying and clustering. Then, using Equation (12), H_t, S_t, and K_t of the spot and futures can be calculated respectively.

5.3. Dynamic hedge ratio and hedging effectiveness. To derive the optimal hedge ratio, it is necessary to specify the utility function in Equation (2). Here, we still adopt CARA utility as before, hence the objective function can be transformed to:

$$U\left(u_{p,t},\sigma^2_{p,t},s^3_{p,t},k^4_{p,t}\right)\approx -\exp\left(-ru_{p,t}\right)\times \times\left(1+\frac{\gamma^2}{2!}\sigma^2_{p,t}-\frac{\gamma^3}{3!}s^3_{p,t}+\frac{\gamma^4}{4!}k^4_{p,t}\right). \quad (16)$$

As shown in Equation (3), the optimization problem is to find a hedge ratio h^*, which can maximize the investor's utility function. Based on the estimated results of H_t, S_t, and K_t in the section 5.2, the $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$ for a certain hedge ratio can be directly yielded by $u_{p,t}=\left[u_{s,t},u_{f,t}\right]\left[1,-h_t\right]^T$ and Equation (4).

Prior to estimating the optimal hedge ratio, we use 0.01 as increments, and define 101 different hedge ratios on the interval [0.5, 1.5]. According to Equation (3), given H_t, S_t, and K_t of the futures and spot, any hedge ratio h can specify a set of central moments of a hedged portfolio, including $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$. Thus, 101 hedge ratios can generate 101 sets of $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$. Then, we take the average of the absolute values of $u_{p,t}$ and $s^3_{p,t}$ firstly within each set and secondly across 101 sets, and can derive the unique average values of $u_{p,t}$ and $s^3_{p,t}$. The unique average values of $\sigma^2_{p,t}$ and $k^4_{p,t}$ can be estimated almost in the same way except for taking the average of $u_{p,t}$ and $s^3_{p,t}$ directly within each set. Table 3 gives the statistics of average values for different central moments.

Table 3. Statistics of average values for different central moments

Average of absolute value of $u_{p,t}$	Average of $\sigma^2_{p,t}$	Average of absolute value of $s^3_{p,t}$	Average of $k^4_{p,t}$
0.0043	3.483e-05	1.4735e-07	4.282e-09

In addition, at any time t, observing how the $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$ change with h, we can easily find that $u_{p,t}$ and $s^3_{p,t}$ are monotonically increasing or decreasing with h, while $\sigma^2_{p,t}$ and $k^4_{p,t}$ display nonlinear relation with h in the shape of a concave up parabola. That means if investors pursue excessively high revenues and positive skewness at expense of huge risk, the optimal hedge ratio h^* will explode to positive or negative infinity. Thus, only if the increasing risk can generate enough negative utility to prevent investors from pursuing excess revenues as h changes will h^* exist. However, as shown in Table 3, the mean absolute value of $u_{p,t}$ is 124 (0.0043/3.483e-05) times larger than mean value of $\sigma^2_{p,t}$. That means if Equation (15) was directly adopted as utility function, h would eventually explode to positive or negative infinity, because utility would be entirely dominated by $u_{p,t}$. Thus, the hedge portfolio revenues need to be adjusted by multiplying a value like 1/124 = 0.008, which can realize the absolute mean value of $u_{p,t}$ equal to the mean value of $\sigma^2_{p,t}$. After this adjustment, it is sure that the expected revenues and risks in the utility function can maintain mutual balance, and h^* can also converge within a reasonable range.

$$U\left(u_{p,t},\sigma^2_{p,t},s^3_{p,t},k^4_{p,t}\right)\approx -\exp\left(-ru_{p,t}\times 0.008\right)\times \times\left(1+\frac{\gamma^2}{2!}\sigma^2_{p,t}-\frac{\gamma^3}{3!}s^3_{p,t}+\frac{\gamma^4}{4!}k^4_{p,t}\right). \tag{17}$$

Using the Genetic Algorithm package in the software Matlab, we can get a set of optimal hedge ratios from Equation (17). However, in practice, the dynamic adjustments of hedge ratios cannot avoid transaction costs. Hedge ratios will not change unless the expected utility increment of a hedge portfolio after adjustments can exceed the transaction costs. In the CSI 300 index futures market, the present transaction costs include the China Financial Futures Exchange fees of $0.25^0/_{000}$ and brokerage firm commissions of $0.75^0/_{000}$-$1^0/_{000}$. Here, we adopt the lowest total transaction costs of $1^0/_{000}$. Thus, the total cost of purchase and sell is of around $2^0/_{000}$, which means the hedged portfolio will be adjusted only the following conditions are met:

$$-\exp\left[-\gamma\left(u_{p,t}-0.0002\right)\times 0.008\right]A_t > > -\exp\left(-\gamma u_{p,t-1}\times 0.008\right)A_{t-1} \tag{18}$$
$$A_t=\left(1+\frac{\gamma^2}{2!}\sigma^2_{p,t}-\frac{\gamma^3}{3!}s^3_{p,t}+\frac{\gamma^4}{4!}k^4_{p,t}\right).$$

Furthermore, in order to investigate how risk aversion coefficients can affect the adjusting frequency of optimal hedge ratios, here we choose a wider range of risk aversion coefficients, $\gamma\in(0.5, 1, 3, 10, 20, 100)$. Then, the optimal adjusting frequencies of the optimal hedge ratios during the 900 trading days are presented in Table 4.

Table 4. Optimal adjusting frequencies of hedge ratios

	γ = 0.5	γ = 1	γ = 3	γ = 10	γ = 20	γ = 100
Adjusting times	429	439	456	482	503	601
Adjusting frequencies	47.2%	48.83%	50.72%	53.62%	55.95%	66.85%

Notes: Adjusting frequencies= Adjusting times/899.

The statistics in Table 4 support the view that the optimal adjusting frequencies will steadily go up, as the risk aversion degrees increase and transaction costs are considered, but the fluctuation range of the adjusting frequencies is narrow. For the most common risk aversion coefficient of γ = 3, the investors need to adjust the hedged portfolio every two days to obtain the maximum utilities. When transaction costs are considered, the series of optimal hedge ratios under different risk aversion coefficients are described in Figure 5.

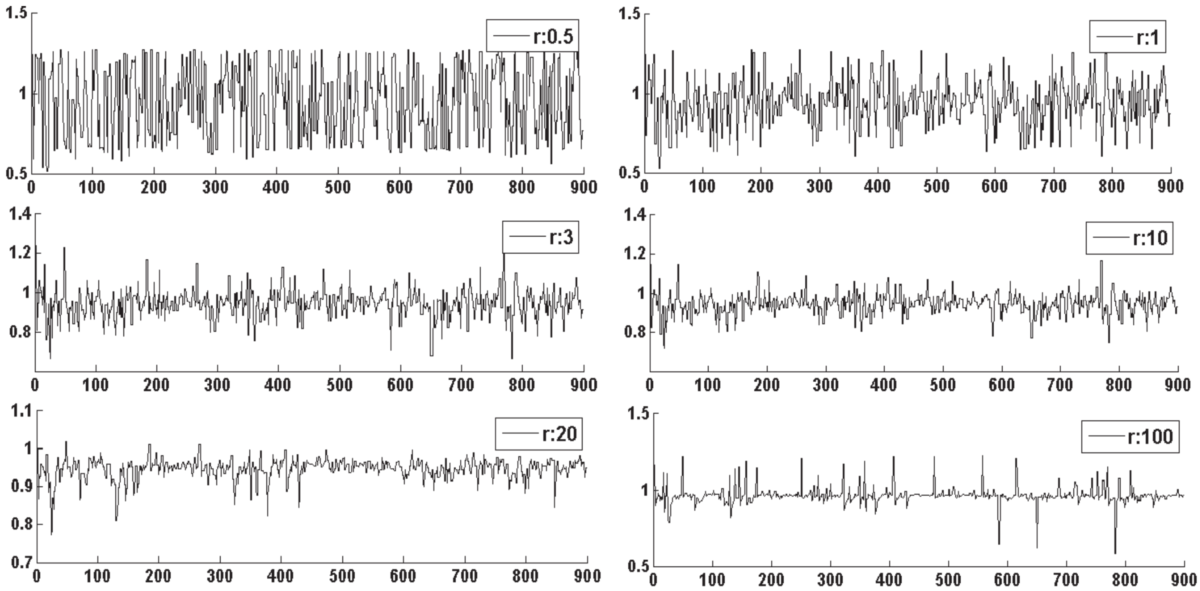

Fig. 5. Series of optimal hedge ratios under various risk aversion coefficients

As shown in Figure 5, the fluctuation of h^* will decrease as risk aversion coefficients increase. This rule can be explained well by Equation (17). Specifically, according to Equation (17), there are four factors, including $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$, that have great impacts on h^*, and γ is capable of controlling the effective of these factors on h^*. Previous analysis has proved that $u_{p,t}$ and $s^3_{p,t}$ will push the value of h^* to the extreme, while, $\sigma^2_{p,t}$ and $k^4_{p,t}$ will promote h^* to converge. In the first scenario of $\gamma < 1$, the higher-order infinitesimal $\gamma^n\left(n\in\left(2,3,4\right)\right)$ can greatly diminish the influence of $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$ on h^*. At this moment, h^*

is mainly determined by $u_{p,t}$. Therefore, the optimal hedge ratios will fluctuate more widely. In the second scenario $\gamma \geq 1$, $\gamma^n\left(n \in (2,3,4)\right)$ can amplify the impacts of $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$ on h^* from low to high. Under this condition, h^* will be mainly affected by $\sigma^2_{p,t}$, which has the largest power of convergence promotions. Therefore, the fluctuation of h^* will be small at this moment. Overall, the risk aversion degrees of investors have great impacts on the range of h^* changes. The higher γ is, the smaller the adjustment range of the dynamic hedge ratio based on the conditional higher moments will be.

To evaluate the performance of the dynamic hedging model with higher moments, we make comparisons of this model with the statistic hedge model based on minimum variance and other dynamic hedge models. The detailed procedure is designed as follows: the first step is to estimate the static optimal hedge ratio h^* by the OLS model, which is $h^* = 0.9387$; secondly, we use the VECM-GARCH model with constant conditional correlation (CCC) to estimate the dynamic hedge ratios; then, substituting different kinds of hedge ratios into Equation (3), we can figure out time-varying $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$, and $k^4_{p,t}$ under each kind of the hedge ratios; finally, Equation (17) and (18) can be used to estimate the time-varying utilities of statistic hedge ratios and dynamic optimal hedge ratios respectively. Table 5 summarizes the average utility levels and statistics of hedge returns under various hedge methods. Here, risk aversion coefficient γ is 3.

Table 5. Estimation of hedging effectiveness

Hedging methods	Utilities	Revenue	Variance	Skweness	Kurtosis
Statistic	-1.00007759	-1.92272e-05	1.8015e-05	0.1403	6.2762
VECM-GARCH	-1.00008015	-1.0420e-06	1.7745e-05	0.0512	6.5327
VECM-GARCHSK	-1.00007592*	6.6884e-04	1.9084e-05	0.4642	7.3755

Notes: Utility under static hedging strategy is estimated by Equation (16). Utility under dynamic hedging strategy using VECM-GARCH and VECM-GARCH-SK model is estimated by Equation (17) . The statistics of revenue, variance, skewness and kurtosis are all average values of time-varying $u_{p,t}$, $\sigma^2_{p,t}$, $s^3_{p,t}$ and $k^4_{p,t}$ with risk aversion coefficient $\gamma = 3$.

As shown in Table 5, the utility maximizing dynamic hedging strategy based on conditional higher moments may properly keep balance between the revenues and risks of hedge portfolio. Although this hedge strategy generates the largest venture, its performance evaluated by the statistic of revenue is much more prominent. Thus, the utility maximizing dynamic hedging strategy based on conditional higher moments brings the highest utility among the three strategies. Although the MV hedge strategies, like the OLS model and dynamic the VECM-GARCH model, can realize the smaller venture as statistics of variance and kurtosis show, they also have to endure a greater reduction of revenue for the decline of venture. Therefore, if hedging effectiveness is evaluated by the utility, which consists of both the venture and revenue, the MV hedge strategy using the VECM-GARCH model performs worst, and the statistic OLS model is next to the worst. The above founding confirms further the previous analysis of the weakness of risk-minimizing hedge strategies.

Conclusion

This study investigates how hedging behavior may change when hedgers consider higher moments of their hedged returns distribution, specifically for skewnesss and kurtosis. This article specifies the risk structure of futures and spot returns by a bivariate GARCH-SK model. Then, in view of problem of "curse of dimensionality" for the GARCH-SK model, we put forward the evaluation method of using ICA to estimate various central moments of hedged portfolio. Based on this, this article also develops a dynamic hedging model which considers transaction costs and time-variation in higher moments. Finally, we choose CSI 300 index futures and spot as samples and conduct empirical studies on the validity of the above model in the Chinese security market. Analysis shows that the unconditional distribution of CSI 300 index futures and underlying spot displays significant features of being skewed and fat-tailed, which are not normal. Thus, when describing the distribution of spot and futures returns, it is essential to consider the importance of conditional variations in moments other than mean and variance; once transaction cost is considered, the optimal adjusting frequencies will steadily go up with an increase in the degree of risk aversion coefficients, but fluctuate in a small size. At last, the dynamic utility maximizing hedging strategy can increase the performance of the hedges in terms of utility maximization, since it does the best to keep balance between the risks and revenue among all the three proposed hedge strategies.

References

1. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedastiticity, *Journal of Econometrics*, 33, pp. 307-327.
2. Brooks, C., Cerny, A. and Miffre, J. (2012). Optimal hedging with higher moments, *Journal of Futures Markets*, 32, pp. 909-944.

3. Brooks, C. and Kat, H. (2002). The statistical Properties of Hedge Fund Index Returns and Their Implications for Investors, *Journal of Alternative Investments*, 5, pp. 26-44.
4. Chen, S.S., Lee, C.F. and Shrestha, K. (2008). Do the pure martingale and joint normality hypotheses hold for futures contracts? Implications for the optimal hedge ratios, *Quarterly Review of Economics and Finance*, 48, pp. 153-174.
5. Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation, *Econometrica*, 50, pp. 987-1007.
6. Gilbert, S., Johes, S.K. and Morris, G.H. (2006). The impact of skewness in the hedging decision, *Journal of Futures Markets*, 26, pp. 503-520.
7. Hansen, B.E. (1994). Autoregressive conditional density estimation, *International Economic Review*, 35, pp. 705-730.
8. Harris, R.D.F. and Shen, J. (2006). Hedging and Value at Risk, *Journal of Futures Markets*, 26, pp. 369-390.
9. Harvey, C.R. and Siddique, A. (1999). Autoregressive conditional skewness, *Journal of Financial and Quantitative Analysis*, 34, pp. 465-487.
10. Hyvarine, Juha, K. and Erkki, O. (1997). *Independent Component Analysis,* New York: Wiley, 324 p.
11. Johnson, L.L. (1960). The theory of hedging and speculation in commodity futures, *Review of Economic Studies*, 27, pp. 139-151.
12. Jing, Y.L. (2012). An empirical study of the impact of skewness and kurtosis on hedging decisions, *Quantitative Finance*, 12, pp. 1827-1837.
13. Leon, A., Rubio, G. and Serna, G. (2005). Autoregressive conditional volatility, skewness and kurtosis, *The Quarterly Review of Economics and Finance*, 45, pp. 599-618.
14. Theodossiou, P. (1998). Financial data and the skewed generalized t distribution, *Management Science*, 44, pp. 1650-1661.
15. Wang, K.L. and Fawson, C. (2001). Modeling Asian stock returns with a more general parametric GARCH specification, *Journal of Financial Studies*, 9, pp. 23-52.
16. Xue, Q.F. and Zhang, S.Y. (2007). Multivariate conditional higher moments volatility modeling, *Journal of Systems Engineering*, 22, pp. 1-8.
17. Zhang, L.B., Wang, C.F. and Fang, Z.M. (2009). Research on Dynamic Hedging Model Considering Conditional Higher-order Moments Risk, *Journal of Industrial Engineering*, 24, pp. 63-67.