Hedging Performance of Chinese Stock Index Futures: An Empirical Analysis Using Wavelet Analysis and Flexible Bivariate GARCH Approaches

Abstract

In this paper, we assess the hedging performance of the newly established CSI 300 stock index futures over some short hedging horizons. We use wavelet analysis as well as conventional models (naïve, ordinary least squares, and error-correction) to compute the constant hedge ratios. The constant conditional correlation (CCC) and dynamic conditional correlation (DCC) bivariate generalised autoregressive conditional heteroskedasticity (BGARCH) specifications are employed to calculate the time-varying hedge ratios. Overall, we find the CSI 300 stock index futures can be an effective hedging tool. Among the constant hedge ratio models, the wavelet analysis yields the best in-sample hedging performance, though its out-of-sample hedging performance is similar to other models. Comparing the time-varying ratio models, CCC BGARCH model is better in terms of in-sample hedging effectiveness while for out-of-sample hedging performance, DCC model is better with short hedging horizons and CCC model is more favourable with long hedging horizons. Finally, the question whether time-varying ratios outperform constant ratios depends on the length of hedging horizon. Short horizons favour BGARCH hedging models while long horizons favour constant hedging ratio models.

Key words: hedge ratio, hedging effectiveness, wavelet analysis, bivariate GARCH.

JEL codes: G15, G19, G32

1. INTRODUCTION AND LITERATURE REVIEW

One of the most important functions of a futures market is to facilitate the activities of hedging the value of the underlying asset against potential losses. A key question for hedge is how to determine the optimal hedge ratio, i.e. how many futures contracts should be held for each unit of asset. The calculation of the optimal hedge ratio[1] has been widely documented and discussed in of the literature. However, empirical results of the best way to obtain the optimal hedge ratio remain controversial.

One classical way in the literature to estimate the optimal hedge ratio is by regressing the spot returns against the futures returns based on the ordinary least squares (OLS) method (see e.g., Johnson, 1960; Stein, 1961; Ederington, 1979; Benet, 1992; Malliaris and Urrutia, 1991). This approach postulates that the objective of hedge is to minimise the variance of the hedged portfolio, thus the hedge ratio that minimises the portfolio variance should be the optimal hedge ratio. In addition to the OLS hedging model, there are other studies using more complex methods to generate the optimal hedge ratio that maximises the expected utility of hedgers or consider the downside risk in the hedged portfolio, such as the mean-Gini coefficient hedging (Chen, Lee, and Shrestha, 2001), the generalised semi-variance hedging (Lien and Tse, 2002), and the lower partial moment (LPM) methodology (Lien and Demirer, 2003). These studies estimate the optimal hedge ratio under the assumption that the joint distribution of the spot and futures prices is time-invariant. Constant optimal hedge ratio is not adjusted continuously on the basis of available information in the past.

However, constant hedge ratio might be inappropriate because the variance and covariance matrix of the spot and futures returns could be conditioned on the past available information. This feature of financial time-series is referred to as the autoregressive conditional

[1] Long and short hedges share the same underpinning calculation method for optimal hedge ratio.

heteroskedasticity (ARCH) (Engle, 1982) and the generalised ARCH (GARCH) (Bollerslev, 1986). Various bivariate generalised autoregressive conditional heteroskedasticity (BGARCH)-type models have been utilised to estimate time-varying hedge ratios (see, e.g., Baillie and Myers, 1991; Brooks, Henry, and Persand, 2002; Lien, Tse, and Tsui, 2002; Miffre, 2004; Yang and Allen, 2004; Cotter and Hanly, 2006; Park and Jei, 2010). Time-varying hedge ratio is distinguished from constant hedge ratio because it is conditioned on the information set available at the previous time period. The hedged portfolio constructed from time-varying hedge ratio is adjusted continuously on a regular basis to reflect information set available at the time the hedging decision is made.

Although the derivation method for time-varying hedge ratio captures the time-varying second moments of financial time series, which constant hedge ratio always ignores, there have been many controversies among various studies whether the conditional (time-varying) hedging model can outperform the unconditional (constant) hedging model. Some studies conclude that the time-varying hedge ratio generates a higher variance reduction than the conventional constant hedge ratio (e.g., Baillie and Myers, 1991; Bera, Garcia, and Roh, 1997; Park and Switzer, 1995). Other studies find that employing a time-varying hedging model does not provide any improvement for the futures hedge compared to a constant hedge model (e.g., Collins, 2000; Lien et al., 2002). Park and Jei (2010) further claim that a time-varying hedging model can make the modest improvement relative to a constant hedging model when the standard deviation of the time-varying hedge ratio is stable and low enough. However, even if such improvement exists, it cannot guarantee that a time-varying hedging model is superior to a constant hedging model. If transaction cost is considered, the benefits of the conditional hedge could be shrunken.

Another important issue on the calculation of optimal hedge ratio is the dependence of the optimal hedge ratio on the hedging horizon. In practice, individual and institution hedgers do not always have the same hedging horizon (Lien and Shrestha, 2007). Chen, Lee and Shrestha (2004) point out that the frequency of data used to estimate optimal hedge ratio must correspond to the length of hedging horizon. Any mismatch could result in incorrect optimal hedge ratio, in turn impacting hedging effectiveness. Besides, Geppert (1995) proposes objective models, which describe the underlying data generating process, to calculate the optimal hedge ratios for different hedging horizons. Some other studies drawing attention to the relationship between the optimal hedge ratio and the hedging horizon include Ederington (1979), Hill and Schneeweis (1982), Benet (1992), Lien and Tse (2000), and so on. Lien and Shrestha (2007) use wavelet analysis to compute the optimal hedge ratios for different hedging horizons. They find that wavelet analysis can efficiently improve the hedging effectiveness of the optimal hedge ratio, compared to other hedging models that calculate constant hedge ratio such as OLS or error-correction (EC) method. Wavelet analysis alleviates the sample reduction problem faced by other constant hedging models to some extent when the length of the hedging horizon is large. However, although there are a few studies addressing the relationship between the length of the hedging horizon and the optimal hedge ratio, most of them limit the issue to the constant hedge ratio[2].

The China Security Index (CSI) 300 futures market has received much attention since it was launched on April 16, 2010. Specifically, as the first stock index futures contract being traded in Chinese mainland, the CSI 300 index futures is paid close attention to by domestic investors for predicting the future spot prices and minimising risk of spot asset. However, the number of empirical studies on the performance of the CSI 300 index futures contracts to hedge risk of spot asset is quite limited until now. Although there are a few studies on the functionality of hedging by using the Chinese index futures contracts (see, e.g. Wen, Wei and Huang, 2011; Wei, Wang, and Huang, 2011), their sample size is quite small and

[2] Yang and Ellen (2004) investigate the hedging effectiveness of time-varying hedge ratio computed by a multivariate GARCH model for 4 different hedging horizons.

methodologies are limited. Besides, their empirical results on the most appropriate way to obtain the optimal hedge ratio are not conclusive. Thus, a thorough investigation on the performance of the newly established CSI 300 index futures contract as a hedging tool is needed.

In this study, based on daily CSI 300 stock index futures data, we examine the hedging performance of both the constant and time-varying hedge ratios by employing variance reduction as the performance benchmark. We focus to investigate the performance of the short-term hedging[3]. First, we use the naïve, ordinary least squares (OLS), error-correction (EC) and wavelet hedging models to estimate constant optimal hedge ratios and compare hedging performance between these constant hedge ratios. Second, two popular flexible bivariate GARCH (BGARCH) model specifications are employed to compute the time-varying hedge ratios. The BGARCH specifications are constant conditional correlation (CCC) and dynamic conditional correlation (DCC) BGARCH. We compare the hedging performance between the two BGARCH hedge ratios. Last, we compare constant and time-varying hedging models on a horizon-by-horizon basis in light of variance reduction. We also shed some light on the debate whether the time-varying hedging models can provide better hedging performance than the constant hedging models.

This study is the first comprehensive empirical investigation on the hedging effectiveness of the Chinese stock index futures and it fills a few gaps in the literature. First, wavelet analysis is used for calculating the optimal hedge ratio for the Chinese stock index futures for the first time. Literature has argued that wavelet hedging has advantages over conventional hedging when long hedging horizon is set. Although the hedging performance of conventional hedging models for the CSI 300 stock index futures has been explored by the literature, the hedging effectiveness of wavelet hedging is unknown.

[3] The maximum hedging horizon is limited to 32 days, due to the availability of data. .

Second, CCC and DCC BGARCH specifications are employed to calculate the optimal time-varying hedge ratios for the Chinese stock index futures. Previous studies limit the application of BGARCH model for hedging of CSI 300 stock index futures to the bivariate BEKK specification only (e.g. Wen et al., 2011). Bivariate Student's *t* density function is adopted for estimating coefficients of the BGARCH models. Student's *t* distribution is more advanced than bivariate normal distribution assumed by previous studies as Student's *t* density function accommodates the fat-tail behaviour of financial time series.

Furthermore, hedging performance of constant and time-varying hedge ratios is calculated and compared based on 6 hedging horizons. Specifically, we investigate the hedging performance of the CCC and DCC BGARCH hedging models for 6 hedging horizons, which has never been explored in the literature before. We also compare the hedging performance between constant and time-varying hedging models horizon by horizon to choose the best-performed hedging model. Previous studies hardly compare the hedging performance of the constant and time-varying hedging models for multiple hedging horizons. Most of their conclusions are drawn in terms of one or two horizons only. In contrast, by considering more hedging horizons, this study could yield a clearer picture whether time-varying hedging models outperform constant hedging models.

This study unveils a few interesting facts about the CSI300 index futures market. Overall, the CSI300 stock index futures market is informationally efficient for short-term hedging purpose. For constant hedging models, the wavelet hedging dominates in terms of in-sample hedging effectiveness. However, the wavelet hedging is similar to the conventional hedging model in terms of out-of-sample forecasted effectiveness.

For the time-varying BGARCH hedging models, the CCC BGARCH hedging model outperforms the DCC one in terms of the in-sample hedging effectiveness in most cases. In the out-of-sample analysis, shorter horizon favours the DCC BGARCH hedging model while longer horizon favours the CCC one.

Third, we find whether time-varying ratios outperform constant hedge ratios depends on the length of hedging horizon. Shorter hedging horizons favour the time-varying BGARCH models. In sharp contrast, longer horizons tend to favour the constant hedging models.

The remainder of the article is organised as follows. Methodology is developed in Section 2. Section 3 describes data used in the study and their descriptive statistics. Empirical results from the estimation and examination of the hedging effectiveness of various constant and time-varying hedging models are reported and discussed in Section 4. Section 5 offers some concluding remarks.

2. METHODOLOGY

The basic concept of hedging is to offset the fluctuations in the value of a spot position by using futures contracts. Consider a portfolio consisting of C_s units of a long spot position and C_f units of a short futures position. Let S_t and F_t denote the natural logarithms of spot and futures prices at the end of period t, respectively. The return on the hedged portfolio over a period, ΔV_H, is given by:

$$\Delta V_H = C_s \Delta S_t - C_f \Delta F_t. \quad (1)$$

where $\Delta S_t = S_t - S_{t-1}$ and $\Delta F_t = F_t - F_{t-1}$.

The optimal hedge ratio is derived by minimising the conditional variance of ΔV_H. Thus it is also called the minimum variance (MV) hedge ratio and is given by:

$$MV\ hedge\ ratio = \frac{C_f}{C_s} = \frac{Cov(\Delta S_t, \Delta F_t | I)}{Var(\Delta F_t | I)}. \quad (2)$$

where I is the appropriate information set (Lien and Shrestha, 2007).

2.1. The Conventional Approach

The simplest approach to obtain the optimal hedge ratio is the naïve hedge ratio where the ratio is set to 1. Another approach which is commonly used is the ordinary least square (OLS) hedge ratio. The ratio is obtained by regressing the spot return against the futures return. Specifically, the regression equation can be written is:

$$\Delta S_t = \alpha + \beta \Delta F_t + e_t. \tag{3}$$

where the MV hedge ratio is given by the estimate of β. In line with Lien and Shrestha (2007), we estimate the optimal hedge ratios for multiple hedging horizons by matching the data frequency with the hedging horizon[4]. For example, if the length of the hedging horizon is k periods[5], the optimal hedge ratio should be calculated for the k-period price change. Nonoverlapping differencing is used to calculate k-period price change for k-period hedging horizon instead of overlapping differencing[6]. Note that the use of regression with autocorrelated errors is not required for estimating the optimal hedge ratio when nonoverlapping differenced price change is used for estimation.

Equation (3) is well understood and simple to estimate. However, it does not consider a co-integrating relationship between the spot and futures prices which is implied in the Cost-of-Carry model. In this article, we follow Lien and Shrestha (2007) to use the following version of error-correction model (ECM) to take account for the cointegration between the spot and futures prices:

[4] In general it is inappropriate to use a one-period price change to estimate the optimal hedge ratio for multi-period hedging horizon unless one-period price changes are serially independent and stationary. In practice, the frequency of data used must correspond to the length of the hedging horizon. Mismatch between the hedging period and differencing period could lead to inaccurate optimal hedge ratios. See detailed discussion in Chen, Lee and Shrestha (2004).

[5] In our study, one period is equal to 1 day.

[6] Nonoverlapping differencing generates independent observations of price change, which is different from overlapping differencing where observations of price change are correlated.

$$\Delta S_t = \beta_1 \Delta F_t + a_0 + a_1 u_{t-1} + \sum_{i=1}^{p-1} a_{si} \, \Delta S_{t-i} + \sum_{j=1}^{q-1} a_{fj} \, \Delta F_{t-j} + \varepsilon_t. \quad (4)$$

where u_t is the error correction term from a cointegrating regression equation given by:

$$S_t = a + bF_t + u_t. \quad (5)$$

In Equation (4), the optimal hedge ratio can be obtained by estimating β_1. This hedge ratio is referred to as the EC hedge ratio which is distinguished from the OLS hedge ratio based on Equation (3). Similar to the OLS hedge ratio, the EC hedge ratios are estimated for different hedging horizons by matching the data frequency with the hedging horizon. Nonoverlapping differencing is used to obtain *k*-period returns for *k*-period hedging horizon. Moreover, parameter a_1 in Equation (4) is an error correction coefficient which examines whether the spot price changes respond to deviations from the long-run equilibrium between spot and futures prices in the short term.

As mentioned earlier, we estimate both OLS and EC hedge ratios by matching the frequency of data with the hedging horizon. Nonoverlapping differencing is used. This will lead to a smaller sample size for a longer horizon. Thus, the estimation of the optimal hedge ratio for a long hedging horizon is difficult when sample size is small. To alleviate this problem, we use the wavelet approach in this article and compare the results with the naïve, OLS and EC methods.

2.2. The Wavelet Approach

Wavelet analysis is recently developed signal processing technique see e.g. Percival and Mofjeld (1997), Gencay, Selcuk, and Witcher (2002), Lien and Shrestha (2007). It transforms the time series from the time domain into different layers of frequency levels and is localised in time and frequency. The basis functions underlying wavelet analysis do not oscillate indefinitely, which differs from the sine and cosine functions that serve as the basis functions underlying Fourier analysis. This allows wavelets to more parsimoniously describe functions with cusps and spikes.

Wavelet analysis involves the projection of the original series onto a sequence of basis functions, which are called wavelets. Wavelets are nonlinear functions that can be scaled and translated to form a basis for the Hilbert space of square integrable functions. Let a given integer J represent the level of resolution or the number of resolution scale. The basis functions $\Psi_{j,k}$ are obtained through scaling and translation of a mother wavelet $\Psi(t)$ as follows:

$$\Psi_{j,k}(t) = 2^{-\frac{j}{2}}\Psi\left(\frac{t-2^{j}k}{2^{j}}\right), \qquad \int \Psi(t)dt = 0, \quad j = 1, 2, \dots, J. \tag{6}$$

where $\lambda_j = 2^{j-1}$ represents scaling and k represents translation or shift. In Equation (6), the basis functions consist of three variables j, k, and time t. Thus, the basis functions are double sequences of functions, which allows us to visualise the process in a way which is not possible using other transform techniques such as Fourier analysis. Furthermore, when the wavelet is stretched or dilated by a factor of 2 its bandwidth will be halved and shifted to the left by a factor of 2. Therefore, each wavelet dilation would cover half of the spectrum. An infinite number of wavelets would be required to cover the whole spectrum. However, there are a finite number of wavelets used in practice. To solve this problem, we use a function called scaling function, $\Phi(t)$, to cover the remaining spectrum that is not covered by the wavelets. $\Phi(t)$ is a low pass filter and satisfies the following admissibility condition:

$$\int \Phi(t)dt = 1. \tag{7}$$

Given a continuous time series *x(t)*, the wavelet coefficients can be obtained by :

$$d_{j,k} = \int x(t)\Psi_{j,k}dt, \quad j = 1,2, \dots, J. \tag{8}$$

$$B_{J,k} = \int x(t)\Phi_{J,k}dt. \quad (9)$$

where

$$\Phi_{J,k}(t) = 2^{-\frac{J}{2}}\Phi\left(\frac{t-2^J k}{2^J}\right).$$

Then, the time series *x(t)* can be represented in terms of the above wavelet coefficients as follows

$$x(t) = \sum_k B_{J,k}\Phi_{J,k}(t) + \sum_k d_{J,k}\Psi_{J,k}(t) + \sum_k d_{J-1,k}\Psi_{J-1,k}(t) + \cdots + \sum_k d_{1,k}\Psi_{1,k}(t),$$

or

$$x(t) = B_J(t) + D_J(t) + D_{J-1}(t) + \cdots + D_1(t). \quad (10)$$

where

$$D_j(t) = \sum_k d_{j,k}\Psi_{j,k}(t) \quad \text{and} \quad B_J(t) = \sum_k B_{J,k}\Phi_{J,k}(t).$$

In Equation (10), the time series *x(t)* is decomposed into different series each associated with a different time scale. This process of decomposition is referred to as multiresolution analysis. In the case of discrete time series, the multiresolution analysis is performed by using discrete wavelet transform. Let the discrete time series be denoted by x_t. The multiresolution analysis for x_t is shown below:

$$x_t = B_{J,t} + D_{J,t} + D_{J-1,t} + \cdots + D_{1,t}. \quad (11)$$

Equation (11) is similar to Equation (10) where the series x_t is decomposed into different series associated with different time scales.

There are two types of discrete wavelet transforms. The first one is the discrete wavelet transform (DWT). DWT uses orthonormal transformation of the original series. Let *N* observations on the discrete time series be denoted by $x_0, \dots, x_{N-1}$, where *N* is assumed to be an integer multiple of 2^J. Under DWT, the $N \times 1$ column vector of discrete wavelet coefficients *m* can be obtained by:

$$m = Mx, \qquad x = [x_0, x_1, \dots, x_{N-1}]^T. \tag{12}$$

where the transform matrix M is an $N \times N$ real-real valued matrix satisfying $M^T M = I$. The coefficient vector m can be divided into subvectors $m_j, j = 1,2, \dots, J$ and b_J as follows:

$$m = \left[m_1^T, m_2^T, \dots, m_J^T, b_J^T\right]^T. \tag{13}$$

where m_j is a $(\frac{N}{2^j}) \times 1$ vector and b_J is a $(\frac{N}{2^J}) \times 1$ vector. m_j represents the wavelet coefficients vector that corresponds to $d_{j,k}$ and b_J represents the scaling coefficient vector that corresponds to $B_{J,k}$.

The second type of discrete wavelet transform is the maximal overlap discrete wavelet transform (MODWT). MODWT involves a highly redundant non-orthogonal transformation, which is different from DWT. MODWT leads to *J* transform coefficient vectors and each coefficient vector has length of *N*. *N* is not required to be of integer multiple of 2. Under MODWT, the coefficient vector of length *(J+1)N* is given by

$$\widetilde{m} = \widetilde{M}x. \tag{14}$$

where $\widetilde{M}$ is an *(J+1)N×N* matrix and

$$\widetilde{m} = \left[\widetilde{m}_1^T, \widetilde{m}_2^T, \dots, \widetilde{m}_J^T, \widetilde{b}_J^T\right]^T. \tag{15}$$

Each $\widetilde{m}_j$ is a wavelet coefficient vector of length *N* and $\widetilde{b}_J$ is the scaling coefficient vector of length *N*.

If the sample frequency is Δt, then $D_{j,t}$ is associated with changes on the time scale of length $2^{j-1}\Delta t$. For example, for daily data, $D_{1,t}$ is associated with daily time scale and $D_{2,t}$ is associated with 2-day time scale and so on. Note that $B_{J,t}$ is associated with the time scale the length of which equals to $2^J\Delta t$ or longer. The integer value of *J* depends on the size of sample. The maximum *J* should be the largest integer which is less than $\log_2(N)$ where *N* is the sample size.

Based on Equation (11), we can decompose the spot and futures return series into different series associated with different time scales. Thus, we have following equations:

$$\Delta S_t = B_{J,t}^S + D_{J,t}^S + D_{J-1,t}^S + \cdots + D_{1,t}^S. \tag{16}$$

$$\Delta F_t = B_{J,t}^F + D_{J,t}^F + D_{J-1,t}^F + \cdots + D_{1,t}^F. \tag{17}$$

Then we can use the *J* decompositions to estimate *J* minimum variance hedge ratios by estimating *J* regressions:

$$D_{j,t}^S = \theta_{j,0} + \theta_{j,1} D_{j,t}^F + e_{j,t}. \tag{18}$$

The minimum variance hedge ratio associated with the *j*th time scale is obtained by $\theta_{j,1}$. It should be noted that by using such wavelet decomposition we estimate each optimal hedge ratio for each hedging horizon without shrinking the sample size. Hence, wavelet approach alleviates the problem that longer hedging horizon leads to smaller sample size in the OLS and EC methods. The above decompositions based on Equations (16) and (17) can be performed by using the discrete wavelet transform (DWT) or the maximal overlap discrete wavelet transform (MODWT). Because MODWT does require the sample size to be an integer multiple of 2^J, we use MODWT as it is more flexible to operate than DWT. Furthermore, the MODWT yields an estimator of the variance of the wavelet coefficients that is statistically more efficient than the corresponding estimator based on the DWT (Percival and Mofjeld, 1997; Lien and Shrestha, 2007). Therefore, We perform multiresolution analysis with 6 levels of time scales using maximal overlap discrete wavelet transform (MODWT) and Daubechies lease asymmetric wavelet with filter length of 8 (also called la8 Daubechies wavelets)[7]. We use la8 Daubechies wavelets because it is a fairly good approximation to an ideal band-pass filter (Gencay et al., 2002; Lien and Shrestha, 2007).

This study is based on daily data. Thus, the time scales are associated with 1-day interval, 2-day interval, 4-day interval, and so on. Since we will compare the in- and out-of-sample hedging performance of wavelet hedge ratios with other constant hedge ratios such as OLS and EC hedge ratios, we calculate k-period return for k-period hedging horizon in the OLS and EC hedging models according to the levels of time scales determined in the wavelet analysis. In other words, k equals to j level of time scale. Because the sample size of price change is 572 (N=572)[8], the maximum level of time scale J is equal to 9. However, as the sample size is limited, we have to do with fact that using 9 levels of time scales in the wavelet analysis would lead to very small sample sizes for out-of-sample performance sub-periods for the OLS and EC methods. This will make the out-of-sample performance comparison highly imprecise (Lien and Shrestha, 2007). Therefore, we choose J to be 6, representing maximum 32-day time interval (32-day period's return). We believe this selection is appropriate for our comparison analysis according to the sample size. Hence, the time intervals include 1-day (j=1), 2-day (j=2), 4-day (j=3), 8-day (j=4), 16-day (j=5), and 32-day (j=6). It should be noted that we deal with hedging performance for the short-term hedging activities in this study as the maximum hedging horizon is not long (32 days).

[7] There are different types of wavelets with different length of filter. Examples include Haar, Daubechies extreme phase and Daubechies least asymmetric wavelets.
[8] Details of sample are illustrated in the third section.

2.3. The Bivariate GARCH Approach

The OLS, error-correction and wavelet hedge ratios are estimated assuming that they are time-invariant and not affected by information in the past. However, sometimes the second moments of spot and futures returns could be conditional on the past information set, which has been examined massively in the literature. As an unconditional OLS hedge ratio is estimated by minimising the variance of hedged portfolio returns, in a similar way, the conditional optimal hedge ratio, β_{t-1}, at time *t* can be written as

$$\beta_{t-1} = \frac{Cov(\Delta S_t, \Delta F_t | \Omega_{t-1})}{Var(\Delta F_t | \Omega_{t-1})} . \qquad (19)$$

where Ω_{t-1} is the available information set at time *t-1*. β_{t-1} is conditional on the information set, Ω_{t-1}. Thus, the optimal hedge ratio is time-varying. A natural and widely used model for estimating Equation (19) is a Bivariate GARCH (BGARCH) model. Since spot and futures prices is potentially cointegrated in the long run as implied in the Cost-of-Carry model, we consider a bivariate vector error correction model (B-VECM) to specify the conditional mean of the BGARCH model. In general, the B-VECM can be written as

$$R_t = \Theta + \sum_{i=1}^{p-1} \Gamma_i R_{t-i} + \Pi u_{t-1} + \varepsilon_t, \qquad (20)$$

$$\varepsilon_t | \Omega_{t-1} \sim F(0, H_t). \qquad (21)$$

where

$R_t = \begin{bmatrix} \Delta S_t \\ \Delta F_t \end{bmatrix}$ is a 2×1 vector for spot and futures returns at time *t*;

$\Theta = \begin{bmatrix} \Theta_s \\ \Theta_f \end{bmatrix}$ is a 2×1 vector for intercepts of the VECM;

$\Gamma_i = \begin{bmatrix} \Gamma_{i,s}^{s} & \Gamma_{i,f}^{s} \\ \Gamma_{i,s}^{f} & \Gamma_{i,f}^{f} \end{bmatrix}$ is a 2×2 matrix with coefficients estimating the short-run predictive power of the lagged returns for the current ones[9];

$u_{t-1} = S_{t-1} - a - bF_{t-1}$ denotes a cointegrating equation which is a transformation of Equation (5) using lagged prices;

$\Pi = \begin{bmatrix} \Pi_s \\ \Pi_f \end{bmatrix}$ is a 2×1 vector with error-correction coefficients for u_{t-1} in both spot and futures equations[10];

$\varepsilon_t = \begin{bmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,t} \end{bmatrix}$ is a 2×1 vector with the error terms of spot and futures equations. F denotes a bivariate conditional distribution, and H_t is a time-varying 2×2 positive-definite conditional covariance matrix.

In Equation (20), when the spot price exceeds the long-run relationship of the spot and futures prices at time *t-1* (i.e. $u_{t-1} > 0$), Π_s and Π_f are supposed to be negative and positive, respectively, so as to maintain the long-run relationship. Likewise, if the spot price falls below the long-run relationship (i.e. $u_{t-1} < 0$), Π_s and Π_f are expected to be negative and positive, respectively.

In this paper, we consider the constant conditional correlation (CCC) and dynamic conditional correlation (DCC) BGARCH models to calculate conditional optimal hedge ratios. Bollerslev (1990) suggest a simple version of H_t in Equation (21) such that the conditional correlation between $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$ is constant over time. Such BGARCH model is named the Constant-Conditional-Correlation (CCC) model and the conditional H_t can be written as

[9] Upper script *s* denotes spot equation while upper script *f* denotes futures equation.
[10] Spot equation is denoted by subscript *s* whereas futures equation is denoted by subscript *f*.

$$H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix} = \begin{bmatrix} \sqrt{h_{11,t}} & 0 \\ 0 & \sqrt{h_{22,t}} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{h_{11,t}} & 0 \\ 0 & \sqrt{h_{22,t}} \end{bmatrix}. \quad (22)$$

where ρ_{12} is the conditional correlation between spot and futures returns and assumed to be constant over time. The individual conditional variances $h_{11,t}$ and $h_{22,t}$ are assumed to follow a standard GARCH (1, 1) process (Bollerslev, 1986):

$$h_{ii,t} = \omega_{i0} + \omega_{i1}\varepsilon_{i,t-1}^2 + \omega_{i2}h_{ii,t-1},\ i = 1,2. \quad (23)$$

where ω_{i1} measures the effect of new shocks and ω_{i2} examines the persistence of old news. The positive definiteness of H_t is assured if $\omega_{i0} > 0, \omega_{i1} > 0, \omega_{i2} > 0,$ and $\omega_{i1} + \omega_{i2} < 1$ (i=1, 2) in Equation (23). These constrains on parameters can be immediately satisfied (Park and Jei, 2010). If ε_t is assumed to follow a bivariate normal distribution, the number of parameter to be estimated in a CCC-GARCH (1, 1) model is limited to only seven. Thus, the CCC-GARCH model is widely used in the literature as a tool to estimate time-varying hedge ratios (e.g. Lien, Tse, and Tsui, 2002). However, the constancy of the conditional correlation coefficient over time is a strong assumption that substantially influences robustness of estimation results of the CCC-GARCH model. It is purely an empirical question whether the correlation between spot and futures returns is constant over time. Therefore, we employ Bera and Kim's (2002) (hereafter referred to as BK) procedure as in Park and Jei (2010) to test the constancy of the conditional correlation assumption in a CCC-BGARCH model. The main features of their test procedure include (Tse, 2000; Park and Jei, 2010): (i) it does not depend on the functional form of the individual conditional variance equation; (ii) a studentized version of the test statistic is available when error distributions are not normal. We expect the BK test procedure to be useful in our study as we deal with more flexible BGARCH models that have individual asymmetric GARCH processes with more general error distributions.

Let u_{it} $(i = 1,2)$ denote the standardised disturbances, then we have $u_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{ii,t}}}$ $(i = 1,2)$. With the null hypothesis that the conditional correlation is constant over time, the BK test statistic is written as

$$TS_1 = \frac{\left[\sum_{t=1}^{T}(\hat{v}_{1t}^{*2}\hat{v}_{2t}^{*2} - 1 - 2\hat{\rho}^2)\right]^2}{4T(1 + 4\hat{\rho}^2 + \hat{\rho}^4)}. \tag{24}$$

where $(v_{1t}^*, v_{2t}^*)' = ((u_{1t} - \rho u_{2t})/\sqrt{1-\rho^2}, (u_{2t} - \rho u_{1t})/\sqrt{1-\rho^2})'$. T is the sample size. ρ is estimated by $\hat{\rho} = \sum_{t=1}^{T} \hat{u}_{1t}\hat{u}_{2t}/T$. The test statistic TS_1 is calculated under the assumption that the error term ε_{it} follows a standard conditional normal distribution. However, this assumption does not capture the heavy tail behaviours of financial time-series. Thus, the test statistic TS_1 is misspecified if the error distribution is not normal. This may result in the null hypothesis being incorrectly rejected. In such case, BK suggests a studentized version of the test statistic TS_1 to accommodate the non-normality of the error term. Letting $\eta_t = \hat{v}_{1t}^{*2}\hat{v}_{2t}^{*2} - 1 - 2\hat{\rho}^2$, a studentized version of Equation (24) is given as

$$TS_2 = \frac{\left[\sum_{t=1}^{T}\eta_t\right]^2}{\sum_{t=1}^{T}(\eta_t - \bar{\eta})^2}. \tag{25}$$

where $\bar{\eta}$ is the estimated mean of η_t.

If the null hypothesis is rejected, there is a time-varying conditional correlation. Thus, the CCC-BGARCH model is misspecified. To deal with the time-varying conditional correlation, we use the Dynamic-Conditional-Correlation (DCC) BGARCH model proposed by Engle (2002). The DCC-BGARCH (1, 1) model can be written as

$$H_t = D_t R_t D_t, \tag{26}$$

$$D_t = diag\{\sqrt{h_{11,t}}, \sqrt{h_{22,t}}\}, \tag{27}$$

$$h_{ii,t} = \omega'_{i0} + \omega'_{i1}\varepsilon_{i,t-1}^2 + \omega'_{i2}h_{ii,t-1},\ i = 1,2, \tag{28}$$

$$R_t = (diag\{Q_t\})^{-1/2} Q_t (diag\{Q_t\})^{-1/2}, \quad (29)$$

$$Q_t = (1 - \delta_1 - \delta_2)\bar{Q} + \delta_1 u_{t-1} u'_{t-1} + \delta_2 Q_{t-1}, \quad (30)$$

$$Q_t = \begin{bmatrix} q_{11,t} & q_{12,t} \\ q_{21,t} & q_{22,t} \end{bmatrix}. \quad (31)$$

where $u_t = (u_{1t}, u_{2t})'$ is a 2×1 vector of standardised residuals denoted by $u_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{ii,t}}}$ $(i = 1,2)$. $h_{ii,t}$ is a standard individual GARCH process. Q_t is a 2×2 symmetric matrix where $q_{11,t}$ and $q_{22,t}$ denote the conditional variances of standardised disturbances u_{1t} and u_{2t} at time *t*, respectively. $q_{12} = q_{21}$ is the conditional covariance between u_{1t} and u_{2t} at time *t*. $\bar{Q} = E[u_t u'_t]$ is a 2×2 unconditional variance-covariance matrix of u_t. δ_1 and δ_2 are scalar parameters, and $\delta_1 \geq 0, \delta_2 \geq 0$, and $\delta_1 + \delta_2 < 1$ guarantee positive definiteness of the conditional variance-covariance matrix Q_t during the optimisation. δ_1 measures the presence of the conditional correlation and δ_2 examines the persistence of the time-varying nature.

It is noteworthy that so far we only take account of past return innovations determining the variance and covariance at time *t* in the CCC- and DCC- BGARCH models. However, the effect from the positive and negative values of past return innovations on the present variance and covariance could be different. In order to examine the asymmetric effect of previous positive and negative shocks separately in the conditional variance-covariance matrix, the Glosten, Jagannathan, and Runkle (1993) (GJR) specification in the individual GARCH process (for CCC and DCC) is employed. The GJR specification of an individual GARCH process for the CCC-BGARCH model can be written as

$$h_{ii,t} = \omega_{i0} + \omega_{i1}\varepsilon_{i,t-1}^2 + \omega_{i2} h_{ii,t-1} + \omega_{i3} I_{i,t-1}\varepsilon_{i,t-1}^2, i = 1,2. \quad (32)$$

where $I_{i,t-1} = 1$ if $\varepsilon_{i,t-1} < 0$ (*i*=1,2) and 0 otherwise. When $\omega_{i3} > 0$, previous negative shocks generate higher volatility than positive ones. This asymmetric effect is called the leverage effect. Similarly, the GJR specification of an individual GARCH process for the DCC-BGARCH model can be written as

$$h_{ii,t} = \omega'_{i0} + \omega'_{i1}\varepsilon_{i,t-1}^2 + \omega'_{i2}h_{ii,t-1} + \omega'_{i3}I_{i,t-1}\varepsilon_{i,t-1}^2, i = 1,2. \quad (33)$$

where $I_{i,t-1} = 1$ if $\varepsilon_{i,t-1} < 0$ (i = 1,2) and 0 otherwise. When $\omega'_{i3} > 0$, previous negative shocks generate higher volatility than positive ones.

Most of applications of BGARCH models for estimating the optimal hedge ratio assume error terms follow a bivariate conditional normal distribution. Bollerslev and Wooldridge (1992) show consistency and asymptotic normality of the quasi-maximum likelihood estimator (QMLE) of the GARCH model. However, QMLE will lose a lot of efficiency if the underlying conditional distribution is not normal (Engle and Gonzalez-Rivera, 1991; Park and Jei, 2010). Such efficiency loss might affect the forecasting of optimal hedge ratio based on the GARCH model estimated by QMLE. A typical distributional feature the financial time-series data have is excess kurtosis. In many applications of the GARCH model it is well known that conditional normality is not enough to explain excess kurtosis in financial data (Park and Jei, 2010). With the above reasons we consider a bivariate Student's *t* distribution for the underlying disturbances of the CCC and DCC BGARCH model specifications proposed above. We believe that Student's *t* distribution is better to accommodate the excess kurtosis and is more efficient for the estimation process than normal distribution (Susmel and Engle, 1994; Tse, 1999).

Parameter estimates of the CCC and DCC BGARCH models are obtained by maximising the log-likelihood function for a bivariate Student's *t* distribution. The contribution of observation *t* to the log-likelihood is expressed in general term as

$$l_t(\Theta) = log\left\{\frac{\Gamma\left(\frac{v+2}{2}\right)v}{(v\pi)\Gamma\left(\frac{v}{2}\right)(v-2)}\right\} - \frac{1}{2}\log(|H_t|) - \frac{1}{2}(v+2)\log\left[1 + \frac{\varepsilon'_t H_t^{-1}\varepsilon_t}{v-2}\right]. \quad (34)$$

where $\Gamma(.)$ is the Gamma function. H_t is a conditional variance-covariance matrix as defined in the BGARCH models. Θ is a parameter vector of the BGARCH models. v denotes the degree of freedom for the bivariate Student's t distribution. v controls the tail behaviour of the conditional distribution and is restricted to be larger than 2. The *t*-distribution approaches to the normality as v increases.

Parameter vector Θ of the BGARCH models can be obtained by maximising the log-likelihood over the sample period, which can be expressed as

$$L(\Theta) = \sum_{t=1}^{T} l_t(\Theta). \tag{35}$$

where *T* is the sample size.

Finally, given the bivariate model of the spot and futures returns, the time-varying hedge ratio can be expressed with the variance-covariance estimates from Equations (22) and (26) for the CCC and DCC BGARCH, respectively, as

$$\hat{\beta}_{t-1} = \frac{\hat{h}_{12,t}}{\hat{h}_{22,t}} = \frac{\hat{h}_{sf,t}}{\hat{h}_{f,t}}. \tag{36}$$

2.4. The Measure of Hedging Performance

The hedging performance measure used in this paper is variance reduction which is extensively employed in the literature. Variance reduction is calculated as the ratio of the variance of return of unhedged position minus variance of return of hedge position over the variance of return of unhedged position. Denoting ΔS_t as return of the unhedged position and ΔV_H as return of the hedged position, variance reduction can be expressed as

$$VR = 1 - \frac{Var(\Delta V_H)}{Var(\Delta S_t)}. \tag{37}$$

where $Var(\Delta V_H)$ and $Var(\Delta S_t)$ denote the variances of returns of the hedged and unhedged positions, respectively. Note that *VR*=1 means the hedge is perfect.

In this study, we estimate constant hedge ratios based on naïve, OLS, EC and wavelet methods and compare hedging performance of them. Because constant hedge ratios derived from OLS and EC methods are estimated by matching data frequency with hedging horizons[11], we show the comparison results of hedging performances based on each hedging horizon. We choose hedging horizons as 1-day, 2-day, 4-day, 8-day, 16-day and 32-day period horizons which correspond to 6 levels of time scales in wavelet analysis. We will compare hedge ratios and hedging performances calculated by naïve, OLS, EC and wavelet methods for each of 6 hedging horizons. Empirical investigation of the in- and out-of-sample hedging performances of the estimated constant hedge ratios is also required. Denoting the constant hedge ratio for *k*-period hedging horizon (*k* =1, 2, 4, 8, 16, 32) as β_k in general[12], the hedged position can be expressed as

$$\Delta_k V_H = \Delta_k S_t - \beta_k \Delta_k F_t. \tag{38}$$

where $\Delta_k V_H$ denotes return of hedged portfolio for *k*-period hedging horizon. $\Delta_k S_t = S_t - S_{t-k}$ and $\Delta_k F_t = F_t - F_{t-k}$, denoting *k*-period differencing of natural logarithms of spot and futures prices for *k*-period hedging horizon, respectively. Thus, we have the hedging performance for *k*-period hedging horizon:

$$VR_k = 1 - \frac{Var(\Delta_k V_H)}{Var(\Delta_k S_t)}. \tag{39}$$

Second, the time-varying hedge ratios are estimated by the CCC and DCC BGARCH models based on daily returns. The dynamic hedged portfolio constructed from time-varying hedge ratios is adjusted daily based on the information of previous day. Thus, the estimation of

[11] That is to say, we use *k*-period differencing for a *k*-period hedging horizon based on daily data.
[12] For the naïve method, $\beta_k = 1$ for *k* = 1, 2, 4, 8, 16, 32.

time-varying hedge ratios is independent of the selection of hedging horizons. Whatever hedging horizon is chosen, hedge ratio is adjusted everyday based on the BGARCH models. Denoting the time-varying hedge ratio as β_{t-1} at date *t-1* when the hedged portfolio is constructed, the return of the hedged portfolio at date *t* can be expressed as

$$\Delta V_{H,t} = \Delta S_t - \beta_{t-1}\Delta F_t. \tag{40}$$

where $\Delta V_{H,t}$ denotes return of dynamic hedged portfolio at date *t*. The return of dynamic hedged portfolio over *k*-period hedging horizon (*k* =1, 2, 4, 8, 16, 32) is represented as

$$\Delta_k V_H = \sum_{n=1}^{k} \Delta V_{H,t+n}. \tag{41}$$

where the hedge starts at date *t*. The variance reduction of dynamic hedged portfolio for *k*-period horizon can be calculated by Equation (39).

Note that we compute both in- and out-of-sample time-varying hedge ratios and corresponding hedging effectiveness. In the out-of-sample forecasting procedure, at first, we follow a recursive procedure to calculate out-of-sample conditional variances and covariances with estimates of the BGARCH models; then we obtain out-of-sample time-varying hedge ratios based on estimated out-of-sample conditional variances and covariances.

We compare the hedging performance of the time-varying hedge ratios estimated by the CCC and DCC BGARCH models for *k*-period horizon. We also make comparisons on hedging performance between constant and time-varying hedge ratios to find out which hedging model delivers the best forecasting performance. Our results help to figure out whether time-varying hedging model outperforms constant hedging model with the use of CSI 300 index futures contracts.

3. DATA AND SAMPLE STATISTICS

China Securities Index 300 (CSI 300), designed and managed by China Securities Index Co., Ltd., was launched on April 8, 2005. CSI 300 index is comprised of 300 stocks listed in Shanghai and Shenzhen stock exchanges, accounting for approximately 70% of market capitalization of both stock exchanges. Thus it largely reflects Chinese A share market in terms of market scale, liquidity, and industry group (Wen et al., 2011). To provide investors with a tool to hedge risk in the stock market, CSI 300 stock index futures was launched on the China Financial Futures Exchange (CFFEX) on April 16, 2010. Details of the contract specifications of CSI 300 index futures are presented in Table 1.

[Insert Table 1 about here]

Daily closing prices for CSI 300 spot index and index futures are obtained from Thomson Reuters Tick History (TRTH). The sample period starts from May 17th, 2010 to September 28th, 2012. The starting date of the sample period is one month after the inception of CSI 300 index futures trading. Data of the first month are excluded for the sample as it takes some time for the newly established index futures market to stabilise.

We construct a continuous time series for futures prices by using prices of the nearby futures contract. We specify the nearby futures contract as the contract with the nearest active trading delivery month (contract month) to the day of trading. From Table 1, there are four CSI 300 index futures contracts being traded simultaneously. Thus, the nearby contract is the contract that expires in the current calendar month. Prices for the nearby futures contract are used until the contract reaches ten working days before its expiration date. Then, prices of the next nearby contract are used. Prices within ten working days around the expiration are abandoned to avoid any unexpected price fluctuation during the delivery event. After matching prices of spot index with prices of index futures, we end up with 573 price observations for each time series.

[Insert Table 2 about here]

The spot and futures daily returns, $R_{s,t}$ and $R_{f,t}$, are calculated by $S_t - S_{t-1}$ and $F_t - F_{t-1}$, respectively, where S_t and F_t are the natural logarithms of the spot and futures prices at date *t*.

To test where these series are stationary, we perform the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests including a constant and linear trend in the regression equation for S_t, F_t, $R_{s,t}$ and $R_{f,t}$. The results for these two tests are presented in Table 2. According to ADF and PP tests, both spot and futures daily prices, are non-stationary. In contrast, for spot and futures daily returns, the null hypotheses of ADF and PP tests are rejected at the 1% level. Thus, spot and futures return series are stationary.

[Insert Table 3 about here]

The results of Johansen cointegration test are reported in Table 3. The Johansen test statistics, including the trace test statistic and the maximum eigenvalue test statistic, reject the null hypothesis of no cointegrating equation between CSI 300 spot and futures prices at the 5% level. However, both test statistics fail to reject the null of one cointegrating equation between spot and futures price series. These results indicate that CSI 300 spot and futures prices are cointegrated with each other. The cointegration equation will be incorporated in an error-correction model to calculate the EC hedge ratio. It will be also considered in a Bivariate-VECM as the conditional mean equation of the BGARCH specifications for the conditional variance-covariance structure of the spot and futures returns.

[Insert Table 4 about here]

The summary statistics for daily returns of CSI 300 spot and futures are presented in Table 4. The mean and standard deviation of two return series are quite similar. Both spot and futures return series are skewed to the left as indicated by positive skewness coefficients (1.3570 and 1.2875, respectively). The sample kurtosis indicates that both returns have excess kurtosis (>3). Both of the Jarque-Bera (JB) statistics are significant at the 1% level, strongly rejecting the null of normality for spot and futures daily returns. This indicates that both return series are highly non-normal. Highly non-normal behaviours of the spot and futures returns might be accommodated by the conditional Student's *t* density function for the BGARCH models. Moreover, the Ljung-Box (LB) test statistics indicate that there is serial correlation in the futures returns but not in the spot returns.

4. EMPIRICAL RESUTLS

In this section, we report the empirical results on constant hedge ratios as well as time-varying hedge ratios.

4.1. Constant Hedge Ratios and Hedging Performance

Table 5 presents the estimation results of the in-sample constant hedge ratios and the corresponding hedging performance for the 6 hedging horizons. First, the EC ratio exceeds the OLS and wavelet ratios for 1-day, 2-day, 16-day and 32-day hedging horizons. Second, as hedging horizon increases, all hedge ratios except the naïve ratio follow a trend that they increase at first, then decrease and increase at last. These results are different from Lien and Shrestha (2007), who find the wavelet ratio increases as hedging horizon increases.

[Insert Table 5 about here]

Now we turn to the in-sample hedging performance of constant hedge ratios in Table 5. First, the hedging performance of the naïve ratio is the worst for all horizons except 1-day horizon.

Second, the hedging effectiveness of the OLS and EC ratios is quite similar for all 6 horizons. Third, as the length of hedging horizon increases, the performance of all hedge ratios improves[13]. More importantly, the wavelet ratio dominates any other ratio in terms of variance reduction for all hedging horizons except the 1-day horizon. It is obvious that the longer the hedging horizon is, the more in-sample variance reduction the wavelet hedging provides. This in-sample result is in line with Lien and Shrestha (2007).

[Insert Table 6 about here]

To perform the out-of-sample evaluation, the whole sample is split into two subsamples. The first subsample is approximately 60% of the whole sample and is used to estimate the hedge ratios. The second subsample, which is remaining 40% of the whole sample, is used to calculate the out-of-sample hedging performance with the use of hedge ratios estimated from the first subsample. Table 6 displays empirical results of the out-of-sample performance for the naïve, OLS, EC, and wavelet hedging models. As can be seen from Table 6, first, the wavelet ratio provides the best hedging performance only for 8-day and 16-day hedging horizons. Second, only the hedging performance of the naïve hedge ratio steadily improves as the hedging horizon increases. Third, the variance reduction of the OLS ratio is still close to that of the EC ratio, but not as much as in Table 5. More importantly, we find that the wavelet hedging does not always provide the best-performing hedge ratios in out-of-sample tests relative to in-sample ones for 6 hedging horizons. As can be seen from Table 6, wavelet analysis yields similar out-of-sample variance reduction as other conventional models for most of horizons. Thus, the out-of-sample results in Table 6 are qualitatively different from in-sample ones in Table 5.

---

[13] The result may apply to the short-term hedging only. It should be noted that we deal with relatively short-term hedging in this study because of the limitation of sample size. The maximum hedging horizon is 32 days. The same comments apply below.

When the hedging horizon is large, both OLS and EC ratios are calculated from a smaller number of observations. Our results show that the wavelet hedging can effectively alleviate this problem only for in-sample tests. Although the wavelet analysis model can produce the best in-sample hedging performance for almost all 6 horizons, there is no clear evidence that the wavelet hedging has the best out-of-sample forecasting power for all the horizons. The out-of-sample result is different from Lien and Shrestha (2007) who find the wavelet hedging has better out-of-sample forecasting performance than conventional hedging models.

Furthermore, results of Table 5 and Table 6 suggest the hedging performance is reasonable and acceptable for hedgers in Chinese stock markets because most of hedging performance is over 92% regardless of what constant hedging models are undertaken. This implies that the CSI 300 index futures market can provide good function for hedgers who are willing to undertake constant hedging ratio models[14].

4.2. Time-varying Hedge Ratios and Hedging Performance

In this subsection, we report the empirical results for the CCC and DCC bivariate GARCH models.

4.2.1. Model Estimation Results

We adopt a two-stage estimation method to simplify the computational process for the conditional variance-covariance equations. In the first stage, we estimate the conditional mean equations specified as the bivariate VECM in Equation (20). In the second stage, various BGARCH models are fitted to the estimated residuals of Equation (20) using the maximum likelihood estimation (MLE) method.

Table 7 shows the estimation results of the bivariate VECM. We select the optimal truncation lag as 6 for Equation (20) based on the AIC criterion. The residual diagnostic

[14] The result may apply to hedgers who take hedging for no more than 32 days.

check confirms the non-normality for residuals of the VECM due to statistically significant Jarque-Bera test statistics. The Ljung-Box test shows that there is no autocorrelation in the residuals. However, we detect heteroskedasticity given significant Q test statistic for squared estimated residuals of the futures equation in the VECM. Therefore, the results of residual diagnosis for the VECM confirm the usage of bivariate GARCH specification and the assumption of non-normal distribution for the error terms.

[Insert Table 7 about here]

The error correction coefficient Π_f is statistically significant at the 5% level whereas Π_s is not significant at any conventional level. The results indicate that it is the futures daily return that responds to deviations from the long-run relationship between spot and futures prices. The spot market is weakly exogenous, which may lead the futures market in the long term. Additionally, Π_f is positive (0.2711) as we expect.

To determine whether the conditional correlation is time-varying, we perform the constancy of the conditional correlation test as in Bera and Kim (2002). The results of the test for spot and futures residuals are shown in Table 8.

[Insert Table 8 about here]

Table 8 shows that the BK test statistic TS_1 under normality assumption is large (68.9503) enough to reject the null hypothesis of the constancy of conditional correlations at the 1% level. However, we need to check whether the null hypothesis can be rejected if the normality assumption does not hold. To this end, we turn to examine the studentized version of TS_1,i.e., TS_2. From Table 8, it is obvious that TS_2 test statistic rejects the null hypothesis of the constancy of conditional correlation at the 1% level. Thus, it is appropriate to use the DCC specification of Engle (2002) to estimate the BGARCH model for the CSI 300 spot and futures series. However, we still need to report the empirical results of the CCC BGARCH model because although the constancy of conditional correlation is statistically rejected, it is not clear whether the DCC specification outperforms the CCC one in terms of hedging performance. Table 9 reports the second-stage estimation results for the CCC and DCC BGARCH models. The residual diagnosis indicates that there are no autocorrelation and heteroskedasticity in the estimated standardised residuals of both BGARACH models. Thus, the CCC and DCC BGARCH models are well specified.

[Insert Table 9 about here]

For the CCC-BGARCH model, ω_{12} is statistically significant at the 1% level, indicating that the persistence of old shocks is evident in the spot market. It is surprise to notice that although ω_{11} is not significant at any conventional level while ω_{13} is statistically significant at the 1% level. This means that the conditional volatility in the spot market can only be explained by the negative lagged shocks. Thus, the variation of the spot market responds to bad news only. In contrast, in the futures market, ω_{21} is statistically significant at the 5% level while ω_{23} is significant at the 10% level. The results indicate: (i) new shocks can affect the conditional volatility; (ii) asymmetry of negative shocks exists in the futures market. The lagged negative shocks can generate higher volatility than the lagged positive ones. We also find the lagged old shocks persist in the futures market given the significant estimate ω_{22}, which is similar to the spot market.

For the DCC specification, both new and old information shocks can help to explain the conditional volatilities in both spot and futures markets given significant estimates ω'_{11}, ω'_{12} , ω'_{21}, and ω'_{22}. However, there is no evidence in the spot and futures markets that the negative shocks can explain the conditional volatility asymmetrically relative to positive ones. Moreover, estimate δ_2 is statistically significant at the 1% level and the value is quite high

(0.9334), which strongly suggests the persistence of time-varying nature of the correlation between spot and futures returns.

Overall, both the CCC and DCC BGARCH models have detected GARCH effects in the spot and futures returns. Time-varying variance-covariance matrix has been reasonably explained by the CCC and DCC specifications. Note that the degree of freedom *v* of the bivariate Student's *t* distribution for error terms is statistically significant at the 1% level for both estimated BGARCH models. This means that the fat tail behaviour of the conditional distribution is well featured in the CCC and DCC BGARCH models proposed in this study. However, the hedging performance of CCC and DCC BGARCH models is still unknown and needs to be further clarified.

4.2.2. Hedging Performance of BGARCH Models

Now we turn to calculate the time-varying hedge ratios by using the estimates of the BGARCH models. In out-of-sample forecasting procedure, we use 60% of the sample to obtain the estimates of the BGARCH models; then we use these estimates to calculate hedge ratios for remaining 40% of the sample. The forecasted variance-covariance matrix is obtained through a recursive procedure. Hedging performance for 6 hedging horizons is calculated by using Equation (39).

[Insert Table 10 about here]

Table 10 summarises the descriptive statistics of the time-varying hedge ratios estimated by the CCC and DCC BGARCH models. For the in-sample estimated hedge ratios, the means of hedge ratio series for both BGARCH models are similar. The mean of CCC (0.9793) is slightly larger than that of DCC (0.9754). However, the difference of the mean of out-of-sample hedge ratio series is large between the two model specifications. The difference is about 0.0393. The out-of-sample mean obtained by the CCC BGARCH is larger than the one obtained by the DCC BGARCH.

Furthermore, we need to pay a special attention to the standard deviation of estimated time-varying hedge ratio series. In Table 10, it is clear that the in-sample hedge ratio series of the CCC BGARCH has the larger standard deviation than the DCC BGARCH. The difference is huge (about 0.0271). For the out-of-sample forecasted hedge ratio series, the standard deviations of the CCC and DCC specifications are similar. The DCC standard deviation is slightly larger than the CCC (difference is about 0.0001).

The variance reduction calculation based on Equation (39) shows that the hedging effectiveness of the BGARCH models could substantially depend on the variance of returns of dynamic hedging portfolio. Because dynamic hedging portfolio is constructed based upon the estimated conditional hedge ratios, the variability of the conditional hedge ratios could substantially affect the hedging effectiveness. Park and Jei (2010) suggest an inverse relationship between hedging performance of the time-varying hedge ratio and its volatility. When hedge ratio is more volatile, i.e., the standard deviation of the estimated hedge ratios is higher, the corresponding hedging performance measured by the variance reduction becomes worse. In our case, the DCC BGARCH model estimates the less volatile in-sample conditional hedge ratio series than the CCC. Thus, it is likely that the in-sample hedging performance of the DCC BGARCH hedging model is better than that of the CCC model. On the other hand, the standard deviations of the CCC and DCC out-of-sample hedge ratios are quite similar. Thus, we expect the out-of-sample hedging performance of the CCC and DCC specifications to be similar. The in- and out-of-sample hedging performance of the CCC and DCC BGARCH models for 6 hedging horizons is shown in Table 11.

[Insert Table 11 about here]

Panel A of Table 11 reports the in-sample hedging performance of the CCC and DCC BGARCH models. First, the CCC BGARCH hedge ratio outperforms the DCC hedge ratio for 1-day, 2-day, 4-day, and 8-day hedging horizons. Second, the DCC hedge ratio performs better than the CCC model in terms of variance reduction for 16-day hedging horizon. Third, for 32-day hedging horizon, the variance reduction of the CCC and DCC hedge ratios is quite similar. Because the CCC hedge ratio is found to outperform the DCC model in 4 hedging horizons, the result disagrees with our expectation that the CCC BGARCH in-sample hedge ratio should have behaved worse due to its larger standard deviation. This could be true as the magnitude of the standard deviation is not the unique factor determining the variance reduction. Park and Jei (2010) claim that there are other factors can determine the value of variance reduction other than the magnitude of the standard deviation such as the smoothness of the standard deviation.

Panel B of Table 11 presents the out-of-sample hedging performance of time-varying hedge ratio. On one hand, the DCC out-of-sample ratio outperforms the CCC one for 1-day, 2-day, and 4-day hedging horizons. On the other hand, the CCC out-of-sample ratio performs better than the DCC one in terms of variance reduction for 8-day, 16-day, and 32-day hedging horizons. The results violate our expectation because none of variance reduction of the CCC and DCC out-of-sample hedge ratios is similar for any of 6 hedging horizons. The DCC specification performs better for shorter horizons whereas the CCC specification performs better for relatively longer horizon.

Furthermore, Table 11 reveals that the hedging performance of the CCC and DCC BGARCH hedge ratios steadily improves along with the increase of hedging horizon. This can be observed in both in- and out-of-sample analyses. Longer the hedging horizon, better performance the BGARCH models can generate.

Overall, the in- and out-of-sample hedging performance of the CCC and DCC BGARCH models has different scenarios. For in-sample hedging performance, the CCC hedging ratio performs better than the DCC ratio for 4 out of 6 hedging horizons. The DCC hedge ratio has larger variance reduction than the CCC one only for 16-day horizon. For out-of-sample hedging performance, the DCC hedge ratio has better performance for short horizons such as 1-day, 2-day and 4-day. In contrast, the CCC hedge ratio has better performance for long horizons such as 8-day, 16-day and 32-day.

Our results do not support the finding by Park and Jei (2010) that the hedging performance of time-varying hedge ratios is negatively related with the volatility of hedge ratios. Moreover, the hedging performance of time-varying hedge ratios steadily improves as the length of hedging horizon increases. Finally, similar to constant hedging models, the hedging performance of the CCC and DCC BGARCH models is over 92%, implying that the CSI 300 index futures market functions well in hedging for hedgers who are in favour of time-varying BGARCH hedging models[15].

4.3. Comparisons between Constant and Time-varying Hedge Ratios

In this section we compare the in- and out-of-sample hedging performance between constant and time-varying hedge ratios for 6 hedging horizons.

First, we compare the in-sample hedging performance of constant and time-varying hedge ratios between the results of Table 5 and Panel A of Table 11. On one hand, the CCC BGARCH hedge ratio has the largest variance reduction for 1-day, 2-day and 4-day hedging horizons. On the other hand, the wavelet hedge ratio beats the other hedge ratios in terms of variance reduction for 8-day, 16-day, and 32-day hedging horizons. Thus, in light of in-

[15] The result may apply to hedgers who take hedging for no more than 32 days.

sample performance of relatively short-term hedging, time-varying hedging model dominates for shorter horizons while constant hedging model dominates for longer horizons.

Second, we compare the out-of-sample hedging performance between results of Table 6 and Panel B of Table 11. On one hand, the DCC BGARCH hedge ratio dominates in terms of variance reduction for 1-day and 4-day hedging horizons. For 2-day hedging horizon, the naïve, EC and DCC BGARCH hedging models have the similar hedging performance which is higher than the other hedging models. On the other hand, constant hedging model outperforms the time-varying one for the rest of hedging horizons. The naïve hedge ratio beats the other hedge ratios for 32-day hedging horizon. The wavelet hedge ratio has the best performance for 8-day and 16-day hedging horizons. The out-of-sample results are qualitatively similar to in-sample ones where shorter horizon favours the time-varying BGARCH hedging model whereas longer horizon favours the constant hedging model.

Previous studies have mixed conclusions whether the BGARCH models can result in better hedging performance than constant hedge ratios. For example, Ballie and Myers (1991) claim that the time-varying hedge ratios calculated from BGARCH models show better hedging performance in terms of variance reduction than the OLS hedge ratio. However, Lien et al. (2002) find that the CCC-BGARCH model cannot outperform the OLS hedge in variance reduction. In this study, the conclusion depends on how long hedging horizon takes. If the hedging horizon is quite short, the BGARCH hedging model may be a better choice than the naïve, OLS, EC and wavelet hedge. However, if the hedging period is longer, constant hedge ratios may be more favourable than the time-varying ones[16].

However, although we find that the BGARCH models make improvements on hedging over constant hedging models for short horizons, these benefits cannot guarantee that a BGARCH

[16] The conclusion is only available for the short-term hedging (about one month). It may change if the hedging horizon is longer (hedging activity takes more than 32 days).

hedging model is superior to a constant hedging model. Because a BGARCH hedge requires hedge ratio to be adjusted on a regular basis (daily basis in our case), one must consider transaction cost involved in the rebalancing procedure. Thus, too frequent adjustments tend to induce too many transaction costs, shrinking the benefits of the conditional hedging models with BGARCH specifications.

5. CONCLUDING REMARKS

Using daily data, this study examines the performance of the CSI 300 stock index futures as a tool for hedging the underlying spot index. The constant and time-varying hedge ratios are estimated and examined for 6 different short hedging horizons up to 32 days. Both in- and out-of-sample analyses are conducted. We use the conventional hedging models including the naïve, ordinary least square (OLS) and error-correction (EC) to compute constant optimal hedge ratios. The wavelet analysis is used to calculate the optimal hedge ratio for the CSI 300 stock index futures. Hedging effectiveness of these constant hedging models for different hedging horizons is investigated. We also explore the relationship between constant hedge ratio and the hedging horizon as well as how the hedging performance of constant hedge ratios behaves as the hedging horizon increases for the short-term hedging activities.

Time-varying hedge ratios are calculated by employing the CCC and DCC bivariate GARCH specifications with a bivariate Student's *t* density function. Descriptive statistics of estimated time-varying hedge ratios are summarised. Hedging effectiveness of the CCC and DCC BGARCH specifications is examined for 6 different hedging horizons, corresponding to the constant hedging models. This enables us to compare hedging performance of time-varying hedge ratios with constant hedge ratios on a horizon-by-horizon basis. Moreover, we also examine a potential relationship between hedging performance and the volatility of time-varying hedge ratios.

Overall, we find the CSI 300 stock index futures contracts can provide an effective hedging performance because almost all the hedging models proposed in this study can generate over 92% variance reduction.

Among the constant hedging models the wavelet hedging dominates in terms of the in-sample hedging effectiveness for 6 hedging horizons. However, in terms of the out-of-sample forecasted effectiveness, wavelet ratios do not dominate for all the horizons. The out-of-sample variance reduction of wavelet analysis is found similar to other conventional models. This implies that the wavelet hedging can effectively alleviates the sample reduction problem faced by the OLS and EC hedging only for in-sample tests. Our finding on the out-of-sample result of wavelet hedge ratio differs from previous literature.

In addition, we find constant hedge ratios follow a non-monotonic trend as the hedging horizon increases. This is not in line with previous studies. The in-sample hedging effectiveness of constant hedge ratios for the short-term hedging improves as the hedging horizon becomes longer.

For the time-varying BGARCH hedging models, the results of the in- and out-of-sample hedging effectiveness are disparate. The CCC BGARCH hedge ratio performs better than the DCC ratio for most of hedging horizons in terms of in-sample hedging effectiveness. However, this situation changes when the out-of-sample hedging performance is accounted for. The DCC ratio beats the CCC ratio for shorter horizons while the CCC ratio outperforms the DCC one for longer horizons.

Our results about the BGARCH hedging models do not show clear evidence to support the proposition of previous literature that the hedging performance of time-varying BGARCH hedge ratio is inversely related with the volatility of it. Additionally, the hedging performance of the BGARCH hedge ratios for the short-term hedging is found to steadily improve as the hedging horizon increases.

We find the choice of best-performing hedging model depends on the length of hedging horizon after comparing hedging effectiveness of constant hedging modes with that of time-varying ones on a horizon-by-horizon basis. When the hedging horizon is short, the time-varying BGARCH hedging models are more favourable. However, as the hedging horizon becomes longer, the constant hedging models are more favourable.

Moreover, it should be noted that although the BGARCH specifications can make improvements on hedging over constant hedging models for the short hedging horizon, transaction cost induced by frequent adjustment of hedged portfolio as required by the time-varying hedging models might shrink these improvements.

This research can be extended in a few ways. First, Lien and Shrestha (2007) claim that the high scale properties of the time series are not likely to be estimated very precisely in the wavelet decomposition, because they are observed few times in the sample. Thus, improvement of computational precision of wavelet analysis for high scales of time series could be left to a future study. Second, because we use daily returns to calculate time-varying hedge ratios, this requires the dynamic hedged portfolio to be adjusted on a daily basis. It is possible that, a time-varying hedging model with lower-frequency adjustment, such as 2-day or 4-day period's adjustment, might have better hedging performance than a model conducting daily adjustment. Since the wavelet analysis can decompose the original return series into several series with different levels of time scales, a future study might use the combination of wavelet decomposition and BGARCH specifications to investigate which adjustment schedule of time-varying hedging models has the best hedging performance. Last, the Student's *t* density function for the BGARCH specifications only captures the excess kurtosis of distributions of financial time series. A future study could extend this density function to account for the large skewness of financial data by incorporating a skewness parameter. Estimation results of the BGARCH models could be improved in this way.

REFERENCES

Baillie, R. T., Myers, R. J., 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. Journal of Applied Econometrics 6, 109–124.

Benet, B. A., 1992. Hedge period length and ex-ante futures hedging effectiveness: The case of foreign-exchange risk cross hedges. Journal of Futures Markets 12, 163–175.

Bera, A. K., Kim, S., 2002. Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. Journal of Empirical Finance 9, 171–195.

Bera, A., Garcia, P., Roh, J., 1997. Estimation of time-varying hedging ratios for corns and soybeans: BGARCH and random coefficient approaches. Sankhya 59, 346–368.

Bollerslev, T.,1986. A generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31,307–327.

Bollerslev, T.,1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. Review of Economics and Statistics 52, 5–59.

Bollerslev, T., Wooldridge, J., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariance. Econometric Reviews 11, 143–172.

Brooks, C., Henry, O., Persand, G., 2002. The effect of asymmetries on optimal hedge ratios. Journal of Business 75, 333–352.

Chen, S. S., Lee, C. F., & Shrestha, K., 2001. On a mean-generalized semivariance approach to determining the hedge ratio. Journal of Futures Markets 21, 581–598.

Chen, S. S., Lee, C. F., & Shrestha, K., 2004. Empirical analysis of the relationship between the hedge ratio and hedging horizon: A simultaneous estimation of the short- and long-run hedge ratios. Journal of Futures Markets 24, 359–386.

Collins, R. A., 2000. The risk management effectiveness of multivariate hedging models in the soy complex. Journal of Futures Markets 20, 189–204.

Cotter, J., Hanly, J., 2006. Reevaluating hedging performance. The Journal of Futures Markets 26(7), 677-702.

Ederington, L. H., 1979. The hedging performance of the new futures markets. The Journal of Finance 34, 157–170.

Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. Econometrica 50, 987–1008.

Engle, R., Gonzalez-Rivera, G., 1991. Semiparametric ARCH model. Journal of Business & Economic Statistics 9, 345–360.

Engle, R., Kroner, K., 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11, 122–150.

Gencay, R., Selcuk, F., Whitcher, B., 2002. An introduction to wavelets and other filtering methods in finance and economics. Academic Press, New York.

Geppert, J. M., 1995. A statistical model for the relationship between futures contract hedging effectiveness and investment horizon length. Journal of Futures Markets 15, 507–536.

Glosten, L., Jagannathan, D., Runkle, D., 1993. On the relation between expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801.

Hill, J., Schneeweis, T., 1982. The hedging effectiveness of foreign currency futures. Journal of Financial Research 5, 95–104.

Johnson, L. L., 1960. The theory of hedging and speculation in commodity futures. Review of Economic Studies 27, 139–151.

Lien, D., Shrestha, K., 2007. An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. The Journal of Futures Markets 27(2), 127-150.

Lien, D., Demirer, R., 2003. Downside risk for short and long hedgers. International Review of Economics and Finance 12, 25–44.

Lien, D., Tse, Y. K., 2000. A note on the length effect of futures hedging. Advances in Investment Analysis and Portfolio Management 7, 131–143.

Lien, D., Tse, Y. K., 2002. Some recent developments in futures hedging. Journal of Economic Surveys 16, 357–396.

Lien, D., Tse, Y. K., Tsui, A. K., 2002. Evaluating the hedging performance of constant-correlation GARCH model. Applied Financial Economics 12, 791–798.

Malliaris, A. G., Urrutia, J. L., 1991. The impact of the lengths of estimation periods and hedging horizons on the effectiveness of a hedge: Evidence from foreign currency futures. Journal of Futures Markets 3, 271–289.

Miffre, J., 2004. Conditional OLS minimum variance hedge ratios. Journal of Futures Markets 24, 945–964.

Park, S.Y., Jei, S. Y., 2010. Estimation and hedging effectiveness of time-varying hedge ratio: flexible bivariate GARCH approaches. The Journal of Futures Markets 30(1), 71-99.

Park, T. H., Switzer, L. N., 1995. Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. Journal of Futures Markets 15, 61–67.

Percival, D. B., Mofjeld, H. O., 1997. Analysis of subtidal coastal sea level fluctuations using wavelets. Journal of the American Statistical Association 92, 868–880.

Stein, J. L., 1961. The simultaneous determination of spot and futures prices. American Economic Review 51, 1012–1025.

Susmel, R., Engle, R. F., 1994. Hourly volatility spillovers between international equity markets. Journal of International Money and Finance 13, 3–25.

Tse, Y., 1999. Price discovery and volatility spillovers in the DJIA index and futures markets. Journal of Futures Markets 19(8), 911-930.

Tse, Y. K., 2000. A test of constant correlations in a multivariate GARCH model. Journal of Econometrics 98, 107–127.

Wei, Y., Wang, Y., Huang, D., 2011. A copula-multifractal volatility hedging model for CSI 300 index futures. Physica A 390, 4260-4272.

Wen, X., Wei, Y., Huang, D., 2011. Speculative market efficiency and hedging effectiveness of emerging Chinese index futures market. Journal of Transactional Management 16, 252-269.

Yang, W., Allen, D.E., 2004. Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets. Accounting and Finance 45, 301-321.

APPENDIX

Table 1. Contract Specifications of CSI 300 index futures

Underlying index	CSI300 index
Contract multiplier	RMB 300 per index point
Unit price	Index point
Minimum fluctuation	0.2 index point
Contract months	The current calendar month (spot month), the next calendar month, and the next two calendar quarterly months (March, June, September, and December)
Trading hour	9.15 a.m.-11.30 a.m. (first trading session Peking time) 1.00 p.m.-3.15p.m. (second trading session Peking time)
Trading hour on last trading day	9.15 a.m.-11.30 a.m. (first trading session Peking time) 1.00 p.m.-3.00p.m. (second trading session Peking time)
Maximum price fluctuation	±10% of settlement price on the previous trading day
Margin requirement	12% of the contract value
Last trading day	Third Friday of the contract month, postponed to the next business day if it falls on a public holiday
Delivery day	Third Friday, same as “Last trading day”
Settlement method	Cash settlement
Transaction code	IF
Exchange	China Financial Futures Exchange (CFFEX)

Table 2. Unit-Root and Stationarity Tests

	ADF		*PP*
CSI 300	*Stat.*	*Lag*	*Stat.*
S_t	-2.6280	0	-2.6259
F_t	-2.7031	0	-2.7227
$R_{s,t}$	-24.1229*	0	-24.1322*
$R_{f,t}$	-24.5832*	0	-24.5760*

Notes: *St* is the log daily price of the spot and *Ft* is the log daily price of the futures. $R_{s,t}$ is the spot daily returns and $R_{f,t}$ is the futures daily returns. ADF is the augmented Dickey-Fuller test with the null hypothesis that time series has a unit root. PP is the Phillips-Perron test with the null hypothesis that time series has a unit root. *Stat.* is the test statistic. *Lag* is the optimal lag truncation chosen by SIC criteria. For PP test bartlett kernel is used and the bandwidth is chosen by Newey-West procedures. * denotes significance at the 1% level.

Table 3. Johansen Cointegration Tests

Null Hypothesis	*CSI 300*	
	λ_{trace}	λ_{max}
$\gamma = 0$	17.8096**	16.9203**
$\gamma = 1$	0.8893	0.8893
	Cointegration Equation	
a	-0.0091	
b	1.0006	
	(0.0097)	

Notes: γ is the number of cointegration equation. The null hypothesis of Johansen cointegration test is that the number of cointegration equation is γ. λ_{trace} is Johansen trace test statistic and λ_{max} is the Johansen maximum eigenvalue test statistic. a is the mean of the cointegration equation and b is the cointegrating coefficient. Standard errors are in the parentheses. ** denotes significance at the 5% level.

Table 4. Descriptive Statistics of CSI 300 spot and futures daily returns

	CSI 300	
	$R_{s,t}$	$R_{f,t}$
Mean	-0.0003	-0.0003
Median	-0.0003	-0.0008
Maximum	0.1398	0.1361
Minimum	-0.0642	-0.0668
Std.	0.0149	0.0153
SK	1.3570	1.2875
KUR	16.8817	14.8310
JB	4768.273*	3494.015*
Q(12)	13.954	19.621***

Notes: $R_{s,t}$ denotes daily returns of CSI 300 spot index. $R_{f,t}$ denotes daily returns of CSI 300 index futures. *Std.* denotes standard deviation. *SK* and *KUR* are coefficients of skewness and kurtosis ($E[\frac{(\Delta R_t-\mu)^2}{\sigma}]$ and $E[\frac{(\Delta R_t-\mu)^4}{\sigma}]$), respectively, where μ is the mean and σ is the standard deviation.) *JB* is the Jarque and Bera test statistic for normality defined as $T[\frac{SK^2}{6}+(KUR-3)^2/24]$, which is asymptotically distributed as $\chi^2(2)$. *Q(12)* denotes the Ljung-Box test statistic at lags 12 days. *, **, *** denote significance at the 1%, 5% and 10% levels, respectively.

Table 5. In-Sample Performance of Constant Hedge Ratios

	Hedging horizon					
	1-day	*2-day*	*4-day*	*8-day*	*16-day*	*32-day*
Naïve						
hedge ratio	1	1	1	1	1	1
VR	0.9168	0.9453	0.9595	0.9687	0.9752	0.9925
OLS						
hedge ratio	0.9341	0.9628	0.9531	0.9309	0.9563	0.9806
VR	0.9214	0.9467	0.9618	0.9740	0.9773	0.9928
EC						
hedge ratio	0.9483	0.9721	0.9614	0.9031	0.9643	0.9864
VR	0.9212	0.9466	0.9617	0.9732	0.9772	0.9928
Wavelet						
hedge ratio	0.8994	0.9676	0.9806	0.9276	0.9405	0.9651
VR	0.8794	0.9475	0.9698	0.9893	0.9965	0.9932

Notes: This table reports estimated constant hedge ratios and in-sample hedging performance. Constant hedge ratios are calculated by naïve, OLS, EC and wavelet methods, respectively. *Naïve* denotes the naïve hedging method. *OLS* denotes the Ordinary Least Square estimation method. *EC* denotes the error-correction model. *Wavelet* denotes the wavelet analysis. *VR* denotes variance reduction which is calculated by Equation (39). Hedging horizon includes 1-day, 2-day, 4-day, 8-day, 16-day and 32-day periods. 6 levels of time scales are used in the wavelet analysis. The optimal lag in the error-correction model is chosen by AIC criteria. White heteroskedasticity-consistent standard errors are used in the estimation of OLS, EC and wavelet hedge ratios.

Table 6. Out-of-Sample Performance of Constant Hedge Ratios

	Hedging horizon					
	1-day	*2-day*	*4-day*	*8-day*	*16-day*	*32-day*
Naïve						
hedge ratio	1	1	1	1	1	1
VR	0.9200	0.9431	0.9593	0.9661	0.9673	0.9946
OLS						
hedge ratio	0.9254	0.9552	0.9594	0.9307	0.9512	0.9569
VR	0.9210	0.9424	0.9632	0.9715	0.9683	0.9910
EC						
hedge ratio	0.9332	0.9655	0.9644	0.9047	0.9121	0.9421
VR	0.9215	0.9430	0.9629	0.9707	0.9655	0.9889
Wavelet						
hedge ratio	0.8902	0.9508	0.9904	0.9299	0.9313	0.9456
VR	0.8995	0.9341	0.9599	0.9918	0.9923	0.9853

Notes: This table reports out-of-sample hedging performance of constant hedge ratios. 60% of the original sample is used to estimate in-sample hedge ratios and remaining 40% of sample data is used to calculate out-of-sample hedging performances. Constant hedge ratios are calculated by naïve, OLS, EC and wavelet methods, respectively. *Naïve* denotes the naïve hedging method. *OLS* denotes the Ordinary Least Square estimation method. *EC* denotes the error-correction model. *Wavelet* denotes the wavelet analysis. *VR* denotes variance reduction which is calculated by Equation (39). Hedging horizon includes 1-day, 2-day, 4-day, 8-day, 16-day and 32-day periods. 6 levels of time scales are used in the wavelet analysis. The optimal lag in the error-correction model is chosen by AIC criteria. White heteroskedasticity-consistent standard errors are used in the estimation of OLS, EC and wavelet hedge ratios.

Table 7. Estimation Results for the Bivariate-VEC Model

Parameters	ΔS_t	Parameters	ΔF_t
Θ_s	-0.0004 (-0.5862)	Θ_f	-0.0003 (-0.4912)
Π_s	0.1371 (1.0206)	Π_f	0.2711** (1.9694)
$\Gamma_{1,s}^s$	-0.4913** (-2.4112)	$\Gamma_{1,s}^f$	0.0022 (0.0106)
$\Gamma_{2,s}^s$	-0.4078*** (-1.8808)	$\Gamma_{2,s}^f$	-0.1239 (-0.5578)
$\Gamma_{3,s}^s$	-0.1137 (-0.5250)	$\Gamma_{3,s}^f$	0.1118 (0.5042)
$\Gamma_{4,s}^s$	0.1170 (0.5523)	$\Gamma_{4,s}^f$	0.3246 (1.4951)
$\Gamma_{5,s}^s$	-0.0704 (-0.3458)	$\Gamma_{5,s}^f$	0.0734 (0.3515)
$\Gamma_{6,s}^s$	-0.0443 (-0.2547)	$\Gamma_{6,s}^f$	0.0724 (0.4064)
$\Gamma_{1,f}^s$	0.4993** (2.5009)	$\Gamma_{1,f}^f$	0.0012 (0.0060)
$\Gamma_{2,f}^s$	0.3936*** (1.8443)	$\Gamma_{2,f}^f$	0.1383 (0.6324)
$\Gamma_{3,f}^s$	0.1288 (0.6082)	$\Gamma_{3,f}^f$	-0.0781 (-0.3598)
$\Gamma_{4,f}^s$	-0.1886 (-0.9102)	$\Gamma_{4,f}^f$	-0.3723*** (-1.7531)
$\Gamma_{5,f}^s$	0.1559 (0.7818)	$\Gamma_{5,f}^f$	0.0381 (0.1863)
$\Gamma_{6,f}^s$	-0.0164 (-0.0961)	$\Gamma_{6,f}^f$	-0.1028 (-0.5886)
JB	4818.625*	*JB*	3989.230*
Q(12)	7.3146	*Q(12)*	7.1730
$Q^2(12)$	1.0862	$Q^2(12)$	93.4666*

Notes: This table reports the estimation results of the bivariate-VEC model based on Equation (20). ΔS_t and ΔF_t are dependent variables of the VECM, respectively. The optimal lag is chosen based on AIC criteria. Original residuals of the VECM are used for diagnostic checks. *JB* denotes the Jarque-Bera test statistic for normality. *Q(12)* and *Q2(12)* are the Ljung-Box *Q* test statistics at lags 12 for residual series and its squares, respectively. Figures in the parenthesis are *t*-statistics. *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

Table 8. Results for the Constancy of Correlation Test

	CSI300
TS_1	68.9503*
$p-Value$	0.0000
TS_2	6.8547*
$p-Value$	0.0088

Notes: TS_1 and TS_2 are calculated by Equations (24) and (25), respectively. These test statistics follow χ^2 distribution with degree of freedom 1. * denotes significance at the 1% level.

Table 9. Estimation Results for CCC and DCC Bivariate GARCH Models

Parameters	*CCC*	*Parameters*	*DCC*
ω_{10}	$2.82\times10^{-5**}$ (2.5612)	ω'_{10}	0.0001^{*} (8.1509)
ω_{11}	-0.0016 (-0.2934)	ω'_{11}	0.0645^{*} (25.3423)
ω_{12}	0.8086^{*} (13.3108)	ω'_{12}	0.3449^{*} (7.0455)
ω_{13}	0.1028^{*} (2.6055)	ω'_{13}	0.1283 (1.5503)
ω_{20}	$9.03\times10^{-6*}$ (3.3402)	ω'_{20}	0.0001^{*} (8.3299)
ω_{21}	0.0130^{**} (2.2046)	ω'_{21}	0.1526^{*} (32.4194)
ω_{22}	0.9181^{*} (40.5328)	ω'_{22}	0.3539^{*} (7.8779)
ω_{23}	0.0420^{***} (1.7210)	ω'_{23}	0.0023 (0.0348)
ρ_{12}	0.9691^{*} (320.6255)	δ_1	-0.0065 (-0.6323)
v	5.8832^{*} (7.3459)	δ_2	0.9334^{*} (14.4799)
		v	5.2150^{*} (7.3114)
$Q_{u_1}(12)$	8.3446	$Q_{u_1}(12)$	7.8154
$Q^2_{u_1}(12)$	1.2988	$Q^2_{u_1}(12)$	1.6930
$Q_{u_2}(12)$	6.1761	$Q_{u_2}(12)$	9.7363
$Q^2_{u_2}(12)$	11.541	$Q^2_{u_2}(12)$	2.4332

Notes: This table reports estimation results for CCC and DCC BGARCH models based on Equations (22) and (26), respectively. Note that estimates of the conditional mean equations of these BGARCH models are given in Table (7). CCC, constant conditional correlation; DCC, dynamic conditional correlation; BGARCH, bivariate generalised autoregressive conditional heteroskedasticity. $Q_{u_i}(12)$ and $Q^2_{u_i}(12)$ denote the Ljung-Box *Q* test statistic at lags 12 for level and squared variables of *ui* , respectively, where *ui* (*i*=1,2) is the standardised residuals of the BGARCH models. Figures in the parenthesis are *t* statistics. *, **, and *** denotes statistical significance at the 1%, 5% and 10% levels, respectively.

Table 10.Summary Statistics of Time-varying Hedge Ratios

Panel A : in-sample hedge ratios				
	Mean	Std.	Minimum	Maximum
CCC-BGARCH	0.9793	0.0921	0.5666	1.1533
DCC-BGARCH	0.9754	0.0650	0.6343	1.1374
Panel B: out-of-sample hedge ratios				
	Mean	Std.	Minimum	Maximum
CCC-BGARCH	0.9973	0.0582	0.8264	1.1211
DCC-BGARCH	0.9580	0.0583	0.6081	1.0795

Notes: This table reports summary statistics of time-varying hedge ratios estimated by CCC and DCC BGARCH models. Panel A reports statistics of in-sample estimated hedge ratios. Panel B reports statistics of out-of-sample forecasted hedge ratios. Note that the whole sample data are used for estimating in-sample time-varying hedge ratios in in-sample analysis. For out-of-sample analysis, 60% of sample data are used to estimate the BGARCH models and remaining 40% of sample data are used to calculate out-of-sample time-varying hedge ratios. CCC-BGARCH, constant conditional correlation bivariate GARCH model; DCC-BGARCH, dynamic conditional correlation bivariate GARCH model; Std. denotes standard deviation.

Table 11. Hedging Performance of Time-varying Hedge Ratios

Panel A: in-sample comparisons of hedging performance

	Hedging horizons					
	1-day	2-day	4-day	8-day	16-day	32-day
CCC-BGARCH	0.9285	0.9499	0.9748	0.9861	0.9895	0.9928
DCC-BGARCH	0.9267	0.9498	0.9675	0.9844	0.9934	0.9928

Panel B: out-of-sample comparisons of hedging performance

	Hedging horizons					
	1-day	2-day	4-day	8-day	16-day	32-day
CCC-BGARCH	0.9308	0.9405	0.9588	0.9888	0.9871	0.9876
DCC-BGARCH	0.9322	0.9430	0.9635	0.9856	0.9858	0.9862

Notes: This table reports in- and out-of-sample hedging performance of time-varying hedge ratios estimated by CCC and DCC BGARCH models. Hedging performance is measured for 6 hedging horizons, including 1-day, 2-day, 4-day, 8-day, 16-day, and 32-day periods. The measure of hedging performance is variance reduction calculated by Equation (39). Panel A reports in-sample performance of time-varying hedge ratios. Panel B reports out-of-sample performance of time-varying hedge ratios. Note that the whole sample data are used for estimating in-sample time-varying hedge ratios in in-sample analysis. For out-of-sample analysis, 60% of sample data are used to estimate the BGARCH models and remaining 40% of sample data are used to calculate out-of-sample time-varying hedge ratios. CCC-BGARCH, constant conditional correlation bivariate GARCH model; DCC-BGARCH, dynamic conditional correlation bivariate GARCH model.