On the effects of static and autoregressive conditional higher order moments on dynamic optimal hedging

Hou, Yang and Holmes, Mark

School of Accounting, Economics and Finance Waikato Management School University of Waikato, School of Accounting, Economics and Finance Waikato Management School University of Waikato

17 October 2017

On the effects of static and autoregressive conditional higher order moments on dynamic optimal hedging

Yang (Greg) Hou
School of Accounting, Economics and Finance
Waikato Management School
University of Waikato
Private Bag 3105, Hamilton, 3240, New Zealand
Email-address: greg.hou@waikato.ac.nz
Telephone number: +64-7-8379402

and

Mark Holmes*
School of Accounting, Economics and Finance
Waikato Management School
University of Waikato
Private Bag 3105, Hamilton, 3240, New Zealand
Email-address: mark.holmes@waikato.ac.nz

* Corresponding author.

On the effects of static and autoregressive conditional higher order moments on dynamic optimal hedging

Abstract

While dynamic optimal hedging is of major interest, it remains unclear as to whether incorporating higher moments of a return distribution leads to better hedging decisions. We examine the effects of introducing a bivariate skew-Student density function with static and autoregressive conditional skewness and kurtosis on dynamic minimum-variance hedging strategies. Static higher order moments improve reductions in variance and value at risk of hedged portfolios. The inclusion of dynamics through an autoregressive component extends these improvements further. These benefits avail for short and long hedging horizons, which is highlighted in the global financial crisis. The static and conditional higher order moments enhance the notion that the size and smoothness of hedge ratios positively relate to hedging effectiveness while volatility does the reverse. Improved effectiveness can be explained given an upgrade of size and smoothness and a downgrade of volatility of hedge ratios attributed to the dynamics of higher order moments.

Keywords: dynamic optimal hedging, multivariate GARCH models, skew-Student density, conditional skewness and kurtosis, hedging effectiveness

The effects of static and autoregressive conditional higher order moments on dynamic optimal hedging

1. Introduction

A departure from the normality of financial return distributions is widely acknowledged in the literature, see for example Harvey (1995), Agarwal and Naik (2004), Brooks and Kat (2002), Harvey and Siddique (2000), and Christie-David and Chaudhry (2001), among others. Investors exhibit a preference for information on higher order moments, other than the mean and variance of asset returns (Scott and Horvarth, 1980; Brook et al., 2012). Such a preference is incorporated via pricing of systematic risk related to skewness (Kraus and Litzenberger, 1976, 1983; Harvey and Siddique, 2000)[1] and optimal asset allocation impacted by kurtosis (Kallberg and Ziemba, 1983; Jondeau and Rockinger, 2006). A consensus has almost been reached insofar as ignoring the effects of higher order moments on asset pricing and portfolio management may lead to sub-optimal investment decisions (Brook et al., 2012).

The effect of higher order moments above the variance on hedging decisions has drawn attention in the recent finance literature. The conventional determination of the optimal hedge ratio (OHR), which is the optimal amount of futures contracts to employ per unit of the cash asset to be hedged (Ederington, 1979), assumes that investors have a two-moment utility function or that the distribution of asset returns is normal. Such an assumption is criticised since return distributions are not normal and investors have additional aversions to negative skewness and positive excess kurtosis (Levy 1969; Brooks et al., 2012). Hence hedging strategies ignoring these facts might lead to sub-optimal hedging decisions (Brooks et al., 2012).

[1] Attempts on the incorporation of skewness are also made with respect to investment decisions (Hong et al., 2007), asset pricing models (Barone-Adesi et al., 2004) and risk aversion (Post et al., 2008).

There is a small group of studies that incorporate the impact of higher order moments into the optimal hedging. The focus thus far has been on minimum variance or maximum utility hedging strategies that generate static OHRs. The effects of skewness and kurtosis of returns of the hedged portfolio with volatility minimised have been taken into account by Harris and Shen (2006) and Cao et al. (2009) where the OHRs comprising the hedged portfolio are derived by a minimisation of value at risk (VaR) and conditional VaR (CVaR) of the hedged portfolio. The effect of the third moment is addressed in minimum variance hedging by Lien and Shrestha (2010) with a skew-normal distribution and Fu (2014) with a multi-objective hedging model. In addition to this, Gilbert et al. (2006) focuses on a partial equilibrium model allowing for skewness only in the hedger's utility function to determine OHRs. Brook et al. (2012) propose a more generalised utility-based framework that accounts for all the moments of return distribution. The general message from these studies is that using information based on moments higher than the variance lead to the better hedging decisions involving the use of static OHRs.

A drawback associated with the static hedging strategy that it leads to sub-optimal decisions in periods of high basis volatility. The reason is that the derivation of the static OHRs ignores the dynamics between the cash and futures returns that are conditional on the past information set. Hence one intuitively expects the OHRs to be time varying, which has been extensively modelled for the minimum-variance hedging strategies by the multivariate generalised autoregressive conditionally heteroscedastic (MGARCH) family of models (see, e.g. Kroner and Sultan, 1993; Bera et al., 1997; Brooks et al., 2002; Baillie et al., 2007; Park and Jei, 2010; Park and Kim, 2016). The derivation of the conditional OHRs substantially depends on the multivariate probability density functions (PDF) of return distributions, the logarithmic likelihood of which is maximised to obtain estimates of the MGARCH model. A question is raised as to whether the higher order moments have impacts on the decisions related to the dynamic optimal hedging since the multivariate return distributions are non-normal (Engle and Gonzalez-Rivera, 1991; Bollerslev, 1987; Baillie and Bollerslev, 1989; Susmel and Engle, 1994; Bauwens and Laurents, 2005). However, to our best knowledge, there is little in the way of studies that have systematically explored this issue[2].

In addition to this, it has been noted that higher moments above the variance can vary over time (Nelson, 1996; Harvey and Siddique, 1999, 2000). When the time-varying feature of skewness and kurtosis is modelled in an autoregressive manner, which is analogous to heteroscedasticity, significant evidence is found for a variety of univariate non-normal financial time series (Hansen, 1994; Harvey and Siddique, 1999; Jondeau and Rockinger, 2003; Brooks et al., 2005; Bali et al., 2008). Hence one would argue that neglecting the latent effects of autoregressive conditional higher order moments on the optimal hedging might result in sub-optimal hedging decisions. This issue, however, has not been tested in the literature.

This paper provides a first study into the effects of the static and autoregressive conditional third and fourth moments on the dynamic optimal hedging[3]. Data obtained for the Standard & Poor (S&P) 500 cash and futures indexes are employed.

A key contribution of this paper is that it assesses whether incorporating the bivariate skew-Student (SKST) return distributions with the static and conditional skewness and kurtosis parameters into the dynamic minimum variance hedging strategies provides higher hedging effectiveness than the same strategies under the assumption of normality. The SKST density is chosen since it can simultaneously capture asymmetry and thickness of tails of a joint distribution (Bauwens and Laurents, 2005). The dynamic hedging strategies proposed in

[2] Park and Jei (2010) assess the hedging performance of several bivariate GARCH hedging strategies under non-normality against the static ordinary least squares (OLS) hedging strategy. They find that the former can provide modest improvements over the latter.

[3] This paper focuses on a short hedge. That is, a long position is taken using futures contracts to hedge against the value of underlying spot assets.

this study are based on the estimation of the bivariate GARCH models specified as the constant-conditional-correlation (CCC), dynamic-conditional-correlation (DCC), asymmetric generalised dynamic-conditional-correlation (AGDCC) and BEKK models, which widely estimated GARCH models. The hedging effectiveness of each strategy in light of variance reduction (VR) and value at risk (VaR) reduction is computed and compared on a horizon-by-horizon basis across the densities of normal, SKST, and SKST with autoregressive conditional higher order moments distributions. This analysis is conducted for both a normal time period and the period of the global financial crisis (GFC).

A second key contribution is that this study investigates how the dynamic OHRs are affected by the static and autoregressive conditional higher order moments. On one hand, the literature has found that a larger size and smaller volatility of dynamic OHRs contribute to higher hedging effectiveness (e.g. Lien and Shrestha, 2007; Lai and Sheu, 2010; Park and Jei, 2010; Lien, 2010; Kim and Park, 2016). This study finds that taking into account the SKST conditional densities with the static and autoregressive conditional higher order moments for the dynamic hedging enhances this characteristic. Also, new evidence is provided where those densities that support that the smoothness of OHRs positively relate to hedging effectiveness[4]. This result enriches the knowledge on the relationship between hedge ratio and hedging effectiveness. It is also found that the magnitude, volatility and smoothness of the dynamic OHRs are significantly affected by the higher order moments. Both static and conditional skewness and kurtosis contribute to an upper level of magnitude and smoothness while at a lower level of volatility. This fact is firstly unveiled in the literature, serving as a rationale for the improvements brought about by the higher order moments and their time-varying feature given the relations between those properties of OHRs and hedging effectiveness.

[4] In this paper, smoothness of a time series is measured by the standard deviation of the first derivatives of this series over time trend such that the lower the standard deviation, the higher the smoothness.

The rest of the paper is organised as follows. Section 2 describes the methodology. Section 3 describes the data and reports on the results. Concluding remarks are offered in Section 4.

2. Methodology

2.1. Time-varying hedge ratios

The concept of minimum-variance hedging strategy is based on reducing the fluctuations in the value of an unhedged position by utilising futures contracts. A portfolio comprised of positions in cash asset and futures contracts is constructed with its variance minimised. Suppose a portfolio consists of C_s units of a long cash position and C_f units of a short futures position. Let S_t and F_t be the natural logarithms of cash and futures prices at the end of date t, respectively. Then the return on the hedged portfolio over one day is shown as

$$\Delta V_{H,t} = C_s \Delta S_t - C_f \Delta F_t. \quad (1)$$

where $\Delta S_t = S_t - S_{t-1}$ and $\Delta F_t = F_t - F_{t-1}$ are the cash and futures returns at date t, respectively.

The optimal hedge ratio (OHR) that minimises the variance of $\Delta V_{H,t}$ is given by

$$h^* = \frac{C_f}{C_s} = \frac{Cov(\Delta S_t, \Delta F_t \,|I)}{Var(\Delta F_t|I)}. \quad (2)$$

where $Cov(.)$ is denote the unconditional covariance and $Var(.)$ denotes the unconditional variance. I is the available information set (Lien and Shrestha, 2007; Hou and Li, 2013). h^* is referred to as the minimum-variance (hereafter referred to as MV) hedge ratio[5].

[5] The MV hedge ratio is for a short hedge. A long hedge with a portfolio consisting of C_s units of a short cash position and C_f units of a long futures position has the same hedge ratio as a short one. The ratio for a short hedge is the major focus of the paper.

It is well known that the second moment of financial data could be conditioned on the past information set, which is extensively specified in an autoregressive conditional heteroscedasticity (ARCH) or a generalised ARCH (GARCH) model (Engle, 1982; Bollerslev, 1986). The GARCH family of models explain the phenomenon of volatility clustering of financial returns where large and small observations are respectively influenced by their lagged values. Analogously, the unconditional MV hedge ratio given by Eq. (2) cannot remain static over time since it is estimated through variance and covariance of cash and futures returns. Henceforth we have the time varying optimal hedge ratio as follows

$$\beta_t = \frac{Cov(\Delta S_t, \Delta F_t | I_{t-1})}{Var(\Delta F_t | I_{t-1})}. \quad (3)$$

where I_{t-1} is the available information set up to date t-1. β_t is conditioned on the information set, I_{t-1}, which is determined by the conditional second moments of cash and futures returns. Hence the MV hedge ratio can be time-varying.

The conditional mean of the cash and future returns is specified in a bivariate vector error correction (VECM) model given the cash and future prices are potentially cointegrated. The error terms of VECM follow the distribution as below

$$\varepsilon_t | I_{t-1} \sim F(0, H_t). \quad (4)$$

where F denotes a bivariate conditional distribution. ε_t is a 2×1 vector of the error terms. H_t is a 2×2 conditional covariance matrix of innovations.

The conditional covariance matrix H_t needs to be forecasted in order to obtain β_t. This paper employs four specifications on bivariate GARCH models to predict the time varying covariance matrix. The conditional variance and covariance series are forecasted in a recursive process using the GARCH model estimates. In particular, the conditional variance, H_t is specified by four bivariate GARCH models that include the constant conditional correlation (CCC) (Bollerslev, 1990), dynamic conditional correlation (DCC) (Engle, 2002), asymmetric generalised dynamic conditional correlation (AGDCC) (Cappiello, Engle, and Sheppard, 2006), and BEKK (Engle and Kroner, 1995). The conditional optimal hedge ratios are respectively obtained for these models. The CCC, DCC and AGDCC have EGARCH variance processes that specify the conditional variance to be a non-linear function of the past shocks and lagged own values. The positivity of the conditional variances is inherently guaranteed due to the logarithmic setting. Hence no restrictions need to be imposed for estimation of parameters. Meanwhile, given the EGARCH model for the individual conditional variances, the positive definiteness of H_t is assured. An appropriate alternative may be the DCC model that models the correlation to be conditioned on the past. H_t. The AGDCC model extends the DCC by capturing asymmetry in the conditional correlation, all else being the same. A diagonal version of the AGDCC is used since it sufficiently reduces the number of parameters that convey little information and thus alleviate the computation burden of estimation process. The BEKK model specifies the conditional covariance matrix of innovations in a way different from the CCC, DCC, and AGDCC models. It can guarantee the positive definiteness of H_t with very few restrictions on the model parameters. The positivity of the conditional variances is inherently secured.

2.2. Flexible multivariate conditional distributions

This paper employs three multivariate probability density functions (PDFs) for the error terms by which four bivariate GARCH models are estimated via the maximum likelihood estimation (MLE). The innovations are henceforth assumed to respectively follow a bivariate conditional normal distribution, a bivariate conditional skewed-Student (SKST) distribution and a bivariate conditional skewed-Student (SKST) distribution with autoregressive conditional skewness and kurtosis parameters.

2.2.1. Conditional normal distribution

The logarithm of the PDF for ε_t following a bivariate conditional normal distribution is shown as

$$l_t(\Theta) = -0.5\{\log(|H_t|) + \varepsilon'_t H_t^{-1} \varepsilon_t + 2\log(2\pi)\}. \quad (5)$$

where l_t denotes the contribution of observation t to the log-likelihood. H_t is a conditional covariance matrix of ε_t, which is specified by the bivariate GARCH models. Θ is a vector of parameters of the bivariate GARCH models.

Estimates of the parameter vector Θ are obtained by maximising the following log-likelihood over the sample path

$$L(\Theta) = \sum_{t=1}^{T} l_t(\Theta). \quad (6)$$

where T is the sample size.

2.2.2. Conditional skew-Student distribution

Although the bivariate normality of conditional distribution is widely applied to the estimation of the bivariate GARCH models, sticking to it can lead to a large efficiency loss of the estimator when the underlying conditional distribution is non-normal (Engle and Gonzalez-Rivera, 1991; Park and Jei, 2010). This efficiency loss may impact the accuracy of estimation results, thus affecting the forecasting ability of the bivariate GARCH models. Indeed, the financial data series have been found to follow non-normal conditional distributions where excess kurtosis and non-zero skewness exist. Returns distributions with fat tails, corresponding to excess kurtosis, are widely observed in the literature (Bollerslev, 1987; Baillie and Bollerslev, 1989). Moreover, the distributions are often skewed and thus it is necessary to take it into account for the estimation process (Park and Jei, 2010). Therefore, a more general conditional distribution capturing both excess kurtosis and non-zero skewness is expected to escalate the predictive power of the bivariate GARCH models, thus improving the estimation accuracy of the optimal hedge ratios (Susmel and Engle, 1994; Tse, 1999; Bauwens and Laurents, 2005). It's intuitive to believe that higher accuracy would lead to higher hedging effectiveness.

We adopt the Bauwens and Laurents (2005) multivariate skew-Student (SKST) density for the standardized innovations ϵ_t, which applies the Fernandez and Steel (1998) skew filter to a multivariate Student's *t* distribution. The contribution of each observation at time *t* to the log-likelihood of a standardized bivariate SKST density can be expressed in general term as

$$l_t(\Theta) = \log(\tfrac{4}{\pi}) + \textstyle\sum_{i=1}^{2} \log(\frac{\xi_i s_i}{1+\xi_i^2}) + \log\left\{\Gamma\left(\frac{v+2}{2}\right)/(\Gamma\left(\frac{v}{2}\right)(v-2))\right\} - (1/2)\,(v+2)\log[1+(\kappa_t^T \kappa_t)/(v-2)], \quad (7)$$

and

$$\kappa_t = (\kappa_{1t}, \kappa_{2t})^T$$

$$\kappa_{it} = (s_i {\epsilon^*}_{it} + m_i)\xi_i^{-I_i}$$

$$m_i = \frac{\Gamma\left(\frac{v-1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\,\Gamma\left(\frac{v}{2}\right)}(\xi_i - \frac{1}{\xi_i})$$

$$s_i^2 = \left(\xi_i^2 + \frac{1}{\xi_i^2} - 1\right) - m_i^2$$

$$I_i = \begin{cases} 1 & \text{if } {\epsilon^*}_{it} \geq -\frac{m_i}{s_i} \\ -1 & \text{if } {\epsilon^*}_{it} < -\frac{m_i}{s_i} \end{cases}.$$

where i = 1, 2. The covariance matrix of ${\epsilon^*}_{it}$ $(i = 1,2)$ is an identity matrix. $\Gamma(.)$ is the gamma function. *v* is the degree of freedom controlling the thickness of tails of the distribution, i.e., the joint kurtosis. Statistically significance of *v* indicates the existence of excess kurtosis. v is restricted to be larger than 2 in order that the covariance matrix exists. If v goes to infinity, the Student t distribution approaches normality.

$m_i(\xi_i, v)$ and $s_i(\xi_i, v)$ are the mean and standard deviation of the non-standardized marginal skewed-t density of Fernandez and Steel (1998). ξ_i is the skewness parameter of the non-standardized marginal density where the sign of the logarithmic ξ_i indicates the sign of the skewness. When $ln\xi_i > 0\ (< 0)$, the skewness is positive (negative) and density is skewed to the right (left). And ξ_i^2 is a measure of skewness of the marginal density.

Estimates of parameter vector Θ are obtained by maximizing Eqs. (6) and (7).

2.2.3. Conditional skew-Student distribution with autoregressive conditional skewness and kurtosis

The skewness and kurtosis parameters of the univariate conditional density for financial returns are conditioned on the past. This feature has been modelled via an autoregressive process by the literature and significant empirical results have been found (see, e.g., Hansen, 1994; Harvey and Siddique, 1999, 2000; Jondeau and Rockinger, 2003; Brooks et al., 2005; Bali et al., 2008). In the spirit of Hansen (1994), Jondeau and Rockinger (2003), and Bali et al. (2008), we model the conditional skewness of marginal densities and joint kurtosis of the standardised bivariate SKST density as follows:

$$\tilde{\xi}_{i,t} = \xi_{0i} + \xi_{1i}\epsilon_{i,t-1} + \xi_{2i}\tilde{\xi}_{i,t-1}. \quad (8)$$

$$\tilde{v}_t = v_0 + v_1\epsilon_{1,t-1} + v_2\epsilon_{2,t-1} + v_3\tilde{v}_{t-1}. \quad (9)$$

where $i = 1,2$. Recall $\epsilon_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{ii,t}}}$ $(i = 1,2)$ where ε_{it} is the error term and $h_{ii,t}$ is the conditional heteroscedastic process. $\tilde{\xi}_{it}$ is the unrestricted skewness parameter and $\tilde{v}_t$ is the unrestricted kurtosis parameter. Since the definition of the SKST density requires the skewness parameter to be positive and the degrees of freedom to exceed 2, we impose the following logistic transformation in the estimation procedure:

$$\xi_{it} = \exp(\tilde{\xi}_{it}),$$

$$v_t = \exp(\tilde{v}_t) + 2.$$

where *exp(.)* denotes the exponential function. Hence the parameters in the Eqs. (8) and (9) are estimated without constraints.

Estimates of the parameter vector Θ that contains the coefficients of the bivariate GARCH models and those of Eqs. (8) and (9) are obtained by the maximization procedure applied to Eqs.(6) and (7). Since the SKST density with autoregressive conditional skewness and kurtosis parameters is more generalized than the other two, it is expected to improve the estimation accuracy of the time varying MV hedge ratios and thus yield higher hedging effectiveness.

2.3. Measurement of hedging effectiveness

We employ two methods to measure hedging effectiveness. The first one is based on variance reduction (VR) which is extensively applied in the literature. VR is a ratio of the difference between the variances of returns of an unhedged position and a hedge portfolio over the variance of returns of the unhedged position. Denoting ΔS_t as returns of an unhedged position and $\Delta V_{H,t}$ as returns of a hedge portfolio, VR can be calculated as

$$HE_{Variance} = 1 - \frac{Var(\Delta V_{H,t})}{Var(\Delta S_t)}. \quad (10)$$

where $Var(.)$ denotes variance. $HE_{Variance}$ refers to hedging effectiveness in terms of variance reduction.

This paper assumes a continuous hedging strategy for a hedger. A hedger starts to hedge against his/her portfolios using futures contracts at date *t*. Given a *k*-day horizon, the hedge stops at date $t+k$. Then the hedger can then start another round of hedging at date $t+k$. The second round ends at date $t+2k$. Then a third round may begin and so on. Multiple rounds of the hedge continue until the date $t+mk$ where m is a positive integer if the hedger does not stop his/her strategy. The continuous hedging strategy makes sense for a hedger in the market since one has to stick to one particular hedged position till the horizon ends. Note that the setting of a continuous hedging strategy requires k-day non-overlapping differencing to obtain returns of unhedged and hedged positions. The non-overlapping differencing would result in a small sample size when k takes large values. Hence the choice of k, i.e., the length of a hedging horizon, depends on the sample size of the sample period used to calculate the hedging effectiveness[6].

We calculate hedging effectiveness of the time varying hedge ratios forecasted by the four bivariate GARCH models under each conditional probability density for multiple hedging horizons. Note since the conditional hedge ratio is updated daily based upon the information of the previous day, its behaviour is independent of the length of hedging horizon. Denoting the conditional hedge ratio as β_{t-1} at date t-1 when a hedge portfolio is constructed, the portfolio returns at date t can be written as

$$\Delta V_{H,t} = \Delta S_t - \beta_{t-1}\Delta F_t. \tag{11}$$

Then the returns over a k-day hedging horizon are expressed as

$$\Delta_k V_{H,t+k} = \sum_{n=1}^{k} \Delta V_{H,t+n}. \tag{12}$$

where $\Delta_k V_{H,t+k}$ denotes the returns over k days of a hedge portfolio starting from date t. The variance reduction for a k-day hedging horizon is calculated by

$$HE_{variance,k} = 1 - \frac{Var(\Delta_k V_{H,t+k}\ \ )}{Var(\Delta_k S_{t+k}\ \ )}. \tag{13}$$

[6] The values of k are given by Section 3.

where $\Delta_k S_{t+k} = S_{t+k} - S_t$ is k-day returns of the unhedged position.

The variance reduction examines the effect of hedging on the second moments of returns distribution. It is also interesting to investigate how a hedging strategy affects the third and fourth moments of returns distribution. The effect reflects to what extent the hedging reduces risks attributed to the tail behaviour. We use value at risk (VaR), which estimates the maximum portfolio loss over a given time span at a given confidence level (Harris and Shen, 2006; Cotter and Hanley, 2006; Lai and Sheu, 2010; Conlon and Cotter, 2012). The VaR at confidence level α is

$$VaR_\alpha = -q_\alpha. \quad (14)$$

where q_α is the negative of the αth percentile of the realisations on returns. The effectiveness in terms of VaR reduction for a k-day hedging horizon is then calculated as

$$HE_{VaR,k} = 1 - \frac{VaR_\alpha(\Delta_k V_{H,t})}{VaR_\alpha(\Delta_k S_t)}. \quad (15)$$

where $VaR_\alpha(\Delta_k V_{H,t})$ and $VaR_\alpha(\Delta_k S_t)$ are the VaR of the k-day returns of the hedge portfolio and unhedged position at confidence level α, respectively.

The effect of the higher moments of a joint return distribution on a MV hedging strategy is attributed to the fact that the third and fourth moments of the hedged portfolio are a function of the optimal hedge ratio constructing that portfolio[7]. Hence, it is expected that the higher accuracy of estimating the optimal hedge ratio by a model results in the higher accuracy of capturing the skewness and kurtosis of the hedge portfolio returns. And thus the higher VaR reduction can be realised by that model.

Since a hedging strategy has to be implemented ex-ante, this paper calculates the out-of-sample hedging effectiveness for the time-varying MV hedge ratios. The in-sample model

[7] See the proof of this in Appendix 1.

estimates on the bivariate GARCH models are obtained first. These estimates are used to forecast the out-of-sample conditional covariance matrix of innovations following a recursive procedure. The conditional hedge ratios are derived based upon the forecasted conditional variance and covariance. The returns of hedge portfolio are computed using the forecasted series of conditional hedge ratios. Finally the hedging effectiveness for different hedging horizons is calculated.

3. Empirical results

3.1. Data and preliminary tests

This paper conducts an empirical analysis of the S&P 500 index futures contracts traded on the Chicago Mercantile Exchange (CME). These series represent highly liquid cash and futures markets and possess a long returns history. The daily closing (settlement) prices of the cash (futures) are collected from Datastream. The sample period runs from July 1, 1982 to March 31, 2009, which covers the period of global finance crisis (GFC)[8]. In order to obtain a continuous price series of the most liquid futures contracts, contracts at the nearest month are selected. Trading volumes of the nearest and the second nearest contracts are compared at the contract month. The nearest contract is switched to the next nearest when the volume of the former is exceeded. After matching the data of cash and futures prices, we end up with 6979 observations for the whole sample.

The out-of-sample hedging effectiveness is examined for two subsample periods. To examine hedging effectiveness for a tranquil period, the subsample path before the GFC from July 1, 1982 to September 28, 2007 is chosen where the in-sample estimation period is from July 1, 1982 to June 30, 1997 (with 3913 observations) and the out-of-sample forecasting period is from July 1, 1997 to September 28, 2007 (with 2674 observations). We refer to

[8] This is also referred to as the sub-prime mortgage crisis.

these two subsamples as Period I and Period II respectively. To examine hedging effectiveness during the crisis, the full sample path is used where the in-sample estimation uses the subsample from July 1, 1982 to September 28, 2007 (with 6587 observations) and the out-of-sample forecasting is done for the subsample from October 1, 2007 and March 31, 2009 (with 392 observations). We refer to these two subsamples as Period III and Period IV respectively.

Given the sample sizes of Periods II and IV, the numbers of hedging horizons are chosen to be 9 and 6 days for the calculation of the out-of-sample hedging effectiveness for the time varying optimal hedge ratios. That is, there are 9 horizons for the analysis on Periods I and II and 6 horizons for the analysis on Periods III and IV. The reason is that the chosen number of horizons assures enough observations to compute the out-of-sample hedging effectiveness for the hedge ratios under question. Henceforth, we have 9 k-day hedging horizons associated with Periods I and II where k = 1, 2, 4, 8, 16, 32, 64, 128, 256, respectively. Meanwhile, 6 k-day (k = 1, 2, 4, 8, 16, 32) hedging horizons associate with Periods III and IV given the small sample size of Period IV. The choice on values of k enables us to examine the hedging effectiveness as the length of horizon doubles. By doing this, the effectiveness of short horizons (when k = 1, 2, 4, 8, and 16) and long horizons (when k = 32, 64, 128 and 256) is simultaneously shown.

As mentioned earlier, we obtain model estimates using samples of Periods I and III and forecast the ratio series using samples of Periods II and IV. The effectiveness of those ratio series is computed for 9 horizons corresponding to Period II and 6 horizons corresponding to Period IV. Cash and futures daily prices are in natural logarithm form. Daily returns are calculated as the first differences of log prices. The k-day non-overlapping differencing is used to calculate k-day returns for cash and hedged portfolio for k-day hedging horizons. This applies to the evaluation of hedging effectiveness.

[Insert Table 1 about here]

The descriptive statistics of daily returns is shown in Table 1. While the means and standard deviations of cash and futures returns are similar across the tranquil sub-periods, the GFC period differs where the mean is negative and standard deviation almost double. This confirms a market downturn and increased volatility during the GFC. Cash and futures returns are left-skewed in all the tranquil sub-periods. The exception is that the futures returns are right-skewed during the GFC reflecting a large probability of negative extreme values of futures returns during this period. An excess kurtosis is revealed for both cash and futures returns for all the sub-periods, revealing the fat-tailed nature of returns distributions. Asymmetries and a large thickness of tails are non-negligible for returns distributions given that all the skewness and kurtosis coefficients are statistically significant (except skewness of cash returns in the GFC). The JB test results confirm the non-normality of asset returns, which will be taken into account by the SKST density function. Moreover, autocorrelation and heteroscedasticity characteristics of asset returns are verified by the Ljung-Box test given the significant Q test statistics. These are explored further through estimation of a bivariate VECM-GARCH model[9].

3.2. Static and conditional skewness and kurtosis

Estimates of static parameters for skewness and joint kurtosis in a bivariate SKST density are reported in Table 2[10]. All the estimates are statistically significant across the GARCH models, verifying the asymmetry and fat-tailed nature of joint distribution of cash and futures returns. The logs of skewness parameters ξ_1 and ξ_2 are both negative, aligning with the result in Table 1 that the cash and futures returns of Period I are both negatively

[9] ADF and PP confirmed that the log index cash and log futures prices are first difference stationary for all the subsamples. Johansen cointegration testing confirmed that the cash and futures prices are cointegrated for the same sub-periods.

[10] In the interests of brevity, results on the CCC, DCC, AGDCC and BEKK GARCH models are not reported but are available upon request.

skewed. The degree of freedom approaches 2, suggesting a large extent of thickness of tails. It is consistent with the high measure of the fourth moments of cash and futures returns in Table 1.

[Insert Table 2 about here]

In terms of the autoregressive processes associated with the skewness of marginal densities and joint kurtosis, the estimates of ξ_{11}, ξ_{21}, ξ_{12} and ξ_{22} are all significant across the GARCH models except ξ_{22} by the BEKK model. The skewness of cash (futures) returns is conditioned on the past shocks and its own lagged values. The result confirms the autoregressive behaviour of degree of asymmetry. Moreover, the signs of ξ_{11}, ξ_{21}, ξ_{12} and ξ_{22} are positive, which is also the case with the CCC, DCC, and AGDCC models. Also, ξ_{11} and ξ_{12} are positive for the BEKK model. The positive autocorrelation of third moment of the S&P 500 cash market is qualitatively consistent with the findings by Jondeau and Rockinger (2003) with regard to the same market, and with Bali et al. (2008) on the CRSP value-weighted index. An exception is the negative ξ_{21} attached with the BEKK, suggesting a negative autocorrelation of cash skewness. The result differs from Jondeau and Rockinger (2003) and Bali et al. (2008); however, it follows the finding by Harvey and Siddique (1999).

The dynamics of the degree of freedom parameter is revealed by significance of v_1, v_2 and v_3 despite insignificant estimates of v_2 and v_3 under the BEKK model. The result suggests the thickness of tails of the joint distribution of cash and futures markets hinges not only on the past shocks arising from the two markets but on its lagged own behaviour. Most of the estimated models agree that the joint kurtosis negatively relates to the past cash shocks but positively relates to the past futures shocks as well as its lagged own values. A significant positive autocorrelation of the joint kurtosis consists with the previous findings on the univariate densities for the similar markets (Jondeau and Rockinger, 2003; Brooks et al., 2005; Bali et al., 2008). Overall, Table 2 reveals the existence of conditional autoregressive processes for skewness and kurtosis parameters in a bivariate density of asset returns[11].

3.3. A comparison of hedging effectiveness

We examine two aspects of the effects of the static and conditional higher order moments on hedging effectiveness of the conditional optimal hedge ratios. First, we examine how hedging effectiveness differs across densities under each GARCH model. Second, the question how the best-performed GARCH model varies across densities is addressed. The out-of-sample hedging effectiveness of the conditional optimal hedge ratios for a normal period (i.e. Period II) and a crisis period (i.e. Period IV) is shown horizon-by-horizon in Table 3 and Table 4, respectively. Note that Period IV is a period of global finance crisis (GFC).

[Insert Table 3 about here]

In Table 3 under the CCC GARCH model, the SKST density provides the best variance reduction in 6 out of 9 horizons most of which have short length. Regarding the reduction in tail risk, the SKST density still performs better than the other two, providing the highest VaR reduction in 5 out of 9 horizons at both 1% and 5% confidence levels. However, we do not find evidence that the SKST density with autoregressive skewness and kurtosis parameters provides benefit for hedging effectiveness under the CCC model. The result from the BEKK model is similar to that for the CCC. The SKST density performs best in terms of both variance reduction and VaR reduction in all the horizons. It also can be seen that the SKST density with autoregressive skewness and kurtosis parameters beats the normal counterpart for the longest hedging horizon in terms of both variance reduction and VaR reduction.

[11] The results from Period III are qualitatively similar to those of Period I. To save space, those results are not reported here. .

Evidence is found under the DCC and AGDCC models in Table 3 that the SKST density with autoregressive skewness and kurtosis parameters benefits hedging effectiveness for the conditional optimal hedge ratios. Under the DCC model, the SKST density with autoregressive skewness and kurtosis parameters performs best in 6 out of 9 horizons in terms of variance reduction. The performance of the same density in terms of VaR reduction is best for 3 horizons at 1% and 4 horizons at 5% confidence levels, respectively. An increasing trend of hedging effectiveness is found for 6 out of 9 horizons across the densities of normal, SKST, and SKST with autoregressive higher order moments. Most of those horizons are short ones, ranging from 1-day to 32-days.

The result from the AGDCC model provides stronger evidence. The SKST density with autoregressive skewness and kurtosis parameters produces the best variance reduction in all the horizons. It provides the best VaR reduction for 8 horizons at both 1% and 5% confidence levels. Those horizons range from short to long periods. For all the horizons, hedging effectiveness increases across the densities of normal, SKST and SKST with autoregressive asymmetry and thickness of tails.

On average, compared with the normal density, the SKST density increases variance reduction by around 0.84%. It also increases VaR reduction at 1% and 5% confidence levels by 2.24% and 2.09%, respectively. The SKST density with autoregressive skewness and kurtosis parameters further increases variance reduction by 0.68% than the SKST density. The former increases VaR reduction at 1% and 5% confidence levels by 2.53% and 2.08% than the latter, respectively.

If we examine how the best-performing model varies across the different densities, the DCC model performs best for most of hedging horizons under the normal density. When we turn to the SKST density, the result changes such that the BEKK model takes the lead, which is true for almost all the horizons. However, when the autoregressive asymmetry and thickness of tails are taken into account, the DCC model performs best again for all the horizons. The result implies that the DCC models are seemingly the optimal choice for the dynamic minimum-variance hedging. It should be cautious that the result is not stable as the best-performed model varies across different assumptions on the shape behaviour of returns distribution.

[Insert Table 4 about here]

Regarding the effects of higher order moments on the conditional optimal hedge ratios, the results for the GFC period are shown in Table 4. A monotonic growth of both variance and VaR reductions is evident across the different density functions using the CCC, DCC, and AGDCC models, and is the case for most of the hedging horizons. Under each GARCH model, the normal density is dominated by the other two density functions for most hedging horizons. Under the BEKK model, an increase of hedging effectiveness is valid only for 2 horizons. However, the SKST density still performs best, which is evidenced in 4 out of 6 hedging horizons. In addition, the SKST density with autoregressive skewness and kurtosis parameters outperforms normality for all the horizons.

On average, it is found that the SKST density function can improve variance reduction by around 3.11% than the Normal counterpart. The former also increases VaR reduction by 6.99% and 6.10% at the 1% and 5% confidence levels over the Normal, respectively. Compared with the static higher order moments, the autoregressive cases further increase variance reduction by 0.80%, VaR reduction at 1% level by 2.69% and VaR reduction at 5% level by 2.46%, respectively. This result indicates the benefit of incorporating the autoregressive behaviour of higher order moments on hedging performance of dynamic minimum-variance hedging strategy is more significant during the GFC period than a normal period. Hence the conditional higher moments are helpful for improving the quality of risk management in light of volatility minimisation and tail risk reduction, especially during a crisis period.

Table 4 reports a different result for the variation of the best-performing model across densities than that reported in Table 3. The best model under normality is the CCC in terms of both variance and VaR reductions for all the horizons. BEKK performs best when the SKST density is employed. The best performance of the BEKK model is evident for most of horizons. However, when allowing the higher order moments to be autoregressive, we find DCC performs best for relatively short horizons (1-day, 2-day, 4-day, and 8-day) whereas BEKK performs best for relatively long ones (16-day and 32-day). Thus, the variation of the best-performed model across densities during GFC is more substantial than during a normal period. Our result suggests that the BEKK and DCC models might be rational choices on the dynamic minimum-variance hedging strategies during a crisis period. However, it should be kept in mind that the pattern of returns distribution has an impact on those choices. Such impact is valid irrespective of when the hedging activities take place.

[Insert Figure 1 about here]

Figure 1 indicates that when the hedging horizon increases, most hedge ratio series exhibit a growing trend of hedging effectiveness which is roughly monotonic. The growth of hedging effectiveness along with an increase of hedging horizon in length is witnessed for both normal and crisis periods, which is consistent with previous studies (e.g. Ederington, 1979; Geppert, 1995; Chen et al., 2004; Lien and Shrestha, 2007; Lai and Sheu, 2010). Moreover, we can consistently observe from the figure that the conditional hedging ratios derived from the densities with the static and conditional higher order moments possess lines running over those from normality. Such observation is more obvious for the crisis period.

3.4. Relations between hedge ratio and hedging effectiveness

In this subsection we examine how the relations between several statistical properties of the conditional hedge ratios and hedging effectiveness differ across different conditional densities. The statistical properties of the conditional hedge ratios under question include the mean, standard deviation, smoothness of ratio series and smoothness of standard deviation. Note the latter three properties contribute towards explaining the stability of series of conditional hedge ratios. The smoothness of ratio series is measured by the standard deviation of the partial first derivatives of hedge ratios over time. Likewise, we use the standard deviation of the partial first derivatives of hedge ratios' standard deviation over time as a proxy for the smoothness of ratio standard deviation[12]. Note that the standard deviation of the partial first derivatives measures the average dispersion of changes of the time series over time. If the standard deviation equals zero, the change is fixed and the series is a linear function of time. Hence one intuitively expects that the lower the standard deviation, the higher the series' smoothness. The properties of the conditional hedge ratios for both Period II and Period IV are summarised in Table 5.

[Insert Table 5 about here]

We examine whether the conditional hedge ratios derived by the best (worst)-performed hedging model under each density possess the shortest (longest) distance of the mean to unity, the smallest (largest) standard deviation, the highest (lowest) ratio smoothness and the highest (lowest) smoothness of ratio standard deviation. Evidence is revealed for Period II and Period IV, respectively.

[12] The partial first derivative over time is an estimate of coefficient of the time trend in a regression model where the hedge ratio and standard deviation series are respectively regressed against a time trend. The series of estimates are obtained by a rolling window method with a window size of 100 observations and step size of 1 observation.

We find from Period II that the best model under normality, i.e. the DCC model, yields ratios the mean of which is closest to unity. However, they have the second smallest standard deviation and second highest smoothness of ratios and standard deviation. The situation gets better when moving to the SKST density where the ratios generated by the best model under that density (BEKK model) possess the smallest standard deviation and highest smoothness of ratios and standard deviation. However, the SKST density with autoregressive higher order moments only confirms that the best model produces the least volatile ratio series. The argument that the ratios that are closest to unity have the best performance is not supported by the densities with static and conditional higher order moments.

The results concerning the properties of ratios derived by the worst models are clearer. Staying with Period II, the worst model under normality (CCC model) produces the ratios with the longest distance of mean to unity, smallest standard deviation and lowest smoothness of ratio and standard deviation series. The finding is confirmed by the densities with static and conditional higher order moments. Thus, results from Period II reveal that the SKST density reinforces the fact of a positive relation between stability of hedge ratios and hedging effectiveness than the normal counterpart.

The results from Period IV reinforce these findings. Results based on all the density functions concur that the ratio series with its mean closest to unity has the best effectiveness. Only the SKST density function finds that the worst performance relates to the longest distance of the mean from unity. The Normal and SKST density function with autoregressive higher order moments supports the case that the least (most) volatile ratios with highest (lowest) smoothness are best (worst) performing. The SKST density function with autoregressive higher order moments further finds that the highest (lowest) smoothness of standard deviation of ratios series contributes to the best (worst) hedging effectiveness. In sum, we find both static and conditional higher order moments reinforce the view that a higher stability of hedge ratios leads to better hedging effectiveness. The result enhances the findings by the previous studies on the relationship between conditional hedge ratio and hedging effectiveness (e.g. Park and Jei, 2010; Lien, 2010; Hou and Li, 2013; Kim and Park, 2016).

[Insert Figure 2 about here]

The relations between different statistical properties of hedge ratios and average hedging effectiveness over all the hedging horizons are plotted in Figure 2[13]. The figure shows how those properties change with average hedging effectiveness across the GARCH models under each density function. As can be seen from the figure, the average hedging effectiveness increases as the hedge ratio means approach unity. This is for all densities in both Period II and Period IV. The volatility of ratio series appears negatively related to hedging effectiveness where the lower the volatility, the better the effectiveness. In addition, there are rough downward trends for the standard deviations of the partial first derivatives of the ratios series and their standard deviations over time, and the average hedging effectiveness. The downward trends indicate a positive relationship between the smoothness of ratio series and volatility and hedging effectiveness. We find those trends are more highlighted under the SKST densities with static and autoregressive higher order moments despite some of them not being monotonic. Indeed, so is the negative relation between volatility and hedging effectiveness. The result based on the analysis of the average hedging effectiveness is consistent with that based upon the best (worst) performed hedging models.

3.5. Hedge ratio and higher order moments

The effects of higher order moments on the statistical properties of conditional hedge ratios are further explored in this subsection. Those effects are qualitatively described by

[13] The properties are plotted against variance reduction in the figure. The plots with VaR reduction on the X axis are similar. These are not reported here, but are available upon request.

Figure 3. This figure depicts how the mean, standard deviation, smoothness of ratio series and their standard deviations vary across different densities under each GARCH model. As mentioned earlier, smoothness is measured by the standard deviation of the partial first derivatives of time series over time.

[Insert Figure 3 about here]

As can be seen from the figure, there is an upward trend in the mean for 6 out of 8 plots. The mean normally converges to unity. It suggests that taking into account the static and conditional higher order moments may monotonically increase the mean of hedge ratios when it is below 1. The standard deviation of ratios moves continuously downward when the densities with the static and conditional higher order moments are chosen. It is confirmed in 5 out of 8 plots. Moreover, a monotonic downward trend of standard deviations of the partial first derivatives of ratios and their standard deviations over time is respectively detected by 5 and 6 out of 8 plots when moving across the X axis from left to right. It suggests smoothness of series may be consistently enhanced by higher order moments and their autoregressive feature.

To explore the quantified effects, we estimate the following equations:

$$\hat{\beta}_t = intercept + \alpha_1 skew_{1t} + \alpha_2 skew_{2t} + \alpha_3 v_t + e_t,$$

$$\hat{\sigma}^{\hat{\beta}_t}{}_t = intercept + \beta_1 skew_{1t} + \beta_2 skew_{2t} + \beta_3 v_t + e_t,$$

$$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t = intercept + \gamma_1 skew_{1t} + \gamma_2 skew_{2t} + \gamma_3 v_t + e_t,$$

$$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t = intercept + \delta_1 skew_{1t} + \delta_2 skew_{2t} + \delta_3 v_t + e_t. \qquad (16)$$

where $\hat{\beta}_t$ denotes the estimated conditional optimal hedge ratios; $\hat{\sigma}^{\hat{\beta}_t}{}_t$ represents the series of standard deviation of $\hat{\beta}_t$. $\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$ and $\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$ denote the time series of standard deviations of $\partial\hat{\beta}_t/\partial t$ and $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$, respectively. $\partial\hat{\beta}_t/\partial t$ and $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ are the partial first derivatives of $\hat{\beta}_t$ and $\hat{\sigma}^{\hat{\beta}_t}{}_t$ over time, respectively. Time variation of $\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$ and $\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$ is obtained by the rolling window mechanism with window size 100 observations and step size 1 observation. $skew_{1t}$ and $skew_{2t}$ are the conditional skewness measures for marginal densities of cash and futures returns in a bivariate SKST distribution, respectively. $skew_{it} = \xi^2{}_{it} (i = 1,2)$ and ξ_{it} is an autoregressive skewness parameter that is an exponential function of $\tilde{\xi}_{i,t}$ defined as in Eq. (8). The sign of $skew_{it}$ is determined by $ln(\xi_{it})$ where $skew_{it}$ is positive (negative) when $ln(\xi_{it}) > 0\ (< 0)$. v_t is the conditional degree of freedom of the bivariate SKST density. v_t is an exponential function of $\tilde{v}_t$ defined as in Eq. (9). Data of Period II and IV are employed for estimating Eq. (16), respectively. Estimation results are separately obtained for the CCC, DCC, AGDCC and BEKK GARCH models for each period, which are shown in Table 6.

[Insert Table 6 about here]

As can be seen from the table, the marginal skewness of futures returns promotes the size of hedge ratio, which is indicated by all the positive significant estimates. The same effect is also rendered by the marginal skewness of cash returns, which is suggested by 4 out of 6 positive significant estimates. The effect from the joint kurtosis is negative given that 3 out of 5 estimates are significant negative. However, the joint effect of skewness and kurtosis on the size of hedge ratio is positive, suggesting that higher order moments of a bivariate distribution escalate the conditional hedge ratios.

Significant effects of higher order moments on the volatility of conditional hedge ratios are found only for Period II. The marginal skewness of futures market decreases the volatility whereas that of cash market does in a reverse way. The argument is supported by most of significant estimates. The effect from the joint kurtosis is mixed given that one significant estimate is positive while the other is negative. However, an agreement can be achieved that the joint effect of skewness and kurtosis is negative if the individual effects are summed. The supporting evidence is found in the estimation under the DCC, AGDCC and BEKK models. The result implies that taking into account the higher order moments helps to gauge less volatile hedge ratio series.

Lastly, the marginal skewness of cash returns downgrades the smoothness of hedge ratio and its volatility whereas that of futures returns elevates it. This is supported by the estimation results under the CCC, DCC and BEKK models. Meanwhile, the joint kurtosis decreases the smoothness of relevant series, suggested by the results under the AGDCC and BEKK models. However, the aggregate effect suggests that the smoothness of both ratio and volatility series is enhanced when both skewness and kurtosis are taken into account.

The finding is useful in explaining why the SKST density with the static and autoregressive skewness and kurtosis parameters improves hedging effectiveness of conditional optimal hedge ratios than the normality. The reasoning is that the higher order moments produce hedge ratios with larger size and higher stability and consequently those ratios possess higher hedging effectiveness. Likewise, the autoregressive higher order moments generate higher effectiveness than the static ones since the hedge ratios estimated by the former possess larger size and higher stability.

3.6. Robustness check on hedging effectiveness

We examine whether the effects of the static and autoregressive higher order moments on hedging effectiveness during a post-GFC period are similar to those of pre- and during-GFC periods[14]. In doing this, we collect a sample of daily closing (settlement) prices of the S&P

[14] A robustness check was also conducted on the effects of the higher order moments on the relations between hedge ratios and hedging effectiveness, and on the statistical properties of hedge ratios. The results of the post-GFC period are qualitatively similar to those of the other examined periods.

500 cash (futures) from April 1, 2009 to March 30, 2017. We end up with 2087 observations for the sample. Following the analysis on the pre-GFC period, the out-of-sample hedging effectiveness of the CCC, DCC, AGDCC, and BEKK GARCH hedge ratios is assessed for 9 hedging horizons[15]. The results are shown in Table 7.

[Insert Table 7 about here]

On one hand, the table shows a monotonic increase in variance reduction when moving across the densities of normality, SKST, and SKST with autoregressive skewness and kurtosis parameters. Such increase is evidenced by all the hedging horizons of the CCC hedge ratios and 8 out of 9 horizons of the AGDCC and BEKK ratios. Likewise, a monotonic growth of VaR reduction exists, which is agreed by 9 horizons of the CCC, 8 horizons of the BEKK, and 6 horizons of the AGDCC. Although the DCC hedge ratios do not witness the superiority of the autoregressive higher order moments in hedging effectiveness, they strongly support that the static higher order moments yield higher effectiveness than the normality for most of hedging horizons. On average, compared to the normality, the static higher order moments increase variance reduction by around 2.52%, VaR reduction at the 1% level by around 9.45%, and VaR reduction at the 5% level by around 3.97%. The autoregressive counterparts further increase variance reduction by around 0.57%, VaR reduction at the 1% level by around 2.50%, and VaR reduction at the 5% level by around 3.45%. The post-GFC results are similar to those of the other periods reported in this paper.

On the other hand, it is found the evidence that the static and autoregressive higher order moments affect the choice on the best-performed hedging model. The CCC GARCH model keeps performing best across all the hedging horizons under normality. This situation varies when turning to the SKST density where the BEKK model is the best for all the horizons. This choice rarely changes when the skewness and kurtosis are time varying. Hence the post-

[15] The results of the GARCH model estimation are available upon request.

GFC evidence consists with that of the pre- and during-GFC periods given that taking into account the static and autoregressive higher order moments impacts the decision on the dynamic hedging model. The result also points to a change of the best-performed model from pre- to post-GFC periods. That is, the DCC model might be the best before the crisis while the BEKK takes the lead after that.

4. Concluding remarks

Although dynamic minimum-variance futures hedging has been widely investigated in the literature, the effects of non-standard tail behaviour of asset returns on conditional hedge ratio remain unclear. This paper explores three aspects important aspects of this issue: (i) whether the bivariate skew-Student density functions with the static and autoregressive third and fourth moments improves hedging effectiveness over the Normal distribution; (ii) whether the density functions matters in terms of the relationship between hedge ratio and hedging effectiveness; and (iii) whether the density functions impact on the time varying features of ratio series.

Compared to the Normal density, there is evidence that the density functions with the static higher order moments increases both variance and VaR reduction. Compared to the static counterpart, the autoregressive higher moments can further increase the same metrics. A monotonic increase in hedging effectiveness from the normal to the non-normal densities is evidenced by at short and long hedging horizons. These findings are obtained for both tranquil and GFC periods. Those improvements imply that taking into account the static and conditional higher order moments of asset returns benefits risk management. Moreover, this paper reveals that the best-performing hedging model varies across different densities. An attention should be paid to the tail behaviour of asset returns for a hedger since it affects the decision on the choice of a best dynamic hedging strategy. The results of the effects of higher order moments on hedging effectiveness are robust in a post-GFC period.

The relations between conditional hedge ratio and hedging effectiveness are revealed that higher magnitude and stability of ratio series lead to better effectiveness. The stability is reflected by volatility and smoothness of ratio and volatility series. This finding is enhanced by the static and autoregressive higher order moments. Furthermore, significant evidence suggests that the static and conditional higher order moments contribute to the size and stability of conditional hedge ratios. This can thus serve as a reason behind on the improvements on hedging effectiveness by the non-normal conditional densities.

Appendix 1 Proofs of a quantitative relationship between the higher moments of a minimum variance hedge portfolio and the optimal hedge ratio

The proof that the third moment of a minimum variance hedge portfolio is a function of the optimal hedge ratio is given by Fu (2014, pp 214).

The proof for the fourth moment stems from Fu (2014). Suppose a minimum variance hedge portfolio is constructed by

$$R_h = R_s - h \quad R_f. \tag{A1.1}$$

where R_h , R_s and R_f are returns of hedge portfolio, unhedged position and futures, respectively. h is the minimum variance hedge ratio. From Eq.(A1.1), the fourth moment of R_h is derived by

$$\begin{aligned} u_4(R_h) &= E[R_h - E(R_h)]^4 \\ &= E\{R_s - E(R_s) - h \quad [R_f - E(R_f)]\}^4 \\ &= E[R_s - E(R_s)]^4 - 2hE\{[R_s - E(R_s)]^3[R_f - E(R_f)]\} + 2h^2E\left\{[R_s - \right. \\ &\quad \left. E(R_s)]^2[R_f - E(R_f)]^2\right\} - 2h^3E\left\{[R_s - E(R_s)] \quad [R_f - E(R_f)]^3\right\} + \\ &\quad h^4E[R_f - E(R_f)]^4. \end{aligned}$$

where the fourth and co-fourth moments of the cash and futures returns are defined by

$u_4(R_s) = E[R_s - E(R_s)]^4$,

$u_4(R_f) = E[R_f - E(R_f)]^4$,

$u_{3,1} = E\{[R_s - E(R_s)]^3[R_f - E(R_f)]\}$,

$u_{2,2} = E\left\{[R_s - E(R_s)]^2[R_f - E(R_f)]^2\right\}$,

$u_{1,3} = E\left\{[R_s - E(R_s)] \quad [R_f - E(R_f)]^3\right\}$.

Then we have

$$u_4(R_h) = u_4(R_s) - 2hu_{3,1} + 2h^2u_{2,2} - 2h^3u_{1,3} + h^4u_4(R_f). \tag{A1.2}$$

Eq.(A1.2) clearly shows that $u_4(R_h)$ is a function of h, the fourth moments and the co-fourth moments.

References

Agarwal, V., and Naik, N. (2004). Risks and portfolio decisions involving hedge funds. Review of Financial Studies, 17, 63–98.

Baillie, R.T., and Bollerslev.T. (1989). The message of daily exchange rates. Journal of Business & Economic Statistics 7, 287-305.

Baillie, R.T., Han, Y.W., Myers, and R.J., Song, J. (2007). Long memory models for daily and high frequency commodity futures returns. Journal of Futures Markets 27 (7), 643–668.

Bali, T.G., Mo, H., and Tang, Y. (2008). The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. Journal of Banking and Finance 32, 269-282.

Barone-Adesi, G., Gagliardini, P., and Urga, G. (2004). Testing asset pricing models with coskewness. Journal of Business and Economic Statistics, 22(4), 474–485.

Bauwens, L., and Laurents, S. (2005). A new class of multivariate skew densities with application to GARCH models. Journal of Business & Economic Statistics 23, 346–354.

Bera, A.K., Garcia, P., and Roh, J.S. (1997). Estimation of time-varying hedging ratios for corn 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. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69,542-547.

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.

Brooks, C., Burke, S.P., Heravi, S., and Persand, G. (2005). Autoregressive conditional kurtosis. Journal of Financial Econometrics 3(3), 399-421.

Brooks, C., Cerny, A., and Miffre, J. (2012). Optimal hedging with higher moments. The Journal of Futures Markets 32 (10), 909-944.

Brooks, C., Henry, O. T., and Persand, G. (2002). The effect of asymmetries on optimal hedge ratios. Journal of Business 75(2), 333–352.

Brooks, C., and Kat, H. (2002). The statistical properties of hedge fund index returns and their implications for investors. Journal of Alternative Investments 5(2), 26–44.

Cao, Z., Harris, R. D. F., and Shen, J. (2009). Hedging and value at risk: A semi-parametric approach. Journal of Futures Markets, 30(8), 780–794.

Cappiello, L., Engle, R.F., and Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial Econometrics 4(4), 537-572.

Chen, S. S., Lee, C. F., and 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.

Christie-David, R., and Chaudhry, M. (2001). Coskewness and cokurtosis in futures markets. Journal of Empirical Finance 8, 55–81.

Conlon, T., and Cotter, J. (2012). An empirical analysis of dynamic multiscale hedging using wavelet decomposition. The Journal of Futures Markets 32(3), 272-299.

Cotter, J., and 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. (2002). Dynamic conditional correlation: a simple class of multivariate generalised autoregressive conditional heteroscedasticity models. Journal of Business and Economic Statistics 20(3), 339-350.

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

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

Fernandez, C., and Steel, M. (1998). On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association 93, 359–371.

Fu, J.H. (2014). Multi-objective hedging model with the third central moment and the capital budget. Economic Modelling 36, 213-219.

Geppert, J. M. (1995). A statistical model for the relationship between futures contract hedging effectiveness and investment horizon length. Journal of Futures Markets 15(5), 507–536.

Gilbert, S., Jones, S. K., and Morris, G. H. (2006). The impact of skewness in the hedging decision. Journal of Futures Markets, 26(5), 503–520.

Hansen, B.E. (1994). Autoregressive conditional density estimation. International Economics Review 35, 705-730.

Harris, R.D.F., and Shen, J. (2006). Hedging and value at risk. Journal of Futures Markets 26(4), 369-390.

Harvey, C. (1995). Predictable risk and returns in emerging markets. Review of Financial Studies 8, 773–816.

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

Harvey, C.R., and Siddique, A. (2000). Conditional skewness in asset pricing tests. The Journal of Finance 55(3), 1263-1295.

Hong, Y., Tu, J., and Zhou, G. (2007). Asymmetries in stock returns: statistical tests and economic evaluation. Review of Financial Studies 20(5), 1547–1581.

Hou, Y., and Li, S. (2013). Hedging performance of Chinese stock index futures: An empirical analysis using wavelet analysis and flexible bivariate GARCH approaches. Pacific-Basin Finance Journal 24, 109-131.

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

Jondeau, E., and Rockinger, M. (2006). Optimal portfolio allocation under higher moments. European Financial Management 12(1), 29–55.

Kallberg, J. G., and Ziemba, W. T. (1983). Comparison of alternative utility functions in portfolio selection problems. Management Science 29(11), 1257–1276.

Kraus, A., and Litzenberger, R. (1976). Skewness preference and the valuation of risky assets. Journal of Finance, 31, 1085–1100.

Kraus, A., and Litzenberger, R. (1983) On the distributional conditions for a consumption-oriented three moment CAPM. Journal of Finance, 38(5), 1381–1391.

Kroner, K.F., and Sultan, J. (1993). Time varying distribution and dynamic hedging with foreign currency futures. Journal of Financial and Quantitative Analysis 28, 535–551.

Lai, Y.-S., and Sheu, H.-J. (2010). The incremental value of a futures hedging using realised volatility. The Journal of Futures Markets 30(9), 874-896.

Levy, H. (1969). A utility function depending on the first three moments. Journal of Finance 24, 715–719.

Lien, D. (2010). A note on the relationship between the variability of the hedge ratio and hedging performance. The Journal of Futures Markets 30 (11), 1100–1104.

Lien, D., and 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., and Shrestha, K. (2010). Estimating optimal hedge ratio: a multivariate skew-normal distribution approach. Applied Financial Economics 20(8), 627-636.

Nelson, D.B. (1991). Conditional heteroscedasticity in asset returns: A new approach. Econometrica 59(2), 347-370.

Nelson, D. (1996). Asymptotic filtering theory for multivariate ARCH models. Journal of Econometrics (71), 1-47.

Park, S.Y., and 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.

Kim, M.J., and Park, S.Y. (2016). Optimal conditional hedge ratio: A simple shrinkage estimation approach. Journal of Empirical Finance 38, 139-156.

Post, T., Vliet, P., and Levy, H. (2008). Risk aversion and skewness preference. Journal of Banking and Finance 32(7), 1178–1187.

Scott, R., & Horvath, P. (1980). On the direction of preference for moments of higher order than the variance. Journal of Finance, 35, 915–919.

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

Table 1. Descriptive statistics of the S&P 500 cash and futures daily returns

	Full Sample (1982.7.1-2009.3.31)		*Period I (1982.7.1-1997.6.30)*		*Period II (1997.7.1-2007.9.28)*		*Period III (1982.7.1-2007.9.28)*		*Period IV (2007.10.1-2009.3.31)*	
	$R_{s,t}$	$R_{f,t}$	$R_{s,t}$	$R_{f,t}$	$R_{s,t}$	$R_{f,t}$	$R_{s,t}$	$R_{f,t}$	$R_{s,t}$	$R_{f,t}$
Mean	2.86×10^{-4}	2.84×10^{-4}	5.36×10^{-4}	5.35×10^{-4}	2.04×10^{-4}	2.05×10^{-4}	4.01×10^{-4}	4.01×10^{-4}	-1.66×10^{-4}	-1.68×10^{-3}
Median	2.22×10^{-4}	2.72×10^{-4}	3.07×10^{-4}	3.31×10^{-4}	1.31×10^{-4}	2.89×10^{-4}	2.65×10^{-4}	3.28×10^{-4}	0.000	0.000
Maximum	0.110	0.177	0.087	0.177	0.056	0.058	0.087	0.177	0.110	0.132
Minimum	-0.229	-0.337	-0.229	-0.337	-0.071	-0.078	-0.229	-0.337	-0.095	-0.104
Std. Dev.	0.011	0.013	0.010	0.012	0.011	0.011	0.010	0.012	0.024	0.024
Skewness	-1.356***	-2.506***	-3.769***	-5.679***	-0.084*	-0.151***	-1.852***	-3.468***	-8.73×10^{-4}	0.267**
Kurtosis	34.899***	90.069***	95.250***	203.384***	6.356***	6.957***	45.900***	125.096***	6.792***	8.503***
JB	2.98×10^{5}***	2.21×10^{6}***	1.40×10^{6}***	6.57×10^{6}***	1.26×10^{3}***	1.76×10^{3}***	5.09×10^{5}***	4.10×10^{6}***	234.878***	499.310***
Q(12)	41.018***	79.406***	26.498***	62.388***	22.682**	21.562**	27.902***	60.169***	28.976***	30.334***
$Q^2(12)$	1157.402***	730.462***	247.310***	332.648***	656.985***	662.053***	457.969***	564.078***	313.207***	307.319***

Notes: This table reports the descriptive statistics of daily log returns of the S&P 500 cash index and index futures. The daily returns are calculated as the first difference of logarithmic prices. $R_{s,t}$, daily cash returns at date t; $R_{f,t}$, daily futures returns at date t. *Period I* denotes the subsample from July 1, 1982 to June 30, 1997; *Period II* denotes the subsample from July 1, 1997 to September 28, 2007; *Period III* denotes the subsample from July 1, 1982 to September 28, 2007; *Period IV* denotes the subsample from October 1, 2007 to March 31, 2009. *Std. Dev.*, standard deviation; *JB*, the Jarque-Bera test statistic for normality. *Q(12)* and $Q^2(12)$ denote the Ljung-Box *Q* test statistics for returns and its squares up to lag order 12, respectively. Results of the skewness/kurtosis tests for normality are reported. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 2. The static and autoregressive skewness and kurtosis parameters

Coefs.	*CCC-SKST*	*CCC-SKST-ARSK*	*Coefs.*	*DCC-SKST*	*DCC-SKST-ARSK*	*Coefs.*	*AGDCC-SKST*	*AGDCC-SKST-ARSK*	*Coefs.*	*BEKK-SKST*	*BEKK-SKST-ARSK*
ξ_1	0.643***		ξ_1	0.734***		ξ_1	0.967***		ξ_1	0.952***	
	(0.0031)			(0.0048)			(0.0129)			(0.0096)	
ξ_2	0.703***		ξ_2	0.762***		ξ_2	0.977***		ξ_2	0.952***	
	(0.0043)			(0.0058)			(0.0130)			(0.0096)	
v	2.004***		v	2.004***		v	2.676***		v	2.000***	
	(0.0005)			(0.0005)			(0.0332)			(2.81×10^{-5})	
ξ_{01}		0.555***	ξ_{01}		0.090***	ξ_{01}		0.105***	ξ_{01}		0.072***
		(0.0085)			(0.0007)			(0.0012)			(0.0018)
ξ_{11}		0.047***	ξ_{11}		0.049***	ξ_{11}		0.039***	ξ_{11}		0.076***
		(0.0007)			(0.0002)			(0.0002)			(0.0015)
ξ_{21}		0.181***	ξ_{21}		0.090***	ξ_{21}		0.091***	ξ_{21}		-0.177***
		(0.0124)			(0.0007)			(0.0062)			(0.0043)
ξ_{02}		0.535***	ξ_{02}		0.098***	ξ_{02}		0.100***	ξ_{02}		0.100***
		(0.0031)			(0.0009)			(0.0010)			(0.0098)
ξ_{12}		0.098***	ξ_{12}		0.052***	ξ_{12}		0.050***	ξ_{12}		0.050***
		(0.0005)			(0.0003)			(0.0003)			(0.0017)
ξ_{22}		0.106***	ξ_{22}		0.044***	ξ_{22}		0.048***	ξ_{22}		0.049
		(0.0046)			(0.0009)			(0.0011)			(0.0972)
v_0		2.813***	v_0		3.284***	v_0		2.963***	v_0		2.803***
		(0.0592)			(0.1348)			(0.0999)			(0.5275)
v_1		-0.201***	v_1		0.050***	v_1		-0.033***	v_1		-0.251***
		(0.0252)			(0.0100)			(0.0109)			(0.0472)
v_2		-0.143***	v_2		0.034***	v_2		0.038***	v_2		0.017
		(0.0151)			(0.0094)			(0.0111)			(0.0386)
v_3		0.092***	v_3		0.068*	v_3		0.063**	v_3		0.051
		(0.0191)			(0.0384)			(0.0310)			(0.1814)

Notes: This table reports estimation results of skewness and kurtosis parameters in Eq. (7) and coefficients governing autoregressive processes of skewness and kurtosis parameters defined in Eqs. (8) and (9) in a skew-Student probability density. Data of subsample Period I (July 1, 1982-June 30, 1997) is used to obtain estimates via maximum likelihood estimation (MLE). The results under CCC, DCC, AGDCC, and BEKK GARCH models are respectively reported. *Coefs.* denotes coefficients. *CCC-SKST*, CCC GARCH model with skew-Student density; *CCC-SKST-ARSK*, CCC GARCH model with skew-Student density with autoregressive skewness and kurtosis parameters. *DCC-SKST*, DCC GARCH model with skew-Student density; *DCC-SKST-ARSK*, DCC GARCH model with skew-Student density with autoregressive skewness and kurtosis parameters. *AGDCC-SKST*, AGDCC GARCH model with skew-Student; *AGDCC-SKST-ARSK*, AGDCC GARCH model with skew-Student density with autoregressive skewness and kurtosis parameters. *BEKK-SKST*, BEKK GARCH model with skew-Student density; *BEKK-SKST-ARSK*, BEKK GARCH model with skew-Student density with autoregressive skewness and kurtosis parameters. Figures in parentheses are estimated standard errors. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Table 3. Out-of-sample hedging effectiveness of time-varying hedge ratios in a normal period

		Time varying hedge ratios											
		CCC			*DCC*			*AGDCC*			*BEKK*		
Hedging horizon	*Hedging effectiveness*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*
1-day	*VR*	0.939	0.942	0.704	0.940	0.942	0.943	0.914	0.913	0.927	0.922	0.944	0.886
	VaR reduction, 1%	0.724	0.723	0.457	0.726	0.720	0.721	0.678	0.658	0.676	0.695	0.724	0.618
	VaR reduction, 5%	0.765	0.769	0.487	0.770	0.771	0.771	0.704	0.731	0.745	0.712	0.778	0.670
2-day	*VR*	0.959	0.963	0.735	0.960	0.963	0.964	0.933	0.933	0.948	0.942	0.965	0.903
	VaR reduction, 1%	0.795	0.794	0.507	0.800	0.788	0.794	0.725	0.748	0.747	0.759	0.810	0.707
	VaR reduction, 5%	0.789	0.800	0.473	0.791	0.805	0.807	0.736	0.761	0.773	0.749	0.807	0.701
4-day	*VR*	0.970	0.973	0.765	0.971	0.974	0.974	0.942	0.949	0.965	0.950	0.976	0.918
	VaR reduction, 1%	0.758	0.785	0.495	0.758	0.778	0.781	0.741	0.738	0.772	0.744	0.796	0.677
	VaR reduction, 5%	0.799	0.795	0.482	0.800	0.804	0.799	0.737	0.758	0.776	0.752	0.812	0.718
8-day	*VR*	0.975	0.978	0.789	0.976	0.979	0.979	0.946	0.951	0.967	0.951	0.981	0.910
	VaR reduction, 1%	0.814	0.834	0.617	0.811	0.835	0.834	0.759	0.753	0.798	0.809	0.852	0.666
	VaR reduction, 5%	0.815	0.817	0.519	0.819	0.838	0.832	0.765	0.794	0.821	0.764	0.826	0.725
16-day	*VR*	0.981	0.982	0.780	0.982	0.983	0.984	0.956	0.962	0.974	0.960	0.986	0.923
	VaR reduction, 1%	0.866	0.884	0.663	0.868	0.876	0.879	0.802	0.794	0.828	0.826	0.895	0.768
	VaR reduction, 5%	0.846	0.856	0.563	0.844	0.870	0.864	0.801	0.842	0.846	0.801	0.880	0.743
32-day	*VR*	0.984	0.985	0.783	0.985	0.985	0.987	0.957	0.974	0.980	0.958	0.988	0.934
	VaR reduction, 1%	0.915	0.922	0.722	0.914	0.923	0.924	0.823	0.906	0.927	0.822	0.927	0.810
	VaR reduction, 5%	0.845	0.854	0.523	0.861	0.846	0.865	0.769	0.838	0.864	0.763	0.879	0.654
64-day	*VR*	0.991	0.990	0.728	0.992	0.989	0.992	0.966	0.972	0.979	0.969	0.994	0.929
	VaR reduction, 1%	0.921	0.925	0.445	0.923	0.875	0.916	0.794	0.829	0.864	0.791	0.918	0.607
	VaR reduction, 5%	0.926	0.918	0.445	0.931	0.930	0.942	0.810	0.893	0.921	0.835	0.937	0.745
128-day	*VR*	0.991	0.990	0.689	0.991	0.984	0.990	0.955	0.954	0.965	0.969	0.995	0.942
	VaR reduction, 1%	0.899	0.898	0.408	0.894	0.844	0.894	0.724	0.768	0.822	0.884	0.934	0.734
	VaR reduction, 5%	0.932	0.893	0.375	0.932	0.830	0.883	0.850	0.715	0.798	0.858	0.935	0.788
256-day	*VR*	0.992	0.990	0.709	0.992	0.985	0.991	0.956	0.980	0.986	0.972	0.996	0.981
	VaR reduction, 1%	0.942	0.897	0.444	0.934	0.857	0.904	0.702	0.866	0.931	0.858	0.929	0.887
	VaR reduction, 5%	0.954	0.894	0.439	0.947	0.863	0.905	0.789	0.893	0.944	0.856	0.936	0.897

Notes: This table reports the out-of-sample hedging effectiveness of time varying hedge ratios forecasted by the bivariate CCC, DCC, AGDCC and BEKK GARCH models. Each model is respectively estimated by MLE on densities of normal, skew-Student and skew-Student with autoregressive skewness and kurtosis parameters. Subsample period for in-sample estimation of the models is Period I (July 1, 1982-June 30, 1997) and that for out-of-sample forecasting on hedge ratios and hedging effectiveness is Period II (July 1, 1997-September 28, 2007). Hedging effectiveness for 9 hedging horizons is reported. *VR*, variance reduction; *VaR reduction, 1%*, reduction of value at risk at 1% confidence level; *VaR reduction, 5%*, reduction of value at risk at 5% confidence level. *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters.

Table 4. Out-of-sample hedging effectiveness of time-varying hedge ratios during the GFC

		Time varying hedge ratios											
		CCC			*DCC*			*AGDCC*			*BEKK*		
Hedging horizon	*Hedging effectiveness*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*
1-day	*VR*	0.968	0.967	0.969	0.962	0.968	0.969	0.891	0.936	0.957	0.928	0.968	0.967
	VaR reduction, 1%	0.813	0.812	0.825	0.781	0.795	0.820	0.675	0.702	0.759	0.714	0.834	0.818
	VaR reduction, 5%	0.825	0.823	0.828	0.808	0.828	0.825	0.721	0.758	0.817	0.747	0.831	0.829
2-day	*VR*	0.979	0.978	0.979	0.972	0.978	0.980	0.901	0.944	0.968	0.937	0.979	0.978
	VaR reduction, 1%	0.873	0.867	0.874	0.843	0.873	0.890	0.707	0.768	0.828	0.763	0.872	0.873
	VaR reduction, 5%	0.872	0.858	0.874	0.833	0.851	0.873	0.761	0.733	0.827	0.749	0.873	0.873
4-day	*VR*	0.987	0.988	0.988	0.982	0.987	0.989	0.890	0.953	0.973	0.945	0.986	0.985
	VaR reduction, 1%	0.932	0.920	0.932	0.893	0.907	0.928	0.691	0.788	0.851	0.752	0.936	0.898
	VaR reduction, 5%	0.889	0.895	0.898	0.853	0.906	0.915	0.704	0.785	0.880	0.839	0.897	0.895
8-day	*VR*	0.989	0.989	0.989	0.983	0.988	0.990	0.848	0.946	0.970	0.937	0.988	0.987
	VaR reduction, 1%	0.914	0.915	0.921	0.893	0.911	0.924	0.660	0.744	0.841	0.735	0.932	0.910
	VaR reduction, 5%	0.884	0.873	0.886	0.815	0.852	0.875	0.643	0.767	0.814	0.698	0.892	0.885
16-day	*VR*	0.993	0.997	0.997	0.987	0.994	0.997	0.816	0.936	0.965	0.925	0.995	0.998
	VaR reduction, 1%	0.935	0.968	0.979	0.891	0.939	0.971	0.637	0.742	0.835	0.725	0.981	0.980
	VaR reduction, 5%	0.951	0.951	0.961	0.904	0.937	0.959	0.671	0.858	0.883	0.794	0.975	0.960
32-day	*VR*	0.994	0.997	0.998	0.989	0.995	0.998	0.910	0.926	0.983	0.921	0.997	0.999
	VaR reduction, 1%	0.934	0.954	0.971	0.898	0.938	0.965	0.757	0.729	0.876	0.721	0.988	0.991
	VaR reduction, 5%	0.944	0.957	0.972	0.913	0.938	0.966	0.762	0.754	0.887	0.734	0.986	0.987

Notes: This table reports the out-of-sample hedging effectiveness of time varying hedge ratios forecasted by the bivariate CCC, DCC, AGDCC and BEKK GARCH models. Each model is respectively estimated by MLE on densities of normal, skew-Student and skew-Student with autoregressive skewness and kurtosis parameters. Subsample period for in-sample estimation of the models is Period III (July 1, 1982 – September 28, 2007) and that for out-of-sample forecasting on hedge ratios and hedging effectiveness is Period IV (October 1, 2007-March 31, 2009). Hedging effectiveness for 6 hedging horizons is reported. *VR*, variance reduction; *VaR reduction, 1%*, reduction of value at risk at 1% confidence level; *VaR reduction, 5%*, reduction of value at risk at 5% confidence level. *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters.

Table 5. Descriptive statistics of time-varying hedge ratios

	$\hat{\beta}_t$			$\partial\hat{\beta}_t/\partial t$			$\partial\hat{\sigma}_t/\partial t$		
	Mean	*Median*	*Std. Dev.*	*Mean*	*Median*	*Std. Dev.*	*Mean*	*Median*	*Std. Dev.*
Panel A: Period II (1997.7.1-2007.9.28)									
CCC									
NORM	0.916	0.913	0.064	2.74×10^{-5}	3.19×10^{-5}	8.79×10^{-4}	-4.37×10^{-6}	1.26×10^{-6}	1.94×10^{-4}
SKST	0.914	0.919	0.041	3.95×10^{-5}	5.15×10^{-6}	5.51×10^{-4}	-3.70×10^{-7}	-6.15×10^{-6}	1.24×10^{-4}
SKST-ARSK	1.479	1.466	0.115	5.73×10^{-5}	1.59×10^{-4}	0.002	3.42×10^{-6}	1.52×10^{-6}	3.99×10^{-4}
DCC									
NORM	0.922	0.920	0.063	2.73×10^{-5}	3.08×10^{-5}	8.72×10^{-4}	-3.93×10^{-6}	-1.91×10^{-6}	1.88×10^{-4}
SKST	0.925	0.937	0.056	4.03×10^{-5}	7.88×10^{-5}	6.27×10^{-4}	-1.11×10^{-6}	-8.61×10^{-6}	2.27×10^{-4}
SKST-ARSK	0.928	0.933	0.046	3.64×10^{-5}	3.71×10^{-5}	5.91×10^{-4}	3.07×10^{-7}	-8.13×10^{-6}	1.57×10^{-4}
AGDCC									
NORM	0.821	0.813	0.092	2.65×10^{-5}	4.19×10^{-5}	9.61×10^{-4}	-1.20×10^{-6}	1.74×10^{-7}	2.06×10^{-4}
SKST	0.960	0.985	0.147	2.14×10^{-5}	1.23×10^{-4}	1.36×10^{-3}	6.15×10^{-6}	-9.40×10^{-7}	2.42×10^{-4}
SKST-ARSK	0.975	1.000	0.116	3.02×10^{-5}	8.17×10^{-5}	9.97×10^{-4}	7.76×10^{-6}	4.60×10^{-6}	1.80×10^{-4}
BEKK									
NORM	0.821	0.823	0.048	-1.93×10^{-5}	1.82×10^{-5}	4.87×10^{-4}	1.29×10^{-6}	1.05×10^{-5}	1.53×10^{-4}
SKST	0.929	0.929	2.33×10^{-4}	1.45×10^{-9}	1.14×10^{-7}	1.28×10^{-6}	-1.40×10^{-7}	-8.29×10^{-8}	1.88×10^{-6}
SKST-ARSK	0.829	0.876	0.159	-5.74×10^{-7}	-3.56×10^{-6}	4.37×10^{-4}	-7.07×10^{-7}	-7.84×10^{-6}	1.50×10^{-4}
Panel B: Period IV (2007.10.1-2009.3.31)									
CCC									
NORM	0.984	0.981	0.054	2.56×10^{-4}	2.25×10^{-4}	8.15×10^{-4}	1.40×10^{-4}	9.22×10^{-5}	1.89×10^{-4}
SKST	0.964	0.971	0.033	1.20×10^{-4}	1.81×10^{-4}	5.82×10^{-4}	1.06×10^{-4}	3.97×10^{-5}	1.12×10^{-4}
SKST-ARSK	0.989	0.995	0.037	1.45×10^{-4}	2.05×10^{-4}	6.46×10^{-4}	1.16×10^{-4}	4.62×10^{-5}	1.13×10^{-4}
DCC									
NORM	0.937	0.932	0.060	1.74×10^{-4}	1.77×10^{-4}	9.11×10^{-4}	9.95×10^{-5}	1.30×10^{-4}	1.76×10^{-4}
SKST	0.946	0.945	0.019	7.30×10^{-5}	1.42×10^{-5}	3.18×10^{-4}	9.82×10^{-5}	9.56×10^{-5}	9.70×10^{-5}
SKST-ARSK	0.981	0.982	0.015	5.23×10^{-5}	2.27×10^{-5}	2.98×10^{-4}	9.78×10^{-5}	8.15×10^{-5}	9.74×10^{-5}
AGDCC									
NORM	1.053	1.108	0.281	2.60×10^{-4}	1.51×10^{-4}	2.04×10^{-3}	1.29×10^{-4}	1.94×10^{-4}	3.64×10^{-4}
SKST	0.836	0.824	0.083	4.74×10^{-4}	3.18×10^{-4}	1.29×10^{-3}	2.50×10^{-4}	1.32×10^{-4}	3.89×10^{-4}
SKST-ARSK	0.987	1.012	0.104	1.44×10^{-4}	2.04×10^{-4}	9.98×10^{-4}	8.37×10^{-5}	9.85×10^{-5}	1.41×10^{-4}
BEKK									
NORM	0.807	0.802	0.054	3.96×10^{-5}	3.46×10^{-5}	8.33×10^{-4}	1.11×10^{-4}	1.19×10^{-4}	6.15×10^{-5}
SKST	0.999	1.000	0.035	3.93×10^{-5}	7.44×10^{-5}	6.69×10^{-4}	1.89×10^{-4}	1.26×10^{-4}	2.59×10^{-4}
SKST-ARSK	1.000	1.000	1.77×10^{-4}	-2.65×10^{-7}	-1.78×10^{-7}	3.35×10^{-6}	8.07×10^{-7}	4.11×10^{-7}	1.16×10^{-6}

Notes: This table reports the descriptive statistics of the time-varying hedge ratios derived by the bivariate GARCH models. Each model is respectively estimated by MLE on densities of normal, skew-Student and skew-Student with autoregressive skewness and kurtosis parameters. Model estimates are obtained using data of Period I (July 1, 1982-June 30, 1997) and Period III (July 1, 1982 – September 28, 2007), respectively. The series of hedge ratios are correspondingly calculated for Period II (July 1, 1997-September 28, 2007) and Period IV (October 1, 2007-March 31, 2009), respectively. $\hat{\beta}_t$ denotes the estimated conditional hedge ratio at date t. $\partial\hat{\beta}_t/\partial t$ denotes the partial first derivatives of $\hat{\beta}_t$ over time trend. It is derived by running the regression $\hat{\beta}_t = a + b * t + e_t$ where t is the time trend and the estimate of b equals $\partial\hat{\beta}_t/\partial t$. The series of $\partial\hat{\beta}_t/\partial t$ are obtained by a rolling window process on the regression with a window size of 100 observations and a step size of 1 observation. $\partial\hat{\sigma}_t/\partial t$ denotes the partial first derivatives of $\hat{\sigma}_t$ over the time trend where $\hat{\sigma}_t$ is the estimated standard deviation of $\hat{\beta}$. The series of $\hat{\sigma}_t$ are obtained by a rolling window process on $\hat{\beta}_t$ with a window size of 100 observations and a step size of 1 observation. The way to derive $\partial\vartheta/\partial t$ is similar to $\partial\hat{\beta}_t/\partial t$ where the series are obtained by a rolling window process. *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters. *Std. Dev.*, standard deviation.

Table 6. The effects of higher order moments on time-varying hedge ratios

Panel A: Period II (1997.7.1 – 2007.9.28)

	CCC				*DCC*			
IVs. / DVs.	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$
$Intercept$	6.724***	0.158	0.005	0.001	1.312***	0.074*	0.001	2.63×10^{-4}
	(0.5613)	(0.1228)	(0.0029)	(0.0010)	(0.1540)	(0.0410)	(0.0008)	(0.0003)
$skew_{1t}$	3.523***	0.347***	0.008***	0.002***	-2.541***	0.725***	0.009***	0.003***
	(0.4362)	(0.0918)	(0.0022)	(0.0007)	(0.3758)	(0.0992)	(0.0019)	(0.0007)
$skew_{2t}$	1.223***	-0.150***	-0.002***	-0.001***	2.097***	-0.794***	-0.010***	-0.004***
	(0.1556)	(0.0343)	(0.0008)	(0.0003)	(0.3212)	(0.0851)	(0.0016)	(0.0006)
v_t	-1.026***	0.004	-3.01×10^{-4}	1.18×10^{-4}	-0.109**	-0.013	-1.61×10^{-4}	-5.09×10^{-5}
	(0.1100)	(0.0243)	(0.0006)	(0.0002)	(0.0448)	(0.0119)	(0.0002)	(0.0001)

	AGDCC				*BEKK*			
IVs. / DVs.	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$
$Intercept$	-36.739***	0.405**	-0.012**	-0.003***	13.810***	-0.008	1.37×10^{-4}	-0.001***
	(1.1373)	(0.1622)	(0.0046)	(0.0009)	(0.3854)	(0.0472)	(0.0003)	(0.0002)
$skew_{1t}$	-45.142***	0.391**	0.003	-3.72×10^{-4}	89.450***	-1.895***	0.005**	0.001
	(1.1533)	(0.1639)	(0.0047)	(0.0009)	(2.1526)	(0.2634)	(0.0019)	(0.0009)
$skew_{2t}$	38.853***	-0.377**	-0.002	4.67×10^{-4}	13.468***	0.431	-0.006***	-0.005***
	(1.0768)	(0.1536)	(0.0044)	(0.0009)	(2.3755)	(0.2919)	(0.0021)	(0.0011)
v_t	11.877***	-0.094*	0.004***	0.001***	-4.209***	0.054***	5.45×10^{-5}	1.87×10^{-4}***
	(0.3589)	(0.0512)	(0.0015)	(0.0003)	(0.1225)	(0.0150)	(0.0001)	(0.0001)

Panel B: Period IV (2007.10.1 – 2009.3.31)

	CCC				*DCC*			
IVs. / DVs.	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$
$Intercept$	0.772*	0.033	1.89×10^{-4}	-0.001	0.616	-2.120	-0.013	-0.041
	(0.4497)	(0.1401)	(0.0034)	(0.0008)	(11.9887)	(8.4366)	(0.1528)	(0.0261)
$skew_{1t}$	0.608	-0.046	-0.001	0.002	2.285***	-0.310	5.25×10^{-5}	0.005**
	(0.7103)	(0.2220)	(0.0051)	(0.0011)	(0.8392)	(0.5932)	(0.0108)	(0.0021)
$skew_{2t}$	-0.274	0.078	0.002	-0.001	4.142	-2.096	-0.013	-0.019
	(0.8184)	(0.2562)	(0.0061)	(0.0013)	(7.0839)	(4.9613)	(0.0892)	(0.0152)
v_t	0.038	4.38×10^{-4}	4.78×10^{-5}	2.15×10^{-4}	0.196	1.043	0.007	0.020
	(0.0768)	(0.0239)	(0.0006)	(0.0001)	(5.8816)	(4.1390)	(0.0749)	(0.0128)

	AGDCC				*BEKK*			
IVs. / DVs.	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$	$\hat{\beta}_t$	$\hat{\sigma}^{\hat{\beta}_t}{}_t$	$\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$	$\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$
$Intercept$	-12.134***	0.090	1.56×10^{-4}	-2.78×10^{-4}	1.051***	0.007	0.001	1.88×10^{-4}*
	(0.8799)	(0.1409)	(0.0062)	(0.0007)	(0.0500)	(0.0249)	(0.0004)	(0.0001)
$skew_{1t}$	2.407*	0.009	-0.001	0.001	0.322	0.066	0.002	4.69×10^{-4}
	(1.2811)	(0.2038)	(0.0085)	(0.0009)	(0.3043)	(0.1565)	(0.0024)	(0.0006)
$skew_{2t}$	7.619***	0.045	-1.04×10^{-4}	-0.001	-0.168	-0.067	0.001	3.41×10^{-4}
	(1.5657)	(0.2507)	(0.0109)	(0.0011)	(0.5166)	(0.2657)	(0.0041)	(0.0010)
v_t	4.356***	0.005	1.64×10^{-4}	1.34×10^{-4}	-0.023	-0.003	-2.60×10^{-4}	-8.23×10^{-5}*
	(0.2922)	(0.0468)	(0.0021)	(0.0002)	(0.0221)	(0.0110)	(0.0002)	(4.29×10^{-5})

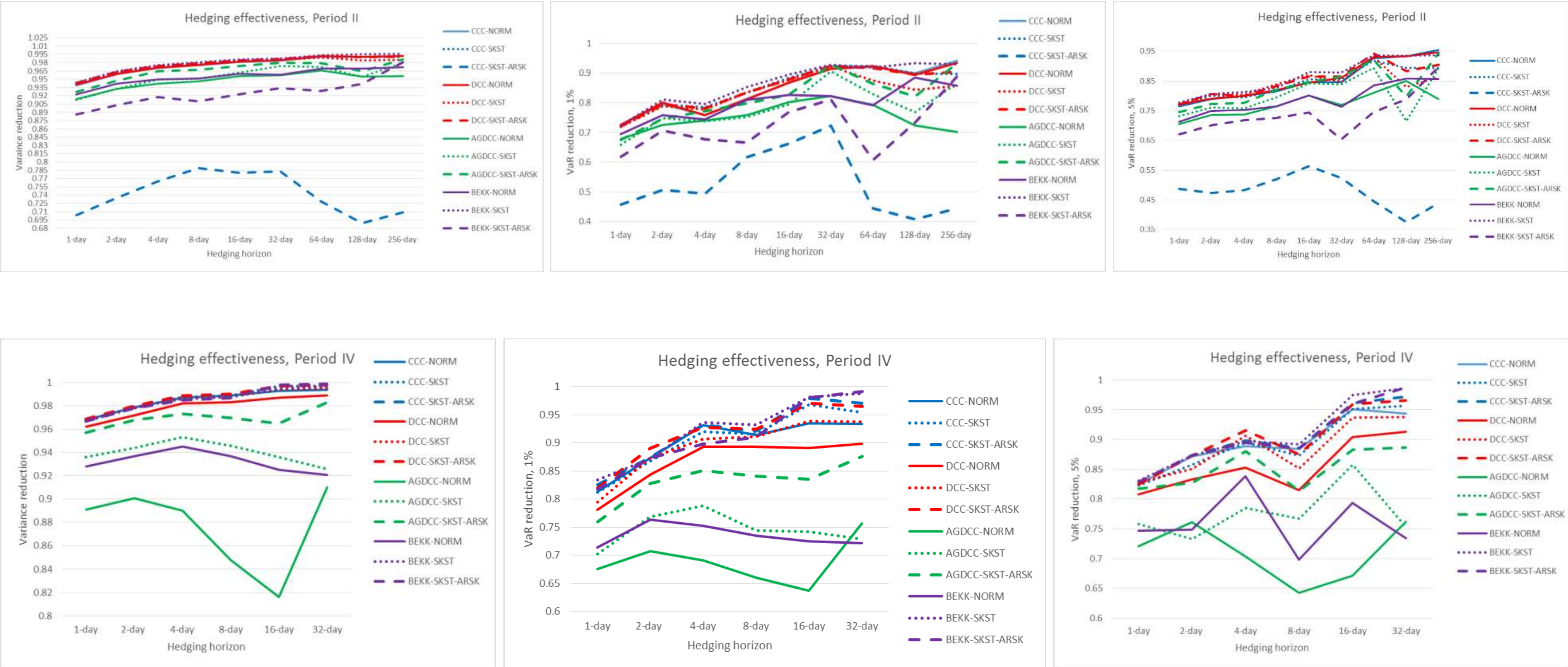
Notes: This table reports the results of the regression analysis on the effects of the conditional skewness and kurtosis on time-varying hedge ratios. Data of subsamples Period II (July 1, 1997-September 28, 2007) and Period IV (October 1, 2007-March 31, 2009) are used for regression. $\hat{\beta}_t$, the estimated conditional MV hedge ratios; $\hat{\sigma}^{\hat{\beta}_t}{}_t$, the series of standard deviation of $\hat{\beta}_t$obtained by the rolling window method with window size 100 observations and step size 1 observation. $\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$, the series of standard deviation of $\partial\hat{\beta}_t/\partial t$ where $\partial\hat{\beta}_t/\partial t$ is the partial first derivatives of $\hat{\beta}_t$ over time trend. It is derived by running the regression $\hat{\beta}_t = a + b * t + e_t$ where t is the time trend and the estimate of b equals $\partial\hat{\beta}_t/\partial t$. $\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$, the series of standard deviation of $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ where $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ is the partial first derivatives of $\hat{\sigma}^{\hat{\beta}_t}{}_t$ over the time trend. $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ is derived by running the regression $\hat{\sigma}^{\hat{\beta}_t}{}_t = c + d * t + e_t$ where t is the time trend and the estimate of d equals $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$. The series of $\partial\hat{\beta}_t/\partial t$, $\hat{\sigma}^{\partial\hat{\beta}_t/\partial t}{}_t$, $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ and $\hat{\sigma}^{\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t}{}_t$ are obtained by the rolling window method similar to $\hat{\sigma}^{\hat{\beta}_t}{}_t$. $skew_{it}$ $(i = 1{,}2)$ is the conditional skewness measure of the marginal densities of a bivariate skew-Student density where $skew_{it} = \xi^2{}_{it}(i = 1{,}2)$ and ξ_{it} is the exponential function of autoregressive skewness parameters defined by Eq. (8). The sign of $skew_{it}$ is determined by $ln(\xi_{it})$. v_t is the conditional degree of freedom of the bivariate skew-Student density. The time –varying process of v_t is defined by Eq. (9). *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *IVs*, independent variables; *DVs,* dependent variables. Figures in parenthesis are estimated standard errors. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 7. Out-of-sample hedging effectiveness of time-varying hedge ratios during the post-GFC period

		Time varying hedge ratios											
		CCC			*DCC*			*AGDCC*			*BEKK*		
Hedging horizon	*Hedging effectiveness*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*	*NORM*	*SKST*	*SKST-ARSK*
1-day	*VR*	0.963	0.966	0.967	0.929	0.937	0.934	0.867	0.925	0.927	0.939	0.966	0.967
	VaR reduction, 1%	0.831	0.838	0.836	0.743	0.764	0.754	0.635	0.752	0.747	0.758	0.839	0.845
	VaR reduction, 5%	0.817	0.824	0.824	0.749	0.757	0.748	0.754	0.743	0.758	0.774	0.827	0.828
2-day	*VR*	0.978	0.980	0.981	0.944	0.952	0.949	0.891	0.942	0.945	0.957	0.981	0.981
	VaR reduction, 1%	0.867	0.869	0.871	0.765	0.782	0.775	0.661	0.734	0.780	0.787	0.867	0.867
	VaR reduction, 5%	0.865	0.872	0.875	0.784	0.793	0.786	0.780	0.779	0.793	0.818	0.882	0.877
4-day	*VR*	0.985	0.987	0.988	0.952	0.960	0.956	0.892	0.948	0.951	0.967	0.988	0.987
	VaR reduction, 1%	0.892	0.886	0.885	0.789	0.785	0.792	0.628	0.773	0.769	0.806	0.893	0.886
	VaR reduction, 5%	0.894	0.894	0.900	0.796	0.798	0.787	0.761	0.769	0.793	0.834	0.897	0.893
8-day	*VR*	0.988	0.991	0.992	0.953	0.963	0.959	0.881	0.949	0.951	0.969	0.992	0.993
	VaR reduction, 1%	0.909	0.918	0.931	0.813	0.819	0.805	0.673	0.788	0.799	0.843	0.940	0.929
	VaR reduction, 5%	0.915	0.919	0.919	0.802	0.814	0.814	0.701	0.781	0.792	0.835	0.923	0.927
16-day	*VR*	0.990	0.993	0.994	0.954	0.964	0.959	0.889	0.950	0.955	0.972	0.995	0.995
	VaR reduction, 1%	0.896	0.917	0.939	0.763	0.803	0.784	0.655	0.728	0.742	0.788	0.940	0.942
	VaR reduction, 5%	0.934	0.944	0.938	0.811	0.816	0.814	0.779	0.811	0.844	0.873	0.944	0.937
32-day	*VR*	0.990	0.993	0.994	0.953	0.964	0.958	0.877	0.951	0.953	0.974	0.995	0.996
	VaR reduction, 1%	0.883	0.906	0.922	0.748	0.807	0.786	0.504	0.733	0.695	0.801	0.943	0.933
	VaR reduction, 5%	0.896	0.913	0.920	0.773	0.793	0.778	0.778	0.770	0.798	0.821	0.933	0.934
64-day	*VR*	0.989	0.993	0.995	0.948	0.960	0.953	0.873	0.947	0.952	0.968	0.996	0.999
	VaR reduction, 1%	0.906	0.910	0.920	0.791	0.815	0.802	0.383	0.757	0.724	0.785	0.951	0.951
	VaR reduction, 5%	0.903	0.922	0.931	0.730	0.817	0.791	0.758	0.728	0.755	0.783	0.965	0.978
128-day	*VR*	0.994	0.997	0.998	0.955	0.964	0.961	0.895	0.962	0.953	0.968	0.999	0.999
	VaR reduction, 1%	0.923	0.949	0.961	0.827	0.818	0.820	0.476	0.831	0.765	0.870	0.973	0.978
	VaR reduction, 5%	0.938	0.962	0.958	0.776	0.813	0.796	0.767	0.797	0.806	0.834	0.971	0.971
256-day	*VR*	0.995	0.997	0.999	0.960	0.964	0.962	0.948	0.964	0.968	0.967	0.998	0.999
	VaR reduction, 1%	0.968	0.974	0.999	0.860	0.812	0.812	0.848	0.873	0.959	0.930	0.964	0.987
	VaR reduction, 5%	0.974	0.975	0.999	0.866	0.811	0.809	0.821	0.885	0.968	0.924	0.952	0.977

Notes: This table reports the out-of-sample hedging effectiveness of time varying hedge ratios forecasted by the bivariate CCC, DCC, AGDCC and BEKK GARCH models. Each model is respectively estimated by MLE on densities of normal, skew-Student and skew-Student with autoregressive skewness and kurtosis parameters. Subsample period for in-sample estimation of the models is the whole sample period (July 1, 1982-March 31, 2009) and that for out-of-sample forecasting on hedge ratios and hedging effectiveness is a post-GFC period from April 1, 2009 to March 30, 2017. Hedging effectiveness for 9 hedging horizons is reported. *VR*, variance reduction; *VaR reduction, 1%*, reduction of value at risk at 1% confidence level; *VaR reduction, 5%*, reduction of value at risk at 5% confidence level. *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters.

Figure 1. Hedging effectiveness across horizons [16]



[16] Period II refers to subsample from July 1, 1997 to September 28, 2007. Period IV refers to subsample from October 1, 2007 to March 31, 2009. VaR reduction, 1% denotes reduction of value at risk at 1% confidence level. VaR reduction, 5% denotes reduction of value at risk at 5% confidence level. *CCC*, constant conditional correlation GARCH model; *DCC*, dynamic conditional correlation GARCH model; *AGDCC*, asymmetric generalised dynamic conditional correlation GARCH model; *BEKK*, BEKK GARCH model. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters.

Figure 2. Relations between hedge ratios and hedging effectiveness for different densities[17]

[17] Variance reduction is averaged out from hedging horizons for each bivariate GARCH model. Period II refers to subsample from July 1, 1997 to September 28, 2007. Period IV refers to subsample from October 1, 2007 to March 31, 2009. *NORM*, normal conditional density; *SKST*, skew-Student conditional density; *SKST-ARSK*, skew-Student conditional density with autoregressive skewness and kurtosis parameters. Smoothness of ratios is measured by standard deviation of $\partial\hat{\beta}_t/\partial t$ where $\hat{\beta}_t$ is estimated hedge ratios and *t* is time. Smoothness of standard deviation is measured by standard deviation of $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ where $\hat{\sigma}^{\hat{\beta}_t}{}_t$ is standard deviation of $\hat{\beta}_t$ and *t* is time.

Figure 3. Properties of hedge ratios across densities[18]

[18] Period II refers to subsample from July 1, 1997 to September 28, 2007. Period IV refers to subsample from October 1, 2007 to March 31, 2009. *CCC,* constant conditional correlation GARCH model*; DCC*, dynamic conditional correlation GARCH model*; AGDCC,* asymmetric generalised dynamic conditional correlation GARCH mode*l; BEKK,* BEKK GARCH model. Smoothness of ratios is measured by standard deviation of $\partial\hat{\beta}_t/\partial t$ where $\hat{\beta}_t$ is estimated hedge ratios and t is time. Smoothness of standard deviation is measured by standard deviation of $\partial\hat{\sigma}^{\hat{\beta}_t}{}_t/\partial t$ where $\hat{\sigma}^{\hat{\beta}_t}{}_t$ is standard deviation of $\hat{\beta}_t$ and t is time.