Forecasting the daily dynamic hedge ratios by GARCH models: evidence from the agricultural futures markets

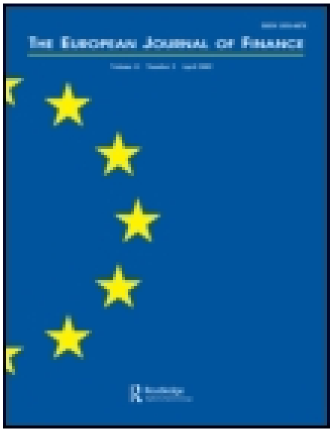
Article *in* European Journal of Finance · September 2013

DOI: 10.1080/1351847X.2013.794744

2 authors:

Yuanyuan Zhang
Southwestern University of Finance and Economics

Taufiq Choudhry
University of Southampton



The European Journal of Finance
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rejf20

Forecasting the daily dynamic hedge ratios by GARCH models: evidence from the agricultural futures markets

Yuanyuan Zhang[a] & Taufiq Choudhry[b]
[a] School of Securities and Futures, Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, People's Republic of China
[b] School of Management, University of Southampton, Southampton SO17 1BJ, UK
Published online: 23 May 2013.

The European Journal of Finance, 2015
Vol. 21, No. 4, 376–399, http://dx.doi.org/10.1080/1351847X.2013.794744

Forecasting the daily dynamic hedge ratios by GARCH models: evidence from the agricultural futures markets

Yuanyuan Zhang[a] and Taufiq Choudhry[b]*

[a]*School of Securities and Futures, Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, People's Republic of China;* [b]*School of Management, University of Southampton, Southampton SO17 1BJ, UK*

(*Received 5 December 2011; final version received 2 April 2013*)

This paper investigates the forecasting ability of six different generalized autoregressive conditional heteroskedasticity (GARCH) models; bivariate GARCH, BEKK GARCH, GARCH-X, BEKK-X, Q-GARCH and GARCH–GJR based on two different distributions (normal and student-*t*). Forecast errors based on four agricultural commodities' futures portfolio return forecasts (based on forecasted hedge ratio) are employed to evaluate the out-of-sample forecasting ability of the six GARCH models. The four commodities under investigation are two storable commodities: wheat and soybean, and two non-storable commodities: live cattle and live hogs. We apply the rolling forecasting method and the Model Confidence Set approach to evaluate and compare the forecasting ability of the six GARCH models. Our results show that the forecasting performances of the six GARCH models are different for storable and non-storable agricultural commodities. We find that the BEKK-type models perform the best in the case of storable products, while the asymmetric GARCH models dominate in the case of non-storable commodities. These results are regardless of the forecast horizon and residual distributions.

Keywords: forecasting; hedge ratio; GARCH; futures market; volatility

JEL Classification: G1; G15

1. Introduction

Lately, there has been much interest in the modelling of optimal hedge ratios (OHRs) and alternative hedging strategies when applied to the commodity and financial futures (see Choudhry 2009 for citations). It is now well known that the principal functions of futures markets are price discovery, hedging, speculation and risk sharing. Hedgers use these markets as a means to avoid the risk associated with adverse price changes in the related cash markets.

The hedge ratio is the number of futures contracts required to minimize the exposure of a unit worth position in the cash market. Previous research in the hedge ratio field has evaluated the relative effectiveness of alternative hedging strategies by examining the in-sample and out-of-sample performance of variance reductions of portfolios of returns in the cash and futures markets (see Lien, Tse, and Tsui 2002; Moon, Yu, and Hong 2009). What is lacking in the literature is the forecasting of the hedge ratio and the evaluation of the forecasting ability and accuracy of different models employed to forecast the hedge ratio. This paper makes an attempt to fulfil this gap in the literature. The forecasting of the hedge ratio is important for understanding the role of the futures markets in asset trading, programme trading, index arbitrage and the development of optimal hedging and trading strategies in portfolio management. The forecasting of OHRs

*Corresponding author. Email: t.choudhry@soton.ac.uk

helps the hedger to choose the most appropriate portfolio and allows for portfolio adjustments in dynamic hedging. Generally speaking, given that hedge ratios of various portfolios are predictable, an investor always prefers a portfolio with a lower financial capital to reach the maximum risk reduction.

This paper empirically investigates the behaviour of dynamic hedge ratios in four agricultural futures markets using alternative variants of the generalized autoregressive conditional heteroskedasticity (GARCH) models, and compares the forecasting performance of OHRs across those GARCH models.[1] More specifically, using daily data from the spot and futures markets of wheat, soybean, live cattle and live hogs, we estimate the time-varying hedge ratios and compare the forecasting performances of the six different GARCH models. This paper applies the rolling forecasting method and the Model Confidence Set (MCS) approach to compare the forecasting ability of the six GARCH models. Of the four commodities under study, wheat and soybeans are storable commodities, and live cattle and hogs are non-storable commodities.[2] According to Covey and Bessler (1995), Yang, Bessler, and Leatham (2001) and Yang and Awokuse (2003), commodity futures markets with different storability characteristics may perform in different manners. Thus, we also investigate the behaviour and forecasting of the hedge ratio based on the storability characteristics of the commodities. The six GARCH models under study are the standard bivariate GARCH, BEKK GARCH, GARCH-X, BEKK-X, asymmetric GARCH–GJR and Q-GARCH models.[3] To our knowledge, no previous study empirically investigates the out-of-sample forecasting by different GARCH models of time-varying hedge ratios for the agricultural futures market and then compares the forecasting performance of these models.[4] This is especially true taking into consideration when applying the rolling method of forecasting and MCS to compare the forecasting ability of different models. All this clearly indicates the substantial contribution this paper makes to the literature. Therefore, results presented in this paper have potentially important implications for academics, researchers, financial practitioners and policy-makers.

Choosing the appropriate forecasting method is an important issue. A large number of different GARCH models have been employed in previous research for forecasting purposes. Reasonable assumptions are essential for forecasting. For example, the relationship between the futures and cash prices for most financial assets is not assured, and if this presupposition is not tenable, forecasts of hedge ratios are not reliable. The time horizon also influences hedge ratio forecasts. Different hedging horizons might affect the forecasting accuracy for various forecasting methods (Chen, Lee, and Shrestha 2004). The longer the forecasting horizon, the more data are included and the more accurate the forecasts are. But the market environment might change, or unexpected events happen which renders these assumptions less reasonable over a longer time horizon. Activities of competitors can also affect forecasting accuracy. The more competition in the market, the more difficult it is to forecast hedge ratios. In a market that has great competition, competitors can change the course of future events after they make forecasts in order to make themselves more competitive, which then makes the forecasts invalid. In the active derivatives' market, decision-making depends on the quality of the forecasts, and hence forecasting of hedge ratios is important and meaningful for hedgers (Park and Antonovitz 1992).

The remainder of this paper is structured as follows. Section 2 describes the OHRs. Section 3 describes the six GARCH models based on normal distribution. Section 4 discusses the data and the basic statistics. Section 5 analyses the GARCH results. The MCS is discussed in Section 6. Section 7 presents and analyses the results of the forecast accuracy tests. Section 8 presents the conclusion.

2. Optimal hedge ratios

The following section describes the OHR, relying heavily on Choudhry (2009), Cecchetti, Cumby, and Figlewski (1988) and Baillie and Myers (1991). The returns on the portfolio of an investor trying to hedge some proportion of the cash position in a futures market can be represented by

$$r_t = r_t^{\mathrm{c}} - \beta_{t-1} r_t^{\mathrm{f}}, \tag{1}$$

where r_t is the return on holding the portfolio of cash and futures positions between $t-1$ and t; r_t^{c} is the return on holding the cash position for the same period; r_t^{f} is the return on holding the futures position for the same period; and β_{t-1} is the hedge ratio. The variance of the return on the hedged portfolio is given by

$$\mathrm{Var}\left(\frac{r_t}{\Omega_{t-1}}\right) = \mathrm{Var}\left(\frac{r_t^{\mathrm{c}}}{\Omega_{t-1}}\right) + \beta_{t-1}^2 \mathrm{Var}\left(\frac{r_t^{\mathrm{f}}}{\Omega_{t-1}}\right) - 2\beta_{t-1}\mathrm{Cov}\left(r_t^{\mathrm{c}}, \frac{r_t^{\mathrm{f}}}{\Omega_{t-1}}\right), \tag{2}$$

where Ω_{t-1} presents the information available over the last period. As indicated by Cecchetti, Cumby, and Figlewski (1988), the return on a hedged position will normally be exposed to the risk caused by unanticipated changes in the relative price between the position being hedged and the futures contract. This ‘basis risk’ ensures that no hedge ratio completely eliminates risk. The hedge ratio that minimizes risk may be obtained by setting the derivative of Equation (2) with respect to β equal to zero. The hedge ratio β_{t-1} can then be expressed as follows:

$$\beta_{t-1} = \mathrm{Cov}\left(r_t^{\mathrm{c}}, \frac{r_t^{\mathrm{f}}}{\Omega_{t-1}}\right) \mathrm{Var}\left(\frac{r_t^{\mathrm{f}}}{\Omega_{t-1}}\right). \tag{3}$$

The value of β_{t-1}, which minimizes the conditional variance of the hedged portfolio return, is the OHR (Baillie and Myers 1991). Usually, the value of the hedge ratio is less than unity; therefore, the hedge ratio that minimizes risk in the absence of basis risk turns out to be dominated by β when basis risk is taken into consideration.[5]

Time-varying OHRs can also be based on utility maximization. According to Myers (1991), under this scenario, an individual investor wants to determine the optimal allocation of initial wealth between two investment opportunities: purchase of a risky asset, and purchase of a risk-free asset. There is a futures market in the risky asset and the investor can, therefore, hedge by selling contracts which mature at or after the period. Using the von Neumann-Morgenstern utility function and a time-varying conditional covariance, Myers (1991) is able to show that the OHR is equal to the one presented by Equation (3). Myers (1991) defines the OHR as the proportion of the long-cash position which should be covered by futures selling. In this model, it is assumed that the OHR is preference-free, but the demand for the asset depends upon investor risk preferences, as well as on the probability distribution of asset price. Thus, the hedge ratio represented by Equation (3) can be based on the risk minimization or utility maximization.

3. GARCH models

In this study, the sample size of the four commodities is moderately large. Based on the central limit theorem, which states that the pattern of large samples approximately follows normal distribution statistically (Parks 1992), the error term (ε_t) in the mean equation of the GARCH models is assumed to be conditionally normal, distributed with mean 0 and conditional variance H_t.

However, financial and economic time series generally have leptokurtic fatter tail distributions than normal distributions for high-frequency data (see Baillie and DeGennaro 1990). Based on the evidence of non-normality, skewness and kurtosis of log level and log-difference of price for all four commodities, the conditional normal ε_t does not fit well with the observations. Additionally, the kurtosis and skewness show up in both the log level and log-difference differences of price for all the commodities. Hu and Kercheval (2010) emphasize the effect of the distribution of the return on portfolio management and they show the advantages of the student-t distribution for the case of daily data. The student-t distribution is equivalent to normal distribution when the degree of freedom approaches a sufficiently large number. In this study, a better choice to model ε_t is employing conditional student-t distribution which is able to capture fatter/thinner tails and higher/lower peaks than normal distributions by changing the degree of freedom. From this perspective, the student-t distribution is more useful than normal distribution when working in the presence of skewness and kurtosis.

Brooks (2008) argues that the non-normality of a time series may be due to the effect of extreme value or high heteroscedasticity when the sample size is large. That is, the rejection of normality in the Jarque-Bera (JB) test does not necessarily imply non-normality for a sufficiently large sample. Hence, forecasting based on both the normality and student-t distribution is reliable. In this paper, we apply both distributions: normal and student-t. In order to conserve space, we only describe the GARCH models with the normal distribution.

3.1 *Bivariate GARCH model*

The time-varying hedge ratios are estimated using the following six different variants of bivariate GARCH models: standard GARCH, GARCH–BEKK, GARCH-X, BEKK–GARCH-X, Q-GARCH and GARCH–GJR.[6] The following bivariate GARCH(p, q) model is applied to returns from the cash and futures markets:

$$y_t = \mu + \delta(\mathrm{z}_{t-1}) + \varepsilon_t, \tag{4}$$

$$\varepsilon_t / \Omega_{t-1} \sim N(0, H_t), \tag{5}$$

$$\mathrm{vech}(H_t) = C + \sum_{i=1}^{p} A_i \ \mathrm{vech}(\varepsilon_{t-i})^2 + \sum_{j=1}^{q} B_j \ \mathrm{vech}(H_{t-j}), \tag{6}$$

where $y_t = (r_t^{\mathrm{c}}, r_t^{\mathrm{f}})$ is a (2×1) vector containing returns from the cash and futures markets, H_t is a (2×2) conditional covariance matrix, C is the (3×1) parameter vector of constant, A_i and B_j are (3×3) parameter matrices, and vech is the column stacking operator that stacks the lower triangular portions of a symmetric matrix. The term (z_t) is the error-correction term from the cointegration relationship between the log of the cash price and the log of the futures price. The error-correction term z_t is equal to $(\mathrm{c}_t - \gamma \mathrm{f}_t)$, where c_t is the log of the cash price index, f_t is the log of the futures prices index and γ is the cointegration coefficient. The error-correction term, which represents the short-run deviations from the long-run cointegrated relationship, has important predictive powers for the conditional mean of the cointegrated series (Engle and Yoo 1987). The long-run relationship between the commodities cash price and the futures price is determined by means of the Engle and Granger (1987) cointegration test.[7] As indicated by Engle and Granger (1987), if two variables are cointegrated, then their bivariate time-series models should include the error-correction term. The error term imposes the long-run relationship between the spot and futures prices, and the GARCH models permits the second moments of the distribution to change through time. Yang, Bessler, and Leatham (2001) claim that prevalent cointegration between cash and futures prices on commodity markets suggests that cointegration should be incorporated into commodity hedging decisions.[8] Even when the GARCH effect is considered, allowance for the existence of cointegration is argued to be an indispensable component when comparing ex-post performance of various hedging strategies. Lee (1994) shows that the GARCH-X model, which incorporates cointegration, provides a better fit for daily-, short- and long-term monthly exchanges and may boost the prediction power of the GARCH model. A significant and positive coefficient (δ) on the error term (Equation (4)) implies that an increase in short-run deviations raises the log difference of cash and/or future prices. The opposite is true if the error term coefficient is negative and significant.

To make the estimation of GARCH(p, q) amenable, Engle and Kroner (1995) have suggested various restrictions be imposed on the parameters of the A_i and B_j matrices. It is necessary to impose restrictions on the parameters to ensure the conditional variance of the individual variable is positive and definite, which can be difficult to do in practice. To resolve these difficulties, researchers have proposed various simplifying assumptions to reduce the number of unknown coefficients in the conditional variance matrix to a manageable level (see Bollerslev 1988 and Bollerslev 1990, for example). As a typically parsimonious representation, the diagonal GARCH model suggested by Bollerslev (1988) assumes that the A and B matrices are diagonal. Thus, replacing all a_{ij} and b_{ij} ($i \neq j$) in A and B with zero. The following equations represent a diagonal vech bivariate GARCH(1,1) conditional variance equation(s):

$$H_{11,t} = C_1 + A_{11}(\varepsilon_{1,t-1})^2 + B_{11}(H_{11,t-1}), \tag{7a}$$

$$H_{12,t} = C_2 + A_{22}(\varepsilon_{1,t-1}, \varepsilon_{2,t-1}) + B_{22}(H_{12,t-1}), \tag{7b}$$

$$H_{22} = C_3 + A_{33}(\varepsilon_{2,t-1})^2 + B_{33}(H_{22,t-1}). \tag{7c}$$

In the bivariate GARCH(1,1) model, the diagonal vech parameterization involves nine conditional variance parameters.

Using the bivariate GARCH model, the time-varying hedge ratio can be computed as follows:

$$h_t^* = \frac{\hat{H}_{12,t}}{\hat{H}_{22,t}}, \tag{8}$$

where $\hat{H}_{12,t}$ is the estimated conditional covariance between the cash and futures returns, and $\hat{H}_{22,t}$ is the estimated conditional variance of futures returns. Since the conditional covariance is time-varying, the optimal hedge would be time-varying too.

3.2 *Bivariate GARCH–BEKK model*

In the BEKK model, as suggested by Engle and Kroner (1995), the conditional covariance matrix is parameterized to

$$\mathrm{vech}(H_t) = C'C + \sum_{k=1}^{k}\sum_{i=1}^{p} A'_{ki}\varepsilon_{t-i}\varepsilon'_{t-1}A_{ki} + \sum_{k=1}^{k}\sum_{j=1}^{q} B'_{kj}H_{t-j}B_{kj}. \tag{9}$$

Equations (4) and (5) also apply to the BEKK model and are defined as before. In Equation (9) A_{ki}, $i = 1, \ldots, q$, $k = 1, \ldots, k$ and B_{kj} $j = 1, \ldots, q$, $k = 1, \ldots, k$ are $N \times N$ matrices. The GARCH–BEKK model is sufficiently general so that it guarantees the conditional covariance matrix, H_t to be positive definite, and renders significant parameter reductions in the estimation. For example, a bivariate BEKK GARCH(1,1) parameterization requires estimates of only 11 parameters in the conditional variance–covariance structure. The time-varying hedge ratio from the BEKK is again represented by Equation (8).[9]

3.3 *Bivariate GARCH-X model*

The GARCH-X model is an extension of the standard GARCH model as it incorporates the square of the error-correction term (z_t) in the conditional covariance matrix. Lee (1994) contends that as short-run deviations from the long-run relationship between the cash and futures prices may affect the conditional variance and covariance, then they will also influence the time-varying OHR. Lien (2004) finds that ignoring cointegration tends to underestimate the value of the hedge ratio. In the GARCH-X model, conditional heteroscedasticity may be modelled as a function of a lagged squared error-correction term in addition to the autoregressive moving average terms in the variance/covariance equations:

$$\mathrm{vech}(H_t) = C + \sum_{i=1}^{p} A_i \,\mathrm{vech}(\varepsilon_{t-i})^2 + \sum_{j=1}^{q} B_j \,\mathrm{vech}(H_{t-j}) + \sum_{k=1}^{k} D_k \,\mathrm{vech}(z_{t-1})^2. \qquad (10)$$

A significant positive effect may imply that the further the series deviate from each other in the short run, the harder they are to predict. Equations (4) and (5) also apply in GARCH-X. The time-varying hedge ratio is again represented in Equation (8). Sultan and Hasan (2008) claim that the GARCH-X and GARCH with error-correction models reduce portfolio risk with a similar performance to dynamic hedging.

3.4 *Bivariate BEKK GARCH-X*

A similar extension can be made to the standard BEKK GARCH linked to an error-correction model of cointegrated series on the second moment of the bivariate distributions of the variables. Such a model is known as the BEKK GARCH-X. The formulation of the BEKK GARCH(1,1)-X model is given by

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B + D'D^2(z_{t-1})^2. \qquad (11)$$

Equations (4) and (5) apply to this model also and the variables are as defined in the BEKK GARCH section. Once again, z_t is the error term from the cointegration tests between the cash and futures prices, and D is the (1×2) matrix of coefficients. The analysis of the size and sign on the error term coefficients are the same as described in the bivariate GARCH-X section. The time-varying hedge ratio from the BEKK GARCH-X is represented in Equation (8), but given the effect of the error term it should differ from the standard BEKK hedge ratio.

3.5 *Bivariate GARCH–GJR*

Along with the leptokurtic distribution of asset returns data, empirical research has shown a negative correlation between current returns and future volatility (Black 1976; Christie 1982). This negative effect of current returns on future variance is sometimes called the leverage effect (Bollerslev, Chou, and Kroner 1992). In the linear (symmetric) GARCH model, the conditional variance is only linked to past conditional variances and squared innovations (ε_{t-1}), and hence the sign of return plays no role in affecting volatilities (Bollerslev, Chou, and Kroner 1992). Glosten, Jagannathan, and Runkle (1993) provide a modification to the GARCH model that allows positive and negative innovations to returns to have different impacts on conditional variance.[10] Glosten, Jagannathan, and Runkle (1993) suggest that the asymmetry effect can be captured simply by incorporating a dummy variable in the original univariate GARCH.

$$H_t = C + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i}^2 I_{t-i} + \sum_{j=1}^{p} \beta_j H_{t-j}, \tag{12}$$

where I_{t-i} is a dummy variable and $I_{t-i} = 1$ if $\varepsilon_{t-1} < 0$; otherwise, $I_{t-1} = 0$. Thus, the ARCH coefficient in a GARCH–GJR model switches between $\alpha + \gamma$ and α, depending on whether the lagged error term is positive or negative. If the estimated value of γ is non-zero, then the leverage effect is present in the data. In other words, the dummy variable is an indicator of the leverage effect in the GARCH model. Notice non-negative constrains should be imposed on parameters to ensure non-negativity of the conditional variance, which will be $C > 0$, $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \gamma \geq 0$.

The GJR model can also be applied to the bivariate version of the GARCH model to capture the conditional variance and covariance. Once again, a parsimonious representation can be obtained by imposing a diagonal restriction on parameter matrices. Consequently, the diagonal bivariate GJR–GARCH(1,1) can be presented by the following equations:

$$H_{11,t} = c_1 + \alpha_1 \varepsilon_{1,t-1}^2 + \beta_1 H_{11,t-1} + \gamma_1 \varepsilon_{1,t-1}^2 I_{t-1}, \tag{13a}$$

$$H_{12,t} = c_2 + \alpha_2 (\varepsilon_{1,t-1} \varepsilon_{2,t-1}) + \beta_2 H_{12,t-1}, \tag{13b}$$

$$H_{22,t} = c_3 + \alpha_3 \varepsilon_{2,t-1}^2 + \beta_3 H_{22,t-1} + \gamma_3 \varepsilon_{2,t-1}^2 I_{t-1}, \tag{13c}$$

where I is the same dummy variable as in the univariate model, α_i, β_i and c_i $(i = 1, 2, 3)$ are parameters of the ARCH term, the GARCH term and the long-term average value. Equations (4) and (5) also apply here. The time-varying hedge ratio based on the GARCH–GJR model is again expressed in Equation (8). Alberg, Shalit, and Yosef (2008) claim that asymmetric GARCH model generally improves forecasting performance. This is particularly true with fat-tail density.

3.6 *Bivariate quadratic GARCH*

As stated earlier more than one version of the GARCH model is capable of capturing the leverage effect. The Q-GARCH also captures the asymmetric effect. This model incorporates the past error as a proxy of the asymmetric effect of negative and positive shocks on conditional variance.

The Q-GARCH (1, 1) is

$$H_t = c + a_1 \varepsilon_{t-1}^2 + b_1 H_{t-1} + d_1 \varepsilon_{t-1}, \tag{14}$$

where $d_1 \varepsilon_{t-1}$ measure the asymmetry, and the generalized Q-GARCH (p, q) is written as

$$H_t = c + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j H_{t-j} + \sum_{i=1}^{q} d_i \varepsilon_{t-i} + 2 \sum_{i=1}^{q} \sum_{j=i+1}^{p} e_{il}\, \varepsilon_{t-i} \varepsilon_{t-j}. \tag{15}$$

In the higher order Q-GARCH model, the term $\sum_{i=1}^{q} \sum_{j=i+1}^{p} e_{il} \varepsilon_{t-i} \varepsilon_{t-j}$ shows that the cross product of the previous error as well as the $\sum_{i=1}^{q} d_i\, \varepsilon_{t-i}$ error term has an impact on the current conditional variance. Similar to GARCH–GJR, the Q-GARCH also ensures the positive conditional covariance. Yet, it is widely accepted that all parameters are assumed to be non-negative to avoid complicated statistical problems.[11]

The methodology used to obtain the optimal forecast of the conditional variance of a time series from a GARCH model is the same as that used to obtain the optimal forecast of the conditional mean.[12] Although many GARCH specifications forecast the conditional variance in a similar way, the forecast function for some extensions of GARCH will be more difficult to derive. For instance, extra forecasts of the dummy variable I is necessary in the GARCH–GJR model. However, following the same framework, it is straightforward to generate forecasts of the conditional variance and covariance, and thus the time-varying hedge ratio using bivariate GARCH models.

Forecasting can be done based on a recursive window or a rolling window. In this paper, we apply the (monthly) rolling window method. A rolling window is one where the length of the in-sample period applied to estimate the model is fixed, so that the start date and end date successfully increase by one observation. This forecasting method estimates and keeps parameters updated for each rolling window. For one-step-ahead forecasting of y, the observations from the first observation to the nth observation are applied for forecasting the y_{n+1}; and to forecast the y_{n+2}, the sample applied is from the second observations to the $n + 1$ observation (Brockwell and Davis 2002). In this manner, the fixed rolling window is updated by one observation at the start and at the end each time to forecast one step ahead. In the rolling window method, not a single optimized estimated parameter is present, but a group of corresponding parameters are involved. In other words, this method avoids the stability (parameter drift) problem of estimated parameters which are updated for each fixed rolling window estimates. The rolling scheme has constant power while the recursive forecasting method does not (Cheung, Chinn, and Pascual 2003). In the modern financial market, business conditions change rapidly. The rolling forecasting method keeps updating estimated parameters to respond to rapid market changes, and thus investors can adjust investment position and budget in time (Lalli 2011). The effectiveness of rolling forecasting makes it a more attractive approach than recursive forecasting.

The methodology of monthly rolling forecasting of time-varying hedge ratios will be carried out in three steps. For example, step 1 for the 1-year forecasting horizon (1 January 2007–1 January 2008 for non-storable commodities) is to estimate the parameters of all the six GARCH models from 1 January 1980 to 31 December 2006 for forecasting the hedge ratio for the following one month period (1 January 2007–31 January 2007). The whole process is then repeated by rolling the sample size by one month starting from 1 February 1980 to 31 January 2007 to forecast for the next month (1 February 2007–28 February 2007). This process is continuously repeated until the hedge ratio has been forecasted for each month during 2007–2008.

In the second step, we predict returns over the forecast horizon based on the forecasted time-varying hedge ratios. In the third and final steps, the empirical results of the performance of various GARCH models will be produced on the basis of hypothesis tests, looking at whether the estimate is significantly different from the real value, and which GARCH model provides the 'BEST' forecast in terms of forecasting accuracy with evidence from the MCSs approach.

4. Data and basic statistics

Commodity prices can be more volatile and unstable compared with exchange rates and interest rates in some periods due to disturbances in demand and supply (Kroner, Kneafsey, and Claessens 2006). As stated earlier, daily log difference of the cash (spot) and futures prices of wheat, soybean, live cattle and live hogs are used in the empirical tests. Live cattle and hogs are non-storable commodities while the remaining two are storable. The storable and non-storable products should provide different results based on differences in terms of storability, volatility of price basis and demand–supply relationship. Though the basis has to be zero at the delivery day, the cash and futures prices for non-storable products move relatively independently without maximum or minimum limit.

The length of the data is different for the storable and non-storable commodities. For the storable commodities, the data range from 1 January 1980 to 23 June 2006 and for the non-storable commodities from 1 January 1980 to 14 January 2008.[13] The futures price for a storable asset is considered equal to the asset's cash price plus the asset's cost-of-carry (Covey and Bessler 1995). A futures price of a non-storable asset is considered equal to the futures market's forecast of the asset's cash price which will be obtained during the delivery period of a particular futures contract at one of its specified delivery locations (Covey and Bessler 1995). All futures price indices are continuous series.[14] The soybean and wheat futures prices are from the Chicago Board of Trade (CBOT). The wheat cash prices are from the CBOT, the soybeans cash price is the price of soybeans in Southeast Iowa, the live cattle cash price is taken from the Commodity research bureau and ICX index, and the hog cash prices are taken from IHX hog index. All data are obtained from *Global Financial Data*.

Basic statistics of log level and log difference of cash and futures prices show excess kurtosis in most of the log of cash and futures prices. The leptokurtosis implies the presence of a fatter tail and higher peak value around the mean than normal distribution for all commodities. In terms of skewness, all price levels show a longer tail on the left-hand side except the live cattle, which has a longer tail on the other side. The log-futures price pattern of live hogs is relatively symmetric as normal distribution. The log difference of cash prices and futures prices show similar results. The significant JB test indicates non-normality for all commodities. These results are quite standard for cash and futures prices in level and log difference and thus are not presented. They are available on request.

The unit root test results indicate that log price levels of all four commodities indicate one unit root at the 1% significant level at all lags. These results are not provided to save space but are available on request. The results of the cointegration test based on the Engle–Granger two-step method shows significant cointegration between cash and futures price indices for all commodities. These results are not provided to save space but are available on request. The cointegration results are required in the application of all the six GARCH models.

5. GARCH results and hedge ratio statistics

The results from the standard bivariate GARCH(1,1), BEKK(1,1), GARCH-X(1,1), BEKK-X(1,1), Q-GARCH(1,1) and GARCH–GJR(1,1) models are quite standard.[15] As an example, we provide results from all the six GARCH models using normal distribution estimates for the live cattle in Table 1. We present the estimates which yield the maximum likelihood value among the 12 monthly estimates during the 1-year rolling procedure. Results show a significant effect of the error-correction term on the mean returns for both the cash and the futures returns. The ARCH and GARCH effects are significant in most of the GARCH models. The GARCH-X model shows the significant effect of the cointegration between log-cash and log-futures prices for the live cattle on volatility. However, the BEKK-X model fails to show significant effect of cointegration. Some evidence of the leverage effect is indicated by the two asymmetric models, GJR and Q-GARCH.

Table 1. The maximum estimated coefficients from the six bivariate GARCH models with normal distribution for live cattle (monthly rolling for 1-year).

$Y_t = \mu + \delta z_{t-1} + \varepsilon_t, \varepsilon_t \vert \Omega_t \sim N(0, H_t)$						
Variable	GARCH	BEKK	GARCH-X	BEKK-X	GJR	Q-GARCH
μ_1	0.00093	0.00064	0.00088	0.00093	−0.00011	0.00046
	(0.000)	(0.000)	(0.000)	(0.000)	(0.796)	(0.000)
δ_1	−0.00564	0.26841	−0.00642	−0.00573	−0.00547	−0.00828
	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.000)
μ_2	0.00015	0.00008	0.00015	0.00008	−0.00006	0.00001
	(0.302)	(0.580)	(0.283)	(0.586)	(0.694)	(0.936)
δ_2	0.00513	0.01924	0.00490	0.00475	0.00598	0.00419
	(0.000)	(0.101)	(0.000)	(0.002)	(0.010)	(0.004)
c_{11}	0.00005	0.00327	0.00005	0.00675	0.00005	0.00005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
a_{11}	1.20125	0.75646	1.19390	0.20827	0.08594	(1.0505)
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
b_{11}	0.04073	0.85186	0.04814	1.09885	(0.04180)	(0.05253)
	(0.000)	(0.000)	(0.000)	(0.000)	[0.321]	[0.000]
d_{11}			−0.00056	0.00001		−0.00781*
			(0.000)	(1.000)		(0.000)
γ_1					1.97826	
					(0.233)	
c_{22}	0.00005	0.00049	0.00004	0.00001	0.00001	0.00001
	(0.000)	(0.000)	(0.000)	(0.999)	(0.110)	(0.000)
a_{22}	0.05950	0.99617	0.04540	0.99535	0.01677	0.00727
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
b_{22}	0.55274	0.06777	0.67866	0.07005	0.98945	0.98156
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_{22}			−0.00021	−0.00001		0.00001*
			(0.000)	(0.999)		(0.986)
γ_2					−0.01839	
					(0.000)	
c_{12}	0.00002	0.00040	0.00001	0.00076	0.00001	0.00002
	(0.000)	(0.000)	(0.149)	(0.000)	(0.233)	(0.000)
a_{12}	−0.00271		0.00229		0.00291	−0.00300
	(0.030)		(0.167)		(0.095)	(0.026)
b_{12}	−0.98599		0.98808		0.98800	−0.98110
	(0.000)		(0.000)		(0.000)	(0.000)
d_{12}			−0.00001	−0.00001		
			(0.096)	(0.999)		
LLF	40069.16	40311.53	40164.04	40228.14	40451.56	40388.38

Note: The d_{11} and d_{22} for Q-GARCH model represent the impacts of asymmetry.

Similar results and conclusions are reached using both the normal and student-*t* distributions over 1- and 2-year forecasting horizons for all commodities. These results are available on request.

Figures 1–4 present the estimated conditional variances of the cash and futures returns and the covariances between them. In order to save space, we only present graphs for wheat and live cattle using all the six models for both distributions for the 1-year rolling forecasting. The most visible feature of the graphs is the difference between the storable wheat and the non-storable live cattle conditional variance and covariance. Other figures are available on request.

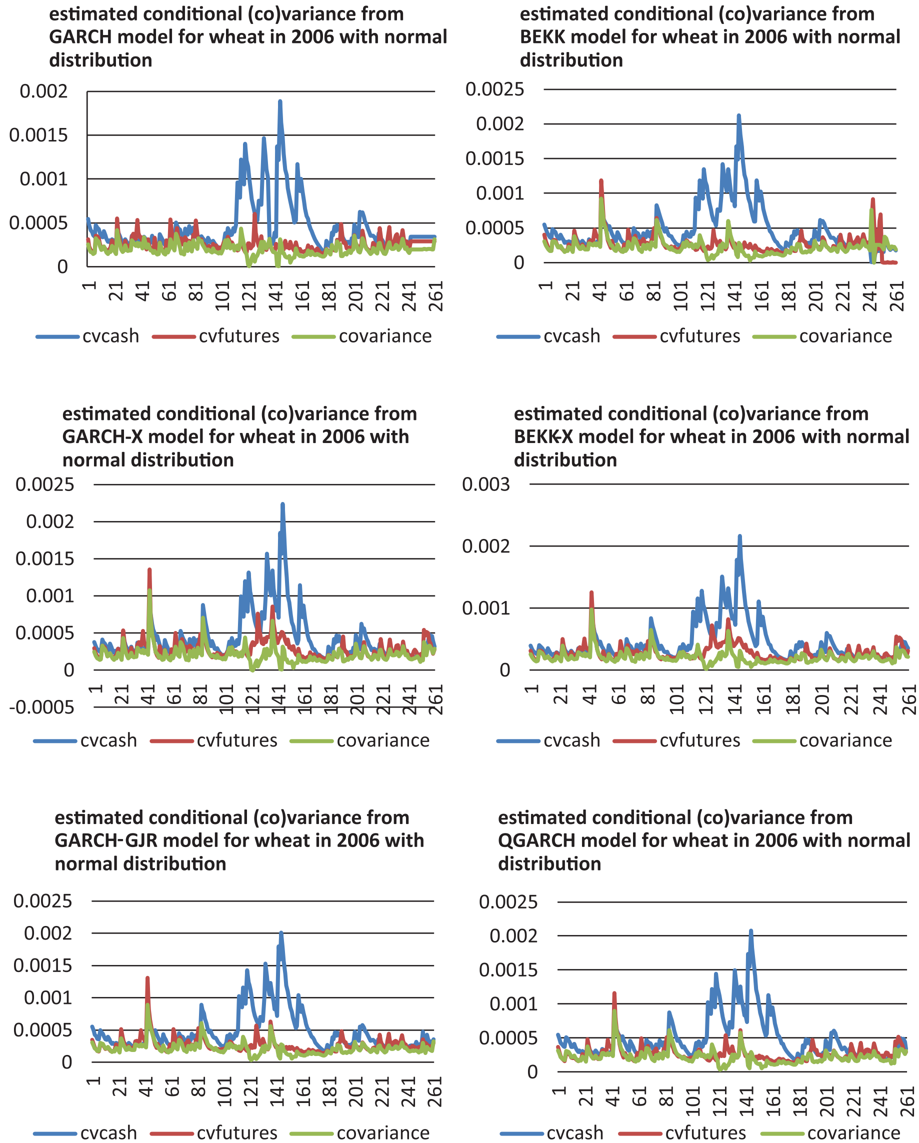


Figure 1. Patterns of conditional (co)variance of last estimation from the six GARCH models based on normal distribution for wheat (1-year).

Table 2 presents the basic statistics of the estimated hedge ratio (OHR) with normal distribution for all commodities over the 1-year horizon. The average OHR for storable products (wheat and soybean) ranges from 0.64 to 0.96, while the hedge ratio is as low as 0.016 for non-storable products (live cattle and live hogs).[16] The volatility of OHR for non-storable produces is smaller than those of storable commodities. Based on the JB test, we find that nearly all the OHR series for the four commodities are non-normally distributed at the 5% significance level; however, the distribution of the OHR series from BEKK-X, GJR and Q-GARCH models for live hogs is not

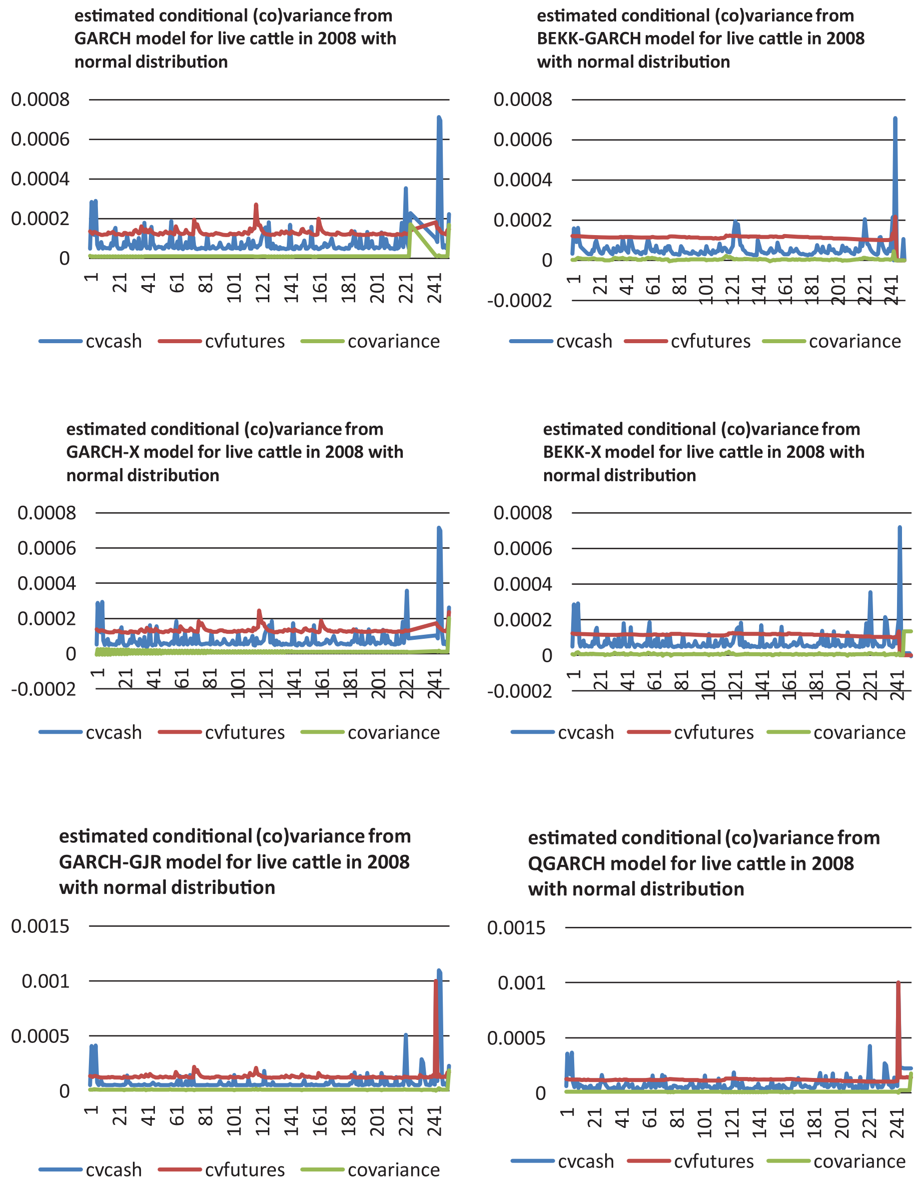


Figure 2. Patterns of conditional (co)variance of last estimation from the six GARCH models based on normal distribution for live cattle (1-year).

significantly different from the normal distribution. For non-normally distributed OHR series, there is no pattern to the significant skewness or kurtosis. Basic statistics from the estimated OHR from other distributions and horizons are similar and are available on request.

Although the point estimation of the hedge ratio generated by the GARCH model is a moderate proxy for the actual hedge ratio value, it is not an appropriate scale to measure a hedge ratio series forecasted with time variation. As a result, the evaluation of forecast accuracy based on

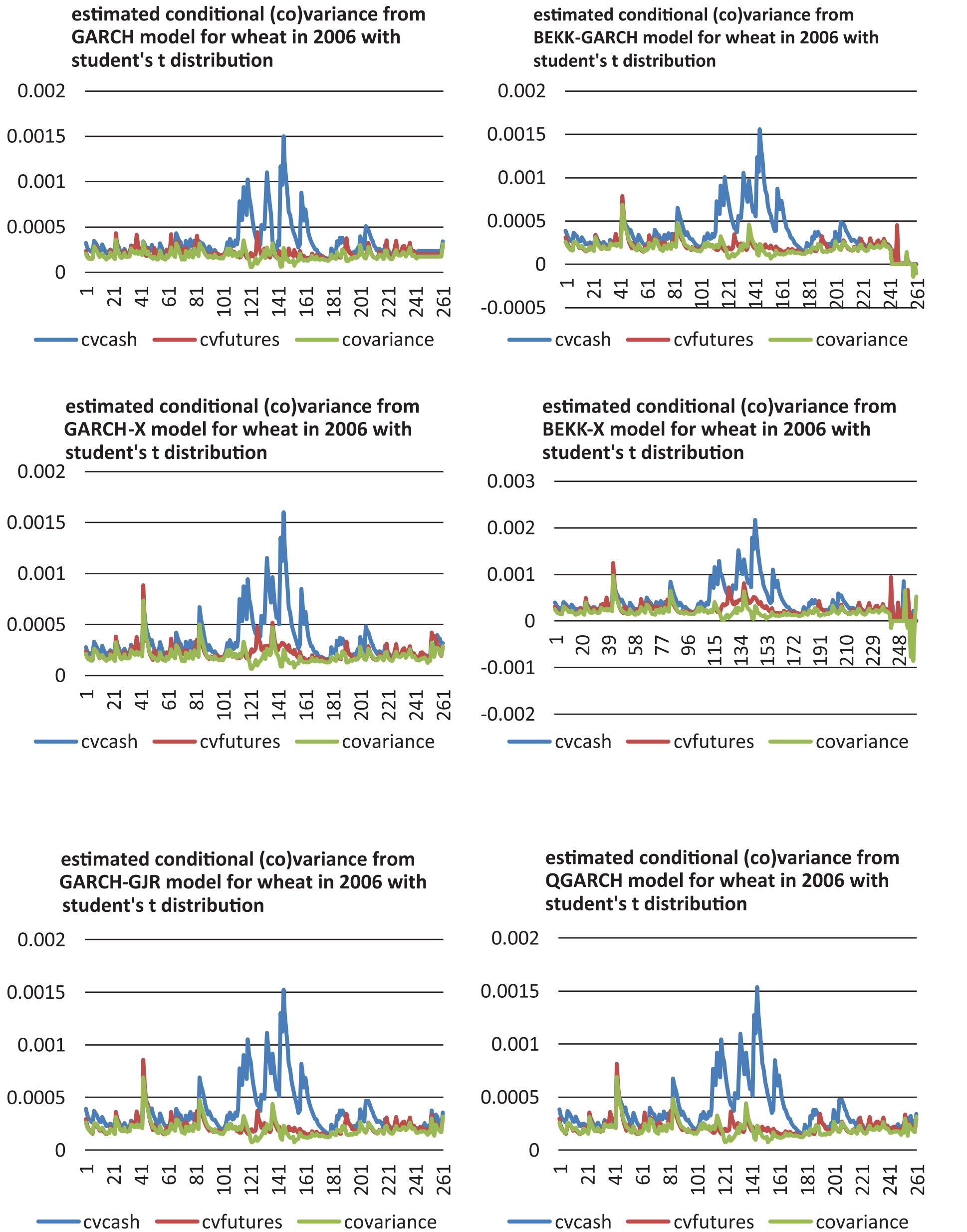


Figure 3. Patterns of conditional (co)variance of last estimation from the six GARCH models based on student-*t* distribution for wheat (1-year).

comparing conditional hedge ratios estimated and forecasted by the same approach cannot provide compelling evidence of the worth of each individual approach. To assess predictive performance, a logical extension is to examine out-of-sample returns. The evaluation of forecast accuracy is thus conducted by forecasting out-of-sample returns of portfolios implied by the computed hedge ratios. The portfolios are constructed as $(r_t^{\mathrm{c}} - \beta_t^* r_t^{\mathrm{f}})$, where r_t^{c} is the log difference of the cash

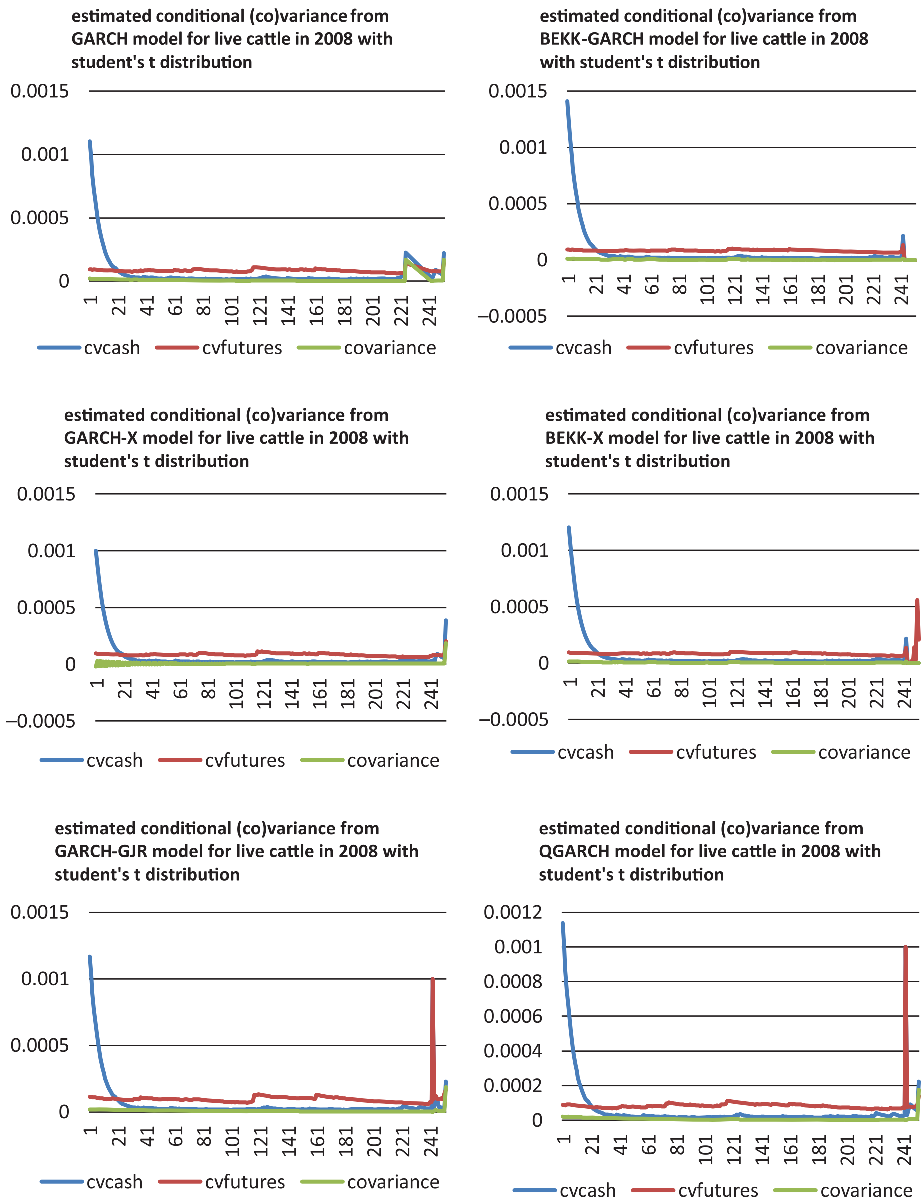


Figure 4. Patterns of conditional (co)variance of last estimation from the six GARCH models based on student-*t* distribution for live cattle (1-year).

(spot) prices, r_t^f is the log difference of the futures prices and β_t^* is the estimated OHR. With the out-of-sample forecasts of time-varying hedge ratios, the out-of-sample forecasts of returns based on the portfolio above can be easily calculated, in which the cash return and futures return are actual returns observed. The relative accuracy of time-varying hedge ratio forecasts can then be assessed by comparing the return forecasts with the actual returns. In this way, the issue of a missing benchmark can be settled.

Table 2. Basic statistics for 1-year forecasted hedge ratio for wheat, soybean, live cattle and live hogs (monthly rolling) based on normal distribution.

Test	Mean	Variance	Skewness	Kurtosis	JB
Wheat					
GARCH	0.79773	0.04506	−0.89515*	2.99036*	122.992*
BEKK	0.82519	0.05153	−0.33703**	0.86755*	13.1262*
GARCH-X	0.77576	0.05366	−0.82546*	1.28341*	47.5528*
BEKK-X	0.78425	0.05069	−0.77735*	1.34833*	46.0565*
GJR	0.83938	0.05283	−0.37443*	0.94584*	15.8274*
Q-GARCH	0.84525	0.05560	−0.52859*	1.08166*	23.1619*
Soybean					
GARCH	0.95718	0.01555	−0.91264*	2.75244*	118.619*
BEKK	0.92290	0.00824	−1.13491*	1.74837*	89.2713*
GARCH-X	0.93112	0.01026	−0.76926*	0.45431	27.9860*
BEKK-X	0.96125	0.01484	−0.87617*	1.06416*	59.7293*
GJR	0.93422	0.01298	−0.64715*	−0.00525	18.2181*
Q-GARCH	0.64648	0.18277	−0.68597*	−1.41735*	42.3156*
Live cattle					
GARCH	0.08273	0.00009	−1.04527*	2.23164*	96.2326*
BEKK	0.05396	0.00150	1.75444*	6.75950*	609.032*
GARCH-X	0.08618	0.00039	1.00214*	0.24731	41.6329*
BEKK-X	0.06057	0.00097	3.06708*	21.8366*	5401.86*
GJR	0.09122	0.00106	0.70896	−0.67259**	25.8599*
Q-GARCH	0.06351	0.00040	1.16424*	1.73150*	88.4090*
Live hogs					
GARCH	0.05577	0.00025	−0.63790*	9.31225*	931.311*
BEKK	0.03033	0.00019	1.37716*	12.8945*	1832.70*
GARCH-X	0.05477	0.00013	−1.63948*	11.2230*	1395.57*
BEKK-X	0.02360	0.00004	−0.15294	−0.31987	2.05669
GJR	0.02913	0.00002	−0.13158	−0.47139	3.06030
Q-GARCH	0.01639	0.00001	−0.13315	−0.03807	0.75981

*Significance at the 5% significant level
**Significance at the 10% significant level

6. Model confidence set

As stated earlier we apply the MCSs to compare the forecasting ability of the six GARCH models. According to Hansen, Lunde, and Nason (2005, 2011), the MCS has several advantages compared with other comparison techniques. First, it takes into consideration the limitations of the data. Informative data will result in a MCS that contains only the best model; however, less informative data make it difficult to distinguish between models and may result in a MCS that contains several models. Second, the MCS procedure makes it possible to make statements about significance that are valid in the traditional sense. This is a property that is not satisfied by the commonly used approach of reporting p-values from multiple pairwise comparisons. Third, the MCS procedure allows for the possibility that more than one model can be the best.

The MCS is a test proposed by Hansen, Lunde, and Nason (2003, 2011) for the model selection, and for comparing the forecasting ability of different models. The MCS selects the best model among a set of competing models with a certain confidence level. The MCS is constructed from a collection of competing models, M^0 and a criterion for evaluating these models empirically

(Hansen, Lunde, and Nason 2005). This method has two procedures, an equivalence test δ_M and an elimination rule e_M.[17] The equivalence test is applied to the set of models $M = M^0$. If the equivalence test δ_M is rejected, there is evidence that the models in M are not equally 'good' and an elimination rule e_M is used to eliminate a model with poor sample performance from M. This procedure is repeated until δ_M is accepted and the MCS is now defined by the set of surviving models. The MCS yields a p-value for each model in M and if the MCS p-value of model i is larger than the significance level α, we say that model i is the 'best' candidate in M^0 at significance level α. In other words, the model with a higher MCS p-value is more likely to survive in M. If we fail to reject the null, it shows that all models in set M are equally good, and we define the $\widehat{M}^*_{1-\alpha} = M$; and if the null is rejected, it is evidence that all forecasting models are not equally good and the elimination rule e_M will be applied to discard the inferior model(s) from M. The $\widehat{M}^*_{1-\alpha}$ is the *MCS* which consists of the best models which are not eliminated from M with confidence level α.

In the case of forecasting, the set M^0 contains forecasting models with index $i = 1, 2, \ldots, m$ and assume that the set $M = M^0$. For equivalence test δ_M, we evaluate the forecasting models in terms of a loss function, such as mean absolute error (MAE), mean square error and root mean square error and we denote the model i in time period t as $L_{i,t} = L(Y_t, \widehat{Y}_{i,t})$, where $\widehat{Y}_{i,t}$ is the forecast of Y_t. First, the bootstrap method is applied to resample the time index for each model, and we define the corresponding bootstrap variable $L^*_{b,i}$. By generating B samples in the sequential testing, we can estimate some parameters and the high-dimensional covariance matrix.

Second, to calculate the test statistic, we let $\bar{d}_{i\cdot} = \bar{L}_i - \bar{L}_{\cdot}$, $\varepsilon^*_{b,i} = \bar{L}^*_{b,i} - \bar{L}_i$ and set

$$\widehat{\mathrm{var}}(\bar{d}_{i\cdot}) = \frac{1}{B}\sum_{b=1}^{B}(\varepsilon^*_{b,i} - \varepsilon^*_{b,\cdot})^2. \tag{16}$$

Here, we define $t_{i\cdot} = \bar{d}_{i\cdot}/\sqrt{\widehat{\mathrm{var}}(\bar{d}_{i\cdot})}$ and the estimated bootstrap distribution $t^*_{b,i\cdot} = (\varepsilon^*_{b,i} - \varepsilon^*_{b,\cdot})/\sqrt{\widehat{\mathrm{var}}(\bar{d}_{i\cdot})}$, and then we can get the test statistics $T_{\max} = \max_i t_{i\cdot}$ and $T^*_{b,\max} = \max_i t^*_{b,i\cdot}$ for $b = 1, \ldots, B$.

Third, under the null hypothesis, we set the p-value of H_0 as

$$P_{H_0} = \frac{1}{B}\sum_{b=1}^{B} I_{\{T_{\max} > T^*_{b,\max}\}}, \tag{17}$$

where $I_{\{T_{\max} > T^*_{b,\max}\}}$ is the indicator function. As stated, if $P_{H_0} > \alpha$, we accept the null hypothesis, and we have the $(1-\alpha)$ MCS $\widehat{M}^*_{1-\alpha} = M$; otherwise, we move on to the elimination rule e_M to find out the 'worst performing model $i^{\triangle}$' which is equal to $i^{\triangle} = \arg\max_{i\in M} \bar{d}_{i\cdot}/\sqrt{\widehat{\mathrm{var}}(\bar{d}_{i\cdot})}$, and remove the model $i^{\triangle}$ from M until we get an acceptance.[18]

7. Forecast error tests results

Table 3 reports the MAE and MCS p-value for each forecasting model for all four commodities using both the normal and student-t distributions over 1- and 2-year forecasting horizons. We find that a low MAE of forecasts is associated with a high MCS p-value, and this result is in line with the principle of the MCS approach that the models with high MCS p-values are more likely to be 'the best' models at a certain level of confidence. The MCS is applied based on two confidence

Table 3. MCS *p*-value for wheat, soybean, live cattle and live hogs (monthly rolling).

	1-year return forecasts				2-year return forecasts			
	Normal		Student's *t*		Normal		Student's *t*	
Models	MAE	*P*-value	MAE	*P*-value	MAE	*P*-value	MAE	*P*-value
Wheat								
GARCH	0.00292	0.145*	0.00205	0.123*	0.00464	0.112*	0.00347	0.151*
BEKK	0.00266	0.250**	0.00193	0.184*	0.00450	0.248*	0.00305	0.426**
GARCH-X	0.00303	0.009	0.00301	0.087	0.00535	0.000	0.00365	0.000
BEKK-X	0.00293	0.096	0.00293	0.094	0.00532	0.000	0.00355	0.000
GJR	0.00265	0.282**	0.00188	0.244*	0.00525	0.053	0.00348	0.151*
Q-GARCH	0.00267	0.241*	0.00182	0.307**	0.00381	0.505**	0.00336	0.234*
Soybean								
GARCH	0.00121	0.088	0.00089	0.014	0.00528	0.001	0.00463	0.0293
BEKK	0.00115	0.240*	0.00077	0.250**	0.00437	1.000**	0.00409	0.449**
GARCH-X	0.00105	0.326**	0.00079	0.182*	0.00481	0.099	0.00430	0.191*
BEKK-X	0.00118	0.229*	0.00083	0.109*	0.00475	0.108*	0.00433	0.178*
GJR	0.00119	0.229*	0.00095	0.000	0.00458	0.216*	0.00425	0.244*
Q-GARCH	0.00208	0.151*	0.00092	0.002	0.00478	0.111*	0.00435	0.098
Live cattle								
GARCH	0.00620	0.190*	0.00625	0.193*	0.00620	0.831**	0.00625	0.337**
BEKK	0.00629	0.025	0.00657	0.000	0.00655	0.008	0.00641	0.095
GARCH-X	0.00618	0.195*	0.00626	0.193*	0.00638	0.097	0.00651	0.026
BEKK-X	0.00625	0.097	0.00648	0.000	0.00653	0.008	0.00627	0.212*
GJR	0.00605	0.696**	0.00621	0.212*	0.00633	0.158*	0.00637	0.179*
Q-GARCH	0.00624	0.096	0.00617	0.279**	0.00641	0.107*	0.00659	0.000
Live hogs								
GARCH	0.00946	0.000	0.00942	0.049	0.01030	0.060	0.01023	0.034
BEKK	0.00939	0.000	0.00929	0.091	0.01027	0.049	0.01005	0.079
GARCH-X	0.00953	0.000	0.00946	0.017	0.01012	0.096	0.01031	0.000
BEKK-X	0.00928	0.112*	0.00930	0.091	0.01011	0.096	0.01019	0.062
GJR	0.00940	0.000	0.00919	0.109*	0.00997	0.488**	0.00949	0.183*
Q-GARCH	0.00915	1.000**	0.00870	0.250**	0.01004	0.179*	0.00945	0.226**

Note: The MAEs and MCS *p*-value for the different forecasts with normal and student-*t* distributions over 1- and 2-year forecasting horizons; the forecasts in $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ are marked by * and **, respectively.

levels, $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ which contain 'best models' at 90% and 75% confidence levels in these sets, respectively. Moreover, the results in $\widehat{M}^*_{75\%} \subset \widehat{M}^*_{90\%}$ are as expected.

For wheat, all forecasting models except the GARCH-X and BEKK-X end up in $\widehat{M}^*_{90\%}$. This is true for both distributions and both forecasting horizons. This result may indicate at least for wheat, that inclusion of cointegration between the cash and futures market may not be important for forecasting. The MCS picks BEKK and GJR models in $\widehat{M}^*_{75\%}$ with normal distribution for the 1-year forecast. With student-*t* distribution at $\widehat{M}^*_{75\%}$, the Q-GARCH significantly outperforms other models at this confidence level. For 2-year forecasts, half the models remain in $\widehat{M}^*_{90\%}$ based on both distributions. However, only the Q-GARCH and BEKK models stay in $\widehat{M}^*_{75\%}$, for normal and student-*t* distributions.

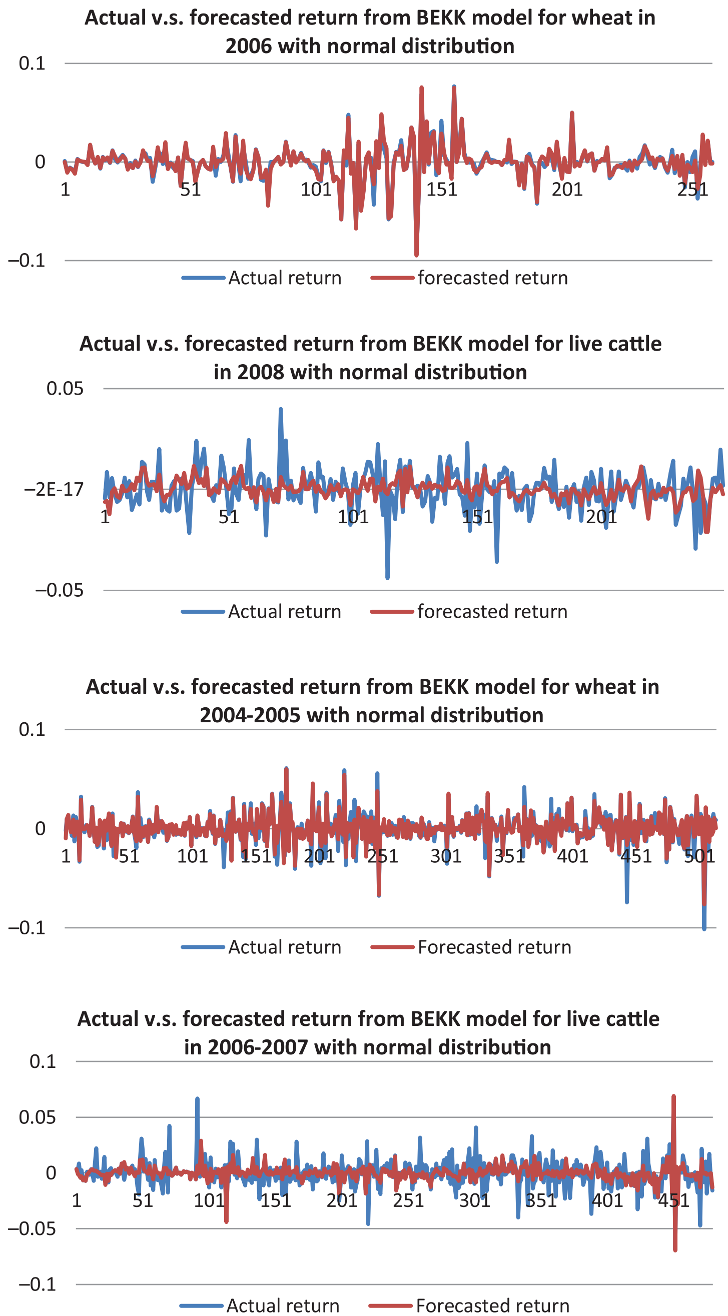


Figure 5. Patterns of actual *vs.* forecasted return from BEKK–GARCH model with normal distribution for wheat and live cattle.

The MCS results for soybean confirm the superiority of the BEKK-family model. The BEKK performs the best in three out of four tests in $\widehat{M}^*_{75\%}$. The MCS $\widehat{M}^*_{90\%}$ returns both the BEKK-family and asymmetric GARCH models. On the other hand, the GARCH model is ruled

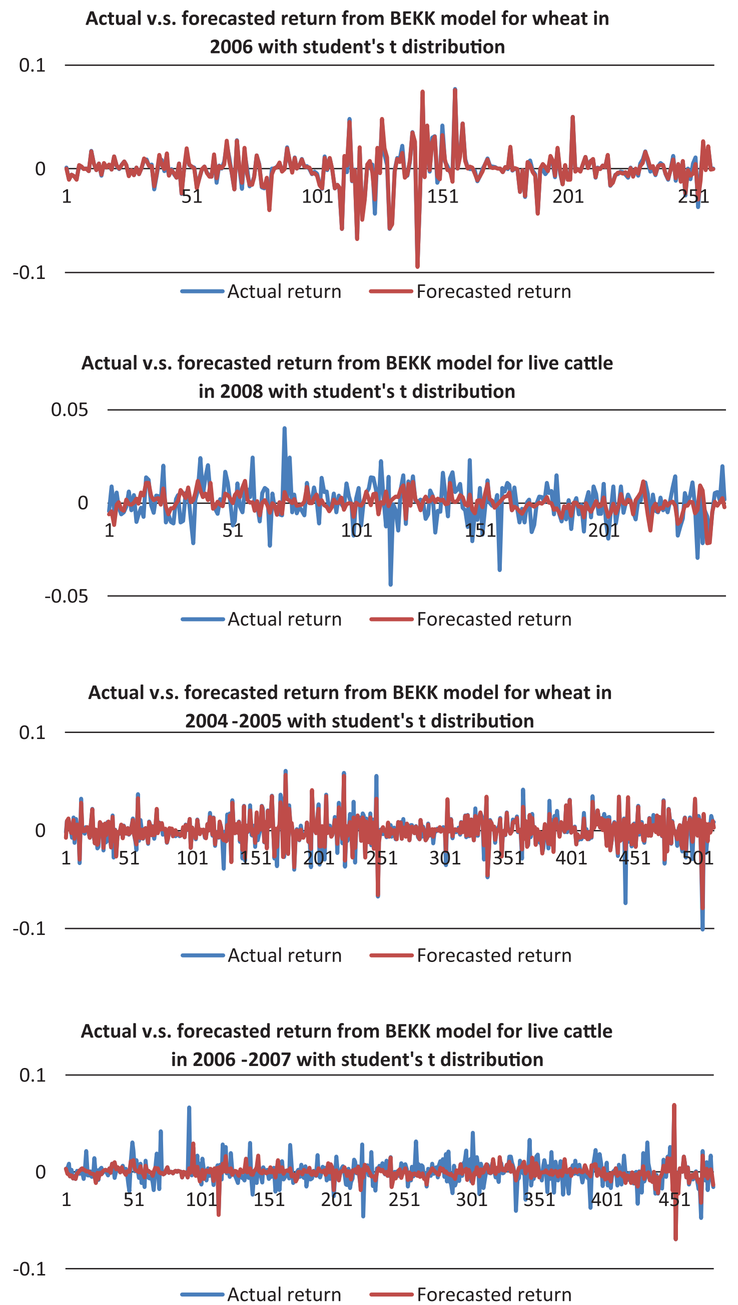


Figure 6. Patterns of actual *vs.* forecasted return from BEKK–GARCH model with student-*t* distribution for wheat and live cattle.

out for each case in $\widehat{M}^{*}_{90\%}$ and $\widehat{M}^{*}_{75\%}$. Both GARCH-X and BEKK-X models perform better for soybean (compared with wheat) indicating the importance of cointegration in forecasting OHR. For storable wheat and soybean, the BEKK-type models appear in both $\widehat{M}^{*}_{90\%}$ and $\widehat{M}^{*}_{75\%}$

in almost all cases. The asymmetric GARCH models survive in two out of eight cases in $\widehat{M}^*_{75\%}$.

The asymmetric GARCH–GJR model performs the best for live cattle during the 1-year forecast periods using normal distribution in $\widehat{M}^*_{75\%}$. For the 1-year horizon, the standard GARCH and the GARCH-X are also picked as one of the best models in $\widehat{M}^*_{90\%}$. Results are similar when comparing the models with the student-*t* distribution. Now the Q-GARCH model performs the best. For the 2-yearforecasts, the MCS $\widehat{M}^*_{75\%}$ contains the GARCH model for the live cattle with both normal and student-*t* distributions. Meanwhile, the asymmetric GJR and Q-GARCH models also provide significantly good forecasts in $\widehat{M}^*_{90\%}$ based on both distributions. In the case of live hogs, the asymmetric GJR and Q-GARCH models overwhelmingly dominate all other models in both $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ with both normal and student-*t* distributions over both horizons. For both non-storable commodities, GARCH-X and BEKK-X do not perform well.

Overall, the MCS for 1- and 2-year forecasts are different for all four commodities in both the $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ confidence levels. The outcome of MCS-$\widehat{M}^*_{75\%}$ reveals that the BEKK model dominates others models for storable wheat and soybean for both forecasting horizons, and the asymmetric GJR and Q-GARCH models have the best forecasting performance for non-storable products, live cattle and live hogs.

Using different distributions of GARCH models and different lengths of forecast horizons does provide different results and conclusions. If the normality of log returns is evident, and if there is either a fatter tail with a lower peak or a thinner tail with a higher peak, the forecasting performance of models with student-*t* distribution is better than those with normal distribution.

Figures 5 and 6, based on the same scale, show the return forecast by the BEKK method (using both distributions) and the actual returns over both forecast horizons for wheat and live cattle. All estimates seem to move together with the actual return, but because of the high frequency of the data, it is difficult to say which method shows the closest correlation. Figures for other methods show similar conclusion and are not provided to save space, but are available on request.

Our results fail to point out any one particular type of GARCH model that has superior ability in forecasting the time-varying hedge ratio in the commodities market. This result backs the claim by Poon and Granger (2003) that no one particular type of GARCH model is superior in forecasting. Superiority of forecasting performance depends upon several different factors. In this paper, results are different based on the storability of the commodity, the length of the forecast horizons and the distribution of the errors.

8. Conclusion

This paper empirically estimates and forecasts the hedge ratios of four commodities from the agricultural futures market; two storable commodities, wheat, soybean, and two non-storable commodities, live cattle and live hogs. The ability to forecast the OHRs/dynamic hedge ratios is important for understanding the role of the futures markets in trading, programme trading, index arbitrage and the development of optimal hedging and trading strategies in fund management. The forecasting of hedge ratios helps the hedger to choose the appropriate portfolio and allows for portfolio adjustment in dynamic hedging.

This paper employs the six different GARCH models to estimate and forecast the hedge ratio using the rolling procedure. The six models applied are the standard bivariate GARCH, bivariate

BEKK, bivariate GARCH-X, bivariate BEKK-X, bivariate Q-GARCH and bivariate GARCH–GJR. We apply both the normal and the student$-t$ distributions of the GARCH models. This paper, thus, also provides a comparison between the forecasting ability of the six GARCH models based on two different distributions. The forecasting is conducted by means of rolling procedure and the MCS approach is applied to compare the forecasting ability of the six GARCH models. The tests are carried out in three steps. In the first step, we forecast the hedge ratio by means of the six GARCH models using monthly rolling procedure. In the second step, the forecasting models are used to forecast returns based on the forecasted time-varying hedge ratios. In the third step, we empirically compare the GARCH models in terms of forecasting accuracy. These will provide evidence for comparative analysis of the merits of the different forecasting models. The evaluation of forecast accuracy is conducted by forecasting out-of-sample returns of portfolios implied by the computed hedge ratios.

The MCS is applied based on two confidence levels, $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$. Overall, the MCS results for 1- and 2-year forecasts are different for all four commodities in both $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ confidence levels. The outcome of MCS-$\widehat{M}^*_{75\%}$ reveals that the BEKK model dominates others models for storable wheat and soybean for both forecasting horizons, and the asymmetric GJR and Q-GARCH models does the best forecasting performance for the non-storable products, live cattle and live hogs.

Our results fail to point out any one particular type of GARCH model that has superior ability over the other models in forecasting the time-varying hedge ratio in the commodities markets. Using different distributions of GARCH models and different lengths of forecast horizons does provide different results and conclusions. The forecasting performance of models with student-*t* distribution is better than those with normal distribution.

Results presented in this paper advocate further research in this field, applying different markets, time periods, length of forecast horizon and methods. There are potential insights to be gained from examining markets with different institutional features.

Acknowledgements

We thank an anonymous referee for his valuable comments and suggestions. We also thank the participants of the FFM Conference 2011 in Marseille, France and the 31st Annual International Symposium on Forecasting (ISF), Prague, Czech Republic, 2011 for comments on an earlier version of the paper. Part of this work was done with the support from the Fundamental Research Funds for the Central Universities.

Notes

1. Numerous articles show the superiority of different GARCH models in estimating hedge ratios in financial markets (see Floros and Vougas 2004, 2006).
2. According to Covey and Bessler (1995), no asset is perfectly storable or non-storable. An asset with minimum storage costs is an asset which does not easily spoil and can be stored cheaply relative to its value.
3. Bollerslev (2008) provides a detailed glossary of the ARCH and GARCH models.
4. In previous studies, different versions of the GARCH models have been used for forecasting volatility, time-varying beta, etc. and then the models compared. See Poon and Granger (2003) and Choudhry and Wu (2008) for citations of some of these previous studies. Poon and Granger (2003) also provide an excellent survey of GARCH and other models forecasting ability.
5. According to Cecchetti et al. (1988), the optimal hedge ratio β can be expressed as $\rho\sigma^{\mathrm{c}}/\sigma^{\mathrm{f}}$, where ρ is the correlation between the futures price and cash price, σ^{c} is the cash standard deviation and σ^{f} is the futures standard deviation. Thus, if the futures have the same or higher price volatility than the cash, the hedge ratio can be no greater than the correlation between them, which will be less than unity.

6. These GARCH models are chosen as they represent the most commonly applied versions of symmetric and asymmetric models in the literature (See Bauwens, Laurent, and Rombouts 2006).
7. The following cointegration relationship is investigated by means of the Engle and Granger (1987) method:

$$\mathrm{c}_t = \eta + \gamma \mathrm{f}_t + z_t,$$

   whereas stated above c_t and f_t are the log of the cash index and futures price index, respectively. The residuals z_t are tested for unit root(s) to check for cointegration between c_t and f_t.
8. Ghosh (1995), Ghosh and Clayton (1996) and Kroner and Sultan (1993) have shown that hedge ratios and hedging performance may change considerably if the cointegration between the cash and futures prices is omitted from the statistical models and estimations.
9. Huang, Su, and Li (2010) find that the BEKK–GRACH model holds advantages over the DCC–GARCH on volatility forecasting in the case of three AAA-rated Euro zero-coupon bonds. Furthermore, when the BEKK–GARCH model is restricted to a scalar BEKK–GARCH, it is superior to DCC for estimating conditional variance and correlation (Caporin and McAleer 2012)
10. There is more than one GARCH model available that is able to capture the asymmetric effect in volatility. According to Engle and Ng (1993) and Glosten et al. (1993), the GARCH–GJR model is the best at parsimoniously capturing this asymmetric effect.
11. Compared with standard GARCH models, the Q-GARCH is better able to capture the asymmetry information impact and higher kurtosis, both important features of economic and financial data. The multivariate Q-GARCH model benefits from a relatively small computational burden when order increasing compared with the standard GARCH model. In the empirical study of Franses and Van Dijk (1996) and Ulu (2005), the Q-GARCH is sensitive to extreme values; however, it performs best in volatility forecasting compared with the random walk, standard GARCH and GJR–GARCH models after filtering extreme observations in the stock market.
12. Harris and Sollis (2003, 247) discuss the methodology in more detail.
13. The different length of the data is due to the availability of the data. The data stop before the start of the current financial crisis. We deliberately wanted to avoid the crisis period because of the large effect of the current financial crisis on asset volatility (Schwert 2011) and to use a period of relatively normal times.
14. The continuous series is a perpetual series of futures prices. It starts at the nearest contract month, which forms the first set of values for the continuous series, either until the contract reaches its expiry date or until the first business day of the actual contract month. At this point, the next trading contract month is taken.
15. To assess the general descriptive validity of the model, a battery of standard specification tests is employed. Specification adequacy of the first two conditional moments is verified through the serial correlation test of white noise. These tests employ the Ljung-Box Q statistics on the standardized (normalized) residuals $(\varepsilon_t/H_t^{1/2})$, standardized squared residuals (ε_t/H_t^2) and the cross-standardized residuals. The latter are the cross-product between the standardized residuals of cash and futures. All series are found to be free of serial correlation (at the 5% level). The absence of serial correlation in the standardized-squared residuals implies there is no necessity to encompass a higher order ARCH process (Giannopoulos 1995).
16. The difference in the hedge ratios between the storable and non-storable commodities could be due to the fact that the price basis for non-storable goods is more volatile than storable goods, and hence the covariance between returns in cash and futures markets of non-storable commodities is lower than that of storable commodities. Consequently, low covariance results in a low hedge ratio.
17. The equivalence test is similar to the equal predictive ability (EPA) test which is proposed by Diebold and Mariano (1995) and Harvey et al. (1997). However, the equivalence test constructs a test statistic which is more efficient than the EPA test on comparing a large number of models.
18. It is worthy noticing that the MCS p-value for model $e_{M_j} \in M^0$ does not necessarily equal the p-value from the null hypothesis, yet it is defined as $\max_{i \leq j} P_{H_0}$.

References

Alberg, D., H. Shalit, and R. Yosef. 2008. "Estimating Stock Market Volatility Using Asymmetric GARCH Models." *Applied Financial Economics* 18 (15): 1201–1208.

Baillie, R., and R. DeGennaro. 1990. "Stock Returns and Volatility." *Journal of Financial and Quantitative Analysis* 25 (2): 203–214.

Baillie, R., and R. Myers. 1991. "Bivariate GARCH Estimates of the Optimal Commodity Futures Hedge." *Journal of Applied Econometrics* 6 (2): 109–124.

Bauwens, L., S. Laurent, and J. Rombouts. 2006. “Multivariate GARCH Models: A Survey.” *Journal of Applied Econometrics* 21 (1): 79–109.

Black, F. 1976. “Studies of Stock Market Volatility Changes.” In *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 177–181.

Bollerslev, T. 1988. “On the Correlation Structure for the Generalized Autoregressive Conditional Heteroscedastic Process.” *Journal of Time Series Analysis* 9 (2): 121–131.

Bollerslev, T. 1990. “Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model.” *The Review of Economics and Statistics* 72 (3): 498–505.

Bollerslev, T. 2008. “Glossary to ARCH (GARCH).” CREATES Research Papers, Department of Economics, Duke University, USA.

Bollerslev, T., R. Chou, and K. Kroner. 1992. “ARCH Modeling in Finance.” *Journal of Econometrics* 52 (1–2): 5–59.

Brockwell, P. J., and R. A. Davis. 2002. *Introduction to Time Series and Forecasting*. New York: Springer.

Brooks, C. 2008. *Introductory Econometrics for Finance*. Cambridge, UK: Cambridge University Press.

Caporin, M., and M. McAleer. 2012. “Do we Really Need both BEKK and DCC? A Tale of Two Multivariate GARCH Models.” *Journal of Economic Surveys* 26 (4): 736–751.

Cecchetti, S. G., R. E. Cumby, and S. Figlewski. 1988. “Estimation of Optimal Futures Hedge.” *Review of Economics and Statistics* 70 (4): 623–630.

Chen, S., C. Lee, and K. Shrestha. 2004. “An Empirical Analysis of Relationship between the Hedge Ratio and Hedging Horizon: A Simultaneous Estimation of a Short- and Long-Run Hedge Ratios.” *Journal of Futures Markets* 24 (4): 359–386.

Cheung, Y. W., M. Chinn, and A. G. Pascual. 2003. *What do we Know About Recent Exchange Rate Models? In-Sample Fit and Out-Sample Performance Evaluated*. Santa Cruz Center for International Economics, Department of Economics, UCSC: UC Santa Cruz.

Choudhry, T. 2009. “Short-Run Deviations and Time-Varying Hedge Ratios: Evidence from Agricultural Futures Markets.” *International Review of Financial Analysis* 18 (1–2): 58–65.

Choudhry, T., and H. Wu. 2008. “Forecasting Ability of GARCH vs Kalman Filter Method: Evidence from Daily UK Time-Varying Beta.” *Journal of Forecasting* 27 (8): 670–689.

Christie, A. 1982. “The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects.” *Journal of Financial Economics* 10 (4): 407–432.

Covey, T., and D. Bessler. 1995. “Asset Storability and the Informational Content of Inter-temporal Prices.” *Journal of Empirical Finance* 2 (2): 103–115.

Diebold, F. X. 2007. *Elements of Forecasting*. Ohio: Thomson/South-Western.

Diebold, F. X., and R. S. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13 (3): 253–263.

Engle, R. F., and C. W. J. Granger. 1987. “Co-Integration and Error Correction: Representation, Estimation, and Testing.” *Econometrica* 55 (2): 251–276.

Engle, R. F., and K. F. Kroner. 1995. “Multivariate Simultaneous Generalized ARCH.” *Econometric Theory* 11: 122–150.

Engle, R. F., and V. Ng. 1993. “Measuring and Testing the Impact of News on Volatility.” *Journal of Finance* 48: 1749–1778.

Engle, R. F., and B. S. Yoo. 1987. “Forecasting and Testing in Co-Integrated Systems*” *Journal of Econometrics* 35 (1): 143–159.

Floros, C., and D. Vougas. 2004. “Hedge Ratios in Greek Stock Index Futures Market.” *Applied Financial Economics* 14 (15): 1125–1136.

Floros, C., and D. Vougas. 2006. “Hedging Effectiveness in Greek Stock Index Futures Market, 1999–2001.” *International Research Journal of Finance and Economics* 5: 7–18.

Franses, P., and D. Van Dijk. 1996. “Forecasting Stock Market Volatility Using (non-linear) GARCH Models.” *Journal of Forecasting* 15 (3): 229–235.

Ghosh, A. 1995. “The Hedging Effectiveness of ECU Futures Contracts: Forecasting Evidence from an Error Correction Model.” *Financial Review* 30 (3): 567–581.

Ghosh, A., and Clayton, R., 1996. “Hedging with International Stock Index Futures: An Intertemporal Error Correction Model.” *Journal of Financial Research* 19 (4): 477–491.

Giannopoulos, K. 1995. “Estimating the Time-Varying Components of International Stock Markets’ Risk.” *European Journal of Finance* 1 (2): 129–164.

Glosten, L., R. Jagannathan, and D. E. Runkle 1993. “On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.” *Journal of Finance* 48 (5): 1779–1801.

Hansen, P. R., A. Lunde, and J. M. Nason. 2003. "Choosing the Best Volatility Models: The Model Confidence Set Approach." *Oxford Bulletin of Economics and Statistics* 65 (s1): 839–861.

Hansen, P. R., A. Lunde, and J. M. Nason. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH (1, 1)?" *Journal of Applied Econometrics* 20 (7): 873–889.

Hansen, P. R., A. Lunde, and J. M. Nason. 2011. "The Model Confidence Set." *Econometrica* 79 (2): 453–497.

Harris, R., and R. Sollis. 2003. *Applied Time Series Modelling and Forecasting*. New York: Wiley.

Harvey, D., S. J. Leybourne, and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting* 13 (2): 281–291.

Hu, W., and A. N. Kercheval. 2010. "Portfolio Optimization for Student *t* and Skewed *t* Returns." *Quantitative Finance* 10 (1): 91–105.

Huang, Y., W. Su, and X. Li. 2010. "Comparison of BEKK GARCH and DCC GARCH Models: An Empirical Study." *Advanced Data Mining and Applications* 99–110.

Kroner, K., K. Kneafsey, and S. Claessens. 2006. "Forecasting Volatility in Commodity Markets." *Journal of Forecasting* 14 (2): 77–95.

Kroner, K. F., and J. Sultan. 1993. "Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures." *Journal of Financial and Quantitative Analysis* 28 (4): 535–551.

Lalli, W. R. 2011. *Handbook of Budgeting*. New Jersey: Wiley.

Lee, T. 1994. "Spread and Volatility in Spot and Forward Exchange Rates." *Journal of International Money and Finance* 13 (3): 375–383.

Lien, D. 2004. "Cointegration and the Optimal Hedge Ratio: The General Case." *Quarterly Review of Economics and Finance* 44 (5): 654–658.

Lien, D., Y. K. Tse, and A. K. Tsui. 2002. "Evaluating the Hedging Performance of the Constant-Correlation GARCH Model." *Applied Financial Economics* 12 (13): 791–798.

Moon, G. H., W. C. Yu, and C. H. Hong. 2009. "Dynamic Hedging Performance with the Evaluation of Multivariate GARCH Models: Evidence from KOSTAR Index Futures." *Applied Economics Letters* 16 (9): 913–919.

Myers, R. 1991. "Estimating Time Varying Hedge Ratios on Futures Markets." *Journal of Futures Markets* 11 (1): 39–53.

Park, T. A., and F. Antonovitz. 1992. "Econometric Tests of Firm Decision Making under Uncertainty: Optimal Output and Hedging Decisions." *Southern Economic Journal* 58 (3): 593–609.

Parks, P. C. 1992. "AM Lyapunov's Stability Theory – 100 Years on." *IMA Journal of Mathematical Control and Information* 9 (4): 275–303.

Poon, S., and C. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41 (2): 478–539.

Schwert, G. 2011. "Stock Volatility During the Recent Financial Crisis." NBER Working Paper, Cambridge, MA: NBER.

Sultan, J., and M. S. Hasan. 2008. "The Effectiveness of Dynamic Hedging: Evidence from Selected European Stock Index Futures." *The European Journal of Finance* 14 (6): 469–488.

Ulu, Y. 2005. "Out-of-Sample Forecasting Performance of the QGARCH Model." *Applied Financial Economics Letters* 1 (6): 387–392.

Yang, J., and T. Awokuse. 2003. "Asset Storability and Hedging Effectiveness in Commodity Futures Markets." *Applied Economics Letters* 10 (8): 487–491.

Yang, J., D. Bessler, and D. Leatham. 2001. "Asset Storability and Price Discovery in Commodity Futures Markets: A New Look." *Journal of Futures Markets* 21 (3): 279–300.