Choudhry, Taufiq, Hasan, Mohammad S and Zhang, Yuanyuan (2019) *Forecasting the Daily Dynamic Hedge Ratios in Emerging European Stock Futures Markets: Evidence from GARCH models.* International Journal of Banking Accounting and Finance, 10 (1). pp. 67-100. ISSN 1755-3830.

Forecasting the Daily Dynamic Hedge Ratios in Emerging European Stock Futures Markets: Evidence from GARCH models

Taufiq Choudhry[a], Mohammad Hasan[b] and Yuanyuan Zhang[c]

[a] *Southampton Business School, University of Southampton, UK;* [b] *Kent Business School, University of Kent, UK;* [c] *Southwestern University of Economics and Finance, P.R. China*

Abstract

This paper empirically estimates and forecasts the hedge ratios of three emerging European and one developed stock futures markets by means of seven different versions of GARCH model. The seven GARCH models applied are bivariate GARCH, GARCH-ECM, BEKK GARCH, GARCH-DCC, GARCH-X, GARCH-GJR and GARCH-JUMP. Daily data during January 2000-July 2014 from Greece, Hungary, Poland and the UK are applied. Forecast errors based on these four stock futures portfolio return forecasts (based on forecasted hedge ratios) are employed to evaluate out-of-sample forecasting ability of the seven GARCH models. The comparison is done by means of Model Confidence Set (MCS) and modified Diebold-Mariano tests. Forecasts are conducted over two non-overlapping out-of-sample periods, a two-year period and a one-year period. MCS results indicate that the GARCH model provides the most accurate forecasts in five cases, while each of the GARCH-ECM, GARCH-X and GARCH-GJR models constitutes model confidence set in four cases at a reasonable confidence level. Models selection based on modified Diebold-Mariano tests further corroborate results of the MCS tests. Differences between the portfolio returns also indicate the high forecasting ability of GARCH-BEKK and GARCH-GJR models.

Keywords: Forecasting; hedge ratio; GARCH; emerging market; volatility

JEL Classification: G1; G15

1. Introduction

Since the seminal research of Working (1953), Johnson (1960) and Ederington (1979), there has been a significant level of interest in the modelling and forecasting of the optimal hedge ratios (OHR) and alternative hedging strategies applied to the commodity and financial futures.[1] It is 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 change in the related cash markets. Careful selection of derivatives' contracts is conditional upon the accuracy of OHR estimates and volatility forecasting techniques. 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).

Given the plethora of literature, there are nevertheless serious gaps in the current research strand in two directions. Firstly, from a risk management perspective, there have been limited attempts to evaluate the forecasting accuracy and performance of the estimated hedge ratios derived from different econometric models.[2] Indeed, knowledge of forecasting ability of optimal hedge ratio/dynamic hedge ratio is important for understanding the role of futures markets in equity trading, program trading, index arbitrage, and the development of optimal hedging and trading strategies in fund management. Secondly, most previous studies confined their attention to more developed and mature financial markets and exchanges. There have only been limited attempts to examine the behaviour of time-varying hedge ratios for emerging markets.[3] Emerging equity markets now account for more than one-fifth of global equity market capitalization. The burgeoning size of the emerging markets has been supported by a growing domestic investor base, including domestic institutional investors and increased financial integration with the rest of the world (Bailey 2010).

This paper takes steps to address these gaps in the literature. We investigate the behaviour of dynamic hedge ratios in three emerging European stock futures markets using alternative variants of GARCH models and compare the forecasting performance of these models. More specifically, using daily data of the spot and futures stock markets of Greece, Hungary and Poland and the following seven variants –GARCH, GARCH-BEKK, GARCH-ECM, GARCH-DCC, GARCH-X, asymmetric GARCH-GJR, and autoregressive jump intensity GARCH (GARCH-JUMP) models – we have estimated the time-varying hedge ratios and compared the forecasting performances of these models of hedge ratios.[4] The forecasting comparison is achieved by means of the Model Confidence Set (MCS) and modified Diebold and Mariano (MDM) tests. We provide further analysis by statistically comparing the difference between the forecasted returns after adjusting for transaction cost from each model during both forecasting horizons. This analysis may provide a way to choose the best model based on application rather than on a statistical criterion, such as MCS. Given the different methods available, the empirical question we address is: which econometric method provides the best forecast? This paper hopes to provide an answer to this question. The forecast is conducted over two non-overlapping different lengths of forecast horizons.[5]

These three markets are chosen for the following reasons. First, all three markets are located in the European Union (EU) with varying degrees of economic and market conditions. Hungary and Poland represent reforming Transition Economies from Eastern Europe while Greece represents a debt-ridden OECD country from the Eurozone which underwent a prolonged period of global financial and the Eurozone crises. The FTSE group classified Hungary and Poland as advanced emerging markets as they represent upper- or lower-middle income gross national income (GNI) countries with advanced market structure or high GNI countries with less developed market infrastructure. Second, these three emerging markets provide ample high-quality data on spot and futures stock prices to conduct a forecasting exercise.

Although it is alleged that some emerging markets are characterised by low liquidity, thin trading, and considerable volatility and possibly with less informed investors with access to information, our sample countries contain reliable and long-spanning data series. These markets underwent substantial changes in regulation and liberalisation which had encouraged wide participation in the market and led to more rapid impounding of information into prices.[6] Furthermore, we have incorporated the experience of a developed market for a meaningful comparison of our results. More specifically, we have added the data series from the UK market to investigate and offer a comparative flavour in our results.

To the best of our knowledge, no previous study empirically investigates the out-of-sample forecasting by different GARCH models of time-varying hedge ratios for emerging European stock futures markets and then compares the forecasting performance of these models. This is particularly true, taking into consideration that we apply the Model Confidence Set (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.

Certain assumptions are important for quality forecasting. These assumptions are the relationships between the cash and futures prices, the length of forecasting horizons, and the level of competition in the market. Forecast of the hedge ratio is not plausible if the relationship between the cash and futures prices is not assured. Further, according to Chen *et al.* (2004), the forecasting accuracy of different methods may be affected by the length of the hedging horizons. A longer length of forecast horizon implies a more accurate forecast due to higher data numbers;[7] thus, we apply two different lengths (one year and two years) of forecast horizons to observe the effect on the forecasting effectiveness of the models. Furthermore, the more competition there is in the market, the more difficult it is to forecast hedge ratios. In a highly competitive market, competitors can change the course of future events after they make forecasts in order to make themselves more competitive, which then invalidates the forecasts.

It is important to point out that the lack of a benchmark is an inevitable weak point for studies on time-varying hedge ratio forecasts. The point estimation of the hedge ratio generated by the GARCH model is only 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. Evaluation of forecast accuracy is thus conducted by forecasting out-of-sample returns of portfolios implied by the forecasted hedge ratios.[8] Summarising our results, following the MCS test, the GARCH model provides the most accurate forecasts in five cases, while each of the GARCH-ECM, GARCH-X and GARCH-GJR models constitutes MCS in four cases at a reasonable confidence level. Models' selection based on MDM tests further corroborates results of the MCS tests. Results from the portfolio returns difference comparison also shows the superior forecasting ability of GARCH-BEKK and GARCH-GJR models.

The remainder of this paper is structured as follows. Section 2 describes the optimal hedge ratios and the seven GARCH models. The Model Confidence Set (MCS) and Diebold-Mariano tests are described in section 3. Section 4 furnishes a brief literature review. Section 5 discusses the data and the basic statistics, while section 6 analyzes the GARCH and MCS and modified Diebold-Mariano results. The comparison between returns is provided in section 7. Section 8 presents the conclusion.

2. Estimation of optimal hedge ratios and the GARCH models

2.1 *The hedge ratio*

Johnson's (1960) risk-minimising hedge ratio h^* is defined as

$$h^* = -\frac{\sigma_{c,f}}{\sigma_f^2} = -\frac{\text{cov}(r_c, r_f)}{\text{var}(r_f)}, \quad (1)$$

where r_c and r_f denote returns on spot and futures indices. The optimal hedge ratio (OHR) then is computed as the slope coefficient of the following regression:

$$r_{ct} = \alpha + \beta r_{ft} + \varepsilon_t, \quad (2)$$

where ε_t is an error term.[9] A $\beta = 0$ implies unhedged position; $\beta = 1$ signifies a fully hedged position; and $\beta < 1$ implies a partial hedge.

It is now well-known in the literature that the conventional hedging model has shortcomings. As the distribution of futures and spot prices are changing through time, h^*, which is expressed as the ratio of covariance between futures returns and cash returns and variance of futures returns, moves randomly through time (Cecchetti *et al.* 1988; Baillie and Myers 1991; Kroner and Sultan 1993). Therefore eq. (2) should be modified as

$$h_T^* = \frac{\text{cov}(r_{T+1}^c, r_{T+1t}^f \text{I} \Omega_T)}{\text{var}(r_{T+1}^f \text{I} \Omega_T)} \quad . \quad (3)$$

In eq. (3), conditional moments are changing as the information set, Ω_T, is updated; consequently, the number of futures contracts held and the optimal hedge ratio will also change over time – hence the *t* subscripts of h_T^*. Under the condition of time-varying distribution, the bivariate GARCH model is utilised to estimate the time-varying hedge ratios to approximate the dynamic hedging strategies.

2.2 *Bivariate GARCH model*

As stated above, the time-varying hedge ratios are estimated from seven variants of bivariate GARCH models: standard GARCH, GARCH-ECM, GARCH-DCC, GARCH-

BEKK, GARCH-GJR, GARCH-X, and GARCH-JUMP.[10,11] The following bivariate GARCH(p, q) model is applied to returns from the stock cash and futures markets:

$$y_t = \mu + \varepsilon_t \tag{4}$$

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

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

where $y_t = (r_t^c, r_t^f)$ is a (2x1) vector containing stock returns from the cash and futures markets, H_t is a (2x2) conditional covariance matrix, C is (3x1) parameter vector of constant, A_i and B_j are (3x3) parameter matrices, and vech is the column-stacking operator that stacks the lower triangular portion of a symmetric matrix. To make the estimation amenable, Engle and Kroner (1995) have suggested imposing various restrictions on the parameters of A_i and B_j matrices. Using the bivariate GARCH model, the time-varying hedge ratio can be computed as

$$h_t^* = \hat{H}_{12,t} / \hat{H}_{22,t}, \tag{7}$$

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.

2.3 *GARCH-ECM model*

When the bivariate GARCH model incorporates the error correction term in the mean equation, it becomes the GARCH-ECM model which is presented as

$$y_t = \mu + \delta(u_{t-1}) + \varepsilon_t \quad . \tag{8}$$

The lagged error-correction term u_{t-1} is retrieved from the cointegration regression between cash and futures stock prices. Therefore, a bivariate GARCH-ECM model is employed to account for the long-run relationship and basis risk (see Kroner and Sultan 1993).[12]

2.4 *Bivariate GARCH-BEKK model*

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

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

Eqs. (4) and (5) also apply to the BEKK model and are defined as before. In eq.(9), A_{ki}, *i = 1,...q, k = 1,...k*, and B_{kj} *j = 1,...q, k = 1,..k* are NxN matrices. The GARCH-BEKK model is sufficiently general in that it guarantees the conditional covariance matrix, H_t to be positive definite, and renders significant parameter reduction in the estimation.[13]

2.5 *Bivariate GARCH-GJR model*

Glosten *et al.* (1993) provide a modification to the GARCH model that allows positive and negative innovations to returns to have different impact on conditional variance.[14] They suggest that the asymmetry effect can also be captured simply by incorporating a dummy variable in the original GARCH,

$$\sigma_t^2 = \alpha_0 + \alpha u_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1} + \beta\sigma_{t-1}^2, \qquad (10)$$

where $I_{t-1} = 1$ if $u_{t-1} > 0$; 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.

2.6 *Bivariate GARCH-X model*

The GARCH-X model is an extension of the GARCH-ECM model as it incorporates the square of error correction term in the conditional covariance matrix (Lee 1994). In the GARCH-X model, conditional heteroscedasticity may be modeled as a function of lagged squared error correction term - in addition to the ARMA terms in the variance/covariance equations:

$$vech(H_t) = C + \sum_{i=1}^{p} A_i vech(\varepsilon_{t-i})^2 + \sum_{j=1}^{q} B_j vech(H_{t-j}) + \sum_{k=1}^{k} D_k vech(u_{t-1})^2 \quad . \qquad (11)$$

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.

2.7 *Bivariate GARCH-DCC*

The preceding variants of the GARCH model assume constant correlation in the conditional covariance matrix. Tse and Tusi (2002) developed the dynamic conditional correlational GARCH (GARCH-DCC) model by allowing the conditional correlation to vary over time.[15] The DCC model is often the most accurate in terms of forecasting depending on the criteria (Engle 2002). The bivariate covariance matrix of DCC can be expressed as

$$H_t = \begin{bmatrix} h_{c,t}^2 & h_{cf,t} \\ h_{cf,t} & h_{f,t}^2 \end{bmatrix} = \begin{bmatrix} h_{c,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \begin{bmatrix} h_{c,f} & 0 \\ 0 & h_{f,t} \end{bmatrix}, \tag{12}$$

where ρ_t is the time-varying conditional correlation coefficient of spot and futures returns at time *t*. The conditional correlation is specified as an autoregressive moving average process

$$\rho_t = (1-\theta_1-\theta_2)\rho + \theta_1\rho_{t-1} + \theta_2\phi_{t-1} . \tag{13}$$

2.8 Autoregressive Jump Intensity GARCH model

The autoregressive jump intensity GARCH (GARCH-JUMP) model was proposed by Chan (2008) and Chan and Young (2006). It provides a framework for incorporating the joint behavior of spot and futures prices with systematic jumps in prices. These jumps can arrive in varying frequencies and can be specified to be common across spot and futures prices. The specification of the bivariate GARCH-JUMP model is defined as follows:

$$\begin{bmatrix} r_t^c \\ r_t^f \end{bmatrix} = \begin{bmatrix} \mu_1 & \delta_s u_{t-1} & \phi_{ss1} r_{t-1}^c & \phi_{sf1} r_{t-1}^f & J_{st} & \varepsilon_{s,t} \\ \mu_2 & \delta_f u_{t-1} & \phi_{fs1} r_{t-1}^s & \phi_{ff1} r_{t-1}^f & J_{ft} & \varepsilon_{f,t} \end{bmatrix} + + + + + . \tag{14}$$

The common jump component J_t is defined as:

$$J_t = \sum_{i=1}^{n1} Y_{t,i} - E_{t-1}\sum_{i=1}^{n1} Y_{t,i} , \tag{15}$$

where J_t has a bivariate normal distribution with zero mean and variance matrix Δ_t. The disturbance term and the common jump component are assumed to be independent.

It is hypothesized that time-varying hedge ratios would be different across different variants of GARCH models. Therefore, the next question arises: which model is more effective in forecasting the stock futures' hedge ratio? In this paper we apply all the above methods to estimate the hedge ratio in four stock futures markets, and compare how effective they are at forecasting performance.

Because the hedge ratio is not an observed entity, the hedge ratio generated by GARCH is not an appropriate scale to measure the hedge ratio series forecasted with time variation. To assure forecast accuracy, examination of out-of-sample portfolio returns is a logical extension. In this paper, evaluation of forecast accuracy is conducted by forecasting out-of-sample portfolio returns implied by the forecasted hedge ratio. The portfolios are constructed as $r_t^c - h_t^* r_t^f$, where r_t^c is the log difference of the cash (spot) prices, r_t^f is the log difference of the futures prices, and h_t^* is the estimated or forecasted optimal hedge ratio. The issue of a missing benchmark to assess the accuracy of time-varying hedge ratio forecast can thus be avoided by comparing the portfolio returns forecasted with the actual returns of the portfolio.

Thus the forecasting and comparison are done via four steps. In the first step, the seven GARCH models are applied to forecast the hedge ratios during the two forecast horizons. In step two these forecasted hedge ratios are used to create the portfolio returns for the two forecast horizons based on the method provided above. Step three involves the application of the MCS and MDM testing procedures to compare the forecasting ability of each model compared to the others. Comparison based on the difference between the transaction cost-adjusted portfolio returns is conducted in step four.

3. 3.1. Model Confidence Set

The Model Confidence Set (MCS) methodology as proposed by Hansen, Lunde and Nason (2011) provides a testing method for the model selection, and for comparing forecasting ability of different models. The purpose of the MCS procedure is to delineate a set of best model(s), say M^*, from a collection of models and, say M^0, based on a data-congruent and user-specified criterion. According to Hansen *et al.* (2005; 2011), the MCS method has several advantages compared with other comparison techniques. First, the procedure does not require a benchmark model to be specified. It allows the flexibility whereby more than one model can be the superior model(s). Second, it takes into consideration the limitations of the data. Informative data will result in a MCS that contains only the best model. Third, the testing method enables one to draw inferences about the significance that is empirically valid in the traditional sense. This is a property that is not satisfied by the commonly used approach of reporting *p*-values from a multiple pairwise comparison. The Monte Carlo experiments demonstrate that the MCS has good small-sample properties.

For expository convenience, the following analysis relies heavily on Hansen *et al.* (2011). The MCS methodology has two step-testing procedures – an equivalence test δ_M and an elimination rule, e_M.[16] 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 the elimination rule e_M is used to eliminate a model with poor performance from M. The procedure is repeated until δ_M is accepted and MCS is now defined by the set of surviving models. The MCS method 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 model 'i' is the best 'candidate' in M^0 at significance level α. That is, through a sequential testing procedure, the set of 'surviving models' satisfies the assertion $\lim_{n\to\infty} P(M_i^* \in M_{i,1-\alpha}^*) \geq 1-\alpha$.

In the case of forecasting, the set of M^0 contains forecasting models with index i = 1, 2,…. m, and assumes 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 (MSE), and root mean square error (RMSE), and we denote the model i in time period t as $L_{i,t} = L(Y_t, \hat{Y}_{i,t})$, where $\hat{Y}_{i,t}$ is the point forecast of Y_t, with t = 1, 2, ….n. We also define the relative performance variables and an auxiliary variable

$$d_{ij,t} = L_{i,t} - L_{j,t} \text{ for all } i, j \in M^0$$

$$\mu_{ij} = E(d_{ij,t}).$$

In a pairwise comparison, the alternative model i is preferred to j if $\mu_{ij} < 0$. Therefore, the set of a superior model is defined by

$$M^* = \{i \in M^0 : \mu_{ij} \leq 0 \text{ for all } j \in M^0\}.$$

The null and alternative hypotheses take the forms

$$H_{o,M} : \mu_{ij} = 0 \qquad H_{A,M} : \mu_{ij} \neq 0 \qquad \text{for all } i, j \in M$$

The hypotheses are tested using two tests which are based on multiple t-statistics. If we define the relative sample loss statistics $\bar{d}_{ij} = n^{-1}\sum_{i=1}^{n} d_{ij,t}$ and $\bar{d}_i = m^{-1}\sum_{j\in m} d_{ij}$, where $\bar{d}_{ij}$ measures the relative sample loss between the i-th and j-th models, while $\bar{d}_i$ is the sample loss to the average across models in M, this enables us to construct the t-statistics

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\hat{\text{var}}(\bar{d}_{ij})}} \qquad \text{and} \qquad t_i = \frac{\bar{d}_i}{\sqrt{\hat{\text{var}}(\bar{d}_i)}} \qquad \text{for } i, j \in M,$$

where $\hat{\text{var}}(\bar{d}_{ij})$ and $\hat{\text{var}}(d_i)$ signify the estimates of $\text{var}(\bar{d}_{ij})$ and $\text{var}(\bar{d}_i)$, respectively. The equivalence test takes the form

$$\mu_{i1} = ... = \mu_{im} \qquad \Leftrightarrow \mu_{ij} = 0 \text{ for all } i, j \in M$$

$$\Leftrightarrow \mu_i = 0 \text{ for all } i \in M.$$

These two formulations of null hypotheses correspond to the test statistics as follows:

$$T_{\max,\,\mu} = \max_{i \in M} t_i \qquad \text{and } T_{R,\,M} = \max_{i,j \in M} |t_{ij}|,$$

which are available to test the hypothesis $H_{o,M}$. The elimination rules corresponding to those two test statistics are

$$e_{\max,\,M} = \arg\max_{i \in M} t_i \qquad e_{R,M} = \arg\max_{i \in M} \sup_{j \in M} t_{ij}.$$

As the distribution of each of the test statistics depends on unknown parameters, a block bootstrap procedure is used to estimate the distribution under the null. For example, under the null hypothesis, we set the p-value of H_0 as

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

where $I_{\{T\max > T*b,\max\}}$ is the indicator function and $T^*_{b,\max}$ is the estimated bootstrap distribution of max t_i. As stated, if $P_{H_0} > \alpha$, we accept the null hypothesis, and we have the $(1-\alpha)MCS$ $\hat{M}^*_{1-\alpha} = M$. In our study, all models that have P_{MCS} higher than 0.10 will belong to this set of best-performing models.

3.2 Diebold and Mariano Pairwise Tests

Diebold and Mariano (1995) develop a test of equal forecast accuracy to test whether two sets of forecast errors, say e_{1t} and e_{2t}, have equal mean value. Using MSE as the measure, the null hypothesis of equal forecast accuracy can be represented as $E[d_t] = 0$, where $d_t = e_{1t}^2 - e_{2t}^2$. Supposing, n, h-step-ahead forecasts have been generated, Diebold and

Mariano (1995) suggest that the mean of the difference between MSEs $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$ has an approximate asymptotic variance of

$$Var(\bar{d}) \approx \frac{1}{n}\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right],$$

where γ_k is the k-th autocovarianceof d_t, which can be estimated as:

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d}).$$

Therefore, the corresponding statistic for testing the equal forecast accuracy hypothesis is $S = \bar{d} / \sqrt{Var(\bar{d})}$, which has an asymptotic standard normal distribution. According to Diebold and Mariano (1995), results of Monte Carlo simulation experiments show that the performance of this statistic is good even for small samples and when forecast errors are non-normally distributed. However, this test is found to be over-sized for small numbers of forecast observations and forecasts of two-steps ahead or greater.

Harvey *et al.* (1997) further develop the test for equal forecast accuracy by modifying Diebold and Mariano's (1995) approach. Since the estimator used by Diebold and Mariano (1995) is consistent but biased, Harvey *et al.* (1997) improve the finite sample performance of the Diebold and Mariano (1995) test by using an approximately unbiased estimator of the variance of $\bar{d}$. The modified test statistic is given by

$$S* = \left[\frac{n + 1 - 2h + n^{-1}h(h-1)}{n}\right]^{1/2} S.$$

Through Monte Carlo simulation experiments, this modified statistic is found to perform much better than the original Diebold-Mariano test at all forecast horizons and when the forecast errors are autocorrelated or have non-normal distribution. In this paper, we apply the modified Diebold-Mariano test (MDM).

Two criteria – MSE and MAE derived from return forecasts – are employed to implement the MDM tests. On each occasion, the tests are conducted to detect superiority between two forecasting models, and thus there are 20 groups of tests for each forecast horizon for each market.

Each MDM test generates two statistics, S_1 and S_2, based on two hypotheses:

1. H_0^1: there is no statistical difference between the two sets of forecast errors.

   H_1^1: the first set of forecasting errors is significantly smaller than the second.

2. H_0^2: there is no statistical difference between the two sets of forecast errors.

   H_1^2: the second set of forecasting errors is significantly smaller than the first.

It is clear that the sum of the *P* values of the two statistics (S_1 and S_2) is equal to unity. If we define the significance of the modified Diebold-Mariano statistics as at least 10% significance level of *t* distribution, adjusted statistics provide three possible answers for superiority between two rival models:

1. If $S_1$is significant, then the first forecasting model outperforms the second.
2. If S_2 is significant, then the second forecasting model outperforms the first.
3. If neither S_1 nor S_2 is significant, then the two models produce equally accurate forecasts.

4. Literature Review

As stated earlier, the hedge ratio is the number of futures contracts needed to minimize the exposure of a unit's worth of position in the cash market. Studies in the late 1970s and the early 1980s employed the traditional regression analysis which assumes that the optimal hedge ratio is time-invariant (see Johnson 1960 and Ederington 1979). However, it is now well established that most asset return distributions are not normal, i.e. return distributions are time-varying with high skewness and high excess kurtosis. As a consequence, the hedge ratio is also changing over time (Sultan and Hasan 2008, p.469). The development of the generalised autoregressive conditional heteroscedastic (GARCH) modelling technique to deal with time-varying volatility has generated tremendous interest in the empirical investigations of the effectiveness of dynamic hedging that allows the hedge ratio to be time varying (Kroner and Sultan 1993). Based on the evidence of time-varying distributions of spot and future prices, the dynamic hedging strategy has been proven superior to any alternative hedging strategy that holds the hedge ratio constant. Therefore, a large body of empirical literature has accumulated since the late 1980s and up to recent years examining the issues of relative effectiveness of a sophisticated hedging method over much simpler and more intuitively appealing traditional hedging methods using currencies, commodities, stock indices, and interest rate products employing ARCH and GARCH specifications (Sultan and Hasan 2008, p. 470).

In previous studies different versions of the GARCH models have been used for forecasting volatility, time-varying beta, and hedge ratio among others, and then the models compared. [See Bauwens, Laurent and Rombouts (2006) for a survey article on GARCH model and Hansen and Lunde (2005) for a comparison on volatility forecast model; Poon and Granger (2003) also provide an excellent survey of GARCH's and other models' forecasting ability.] Given the dearth of literature, this section has drawn only from the area of hedge ratio forecasting using the GARCH class of model to furnish readers with an overview of the current state-of-the-art research.

Laurent *et al.* (2012) examine forecasting accuracy of 125 variants of GARCH models using 10 assets from the New York Stock Exchange employing one, five and twenty-day ahead conditional variance forecasts over a period of 10 years using model confidence set (MCS) and superior predictive test (SPA) tests. The study reports that during unstable periods such as the dot-com bubble, the superior models consist of sophisticated GARCH specifications such as orthogonal and dynamic conditional correlation (DCC) embedded with the leverage effect. During tranquil periods, GARCH with simple specifications such as constant conditional correlation and symmetry in the variance perform well. Finally, during the 2007-2008 financial crises, GARCH specification with non-stationarity in the conditional variance process generates superior forecast.

Zhang and Choudhry (2015) investigate the forecasting ability of five different GARCH models – bivariate GARCH, GARCH-BEKK, GARXH-X, BEKK-X, Q-GARCH based on four commodities – wheat, soybean, live cattle and live hogs. Their results indicate that BEKK-type models perform best in the cases of storable products such as wheat and soybean. The GARCH-GJR performs the best in the case of non-storable commodities, such as live cattle and hogs.

Zhang and Choudhry (2016) examine the forecasting performance of four variants of GARCH models – the GARCH-BEKK, GARCH-DCC, GARCH-MIDAS and Gaussian Copula GARCH –and the Kalman filter method during the pre-financial crisis and crisis periods using MCS to estimate time-varying betas. Their results are based on daily stock prices' data of two large banks from Austria, Belgium, Greece, Holland, Iceland, Italy, Portugal and Spain. Empirical results indicate that GARCH-BEKK performs the best during the pre-crisis period and the Kalman filter outperforms the GARCH models during the crisis period.

Choudhry and Hasan (2011) investigate the forecasting ability of five different variants of GARCH models – namely the bivariate GARCH, GARCH-ECM, GARCH-BEKK, GARCH-X and GARCH-GJR – using daily stock indices' futures data from December 1999 to December 2009 from Brazil, Hungary, South Africa and South Korea. Their results show that the BEKK model outperforms other models during the two-year forecast horizon and no model truly dominates in the shorter one-year forecast horizon. Overall, the bivariate GARCH model performs the worst. Their results also imply that the forecasting superiority of the model applied depends on the underlying market and the length of the forecast horizon.

The general impression from the foregoing discussion is that the forecasting accuracy of the optimal hedge ratio obtained from different variants of the GARCH model and the effectiveness of dynamic hedging is an issue of ongoing research to the financial practitioners and researchers. Given the paucity of research regarding the forecasting performance of time-varying hedge ratios based on alternative variants of the GARCH model, we have re-examined the issue using seven variants of the GARCH model to offer a more parsimonious time series approach using a longer time span and more recent data of both emerging and developed markets.

5. Data and diagnostics

The models are estimated using daily data spanning the period from January 2000 to July 2014 on stock cash indices and their counterpart futures contracts from Greece, Hungary, Poland and the UK. To avoid the sample effect and overlapping issue, two non-overlapping forecast horizons are considered – a one-year forecast horizon (July 2011-June 2012) and a two-year forecast horizon (July 2012-June 2014). All seven models are estimated for the periods January 2001 to June 2011 and January 2001 to June 2012, and the estimated parameters are applied for forecasting over the forecast samples 2011-2012 and 2012-2014, respectively. If the two forecasting periods are overlapping, the data for in-sample estimation are contaminated due to a certain amount of mutual data which may result in non-robust forecasting. According to Harri and Brorsen (2002) non-overlapping subsamples can enhance the reliability and robustness of outcome, and simplify the interpretation of results.

The FTSE/ASE Large Cap Index consists of 25 of the largest and most liquid stocks that trade on the Athens Stock Exchange. It was developed in September 1997 out of a partnership between the Athens Stock Exchange and FTSE International. The BUX index is the official capitalisation-weighted index of the blue-chip shares listed on the Budapest Stock Exchange (BSE).[17] Its futures and option products are available in the BSE derivatives section. The Warsaw Stock Exchange WIG20 index is the blue-chip index and consists of 20 of the biggest and the most liquid companies on the Warsaw Stock Exchange main list. In March 2014 the market value of all the companies on WIG20 amounted to 43.6% of the Warsaw Stock Exchange main list total capitalisation. The FTSE100 is a share index of the 100 companies listed on the London Stock Exchange with the highest market capitalisation. All futures price indices are continuous series.[18] The data are collected from Datastream International.

Descriptive statistics of the distribution of cash and futures returns indicate that the density function is negatively skewed for both cash and futures returns for all markets except Greece. The values of the excess kurtosis statistic are greater than 2 for all countries, which suggests that the density function for each country has a fat tail. The values of the Jarque-Bera statistic are high, suggesting that the returns are not normally distributed. Judging by the skewness, excess kurtosis and Jarque-Bera statistics, it can be inferred that the returns exhibit 'fat-tails' in all markets. The data series have also been checked for stationarity using the Augmented Dickey-Fuller (ADF) unit root test. The ADF test results indicate that each of the returns series has no unit root. Tests for autocorrelation in the first moments using the $Q(20)$ statistic indicate that none is present in any of the returns. Finally, tests for ARCH using Engle's (1982) LM statistic generally support the hypothesis of time-varying variances. These results are available from the authors on request.

6. GARCH, MCS and modified Diebold-Mariano tests results

Given that the GARCH results are quite standard, we only provide the GARCH results for Greece as an example. The remaining GARCH results are available on request. Table 1 reports the estimated coefficients of GARCH, BEKK, GARCH-X, GARCH-ECM, GARCH-GJR, GARCH-DCC and GARCH-JUMP models for Greece from 2000 to 2014.[19] All coefficients are significant at the 5% significance level (except γ_1 at the 10% level), while coefficients μ_1 and a_{22} in the GARCH-JUMP model are insignificant. In other words, the ARCH effect is insignificant in the futures equation of the GARCH-JUMP model. For all GARCH models except in the GARCH-JUMP futures equation, the sums of parameters of ARCH (a_{11} and a_{22}) and GARCH (b_{11} and b_{22}) terms are close to unity, which indicates that the impact of ARCH and GARCH on current conditional variance is persistent and that the volatility clustering dies out slowly. The unconditional variance/covariance terms (c_{11}, c_{12} and c_{22}) are small and significant, which suggests there are positive and significant interactions between the two log-prices. The GARCH-ECM model includes error correction in the mean equation, and the error term is found to be significant for Greece. The GARCH-GJR model captures asymmetric information effects for Greece with small positive and negative effects on cash and futures markets, respectively. The significance of square error correction of the GARCH-X model indicates that there exists cointegration between log-cash and log-futures prices. When $A+B<1$ for the GARCH-DCC model, this implies that it is mean-reverting. We obtain similar results for Hungary and Poland.

Table 2 reports basic statistics of forecasted hedge ratio (OHR) from the seven GARCH models for all three markets from July 2011 to June 2012 in Panel A and from July 2012 to June 2014 in Panel B, respectively. In Panel A, the average OHR for Poland ranges from -0.01722 (minimum) to 0.001258 (maximum), but the means of OHR are close to 1 for Greece, Hungary, and the UK. In other words, it is risker to hedge in Greece, Hungry, and the UK based on the high OHRs in these three countries. The variance of OHRs for Greece, Hungary and the UK are also higher than that of Poland. The skewness, kurtosis and J-B tests results indicate that the OHR series is non-normally distributed and slightly left-skewed, with a sharper peak for Poland. For the cases of Greece, Hungary and the UK, the OHR series have higher peaks and fatter tails than normal distribution in the one-year forecasting. In Panel B, the two-year forecasts of OHR series tell a very similar story to that in Panel A; but the OHR series of Greece is less skewed, and with more left-skewed OHR for Hungary and the UK.

In order to compare forecast accuracy of different variants of GARCH model, we employ both MCS and MDM Tests. Table 3 reports the RMSE's and the MCS *p*-values of the Model Confidence Set (MCS) on the forecasting accuracy of the seven GARCH models at both $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ confidence levels for all three markets. Since the MAD (mean absolute deviation) criterion is less sensitive to outliers which may lead to greater mis-predictions than the MSE-type method (Hansen *et al.* 2003), and because the RMSE method produces larger values than the MSE when the mean square error is smaller than unity, we report the Root Mean Square Error (RMSEs) and MCS *p*-values for each forecasting model. We find that a low RMSE 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. Moreover, the results in $\widehat{M}^*_{75\%} \subset \widehat{M}^*_{90\%}$ are as expected.

In the case of Poland, the GARCH-X, GARCH, and GARCH-JUMP are the top three models with smallest forecast errors at a 75% confidence level for one-year prediction. Among these three models, GARCH-X performs the best. In addition, during the two-year forecast horizon, the GARCH-GJR, GARCH-BEKK, GARCH-ECM and GARCH-X models outperform other models in the set. The GARCH-GJR model performs best among all four in the set with a 75% significance level. In the case of Greece, the GARCH-DCC performs the best during one-year forecast horizons. During the one-year forecast horizon, the $\widehat{M}^*_{75\%}$ contains some more models, i.e., GARCH-ECM and GARCH-JUMP. During the two-year horizon, GARCH-GJR and GARCH-BEKK models are significant within the $\widehat{M}^*_{75\%}$ confidence level, with GARCH-GJR being the best model in the set. The GARCH-DCC model outperforms all other models' forecasts for Hungary during the one-year horizon and the GARCH-BEKK produces the best forecasts for two-year prediction. In the case of the UK, GARCH-ECM and GARCH-X are the superior models for one-year and two-year predictions, respectively. In addition, GARCH-ECM, GARCH and GACRH-GJR constitute the set of surviving models at a 75% significance level in the one-year forecast; GARCH and GARCH-GJR belong to the surviving set at a 90% confidence level in the longer forecast horizon.

Generally, the GARCH model provides the most accurate forecasts given that it outperforms the other variants of GARCH model in four cases at the 75% confidence level and in five cases at the 90% confidence level. Each of the GARCH-ECM, GARCH-X and GARCH-GJR models constitutes model confidence set in four cases at the 90% confidence level (each with three cases at the 75% confidence level). GARCH-BEKK appeared in the model confidence set in three cases at the 75% level for the two-year forecast (Poland, Greece and Hungary). The GARCH-ECM and GARCH-DCC are ranked as the second and third best forecasting models for the one-year forecast horizon, respectively. For the two-year forecast horizon, GARCH-GJR appears to be the best model in two cases (Poland and

Greece). GARCH-X performs best in the two-year forecast horizon for Hungary and UK. The MCS results fail to point out any one particular type of GARCH model that has superior ability in forecasting the time-varying hedge ratio in these European stock futures markets.

In Table 4, we present the results of the MDM to compare prediction accuracy between any two GARCH models for Greece. Both MSE and MAE measurements are applied. The MAE is stricter than the MSE method since the MAE produces "better" or "worse" when the MSE yields an insignificant difference between two models, and we say they are "equally" good. For the one-year forecast, the GARCH-ECM is the superior model with the MSE method, while it is the second best after the DCC model under MAE. For the two-year forecast, both methods prefer the GARCH-GJR model, and the GARCH-BEKK is the second best under both methods.

A summary of the modified Diebold-Mariano and MCS tests results is presented in Table 5. We find that the best models selected from the modified Diebold-Mariano test are all included in the MCS test. In other words, these two test results are in line with each other and hence our findings are more persuasive.

Generally speaking, GARCH-GJR and GARCH-BEKK models could be the first and second best candidates for two-year forecast of OHR in emerging European markets. GARCH-X and GARCH-ECM outperform other competing models for Poland and UK in the one-year prediction, respectively. This result backs the claim by Poon and Granger (2003) that no one type of GARCH model is superior in forecasting; rather, superiority of forecasting performance depends upon several different factors. In this paper, results are different based on the market under consideration, and the length of the forecast horizons.

Figures 1 and 2 show the returns based on the forecast hedge ratios by all seven GARCH models and the actual returns over both forecast horizons for all four markets. 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.

7. Comparison of Returns

We provide further analysis by statistically comparing the difference between the forecasted returns after adjusting for transaction cost from each model during both forecasting horizons. This analysis may provide a way to choose the best model based on application rather than on a statistical criterion, such as MCS. The forecasted returns adjusted for transaction cost (c) are based on Kroner and Sultan's (1993) method which had assumed a transaction cost of 0.01%.[20] For example, the forecasted return from the GARCH model if we rebalance the futures position is $R = r_t^c - \hat{h}^*_{garch,t} r_t^f - c$, where r_t^c is the cash returns, r_t^f is the futures returns, $\hat{h}^*_{garch,t}$ is the hedge ratio forecasted by the GARCH model, and *c* is the transaction cost. We rebalance the futures positions if and only if the balanced position after accounting for the transaction cost yields higher return than the most recent balanced futures position. In this way, we achieve all the transaction cost-adjusted returns series during both forecasting horizons.

We test whether the transaction cost-adjusted returns from different models are significantly different from each other based on the MAE and MSE tests. For example, the difference between GARCH and BEKK models' forecasted returns is tested as $| R_{garch} - R_{bekk} |$ and $(R_{garch} - R_{bekk})^2$ being statistically different from zero, which is the paradigm of MAE and MSE tests. If the statistical value of MAE and MSE tests is significant, we can conclude that the return series from two models are statistically different from each other. Table 6 presents the results from the MAE and MSE tests. For each country, 21 different MAE and MSE tests are run for each of the two forecast horizons. All returns series are statistically different from each other at the 5% significance level during both the one-year and two-year forecast horizons. This result clearly indicates the importance of selecting the right model to forecast the returns.

Given that the differences between the returns are significant, Table 7 presents the mean value of the transaction cost-adjusted forecasted returns from the seven GARCH models for all countries during both forecast horizons. For Poland, GARCH-BEKK and GARCH-X provide the highest returns during one-year and two-year forecast horizons, respectively. GARCH-X and GARCH-BEKK provide the highest returns for Greece during the one-year and two-year forecast periods, respectively. For Hungary, GARCH-GJR provides the highest return during both the one-year and two-year periods. For UK, GARCH-BEKK yields highest returns in both horizons.

These results also fail to point out any one particular type of GARCH model that has superior ability in forecasting the time-varying hedge ratio in these European stock futures markets. In summary, the GARCH-BEKK and GARCH-GJR indicate the greatest effectiveness in terms of high forecasted returns. This result backs the forecasting accuracy of these models provided earlier by the modified Diebold-Mariano tests for longer forecast horizon.

8. Conclusion

This paper investigates the behaviour of dynamic hedge ratios in three emerging European stock futures markets using alternative variants of GARCH models and compares the forecasting performance of these GARCH models. Using daily data of the spot and futures markets of Greece, Hungary, Poland and the UK and the following seven models, GARCH, GARCH-BEKK, GARCH-ECM, GARCH-DCC, GARCH-X, GARCH-GJR and GARCH-JUMP, we have estimated the time-varying hedge ratios and compared the hedge ratio forecasting performances of these models. To the best of our knowledge no other paper has forecast the hedge ratios of emerging European stock futures markets. Ability to forecast the optimal hedge ratios/dynamic hedge ratios is important for understanding the role of the futures markets in trading, program trading, index arbitrage, and the development of optimal hedging and trading strategies in fund management. The forecasting of hedge ratios guides the hedger to choose the most appropriate portfolio and allows for portfolio adjustment in dynamic hedging.

The tests are carried out in three steps. In the first step we forecast the hedge ratio by means of the seven GARCH models. In the second step, we create the out-of-sample portfolio returns based on the forecasted hedge ratio by the seven models, and 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 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. Evaluation of forecast accuracy is conducted by forecasting out-of-sample returns of portfolios implied by the forecasted hedge ratios. The Model Confidence Set (MCS) and modified Diebold-Mariano tests are applied to compare the forecasting ability of the seven GARCH models. The MCS is applied based on two confidence levels, $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$. Application of the MCS approach makes this paper more unique in the literature.

Results from the MCS 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 these three emerging and one developed European futures markets. In summarising the MCS results, the GARCH model provides the most accurate forecasts given that it outperforms the other variants of GARCH model in four cases at the 75% confidence level and in five cases at the 90% confidence level. Each of the GARCH-ECM, GARCH-X and GARCH-GJR models constitutes model confidence set in four cases at the 90% confidence level. The GARCH-ECM and GARCH-DCC are ranked as the second and third best forecasting models for the 1-year forecast horizon, respectively. For the two-year forecast horizon, GARCH-GJR appears to be the best model in the cases of Poland and Greece, and GARCH-X performs best in the 2-year forecast horizon for the UK. The GARCH-BEKK model appeared in the model confidence set in three cases (Poland, Greece and Hungary) at the 75% level for the two-year forecast with the most superior model in the case of Hungary. The MCS results fail to point out any one particular type of GARCH model that has superior ability in forecasting the time-varying hedge ratio in these European stock futures markets. The modified Diebold-Mariano test results also indicate that the models selected from the MDM test are incorporated in the model confidence set and therefore accord well with the MCS test results.

We provide further analysis by statistically comparing the difference between the forecasted returns after adjusting for transaction cost from each model during both forecasting horizons. This analysis may provide a way to choose the best model based on application rather than on a statistical criterion, such as the MCS. In summary of these results, the GARCH-BEKK and GARCH-GJR models indicate the greatest effectiveness in terms of high forecasted returns. This result backs the forecasting accuracy of these models provided earlier by the modified Diebold-Mariano tests for a longer forecast horizon.

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

Acknowledgements

The authors would like to thank the editor, special issue editors, and two anonymous referees of the journals, as well as the participants of the 2016 Portsmouth-Fordham Conference on Banking and Finance at the University of Portsmouth for comments on an earlier version of the paper. Any remaining errors are the sole responsibility of the authors.

References

Alexander, C. and Barbosa, A. (2007) 'Effectiveness of minimum-variance hedging: the impact of electronic trading and exchange-traded funds', *Journal of Portfolio Management*, Vol. 33, No. 2, pp. 46-59.

Anderson, R. W. and Danthine, J. P. (1980) 'Hedging and joint production: theory and illustrations', *Journal of Finance*, Vol. 35, No. 2, pp. 489-497.

Bailey, O. (2010) *Developments in Emerging Equity Markets*, Bulletin-December Quarter 2010, Reserve Bank of Australia.

Baillie, R. and Myers, R. (1991) 'Bivariate GARCH estimates of the optimal commodity futures hedge', *Journal of Applied Econometrics*, Vol. 6, No. 2, pp. 109-124.

Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006) 'Multivariate GARCH models: a survey', *Journal of Applied Econometrics*, Vol. 21, pp. 79-109.

Cecchetti, S. G., Cumby, R. E. and Figlewski, S. (1988.) 'Estimation of optimal futures hedge', *Review of Economics and Statistics*, Vol. 70, No. 4, pp. 623-630.

Chan, W. H. (2008) 'Dynamic hedging with foreign currency futures in the presence of jumps', *Studies in Nonlinear Dynamic and Econometrics*, Vol. 12, Article 4.

Chan, W. H. and Young, D. (2006) 'Jumping hedges: an examination of movements in cooper spot and futures markets', *Journal of Futures Markets*, Vol. 26, pp. 169-188.

Chen, S., Lee, C. and Shrestha, K. (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*, Vol. 24, No. 4, pp. 359-386.

Choudhry, T. and Hasan, M. S. (2011) 'Forecasting the daily dynamic hedge ratios in emerging stock futures markets: evidence from the GARCH model', Paper presented at the Annual Meeting of the European Financial Management Association, Braga.

Choudhry, T. and Wu, H. (2008) 'Forecasting ability of GARCH vs Kalman filter method: evidence from daily UK time-varying beta', *Journal of Forecasting*, Vol. 27, No. 8, pp. 670-689.

Diebold, F. X. and Mariano, R. S. (1995) 'Comparing predictive accuracy', *Journal of Business and Economic Statistics*, Vol. 13, No. 3, pp. 253-263.

Ederington, L. (1979) 'The hedging performance of the new futures markets', *Journal of Finance*, Vol. 34, No. 1, pp. 157-170.

Engle, R. F. (2002) 'Dynamic conditional correlation: a simple class of multivariate GARCH models', *Journal of Business and Economics Statistics*, Vol. 20, pp. 339-350.

Engle, R. F. and Kroner, K. (1995) 'Multivariate simultaneous generalized ARCH', *Econometric Theory*, Vol. 11, No. 1, pp. 122-150.

Engle, R. F. and Ng, V. (1993) 'Measuring and testing the impact of news on volatility', *Journal of Finance*, Vol. 48, pp. 1749-1778.

Glosten, L., Jagannathan, R. and Runkle, D. E. (1993) 'On the relationship between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance*, Vol. 48, No. 5, pp. 1779-1801.

Hansen, P. R., Lunde, A. and Nason, J. M. (2003) 'Choosing the best volatility models: the model confidence set approach', *Oxford Bulletin of Economics and Statistics*, Vol. 65, S1, pp. 839-861.

Hansen, P. R, Lunde, A. and Nason, J. M. (2005) 'A forecast comparison of volatility models: does anything beat a GARCH(1,1)?', *Journal of Applied Econometrics*, Vol. 20, No. 7, pp. 873-889.

Hansen, P. R., Lunde, A. and Nason, J. M. (2011) 'The model confidence set', *Econometrica*, Vol. 79, No. 2, pp. 453-497.

Harri, A. and Brorsen, B. W. (2002) 'The overlapping data oroblem', Unpublished Working Paper, Oklahoma State University.

Harvey, D., Leybourne, S. J. and Newbold, P. (1997) 'Testing the equality of prediction mean squared errors', *International Journal of Forecasting*, Vol. 13, No. 2, pp. 281-291.

Hasan, M. S. and Choudhry, T. (2013) 'On the effectiveness of dynamic stock index portfolio hedging: evidence from emerging markets futures', in Batten, J. A., Mackay, P. and Wagner, N. (Eds): *Advances in Financial Risk Management-Corporates, Intermediaries and Portfolios*, pp. 364-390, Palgrave Macmillan, Basingstoke.

Johnson, L. (1960) 'The theory of hedging and speculation in commodity futures', *Review of Economic Studies*, Vol. 27, No. 3, pp. 139-151.

Kenourgios, D., Samitas, A. and Drosos, P. (2008) 'Hedge ratio estimation and hedging effectiveness: the case of the S&P 500 stock index futures contract', *International Journal of Risk Assessment and Management*, No. 9, No. ½, pp. 121-134.

Kroner, K. F. and Sultan, J. (1993) 'Time-varying distributions and dynamic hedging with foreign currency futures', *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 4, pp. 535-551.

Lai, Y., Chen, C. W. S. and Gerlach, R. (2009) 'Optimal dynamic hedging via Copula-threshold-GARCH models', *Mathematics and Computers in Simulation*, Vol. 79, No. 8, pp. 2609-2624.

Laurent, S., Rombouts, J. V. K. and Violante, F. (2012) 'On the forecasting accuracy of multivariate GARCH models', *Journal of Applied Econometrics*, Vol. 27, pp. 934-955.

Lee, T. (1994) 'Spread and volatility in spot and forward exchange rates', *Journal of International Money and Finance*, Vol. 13, No. 3, pp. 375-383.

Moon, G. H., Yu, W. C. and Hong, C. H. (2009) 'Dynamic hedging performance with the evaluation of multivariate GARCH models: evidence from KOSTAR index futures', *Applied Economics Letters*, Vol 16, No. 9, pp. 913-919.

Park, T. A. and Antonovitz, F. (1992) 'Econometric tests of firm decision making under uncertainty: optimal output and hedging decisions', *Southern Economic Journal*, Vol. 58, No. 3, pp. 593-609.

Parks, P. C. (1992) 'AM Lyapunov's stability theory-100 years On', *IMA Journal of Mathematical Control and Information*, Vol. 9, No. 4, pp. 275-303.

Peters, T. (2008) *Forecasting the Covariance Matrix with the DCC GARCH Model*, Mathematical Statistics Stockholm University dissertation paper.

Poon, S. and Granger, C. W. J. (2003) 'Forecasting volatility in financial markets: a review', *Journal of Economic Literature*, Vol. 41, No. 2, pp. 478-539.

Stengos, T. and Panas, E. (1992) 'Testing the efficiency of the Athens Stock Exchange: some results from the banking sector', *Empirical Economics*, Vol. 17, No. 2, pp. 239-252.

Sultan, J. and Hasan, M. S. (2008) 'The effectiveness of dynamic hedging: evidence from selected European stock index futures', *European Journal of Finance*, Vol. 14, No.6, pp. 469-488.

Tse, Y. K. and Tusi, A. K. C. (2002) 'A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations', *Journal of Business and Economics Statistics*, Vol. 20, pp. 351-362.

Working, H. (1953) 'Futures trading and hedging', *American Economic Review*, Vol. 43, No. 3, pp. 314-343.

Yang, M. J. and Lai, Y. C. (2009) 'An out-of-sample comparative analysis of hedging performance of stock index futures: dynamic versus static hedging', *Applied Economics*, No. 19, No. 13, pp. 1059-1072.

Zhang, Y. and Choudhry, T. (2015) 'Forecasting the daily dynamic hedge ratios by GARCH models: evidence from the agricultural futures markets', *European Journal of Finance*, Vol. 21, No. 4, pp. 376-399.

Zhang, Y. and Choudhry, T. (2016) 'Forecasting the daily time-varying beta of European banks during the crisis period: comparison between GARCH models and the Kalman filter', *Journal of Forecasting*, DOI: 10.1002/for.2442.

Notes

1. The hedge ratio is simply the number of futures contracts needed to minimise the exposure of a unit's worth position in the cash market.
2. Zhang and Choudhry (2015) is an exception. They estimate and forecast the hedge ratios in commodities' futures markets.
3. However, studies of Alexander and Barbosa (2007), Hasan and Choudhry (2013), Lai *et al.* (2009) and Moon *et al.* (2009) are a few exceptions, but do not forecast the hedge ratios.
4. In the post-GARCH era, the issue of dynamic hedging received much attention and acceptance due to the ability of the GARCH models to account for nonlinearity, volatility cluster, non-normality and time dependency in variance/covariance of portfolio and futures returns. For this reason, we have chosen to study the forecasting accuracy of the dynamic hedge ratios using a framework of the GARCH class of models.
5. Forecasting in this paper is conducted based on the rolling forecast.
6. Stengos and Panas (1992) investigate the efficient market hypothesis in the Athens Stock Exchange for a number of selected stocks from the banking sector. The study finds support for the 'weak' and 'semi-strong' versions of the efficient market hypothesis.
7. Longer period also includes changes in the market environment or unexpected events which render these assumptions less reasonable over a longer time horizon.
8. Choudhry and Wu (2008) and Zhang and Choudhry (2015) also apply the same procedure.
9. The OLS estimation of the hedge ratio from equation (2) is based on the assumption of time invariant asset distributions suggested by Ederington (1979) and Anderson and Danthine (1980).
10. In this study, the sample size of the four futures 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.
11. This section has drawn extensively from Hasan and Choudhry (2013).
12. Kenourgios *et al.* (2008) show that ECM-GARCH outperforms simple error correction representation, GARCH and EGARCH models, in capturing properties of the hedge ratios on S&P500 stock index futures.
13. For example, a bivariate BEKK GARCH(1,1) parameterisation requires the estimation of only 11 parameters in superiority of BEKK over the GARCH-DCC.
14. There is more than one GARCH model available that is able to capture the asymmetric effect in volatility. According to Engle and Ng (1993), the Glosten *et al.* (1993) model is the best at parsimoniously capturing this asymmetric effect.
15. Peters (2008) shows that the DCC model outperforms naive sample on predicting a covariance matrix during in a short-run frame with high persistence.
16. The equivalence test is similar to the equal predictive ability (EPA) 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.
17. The Budapest Commodity Exchange (BCE) and the Budapest Stock Exchange (BSE) merged in October 2005, which made the BSE one of the main derivatives' centres in Central Europe. The BSE played a significant role in the privatisation of many leading Hungarian companies. The BSE was one of the first in the world which started to use free-float capitalisation weighting instead of the traditional market capitalisation weighting in October 1999.
18. The continuous series is a perpetual series of futures prices. It starts at the nearest contract month, which forms the first 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.

19. The BHHH algorithm is used as the optimisation method to estimate the GARCH, ECM-GARCH and GARCH-GJR models, and the BFGS algorithm is applied for the rest of the models.
20. Yang and Lai (2009) noted that the transaction cost ranges between 0.005% and 0.01% in the major global exchanges which are trading financial contracts of DJIA, S&P500,, NASDAQ100, FTSE100, CAC40, DAX30 and Nikkei225.

Table 1 GARCH models Results for Greece from 1 January 2002 -30 June 2014

Variable	GARCH	BEKK	GARCH-X	ECM	GJR	DCC	JUMP	
μ_1	6.22e-004 [0.0091]	0.00056 [0.0244]	-4.92e-004 [0.0000]	6.15e-004 [0.0028]	5.95e-004 [0.0129]	6.13e-004 [0.0115]	0.03943 [0.1296]	
δ_1			0.07430 [0.0000]	6.29e-004 [0.0030]			0.07896 [0.00000]	
μ_2	6.53e-004 [0.0098]	0.00058 [0.0362]	-2.69e-004 [0.0000]	0.16964 [0.0000]	5.60e-004 [0.0285]	7.17e-004 [0.0041]		0.04992 [0.0614]
δ_2			-0.07530 [0.0000]	0.19573 [0.0000]				0.03259 [0.0813]
c_{11}	2.36e-006 [0.0000]	0.00113 [0.0000]	4.62e-004 [0.0000]	2.54e-006 [0.0000]	2.36e-006 [0.0000]	1.53e-006 [0.0000]	0.01765 [0.0009]	0.02369 [0.0001]
a_{11}	0.06518 [0.0000]	0.97206 [0.0000]	0.50137 [0.0000]	0.06492 [0.0000]	0.06600 [0.0000]	0.07983 [0.0000]	0.07162 [0.0000]	0.07307 [0.0000]
b_{11}	0.92877 [0.0000]	0.23207 [0.0000]	0.50072 [0.0000]	0.92828 [0.0000]	0.92930 [0.0000]	0.92019 [0.0000]	0.91579 [0.0000]	0.91378 [0.0000]
$d_{11}(\gamma_1)$			1.40e-003 [0.0000]		-2.95e-003 [0.0927]			
c_{22}	2.78e-006 [0.0000]	-0.00047 [0.0000]	4.62e-004 [0.0000]	2.853e-006 [0.0000]	2.88e-006 [0.0000]	1.65e-006 [0.0456]	3.59188 [0.0000]	4.57732 [0.0000]
a_{22}	0.06996 [0.0000]	0.96833 [0.0000]	0.50130 [0.0000]	0.06791 [0.0000]	0.06610 [0.0000]	0.07605 [0.0000]	0.08125 [0.8818]	0.07704 [0.9341]
b_{22}	0.92529 [0.0000]	0.24838 [0.0000]	0.50080 [0.0000]	0.92657 [0.0000]	0.92480 [0.0000]	0.92482 [0.0000]	0.03188 [0.0302]	0.02445 [0.0039]
$d_{22}(\gamma_2)$			1.44e-003 [0.0000]		7.28e-003 [0.0000]			
c_{12}	2.46e-006 [0.0000]	0.00123 [0.0000]	4.60e-004 [0.0000]	2.60e-006 [0.0000]	2.50e-006 [0.0000]			
a_{12}	0.06653 [0.0000]		0.49930 [0.0000]	0.06535 [0.0000]	0.06610 [0.0000]			
b_{12}	0.92746 [0.0000]		0.49960 [0.0000]	0.92793 [0.0000]	0.92760 [0.0000]			
d_{12}			-7.13e-004 [0.0000]					
A						0.04223 [0.0000]	0.04115 [0.0000]	
B						0.95302 [0.0000]	0.94943 [0.0000]	
LLF	20564.16	20572.56	19308.61	20651.98	20579.19	-5805.58	-6361.57	

Note: the *p*-value in square parentheses. The $d_{11}(\gamma_1)$ and $d_{22}(\gamma_2)$ for GARCH-X and GARCH-GJR models represent the short-run deviations from the long-run relationship between the cash and futures prices and asymmetric impacts of information respectively; A and B are parameters of dynamic conditional correlations. In the DCC-JUMP model, we use a two-step procedure in which univariate GARCH models are estimated for cash and futures return with possion-distributed jump, and then we estimate dynamic conditional correlation between these two returns.

Table 2 Basic Statistics of forecasted OHR

Models	Mean	Variance	Skewness	Kurtosis	J-B
Panel A Forecasted OHR from July 2011 to June 2012					
Poland					
GARCH	0.01258	0.00105	-0.85529*	15.9507*	2809.43*
BEKK-GARCH	0.00292	0.01830	0.71457*	0.21339*	1.62235*
ECM-GARCH	0.00325	0.00103	-0.01223	-0.18987	0.40011
GARCH-X	-0.01722	0.00049	-0.53712*	1.14013*	26.7885*
GARCH-GJR	0.00069	0.00222	-0.05066	2.50911*	68.5768*
GARCH-DCC	-0.00573	.000004	-0.82199*	-0.33889	30.6410*
GARCH-JUMP	0.01159	0.00002	1.64984*	3.50773*	2394.59*
Greece					
GARCH	0.89595	0.00356	-0.69475*	7.49785*	634.788*
BEKK-GARCH	-0.47702	59.6145	0.35635*	1.60233*	33.5736*
ECM-GARCH	0.84978	0.00270	-2.33071*	9.39098*	1199.95*
GARCH-X	0.90857	0.00462	-1.63878*	3.01181*	216.378*
GARCH-GJR	2.02532	0.02817	-1.23526*	2.24682*	121.740*
GARCH-DCC	0.73063	0.08615	1.55886*	4.03170*	283.558*
GARCH-JUMP	0.75741	0.07538	1.20128*	2.27577*	119.553*
Hungary					
GARCH	0.99542	0.00235	2.30019*	6.87723*	747.353*
BEKK-GARCH	0.12629	8.54957	-0.20181	1.44792*	24.4766*
ECM-GARCH	0.99623	.000007	-2.51187*	8.05249*	983.381*
GARCH-X	1.01395	0.00101	3.83529*	14.6690*	2991.37*
GARCH-GJR	1.09312	0.00393	0.54742*	1.26255*	30.4872*
GARCH-DCC	0.98493	0.08950	0.72123*	-0.42661	24.7012*
GARCH-JUMP	1.17393	0.12987	0.67658	-0.54765**	23.0854*
UK					
GARCH	0.97218	0.00043	-2.26954*	8.10985*	939.306*
BEKK-GARCH	1.00852	0.04465	16.0333*	258.384*	737227*
ECM-GARCH	0.90776	0.00028	1.33786*	1.58510*	105.183*
GARCH-X	0.99975	0.00016	1.09365*	5.50986*	382.179*
GARCH-GJR	0.98313	0.00014	-2.37271*	6.39869*	690.152*
GARCH-DCC	0.99040	0.91916	1.62637*	1.82391*	151.239*
GARCH-JUMP	1.41427	3.07472	2.69715*	7.29013*	897.838*
Panel B Forecasted OHR from July 2012 to June 2014					
Poland					
GARCH	0.01954	0.00083	0.04072	7.79323*	1318.58*
BEKK-GARCH	-0.00593	0.00912	0.12424	-0.08811	-0.08811
ECM-GARCH	0.00672	0.00042	-0.12855*	1.43648*	46.2299*
GARCH-X	-0.01203	0.00309	-0.28278*	8.49466*	1573.40*
GARCH-GJR	-0.00164	0.00005	0.17601	3.12839*	215.146*
GARCH-DCC	-0.00475	0.00000	-2.13379*	8.09088*	1816.43*
GARCH-JUMP	-0.00849	.000004	-1.05989*	0.90741*	115.421*
Greece					
GARCH	1.71951	5.00114	2.34169*	6.22534*	1317.45*
BEKK-GARCH	0.83166	0.00409	-0.06966	2.54649*	141.192*
ECM-GARCH	0.89378	0.00046	-1.49010*	2.12457*	290.792*
GARCH-X	0.89076	0.00638	-0.07183	-0.07183*	19.7767*
GARCH-GJR	0.55882	0.02636	0.56104*	0.24728	28.6596*

GARCH-DCC	1.18343	0.11354	0.79800*	-0.27880	56.9843*
GARCH-JUMP	1.19332	0.11326	0.11326*	-0.45387*	47.7996*
Hungary					
GARCH	0.85084	0.00536	-1.71627*	4.13219*	626.445*
BEKK-GARCH	0.56015	0.33198	-0.80966*	3.74755*	361.798*
ECM-GARCH	1.00109	.000001	-3.52009*	19.5979*	9413.64*
GARCH-X	0.99694	0.00078	-19.4113*	405.419*	3600803*
GARCH-GJR	1.17351	0.00312	-7.17938*	82.8376*	153439*
GARCH-DCC	1.01520	0.04262	1.72828*	3.74724*	564.194*
GARCH-JUMP	1.02564	0.04196	1.68353*	3.52642*	516.067*
UK					
GARCH	0.99308	0.000006	-4.32020*	26.1339*	16447.1*
BEKK-GARCH	0.99910	0.00001	-0.86582*	4.72799*	549.303*
ECM-GARCH	1.00015	0.000004	-2.68025*	9.97747*	2784.85*
GARCH-X	0.97142	0.00129	-0.92494*	2.18821*	178.233*
GARCH-GJR	0.98888	0.00003	-4.64886*	32.2367	24436.1*
GARCH-DCC	1.49113	0.19431	0.41312*	-0.08226	14.9662*
GARCH-JUMP	1.00586	1.00586	-0.12474*	1.95930*	529.089*

Note: * and ** imply significance at the 5% and 10% levels, respectively.

Table 3

The MCS result of forecasted returns from six GARCH models from 1 July 2011 to 30 June 2012 and from 1 July 2012 to 30 June 2014 for Poland, Greece and Hungary

Horizons	01 July 2011- 30 June 2012		01 July 2012- 30 June 2014	
	RMSE	MCS *P*-value	RMSE	MCS *P*-value
Poland				
GARCH	1.62218	0.9995**	3.43027	0.0355
BEKK-GARCH	1.62854	3.0e-04	2.26361	0.2380*
ECM-GARCH	1.62484	5.0e-04	2.28748	0.1530*
GARCH-X	1.62041	1.0000**	2.28747	0.1530*
GARCH-GJR	1.62693	4.0e-04	2.20626	1.0000**
GARCH-DCC	1.62403	0.0018	2.42847	0.0405
GARCH-JUMP	1.62235	0.9918**	2.43228	0.0405
Greece				

GARCH	3.55601	0.0210	3.43027	0.0000
BEKK-GARCH	8.17838	0.0024	2.26361	0.9808**
ECM-GARCH	3.54233	0.9970**	2.28748	0.0135
GARCH-X	3.55782	0.0210	2.28747	0.0210
GARCH-GJR	4.08702	0.0016	2.20626	1.0000**
GARCH-DCC	3.54035	1.0000**	2.42847	0.0132
GARCH-JUMP	3.54681	0.3867**	2.43228	0.0000
Hungary				
GARCH	2.10959	0.8182**	1.30593	0.3933**
BEKK-GARCH	3.31353	0.0130	1.30108	1.0000**
ECM-GARCH	2.11150	0.9332**	1.40640	4.0e-04
GARCH-X	2.12103	0.0733	1.40359	0.2064**
GARCH-GJR	2.16012	0.0733	1.53104	0.0000
GARCH-DCC	2.10850	1.0000**	1.43820	4.0e-04
GARCH-JUMP	2.21896	0.0733	1.44543	0.0689
UK				
GARCH	1.67276	0.7962**	1.19292	0.1175*
BEKK-GARCH	1.69792	0.0000	1.19888	0.0121
ECM-GARCH	1.63447	1.0000**	1.19865	0.0186
GARCH-X	1.69007	0.0973	1.17536	1.0000**
GARCH-GJR	1.68040	0.7962**	1.18957	0.1221*
GARCH-DCC	1.96301	0.0000	1.67403	0.0000
GARCH-JUMP	2.68952	0.0000	1.20425	0.0000

Note: The MCS *p*-values that are marked with * and ** are those in $\widehat{M}^*_{90\%}$ and $\widehat{M}^*_{75\%}$ correspondingly.

Table 4

MDM(Modified Diebold Mariano) test of forecasted error from GARCH models with normal distribution for Greece.

(M)DM test of forecasted return for Greece				
Measurement Models	one-year forecast		two-year forecast	
	MSE	MAE	MSE	MAE
GARCH vs. BEKK	>	>	<	<
GARCH vs. GARCH-X	=	=	<	<
GARCH vs. GARCH-ECM	<	<	<	<
GARCH vs. GARCH-GJR	>	>	<	<
GARCH vs.GARCH-DCC	=	<	<	<
GARCH vs. GARCH-JUMP	=	<	<	<
BEKK vs. GARCH-X	<	<	>	>
BEKK vs. GARCH-ECM	<	<	>	>
BEKK vs. GARCH-GJR	<	<	<	<
BEKK vs.GARCH-DCC	<	<	>	>
BEKK vs. GARCH-JUMP	<	<	>	>
GARCH-Xvs.GARCH-ECM	<	<	=	=
GARCH-X vs.GARCH-GJR	>	>	<	<
GARCH-X vs.-GARCH-DCC	=	<	>	>
GARCH-X vs. GARCH-JUMP	=	<	>	>
GARCH-ECM vs. GARCH-GJR	>	>	<	<
GARCH-ECM vs. GARCH-DCC	=	=	>	>

GARCH-ECM vs. GARCH-JUMP	=	=	>	>
GARCH-GJR vs. GARCH-DCC	<	<	>	>
GARCH-GJR vs. GARCH-JUMP	<	<	>	>
GARCH-DCC vs. GARCH-JUMP	=	>	=	>

Note: >, < and = represent that the prior model outperform, underperform the latter one and it is insignificantly different from each other, respectively.

Table 5 Summary of MDM and MCS tests

Panel A MDM test				
Market	one-year forecast of return		two-year forecast of return	
	MSE	MAE	MSE	MAE
Poland	X	X	GJR	GJR
Greece	ECM	DCC, ECM	GJR	GJR
Hungary	DCC	DCC	BEKK	BEKK
UK	ECM	ECM	X, GJR	X,GJR
Panel B MCS (Model confidence sets)				
Market	one-year forecast of return		two-year forecast of return	
Poland	X,GARCH,JUMP		GJR,BEKK,ECM,X	
Greece	DCC, ECM		GJR, BEKK	
Hungary	DCC,GARCH,ECM		BEKK,GARCH	
UK	ECM,GARCH,GJR		X,GJR,GARCH	

Table 6

The *t*-statistic values of one-year and two-year return differences between any two models based on MAE and MSE approaches with transaction costs

Horizon	one-year returns		two2-year returns	
model	MAE	MSE	MAE	MSE
Poland				
GARCH vs. BEKK	8.06399*	2.94199*	10.76057*	2.65828*
GARCH vs. GARCH-X	6.60445*	1.85906**	9.98922*	2.75284*
GARCH vs. GARCH-ECM	7.43787*	1.93949**	14.41874*	5.11256*
GARCH vs. GARCH-GJR	6.37187*	1.60611	10.22379*	2.69770*
GARCH vs. GARCH-DCC	6.36599*	1.97172*	10.70681*	2.77652*
GARCH vs. GARCH-JUMP	5.10348*	1.77783**	11.53454*	2.94018*
BEKK vs. GARCH-X	8.57518*	3.47916*	12.43313*	3.71529*
BEKK vs. GARCH-ECM	8.53660*	3.50507*	10.87245*	2.74619*
BEKK vs. GARCH-GJR	8.44514*	3.29570*	12.28435*	3.54855*
BEKK vs. GARCH-DCC	8.41961*	3.53539*	12.14018*	3.44253*
BEKK vs. GARCH-JUMP	8.74055*	3.63889*	12.07166*	3.39689*
GARCH-Xvs.GARCH-ECM	11.04974*	4.20251*	10.27734*	3.13906*
GARCH-X vs.GARCH-GJR	7.17681*	1.71367*	14.57496*	5.18786*
GARCH-X vs. GARCH-DCC	8.91640*	2.68533*	14.49757*	5.26539*
GARCH-X vs. GARCH-JUMP	9.28014*	3.08725*	14.71285*	5.25401*
GARCH-ECMvs.GARCH-GJR	8.01653*	1.91181**	9.63062*	3.02722*
GARCH-ECM vs. GARCH-DCC	10.81146*	3.94871*	9.59839*	2.99737*
GARCH-ECM vs. GARCH-JUMP	12.80637*	5.87763*	9.61311*	3.00033*

GARCH-GJR vs. GARCH-DCC	7.65791*	1.84811*	13.92835*	5.45992*
GARCH-GJR vs. GARCH-JUMP	7.75896*	2.25029*	14.81834*	5.17742*
DCC-GARCH vs. GARCH-JUMP	12.75868*	5.40987*	16.58513*	6.29343*
Greece				
GARCH vs. BEKK	9.15712*	4.22039*	9.15712*	4.22039*
GARCH vs. GARCH-X	9.12364*	3.96811*	9.12364*	3.96811*
GARCH vs. GARCH-ECM	9.98748*	3.81031*	9.98748*	3.81031*
GARCH vs. GARCH-GJR	15.48602*	5.69462*	15.48602*	5.69462*
GARCH vs. GARCH-DCC	7.93239*	2.43865*	7.93239*	2.43865*
GARCH vs. GARCH-JUMP	8.22864*	2.77665*	8.22864*	2.77665*
BEKK vs. GARCH-X	9.12612*	4.18990*	9.12612*	4.18990*
BEKK vs. GARCH-ECM	9.15796*	4.22137*	9.15796*	4.22137*
BEKK vs. GARCH-GJR	9.68935*	4.40805*	9.68935*	4.40805*
BEKK vs. GARCH-DCC	9.24522*	4.39673*	9.24522*	4.39673*
BEKK vs. GARCH-JUMP	9.23928*	4.37077*	9.23928*	4.37077*
GARCH-Xvs.GARCH-ECM	11.16933*	4.55829*	11.16933*	4.55829*
GARCH-X vs.GARCH-GJR	15.44705*	5.87776*	15.44705*	5.87776*
GARCH-X vs. GARCH-DCC	7.41126*	2.23584*	7.41126*	2.23584*
GARCH-X vs. GARCH-JUMP	7.77408*	2.51145*	7.77408*	2.51145*
GARCH-ECMvs.GARCH-GJR	15.52706*	5.77295*	15.52706*	5.77295*
GARCH-ECM vs. GARCH-DCC	8.31743*	2.63859*	8.31743*	2.63859*
GARCH-ECM vs. GARCH-JUMP	8.67490*	3.02593*	8.67490*	3.02593*
GARCH-GJR vs. GARCH-DCC	14.90226*	4.75762*	14.90226*	4.75762*
GARCH-GJR vs. GARCH-JUMP	14.71840*	4.64995*	14.71840*	4.64995*
DCC-GARCH vs. GARCH-JUMP	8.38648*	2.26182*	8.38648*	2.26182*
Hungary				
GARCH vs. BEKK	9.93271*	4.86303*	13.07326*	2.96889*
GARCH vs. GARCH-X	8.47295*	3.01724*	16.92977*	5.25520*
GARCH vs. GARCH-ECM	9.76620*	3.63893*	15.69008*	4.69041*
GARCH vs. GARCH-GJR	13.02845*	5.69400*	21.38545*	7.38965*
GARCH vs. GARCH-DCC	12.78613*	6.20229*	10.57790*	3.78992*
GARCH vs. GARCH-JUMP	6.20229*	5.69487*	10.71311*	3.84708*
BEKK vs. GARCH-X	10.02962*	4.93780*	13.46949*	2.78617*
BEKK vs. GARCH-ECM	10.03614*	4.94215*	13.46037*	2.78577*
BEKK vs. GARCH-GJR	9.99552*	4.90064*	14.90500*	3.02254*
BEKK vs. GARCH-DCC	9.76999*	4.71208*	11.89416*	2.56707

BEKK vs. GARCH-JUMP	9.81453*	4.68105*	11.98361*	2.58093
GARCH-Xvs.GARCH-ECM	7.75553*	3.18634*	1.50237	1.00207
GARCH-X vs.GARCH-GJR	11.44965*	4.68776*	23.36169*	10.89808*
GARCH-X vs. GARCH-DCC	12.26972*	5.54574*	11.59623*	3.84166*
GARCH-X vs. GARCH-JUMP	10.22659*	5.44041*	11.48575*	3.86390*
GARCH-ECMvs.GARCH-GJR	11.39013*	4.66721*	19.37696*	4.14699*
GARCH-ECM vs. GARCH-DCC	12.55197*	5.66252*	11.29832*	3.95019*
GARCH-ECM vs. GARCH-JUMP	10.34588*	5.43056*	11.19112*	3.97061*
GARCH-GJR vs. GARCH-DCC	14.37652*	6.75520*	17.84009*	7.10806*
GARCH-GJR vs. GARCH-JUMP	11.13233*	5.57021*	17.58360*	6.99232*
GARCH-DCC vs. GARCH-JUMP	14.32880*	6.91991*	22.41342*	9.85487*
UK				
GARCH vs. BEKK	2.61483*	1.04029	18.86601*	7.97946*
GARCH vs. GARCH-X	16.42792*	7.89637*	23.93676*	11.39474*
GARCH vs. GARCH-ECM	9.65045*	3.36398*	13.57086*	4.40393*
GARCH vs. GARCH-GJR	7.66165*	2.62679*	14.77752*	5.49142*
GARCH vs. GARCH-DCC	9.28450*	3.38804*	16.02728*	6.52568*
GARCH vs. GARCH-JUMP	6.13515*	2.89315*	16.60063*	6.05627*
BEKK vs. GARCH-X	5.38538*	1.13996*	14.35796*	6.45847*
BEKK vs. GARCH-ECM	1.61520*	1.00593*	14.16196*	4.62882*
BEKK vs. GARCH-GJR	2.06868*	1.01681*	17.72814*	7.34491*
BEKK vs. GARCH-DCC	9.47420*	3.46555*	15.93696*	6.48621*
BEKK vs. GARCH-JUMP	6.22617*	2.89565*	15.55821*	5.33176*
GARCH-Xvs.GARCH-ECM	17.71236*	9.18596*	14.33977*	4.58551*
GARCH-X vs.GARCH-GJR	18.01952*	9.56535*	20.82805*	8.89108*
GARCH-X vs. GARCH-DCC	8.86651*	3.34732*	15.97773*	6.49989*
GARCH-X vs. GARCH-JUMP	6.04884*	2.89900*	15.27657*	5.01483*
GARCH-ECMvs.GARCH-GJR	9.12039*	3.87387*	13.56677*	4.56190*
GARCH-ECM vs. GARCH-DCC	9.40777*	3.35735*	16.05631*	6.50027*
GARCH-ECM vs. GARCH-JUMP	6.14838*	2.86998*	17.09433*	6.47736*
GARCH-GJR vs. GARCH-DCC	9.30333*	3.36408*	16.06653*	6.54984*
GARCH-GJR vs. GARCH-JUMP	6.13009*	2.87360*	17.61973*	7.15780*
GARCH-DCC vs. GARCH-JUMP	5.62051*	2.81245*	15.82613*	6.40764*

Note: We test the return difference between the models by means of MAE and MSE statistics. For example we test GARCH and BEKK models by testing if $|R_{garch} - R_{bekk}|$ and $(R_{garch} - R_{bekk})^2$ are statistically different from zero, which is the paradigm of MAE and MSE tests. If the statistical values of MAE and MSE tests are significant, we can conclude that the return series from the two models are statistically different from each other.

The *t*-statistics of the MAE and MSE tests with null hypothesis of difference between return series are 0 and the rejection of null indicates that the return difference between two models is statistically significantly different from zero. '*' and '**' represent that the statistical value are significant at 5% or 10% levels, respectively.

Table 7

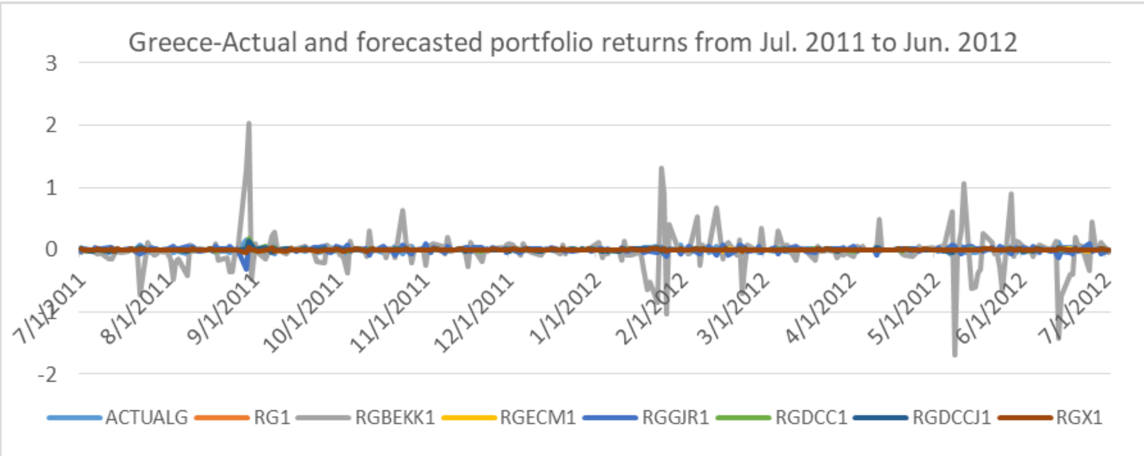
The mean value of one-year and two-year forecasted return series from seven GARCH models with transaction costs (TCs)

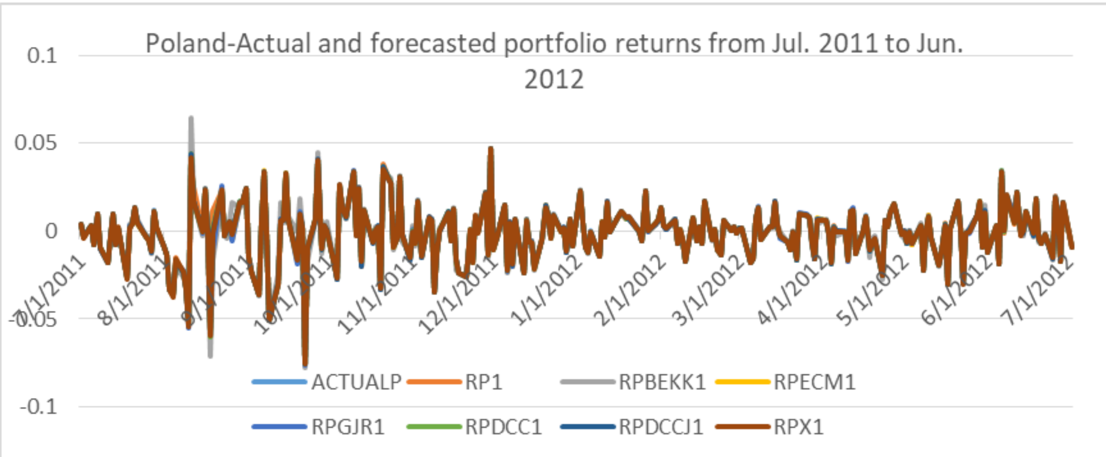
Models / Country Horizon		Mean of forecasted returns from the following models						
		GARCH	BEKK	ECM	GARCH-X	GJR	DCC	JUMP
Poland	1-year	-8.8e-04 (51)	1.6e-04 (21)	-9.1e-04 (88)	-9.3e-04(39)	-8.9e-04(54)	-9.0e-04 (7)	-8.9e-04 (30)
	2-year	3.1e-05 (30)	1.3e-05(29)	-7.0e-06(27)	7.8e-05(60)	3.0e-06(51)	9.0e-06(25)	-1.8e-05 (66)
Greece	1-year	-1.6e-04 (28)	-3.9e-04(47)	-1.2e-04(50)	-8.7e-05(87)	-1.2e-04(66)	-5.2e-04(11)	-5.0e-03 (82)
	2-year	-9.9e-05 (12)	4.0e-06(57)	-9.5e-05(34)	-5.9e-04(22)	-1.5e-04(13)	3.0e-06(9)	2.6-04 (22)
Hungary	1-year	-7.4e-04 (60)	1.0e-04(32)	-8.9e-04(9)	-8.3e-04(89)	2.8e-03(28)	1.1e-04(21)	-1.8e-02 (71)
	2-year	-3.2e-03 (37)	-2.8e-04(20)	9.0e-06(22)	1.3e-05(67)	4.5e-04(72)	-2.5e-04(42)	2.80e-05(15)
UK	1-year	-1.1e-04 (55)	1.4e-03(44)	-1.3e-04(81)	-8.3e-05(15)	-8.6e-05(33)	4.1e-04(69)	-2.6e-04(44)
	2-year	-1.1e-04 (6)	-8.5e-05(31)	-1.1e-04(45)	-1.1e-04(38)	-1.0e-04(3)	-3.7e-04(56)	-1.1e-04(90)

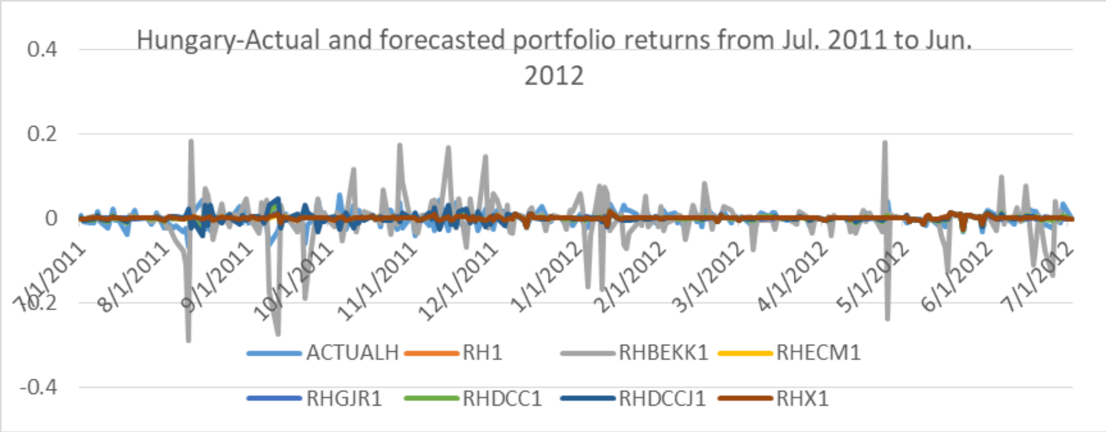
Note: 1. Following Kroner and Sultan (1993), with transaction costs, the forecasted return from GARCH model is $R_{garch} = r_t^c - \hat{h}^*_{garch,t} r_t^f - c$ if we rebalance the futures position, and the return is $R_{garch} = r_t^c - \hat{h}^*_{garch,t} r_t^f$ without rebalancing, where $\hat{h}^*_{garch}$ is the hedge ratio from the most recent re-balancing at time t' and we balance futures position if and only if the balanced position yields higher return than the previous futures position, i.e., $r_t^c - \hat{h}^*_{garch,t} r_t^f - c > r_t^c - \hat{h}^*_{garch,t} r_t^f$. In the same manner, we obtain other return series and the average value of each forecasted return series from six models, to test which model provides the highest return when transaction costs are considered.

2. The times of rebalancing futures positions are included in parentheses for each country from six GARCH models when transaction costs are incorporated.

Figure 1 Graphs of one-year Out-of-Sample return forecasts







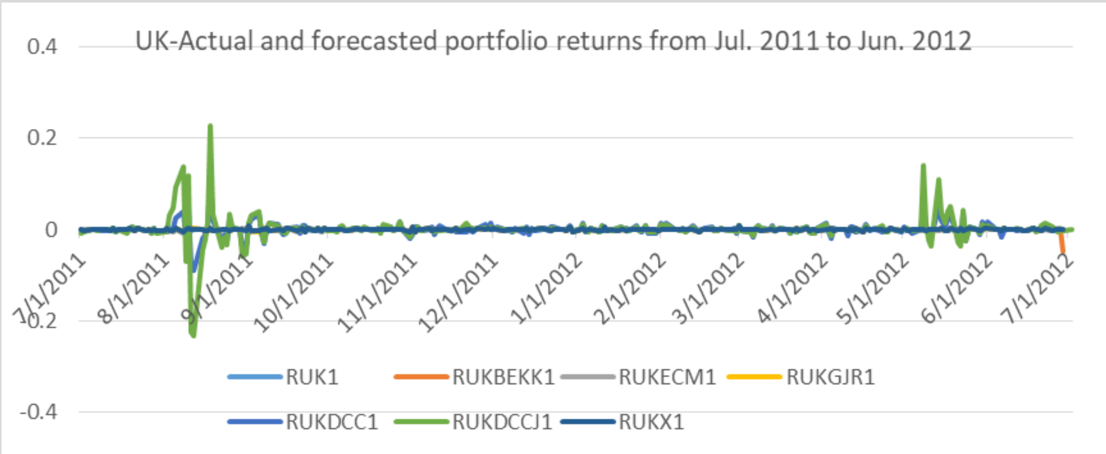
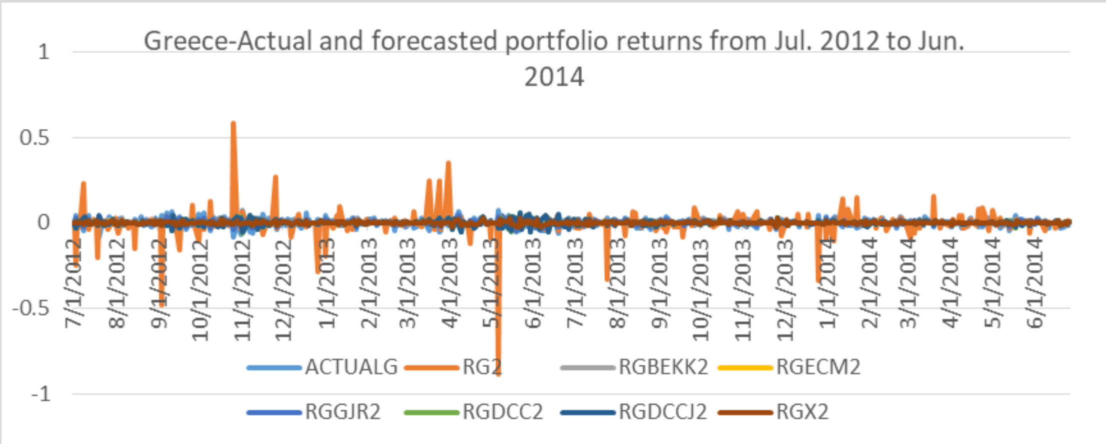
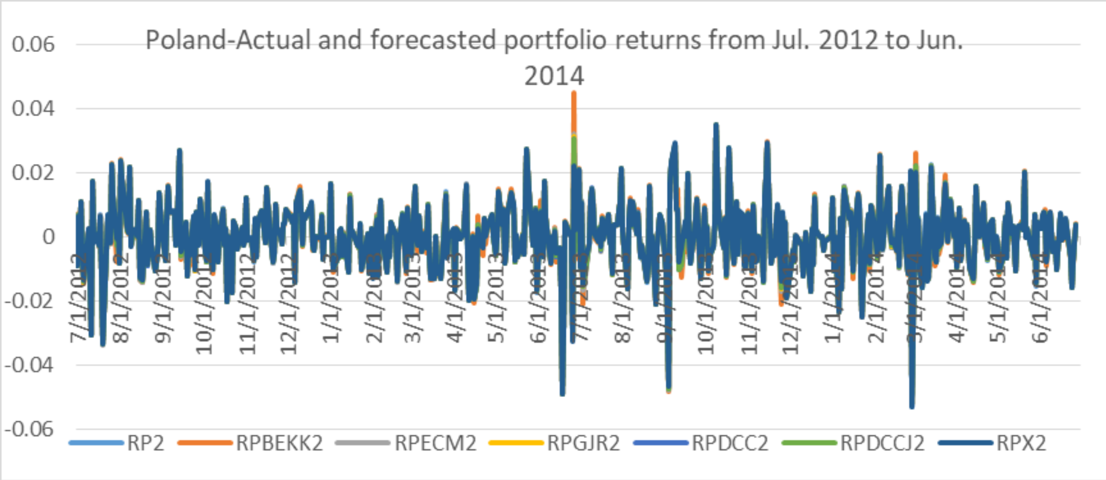
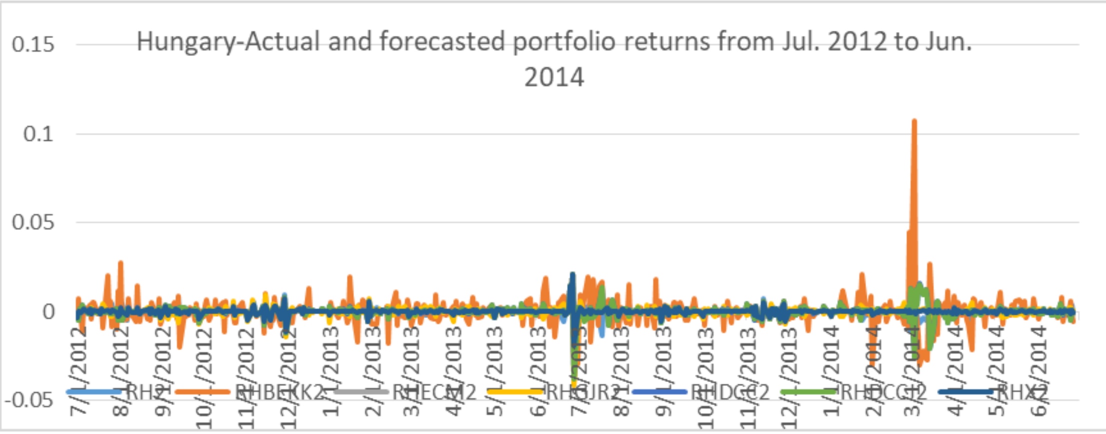
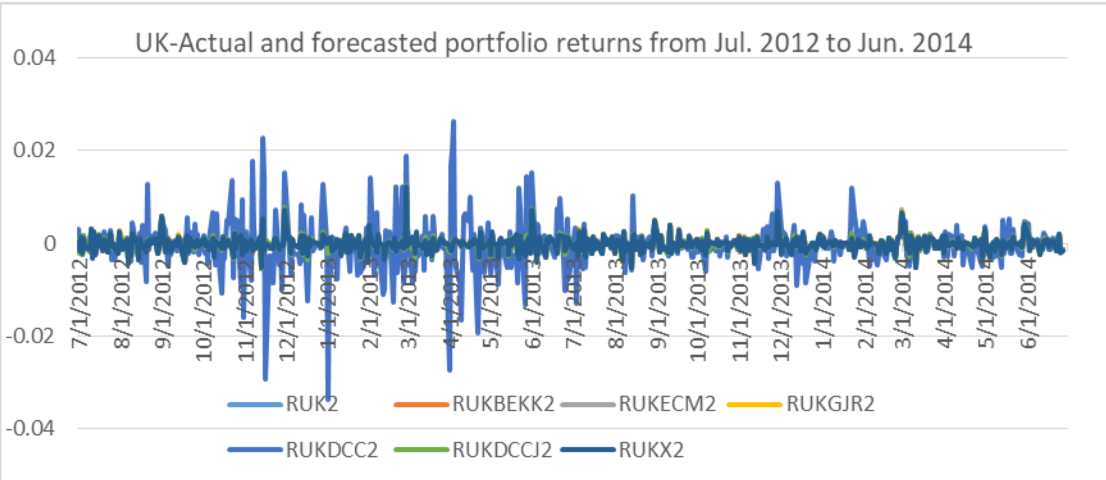


Figure 2 Graphs of two-year Out-of-Sample return forecasts




Poland-Actual and forecasted portfolio returns from Jul. 2012 to Jun. 2014
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
RP2
RPBEKK2
RPECM2
RPGJR2
RPDCC2
RPDCCJ2
RPX2


Hungary-Actual and forecasted portfolio returns from Jul. 2012 to Jun. 2014
0.15
0.1
0.05
0
-0.05


UK-Actual and forecasted portfolio returns from Jul. 2012 to Jun. 2014
0.04
0.02
0
-0.02
-0.04
7/1/2012
8/1/2012
9/1/2012
10/1/2012
12/1/2012
2/1/2013
3/1/2013
5/1/2013
6/1/2013
7/1/2013
8/1/2013
9/1/2013
10/1/2013
11/1/2013
12/1/2013
1/1/2014
2/1/2014
3/1/2014
4/1/2014
5/1/2014
6/1/2014
RUK2
RUKBEKK2
RUKECM2
RUKGJR2
RUKDCC2
RUKDCCJ2
RUKX2