High-low range in GARCH models of stock return volatility

forthcoming in Applied Economics

Peter Molnár*

March 22, 2016

Abstract

We suggest a simple and general way to improve the GARCH volatility models using the intraday range between the highest and the lowest price to proxy volatility. We illustrate the method by modifying a GARCH(1,1) model to a Range-GARCH(1,1) model. Our empirical analysis conducted on stocks, stock indices and simulated data shows that the Range-GARCH(1,1) model performs significantly better than the standard GARCH(1,1) model both in terms of in-sample fit and out-of-sample forecasting ability.

JEL Classification: C22, G17

Key words: volatility, high, low, range, GARCH

1 Introduction

Changes of asset prices (returns) and their variances belong to the fundamental variables in finance. Even though returns of most financial assets are to a large extent unpredictable, their variances display high temporal dependency and are predictable. Starting with the work of Engle (1982) and Bollerslev (1986), the ARCH and GARCH classes of models

*Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway, Tel.: +4773593576, E-mail address: Peto.Molnar@gmail.com. I would like to thank Jonas Andersson, Milan Bašta, Ray Chou, Stein-Erik Fleten, Ove Hetland, Einar Cathrinus Kjenstad, Svein Olav Krakstad, Lukáš Laffèrs, Gudbrand Lien, Kjell Nyborg for helpful comments.

have become standard tools to for volatility modelling and forecasting, see Andersen et al. (2006).

In GARCH type of models, demeaned[1] squared returns serve as a way to calculate innovations to the volatility. Rewriting the GARCH(1,1) model in terms of observed variables (returns) only shows that the GARCH(1,1) model in fact calculates volatility as a weighted moving average of past squared returns. If volatility is changing gradually over time, the GARCH model will work simply because squared returns are daily volatility estimates and therefore the GARCH model essentially calculates volatility as a weighted moving average of the past volatilities.

This intuition has interesting implications. Most importantly, replacing the squared returns by more precise volatility estimates will produce better GARCH models, regarding both in-sample fit and out-of-sample forecasting performance. Additionally, coefficients of GARCH models based on volatility estimates more precise than squared returns will be changed in such a way that they will put more weight on more recent observations. We examine both these implications.

To test our idea, we estimate a GARCH(1,1) model using both squared returns and a more precise volatility proxy, in particular the Parkinson (1980) volatility estimator based on range (the difference between high and low). The results confirm our expectations.

Our work is related to other range-based volatility models, namely Alizadeh et al. (2002), Chou (2005) and Brandt and Jones (2006) and more recently Miralles-Marcelo et al. (2013). However, standard GARCH models are estimated to fit the conditional distribution of returns, whereas the previously mentioned models are estimated to fit the conditional distribution of range (log-range). This means that only our model can be estimated directly in standard econometric software without any programming. For a review of range-based volatility models see Chou et al. (2015). Range-based volatility estimators are compared in Molnár (2012) and some of the recent applications of range

[1]For most of the assets, mean daily return is much smaller than its standard deviation and therefore can be considered equal to zero. In this paper we assume that it is indeed zero. This assumption not only makes further analysis simpler, but it actually helps to estimate volatility more precisely. In the words of Poon and Granger (2003): "The statistical properties of sample mean make it a very inaccurate estimate of the true mean, especially for small samples, taking deviations around zero instead of the sample mean typically increases volatility forecast accuracy."

are Awartani and Maghyereh (2013), Lucey et al. (2014) and Lyócsa (2014).

Our contribution is threefold. First, we construct a range-based GARCH model (RGARCH). This model is a simple modification of the standard widely used GARCH(1,1) model, but still outperforms it significantly. Second, our paper should be viewed as an illustration of how the existing GARCH models can be easily improved by using more precise volatility proxies. Even though this paper devotes most of the space to compare the RGARCH(1,1) model with the standard GARCH(1,1) model, our main goal is not to convince the reader that our model is the best one. On the contrary, since leverage effect is a well-documented phenomenon, an asymmetric RGARCH model is very likely to outperform the RGARCH(1,1). However, we focus on the GARCH(1,1) model, as it is arguably the most fundamental volatility model and the incorporation of the range into this model illustrates the general idea well. Third, we confirm that GARCH models should indeed be considered just filtering devices, not models for data generating processes.

The rest of the paper is organized in the following way: Section 2 provides a basic introduction to volatility modelling and an overview of existing range-based volatility estimators. Section 3 describes the data, methodology and results. Finally, Section 4 concludes.

2 Theoretical background

2.1 GARCH models

Let P_t be the price of a speculative asset at the end of day t. Define return r_t as

$$r_t = \log(P_t) - \log(P_{t-1}). \quad (1)$$

Daily returns are known to be basically unpredictable and their expected value is very close to zero. On the other hand, variance of daily returns changes significantly over time. We assume that daily returns are drawn from a normal distribution with a zero mean and time-varying variance:

$$r_t \sim N\left(0, \sigma_t^2\right). \quad (2)$$

Both the zero mean and normal distribution assumptions are not necessary and can be abandoned without any difficulty. For the sake of exposition, we maintain these assumptions throughout the whole paper. This allows us to focus on modelling of conditional variance (volatility) only. The first model to capture the time variation of volatility is Engle's (1982) Auto Regressive Conditional Heteroskedasticity (ARCH) model. The ARCH(p) has the form

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2, \tag{3}$$

where r_t is a return in day t, σ_t^2 is an estimate of the volatility in day t and ω and α_i's are positive constants. The Generalized ARCH model was afterwards introduced by Bollerslev (1986). The GARCH(p,q) has the following form:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{4}$$

where the β_i's are positive constants. The GARCH model has become more popular, because with just a few parameters it can fit data better than a more parametrized ARCH model. Particularly popular is its simplest version, the GARCH(1,1) model[2]:

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2. \tag{5}$$

Estimation of the GARCH(1,1) typically yields the following results. Parameter ω is very small (e.g. 0.0006), $\alpha + \beta$ is close to one, but smaller than one. Moreover, most of the weight is on the β coefficient, e.g. $\alpha = 0.04$, $\beta = 0.95$. In other words, the estimated GARCH(1,1) model is usually very close to its reduced form, the Exponential Weighted Moving Average (EMWA) model

$$\sigma_t^2 = \alpha r_{t-1}^2 + (1 - \alpha)\, \sigma_{t-1}^2. \tag{6}$$

The EMWA model is useful particularly for didactic purposes. In this model the new volatility estimate is estimated as a weighted average of the most recently observed volatility proxy (squared returns) and the last estimate of the volatility. Loosely speaking, we gradually update our belief about the volatility as new information (noisy volatility proxy) becomes available. If the new information indicates that the volatility was larger than our previous belief about it, we update our belief upwards and vice versa. The coefficient α tells us how much weight we put on the new information. If we use a less noisy volatility

[2]Even though the GARCH(1,1) is a very simple model, it still works surprisingly well in comparison with much more complex volatility models (Hansen and Lunde (2005)).

proxy instead of squared returns, the optimal α should be larger and the performance of the model should be better.

The same intuition applies to GARCH models too. This naturally leads to the proposal of the modified GARCH(1,1)

$$\sigma_t^2 = \omega + \alpha \widehat{\sigma_{proxy,t-1}^2} + \beta \sigma_{t-1}^2 \tag{7}$$

where $\widehat{\sigma_{proxy,t-1}^2}$ is the less noisy volatility proxy.

Next we need to decide upon what should be used as a better (less noisy) volatility proxy. Generally, the better the proxy we use, the better should the model work. Therefore, the natural candidate would be realized volatility. This would lead to models related to Shephard and Sheppard (2010) and Hansen et al. (2011). However, despite the attractiveness of the realized variance we do not use it as a volatility proxy. Realized variance must be calculated from high frequency data and these data are in many cases not available at all or available only over shorter time horizons and costly to obtain and work with. Moreover, due to market microstructure effects the estimation of volatility from high frequency data is a rather complex issue (see Dacorogna et al. (2001)). Contrary to high frequency data, high (H) and low (L) prices, which are usually widely available, can be used to estimate volatility (Parkinson (1980)):

$$\widehat{\sigma_P^2} = \frac{[\ln(H/L)]^2}{4 \ln 2}. \tag{8}$$

This estimator is derived under the assumption that, during the day, the logarithm of the price follows a Brownian motion with a zero drift. Even though this is not always true, Parkinson's volatility estimator performs very well with the real world data (Chou et al. (2010)).

An alternative volatility proxy we could use is the Garman and Klass (1980) volatility estimator, which utilizes additional open (O) and close (C) data:

$$\widehat{\sigma_{GK}^2} = 0.5 [\ln(H/L)]^2 - (2 \ln 2 - 1) [\ln(C/O)]^2 . \tag{9}$$

Under ideal conditions (Brownian motion with zero drift) this estimator is less noisy than the Parkinson volatility estimator, because it utilizes open and close prices too. However, in this paper we use Parkinson's volatility estimator ($\sigma_{proxy}^2 = \sigma_P^2$). We have done all the calculations for the Garman-Klass volatility estimator too and found out that for this particular purpose the Garman-Klass estimator does not improve the results more than Parkinson estimator. Moreover, for the same data sets where high and low prices are available, open price is sometimes not available.

In this paper we therefore study the following model

$$\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2, \tag{10}$$

which we denote as RGARCH(1,1) (range GARCH) model. This model can obviously be extended to the RGARCH(p,q) model

$$\sigma_t^2 = \omega + \sum_{i=1}^{p}\alpha_i\widehat{\sigma_{P,t-i}^2} + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2. \tag{11}$$

Since it is generally known that GARCH(p,q) of order higher than (1,1) is seldom useful (see e.g. Hansen and Lunde (2005)), we study the RGARCH model only in its simplest version (10), i.e. the RGARCH(1,1) model. Most of the paper is devoted to the comparison of the standard GARCH(1,1) model (5) and the RGARCH(1,1) model (10). Since we do not study GARCH and RGARCH models of higher orders, we sometimes refer to GARCH(1,1) and RGARCH(1,1) models simply as GARCH and RGARCH models.

Our hypotheses are the following:

Hypothesis 1 *An RGARCH(1,1) outperforms the standard GARCH(1,1) model, both in sense of the in sample fit and out of sample forecasting performance.*

Additionally, as previously explained, we expect that the estimated coefficients of the GARCH models will be changed in such a way that more weight will be put on the recent observation(s) of the volatility proxy. This leads us to the second hypothesis.

Hypothesis 2 *If we modify the GARCH(1,1) to the RGARCH(1,1) model, we expect α to increase and β to decrease.*

Since the RGARCH(1,1) model puts more weight on the most recent observation of the volatility, this model will provide largest improvement in those situations when the recent observation tells us much more about the future volatility then the past observations. This leads us to the following hypothesis.

Hypothesis 3 *The superiority of the RGARCH(1,1) model over the GARCH(1,1) model is the strongest when day-to-day changes in volatility are large.*

However, this does not mean that GARCH should be better model in situations when changes in volatility are small. We expect RGARCH model to be superior in both situations, but its superiority should be largest in situations when volatility changes a lot.

Even though we formulated 3 hypotheses, the central one is Hypothesis 1. The purpose of Hypothesis 2 and Hypothesis 3 is mostly to provide some additional insights why and when RGARCH model works better than standard GARCH model.

To evaluate the usefulness of the RGARCH model, we briefly compare it not only with the basic GARCH(1,1) model, but with the other commonly used GARCH models too. GARCH models we compare the RGARCH to are the following:

The GJR-GARCH of Glosten et al. (1993):

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma r_{t-1}^2 I_{t-1}. \tag{12}$$

where $I_t = 1$ if $r_t < 0$ and zero otherwise.

The Exponential GARCH (EGARCH) of Nelson (1991):

$$\log\left(\sigma_t^2\right) = \omega + \alpha \left|\frac{r_{t-1}}{\sigma_{t-1}}\right| + \beta \log\left(\sigma_{t-1}^2\right) + \gamma \frac{r_{t-1}}{\sigma_{t-1}}. \tag{13}$$

The standard deviation GARCH of Taylor (1986), denoted in this paper as stdGARCH, both in its symmetric version:

$$\sigma_t = \omega + \alpha |r_{t-1}| + \beta \sigma_{t-1} \tag{14}$$

and in the asymmetric version, similar to (12), taking into account the leverage effect (astdGARCH):

$$\sigma_t = \omega + \alpha |r_{t-1}| + \beta \sigma_{t-1} + \gamma |r_{t-1}| I_{t-1}. \tag{15}$$

The last model we use is the component GARCH (cGARCH):

$$\sigma_t^2 - m_t = \overline{\omega} + \alpha\left(r_{t-1}^2 - m_t\right) + \beta\left(\sigma_{t-1}^2 - m_t\right) \tag{16}$$

$$m_t = \omega + \rho\left(m_t - \omega\right) + \phi\left(r_{t-1}^2 - \sigma_{t-1}^2\right). \tag{17}$$

2.2 Estimation

All the GARCH models, including the models (5), (12) - (15) in our paper are estimated via Maximum Likelihood. Since the RGARCH model changes only the specification of the variance equation (equation (10) instead of (5)), we do not need to derive a new likelihood function for estimation of this model. This in turns mean that our model can be estimated without any programming in widely available econometric packages which allow to include exogenous variables in the variance equation, e.g. EViews, R or OxMetrics. We simply specify that we want to estimate a GARCH(0,1) model with an exogenous variable $\widehat{\sigma^2_{P,t-1}}$.

As mentioned earlier, we assume returns to be normally distributed with zero mean (equation (2)) and variance evolving according to a given GARCH model. However, there are alternative distributions for residuals to consider (e.g. Student's t-distribution or GED distribution). We did the calculations for alternative distributions too, but found that comparison of the. RGARCH model with the standard GARCH model is unaffected by the assumption of the residuals' distribution as long as the return distribution is the same for both models. For the sake of brevity, we report only the results for normally distributed residuals.

Two most closely related models are the Conditional Autoregressive Range model (CARR) of Chou (2005) and Range-Based EGARCH model (REGARCH) of Brandt and Jones (2006). A common feature of these models with the standard GARCH models is the variance equation. The variance equation for RGARCH model is created by a modification of the GARCH(1,1) (5), the variance equation of the CARR model is a modification of the GJR-GARCH (12) and the variance equation of the REGARCH is a modification of EGARCH (13).

However, CARR and REGARCH are otherwise significantly different from RGARCH and other GARCH models. Standard GARCH models as well as our RGARCH model are estimated by fitting the conditional distribution of returns. On contrary, estimation of the CARR and the REGARCH models is based on the distribution of the range. Denote range as

$$D_t = \ln\left(H_t/L_t\right). \tag{18}$$

The REGARCH model is estimated by fitting the conditional distribution of log-range:

$$\ln\left(D_t\right) \sim N(0.43 + \ln\left(\sigma_t\right), 0.29^2), \tag{19}$$

and the CARR model is estimated by fitting the conditional distribution of range

$$D_t = \lambda_t \varepsilon_t, \tag{20}$$

where λ_t is the conditional mean of the range (varying according to equation similar to (12)) and ε_t is distributed according to either the exponential or the Weibull distribution.

In other words, these models are not estimated to capture the conditional distribution of the returns, but the conditional distribution of range instead. Since these estimations are not implemented in standard econometric software, CARR and RGARCH models must be programmed first.

On the contrary, RGARCH model combines the ease of estimation of the standard GARCH models with the precision of the range-based models.

Now we evaluate the performance of the RGARCH model (10) by comparing it with the standard GARCH(1,1) model (5), because these two models are very closely related and their direct comparison is very intuitive. We compare both in-sample fit and out-of-sample forecasting performance. The analysis of the in-sample fit will give us some insights about how these models work. Since the forecasting ability is typically the most important feature of a volatility model, we focus mostly on its forecasting ability.

2.3 In-sample comparison

We start the in-sample comparison between RGARCH(1,1) and standard GARCH(1,1) models by an estimation of equations (5) and (10). This allows us to see whether the coefficients change according to our Hypothesis 2. To evaluate which model is a better fit for the data, we use Akaike Information Criterion (AIC). However, as we are comparing models with an equal number of parameters, any information criterion would necessary produce the same raking of these models. We believe that in our particular case, when we are comparing two very closely related models (the conditional distribution of returns is the same, models differ in specification of variance equation only), AIC is a sensible criterion.

Moreover, we estimate the combined GARCH(1,1) model

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \alpha_2 \widehat{\sigma_{P,t-1}^2} + \beta \sigma_{t-1}^2 \tag{21}$$

too. This allows us to better understand which volatility proxy: squared returns r_{t-1}^2 or the Parkinson volatility proxy $\widehat{\sigma_{P,t-1}^2}$, is a more relevant variable in the variance equation.

2.4 Out-of-sample forecasting evaluation

To evaluate forecasting performance of two competing models, we first create forecasts from these models and afterwards evaluate which of these forecasts is on average closer to the true volatility (we explain later what is meant by "true volatility").

To do this, we must first decide how to create the forecasts, particularly how much data to use for the forecasting. If we use too little data, the model will be estimated imprecisely and the forecasting will not be very good. On the other hand, if we use too much data, we can estimate the model precisely, but when the dynamics of the true volatility changes, our model will adapt to this change too slowly. To avoid this problem, we use rolling window forecasting[3] with four different window sizes: 300, 400, 500 and 600 trading days. These numbers are obviously somewhat arbitrary, but we are focused on the comparison of different volatility models, not on the search for the optimal forecasting window. Due to space limitations, we restrict our attention to one-day-ahead forecasts.

Next we decide which benchmark to use (as the "true" volatility). The most common benchmark is squared returns. Squared returns are widely used due to the data availability. However, squared returns are a very noisy volatility proxy. Therefore we use the Parkinson volatility estimator and the realized variance too. Due to space limitations, we do not report results when the Parkinson volatility estimator is used as a benchmark, though the results are even more convincing than for squared returns. However, whenever the data on the realized variance is available, we use it as a benchmark.

To evaluate which forecast is closer to the true value, we must next decide on the loss function. We use the Mean Squared Error (MSE) as a loss function. For the sake of exposition, we report Root Mean Squared Error (RMSE) instead of MSE in all the tables. MSE is not only the most common loss function, but it has many other convenient properties, particularly the robustness. Since we are using imperfect volatility proxies, the choice of an arbitrary loss function (e.g. Mean Absolute Error or Mean Percentage Error) could lead to problems, particularly to the inconsistent ranking of different models (see Hansen and Lunde (2006) and Patton (2011)).

Next we want to know whether the MSE for two different models are statistically

[3]By rolling window forecasting with window size 100 we mean that we use the first 100 observations to forecast volatility on the 101, then we use observations 2 to 101 to forecast volatility for day 102 and so on.

different. We adopt the Diebold and Mariano (1995) test for this purpose. The Diebold-Mariano test statistic (DM) is computed in the following way: denote two competing forecasts as $\widehat{\sigma_{1,t}^2}$ and $\widehat{\sigma_{2,t}^2}$ and the true volatility as $\sigma_{true,t}^2$. In our case $\widehat{\sigma_{1,t}^2} = \widehat{\sigma_{RGARCH,t}^2}$ and $\widehat{\sigma_{2,t}^2}$ is the competing model; in the majority of this paper it is the GARCH(1,1) model. First we construct the vector of differences in squared errors

$$d_t = \left(\widehat{\sigma_{1,t}^2} - \sigma_{true,t}^2\right)^2 - \left(\widehat{\sigma_{2,t}^2} - \sigma_{true,t}^2\right)^2. \tag{22}$$

Next we construct the Diebold-Mariano test statistic

$$DM = \frac{\overline{d}}{\sqrt{\widehat{V}\left(\overline{d}\right)}}, \tag{23}$$

where $\overline{d}$ denotes the sample mean of d_t and $\widehat{V}\left(\overline{d}\right)$ is variance of the sample mean. DM is assumed to have a standard normal distribution. Later in the results we denote by asterisk * (**) cases when the DM test statistics lies below 5-percentile (1-percentile), i.e. the cases where we can reject at 5% (1%) confidence level the hypothesis that the competing model has smaller MSE than the RGARCH(1,1) model.[4]

2.5 Opening jump

In the previous discussion we assumed that all the models are estimated on the close-to-close returns defined by equation (1). This is typically the case for the standard GARCH models. On the other hand, a common approach in the literature dealing with high frequency data is to model open-to-close returns

$$r_t = \log\left(P_t\right) - \log\left(O_t\right). \tag{24}$$

The volatility for the trading period (from open to close of the market) can be estimated quite precisely, whereas this precision is not available for estimation of the period over the night (the in opening jump). Moreover, the dynamics of the opening jumps is arguably different from the dynamics of the volatility of the trading part of the day. Since Parkinson volatility estimator (8) estimates open-to-close volatility only, we face the same problem. We follow the standard approach in the realized variance literature and the models presented in this paper are estimated on the open-to-close returns. We have done the same

[4]In our data the DM test statistic never lies above 95-percentile.

estimations for close-to-close returns. Main results remain the same (RGARCH model outperforms GARCH model).[5] These results are available in previous versions of the paper or upon request from the author.

3 Data and results

To show the generality of our idea we study a wide class of assets, particularly 30 individual stocks, 6 stock indices and simulated data. Due to space limitations, our analysis cannot be as detailed as it would be if we studied a single asset. We believe that the analysis of the main features of the problem on the broad data set is more convincing than very detailed analysis based on a small data set. We use daily data, particularly the highest, the lowest, the opening and the closing price of the day.

[5]However, there is on difference worth mentioning. The Parkinson volatility estimator estimates the volatility only for the open-to-close period. If we estimate RGARCH model on close-to-close returns, we must be careful with interpretation of the α coefficient in the RGARCH model. As long as opening jumps are present, the Parkinson volatility estimator underestimates volatility of daily returns,

$$E\left(\widehat{\sigma_P^2}\right) < E\left(r^2\right) = \sigma^2. \tag{25}$$

As a result, the estimated coefficient α will be larger to balance this bias in $\widehat{\sigma_P^2}$. This intuition can explain one seemingly surprising result. The RGARCH model estimated on the close-to-close data typically yield coefficients α and β such that $\alpha + \beta > 1$, even though estimation of the standard GARCH(1,1) model yields coefficients α and β such that $\alpha + \beta < 1$. However, as we just explained, these α coefficients are not directly comparable in presence of opening jumps. We illustrate this on a simple example. If we specify GARCH(1,1) in the following form

$$\sigma_t^2 = \omega + \alpha \frac{r_{t-1}^2}{2} + \beta \sigma_{t-1}^2,$$

then the estimated coefficient α will be exactly twice as large as when we estimate equation (5). Therefore, if the RGARCH model is estimated on the close-to-close returns, the coefficient α does not have the same interpretation as in standard GARCH models. Even though we expect α to increase and β to decrease, if we use close-to-close returns, we must focus on the coefficient β only. The coefficient β will change only because a less noisy volatility proxy is used, whereas change in coefficient α is caused by both high precision and bias of the Parkinson volatility estimator.

3.1 Stocks

We study the components[6] of the Dow Jones Industrial Average, namely the stocks with tickers AA, AXP, BA, BAC, C, CAT, CVX, DD, DIS, GE, GM, HD, HPQ, IBM, INTC, JNJ, JPM, CAG[7], KO, MCD, MMM, MRK, SFT, PFE, PG, T, UTX, VZ and WMT. Data were obtained from the CRSP database and consist of 4423 daily observations of high, low and close prices from June 15, 1992 to December 31, 2010.

3.1.1 In-sample analysis

Table 1 presents estimated coefficients for the equations (5) and its modified version (7) together with values of Akaike Information Criterion (AIC).

For every single stock, the coefficients in the RGARCH(1,1) have changed in exactly the same way we expected. Additionally, according to AIC, modified GARCH(1,1) is superior to its standard counterpart for every single stock in our sample.

Next we estimate the Combined GARCH(1,1) model (eq. 21). Results of this estimation (reported in Table 2 together with respective p-values) show that coefficients α_2 is always highly significant, the coefficient α_1 is insignificant in most of the cases. and even when it is statistically significant, it is rather small. This confirms that σ_P^2 is a better volatility proxy than r^2 and when we have the first one available, the inclusion of the second one can improve the model only marginally. Note that the coefficient α_1 is negative in most cases. This is expected, since an optimal volatility estimator (9) combines the Parkinson volatility estimator with squared returns in such a way that squared returns have negative weight. We discuss this more in the subsection with simulated data.

3.1.2 Out-of-sample forecasting performance

As seen in the previous subsection, the RGARCH model outperforms the standard GARCH model in the in-sample fit of the data. The next obvious question is the comparison of the predictive ability of these models. To answer this question, we compare one-day ahead forecasts of the models (5) and (7) with squared returns as a benchmark. Results are presented in the Table 3.

[6] Components of stock indices change over time. These stocks were DJI components on January 1, 2009.

[7] Since historical data for KFT (component of DJI) are not available for the complete period, we use its competitor CAG instead.

As we can see from Table 3, the RGARCH(1,1) model outperforms GARCH(1,1). All the cases (stock-estimation window pairs) when the difference is statistically significant favour the RGARCH model. The reason the difference is often insignificant is a very noisy volatility benchmark (squared returns). Therefore we should postpone the evaluation of size of the improvement of RGARCH(1,1) model over GARCH(1,1) model until next subsections, where we use realized variance as a less noisy benchmark.

The next obvious question is how our RGARCH performs relative to other more complicated GARCH models. Even though a detailed answer to this question is beyond the scope of this paper, we provide some basic comparison. We compare the RGARCH model (10) not only with the basic GARCH model (5), but with its other versions (12)-(16) as well. We chose an estimation window equal to 400. A shorter estimation window would favour the RGARCH model even more. A too long estimation window is not desirable, because, as Table 3 documents, volatility forecasting becomes less precise when we use a too long estimation window.

As we can see from Table 4, the comparison of the RGARCH model with other GARCH models is very similar to the previous comparison, the RGARCH model outperforms other GARCH models. When we consider the cases where the difference is statistically significant, the RGARCH model always outperforms all other studied GARCH models. In rest of the cases, when the difference is not statistically significant, the RGARCH model outperforms other studied GARCH models most of the time. Remember that we do not argue that RGARCH model is the best volatility model. It is clearly not, as it does not take into account e.g. leverage effect. Therefore, the comparison of the RGARCH model with other GARCH models serves mostly the illustrative purposes, particularly to show that even such a simple model (but based on more precise data) can outperform more complicated models.

The results summarized in Tables 3 and 4 show the superior performance of the RGARCH model. The improvement in the RGARCH model in comparison to the basic GARCH(1,1) model seems to be rather small at the first glance. Even though the RGARCH model outperforms the basic GARCH(1,1) model in most cases, the average improvement of the RMSE reported in Table 3 is about 1.2%. This could give us a first impression that the improvement of the RGARCH(1,1) model over the GARCH(1,1) model is rather small.

However, there is a potential problem with this standard evaluation procedure, where we compare the forecasted volatility with the squared returns. Even though the squared returns are unbiased estimates of the volatility, they are very noisy[8]. The most natural solution to this problem is to use the true volatility as a benchmark, or, if unavailable, some other less noisy volatility proxy. Following subsections use less noisy volatility proxies (realized variance for the stock indices and true volatility for simulated data). However, due to stock data limitations, we suggest an alternative way to compare the basic GARCH(1,1) model and the RGARCH(1,1) model. Instead of comparing squared returns with volatility forecast directly, we can compare the likelihood that the returns were drawn from the distribution parametrized by the given volatility. This approach is not perfect either, because the calculated probability depends on the specification of the distribution of the stock returns. Since we compare two models with the same specification of the conditional distribution of returns, $N\left(0,\sigma_{t,1}^2\right)$ and $N\left(0,\sigma_{t,2}^2\right)$, which differ only in the specification of the variance equation, this is not a problem. We now compare the basic GARCH(1,1) model with the RGARCH model in terms of the value of the log-likelihood function. The log-likelihood is calculated according to the following formula:

$$LLF = -\frac{n}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{n}\ln\left(\widehat{\sigma_t^2}\right) - \frac{1}{2}\sum_{t=1}^{n}\frac{r_t^2}{\widehat{\sigma_t^2}}, \tag{26}$$

where σ_t^2 is the volatility forecasted from the studied volatility model (using past information only).

Table 5 confirms our previous comparison between the RGARCH model and the standard GARCH model. The RGARCH model outperforms the standard GARCH(1,1) model for basically every stock and every estimation window.

[8]A comparison of the forecasted volatility with squared returns will penalize the volatility forecast whenever the squared return and volatility forecast differ, even if the volatility forecast was perfect. Moreover, when we have two models and one of them forecasts volatility to be $\sigma^2 = 0.1^2$ on the day when the stock return is $r = 1$ and the second model forecasts volatility to be $\sigma^2 = 3^2$ on the day when stock return is $r = \sqrt{10}$, then MSE (RMSE) will favour the first model ($\left(0.1^2 - 1^2\right)^2 < \left(10 - 3^2\right)^2$), even though the probability of the return $r = 1$ being drawn from the distribution $N\left(0, 0.1^2\right)$ is more than 10^{40}-times smaller than probability of the return $r = \sqrt{10}$ being drawn from the distribution $N\left(0, 3^2\right)$.

3.2 Stock indices

In addition to the individual stocks of the Dow Jones Industrial Average stock index we decided to compare the performance of the RGARCH model to the standard GARCH model on the major world indices (French CAC 40, German DAX, Japanese Nikkei 225, Britain's FTSE 100 and American DJI and NASDAQ 100). There are two reasons for this. First, volatility dynamics is generally different for individual stocks and for the whole stock markets. Second, estimates of realized variance, which is a proxy for the true variance, are publicly available for these indices[9]. Open, high, low and close prices are downloaded from finance.yahoo.com. Data covers the period January 3, 1993 - April 27, 2009 for open, high and low prices and the period January 3, 1996 - April 27, 2009 for the realized variance. Due to small differences in trading days in different markets, the number of observations varies accordingly.

For the in-sample analysis we use the data ranging from January 3, 1993 to April 27, 2009. For the out of sample comparison we use the volatilities forecasted for the period January 3, 1996 - April 27, 2009. However, estimates of realized variance are not available for some trading days. These days are included in the volatility forecast comparison when squared returns are used as a benchmark, but excluded when the benchmark is realized variance.

3.2.1 In-sample analysis

Table 6 presents estimated coefficients for the GARCH model (5) and the RGARCH model (10) together with the values of Akaike Information Criterion (AIC). The results are again in line with those in Table 1. RGARCH model performs better than the standard GARCH model for every index. The coefficients in the RGARCH are changed as expected - coefficient α is increased and coefficient β is decreased for all the indices.

Now we estimate the combined GARCH model (21). The results (presented in Table 7) are consistent with those in Table 2.

[9]Heber, Gerd, Asger Lunde, Neil Shephard and Kevin K. Sheppard (2009) "Oxford-Man Institute's Realized Library", Oxford-Man Institute, University of Oxford

3.2.2 Out–of-sample forecasting performance

Now we compare the forecasting performance of the RGARCH model and the standard GARCH model against both squared returns (r^2) and realized variance (RV) used as a benchmark. Results are in Table 8.

This table provides the strongest evidence for the superiority of the RGARCH model over the standard GARCH model. For every single index and for every single estimation window size, the RGARCH model outperforms the standard GARCH model. The difference in the forecasting performance of these two models is much more obvious when we use realized variance as a benchmark (since it is much less noisy than squared returns).

3.3 Simulated data

In reality, we can never know for sure what the true volatility was. However, if we simulate the data, we know the true volatility exactly. Simulation therefore provides a convenient tool to study different volatility models. We can compare not only the overall performance of different models, but we can study under which conditions these models perform particularly good or bad. On the other hand, it is always questionable how close the simulated data are to the real world. In order to convince the reader that the simulated data are close to reality (and we did not construct them deliberately to show superiority of our model), we borrow the credibility of Alizadeh et al. (2002). They simulate the data in the following way. First we simulate the volatility process

$$\ln \sigma_t = \ln \overline{\sigma} + \rho_H \left(\ln \sigma_{t-1} - \ln \overline{\sigma} \right) + \mu_1 \varepsilon_{t-1} \tag{27}$$

with parameters $\ln(\overline{\sigma}) = -2.5$, $\rho_H = 0.985$ and $\mu_1 = 0.75/\sqrt{257} = 0.048$ and residuals ε drawn from standard normal distribution. For every day $t = 1, 2, ..., 100000$ we simulate a Brownian motion[10] with zero drift term and diffusion term equal to σ_t. Save the highest, the lowest and the final value of this Brownian motion. According to Alizadeh et al. (2002), volatility dynamics (27) together with mentioned parameters is broadly consistent with literature on stochastic volatility.

The volatility process (27) does not favour directly either of the competing models GARCH (5) and RGARCH (10). Volatility evolves over the time, and neither past returns

[10] We use 100000 discrete steps for the approximation of the continuous Brownian motion.

nor past high or low prices influence the future volatility in any way. Note that there are no opening jumps in this these simulated data.

In addition to data simulated according to (27) with parameter $\mu_1 = 0.75/\sqrt{257}$, we simulate the data for two other parameter values too, $\mu_{0.5} = 0.5\mu_1$ and $\mu_2 = 2\mu_1$. Parameter μ_1 represents a case with medium daily changes in volatility and parameters $\mu_{0.5}$ and μ_2 represent cases with small and large changes in daily volatility.

3.3.1 In-sample analysis

Table 9 presents estimated coefficients for the standard GARCH model (5) and the RGARCH model (10) together with the values of Akaike Information Criterion (AIC). As expected, the RGARCH model performs better than the standard GARCH model.

Coefficients in the RGARCH are changed in exactly the same way as in the previous section - coefficient α is increased and coefficient β is decreased. Note that $\alpha+\beta$ is smaller than one for both GARCH and RGARCH model (implying stationarity) and $\alpha+\beta$ is approximately the same (around 0.98) for both models. This means that both GARCH and RGARCH models imply the same (high) volatility persistence. This is very natural, since we simulate volatility as a highly persistent process. Note that when volatility changes more rapidly (μ increases), more weight is put on the recently observed volatility proxy (α increases) and less weight is put on the past observation of volatility (β decreases).

Now we estimate the combined GARCH model (21). As we can see (Table 10), the results are generally consistent with those in Table 2.

The main difference is that the negative coefficient α_1 is now clearly significant. As Garman and Klass (1980) showed, the optimal volatility forecast based on open, high, low and close price is (9). It is a weighted average of the Parkinson volatility estimator (8) and squared open-to-close returns, where squared returns have negative weight. This is the reason why coefficient α_1 is negative. Note that the ration between the coefficients α_1 and α_2 is very close to the ratio predicted from the Garman-Klass formula. As previously mentioned, we use the Parkinson volatility estimator (8) instead of Garman and Klass (9) volatility estimator because of the data concerns (open prices are often not available).

3.3.2 Out–of-sample forecasting performance

Now we compare the forecasting performance of the RGARCH model and the standard GARCH model on the simulated data. Results are shown in Table 11.

These results illustrate the benefit of using simulated data. Now we know exactly what the true volatility is and we can use it as a benchmark. Additionally, simulation allows us to have much larger data sample (100000 observations of the simulated data instead of 4423 observations of the real data), which in turns mean that all the results are highly statistically significant.

First note that the results obtained from the simulated data (Table 11) are consistent with results in Table 3 and Table 8. Table 3 and Table 8 show that the RGARCH model outperforms the standard GARCH model most of the time. Since the simulated data are much larger, we basically got rid of the noise and now we can see (in Table 11) exactly how much better the RGARCH performs. Let us focus for now primarily on the data simulated with the parameter μ_1, which is arguably closest to the real world. The improvement seems to be small, just around 1% decrease in RMSE, when we use squared returns as a benchmark. However, use of the true volatility as a benchmark shows that the real improvement of the RGARCH in comparison to the standard GARCH model is much larger, around 20%.

In fact, the mean squared error (MSE) between the forecasted volatility ($\widehat{\sigma^2}$) and a noisy volatility proxy (σ^2_{noisy}) can be rewritten in the following way:

$$MSE\left(\widehat{\sigma^2}, \sigma^2_{noisy}\right) = MSE\left(\widehat{\sigma^2}, \sigma^2_{true}\right) + MSE\left(\sigma^2_{true}, \sigma^2_{noisy}\right) \tag{28}$$

where σ^2_{true} is the true volatility. This means that part of the MSE is due to the model imperfection (first term) and second part is due to the noisiness of the volatility proxy. When squared returns are used as a benchmark, then the second term typically dominates and it is therefore difficult to choose between competing volatility models based on the MSE (RMSE).

To understand when the RGARCH model provides the largest improvement over GARCH model (Hypothesis 3), let us look at Table 11. As we cam see, the larger the day-to-day changes in volatility, the larger the improvement of the RGARCH model (relatively to the GARCH model). The decrease in RMSE (with the true volatility as a benchmark) when we use RGARCH instead of GARCH is 6%–9% in case of small day-to-day changes in volatility, 16%–22% for moderate changes in volatility and 23%–24% for large changes in volatility. This confirms our Hypothesis 3.

4 Summary

We demonstrate that a simple way of incorporating range, the intraday difference between the highest and the lowest price, into the standard GARCH framework of volatility models results in superior empirical performance. We illustrate the method by modifying a GARCH(1,1) model to a Range-GARCH(1,1) model. Empirical tests performed on 30 stocks, 6 stock indices and simulated data show that the Range-GARCH model outperforms the standard GARCH model, both in terms of in-sample fit and out-of-sample forecasting.

The intuition of this result is the following. The Range-GARCH model not only replaces squared returns by a more precise volatility proxy in the form of range, but it also puts more importance of the most recent volatility estimate and therefore performs particularly well when the level of volatility changes quickly. This is a desirable feature, because volatility forecasting is especially important in situations of rapidly changing volatility levels.

Range-GARCH combines the high precision of range with the simplicity and ease of estimation of the standard GARCH models. High and low prices are typically widely available and the model itself can be easily estimated using standard econometric software. The Range-GARCH model proposed in this paper should therefore be of significant interest both to academics and practitioners alike.

References

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002): "Range-Based Estimation of Stochastic Volatility Models," *The Journal of Finance*, 57, 1047–1091.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2006): "Volatility and Correlation Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. J. Granger, and A. Timmermann, Elsevier, vol. 1, chap. 15, 777–878.

Awartani, B. and A. I. Maghyereh (2013): "Dynamic spillovers between oil and stock markets in the Gulf Cooperation Council Countries," *Energy Economics*, 36, 28 – 42.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Brandt, M. W. and C. S. Jones (2006): "Volatility Forecasting With Range-Based EGARCH Models," *Journal of Business and Economic Statistics*, 24, 470–486.

Chou, R. Y. (2005): "Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model," *Journal of Money, Credit, and Banking,*, 37, 561–582.

Chou, R. Y., H. Chou, and N. Liu (2010): "Range Volatility Models and Their Applications in Finance," in *Handbook of Quantitative Finance and Risk Management*, ed. by C.-F. Lee and J. Lee, Springer, chap. 83.

——— (2015): "Range Volatility: A Review of Models and Empirical Studies," in *Handbook of Financial Econometrics and Statistics*, Springer, 2029–2050.

Dacorogna, M. M., R. Gencay, U. Muller, R. B. Olsen, and O. V. Olsen (2001): *An introduction to high frequency finance*, Academic Press, New York.

Diebold, F. X. and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–265.

Engle, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

Garman, M. B. and M. J. Klass (1980): "On the Estimation of Security Price Volatilities from Historical Data," *The Journal of Business*, 53, 67–78.

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

Hansen, P. R., Z. Huang, and H. H. Shek (2011): "Realized GARCH: a joint model for returns and realized measures of volatility," *Journal of Applied Econometrics*, forthcoming.

Hansen, P. R. and A. Lunde (2005): "A forecast comparison of volatility models: does anything beat a GARCH(1,1)?" *Journal of Applied Econometrics*, 20, 873–889.

——— (2006): "Consistent ranking of volatility models," *Journal of Econometrics*, 131, 97–121.

Lucey, B. M., C. Larkin, and F. O'Connor (2014): "Gold markets around the world – who spills over what, to whom, when?" *Applied Economics Letters*, 21, 887–892.

Lyócsa, S. (2014): "Growth-returns nexus: Evidence from three Central and Eastern European countries," *Economic Modelling*, 42, 343 – 355.

Miralles-Marcelo, J. L., J. L. Miralles-Quirs, and M. d. M. Miralles-Quirs (2013): "Improving the CARR model using extreme range estimators," *Applied Financial Economics*, 23, 1635–1647.

Molnár, P. (2012): "Properties of range-based volatility estimators," *International Review of Financial Analysis*, 23, 20–29.

Nelson, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Parkinson, M. (1980): "The Extreme Value Method for Estimating the Variance of the Rate of Return," *The Journal of Business*, 53, 61–65.

Patton, A. J. (2011): "Volatility forecast comparison using imperfect volatility proxies," *Journal of Econometrics*, 160, 246–256.

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

Shephard, N. and K. Sheppard (2010): "Realising the future: forecasting with high-frequency-based volatility (HEAVY) models," *Journal of Applied Econometrics*, 25, 197–231.

Taylor, S. J. (1986): *Modelling Financial Time Series*, John Wiley and Sons, 2nd ed.

A Appendix

Table 1: Estimated coefficients of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$, reported together with the values of Akaike Information Criterion (AIC) of the respective equations.

Ticker	GARCH(1,1)				RGARCH(1,1)			
	ω	α	β	AIC	ω	α	β	AIC
AA	1.61E-06	0.036	0.960	-5.121	4.21E-06	0.066	0.926	-5.131
AXP	1.61E-06	0.071	0.927	-5.320	2.26E-06	0.160	0.842	-5.348
BA	2.67E-06	0.057	0.934	-5.497	5.20E-06	0.148	0.830	-5.520
BAC	1.69E-06	0.080	0.917	-5.508	1.77E-06	0.197	0.816	-5.529
CAT	2.78E-06	0.045	0.947	-5.303	1.11E-05	0.145	0.826	-5.325
CSCO	2.98E-06	0.078	0.921	-4.756	4.04E-06	0.184	0.814	-4.787
CVX	3.29E-06	0.066	0.917	-5.838	5.20E-06	0.134	0.840	-5.854
DD	1.04E-06	0.038	0.959	-5.551	2.53E-06	0.088	0.901	-5.573
DIS	2.57E-06	0.053	0.939	-5.460	5.51E-06	0.107	0.867	-5.494
GE	8.38E-07	0.062	0.937	-5.742	2.54E-06	0.180	0.811	-5.765
HD	2.82E-06	0.053	0.939	-5.313	7.22E-06	0.121	0.852	-5.334
HPQ	2.15E-06	0.035	0.961	-4.997	3.06E-06	0.054	0.941	-5.008
IBM	8.21E-07	0.054	0.946	-5.552	6.67E-07	0.153	0.860	-5.574
INTC	2.60E-06	0.054	0.942	-4.943	4.52E-06	0.142	0.855	-4.966
JNJ	1.28E-06	0.069	0.926	-6.021	1.47E-06	0.170	0.824	-6.044
JPM	1.82E-06	0.080	0.919	-5.273	1.86E-06	0.158	0.841	-5.307
CAG	1.80E-06	0.057	0.936	-5.815	5.68E-06	0.238	0.740	-5.843
KO	5.68E-07	0.044	0.954	-5.965	6.22E-07	0.114	0.883	-5.980
MCD	1.84E-06	0.046	0.947	-5.654	2.28E-06	0.091	0.898	-5.673
MMM	1.57E-06	0.033	0.959	-5.890	8.19E-06	0.136	0.814	-5.911
MRK	6.02E-06	0.058	0.920	-5.513	1.17E-05	0.124	0.826	-5.533
MSFT	1.05E-06	0.062	0.937	-5.392	6.69E-07	0.195	0.809	-5.408
PFE	1.80E-06	0.046	0.948	-5.509	6.52E-06	0.177	0.805	-5.520
PG	1.69E-06	0.057	0.934	-5.953	4.79E-06	0.213	0.764	-5.989
T	1.27E-06	0.057	0.940	-5.621	2.36E-06	0.109	0.881	-5.629
TRV	3.95E-06	0.074	0.913	-5.544	9.41E-06	0.198	0.782	-5.586
UTX	2.44E-06	0.074	0.918	-5.700	5.05E-06	0.198	0.788	-5.723
VZ	1.46E-06	0.052	0.943	-5.695	4.34E-06	0.159	0.826	-5.704
WMT	1.39E-06	0.058	0.939	-5.617	1.91E-06	0.127	0.861	-5.638
XOM	2.70E-06	0.074	0.912	-5.922	5.32E-06	0.164	0.807	-5.949

Table 2: Estimated coefficients and p-values for the combined GARCH(1,1) model $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \alpha_2 \widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$.

Ticker	combined GARCH(1,1)							
	ω	p-value	α_1	p-value	β	p-value	α_2	p-value
AA	4.37E-06	0.000	-0.002	0.811	0.925	0.000	0.069	0.000
AXP	2.33E-06	0.004	-0.041	0.003	0.827	0.000	0.218	0.000
BA	6.07E-06	0.000	-0.028	0.013	0.810	0.000	0.191	0.000
BAC	1.76E-06	0.002	0.007	0.546	0.819	0.000	0.187	0.000
CAT	1.47E-05	0.000	-0.052	0.000	0.783	0.000	0.231	0.000
CSCO	3.82E-06	0.015	-0.025	0.058	0.812	0.000	0.211	0.000
CVX	5.67E-06	0.000	-0.018	0.135	0.829	0.000	0.161	0.000
DD	2.78E-06	0.000	-0.025	0.002	0.896	0.000	0.117	0.000
DIS	5.88E-06	0.000	-0.034	0.001	0.864	0.000	0.140	0.000
GE	2.56E-06	0.000	-0.005	0.704	0.809	0.000	0.186	0.000
HD	8.19E-06	0.000	-0.018	0.095	0.837	0.000	0.150	0.000
HPQ	3.01E-06	0.000	0.001	0.849	0.941	0.000	0.053	0.000
IBM	6.69E-07	0.353	-0.010	0.178	0.853	0.000	0.171	0.000
INTC	4.90E-06	0.012	-0.032	0.006	0.842	0.000	0.187	0.000
JNJ	1.47E-06	0.000	0.005	0.598	0.826	0.000	0.162	0.000
JPM	1.90E-06	0.017	-0.030	0.013	0.829	0.000	0.200	0.000
CAG	6.83E-06	0.000	-0.042	0.002	0.699	0.000	0.315	0.000
KO	6.15E-07	0.046	-0.002	0.773	0.882	0.000	0.117	0.000
MCD	4.61E-06	0.000	-0.041	0.000	0.841	0.000	0.178	0.000
MMM	9.43E-06	0.000	-0.092	0.000	0.790	0.000	0.242	0.000
MRK	1.41E-05	0.000	-0.029	0.009	0.796	0.000	0.173	0.000
MSFT	5.69E-07	0.534	-0.018	0.240	0.798	0.000	0.224	0.000
PFE	6.28E-06	0.000	0.007	0.496	0.813	0.000	0.163	0.000
PG	5.18E-06	0.000	-0.061	0.000	0.733	0.000	0.303	0.000
T	1.95E-06	0.001	0.026	0.000	0.894	0.000	0.072	0.000
TRV	1.03E-05	0.000	-0.041	0.000	0.768	0.000	0.252	0.000
UTX	5.49E-06	0.000	-0.020	0.107	0.773	0.000	0.232	0.000
VZ	3.96E-06	0.000	0.018	0.009	0.840	0.000	0.129	0.000
WMT	1.97E-06	0.003	-0.010	0.338	0.855	0.000	0.142	0.000
XOM	5.75E-06	0.000	-0.030	0.021	0.794	0.000	0.204	0.000

Table 3: Comparison of the forecasting performance of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$. Numbers in this table are 1000×RMSE of the one-day-ahead rolling window forecast reported for different window sizes w. An asterisk * (**) indicates when the difference is significant at the 5% (1%) level.

Ticker	GARCH(1,1)				RGARCH(1,1)			
	w=300	w=400	w=500	w=600	w=300	w=400	w=500	w=600
AA	1.277	1.296	1.309	1.322	1.268	1.281	1.291	1.305
AXP	1.167	1.177	1.189	1.202	1.179	1.199	1.203	1.215
BA	0.656	0.657	0.657	0.662	0.649	0.650	0.651	0.657
BAC	2.594	2.621	2.646	2.673	2.791	2.824	2.701	2.761
CAT	0.710	0.717	0.722	0.731	0.694*	0.701	0.710	0.719
CSCO	1.749	1.761	1.781	1.806	1.700	1.708*	1.736*	1.747*
CVX	0.643	0.648	0.657	0.662	0.634	0.635	0.642	0.647
DD	0.675	0.679	0.686	0.692	0.660*	0.665**	0.671**	0.677**
DIS	0.684	0.688	0.696	0.703	0.665*	0.669*	0.678*	0.682*
GE	0.869	0.870	0.879	0.888	0.882	0.865	0.862	0.871
HD	0.794	0.801	0.809	0.815	0.789	0.800	0.800	0.844
HPQ	1.050	1.058	1.070	1.083	1.043	1.057	1.063	1.077
IBM	0.631	0.635	0.641	0.648	0.624*	0.629*	0.637	0.643
INTC	1.194	1.195	1.205	1.218	1.161*	1.169*	1.180*	1.193*
JNJ	0.359	0.358	0.356	0.357	0.350*	0.349*	0.350	0.351
JPM	1.757	1.787	1.805	1.817	1.711	1.724*	1.736**	1.758**
CAG	0.534	0.537	0.538	0.543	0.514	0.531	0.536	0.542
KO	0.496	0.495	0.497	0.500	0.488	0.488	0.491	0.496
MCD	0.670	0.670	0.676	0.678	0.665	0.667	0.682	0.694
MMM	0.446	0.446	0.451	0.455	0.444	0.445	0.449	0.452
MRK	0.642	0.649	0.653	0.660	0.632*	0.636**	0.639*	0.649**
MSFT	0.676	0.683	0.688	0.696	0.676	0.673*	0.675**	0.684**
PFE	0.540	0.546	0.545	0.553	0.546	0.547	0.552	0.555
PG	0.505	0.508	0.509	0.510	0.493*	0.493**	0.498	0.498
T	0.612	0.614	0.619	0.626	0.597	0.601*	0.608*	0.613*
TRV	1.161	1.169	1.177	1.190	1.180	1.178	1.185	1.188
UTX	0.689	0.698	0.701	0.710	0.681*	0.686**	0.695*	0.702**
VZ	0.570	0.573	0.577	0.583	0.561**	0.563**	0.569*	0.575**
WMT	0.625	0.628	0.633	0.640	0.612	0.618	0.619	0.628
XOM	0.610	0.612	0.614	0.621	0.588**	0.590**	0.597*	0.604*

Table 4: Comparison of the forecasting performance of the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma^2_{P,t-1}} + \beta\sigma^2_{t-1}$and several different GARCH models. Numbers in this table are 1000×RMSE of the one-day-ahead rolling window forecast with forecasting window equal to 400.

ticker	RGARCH	GARCH	GJR	EGARCH	stdGARCH	astdGARCH	cGARCH
AA	1.281	1.296	1.286	1.277	1.294	1.270	1.309
AXP	1.199	1.177	1.189	1.174	1.173	1.178	1.177
BA	0.648	0.655	0.647	0.659	0.655	0.650	0.650
BAC	2.825	2.623	2.654	2.549	2.631	2.595	2.550
CAT	0.705	0.720	0.716	0.718	0.722*	0.716	0.723*
CSCO	1.881	1.928**	1.963	1.895	1.909*	1.888	1.937*
CVX	0.633	0.646*	0.628	0.630	0.653	0.632	0.662**
DD	0.663	0.678**	0.676*	0.683**	0.678**	0.680**	0.678**
DIS	0.668	0.688*	0.685	0.688	0.690	0.689	0.690*
GE	0.863	0.869	0.862	0.855	0.866	0.863	0.887
HD	0.803	0.804	0.799	0.799	0.807	0.799	0.803
HPQ	1.057	1.058	1.056	1.059	1.058	1.056	1.071*
IBM	0.639	0.645	0.633	0.635	0.642	0.633*	0.650*
INTC	1.170	1.196*	1.160	1.158	1.175	1.156	1.207*
JNJ	0.347	0.355*	0.351	0.351	0.353	0.351	0.355*
JPM	1.724	1.786*	1.711	1.715	1.782*	1.730	1.761
CAG	0.531	0.537	0.536	0.533	0.530	0.532	0.529
KO	0.485	0.492	0.505	0.492	0.487	0.488	0.491
MCD	0.669	0.672	0.695	0.824	0.663	0.663	0.668
MMM	0.442	0.443	0.444	0.441	0.442	0.442	0.447
MRK	0.635	0.648**	0.652**	0.648*	0.647*	0.647*	0.653**
MSFT	0.674	0.684*	0.675	0.676	0.686*	0.677	0.686*
PFE	0.562	0.561	0.567	0.555	0.556	0.554	0.560
PG	0.492	0.507**	0.507**	0.503*	0.503*	0.502*	0.508**
T	0.601	0.613*	0.607	0.611	0.613	0.609	0.613
TRV	1.176	1.167	1.174	1.173	1.176	1.175	1.171
UTX	0.685	0.697**	0.697	0.695	0.697**	0.691	0.703
VZ	0.562	0.571**	0.569	0.569	0.570**	0.566	0.574*
WMT	0.621	0.632	0.625	0.629	0.626	0.624	0.633
XOM	0.588	0.609**	0.595	0.594	0.613	0.600	0.618**

Table 5: Comparison of forecasting performance GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$. Numbers in this table are the log-likelihood function (26) of the returns r_t being drawn from the distributions $N\left(0, \widehat{\sigma_t^2}\right)$, where $\widehat{\sigma_t^2}$ is a one-day-ahead rolling window volatility forecast reported for different window sizes w.

Ticker	GARCH(1,1)				RGARCH(1,1)			
	w=300	w=400	w=500	w=600	w=300	w=400	w=500	w=600
AA	9803	9580	9267	9020	9873	9597	9320	9032
AXP	10377	10166	9881	9595	10502	10242	9969	9688
BA	10434	10225	9809	9660	10500	10258	9993	9708
BAC	10687	10451	10154	9875	10783	10527	10236	9949
CAT	10105	9916	9631	9342	10202	9950	9675	9385
CSCO	9309	9080	8825	8528	9478	9237	8955	8646
CVX	11371	11017	10853	10576	11440	11145	10882	10599
DD	10853	10593	10321	10050	10916	10641	10377	10095
DIS	10535	10298	10024	9747	10681	10411	10142	9859
GE	11086	10860	10567	10258	11176	10902	10617	10325
HD	10266	10024	9729	9479	10372	10084	9809	9542
HPQ	9587	9255	9076	8792	9715	9415	9174	8869
IBM	10813	10575	10247	9972	10986	10716	10378	10130
INTC	9420	9208	8937	8665	9484	9278	9001	8735
JNJ	12013	11776	11492	11230	12063	11126	11522	11264
JPM	10158	10014	9730	9464	10345	10113	9830	9554
CAG	11421	11192	10939	10681	11563	11301	10994	10722
KO	11682	11454	11155	10924	11782	11517	11250	10980
MCD	11058	10846	10564	10288	11129	10871	10596	10320
MMM	11238	11105	10819	10542	11377	11153	10878	10599
MRK	10131	9775	9632	9294	10348	10120	9813	9570
MSFT	10234	10038	9724	9478	10396	10171	9867	9611
PFE	10741	10499	10222	9959	10827	10564	10269	10004
PG	11512	11236	10962	10709	11571	11369	11081	10777
T	10948	10704	10459	10172	11002	10744	10473	10206
TRV	10801	10614	10312	10069	10899	10678	10395	10111
UTX	11013	10790	10477	10179	11054	10840	10556	10269
VZ	11132	10892	10605	10329	11198	10930	10645	10361
WMT	11004	10778	10510	10109	11130	10860	10558	10276
XOM	11464	11223	10947	10657	11567	11294	11014	10729

Table 6: Estimated coefficients of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and it modified version RGARCH(1,1) $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$, reported together with the values of Akaike Information Criterion (AIC) of the respective equations for the simulated data.

Index	GARCH(1,1)				RGARCH(1,1)			
	ω	α	β	AIC	ω	α	β	AIC
CAC40	1.03E-06	0.075	0.920	-6.327	1.80E-06	0.182	0.821	-6.352
DAX	6.16E-07	0.088	0.911	-6.417	1.28E-06	0.174	0.842	-6.446
DJI	9.39E-07	0.083	0.910	-6.674	-1.77E-06	0.128	0.717	-6.645
FTSE	7.64E-07	0.085	0.910	-6.581	1.47E-06	0.188	0.837	-6.598
NASDAQ	9.43E-07	0.056	0.942	-5.534	4.30E-07	0.135	0.893	-5.561
NIKKEI	3.20E-06	0.093	0.890	-6.084	1.64E-06	0.179	0.854	-6.113

Table 7: Estimated coefficients and p-values for the combined GARCH(1,1) model $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \alpha_2\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$.

Index	combined GARCH(1,1)							
	ω	p-value	α_1	p-value	β	p-value	α_2	p-value
CAC40	1.93E-06	0.000294	-0.064	7.02E-05	0.789	0	0.286	0
DAX	1.61E-06	4.00E-15	-0.064	2.74E-05	0.815	0	0.276	0
DJI	5.69E-07	0.003	0.080	0	0.896	0	0.008	1.45E-04
FTSE	1.51E-06	1.33E-05	-0.005	0.723	0.834	0	0.198	5.06E-11
NASDAQ	3.07E-07	0.553	-0.050	2.97E-05	0.891	0	0.204	0
NIKKEI	1.11E-06	0.031043	-0.088	1.72E-10	0.837	0	0.319	0

Table 8: Comparison of the forecasting performance of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$. Numbers in this table are 1000×RMSE of the one-day-ahead rolling window forecasts reported for different window sizes w and different benchmarks (squared returns r^2 and the realized variance RV) for the stock indices.

Index	Bench	GARCH(1,1)				RGARCH(1,1)			
		w=300	w=400	w=500	w=600	w=300	w=400	w=500	w=600
CAC40	r^2	0.335	0.339	0.342	0.346	0.331	0.335	0.338	0.342
	RV	0.185	0.181	0.179	0.180	0.172**	0.169**	0.167**	0.167**
DAX	r^2	0.474	0.477	0.481	0.488	0.446**	0.454**	0.461*	0.469*
	RV	0.252	0.242	0.236	0.235	0.212**	0.208**	0.207**	0.207**
DJI	r^2	0.353	0.355	0.362	0.367	0.336	0.341	0.347	0.350
	RV	0.174	0.172	0.176	0.179	0.142**	0.142**	0.141**	0.139**
FTSE	r^2	0.376	0.382	0.385	0.390	0.364*	0.368**	0.372**	0.377**
	RV	0.201	0.226	0.212	0.209	0.196	0.202*	0.189*	0.186*
NASDAQ	r^2	0.931	0.939	0.949	0.963	0.908**	0.917**	0.929**	0.942**
	RV	0.464	0.452	0.440	0.446	0.431*	0.423*	0.426	0.432
NIKKEI	r^2	0.467	0.475	0.478	0.478	0.456*	0.461	0.467	0.470
	RV	0.237	0.283	0.269	0.249	0.196**	0.188**	0.177**	0.173**

Table 9: Estimated coefficients of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$, reported together with the values of Akaike Information Criterion (AIC) of the respective equations for the simulated data.

	GARCH(1,1)				RGARCH(1,1)			
	ω	α	β	AIC	ω	α	β	AIC
$\mu_{0.5}$	1.77E-04	0.016	0.958	-2.143	1.76E-04	0.053	0.922	-2.149
μ_1	1.73E-04	0.044	0.933	-2.112	1.61E-04	0.122	0.857	-2.133
μ_2	1.50E-04	0.114	0.875	-2.037	1.20E-04	0.274	0.723	-2.101

Table 10: Estimated coefficients and p-values for the combined GARCH(1,1) model $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \alpha_2\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$ for the simulated data.

	combined GARCH(1,1)			
	ω	α_1	β	α_2
$\mu_{0.5}$	1.71E-04	-0.027	0.908	0.094
μ_1	1.53E-04	-0.057	0.834	0.204
μ_2	1.13E-04	-0.119	0.686	0.431

Table 11: Comparison of the forecasting performance of the GARCH(1,1) model $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$. Numbers in this table are 1000×RMSE of the one-day-ahead rolling window forecasts reported for different window sizes w and different benchmarks squared returns (r^2) and the realized variance (RV) for the simulated data. The differences in MSE are significant at any significance level (due to very large number of observations).

	GARCH(1,1)				RGARCH(1,1)				σ_{true}^2
	w=300	w=400	w=500	w=600	w=300	w=400	w=500	w=600	
r^2 as a benchmark									
$\mu_{0.5}$	10.16	10.14	10.12	10.11	10.15	10.12	10.10	10.09	9.99
μ_1	11.90	11.86	11.84	11.83	11.78	11.74	11.72	11.71	11.49
μ_2	20.31	20.22	20.11	20.07	19.78	19.71	19.63	19.60	18.98
σ_{true}^2 as a benchmark									
$\mu_{0.5}$	1.81	1.71	1.63	1.57	1.71	1.59	1.49	1.43	0
μ_1	3.00	2.88	2.80	2.75	2.52	2.32	2.21	2.15	0
μ_2	7.15	6.97	6.84	6.72	5.52	5.30	5.22	5.15	0