UNIVERSITY OF ECONOMICS IN PRAGUE

FACULTY OF FINANCE AND ACCOUNTING

HABILITATION THESIS

Volatility modeling and forecasting: utilization of realized volatility, implied volatility and the highest and lowest price of the day.

Author:

PETER MOLNÁR

2020

Acknowledgements

I would like to express an immense gratitude to doc. Karel Brůna for his invaluable help not only with this habilitation, but also at many other occasions.

Contents

Chapter 1

Introduction

This habilitation thesis is based on my published articles about volatility modelling and forecasting. Articles therefore constitute the main part of the thesis. In addition, an introduction of this habilitation serves as both introduction to this topic, as well as a brief nontechnical summary of the work I have done in this field.

1.1 Volatility in financial markets

Financial markets play a very important role in the modern society, as they significantly contribute to an efficient capital allocation. Therefore, their understanding is of utmost importance. The most fundamental variables in studies about financial markets are returns and volatility.

Return r_t, a relative change in the price P of an asset from time $t-1$ to t can be defined as follows:

$$P_t = ln(P_t) - ln(P_t) \tag{1.1}$$

Volatility is a measure of variability of returns. Volatility can be defined as a standard deviation of returns. However, the term volatility is sometimes also used to denote variance of returns, or logarithm of variance/standard deviation. In many cases, it does not matter which of these definitions is used, as these measures are monotonic transformations of each other. For example, a sentence like "Volatility is increased around major macroeconomic announcements." means the same no matter which definition of volatility has reader in mind. Therefore, I will use the term volatility when it is not important to distinguish between these definitions, and I will be more specific about standard deviation / variance / logarithmic variance whenever it is important.

In a simplest case, for example an estimation of monthly volatility from daily returns, it is easy to estimate volatility directly from its definition. However, volatility is usually changes from day to day, and therefore, some volatility models are needed.

1.2 Traditional models based on daily returns

The era of volatility modeling started with Engle (1982), whose idea was generalized by Bollerslev (1986). Daily asset returns r_t can be described in the following way:

$$r_t = \mu_t + \sigma_t Z_t \tag{1.2}$$

The drift μ_t can be very often approximated by an autoregressive process of first order, and often even by a constant, or zero. The Z_t is an innovation distributed according to some given distribution, and σ_t is the volatility process. The Generalized autoregressive conditional heteroskedacity (GARCH) model of Bollerslev (1986) specifies the σ_t process in the following way:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.3}$$

where p is the number of lags of the conditional variance and q is the number of lags of the squared errors. This model also imposes the following restrictions on the estimated parameters:

$$\alpha_0 > 0, \qquad \alpha_i \geq 0, \qquad \beta_i \geq 0 \tag{1.4}$$

The GARCH model of Bollerslev (1986) has been extended and modified in many ways. In equity markets, negative returns are usually followed by increased volatility. This effect can be captured e.g. by models of Glosten et al. (1993) and Nelson (1991).

For example, the conditional variance in the Glosten, Jagannathan, Runkle GARCH (GJR-GARCH(p,q)) model is specified as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \left(\alpha_i + \gamma_i I_{t-i}\right) Z_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.5}$$

with the indicator function I_{t-i} specified as follows:

$$I_{t-i} = \begin{cases} 1 & \text{if } Z_{i-t} < 0, \quad \text{for } i = 1, 2, \ldots, q; \\ 0 & \text{if } Z_{t-i} \geq 0, \quad \text{for } i = 1, 2, \ldots, q. \end{cases}$$

and restrictions on the estimated parameters, where γ_i is the asymmetry parameter:

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad \beta_i \geq 0, \quad \alpha_i + \gamma_i \geq 0. \tag{1.6}$$

It should be mentioned that GARCH models are not the only type of volatility models based on daily returns. Taylor (1982) introduced stochastic volatility models. However, these models require more advanced estimation techniques than the GARCH models (Ghysels et al., 1996; Harvey et al., 1994; Harvey & Shephard, 1996), and therefore have never become as popular as GARCH models.

GARCH models has been extended in various ways, and some of them can be quite complex. However, what have all the standard GARCH models in common is the data which are used in their estimation - past returns, usually past daily returns. Past daily returns contain only limited information about past daily volatility. Therefore, if we use data more informative about volatility, it should be possible to construct volatility models which perform better. This is the field where I have been active, but obviously many other researchers has explored this field too.

There are three most natural candidates for additional data in volatility models:

- to utilize not just closing price of the day, but also opening price, the highest price of the day and the lowest price of the day
- to utilize intraday high-frequency data
- to extract volatility implied by option prices

The main benefit of daily high, low and opening price is that they are almost always available, and therefore it is easy to use them. At the same time, they provide very large improvements in the accuracy of volatility estimates relatively to squared returns.

Intraday high-frequency data can provide even more information about intraday volatility. Therefore, whenever these data are available, and high precision of volatility estimates is needed, these data should be used. However, they are usually not publicly available without a paid subscription, exist only for some assets, and only over approximately last 20 years. Moreover, working with them requires special considerations, as they are subject to market micro-structure noise, such as bid-ask spread.

Volatility implied by option prices has huge advantage that it is a forward-looking volatility measure. However, it can be utilized only for those assets where options with this asset as an underlying exist.

1.3 Daily high-low range

Denote the opening price of the day t as O_t, the highest price H_t, the lowest price L_t, and the closing C_t. The set of these four prices for each day should contain more information about the volatility than daily returns. Intuitively, particularly the highest and the lowest price of the day should be very useful in volatility estimation. This idea was first formalized by Parkinson (1980) and Garman and Klass (1980). In a similar manner as prices are usually transformed into returns, it is useful to transform these prices into their return form in the following way. Define:

$$c_t = ln(C_t) - ln(O_t) \tag{1.7}$$

$$h_t = ln(H_t) - ln(O_t) \tag{1.8}$$

$$l_t = ln(L_t) - ln(O_t) \tag{1.9}$$

Then the Parkinson (1980) estimator is defined as:

$$PK_t = \frac{(h_t - l_t)^2}{4ln2} \tag{1.10}$$

while the Garman and Klass (1980) estimator is defined as:

$$GK_t = 0.511\,(h_t - l_t)^2 - 0.019\,(c_t(h_t + l_t) - 2h_t l_t) - 0.383c_t^2 \tag{1.11}$$

The difference $h_t - l_t$ is sometimes called range, and therefore estimators based on this variable are sometimes called range-based volatility estimators. Here belong also e.g. Rogers and Satchell (1991) estimator:

$$RS_t = h_t(h_t - c_t) + l_t(l_t - c_t) \tag{1.12}$$

In the first paper of this habilitation, Molnár (2012), I study properties of range-based volatility estimators. Range-based estimators can be conveniently utilized in any study where more precise volatility estimates are desirable, see e.g. Kim et al. (2019).

It is quite intuitive that highest and lowest prices of the day could improve GARCH models based on returns calculated from closing prices. In particular, squared residuals in GARCH model serve as a proxy for variance in that particular day. Therefore, simple replacement of squared residuals by the variance estimated from the highest and lowest price of the day should lead to better volatility model. This intuition is formalized and confirmed in the second paper of this thesis, Molnár (2016), where I suggest and evaluate a Range-GARCH model. Molnár (2016) is not the only paper which utilizes high-low range, see e.g. Chou (2005) and Brandt and Jones (2006). However, both these papers model the high-low range, while Range-GARCH model is a simple extension of the GARCH(1,1) model. The advantage of this approach is that is can be easily estimated in any software that can estimate the GARCH model.

Early GARCH models were univariate models, they have been quickly generalized to multivariate models. Multivariate GARCH volatility models suffer from dimensionality problem, as the number of estimated parameters increases quadraticaly with the number of assets. Therefore, probably the most popular multivariate GARCH model is the dynamic conditional correlation (DCC) model introduced independently by Engle (1982) and Tse and Tsui (2002).

In the third paper of this habilitation, Fiszeder et al. (2019), we incorporate Parkinson (1980) volatility estimator in the DCC model in a similar way as in Molnár (2016) and found that the Range-GARCH DCC model outperforms the standard GARCH DCC model not only in statistical sense, but also in terms of Value-at-Risk forecasts. Moreover, we find that the DCC model extended with the Range-GARCH model works better than the DCC model extended with the model of Chou (2005).

1.4 Realized volatility

Soon after the intraday high-frequency data became available for research, several people understood the benefits of these data, and the idea of the realized volatility introduced (Andersen et al., 2001a; Andersen et al., 2001b; Barndorff-Nielsen & Shephard, 2002).

In principle, realized variance RV during the day t can be estimated in the following way:

$$RV_t = \sum_{j=1}^{m} r_{t,j}^2 \tag{1.13}$$

where $r_{t,j}$ is the j^{th} intraday return on day t, m is the number of intraday returns.

The formula above captures the idea behind the realized variance. In reality, this formula is often extended to account for example for the autocorrelation in intraday returns (Liu et al., 2015; Patton & Sheppard, 2009):

$$RV_{AC,t} = \left[\sum_{j=1}^{m} r_{t,j}^2 + 2 \times \sum_{j=1}^{m-1} r_{t,j+1} r_{t,j} \right] \tag{1.14}$$

or the presence of jumps in the price, leading to estimators of bipower estimator of Barndorff-Nielsen and Shephard, 2004:

$$RV_{BV,t} = \frac{1\pi}{2} \times \sum_{j=1}^{m-1} |r_{t,j}||r_{t,j+1}| \tag{1.15}$$

or median realized variance of Andersen et al., 2012:

$$RV_{MV,t} = \frac{m\pi}{(m-1)(6-4\sqrt{3}+\pi)} \times \sum_{j=2}^{m-1} \left(med \left(|r_{t,j-1}|, |r_{t,j}|, |r_{t,j+1}| \right) \right)^2 \tag{1.16}$$

Even though estimation of realized volatility requires some effort, once we have the estimates of realized volatility, it is usually very easy to work with it. Probably the first volatility model based on realized variance is (Andersen et al., 2003). They find that models based on realized variance perform better than standard GARCH models. This conclusion is reached also by Horpestad et al. (2019). Volatility models based on realized variance have become very popular. I have used them in several of my papers, for example Aalborg et al. (2019), Haugom et al. (2014), Lyócsa and Molnár (2017, 2016), Lyócsa et al. (2016), Lyócsa et al. (2017).

In the fifth paper of this habilitation, Lyócsa and Molnár (2018), we study the impact of monetary policy announcements on stock market volatility in the U.S., Canada, Japan, the U.K., Germany, France and Italy. We find that volatility is increased on the day of interest rate announcements, and decreased for several days after the announcement. It would be more much difficult to arrive to clear conclusions without using realized volatility.

Realized volatility can of course used also in multivariate volatility models. Example of it is the fifth paper of this dissertation, Lyócsa et al. (2019). In this case we study volatility forecasting for crude oil and natural gas. However, in this particular case, the multivariate model does not perform better than univariate models.

1.5 Implied volatility

Price of the option depends on several parameters. One of these parameters is the volatility of the underlying asset. Since other parameters, such as interest rate or time to maturity, option pricing formula can be used to calculate a price for a given volatility. Moreover, since the relationship between the option price and volatility is monotonous, also the opposite is possible: for a given option price, it is possible to calculate volatility implied by the option.

The famous VIX index is an implied volatility, just not from one option, but from basket of options chosen and weighted in such a way that it is a 30-day implied volatility. The VIX index is often called the fear index, and it is one of the most followed financial indicators in the world. As a result, similar indices has been constructed also for many other markets, see Siriopoulos and Fassas (2019) for an overview. I have also contributed in this area (Birkelund et al., 2015; Bugge et al., 2016).

Since implied volatility is forward-looking, it is a very useful measure for various purposes. Even though implied volatility should not be interpreted directly as expected volatility (since it contains also risk premium), it is still very useful in volatility forecasting, as have been confirmed by various studies (Christensen & Prabhala, 1998; Prokopczuk & Simen, 2014), including studies in which I participated (Bugge et al., 2016; Haugom et al., 2014).

The volatility is such an important information about financial markets that nowadays exists even derivatives on volatility, and this market is very active and growing. More specifically, there exist option and futures contracts with the VIX index as an underlying asset, and these derivatives have very interesting characteristics, see Bašta and Molnár (2019), Zhang and Zhu (2006) and Bordonado et al. (2017).

The last paper of this habilitation utilizes implied volatility and historical volatility estimated via Parkinson (1980) estimator. In particular, we study comovement in the stock market volatility and the oil market volatility, considering these both volatility measures. We find that the implied volatility of the stock market slightly leads the implied volatility of the oil market, while this feature is weaker between realized volatilities.

Chapter 2

Properties of range-based volatility estimators

International Review of Financial Analysis 23 (2012) 20–29

Contents lists available at ScienceDirect

International Review of Financial Analysis

Properties of range-based volatility estimators

Peter Molnár *

Norwegian School of Economics, Department of Finance and Management Science, Helleveien 30, 5045 Bergen, Norway
Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management, 7491 Trondheim, Norway

A R T I C L E I N F O

Article history:
Received 16 November 2010
Received in revised form 8 June 2011
Accepted 30 June 2011
Available online 13 July 2011

JEL classification:
C58
G17
G32

Keywords:
Volatility
High
Low
Range

A B S T R A C T

Volatility is not directly observable and must be estimated. Estimator based on daily close data is imprecise. Range-based volatility estimators provide significantly more precision, but still remain noisy volatility estimates, something that is sometimes forgotten when these estimators are used in further calculations.
First, we analyze properties of these estimators and find that the best estimator is the Garman–Klass (1980) estimator. Second, we correct some mistakes in existing literature. Third, the use of the Garman–Klass estimator allows us to obtain an interesting result: returns normalized by their standard deviations are approximately normally distributed. This result, which is in line with results obtained from high frequency data, but has never previously been recognized in low frequency (daily) data, is important for building simpler and more precise volatility models.

1. Introduction

Asset volatility, a measure of risk, plays a crucial role in many areas of finance and economics. Therefore, volatility modelling and forecasting become one of the most developed parts of financial econometrics. However, since the volatility is not directly observable, the first problem which must be dealt with before modelling or forecasting is always a volatility measurement (or, more precisely, estimation).

Consider stock price over several days. From a statistician's point of view, daily relative changes of stock price (stock returns) are almost random. Moreover, even though daily stock returns are typically of a magnitude of 1% or 2%, they are approximately equally often positive and negative, making average daily return very close to zero. The most natural measure for how much stock price changes is the variance of the stock returns. Variance can be easily calculated and it is a natural measure of the volatility. However, this way we can get only an average volatility over an investigated time period. This might not be sufficient, because volatility changes from one day to another. When we have daily closing prices and we need to estimate volatility on a daily basis, the only estimate we have is squared (demeaned) daily return. This estimate is very noisy, but since it is very often the only one we have, it is commonly used. In fact, we can look at most of the volatility models (e.g. GARCH class of models or stochastic volatility models) in such a way that daily volatility is first estimated as squared returns and consequently processed by applying time series techniques.

When not only daily closing prices, but intraday high frequency data are available too, we can estimate daily volatility more precisely. However, high frequency data are in many cases not available at all or available only over a shorter time horizon and costly to obtain and work with. Moreover, due to market microstructure effects the volatility estimation from high frequency data is rather a complex issue (see Dacorogna, Gencay, Müller, Olsen, & Pictet, 2001).

However, closing prices are not the only easily available daily data. For the most of financial assets, daily open, high and low prices are available too. Range, the difference between high and low prices is a natural candidate for the volatility estimation. The assumption that the stock return follows a Brownian motion with zero drift during the day allows Parkinson (1980) to formalize this intuition and derive a volatility estimator for the diffusion parameter of the Brownian motion. This estimator based on the range (the difference between high and low prices) is much less noisy than squared returns. Garman and Klass (1980) subsequently introduce estimator based on open, high, low and close prices, which is even less noisy. Even though these estimators have existed for more than 30 years, they have been rarely used in the past by both academics and practitioners. However, recently the literature using the range-based volatility estimators started to grow (e.g. Alizadeh, Brandt, and Diebold (2002), Brandt and Diebold (2006), Brandt and Jones (2006), Chou (2005), Chou (2006), Chou and Liu (2010)). For an overview see Chou et al. (2010).

* Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management, 7491 Trondheim, Norway.
E-mail address: Peter.Molnar@iot.ntnu.no.

1057-5219/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.irfa.2011.06.012

Despite increased interest in the range-based estimators, their properties are sometimes somewhat imprecisely understood. One particular problem is that despite the increased accuracy of these estimators in comparison to squared returns, these estimators still only provide a noisy estimate of volatility. However, in some manipulations (e.g. division) people treat these estimators as if they were exact values of the volatility. This can in turns lead to flawed conclusions, as we show later in the paper. Therefore we study these properties.

Our contributions are the following. First, when the underlying assumptions of the range-based estimators hold, all of them are unbiased. However, taking the square root of these estimators leads to biased estimators of standard deviation. We study this bias. Second, for a given true variance, distribution of the estimated variance depends on the particular estimator. We study these distributions. Third, we show how the range-based volatility estimators should be modified in the presence of opening jumps (stock price at the beginning of the day typically differs from the closing stock price from the previous day).

Fourth, the property we focus on is the distribution of returns standardized by standard deviations. A question of interest is how this is affected when the standard deviations are estimated from range-based volatility estimators. The question whether the returns divided by their standard deviations are normally distributed has important implications for many fields in finance. Normality of returns standardized by their standard deviations holds promise for simple-to-implement and yet precise models in financial risk management. Using volatility estimated from high frequency data, Andersen, Bollerslev, Diebold, and Labys (2000), Andersen, Bollerslev, Diebold, and Ebens (2001), Forsberg and Bollerslev (2002) and Thomakos and Wang (2003) show that standardized returns are indeed Gaussian. Contrary, returns scaled by standard deviations estimated from GARCH type of models (which are based on daily returns) are not Gaussian, they have heavy tails. This well-known fact is the reason why heavy-tailed distributions (e.g. t-distribution) were introduced into the GARCH models. We show that when properly used, range-based volatility estimators are precise enough to replicate basically the same results as those of Andersen et al. (2001) obtained from high frequency data. To our best knowledge, this has not been previously recognized in the daily data. Therefore volatility models built upon high and low data might provide accuracy similar to models based upon high frequency data and still keep the benefits of the models based on low frequency data (much smaller data requirements and simplicity).

The rest of the paper is organized in the following way. In Section 2, we describe existing range-based volatility estimators. In Section 3, we analyze properties of range-based volatility estimators, mention some caveats related to them and correct some mistakes in the existing literature. In Section 4 we empirically study the distribution of returns normalized by their standard deviations (estimated from range-based volatility estimators) on 30 stock, the components of the Dow Jones Industrial Average. Section 5 concludes.

2. Overview

Assume that price P follows a geometric Brownian motion such that log-price $p=ln(P)$ follows a Brownian motion with zero drift and diffusion σ.

$$dp_t = \sigma dB_t \tag{1}$$

Diffusion parameter σ is assumed to be constant during one particular day, but can change from one day to another. We use one day as a unit of time. This normalization means that the diffusion parameter in (1) coincides with the daily standard deviation of returns and we do not need to distinguish between these two quantities. Denote the price at the beginning of the day (i.e. at the time $t=0$) O (open), the price in the end of the day (i.e. at the time $t=1$) C (close), the highest price of the day H, and the lowest price of the day L. Then we can calculate open-to-close, open-to-high and open-to-low returns as

$$c = ln(C) - ln(O) \tag{2}$$

$$h = ln(H) - ln(O) \tag{3}$$

$$l = ln(L) - ln(O) \tag{4}$$

Daily return c is obviously a random variable drawn from a normal distribution with zero mean and variance (volatility) σ^2

$$c \sim N\left(0, \sigma^2\right) \tag{5}$$

Our goal is to estimate (unobservable) volatility σ^2 from observed variables c, h and l. Since we know that c^2 is an unbiased estimator of σ^2,

$$E\left(c^2\right) = \sigma^2 \tag{6}$$

we have the first volatility estimator (subscript s stands for "simple")

$$\sigma_s^2 = c^2 \tag{7}$$

Since this simple estimator is very noisy, it is desirable to have a better one. It is intuitively clear that the difference between high and low prices tells us much more about volatility than close price. High and low prices provide additional information about volatility. The distribution of the range $d \equiv h - l$ (the difference between the highest and the lowest value) of Brownian motion is known (Feller (1951)). Define $P(x)$ to be the probability that $d \leq x$ during the day. Then

$$P(x) = \sum_{n=1}^{\infty} (-1)^{n+1} n \left\{ Erfc\left(\frac{(n+1)x}{\sqrt{2}\sigma}\right) - 2Erfc\left(\frac{nx}{\sqrt{2}\sigma}\right) + Erfc\left(\frac{(n-1)x}{\sqrt{2}\sigma}\right) \right\} \tag{8}$$

where

$$Erfc(x) = 1 - Erf(x) \tag{9}$$

and $Erf(x)$ is the error function. Using this distribution Parkinson (1980) calculates (for $p \geq 1$)

$$E(d^p) = \frac{4}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right)\left(1 - \frac{4}{2^p}\right)\zeta(p-1)\left(2\sigma^2\right) \tag{10}$$

where $\Gamma(x)$ is the gamma function and $\zeta(x)$ is the Riemann zeta function. Particularly for $p=1$

$$E(d) = \sqrt{8\pi}\sigma \tag{11}$$

and for $p=2$

$$E\left(d^2\right) = 4\, ln(2)\sigma^2 \tag{12}$$

Based on formula (12), he proposes a new volatility estimator:

$$\widehat{\sigma_P^2} = \frac{(h-l)^2}{4\, ln\, 2} \tag{13}$$

Garman and Klass (1980) realize that this estimator is based solely on quantity $h-l$ and therefore an estimator which utilizes all the available information c, h and l will be necessarily more precise. Since search for the minimum variance estimator based on c, h and l is an infinite dimensional problem, they restrict this problem to analytica estimators, i.e. estimators which can be expressed as an analytical function of c, h and l. They find that the minimum variance analytical estimator is given by the formula

$$\widehat{\sigma^2_{GKprecise}} = 0.511(h-l)^2 - 0.019(c(h+l)-2hl) - 0.383c^2 \qquad (14)$$

The second term (cross-products) is very small and therefore they recommend neglecting it and using more practical estimator:

$$\widehat{\sigma^2_{GK}} = 0.5(h-l)^2 - (2\ln 2 - 1)c^2 \qquad (15)$$

We follow their advice and further on when we talk about Garman–Klass volatility estimator (GK), we refer to Eq. (15). This estimator has additional advantage over Eq. (14) – it can be simply explained as an optimal (smallest variance) combination of simple and Parkinson volatility estimator.

Meilijson (2009) derives another estimator, outside the class of analytical estimators, which has even smaller variance than GK. This estimator is constructed as follows.

$$\widehat{\sigma^2_M} = 0.274\sigma_1^2 + 0.16\sigma_s^2 + 0.365\sigma_3^2 + 0.2\sigma_4^2 \qquad (16)$$

where

$$\sigma_1^2 = 2\left[(h'-c')^2 + l'\right] \qquad (17)$$

$$\sigma_3^2 = 2(h'-c'-l')c' \qquad (18)$$

$$\sigma_4^2 = -\frac{(h'-c')l'}{2\ln 2 - 5/4} \qquad (19)$$

where $c'=c$, $h'=h$, $l'=l$ if $c>0$ and $c'=-c$, $h'=-l$, $l'=-h$ if $c<0$.[1]

Rogers and Satchell (1991) derive an estimator which allows for arbitrary drift.

$$\widehat{\sigma^2_{RS}} = h(h-c) + l(l-c) \qquad (20)$$

There are two other estimators which we should mention. Kunitomo (1992) derives a drift-independent estimator, which is more precise than all the previously mentioned estimators. However "high" and "low" prices used in his estimator are not the highest and lowest price of the day. The "high" and "low" used in this estimator are the highest and the lowest price relative to the trend line given by open and high prices. These "high" and "low" prices are unknown unless we have tick-by-tick data and therefore the use of this estimator is very limited.

Yang and Zhang (2000) derive another drift-independent estimator. However, their estimator can be used only for estimation of average volatility over multiple days and therefore we do not study it in our paper.

Efficiency of a volatility estimator $\widehat{\sigma}^2$ is defined as

$$Eff\left(\widehat{\sigma^2}\right) \equiv \frac{var\left(\sigma_s^2\right)}{var\left(\widehat{\sigma^2}\right)} \qquad (21)$$

Simple volatility estimator has by definition efficiency 1, Parkinson volatility estimator has efficiency 4.9, Garman–Klass 7.4 and Meilijson 7.7. Rogers, Satchell has efficiency 6.0 for the zero drift and larger than 2 for any drift.

[1] This estimator is not analytical, because it uses different formula for days when $c>0$ than for days when $c<0$.

Remember that all of the studied estimators except for Rogers, Satchell are derived under the assumption of zero drift. However, for most of the financial assets, mean daily return is much smaller than its standard deviation and can therefore be neglected. Obviously, this is not true for longer time horizons (e.g. when we use yearly data), but this is a very good approximation for daily data in basically any practical application.

Further assumptions behind these estimators are continuous sampling, no bid–ask spread and constant volatility. If prices are observed only infrequently, then the observed high will be below the true high and observed low will be above the true low, as was recognized already by Garman and Klass (1980). Bid–ask spread has the opposite effect: observed high price is likely to happen at ask, observed low price is likely to happen at the low price and therefore the difference between high and low contains in addition bid–ask spread. These effects work in the opposite direction and therefore they will at least partially cancel out. More importantly, for liquid stocks both these effects are very small. In this paper we maintain the assumption of constant volatility within the day. This approach is common even in stochastic volatility literature (e.g. Alizadeh et al., 2002) and assessing the effect of departing from this assumption is beyond the scope of this paper. However, this is an interesting avenue for further research.

3. Properties of range-based volatility estimators

The previous section provided an overview of range-based volatility estimators including their efficiency. Here we study their other properties. Our main focus is not their empirical performance, as this question has been studied before (e.g. Bali and Weinbaum (2005)). We study the performance of these estimators when all the assumptions of these estimators hold perfectly. This is more important than it seems to be, because this allows us to distinguish between the case when these estimators do not work (assumptions behind them do not hold) and the case when these estimators work, but we are misinterpreting the results. This point can be illustrated in the following example. Imagine that we want to study the distribution of returns standardized by their standard deviations. We estimate these standard deviations as a square root of the Parkinson volatility estimator Eq. (13) and find that standardized returns are not normally distributed. Should we conclude that true standardized returns are not normally distributed or should we conclude that the Parkinson volatility estimator is not appropriate for this purpose? We answer this and other related questions.

To do so, we ran 500,000 simulations, one simulation representing one trading day. During every trading day log-price p follows a Brownian motion with zero drift and daily diffusion $\sigma=1$. We approximate continuous Brownian motion by $n=100,000$ discrete intraday returns, each drawn from $N(0, 1/\sqrt{n})$.[2] We save high, low and close log-prices h, l, c for every trading day.[3]

3.1. Bias in σ

All the previously mentioned estimators are unbiased estimators of σ^2. Therefore, square root of any of these estimators will be a biased estimator of σ. This is direct consequence of well known fact that for a random variable x the quantities $E(x^2)$ and $E(x)^2$ are generally different. However, as I document later, using $\sqrt{\widehat{\sigma^2}}$ as $\widehat{\sigma}$, as an estimator of σ, is not uncommon. Moreover, in many cases the objects of our

[2] Such a high n allows us to have almost perfectly continuous Brownian motion and having so many trading days allow us to know the distributions of range based volatility estimators with very high precision. Simulating these data took one months on an ordinary computer (Intel Core 2 Duo P8600 2.4 GHz, 2 GB RAM). Note that we do not derive analytical formulas for the distributions of range-based volatility estimators. Since these formulas would not bring additional insights into the questions we study, their derivation is behind the scope of this paper.

[3] Open log-price is normalized to zero.

interests are standard deviations, not variances. Therefore, it is important to understand the size of the error introduced by using $\sqrt{\widehat{\sigma^2}}$ instead of $\hat{\sigma}$ and potentially correct for this bias. Size of this bias depends on the particular estimator.

As can be easily proved, an unbiased estimator $\widehat{\sigma_s}$ of the standard deviation σ based on $\sqrt{\widehat{\sigma_s^2}}$ is

$$\widehat{\sigma_s} = \sqrt{\widehat{\sigma_s^2}} \times \sqrt{\frac{\pi}{2}} = |c| \times \sqrt{\pi/2} \qquad (22)$$

Using the results Eqs. (11) and (13) we can easily find that an estimator of standard deviation based on range is

$$\widehat{\sigma_P} = \frac{h-l}{2} \times \sqrt{\frac{\pi}{2}} = \sqrt{\widehat{\sigma_P^2}} \times \sqrt{\frac{\pi \ln 2}{2}} \qquad (23)$$

Similarly, when we want to evaluate the bias introduced by using $\sqrt{\widehat{\sigma^2}}$ instead of $\hat{\sigma}$ for the rest of volatility estimators, we want to find constants c_{GK}, c_M and c_{RS} such that

$$\widehat{\sigma_{GK}} = \sqrt{\widehat{\sigma_{GK}^2}} \times c_{GK} \qquad (24)$$

$$\widehat{\sigma_M} = \sqrt{\widehat{\sigma_M^2}} \times c_M \qquad (25)$$

$$\widehat{\sigma_{RS}} = \sqrt{\widehat{\sigma_{RS}^2}} \times c_{RS} \qquad (26)$$

From simulated high, low and close log-prices h, l, c we estimate volatility according to Eqs. (7), (13), (15), (16), (20) and calculate mean of the square root of these volatility estimates. We find that $c_s = 1.253$, $c_P = 1.043$ (what is in accordance with theoretical values $\sqrt{\pi/2} = 1.253$ and $\sqrt{\pi \ln 2/2} = 1.043$) and $c_{GK} = 1.034$, $c_M = 1.033$ and $c_{RS} = 1.043$. We see that the square root of the simple volatility estimator is a severely biased estimator of standard deviation (bias is 25%), whereas bias in the square root of range-based volatility estimators is rather small (3%–4%).

Even though it seems obvious that $\sqrt{\widehat{\sigma^2}}$ is not an unbiased estimator of σ, it is quite common even among researchers to use $\sqrt{\widehat{\sigma^2}}$ as an estimator of σ.I document this in two examples.

Bali and Weinbaum (2005) empirically compare range-based volatility estimators. The criteria they use are: mean squared error

$$MSE(\sigma_{estimated}) = E\left[(\sigma_{estimated} - \sigma_{true})^2\right] \qquad (27)$$

mean absolute deviation

$$MAD(\sigma_{estimated}) = E[|\sigma_{estimated} - \sigma_{true}|] \qquad (28)$$

and proportional bias

$$Prop.Bias(\sigma_{estimated}) = E[(\sigma_{estimated} - \sigma_{true})/\sigma_{true}] \qquad (29)$$

For daily returns they find:

> "The traditional estimator [Eq. (7) in our paper] is significantly biased in all four data sets. [...] it was found that squared returns do not provide unbiased estimates of the ex post realized volatility. Of particular interest, across the four data sets, extreme-value volatility estimators are almost always significantly less biased than the traditional estimator."

This conclusion sounds surprising only until we realize that in their calculations $\sigma_{estimated} \equiv \sqrt{\widehat{\sigma^2}}$, which, as just shown, is not an unbiased estimator of σ. Actually, it is severely biased for a simple volatility estimator. Generally, if our interest is unbiased estimate of the standard deviation, we should use formulas (22)–(26).

A similar problem is in Bollen and Inder (2002). In testing for the bias in the estimators of σ, they correctly adjust $\sqrt{\widehat{\sigma_s^2}}$ using formula (22), but they do not adjust $\sqrt{\widehat{\sigma_P^2}}$ and $\sqrt{\widehat{\sigma_{GK}^2}}$ by constants c_P and c_{GK}.

3.2. Distributional properties of range-based estimators

Daily volatility estimates are typically further used in volatility models. Ease of the estimation of these models depends not only on the efficiency of the used volatility estimator, but on its distributional properties too (Broto and Ruiz (2004)). When the estimates of relevant volatility measure (whether it is σ^2, σ or $ln\sigma^2$) have approximately normal distribution, the volatility models can be estimated more easily.[4] We study the distributions of $\widehat{\sigma^2}$, $\sqrt{\widehat{\sigma^2}}$ and $ln\widehat{\sigma^2}$, because these are the quantities modelled by volatility models. Most of the GARCH models try to capture time evolution of σ^2, EGARCH and stochastic volatility models are based on time evolution of $ln\sigma^2$ and some GARCH models model time evolution of σ.

Under the assumption of Brownian motion, the distribution of absolute value of return and the distribution of range are known (Karatzas and Shreve (1991), Feller (1951)). Using their result, Alizadeh et al. (2002) derive the distribution of log absolute return and log range. Distribution of $\widehat{\sigma^2}$, $\sqrt{\widehat{\sigma^2}}$ and $ln\widehat{\sigma^2}$ is unknown for the rest of the range-based volatility estimators. Therefore we study these distributions. To do this, we use numerical evaluation of h, l and c data, which are simulated according to the process Eq. (1).[5]

First we study the distribution of $\widehat{\sigma^2}$ for different estimators. These distributions are plotted in Fig. 1. Since all these estimators are unbiased estimators of σ^2, all have the same mean (in our case one). Variance of these estimators is given by their efficiency. From the inspection of Fig. 1, we can observe that the density function of $\widehat{\sigma^2}$ is approximately lognormal for range-based estimators. On the other hand, distribution of squared returns, which is χ^2 distribution with one degree of freedom, is very dispersed and reaches maximum at zero. Therefore, for most of the purposes, distributional properties of range-based estimators are more appropriate for further use than the squared returns. For the range, this was already noted by Alizadeh et al. (2002). However, this is true for all the range-based volatility estimators. The differences in distributions among different range-based estimators are actually rather small.

The distributions of $\sqrt{\widehat{\sigma^2}}$ are plotted in Fig. 2. These distributions have less weight on the tails than the distributions of $\widehat{\sigma^2}$. This is not surprising, since the square root function transforms small values (values smaller than one) into larger values (values closer towards one) and it transforms large values (values larger than one) into smaller values (values closer to one). Again, the distributions of $\sqrt{\widehat{\sigma^2}}$ for range range-based estimators have better properties than the distribution of the absolute returns. To distinguish the difference between different range-based volatility estimators, we calculate the summary statistics and present them in Table 1.

[4] E.g. Gaussian quasi-maximum likelihood estimation, which plays an important role in estimation of stochastic volatility models, depends crucially on the near-normality of log-volatility.

[5] The fact that we do not search for analytical formula is not limiting at all. The analytical form of density function for the simplest range-based volatility estimator, range itself, is so complicated (it is an infinite series) that in the end even skewness and kurtosis must be calculated numerically.

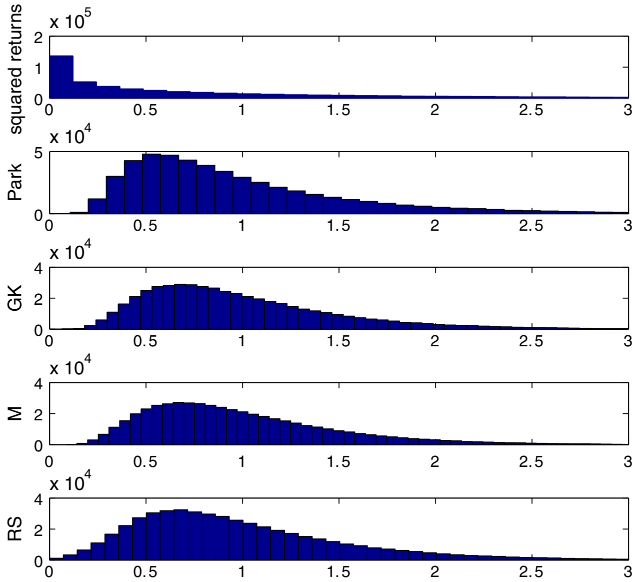

Fig. 1. Distribution of variances estimated as squared returns and from Parkinson, Garman–Klass, Meilijson and Rogers–Satchell formulas.

No matter whether we rank these distributions according to their mean (which should be preferably close to 1) or according to their standard deviations (which should be the smallest possible), ranking is the same as in the previous case: the best is Meilijson volatility estimator, then Garman–Klass, next Roger–Satchell, next Parkinson and the last is the absolute returns.

In many practical applications, the mean squared error (MSE) of an estimator $\hat{\theta}$

$$\mathrm{MSE}\left(\hat{\theta}\right) = E\left[\left(\hat{\theta} - \theta\right)^2\right] \tag{30}$$

is the most important criterion for the evaluation of the estimators, since MSE quantifies the difference between values implied by an estimator and the true values of the quantity being estimated. The

Table 1
The summary statistics for the square root of the volatility estimated as absolute returns and as a square root of the Parkinson, Garman–Klass, Meilijson and Rogers–Satchell formulas.

	Mean	Std	Skewness	Kurtosis
$\lvert r \rvert$	0.80	0.60	1.00	3.87
$\sqrt{\widehat{\sigma_P^2}}$	0.96	0.29	0.97	4.24
$\sqrt{\widehat{\sigma_{GK}^2}}$	0.97	0.24	0.60	3.40
$\sqrt{\widehat{\sigma_M^2}}$	0.97	0.24	0.54	3.28
$\sqrt{\widehat{\sigma_{RS}^2}}$	0.96	0.28	0.46	3.44

MSE is equal to the sum of the variance and the squared bias of the estimator

$$\mathrm{MSE}\left(\hat{\theta}\right) = \mathrm{Var}\left(\hat{\theta}\right) + \left(\mathrm{Bias}\left(\hat{\theta}, \theta\right)\right)^2 \tag{31}$$

and therefore in our case (when estimator with smallest variance has smallest bias) is the ranking according to MSE identical with the ranking according to bias or variance.

In the end, we investigate the distribution of $\ln\widehat{\sigma^2}$ (see Fig. 3). As we can see, the logarithm of the squared returns is highly nonnormally distributed, but the logarithms of the range-based volatility estimators have distributions similar to the normal distribution. To see the difference among various range-based estimators, we again calculate their summary statistics (see Table 2).

Note that the true volatility is normalized to one. Normality of the estimator is desirable for practical reasons and therefore the ideal estimator should have mean and skewness equal to zero, kurtosis close to three and standard deviation as small as possible. We see that from the five studied estimators the Garman–Klass and Meilijson volatility estimators, in addition to being most efficient, have best distributional properties.

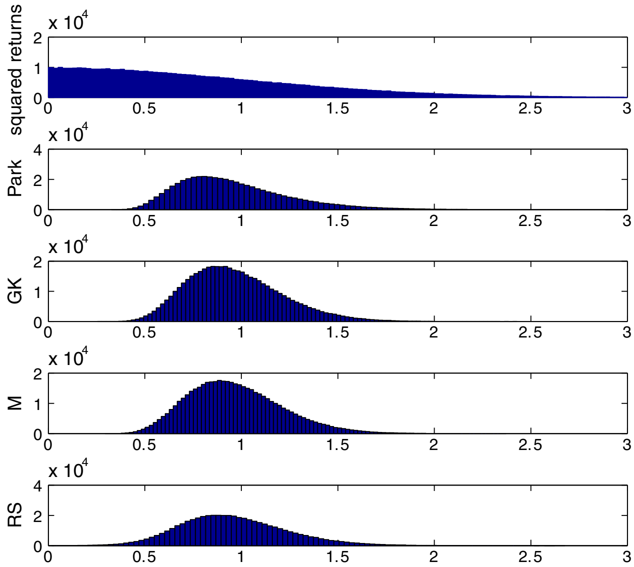

Fig. 2. Distribution of square root of volatility estimated as squared returns and from Parkinson, Garman–Klass, Meilijson and Rogers–Satchell formulas.

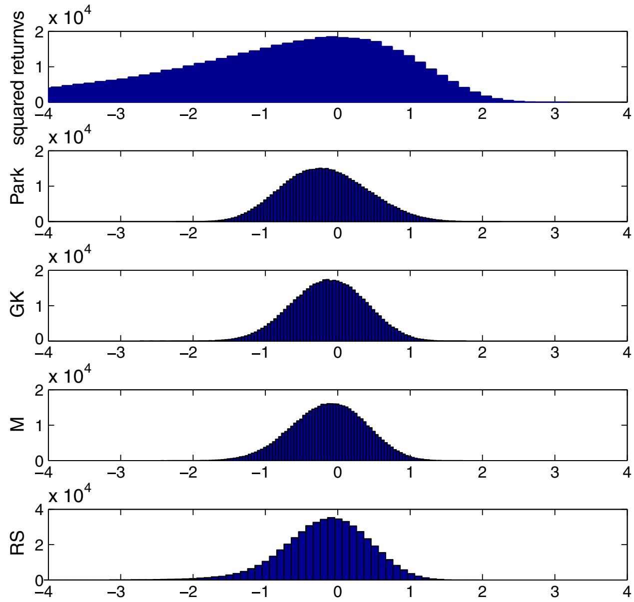

Fig. 3. Distribution of the logarithm of volatility estimated as squared returns and from the Parkinson, Garman–Klass, Meilijson and Rogers–Satchell formulas.

Table 2
Summary statistics for logarithm of volatility estimated as a logarithm of squared returns and as a logarithm of Parkinson, Garman–Klass, Meilijson and Rogers–Satchell volatility estimators.

	Mean	Std	Skewness	Kurtosis
$ln(r^2)$	−1.27	2.22	−1.53	6.98
$ln\left(\widehat{\sigma_P^2}\right)$	−0.17	0.57	0.17	2.77
$ln\left(\widehat{\sigma_{GK}^2}\right)$	−0.13	0.51	−0.09	2.86
$ln\left(\widehat{\sigma_M^2}\right)$	−0.13	0.50	−0.14	2.86
$ln\left(\widehat{\sigma_{RS}^2}\right)$	−0.17	0.61	−0.71	5.41

3.3. Normality of normalized returns

As was empirically shown by Andersen et al. (2000), Andersen et al. (2001), Forsberg and Bollerslev (2002) and Thomakos and Wang (2003) on different data sets, standardized returns (returns divided by their standard deviations) are approximately normally distributed. In other words, daily returns can be written as

$$r_i = \sigma_i z_i \tag{32}$$

where $z_i \sim N(0,1)$. This finding has important practical implications too. If returns (conditional on the true volatility) are indeed Gaussian and heavy tails in their distributions are caused simply by changing volatility, then what we need the most is a thorough understanding of the time evolution of volatility, possibly including the factors which influence it. Even though the volatility models are used primarily to capture time evolution of volatility, we can expect that the better our volatility models, the less heavy-tailed distribution will be needed for modelling of the stock returns. This insight can contribute to improved understanding of volatility models, which is in turn crucial for risk management, derivative pricing, portfolio management etc.

Intuitively, normality of the standardized returns follows from the Central Limit Theorem: since daily returns are just a sum of high-frequency returns, daily returns will be drawn from normal distribution.[6]

Since both this intuition and the empirical evidence of the normality of returns standardized by their standard deviations is convincing, it is appealing to require that one of the properties of a "good" volatility estimator should be that returns standardized by standard deviations obtained from this estimator will be normally distributed (see e.g. Bollen and Inder (2002)). However, this intuition is not correct. As I now show, returns standardized by some estimate of the true volatility do not need to, and generally will not, have the same properties as returns standardized by the true volatility. Therefore we need to understand whether the range-based volatility estimators are suitable for standardization of the returns. There are two problems associated with these volatility estimators: they are noisy and their estimates might be (and typically are) correlated with returns. These two problems might cause returns standardized by the estimated standard deviations not to be normal, even when the returns standardized by their true standard deviations are normally distributed.

3.3.1. Noise in volatility estimators

We want to know the effect of noise in volatility estimates $\widehat{\sigma_i}$ on the distribution of returns normalized by these estimates ($\widehat{z_i} = r_i / \widehat{\sigma_i}$) when true normalized returns $z_i = r_i/\sigma_i$ are normally distributed. Without loss of generality, we set $\sigma_i = 1$ and generate one million observations of r_i, $i \in \{1,..., 1,000,000\}$, all of them are iid N(0,1). Next we generate $\widehat{\sigma_{i,n}}$ in such a way that $\widehat{\sigma}$ is unbiased estimator of σ, i.e. $E\left(\widehat{\sigma_{i,n}}\right) = 1$ and n represents the level of noise in $\widehat{\sigma_{i,n}}$. There is no noise for $n = 0$ and therefore $\widehat{\sigma_{i,0}} = \sigma_i = 1$. To generate $\widehat{\sigma_{i,n}}$ for $i > 0$

[6] given the limited time-dependence and some conditions on existence of moments.

Table 3
Summary statistics for a random variable obtained as ratio of normal random variable with zero mean and variance one and lognormal random variable with constant mean equal to one and variance increasing from 0 to 0.8.

$n = \mathrm{Var}\left(\widehat{\sigma_i}\right)$	$\mathrm{Mean}\left(\widehat{z_{i,n}}\right)$	$\mathrm{Std}\left(\widehat{z_{i,n}}\right)$	$\mathrm{Skewness}\left(\widehat{z_{i,n}}\right)$	$\mathrm{Kurtosis}\left(\widehat{z_{i,n}}\right)$
0.0	0.0001	1.00	0.00	3.00
0.2	0.0003	1.32	0.02	6.22
0.4	0.0013	1.66	−0.01	11.80
0.6	−0.0007	2.03	0.03	19.76
0.8	0.0025	2.43	0.01	34.60

we must decide upon distribution of $\widehat{\sigma_{i,n}}$. Since we know from the previous section that range-based volatility estimates are approximately lognormally distributed, we draw estimates of the standard deviations from lognormal distributions. We set the parameters μ and s^2 of lognormal distribution in such a way that $E\left(\widehat{\sigma_{i,n}}\right) = 1$ and Var $\left(\widehat{\sigma_{i,n}}\right) = n$, particularly $\mu = -\frac{1}{2} ln(1+n)$, $s^2 = ln(1+n)$. For every n, we generate one million observations of $\widehat{\sigma_{i,n}}$. Next we calculate normalized returns $\widehat{z_{i,n}} = r_i / \widehat{\sigma_{i,n}}$. Their summary statistics is in the Table 3.

Obviously, $\widehat{z_{i,0}}$, which is by definition equal to r_i, has zero mean, standard deviation equal to 1, skewness equal to 0 and kurtosis equal to 3. We see that normalization by $\widehat{\sigma}$, a noisy estimate of σ, does not change $E\left(\widehat{z}\right)$ and skewness of $\widehat{z}$. This is natural, because r_i is distributed symmetrically around zero. On the other hand, adding noise increases standard deviation and kurtosis of $\widehat{z}$. When we divide normally distributed random variable r_i by random variable $\widehat{\sigma_i}$, we are effectively adding noise to r_i, making its distribution flatter and more dispersed with more extreme observations. Therefore, standard deviation increases. Since kurtosis is influenced mostly by extreme observations, it increases too.

3.3.2. Bias introduced by normalization of range-based volatility estimators

Previous analysis suggests that the more noisy volatility estimator we use for the normalization of the returns, the higher the kurtosis of the normalized returns will be. Therefore we could expect to find the highest kurtosis when using the Parkinson volatility estimator Eq. (13). As we will see later, this is not the case. Returns and estimated standard deviations were independent in the previous section, but this is not the case when we use range-based estimators.

Let us denote $\sigma_{PARK} \equiv \sqrt{\widehat{\sigma_{PARK}^2}}$, $\sigma_{GK} \equiv \sqrt{\widehat{\sigma_{RS}^2}}$, $\sigma_M \equiv \sqrt{\widehat{\sigma_M^2}}$ and $\sigma_{RS,t} \equiv \sqrt{\widehat{\sigma_{RS}^2}}$. We study the distributions of $\widehat{z}_{PARK,i} \equiv r_i / \sigma_{PARK,i}$, $\widehat{z}_{GK,i} \equiv r_i / \sigma_{GK,i}$, $\widehat{z}_{M,i} \equiv r_i / \sigma_{M,i}$, $\widehat{z}_{RS,t} \equiv r_i / \sigma_{RS,i}$. Histograms for these distributions are shown in Fig. 4 and corresponding summary statistics are in Table 4.

The true mean and skewness of these distributions are zero, because returns are symmetrically distributed around zero, triplets (h, l, c) and $(-l, -h, -c)$ are equally likely and all the studied estimators are symmetric in the sense that they produce the same estimates for the log price following the Brownian motion $B(t)$ and for the log price following Brownian motion $-B(t)$, particularly $\widehat{\sigma_{PARK}}(h,l,c) = \widehat{\sigma_{PARK}}(-l,-h,c)$, $\widehat{\sigma_{GK}}(h,l,c) = \widehat{\sigma_{GK}}(-l,-h,c)$, $\widehat{\sigma_M}(h,l,c) = \widehat{\sigma_M}(-l,-h,c)$ and $\widehat{\sigma_{RD}}(h,l,c) = \widehat{\sigma_{RS}}(-l,-h,c)$.

However, it seems from Table 4 that distribution of $\widehat{z}_{RS,i}$ is skewed. There is another surprising fact about $\widehat{z}_{RS,i}$. It has very heavy tails. The reason for this is that the formula (20) is derived without the assumption of zero drift. Therefore, when stock price performs one-way movement, this is attributed to the drift term and volatility is estimated to be zero. (If movement is mostly in one direction, estimated volatility will be nonzero, but very small). Moreover, this is exactly the situation when stock returns are unusually high. Dividing the largest returns by the smallest estimated standard deviations causes a lot of extreme observations and therefore very heavy tails. Due to these extreme

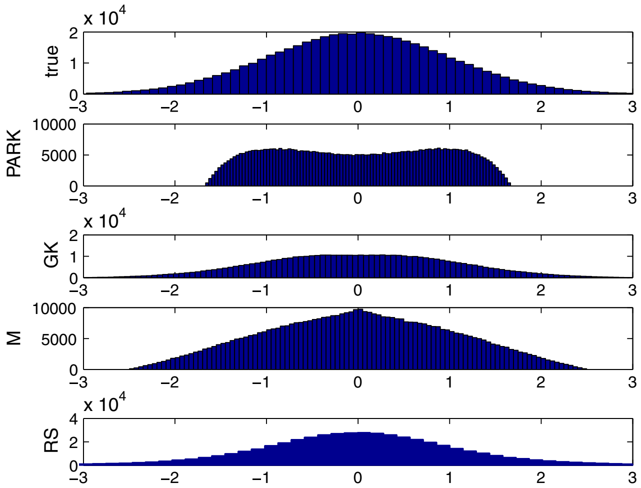

Fig. 4. Distribution of normalized returns. "true" is the distribution of the stock returns normalized by the true standard deviations. This distribution is by assumption N(0,1). PARK, GK, M and RS refer to distributions of the same returns after normalization by volatility estimated using the Parkinson, Garman–Klass, Meilijson and Rogers–Sanchell volatility estimators.

observations the skewness of the simulated sample is different from the skewness of the population, which is zero. This illustrates that the generality (drift independence) of the Rogers and Satchell (1991) volatility estimator actually works against this estimator in cases when the drift is zero.

When we use the Parkinson volatility estimator for the standardization of the stock returns, we get exactly the opposite result. Kurtosis is now much smaller than for the normal distribution. This is in line with empirical finding of Bollen and Inder (2002). However, this result should not be interpreted that this estimator is not working properly. Remember that we got the result of the kurtosis being significantly smaller than 3 under ideal conditions, when the Parkinson estimator works perfectly (in the sense that it works exactly as it is supposed to work). Remember that this estimator is based on the range. Even though the range, which is based on high and low prices, seems to be independent of return, which is based on the open and close prices, the opposite is the case. Always when return is high, range will be relatively high too, because range is always at least as large as absolute value of the return. $|r|/\sigma_{PARK}$ will never be larger than $\sqrt{4\ ln\ 2}$, because

$$\frac{|r|}{\sigma_{PARK}} = \frac{|r|}{\frac{h-l}{\sqrt{4\ ln\ 2}}} = \sqrt{4\ ln\ 2}\frac{|r|}{h-l} \leq \sqrt{4\ ln\ 2} \tag{33}$$

The correlation between $|r|$ and σ_{PARK} is 0.79, what supports our argument. Another problem is that the distribution of $\widehat{z}_{P,i}$ is bimodal.

As we can see from the histogram, distribution of $\widehat{z}_{M,i}$ does not have any tails either. This is because the Meilijson volatility estimator suffers from the same type of problem as the Parkinson volatility estimator, just to a much smaller extent.

The Garman–Klass volatility estimator combines the Parkinson volatility estimator with simple squared return. Even though both, the Parkinson estimator and squared return are highly correlated with size of the return, the overall effect partially cancels out, because these two quantities are subtracted. Correlation between $|r|$ and σ_{GK} is indeed only 0.36. $\widehat{z}_{GK,i}$ has approximately normal distribution, as the effect of noise and the effect of correlation with returns to large extent cancels out.

We conclude this subsection with the appeal that we should be aware of the assumptions behind the formulas we use. As range-based volatility estimators were derived to be as precise volatility estimators as possible, they work well for this purpose. However, there is no reason why all of these estimators should work properly when used for the standardization of the returns. We conclude that from the studied estimators the only estimator appropriate for standardization of returns is the Garman–Klass volatility estimator. We use this estimator later in the empirical part.

Table 4
Summary statistics for returns nomalized by different volatility estimates: $\widehat{z}_{PARK,i} \equiv r_i / \sigma_{PARK,i}$, $\widehat{z}_{GK,i} \equiv r_i / \sigma_{GK,i}$, $\widehat{z}_{M,i} \equiv r_i / \sigma_{M,i}$, $\widehat{z}_{RS,t} \equiv r_i / \sigma_{RS,i}$.

	Mean	Std	Skewness	Kurtosis
$z_{true,\ i}$	0.00	1.00	0.00	3.00
$\widehat{z_{P,i}}$	0.00	0.88	−0.00	1.79
$\widehat{z_{GK,i}}$	0.00	1.01	0.00	2.61
$\widehat{z_{M,i}}$	0.00	1.02	0.00	2.36
$\widehat{z_{RS,i}}$	0.01	1.35	1.62	123.96

3.4. Jump component

So far in this paper, returns and volatilities were related to the trading day, i.e. the period from the open to the close of the market. However, most of the assets are not traded continuously for 24 h a day. Therefore, opening price is not necessarily equal to the closing price from the previous day. We are interested in daily returns

$$r_i = ln(C_i) - ln(C_{i-1}) \tag{34}$$

simply because for the purposes of risk management we need to know the total risk over the whole day, not just the risk of the trading part of the day. If we do not adjust range-based estimators for the presence of opening jumps, they will of course underestimate the true volatility. The Parkinson volatility estimator adjusted for the presence of opening jumps is

$$\widehat{\sigma_P^2} = \frac{(h-l)^2}{4\ ln\ 2} + j^2 \tag{35}$$

where $j_i = ln(O_i) - ln(C_{i-1})$ is the opening jump. The jump-adjusted Garman–Klass volatility estimator is:

$$\widehat{\sigma_{GK}^2} = 0.5(h-l)^2 - (2\ ln\ 2-1)c^2 + j^2 \tag{36}$$

Other estimators should be adjusted in the same way. Unfortunately, including opening jump will increase variance of the estimator when opening jumps are significant part of daily returns.[7] However, this is the only way how to get unbiased estimator without imposing some additional assumptions. If we knew what part of the overall daily volatility opening jumps account for, we could find optimal weights for the jump volatility component and for the volatility within the trading day to minimize the overall variance of the composite estimator. This is done in Hansen and Lunde (2005), who study how to combine opening jump and realized volatility estimated from high frequency data into the most efficient estimator of the whole day volatility. However, the relation of opening jump and the trading day volatility can be obtained only from data. Moreover, there is no obvious reason why the relationship from the past should hold in the future. Simply adding jump component makes range-based estimators unbiased without imposing any additional assumption.[8]

Adjustment for an opening jump is not as obvious as it seems to be and even researchers quite often make mistakes when dealing with this issue. The most common mistake is that the range-based volatility estimators are not adjusted for the presence of opening jumps at all

[7] Jump volatility is estimated with smaller precision than volatility within trading day.

[8] These assumptions could be based on past data, but they would still be just assumptions.

P. Molnár / International Review of Financial Analysis 23 (2012) 20–29 27

(see e.g. Parkinson volatility estimator in Bollen and Inder (2002)). A less common mistake, but with worse consequences is an incorrect adjustment for the opening jumps. E.g. Bollen and Inder (2002) and Fiess and MacDonald (2002) refer to the following formula

$$\sigma^2_{GKwrong,i} = 0.5(\ln H_i - \ln L_i)^2 - (2\ln 2 - 1)(\ln C_i - \ln C_{i-1})^2 \qquad (37)$$

as Garman–Klass formula. This "Garman–Klass volatility estimator" will on average be even smaller than a Garman–Klass estimator not adjusted for jumps. Moreover, it sometimes produces negative estimates for volatility (variance σ^2).

4. Normalized returns –empirics

Andersen et al. (2001) find that although the unconditional daily return distributions are leptokurtic, the daily returns normalized by the realized standard deviations are close to normal. Their conclusion is based on standard deviations obtained these from high frequency data. We study whether (and to what extend) this result is obtainable when standard deviations are estimated from daily data only.

We study stocks which were the components of the Dow Jow Industrial Average on January 1, 2009, namely AA, AXP, BA, BAC, C, CAT, CVX, DD, DIS, GE, GM, HD, HPQ, IBM, INTC, JNJ, JPM, CAG,[9] KO, MCD, MMM, MRK, SFT, PFE, PG, T, UTX, VZ and WMT. We use daily open, high, low and close prices. The data covers years 1992 to 2008. Stock prices are adjusted for stock splits and similar events. We have 4171 daily observations for every stock. These data were obtained from the CRSP database. We study DJI components to make our results as highly comparable as possible with the results of Andersen et al. (2001).

For brevity, we study only two estimators: the Garman–Klass estimator Eq. (15) and the Parkinson estimator Eq. (13). We use the Garman–Klass volatility estimator because our previous analysis shows that it is the most appropriate one. We use the Parkinson volatility estimator to demonstrate that even though this estimator is the most commonly used range-based estimator, it should not be used for normalization of returns. Moreover, we study the effect of including or excluding a jump component into range-based volatility estimators.

First of all, we need to distinguish the daily returns and the trading day returns. By the daily returns we mean close-to-close returns, calculated according to formula (34). By the trading day returns we mean returns during the trading hours, i.e. open-to-close returns, calculated according to formula (2). We estimate volatilities accordingly: volatility of the trading day returns from Eqs. (13) to (15) and the volatility of the daily returns using Eqs. (35) and (36). Next we calculate standardized returns. We calculate standardized returns in three different ways: trading day returns standardized by trading day standard deviations (square root of trading day volatility), daily returns standardized by daily standard deviation and daily returns standardized by trading day standard deviation. Why do we investigate daily returns standardized by trading day standard deviations too? Theoretically, this does not make much sense because the return and the standard deviations are related to different time intervals. However, it is still quite common (see e.g. Andersen et al. (2001)), because people are typically interested in daily returns, but the daily volatility cannot be estimated as precisely as trading day volatility. The volatility of the trading part of the day can be estimated very precisely from the high frequency data, whereas estimation of the daily volatility is always less precise because of the necessity of including the opening jump component. Therefore, trading day volatility is commonly used as a proxy for daily volatility. This approximation is satisfactory as long as the opening jump is small in comparison to trading day volatility, which is typically the case.

Now we calculate summary statistics for the different standardized returns as well as returns themselves. Results for the standard deviations are presented in Table 5 and results for the kurtosis are presented in Table 6. We do not put similar tables for mean and kurtosis into this paper, because these results are less interesting and can be summarized in one sentence: Mean returns are always very close to zero, independent of which standardization we used. Skewness is always very close to zero too.

The results for standard deviations and kurtosis are generally in line with the predictions from our simulations too. First let us discuss the standard deviations of the standardized returns. As Table 5 documents, normalization by standard deviations obtained from the Parkinson volatility estimator results in standard deviation smaller than one, approximately around 0.9 whereas normalization by standard deviation obtained from the Garman–Klass volatility estimator results in standard deviations larger than one, around 1.05. Normalization by standard deviations estimated from GARCH model is approximately 1.1. This is expected as well, because division by a noisy random variable increases the standard deviation.

Results for the kurtosis of standardized returns (see Table 6) are in line with the predictions from our simulations too. Return distributions have heavy tails (kurtosis significantly larger than 3). Second, the daily returns normalized by the standard deviations calculated from Garman–Klass formula are close to normal (kurtosis is close to 3). Third, the daily returns normalized by the standard deviations calculated from Parkinson formula have no tails (kurtosis is significantly smaller

Table 5
Standard deviations of the stock returns. r_{td} is an open-to-close return, r_d is a close-to-close return. $\hat{\sigma}_{GK,td}$ ($\hat{\sigma}_{P,td}$) is square root of Garman–Klass (Parkinson) volatility estimate without opening jump component. $\hat{\sigma}_{GK,d}$ ($\hat{\sigma}_{P,d}$) is square root of Garman–Klass (Parkinson) volatility estimate including opening jump component. $\hat{\sigma}_{garch}$ is standard deviation estimated from GARCH(1,1) model based on daily returns.

	Trading day returns			Daily returns					
	r_{td}	$\frac{r_{td}}{\hat{\sigma}_{GK,td}}$	$\frac{r_{td}}{\hat{\sigma}_{P,td}}$	r_d	$\frac{r_d}{\hat{\sigma}_{GK,d}}$	$\frac{r_d}{\hat{\sigma}_{P,d}}$	$\frac{r_d}{\hat{\sigma}_{GK,td}}$	$\frac{r_d}{\hat{\sigma}_{P,td}}$	$\frac{r_d}{\hat{\sigma}_{garch}}$
AA	0.02	1.14	0.94	0.02	1.11	0.96	1.00	1.28	1.12
AXP	0.02	1.11	0.92	0.02	1.07	0.94	1.00	1.26	1.11
BA	0.02	1.04	0.89	0.02	1.02	0.92	1.00	1.20	1.10
BAC	0.02	1.12	0.93	0.02	1.08	0.94	1.00	1.26	1.12
C	0.02	1.11	0.91	0.03	1.05	0.92	1.01	1.26	1.12
CAT	0.02	1.10	0.92	0.02	1.08	0.95	1.00	1.28	1.13
CVX	0.01	1.11	0.92	0.02	1.08	0.95	1.00	1.25	1.09
DD	0.02	1.07	0.90	0.02	1.02	0.91	1.00	1.18	1.06
DIS	0.02	1.03	0.88	0.02	0.99	0.90	1.00	1.18	1.09
GE	0.02	1.07	0.91	0.02	1.03	0.93	1.00	1.20	1.09
GM	0.02	1.10	0.92	0.03	1.08	0.95	1.00	1.27	1.13
HD	0.02	1.06	0.90	0.02	1.02	0.92	1.00	1.20	1.10
HPQ	0.02	1.08	0.91	0.03	1.04	0.92	1.00	1.23	1.11
IBM	0.02	1.07	0.91	0.02	1.04	0.93	1.00	1.25	1.13
INTC	0.02	1.08	0.92	0.03	1.06	0.95	1.00	1.31	1.19
JNJ	0.01	1.06	0.89	0.02	1.00	0.90	1.00	1.17	1.06
JPM	0.02	1.06	0.90	0.02	1.03	0.92	1.00	1.22	1.10
CAG	0.01	1.09	0.89	0.02	0.98	0.87	1.00	1.15	1.01
KO	0.01	1.03	0.88	0.02	0.99	0.89	1.00	1.15	1.04
MCD	0.02	1.04	0.89	0.02	0.99	0.89	1.00	1.15	1.05
MMM	0.01	1.05	0.89	0.02	1.02	0.90	1.00	1.16	1.04
MRK	0.02	1.05	0.89	0.02	1.01	0.91	1.00	1.20	1.09
MSFT	0.02	1.04	0.90	0.02	1.03	0.93	1.00	1.24	1.14
PFE	0.02	1.08	0.91	0.02	1.04	0.92	1.00	1.22	1.10
PG	0.01	1.07	0.90	0.02	1.01	0.90	1.00	1.17	1.05
T	0.02	1.09	0.91	0.02	1.05	0.92	1.00	1.20	1.06
UTX	0.02	1.08	0.91	0.02	1.05	0.93	1.00	1.22	1.09
VZ	0.02	1.08	0.91	0.02	1.04	0.92	1.00	1.21	1.08
WMT	0.02	1.04	0.88	0.02	1.01	0.90	1.00	1.20	1.08
XOM	0.01	1.08	0.91	0.02	1.06	0.94	1.00	1.22	1.08
Mean	0.02	1.07	0.90	0.02	1.04	0.92	1.00	1.22	1.09

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

Table 6
Kurtosis of the stock returns. r_{td} is an open-to-close return, r_d is a close-to-close return. $\hat{\sigma}_{GK,td}$ ($\hat{\sigma}_{P,td}$) is square root of Garman–Klass (Parkinson) volatility estimate without opening jump component. $\hat{\sigma}_{GK,d}$ ($\hat{\sigma}_{P,d}$) is square root of Garman–Klass (Parkinson) volatility estimate including opening jump component. $\hat{\sigma}_{garch}$ is standard deviation estimated from GARCH(1,1) model based on daily returns.

	Trading day returns			Daily returns					
	r_{td}	$\frac{r_{td}}{\hat{\sigma}_{GK,td}}$	$\frac{r_{td}}{\hat{\sigma}_{P,td}}$	r_d	$\frac{r_d}{\hat{\sigma}_{GK,d}}$	$\frac{r_d}{\hat{\sigma}_{P,d}}$	$\frac{r_d}{\hat{\sigma}_{GK,td}}$	$\frac{r_d}{\hat{\sigma}_{P,td}}$	$\frac{r_d}{\hat{\sigma}_{garch}}$
AA	9.63	2.84	1.76	11.63	2.73	1.87	3.48	2.56	4.62
AXP	8.46	3.03	1.81	9.62	2.84	1.91	4.10	2.70	5.00
BA	6.42	2.99	1.81	10.76	2.75	1.91	3.12	2.62	6.82
BAC	19.47	2.87	1.78	26.81	2.78	1.91	3.50	2.85	8.69
C	34.05	3.12	1.82	38.79	2.95	1.96	3.62	2.70	6.70
CAT	5.71	2.93	1.80	7.31	2.78	1.90	3.88	2.78	6.97
CVX	11.28	2.99	1.80	13.44	2.80	1.90	3.89	2.43	3.77
DD	7.07	2.98	1.81	7.53	2.84	1.95	3.54	2.63	5.23
DIS	6.75	2.93	1.81	11.04	2.76	1.94	4.23	3.78	9.88
GE	10.29	2.80	1.77	10.07	2.71	1.93	3.20	2.68	4.95
GM	43.27	2.93	1.82	26.30	2.74	1.89	3.73	2.80	7.41
HD	6.43	2.93	1.80	19.21	2.70	1.90	3.23	2.73	10.84
HPQ	7.63	2.92	1.80	9.29	2.77	1.93	3.30	2.81	9.73
IBM	6.87	2.82	1.78	9.44	2.75	1.91	3.74	3.90	8.17
INTC	6.45	2.62	1.76	8.59	2.57	1.87	3.89	4.56	6.86
JNJ	5.70	3.02	1.83	10.56	2.88	1.97	3.17	2.63	4.98
JPM	14.60	3.00	1.83	12.05	2.80	1.96	3.46	2.79	4.89
CAG	8.64	3.54	1.93	16.43	3.37	2.08	4.42	2.78	10.33
KO	7.81	3.12	1.86	8.56	2.94	1.98	3.41	2.57	6.66
MCD	8.56	3.05	1.84	7.48	2.84	1.95	3.14	2.48	5.26
MMM	6.86	3.08	1.84	7.60	3.01	1.99	3.59	2.68	8.72
MRK	6.64	2.96	1.82	24.22	2.78	1.92	4.35	4.83	42.92
MSFT	5.22	2.63	1.78	8.61	2.55	1.91	5.81	9.43	9.24
PFE	5.36	2.83	1.78	6.17	2.74	1.90	3.40	2.86	6.57
PG	8.22	2.96	1.83	75.61	2.89	1.97	3.46	3.11	17.85
T	6.23	3.00	1.81	7.40	2.90	1.96	3.32	2.42	4.29
UTX	9.11	3.01	1.79	32.55	2.81	1.91	3.83	3.01	28.66
VZ	6.89	2.98	1.79	7.80	2.88	1.93	4.67	2.61	4.52
WMT	6.59	3.19	1.86	5.98	2.99	1.97	3.72	3.19	4.41
XOM	11.30	2.91	1.77	12.62	2.81	1.91	3.21	2.44	4.11
Mean	10.25	2.97	1.81	15.45	2.82	1.93	3.71	3.15	8.97

than 3). Fourth, normalization of daily returns by standard deviation estimated for trading day only, will cause upward bias in kurtosis. This is a consequence of the standardization by an incorrect standard deviation – sometimes (particularly in a situation when the opening jump is large), returns are divided by too small standard deviation, which will cause too many large observations for normalized returns.

The last column of Table 6 reports kurtosis of returns normalized by standard deviations estimated from GARCH (1,1) model with mean return fixed to zero. As we can see, these normalized returns are not Gaussian, they have fat tails. This is consistent with the fact that GARCH models with fat-tailed conditional distribution of returns fit data better than GARCH models with conditionally normally distributed returns. However, as is clear from this paper, this is the case simply because GARCH models always condition return distribution on the estimated volatility, which is only a noisy proxy of the true volatility. Therefore, even when distribution of returns conditional on the true volatility is Gaussian, distribution of returns conditional on estimated volatility will have heavy tails. This result has an important implication for volatility modelling: the more precisely we can estimate the volatility, the closer will be the conditional distribution of returns to the normal distribution.

5. Conclusion

Range-based volatility estimators provide significant increase in accuracy compared to simple squared returns. Even though efficiency of these estimators is known, there is some confusion about other properties of these estimators. We study these properties. Our main focus is the properties of returns standardized by their standard deviations.

First, we correct some mistakes in existing literature. Second, we study different properties of range-based volatility estimators and find that for most purposes, the best volatility estimators is the Garman–Klass volatility estimator. The Meilijson volatility estimator improves its efficiency slightly, but it is based on a significantly more complicated formula. However, performance of all the range-based volatility estimators is similar in most cases except for the case when we want to use them for standardization of the returns.

Returns standardized by their standard deviations are known to be normally distributed. This fact is important for the volatility modelling. This result was possible to obtain only when the standard deviations were estimated from the high frequency data. When the standard deviations were obtained from volatility models based on daily data, returns standardized by these standard deviations are not Gaussian anymore, they have heavy tails. Using simulations we show that even when returns themselves are normally distributed, returns standardized by (imprecisely) estimated volatility are not normally distributed; their distribution has heavy tails. In other words: the fact that standard volatility models show that even conditional distribution of returns has heavy tails does not mean that returns are not normally distributed. It means that these models cannot estimate volatility precisely enough and the noise in the volatility estimates causes the heavy tails.

It is not obvious whether range-based volatility estimators can be used for the standardization of the returns. Using simulations we find that for the purpose of returns standardization there are large differences between these estimators and we find that the Garman–Klass volatility estimator is the only one appropriate for this purpose. Putting all the results together, we rate the Garman–Klass volatility estimator as the best volatility estimator based on daily (open, high, low and close) data. We test this estimator empirically and we find that we can indeed obtain basically the same results from daily data as Andersen et al. (2001) obtained from high-frequency (transaction) data. This is important, because the high-frequency data are very often not available or available only for a shorter time period and their processing is complicated. Since returns scaled by standard deviations estimated from GARCH type of models (based on daily returns) are not Gaussian (they have fat tails), our results show that the GARCH type of models cannot capture the volatility precisely enough. Therefore, in the absence of high-frequency data, further development of volatility models based on open, high, low and close prices is recommended.

Acknowledgements

I would like to thank to Jonas Andersson, Milan Bašta, Ove Rein Hetland, Lukáš Laffers, Michal Zdeněk and anonymous referees for their helpful comments.

References

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, *57*, 1047–1091.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*, 43–76.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2000). Exchange rate returns standardized by realized volatility are nearly Gaussian. *Multinational Finance Journal*, *4*, 159–179.
Bali, T. G., & Weinbaum, D. (2005). A comparative study of alternative extreme-value volatility estimators. *Journal of Futures Markets*, *25*, 873–892.
Bollen, B., & Inder, B. (2002). Estimating daily volatility in financial markets. *Journal of Empirical Finance*, *9*, 551–562.
Brandt, M. W., & Diebold, F. X. (2006). A no-arbitrage approach to range based estimation of return covariances and correlations. *Journal of Business*, *79*, 61–74.
Brandt, M. W., & Jones, Ch. S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business and Economic Statistics*, *24*(4), 470–486.
Broto, C., & Ruiz, E. (2004). Estimation methods for stochastic volatility models: a survey. *Journal of Economic Surveys*, *18*, 613–649.

Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit and Banking*, *37*(3), 561–582.

Chou, R. Y. (2006). Modeling the asymmetry of stock movements using price ranges. *Advances in Econometrics*, *20*, 231–258.

Chou, R. Y., Chou, H., & Liu, N. (2010). *Range volatility models and their applications in finance.* Handbook of Quantitative Finance and Risk Management. Springer Chapter 83.

Chou, R. Y., & Liu, N. (2010). The economic value of volatility timing using a range-based volatility model. *Journal of Economic Dynamics & Control*, *34*, 2288–2301.

Dacorogna, M. M., Gencay, R., Müller, U. A., Olsen, R. B., & Pictet, O. V. (2001). *An introduction to high-frequency finance.* San Diego: Academic Press.

Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *The Annals of Mathematical Statistics*, *22*, 427–432.

Fiess, N. M., & MacDonald, R. (2002). Towards the fundamentals of technical analysis: analysing the information content of High, Low and Close proces. *Economic Modelling*, *19*, 354–374.

Forsberg, L., & Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model. *Journal of Applied Econometrics*, *17*, 535–548.

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *The Journal of Business*, *53*(1), 67–78.

Hansen, P. R., & Lunde, A. (2005). A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics*, *3*, 525–554.

Karatzas, I., & Shreve, S. E. (1991). *Brownian motion and stochastic calculus.* New York: Springer-Verlag.

Kunitomo, N. (1992). Improving the Parkinson method of estimating security price volatilities. *Journal of Business*, *65*, 295–302.

Meilijson, I. (2009). The Garman–Klass volatility estimator revisited. working paper available at:. http://arxiv.org/PS_cache/arxiv/pdf/0807/0807.3492v2.pdf

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, *53*, 61–65.

Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability*, *1*, 504–512.

Thomakos, D. D., & Wang, T. (2003). Realized volatility in the futures markets. *Journal of Empirical Finance*, *10*, 321–353.

Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, *73*, 477–491.

Chapter 3

High-low range in GARCH models of stock return volatility

APPLIED ECONOMICS, 2016
VOL. 48, NO. 51, 4977–4991
http://dx.doi.org/10.1080/00036846.2016.1170929

Routledge
Taylor & Francis Group

High-low range in GARCH models of stock return volatility

Peter Molnár

Department of Industrial economics and Technology Management, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

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.

KEYWORDS
Volatility; high; low; range; GARCH

JEL CLASSIFICATION
C22; G17

I. 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 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. In addition, 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, Brandt, and Diebold (2002), Chou (2005) and Brandt and Jones (2006) and more recently Miralles-Marcelo, Miralles-Quirós, and Miralles-Quirós (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, Chou, and Liu (2015). Range-based volatility estimators are compared in Molnár (2012) and some of the recent applications of range are Awartani and Maghyereh (2013), Lucey, Larkin, and O'Connor (2014) and Lyócsa (2014).

CONTACT Peter Molnár peto.molnar@gmail.com Department of Industrial economics and Technology Management, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway

[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 article, 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'.

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 article 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 article 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 article is organized in the following way: Section II provides a basic introduction to volatility modelling and an overview of existing range-based volatility estimators. Section III describes the data, methodology and results. Finally, Section IV concludes.

II. Theoretical background

GARCH models

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

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

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:

$$r_t \sim N(0, \sigma_t^2). \tag{2}$$

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 article. 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) autoregressive conditional Heteroscedasticity (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 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 1, but smaller than 1. 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 proxy instead of squared

[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).

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)

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

where $\widehat{\sigma^2_{\text{proxy},t-1}}$ 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, Huang, and Shek (2012). 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, Chou, and Liu 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 article, we use Parkinson's volatility estimator ($\sigma^2_{\text{proxy}} = \sigma^2_P$). 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 article, we therefore study the following model:

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

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

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

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 article 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.*

In addition, 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 three 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. We compare the RGARCH to the following models:

The GJR-GARCH of Glosten, Jagannathan, and Runkle (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):

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

The standard deviation GARCH of Taylor (1986), denoted in this article 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 Equation (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):

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

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

Estimation

All the GARCH models, including the models (5), (12)–(15) in our article, 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 the estimation of this model. This in turns means 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 $\widehat{\sigma_{P,t-1}^2}$.

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 generalized error 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 the contrary, estimation of the CARR and the REGARCH models is based on the distribution of the range. Denote range as

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

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

$$\ln(D_t) \sim N(0.43 + \ln(\sigma_t), 0.29^2), \quad (19)$$

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

$$D_t = \lambda_t \varepsilon_t, \quad (20)$$

where λ_t is the conditional mean of the range [varying according to equation similar to Equation (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.

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 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

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

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.

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.

[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 2–101 to forecast volatility for day 102 and so on.

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; Patton 2011).

Next, we want to know whether the MSE for two different models are statistically 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 $\widehat{\sigma^2_{1,t}}$ and $\widehat{\sigma^2_{2,t}}$ and the true volatility as σ^2_{true} In our case $\widehat{\sigma^2_{1,t}} = \widehat{\sigma^2_{\text{RGARCH},t}}$ and $N(0, 0.1^2)$ is the competing model; in the majority of this article, it is the GARCH(1,1) model. First, we construct the vector of differences in squared errors

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

Next, we construct the Diebold–Mariano test statistic

$$\text{DM} = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}, \tag{23}$$

where $\bar{d}$ denotes the sample mean of d_t, and $\hat{V}(\bar{d})$ 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]

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

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

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 article are estimated on the open-to-close returns. We have done the same estimations for close-to-close returns. Main results remain the same (RGARCH model out-performs GARCH model).[5] These results are available in previous versions of the article or upon request from the author.

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

[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,

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

As a result, the estimated coefficient α will be larger to balance this bias in $\widehat{\sigma^2_P}$. 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:

$$\sigma^2_t = \omega + \alpha \frac{r^2_{t-1}}{2} + \beta\sigma^2_{t-1},$$

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.

III. 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.

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 15 June 1992 to 31 December 2010.

In-sample analysis

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

For every single stock, the coefficients in the RGARCH(1,1) have changed in exactly the same way we expected. In addition, 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 [Equation (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.

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.

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 article, 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

[6]Components of stock indices change over time. These stocks were DJI components on 1 January 2009.
[7]Since historical data for KFT (component of DJI) are not available for the complete period, we use its competitor CAG instead.

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 AIC of the respective equations.

	GARCH(1,1)				RGARCH(1,1)			
Ticker	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

statistically significant, the RGARCH model out performs 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

[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 $\left((0.1^2 - 1^2)^2 < (10 - 3^2)^2\right)$, even though the probability of the return $r = 1$ being drawn from the distribution $N(0, 0.1^2)$ is more than 10^{4o}-times smaller than probability of the return $r = \sqrt{10}$ being drawn from the distribution $N(0, 3^2)$.

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

	Combined GARCH(1,1)							
Ticker	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

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:

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

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.

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 3 January 1993–27 April 2009 for open, high and low prices and the period 3 January 1996–27 April 2009 for the realized variance. Due to small differences in trading days in different markets, the number of observations varies accordingly.

[9]Heber et al. (2009).

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

	GARCH(1,1)				RGARCH(1,1)			
Ticker	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*

Numbers in this table are 1000 × RMSE of the 1-day-ahead rolling window forecast reported for different window sizes w. Asterisk * and** indicate when the difference is significant at the 5% and1% level.

For the in-sample analysis we use the data ranging from 3 January 1993 to 27 April 2009. For the out of sample comparison, we use the volatilities forecasted for the period 3 January 1996–27 April 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.

In-sample analysis

Table 6 presents estimated coefficients for the GARCH model (5) and the RGARCH model (10) together with the values of AIC. The results are again in line with those in Table 1. GARCH 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.

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 given 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).

Table 4. Comparison of the forecasting performance of the RGARCH(1,1) model $\sigma_t^2 = \omega + \alpha\widehat{\sigma^2_{P,t-1}} + \beta\sigma_{t-1}^2$ and several different GARCH models.

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**

Numbers in this table are 1000 × RMSE of the 1-day-ahead rolling window forecast with forecasting window equal to 400.

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, Brandt, and Diebold (2002). They simulate the data in the following way. First, we simulate the volatility process

$$\ln\sigma_t = \ln\bar{\sigma} + \rho_H(\ln\sigma_{t-1} - \ln\bar{\sigma}) + \mu_1\varepsilon_{t-1} \qquad (27)$$

with parameters $\ln(\bar{\sigma}) = -2.5$, $\rho_H = 0.985$ and $\mu_1 = 0.75/\sqrt{257} = 0.048$ and residuals ε drawn from standard normal distribution. For every day $t = 1, 2, \ldots, 100000$, 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, Brandt, and Diebold (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 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 Equation (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.

[10]We use 100,000 discrete steps for the approximation of the continuous Brownian motion.

Table 5. Comparison of forecasting performance GARCH(1,1) model $\sigma_t^2 = \omega + a r_{t-1}^2 + \beta \sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + a \widehat{\sigma_{P,t-1}^2} + \beta \sigma_{t-1}^2$.

	GARCH(1,1)				RGARCH(1,1)			
Ticker	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	10,377	10,166	9881	9595	10,502	10,242	9969	9688
BA	10,434	10,225	9809	9660	10,500	10,258	9993	9708
BAC	10,687	10,451	10,154	9875	10,783	10,527	10,236	9949
CAT	10,105	9916	9631	9342	10,202	9950	9675	9385
CSCO	9309	9080	8825	8528	9478	9237	8955	8646
CVX	11,371	11,017	10,853	10,576	11,440	11,145	10,882	10,599
DD	10,853	10,593	10,321	10,050	10,916	10,641	10,377	10,095
DIS	10,535	10,298	10,024	9747	10,681	10,411	10,142	9859
GE	11,086	10,860	10,567	10,258	11,176	10,902	10,617	10,325
HD	10,266	10,024	9729	9479	10,372	10,084	9809	9542
HPQ	9587	9255	9076	8792	9715	9415	9174	8869
IBM	10,813	10,575	10,247	9972	10,986	10,716	10,378	10,130
INTC	9420	9208	8937	8665	9484	9278	9001	8735
JNJ	12,013	11,776	11,492	11,230	12,063	11,126	11,522	11,264
JPM	10,158	10,014	9730	9464	10,345	10,113	9830	9554
CAG	11,421	11,192	10,939	10,681	11,563	11,301	10,994	10,722
KO	11,682	11,454	11,155	10,924	11,782	11,517	11,250	10,980
MCD	11,058	10,846	10,564	10,288	11,129	10,871	10,596	10,320
MMM	11,238	11,105	10,819	10,542	11,377	11,153	10,878	10,599
MRK	10,131	9775	9632	9294	10,348	10,120	9813	9570
MSFT	10,234	10,038	9724	9478	10,396	10,171	9867	9611
PFE	10,741	10,499	10,222	9959	10,827	10,564	10,269	10,004
PG	11,512	11,236	10,962	10,709	11,571	11,369	11,081	10,777
T	10,948	10,704	10,459	10,172	11,002	10,744	10,473	10,206
TRV	10,801	10,614	10,312	10,069	10,899	10,678	10,395	10,111
UTX	11,013	10,790	10,477	10,179	11,054	10,840	10,556	10,269
VZ	11,132	10,892	10,605	10,329	11,198	10,930	10,645	10,361
WMT	11,004	10,778	10,510	10,109	11,130	10,860	10,558	10,276
XOM	11,464	11,223	10,947	10,657	11,567	11,294	11,014	10,729

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 1-day-ahead rolling window volatility forecast reported for different window sizes w.

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

	GARCH(1,1)				RGARCH(1,1)			
Index	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

In-sample analysis

Table 9 presents estimated coefficients for the standard GARCH model (5) and the RGARCH model (10) together with the values of 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.

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

	Combined GARCH(1,1)							
Index	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 $\sigma_t^2 = \omega + ar_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + a\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$. 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.

		GARCH(1,1)				RGARCH(1,1)			
Index	Bench	*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 $\sigma_t^2 = \omega + ar_{t-1}^2 + \beta\sigma_{t-1}^2$ and the RGARCH(1,1) model $\sigma_t^2 = \omega + a\widehat{\sigma_{P,t-1}^2} + \beta\sigma_{t-1}^2$, reported together with the values of AIC of the respective equations for the simulated data.

	GARCH(1,1)				RGARCH(1,1)			
	ω	a	β	AIC	ω	a	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

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).

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. In addition, simulation allows us to have much larger data sample (100,000 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 Tables 3 and 8. Tables 3 and 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 (Table 11) exactly how

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

	GARCH(1,1)				RGARCH(1,1)				
	w = 300	w = 400	w = 500	w = 600	w = 300	w = 400	w = 500	w = 600	σ^2_{true}
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^2_{true} 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

Numbers in this table are 1000 × RMSE of the 1-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).

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

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

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 can 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.

IV. 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 out performs 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 article should therefore be of significant interest both to academics and practitioners alike.

Acknowledgements

I would like to thank Jonas Andersson, Milan Bašta, Ray Chou, Stein-Erik Fleten, Ove Hetland, Einar Cathrinus

Kjenstad, Svein Olav Krakstad, Lukáš Lafférs, Gudbrand Lien and Kjell Nyborg for helpful comments.

Disclosure statement

No potential conflict of interest was reported by the author.

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. doi:10.1111/1540-6261.00454.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold. 2006. "Volatility and Correlation Forecasting." In *Handbook of Economic Forecasting*, edited by Elliott, G., C. W. J. Granger, and A. Timmermann, 777–878. Vol. 1, Chap. 15. Amsterdam: Elsevier.

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. doi:10.1016/j.eneco.2012.11.024.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31: 307–327. doi:10.1016/0304-4076(86)90063-1.

Brandt, M. W., and C. S. Jones. 2006. "Volatility Forecasting With Range-Based EGARCH Models." *Journal of Business & Economic Statistics* 24: 470–486. doi:10.1198/073500106000000206.

Chou, R. Y.-T. 2005. "Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model." *Journal of Money, Credit, and Banking* 37: 561–582. doi:10.1353/mcb.2005.0027.

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*, edited by Lee, C.-F. and J. Lee. Chap. 83. New York, NY: Springer.

Chou, R. Y., H. Chou, and N. Liu. 2015. "Range Volatility: A Review of Models and Empirical Studies." In *Hand-book of Financial Econometrics and Statistics*, 2029–2050. New York, NY: Springer.

Dacorogna, M. M., R. Gencay, U. Muller, R. B. Olsen, and O. V. Olsen. 2001. *An Introduction to High Frequency Finance*. New York: Academic Press.

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. doi:10.2307/1912773.

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. doi:10.1086/jb.1980.53.issue-1.

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. doi:10.1111/j.1540-6261.1993.tb05128.x.

Hansen, P. R., Z. Huang, and H. H. Shek. 2012. "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility." *Journal of Applied Econometrics* 27 (6): 877–906.

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. doi:10.1002/(ISSN)1099-1255.

Hansen, P. R., and A. Lunde. 2006. "Consistent Ranking of Volatility Models." *Journal of Econometrics* 131: 97–121. doi:10.1016/j.jeconom.2005.01.005.

Heber, G., A. Lunde, N. Shephard, and K. K. Sheppard. 2009. *Oxford-Man Institute's Realized Library*. Oxford-Man Institute, University of Oxford. http://realized.oxford-man.ox.ac.uk/

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. doi:10.1080/13504851.2014.896974.

Lyócsa, S. 2014. "Growth-Returns Nexus: Evidence from Three Central and Eastern European Countries." *Economic Modelling* 42: 343–355. doi:10.1016/j.econmod.2014.07.023.

Miralles-Marcelo, J. L., J. L. Miralles-Quirós, and M. d. M. Miralles-Quirós. 2013. "Improving the CARR Model Using Extreme Range Estimators." *Applied Financial Economics* 23: 1635–1647. doi:10.1080/09603107.2013.844325.

Molnár, P. 2012. "Properties of Range-Based Volatility Estimators." *International Review of Financial Analysis* 23: 20–29. doi:10.1016/j.irfa.2011.06.012.

Nelson, D. B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59: 347–370. doi:10.2307/2938260.

Parkinson, M. 1980. "The Extreme Value Method for Estimating the Variance of the Rate of Return." *The Journal of Business* 53: 61–65. doi:10.1086/jb.1980.53.issue-1.

Patton, A. J. 2011. "Volatility Forecast Comparison Using Imperfect Volatility Proxies." *Journal of Econometrics* 160: 246–256. doi:10.1016/j.jeconom.2010.03.034.

Poon,, S.-H., and C. W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41: 478–539. doi:10.1257/.41.2.478.

Shephard,, N., and K. Sheppard. 2010. "Realising the Future: Forecasting with High-Frequency-Based Volatility (HEAVY) Models." *Journal of Applied Econometrics* 25: 197–231. doi:10.1002/jae.1158.

Taylor, S. J. 1986. *Modelling Financial Time Series*. 2nd ed. New York, NY: John Wiley and Sons.

Chapter 4

Range-based DCC models for covariance and value-at-risk forecasting

Journal of Empirical Finance 54 (2019) 58–76

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Range-based DCC models for covariance and value-at-risk forecasting☆

Piotr Fiszeder[a], Marcin Fałdziński[a], Peter Molnár[a,b,c,*]

[a] Faculty of Economics Sciences and Management, Nicolaus Copernicus University in Torun, Torun, Poland
[b] UiS Business School, University of Stavanger, Stavanger, Norway
[c] Faculty of Finance and Accounting, University of Economics, Prague, Czech Republic

ARTICLE INFO

Keywords:
Volatility
Dynamic conditional correlation
High-low range
Covariance forecasting
Value-at-risk

ABSTRACT

The dynamic conditional correlation (DCC) model by Engle (2002) is one of the most popular multivariate volatility models. This model is based solely on closing prices. It has been documented in the literature that the high and low prices of a given day can be used to obtain an efficient volatility estimation. We therefore suggest a model that incorporates high and low prices into the DCC framework. We conduct an empirical evaluation of this model on three datasets: currencies, stocks, and commodity exchange traded funds. Regardless of whether we consider in-sample fit, covariance forecasts or value-at-risk forecasts, our model outperforms not only the standard DCC model, but also an alternative range-based DCC model.

1. Introduction

Models that can describe the dynamic properties of two or more asset returns play an important role in financial econometrics. Multivariate volatility models have been used to understand and predict the temporal dependence in second order moments of asset returns. These models can explain how covariances change over time and therefore describe temporal dependencies among assets. Such relations are vital in many financial applications, such as asset pricing, portfolio optimization, risk management, the estimation of systemic risk in banking, value-at-risk estimation, asset allocation and many others.

One of the most popular multivariate volatility models is the dynamic conditional correlation (DCC) model introduced independently by Engle (2002) and Tse and Tsui (2002). The latter representation however has attracted considerably less interest in the literature. The advantages of the DCC model are the positive definiteness of the conditional covariance martices and the ability to describe time-varying conditional correlations and covariances in a parsimonious way. The parameters of the DCC model can be estimated in two stages, which makes this approach relatively simple and possible to apply even for very large portfolios. The DCC model has become extremely popular and has been widely applied and modified (e.g. Heaney and Sriananthakumar, 2012; Lehkonen and Heimonen, 2014; Bouri et al., 2017; Bernardi and Catania, 2018; Dark, 2018; Karanasos et al., 2018).

Most volatility models are return-based models, i.e. they are estimated on returns, which are calculated based only on closing prices. Meanwhile, the use of daily low and high prices leads to more accurate estimates and forecasts of variances (see e.g. Chou, 2005; Brandt and Jones, 2006; Lin et al., 2012; Fiszeder and Perczak, 2016; Molnár, 2016) and covariances (see e.g. Chou et al., 2009; Fiszeder, 2018). Daily low and high prices are almost always available alongside closing prices in financial series. Therefore, making use of them in volatility models is very important from a practical viewpoint. DCC models formulated with the usage of

☆ This research was supported by the National Science Centre project number 2016/21/B/HS4/00662 entitled “Multivariate volatility models - the application of low and high prices”. The authors would like to thank anonymous reviewer for his/her constructive comments.
* Corresponding author at: University of Stavanger, UiS Business School, 4036 Stavanger, Norway.
E-mail address: peto.molnar@gmail.com (P. Molnár).

https://doi.org/10.1016/j.jempfin.2019.08.004
Received 13 July 2018; Received in revised form 12 June 2019; Accepted 10 August 2019
Available online 21 August 2019
0927-5398/

low and high prices have already been proposed in the literature, including the range-based DCC by Chou et al. (2009) and the range-based regime-switching DCC by Su and Wu (2014). These models, however, are based on modelling the time evolution of price range and it is not possible to compare them directly with the return-based DCC model. We propose a DCC model constructed using the Range-GARCH model by Molnár (2016), which is formulated with the usage of low and high prices but also based on returns calculated from closing prices.

Our contribution is threefold. First, we construct a new specification of the DCC model based on the Range-GARCH model by Molnár (2016), which we refer to as the DCC-Range-GARCH model (denoted by DCC-RGARCH). The model itself is very similar to the DCC model by Engle (2002). Squared errors in the univariate GARCH model are replaced by the Parkinson (1980) volatility estimator, but the parametrization of the covariance matrix remains the same. Second, we show using low and high prices in the formulation of the DCC model improves the estimation of the covariance matrix of returns and increases the accuracy of covariance and VaR forecasts based on this model, compared with the standard DCC model based on closing prices. Since both models, DCC and DCC-RGARCH, share the same structure in the correlation component, achieving more precise volatility estimates improves the covariance forecasts. Third, we demonstrate that covariance forecasts based on our proposed model are more accurate than those obtained using the range-based DCC model by Chou et al. (2009). That is an important conclusion, because the range-based DCC model is also formulated using low and high prices and is the main competitor for the DCC-RGARCH model in this class of models.

The rest of the paper is organized in the following way. Section 2 provides a description of applied models and methods. Section 3 presents data: three currency pairs -EUR/USD, USD/JPY and GBP/USD, three commodity exchange traded funds (ETFs) - United States Oil Fund, United States Natural Gas Fund and Energy Select Sector SPDR Fund and five U.S. stocks - Amazon, Apple, Goldman Sachs, Google and IBM. In Section 4.1 the parameters of the return-based DCC, range-based DCC and DCC-RGARCH models are estimated and compared. Section 4.2 evaluates the forecasts of the variance of returns from the GARCH, CARR and RGARCH models. In Section 4.3 the accuracy of covariance forecasts based on the DCC-GARCH and DCC-CARR models is compared with the forecasts from the DCC-RGARCH model. Section 4.4 evaluates the VaR forecasts based on all considered DCC models. Section 5 concludes.

2. Theoretical background

2.1. The DCC-GARCH model

In this paper we extend the DCC model by Engle (2002) by introducing the range (the difference between low and high prices) to the model. First, we present the standard DCC model based on closing prices. In order to better distinguish this model from its competitors used in the paper, which are based on different univariate models, we will refer to it as the DCC-GARCH model.

Let us assume that $\boldsymbol{\varepsilon}_t$ ($N \times 1$ vector) is the innovation process for the conditional mean (or in a particular case the multivariate return process) and can be written as:

$$\boldsymbol{\varepsilon}_t | \psi_{t-1} \sim Normal\ (0, \mathbf{cov}_t), \tag{1}$$

where ψ_{t-1} is the set of all information available at time $t-1$, *Normal* is the multivariate normal distribution and $\mathbf{cov}_t$ is the $N \times N$ symmetric conditional covariance matrix.

The DCC(P, Q)-GARCH(p, q) model by Engle (2002) can be presented as:

$$\mathbf{cov}_t = \mathbf{D}_t \mathbf{cor}_t \mathbf{D}_t, \tag{2}$$

$$\mathbf{cor}_t = \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1}, \tag{3}$$

$$\mathbf{Q}_t = \left(1 - \sum_{i=1}^{Q} \zeta_i - \sum_{j=1}^{P} \theta_j\right) \mathbf{S} + \sum_{i=1}^{Q} \zeta_i (\mathbf{z}_{t-i} \mathbf{z}'_{t-i}) + \sum_{j=1}^{P} \theta_j \mathbf{Q}_{t-j}, \tag{4}$$

where $\mathbf{D}_t = \text{diag}(h_{1t}^{1/2}, h_{2t}^{1/2} \ldots, h_{Nt}^{1/2})$, conditional variances h_{kt} (for $k = 1, 2, \ldots, N$) are described as univariate GARCH models (Eqs. (5)–(6)), $\mathbf{z}_t$ is the standardized $N \times 1$ residual vector assumed to be serially independently distributed given as $\mathbf{z}_t = \mathbf{D}_t^{-1} \boldsymbol{\varepsilon}_t$, $\mathbf{cor}_t$ is the time varying $N \times N$ conditional correlation matrix of $\mathbf{z}_t$, $\mathbf{S}$ is the unconditional $N \times N$ covariance matrix of $\mathbf{z}_t$ (according to Engle, 2002) and $\mathbf{Q}_t^*$ is the diagonal $N \times N$ matrix composed of the square root of the diagonal elements of $\mathbf{Q}_t$. The parameters ζ_i (for $i = 1, 2, \ldots, Q$), θ_j(for $j = 1, 2, \ldots, P$) are nonnegative and satisfy the condition $\sum_{i=1}^{Q} \zeta_i + \sum_{j=1}^{P} \theta_j < 1$.

The univariate GARCH(p, q) model applied in the DCC-GARCH model can be written as:

$$\varepsilon_{kt} | \psi_{t-1} \sim Normal\ (0, h_{kt}), \quad k = 1, 2, \ldots, N, \tag{5}$$

$$h_{kt} = \alpha_{k0} + \sum_{i=1}^{q} \alpha_{ki} \varepsilon_{k\,t-i}^2 + \sum_{j=1}^{p} \beta_{kj} h_{k\,t-j}, \tag{6}$$

where $\alpha_{k0} > 0, \alpha_{ki} \geq 0, \beta_{kj} \geq 0$ (for $k = 1, 2, \ldots, N$; $i = 1, 2, \ldots, q$; $j = 1, 2, \ldots, p$), weaker conditions for non-negativity of the conditional variance can be assumed (see Nelson and Cao, 1992). The requirement for covariance stationarity of ε_{kt} is $\sum_{i=1}^{q} \alpha_{ki} + \sum_{j=1}^{p} \beta_{kj} < 1$.

A nice feature of the DCC-GARCH model is that its parameters can be estimated by the quasi-maximum likelihood method using a two-stage approach (see Engle and Sheppard, 2001). Let the parameters of the model $\boldsymbol{\Theta}$ be written in two groups $\boldsymbol{\Theta}' = (\boldsymbol{\Theta}'_1, \boldsymbol{\Theta}'_2)$, where $\boldsymbol{\Theta}_1$ is the vector of parameters of conditional means and variances and $\boldsymbol{\Theta}_2$ is the vector of parameters of the correlation part of the model. The log-likelihood function can be written as the sum of two parts:

$$L(\boldsymbol{\Theta}) = L_{Vol}(\boldsymbol{\Theta}_1) + L_{Corr}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1), \tag{7}$$

where $L_{Vol}(\boldsymbol{\Theta}_1)$ represents the volatility part:

$$L_{Vol}(\boldsymbol{\Theta}_1) = -\frac{1}{2}\sum_{t=1}^{n}\left(N\ln(2\pi) + ln\left|\mathbf{D}_t\right|^2 + \boldsymbol{\varepsilon}'_t\mathbf{D}_t^{-2}\boldsymbol{\varepsilon}_t\right), \tag{8}$$

while $L_{Corr}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1)$ can be viewed as the correlation component:

$$L_{Corr}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1) = -\frac{1}{2}\sum_{t=1}^{n}\left(\ln\left|\mathbf{cor}_t\right| + \mathbf{z}'_t\mathbf{cor}_t^{-1}\mathbf{z}_t - \mathbf{z}'_t\mathbf{z}_t\right). \tag{9}$$

$L_{Vol}(\boldsymbol{\Theta}_1)$ can be written as the sum of log-likelihood functions of N univariate GARCH models:

$$L_{Vol}(\boldsymbol{\Theta}_1) = -\frac{1}{2}\sum_{k=1}^{N}\left(n\ln(2\pi) + \sum_{t=1}^{n}\left(ln(h_{kt}) + \frac{\epsilon_{kt}^2}{h_{kt}}\right)\right). \tag{10}$$

This means that in the first stage the parameters of univariate GARCH models can be estimated separately for each of the assets and the estimates of h_{kt} can be obtained. In the second stage residuals transformed by their estimated standard deviations are used to estimate the parameters of the correlation part ($\boldsymbol{\Theta}_2$) conditioning on the parameters estimated in the first stage ($\hat{\boldsymbol{\Theta}}_1$).

2.2. The CARR model

The second benchmark to compare with our new model is the range-based DCC model. This is based on the CARR model by Chou (2005), which we describe now.

Let assume that H_t and L_t are high and low prices over a fixed period such as day, week or month and the observed price range is given as $R_t = ln\left(H_t\right) - ln\left(L_t\right)$. The CARR$(p, q)$ model can be described as:

$$R_t = \lambda_t u_t, \tag{11}$$

$$u_t | \psi_{t-1} \sim \exp\left(1, \xi_t\right), \tag{12}$$

$$\lambda_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i R_{t-i} + \sum_{j=1}^{p}\beta_j\lambda_{t-j}, \tag{13}$$

where λ_t is the conditional mean of the range and u_t is the disturbance term.
The exponential distribution is a natural choice for the conditional distribution of u_t because it takes positive values. To ensure the positivity of λ_t the parameters of the CARR model have to meet conditions analogous to those in the GARCH model (see Nelson and Cao, 1992). The process is covariance stationary if the following condition is met:

$$\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1. \tag{14}$$

It is worth emphasizing that the CARR model describes the dynamics of the conditional mean of the price range, not the conditional variance of returns as in the case of the GARCH model.

The parameters of the CARR model can be estimated by the quasi-maximum likelihood method. The log-likelihood function can be written as:

$$L(\boldsymbol{\varsigma}) = -\sum_{t=1}^{n}\left(\ln\lambda_t + \frac{R_t}{\lambda_t}\right), \tag{15}$$

where $\boldsymbol{\varsigma}$ is a vector containing unknown parameters of the model. The estimators obtained by the quasi-maximum likelihood method are consistent (see Engle and Russell, 1998; Engle, 2002; Chou, 2005).

2.3. The DCC-CARR model

In this paper the new DCC-RGARCH model is compared not only with the DCC-GARCH model, formulated on closing prices, but also with the range-based DCC model which, like the proposed model, is formulated using low and high prices. Chou et al. (2009) combined the CARR model by Chou (2005) with the DCC model by Engle (2002) to propose the range-based DCC model, which we refer to as the DCC-CARR model in this paper. The CARR model describes the dynamics of the conditional mean of the price range, and so in order to estimate values of the conditional standard deviation of returns the conditional price range has to be scaled according to the formula: $\lambda_{kt}^* = \text{adj}_k\lambda_{kt}$ for $k = 1, 2, \ldots, N$, where $\text{adj}_k = \overline{\sigma}_k / \overline{\lambda}_k$. The scaling factor adj_k is estimated as the quotient of unconditional standard deviation of returns by the sample mean of the conditional range.

The DCC(P, Q)-CARR(p, q) model can be expressed as:

$$\boldsymbol{\varepsilon}_t | \psi_{t-1} \sim Normal\ (0, \mathbf{cov}_t), \tag{16}$$

$$\mathbf{cov}_t = \mathbf{D}_t \mathbf{cor}_t \mathbf{D}_t, \tag{17}$$

$$\mathbf{cor}_t = \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1}, \tag{18}$$

$$\mathbf{Q}_t = \left(1 - \sum_{i=1}^{Q} \zeta_i - \sum_{j=1}^{P} \theta_j\right) \mathbf{S} + \sum_{i=1}^{Q} \zeta_i (\mathbf{z}_{t-i}^{CARR} (\mathbf{z}_{t-i}^{CARR})') + \sum_{j=1}^{P} \theta_j \mathbf{Q}_{t-j}, \tag{19}$$

where $\mathbf{D}_t = \text{diag}\left(\lambda_{1t}^*, \lambda_{2t}^*, \dots, \lambda_{Nt}^*\right)$, $\mathbf{z}_t^{CARR}$ is the standardized $N \times 1$ residual vector which contains the standardized residuals z_{kt}^{CARR} calculated from the CARR model (Eqs. (11)–(13)) as $z_{kt}^{\text{CARR}} = \varepsilon_{kt} / \lambda_{kt}^*$, the other variables are defined in the same way as in the DCC-GARCH model.

The parameters of the DCC-CARR model can be estimated by the quasi-maximum likelihood method using a two-stage approach. The log-likelihood function can be written as the sum of two parts, the volatility part and the correlation part:

$$L^{DCC-CARR}(\boldsymbol{\Theta}) = L_{Vol}^{DCC-CARR}(\boldsymbol{\Theta}_1) + L_{Corr}^{DCC-CARR}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1), \tag{20}$$

$$L_{Vol}^{DCC-CARR}(\boldsymbol{\Theta}_1) = -\frac{1}{2} \sum_{k=1}^{N} \left(n \ln(2\pi) + \sum_{t=1}^{n} \left(2ln(\lambda_{kt}^*) + \frac{\epsilon_{kt}^2}{\lambda_{kt}^{*2}} \right) \right) \tag{21}$$

$$L_{Corr}^{DCC-CARR}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1) = -\frac{1}{2} \sum_{t=1}^{n} \left(\ln |\mathbf{cor}_t| + (\mathbf{z}_t^{CARR})' \mathbf{cor}_t^{-1} \mathbf{z}_t^{CARR} - (\mathbf{z}_t^{CARR})' \mathbf{z}_t^{CARR} \right). \tag{22}$$

This means that in the first stage the parameters of the CARR models can be estimated separately for each of the assets. In the second stage the standardized residuals z_{kt}^{CARR} are used to maximize Eq. (22) in order to estimate the parameters of the correlation component.

2.4. The Range-GARCH model

In the new specification of the DCC-RGARCH model we use the Range-GARCH model introduced by Molnár (2016). The RGARCH(p, q) model can be formulated as:

$$\varepsilon_t | \psi_{t-1} \sim Normal\ (0, h_t), \tag{23}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \sigma_{Pt-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}, \tag{24}$$

where σ_{Pt}^2 is the Parkinson (1980) estimator calculated as $\sigma_{Pt}^2 = [ln(H_t/L_t)]^2/(4\ln 2)$.
In this formulation other variance estimators based on low, high and opening or closing prices, like the Garman and Klass (1980) or Rogers and Satchell (1991) estimators, can be applied instead of the Parkinson estimator. For an overview of range-based volatility estimators see Molnár (2012), Fiszeder and Perczak (2013).

To ensure the positivity of h_t the parameters of the RGARCH model must meet conditions analogous to those in the GARCH model (see Nelson and Cao, 1992). The RGARCH process is covariance stationary if the following condition is met:

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1. \tag{25}$$

It is worth emphasizing that the RGARCH model describes the dynamics of the conditional variance of returns, not the conditional mean of the price range, as in the case of the CARR model. The parameters of the RGARCH model can be estimated by the quasi-maximum likelihood method and the likelihood function is the same as in the return-based GARCH model.

2.5. The DCC-Range-GARCH model

In this subsection we introduce our new DCC-Range-GARCH model (denoted by DCC-RGARCH). The DCC(P, Q)-RGARCH(p, q) model can be presented as:

$$\boldsymbol{\varepsilon}_t | \psi_{t-1} \sim Normal\ (0, \mathbf{cov}_t), \tag{26}$$

$$\mathbf{cov}_t = \mathbf{D}_t \mathbf{cor}_t \mathbf{D}_t, \tag{27}$$

$$\mathbf{cor}_t = \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1}, \tag{28}$$

$$\mathbf{Q}_t = \left(1 - \sum_{i=1}^{Q} \zeta_i - \sum_{j=1}^{P} \theta_j\right) \mathbf{S} + \sum_{i=1}^{Q} \zeta_i (\mathbf{z}_{t-i}^{RGARCH} (\mathbf{z}_{t-i}^{RGARCH})') + \sum_{j=1}^{P} \theta_j \mathbf{Q}_{t-j}, \tag{29}$$

where $\mathbf{D}_t = \text{diag}\left((h_{1t}^{RGARCH})^{1/2}, (h_{2t}^{RGARCH})^{1/2} \ldots, (h_{Nt}^{RGARCH})^{1/2}\right)$, conditional variances h_{kt}^{RGARCH} (for $k = 1, 2, \ldots, N$) are described as for the RGARCH model (Eqs. (23)–(24)), $\mathbf{z}_t^{RGARCH}$ is the standardized $N \times 1$ residual vector which contains the standardized residuals z_{kt}^{RGARCH} calculated from the RGARCH model as $z_{kt}^{\text{RGARCH}} = \varepsilon_{kt} / \left(h_{kt}^{RGARCH}\right)^{1/2}$, the other variables are defined in the same way as in the DCC-GARCH model.

The parameters of the DCC-R-GARCH model can be estimated by the quasi-maximum likelihood method using a two-stage approach. The log-likelihood function can be written as the sum of two parts, the volatility part and the correlation part:

$$L^{DCC-RGARCH}(\boldsymbol{\Theta}) = L_{Vol}^{DCC-RGARCH}(\boldsymbol{\Theta}_1) + L_{Corr}^{DCC-RGARCH}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1), \tag{30}$$

$$L_{Vol}^{DCC-RGARCH}(\boldsymbol{\Theta}_1) = -\frac{1}{2}\sum_{k=1}^{N}\left(n\ln(2\pi) + \sum_{t=1}^{n}\left(ln(h_{kt}) + \frac{\epsilon_{kt}^2}{h_{kt}}\right)\right) \tag{31}$$

$$L_{Corr}^{DCC-RGARCH}(\boldsymbol{\Theta}_2 | \boldsymbol{\Theta}_1) = -\frac{1}{2}\sum_{t=1}^{n}\Big(\ln|\mathbf{cor}_t| + (\mathbf{z}_t^{RGARCH})'\mathbf{cor}_t^{-1}\mathbf{z}_t^{RGARCH} - (\mathbf{z}_t^{RGARCH})'\mathbf{z}_t^{RGARCH}\Big), \tag{32}$$

This means that in the first stage the parameters of univariate RGARCH models can be estimated separately for each of the assets. In the second stage the standardized residuals z_{kt}^{RGARCH} are used to maximize Eq. (32) in order to estimate the parameters of the correlation component.

3. Data

We apply the proposed model and its competitors to three different sets of data: three currency rates, three commodity exchange traded funds and five stocks. The currency rates are the three most heavily traded currency pairs in the Forex market, namely: EUR/USD, USD/JPY and GBP/USD.

The second set are three exchange-traded funds (ETF) listed on the New York Stock Exchange Arca, namely (the names given in the brackets will be used later in tables): the United States Oil Fund (Oil), the United States Natural Gas Fund (Natural Gas) and the Energy Select Sector SPDR Fund (Energy). Commodity exchange traded funds provide investors with the convenience of commodity exposure without a commodity futures account. The first two ETFs offer exposure to a single commodity (oil/gas), whereas the third ETF tracks the price and performance of the Standard and Poor's Energy Select Sector Index.

The third set of data consists of five selected U.S. stocks, namely: Amazon, Apple, Goldman Sachs, Google and IBM. Since there are many stocks that could be chosen for this purpose, we decided to follow CBOE and select the stocks for which CBOE calculates implied volatility indices (even though implied volatility indices are not used in this paper).

We evaluate the models considered for daily data in the nine-year period from January 2, 2008, to December 30, 2016. This is a relatively long period, which includes both very volatile periods – the collapse of Lehman Brothers, the worst phase of the global financial crisis, the European sovereign debt crisis and Brexit – but also tranquil periods with low volatility.

The descriptive statistics for the percentage returns calculated as $r_t = 100\ln(p_t/p_{t-1})$, where p_t is the closing price at time t, are presented in Table 1. The means of returns are positive for stocks and the Energy Select Sector SPDR Fund and negative for currencies and the other ETFs. The standard deviation of returns is significantly lower for currencies. Most distributions of returns are asymmetric, and all display high leptokurtosis.

4. Results

We consider three DCC models in the analysis:

(1) The DCC-GARCH model by Engle (2002) summarized by Eqs. (1)–(6), where parameters are estimated based only on closing prices.

(2) The DCC-CARR model by Chou et al. (2009), see Eqs. (16)–(19). In this specification the CARR model (Eqs. (11)–(13)) is applied in the DCC model instead of the univariate GARCH model.

(3) The proposed DCC-RGARCH model summarized by Eqs. (26)–(29). In this specification the RGARCH model described by Eqs. (23)–(24) is applied in the DCC model instead of the univariate GARCH model.

We also consider a DCC model using two asymmetric GARCH models, i.e. the EGARCH (Nelson, 1991) and GJR (Glosten et al., 1993) models, instead of the standard GARCH model. These models are able to capture often-reported asymmetric responses to positive and negative shocks in the conditional variance. However we find that covariance forecasts based on the DCC-EGARCH and DCC-GJR models are not significantly better than forecasts from the DCC-GARCH model for any of the currencies and ETFs considered, or for most stocks (the results are given in Tables A.1 and A.2 in the Appendix), and so we do not extend our models to describe the effect of asymmetry in variance.

The considered exchange rates, ETFs and stocks are not cointegrated (according to the Johansen test). Mean equations for returns are very simple: each mean equation is a constant, because in our data the sample return of any asset is not dependent on its own past returns nor on the past returns of other assets.

We first compare the fit of the models estimated on the whole sample of data, and then compare the forecasts from these models. We analyse forecasts of variances and forecasts of covariances separately, because models for variances already exist whereas forecasting covariances is our main contribution.

Table 1
Summary statistics of daily returns.

Assets	Mean × 10^2	Minimum	Maximum	Standard deviation	Skewness	Excess kurtosis
Currency rates						
EUR/USD	−1.401	−2.554	3.503	0.657	0.116*	4.825*
JPY/USD	−0.198	−5.448	3.779	0.692	−0.008	7.670*
GBP/USD	−2.037	−8.322	2.870	0.641	−1.245*	17.043*
Exchange-traded funds						
Oil	−8.234	−11.439	9.199	2.286	−0.133*	5.256*
Natural Gas	−15.146	−9.745	13.942	2.651	0.172*	4.173*
Energy	0.463	−19.033	18.051	1.965	−0.408*	15.390*
Stocks						
Amazon	9.220	−13.640	23.768	2.482	0.548*	11.837*
Apple	6.631	−19.128	12.577	2.039	−0.499*	10.454*
Goldman Sachs	0.984	−22.022	23.245	2.538	0.054	18.504*
Google	3.658	−10.271	18.231	1.894	0.752*	14.891*
IBM	2.801	−8.799	11.035	1.443	−0.215*	8.928*

The sample period is January 2, 2008, to December 30, 2016.
*Indicates that the null hypothesis (the skewness or excess kurtosis is equal to zero) was rejected at the 10% significance level.

4.1. In-sample comparison of models

The parameters of the considered models are estimated using the quasi-maximum likelihood method. The results of the estimation are presented in Tables 2–4 separately for exchange rates, ETFs and stocks.

The estimation of parameters for the GARCH, R-GARCH and CARR models is based on different kinds of data: on closing prices for the first two models[1] and on range data for the third model. However, for the DCC-CARR, which uses the CARR model, it is possible to calculate the likelihood function based on the scaled conditional price range according to formula (21). Thanks to this, it is possible to evaluate all the DCC models based on the whole likelihood function, including both the volatility and correlation parts. In order to assess whether the differences between values of likelihood function are statistically significant, we apply the Rivers and Vuong (2002) and Clarke (2007) tests for non-nested model selection. The values of the likelihood function are higher for the DCC-RGARCH model than for the benchmark DCC-GARCH model for all analysed data sets, which means that the DCC-RGARCH model better describes the considered time series. The results for the DCC-CARR model are ambiguous and depend on the type of test applied.

The application of range data changes the parameter estimates for the considered models significantly. Specifically, the estimates of the parameters α_{k1} are much higher and the estimates of the parameters β_{k1} much lower in the CARR and RGARCH models compared with the GARCH model. This is important in terms of both modelling and forecasting volatility, because for the CARR and RGARCH models the shocks in the previous period have a stronger impact on the current volatility than the impact you observe for the GARCH model. Thus models formulated with range data respond more quickly to changing market conditions. Slow response to abrupt changes in the market is widely cited as one of the greatest weaknesses of GARCH-type models formulated based on closing prices (e.g. Andersen et al., 2003; Hansen et al., 2012).

Direct comparison of the parameters of the CARR model with the parameters of the GARCH and RGARCH models is, however, difficult, because they describe different measures of volatility. The CARR model describes the dynamics of the conditional mean of the price range, while the GARCH and RGARCH models describe the conditional variance of returns.

One can also notice that the sum of the estimates of the parameters α_{k1} and β_{k1} in the RGARCH model is higher than one for ETFs and stocks. However, this does not mean that the analysed processes are covariance non-stationary. It results from the fact that the Parkinson estimator underestimates the volatility of returns in the presence of opening jumps (such jumps do not occur in the Forex market since it does not close overnight), causing an increase in the estimate of the parameter α_{k1} (see Molnár, 2016).

On the other hand, there are no considerable differences between the considered models in the estimates of parameters for the correlation component. Thus, the main differences in the behaviour of the time-varying covariances from those models results from the usage of the different standardized residuals z_{kt}, z_{kt}^{CARR} and z_{kt}^{RGARCH} in Eqs. (4), (19) and (29) of the DCC-GARCH, DCC-CARR and DCC-RGARCH models, respectively.

4.2. Comparison of variance forecasts

In this section we compare the forecasting performance of the three univariate models, which are used in the DCC models. We formulate out-of-sample one-day-ahead forecasts of variance based on the GARCH, CARR and RGARCH models, where parameters are estimated separately each day based on a rolling sample of a fixed size of 500 (approximately a two-year period; the first

[1] In the R-GARCH model, the Parkinson estimator with the high-low range is used as an explanatory variable but the likelihood function is formulated based on closing prices.

Table 2
Results of parameter estimation for currency rates.

Parameter	DCC-GARCH		DCC-CARR		DCC-RGARCH	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
γ_{10}	−0.011	0.011	–	–	−0.019	0.011
α_{10}	0.001	0.001	0.007	0.003	0.002	0.002
α_{11}	0.037	0.006	0.093	0.012	0.052	0.013
β_{11}	0.960	0.006	0.901	0.013	0.943	0.015
γ_{20}	−0.015	0.013	–	–	−0.005	0.012
α_{20}	0.006	0.004	0.019	0.007	0.010	0.006
α_{21}	0.055	0.019	0.134	0.023	0.133	0.040
β_{21}	0.933	0.024	0.847	0.029	0.843	0.046
γ_{30}	−0.009	0.011	–	–	−0.016	0.010
α_{30}	0.003	0.002	0.006	0.003	0.004	0.002
α_{31}	0.076	0.030	0.110	0.014	0.116	0.049
β_{31}	0.921	0.026	0.883	0.014	0.871	0.041
ζ_1	0.044	0.006	0.048	0.007	0.044	0.006
θ_1	0.922	0.011	0.923	0.012	0.921	0.011
ln L	−5694.139		−5649.297		−5648.297	
Rivers–Vuong	–		2.796 (0.003)		2.563 (0.005)	
Clarke	–		−2.028 (0.979)		6.414 (0.000)	

The sample period is January 2, 2008, to December 30, 2016, the parameters γ_{10}, γ_{20}, γ_{30} are constants, α_{k0} α_{k1} β_{k1} are the parameters of the univariate GARCH model (Eq. (6)), the CARR model (Eq. (13)) and the RGARCH model (Eq. (24)), $k = 1, 2, 3$ for EUR/USD, JPY/USD and GBP/USD, respectively, ζ_1, θ_1 are the parameters of the correlation part (Eqs. (4), (19) and (29) for the DCC-GARCH, DCC-CARR and DCC-RGARCH models, respectively), ln L is the logarithm of the likelihood function, the Rivers–Vuong and Clarke are test statistics for model selection, where comparisons are made with the DCC-GARCH model, p-values are given in brackets. A low p-value means that the indicated model is superior to the benchmark DCC-GARCH model.

Table 3
Results of parameter estimation for exchange-traded funds.

Parameter	DCC-GARCH		DCC-CARR		DCC-RGARCH	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
γ_{10}	−0.127	0.051	–	–	−0.127	0.051
α_{10}	0.090	0.034	0.049	0.015	0.111	0.052
α_{11}	0.056	0.009	0.096	0.011	0.154	0.026
β_{11}	0.932	0.011	0.887	0.014	0.897	0.018
γ_{20}	−0.017	0.036	–	–	−0.055	0.035
α_{20}	0.020	0.011	0.017	0.007	0.031	0.022
α_{21}	0.065	0.014	0.140	0.017	0.236	0.069
β_{21}	0.933	0.014	0.854	0.019	0.864	0.039
γ_{30}	0.058	0.026	–	–	0.016	0.026
α_{30}	0.024	0.009	0.048	0.013	0.019	0.016
α_{31}	0.090	0.015	0.256	0.023	0.382	0.075
β_{31}	0.904	0.015	0.719	0.026	0.748	0.047
ζ_1	0.014	0.003	0.017	0.003	0.013	0.003
θ_1	0.980	0.004	0.980	0.004	0.982	0.005
ln L	−13 419.952		−13 445.131		−13 358.665	
Rivers-Vuong	–		−0.553 (0.710)		3.143 (0.001)	
Clarke	–		−11.344 (1.000)		4.117 (0.000)	

The sample period is January 2, 2008, to December 30, 2016, the parameters γ_{10}, γ_{20}, γ_{30} are constants, α_{k0} α_{k1} β_{k1} are the parameters of the univariate GARCH model (Eq. (6)), the CARR model (Eq. (13)) and the RGARCH model (Eq. (24)), $k = 1, 2, 3$ for Natural Gas, Oil and Energy, respectively, ζ_1, θ_1 are the parameters of the correlation part (Eqs. (4), (19) and (29) for the DCC-GARCH, DCC-CARR and DCC-RGARCH models, respectively). Ln L is the logarithm of the likelihood function, the Rivers–Vuong and Clarke are test statistics for model selection, where comparisons are made with the DCC-GARCH model, p-values are given in brackets. A low p-value means that the indicated model is superior to the benchmark DCC-GARCH model.

in-sample period is from January 2, 2008 to December 31, 2009). We evaluate forecasts for the seven-year period from January 4, 2010, to December 30, 2016.

The sum of squares of 15-min returns (the realized variance) is used as a proxy of the daily variance. The forecasts from the models are evaluated based on two primary measures, namely, the mean squared error (MSE) and the mean absolute error (MAE). In order to evaluate the statistical significance of the results the Diebold–Mariano test (Diebold and Mariano, 1995) corrected for small-sample bias (Harvey et al., 1997) is applied.

A pairwise comparison is performed and the results for the RGARCH model are presented with respect to the two benchmarks: first the GARCH model and second the CARR model. The GARCH and CARR models are the most popular univariate volatility

Table 4
Results of parameter estimation for stocks.

Parameter	DCC-GARCH		DCC-CARR		DCC-RGARCH	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
γ_{10}	0.118	0.045	–	–	0.119	0.043
α_{10}	0.023	0.039	0.057	0.019	0.400	0.138
α_{11}	0.014	0.010	0.187	0.024	0.396	0.087
β_{11}	0.982	0.016	0.793	0.028	0.684	0.071
γ_{20}	0.155	0.038	–	–	0.086	0.034
α_{20}	0.132	0.042	0.128	0.035	0.189	0.069
α_{21}	0.098	0.024	0.264	0.038	0.250	0.055
β_{21}	0.868	0.027	0.679	0.051	0.783	0.046
γ_{30}	0.053	0.037	–	–	0.029	0.033
α_{30}	0.062	0.047	0.064	0.017	0.063	0.036
α_{31}	0.115	0.063	0.241	0.030	0.243	0.054
β_{31}	0.879	0.062	0.734	0.034	0.821	0.037
γ_{40}	0.055	0.035	–	–	0.057	0.032
α_{40}	0.127	0.085	0.081	0.018	0.339	0.140
α_{41}	0.083	0.060	0.240	0.025	0.595	0.193
β_{41}	0.885	0.070	0.720	0.031	0.556	0.143
γ_{50}	0.036	0.026	–	–	0.019	0.024
α_{50}	0.126	0.041	0.061	0.032	0.134	0.033
α_{51}	0.124	0.037	0.221	0.048	0.404	0.065
β_{51}	0.814	0.046	0.741	0.058	0.663	0.049
ζ_1	0.003	0.001	0.006	0.002	0.003	0.000
θ_1	0.993	0.003	0.989	0.004	0.991	0.001
ln L	−21 205.733		−21 055.408		−20 920.942	
Rivers-Vuong	–		2.538 (0.006)		4.910 (0.000)	
Clarke	–		−3.255 (0.999)		12.497 (0.000)	

The sample period is January 2, 2008, to December 30, 2016, the parameters γ_{10}, γ_{20}, γ_{30} are constants, α_{k0} α_{k1} β_{k1} are the parameters of the univariate GARCH model (Eq. (6)), the CARR model (Eq. (13)) and the RGARCH model (Eq. (24)), $k = 1,2,3,4,5$ for Amazon, Apple, Goldman Sachs, Google and IBM, respectively, ζ_1, θ_1 are the parameters of the correlation part (Eqs. (4), (19) and (29) for the DCC-GARCH, DCC-CARR and DCC-RGARCH models, respectively), ln L is the logarithm of the likelihood function, the Rivers–Vuong and Clarke are test statistics for model selection, where comparisons are made with the DCC-GARCH model, p-values are given in brackets. A low p-value means that the indicated model is superior to the benchmark DCC-GARCH model.

models formulated based on returns constructed on closing prices and price range, respectively. The forecasting performance results are presented in Tables 5 and 6 for the MSE and MAE criteria, respectively.

According to the MSE criterion, the forecasts of variance from the RGARCH model are more accurate for currencies and the Energy Select Sector SPDR Fund. For the other ETFs and stocks, the results are mixed. However, there are large outliers in the data set, which affect the MSE measure. Such outliers are present for ETFs and stocks (see e.g. minimum and maximum returns in Table 1). A quite different picture emerges from the MAE criterion. According to this measure the best forecasts are formulated based on the RGARCH (except Amazon and Apple stocks) and, in almost all cases, the higher forecasting accuracy of this model is statistically significant at the 10% significance level (the exceptions are the GBP/USD currency pair and Google's stock with respect to the CARR benchmark model). The CARR and RGARCH models' forecasting superiority over the GARCH model has already been documented by Chou (2005) and Molnár (2016), respectively. Higher forecast accuracy based on the RGARCH model in comparison to the CARR model has not previously been demonstrated in the literature.

In order to check the robustness of the results, we also consider 5-min returns instead of 15-min returns and three additional evaluation measures (the coefficient of determination, the logarithmic loss function and the linear exponential loss function). The results for the MSE and MAE criteria for 5-min returns are presented in Table A.3 in Appendix. The conclusions are very similar to those presented for 15-min returns.

The first additional measure is the coefficient of determination from the Mincer–Zarnowitz regression. A proxy of volatility is regressed on a constant and the forecast of volatility. It is a very simple and popular way to evaluate the forecasting performance of volatility models (see e.g. Poon and Granger, 2003). The values of the coefficient of determination for the competing models are presented in Table 7. These results are in accordance with those for the MSE measure.

To reduce the impact of outliers, we also use the logarithmic loss function. This is calculated similarly to the MSE measure, but the logarithm of a volatility proxy and the logarithm of the volatility forecast are applied (see Pagan and Schwert, 1990). The estimates of the logarithmic loss function are given in Table 8. These results are very similar to those for the MAE criterion and indicate that the forecasts from the RGARCH model are superior.

Additionally, we apply a linear exponential loss function (LINEX). For the positive coefficient a of the LINEX, the function is approximately linear for over-prediction errors and exponential for under-prediction errors. This means that under-prediction errors have a higher impact on the loss function than over-prediction errors. For the negative coefficient a the situation is exactly the opposite. The values of the LINEX function for $a = -1$ and $a = 1$ are presented in the Appendix in Tables A.4 and A.5 respectively. The results for all currency rates indicate that the variance forecasts based on the RGARCH model are more accurate than the

Table 5
Evaluation of variance forecasts: the MSE criterion.

Assets	GARCH	CARR	RGARCH	GARCH vs. RGARCH	CARR vs. RGARCH
	MSE			P-value of DM test	
Currency rates					
EUR/USD	0.112	0.120	**0.098**	0.010	0.004
GBP/USD	0.811	1.134	**0.560**	0.062	0.197
JPY/USD	0.426	0.485	**0.330**	0.022	0.049
Exchange-traded funds					
Energy	9.133	9.493	**7.558**	0.019	0.004
Oil	**14.049**	19.470	15.005	0.973	0.000
Natural Gas	**22.402**	26.507	23.383	0.960	0.000
Stocks					
Amazon	**164.230**	183.148	181.768	0.978	0.313
Apple	122.262	**94.508**	98.246	0.177	0.857
Goldman Sachs	11.917	11.986	**11.365**	0.264	0.172
Google	**50.899**	58.700	58.730	0.760	0.521
IBM	**11.586**	13.727	13.208	0.834	0.069

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The lowest values of MSE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the GARCH and CARR models. A p-value lower than the significance level means that the forecasts of variance from the RGARCH model are more accurate than the forecasts from a benchmark model (here GARCH or CARR).

Table 6
Evaluation of variance forecasts: the MAE criterion.

Assets	GARCH	CARR	RGARCH	GARCH vs. RGARCH	CARR vs. RGARCH
	MAE			P-value of DM test	
Currency rates					
EUR/USD	0.166	0.169	**0.155**	0.000	0.000
GBP/USD	0.167	0.161	**0.147**	0.000	0.165
JPY/USD	0.230	0.228	**0.206**	0.000	0.000
Exchange-traded funds					
Energy	1.190	1.292	**1.040**	0.000	0.000
Oil	2.213	2.485	**2.137**	0.002	0.000
Natural Gas	3.233	3.527	**3.198**	0.095	0.000
Stocks					
Amazon	3.704	**3.265**	3.322	0.000	0.974
Apple	2.410	**2.220**	2.263	0.011	0.915
Goldman Sachs	1.752	1.854	**1.682**	0.015	0.000
Google	2.001	1.861	**1.844**	0.013	0.172
IBM	1.064	1.043	**1.007**	0.003	0.000

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The lowest values of MAE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the GARCH and CARR models. A p-value lower than the significance level means that the forecasts of variance from the RGARCH model are more accurate than the forecasts from a benchmark model (here GARCH or CARR).

forecasts from the competing models. The outcomes for other assets are ambiguous, but they depend heavily on outliers. When the highest 1% of values are excluded, the values of the LINEX loss function are much smaller and more often indicate the RGARCH model as the best forecasting model.

4.3. Comparison of covariance forecasts

In this section, we compare out-of-sample one-day-ahead forecasts of covariance from the DCC-GARCH and DCC-CARR models with the forecasts from the DCC-RGARCH model. We use the same estimation and forecasting samples as for variances in Section 4.2. The sum of products of 15-min returns (the realized covariance) is employed as a proxy of the daily covariance for the evaluation of the forecasts. We use the same evaluation measures as in the previous section. We perform a pairwise comparison by the Diebold–Mariano test for the DCC-RGARCH model with respect to the two benchmarks: first the DCC-GARCH model and second the DCC-CARR model.

The forecasting performance results for the covariance of returns are presented in Tables 9 and 10 for the MSE and MAE criteria, respectively. For all analysed relations except the one between the United States Oil and United States Natural Gas Funds based

Table 7
Evaluation of variance forecasts: the coefficient of determination.

Assets	GARCH	CARR	RGARCH
Currency rates			
EUR/USD	0.254	0.217	**0.355**
GBP/USD	0.305	0.034	**0.513**
JPY/USD	0.200	0.080	**0.417**
Exchange-traded funds			
Energy	0.318	0.290	**0.453**
Oil	**0.405**	0.315	0.372
Natural Gas	**0.253**	0.138	0.216
Stocks			
Amazon	**0.307**	0.084	0.100
Apple	0.089	**0.395**	0.302
Goldman Sachs	0.380	0.390	**0.391**
Google	**0.244**	0.141	0.149
IBM	**0.378**	0.128	0.154

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The highest values of R^2 are marked in bold.

Table 8
Evaluation of variance forecasts: the logarithmic loss function.

Assets	GARCH	CARR	RGARCH
Currency rates			
EUR/USD	0.326	0.332	**0.294**
GBP/USD	0.249	0.262	**0.208**
JPY/USD	0.487	0.479	**0.404**
Exchange-traded funds			
Energy	0.485	0.552	**0.357**
Oil	0.617	0.631	**0.532**
Natural Gas	0.561	0.641	**0.546**
Stocks			
Amazon	1.039	**0.750**	0.770
Apple	0.974	**0.815**	0.867
Goldman Sachs	0.594	0.605	**0.557**
Google	0.883	**0.723**	0.742
IBM	0.735	0.681	**0.628**

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The lowest values of the logarithmic loss function are marked in bold.

on the MSE measure, the lowest values of loss functions are found for the DCC-RGARCH model. In most cases, this model's higher forecasting accuracy is statistically significant.[2] In the MAE measure less weight is assigned to outliers and the results for this measure clearly indicate that the DCC-RGARCH model is the best forecasting model.

The forecasts formulated based on the DCC-RGARCH are more precise than the forecasts from both the benchmark models. The first benchmark, DCC-GARCH, is based on returns formulated on the closing prices. This result shows that the application of range data in the standard univariate GARCH model increases the accuracy of covariance forecasts based on the DCC model. The second benchmark, DCC-CARR, is based on range data. This means that the way in which range data is utilized in the univariate volatility model is decisive in determining the forecasting accuracy of the DCC model. Since both benchmarks, i.e. the DCC-GARCH and DCC-CARR models, share the same structure in the correlation component as the DCC-RGARCH model, our results clearly show that more precise volatility estimates improve covariance forecasts.

The DCC-CARR model, which can be treated as the main benchmark model for models constructed based on range data, was not only inferior to the DCC-RGARCH model for most assets, but also inferior to the DCC-GARCH model for currencies and ETFs.

To check the robustness of the results, we also consider 5-min returns instead of 15-min returns and two other loss functions (the coefficient of determination and the LINEX loss function). The results for the MSE and MAE criteria for 5-min returns are presented in Table A.6 in Appendix. The outcomes are very similar to those presented for 15-min returns.

[2] Under the MSE criterion the difference between the loss function of the DCC-RGARCH model and the benchmark model is not statistically significant for EUR/USD-GBP/USD, JPY/USD-GBP/USD, Apple-IBM (with both benchmark models), Oil-Energy (with the DCC-GARCH benchmark) and Amazon-Apple, Amazon-Goldman Sachs, Apple-Google (with the DCC-CARR benchmark). Under the MAE measure there are only two relations for which there is no evidence to reject the null hypothesis of equal predictive ability. These are JPY/USD-GBP/USD (with both benchmark models) and Amazon-Apple (with the DCC-CARR benchmark).

Table 9
Evaluation of covariance forecasts: the MSE criterion.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH vs. DCC-RGARCH	DCC-CARR vs. DCC-RGARCH
	MSE			P-value of DM test	
Currency rates					
EUR/USD-JPY/USD	0.654	0.748	**0.561**	0.044	0.044
EUR/USD-GBP/USD	0.698	0.941	**0.508**	0.149	0.167
JPY/USD-GBP/USD	1.334	2.075	**1.016**	0.233	0.193
Exchange-traded funds					
Oil-Natural Gas	**62.800**	64.876	63.108	0.750	0.000
Oil-Energy	61.59	103.049	**60.347**	0.182	0.000
Natural Gas-Energy	30.734	81.546	**30.230**	0.036	0.000
Stocks					
Amazon-Apple	198.390	191.357	**172.797**	0.038	0.109
Amazon-Goldman Sachs	70.799	73.782	**70.702**	0.008	0.186
Amazon-Google	169.973	160.102	**148.522**	0.001	0.088
Amazon-IBM	48.763	45.834	**42.501**	0.005	0.088
Apple-Goldman Sachs	99.443	99.274	**87.165**	0.023	0.048
Apple-Google	268.149	265.364	**227.639**	0.097	0.113
Apple-IBM	148.476	145.042	**114.888**	0.122	0.145
Goldman Sachs-Google	66.727	63.389	**56.076**	0.000	0.032
Goldman Sachs-IBM	41.518	40.048	**36.398**	0.000	0.006
Google-IBM	56.165	53.489	**47.270**	0.021	0.076

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of MSE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the DCC-GARCH and DCC-CARR models. A p-value lower than the significance level means that the forecasts of covariance from the DCC-RGARCH model are more accurate than the forecasts from a benchmark model (here DCC-GARCH or DCC-CARR).

Table 10
Evaluation of covariance forecasts: the MAE criterion.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH vs. DCC-RGARCH	DCC-CARR vs. DCC-RGARCH
	MAE			P-value of DM test	
Currency rates					
EUR/USD-JPY/USD	0.106	0.112	**0.104**	0.002	0.000
EUR/USD-GBP/USD	0.098	0.099	**0.092**	0.000	0.006
JPY/USD-GBP/USD	0.088	0.092	**0.086**	0.166	0.120
Exchange-traded funds					
Oil-Natural Gas	1.453	1.478	**1.446**	0.028	0.000
Oil-Energy	1.291	1.494	**1.204**	0.000	0.000
Natural Gas-Energy	1.046	1.997	**1.024**	0.000	0.000
Stocks					
Amazon-Apple	1.299	1.157	**1.142**	0.000	0.103
Amazon-Goldman Sachs	1.210	1.134	**1.106**	0.000	0.001
Amazon-Google	1.506	1.299	**1.255**	0.000	0.001
Amazon-IBM	0.875	0.793	**0.765**	0.000	0.000
Apple-Goldman Sachs	1.060	1.037	**0.978**	0.000	0.000
Apple-Google	1.157	1.065	**1.015**	0.000	0.000
Apple-IBM	0.820	0.767	**0.716**	0.000	0.000
Goldman Sachs-Google	1.093	1.050	**0.971**	0.000	0.000
Goldman Sachs-IBM	0.841	0.813	**0.752**	0.000	0.000
Google-IBM	0.743	0.689	**0.651**	0.000	0.000

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of MAE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the DCC-GARCH and DCC-CARR models. A p-value lower than the significance level means that the forecasts of covariance from the DCC-RGARCH model are more accurate than the forecasts from a benchmark model (here DCC-GARCH or DCC-CARR).

Table 11 presents the coefficient of determination values from the Mincer–Zarnowitz regression. A proxy of covariance is regressed on a constant and the forecast of covariance. We are unable to calculate the logarithmic loss function (see Section 4.2) because some covariances are negative.

Table 11
Evaluation of covariance forecasts: the coefficient of determination.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH
Currency rates			
EUR/USD-JPY/USD	0.224	0.098	**0.364**
EUR/USD-GBP/USD	0.320	0.097	**0.513**
JPY/USD-GBP/USD	0.370	0.016	**0.507**
Exchange-traded funds			
Oil-Natural Gas	**0.023**	0.005	0.016
Oil-Energy	0.373	0.083	**0.392**
Natural Gas-Energy	0.029	0.012	**0.031**
Stocks			
Amazon-Apple	0.033	0.070	**0.215**
Amazon-Goldman Sachs	0.054	0.103	**0.154**
Amazon-Google	0.050	0.115	**0.256**
Amazon-IBM	0.049	0.116	**0.251**
Apple-Goldman Sachs	0.070	0.076	**0.208**
Apple-Google	0.045	0.050	**0.279**
Apple-IBM	0.025	0.047	**0.404**
Goldman Sachs-Google	0.077	0.122	**0.248**
Goldman Sachs-IBM	0.084	0.119	**0.250**
Google-IBM	0.050	0.108	**0.327**

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance based on 15-min returns is used as a proxy of covariance. The highest values of R^2 are marked in bold.

For all covariances except the relation between the United States Oil and United States Natural Gas Funds the highest R^2 values are obtained for the DCC-RGARCH model. In most cases the superiority of this model is considerable.

We obtain different results for the asymmetric loss function LINEX. The values of the function for $a = -1$ and $a = 1$ are presented in the Appendix in Tables A.7 and A.8, respectively. The results for all relations between currencies rates indicate that the covariance forecasts based on the DCC-RGARCH model are more accurate than the forecasts from the competing DCC models. The outcomes for other assets are mixed but outliers have considerable influence on the evaluation. After excluding the highest 1% of values the results depend on the valuation of the over- and under-prediction errors. For $a = -1$, i.e. when over-prediction errors have a higher impact on the loss function, then the best forecasts are based on the DCC-RGARCH model, whereas for $a = 1$, i.e. when under-prediction errors have a greater influence on the LINEX, then the DCC-CARR is better according to this criterion.

4.4. Forecasting value-at-risk

Covariance forecasting is crucial for most multivariate financial applications, such as portfolio construction, valuation of assets, risk management and contagion effect. More accurate covariance forecasts give an advantage in various financial applications. That is why covariance forecasting, similarly like volatility forecasting, has not only statistical but also economic consequences.

In this subsection we apply the considered DCC models to one such application, namely the evaluation of risk, using the value-at-risk (VaR) measure. VaR was developed by financial practitioners as an easily interpretable number which encodes information about a portfolio's risk. Despite being a single number, VaR enables managers to interpret the cost of risk and allocate capital efficiently. We formulate daily forecasts of VaR for three separate portfolios of currency rates, commodity exchange traded funds and stocks. All the portfolios are constructed with equal weights. The same assets and forecasting period are assumed as in the analysis of variances and covariances in Sections 4.2 and 4.3 We construct VaR forecasts for the 95% and 99% confidence levels.

Our evaluation of the forecasts is based on two approaches: the first involves testing the competing VaR models for statistical accuracy, while the second pertains to measuring the loss to the economic agent as a result of using the model. We test the statistical adequacy of the forecasts based on: the unconditional coverage test by Kupiec (1995), the independence and conditional coverage tests by Christoffersen (1998), and the unconditional coverage, independence and conditional coverage tests by Candelon et al. (2011). The results of these tests for the 95% VaR forecasts are presented in Table 12 (the outcomes for the 99% confidence level are given in Table A.9 in the Appendix). The results for the Candelon et al. (2011) tests are presented for 5 moments, but we also obtained very similar results for 1, 2, 3, 4 and 6 moments.

We do not obtain fully satisfactory results for all portfolios for any of the models, but the outcomes depend heavily on the kind of assets and tests applied. The statistical test results do not differ sufficiently between the competing models to clearly indicate which is a better model.

In the second approach, we perform an economic evaluation of the models based on loss functions. We concentrate on firm loss functions. This approach emphasizes the role of the utility function of risk managers, who have to consider their firms' profitability, and therefore prefer smaller scaled VaR measures for efficient capital allocation. In order to assess whether the differences between loss functions are statistically significant, we apply the Diebold–Mariano test. The results for the 95% VaR forecasts are given in Table 13 (the outcomes for the 99% confidence level are presented in Table A.10 in the Appendix).

Table 12
Evaluation of 95% VaR forecasts: unconditional coverage and independence tests.

Statistic	DCC-GARCH		DCC-CARR		DCC-RGARCH	
	Value	P-value	Value	P-value	Value	P-value
Currency rates						
LR_{UC}	0.549	0.489	0.436	0.509	0.927	0.336
LR_{IND}	1.207	0.272	2.783	0.095	1.019	0.313
LR_{CC}	1.756	0.416	3.219	0.200	1.946	0.378
J_{UC}	0.417	0.532	0.557	0.494	0.772	0.353
J_{IND}	5.049	0.095	0.318	0.929	16.040	0.010
J_{CC}	7.944	0.082	0.736	0.905	37.134	0.010
Exchange-traded funds						
LR_{UC}	3.294	0.070	0.000	0.991	0.618	0.432
LR_{IND}	0.368	0.544	0.548	0.459	1.041	0.308
LR_{CC}	3.662	0.160	0.548	0.760	1.660	0.436
J_{UC}	3.288	0.063	0.010	0.924	0.478	0.486
J_{IND}	4.312	0.130	2.181	0.363	5.593	0.075
J_{CC}	10.308	0.048	2.192	0.567	6.879	0.105
Stocks						
LR_{UC}	2.202	0.139	20.416	0.000	4.869	0.027
LR_{IND}	0.002	0.968	1.433	0.231	0.251	0.616
LR_{CC}	2.203	0.332	21.850	0.000	5.120	0.077
J_{UC}	2.269	0.139	28.409	0.000	4.590	0.031
J_{IND}	3.307	0.202	69.453	0.002	6.289	0.060
J_{CC}	5.880	0.146	19970.690	0.000	11.226	0.036

The evaluation period is January 4, 2010, to December 30, 2016, LR_{UC} is the statistic for the Kupiec (1995) unconditional coverage test, LR_{IND} is the statistic for the Christoffersen (1998) independence test, LR_{CC} is the statistic for the Christoffersen (1998) conditional coverage test, J_{UC} is the statistic for the Candelon et al. (2011) unconditional coverage test, J_{IND} is the statistic for the Candelon et al. (2011) independence test for up to five lags, J_{CC} is the statistic for the Candelon et al. (2011) conditional coverage test with the number of moments fixed to 5, p-values for J_{UC}, J_{IND}, J_{CC} were corrected by Dufour's (2006) Monte Carlo procedure.

Table 13
Evaluation of 95% VaR forecasts: firm loss functions tests.

Loss function	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH vs. DCC-RGARCH	DCC-CARR vs. DCC-RGARCH
	Value of loss function × 10			P-value of DM test	
Currency rates					
FLF(STS)	0.371	0.394	**0.369**	0.158	0.014
FLF(C1)	5.991	6.091	**5.972**	0.033	0.000
FLF(C2)	3.079	3.204	**3.055**	0.127	0.000
FLF(C3)	7.189	7.203	**7.161**	0.153	0.032
Exchange-traded funds					
FLF(STS)	**1.500**	1.689	1.566	0.975	0.001
FLF(C1)	5.829	5.912	**5.804**	0.044	0.000
FLF(C2)	9.776	10.144	**9.627**	0.016	0.000
FLF(C3)	23.725	23.918	**22.428**	0.001	0.000
Stocks					
FLF(STS)	1.533	1.806	**1.480**	0.077	0.000
FLF(C1)	6.018	6.811	**5.956**	0.001	0.000
FLF(C2)	7.784	14.668	**7.472**	0.000	0.000
FLF(C3)	18.911	27.205	**18.612**	0.003	0.000

The evaluation period is January 4, 2010, to December 30, 2016, FLF(STS) is the loss function by Sarma et al. (2003), FLF(C1), FLF(C2), FLF(C3) are three loss functions by Caporin (2008). The lowest values of loss functions are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the DCC-GARCH and DCC-CARR models. A p-value lower than the significance level means that economic losses for the DCC-RGARCH model are lower than losses for a benchmark model (here DCC-GARCH or DCC-CARR).

For most of the considered loss functions, significantly more accurate VaR forecasts are constructed based on the DCC-RGARCH model than the DCC-GARCH or DCC-CARR models. This means that risk managers should prefer the DCC-RGARCH model for the estimation of their VaR forecasts. The results are very similar for both commonly employed confidence levels, 95% and 99%.

Table A.1
Evaluation of covariance forecasts based on DCC-EGARCH and DCC-GJR models: the MSE criterion.

Assets	DCC-GARCH	DCC-EGARCH	DCC-GJR	DCC-GARCH vs. DCC-EGARCH	DCC-GARCH vs. DCC-GJR
	MSE			P-value of DM test	
Currency rates					
EUR/USD-JPY/USD	**0.654**	0.756	0.755	0.953	0.946
EUR/USD-GBP/USD	**0.698**	0.939	0.923	0.809	0.797
JPY/USD-GBP/USD	**1.334**	2.033	2.052	0.812	0.819
Exchange-traded funds					
Oil-Natural Gas	**62.800**	64.495	64.842	0.997	1.000
Oil-Energy	**61.59**	67.994	67.968	0.999	0.995
Natural Gas-Energy	**30.734**	31.307	31.530	0.984	1.000
Stocks					
Amazon-Apple	**198.291**	198.909	200.043	0.682	0.792
Amazon-Goldman Sachs	77.756	76.526	**75.877**	0.086	0.067
Amazon-Google	**169.777**	175.503	175.194	0.926	0.914
Amazon-IBM	48.742	**47.569**	48.025	0.005	0.092
Apple-Goldman Sachs	**99.387**	103.087	102.637	0.989	0.989
Apple-Google	**267.998**	279.135	274.478	1.000	0.853
Apple-IBM	**148.393**	150.017	150.018	0.987	0.990
Goldman Sachs-Google	**66.695**	70.884	69.567	1.000	0.995
Goldman Sachs-IBM	**41.495**	41.751	41.582	0.789	0.654
Google-IBM	**56.133**	56.646	56.662	0.873	0.878

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of MSE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the benchmark the DCC-GARCH model. A p-value lower than the significance level means that the forecasts of covariance from the DCC-EGARCH or DCC-GJR models are more accurate than the forecasts from the DCC-GARCH model.

Table A.2
Evaluation of covariance forecasts based on DCC-EGARCH and DCC-GJR models: the MAE criterion.

Assets	DCC-GARCH	DCC-EGARCH	DCC-GJR	DCC-GARCH vs. DCC-EGARCH	DCC-GARCH vs. DCC-GJR
	MAE			P-value of DM test	
Currency rates					
EUR/USD-JPY/USD	**0.106**	0.113	0.111	1.000	1.000
EUR/USD-GBP/USD	**0.098**	0.107	0.100	1.000	0.801
JPY/USD-GBP/USD	**0.088**	0.090	0.089	0.754	0.639
Exchange-traded funds					
Oil-Natural Gas	**1.453**	1.466	1.468	0.992	1.000
Oil-Energy	**1.291**	1.297	1.302	0.650	0.734
Natural Gas-Energy	1.046	**1.043**	1.051	0.232	0.842
Stocks					
Amazon-Apple	1.299	**1.258**	1.289	0.000	0.134
Amazon-Goldman Sachs	1.209	1.134	**1.145**	0.000	0.000
Amazon-Google	1.506	1.333	**1.329**	0.000	0.000
Amazon-IBM	0.875	0.820	**0.829**	0.000	0.000
Apple-Goldman Sachs	**1.060**	1.115	1.158	1.000	1.000
Apple-Google	**1.157**	1.483	1.483	1.000	1.000
Apple-IBM	**0.820**	0.866	0.891	1.000	1.000
Goldman Sachs-Google	**1.093**	1.134	1.120	1.000	1.000
Goldman Sachs-IBM	0.841	0.828	**0.839**	0.005	0.355
Google-IBM	**0.742**	0.755	0.755	0.993	0.991

The realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of MAE are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the benchmark DCC-GARCH model. A p-value lower than the significance level means that the forecasts of covariance from the DCC-EGARCH or DCC-GJR models are more accurate than the forecasts from the DCC-GARCH model.

5. Conclusion

The DCC-GARCH model is one of the most popular multivariate volatility models, due to its simplicity and ease of estimation. However, its parameters are usually estimated based only on closing prices, even though high and low prices contain more information about volatility. In this study, we have proposed a new specification of the DCC model called the DCC-Range-GARCH

Table A.3
Evaluation of variance forecasts: the MSE and MAE criteria for realized variance calculated based on 5-min returns.

Assets	GARCH	CARR	RGARCH	GARCH	CARR	RGARCH
	MSE			MAE		
Currency rates						
EUR/USD	0.106	0.111	**0.093**	0.157	0.158	**0.147**
GBP/USD	0.721	1.025	**0.486**	0.158	0.149	**0.139**
JPY/USD	0.411	0.462	**0.317**	0.219	0.215	**0.194**
Exchange-traded funds						
Energy	7.695	8.086	**6.581**	1.138	1.232	**0.991**
Oil	**13.154**	18.126	14.142	2.119	2.365	**2.043**
Natural Gas	**20.715**	24.516	21.849	3.061	3.350	**3.038**
Stocks						
Amazon	199.106	**192.202**	194.809	3.941	**2.347**	3.393
Apple	**96.345**	97.437	102.881	2.504	3.315	**2.404**
Goldman Sachs	18.464	17.595	**17.450**	1.901	1.953	**1.806**
Google	68.873	**59.054**	60.641	2.174	**1.898**	1.910
IBM	15.504	**14.592**	14.935	1.180	**1.078**	1.099

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 5-min returns. The lowest values of MSE and MAE are marked in bold.

Table A.4
Evaluation of variance forecasts: the LINEX function with a = −1.

Assets	GARCH	CARR	RGARCH	GARCH	CARR	RGARCH
	Full sample			Excluding 1% of upper outliers		
Currency rates						
EUR/USD	0.301	0.407	**0.210**	0.019	0.021	**0.018**
GBP/USD	1015.023	1.500e+04	**59.152**	0.013	0.015	**0.012**
JPY/USD	5.190e+07	8.880e+13	**1.012e+05**	0.036	0.033	**0.029**
Exchange-traded funds						
Energy	**4.842**	275.893	12.034	1.410	2.584	**0.872**
Oil	**7784.863**	3.710e+08	1.185e+04	77.673	148.632	**54.823**
Natural Gas	2379.619	3005.722	**2379.228**	**34.479**	1486.421	62.574
Stocks						
Amazon	698.416	**32.762**	62.284	2.094	1.801	**1.483**
Apple	**328.116**	4.120e+13	3.530e+16	13.742	**5.866**	7.255
Goldman Sachs	7.469e+05	**5.362e+04**	7.375e+04	**2.044**	3.335	2.046
Google	1.120e+18	52.643	**29.517**	1.625	1.297	**1.042**
IBM	1.396	**0.880**	1.176	0.375	0.334	**0.324**

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The lowest values of the LINEX function are marked in bold.

Table A.5
Evaluation of variance forecasts: the LINEX function with a = 1.

Assets	GARCH	CARR	RGARCH	GARCH	CARR	RGARCH
	Full sample			Excluding 1% of upper outliers		
Currency rates						
EUR/USD	0.301	0.407	**0.210**	2,206e−02	2,283e−02	**2,031e−02**
GBP/USD	5,190e+07	8,881e+13	**1,012e+05**	1,504e−02	1,504e−02	**1,285e−02**
JPY/USD	1,015e+03	1,500e+04	**5,915e+01**	4,273e−02	4,170e−02	**3,617e−02**
Exchange-traded funds						
Energy	7,661e+24	6,502e+25	**3,173e+24**	101.585	26.510	**24.855**
Oil	**7,990e+16**	2,846e+17	1,817e+17	**2,463e+04**	8,599e+04	6,073e+04
Natural Gas	**1,206e+21**	1,793e+23	1,449e+23	2732.795	4001.910	**2715.055**

(*continued on next page*)

Table A.5 (*continued*).

Assets	GARCH	CARR	RGARCH	GARCH	CARR	RGARCH
	Full sample			Excluding 1% of upper outliers		
Stocks						
Amazon	9,198e+135	4,589e+138	**3,725e+135**	2,250e+11	**1,790e+10**	1,416e+11
Apple	4,573e+128	**1,856e+95**	1,600e+97	**2,606e+07**	8,760e+07	1,921e+08
Goldman Sachs	8,917e+15	1,104e+16	**2,803e+14**	405.994	102.403	**243.711**
Google	6,214e+67	**4,470e+67**	8,280e+67	**3,076e+03**	3,737e+04	4,420e+04
IBM	**7,689e+22**	1,113e+30	8,391e+29	51.523	50.826	**41.546**

The evaluation period is January 4, 2010, to December 30, 2016, the realized variance is used as a proxy of variance and estimated as the sum of squares of 15-min returns. The lowest values of the LINEX function are marked in bold.

Table A.6
Evaluation of covariance forecasts: the MSE and MAE criteria for realized covariance calculated based on 5-min returns.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH	DCC-CARR	DCC-RGARCH
	MSE			MAE		
Currency rates						
EUR/USD-JPY/USD	0.041	0.048	**0.035**	0.096	0.101	**0.093**
EUR/USD-GBP/USD	0.060	0.082	**0.043**	0.094	0.093	**0.087**
JPY/USD-GBP/USD	0.071	0.122	**0.055**	0.080	0.083	**0.078**
Exchange-traded funds						
Oil-Natural Gas	**5.940**	6.146	5.979	1.392	1.417	**1.383**
Oil-Energy	5.714	9.734	**5.640**	1.247	1.451	**1.161**
Natural Gas-Energy	2.923	8.124	**2.869**	1.002	2.005	**0.976**
Stocks						
Amazon-Apple	15.597	14.959	**13.280**	1.202	1.059	**1.051**
Amazon-Goldman Sachs	6.657	6.230	**5.994**	1.140	1.051	**1.034**
Amazon-Google	14.906	13.894	**12.957**	1.441	1.222	**1.193**
Amazon-IBM	4.315	4.038	**3.733**	0.826	0.739	**0.714**
Apple-Goldman Sachs	7.480	7.534	**6.677**	0.993	0.974	**0.917**
Apple-Google	17.530	17.337	**14.424**	1.075	0.990	**0.942**
Apple-IBM	11.066	10.786	**8.328**	0.781	0.728	**0.680**
Goldman Sachs-Google	5.318	4.966	**4.360**	1.027	0.980	**0.899**
Goldman Sachs-IBM	3.404	3.272	**2.999**	0.791	0.770	**0.709**
Google-IBM	4.630	4.360	**3.827**	0.705	0.649	**0.614**

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 5-min returns. The lowest values of MSE and MAE are marked in bold.

Table A.7
Evaluation of covariance forecasts: the LINEX function with $a = -1$.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH	DCC-CARR	DCC-RGARCH
	Full sample			Excluding 1% of upper outliers		
Currency rates						
EUR/USD-JPY/USD	0.149	0.338	**0.076**	8,744e−03	6,985e−03	**6,085e−03**
EUR/USD-GBP/USD	0.020	0.019	**0.021**	8,755e−03	9,538e−03	**8,503e−03**
JPY/USD-GBP/USD	334.772	3,301e+04	**11.909**	5,310e−03	5,535e−03	**5,004e−03**
Exchange-traded funds						
Oil-Natural Gas	2,590e+06	**1,771e+06**	2,001e+06	8.122	8.769	**7.664**
Oil-Energy	2.190	**1.072**	1.585	1.368	0.870	**1.072**
Natural Gas-Energy	1716.904	3,315e+05	**1000.640**	2.278	107.617	**2.016**
Stocks						
Amazon-Apple	7075.788	**2292.080**	3985.507	0.722	**0.487**	0.518
Amazon-Goldman Sachs	5106.527	**2016.502**	5924.394	0.635	**0.556**	0.565
Amazon-Google	1,280e+08	1,570e+08	**1,170e+08**	1.008	0.670	**0.665**
Amazon-IBM	**51.210**	53.546	74.122	0.313	0.234	**0.230**

(*continued on next page*)

Table A.7 (*continued*).

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH	DCC-CARR	DCC-RGARCH
	Full sample			Excluding 1% of upper outliers		
Apple-Goldman Sachs	2,750e+05	**2,350e+05**	2,446e+05	0.496	0.472	**0.428**
Apple-Google	**13.275**	111.506	1104.156	0.641	0.521	**0.493**
Apple-IBM	**1.060**	1.310	26.865	**0.258**	0.578	0.545
Goldman Sachs-Google	3,820e+08	6,970e+08	**3,480e+08**	0.582	0.578	**0.545**
Goldman Sachs-IBM	706.926	**523.911**	869.393	0.270	0.252	**0.235**
Google-IBM	1,367e+06	6,258e+05	**1,188e+06**	0.190	0.168	**0.162**

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of the LINEX function are marked in bold.

Table A.8
Evaluation of covariance forecasts: the LINEX function with a = 1.

Assets	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH	DCC-CARR	DCC-RGARCH
	Full sample			Excluding 1% of upper outliers		
Currency rates						
EUR/USD-JPY/USD	0.082	0.094	**0.053**	9,379e–03	**1,038e–02**	9,186e–03
EUR/USD-GBP/USD	2.832	33.415	**0.322**	7,551e–03	7,295e–03	**6,649e–03**
JPY/USD-GBP/USD	0.043	0.035	**0.100**	5,626e–03	5,913e–03	**5,303e–03**
Exchange-traded funds						
Oil-Natural Gas	**4,393e+06**	7,413e+06	8,555e+06	**7.022**	8.507	8.077
Oil-Energy	8,996e+12	2,484e+14	**7,500e+12**	68.685	1194.447	**54.811**
Natural Gas-Energy	3,420e+04	**4,503e+03**	4,924e+04	**1.060**	1.298	1.128
Stocks						
Amazon-Apple	**8,926e+54**	2,283e+55	1,786e+55	20.741	**12.284**	15.554
Amazon-Goldman Sachs	**3,554e+23**	6,687e+23	4,391e+23	16.859	**14.397**	23.327
Amazon-Google	**1,992e+33**	1,023e+34	6,252e+33	1217.696	**342.058**	744.859
Amazon-IBM	4,911e+16	5,291e+16	**4,548e+16**	5.500	**3.768**	5.002
Apple-Goldman Sachs	**1,213e+38**	4,439e+38	3,180e+38	3.709	**2.987**	3.258
Apple-Google	8,100e+11	**2,099e+10**	3,360e+10	976.587	1301.767	**935.519**
Apple-IBM	1,485e+07	**1,252e+07**	1,533e+07	5.479	**3.032**	3.542
Goldman Sachs-Google	9,756e+22	6,748e+23	**1,957e+09**	14.760	**6.429**	10.240
Goldman Sachs-IBM	**6,722e+12**	9,757e+12	7,831e+12	1.973	**1.700**	1.848
Google-IBM	7,565e+22	7,253e+22	**2,276e+16**	1.404	1.489	**1.322**

The evaluation period is January 4, 2010, to December 30, 2016, the realized covariance is used as a proxy of covariance and estimated as the sum of products of 15-min returns. The lowest values of the LINEX function are marked in bold.

model, which is a combination of the DCC model by Engle (2002) and the Range-GARCH model by Molnár (2016). The DCC-Range-GARCH model is very similar to the DCC model by Engle but it is based on a much more efficient volatility estimator formulated on the daily range, the log-difference between the high and low prices.

We have compared our DCC-Range-GARCH model to the DCC-GARCH model by Engle (2002) and the DCC-CARR model by Chou et al. (2009). All these three models are very similar in their correlation part, but differ in their specification for conditional variances. The DCC-GARCH model is based on the GARCH model, the DCC-Range-GARCH model is formulated on the Range-GARCH model and the DCC-CARR model is based on the CARR model by Chou (2005). We have evaluated these models on three data sets: currencies, exchange traded funds and stocks.

The univariate range-based models, CARR and Range-GARCH, had not been previously compared. We therefore first compare forecasting accuracy of these models. We found that the CARR model is outperformed by the Range-GARCH model. Surprisingly, the CARR model is often inferior even to the standard GARCH model, whereas the Range-GARCH model outperforms it in most cases. We then turned our attention to multivariate models and the comparison of covariance forecasts, which were the main focus of this paper. We found that the proposed DCC-Range-GARCH model is superior not only to the standard DCC-GARCH model but also to the DCC-CARR model.

Our results illustrate that the use of range data in the DCC model can improve the estimation of covariances of returns and increase the accuracy of covariance and VaR forecasts based on this model, compared with using closing prices only. Moreover, the way the range is utilized matters, as our proposed model outperforms the DCC-CARR model, which is also based on range. Therefore, other multivariate range-based volatility models such as the double smooth transition conditional correlation CARR model by Chou and Cai (2009), the range-based copula models by Chiang and Wang (2011) and Wu and Liang (2011) and the range-based regime-switching dynamic conditional correlation model by Su and Wu (2014) would probably also benefit from using the Range-GARCH model in place of the CARR specification.

Table A.9
Evaluation of 99% VaR forecasts: unconditional coverage and independence tests.

Statistic	DCC-GARCH		DCC-CARR		DCC-RGARCH	
	Value	P-value	Value	P-value	Value	P-value
Currency rates						
LR_{UC}	0.270	0.603	4.621	0.032	0.076	0.782
LR_{IND}	0.285	0.594	0.567	0.452	0.321	0.571
LR_{CC}	0.555	0.758	5.188	0.075	0.398	0.820
J_{UC}	0.080	0.746	4.093	0.042	0.001	0.911
J_{IND}	3.275	0.130	3.308	0.129	1.858	0.310
J_{CC}	3.728	0.253	6.720	0.087	2.012	0.520
Exchange-traded funds						
LR_{UC}	2.039	0.153	0.008	0.928	0.414	0.520
LR_{IND}	0.165	0.685	0.372	0.542	0.258	0.612
LR_{CC}	2.204	0.332	0.380	0.827	0.672	0.715
J_{UC}	1.658	0.206	0.101	0.739	0.166	0.652
J_{IND}	2.770	0.181	0.508	0.788	0.323	0.883
J_{CC}	16.695	0.019	0.545	0.990	0.559	0.915
Stocks						
LR_{UC}	26.337	0.000	1.514	0.219	28.175	0.000
LR_{IND}	0.721	0.396	1.050	0.306	0.631	0.427
LR_{CC}	27.059	0.000	2.564	0.278	28.806	0.000
J_{UC}	15.976	0.001	1.713	0.159	16.827	0.001
J_{IND}	8.888	0.013	83.216	0.000	5.052	0.055
J_{CC}	27.499	0.008	31.612	0.008	26.992	0.007

The evaluation period is January 4, 2010, to December 30, 2016, LR_{UC} is the statistic for the Kupiec (1995) unconditional coverage test, LR_{IND} is the statistic for the Christoffersen (1998) independence test, LR_{CC} is the statistic for the Christoffersen (1998) conditional coverage test, J_{UC} is the statistic for the Candelon et al. (2011) unconditional coverage test, J_{IND} is the statistic for the Candelon et al. (2011) independence test for up to five lags, J_{CC} is the statistic for the Candelon et al. (2011) conditional coverage test with the number of moments fixed to 5, p-values for J_{UC}, J_{IND}, J_{CC} were corrected by Dufour's (2006) Monte Carlo procedure.

Table A.10
Evaluation of 99% VaR forecasts: firm loss functions tests.

Loss function	DCC-GARCH	DCC-CARR	DCC-RGARCH	DCC-GARCH vs. DCC-RGARCH	DCC-CARR vs. DCC-RGARCH
	Value of loss function $\times 10^1$			P-value of DM test	
Currency rates					
FLF(STS)	0.506	0.514	**0.502**	0.066	0.021
FLF(C1)	6.913	6.927	**6.900**	0.080	0.012
FLF(C2)	5.322	5.405	**5.280**	0.110	0.000
FLF(C3)	9.981	9.988	**9.939**	0.149	0.076
Exchange-traded funds					
FLF(STS)	**1.744**	1.849	1.751	0.599	0.001
FLF(C1)	6.865	6.862	**6.800**	0.000	0.000
FLF(C2)	17.239	17.239	**16.764**	0.000	0.000
FLF(C3)	33.010	33.260	**32.392**	0.000	0.000
Stocks					
FLF(STS)	1.474	2.029	**1.416**	0.014	0.000
FLF(C1)	6.811	7.535	**6.734**	0.000	0.000
FLF(C2)	13.532	23.767	**12.907**	0.000	0.000
FLF(C3)	26.455	37.688	**25.790**	0.000	0.000

The evaluation period is January 4, 2010, to December 30, 2016, FLF(STS) is the loss function by Sarma et al. (2003), FLF(C1), FLF(C2), FLF(C3) are three loss functions by Caporin (2008). The lowest values of loss functions are marked in bold. The p-values of the Diebold–Mariano test are presented for pairs of models with respect to the two benchmarks: the DCC-GARCH and DCC-CARR models. A p-value lower than the significance level means that economic losses for the DCC-RGARCH model are lower than those for a benchmark model (here DCC-GARCH or DCC-CARR).

Appendix

See Tables A.1–A.10.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.
Bernardi, M., Catania, L., 2018. Portfolio optimisation under flexible dynamic dependence modelling. J. Empir. Financ. 48, 1–18.
Bouri, E., Molnár, P., Azzi, G., Roubaud, D., Hagfors, L.I., 2017. On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? Finance Res. Lett. 20, 192–198.
Brandt, M., Jones, C., 2006. Volatility forecasting with range-based EGARCH models. J. Bus. Econom. Statist. 24, 470–486.
Candelon, B., Colletaz, G., Hurlin, C., Tokpavi, S., 2011. Backtesting value-at-risk: A GMM duration-based test. J. Financ. Econ. 9 (2), 314–343.
Caporin, M., 2008. Evaluating value-at-risk measures in the presence of long memory conditional volatility. J. Risk 10 (3), 79–110.
Chiang, M.H., Wang, L.M., 2011. Volatility contagion: A range-based volatility approach. J. Econometrics 165 (2), 175–189.
Chou, R.Y., 2005. Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. J. Money Credit Bank. 37 (3), 561–582.
Chou, R.Y., Cai, Y., 2009. Range-based multivariate volatility model with double smooth transition in conditional correlation. Glob. Financ. J. 20 (2), 137–152.
Chou, R.Y., Wu, C.C., Liu, N., 2009. Forecasting time-varying covariance with a range-based dynamic conditional correlation model. Rev. Quant. Financ. Account. 33, 327–345.
Christoffersen, P.F., 1998. Evaluating interval forecasts. Internat. Econom. Rev. 39 (4), 841–862.
Clarke, K.A., 2007. A simple distribution-free test for nonnested model selection. Polit. Anal. 15 (3), 347–363.
Dark, J., 2018. Multivariate models with long memory dependence in conditional correlation and volatility. J. Empir. Financ. 48, 162–180.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econom. Statist. 13, 253–263.
Dufour, J.-M., 2006. Monte Carlo tests with nuisance parameters: A general approach to finite sample inference and nonstandard asymptotics. J. Econometrics 127 (2), 443–477.
Engle, R.F., 2002. Dynamic conditional correlation – a simple class of multivariate GARCH models. J. Bus. Econom. Statist. 20, 339–350.
Engle, R.F., Russell, J., 1998. Autoregressive conditional duration: A new model for irregularly spaced transaction data. Econometrica 66 (5), 1127–1162.
Engle, R.F., Sheppard, K., Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, NBER Working Paper No. 8554, 2001.
Fiszeder, P., 2018. Low and high prices can improve covariance forecasts: The evidence based on currency rates. J. Forecast. 37 (6), 641–649.
Fiszeder, P., Perczak, G., 2013. A new look at variance estimation based on low, high and closing prices taking into account the drift. Stat. Neerl. 67 (4), 456–481.
Fiszeder, P., Perczak, G., 2016. Low and high prices can improve volatility forecasts during the turmoil period. Int. J. Forecast. 32 (2), 398–410.
Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53 (1), 67–78.
Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48 (5), 1779–1801.
Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: a joint model for returns and realized measures of volatility. J. Appl. Econom. 27 (6), 877–906.
Harvey, D.I., Leybourne, S.J., Newbold, P., 1997. Testing the equality of prediction mean squared errors. Int. J. Forecast. 13 (2), 281–291.
Heaney, R., Sriananthakumar, S., 2012. Time-varying correlation between stock market returns and real estate returns. J. Empir. Financ. 19 (4), 583–594.
Karanasos, M., Menla Ali, F., Margaronis, Z., Nath, R., 2018. Modelling time varying volatility spillovers and conditional correlations across commodity metal futures. Int. Rev. Financ. Anal. 57, 246–256.
Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. J. Derivatives 3, 73–84.
Lehkonen, H., Heimonen, K., 2014. Timescale-dependent stock market comovement: BRICs vs. Developed markets. J. Empir. Financ. 28, 90–103.
Lin, E.M.H., Chen, C.W.S., Gerlach, R., 2012. Forecasting volatility with asymmetric smooth transition dynamic range models. Int. J. Forecast. 28 (2), 384–399.
Molnár, P., 2012. Properties of range-based volatility estimators. Int. Rev. Financ. Anal. 23, 20–29.
Molnár, P., 2016. High-low range in GARCH models of stock return volatility. Appl. Econ. 48 (51), 4977–4991.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59 (2), 347–370.
Nelson, D.B., Cao, C.Q., 1992. Inequality constraints in the univariate GARCH model. J. Bus. Econom. Statist. 10, 229–235.
Pagan, A.R., Schwert, G.W., 1990. Alternative models for conditional stock volatility. J. Econometrics 45, 267–290.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53 (1), 61–65.
Poon, S.-H., Granger, C., 2003. Forecasting volatility in financial markets: A review. J. Econ. Lit. 41, 478–539.
Rivers, D., Vuong, Q., 2002. Model selection tests for nonlinear dynamic models. Econom. J. 5 (1), 1–39.
Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. Ann. Appl. Probab. 1 (4), 504–512.
Sarma, M., Thomas, S., Shah, A., 2003. Selection of value-at-risk models. J. Forecast. 22 (4), 337–358.
Su, Y.K., Wu, C.C., 2014. A new range-based regime-switching dynamic conditional correlation model for minimum-variance hedging. J. Math. Finance 4 (3), 207–219.
Tse, Y.K., Tsui, A.K.C., 2002. A multivariate GARCH model with time-varying correlations. J. Bus. Econom. Statist. 20, 351–362.
Wu, C.C., Liang, S., 2011. The economic value of range-based covariance between stock and bond returns with dynamic copulas. J. Empir. Financ. 18 (4), 711–727.

Chapter 5

Central bank announcements and realized volatility of stock markets in G7 countries

J. Int. Financ. Markets Inst. Money 58 (2019) 117–135

Contents lists available at ScienceDirect

Journal of International Financial Markets, Institutions & Money

journal homepage: www.elsevier.com/locate/intfin

Central bank announcements and realized volatility of stock markets in G7 countries ☆

Štefan Lyócsa [a,b,*], Peter Molnár [c,d], Tomáš Plíhal [b]

[a] University of Presov, Faculty of Management, Presov, Slovakia
[b] Masaryk University, Faculty of Economics and Administration, Brno, Czech Republic
[c] University of Stavanger, UiS Business School, Stavanger, Norway
[d] University of Economics, Faculty of Finance and Accounting, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 8 December 2017
Accepted 28 September 2018
Available online 29 September 2018

JEL classification:
C22
E44
G14
G15

Keywords:
Target interest rate
Realized volatility
Central banks
High-frequency data
Monetary policy

ABSTRACT

We investigate the impact of monetary policy announcements on stock market volatility in the U.S., Canada, Japan, the U.K., Germany, France and Italy during the 2006–2016 period. More specifically, we study the impact of policy rate and quantitative easing announcements of domestic and foreign central banks on realized volatility before, during, and after the event. We find that on the day of an interest rate announcement of the domestic central bank, volatility increases in a manner that is both statistically and economically significant. We also find a decline in volatility five days after an interest rate announcement across all countries in our sample. We further find that quantitative easing announcements have no impact on stock market volatility not only at but also five days before and five days after the announcement date.

1. Introduction

Participants in financial markets closely follow announcements of macroeconomic news. The macroeconomic situation is directly relevant to the valuation of financial assets; therefore, the news often tells market participants whether they should update prices upwards or downwards. Due to the major importance of this topic, a large body of literature assesses the impact of macroeconomic news announcement on financial markets. Stock market reaction was studied by Flannery and Protopapadakis (2002), Ehrmann and Fratzscher (2004), Bernanke and Kuttner (2005), Bekaert and Engstrom (2010), and Hussain (2011). The response of foreign exchange markets was investigated by Almeida et al. (1998), Andersen et al. (2003), Ehrmann and Fratzscher (2005), Evans and Speight (2010a), Omrane and Hafner (2015), Petralias and Dellaportas (2015), and El Ouadghiri and Uctum (2016); see Neely and Dey (2010) for a review. The impact of macroeconomic announcements on government bond markets has been investigated by Fleming and Remolona (1997, 1999), Christie-David et al.

☆ Lyócsa appreciates the support provided by the Slovak Grant Agency under Grant No. 1/0406/17. Plíhal and Lyócsa also appreciate the support provided by the Czech Grant Agency under Grant No. 18-05829S.
* Corresponding author.
E-mail addresses: stefan.lyocsa@econ.muni.cz (Š. Lyócsa), tomas.plihal@econ.muni.cz (T. Plíhal).

https://doi.org/10.1016/j.intfin.2018.09.010
1042-4431/

(2002), Balduzzi et al. (2001), Gürkaynak et al. (2005), Beechey and Wright (2009), and El Ouadghiri et al. (2016), whereas other authors have studied several classes of assets, see Boyd et al. (2005), Faust et al. (2007), Bartolini et al. (2008).

Central bank announcements are one of the most important types of macroeconomic announcements. Market participants closely focus on monetary policy decisions; see Thorbecke (1997) and Thornton (1998). Since an understanding of volatility is important for market participants, the impact of news announcements on volatility has been investigated. Harvey and Huang (1991) and Ederington and Lee (1993) find that volatility increases around macroeconomic announcements. This topic was further studied by Dominguez (1998), Nikkinen and Sahlström (2001), Bauwens et al. (2005), Dominguez (2006), Nikkinen et al. (2006), Beine et al. (2009) and many others.

In this study we give answer to the question of how volatility on equity markets reacts to central banks' key scheduled monetary policy announcements. If market volatility decreases, we say that news announcement has a stabilizing impact on equity markets and vice versa. As market's response is likely to vary before, at or after the news announcement, we study all three cases separately. The previous literature does not offer a satisfactory answer to our question. For example, the previous literature documents an increase in volatility around news announcements, this cannot be interpreted as finding that announcements have a destabilizing impact on financial markets. The information content of news is usually high; therefore, it is only natural that volatility is high around earnings announcements. For example, Chae (2005) argues that if the information content of news is high, it will necessarily increase trading volume (and thus volatility) in the market because market participants adjust their views based on the news provided. Thus, the consensus of market participants increases on the day of the news announcement.

One way to evaluate whether news announcements have a stabilizing or destabilizing impact on financial markets is to investigate whether volatility increases or decreases several days after the announcement. This question has been addressed before; however, we argue that it was not answered satisfactorily for two reasons.

First, this question was originally addressed within the generalized autoregressive conditional heteroscedasticity (GARCH) framework based on daily data; see Bomfim (2003), Kim et al. (2004) and Bauwens et al. (2005). However, the question whether volatility increased on a particular day can be answered more accurately with help of more precise volatility data, for example from realized volatility calculated from high-frequency data.

Second, researchers began to utilize implied volatility calculated from option prices, see Ederington and Lee (1993), Nikkinen and Sahlstrőm (2004), Äijö (2008), Aktas (2011), Füss et al. (2011), Jiang et al. (2012), Marshall et al. (2012) and Krieger et al. (2015). Studies based on implied volatility usually find that volatility decreases after the announcement of scheduled news.

We instead use realized volatility calculated from high-frequency data for the 2006–2016 period. High-frequency data have previously been used to evaluate the impact of news announcements but usually have been used to study market reaction immediately before and immediately after an announcement; see Balduzzi et al. (2001) and Evans and Speight (2010b). These studies typically find that the direct impact of news announcements is concentrated in a very short time window around the announcement. We study whether stock market volatility increases or decreases 5 days before, the day of and 5 days after monetary policy announcements by central banks in G7 countries. Our contribution to the literature is threefold. *First*, most of the papers studying the impact of monetary policy announcements on financial markets study foreign exchange markets, not stock markets. However, interest rates have a profound relationship to stock markets (e.g. Sweeney and Warga, 1986; Akhtaruzzaman et al., 2014). For example, given the present value cash flow discounting model to stock valuation, the relationship between interest rates and stock markets should be negative. Monetary policy announcements are therefore likely to have broad impact on stocks markets and we contribute to this literature. *Second*, we use a multi-country sample, five central banks and their impact on the volatility of eight stock market indices. Moreover, we control for the news announcements related to interest rate and quantitative easing of foreign central banks as well. This way we control for the possible interconnectedness and/or coordination of monetary policy[1]. *Third*, the literature in this field is based on implied volatility, and there are large differences in the behavior of implied and realized volatility around scheduled announcements, as we explain in the following section. By using realized instead of the implied volatility, our approach offers a new and different insights into the role of monetary policy announcements on the stock market volatility. Employing realized volatility allows us to take a look at the role of news announcement on the actual (realized) price variation of stock markets not on the expectations.

Market level analysis reveals that realized volatility of equity markets is increased on the day when the domestic central bank announces its policy toward the target interest rate and declines during the subsequent five days. Evidence from dynamic common correlated effect panel models shows that announcements of foreign central banks related to target interest rates tend to increase market volatility (5 days) before and at the news announcements and lead to a mild decline (5 days) after the news announcement. We control for many factors, including: movements on the foreign exchange and money markets, volatility persistence, day-of-the-week effects, uncertainty about the outcome of a policy meeting, the surprise effect of a policy announcement, the effect of quantitative easing announcements and, importantly, the news released by the foreign central banks of the remaining markets in our sample.

[1] A leading example is the joint announcement of interest rate cuts on 8th October 2008 by FED, ECB, Bank of Canada, Bank of England, Swiss National Bank, and The People's Bank of China.

The remainder of this paper is organized as follows. In Section 2 we briefly discuss differences between implied and realized volatility in a context of an event study. Section 3 describes our data, and the econometric methodology is presented in Section 4, which also includes a detailed description of the explanatory and control variables. In Section 5, we present our results, and Section 6 concludes.

2. Implied and realized volatility in an event study

Implied and realized volatility might behave similarly over the longer time period but more differently around certain events that last for only several of days. First of all, realized volatility is an estimate of price variation over a particular period, say a day. Implied volatility typically used in the literature is calculated from option prices with maturity close to the next 30 calendar days. Assume that everybody (correctly) expect that stock market volatility on one particular day will be unusually high. In this case, realized volatility before that day will not pick up that information (as it is historical) and on that particular day it will be unusually high. However, implied volatility before this particular day will already be higher, but presumably not that much, as the maturity of underlying options also includes other 29 calendar days. Therefore, changes in realized volatility will be more pronounced than changes in implied volatility. Moreover, this 30-day averaging has an impact on how to interpret changes in implied volatility, as we discuss later.

Second, implied volatility is not simply an expected volatility. Implied volatility is usually significantly higher than what volatility turns out to be ex post (see e.g. Birkelund et al., 2015). Implied volatility can be considered as a sum of expected volatility and volatility risk premium. Therefore, it is difficult to disentangle whether change in implied volatility should be attributed to change in expected volatility, or change in volatility risk premium.

The different behavior of implied volatility and realized volatility around the event day is illustrated in Fig. 1, in which we plot daily percentage changes in realized and implied volatility on the S&P 500 (*VIX*) around an interest rate announcement for the U.S. stock market. As can be observed, on the event date, realized volatility increases, whereas implied volatility decreases. This result is likely caused by the fact that realized volatility is calculated during the entire trading period of the day, whereas implied volatility is a closing value from the end of the day, and it therefore reflects the market expectations for the next 30 days that excludes news announcements made earlier during or before trading hours of that particular day. If news announcement day is a day with higher volatility, implied volatility the day before the news announcement should be higher than at the news announcement day. In Fig. 1, our data support this reasoning. In other words, there is a tendency for implied volatility to decrease after the highly volatile announcement day is excluded from the 30-day time window.

As the news might also contain information that in the future, price variation will be low or high, changes in implied volatility might consist of two parts: decline due to exclusion of volatile day from a 30-day maturity window and increase/decrease in the future levels of price variation due to new information. Therefore, simply observing decrease in implied volatility does not tell us that an announcement had a calming effect on financial markets.

Instead, we utilize realized volatility calculated from high-frequency data. Consequently, we have an estimate of price variation for each individual day, and we can easily evaluate whether volatility increases or decreases before, during and after a monetary policy announcement. Altogether, realized volatility convey different type of information than implied volatility, and therefore our study complement existing literature.

3. Data

Our sample begins in January 2006 and ends in November 2016. We cover the G7 countries, which are a group of the seven major advanced economies according to the International Monetary Fund. This group consists of Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. Our data can be divided into three main categories: stock market indices, news announcements from central banks, and exchange rate and interest rate data. Because France, Germany and Italy are part of the European monetary union, their stock markets share the same set of news announcements from the European Central Bank (ECB).

3.1. Stock market indices

We study the effect of the news announcements of central banks on seven stock market indices: S&P 500 (United States), FTSE 100 (United Kingdom), TSX (Canada), NIKKEI 225 (Japan), STOXX 50 (Europe), DAX (Germany), CAC (France), and MIB (Italy). The measures of volatility are precalculated for the given sample period and are downloaded directly from the Oxford-Man Institute of Quantitative Finance Realized Library[2].

[2] http://realized.oxford-man.ox.ac.uk/data/download2

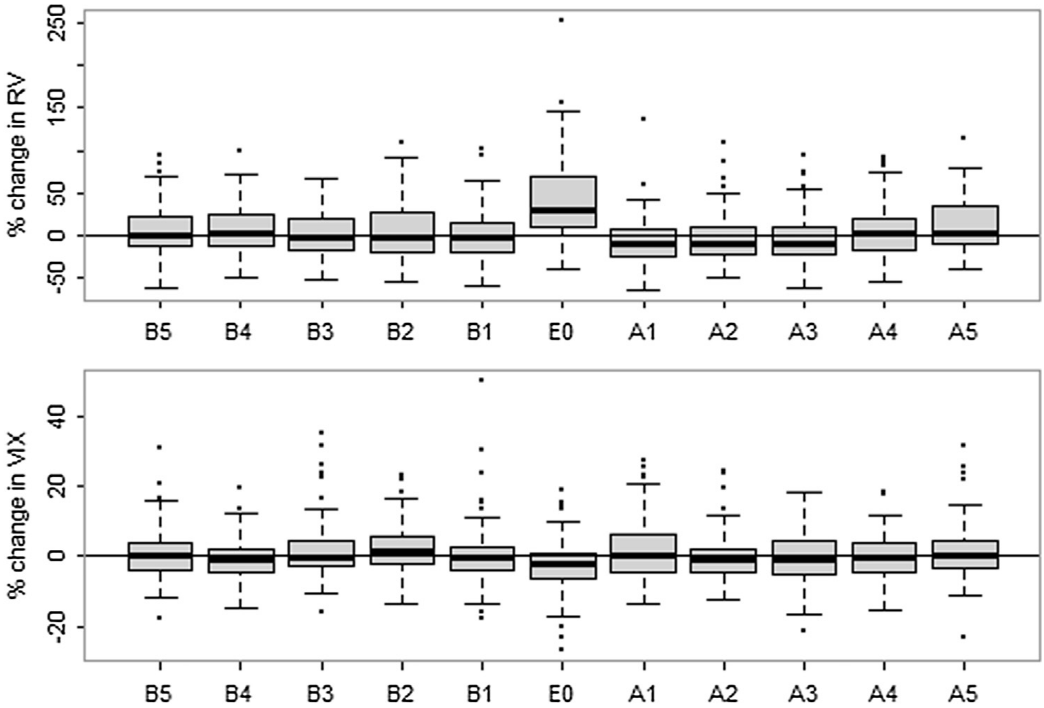

Fig. 1. Target rate announcements and the realized and implied volatilities of the S&P 500. Notes: The values on the y-axis are percentage changes in either the realized variance (RV) or the value of the VIX index. B5, B4, . . ., B1 denote 5, 4, . . ., 1 day before the target rate announcement, respectively. E0 is the day of the target rate announcement. A1, A2, . . ., A5 denote 1, 2, . . ., 5 days after the target rate announcement, respectively.

3.2. News announcements

We focus on the most important macroeconomic news announcements related to the central banks: target interest rates and quantitative easing. The target interest rate is an essential part of monetary policy strategy. Many central banks set a target interest rate in an effort to influence short-term interest rates. The data were collected from Bloomberg and are described in Table 1.

For Canada and the United Kingdom, the data about the target interest rate announcement required no adjustments and are easily available for the entire observed period. The European Central Bank reports the ECB Main Refinancing Rate, the ECB Deposit Facility Rate, and the ECB Marginal Lending Facility. We refer only to the first interest rate noted because it is the most important rate, which unlike the other two rates, was reported during the entire observed period. Moreover, all these rates are announced at the same time. In the United States, the federal funds rate is targeted by the Federal Reserve's Federal Open Market Committee (FOMC). In December 2008, the target interest rate was replaced by a target range. We use the FOMC Rate Decision (Upper Bound) because Bloomberg also uses this label for the interest rate reported before the introduction of the target range. Therefore, this variable covers the entire observed period. Conversely, the variable representing the lower bound of the interest rate was introduced only after December 2008.

The situation in Japan is slightly more complex. On April 4, 2013, the Bank of Japan shifted its monetary policy focus to a targeted monetary base via Japanese government bond purchases.[3] On January 29, 2016, the Bank of Japan employed a new approach to its monetary policy known as "Quantitative and Qualitative Monetary Easing with a Negative Interest Rate."[4] As a result of these two major changes in monetary policy, there is a gap in our data from April 4, 2013, to January 29, 2016, because no interest rate news was announced. During this period, the Bank of Japan targeted a monetary base instead of the interest rate.

The second category of news announcements is related to quantitative easing. This category is an unconventional monetary policy used by central banks to stimulate their economies when conventional monetary policy is no longer effective. This category is often used in situations in which the central bank's interest rates are currently near zero, and there is little room for a greater decline. Quantitative easing usually consists of the purchases of long-term financial assets from banks and other financial institutions. This policy creates additional money in the economy and should lower long-term interest rates.

As days of quantitative easing, we select the days when the introduction of quantitative easing or any change in this policy were announced. During our sample period all the selected countries except Canada have experiences with quantitative easing. The relevant data were collected manually from the official sites of the central banks.[5]

[3] http://www.boj.or.jp/en/announcements/release_2013/k130404a.pdf3

[4] http://www.boj.or.jp/en/announcements/release_2016/k160129a.pdf4

[5] http://www.boj.or.jp/en/mopo/mpmsche_minu/index.htm/#p01https://www.ecb.europa.eu/press/govcdec/html/index.en.htmlhttps://www.federalreserve.gov/newsevents/press/monetary/2016monetary.htmhttp://www.bankofengland.co.uk/monetarypolicy/Pages/decisions.aspx.

Table 1
Interest rate news announcements. Source: https://www.bloomberg.com.

Country	Event name from Bloomberg	Ticker
Canada	Bank of Canada Rate Decision	CABROVER Index
Eurozone	ECB Main Refinancing Rate	EURR002W Index
Japan	BOJ Target Rate	BOJDTR Index
United Kingdom	Bank of England Bank Rate	U.K. BRBASE Index
United States	FOMC Rate Decision (Upper Bound)	FDTR Index

Before our analysis begins, we attribute each news to a given event window, i.e., whether the news became public before, during, or after the trading hours of markets in our sample. For example, on January 12, 2006, at 12:00:00, the scheduled target rate announcement in the U.K. occurred during trading hours in the U.K. and EU markets but before trading began in the U.S. and Canada and after trading hours in Japan. Thus, we synchronized each event separately for each market.

3.3. Exchange rate and interest rate data

Our dataset also contains information about each country's short-term interest rates and exchange rates. The data are at daily frequencies and cover the same time period as the macroeconomic news announcements. The interest rate is represented by the short-term 3-month interest rate for each country. Furthermore, we utilize the so-called "effective exchange rate." An effective exchange rate is calculated as a weighted average of the individual exchange rates of a particular country with its main trading partners. This rate is also known as a trade-weighted exchange rate because the weights are set according to the importance of each partner country's share of trade with the reporting country.

4. Methodology

In this section, we first describe how we measure the volatility of stock markets. Next, we describe our methodology, and ultimately, we explain all the explanatory variables used in the analysis.

4.1. Volatility estimators

The impact of news announcements on stock market variance is based on modeling the realized measures of daily volatility, which is the variable of interest in this study. The most common estimator of daily volatility in the literature is given as follows:

$$\sum_{j=1}^{N} r_{i,t,j}^2 \quad (1)$$

In (1), $r_{i,t,j}$ is the j^{th} intraday return of i^{th} stock market index at day t. N denotes the number of intraday returns, which is a function of the length of the trading hours and the sampling frequency. However, in the presence of microstructure noise, the estimator provided in (1) is biased. Alternatively, one could employ one of several estimators, which are consistent in the presence of a form of microstructure noise (e.g., Barndorff-Nielsen and Shepard, 2004; Hansen and Lunde, 2006; Shephard and Sheppard, 2010; Andersen et al., 2012). A different empirical strategy that we follow in this study was suggested by Patton and Sheppard (2009), who show that different measures may encompass different information; they also advocate the use of a combination of realized measures. Motivated by these considerations, we report the results for a simple arithmetic average of the following eight realized measures of volatility (Barndorff-Nielsen and Shepard, 2004; Hansen and Lunde, 2006; Shephard and Sheppard, 2010; Andersen et al., 2012):

- 5-min realized volatility.
- 10-min realized volatility.
- 5-min realized kernel.
- 5-min realized volatility with 1-min sub-sampling.
- 10-min realized volatility with 1-min sub-sampling.
- 5-min bi-power volatility.
- 5-min bi-power volatility with 1-min sub-sampling.
- 5-min median-truncated volatility.

In all our calculations, we use the logarithm of the resulting mean realized variance, i.e., realized volatility, because the distribution of the realized variances tends to be skewed to the right and is subject to outliers. We denote the log of the combination of realized measures simply as RV_t and refer to it as realized volatility in the subsequent text.

4.2. Econometric models

4.2.1. Univariate model

The impact of news announcement on eight stock market indices is evaluated for each index separately by modeling the Difference of the RV_t using a Heterogeneous AutoRegressive model with eXogenous variables and Generalized AutoRegressive Conditional Heteroskedastic errors (DHARX-GARCH henceforth). The choice to model differences in the RV_t is motivated by the observation that ΔRV_t is less persistent and thus more suitable for time-series modeling. The choice is also in accordance with the existing literature. For example, Marshall et al. (2012) have studied the log difference in implied volatilities. For the sake of robustness, we estimate not only models for change in realized volatility but also models in which the dependent variable is the level of realized volatility (see Section 5.4). The specification is as follows:

$$\begin{aligned} \Delta RV_t &= \mu_0 + \mu_1 RV_{t-1} + \mu_2 RV_{t-1,t-5} + \mu_2 RV_{t-1,t-22} + \sum_{s=1}^{S(i)} \kappa_s EV_s + \sum_{c=1}^{C(i)} \delta_c CV_c + z_t \\ z_t &= (1+\theta_1 L^1)\varepsilon_t \\ \varepsilon_t &= \sigma_t \eta_t, \quad \eta_t \sim iid(0,\ 1) \end{aligned} \tag{2}$$

In model (2), μ, κ, δ, and θ are model parameters. By EV_s, we denote event variables, and by CV_c, we denote control variables (described in more detail in Section 3.3). The variable RV_{t-1} is the lagged realized volatility, and $RV_{t-1,t-5}$, $RV_{t-1,t-22}$ are the average realized volatilities across the previous five and twenty-two trading days, respectively. The lagged realized volatilities tend to capture both weekly and monthly volatility movements, and all three should reflect the heterogeneity in investors' dealing frequencies and investment time horizons (Müller et al., 1997; Corsi, 2009).

Although the lagged realized volatilities and other exogenous variables capture most of the dynamics of market volatility, the error term z_t may remain subject to autocorrelation and conditional heteroscedasticity. We therefore model the term z_t as a moving average process (L is the lag operator), whereas we allow the evolution of $\sigma^2{}_t$ to follow a suitable GARCH model. Two GARCH models are considered: the standard model of Bollerslev (1986),

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1} + \beta_1 \sigma_{t-1}^2 \tag{3}$$

and Nelson's (1991) exponential GARCH model,

$$\ln(\sigma_t^2) = \omega + \alpha_1 s_{t-1} + \gamma_1(|s_{t-1}| - E[|s_{t-1}|]) + \beta_1 \ln(\sigma_{t-1}^2) \tag{4}$$

In specification (4), s_t denotes standardized innovations, and α_i and γ_i control for the leverage and sign effects, respectively.

In addition to the standard normal distribution, we considered a possibility that η_t follows a distribution that can capture the possible asymmetric and leptokurtic properties of the residuals manifested in the behavior of the random variable η_t. Therefore, we also considered the SU-normal distribution of Johnson (1949a,b) with the probability density function defined as follows:

$$f(x) = (2\pi)^{-1/2} J e^{-\frac{z^2}{2}} \tag{5}$$

where $z = \varsigma^{-1}(sinh^{-1}(x) - \lambda)$ and $J = \varsigma^{-1}(x^2 + 1)^{1/2}$. λ and ς are shape parameters that specify the skewness and kurtosis of the distribution.

Different choices for the process of $\sigma^2{}_t$ and η_t lead to four models[6]. We prefer models in which the resulting standardized residuals do not display autocorrelation and conditional heteroscedasticity, as indicated by the Escanciano and Lobato (2009) test. If more suitable models remain, we report a specification that was preferred according to the Bayesian information criterion (BIC) (Schwartz, 1978).

4.2.2. Dynamic common correlated effect model

Estimating individual models allows us to observe market-level heterogeneities; however, it is obvious that due to common global factors, the realized volatilities of developed stock markets will be cross-correlated. Alternatively, we have considered an estimation of the dynamic common correlated effect model along the lines of Chudik and Pesaran (2015). We considered estimating a dynamic panel that follows the mean equation specification of Eq. (2); in addition, we acted in accordance with the approach as described in Ditzen (2016). To achieve consistent parameter estimates in the panel setting, Chudik and Pesaran (2015) show that specifications such as that in Eq. (2) should be augmented with at least $(T)^{1/3}$ cross-sectional means (for up to 2770 observations, that is 14 additional variables). Specifically, we use the following augmentation:

$$\Delta RV_{i,t} = \mu_{i,0} + \mu_{i,1} RV_{i,t-1} + \mu_{i,2} RV_{i,(t-1,t-5)} + \mu_2 RV_{i,(t-1,t-22)} + \sum_{s=1}^{S(i)} \kappa_{i,s} EV_{i,s} + \sum_{c=1}^{C(i)} \delta_{i,c} CV_{i,c} + \sum_{l=0}^{3} \phi_{i,l} \bar{\mathbf{Z}}_{i,t-l} + z_{i,t} \tag{6}$$

[6] DHAR-GARCH with η_t following normal distribution, DHAR-GARCH with η_t in accordance with Johnson's SU distribution, DHAR-EGARCH with η_t following normal distribution, and DHAR-EGARCH with η_t in accordance with Johnson's SU distribution.

where index i is a given market, and the augmented variables that we assume are behind the cross-sectional dependence of residuals in Eq. (2) are $RV_{i,t-1}$, $RV_{i,(t-1,\ t-5)}$, $RV_{i,(t-1,\ t-22)}$ and $\Delta FX_{i,t}$, i.e. cross sectional averages of these four variables are stacked into the vector $\bar{\mathbf{Z}}_{i,t-l}$. After stacking the coefficients $\mu_{i,1}$, $\mu_{i,2}$, $\mu_{i,3}$, $\boldsymbol{\kappa}_{i,s}$, and $\boldsymbol{\delta}_{i,c}$ into a vector, $\boldsymbol{\pi}_i = (\mu_{i,1}, \mu_{i,2}, \mu_{i,3}, \boldsymbol{\kappa}_{i,s}, \boldsymbol{\delta}_{i,c})$. The mean group estimates are as follows:

$$\hat{\boldsymbol{\pi}}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\pi}}_i \tag{7}$$

We further use the following estimator of the asymptotic variance (Chudik and Pesaran, 2015):

$$D(\hat{\pi}_{MG}) = \frac{\hat{\boldsymbol{\Sigma}}_{\mathbf{MG}}}{N} = \frac{1}{N(N-1)}\sum_{i=1}^{N}(\hat{\pi}_i - \hat{\boldsymbol{\pi}}_{MG})(\hat{\pi}_i - \hat{\boldsymbol{\pi}}_{MG})' \tag{8}$$

where under that null hypothesis that the coefficients are equal to 0, the mean group estimates have the following asymptotic distribution:

$$\sqrt{N}(\hat{\pi}_i - \pi) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_{MG}) \tag{9}$$

4.3. Explanatory variables

In this study, we investigate the impact of most relevant news announcements from central bank(s) on volatility of stock markets. In Eqs. (2) and (6), κ_s coefficients measure the impact of a news announcement on market volatility. For each stock market index, we have considered several classes of variables that target capturing the reactions in the market before, during, and after the event while also considering uncertainty about the announcement and the possible magnitude of the unexpected portion of the news announcement.

4.3.1. Key interest rate

Before the news announcement date To indicate the period before the news announcement, we introduce a dummy variable $BN_{tb(k)}$ that returns a value of 1 if tb belongs to an n-day window prior to the scheduled announcement of the event k and 0 otherwise. We report the results for $n = 5$; however, in Section 5.4, we also consider other choices. As volatility levels cluster, and tranquil and calm market periods tend to change, we multiply event variables by $RV_{tb(k)-n-1,tb(k)-2n}$, which is the average level of volatility over the n-days prior to the news announcement window of event k.

- $BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$. The interaction captures changes in volatility before the news announcement.
- $BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n} \times VAE_{tb(k)}$. The $VAE_{tb(k)}$ captures the variance of analysts' estimates of the target rate. Thus, we control for the uncertainty over the central bank's next policy steps.

At the news announcement date To indicate the day at the news announcement, we introduce a dummy variable $N_{tn(k)}$ that returns a value of 1 if $tn(k)$ belongs to a day at the anticipated target rate announcement k and 0 otherwise. The extent of the change in volatility on the announcement day may vary with respect to the level of volatility. During periods of higher market volatility, a change in realized volatility of size 1 is relatively smaller than the same change during calmer volatility periods. Therefore, as before, we multiply event variables by $RV_{tn(k)-1}$, e.g., the level of volatility one day before news announcement day k.

- $N_{tn(k)} \times RV_{tn(k)-1}$. An extensive body of literature suggests that on the day of important news, asset volatility should increase. Therefore, we expect a positive coefficient. However, the literature is salient on the comparison of the size of the effect across countries. By using new data that cover the period of unconventional monetary policy, we also provide evidence that controls for announcements; this also indicates quantitative easing. This provision is important because most of the time, the news related to quantitative easing is announced during the scheduled target rate announcement days.
- $N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}{}^{-}$. Since Giovannini and Jorion (1989), it has been understood that the direction of interest rate changes may have a different impact on market volatility. However, our dataset contains historically extremely low interest rates that are not present in the data sample of Giovannini and Jorion (1989). Therefore, we consider the direction of the surprise the news provides relative to the average expected by the analyst. The variable is $S_{tn(k)} = 100\% \times$ (new – expected)/expected, where "new" and "expected" refer to the revised target interest rate and the median level of the target interest rate expected by analysts, respectively. To disentangle the direction, we define $S_{tn(k)}{}^{-} = S_{tn(k)} \times$ I[new < expected], where I[.] is the indicator function returning 1 if the condition is true and 0 otherwise. Finally, to consider the changing level of market volatility, we multiply the surprise variable $S_{tn(k)}{}^{-}$ by the volatility level the day before the event. Because surprises should not be priced in the market, we expect that they should increase market volatility. Because $S_{tn(k)}{}^{-}$ is negative, we expect that the coefficient at the interaction terms will be negative such that the overall effect on volatility is positive.

- $N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$. As before, we control for the size and direction of the surprise element on the market, where $S_{tn(k)}^{+} = S_{tn(k)} \times$ I[new > expected]. Because $S_{tn(k)}^{+}$ is positive, we expect that the coefficient at the interaction terms will be positive such that the overall effect on volatility is positive. Comparing the two coefficients can be used to assess the market's asymmetric reactions to surprises regarding the monetary policy announcement.

After the news announcement date Obviously, if news at event k contained valuable information for the market, volatility should increase on the new announcement day, whereas the next day, volatility should bounce back if there is no news of such importance for the market to process. Because it is safe to assume that this occurrence is often the case, we should control for such (nearly determined) decreases in volatility the day after the announcement by introducing the $DR_{ta(k)}$ dummy variable with a value of 1 if $ta(k)$ is a date one day after the announcement of the event k and 0 otherwise. After controlling for sudden decreases in volatility, our key variable of interest is $AN_{ta(k)+1}$, which returns a value of 1 if $ta(k) + 1$ belongs to an n-day window after the scheduled announcement of the event k and 0 otherwise. Note that we are capturing what occurs from one day after the news announcement (ta + 1) until n-days after the news announcement. In the main text of this paper, we report results for $n = 5$ days. Thus, we can more directly control for the overall net effect of the monetary policy announcement on the stock market. As before, we also control for the level of volatility. As in the previous case for variables on the announcement day, we again multiply the event variables by the level of volatility. To enable comparisons against volatility before the news announcement, we multiply the event variables by the level of volatility one day before the news announcement $RV_{tn(k)-1}$, which was previously defined.

- $DR_{ta(k)} \times RV_{tn(k)-1}$. The variable controls for a sudden decrease in the level of volatility the day after the news announcement.
- $AN_{ta(k)+1} \times RV_{tn(k)-1}$. This variable is of central importance to the study because we are testing whether news announcements led to a decrease or an increase in the overall level of market uncertainty. If the coefficient is negative, the news announcement led to a short-term stabilization of the equity markets.
- $AN_{ta(k)+1} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$. The effect of a news announcement on market volatility during subsequent days may differ with respect to the magnitude and direction of the surprise. This variable checks whether an unexpected reduction in the target rate led to an increased (de)stabilization of volatility.
- $AN_{ta(k)+1} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$. In this case, we check whether an unexpected increase in the target rate led to changes in volatility.

International development of rates News announcements by other relevant central banks may also influence stock market volatility, both due to the direct effect and because monetary policy of central banks in one country may indicate future monetary policy of other central banks. Controlling for the international development of target rates, i.e., whether foreign central bank of interest have made announcements of target rates (or quantitative easing), leads to several forward challenges that lead to careful synchronization of the data. Although the news announcement of a foreign central bank was made on the same calendar day as the local stock market is open, we consider the exact time when the news was released. The goal of such synchronization is that we want the news to be attributed to the correct event. As before, we are interested in the days before, on, and after the news announcement. We introduce count variables that sum the number of events across foreign central banks.

- $BNW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$. Here, $BNW_{tb(k)}$ sums $BN_{tb(k)}$ across all remaining foreign central banks in the sample, if those news occurred during the period n-days before event k on the local market. For example, let $tb(k)$ be one day before the anticipated news announcement on the U.S. stock market. If at the same time, $tb(k)$ corresponds to one and three days before the news announcement of the *ECB* and *BOE*, respectively, the variable $BNW_{tb(k)}$ is equal to 2.
- $NW_{tn(k)} \times RV_{tn(k)-1}$. The variable $NW_{tn(k)}$ is a count variable, which sums $N_{tn(k)}$ across all other markets but only if rate announcements on other markets occurred during the trading hours of a given local market. To explain the market level of volatility, the dummy variable is multiplied by $RV_{tn(k)-1}$. Note that we expect a smaller positive coefficient on $NW_{tn(k)} \times RV_{tn(k)-1}$ compared to $N_{tn(k)} \times RV_{tn(k)-1}$ because monetary policy in other countries may be important but not as important as the local policy.
- $DRW_{ta(k)} \times RV_{tn(k)-1}$. Volatility before the event announcement is multiplied by a count variable, which sums $DR_{ta(k)}$ across all other markets.
- $ANW_{ta(k)+1} \times RV_{tn(k)-1}$. Volatility before the event announcement is multiplied by a count variable, which sums $AN_{ta(k)}$ across all other markets.

4.3.2. Quantitative easing

The effect of an unconventional monetary policy on the stock markets remains largely unknown. If central banks indicate or engage in quantitative easing policies, markets may react, leading to increased market volatility. However, such events are rare and not present for all markets in our sample. Nevertheless, these events can have a significant effect; we introduce several event variables to control for these policy actions.

4.3.2.1. Local monetary policy indication of quantitative easing.

- $BQ_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$. $QB_{tb(k)}$ is a dummy variable with value of 1 if $tb(k)$ is a date in an n-day window before news announcement k, related to quantitative easing. As before, we report $n = 5$ in the main body of this research paper.
- $Q_{tn(k)} \times RV_{tn(k)-1}$. $Q_{tn(k)}$ is a dummy variable with value of 1 if $tn(k)$ is a date of news announcement k but only if it occurred during the trading hours of the given stock market. We expect positive coefficients loading on the $Q_{tn(k)}$ variable because quantitative easing may be perceived as a significant monetary policy direction for the whole economy.
- $DRQ_{ta(k)} \times RV_{tn(k)-1}$. To explain the sudden decrease in realized volatility, we included the dummy variable $DRQ_{ta(k)}$, which is equal to 1 one day after the event and 0 otherwise.
- $AQ_{ta(k)+1} \times RV_{tn(k)-1}$. To observe whether news related to QE has led to the decrease of realized volatility, we used a dummy variable $AQ_{ta(k)+1}$ with a value of 1 if t is a date in an n-day window after the news announcement k related to quantitative easing.

4.3.2.2. International monetary policy on quantitative easing.

- $BQW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$. $BQW_{tb(k)}$ is a count variable, which sums $BQ_{tb(k)}$ across all other markets. As before, the main idea is that monetary policy in other developed countries may indicate future monetary policy in the given country; therefore, relevant news announcements of other central banks may have an effect on the local stock market.
- $QW_{tn(k)} \times RV_{tn(k)-1}$. $QW_{tn(k)}$ is a count variable, which sums $Q_{tn(k)}$ across all other markets, but only if the rate announcement in other markets occurred during the trading hours of a given market.
- $DRQW_{ta(k)} \times RV_{tn(k)-1}$. The count variable $DRQW_{ta(k)}$ sums $DRQ_{ta(k)}$ across all other markets.
- $AQW_{ta(k)+1} \times RV_{tn(k)-1}$. The count variable $AQW_{ta(k)}$ sums $AQ_{ta(k)}$ across all other markets.

Certain variables were not used for all markets. For example, during the trading hours of the Japanese stock market, the other stock markets in our sample are inactive. Therefore, the specification modeling realized volatility on the NIKKEI 225 excludes several news announcement variables of other markets (namely, $NW_{tn(k)}$ and $QW_{tn(k)}$). Similarly, when modeling the Canadian TSX, we do not have variables related to the *QE* in Canada, because during our sample period, *QE* was only considered in Canada, not actually employed.

4.3.3. Control variables

In the specification of Eqs. (2) and (6), δ_c coefficients are related to a set of control variables. The motivation for including control variables is that other relevant events may be influencing the level of market volatility on a given day t. Therefore, all control variables are not lagged, and it is assumed that they are exogenous with respect to market volatility.

Since changes in short-term interest rates and exchange rates can be perceived as a signal of monetary policy, we included simple differences[7] of 90-day money market interest rates denoted as ΔIRt and a logarithmic difference of a currency index, which measures a currency's appreciation/depreciation relative to the currencies of a country's main trading partners and is denoted as ΔFXt. We also included the squared return (ΔFX2t) to proxy the uncertainty levels on the foreign exchange market, which may spill over to the equity market.

We have also included dummy variables for days of the week. In an influential paper, Andersen and Bollerslev (1998) argue that the day-of-the-week effect on the foreign exchange market may be partly responsible for the clustering of volatility due to the clustering of macroeconomic news announcements on specific weekdays. We therefore introduce dummy variables that are multiplied by the lagged realized volatility, i.e., RV_{t-1}.

5. Results

5.1. Data and model characteristics

The time series of realized volatility for all the stock markets are plotted in Fig. 2. As we can observe, the volatilities of different stock markets exhibit similar time patterns with increased volatility during the financial crisis and the debt crisis. Note that downward spikes for the European STOXX-50 index can be explained by the non-trading of most weighted markets in Europe, i.e., the low levels of volatility are caused by smaller markets trading when big markets have a non-trading day.

Descriptive statistics of the key variables are summarized in Tables 2.1 and 2.2. It is worth noting that the logarithmic transformation of realized volatility was clearly useful. Although realized volatility is highly skewed (not reported here), the logarithm of realized volatility (denoted as *RV* in Tables 2.1 and 2.2 and in the remainder of the paper) is less so, with skewness not very different from zero. As is often the case in the finance literature, realized volatility shows a high level of persistence: the lowest first-order autocorrelation was found for Japan, at 0.770. This finding supports our choice for modeling volatility via an autoregressive model: the DHAR-GARCH model.

[7] The use of simple differences instead of percentage changes is caused by the period of negative rates.

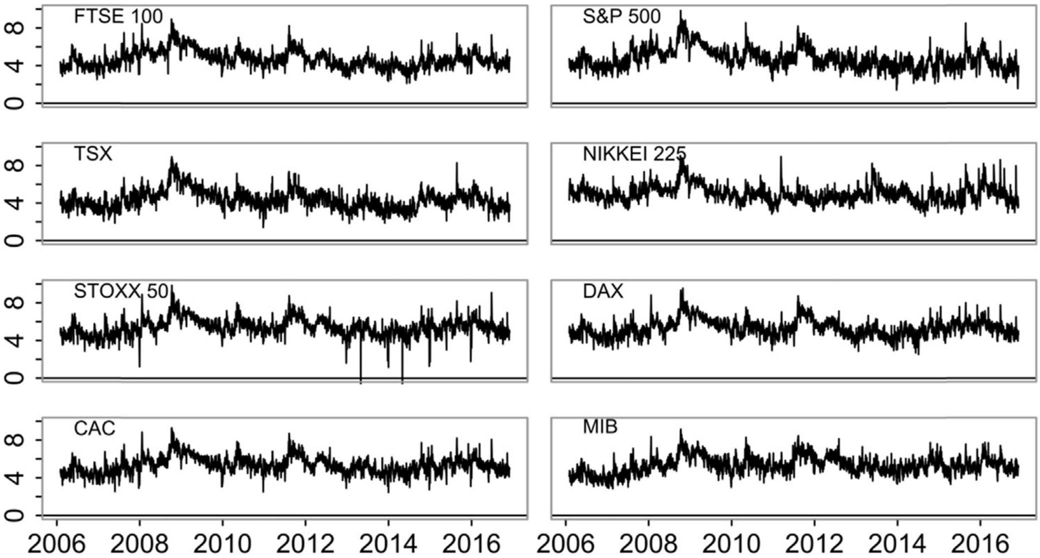

Fig. 2. Realized volatility.

As expected, the highest number of analysts covering rate announcements is for the FED. However, the variance of analyst estimates appears to be small and based on the raw data; analysts often make the same predictions. In reality, surprises are rare, as shown in Tables 2.1 and 2.2, column "#," which shows the number of non-zero surprises. The surprises are reported in a $x100\%$ scale and vary greatly primarily because of the low levels of interest rates at the end of our sample period, in which a small deviation from the average of analyst's expectation leads to large percentage changes.

5.2. Results for panel of all countries

Rate announcements do not occur often. For example, in Canada, the proportion of trading days when an announcement event occurred was only 3.2%. In our sample period, quantitative easing announcements are naturally rare (see Tables 2.1 and 2.2). Consequently, the per-market estimators may not have the proper power to detect changes in realized volatility. Instead, motivated by the mean group estimation literature, we first evaluate the results across all markets together. In Table 3, we report the averages of coefficients and their significances using two aggregation techniques: the dynamic common correlated error model (DCCE, see Section 4.2.2) and the simple mean group (MG) estimator. The latter estimator is a simple average, and the t-ratio is calculated from the per-market models (reported in Tables 4 and 5).

The results between the two estimation methods often differ; one notable exception is the results for the period before the news announcement and for the effect of news related to quantitative easing. The results from the DCCE model led to a more intuitive finding that increased uncertainty over the outcome of a news announcement that led to higher market volatility before that news announcement. Simultaneously, volatility before a news announcement tends to be generally smaller ("calm before the storm" effect). The MG estimator led to insignificant coefficients with opposing signs. However, the size of the effect estimated via DCCE is small. Because our dependent variable is the log of the realized variance, the decrease of approximately 0.3% (before the news announcement) translates into an approximately 1.3% decrease in the realized variance[8].

At the news announcement, volatility increases, particularly if coupled with larger surprises, i.e., if there are a deviation from market-wide expectations. The size of the effect is large because it corresponds to a 5.3% increase in the log of the realized variance, i.e., ≈30.5% increase in the realized variance. Furthermore, surprises appear to be asymmetric, with market volatility being much higher for rates below the anticipated values.

Based on the DCCE estimator, five days after the announcement, the estimated decline in the log of realized variance is approximately 1%, which corresponds to an approximately 5% decrease in the realized variance. The decrease was not captured via the mean group estimator. In contrast to the observation made at the announcement date, surprises tend to be associated with lower market volatility. However, the effect on volatility five days after the announcement is much smaller than that during the announcement day.

When foreign central banks make announcements, our estimates suggest that volatility tends to increase slightly before and on the news announcement days. After a news announcement, volatility tends to decline. These effects, although occasionally significant, are small. For example, (for DCCE estimator) at the announcement day, the realized volatility increase is approximately 0.6% (≈ 3% in realized variance); during the five-day period after the announcement, the decline tendency caused by foreign central bank announcements is approximately 0.2% (≈ 1% in realized variance).

[8] In all these transformations, we are assuming an average level of realized volatility across all markets, i.e., 5.035.

Table 2.1
Descriptive statistics of key variables for United States, United Kingdom, Japan, and Canada.

United States		Rate announcements 88, Quantitative easing indication 8						
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV		4.754	1.117	0.691	3.724	1.437	9.876	0.847
ΔRV		0.000	0.617	0.186	3.677	−2.420	2.647	−0.342
$\Delta IR \times 100$		−0.106	1.171	−2.114	25.902	−11.800	9.500	0.724
$\Delta FX \times 100$		0.006	0.344	0.142	6.624	−2.202	2.207	−0.003
$\Delta FX^2 \times 100$ (annualized)		3.960	3.751	2.271	12.143	0.000	35.028	0.210
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.059	0.088	1.822	5.048	0.000	0.297	0.439
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	2	−4.225	37.356	−9.168	85.347	−350.005	0.000	−0.013
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	1	6.323	59.314	9.220	86.012	0.000	556.412	−0.012
United Kingdom		Rate announcements 127, Quantitative easing indication 9						
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV		4.713	0.945	0.679	3.639	2.142	8.957	0.835
ΔRV		0.000	0.542	0.079	4.677	−3.343	3.065	−0.361
$\Delta IR \times 100$		−0.155	2.626	−25.171	1003.630	−106.500	21.250	0.316
$\Delta FX \times 100$		−0.008	0.484	−1.248	17.273	−6.200	2.185	0.097
$\Delta FX^2 \times 100$ (annualized)		5.279	5.576	3.979	41.140	0.000	98.426	0.273
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.072	0.109	2.210	6.983	0.003	0.434	0.645
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	2	−1.490	14.935	−10.866	120.693	−167.000	0.000	−0.010
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	3	3.879	39.914	11.059	123.844	0.000	449.097	−0.010
Japan		Rate announcements 100, Quantitative easing indication 5						
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV		4.983	0.900	0.801	4.398	2.605	9.042	0.768
ΔRV		0.000	0.612	0.567	6.314	−2.916	4.144	−0.351
$\Delta IR \times 100$		−0.094	1.155	−2.067	27.123	−11.800	9.500	0.718
$\Delta FX \times 100$		0.003	0.673	0.363	7.877	−3.674	4.807	−0.017
$\Delta FX^2 \times 100$ (annualized)		7.494	7.619	2.523	13.852	0.000	76.314	0.211
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.022	0.025	1.002	2.568	0.000	0.073	0.614
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	2	−6.543	49.389	−8.123	70.646	−453.092	0.000	−0.018
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	2	5.701	43.072	8.136	70.875	0.000	395.532	−0.018
Canada		Rate announcements 87, Quantitative easing indication 0						
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV		4.383	1.015	0.767	4.118	1.408	8.943	0.808
ΔRV		0.000	0.629	0.147	3.806	−2.378	3.294	−0.371
$\Delta IR \times 100$		−0.105	1.168	−2.081	25.974	−11.800	9.500	0.718
$\Delta FX \times 100$		−0.007	0.586	−0.110	6.476	−3.082	3.966	−0.048
$\Delta FX^2 \times 100$ (annualized)		6.674	6.485	2.219	10.870	0.000	62.955	0.242
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.066	0.068	2.400	8.782	0.008	0.302	−0.164
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	5	−6.026	32.563	−6.514	48.024	−262.617	0.000	−0.035
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	2	1.549	10.797	7.413	58.997	0.000	91.365	−0.021

Notes: RV denotes realized volatility, ΔRV denotes the difference between two consecutive realized volatilities, ΔIR is the difference in the short-term interest rate, and ΔFX is the log difference of the currency index. Column "#" denotes the number of non-zero surprises. AC1 is the level of the first-order autocorrelation coefficient. Statistics for variables $BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$; $N_{tn(k)} \times RV_{tn(k)-1} \times S^{-}_{tn(k)}$; $N_{tn(k)} \times RV_{tn(k)-1} \times S^{+}_{tn(k)}$ are calculated only for event days, i.e., when $BN_{tb(k)} = 1$ or when $N_{tn(k)} = 1$.

Interestingly, news related to quantitative easing led to a small decrease in realized volatility before and on the day of the announcement (DCEE estimator) but only when reported by the domestic central bank. Conversely, if foreign central banks reported initialization or continuation of the QE, volatility increased by 3.6%, which corresponds to a nearly 20% increase in the realized variance; the latter effect has the same direction for both DCCE and MG estimators.

5.3. Results for individual countries

The individual market-level results are reported in Tables 4 and 5, and the key results are illustrated in Fig. 3. The estimated results from the DHAR-GARCH models show that variables RV_{t-1}, $RV_{t-1,t-5}$ and $RV_{t-1,t-22}$ explain the behavior of volatility changes well, because they are often significant with strong effects. The results of diagnosis tests showed that the error term should be modeled as an MA(1) process. This modeling is captured by the term ε_{t-1}. Moreover, the error term exhibited conditional heteroscedasticity; therefore, we model it as a GARCH model; see the variance equation in the lower part of each table. For all stock markets except the U.S. and Canada, the GARCH(1,1) model is sufficient using our modeling framework (see Section 4.2). The results of diagnostic tests (in the bottom of both tables) show that in most cases, residuals from our models have satisfactory properties; one exception is the French CAC and Italy's MIB market index's volatility, for

Table 2.2
Descriptive statistics of key variables for Europe, Germany, France and Italy.

Europe		Rate announcements 121, Quantitative easing indication 2						
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV-STOXX		5.402	0.957	−0.142	7.587	−2.885	9.879	0.737
ΔRV-STOXX		0.000	0.694	0.645	29.939	−7.502	8.686	−0.396
$\Delta IR \times 100$		−0.105	1.168	−2.112	25.899	−11.800	9.500	0.725
$\Delta FX \times 100$		0.000	0.371	−0.322	8.521	−3.106	2.541	−0.020
$\Delta FX^2 \times 100$ *(annualized)*		4.156	4.165	2.640	16.440	0.000	49.299	0.199
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.033	0.044	2.172	7.429	0.001	0.189	0.304
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	5	−10.759	63.236	−6.911	53.510	−553.172	0.000	−0.029
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	3	5.743	37.181	6.639	47.236	0.000	304.355	−0.024
Germany								
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV-DAX		5.334	0.924	0.518	3.743	2.522	9.557	0.806
ΔRV-DAX		0.000	0.576	0.168	4.048	−2.258	2.567	−0.363
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.033	0.045	2.320	8.287	0.001	0.204	0.300
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	5	−10.500	62.736	−7.074	55.662	−553.271	0.000	−0.029
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	3	5.548	36.018	6.649	47.345	0.000	295.464	−0.024
France								
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV-CAC		5.351	0.902	0.457	3.646	2.460	9.302	0.796
ΔRV-CAC		0.000	0.576	0.233	5.146	−2.278	3.497	−0.365
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.033	0.045	2.292	8.139	0.001	0.201	0.288
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	5	−10.359	60.990	−6.916	53.738	−536.778	0.000	−0.029
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	3	5.668	36.949	6.661	47.340	0.000	301.242	−0.024
Italy								
	#	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum	AC1
RV-MIB		5.358	0.930	0.306	3.154	2.847	9.156	0.804
ΔRV-MIB		0.000	0.582	0.002	3.582	−2.060	2.345	−0.358
$BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$		0.033	0.044	2.173	7.478	0.001	0.194	0.288
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{-}$	5	−10.663	63.161	−7.004	55.027	−558.770	0.000	−0.029
$N_{tn(k)} \times RV_{tn(k)-1} \times S_{tn(k)}^{+}$	3	5.819	37.687	6.576	45.901	0.000	300.283	−0.024

Notes: RV denotes realized volatility, ΔRV denotes the difference between two consecutive realized volatilities, ΔIR is the difference in the short-term interest rate, and ΔFX is the log difference of the currency index. Column "#" denotes the number of non-zero surprises. AC1 is the level of the first-order autocorrelation coefficient. Statistics for variables $BN_{tb(k)} \times RV_{tb(k)-n-1,t-2n} \times VAE_{tb(k)}$; $N_{tn(k)} \times RV_{tn(k)-1} \times S^{-}_{tn(k)}$; $N_{tn(k)} \times RV_{tn(k)-1} \times S^{+}_{tn(k)}$ are calculated only for event days, i.e., when $BN_{tb(k)} = 1$ or when $N_{tn(k)} = 1$.

which Escanciano and Lobato (2009) test the suggested presence of conditional heteroskedasticity. However, the first-order autocorrelation of squared residuals was very mild, at only 0.057 for the French market and 0.022 for the Italian market.

First, we discuss the impact of interest rate announcements in their respective countries. In Fig. 3, we illustrate the main results, where it appears that during the 5-day period before the announcement, most of the time-realized volatility decreases, except for the Japanese, Italian and European markets. However, except for Japan, changes in volatility before news announcements do not appear significant.

On the announcement day, we observe that volatility of the stock market is increased in all the countries. This finding was expected, and the effects are strong. The coefficients range from a 2.2% (U.K.) to a 9.4% (U.S.) increase in the log of the realized variance compared to the levels of volatility one day before. Given the average level of the realized variances, this result corresponds to 10.9% (U.K.) and 54.8% (U.S.) increases in the realized variance. We further investigate the importance of the interest rate surprise. To allow an asymmetric response to positive and negative surprises, we included the interaction terms $N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{-}$ and $N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{+}$. If surprises increase volatility, the coefficient for the first term should be negative, and the coefficient for the second term should be positive. The expected signs were found for all countries except for the U.S. and Canada. For the U.S., volatility decreases the greater the surprise effect and regardless of the direction of the surprise, although surprises below expectations led to a larger volatility decrease. For Canada, target rates above expectations decrease market volatility, whereas target rates below expectations increase market volatility (as expected). The overall effect is illustrated in Fig. 3 and shows that for all the markets considered, volatility increases on the day of the news announcement; this effect is largest for the U.S.

The first term in the panel "*After interest rate meeting*" in Tables 4 and 5, $DRA_{ta(k)} \times RV_{tn(k)-1}$, captures decreases in volatility immediately after the announcement day. We expected the coefficient to be negative and possibly close to the magnitude of the coefficient capturing volatility changes at the day of the interest rate meeting. Although the coefficient is negative across all markets, it is significant only for the U.S. and significant at the 10% level for Canada and Italy. However, we interpret the negative coefficients to support our specification because failing to control for these declines may result in an overestimation of the decline in market volatility after the news announcement. In other words, a decrease in volatility from the

Table 3
Dynamic common correlated effects and mean group estimation.

		DCCE	MG
Constant		**0.327*****	**0.018*****
ε_{t-1}	MA term		**−0.278*****
RV_{t-1}	Lagged real. vol.	**−0.883*****	**−0.300*****
$RV_{t-1,t-5}$	Lagged weakly real. vol.	**0.390*****	**0.135*****
$RV_{t-1,t-22}$	Lagged monthly real. vol.	**0.378*****	**0.153*****
ΔIR_{t-1}	Lagged changes in inter. rate	**−0.002*****	0.001
ΔFX_{t-1}	Lagged FX returns	**0.019*****	**0.031***
ΔFX^2_{t-1}	Lagged squared FX returns	**−0.002*****	0.001
$Mon_t \times RV_{t-1}$	Monday effect	**−0.037*****	**−0.037*****
$Tue_t \times RV_{t-1}$	Tuesday effect	**0.005*****	**0.012*****
$Thu_t \times RV_{t-1}$	Thursday effect	**0.008*****	**0.009*****
$Fri_t \times RV_{t-1}$	Friday effect	0.002	0.000
Before interest rate meeting			
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n} \times VAE_{tb(k)}$	Uncertainty of the outcome	**0.301*****	−0.099
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	**−0.003*****	0.001
At interest rate meeting			
$N_{ta(k)} \times RV_{ta(k)-1}$	Effect at the announcement	**0.053*****	**0.045*****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	**−0.201*****	**−0.171*****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	**0.042*****	−0.011
After interest rate meeting			
$DR_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	0.001	**−0.020*****
$AN_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	**−0.010*****	0.000
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	**0.054*****	**0.049***
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	−0.024	−0.021
International development of rates			
$BNW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	**0.003*****	**0.003*****
$NW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	**0.006****	0.007
$DRW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	0.001	0.001
$ANW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	**−0.002*****	0.000
Quantitative easing (QE)			
$BQ_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	**−0.029*****	**−0.010****
$Q_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	**−0.004*****	**0.032*****
$DRQ_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	**0.040*****	−0.057
$AQ_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	**0.009*****	−0.005
International QE			
$BQW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	n-day effect before the announcement	**0.015*****	0.005
$QW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	**0.036*****	**0.018***
$DRQW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	**−0.007****	**−0.016*****
$AQW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	0.000	**−0.004****
Average R^2		31.5%	

Note: Subscripts ***, **, *, denote significance at the 1%, 5%, and 10% significance levels, respectively. Bold values denote significance at least at the 10% level.

event day to a subsequent day does not mean that volatility on the subsequent day is low; it is simply caused by a volatility increase on the event day.

The possible stabilizing or destabilizing effect of a monetary policy news announcement on the stock market after the news announcement is studied via the remaining three terms in the panel "After interest rate meeting" (Tables 4 and 5). The coefficients for the term $AN_{ta(k)} \times RV_{tn(k)-1}$ are negative but insignificant for most countries. In other words, on average, compared to the level of volatility before the event day, volatility 5 days after the interest rate announcement has not decreased. Therefore, it appears that a monetary policy news announcement affects volatility only on the day of the announcement and not during the period after the announcement. The coefficients related to the remaining two interaction terms $AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{-}$ and $AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{+}$ capture whether the response depends on the direction and magnitude of the interest rate surprises. We find that surprises matter; however, their effect differs with respect to any given market. For all markets except the U.K., interest rate surprises below expectations led to the stabilization of the market. Conversely, interest rates above expectations increased market volatility for the European markets, including the U.K., while tending to reduce market volatility for the U.S., Canada, and Japan. Furthermore, the effect of surprises is small. For example, the coefficient for surprises below expectations is estimated to be −0.116 for the U.K.; in addition, given the averages, the effect on realized volatility is a decline of approximately 0.7%.

To answer our core research question of whether central banks' announcements have a stabilizing or destabilizing effect on stock markets, we have graphically summarized our results in Fig. 3. The average impact of central bank announcements on volatility before and after the announcement appears to be mixed. In addition, note that compared to the day of the announcement, the magnitude of the effect is much smaller for the before and after periods.

Table 4
Individual models for the U.S., the U.K., Canada and Japan: DHAR-GARCH.

		U.S.	U.K.	JP	CA
Constant		0.028	0.020	0.017	0.020
ε_{t-1}	MA term	**−0.379*****	**−0.310*****	**−0.410*****	**−0.278*****
RV_{t-1}	Lagged real. vol.	**−0.138*****	**−0.277****	**−0.186****	**−0.351*****
$RV_{t-1,t-5}$	Lagged weakly real. vol.	0.012	0.125	0.013	**0.160*****
$RV_{t-1,t-22}$	Lagged monthly real. vol.	**0.119*****	**0.138*****	**0.155*****	**0.183*****
ΔIR_{t-1}	Lagged changes in inter. rate	−0.005	**0.008*****	0.001	0.006
ΔFX_{t-1}	Lagged FX returns	**0.066****	0.005	**0.106*****	−0.051
ΔFX^2_{t-1}	Lagged squared FX returns	−0.000	0.002	**0.004****	−0.000
$Mon_t \times RV_{t-1}$	Monday effect	**−0.035*****	**−0.046*****	**−0.014****	−0.026
$Tue_t \times RV_{t-1}$	Tuesday effect	**0.016*****	**0.014****	−0.001	0.004
$Thu_t \times RV_{t-1}$	Thursday effect	0.000	**0.017*****	0.008	−0.008
$Fri_t \times RV_{t-1}$	Friday effect	**−0.018****	0.008	0.005	**−0.028****
Before interest rate meeting					
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n} \times VAE_{tb(k)}$	Uncertainty of the outcome	0.298	−0.185	−1.036	0.247
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	−0.000	−0.002	**0.007***	−0.002
At interest rate meeting					
$N_{ta(k)} \times RV_{ta(k)-1}$	Effect at the announcement	**0.094*****	**0.022***	**0.022****	**0.038****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	**0.155****	**−0.376*****	−0.016	**−0.131****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	**−0.110*****	**0.080*****	0.003	−0.199
After interest rate meeting					
$DR_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	**−0.038*****	−0.018	−0.012	**−0.022***
$AN_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.006	−0.003	−0.000	0.006
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	0.102	**−0.116****	**0.110*****	**0.073*****
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	**0.025***	**−0.090*****	**0.041****	0.142
International development of rates					
$BNW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	0.003	0.002	0.003	0.003
$NW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	−0.031	**0.014***		**0.060*****
$DRW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	−0.001	0.003	−0.002	−0.008
$ANW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	0.004	0.000	0.001	0.000
Quantitative easing (QE)					
$BQ_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	n-day effect before the announcement	−0.001	0.004	0.007	
$Q_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	0.032	0.022	**0.189*****	
$DRQ_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	**−0.043****	0.012	−0.035	
$AQ_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.004	−0.009	−0.0165	
International QE					
$BQW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	**0.012***	0.001	**0.010***	0.010
$QW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement		0.024		**0.082****
$DRQW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	0.007	−0.001	**−0.040****	−0.021
$AQW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.011	0.001	−0.003	0.002
GARCH					
ω	Constant	−0.034	0.049	0.001	**−0.128****
α	Effect of the lagged error term	**0.031****	**0.088****	**0.040****	**0.063*****
β	Effect of the lagged latent volatility	**0.974*****	**0.676*****	**0.923*****	**0.800*****
γ	Asymmetric coefficient	0.087			**0.058***
λ	Skewness parameter	**0.940*****	**0.257****	**0.883*****	**0.329*****
ς	Kurtosis parameter	**2.859*****	**2.062*****	**2.009*****	**2.370*****
Diagnostics					
Log-likelihood		−2108.90	−1653.54	−1878.51	−2098.79
Bayesian information criterion		1.66	1.32	1.54	1.65
Sign bias test (joint effect)		0.78	0.23	0.43	0.72
Escanciano and Lobato (2009) test of autocorr. of std. res. (up to 5 lags)		0.89	0.80	0.49	0.80
Escanciano and Lobato (2009) test of autocorr. of squared std. res. (up to 5 lags)		0.55	0.38	0.14	0.15
Correlation between fitted and observed values		0.56	0.49	0.53	0.50
Number of observations		2747	2719	2728	2626

Note: Subscripts ***, **, and * denote significance at the 1%, 5%, and 10% significance levels, respectively. Significances are based on standard errors, as in White (1982). Bold values denote significance at least at the 10% level.

Realized volatility in the stock markets also appears to be increasing during the days when interest rate announcements are made by central banks in other countries. However, this effect is significant only in the U.K. and Canada. This result is very intuitive. Canada has strong ties to the U.S., and the U.K. has strong ties to the EU; therefore, one would expect that these countries will respond to announcement of foreign central banks. Most likely, Canada is strongly responding to FED announcements, and the U.K. is strongly responding to ECB announcements.

Table 5
Individual models for European markets: DHAR-GARCH.

		EU	DE	FR	IT
Constant		0.004	0.007	0.012	0.036
ε_{t-1}	MA term	**−0.206***	**−0.252*****	**−0.180***	**−0.208****
RV_{t-1}	Lagged real. vol.	**−0.369*****	**−0.318*****	**−0.405*****	**−0.358*****
$RV_{t-1,t-5}$	Lagged weakly real. vol.	**0.204***	**0.146***	**0.235****	**0.184****
$RV_{t-1,t-22}$	Lagged monthly real. vol.	**0.153*****	**0.162*****	**0.158*****	**0.159*****
ΔIR_{t-1}	Lagged changes in inter. rate	−0.004	−0.002	−0.000	−0.004
ΔFX_{t-1}	Lagged FX returns	0.027	0.026	**0.044***	0.028
ΔFX^{2}_{t-1}	Lagged squared FX returns	−0.001	0.000	−0.001	0.001
$Mon_t \times RV_{t-1}$	Monday effect	**−0.040*****	**−0.040*****	**−0.042*****	**−0.048*****
$Tue_t \times RV_{t-1}$	Tuesday effect	0.009	**0.013****	**0.015****	**0.022*****
$Thu_t \times RV_{t-1}$	Thursday effect	**0.016*****	**0.013****	**0.014*****	0.008
$Fri_t \times RV_{t-1}$	Friday effect	0.008	0.005	**0.013****	0.009
Before interest rate meeting					
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n} \times VAE_{tb(k)}$	Uncertainty of the outcome	0.027	0.077	0.084	−0.303
$BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	0.001	−0.001	−0.000	0.003
At interest rate meeting					
$N_{ta(k)} \times RV_{ta(k)-1}$	Effect at the announcement	**0.044*****	**0.049*****	**0.040*****	**0.052*****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	**−0.276*****	**−0.200****	**−0.290*****	**−0.234*****
$N_{ta(k)} \times RV_{ta(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	0.060	0.000	0.060	0.009
After interest rate meeting					
$DR_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	−0.018	−0.015	−0.013	**−0.024***
$AN_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.001	0.002	0.001	−0.005
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{-}/100$	Effect of a surprisingly lower rate	**0.055*****	**0.073*****	**0.073*****	**0.027***
$AN_{ta(k)} \times RV_{tn(k)-1} \times S_{ta(k)}^{+}/100$	Effect of a surprisingly higher rate	**−0.078****	**−0.071*****	**−0.080*****	**−0.057****
International development of rates					
$BNW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	0.003	0.002	0.003	0.004
$NW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	0.004	0.002	0.003	0.001
$DRW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	0.005	0.004	0.002	0.005
$ANW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.002	−0.000	−0.002	−0.003
Quantitative easing (QE)					
$BQ_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	−0.019	−0.016	**−0.019***	**−0.024*****
$Q_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	0.007	0.043	0.038	0.049
$DRQ_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	−0.086	−0.155	−0.135	**−0.179*****
$AQ_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.005	0.014	−0.000	0.019
International QE					
$BQW_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n}$	5-day effect before the announcement	0.009	0.009	0.009	0.007
$QW_{tn(k)} \times RV_{tn(k)-1}$	Effect at the announcement	−0.001	−0.003	−0.002	0.014
$DRQW_{ta(k)} \times RV_{tn(k)-1}$	1-day effect after the announcement	−0.014	−0.013	−0.013	−0.030
$AQW_{ta(k)} \times RV_{tn(k)-1}$	5-day effect after the announcement	−0.008	−0.004	−0.009	0.001
GARCH					
ω	Constant	**0.105***	0.018	0.020	**0.015****
α	Effect of the lagged error term	**0.169*****	**0.063*****	**0.067*****	**0.046*****
β	Effect of the lagged latent volatility	**0.458***	**0.863*****	**0.848*****	**0.894*****
γ	Asymmetric coefficient				
λ	Skewness parameter	0.113	**0.560*****	**0.224****	**0.662*****
ς	Kurtosis parameter	**1.588*****	**2.392*****	**1.929*****	**2.866*****
Diagnostics					
Log-likelihood		−1966.24	−1851.22	−1799.26	−1889.68
Bayesian information criterion		1.54	1.45	1.40	1.49
Sign bias test (joint effect)		0.36	0.28	0.02	0.01
Escanciano and Lobato (2009) test of autocorr. of std. res. (up to 5 lags)		0.89	0.92	0.99	0.71
Escanciano and Lobato (2009) test of autocorr. of squared std. res. (up to 5 lags)		0.55	0.52	0.04	0.03
Correlation between fitted and observed values		0.56	0.56	0.53	0.54
Number of observations		2747	2747	2744	2770

Note: Subscripts ***, **, and * denote significance at the 1%, 5%, and 10% significance levels, respectively. Significances are based on standard errors, as in White (1982). Bold values denote significance at least at the 10% level.

The volatility five days before foreign central banks' announcements is not influenced by these announcements except for the U.S., which has a slight increase in market volatility during the 5-day window before the event. Similarly, we do not observe a change in volatility five days after the announcement, except for the Eurozone, where volatility decreases.

Next, we discuss how news related to policy announcements about quantitative easing influenced stock market volatility. On the day of the announcements, realized volatility increased; although in many cases, the size of the coefficients was comparable to the size estimated for interest rate announcements, a significant increase is found only for Japan. This result is in

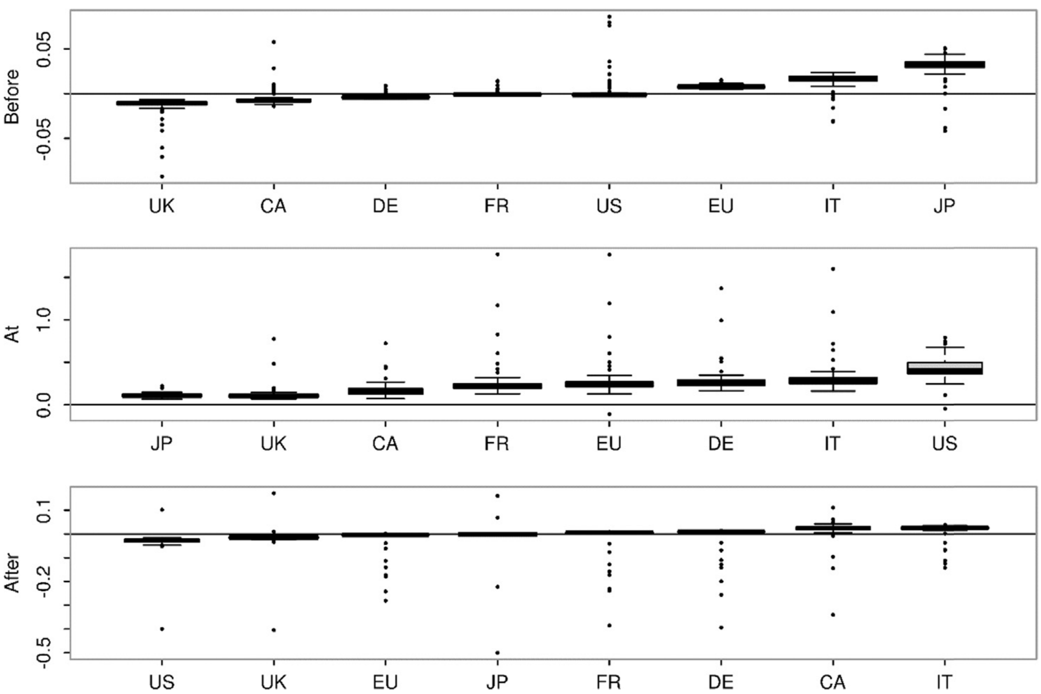

Fig. 3. Effects before/the day of/after the interest rate announcement on the ΔRV_t. Notes: All effects are calculated only for the corresponding event period using Eq. (2) and the estimated coefficients from Tables 4 and 5. In the case of the before period, we calculated $\kappa_1^* \times (BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n} \times VAE_{tb(k)})$ + $\kappa_2^* \times (BN_{tb(k)} \times RV_{tb(k)-n-1,tb(k)-2n})$ only for dates that fall into the "before interest rate meeting," where κ_j^* are estimated coefficients. The resulting effects are plotted as boxplots. At the event day, the plotted values originate from $\kappa_3^* \times (N_{ta(k)} \times RV_{ta(k)-1}) + \kappa_4^* \times (N_{ta(k)} \times RV_{ta(k)-1} \times S^-_{ta(k)}) + \kappa_5^* \times (N_{ta(k)} \times RV_{ta(k)-1} \times S^+_{ta(k)})$, and for after the event day period, the values are $\kappa_6^* \times (AN_{ta(k)} \times RV_{tn(k)-1}) + \kappa_7^* \times (AN_{ta(k)} \times RV_{tn(k)-1} \times S^-_{ta(k)}) + \kappa_8^* \times (AN_{ta(k)} \times RV_{tn(k)-1} \times S^+_{ta(k)})$.

contrast to the results achieved via the DCEE model. Therefore, it appears that when controlling for common factors, the effect of the QE on market volatility diminishes. The quantitative easing of foreign central banks increases volatility in Canada. After controlling for possible one day declines in realized volatility ($DRQ_{ta(k)} \times RV_{tn(k)-1}$ variable), we did not observe a significant increase or decrease in volatility after quantitative easing announcements; the only difference was the reaction of the U.S. stock market.

5.4. Sensitivity checks

To determine whether our results are robust with regard to the model and the variable specification, we performed two sensitivity checks. First, we estimated the HAR-GARCH models that differ with regard to Eq. (2) only in that the dependent variable is the level of realized volatility instead of the difference. Second, our choice for studying 5 days before and 5 days after the news announcement is arbitrary. Therefore, we should study our key results while considering other choices for specifying before and after event windows. Thus, we also increase our understanding of the effect of the target interest rate announcement on stock market volatility. We study the following choices of n, $n = \{3, 7, 9\}$.[9]

First, the HAR-GARCH model provided qualitatively very similar results to the baseline DHAR-GARCH model. More variability in the results is observed when the length of the period before and after the event changes. For the shorter period of $n = 3$ days, the effects become slightly stronger. There is more evidence that volatility declines before and after the news announcement. However, the effects are small. For example, in the DCCE model, realized volatility declines 0.43%, which corresponds to an approximately 2% decline in the realized variance. In the period after the news announcement, volatility declined 1.18%, i.e., 5.8% in the realized variance. For longer time periods, $n = \{7, 9\}$ the effects declined in magnitude.

Similar findings were observed when the DHAR-GARCH model was fitted for individual markets. For example, small significant declines in realized volatility for the before and after period were now estimated for the U.S.; for the EU and the German stock markets, significant declines in realized volatility were found for the after period.

5.5. Results for implied volatility

In order to check whether our reasoning about the differences between realized and implied volatility is supported in an empirical setting, we re-estimate the same country-level models as those reported in Tables 4 and 5. The dependent variable,

[9] The results from Sections 5.4 and 5.5 are available upon request.

the difference in realized volatility is substituted by the difference in implied volatility. We hypothesized that implied volatility should decline at the news announcement day. The results show a mild decrease in implied volatility on announcement day for the U.S., U.K., and Canada, while for other countries the effect is insignificant. After the announcement day, the results are mixed. For sake of brevity are these results not reported in the paper.

6. Conclusion

In this paper, we investigated the impact of monetary policy announcements on the volatility of stock markets in G7 countries. Monetary policy announcements are part of the most important macroeconomic news. Unsurprisingly, significant research has been devoted to the study of these events. However, the research analyzing the impact of these events on stock-market volatility is relatively scarce, particularly with respect to cross-country studies and studies that utilize the precise volatility measure of the realized volatility that is calculated from high-frequency data. Therefore, we analyze a broad data sample covering five central banks and stock markets in the U.S., Canada, Japan, the U.K., Germany, France, Italy and the Eurozone as a whole.

Previous studies have analyzed the impact of news announcements on the volatility of financial markets utilizing either GARCH models or implied volatility. However, GARCH models are based on one price observation per day, and therefore, they cannot precisely estimate volatility. Implied volatility is unsuitable for this purpose for another reason. On the day of a scheduled macroeconomic announcement, daily values of implied volatility should decrease simply because they capture volatility over the ensuing 30-day time period, i.e. news is excluded as it was announced before the closing value of the implied volatility index. Therefore, if the event day is associated with increased volatility, by observing a decrease in implied volatility, we cannot conclude whether this event has a stabilizing or a destabilizing impact on volatility during the subsequent days.

To avoid both these obstacles, we utilize realized volatility calculated from high-frequency data. High-frequency data have previously been used in connection with monetary policy announcements to evaluate the immediate impact of these announcements. In addition to immediate one-day impact, we study also the impact of these events on realized volatility five days before, five days after and the day of an announcement. We control for fluctuations in the short-term money market, the foreign exchange market, and the day-of-the week effects; we also control for the uncertainty of the incoming target rate announcement, the effect of surprises, news related to the initiation or continuation of quantitative easing, and announcements related to foreign central banks' quantitative easing and target interest rates.

Some previous studies have found either a decline (Bomfim, 2003) or an increase (Bauwens et al., 2005) in market volatility before the news announcement. Our results do not support any of these claims. We assumed that volatility on the day of the announcement should be high, and our models confirm this assumption with significant results. This finding suggests that the information content of interest rate news is high (Chae, 2005). This result is in accordance with Balduzzi, Elton, and Green (2001), Evans and Speight (2010b), and Hussain (2011). With regard to volatility behavior after the announcement day, we found evidence of a small stabilizing effect (i.e. volatility decrease) of scheduled interest rate announcements on stock markets; thus, our results contrast with the existing literature (that uses implied volatility), which found strong stabilizing effects. We also contribute to the literature in that we control for the effect of monetary policy interconnectedness or coordination by including news announcement from foreign central banks (interest rate and quantitative easing). Here our results regarding the effect of scheduled news announcements from foreign central banks are also somewhat weaker compared with the earlier findings of Nikkinen and Sahlström (2004) and Nikkinen et al. (2006), as we found that this relationship is usually not very significant at the country level, only when a dynamic common correlated effect model is used. Specifically, we can derive the following conclusions that hold in general:

- Volatility increases on the day of an announcement. Such increases are significant not only in statistical terms but also in economic terms as the increases range from 2.2% (U.K.) to a 9.4% (U.S.). Using the DCCE model, volatility was estimated to be 5.3% in the log of the realized variance (≈10.9%, 54.8% and 30.5% increases in realized variance, respectively). The volatility increases higher when the surprise for the market is greater.
- There is some evidence that monetary policy announcements related to target interest rates have a stabilizing effect on the market during the 5-day period after an announcement. Using the DCCE model, the decline of the log of the realized variance was estimated at approximately 1%, which corresponds to approximately 5.1% in the realized variance. At the individual market level, the changes of the log of the realized variance range from +0.6% (Italy, Canada) to −0.6% (U. S.). When the surprise effect on the policy announcement day is larger, market volatility tends to decline more in the days following the announcement.
- Foreign central banks' announcements about interest rates are positively associated with changes in the local market's volatility (DCCE panel model); however, at the individual market level, these changes are rarely significant (exceptions are the U.K. and Canada).
- We have also studied news related to the initialization or continuation of quantitative easing policies. The results indicate that on average, quantitative easing announcements did not have an impact on the volatility of stock markets in G7 countries.

References

Äijö, J., 2008. Impact of US and UK macroeconomic news announcements on the return distribution implied by FTSE-100 index options. Int. Rev. Financ. Anal. 17 (2), 242–258.
Akhtaruzzaman, M., Shamsuddin, A., Easton, S., 2014. Dynamic correlation analysis of spill-over effects of interest rate risk and return on Australian and US financial firms. J. Int. Financ. Markets Inst. Money 31, 378–396.
Aktas, E., 2011. Systematic factors, information release and market volatility. Appl. Financ. Econ. 21 (6), 415–420.
Almeida, A., Goodhart, C., Payne, R., 1998. The effects of macroeconomic news on high frequency exchange rate behavior. J. Financ. Quant. Anal. 33 (03), 383–408.
Andersen, T.G., Bollerslev, T., 1998. Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. J. Finance 53 (1), 219–265.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2003. Micro effects of macro announcements: real-time price discovery in foreign exchange. Am. Econ. Rev. 93 (1), 38–62.
Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. J. Econom. 169 (1), 75–93.
Balduzzi, P., Elton, E.J., Green, T.C., 2001. Economic news and bond prices: evidence from the US Treasury market. J. Financ. Quant. Anal. 36 (04), 523–543.
Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. J. Financ. Econom. 2 (1), 1–37.
Bartolini, L., Goldberg, L.S., Sacarny, A., 2008. How economic news moves markets. Curr. Issues Econ. Finance 14 (6).
Bauwens, L., Omrane, W.B., Giot, P., 2005. News announcements, market activity and volatility in the euro/dollar foreign exchange market. J. Int. Money Finance 24 (7), 1108–1125.
Beechey, M.J., Wright, J.H., 2009. The high-frequency impact of news on long-term yields and forward rates: is it real? J. Monet. Econ. 56 (4), 535–544.
Beine, M., Laurent, S., Palm, F.C., 2009. Central bank FOREX interventions assessed using realized moments. J. Int. Financ. Markets Inst. Money 19 (1), 112–117.
Bekaert, G., Engstrom, E., 2010. Inflation and the stock market: understanding the "Fed Model". J. Monet. Econ. 57 (3), 278–294.
Bernanke, B.S., Kuttner, K.N., 2005. What explains the stock market's reaction to Federal Reserve policy? J. Finance 60 (3), 1221–1257.
Birkelund, O.H., Haugom, E., Molnár, P., Opdal, M., Westgaard, S., 2015. A comparison of implied and realized volatility in the Nordic power forward market. Energy Econ. 48, 288–294.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31 (3), 307–327.
Bomfim, A.N., 2003. Pre-announcement effects, news effects, and volatility: monetary policy and stock market. J. Bank. Finance 27 (1), 133–151.
Boyd, J.H., Hu, J., Jagannathan, R., 2005. The stock market's reaction to unemployment news: why bad news is usually good for stocks. J. Finance 60 (2), 649–672.
Chae, J., 2005. Trading volume, information asymmetry, and trading information. J. Finance 60 (1), 413–442.
Christie-David, R., Chaudhry, M., Khan, W., 2002. News releases, market integration, and market leadership. J. Financ. Res. 25 (2), 223–245.
Chudik, A., Pesaran, M.H., 2015. Common correlated effects estimation of heterogenous dynamic panel data models with weakly exogenous regressors. J. Econom. 188 (2), 393–420.
Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 7 (2), 174–196.
Ditzen, J., 2016. xtdcce: estimating dynamic common correlated effects in Stata. United Kingdom Stata Users' Group Meetings 2016 (No. 08). Stata Users Group.
Dominguez, K.M., 1998a. Central bank intervention and exchange rate volatility. J. Int. Money Finance 17 (1), 161–190.
Dominguez, K.M., 1998b. The European Central Bank, the Euro, and global financial markets. J. Econ. Perspect. 20 (4), 67–88.
Dominguez, K.M., 2006. When do central bank interventions influence intra-daily and longer-term exchange rate movements? J. Int. Money Finance 25 (7), 1051–1071.
Ederington, L.H., Lee, J.H., 1993. How markets process information: news releases and volatility. J. Finance 48 (4), 1161–1191.
Ehrmann, M., Fratzscher, M., 2005. Exchange rates and fundamentals: new evidence from real-time data. J. Int. Money Finance 24 (2), 317–341.
Ehrmann, Michael, Fratzscher, Marcel, 2004. Taking stock: Monetary policy transmission to equity markets. J. Money Credit Bank. 36, 719{737.
Escanciano, J.C., Lobato, I.N., 2009. An automatic Portmanteau test for serial correlation. J. Econ. 151 (2), 140–149.
Evans, K., Speight, A., 2010a. International macroeconomic announcements and intraday euro exchange rate volatility. J. Jpn. Int. Econ. 24 (4), 552–568.
Evans, K.P., Speight, A.E., 2010b. Dynamic news effects in high frequency Euro exchange rates. J. Int. Financ. Markets Inst. Money, 20238–20258.
Faust, J., Rogers, J.H., Wang, S.Y.B., Wright, J.H., 2007. The high-frequency response of exchange rates and interest rates to macroeconomic announcements. J. Monet. Econ. 54 (4), 1051–1068.
Flannery, M.J., Protopapadakis, A.A., 2002. Macroeconomic factors do influence aggregate stock returns. Rev. Financ. Stud. 15 (3), 751–782.
Fleming, M.J., Remolona, E.M., 1997. What moves the bond market? Econ. Policy Rev. 3 (4).
Fleming, M.J., Remolona, E.M., 1999. Price formation and liquidity in the US Treasury market: the response to public information. J. Finance 54 (5), 1901–1915.
Füss, R., Mager, F., Wohlenberg, H., Zhao, L., 2011. The impact of macroeconomic announcements on implied volatility. Appl. Financ. Econ. 21 (21), 1571–1580.
Giovannini, A., Jorion, P., 1989. Interest rates and risk premia in the stock market and in the foreign exchange market. J. Int. Money Finance 6 (1), 107–123.
Gürkaynak, R.S., Sack, B., Swanson, E., 2005. The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models. Am. Econ. Rev. 95 (1), 425–436.
Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. J. Bus. Econ. Stat. 24 (2), 127–161.
Harvey, C.R., Huang, R.D., 1991. Volatility in the foreign currency futures market. Rev. Financ. Stud. 4 (3), 543–569.
Hussain, S.M., 2011. Simultaneous monetary policy announcements and international stock markets response: an intraday analysis. J. Bank. Finance 35 (3), 752–764.
Jiang, G.J., Konstantinidi, E., Skiadopoulos, G., 2012. Volatility spillovers and the effect of news announcements. J. Bank. Finance 36 (8), 2260–2273.
Johnson, N.L., 1949a. Systems of frequency curves generated by method of translation. Biometrika 36 (1/2), 149–176.
Johnson, N.L., 1949b. Bivariate distributions based on simple translation systems. Biometrika 36 (3/4), 297–304.
Kim, S.J., McKenzie, M.D., Faff, R.W., 2004. Macroeconomic news announcements and the role of expectations: evidence for US bonds, stock and foreign exchange markets. J. Multinatl. Financ. Manage. 14 (3), 217–232.
Krieger, K., Mauck, N., Vazquez, J., 2015. Comparing US and European market volatility responses to interest rate policy announcements. Int. Rev. Financ. Anal. 39, 127–136.
Marshall, A., Musayev, T., Pinto, H., Tang, L., 2012. Impact of news announcements on the foreign exchange implied volatility. J. Int. Financ. Markets Inst. Money 22 (4), 719–737.
Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V., 1997. Volatilities of different time resolutions – analyzing the dynamics of market components. J. Empirical Finance 4, 213–239.
Neely, C.J., Dey, S.R., 2010. A survey of announcement effects on foreign exchange returns. Review, 92.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59 (2), 347–370.
Nikkinen, J., Sahlstrőm, P., 2004. Scheduled domestic and US macroeconomic news and stock valuation in Europe. J. Multinatl. Financ. Manage. 13 (3), 201–215.
Nikkinen, J., Sahlström, P., 2001. Impact of scheduled US macroeconomic news on stock market uncertainty: a multinational perspective. Multinatl. Finance J. 5 (2), 129–148.

Nikkinen, J., Omran, M., Sahlström, P., Äijö, J., 2006. Global stock market reactions to scheduled US macroeconomic news announcements. Global Finance J. 17 (1), 92–104.
Omrane, B.W., Hafner, C., 2015. Macroeconomic news surprises and volatility spillover in foreign exchange markets. Empirical Econ. 48 (2), 577–607.
El Ouadghiri, I., Uctum, R., 2016. Jumps in equilibrium prices and asymmetric news in foreign exchange markets. Econ. Model. 54 (April), 218–234.
El Ouadghiri, I., Mignon, V., Boitout, N., 2016. On the impact of macroeconomic news surprises on Treasury-bond returns. Ann. Finance 12 (1), 29–53.
Patton, A.J., Sheppard, K., 2009. Optimal combinations of realized volatility estimators. Int. J. Forecast. 25 (2), 218–238.
Petralias, A., Dellaportas, P., 2015. Volatility prediction based on scheduled macroeconomic announcements. Can. J. Stat. 43 (2), 199–223.
Schwartz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6 (2), 461–464.
Shephard, N., Sheppard, K., 2010. Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. J. Appl. Econom. 24 (2), 197–231.
Sweeney, R.J., Warga, A.D., 1986. The pricing of interest-rate risk: evidence from the stock market. J. Finance 41 (2), 393–410.
Thorbecke, W., 1997. On stock market returns and monetary policy. J. Finance 52 (2), 635–654.
Thornton, D.L., 1998. The information content of discount rate announcements: what is behind the announcement effect? J. Bank. Finance 22 (1), 83–108.
White, H., 1982. Maximum likelihood estimation of misspecified models. Econom.: J. Econom. Soc., 1–25

Chapter 6

Exploiting dependence: Day-ahead volatility forecasting for crude oil and natural gas exchange-traded funds

Energy 155 (2018) 462–473

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Exploiting dependence: Day-ahead volatility forecasting for crude oil and natural gas exchange-traded funds

Štefan Lyócsa [a, b, *], Peter Molnár [c]

[a] Institute of Economic Research, Slovak Academy of Sciences, Slovakia
[b] Institute of Financial Complex Systems, Masaryk University, Brno, Czech Republic
[c] University of Stavanger, UiS Business School, Stavanger, Norway

ARTICLE INFO

Article history:
Available online 3 May 2018

JEL Classification:
C53
Q02

Keywords:
Oil
Natural gas
Volatility forecasting
High-frequency data
ETF

ABSTRACT

This paper investigates volatility forecasting for crude oil and natural gas. The main objective of our research is to determine whether the heterogeneous autoregressive (*HAR*) model of Corsi (2009) can be outperformed by harnessing information from a related energy commodity. We find that on average, information from related commodity does not improve volatility forecasts, whether we consider a multivariate model, or various univariate models that include this information. However, superior volatility forecasts are produced by combining forecasts from various models. As a result, information from the related commodity can be still useful, because it allows us to construct wider range of possible models, and averaging across various models improves forecasts. Therefore, for somebody interested in precise volatility forecasts of crude oil or natural gas, we recommend to focus on model averaging instead of just including information from related commodity in a single forecast model.

1. Introduction

Crude oil and natural gas have been known to mankind for a few thousand years. However, their importance was negligible until the middle of the nineteenth century, when crude oil refinement was developed. Since this time, the importance of these commodities has significantly increased.

Because it can be easily transported, crude oil has rapidly become a globally traded commodity. Conversely, the transportation of natural gas is more complicated. As a result, several fragmented markets instead of a single market exist for natural gas. The price of natural gas varies significantly by location; its price is often set in relation to the price of oil. Due to technological progress in the transportation and extraction of natural gas from unconventional sources (shale gas), natural gas is becoming an increasingly important global commodity. Since worldwide reserves of natural gas are estimated to persist for a substantially longer period of time than oil reserves, natural gas may become a more important commodity than crude oil.

Understanding and forecasting the volatility of oil and gas is crucial for risk management, hedging purposes, cost management for oil and gas consumer/customers, and policy makers. For example, investors may target certain risk levels, which are influenced by the forecasted level of volatility. Volatility is important for pricing financial derivatives. Risk managers may monitor levels of expected value at risk, which can be calculated based on forecasted volatility. Commodities are often uncorrelated or even negatively correlated with equity markets [1–6]; and [7] and therefore can be employed to diversify investment portfolios. As a result, investors are becoming increasingly more interested in understanding and forecasting the volatility of commodities. Price volatility impacts the costs of companies. When the volatility of prices is greater, the financial planning is less reliable. Consequently, oil price shocks or the volatility of oil prices may influence stock markets [8–10]. Policy makers are interested in energy price volatilities because shocks to these important commodities tend to influence the real economy; for some countries, the total output of the economy is substantially dependent on oil and gas prices and their volatility [11,12].

The main research objective of this study is to determine whether volatility forecasts of oil and natural gas prices based on the heterogeneous autoregressive (HAR) model of Corsi [13] which is known to perform very well, can be outperformed by harnessing

* Corresponding author. Institute of Economic Research (EU SAV), Slovak Academy of Sciences, Sancova No. 56, 811 05, Slovakia.
E-mail addresses: stefan.lyocsa@savba.sk (Š. Lyócsa), peter.molnar@uis.no (P. Molnár).

https://doi.org/10.1016/j.energy.2018.04.194
0360-5442/

information from a related energy commodity. We perform this evaluation using the two commodities. We either include information from the previous volatility of the other commodity in the volatility forecasting models (univariate models) or we jointly forecast the two volatilities (multivariate model) by exploring their likely contemporaneous dependence. Because practitioners face model choice uncertainty, we also compare our forecasts with the forecasts provided by combined forecasts.

Common economic sense dictates that natural gas and oil prices should be related. In many instances, they are employed for similar purposes; therefore, they are substitutes. However, natural gas cannot be transported and stored as easily and as cheaply as oil. For a long time, oil used to be a global commodity, whereas natural gas used to be traded and priced locally, often in a relation to the oil price. However, approximately 15 years ago, large improvements in hydraulic fracking emerged. Hydraulic fracking is a technology which allows extraction of natural gas from resources where it was previously not possible. Large increase in supply of natural gas caused decoupling of oil and gas prices and transformed the gas market tremendously, from a highly fragmented market to a global market. Since natural gas is becoming a global commodity, understanding and forecasting its volatility have become increasingly important. However, whether information from oil market can improve forecasts of the gas volatility, and vice versa, is an empirical question which has not been addressed in the literature adequately.

Several studies have investigated the empirical relationship between the price of natural gas and the price of oil. According to Batten et al. [14]; the relationship between oil and natural gas is unstable; however, the majority of empirical studies report that oil leads natural gas. Batten et al. [14] have discovered that the relationship between oil and natural gas has weakened particularly after 2006/2007. They obtained somewhat surprising results—natural gas tends to lead the oil market. The long-term relationship between natural gas and oil was examined by Brigida [15] using a regime-switching model that allows changes in the long-term relationship (cointegration) between natural gas prices and oil prices. Brigida [15] argues that the time-varying relationship between oil and natural gas may be attributed to changes in the equilibrium states between the prices of the two commodities. Similarly, Caporin and Fontini [16] explore the long-run relationship between oil and natural gas and the changes in the long-term relationship when shale gas production arrived in the market. They discover not only that shale gas has a negative impact (as expected) on gas prices but also that the role of oil in the price formation of natural gas increased. Ramberg and Parson [17] challenge the existence of the long-term relationship between natural gas and oil and attribute the changes in their relationship to the fact that natural gas exhibits large short-term volatility. In a recent study, Bunn et al. [18] explored the dependence on the development of prices of U.S. oil and gas futures and discovered that the two commodity futures are connected and their connectedness increases when investors tend to make speculative bets and decreases when hedging activity increases.

Due to its importance, oil volatility has been the focus of many studies. The most common approach is to use a generalized autoregressive conditional heteroscedasticity (*GARCH*) type model that is based on daily data [19–23] to forecast oil volatility. Lin and Wesseh [24] studied the volatility of natural gas within a generalized autoregressive conditional heteroscedasticity (*GARCH*) framework and linked volatility shocks to major demand and supply shocks related to geopolitical, technological and natural events. Chan and Grant [25] modeled the volatility of several commodities, such as a series of oil and natural gas price returns, using *GARCH* and stochastic volatility models. They revealed that stochastic volatility models tend to outperform *GARCH* models, asymmetric volatility is important for the oil price volatility, and models that control for sudden changes in returns (jumps) are important for both oil and natural gas volatility models. GARCH models are also utilized by van Goor and Scholtens [26] who assess whether market fundamentals can explain the volatility of natural gas.

The majority of existing studies about modeling or forecasting volatility on energy markets is based on multivariate *GARCH* models, which are based on daily price data, e.g., Chang et al. [27,28]; Efimova and Serletis [29]. Alternatively, Vo [30] utilizes a stochastic volatility model and discovers that previous volatility on the stock market (oil market) improves volatility predictions on the oil market (stock market). However, compared with the *GARCH* models that are based on daily data, the realized volatility (*RV*) calculated from high-frequency data [31] generates more precise volatility models.[1] These models have been applied not only to stock markets [32] and exchange rates [33,34] but also to commodities [35], particularly oil and natural gas [12,36–38]. However, the interplay between the volatility of natural gas and the volatility of oil has not been explored. One exception is Degiannakis and Filis [39]; who suggest that *HAR* models of oil volatility augmented with exogenous volatilities from different asset classes usually outperform the standard *HAR* model. However, natural gas volatility proved to have minimal impact on the volatility of oil.

In this study, we contribute to this new strand of literature as we empirically explore whether the link between oil and natural gas can be exploited when forecasting the day-ahead volatility of these two energy commodities. We employ models that exploit not only the cross-lagged connectedness between related assets but also their contemporaneous dependence. We utilize the concept of realized volatility to forecast the volatility of crude oil and natural gas and determine whether the multivariate Generalized Heterogenous Autoregressive (*GHAR*) model of Čech and Baruník [40]; which includes the joint estimation of oil and gas volatilities and dependence, performs better than its univariate counterpart—the *HAR* model of Corsi [13].

Our second contribution is our focus on the combined forecasts from several models, whereas the majority of existing studies either apply a particular volatility model or compare several models. We assess whether the previous volatility (or its component) of a related commodity helps to increase the accuracy of the volatility forecasts. Volatility components, such as semi-variances, continuous and jump components, asymmetric volatility components, non-trading day terms, and measurement errors of the realized volatility are utilized.

Our third contribution is our study of the volatility of the two largest oil and gas exchange traded funds (ETF)—the United States Oil Fund and the United States Natural Gas Fund—although the majority of the existing literature concerns the volatility of oil and gas futures. Both of these funds have existed for almost ten years; their popularity and importance are rapidly increasing. A study based on exchange traded products complements the existing literature. Since both ETFs that we study primarily invest in futures contracts, the direct use of ETFs and futures contracts would likely lead to similar results. However, ETFs have two advantages. First, unlike futures contracts, ETFs have no pre-specified expiration date and are continuously traded. In the case of a futures time series, the data over multiple years are not one-time series. This long-term time series is always created from several futures contracts,

[1] Another alternative approach to GARCH and stochastic volatility models, which are both based on daily observations of prices, is the use of implied volatility, as suggested by e.g. Liu et al. [78]. However, an implied volatility index for natural gas does not exist.

usually the most liquid contracts. A researcher often makes arbitrary choices about rollover dates. Therefore, some artificial effects are possible, for example, volatility may increase of decrease prior to a rollover. The use of exchange traded funds is more convenient and removes an arbitrary choice of rollover dates from a researcher.

Our results have implications for the volatility transmission literature as we study how lagged volatility of the related commodity is related to out-of-sample volatility forecasts of another commodity. Our results suggest that the two markets are related, but harnessing information from related commodity does not improve volatility forecasts much. Contemporaneous dependence among volatilities is not useful in an out-of-sample framework. However, significant and consistent improvements in forecasting accuracy are recorded when combination forecasts are considered.

The remainder of the paper is organized as follows: section 2 describes the data, section 3 explains the methodology, section 4 presents the results, and section 5 presents the conclusion.

2. Data

We employ high-frequency data from the Trade and Quote (TAQ) database (consolidated trades) to forecast the realized volatility from 9:30 a.m. to 4:00 p.m. for two commodity ETFs: Natural Gas (United States Natural Gas - UNG) and Oil (United States Oil - USO). Both commodities were issued by the United States Commodity Funds. In terms of size, both ETFs are the largest among their competing peers. The UNG is an ETF that offers exposure to natural gas futures (near-month future contracts), and the USO is exposed to WTI crude oil futures.

Our sample covers the period from January 2008 to December 2014. The start of our sample was chosen because the inception of the UNG ETF was in 2007, and the inception of the USO was in 2006. The end of our dataset corresponds to the available data in the TAQ database at the time of our analysis. Our forecasting analysis begins with the first out-of-sample data from early 2010 because previous data are used to estimate initial forecasting models in the rolling window scheme employed in this study.

High-frequency data may be subject to unknown errors (wrong entries) or missing data, which may influence our analysis. Therefore, our first step was performing the following data-cleaning procedures:

1) From the downloaded data, all entries outside of the trading window, i.e., before 9:30 a.m. and after 4:00 p.m. on a given calendar day, were removed.
2) Zero price entries were removed.
3) Ten-minute sampling frequency was performed; the last price in a given sampling window (40 price observations per trading window) was stored.
4) If more than 20% of the observations in a given trading day were missing, the trading day was removed.
5) If more than 10% of the consecutive observations in a given trading day were missing, the trading day was removed.
6) Missing values in a sampling window from 9:30:00 a.m. until 9:39:59 a.m. were inputted using the first known price on a given trading day. Missing values in a sampling window from 3:50:00 p.m. until 3:59:59 p.m. were inputted using the last known price on a given trading day.
7) All other remaining missing values were inputted using linear interpolation methods.
8) Prior to the analysis, both ETFs (UNG, USO) were synchronized across trading days via a list-wise deletion to ensure that our results can be compared across univariate and multivariate models.

The trading day constitutes of 6.5 h of trading. We employ a 10-min sampling frequency that yields $N = 39$ sampling windows. Recording the opening price (at $j = 1$) and subsequent last prices of the 39 sampling windows generates 40 prices for a trading day. Let $P_{t,j}$ denote the value of the ETF, where t is a given day, and j is a given intraday sampling window. The continuous intraday return is defined as

$$r_{t,j} = \ln\left(\frac{P_{t,j}}{P_{t,j-1}}\right) \tag{1}$$

where $P_{t,j}$ is the price of the j-th sampling window, and $j = 0,1, \ldots,39$. The use of 10-min sampling windows is a compromise between the possible effect of market microstructure noise on our volatility estimates and the precision of our estimate. Our choice of a 10-min sampling window was motivated by the fact that ETFs are relatively new financial assets, with a recent inception date, which might cause smaller liquidity, particularly at the beginning of our series. The employment of the more popular 5-min sampling frequency would therefore produce many intraday “data-holes”, particularly at the beginning of our sample period of the natural gas ETF.

Our main quantity of interest is the intraday realized volatility, which is defined as follows:

$$RV_t = \sqrt{252\sum_{j=1}^{N} r_{t,j}^2} \tag{2}$$

We model the annualized (multiplication coefficient of 252) realized volatility instead of the realized variance (the square of realized volatility). Due to the extreme skewness and non-normality, the empirical literature usually prefers direct modeling the log of realized variance instead of realized variance. Our choice of realized volatility is attributed to the fact that factors from decomposed Cholesky variance-covariance matrices, which are similar in scale to the realized volatility, are being modeled in the *GHAR* model of Čech and Baruník [40].

The data-cleaning and synchronization procedures yielded data for 1753 trading days (68,367 return observations). After calculating the variables of interest and estimating initial forecasting models, 627 out-of-sample forecasts remained.

3. Methodology

One-day-ahead volatility forecasts for USO and UNG are investigated within the univariate [13] and multivariate *HAR* models [40]. First, in the empirical part, we demonstrate that our benchmark *HAR* model was rarely outperformed by more sophisticated univariate models. Even inclusion of the lagged realized volatility from the strategically linked commodity market has not improved volatility forecasts. Second, we demonstrate that multivariate *HAR* models outperform univariate models, including the benchmark. Last, we construct combination forecasts from univariate models and volatility forecasts that are not worse than the multivariate *HAR* forecasts. The following subsections describe our forecasting equations and variables, including criteria (loss functions) and procedures for comparing competing volatility forecasts.

3.1. Univariate heterogeneous autoregressive models

3.1.1. The benchmark model

Our benchmark model is the *HAR* model of Corsi [13]; which is extensively employed in the literature because it is significantly simpler than the *GARCH* model of Bollerslev [41]. Due to utilization of realized volatility calculated from high-frequency data, its forecasts of intraday volatility tend to be significantly better. The *HAR* model has the following structure

$$RV_{t+1} = \beta_1 + \beta_2 RV_t + \beta_3 RV_{t,t-4} + \beta_4 RV_{t,t-21} + e_t \quad (3)$$

where RV_{t+1} is the realized volatility at the $t + 1$ trading day, regressed on the realized volatility of the previous trading day RV_t, the average weekly realized volatility $RV_{t,t-4} = (RV_t + RV_{t-1} + \ldots + RV_{t-4})/5$, and the average monthly realized volatility $RV_{t,t-21} = (RV_t + RV_{t-1} + \ldots + RV_{t-21})/22$. According to Corsi [13]; the inclusion of weekly and monthly average realized volatilities should conveniently capture the long-memory dependence, which is often observed for volatility, including oil price volatility [42]. Compared with the *GARCH* models, *HAR* models are simply estimated via the ordinary least squares method.

3.1.2. Good news and bad news model: realized semi-variances

According to several papers, considering positive and negative semi-variances (defined below) may improve the in- or out-of-sample forecasting performance [23,43]. To capture this asymmetric volatility, we follow the recent work of Patton and Sheppard and employ positive and negative semi-variances:

$$\begin{aligned} RV_t^{2(-)} &= \sum_{j=1}^{N} r_{t,j}^2 \times I(r_{t,j} < 0) \\ RV_t^{2(+)} &= \sum_{j=1}^{N} r_{t,j}^2 \times I(r_{t,j} \geq 0) \end{aligned} \quad (4)$$

In Eq. (4), $RV_t^{(-)}$ is the negative semi-variance and $RV_t^{(+)}$ is the positive semi-variance, where I(.) is the indicator function that returns the value of 1 if the condition applies and returns the value of 0 otherwise. One method of using semi-variances in *HAR* models is to add the signed jump, which is the difference between a positive semi-variance and a negative semi-variance [44]. Because we are using realized volatilities, our measure of signed jumps SJ_t is defined as follows:

$$SJ_t = \begin{cases} \sqrt{RV_t^{2(+)} - RV_t^{2(-)}} & , RV_t^{2(+)} - RV_t^{2(-)} > 0 \\ -\sqrt{RV_t^{2(-)} - RV_t^{2(+)}} & , RV_t^{2(+)} - RV_t^{2(-)} < 0 \end{cases} \quad (5)$$

Adding lagged signed jumps in Eq. (3) yields the following model, which is denoted by *HAR-SJ*:

$$RV_{t+1} = \beta_1 + \beta_2 RV_t + \beta_3 RV_{t,t-4} + \beta_4 RV_{t,t-21} + \beta_5 SJ_{t-1} + e_t \quad (6)$$

3.1.3. Leverage model

Motivated by the *HAR* model and extensions by McAleer and Medeiros [45]; Corsi and Reno [46]; and a study of Chan and Grant [25] we add to the *HAR* model variables constructed from past returns. We include the return in absolute value and the interaction term of an absolute return and the indicator function I(.) that returns 1 if the intraday return r_{t-1} is negative and 0 otherwise. We denote this model as the *L-HAR*-model:

$$\begin{aligned} RV_{t+1} = {} & \beta_1 + \beta_2 RV_t + \beta_3 RV_{t,t-4} + \beta_4 RV_{t,t-21} + \beta_5 |r_{t-1}| \\ & + \beta_6 |r_{t-1}| I(r_{t-1} < 0) + e_t \end{aligned} \quad (7)$$

If we include the term $|r_{t-1}|I(r_{t-1}<0)$ only, the β_5 coefficient may be overestimated due to returns that are high in absolute value, which usually occurs during periods of high volatility. We control for this possibility by including the term $|r_{t-1}|$.

3.1.4. Model with continuous and jump components

The total variation of the log price process is the sum of the continuous and discontinuous (jump) components [47]. To disentangle the two components from the total variation, Barndorff-Nielsen and Shephard [47] proposed a consistent estimator of the continuous component, even in the presence of jumps—*bi-power variation*. In this study, we apply a more recent approach by Andersen et al. [48]; who proposed estimating the variance of the continuous component via the *median realized variance* (*MRV*), which should yield an estimate with improved finite-sample properties in the presence of small returns and jumps:

$$MRV_t^2 = \left(\frac{\pi}{6 - 4\sqrt{3} + \pi}\right)\left(\frac{N}{N-2}\right)\sum_{j=2}^{N-2}\left(med\{|r_{t,j-1}|, |r_{t,j}|, |r_{t,j+1}|\}\right)^2 \quad (8)$$

[49] argued that estimating the jump component as the difference between the realized variance and the median realized variance may produce non-zero estimates in empirical applications, even if no jumps exist. As a remedy, we determine the presence of a jump component using a test by Andersen et al. [48]. The size of the jump component (JC_t) and continuous component (CC_t) for the trading day t is estimated as

$$JC_t = \sqrt{\max\left\{0, \left(RV_t^2 - MRV_t^2\right) \times I(JT_t > \phi_{1-\alpha})\right\}} \quad (9)$$

$$CC_t = \sqrt{MRV_t^2 \times I(JT_t > \phi_{1-\alpha}) + RV_t^2 \times I(JT_t \leq \phi_{1-\alpha})} \quad (10)$$

where $\phi_{1-\alpha}$ denotes the critical value of the standard normal distribution ($\alpha = 0.05$), I(.) is the indicator function that returns the value of 1 if the condition applies and 0 otherwise, and JT_t is the test statistic defined in Andersen et al. [48].

Sudden price changes were considered in volatility models by Chan and Grant [25]. In our setting, we substitute the lagged jumps and the continuous components into Eq. (3) instead of the realized volatilities; we obtain the *HAR-CJ* model:

$$RV_{t+1} = \beta_1 + \beta_2 JC_t + \beta_3 CC_t + \beta_4 RV_{t,t-4} + \beta_5 RV_{t,t-21} + e_t \quad (11)$$

3.1.5. Model with a non-trading day component

According to Lyócsa and Molnár [50]; use of the simple autoregessive volatility models causes underestimated volatility persistence because weekends and holidays cause gaps in consecutive trading days. Using data from equity markets around the world, their empirical findings suggest that a simple adjustment of the *HAR* model improves the in- and out-of-sample model fit. We follow their recommendation and consider the *HAR-NT* model

$$\begin{aligned} RV_{t+1} = {} & \beta_1 + \beta_2 RV_t + \beta_3 RV_t \times I(\Delta(t, t-1) \geq 3) + \beta_4 RV_{t,t-4} \\ & + \beta_5 RV_{t,t-21} + e_t \end{aligned} \quad (12)$$

where $\Delta(t,t-1)$ denotes the calendar-day difference between two consecutive trading days. In almost all instances, the value of three corresponds to trading gaps due to weekends. The inclusion of the interaction terms should produce higher estimates of the β_2 coefficients in Eq. (12) compared with the standard *HAR* model defined in Eq. (3).

3.1.6. Model with components that depend on measurement error

Bollerslev et al. [51], have argued that although it is generally recognized that realized volatility is estimated with measurement error, empirical studies assume that the measurement error is constant. Bollerslev et al. [51], demonstrated that a higher variance of measurement error causes less predictable realized volatility. The *HAR-Q* model of Bollerslev et al. [51], gives higher weight to realized volatility during days of lower measurement error and vice versa. Our version of the *HAR-Q* that models the realized volatility is as follows:

$$RV_{t+1} = \beta_1 + RV_t\left(\beta_2 + \beta_3\sqrt{RQ_t}\right) + \beta_4 RV_{t,t-4} + \beta_5 RV_{t,t-21} + e_t \quad (13)$$

The term RQ_t is the realized quarticity that estimates the uncertainty of the estimator of realized volatility (see Ref. [51]:

$$RQ_t = \frac{N}{3}\sum_{j=1}^{N} r_{t,j}^4 \quad (14)$$

3.1.7. The effect of related lagged realized volatilities

In the subsequent analysis (Table 2), the *HAR* (Eq. (3)), *HAR-SJ* (Eq. (6)), *L-HAR* (Eq. (7)), *HAR-CJ* (Eq. (11)), *HAR-NT* (Eq. (12)), and *HAR-Q* (Eq. (13)) models are denoted by standard models. Each of the six models are augmented by adding the realized volatility component of the other commodity and are denoted by *HAR-X*, *HAR-SJ-X*, *L-HAR-X*, *HAR-CJ-X*, *HAR-NT-X*, and *HAR-Q-X*. For example, in the case of the *HAR* model that forecasts the realized volatility of the natural gas ETF, the *HAR-X* model includes the lagged realized volatility of the oil ETF. Similarly, the *HAR-SJ-X* model includes the signed jump component of the volatility of the oil. These simple augmentations produce a total of twelve univariate *HAR* models.

The augmentation is motivated by the fact that the two commodities may be perceived as imperfect substitutes either in real life or financial markets for speculative or hedging purposes although they may differ. Therefore, volatility spillovers between the two markets may occur in both directions, and thus, adding lagged volatilities may improve the forecasting performance (e.g., [20,29]. For example, Lin and Li [53] obtained evidence of volatility spillover from the oil to the natural gas market in the United States and Europe but not in Japan. A similar augmentation was recently employed by Degiannakis and Filis [39]; who augmented the *HAR* model with three volatility components of the related commodity, namely, the lagged *RV*, weakly average *RV* and monthly average *RV*.

3.2. Generalized heterogeneous autoregressive model

To model the joint dynamics and volatility of a multiple time series, Čech and Baruník [40] proposed the generalized heterogeneous autoregressive (*GHAR* model. This modeling strategy might be advantageous if the evolution of a multiple time series is determined by some common factors. Additionally, Čech and Baruník (2017) proposed modeling elements of a suitable variance-covariance matrix within the seemingly unrelated regression (SUR) framework to exploit cross-sectional dependence among regression errors.

The *GHAR* system models the individual daily “volatilities” and the dependency between two assets—natural gas and oil—and produces a system of three equations. In the following subsections, we describe the *GHAR* model in a more general setting for *N* assets.

We follow the work of Barndorff-Nielsen et al. [54] and estimate the daily realized variance-covariance matrix via the semi-definite Multivariate Realized Kernel (*MRK*) estimator.

$$\widehat{\boldsymbol{\Sigma}}_t^{MRK} = \sum_{h=-N}^{N} k\left(\frac{h}{H}\right)\boldsymbol{\Gamma}_{t,h} \quad (15)$$

were $k(.)$ is the kernel-weighting function. We employed the recommended Parzen scheme:

$$\boldsymbol{\Gamma}_{t,h} = \sum_{j=h+1}^{N} \mathbf{r}_{t,j}\mathbf{r}_{t,j-h}^{\mathrm{T}}, \quad h \geq 0 \quad (16)$$

where $\mathbf{r}_{t,j}$ denotes a column vector of intraday continuous returns of a given set of M assets at day t, and h is the bandwidth parameter. For $h<0$, $\boldsymbol{\Gamma}_{t,h} = \boldsymbol{\Gamma}_{t,-h}^{\mathrm{T}}$. Following Chiriac and Voev [55]; Čech and Baruník [40] did not directly model the elements of the $M \times M$ $\widehat{\boldsymbol{\Sigma}}_t^{MRK}$ matrix. To ensure positive semi-definiteness of the resulting forecasts, we decompose the $\widehat{\boldsymbol{\Sigma}}_t^{MRK}$ matrix via Cholesky factorization $\mathbf{P}_t\mathbf{P}_t^{\mathrm{T}} = \widehat{\boldsymbol{\Sigma}}_t^{MRK}$ and define a $q(q+1)/2$ column vector:

$$\mathbf{X}_t = vech(\mathbf{P}_t) \quad (17)$$

where vech(.) creates a column vector whose elements correspond to the lower triangular elements of $\mathbf{P}_t$. The system of forecasting regressions is given by:

$$\begin{pmatrix} \mathbf{X} \\ . \\ \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_{1,t} & 0 & 0 \\ 0 & . & 0 \\ 0 & 0 & \mathbf{Z}_{m,t} \end{pmatrix}\begin{pmatrix} \boldsymbol{\beta} \\ . \\ \boldsymbol{\beta} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon} \\ . \\ \boldsymbol{\varepsilon} \end{pmatrix} \quad (18)$$

Vector $\mathbf{Z}_{i,t}$ consists of $(\mathbf{e}\ \mathbf{X}_{i,t}\ \mathbf{X}_{i,t,t-4}\ \mathbf{X}_{i,t,t-21})$, where the elements of $\mathbf{X}_{i,t,t-4}$ and $\mathbf{X}_{i,t,t-21}$ correspond to the average values over the previous 5 trading days and 22 trading days, respectively; $\mathbf{e}$ is a vector of ones; $\boldsymbol{\beta}$ is a column vector of coefficients (including the constant); and $\boldsymbol{\varepsilon}_i$ is the vector of error terms. Following Čech and Baruník [40]; we estimate this system within the SUR framework with the generalized least squares estimation method.

Compared with the automatic bandwidth parameter procedure employed by Čech and Baruník [40]; we apply a more heuristic approach to estimate the bandwidth parameter h. We consider bandwidths in the range from one to eight. Each forecast that employs a *GHAR* model is estimated with the bandwidth parameter h, which produced the lowest mean square forecast error in a sample of the most recent 126 out-of-sample forecasts.

To produce an out-of-sample forecast with univariate *HAR* models, we do not rely on these procedures, and therefore, we have 126 more out-of-sample forecasts for the univariate *HAR* models than for the *GHAR* models. To facilitate the comparison between univariate models and multivariate *HAR* models, we match out-of-sample forecasts to ensure that they correspond to the same trading days.

3.3. Forecast combination

Forecasters face uncertainty about the variables to use or the specifications to choose. In practice, these uncertainties create several competing models that produce different forecasts. Using ideas originated by Bates and Granger [56]; Timmermann [57] argued that if individual forecasts are unbiased and not highly correlated, some combination of these forecasts will produce unbiased forecasts with a smaller forecast error, i.e., an even better forecast.

Our choice of the combination function is the simple average. The first combination forecast $F_{Wo,t+1}$ averages forecasts from standard six *HAR* models:

$$F_{Wo,t+1} = |K|^{-1}\sum_{k\in K} F_{k,t+1} \tag{19}$$

where K denotes the set of six univariate models, $|.|$ denotes the cardinality of the set, and $F_{k,t+1}$ denotes the kth individual forecast. Combination forecasts of natural gas realized volatility from Eq. (19) are denoted by *WoO* (without oil) and combination forecasts of realized volatility of Oil are denoted by *WoG* (without gas).

The second combination forecast $F_{W,t+1}$ averages forecasts from a univariate six HAR model that are augmented by the lagged volatility component(s) of related commodity's realized volatility:

$$F_{W,t+1} = |L|^{-1}\sum_{l\in L} F_{l,t+1} \tag{20}$$

were L denotes the set of six augmented univariate *HAR* models. Combination forecasts of natural gas realized volatility from Eq. (20) are denoted by *WO* (*with oil*), and combination forecasts of the realized volatility of oil are denoted by *WG* (with gas).

The third combination forecast $F_{all,t+1}$ averages the previous two combination forecasts and the multivariate forecast produced from the *GHAR* model:

$$F_{all,t+1} = \left(F_{W,t+1} + F_{Wo,t+1} + F_{GHAR,t+1}\right)/3 \tag{21}$$

3.4. Forecast evaluation

We evaluate forecasts based on a distance measure between the forecast of volatility and the proxy of the volatility. The distance function is referred to as a loss function. The forecasting ability is evaluated using three loss functions: Squared Forecast Error (*SFE*), *QLIKE*, and absolute forecast error (*AFE*):

$$L^{SFE}_{t+1} = \left(F_{t+1} - RV_{t+1}\right)^2 \tag{22}$$

$$L^{QLIKE}_{t+1} = \frac{RV_{t+1}}{F_{t+1}} - \ln\left(\frac{RV_{t+1}}{F_{t+1}}\right) - 1 \tag{23}$$

$$L^{AFE}_{t+1} = \left|F_{t+1} - RV_{t+1}\right| \tag{24}$$

where RV_{t+1} is the 10-min realized volatility (Eq. (2)), which is our proxy for the true unobserved market volatility [58], and F_{t+1} is a given forecast of volatility. We use the *QLIKE* loss functions in Christoffersen [59]; p. 85).

We employ all three loss functions because they give weight to different aspects of forecast errors. In risk management, large forecast errors may be disproportionally penalized, and therefore, the *SFE* is often used in practice. Rare extreme forecast errors are often observed in studies that address forecasting of market volatility, which might be attributed to the nature of markets, where unexpected events might cause extreme price movements that cannot be captured by forecasting models. *AFE* is more balanced because it is more robust in the presence of outliers. Patton [60] advocates the use of *QLIKE* and *SFE* because it provides consistent model rankings even when a forecasted variable is measured with noise. We include *QLIKE* in our set of loss functions. Similar choices for loss functions are found in Hansen and Lunde [58]; Borovkova and Mahakena [61]; and Golosnoy and Okhrin [62].

Let $L_{A,t}$ represent the value of a given loss function for model A at day t and $L_{B,t}$ represent the value of a given loss function for model B at day t. The loss differential is defined as follows:

$$d_{A,B,t} = L_{A,t} - L_{B,t} \tag{25}$$

The size of the average loss differential is statistically evaluated using the Hansen et al. [52], model confidence set, which is a test of the superior predictive ability of one or more models within a given set of models. We test the following hypotheses:

$$\begin{aligned} H_0 &: E\left[d_{A,B,t}\right] = 0 \\ H_0 &: E\left[d_{A,B,t}\right] \neq 0 \end{aligned} \tag{26}$$

We always compare two forecasting models. The first forecasting model is the benchmark forecast, which is given by the standard *HAR* model; the second forecasting model is given by the competing model (or combination of forecasts). The Hansen et al. [52], model confidence set runs the following test statistics[2]:

$$t_{A,B} = \frac{\overline{d}_{A,B}}{\sqrt{\widehat{V}\left(\overline{d}_{A,B}\right)}} \tag{27}$$

where $\overline{d}_{A,B}$ is the average value of the loss differential, and the volatility of $\overline{d}_{A,B}$ in the denominator is estimated via the block-bootstrap. The block length p is set to the maximum number of significant parameters of a simple $AR(p)$ model of $d_{A,B,t}$ [52]. Using the bootstrap with 10,000 bootstrap samples, we estimate the distribution of the test statistics under the null hypothesis and derive the p-values for each model comparison test.

3.5. Forecasting procedure

The forecasting results presented in sections 4.2–4.4 are based on a rolling forecast scheme with an estimation window of 1000 observations. Using the first 1000 observations, we estimate all univariate *HAR* models ($12 \times 2 = 24$) and *GHAR* models (eight for each bandwidth parameter). Using the estimated coefficients and the last thousand observations, we forecast the realized volatility for the 1001st observation. The combination forecasts are also calculated from these forecasts. We repeat the procedure and roll the estimation window one observation ahead; i.e., we use the 2nd to 1001st observations to estimate the univariate and multivariate models that are employed to forecast the realized volatility of the 1002nd observation. We continue until we reach the end of our sample. However, our forecast evaluation starts with the 127th forecast because the first 126 forecasts are used to choose a *GHAR* model with a given bandwidth parameter that yielded the lowest average squared forecast error over the previous 126 forecasts (refer to the end of section 3.2 for explanation).

Our choice of 1000 observations in the estimation window is motivated by the fact that natural gas and oil are subject to seasonal demands. Using the estimation window that is less than a year is problematic. We assumed that several years are needed to mitigate a possible effect of seasonality on the oil and natural gas market. The argument is similar to the argument of Narayan and Gupta [63]; who used 50% of the total sample to create the first forecast, which considered large oil shocks. Previous empirical literature on forecasting volatility also employed similarly sized estimation windows (e.g., Sevi [30] [3].

[2] This test was performed using the procedures developed by Bernardi and Catania [79] for program R.

[3] We run our analysis also with estimation windows of size 500 and 750. Qualitatively, the results were similar. However, forecasts were more accurate (lower average forecast errors) with a window size of 1000. The results are available upon request.

The choice for the rolling forecast scheme is motivated by the fact that this approach is better suited when parameter changes are expected [64]. As shown in previous studies (e.g. [15], the relationships between natural gas and oil are unstable, and a rolling window forecasting scheme is more suitable to capture structural changes than an expanding window forecasting scheme. An evaluation of forecasts relies on the test by Narayan and Gupta [52]; which explicitly assumes a rolling forecast scheme.

3.6. Accounting for possibility of structural changes

Recent developments suggest that accounting for structural breaks makes forecasts more precise (e.g. Refs. [65–71]. First, we estimate all univariate *HAR* models through ordinary least squares (OLS) ignoring possible existence of breaks in parameters. To account for a possibility of a structural break with an unknown date, we use an approach suggested by Ref. [68]. We estimate the *HAR* model through weighted least squares (WLS), where we use robust weights (see Eq. 44 and Eq. 48 in Ref. [68]. The weights are a monotonically increasing function of time, where most recent observation has the highest weight. Comparison of the OLS and WLS forecasts revealed that this strategy is advantageous for all loss functions. For the gas volatility, the forecast errors from WLS were on average 1.6% smaller than with OLS and for the oil, forecast errors were in average 3.0% smaller. These improvements are not negligible, and we therefore decided to report results only for the WLS estimation.

4. Results

4.1. Characteristics of the realized volatility of natural gas and oil

In this section, we present some features and stylized facts of our data. In Table 1 we find calculated summary statistics of the realized volatility for the crude oil and natural gas ETFs. In Fig. 1, we observe that the realized volatility declines because the beginning of our sample includes the global financial crisis that produced turbulent periods across asset classes around the world.

Table 1 shows notable differences between oil and gas. Existing literature revealed that natural gas is subject to higher short-term volatility. For example, Misund and Oglend [73] found that already daily fluctuations in gas demand might impact price volatility of the gas. Previous studies have indicated that the volatility of natural gas is larger than the volatility of oil (e.g. Ref. [17,24]. We confirm that this finding also holds for the volatility of ETFs because the mean, median and lower and upper quartiles of realized volatility are larger for natural gas than for oil. In addition, the jump component of volatility is 82.5% larger for natural gas than for oil, whereas approximately 19.7% of the total volatility can be attributed to the jump component for natural gas but only 14.6% in the case of oil. As shown in Fig. 1, price variations due to jumps frequently occur. Both jump components show minimal persistence, similarly as in Andersen et al., [74]; who observed small persistence for foreign exchange, S&P 500 and 30-year Treasury bond volatility and Lyócsa and Molnár [75]; who observed minimal persistence in jump components in the gold and silver ETFs or jump components of realized volatility of future contracts of non-ferrous metals, e.g., nickel, zinc, aluminum, copper and lead [76]. The first-order autocorrelation for the volatility of natural gas (0.35) is significantly smaller than the first-order autocorrelation for the volatility of oil (0.81).

The key point is that the predictability of the volatility of oil and natural gas is mitigated by the presence of jumps because jumps reveal minimal persistence. Because jumps are relatively more important for natural gas, the results in Table 1 suggest that forecasting volatility for natural gas will be more difficult than forecasting crude oil. Our study confirmed this suggestion.

4.2. Forecasts of the realized volatility of natural gas

On the left side of Fig. 2, we plotted the realized volatility (black line) and forecasted volatility from the model (Eq. (20), red line). The realized and forecasted values reveal that forecasts are unable to catch sudden spikes in volatility of the natural gas. This finding suggests a positive association between forecast error and the size of the realized volatility, which is not surprising because *HAR* and

Table 1
Summary statistics for realized volatility.

	Mean	Median	SD	IQ	IIIQ	Kurt.	Skew.	ρ(1)	
Panel A: Natural gas									
R	−0.06	−0.05	2.03	−1.18	1.07	5.71	0.29	−0.07	***
RV	29.71	26.65	14.10	20.32	35.45	14.32	2.21	0.35	***
CC	27.16	24.84	11.57	18.98	32.77	6.39	1.45	0.54	***
JC	5.84	0.00	13.25	0.00	0.00	25.78	3.66	−0.04	**
SJ	−1.47	−4.58	20.47	−13.9	11.82	7.51	0.41	−0.07	***
AAR	0.78	0.05	1.22	0.00	1.18	7.71	2.07	−0.04	
NT	5.62	0.00	11.99	0.00	0.00	7.83	2.20	−0.22	***
RV_{RQ}	328.77	192.86	561.40	109.24	360.12	304.26	13.1	0.07	***
Panel B: Oil									
R	0.00	0.06	1.60	−0.76	0.80	7.93	−0.07	0.01	
RV	21.98	18.52	13.07	13.37	25.82	7.79	1.98	0.81	***
CC	20.82	17.49	12.36	12.58	24.55	7.59	1.92	0.81	***
JC	3.20	0.00	7.59	0.00	0.00	19.21	3.51	0.05	
SJ	−0.45	0.82	14.34	−9.25	8.68	4.82	−0.10	0.00	
AAR	0.56	0.00	1.01	0.00	0.76	14.19	2.87	0.12	***
NT	4.33	0.00	10.20	0.00	0.00	14.23	3.07	−0.18	***
RV_{RQ}	181.18	91.88	277.77	48.02	186.82	25.60	4.17	0.69	***

Notes: The results in this table were produced using 1753 synchronized observations. In the columns, IQ and IIIQ denote the first quartile and third quartile, respectively; Kurt. And Skew. Are the kurtosis and skewness, respectively; ρ(1) is the value of the first-order autocorrelation coefficient; and *** and ** denote the statistical significance at the 1% significance level and 5% significance level, respectively, based on the Escanciano and Lobato [72] serial correlation test. In the row, R denotes the intraday return; RV denotes the realized volatility (Eq. (2)); CC is the continuous component (Eq. (10)) of the realized volatility; JC is the jump component (Eq. (9)) of the realized volatility; SJ is the signed jump (Eq. (5)); AAR is the absolute value of the intraday return for days when returns are negative and 0 otherwise; the asymmetric component NT is the non-trading component, which is the annualized realized volatility after a non-trading weekend and 0 otherwise, and RV_{RQ} is the interaction between the realized volatility and the realized quarticity.

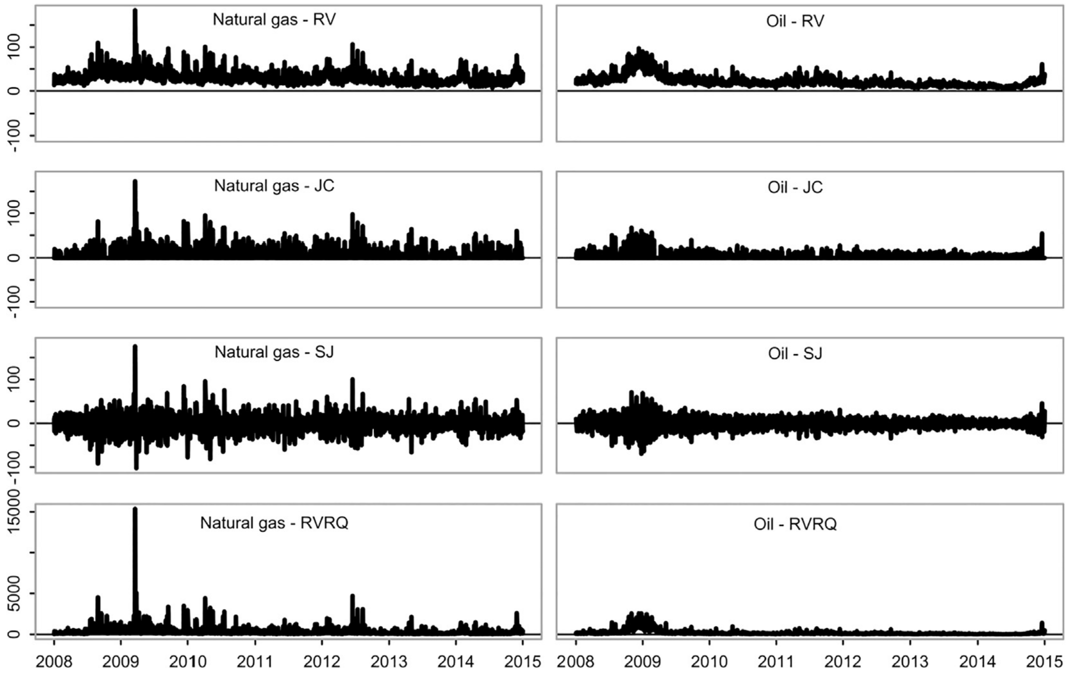

Fig. 1. Evolution of realized volatility (*RV*), jump component (*JC*), signed jump (*SJ*) and interaction between realized volatility and realized quarticity (RV_{RQ}).

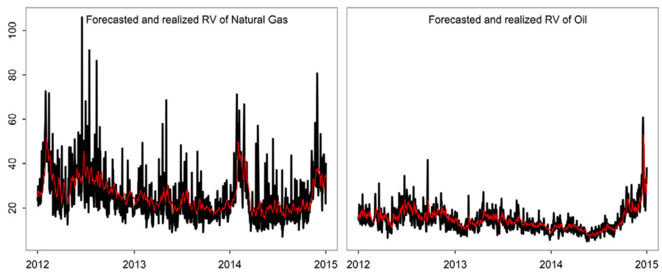

Fig. 2. Forecasted (red) and realized (black) values from combination forecasts (*WO* – left – and *WG* models – right).

GHAR models are based on historical volatility using linear autoregessive volatility specification. Thus, when the true volatility is unusually high, forecasts from these models will be unusually low.

The takeaway from Fig. 2 is that the evaluation of forecasts needs to be performed across different evaluation functions. As previously noted in section 3.3, the *SFE* will penalize extreme errors more than the *AFE*, and the *QLIKE* loss function uses ratios of the actual and forecasted realized volatility and may mitigate the weight of these extreme forecast errors when evaluating forecasts.

In Table 2, we present the average values of the loss functions. For example, the value of 93.19 (HAR model in *Panel A* and column *MSFE*) is the average of the squared forecast error between the forecast produced using the benchmark HAR model (Eq. (3)) and the proxy of the true realized volatility (Eq. (2)). The statistical significance corresponds to a test that assesses whether an alternative forecast outperforms the benchmark forecast represented by the previously mentioned *HAR* model. For example, in *Panel A*, we observe that *L-HAR* and *HAR-Q* produced forecasts with lower forecast error regardless of the employed loss function. Several of these improvements were also statistically significant. Taking the average of the forecasts from all standard models yields the combination forecast (Eq. (19)), which also produced statistically improved forecasts.

Does the lagged realized volatility of the oil contain valuable information for day-ahead forecasts of the volatility of natural gas? In *Panel B*, we observe that results are mixed as compared to the benchmark *HAR* model forecasts were improved in above half of the cases. Forecast improvements appear to depend on the type of the *HAR* model. For example, a comparison of the *L-HAR* and *L-HAR-X* model reveals that the average loss functions of the *L-HAR-X* model are lower than the average loss functions of the *L-HAR* model. Note that the statistical test, whose results are reported in Table 2, compares each model with the benchmark *HAR* model (first row). Therefore, we also present the results in *Panel E*, where we compare the results from two combination forecasts: *WoO* does not include the volatility of the oil, and *WO* does include the volatility of oil. This finding enables us to form general conclusions about the usefulness of lagged realized volatility of oil for forecasting purposes of the next-day realized volatility of natural gas. The ratio is lower than 1, which indicates that *WO* produced a lower average loss function, but the improvement is statistically significant only for the *AFE* loss function. It therefore appears that the lagged realized volatility of oil has only limited value for predicting day-ahead realized volatility of natural gas.

Does contemporaneous dependence between the volatility of oil and the volatility of gas improve day-ahead forecasts of the volatility of natural gas? We refer to the multivariate *GHAR* model that jointly models the volatility of oil and gas. Compared with the benchmark forecasts of the *HAR* model, the values of *MSFE* and *QLIKE* deteriorated (In Panel C of Table 2). Improvements were obtained for *MAFE*, but were not statistically significant. The multivariate framework of the *GHAR* model efficiently estimates the common factor that is responsible for part of the dynamics of the volatilities of both commodities (joint estimation via SUR and inclusion of the covariance equation) compared with univariate models. When these common dynamics drive volatility dynamics, forecasts may be improved. As Fig. 1 suggests, larger forecast errors are associated with events when the volatility is large. During these periods, volatility is predominantly driven by idiosyncratic factors, and therefore, *GHAR* might generate less efficient forecasts. Because

Table 2
Forecasting performance of *HAR*, *GHAR* and combination forecasts.

	Natural gas				Oil		
	MSFE	QLIKE	MAFE		MSFE	QLIKE	MAFE
Panel A: Standard models							
HAR	93.19	0.0633	6.81	*HAR*	18.29	0.0326	3.02
HAR-CJ	**91.18***	0.0621	6.75	*HAR-CJ*	18.06	0.0325	3.01
HAR-SJ	92.86	0.0632	**6.78****	*HAR-SJ*	**17.76*****	**0.0319*****	**2.98*****
L-HAR	**92.07***	**0.0628***	**6.72*****	*L-HAR*	17.59	0.0322	**2.97****
HAR-NT	94.08	0.0638	6.9	*HAR-NT*	18.02	0.0329	3.03
HAR-Q	**91.86****	0.0627	**6.76***	*HAR-Q*	18.42	0.0326	3.03
Averaged forecast							
Without OIL (WoO)	**91.91*****	**0.0625*****	**6.76*****	*Without GAS (WoG)*	**17.87****	**0.0322*****	**2.99*****
Panel B: Models that utilize another commodity							
HAR-X	93.26	0.0635	**6.77***	*HAR-X*	18.25	0.0325	3.02
HAR-CJ-X	91.78	0.0628	6.77	*HAR-CJ-X*	**17.84***	**0.0321****	**2.99****
HAR-SJ-X	92.81	0.0634	**6.78***	*HAR-SJ-X*	**17.76*****	**0.0319*****	**2.98*****
L-HAR-X	91.64	0.0627	**6.70*****	*L-HAR-X*	17.61	0.0322	**2.98****
HAR-NT-X	93.75	0.0637	6.91	*HAR-NT-X*	17.96	0.0331	3.04
HAR-Q-X	92.05	0.0630	**6.75***	*HAR-Q-X*	18.33	**0.0324****	3.02
Averaged forecast							
With OIL (WO)	**91.61*****	**0.0624*****	**6.74*****	*With GAS (WG)*	**17.79****	**0.0321*****	**2.99*****
Panel C: Multivariate model							
GHAR	95.94	0.0653	6.96	*GHAR*	19.00	0.0344	3.02
Panel D: Average across all models							
All models	92.71	0.063	6.80	*All models*	17.94	0.0325	**2.99****
Panel E: Comparing comb. fore.							
WO/WoO	0.997	0.998	**0.996***	*WG/WoG*	**0.995*****	**0.995*****	0.998
All models/WoO	1.009	1.008	1.005	*All models/WoG*	1.004	1.008	0.999

Note: In Panels A–E *, **, *** denotes the statistical significance of the Hansen et al. [52] model encompassing test at the 10%, 5%, and 1% significance levels, where we compare the corresponding model with the benchmark represented by the standard *HAR* model. In *Panel E* the Hansen et al.'s (2011) test predictive equality of predictive ability of two combination forecast models, where the bold values less than one correspond to a statistically significant improvement of the model in the numerator compared with the model in the denominator. *HAR-CJ* represents the model with continuous and jump components, *HAR-SJ* with signed jump, *L-HAR* leverage model, *HAR-NT* model with non-trading component, and the *HAR-Q* model with realized quarticity. *GHAR* represents the multivariate *HAR*. *WG* and *WO* denotes forecasting model that averages forecasts from standard models. *WoG* and *WoO* denotes the forecasting model that averages forecasts from models enhanced with natural gas or oil volatility component. *GHAR* represents the multivariate *HAR*. *All models* denote forecast that average either *WG*, *WoG*, and *GHAR* forecasts (for oil) or *WO*, *WoO*, and *GHAR* forecasts (for natural gas).

SFE gives more weight to forecast errors than *AFE*, *GHAR* has not produced improved forecasts for *SFE* but has produced improved forecast errors for the *AFE* loss function. Because the results are mixed, we conclude that exploiting the contemporaneous dependence between the volatility of oil and the volatility of natural gas does not improve the forecasts of natural gas.

Although forecasts from the *GHAR* do not appear to be particularly useful by themselves, the inclusion of these forecasts into a combination forecast may produce improved results. The *all models* (Eq. (21)) averages two combination forecasts and the *GHAR* forecast. The results are reported in *Panel D* of Table 2. Comparison of forecast errors between *WoO* (Eq. (19)) and *WO* (Eq. (20)) suggests that the inclusion of the *GHAR* into the combination forecast *all models* has not helped. In Panel E, we directly test the *WoO* forecast that disregards the realized volatility of oil with the *all models* that includes the realized volatility of oil. The ratio exceeds 1; thus, *all models* does not exhibit superior predictive ability compared to the *WoO*.

4.3. Forecasts of the realized volatility of oil

In the right panel of Fig. 2, we plotted the forecasted and realized volatility of oil (*WG* model). The forecasts track the general movement of volatility, and forecast errors tend to be larger for days with higher market volatility than for days with lower market volatility. The first striking difference between oil and natural gas volatility forecasts is that forecast errors for the oil (refer to Table 2) are substantially smaller than the forecast errors of natural gas regardless of the loss function. Table 2 also reveals that the alternative model specifications appear to be less relevant than the alternative model specifications for natural gas. *L-HAR* and *HAR-Q* models, which appeared to have some merit for predicting the volatility of natural gas, does not have the same merit when the volatility of oil is concerned. Only *HAR-SJ* (Eq. (6)) shows consistent and statistically significant improvements across loss functions. The choice of the 'right' model specification depends on the given asset. From a forecaster's perspective, a safer approach is to rely on combination forecasts. The *WG* model produces better forecasts than the forecasts of the benchmark model but does not improve the *HAR-SJ* models.

Does the lagged realized volatility of the natural gas contain valuable information for day-ahead forecasts of the volatility of oil? In *Panel B*, we observe that forecasts have improved compared with the benchmark *HAR* model. For example, *HAR-SJ* and *HAR-SJ-X* yielded the same average loss function; i.e., adding lagged *signed jump* of the natural gas to the *HAR-SJ* model of the volatility of oil produced comparable results. A comparison of the *WG* and *WoG* models (refer to *Panel E*) shows statistically significant improvement. We conclude that the lagged realized volatility of natural gas does produce improved forecasts of the day-ahead realized volatility of oil. However, the improvements are small. For example, comparing two combination forecasts, one without the information about the volatility on the natural gas market, with a forecast that incorporates information about the volatility on the natural gas market, led to a 0.5% improvement in forecast accuracy. The sublime role of the natural gas volatility is consistent with the findings of Degiannakis and Filis [39]; who found almost non-existent improvements of forecasting accuracy in a similar, augmented *HAR* model.

Does the contemporaneous dependence between the volatility of oil and the volatility of gas improve day-ahead forecasts of the volatility of oil? The *GHAR* model has not improved volatility forecasts when evaluated via *SFE* and *QLIKE* (refer to Panel C). Multivariate model led to similar forecasts as its univariate counterpart when forecasts were evaluated through the *MAFE* loss function. Given the *AFE* loss function, the combination of the *GHAR* with *WG* and *WoG* (model *all models*, Eq. (21)) also does not seem to improve forecasts that are solely based on *WG* (Panel E). The general contemporaneous dependence between natural gas volatility and oil volatility does not improve forecasts of oil volatility.

4.4. Discussion of our results based on existing research

Although numerous studies have focused on the connection between the oil market and the gas market general (refer to the introduction), we are not aware of any research that analyzes whether the connection between oil and gas can improve forecasts of realized volatility of these two commodities. An exception is a recent study by Degiannakis and Filis [40], who studied the role of the volatility of natural gas for predicting the volatility of oil. Bollerslev et al. [42] conclude that standardized realized volatility exhibit very similar patterns across various assets. Therefore, we compare our results with results from papers that explored questions similar to ours, even though the commodities (or assets) differed.

The research question in Lyócsa and Molnár [75] is similar to our research question; however, they study realized volatility forecasting for gold and silver. Similarly, they discover that the utilization of information from other commodities may improve volatility forecasts. A similar conclusion is obtained by Ref. [34] for currencies. Regarding the realized volatility forecasts of oil, Souček and Todorova [37] suggest that these forecasts can be improved by incorporating information from equity markets. Degiannakis and Filis [39] obtain similar results; however, they consider not only equities but also other asset classes, including commodities and currencies.

Our results, which indicate that averaging forecasts from various models produces better forecasts, are consistent with the results from previous studies (refer to e.g., Liu and Maheu [77,75]; Lyócsa et al., [76]; and also consistent with our expectations based on existing literature, but our study is the first to address the realized volatility of oil and gas.

5. Concluding remarks

We study the volatility forecasts of crude oil and natural gas for the period from 2008 to 2014. Our central question is whether the link between these two strategically linked commodities gas can be utilized when forecasting day-ahead volatility. In this study, we provide empirical evidence to this question using high-frequency data, twelve univariate models, one multivariate model, and three combination forecasting models, which belong to the class of *HAR* models. In addition to the benchmark model of Corsi [13]; we employ models that utilize signed jumps [44] or decomposition of the realized volatility into continuous and jump components [47,48], models that account for the measurement error of realized variance [51], models that seek to exploit leverage or the asymmetric volatility effect, and models that account for breaks in the time series due to non-trading days. Each of these specifications is extended by including lagged components of realized volatility from the related commodity—either natural gas (when forecasting volatility of oil) or oil (when forecasting volatility of gas). The multivariate model is based on the modeling of the Cholesky factors of the realized kernel-based covariance matrix within a seemingly unrelated regression framework, which was proposed by Čech and Baruník [40]. Our forecasting system of equations incorporates the two related commodities. Combination forecasts are also employed. The first combination forecast averages forecasts across six univariate *HAR* models that do not include volatility components from related commodity. The second combination forecasts averages forecasts across univariate models that include the volatility component from the related commodity. The third combination forecast averages the multivariate forecast with the two previous combination forecasts.

Our key result is that adding volatility components of the related commodity alone can hardly improve volatility forecasts. Forecast improvements differ across volatility models and two considered commodities. However, we found that combination of forecasts from various models tend to perform better than forecasts from individual models. We summarize our findings as follows:

- Adding a lagged volatility component of the related commodity (gas) is more advantageous when forecasting the volatility of oil.
- Exploiting the contemporaneous dependence between oil and natural gas in a multivariate (*GHAR*) model does not improve day-ahead volatility forecasting.
- Combination forecasts produce improved forecasts.
- It is still possible to harness information on volatility from related commodity when used within a framework of a combination forecast.

This paper can serve as a guideline for researchers who must precisely forecast the volatility of crude oil or natural gas. Although previous studies have demonstrated that the *HAR* model is a very good model for oil and gas, our results suggest that the volatilities of the two commodities are distinct enough, in that it seems difficult to exploit information about the volatility of the other commodity to improve volatility forecasts. Tangible, although relative small performance improvements can be exploited, only when forecast are combined from various forecasting models that also include information from the related commodity.

Acknowledgements

Mr. Lyócsa appreciates the support provided by the Slovak Research and Development Agency under contract No. APVV-14-0357 and the funding support of the GACR 18-05829S.

References

[1] Abanomey WS, Mathur I. International portfolios with commodity futures and currency forward contracts. J Invest 2001;10(3):61–8. https://doi.org/10.3905/joi.2001.319474.
[2] Erb CB, Harvey CR. The strategic and tactical value of commodity futures. Financ Anal J 2006;62(2):69–97. https://doi.org/10.2469/faj.v62.n2.4084.
[3] Georgiev G. Benefits of commodity investment. J Altern Invest 2001;4(1): 40–8. https://doi.org/10.3905/jai.2001.318997.
[4] Gorton G, Rouwenhorst KG. Facts and fantasies about commodity futures. Financ Anal J 2006;62(2):47–68. https://doi.org/10.2469/faj.v62.n2.4083.
[5] Miffre J, Rallis G. Momentum strategies in commodity futures markets. J Bank Finance 2007;31(6):1863–86. https://doi.org/10.1016/j.jbankfin.2006.12.005.
[6] Nijman T, Swinkels LLAP. Strategic and tactical allocation to commodities for retirement savings schemes. 2003. https://doi.org/10.2139/ssrn.425320.
[7] Fuertes AM, Miffre J, Rallis G. Tactical allocation in commodity futures markets: combining momentum and term structure signals. J Bank Finance 2010;34(10):2530–48. https://doi.org/10.1016/j.jbankfin.2010.04.009.
[8] Bouri E. Return and volatility linkages between oil prices and the Lebanese stock market in crisis periods. Energy 2015;89:365–71. https://doi.org/10.1016/j.energy.2015.05.121.
[9] Chiou JS, Lee YH. Jump dynamics and volatility: oil and the stock markets. Energy 2009;34(6):788–96. https://doi.org/10.1016/j.energy.2009.02.011.
[10] Wei Y, Guo X. Oil price shocks and China's stock market. Energy 2017. https://doi.org/10.1016/j.energy.2017.07.137.
[11] Ferderer JP. Oil price volatility and the macroeconomy. J Macroecon 1996;18(1):1–26. https://doi.org/10.1016/S0164-0704(96)80001-2.
[12] Rafiq S, Salim R, Bloch H. Impact of crude oil price volatility on economic activities: an empirical investigation in the Thai economy. Resour Pol 2009;34(3):121–32. https://doi.org/10.1016/j.resourpol.2008.09.001.

[13] Corsi F. A simple approximate long-memory model of realized volatility. J Financ Econom 2009;7(2):174–96. https://doi.org/10.1093/jjfinec/nbp001.
[14] Batten JA, Ciner C, Lucey BM. The dynamic linkages between crude oil and natural gas markets. Energy Econ 2017;62:155–70. https://doi.org/10.1016/j.eneco.2016.10.019.
[15] Brigida M. The switching relationship between natural gas and crude oil prices. Energy Econ 2014;43:48–55. https://doi.org/10.1016/j.eneco.2014.01.014.
[16] Caporin M, Fontini F. The long-run oil–natural gas price relationship and the shale gas revolution. Energy Econ 2017;64:511–9. https://doi.org/10.1016/j.eneco.2016.07.024.
[17] Ramberg DJ, Parsons JE. The weak tie between natural gas and oil prices. Ramberg. In: David J, Parsons, John E, editors. The weak tie between natural gas and oil prices (November 9, 2010). Center for Energy and Environmental Policy Research (CEEPR); 2010. https://doi.org/10.2139/ssrn.1710686. No. 10-017.
[18] Bunn DW, Chevallier J, Le Pen Y, Sevi B. Fundamental and financial influences on the co-movement of oil and gas prices. Energy J 2017;38(2). https://doi.org/10.5547/01956574.38.2.dbun.
[19] Arouri MEH, Lahiani A, Lévy A, Nguyen DK. Forecasting the conditional volatility of oil spot and futures prices with structural breaks and long memory models. Energy Econ 2012;34(1):283–93. https://doi.org/10.1016/j.eneco.2011.10.015.
[20] Ewing BT, Malik F, Ozfidan O. Volatility transmission in the oil and natural gas markets. Energy Econ 2002;24(6):525–38. https://doi.org/10.1016/S0140-9883(02)00060-9.
[21] Kang SH, Kang SM, Yoon SM. Forecasting volatility of crude oil markets. Energy Econ 2009;31(1):119–25. https://doi.org/10.1016/j.eneco.2008.09.006.
[22] Sadorsky P. Modeling and forecasting petroleum futures volatility. Energy Econ 2006;28(4):467–88. https://doi.org/10.1016/j.eneco.2006.04.005.
[23] Wei Y, Wang Y, Huang D. Forecasting crude oil market volatility: further evidence using GARCH-class models. Energy Econ 2010;32(6):1477–84. https://doi.org/10.1016/j.eneco.2010.07.009.
[24] Lin B, Wesseh PK. What causes price volatility and regime shifts in the natural gas market. Energy 2013;55:553–63. https://doi.org/10.1016/j.energy.2013.03.082.
[25] Chan JC, Grant AL. Modeling energy price dynamics: GARCH versus stochastic volatility. Energy Econ 2016;54:182–9. https://doi.org/10.1016/j.eneco.2015.12.003.
[26] van Goor H, Scholtens B. Modeling natural gas price volatility: the case of the UK gas market. Energy 2014;72:126–34. https://doi.org/10.1016/j.energy.2014.05.016.
[27] Chang CL, McAleer M, Tansuchat R. Crude oil hedging strategies using dynamic multivariate GARCH. Energy Econ 2011;33(5):912–23. https://doi.org/10.1016/j.eneco.2011.01.009.
[28] Wang Y, Wu C. Forecasting energy market volatility using GARCH models: can multivariate models beat univariate models? Energy Econ 2012;34(6):2167–81. https://doi.org/10.1016/j.eneco.2012.03.010.
[29] Efimova O, Serletis A. Energy markets volatility modelling using GARCH. Energy Econ 2014;43:264–73. https://doi.org/10.1016/j.eneco.2014.02.018.
[30] Vo M. Oil and stock market volatility: a multivariate stochastic volatility perspective. Energy Econ 2011;33(5):956–65. https://doi.org/10.1016/j.eneco.2011.03.005.
[31] Andersen TG, Bollerslev T. Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. J Finance 1998;53(1):219–65. https://doi.org/10.1111/0022-1082.85732.
[32] Christoffersen P, Jacobs K, Mimouni K. Volatility dynamics for the S&P500: evidence from realized volatility, daily returns, and option prices. Rev Financ Stud 2010;23(8):3141–89. https://doi.org/10.1093/rfs/hhq032.
[33] Andersen TG, Bollerslev T, Diebold FX, Labys P. The distribution of realized exchange rate volatility. J Am Stat Assoc 2001;96(453):42–55. https://doi.org/10.1198/016214501750332965.
[34] Lyócsa S, Molnár P, Fedorko I. Forecasting exchange rate volatility: the case of the Czech republic, Hungary and Poland. Czech J Econ Finan 2016;66(5):453.
[35] Birkelund OH, Haugom E, Molnár P, Opdal M, Westgaard S. A comparison of implied and realized volatility in the Nordic power forward market. Energy Econ 2015;48:288–94. https://doi.org/10.1016/j.eneco.2014.12.021.
[36] Haugom E, Langeland H, Molnár P, Westgaard S. Forecasting volatility of the US oil market. J Bank Finance 2014;47:1–14. https://doi.org/10.1016/j.jbankfin.2014.05.026.
[37] Souček M, Todorova N. Realized volatility transmission between crude oil and equity futures markets: a multivariate HAR approach. Energy Econ 2013;40:586–97. https://doi.org/10.1016/j.eneco.2013.08.011.
[38] Wang T, Wu J, Yang J. Realized volatility and correlation in energy futures markets. J Futures Market 2008;28(10):993–1011. https://doi.org/10.1002/fut.20347.
[39] Degiannakis S, Filis G. Forecasting oil price realized volatility using information channels from other asset classes. J Int Money Finance 2017;76:28–49. https://doi.org/10.1016/j.jimonfin.2017.05.006.
[40] Čech F, Baruník J. On the modelling and forecasting of multivariate realized volatility: generalized heterogeneous autoregressive (GHAR) model. J Forecast 2017;36(2):181–206. https://doi.org/10.1002/for.2423.
[41] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. J Econom 1986;31(3):307–27. https://doi.org/10.1016/0304-4076(86)90063-1.
[42] Gil-Alana LA, Gupta R, Olubusoye OE, Yaya OS. Time series analysis of persistence in crude oil price volatility across bull and bear regimes. Energy 2016;109:29–37. https://doi.org/10.1016/j.energy.2016.04.082.
[43] Sévi B. Forecasting the volatility of crude oil futures using intraday data. Eur J Oper Res 2014;235(3):643–59. https://doi.org/10.1016/j.ejor.2014.01.019.
[44] Patton AJ, Sheppard K. Good volatility, bad volatility: signed jumps and the persistence of volatility. Rev Econ Stat 2015;97(3):683–97. https://doi.org/10.1162/REST_a_00503.
[45] McAleer M, Medeiros MC. A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. J Econom 2008;147(1):104–19. https://doi.org/10.1016/j.jeconom.2008.09.032.
[46] Corsi F, Reno R. HAR volatility modelling with heterogeneous leverage and jumps. 2009. Available at: SSRN 1316953.
[47] Barndorff-Nielsen OE, Shephard N. Power and bipower variation with stochastic volatility and jumps. J Financ Econom 2004;2(1):1–37. https://doi.org/10.1093/jjfinec/nbh001.
[48] Andersen TG, Dobrev D, Schaumburg E. Jump-robust volatility estimation using nearest neighbor truncation. J Econom 2012;169(1):75–93. https://doi.org/10.1016/j.jeconom.2012.01.011.
[49] Beine M, Lahaye J, Laurent S, Neely CJ, Palm FC. Central bank intervention and exchange rate volatility, its continuous and jump components. Int J Finance Econ 2007;12(2):201–23. https://doi.org/10.1002/ijfe.330.
[50] Lyócsa Š, Molnár P. The effect of non-trading days on volatility forecasts in equity markets. Finance Res Lett 2017. https://doi.org/10.1016/j.frl.2017.07.002.
[51] Bollerslev T, Patton AJ, Quaedvlieg R. Exploiting the errors: a simple approach for improved volatility forecasting. J Econom 2016;192(1):1–18. https://doi.org/10.1016/j.jeconom.2015.10.007.
[52] Hansen PR, Lunde A, Nason JM. The model confidence set. Econometrica 2011;79(2):453–97. https://doi.org/10.3982/ECTA5771.
[53] Lin B, Li J. The spillover effects across natural gas and oil markets: based on the VEC–MGARCH framework. Appl Energy 2015;155:229–41. https://doi.org/10.1016/j.apenergy.2015.05.123.
[54] Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. J Econom 2011;162(2):149–69. https://doi.org/10.1016/j.jeconom.2010.07.009.
[55] Chiriac R, Voev V. Modelling and forecasting multivariate realized volatility. J Appl Econom 2011;26(6):922–47. https://doi.org/10.1002/jae.1152.
[56] Bates JM, Granger CWJ. The combination of forecasts. J Oper Res Soc 1969;20:451–68. https://doi.org/10.1057/jors.1969.103.
[57] Timmermann A. Forecast combinations. In: Elliott G, Granger C, Timmermann A, editors. Handb Econ Forecast, 1; 2006. p. 136–96.
[58] Hansen PR, Lunde A. Realized variance and market microstructure noise. J Bus Econ Stat 2006;24(2):127–61. https://doi.org/10.1198/073500106000000071.
[59] Christoffersen PF. Elements of financial risk management. second ed. Oxford: Academic Press; 2012.
[60] Patton AJ. Volatility forecast comparison using imperfect volatility proxies. J Econom 2011;160(1):246–56. https://doi.org/10.1016/j.jeconom.2010.03.034.
[61] Borovkova S, Mahakena D. News, volatility and jumps: the case of natural gas futures. Quant Finance 2015;15(7):1217–42. https://doi.org/10.1080/14697688.2014.986513.
[62] Golosnoy V, Okhrin Y. Using information quality for volatility model combinations. Quant Finance 2015;15(6):1055–73. https://doi.org/10.1080/14697688.2012.739728.
[63] Narayan PK, Gupta R. Has oil price predicted stock returns for over a century? Energy Econ 2015;48:18–23. https://doi.org/10.1016/j.eneco.2014.11.018.
[64] West KD. Forecast evaluation. In: Elliott G, Granger CWJ, Timmerman A, editors. Handbook of economic forecasting, vol. 1. Elsevier; 2006. p. 99–134.
[65] Giacomini R, Rossi B. Detecting and predicting forecast breakdowns. Rev Econ Stud 2009;76(2):669–705. https://doi.org/10.1111/j.1467-937X.2009.00545.x.
[66] Paye BS, Timmermann A. Instability of return prediction models. J Empir Finance 2006;13(3):274–315. https://doi.org/10.1016/j.jempfin.2005.11.001.
[67] Pesaran MH, Timmermann A. Selection of estimation window in the presence of breaks. J Econom 2007;137(1):134–61. https://doi.org/10.1016/j.jeconom.2006.03.010.
[68] Pesaran MH, Pick A, Pranovich M. Optimal forecasts in the presence of structural breaks. J Econom 2013;177(2):134–52. https://doi.org/10.1016/j.jeconom.2013.04.002.
[69] Rapach DE, Wohar ME. Structural breaks and predictive regression models of aggregate US stock returns. J Financ Econom 2006;4(2):238–74. https://doi.org/10.1093/jjfinec/nbj008.
[70] Wen F, Gong X, Cai S. Forecasting the volatility of crude oil futures using HAR-type models with structural breaks. Energy Econ 2016;59:400–13. https://doi.org/10.1016/j.eneco.2016.07.014.
[71] Yin A. Forecasting and model averaging with structural breaks. Graduate Theses Dissert 2015;14720. http://lib.dr.iastate.edu/etd/14720.
[72] Escanciano JC, Lobato IN. An automatic portmanteau test for serial correlation. J Econom 2009;151(2):140–9. https://doi.org/10.1016/j.jeconom.2009.03.001.
[73] Misund B, Oglend A. Supply and demand determinants of natural gas price volatility in the UK: a vector autoregression approach. Energy 2016;111:178–89. https://doi.org/10.1016/j.energy.2016.05.124.

[74] Andersen TG, Bollerslev T, Diebold FX. Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Rev Econ Stat 2007;89(4):701–20. https://doi.org/10.1162/rest.89.4.701.

[75] Lyócsa Š, Molnár P. Volatility forecasting of strategically linked commodity ETFs: gold-silver. Quant Finance 2016;16(12):1809–22. https://doi.org/10.1080/14697688.2016.1211799.

[76] Lyócsa Š, Molnár P, Todorova N. Volatility forecasting of non-ferrous metal futures: covariances, covariates or combinations? J Int Financ Market Inst Money 2017. https://doi.org/10.1016/j.intfin.2017.08.005.

[77] Liu C, Maheu JM. Forecasting realized volatility: a Bayesian model-averaging approach. J Appl Econom 2009;24(5):709–33. https://doi.org/10.1002/jae.1070.

[78] Liu ML, Ji Q, Fan Y. How does oil market uncertainty interact with other markets? An empirical analysis of implied volatility index. Energy 2013;55:860–8. https://doi.org/10.1016/j.energy.2013.04.037.

[79] Bernardi M, Catania L. The model confidence set package for R. 2014. arXiv:1410.85041.

Chapter 7

Oil market volatility and stock market volatility

Finance Research Letters 26 (2018) 204–214

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Oil market volatility and stock market volatility

Milan Bašta[a], Peter Molnár[*,b,c]

[a] *Faculty of Informatics and Statistics, Department of Statistics and Probability, University of Economics, Prague, Czech Republic*
[b] *University of Stavanger, UiS Business School, Stavanger, Norway*
[c] *Faculty of Finance and Accounting, Department of Monetary Theory and Policy, University of Economics, Prague, Czech Republic*

ARTICLE INFO

Keywords:
Volatility
Oil market
Stock market
VIX
OVX
Wavelets

ABSTRACT

This paper studies the comovement between volatility of the equity market and the oil market, both for implied and realized volatilities. The wavelet methodology enables us to study this relationship on various time scales. We find that there is a strong comovement between the volatilities of the two markets. However, this comovement is time-varying and depends on the time scale. It is strong at yearly horizon, but much weaker at horizons of a few days. Moreover, implied volatility of the stock market leads the implied volatility of the oil market, whereas no such relationship is observed for realized volatilities.

1. Introduction

Participants in financial markets are subject to the volatility of their investments. Therefore, volatility plays a crucial role in portfolio optimization, risk management, hedging, derivative pricing and particularly option pricing. In 1993, the CBOE introduced a volatility index for the US stock market, the VIX index. The VIX index, often called a fear index, has become one of the most followed indicators in the financial markets.

Due to the huge success and importance of the VIX index, similar indices have been introduced not only for other equity markets (Bugge et al., 2016), but also for commodities (Birkelund et al., 2015). The unique feature of implied volatility is that it is forward-looking, whereas volatility models based on historical data are backward-looking. The implied volatility is a measure of a market risk, and it is forward-looking since it contains investors expectation of future market changes. Hence, studies based on implied volatility can help us understand how the expectations about risk are transferred from one market to another. Such studies have become increasingly popular, for example, Sari et al. (2011) found that VIX had a significantly suppressing effect on oil prices in the long run and Qadan and Yagil (2012) concluded that VIX index has significant impact on gold prices.

In recent years, commodities have become a more and more important part of many portfolios. Crude oil is probably the most important commodity in the world. Oil has a weight above 50% in the general commodity index. Moreover, oil prices have a strong impact on many other commodities. Understanding oil price volatility is important for several reasons. Not only has oil price volatility an impact on company investments (Henriques and Sadorsky, 2011) and other macroeconomic variables (Rafiq et al., 2009), but it also has a direct impact on the economies of oil importing and oil exporting countries.

Since the stock market can be considered as a proxy for the general economy, and oil is the most important commodity, we therefore study the relationship between implied volatility for the equity market (the VIX index) and implied volatility for crude oil (the OVX index). There are several papers related to our study. Ji and Fan (2012) found that the crude oil market has significant volatility spillover effects on non-energy commodity markets and that the overall level of correlation strengthened after the crisis. Guo and Ji (2013) found a significant impact of Google search query volumes on oil volatility. Haugom et al. (2014) found that

* Corresponding author.
E-mail address: peter.molnar@uis.no (P. Molnár).

https://doi.org/10.1016/j.frl.2018.02.001
Received 10 August 2017; Received in revised form 5 February 2018; Accepted 5 February 2018
Available online 07 February 2018
1544-6123/

volatility model for oil is improved when implied volatility is included in the model. Dutta et al. (2017) found that the OVX index predicts volatility of Middle East and African stock markets. Luo and Qin (2017) studied the impact of oil volatility on the Chinese stock market and confirmed the importance of forward-looking implied volatility by finding that implied volatility has significant and negative effects on the Chinese stock market while the impact of realized volatility shocks is negligible.

Papers most closely related to our work are those by Liu et al. (2013) and Maghyereh et al. (2016). Liu et al. (2013) studied transmission between the oil implied volatility (OVX) and stock market implied volatility (VIX), euro/dollar exchange rate implied volatility (EVZ) and gold price implied volatility (GVZ). Maghyereh et al. (2016) analysed the relationship between implied volatility of oil and implied volatility of various stock markets. Both these papers are based on variance decomposition and analyse the relationships on daily frequency.

We utilize the wavelet methodology, which has the advantage of enabling us to investigate the relationship between variables on various time scales. In other words, we investigate the comovement and the lead-lag relationship between VIX and OVX not just on one arbitrary time scale (e.g. daily), but on various time scales from daily to yearly. We find that the implied volatility of the equity market (VIX) and the implied volatility of the oil market (OVX) are highly correlated.

Our main contribution is the finding that the relationship between implied volatility of oil and implied volatility of stock market depends on time scale. There is only contemporary correlation (no lead/lag relationship) between VIX and OVX on short time scales (high frequencies), but there is a significant lead/lag relationship on longer time scales (lower frequencies). Our results have implications both for general understanding of financial markets as well as for traders and other market participants exposed to the volatility of the oil market.

The rest of the paper is organized as follows. Section 2 introduces the continuous wavelet transform, wavelet power, squared wavelet coherence, relative phase and cross-wavelet gain. Section 3 presents the data used in the analysis. Section 4 reports and discusses the results of the analysis. Section 5 concludes.

2. Wavelet characteristics

The continuous wavelet transform (CWT) and measures derived from the CWT, such as the cross-wavelet transform, cross-wavelet power, squared wavelet coherence and relative phase (see e.g., Torrence and Compo, 1998 or Grinsted et al., 2004) provide a favorite set of tools used to explore time-varying relationships between two time series.

While introducing the continuous wavelet transform and related measures, we make use of the introduction available in Bašta et al. (2017). Specifically, we assume the Morlet wavelet defined as

$$\psi_0(\eta) = \pi^{-1/4}\exp(i\omega_0\eta)\exp\left(-\frac{1}{2}\eta^2\right), \tag{1}$$

where ω_0 is equal to 6 and where η is a dimensionless time (Grinsted et al., 2004).

The continuous wavelet transform of a time series $\{X_t: t = 0, ..., N-1\}$ of length N at time t and at scale $s > 0$ is defined as (Grinsted et al., 2004)

$$W_{t,s}^X = \frac{1}{\sqrt{s}}\sum_{k=0}^{N-1} X_k \psi_0^*\left(\frac{k-t}{s}\right), \tag{2}$$

where * denotes complex conjugation. The wavelet coefficient $W_{t,s}^X$ captures how scale s contributes to the dynamics of $\{X_t\}$ at time t. Since $W_{t,s}^X$ is complex, it is advantageous to introduce the wavelet power spectrum at time t and scale s defined as $|W_{t,s}^X|^2$, i.e. as the square of the modulus of $W_{t,s}^X$. $|W_{t,s}^X|^2$ is related to the variability of $\{X_t\}$ at time t and scale s, large/small values of $|W_{t,s}^X|^2$ meaning large/small variability of $\{X_t\}$ at time t and scale s. Since $|W_{t,s}^X|^2$ is biased in favour of large-scale features, Liu et al. (2007) suggested to use the *corrected* wavelet power spectrum given as $|W_{t,s}^X|^2/s$.

To study the comovement of two time series $\{X_t: t = 0, ..., N-1\}$ and $\{Y_t: t = 0, ..., N-1\}$, the cross-wavelet transform between $\{X_t\}$ and $\{Y_t\}$ can be used, being defined as (Grinsted et al., 2004)

$$W_{t,s}^{XY} = W_{t,s}^X W_{t,s}^{Y*}. \tag{3}$$

Generally, $W_{t,s}^{XY}$ is a complex number. The modulus of $W_{t,s}^{XY}$ is called the cross-wavelet power and can be considered as the absolute covariance between $\{X_t\}$ and $\{Y_t\}$ at time t and scale s. The argument of $W_{t,s}^{XY}$ is called the relative phase and can take any value from $-\pi$ to π. It captures the lead/lag relationship between $\{X_t\}$ and $\{Y_t\}$ at time t and scale s. More specifically, if relative phase is zero, no lead or lag is present. On the other hand, positive values of the relative phase imply that $\{X_t\}$ leads $\{Y_t\}$, while negative values imply that $\{X_t\}$ lags behind $\{Y_t\}$. The relative phase can be converted to time ΔT by which $\{X_t\}$ leads $\{Y_t\}$ (if $\Delta T > 0$), or lags behind it ($\Delta T < 0$). Specifically,

$$\Delta T = \frac{\arg(W_{t,s}^{XY})}{2\pi f}, \tag{4}$$

where f denotes the Fourier frequency associated with scale s which is, for the Morlet wavelet assumed in Eq. (1), given as (Torrence and Compo, 1998) $f = 1/(1.03\ s)$.

The squared wavelet coherence between two time series $\{X_t: t = 0, ..., N-1\}$ and $\{Y_t: t = 0, ..., N-1\}$ at time t and scale s is defined as (Grinsted et al., 2004)

$$R^2_{t,s} = \frac{\left| S\left(\frac{1}{s} W^{XY}_{t,s}\right) \right|^2}{S\left(\frac{1}{s} |W^X_{t,s}|^2\right) S\left(\frac{1}{s} |W^Y_{t,s}|^2\right)}, \tag{5}$$

where the S operator in Eq. (5) defines smoothing in time and scale. More specifically, the operator is defined as (Grinsted et al., 2004)

$$S(W_{t,s}) = S_{scale}(S_{time}(W_{t,s})), \tag{6}$$

where

$$S_{time}(W_{t,s})|_s = \left(W_{t,s} * c_1^{-\frac{t^2}{2s^2}} \right)\Bigg|_s, \tag{7}$$

$$S_{scale}(W_{t,s})|_t = (W_{t,s} * c_2 \Pi(0.6s))|_t, \tag{8}$$

where c_1 and c_2 are normalization constants and Π is the boxcar function. In Eqs. (7) and (8), the * operator denotes convolution. Squared wavelet coherence can attain any value in the range from 0 to 1 and can be considered as a local (at time t and scale s) squared correlation between the time series $\{X_t\}$ and $\{Y_t\}$. If the squared wavelet coherence is close to one, a strong linear relationship is suggested at the given time t and scale s. On the other hand, squared wavelet coherence close to zero denotes a very weak linear relationship between the time series at time t and scale s.

Ge (2008) suggested that relative phase, i.e. the argument of the cross-wavelet transform, should be explored (only) in those regions (in the time-scale plane) for which the values of the squared wavelet coherence are rather high.

Mandler and Scharnagl (2014) propose to use the cross-wavelet gain defined as

$$G_{t,s} = \frac{\left| S\left(\frac{1}{s} W^{XY}_{t,s}\right) \right|}{S\left(\frac{1}{s} |W^X_{t,s}|^2\right)}, \tag{9}$$

which can be interpreted as a local absolute value of the regression coefficient of $\{Y_t\}$ on $\{X_t\}$.

Since wavelet coefficients are calculated by linearly filtering the time series using a non-causal linear filter (see Eq. (2)), the coefficients, the (corrected) wavelet power spectrum, wavelet coherence, relative phase and cross-wavelet gain cannot be directly obtained for regions close to the beginning and the end of the time series. As a result, artificial boundary conditions have to be introduced so that all the characteristics associated with times corresponding to the beginning and the end of the time series can be calculated. The region where the values of these characteristics are affected to a non-negligible extent by these artificial boundary conditions is called the *cone of influence* and the results in the cone must be interpreted with caution since they may not reflect the true underlying dynamics.

3. Data

The data were downloaded from finance.yahoo.com. As previously mentioned, the VIX index records 30-day volatility implied by options written on the S&P 500 index. The OVX index records 30-day volatility implied by options written on the United States Oil Fund.

The United States Oil Fund (USO) is an exchange traded fund that seeks to provide investors with easy exposure to the oil market. Since investing in physical oil would be too costly, it invests in oil-related financial instruments, mostly oil futures. The USO's investment objective is to track percentage changes in the price of West Texas Intermediate (WTI) light, sweet crude oil delivered to Cushing, Oklahoma, as measured by the changes in the price of the futures contract for light, sweet crude oil traded on the New York Mercantile Exchange (the "NYMEX"), less the USO's expenses. The USO is the most actively traded commodity exchange traded fund, and is the 7th most traded of all the exchange traded funds.[1] We consider the daily time series of VIX and OVX for the period from May 10, 2007 till July 28, 2016.

Besides the implied volatility indices VIX and OVX, we also consider the time series of realized volatility of S&P 500 and USO in the period from May 10, 2007 till July 28, 2016. Even though our main interest is the comovement between the implied volatility of the oil and stock market, we conduct the same analysis for realized volatilities. Implied volatility reflects primarily the expectations about the future, whereas realized volatility reflects what really happened. Therefore, investigating both types of volatility provides us with a more complete picture of comovement between volatility of oil and stock market.

The Garman and Klass (1980) estimate of realized volatility (variance) for trading day t is adjusted for opening jump (Molnár, 2012) and calculated as:

$$GK_t = 0.5(h_t - l_t)^2 - (2\log 2 - 1)c_t^2 + j_t^2, \tag{10}$$

[1] According to http://etfdb.com/compare/volume/, accessed on July 30, 2017.

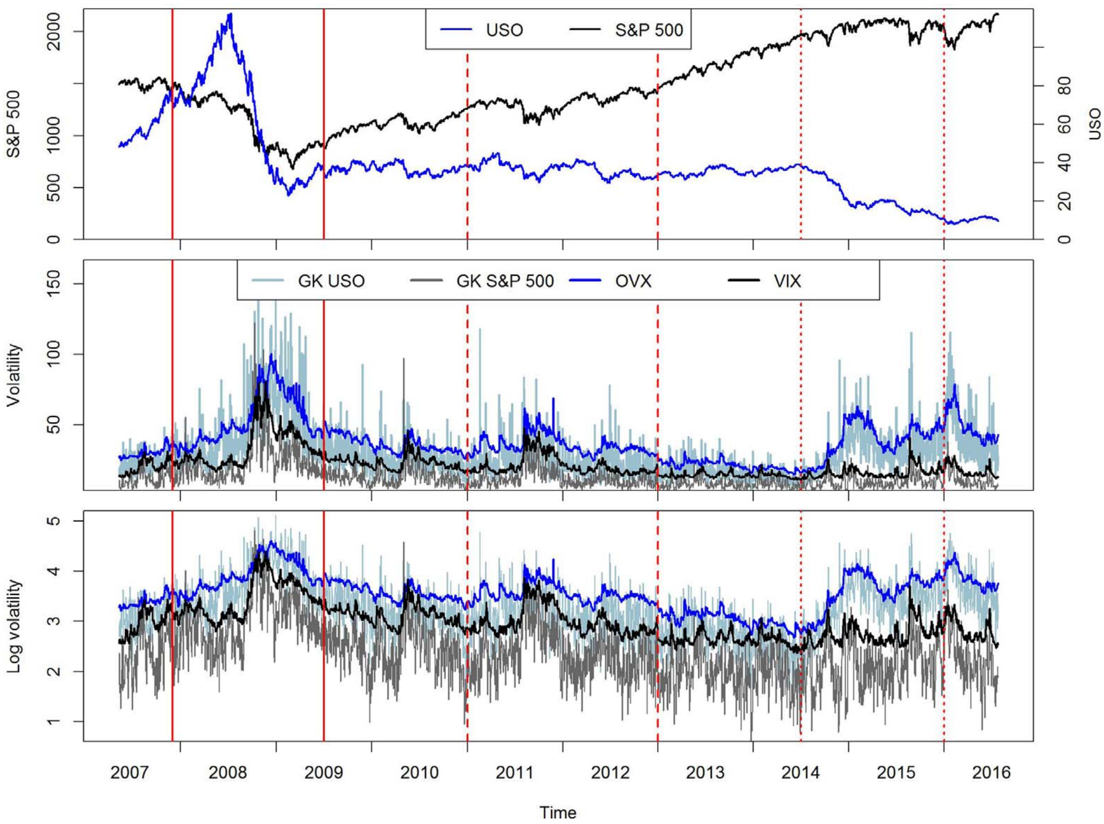

Fig. 1. *Top plot:* S&P 500 (black) and USO (blue). *Middle plot*: VIX (black), GK S&P 500 (gray), OVX (blue) and GK USO (lightblue). *Bottom plot*: LVIX (black), LGK S&P 500 (gray), LOVX (blue) and LGK USO (lightblue). The Great Recession, the Arab spring and the Oil price fall are depicted by the solid, dashed and dotted red vertical lines. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

where $c_t = \log(C_t) - \log(O_t)$, $h_t = \log(H_t) - \log(O_t)$, $l_t = \log(L_t) - \log(O_t)$ and $j_t = \log(O_t) - \log(C_{t-1})$, where log denotes the natural logarithm and C_t, O_t, H_t and L_t are the close, open, high and low price of the asset (S&P 500 or USO) for day t. Implied volatility measures volatility over a 30-day period over the whole 24 hours of each day, not just over periods from open to close. Therefore, opening jump needs to be included in the estimate of realized volatility to make it comparable with implied volatility.[2]

Implied volatility indices VIX and OVX are annualized standard deviations of the underlying assets (S&P 500 and USO) quoted as percentages, whereas the Garman–Klass estimates of realized volatilities of S&P 500 and USO calculated according to Eq. (10) are variances on a daily scale. Consequently, we take the square root of the time series of the Garman–Klass estimates of realized volatilities of S&P 500 and USO, multiply the results by the square root of 252 and further by 100, and refer to these newly obtained variables (time series) as GK S&P 500 and GK USO. GK S&P 500 and GK USO are comparable to VIX and OVX since they are measured on the same scale.

In the top plot of Fig. 1 S&P 500 and USO are plotted, in the middle plot VIX, OVX, GK S&P 500 and GK USO are displayed, and in the bottom plot of the figure the natural logarithms of VIX, OVX, GK S&P 500 and GK USO are presented, hereafter denoted as LVIX, LOVX, LGK S&P 500 and LGK USO. The summary statistics for the four time series in the middle plot of Fig. 1 are given in Table 1, whereas the summary statistics for the four time series in the bottom plot of Fig. 1 are given in Table 2.

We can observe that GK S&P 500 and GK USO (or LGK S&P 500 and LGK USO) are rather noisy compared to VIX and OVX (or LVIX and LOVX). This is not surprising, since GK are estimates of volatility for a particular day, whereas implied volatility is calculated for a 30-day period, therefore effectively averaging out differences between individual days.

From Tables 1 and 2 we can also see that the realized volatility time series have lower mean values compared to the implied volatility time series. The reason for this is that even though implied volatility is often considered as an expectation of future volatility, it is not expected volatility. Implied volatility usually reflects both expected volatility and volatility risk premium. In other words, an option trader is willing to sell options to buyers only as long as he, on average, earns some profit by doing so. If he would price options in such a way that implied volatility would be equal to expected volatility, he would, on average, not earn any profit from it. Therefore, options are usually priced in such a way that implied volatility is higher than expected volatility. It is also obvious that OVX (LOVX) is generally higher than VIX (LVIX), which implies that oil market is generally more volatile than stock market.

The period of Great Recession from the end of 2007 till the middle of 2009 is depicted by the two red vertical solid lines in Fig. 1.

[2] The necessity to include the opening jump is the reason why we do not utilize realized volatility calculated from high-frequency data. High-frequency data allow for more precise estimation of volatility, but only during the trading part of the day. Since we need to include the opening jump, high-frequency data would not be of much help to us.

Table 1
Summary statistics (mean, median, standard deviation, interquartile range, minimum, maximum, skewness and excess kurtosis) for VIX, OVX, GK S&P 500 and GK USO.

	mean	median	st. dev.	IQR	min	max	skew	ex. k.
VIX	21.4	18.5	9.8	9.5	10.3	80.9	2.3	6.8
OVX	37.6	34.6	14.5	15.8	14.5	100.4	1.3	2.2
GK S&P	13.0	10.3	10.1	8.2	1.9	124.4	3.6	21.9
GK USO	30.7	25.8	18.9	19.7	3.1	161.5	1.9	5.1

Table 2
Summary statistics (mean, median, standard deviation, interquartile range, minimum, maximum, skewness and excess kurtosis) for LVIX, LOVX, LGK S&P 500 and LGK USO.

	mean	median	st. dev.	IQR	min	max	skew	ex. k.
LVIX	2.98	2.92	0.37	0.49	2.33	4.39	0.99	0.88
LOVX	3.56	3.54	0.37	0.44	2.67	4.61	0.08	-0.01
LGK S&P	2.37	2.33	0.59	0.77	0.65	4.82	0.42	0.48
LGK USO	3.26	3.25	0.56	0.75	1.13	5.08	0.13	-0.07

This period is closely related to the period of financial crisis of 2007–2008 and the period of the subprime mortgage crisis of 2007–2009. During these periods, the stock market was highly volatile. The decreased economic growth in many sectors (manufacturing, transportation etc.) also decreased the demand for energy products, which led to falling and volatile oil prices.

Recovery from the Great Recession started in 2009 but the Arab Spring of 2011–2012 was a major geo-political event which played its role in the oil market due to the uncertainty in the level of oil production in the affected countries. This period is depicted by the two red vertical dashed lines in Fig. 1.

The period of oil price fall from the middle of 2014 till 2015 due to (or accompanied by) a long-term slowdown of several major economies such as China, Russia etc., alternative ways of reaching oil resources (hydraulic fracking) and OPEC's members decision as of November 2014 to maintain oil production at the usual levels is depicted by the two red vertical dotted lines in Fig. 1. During this period, volatility increased slightly in the stock market and more profoundly in the oil market. This period also overlaps with the Russian financial crisis (2014–2017) and with the Chinese stock market crisis (2015–2016).

4. Results

From Fig. 1 and Tables 1 and 2, it is obvious that LVIX, LOVX, LGK S&P 500 and LGK USO are less skewed and more Gaussian compared to VIX, OVX, GK S&P 500 and GK USO. Consequently, the logarithmic time series will be used in further analysis. Using the logarithmic time series also offers an appealing interpretation since changes in these time series can be directly interpreted as percentage changes in the original time series.

The corrected wavelet power spectra[3] for LVIX and LOVX are presented in Fig. 2. Time is given on the horizontal axis, while Fourier period[4] P in (trading) days is depicted on the vertical axis. The colour in the plots captures the values of the corrected wavelet power spectrum (see the colour bars on the right side of the plots) and the cone of influence is separated by the U-shaped white curve.

Fig. 2 reveals that major variability in LVIX occurs at Fourier periods larger than 32 days. The variability at these Fourier periods is pronounced especially during the Great Recession, during the start of the recovery from the Great Recession, and during the Arab Spring. Noticeable variability in LVIX at Fourier periods shorter than 32 days is present during the Oil price fall. In LVOX, a pronounced variability at Fourier periods larger than 64 days occurs during the Oil price fall, but is also present during the Great Recession and the Arab Spring. Not much variability is present at Fourier periods shorter than 64 days.

Because of the noise present in LGK S&P 500 and LGK USO, corrected wavelet power spectra for LGK S&P 500 and LGK USO (not presented in any figures) exhibit major variability not only at large Fourier periods (64 days and larger), but also at short ones (2–16 days).

Squared wavelet coherence between LVIX and LOVX is depicted in Fig. 3. The red colour in the figure denotes values of wavelet coherence close to one, whereas the blue colour denotes values of wavelet coherence close to zero (see the colour bar on the right side of the figure). Relative phase is depicted by arrows only in those regions - in agreement with the suggestion of Ge (2008), see Section 2 - where the corresponding squared wavelet coherence is larger or equal to the 70th percentile of the distribution of the squared wavelet coherence. Generally, the value of relative phase is equal to the angle formed by the positive horizontal axis (the initial side) and the line of the arrow (the terminal side). The angle can take any value from $-\pi$ to and π. For example, if the arrow points to the right, relative phase is equal to zero and no delay is present between the time series. If the arrow points upwards, relative

[3] The R software (R Core Team, 2017) and the biwavelet R package (Gouhier et al., 2016) have been used to obtain various wavelet-analysis figures presented in the text.
[4] The Fourier period P is given as $P = 1/f$.

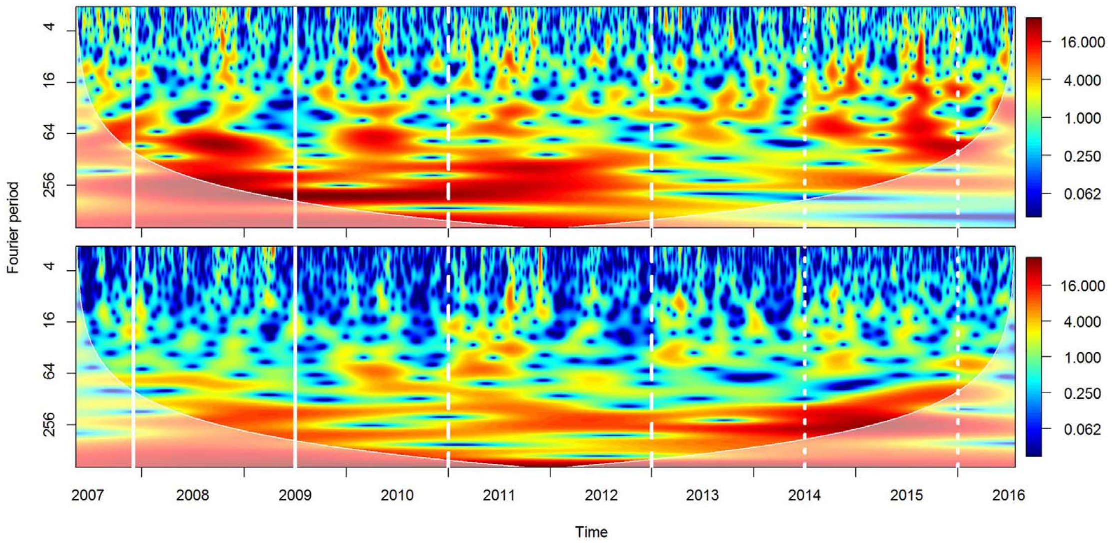

Fig. 2. Corrected wavelet power spectra for LVIX (top plot) and LOVX (bottom plot). The value of the corrected wavelet power spectrum is depicted in colour. The semi transparent regions at the left and right boundary of the plots separated by the white U-shaped curves are the cones of influence. The Great Recession, the Arab spring and the Oil price fall are depicted by the solid, dashed and dotted white vertical lines. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

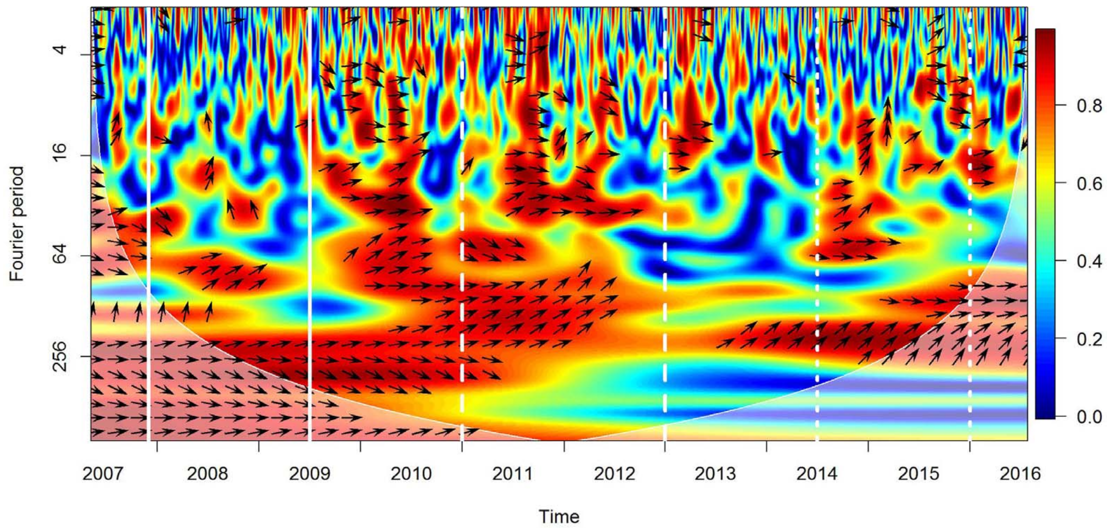

Fig. 3. Squared wavelet coherence and relative phase between LVIX and LOVX. The value of squared wavelet coherence is depicted in colour and the value of relative phase by arrows. The semi transparent region at the left and right boundary of the figure separated by the white U-shaped curve is the cone of influence. The Great Recession, the Arab spring and the Oil price fall are depicted by the solid, dashed and dotted white vertical lines. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

phase is $\pi/2$ and LVIX leads LOVX by $\pi/2$ in phase. If the arrow points downwards, relative phase is $-\pi/2$ and LVIX lags behind LOVX by $\pi/2$ in phase. Results of squared wavelet coherence and relative phase for LGK S&P 500 and LGK USO are presented in Fig. 4.

Let us discuss the results for LVIX and LOVX at first. The relationship is time-varying. LVIX and LOVX are strongly correlated at most Fourier periods from the middle of the year 2009 till the middle of the year 2012 (which covers the start of the recovery from the Great Recession and the start of the Arab Spring), whereas outside these dates the correlation is mostly strong at Fourier periods larger than 32 days and mostly not so strong at shorter ones. This conclusion holds in general - specifically, the squared wavelet coherence mostly increases with Fourier period, which can be observed in the top plot of Fig. 5 where the squared wavelet coherence averaged *over time* is depicted by the black line. The averaged squared wavelet coherence peaks at Fourier periods close to 1 year and decreases as we move towards shorter Fourier periods, such as 1 month or 1 week, or towards Fourier periods longer than 1 year. This suggests that in the long run (i.e. in the dynamics associated with approx. 1-year Fourier periods) LVIX and LOVX are much more tightly connected than in the medium or short run (i.e. in the dynamics associated with approx. 1-month or 1-week Fourier periods).

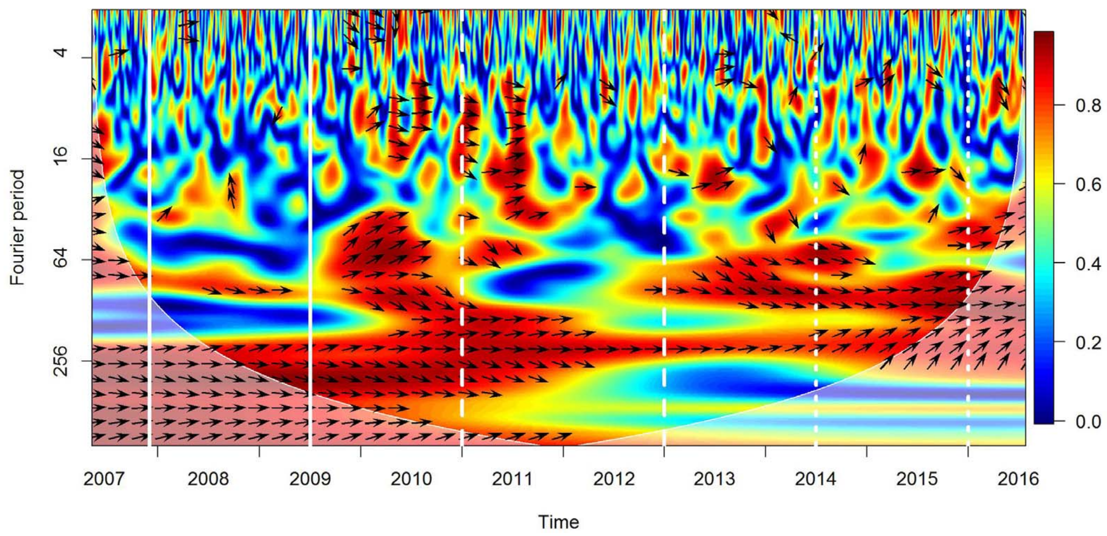

Fig. 4. Squared wavelet coherence and relative phase between LGK S&P 500 and LGK USO. The meaning of the colours, arrows, the U-shaped region and the white vertical lines is analogous to that of Fig. 3.

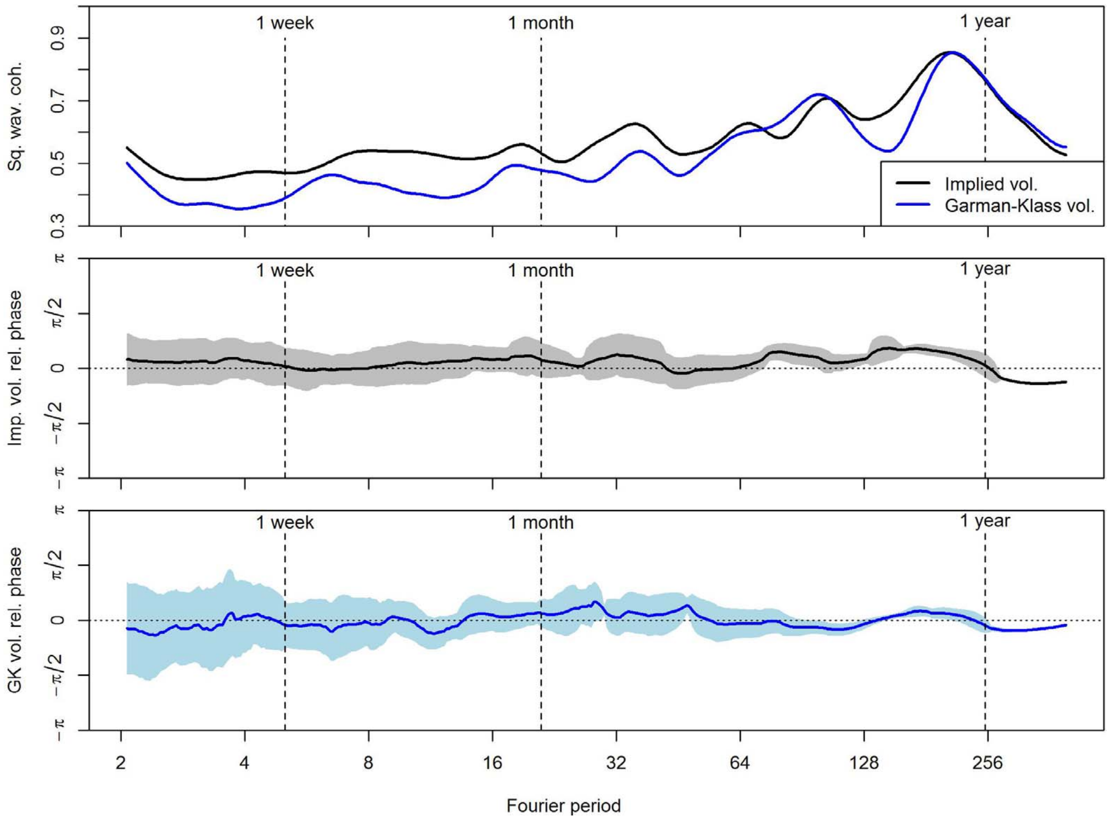

Fig. 5. *Top plot*: Squared wavelet coherence between LVIX and LOVX (black curve) and between LGK S&P 500 and LGK USO averaged over time (blue curve). Note that the range of the y-axis is from 0.3 to 1. *Middle plot*: Relative phase between LVIX and LOVX averaged over time (black curve) ± the corresponding standard deviation (the gray region). *Bottom plot*: Relative phase between LGK S&P 500 and LGK USO averaged over time (the blue curve) ± the corresponding standard deviation (the light blue region). *Note*: Values of the squared wavelet coherence in the cone of influence are excluded from the calculation of the average in the top plot. Values of relative phase in the cone of influence as well as those which correspond to regions with squared wavelet coherence below the 70th percentile of the distribution of squared wavelet coherence are excluded from the calculation of average phase and from the calculation of the corresponding standard deviation in the middle and bottom plot. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

Relative phase (in regions with sufficiently large squared wavelet coherence) seems to be very close to zero at all Fourier periods with some events in the time-frequency plane where LVIX seems to slightly lead LOVX.[5] This is confirmed by the middle plot of Fig. 5 where relative phase is averaged over time[6] and the average is plotted against Fourier period; see the black curve in the middle plot of Fig. 5. The boundaries of the gray regions in the plot correspond to the average relative phase ± the standard deviation[7] of the relative phase. The width of the gray region at a particular Fourier period tells us how the distribution of relative phase is variable. Taking the standard deviation of relative phase into account, we can see that relative phase is on average effectively zero at all Fourier periods. The only exceptions are Fourier periods slightly shorter than 100 days or 200 days where the average values of relative phase are above zero and the standard deviation is relatively small. These exceptions are associated with the following regions in Fig. 3: a.) the Great Recession and Fourier periods slightly shorter than 100 days, b.) the Arab Spring and Fourier periods a little shorter than 200 days, and c.) the epoch starting approx. one year before the Oil price fall and Fourier periods round 200 days. In all these regions LVIX seems to slightly lead LOVX.

A likely reason why the VIX index slightly leads the OVX index is the fact that the VIX index is in general much more important. VIX index is calculated from options which are more liquid than options used to calculate the OVX index. Moreover, neither VIX nor OVX are tradable (nobody can buy either of these indices directly). However, various derivatives (futures, options, exchange traded funds and exchange traded notes) with the VIX index as an underlying asset are actively traded (Bordonado et al., 2017), whereas no such derivatives exist for the OVX index as an underlying asset. All these factors may cause that the VIX index reacts to new information faster than the OVX index, and therefore leads the OVX index.

By comparing Fig. 4 with Fig. 3, it can easily be discerned that concerning the dynamics at Fourier periods *larger than 64 days*, the squared wavelet coherence and relative phase between LKG S&P 500 and LGK USO are very similar to those between LVIX and LOVX, the difference being that LKG S&P 500 and LGK USO seem to be a little less correlated during the Arab Spring and slightly more correlated during the Oil price fall and the epoch prior to this fall when compared to the correlation between LVIX and LOVX. Moreover, there is a smaller total area in the time-frequency plane where LKG S&P 500 leads LGK USO compared to the area where LVIX leads LOVX. This could indicate that in some cases, expectations about future volatility, which were transmitted from equity to oil market, did not materialize in the actual (realized) volatility.

We can also note that LGK USO is correlated with and leads LGK S&P 500 at the scale of approximately 64 days from the beginning of 2013 through the beginning of 2014 (see Fig. 4), while such a relationship is not observed between implied volatilities (Fig. 3). In order to interpret this, we have to again remember that implied volatilities reflect the expectations about the future. This means that the behavior which happened in realized volatilities was not anticipated beforehand by implied volatilities. Further, a possible reason why realized volatility of the oil market was correlated with and leading the realized volatility of the stock market could be that the oil price played a more important role during the calm year 2013 than in other years.

At Fourier periods *shorter than 64 days* LGK S&P 500 and LGK USO are seen to be less correlated and generally not so steady in relative phase when compared to LVIX and LOVX. This is also confirmed by Fig. 5 where the time-averaged squared wavelet coherence and the time-averaged relative phase are plotted as a function of Fourier period for both pairs of time series. The reason for the lower squared wavelet coherence and for the "unsteady" relative phase between LKG S&P 500 and LGK USO (when compared to the case of LVIX and LOVX) can presumably be explained by noticing that both the time series of Garman–Klass estimates of realized volatility are rather noisy and the noise dominates the dynamics at Fourier periods of 32 days and less. The presence of the noise generally leads to a decrease in the correlation between LKG S&P 500 and LGK USO and to an unsteady relative phase at Fourier periods lower than 1 week.

As documented in Fig. 5, the strongest comovement between the implied as well as between the realized volatility time series occurs round the Fourier period of 210 days. Consequently, in Fig. 6 we plot the wavelet squared coherence and relative phase for the Fourier period of one year (252 trading days) as a function of time. It is obvious that the relationship between the implied as well as between the realized volatility time series is time-varying at this Fourier period. Further, LVIX seems to slightly lead LOVX at this Fourier period, while the lead of LGK S&P 500 before LGK USO is less pronounced. This is in accordance with what we observed previously.

It should also be noted that missing arrows in Fig. 3 do *not* imply that *no* lead/lag relationship is present since arrows are plotted only in those regions where the corresponding squared wavelet coherence is sufficiently large as explained in the previous text. This can be demonstrated, for example, for the Fourier period of one year at the end of 2012 and the beginning of 2013 where no arrows are plotted in Fig. 3 since the squared wavelet coherence has decreased a lot, but LVIX still leads LOVX (see Fig. 6). The likely reason for the decrease in the squared wavelet coherence is the aftermath of the Arab Spring, during which oil price was subject to more idiosyncratic shocks which translated into weaker comovement with the stock market.

Next we look at the comovement between the VIX and OVX from a viewpoint of VIX being the independent, and OVX being the dependent variable. Cross-wavelet gain between LOVX (response variable) and LVIX (explanatory variable) is depicted in Fig. 7. Regions associated with the green colour correspond to events where a change in LVIX is accompanied by a change in LOVX of a similar size. Light/dark blue colours correspond to regions where a change in LVIX is accompanied by a slightly smaller/much smaller change in LOVX. On the other hand, light/dark red colours correspond to regions where a change in LVIX is accompanied by a

[5] There are also some events in the time-frequency plane where LOVX seems to lead LVIX. However, these events seem to be less frequent than the events where LVIX leads LOVX.

[6] Circular average is used to calculate the average relative phase.

[7] Circular standard deviation is used to calculate the standard deviation of relative phase.

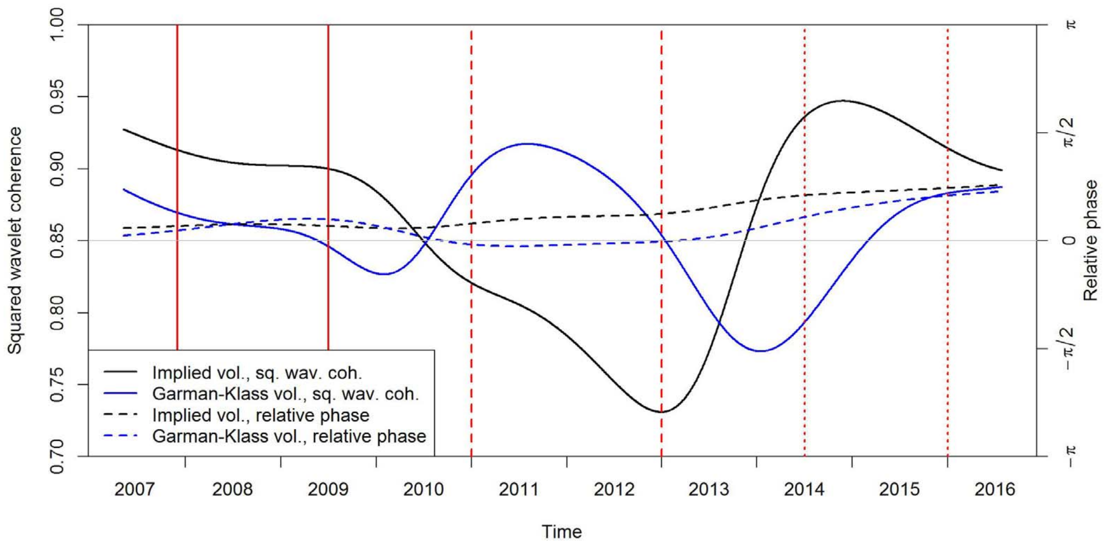

Fig. 6. Squared wavelet coherence and relative phase between LVIX and LOVX and between the LGK S&P 500 and LGK USO for the Fourier period of 252 days. Note that the left vertical axis ranges from 0.7 to 1. The Great Recession, the Arab spring and the Oil price fall are depicted by the solid, dashed and dotted red vertical lines. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

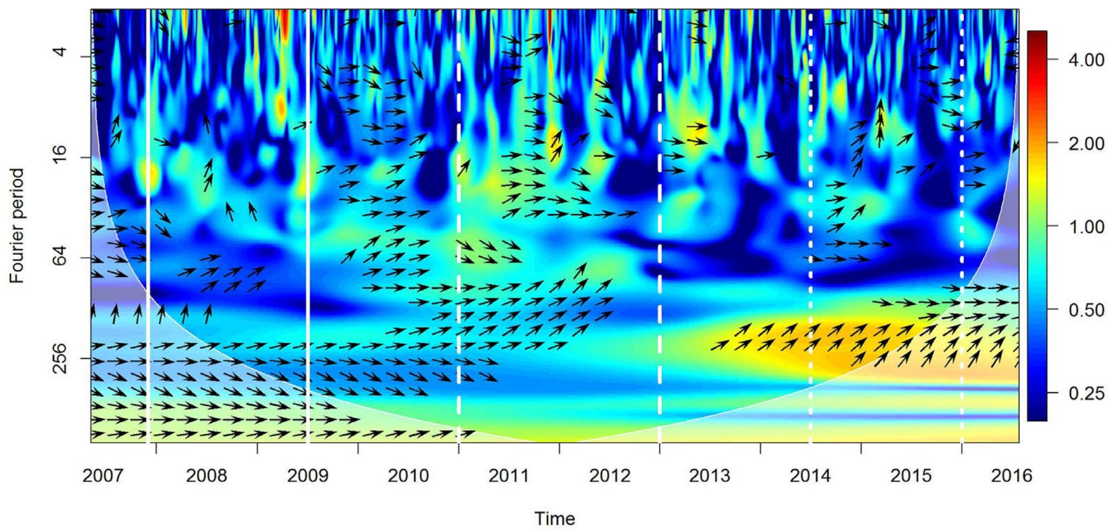

Fig. 7. Cross-wavelet gain between LOVX (response variable) and LVIX (explanatory variable). The cross-wavelet gain is depicted in colour. The value of relative phase is depicted by arrows in exactly the same regions as in Fig. 3, i.e. in the regions with a high value of squared wavelet coherence. The semi transparent region at the left and right boundary of the figure separated by the white U-shaped curve is the cone of influence. The Great Recession, the Arab spring and the Oil price fall are depicted by the solid, dashed and dotted white vertical lines. (For interpretation of colours in this figure, the reader is referred to the web version of this article.)

slightly larger/much larger change in LOVX. Analogously to Fig. 3, relative phase is depicted by arrows in the regions where the squared wavelet coherence is larger or equal to the 70th percentile of the distribution of the squared wavelet coherence - the value of the cross-wavelet gain will be interpreted and discussed only in these regions since this helps us to understand the nature of the comovement in the regions where the time series are strongly correlated.

Since changes in LVIX and LOVX can be directly interpreted as percentage changes of VIX and OVX, the interpretation of Fig. 7 is straightforward. Specifically, during the dates from the start of the year 2010 till the end of the year 2012 (which covers the start of the recovery from the Great Recession, and also the Arab spring) percentage changes of VIX in regions with large squared wavelet coherence (which correspond mostly to Fourier periods of 16 days and larger) are accompanied by percentage changes of OVX of a similar size. This is in agreement with Fig. 1 where medium and long-run changes in LVIX and LOVX are correlated and comparable in size during these dates. During the Great Recession, percentage changes of VIX in regions with large squared wavelet coherence (and especially those regions where Fourier periods are larger than 16 days) are generally accompanied by slightly smaller percentage changes of OVX. This is also in agreement with Fig. 1 where long-run changes of LVIX are correlated with long-run changes of LOVX of smaller amplitudes. Approximately one year before the Oil price fall, percentage changes of VIX in regions with large squared wavelet coherence and Fourier periods of 64 days and larger are accompanied by percentage changes of OVX with larger amplitudes.

We present the cross-wavelet gain only for the case where VIX is the independent and OVX the dependent variable and not vice versa since, as follows from Eqs. (5) and (9), the cross-wavelet gain for the latter case is equal to the squared wavelet coherence divided by the cross-wavelet gain for the former case. Consequently, the two cross-wavelet gains are complementary, the latter cross-wavelet gain bringing no new information which would not be included in the former cross-wavelet gain. Our choice of VIX being the independent and OVX the dependent variable is natural in the view of the previous results where VIX was seen to lead OVX.

5. Conclusions

We have employed wavelet analysis to study the relationship between the stock market and oil market volatility. We used the logarithm of volatility of the S&P 500 equity index and USO oil fund and of their implied volatilities VIX and OVX.[8] Wavelet analysis allowed us to study both the frequency as well as the temporal aspect of the relationship.

As expected, our findings show that the implied volatility of the equity market (VIX) and the implied volatility of the oil market (OVX) are highly correlated. This result is in accordance with Liu et al. (2013) who also find strong relationship between the OVX and VIX. However, the correlation between the stock and oil market volatility, whether measured as implied or realized volatility, is time-varying and depends on the time scale. It is strongest at Fourier periods round 210 days and gets weaker as we decrease or increase the Fourier period. Moreover, the VIX index slightly leads the OVX index, while this feature is weaker between realized volatilities.

We also have several more specific findings. Namely, during the Great Recession LVIX was correlated with LOVX especially at Fourier periods larger than 64 days. LVIX is suggested to have slightly led LOVX. A one percentage change of VIX was generally accompanied by a percentage change of OVX of a *smaller* size. This suggests that the Great Recession primarily affected the stock market, the impact on the oil market happening simultaneously or with a slight delay, and with lower amplitude.

Further, from the middle of 2009 till 2012 (which covers the start of the recovery from the Great Recession and the Arab Spring) LVIX generally exhibited strong correlation with LOVX at most of the explored Fourier periods, while slightly leading LOVX at some events in the time-frequency plane. A one percentage change of VIX was generally accompanied by an approximately one percentage change of OVX at most Fourier periods. This suggests that the recovery from the Great Recession demonstrated itself very similarly in both the markets with the stock market slightly leading the oil market.

Further, the Oil price fall (and the overlapping epochs of the Russian financial crisis and the Chinese stock market crisis) resulted in the correlation of LVIX and LOVX at several Fourier periods and impacted the oil market more strongly than the stock market despite the fact that the stock market led the oil market at some events.

Our results have implications both for general understanding of financial markets as well as for traders and other market participants exposed to the volatility of the oil and stock market. We found that volatilities of the oil and stock market are more strongly correlated on yearly horizon than on horizons of several days. This implies that diversification benefits for traders or investors exposed to volatility of both the oil and the stock market in the horizon of a few days are higher than for traders or investors on the yearly horizon.

Acknowledgment

We thank the anonymous reviewers whose comments helped improve the content of the paper.

References

Bašta, M., Molnár, P., Helman, K., 2017. Exploring relationship between developed and developing financial markets by wavelets. Proceedings of The 11th International Days of Statistics and Economics, September 14–16, 2017, Prague.
Birkelund, O.H., Haugom, E., Molnár, P., Opdal, M., Westgaard, S., 2015. A comparison of implied and realized volatility in the nordic power forward market. Energy Econ. 48, 288–294.
Bordonado, C., Molnár, P., Samdal, S.R., 2017. Vix exchange traded products: price discovery, hedging, and trading strategy. J. Futures Markets 37 (2), 164–183.
Bugge, S.A., Guttormsen, H.J., Molnár, P., Ringdal, M., 2016. Implied volatility index for the norwegian equity market. Int. Rev. Financ. Anal. 47, 133–141.
Dutta, A., Nikkinen, J., Rothovius, T., 2017. Impact of oil price uncertainty on middle east and african stock markets. Energy 123, 189–197.
Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 67–78.
Ge, Z., 2008. Significance tests for the wavelet cross spectrum and wavelet linear coherence. Annales Geophysicae 26 (12), 3819.
Gouhier, T., Grinsted, A., Simko, V., 2016. biwavelet: Conduct univariate and bivariate wavelet analyses (version 0.20.10).
Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Process. Geophys. 11 (5/6), 561–566.
Guo, J.-F., Ji, Q., 2013. How does market concern derived from the internet affect oil prices? Appl. Energy 112, 1536–1543.
Haugom, E., Langeland, H., Molnár, P., Westgaard, S., 2014. Forecasting volatility of the us oil market. J. Banking Finance 47, 1–14.
Henriques, I., Sadorsky, P., 2011. The effect of oil price volatility on strategic investment. Energy Econ. 33 (1), 79–87.
Ji, Q., Fan, Y., 2012. How does oil price volatility affect non-energy commodity markets? Appl. Energy 89 (1), 273–280.
Liu, M.-L., Ji, Q., Fan, Y., 2013. How does oil market uncertainty interact with other markets? an empirical analysis of implied volatility index. Energy 55, 860–868.
Liu, Y., San Liang, X., Weisberg, R.H., 2007. Rectification of the bias in the wavelet power spectrum. J. Atmos. Oceanic Technol. 24 (12), 2093–2102.
Luo, X., Qin, S., 2017. Oil price uncertainty and chinese stock returns: new evidence from the oil volatility index. Finance Res. Lett. 20, 29–34.
Maghyereh, A.I., Awartani, B., Bouri, E., 2016. The directional volatility connectedness between crude oil and equity markets: new evidence from implied volatility indexes. Energy Econ. 57, 78–93.
Mandler, M., Scharnagl, M., 2014. Money growth and consumer price inflation in the euro area: a wavelet analysis.
Molnár, P., 2012. Properties of range-based volatility estimators. Int. Rev. Financ. Anal. 23, 20–29.
Qadan, M., Yagil, J., 2012. Fear sentiments and gold price: testing causality in-mean and in-variance. Appl. Econ. Lett. 19 (4), 363–366.

[8] Realized volatility is estimated from Garman–Klass formula (Garman and Klass, 1980) and its logarithm is denoted as LGK S&P 500 and LGK USO, whereas logarithm of the VIX and OVX indices are denoted as LVIX and LOVX.

R Core Team, 2017. R: a language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.
Rafiq, S., Salim, R., Bloch, H., 2009. Impact of crude oil price volatility on economic activities: an empirical investigation in the thai economy. Resour. Policy 34 (3), 121–132.
Sari, R., Soytas, U., Hacihasanoglu, E., 2011. Do global risk perceptions influence world oil prices? Energy Econ. 33 (3), 515–524.
Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78.

Bibliography

Aalborg, H. A., Molnár, P., & de Vries, J. E. (2019). What can explain the price, volatility and trading volume of bitcoin? *Finance Research Letters*, *29*, 255–265.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42–55.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, *169*(1), 75–93.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, *17*(5), 457–477.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Bašta, M., & Molnár, P. (2019). Long-term dynamics of the vix index and its tradable counterpart vxx. *Journal of Futures Markets*, *39*(3), 322–341.

Birkelund, O. H., Haugom, E., Molnár, P., Opdal, M., & Westgaard, S. (2015). A comparison of implied and realized volatility in the nordic power forward market. *Energy Economics*, *48*, 288–294.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bordonado, C., Molnár, P., & Samdal, S. R. (2017). Vix exchange traded products: Price discovery, hedging, and trading strategy. *Journal of Futures Markets*, *37*(2), 164–183.

Brandt, M. W., & Jones, C. S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business & Economic Statistics*, *24*(4), 470–486.

Bugge, S. A., Guttormsen, H. J., Molnár, P., & Ringdal, M. (2016). Implied volatility index for the norwegian equity market. *International Review of Financial Analysis*, *47*, 133–141.

Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (carr) model. *Journal of Money, Credit and Banking*, *37*(3), 561–582.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, *50*(2), 125–150.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fiszeder, P., Fałdziński, M., & Molnár, P. (2019). Range-based dcc models for covariance and value-at-risk forecasting. *Journal of Empirical Finance*, *54*, 58–76.

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of business*, 67–78.

Ghysels, E., Harvey, A. C., & Renault, E. (1996). 5 stochastic volatility. *Handbook of statistics*, *14*, 119–191.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801.

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, *61*(2), 247–264.

Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, *14*(4), 429–434.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014). Forecasting volatility of the us oil market. *Journal of Banking & Finance*, *47*, 1–14.

Horpestad, J. B., Lyócsa, Š., Molnár, P., & Olsen, T. B. (2019). Asymmetric volatility in equity markets around the world. *The North American Journal of Economics and Finance*, *48*, 540–554.

Kim, N., Lučivjanská, K., Molnár, P., & Villa, R. (2019). Google searches and stock market activity: Evidence from norway. *Finance Research Letters*, *28*, 208–220.

Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics*, *187*(1), 293–311.

Lyócsa, Š., & Molnár, P. (2018). Exploiting dependence: Day-ahead volatility forecasting for crude oil and natural gas exchange-traded funds. *Energy*, *155*, 462–473.

Lyócsa, Š., & Molnár, P. (2017). The effect of non-trading days on volatility forecasts in equity markets. *Finance Research Letters*, *23*, 39–49.

Lyócsa, Š., & Molnár, P. (2016). Volatility forecasting of strategically linked commodity etfs: Gold-silver. *Quantitative Finance*, *16*(12), 1809–1822.

Lyócsa, S., Molnár, P., & Fedorko, I. (2016). Forecasting exchange rate volatility: The case of the czech republic, hungary and poland. *Finance a Uver*, *66*(5), 453.

Lyócsa, Š., Molnár, P., & Plíhal, T. (2019). Central bank announcements and realized volatility of stock markets in g7 countries. *Journal of International Financial Markets, Institutions and Money*, *58*, 117–135.

Lyócsa, Š., Molnár, P., & Todorova, N. (2017). Volatility forecasting of non-ferrous metal futures: Covariances, covariates or combinations? *Journal of International Financial Markets, Institutions and Money*, *51*, 228–247.

Molnár, P. (2016). High-low range in garch models of stock return volatility. *Applied Economics*, *48*(51), 4977–4991.

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: Journal of the Econometric Society*, 347–370.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 61–65.

Patton, A. J., & Sheppard, K. (2009). Optimal combinations of realised volatility estimators. *International Journal of Forecasting*, *25*(2), 218–238.

Prokopczuk, M., & Simen, C. W. (2014). The importance of the volatility risk premium for volatility forecasting. *Journal of Banking & Finance*, *40*, 303–320.

Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 504–512.

Siriopoulos, C., & Fassas, A. (2019). Implied volatility indices–a review. *Available at SSRN 1421202*.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series analysis: theory and practice*, *1*, 203–226.

Tse, Y. K., & Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, *20*(3), 351–362.

Zhang, J. E., & Zhu, Y. (2006). Vix futures. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *26*(6), 521–531.