Forecasting oil price realized volatility: A new approach

Degiannakis, Stavros and Filis, George

Panteion University of Social and Political Sciences, Panteion University of Social and Political Sciences

29 January 2016

Forecasting oil price realized volatility: A new approach

Stavros Degiannakis[1,2] and George Filis[1,*]

[1]Department of Economics and Regional Development, Panteion University of Social and Political Sciences, 136 Syggrou Avenue, 17671, Greece.
[2]Postgraduate Department of Business Administration, Hellenic Open University, Aristotelous 18, 26 335, Greece.

*Corresponding author: email: gfilis@panteion.gr

Abstract

This paper adds to the extremely limited strand of the literature focusing on the oil price realized volatility forecasting. More specifically, we evaluate the information content of four different asset classes' volatilities when forecasting the oil price realized volatility for 1-day until 66-day ahead. To do so, we concentrate on the Brent crude oil and fourteen other assets, which are grouped into four different asset classes, based on Heterogeneous AutoRegressive (HAR) framework. Our out-of-sample forecasting results can be summarised as follows. (i) The use of exogenous volatilities statistically significant improves the forecasting accuracy at all forecasting horizons. (ii) The HAR model that combines volatilities from multiple asset classes is the best performing model. (iii) The Direction of Change suggests that all HAR models are highly accurate in predicting future movements of oil price volatility. (iv) The forecasting accuracy of the models is better gauged using the Median Absolute Error and the Median Squared Error. (v) The findings are robust even during turbulent economic periods. Hence, different asset classes' volatilities contain important information which can be used to improve the forecasting accuracy of oil price volatility.

Keywords: Volatility forecasting, realized volatility, crude oil futures, Brent crude oil, HAR, MCS.
JEL: C22, C53, G13, Q02, Q47.

1. Introduction and brief review of the literature

Crude oil price movements are of major importance for the global economy. Elder and Serletis (2010) opine that oil price uncertainty exercises significant impact on the economy. It is no coincidence that since the second half of 2015 the plunge of oil prices and its economic effects have monopolised media attention from the most widely circulated financial press. Even more, this fall in oil prices has resulted in increased oil price volatility, which is an essential input in many macroeconomic models, as well as, in option pricing and value at risk.

Furthermore, oil price volatility forecasts are particularly important nowadays due to the fact that the increased participation of hedge funds in the oil market over the last decade or so, has results in the financialisation of the market (Buyuksahin *et al.*, 2010; Silvennoinen and Thorp, 2010; Tang and Xiong, 2010; Buyuksahin and Robe, 2011; Hamilton and Wu, 2014; Sadorsky, 2014; Buyuksahin and Robe, 2014). In addition, we observe that financial institutions are now considering the oil market as a profitable alternative investment for their portfolios (see, for example, Kat and Oomen, 2007, and Silvennoinen and Thorp, 2010).

Thus, accurate forecasts of oil price volatility are both timely and essential for policy makers, oil traders, as well as, researchers. However, despite the importance of oil price volatility forecasts, this strand of the literature is rather under-researched and our aim is to extend this line of enquiry.

It is interesting to note a paradox in the field of oil price volatility forecasting. Despite the fact that the importance of oil price fluctuations and volatility on the economy and financial markets have long been established[1] and researchers forecast asset market volatility since the 80s[2], the earlier study in the field of oil volatility forecasting dates as recent as 2006 by Sadorsky.

Sadorsky (2006) forecasts the squared daily returns of oil futures prices (as a proxy of volatility) using GARCH, TGARCH and Exponential Smoothing, VAR and BEKK models. The VAR and BEKK models include also the squared returns of other

[1] See, for instance, Hamilton, 1983; Burbidge and Harrison (1984), Jones and Kaul (1996), Huang *et al.* (1996), Bernanke *et al.* (1997), Ferderer (1997), Sadorsky (1999), Brown and Yucel (2002), Kilian and Park (2009), Malik and Ewing (2009), Filis (2010), Arouri *et al.* (2011), Filis *et al.* (2011), Masih *et al.* (2011), Degiannakis *et al.* (2013), Rahman and Serletis (2011), Baumeister and Peersman (2013), Filis and Chatziantoniou (2014).

[2] See, Akgiray (1989), Bollerslev *et al.* (1992), West and Cho (1995), Frances and van Dijk (1996), McKenzie and Mitchell (2002), Brooks and Persand (2002, 2003), Degiannakis (2004), Andersen *et al.* (2003, 2005), Hansen and Lunde (2005), Degiannakis (2008), Angelidis and Degiannakis (2008), Ana-Maria Fuertes *et al.* (2009), Frijns *et al.* (2010) among others.

petroleum futures (including the heating oil, gasoline and natural gas). He finds that the GARCH-family models are able to outperform the random walk model, which is used as the benchmark. Sadorsky and McKenzie (2008) seconds Sadorsky's (2006) findings, showing that the GARCH-type models produce more accurate forecasts than any other competing model, although only in the longer-horizons. They claim that in shorter-horizons, it is the power autoregressive model that produces the best forecasts of oil price volatility.

Following Sadorsky (2006) and Sadorsky and McKenzie (2008), an increasing number of authors has turned their attention to oil price volatility forecasting. For example, Kang *et al.* (2009), uses daily oil spot prices in order to forecast the 1-day, 5-days and 20-days ahead conditional volatilities by means of CGARCH, FIGRACH and IGARCH models. Their findings suggest that the CGARCH and FIGARCH models are more useful in modelling and forecasting the volatility in the crude oil prices.

More recently, Nomikos and Pouliasis (2011) and Kang and Yoon (2013) consider oil futures prices to estimate and forecast oil price conditional volatility. Nomikos and Pouliasis (2011) use Mix-GARCH and MRS-GARCH models to forecast the 1-day-ahead oil price volatility and find that both models are able to outperform the forecasts of the simple GARCH model. Kang and Yoon (2013), on the other hand, combine ARFIMA models with GARCH models to produce 1-day, 5-days and 20-days ahead forecasts. They claim that the ARFIMA-FIGARCH models are better in modelling oil price conditional volatility. Nevertheless they maintain that no model consistently outperforms all other competing ones.

Similarly, several other authors model the conditional volatility of oil prices and forecast these volatilities, using univariate models such as the FIAPARCH, HYGARCH, EGARCH, FIGARCH, APARCH, as well as, multivariate models such as BEKK, VAR and Risk Metrics (see, Agnolucci, 2009; Wei *et al.*, 2010; Arouri *et al.*, 2012; Hou and Suardi, 2012; Chkili *et al.*, 2014)[3]. For the multivariate models, they consider conditional volatilities of other energy commodities, similar to those of Sadorsky (2006). The general consensus is that the univariate GARCH-type models are able to produce more accurate forecasts than any other competing models. It is

[3] Wang and Wu (2012) is the only paper that considers weekly, rather than daily, oil prices.

worth noting that the majority of these papers are evaluating the forecasting accuracy of their models in 1-day, 5-days and 20-days ahead horizons.

A study that is quite distinct is this of Efimova and Serletis (2014). Similar to the previous studies, they also use oil spot prices to model and forecast the 1-day ahead oil conditional volatility using univariate GARCH-type models, as well as, multivariate models such as BEKK, DCC and VARMA-GARCH. Nevertheless, it is the first paper to consider the inclusion of an additional asset class in order to assess if this yields better forecasts for the oil price volatility. More specifically, all previous papers which have estimated multivariate models have considered prices only from other energy markets (e.g. heating oil, gasoline, etc). By contrast, Efimova and Serletis (2014) include the S&P500 daily returns to their models. Their findings corroborate with these of the previous literature, suggesting that the univariate models are able to produce more accurate forecasts and that the inclusion of the S&P500 daily returns did not produce better forecasts.

All aforementioned papers use daily oil prices and forecast the conditional oil price volatility. Nevertheless, empirical evidence (primarily from the finance literature) has long suggested that intraday (ultra-high frequency) data are more information rich and thus they can produce more accurate estimates of the daily volatility (see, *inter alia*, Oomen, 2001; Andersen *et al.*, 2001, 2003, 2010; McAleer and Medeiros, 2008). More specifically, Andersen and Bollerslev (1998) introduce an alternative measure of daily volatility, which considers intraday data, namely the Realized Volatility. Realized volatility is based on the idea of using the sum of squared intraday returns to generate more accurate daily volatility measures.

Numerous studies have shown that the intraday data are able to produce better forecasts, compared to the daily data (see, for instance, Hansen and Lunde, 2005; Engle and Sun, 2007; Tay *et al.*, 2009).

However, until very recently the use of ultra-high frequency data for volatility forecasting has concentrated only for stock market and exchange rate volatilities (see, among others, Akgiray, 1989; Bollerslev *et al.*, 1992; West and Cho, 1995; McKenzie and Mitchell, 2002; Brooks and Persand, 2002, 2003; Degiannakis, 2004; Andersen *et al.*, 2003, 2005; Hansen and Lunde, 2005; Degiannakis, 2008; Angelidis and Degiannakis, 2008).

It is only since 2014 that studies try to forecast oil price volatility using ultra-high frequency data. One of the early studies is this by Haugom *et al.* (2014) who construct the realized volatility in order to forecast the Value-at-Risk (VaR) for the Brent crude oil futures. The authors use data from 3rd January, 2006 to 31st March, 2012 of the Brent Crude oil futures, considering the front-month futures contracts only. The authors use the Heterogeneous AutoRegressive (HAR) model of Corsi (2009) to forecast the realized oil volatility, given its superiority in forecasting this volatility measure (see, *inter alia*, Andersen *et al.*, 2007; Corsi, 2009; Busch *et al.*, 2011; Fernandes *et al.*, 2014)[4].

Sévi (2014) also forecasts the realized volatility of oil futures prices for the front-month futures contracts. More specifically, the author considers 5min intraday oil price returns to construct the daily realized volatility. He then uses several extensions of the HAR model in order to consider the jump component, semivariances, leverage effects, as well as, asymmetries in these components. The data range from January, 1987 to December, 2010. Despite the fact that Sévi (2014) considers in total nine different HAR models, he concludes that none of these models is able to outperform the forecasting accuracy of the simple HAR model, which is based only on the oil realized volatility (HAR-RV), in any forecasting horizons (i.e. 1-day to 66-days ahead).

More recently, Prokopczuk *et al.* (2015) use intraday data to forecast the realized volatility of crude oil prices, as well as, of gasoline, heating oil and the natural gas for three forecasting horizon, namely 1-day, 5-days and 22-days ahead. Their data span from January 2007 until June 2012. In order to construct their realized volatilities for the three time-series, the authors choose a sampling frequency of 15min. As in Haugom *et al.* (2014) and Sévi (2014), Prokopczuk *et al.* (2015) also use a HAR model for their forecasting exercise. Similarly with Sévi (2014), they also consider several extensions of the HAR-RV model, in order to capture whether the jump detection produces better forecasts. Their findings corroborate those of Sévi (2014), showing that the modelling of jumps does not improve the forecast accuracy of the simple HAR-RV model.

[4] The HAR model considers information of the previous day's, week's and month's volatility and thus, it is able to accommodate the heterogeneous beliefs of traders in the oil market. Bollerslev and Wright (2001) maintain that any volatility series exhibits long-memory behaviour and thus a model which considers this stylized fact (such as the HAR model) is able to produce better forecasts. Andersen *et al.* (2007), Corsi (2009), Busch *et al.* (2011) and Fernandes *et al.* (2014) also argue that HAR models are more successful in forecasting asset price volatility as they are parsimonious and they capture the long-memory that is observed in asset price volatility.

Phan *et al.* (2015), on the other hand, examine whether the S&P500 volatility improves the oil price volatility forecasts. The authors consider 5min intraday data to construct the realized volatility measure; nevertheless, they use an EGARCH(1,1) model rather than HAR-RV. They report that the cross-market volatility interaction improves the forecasts for the oil price volatility. Finally, Chatrath *et al.* (2015) also forecast the oil price volatility, using a sampling frequency of 5min to construct their realized volatility measure. The authors employ similar regressions to those by Christensen and Prabhala (1998) and Jiang and Tian (2005) and find that the incorporation of the crude oil implied volatility improves the forecasting of realized volatility.

Our paper directly extends the previous contributions of Haugom *et al.* (2014), Sévi (2014) and Prokopczuk *et al.* (2015). More specifically, we add to this extremely limited strand of the literature focusing on the oil price realized volatility forecasting, using the standard forecasting HAR-RV model[5]; however, we extend the current state-of-the-art in a number of ways. (i) We consider 14 exogenous variables (using various HAR-RV-X models), which belong to four different asset classes (*stocks*, *foreign exchange*, *commodities* and *macroeconomics*) and we investigate whether their realized volatilities improve the oil volatility forecasts. (ii) We clearly explain how to handle exogenous variables in a HAR model in order to proceed with the forecasts. (iii) We assess the forecasting accuracy of the HAR-RV-X models based on each individual asset class, their combined forecasts, as well as, the forecast-averaging. (iv) We assess the forecasting accuracy of our models during economic turbulent periods, such as the Global Financial Crisis of 2007-09. (v) We use the newly developed Model Confidence Set and the Direction-of-Change (DoC) to evaluate the forecasting accuracy of the competing models. (vi) We assess whether it is more appropriate to evaluate forecasts using the Median Absolute Error and the Median Squared Error, given that the Mean Absolute and Squared Errors are highly asymmetrically distributed. (vii) Our forecasting horizons range from 1-day to 66-days-ahead, given that different stakeholders have different predictive needs.

In short, we report the following regularities. (i) The HAR-RV-X models outperform the forecasting accuracy of the HAR-RV at all forecasting horizons. (ii) The HAR-RV-X models that combine multiple asset classes are the best performing

[5] We do not consider the jump components in this study, given that the existing literature provides evidence that their inclusion does not produce better forecasts.

models, since they capture the different effects that the oil price volatility receives from each asset class at different times. (iii) The DoC suggests that all HAR models are highly accurate in predicting the movements of oil price volatility. Thus, we maintain that HAR-RV-X models should be used from stakeholders who are interested in the accuracy of the forecasts, whereas those interested only in the movement of oil price volatility can be limited to HAR-RV. (iv) The forecasting accuracy of the models is better gauged using the Median Absolute Error and the Median Squared Error. (v) The findings are robust even when we concentrate only on turbulent economic periods.

The rest of the paper is structured as follows. Section 2 provides a detailed description of the data. Section 3 explains the construction of the realized volatility, whereas Section 4 describes the econometric approach employed in this paper. Section 5 explains the forecasting strategy that is followed and Section 6 presents the forecasting evaluation techniques. Section 7 analyses the findings of the study and Section 8 includes the robustness checks. Section 9 concludes the study.

2. Data Description

In this study we use tick by tick data of the front-month futures contracts for the following series: Brent Crude Oil (ICE Futures Europe), GBP/USD (CME Group), CAD/USD (CME Group), EUR/USD (CME Group), FTSE100 (ICE Futures Europe), S&P500 (CME Group), Hang Seng (Hong Kong Stock Exchange), Euro Stoxx 50 (Eurex), Gold (CME Group), Copper (CME Group), Natural Gas (CME Group), Palladium (CME Group), Silver (CME Group) and the US 10yr T-bills (CME Group). All data are obtained from TickData. We use an additional US macroeconomic volatility indicator, which is available in daily frequency, namely the Economic Policy Uncertainty (EPU)[6] Index by Baker *et al.* (2013). The period of our study spans from 1st of August, 2003 to 5th of August, 2015 and it is dictated by the availability of intraday data for the Brent Crude oil futures contracts.

The choice of variables is justified by the fact that there is a growing literature that confirms the cross-market transmission effects (either of returns or volatilities)

[6] As indicated by Baker *et al.* (2013), EPU index is constructed based on three types of underlying components. The first component quantifies newspaper coverage of policy-related economic uncertainty. The second component reflects the number of federal tax code provisions set to expire in future years. The third component uses disagreement among economic forecasters as a proxy for uncertainty. For more information the reader is directed to http://www.policyuncertainty.com/.

between oil and four main asset classes (i.e., commodities, exchange rates (Forex), stock markets and macroeconomic indicators) (see, *inter alia*, Hammoudeh *et al.*, 2004; Ågren, 2006; Aloui and Jammazi, 2009; Malik and Ewing, 2009; Sari *et al.*, 2010; Arouri *et al.*, 2011; Mensi *et al.*, 2013, 2014; Beckmann and Czudaj, 2013; Antonakakis *et al.*, 2014; Fratzscher *et al.*, 2014; Guesmi and Fattoum, 2014; Sadorsky, 2014; Soucek and Todorova, 2013, 2014; Antonakakis and Kizys, 2015; Phan *et al.*, 2015; IEA, 2015). Given these interaction, we posit that these four asset classes contain information for the future movements of the oil price volatility.

Furthermore, we consider the specific variables (among the four asset classes) as they are among the most tradable futures contracts globally[7]. Nevertheless, this choice of variables also serves the following purpose.

Specifically for the stock market indices, we choose among the key US, EU and Asian indices as (i) their combined trading spans across the full day and (ii) they represent the stock market indices of the largest economies in the world. However, we also include the FTSE100 index futures given that we forecast the Brent crude oil volatility.

As far the foreign exchange variables are concerned, we maintain that the EUR/USD is the main currency that exercises an impact on oil fluctuations, whereas the use of the GBP/USD futures is incontestable, given that it is related to the Brent crude oil. Finally, the choice of the CAD/USD is motivated by Chen *et al.* (2010) who maintain that currencies of commodity exporters contain important information for the future movements of commodity prices.

Finally, we use the US 10yr T-bill futures and the US EPU as recent studies have shown that oil price volatility are responsive to change in the economic conditions (see, for instance, Antonakakis *et al.*, 2014). We treat both the US 10yr T-bill and the US EPU as variables that approximate global economic developments, given the importance of the US in the global economy.

Important milestones for the construction of the intra-day time series are the following:

[7] Although, this is not the case for the Hang Seng index, given that the most traded Chinese index futures is the CSI 300. Nevertheless, intraday data for the CSI 300 index futures are available after 2008 and thus we had to replace this index with Hang Seng, which is among the most traded index futures in the Asian region.

(i) Trading day: In our paper we define as trading day the period between 21:01 GMT the night before until 21:00 GMT that evening. The particular definition of the trading day is motivated by Andersen *et al.* (2001, 2003, 2007).

(ii) Holidays and short trading days: We exclude from our series several fixed and moving holidays, such as Christmas, Martin Luther King day, Washington birthday day, Good Friday, Easter Monday, Memorial day, July 4th, Labour day and Thanksgiving and the day after.

(iii) Non-trading hours: We remove any trading that takes place between Friday 21:01 GMT until Sunday 21:00 GMT.

(iv) Brent Crude Oil 2-hours Sunday trading session: We use two approaches for the additional 2-hour trading session that occurs in the Brent Crude Oil futures on Sundays. The first approach is to disregard these observations, whereas the second approach is to incorporate these observations to the Monday's trading day. The results of our forecasting exercise are not affected by the choice of the approach. Given the indifference in the results, we have decided to follow the second approach as it is more instructive to consider all available information in the construction of the realized volatility measure.

(v) Calendar or business-time sampling: We choose the calendar sampling as it is most commonly used in the literature and thus, allows for comparability of the results. Furthermore, as Sévi (2014) explains, the use of business-time sampling is not recommended as its asymptotic properties are less well-known.

(vi) Common sample: Finally, to arrive to a common sample across all series, we have considered the trading days when the Brent Crude Oil is traded[8].

After the aforementioned considerations, our final sample consists of a total of 56.71 million 1min observations for $T = 3028$ trading days.

3. Realized volatility

According to Andersen and Bollerslev (1998) the daily realized volatility is estimated as the sum of squared intra-day returns, as shown in eq.1:

$$DRV_t^{(\tau)} = \sqrt{\sum_{j=1}^{\tau} \left(logP_{t_j} - logP_{t_{j-1}}\right)^2}, \tag{1}$$

[8] If in any given day we have an observation for the oil but it is not a trading day for one of the other variables, then we use the value that this variable had the day before.

where P_t are the observed prices of the asset at trading day *t*, and τ are the equidistant time intervals.

The realized volatility converges to the integrated volatility as the sampling frequency (m) goes to zero and the number of time intervals (τ) approach infinity. Nevertheless, more noise is added to the estimated volatility when the sampling frequency verges on zero, due to microstructure frictions. Thus, there is a trade-off between the bias that is inserted in the in the realized volatility measure and its accuracy. Andersen *et al.* (2006) suggested the construction of the volatility signature plot, which depicts the average realized volatility against the sampling frequency. Based on the volatility signature plot, the optimal sampling frequency is the one where the average realized volatility starts to stabilise. In order to identify the point where the realized volatility appears to stabilise, we decompose the inter-day variance $(logP_t - logP_{t-1})^2$ into the intra-day variance $\left(\left(DRV_t^{(\tau)}\right)^2\right)$ and intra-day autocovariances $\left(\sum_{j=1}^{\tau-1}\sum_{i=j+1}^{\tau} y_{t_i} y_{t_{i-j}}\right)$, as in eq.2:

$$(logP_t - logP_{t-1})^2 = \left(DRV_t^{(\tau)}\right)^2 + 2\sum_{j=1}^{\tau-1}\sum_{i=j+1}^{\tau}\left(logP_{t_i} - logP_{t_{i-j}}\right). \quad (2)$$

The $\sum_{j=1}^{\tau-1}\sum_{i=j+1}^{\tau}\left(logP_{t_i} - logP_{t_{i-j}}\right)$ represents the bias that is inserted in the realized volatility measure, with $E\left(logP_{t_i} - logP_{t_{i-j}}\right) = 0$, for $j \neq 0$. Thus, the optimal sampling frequency (m) is the highest frequency that minimises the autocovariance bias. Table 1 shows the optimal sampling frequencies for our series.

[TABLE 1 HERE]

Furthermore, it is well established that when markets are closed, i.e. during the overnight periods, holidays and weekends, information still flows. The existing literature has proposed different approaches to dealing with this issue. For instance, authors such as, Andersen *et al.* (2001), Thomakos and Wang (2003) or Wu (2011), opine that overnight periods and weekends could be ignored from the construction of the realized volatility. By contrast, Hansen and Lunde (2005) maintain that a good proxy of the true volatility should accommodate the fact that information flows when markets are closed and thus, they proposed to adjust the intra-day volatility with the close-to-open inter-day volatility, as shown in eq.3:

$$DRV_{(HL),t}^{(\tau)} = \sqrt{\omega_1\left(logP_{t_1} - logP_{t-1_\tau}\right)^2 + \omega_2 \sum_{j=2}^{\tau}\left(logP_{t_j} - logP_{t_{j-1}}\right)^2}, \quad (3)$$

where the weights ω_1 and ω_2 are such that minimise the difference between the realized volatility and the integrated volatility, i.e. to minimise the variance of the realized volatility $\left(\min V\left(DRV_{t,(HL)}^{(\tau)}\right)\right)$. In this paper we are in line with Hansen and Lunde (2005) and thus we choose the second approach. Table 2 presents the descriptive statistics of our annualised realized volatility series $RV_{(HL),t}^{(\tau)}$:

$$RV_{(HL),t}^{(\tau)} = \sqrt{252} \times DRV_{(HL),t}^{(\tau)}, \quad (4)$$

for all variables and Figure 1 portrays their plots over the sample period.

[TABLE 2 HERE]

[FIGURE 1 HERE]

From Table 2 we notice that EPU has the highest average value and that it is very volatile, given its maximum, minimum and standard deviation values. From the realized volatilities, it is the natural gas (NG) that exhibits the highest average volatility, followed by the palladium (PA), silver (SV) and oil (CO). On the contrary, the lowest average volatilities are observed in the T-bills (TY) and the three exchange rate volatilities (BP, CD and EC). It is also evident that none of the series under consideration are normally distributed, where they exhibit excess kurtosis and positive skewness. Another interesting point is the average number of 1min observations that each series has, with the Eurostoxx 50 (XX), FTSE100 (FT) and Hang Seng (HI) to show the lowest figures, due to the shorter trading sessions that these markets have. The unit root test results support the hypothesis of stationary realized volatilities.

Furthermore, as it is apparent from Figure 1, volatility clustering of high values is observed for all series during the Global Financial Crisis (GFC) of 2007-09, although additional clusters of high volatility are evident in other periods for each series. Focusing on the Brent Crude Oil volatility, a second cluster of high volatility appears in the late 2014 – early 2015 period, mainly due to the plunge of the oil prices. Finally, we should mention that all autocorrelations (not shown here for brevity) decrease monotonically, suggest long-memory processes for our series.

4. Econometric specifications

4.1. Naïve models

We consider two naïve models, namely a simple Random Walk (RW) without a drift and an Autoregressive model of order 1, or AR(1), as shown in eqs. 5 and 6, respectively:

$$log\left(RV^{(\tau)}_{(HL),t}\right) = log\left(RV^{(\tau)}_{(HL),t-1}\right) + \varepsilon_t, \quad (5)$$

$$log\left(RV^{(\tau)}_{(HL),t}\right) = \varphi_0^{(t)}\left(1 - \varphi_1^{(t)}\right) + \varphi_1^{(t)} log\left(RV^{(\tau)}_{(HL),t-1}\right) + \varepsilon_t, \quad (6)$$

where $RV^{(\tau)}_{(HL),OIL,t}$ is the annualised realized volatility of the Brent crude oil at time *t*, $\varphi_0^{(t)}$, $\varphi_1^{(t)}$ are coefficients to be estimated and ε_t is a white noise.

4.2. HAR-RV model

We employ the HAR model by Corsi (2009), which is recently implemented in Haugom *et al.* (2014), Sévi (2015) and Prokopczuk *et al.* (2015). Eq. 7 presents the HAR-RV model

$$log\left(RV^{(\tau)}_{(HL),OIL,t}\right) = w_0^{(t)} + w_1^{(t)} log\left(RV^{(\tau)}_{(HL),OIL,t-1}\right) + w_2^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV^{(\tau)}_{(HL),OIL,t-k}\right)\right) + w_3^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV^{(\tau)}_{(HL),OIL,t-k}\right)\right) + \varepsilon_t, \quad (7)$$

where $RV^{(\tau)}_{(HL),OIL,t}$ is the annualised realized volatility of the Brent crude oil at time *t* and $w_0^{(t)}, w_1^{(t)}, w_2^{(t)}, w_3^{(t)}$ are parameters to be estimated. The HAR-RV model relates the current trading day's realized volatility of the Brent crude oil with the daily, weekly and monthly realized volatilities of the same asset.

4.3. HAR-RV-X model

We extend the HAR-RV model to incorporate exogenous variables, as discussed in Section 2. The HAR-RV-X model is shown in the following equation:

$$log\left(RV^{(\tau)}_{(HL),OIL,t}\right) = w_0^{(t)} + w_1^{(t)} log\left(RV^{(\tau)}_{(HL),OIL,t-1}\right) + w_2^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV^{(\tau)}_{(HL),OIL,t-k}\right)\right) + \quad (8)$$

$$w_3^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV_{(HL),OIL,t-k}^{(\tau)}\right)\right) + w_{(a),4}^{(t)} log\left(RV_{(HL),X_{(a)},t-1}^{(\tau)}\right) +$$
$$w_{(a),5}^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV_{(HL),X_{(a)},t-k}^{(\tau)}\right)\right) +$$
$$w_{(a),6}^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV_{(HL),X_{(a)},t-k}^{(\tau)}\right)\right) + \varepsilon_t,$$

where the (a) denotes the alternative fourteen (14) exogenous realized volatilities that are used in this paper. This model is extended to accommodate more than a single exogenous variable[9].

4.4. Forecasting realized volatility

Equations 7 and 8 are estimated in the natural logarithms of the realized volatilities. However, we are interested in forecasting the realized volatility (rather than its logarithm), which is variable of interest for traders, portfolio managers and policy makers. Thus, in our forecasts we concentrate on the estimator of the $RV_{(HL),OIL,t}^{(\tau)}$, which is the $exp\left(log\left(RV_{(HL),OIL,t}^{(\tau)}\right) + {}^{1}/_{2}\,\hat{\sigma}_{\varepsilon}^{2}\right)$. The HAR-RV 1-day-ahead forecast is as follows:

$$\begin{aligned} RV_{(HL),OIL,t+1|t}^{(\tau)} = exp\Bigg(& \widehat{w}_0^{(t)} + \widehat{w}_1^{(t)} log\left(RV_{(HL),OIL,t}^{(\tau)}\right) \\ & + \widehat{w}_2^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV_{(HL),OIL,t-k}^{(\tau)}\right)\right) \\ & + \widehat{w}_3^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV_{(HL),OIL,t-k}^{(\tau)}\right)\right) + {}^{1}/_{2}\,\hat{\sigma}_{\varepsilon}^{2}\Bigg) \end{aligned} \quad (9)$$

Equivalently, the HAR-RV-X model one-day-ahead forecast is shown in eq.10:

[9] We do not consider a multivariate HAR model for the following reason. The idea of a multivariate HAR model is to capture bidirectional effects between variables. However, given that in our model we consider the lagged values of the exogenous variables, we maintain that any effects the oil RV exercises in these variables, it is reflected in their lagged values. Thus, we remove part of the complexity of the model, without losing any significant information.

$$
\begin{aligned}
RV_{(HL),OIL,t+1|t}^{(\tau)} = exp\Bigg(& \widehat{w}_0^{(t)} + \widehat{w}_1^{(t)} log\left(RV_{(HL),OIL,t}^{(\tau)}\right) \\
& + \widehat{w}_2^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV_{(HL),OIL,t-k}^{(\tau)}\right)\right) \\
& + \widehat{w}_3^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV_{(HL),OIL,t-k}^{(\tau)}\right)\right) \\
& + \widehat{w}_{(a),4}^{(t)} log\left(RV_{(HL),X_{(a)},t}^{(\tau)}\right) \\
& + \widehat{w}_{(a),5}^{(t)}\left(5^{-1}\sum_{k=1}^{5} log\left(RV_{(HL),X_{(a)},t-k}^{(\tau)}\right)\right) \\
& + \widehat{w}_{(a),6}^{(t)}\left(22^{-1}\sum_{k=1}^{22} log\left(RV_{(HL),X_{(a)},t-k}^{(\tau)}\right)\right) + {}^{1}\!/_{2}\,\hat{\sigma}_{\varepsilon}^{2}\Bigg)
\end{aligned} \tag{10}
$$

The s-days-ahead forecasts ($s = 2, \dots, 66\ days$) are estimated in a similar fashion. More specifically, the s-days-ahead forecast of the HAR-RV model, for horizon ($s \geq 2$)

$$
\begin{gathered}
RV_{(HL),OIL,t+s|t}^{(\tau)} = \\
exp\left(\widehat{w}_0^{(t)} + \widehat{w}_1^{(t)} log\left(RV_{(HL),OIL,t+s-1|t}^{(\tau)}\right) + \right. \\
\widehat{w}_2^{(t)}\left(s^{-1}\textstyle\sum_{k=1}^{s} log\left(RV_{(HL),OIL,t-k+s|t}^{(\tau)}\right) + \right. \\
\left.(5-s)^{-1}\textstyle\sum_{k=s+1}^{5} log\left(RV_{(HL),OIL,t-k+s}^{(\tau)}\right)\right) + \\
\widehat{w}_3^{(t)}\left(s^{-1}\textstyle\sum_{k=1}^{s} log\left(RV_{(HL),OIL,t-k+s|t}^{(\tau)}\right) + \right. \\
\left.\left.(22-s)^{-1}\textstyle\sum_{k=s+1}^{22} log\left(RV_{(HL),OIL,t-k+s}^{(\tau)}\right)\right) + {}^{1}\!/_{2}\,\hat{\sigma}_{\varepsilon}^{2}\right)
\end{gathered} \tag{11}
$$

Finally, the s-days-ahead forecast of the HAR-RV-X model, horizon ($s \geq 2$)

$$
\begin{gathered}
RV_{(HL),OIL,t+s|t}^{(\tau)} = \\
exp\left(\widehat{w}_0^{(t)} + \widehat{w}_1^{(t)} log\left(RV_{(HL),OIL,t+s-1|t}^{(\tau)}\right) + \right. \\
\widehat{w}_2^{(t)}\left(s^{-1}\textstyle\sum_{k=1}^{s} log\left(RV_{(HL),OIL,t-k+s|t}^{(\tau)}\right) + \right. \\
\left.(5-s)^{-1}\textstyle\sum_{k=s+1}^{5} log\left(RV_{(HL),OIL,t-k+s}^{(\tau)}\right)\right) + \\
\widehat{w}_3^{(t)}\left(s^{-1}\textstyle\sum_{k=1}^{s} log\left(RV_{(HL),OIL,t-k+s|t}^{(\tau)}\right) + \right. \\
\left.(22-s)^{-1}\textstyle\sum_{k=s+1}^{22} log\left(RV_{(HL),OIL,t-k+s}^{(\tau)}\right)\right) + \widehat{w}_{(a),4}^{(t)} log\left(RV_{(HL),X_{(a)},t-k+s|t}^{(\tau)}\right) +
\end{gathered} \tag{12}
$$

$$\hat{w}_{(a),5}^{(t)} \left(s^{-1} \sum_{k=1}^{s} log \left(RV_{(HL),X_{(a)},t-k+s|t}^{(\tau)} \right) + \right.$$
$$\left. (5-s)^{-1} \sum_{k=s+1}^{5} log \left(RV_{(HL),X_{(a)},t-k+s}^{(\tau)} \right) \right) +$$
$$\hat{w}_{(a),6}^{(t)} \left(s^{-1} \sum_{k=1}^{s} log \left(RV_{(HL),X_{(a)},t-k+s|t}^{(\tau)} \right) + \right.$$
$$\left. (22-s)^{-1} \sum_{k=s+1}^{22} log \left(RV_{(HL),X_{(a)},t-k+s}^{(\tau)} \right) \right) + {}^{1}\!/_{2}\, \hat{\sigma}_{\varepsilon}^{2} \Big)$$

The exact forecasting strategy is detailed in the Section 5.

It is important to explain here how we proceed with the out-of-sample forecasts of the 1-day ahead until the 66-days ahead, as far as the HAR-RV and HAR-RV-X models are concerned. For the 1-day ahead forecast of the Brent Crude oil the models use data that belong to the information set at time *t* and thus, they are known to the forecaster at the time of the forecasting exercise. Nevertheless, from the 2-days ahead forecasts onwards (i.e. $t+2, \dots, t+66$), the forecast of the HAR-RV-X model of eq. (10) requires the use of future data that do not belong to the information set at time t. For example, for the $t+2$ forecast we need to know the $t+1$ volatility values of all variables. As far as the Brent Crude oil volatility is concerned, there is not an issue as the model uses the 1-day ahead forecast, i.e. at $t+1$. Turning to the exogenous variables, there are three possible choices to overcome the issue of using future data that do not belong to the information set at time *t*.

The first choice is to assume a zero value from $t+1$ onwards for the volatility(ies) of the exogenous variable(s), since the information is not available.

The second choice is to assume that at time $t+1$ onwards the volatility of the exogenous variable remains constant, i.e. $E\left(RV_{(HL),X_{(a)},t+1}^{(\tau)}\right) = RV_{(HL),X_{(a)},t}^{(\tau)}$. The concept that the best forecast of the next days' volatility value is today's value (plus a random component) is referred to as the random walk and it is based on the Efficient Market Hypothesis.

The third choice is to forecast volatilities of the exogenous variables and any data that are required for the estimation of the $t+2, \dots, t+66$ forecasts of the Brent crude oil volatility (which are not available at time *t*), they are taken from the forecasted values of the exogenous volatilities.

The first alternative is clearly rejected on the grounds that the second alternative is closely related to the finance literature and, thus, preferred. To proceed with the second choice, though, we would need to confirm that the RW generates the most accurate forecasts for the exogenous variables and thus confirms the EMH. To do so, we forecast each of the realized volatilities of the exogenous variables, using both a RW model and the HAR-RV model of eq. 7. Our results (not shown here for brevity but they are available upon request) reveal that the HAR-RV model is able to outperform the RW for each of the fourteen exogenous variables. Thus, we reject the second choice and we proceed with the Brent Crude oil forecasts based on the third choice. The third choice is shown in eq. 12, where we denote the information of the previous week's and previous month's exogenous volatilities as $\left(s^{-1}\sum_{k=1}^{s} log\left(RV_{(HL),X_{(a)},t-k+s|t}^{(\tau)}\right) + (5-s)^{-1}\sum_{k=s+1}^{5} log\left(RV_{(HL),X_{(a)},t-k+s}^{(\tau)}\right)\right)$ and $\left(s^{-1}\sum_{k=1}^{s} log\left(RV_{(HL),X_{(a)},t-k+s|t}^{(\tau)}\right) + (22-s)^{-1}\sum_{k=s+1}^{22} log\left(RV_{(HL),X_{(a)},t-k+s}^{(\tau)}\right)\right)$, respectively. The first term represents the information from the forecasted exogenous volatilities, where the second term indicates the information from the constructed realized exogenous volatilities.

This is an important innovation in our procedure. The existing literature either ignores this particular procedure and, thus, the forecasting accuracies of these papers can be put into question, or they fail to explain this.

5. Forecasting strategy

It is important to clearly explain the forecasting strategy that we follow, which is divided in 7 steps.

Step 1: We forecast the Brent crude oil realized volatility using the two naïve models (RW and AR(1)) and the HAR-RV and we assess which is the best performing model.

Step 2: We forecast the Brent crude oil realized volatility using the HAR-RV-X model, for each of the fourteen exogenous volatilities, although we group these variables into four asset classes (namely, *Stocks*, *Foreign exchange*, *Commodities* and *Macro*). The aim of this step is to identify the best HAR-RV-X model for each asset class. Taking for example the *Stocks*, we estimate four HAR-RV-X models, one for each stock index in our sample. We then compare the forecast of each HAR-RV-X model with the best performing model from *Step 1*. If any HAR-RV-X model performs better than the best model from *Step 1*, then we proceed with *Step 3*, otherwise we exclude this asset class from the remaining exercise.

Step 3: We forecast the Brent Crude oil realized volatility using the best HAR-RV-X from *Step 2*, adding each one of the remaining asset volatilities of the particular asset class. Continuing our example with *Stocks*, assuming that the best model from *Step 2* is the HAR-RV-SP, then we add to this model the realized volatility of the FTSE100, Hang Send and Euro Stoxx 50 (although one at a time), which gives us three new models at this step for the *Stocks* asset class. We then compare the forecast of each HAR-RV-X of *Step 3*, with the best model from *Step 2*. If a HAR-RV-X model from *Step 3* outperforms the best model from *Step 2*, then we proceed to *Step 4*, otherwise we stop and we claim that for this particular asset class, the best model is the one from *Step 2* (e.g. the HAR-RV-SP, in our example).

Step 4: We follow the same pattern as in *Step 3*, adding to the new best model each one of the remaining realized volatilities of the particular asset class. The same procedure is employed for all four asset classes and it is iterated as many times as it is required to reach at the best forecasting model from each asset class. Thus, at the end of this procedure we have four best models, which we name according to their respective asset class, i.e. HAR-RV-STOCKS, HAR-RV-FOREX, HAR-RV-COMMODITIES and HAR-RV-MACRO.

Step 5: In this step we proceed with the combined forecasts (HAR-RV-COMBINED) in order to assess whether the inclusion of more than one asset class could provide even better forecasts for the Brent crude oil volatility. To do so, we follow the same procedure as in the previous steps. More specifically, we use as our benchmark the best HAR-RV-X model that is identified from the previous steps (let us assume that the best model was the HAR-RV-STOCKS) and we add each one of the remaining three best models from each of the other asset classes (i.e. the HAR-RV-FOREX, HAR-RV-COMMODITIES and HAR-RV-MACRO). We then compare each of the three new models with the best model from *Step 4*. If any of the three new models from Step 5 is performing better than the best model from *Step 4*, then we proceed with *Step 6*, otherwise we provide evidence that the combined models do not offer any superior forecasts.

Step 6: At this step we proceed with our forecast using the best combined model from *Step 5*, adding each of the HAR-RV-X models of the remaining two asset classes. Once again if any of the two new models perform better than the best model from *Step 5*, we proceed by adding the last HAR-RV-X model of the last remaining asset class, otherwise we stop and we report the best combined forecasting model. The best model of this step is denoted as HAR-RV-COMBINED.

Step 7: The final step of our procedure is to produce model-averaged forecasts (HAR-RV-AVERAGE). The literature suggests that model-averaged forecasts could improve the forecasting accuracy, with equal weight averaging to work particularly well (see, for instance, Aiolfi and Favero, 2005; Timmermann, 2006; Samuels and Sekkel, 2013).

An indicative flow chart with the forecasting strategy is shown in Figure 2. As explained, the same procedure is applied for all asset classes and combined forecasts.

[FIGURE 2 HERE]

In total we estimate and evaluate 34 HAR-RV-X models in *Steps 2-4*, in order to finalise the best competing model from each asset class. We then estimate and evaluate 5 additional HAR-RV-X models, which correspond to *Steps 5* and *6*. At the final Step (*Step 7*) we estimate one additional model, which is the HAR-RV-AVERAGE. The forecast evaluation is described in Section 6.

The choice of this strategy is motivated by the fact that we want to extract the highest level of information from the exogenous variables, so that we can achieve the highest forecasting accuracy. In particular, based on the empirical research presented in Section 2 we have established that oil price volatility is impacted by four different "channels" (namely, *Stocks*, *Foreign exchange*, *Commodities* and *Macro*), which possibly transmit different information. In order to capture these different "channels" we need first to separate the variables according to their asset class. Furthermore, to assess whether the information flow from more than a single "channel" provides better forecasting accuracy, we proceed with estimation of the HAR-RV-COMBINED and HAR-RV-AVERAGE models.

6. Forecast evaluation

The initial sample period is $\tilde{T} = 1000$ days and we use the remaining $\breve{T} = 2028$ for our out-of-sample forecasting period. For the first out-of-sample forecast for the 1-day until 66-days ahead, we use the initial sample period $\tilde{T} = 1000$. For each subsequent forecast, we use a rolling window approach with fixed length of 1000 days. Engle *et al.* (1993), Angelidis *et al.* (2004) and Degiannakis *et al.* (2008)

maintain that the use of restricted samples are capable of capturing changes in the market activity better.

The forecasting accuracy of the models explained in Section 4 is initially evaluated using two well established evaluation functions, namely the Mean Squared Predicted Error (MSE) and the Mean Absolute Predicted Error (MAE):

$$MSE^{(s)} = T^{-1} \sum_{t=1}^{T} \left(RV_{(HL),OIL,t+s|t}^{(\tau)} - RV_{(HL),OIL,t+s}^{(\tau)} \right)^2, \tag{13}$$

and

$$MAE^{(s)} = T^{-1} \sum_{t=1}^{T} \left| RV_{(HL),OIL,t+s|t}^{(\tau)} - RV_{(HL),OIL,t+s}^{(\tau)} \right|, \tag{14}$$

where $RV_{(HL),OIL,t+s|t}^{(\tau)}$ is the Brent Crude oil realized volatility forecast, whereas $RV_{(HL),OIL,t+s}^{(\tau)}$ is the actual realized volatility.

Nevertheless, we depart from the standard setup of the forecasting evaluation, as this is presented in the previous works. The majority of the papers presented here compare the forecasts from a variety of models against a benchmark model, using the Diebold-Mariano test (Diebold and Mariano, 1995). In this study, however, we employ the newly established Model Confidence Set (MCS) procedure by Hansen *et al.* (2011), which identifies the set of the best models, as these are defined in terms of a specific loss function, without an *a priori* choice of a benchmark model. In our case, the two loss functions are the MSE and MAE.

The MCS explores the predictive ability of an initial set of M^0 models and investigates, at a predefined level of significance, which group of models survive an elimination algorithm. Let us define as $\Psi_{n,t}$ the evaluation function of model n at day t, and $d_{n,n^*,t} = \Psi_{n,t} - \Psi_{n^*,t}$ is the evaluation differential for $n, n^* \in M^0$. For example, the evaluation function may be the Mean Absolute Error, so $\Psi_{n,t} \equiv \left(RV_{(HL),OIL,t+s|t}^{(\tau)} - RV_{(HL),OIL,t+s}^{(\tau)} \right)^2$, where $RV_{(HL),OIL,t+s|t}^{(\tau)}$ is the s-days-ahead oil realized volatility forecast. The hypotheses that are being tested are:

$$H_{0,M}: E\left(d_{n,n^*,t}\right) = 0, \tag{15}$$

for $\vee\ n, n^* \in M$, $M \subset M^0$ against the alternative hypothesis $H_{1,M}: E\left(d_{n,n^*,t}\right) \neq 0$, for some $n, n^* \in M$. The elimination algorithm based on an equivalence test and an elimination rule, employs the equivalence test for investigating the $H_{0,M}$ for $\forall\ M \subset M^0$ and the elimination rule to identify the model n to be removed from M in the case that $H_{0,M}$ is rejected.

Finally, we consider the Direction-of-Change (DoC) an additional forecasting evaluation technique. The DoC is particularly important for market timing, which is essential for asset allocation and trading strategies. The DoC reports the proportion of forecasts that have correctly predicted the direction (up or down) of the volatility movement. Let us denote as $P_{n,i}^{(s)}$ a dummy variable that takes the value of 1 for each trading day i that model n correctly predicts the direction of the volatility movement s trading days ahead, and zero otherwise, i.e.:

$$P_{n,i}^{(s)} = \begin{cases} 1 \quad if \quad RV_{(HL),OIL,t+s}^{(\tau)} > RV_{(HL),OIL,t}^{(\tau)} \; and \; RV_{(HL),OIL,t+s|t}^{(\tau)} > RV_{(HL),OIL,t}^{(\tau)} \\ 1 \quad if \quad RV_{(HL),OIL,t+s}^{(\tau)} < RV_{(HL),OIL,t}^{(\tau)} \; and \; RV_{(HL),OIL,t+s|t}^{(\tau)} < RV_{(HL),OIL,t}^{(\tau)} \\ 0 \qquad\qquad\qquad otherwise \end{cases} \tag{16}$$

Then, the % proportion of forecasted values that have corrected predicted the direction of the volatility movement $\left(DoC_{prop}\right)$ is shown in eq. 15:

$$DoC_{prop}^{(s)} = \frac{\sum_{i=1}^{\check{T}} P_{n,i}^{(s)}}{\check{T}} \times 100, \tag{17}$$

where $\check{T}$ is the number of out-of-sample forecasted values. A standard x^2-test is applied to assess the significance of the $DoC_{prop}^{(s)}$.

7. Empirical results

7.1. MAE and MSE

We evaluate the forecasting accuracy of our models for 1-day until 66-days ahead, although we report six different horizons, namely 1-day, 5-days, 10-days, 22-day, 44-days and 66-days ahead. The results for the MAE and MSE are shown in Table 3.

[TABLES 3 HERE]

The first observation that we report from Table 3 is that the asset(s) that generate the best HAR-RV-X models for each of the asset classes, remains unchanged for all forecasting horizons in the cases of the *Commodities* and *Macro*. More specifically, it is the inclusion of both the Natural Gas (NG) and Silver (SV) realized volatilities that improve the simple HAR-RV forecasting accuracy regarding the Commodities asset class, whereas the HAR-RV-TY is the best performing model for the *Macro*.

By contrast, the assets' volatilities from the *Forex* and *Stocks* that contribute to the improvement of the HAR-RV model are different at the different forecasting horizons. In particular, in the case of *Stocks* we observe the HAR-RV-SP is the best model, although in the medium-run horizons (5-days to 22-days ahead) it is the HAR-RV-SP-XX, whereas for the longer-run horizons (i.e. 44-days and 66-days ahead) the best model is the HAR-RV-SP-FT. As far as the *Forex* is concerned we notice that until the 22-days ahead the best model is the HAR-RV-BP and it changes to HAR-RV-EC for the 44-days and 66-days ahead.

Focusing on the performance of each individual asset class, it is interesting to note that it is the HAR-RV-STOCK that provide the most accurate forecasts for the short- to medium-run horizons (until 22-days ahead), whereas the HAR-RV-MACRO assumes the role of the best performing model for the long-run horizons (i.e. 44-days and 66-days ahead). Furthermore, it is evident that the worse performing models are the two naïve models (RW and AR(1)), as well as, the HAR-RV model.

In terms, though, of the model that outperforms all others, this is clearly the HAR-RV-COMBINED, which includes in the same HAR-RV-X model the volatilities of more than a single asset class. A plausible explanation as to why the HAR-RV-COMBINED is the best performing model lies to the fact that oil price volatility is not influenced by a single asset class throughout the sample period, but rather it receives impact from different asset classes. Interestingly enough, the HAR-RV-AVERAGE model does not manage to improve further the forecasting accuracy.

This finding is of particular importance as the existing literature on the forecast of the oil realized volatility suggests that the HAR-RV generates the most accurate forecasts. In this paper we manage to provide superior forecasts compared to the HAR-RV model.

7.2. MCS procedure

Next, we discuss the results from the MCS procedure, reported in Table 4. The results from Table 3 may suggest that the HAR-RV-X models outperform the HAR-RV model, nevertheless it is vital to assess whether the HAR-RV could be included among the best performing models before we make any final conclusions.

[TABLES 4 HERE]

From Table 4 we can make the following observations. First and foremost, the HAR-RV-X models are always included in the set with the best performing models, for one or more forecasting horizons, whereas the two naïve models and the HAR-RV are never among the best performing models[10]. We also note that the highest probability is assigned to the HAR-RV-COMBINED across all horizons, with the only exception the p-values for the 1-day ahead based on the MSE loss function. Another very important finding, from Table 4, is the fact that as we move further out to the forecasting horizon it is only the HAR-RV-COMBINED model that belongs to the set of the best models.

7.3. Direction-of-Change

The DoC results are shown in Tables 5 and 6. Table 5 reports the proportion of forecasted values that have corrected predicted the direction of the volatility movement, whereas Table 6 compares the DoC performance of each HAR-RV- model against the HAR-RV.

[Tables 5 and 6 HERE]

Tables 5 and 6 show that all HAR models exhibit high accuracy in predicting the direction of the oil volatility movements. Interestingly enough, even though the HAR-RV model is not included among the best models (especially for the medium- and long-run forecasts, as suggested by the MCS test), its ability to predict the direction of change is comparable with all HAR-RV-X models. From Table 6 we notice more clearly that all HAR-RV-X models are performing marginally better compared to the HAR-RV and this holds for almost all forecasting horizons.

Overall, evidence suggests that the use of the exogenous volatilities of different asset classes results in the substantial improvement in the forecasting accuracy of the Brent Crude oil volatility. More importantly, though, we highlight that as we move towards longer-run forecasting horizons, where accurate forecasts are harder to be made, the set of the best performing models shrinks, leaving only the HAR-RV-COMBINED. On the other hand, focusing on the DoC we maintain that all models are highly accurate in predicting the direction of the oil volatility movements. Thus, the combination of the MCS and the DoC results reveals a very important finding, which has not been previously discussed in this strand of the literature.

More specifically, the findings reveal that for those stakeholders who are interested in the future movement of oil price volatility the simple HAR-RV model is adequate. Nevertheless, those stakeholders who put more emphasis on the accuracy of

[10] The only exception is the HAR-RV at the 1-day ahead forecast, based on the MSE loss function.

the forecasts, they should use the HAR-RV-X models and more specifically the HAR-X-COMBINED model. Finally, the fact that the HAR-RV-COMBINED outperforms all other models provides support to our claim that different asset classes provide different information to oil price volatility and thus, their combination improves the forecasting accuracy.

8. Robustness

Our first robustness check is related to the distribution of the forecast errors. More specifically, the squared difference and the absolute difference between the oil realized volatility forecast $\left(RV^{(\tau)}_{(HL),OIL,t+s|t}\right)$ and the realized volatility $\left(RV^{(\tau)}_{(HL),OIL,t+s}\right)$ is highly asymmetric. This suggests that the use of the median deviation may report a more accurate picture of the forecasting errors, not in terms of which is the best model, but rather on their magnitude. Thus, for example, even though the HAR-X-COMBINED model undoubtedly exhibits the higher forecasting accuracy, the actual deviation between the model's predicted volatilities and actual values may be lower than the reported ones from MSE and MAE. To illustrate this, we first present the distribution of the absolute and squared deviations between the forecasted values from HAR-RV-COMBINED and the actual oil realized volatility (see Figure 3).

[FIGURE 3 HERE]

As evident from Figure 3, the distribution of the deviations is highly skewed, which provides support to our claim that it is instructive to use the median deviations (i.e. the Median Absolute Error – MeAE or the Median Squared Error - MeSE), as they may assess better the magnitude of the prediction error.

[TABLES 7 HERE]

[FIGURE 4 HERE]

From Table 7 and Figure 4 we observe that as the forecasting horizon increases, the magnitude of the prediction errors differs greatly between the mean and median deviations. For example, the MAE (MSE) for the 1-day ahead forecasts is reported to be 5.3737 (69.3438), whereas the MeAE (MeSE) is estimated as 3.6294 (13.1724). Equivalently, for the 66-days ahead, even though the MAE (MSE) reports values of the magnitude of 9.2928 (207.3578), the MeAE (MeSE) are only 6.0662 (36.7994).

As a further robustness check we assess the validity of our findings in extreme economic conditions, such as the Global Financial Crisis of 2007-09. We follow the same forecasting evaluation procedure and we evaluate our forecasts only for the period August 2007 until June 2009. For brevity, we only present the results from the MCS procedure (see, Table 8).

[TABLE 8 HERE]

Table 8 suggests that the HAR-RV-X models are able to outperform the HAR-RV model, even during turbulent times. More specifically, the HAR-RV is not included in the set of the best performing models at any forecasting horizon, with the exception being the 1-day ahead, based only on the MSE loss function. Furthermore, it is evident that the best performing model is the HAR-RV-COMBINED, especially in the longer-run forecasting horizons. Overall, the MCS results shown in Table 8 corroborate the findings from Table 4. Therefore, the evidence provided by the robustness validates the proposed forecasting strategy plan, as it is effective even under extreme economic conditions.

9. Conclusion

The aim of this paper is to contribute to the limited but growing literature on oil price realized volatility forecasting. To do so we use tick by tick data of the front-month futures contracts for 14 asset prices. The period of our study spans from 1st of August, 2003 to 5th of August, 2015, which provides us with a total of 56.71 million 1min observations for 3028 trading days. Our forecasting horizons range from 1-day to 66-days-ahead, given that different stakeholders have different predictive needs.

The current consensus provides evidence that the HAR-RV model outperforms all other competing forecasting models (see, Haugom *et al.*, 2014; Sévi, 2014; Prokopczuk *et al.*, 2015). Our paper builds upon these previous contributions and extents them in multiple ways.

In short, our out-of-sample results suggest that the HAR-RV models with the exogenous volatilities from different asset classes (i.e. HAR-RV-X) outperform the forecasting accuracy of the HAR-RV at all forecasting horizons, contrary to the current consensus. In particular, we show that the HAR-RV-X models that combine multiple asset classes are the best performing models. Interestingly, enough the Direction of Change suggests that all HAR models are highly accurate in predicting the movements of oil price volatility. Thus, we maintain that HAR-RV-X models should be used from stakeholders who are interested in the accuracy of the forecasts, whereas those interested only in the movement of oil price volatility should be limited to HAR-RV. Our robustness section provides evidence that the forecasting accuracy of the models is better gauged using the Median Absolute Error and the Median Squared Error. Finally, it is important to note that our findings are robust even when we concentrate only on turbulent economic periods, such as the Global Financial Crisis of 2007-09.

More importantly, the fact that HAR-RV-X models that combine multiple asset classes' volatilities are the best performing models, provides strong support to our argument that different asset classes' volatilities provide important information for the forecast of oil price volatility, given that there different "channels" through which every asset class could impact oil price volatility.

An interesting avenue for further research is the use of our forecasting strategy for the prediction of other assets.

References

Agnolucci, P. (2009). Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics*, *31*(2), 316-321.

Ågren, M. (2006). Does oil price uncertainty transmit to stock markets? (No. 2006: 23). Working Paper, Department of Economics, Uppsala University.

Aiolfi, M., & Favero, C. A. (2005). Model uncertainty, thick modelling and the predictability of stock returns. *Journal of Forecasting*, *24*(4), 233-254.

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of Business*, 55-80.

Aloui, C., & Jammazi, R. (2009). The effects of crude oil shocks on stock market shifts behaviour: a regime switching approach. *Energy Economics*, *31*(5), 789-799.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 885-905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42-55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71, 579–625.

Andersen, T. G., Bollerslev, T., & Meddahi, N. (2005). Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica*, *73*(1), 279-296.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006). Volatility and correlation forecasting. *Handbook of Economic Forecasting*, *1*, 777-878.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, *89*(4), 701-720.

Andersen, T. G., Bollerslev, T., Frederiksen, P., & Ørregaard Nielsen, M. (2010). Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. *Journal of Applied Econometrics*, *25*(2), 233-261.

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology*, *1*(1), 105-128.

Angelidis, T., & Degiannakis, S. (2008). Volatility forecasting: Intra-day versus inter-day models. *Journal of International Financial Markets, Institutions and Money*, *18*(5), 449-465.

Antonakakis, N., Chatziantoniou, I., & Filis, G. (2014). Dynamic spillovers of oil price shocks and economic policy uncertainty. *Energy Economics*, 44, 433-447.

Antonakakis, N., & Kizys, R. (2015). Dynamic spillovers between commodity and currency markets. *International Review of Financial Analysis*.

Arouri, M. E. H., Jouini, J., & Nguyen, D. K. (2011). Volatility spillovers between oil prices and stock sector returns: implications for portfolio management. *Journal of International Money and Finance*, *30*(7), 1387-1405.

Arouri, M. E. H., Lahiani, A., Lévy, A., & Nguyen, D. K. (2012). Forecasting the conditional volatility of oil spot and futures prices with structural breaks and long memory models. *Energy Economics*, *34*(1), 283-293.

Baker, S. R., Bloom, N., & Davis, S. J. (2013). Measuring economic policy uncertainty. *Chicago Booth research paper*, (13-02).

Baumeister, C., & Peersman, G. (2013). Time-Varying Effects of Oil Supply Shocks on the US Economy. *American Economic Journal: Macroeconomics*, *5*(4), 1-28.

Beckmann, J., & Czudaj, R. (2013). Oil prices and effective dollar exchange rates. *International Review of Economics & Finance*, 27, 621-636.

Bernanke, B. S., & Mishkin, F. S. (1997). *Inflation targeting: a new framework for monetary policy?* (No. w5893). National Bureau of Economic Research.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1), 5-59.

Bollerslev, T., & Wright, J. H. (2001). High-frequency data, frequency domain inference, and volatility forecasting. *Review of Economics and Statistics*, *83*(4), 596-602.

Brooks, C., & Persand, G. (2002). Model choice and value-at-risk performance. *Financial Analysts Journal*, *58*(5), 87-97.

Brooks, C., & Persand, G. (2003). Volatility forecasting for risk management. *Journal of Forecasting*, *22*(1), 1-22.

Brown, S. P., & Yücel, M. K. (2002). Energy prices and aggregate economic activity: an interpretative survey. *The Quarterly Review of Economics and Finance*, *42*(2), 193-208.

Burbidge, J., & Harrison, A. (1984). Testing for the effects of oil-price rises using vector autoregressions. *International Economic Review*, 459-484.

Busch, T., Christensen, B. J., and Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, 160, 48–57.

Buyuksahin, B., & Robe, M. A. (2011). Does' paper oil'matter? Energy markets' financialization and equity-commodity co-movements. *Energy Markets' Financialization and Equity-Commodity Co-Movements (July 28, 2011)*.
Buyuksahin, B., Haigh, M. S., & Robe, M. A. (2010). Commodities and equities: Ever a market of one?. *Journal of Alternative Investments, Winter*.
Buyuksahin, B., & Robe, M. A. (2014). Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, *42*, 38-70.
Chatrath, A., Miao, H., Ramchander, S., & Wang, T. (2015). The Forecasting Efficacy of Risk-Neutral Moments for Crude Oil Volatility. *Journal of Forecasting*, *34*(3), 177-190.
Chen, Y-C., Rogoff, S.K., & Rossi, B. (2010). Can exchange rates forecast commodity prices?. *The Quarterly Journal of Economics*, *125*(3), 1145-1194.
Chkili, W., Hammoudeh, S., & Nguyen, D. K. (2014). Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Economics*, *41*, 1-18.
Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, *50*(2), 125-150.
Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7, 174–196.
Degiannakis, S. (2004). Volatility forecasting: evidence from a fractional integrated asymmetric power ARCH skewed-t model. *Applied Financial Economics*, *14*(18), 1333-1342.
Degiannakis, S. (2008). Forecasting VIX. *Journal of Money, Investment and Banking*, 4, 5-19.
Degiannakis, S., Livada, A., & Panas, E. (2008). Rolling-sampled parameters of ARCH and Levy-stable models. *Applied Economics*, *40*(23), 3051-3067.
Degiannakis, S., Filis, G., & Floros, C. (2013). Oil and stock returns: Evidence from European industrial sector indices in a time-varying environment. *Journal of International Financial Markets, Institutions and Money*, *26*, 175-191.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13, 253-263.
Efimova, O., & Serletis, A. (2014). Energy markets volatility modelling using GARCH. *Energy Economics*, *43*, 264-273.
Elder, J., & Serletis, A. (2010). Oil price uncertainty. *Journal of Money, Credit and Banking*, *42*(6), 1137-1159.
Engle, R.F., Hong, C.H., Kane, A. and Noh, J. (1993). Arbitrage Valuation of Variance Forecasts with Simulated Options, Advances in Futures and Options Research, 6, 393-415.
Engle, E., & Sun, Z. (2007). When is noise not noise–a microstructure estimate of realized volatility.
Ferderer, J. P. (1997). Oil price volatility and the macroeconomy. *Journal of Macroeconomics*, *18*(1), 1-26.
Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, *40*, 1-10.
Filis, G. (2010). Macro economy, stock market and oil prices: Do meaningful relationships exist among their cyclical fluctuations?. *Energy Economics*, *32*(4), 877-886.
Filis, G., Degiannakis, S., & Floros, C. (2011). Dynamic correlation between stock market and oil prices: The case of oil-importing and oil-exporting countries. *International Review of Financial Analysis*, *20*(3), 152-164.

Filis, G., & Chatziantoniou, I. (2014). Financial and monetary policy responses to oil price shocks: evidence from oil-importing and oil-exporting countries. *Review of Quantitative Finance and Accounting*, *42*(4), 709-729.

Fratzscher, M., Schneider, D., & Van Robays, I. (2014). Oil prices, exchange rates and asset prices.

Frijns, B., Tallau, C., & Tourani-Rad, A. (2010). The information content of implied volatility: evidence from Australia. *Journal of Futures Markets*, *30*(2), 134-155.

Fuertes, A. M., Izzeldin, M., & Kalotychou, E. (2009). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, *25*(2), 259-281.

Guesmi, K., & Fattoum, S. (2014). Return and volatility transmission between oil prices and oil-exporting and oil-importing countries. *Economic Modelling*, *38*, 305-310.

Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *The Journal of Political Economy*, 228-248.

Hamilton, J. D., & Wu, J. C. (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance*, *42*, 9-37.

Hammoudeh, S., Dibooglu, S., & Aleisa, E. (2004). Relationships among US oil prices and oil industry equity indices. *International Review of Economics & Finance*, *13*(4), 427-453.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. *Journal of Applied Econometrics*, *20*(7), 873-889.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453-497.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014). Forecasting volatility of the US oil market. *Journal of Banking & Finance*, *47*, 1-14.

Hou, A., & Suardi, S. (2012). A nonparametric GARCH model of crude oil price return volatility. *Energy Economics*, *34*(2), 618-626.

Huang, R. D., Masulis, R. W., & Stoll, H. R. (1996). Energy shocks and financial markets. *Journal of Futures Markets*, *16*(1), 1-27.

IEA (2015). What drives crude oil prices? US International Energy Administration, July 07. https://www.eia.gov/finance/markets/spot_prices.cfm

Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *Review of Financial Studies*, *18*(4), 1305-1342.

Jones, C. M., & Kaul, G. (1996). Oil and the stock markets. *The Journal of Finance*, 51(2), 463-491.

Kat, H. M., & Oomen, R. C. (2007). What every investor should know about commodities Part II: multivariate return analysis. *Journal of Investment Management*, *5*(3), 1–25.

Kang, S. H., Kang, S. M., & Yoon, S. M. (2009). Forecasting volatility of crude oil markets. *Energy Economics*, *31*(1), 119-125.

Kang, S.-H., & Yoon, S.-M. (2013). Modelling and forecasting the volatility of petroleum futures prices. *Energy Economics*, 36, 354–362.

Kilian, L., & Park, C. (2009). The impact of oil price shocks on the US stock market. *International Economic Review*, *50*(4), 1267-1287.

Malik, F., & Ewing, B. T. (2009). Volatility transmission between oil prices and equity sector returns. *International Review of Financial Analysis*, *18*(3), 95-100.

Masih, R., Peters, S., & De Mello, L. (2011). Oil price volatility and stock price fluctuations in an emerging market: evidence from South Korea. *Energy Economics*, 33(5), 975-986.

McAleer, M., & Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, 27(1-3), 10-45.

McKenzie, M., & Mitchell, H. (2002). Generalized asymmetric power ARCH modelling of exchange rate volatility. *Applied Financial Economics*, *12*(8), 555-564.

Mensi, W., Beljid, M., Boubaker, A., & Managi, S. (2013). Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. *Economic Modelling*, *32*, 15-22.

Mensi, W., Hammoudeh, S., Nguyen, D. K., & Yoon, S. M. (2014). Dynamic spillovers among major energy and cereal commodity prices. *Energy Economics*, *43*, 225-243.

Nomikos, N. K., & Pouliasis, P. K. (2011). Forecasting petroleum futures markets volatility: The role of regimes and market conditions. *Energy Economics*, *33*(2), 321-337.

Oomen, R. (2001). Using High Frequency Data to Calculate, Model and Forecast Realized Volatility (No. 75). Society for Computational Economics.

Phan, D. H. B., Sharma, S. S., & Narayan, P. K. (2015). Intraday volatility interaction between the crude oil and equity markets. *Journal of International Financial Markets, Institutions and Money*, forthcoming.

Prokopczuk, M., Symeonidis, L., & Wese Simen, C. (2015). Do Jumps Matter for Volatility Forecasting? Evidence from Energy Markets. *Journal of Futures Markets*, 1-35.

Rahman, S., & Serletis, A. (2011). The asymmetric effects of oil price shocks. *Macroeconomic Dynamics*, *15*(S3), 437-471.

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics*, 21, 449–469.

Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. *Energy Economics*, 28, 467–488.

Sadorsky, P. (2014). Modelling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat. *Energy Economics*, 43, 72-8.

Sadorsky, P., & McKenzie, M. D. (2008). Power transformation models and volatility forecasting. *Journal of Forecasting*, 27, 587–606.

Samuels, J. D., & Sekkel, R. M. (2013). *Forecasting with many models: Model confidence sets and forecast combination* (No. 2013-11). Bank of Canada Working Paper.

Sari, R., Hammoudeh, S., & Soytas, U. (2010). Dynamics of oil price, precious metal prices, and exchange rate. *Energy Economics*, *32*(2), 351-362.

Sévi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research*, *235*(3), 643-659.

Silvennoinen, A., & Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, *24*, 42-65.

Souček, M., & Todorova, N. (2013). Realized volatility transmission between crude oil and equity futures markets: A multivariate HAR approach. *Energy Economics*, *40*, 586-597.

Souček, M., & Todorova, N. (2014). Realized volatility transmission: The role of jumps and leverage effects. *Economics Letters*, *122*(2), 111-115.
Tang, K., & Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, *68*(5), 54-74.
Tay, A., Ting, C., Tse, Y. K., & Warachka, M. (2009). Using high-frequency transaction data to estimate the probability of informed trading. *Journal of Financial Econometrics*, 7(3), 288-311.
Thomakos, D. D., & Wang, T. (2003). Realized volatility in the futures markets. *Journal of Empirical Finance*, *10*(3), 321-353.
Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, *1*, 135-196.
Wei, Y., Wang, Y., & Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. *Energy Economics*, 32, 1477–1484.
West, K. D., & Cho, D. (1995). The predictive ability of several models of exchange rate volatility. *Journal of Econometrics*, *69*(2), 367-391.
Wu, L. (2011). Variance dynamics: Joint evidence from options and high-frequency returns. *Journal of Econometrics*, *160*(1), 280-287.

Tables

Table 1: Optimal Sampling frequencies for the realized volatility construction.

Variable	Futures ticker	Market	*m*
Brent Crude Oil	CO	ICE Futures Europe	23
GBP/USD	BP	CME Group	30
CAD/USD	CD	CME Group	25
EUR/USD	EC	CME Group	16
FTSE100	FT	ICE Futures Europe	1
S&P500	SP	CME Group	6
Hang Seng	HI	Hong Kong Stock Exchange	60
Euro Stoxx 50	XX	Eurex	3
Gold	GC	CME Group	15
Copper	HG	CME Group	20
Natural Gas	NG	CME Group	10
Palladium	PA	CME Group	90
Silver	SV	CME Group	28
US T-bill 10yr	TY	CME Group	15

Note: *m* denotes the optimal sampling frequency.

Table 2: Descriptive statistics of the series under investigation. The sample runs from 1st August 2003 to 5th August 2015.

	OIL	*FOREX*			*STOCK MARKETS*			
	CO	BP	CD	EC	SP	XX	FT	HI
Mean	28.670	8.841	9.558	9.402	15.826	20.849	16.595	18.807
Maximum	164.219	49.515	47.253	35.554	147.168	150.027	143.174	158.970
Minimum	5.075	2.187	2.396	2.258	4.435	6.324	4.302	2.295
Std. Dev.	14.911	4.352	4.313	3.914	11.383	11.585	11.192	13.899
Skewness	2.447	2.509	1.982	1.675	3.850	3.231	3.485	3.202
Kurtosis	13.119	13.146	10.641	8.200	26.030	20.522	22.222	19.995
ADF *(p-value)*	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000
PP *(p-value)*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Jarque-Bera *(p-value)*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Average 1min obs/day	1308	1362	1362	1364	1332	752	663	453
Daily obs	3028	3028	3028	3028	3028	3028	3028	3028

	COMMODITIES					*MACRO*	
	GC	HG	NG	PA	SV	TY	EPU
Mean	17.782	27.337	46.720	32.653	31.118	5.937	100.953
Maximum	98.135	144.822	424.579	141.978	200.658	48.187	626.028
Minimum	4.247	7.133	14.861	4.470	8.418	1.874	7.396
Std. Dev.	8.606	14.178	21.813	18.956	15.995	2.957	68.328
Skewness	2.616	2.355	4.301	1.902	2.946	3.075	1.780
Kurtosis	15.390	11.903	53.480	8.278	19.073	25.931	8.343
ADF *(p-value)*	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PP *(p-value)*	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Jarque-Bera *(p-value)*	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Average 1min obs/day	1211	1210	1229	1205	1207	1219	
Daily obs	3028	3028	3028	3028	3028	3028	3028

Note: The values here are based on the annualised realized volatilities that have been scaled according to the Hansen and Lunde (2005) approach. CO=Brent Crude oil, BP=GBP/USD, CD=CAD/USD, EC=EUR/USD, SP=S&P500, XX=Euro Stoxx 50, FT=FTSE100, HI=Hang Seng, GC=Gold, HG=Copper, NG=Natural gas, PA=Palladium, SV=Silver, TY=US 10yr T-Bills, EPU=Economic Policy Uncertainty.

Table 3: MAE and MSE results

MAE results

Days ahead	RW	AR(1)	HAR-RV	HAR-RV-STOCKS	HAR-RV-FOREX	HAR-RV-COMMODITIES	HAR-RV-MACRO	HAR-RV-COMBINED	HAR-RV-AVERAGE
1	11.3344	6.7866	5.4373	5.3734	5.4153	5.4141	5.4074	5.3737	5.3824
5	11.3832	9.9281	6.1556	6.0386	6.1041	6.1106	6.1046	6.0262	6.0520
10	11.4482	11.0303	6.6644	6.4666	6.5919	6.5474	6.5195	6.4634	6.4809
22	11.5643	11.5289	7.6183	7.2579	7.4730	7.3518	7.3318	7.2301	7.2527
44	11.7839	11.7682	9.1349	8.6524	8.7513	8.6485	8.5070	8.3560	8.4734
66	12.0567	12.0409	10.3106	9.7829	9.6814	9.7320	9.4164	9.2928	9.4696

MSE results

Days ahead	RW	AR(1)	HAR-RV	HAR-RV-STOCKS	HAR-RV-FOREX	HAR-RV-COMMODITIES	HAR-RV-MACRO	HAR-RV-COMBINED	HAR-RV-AVERAGE
1	289.4508	120.0570	70.6880	69.5096	69.6735	69.5644	69.7509	69.3438	69.2898
5	290.9137	245.6501	90.1713	86.5421	87.3805	87.9248	87.5526	85.8934	86.5780
10	292.9822	281.2121	104.7010	97.9279	99.4978	99.4999	99.1631	96.5292	97.7476
22	296.8560	295.7655	140.0061	123.6066	125.5833	127.3616	125.7728	117.7922	123.1893
44	303.9693	303.4793	202.6877	179.2979	173.0563	176.6120	171.8393	163.1676	171.1347
66	312.2399	311.8076	264.7186	237.6645	220.5482	233.7457	218.1795	207.3578	223.9654

Note: The HAR-RV-STOCKS model for the 1-day ahead is the HAR-RV-SP, for the 5-days until 22-days ahead is the HAR-RV-SP-XX and for the 44-days and 66-days ahead is the HAR-RV-SP-FT. The HAR-RV-FOREX model for the 1-day until 22-days ahead is the HAR-RV-BP and for the 44-days and 66-days ahead is the HAR-RV-EC. The HAR-RV-COMMODITIES model is the HAR-RV-NG-SV for all horizons. The HAR-RV-MACRO model is the HAR-RV-TY for all horizons. The HAR-RV-COMBINED is the HAR-RV-SP-BP for the 1-day ahead, the HAR-RV-SP-XX-BP for the 5-days until 22-days ahead, the HAR-RV-NG-SV-TY for the 44-days ahead and the HAR-RV-EC-TY for the 66-days ahead.

Table 4: Comparison between MAE (MSE) and MeAE (MeSE)

	1-day	5-days	10-days	22-days	44-days	66-days
MAE	5.3737	6.0262	6.4634	7.2301	8.3561	9.2928
MeAE	3.6294	4.0780	4.3694	4.7989	5.5915	6.0662
MSE	69.3438	85.8935	96.5292	117.7922	163.1676	207.3578
MeSE	13.1724	16.6305	19.0917	23.0292	31.2649	36.7994

Note: These values correspond to the HAR-RV-COMBINED model.

Table 5: MCS p-values for 1-day to 66-days ahead

Loss function: MAE	Forecasting Horizon					
	1-day	5-days	10-days	22-days	44-days	66-days
	p-values are reported					
RW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR(1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV	0.0116	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV-STOCKS	**1.0000**	**0.6740**	**0.9320**	**0.8591**	0.0046	0.0000
HAR-RV-FOREX	**0.1385**	0.0578	0.0028	0.0000	0.0000	0.0000
HAR-RV-COMMODITIES	**0.4810**	0.0589	**0.1626**	0.0814	0.0000	0.0004
HAR-RV-MACRO	**0.1745**	0.0200	**0.3003**	0.0802	0.0087	0.0350
HAR-RV-COMBINED	**0.9844**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
HAR-RV-AVERAGE	**0.8259**	**0.6740**	**0.8805**	**0.8591**	0.0182	0.0236
Loss function: MSE	Forecasting Horizon					
	1-day	5-days	10-days	22-days	44-days	66-days
	p-values are reported					
RW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR(1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV	**0.2293**	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV-STOCKS	**0.9493**	**0.6142**	**0.4368**	0.0303	0.0000	0.0000
HAR-RV-FOREX	**0.8862**	**0.2289**	0.0426	0.0002	0.0000	0.0000
HAR-RV-COMMODITIES	**0.9493**	**0.1006**	**0.1076**	0.0030	0.0000	0.0000
HAR-RV-MACRO	**0.3969**	**0.1071**	**0.1076**	0.0303	0.0001	0.0001
HAR-RV-COMBINED	**0.9493**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
HAR-RV-AVERAGE	**1.0000**	**0.6142**	**0.4368**	0.0303	0.0000	0.0000

Note: figures in bold denote that the model belongs to the confidence set of the best performing models. The interpretation of the MCS p-value is analogous to that of a classical p-value; a $(1-a)$ confidence interval that contains the 'true' parameter with a probability no less than $(1-a)$.

Table 6: Direction-of-Change results

Days ahead	HAR-RV	HAR-RV-STOCKS	HAR-RV-FOREX	HAR-RV-COMMODITIES	HAR-RV-MACRO	HAR-RV-COMBINED	HAR-RV-AVERAGE
1	0.6835*	0.6835*	0.6881*	0.6794*	0.6840*	0.6865*	0.6879*
5	0.6702*	0.6718*	0.6779*	0.6672*	0.6753*	0.6764*	0.6726*
10	0.6651*	0.6702*	0.6606*	0.6697*	0.6641*	0.6651*	0.6701*
22	0.6493*	0.6463*	0.6366*	0.6488*	0.6529*	0.6483*	0.6446*
44	0.6162*	0.6177*	0.6249*	0.6244*	0.6228*	0.6249*	0.6272*
66	0.6009*	0.6086*	0.6091*	0.6070*	0.6142*	0.6096*	0.6079*

Note: The HAR-RV-STOCKS model for the 1-day ahead is the HAR-RV-SP, for the 5-days until 22-days ahead is the HAR-RV-SP-XX and for the 44-days and 66-days ahead is the HAR-RV-SP-FT. The HAR-RV-FOREX model for the 1-day until 22-days ahead is the HAR-RV-BP and for the 44-days and 66-days ahead is the HAR-RV-EC. The HAR-RV-COMMODITIES model is the HAR-RV-NG-SV for all horizons. The HAR-RV-MACRO model is the HAR-RV-TY for all horizons. The HAR-RV-COMBINED is the HAR-RV-SP-BP for the 1-day ahead, the HAR-RV-SP-XX-BP for the 5-days until 22-days ahead, the HAR-RV-NG-SV-TY for the 44-days ahead and the HAR-RV-EC-TY for the 66-days ahead.
* denotes significance at 1% level.

Table 7: Direction-of-Change results: Comparing models with HAR-RV.

Days ahead	$\frac{HAR-RV-STOCKS}{HAR-RV}$	$\frac{HAR-RV-FOREX}{HAR-RV}$	$\frac{HAR-RV-COMMODITIES}{HAR-RV}$	$\frac{HAR-RV-MACRO}{HAR-RV}$	$\frac{HAR-RV-COMBINED}{HAR-RV}$	$\frac{HAR-RV-AVERAGE}{HAR-RV}$
1	1.0000	1.0067	0.9940	1.0007	1.0045	1.0065
5	1.0023	1.0114	0.9954	1.0076	1.0091	1.0036
10	1.0077	0.9931	1.0069	0.9985	1.0000	1.0074
22	0.9953	0.9804	0.9992	1.0055	0.9984	0.9927
44	1.0025	1.0141	1.0132	1.0108	1.0141	1.0179
66	1.0127	1.0136	1.0102	1.0221	1.0144	1.0115

Note: The HAR-RV-STOCKS model for the 1-day ahead is the HAR-RV-SP, for the 5-days until 22-days ahead is the HAR-RV-SP-XX and for the 44-days and 66-days ahead is the HAR-RV-SP-FT. The HAR-RV-FOREX model for the 1-day until 22-days ahead is the HAR-RV-BP and for the 44-days and 66-days ahead is the HAR-RV-EC. The HAR-RV-COMMODITIES model is the HAR-RV-NG-SV for all horizons. The HAR-RV-MACRO model is the HAR-RV-TY for all horizons. The HAR-RV-COMBINED is the HAR-RV-SP-BP for the 1-day ahead, the HAR-RV-SP-XX-BP for the 5-days until 22-days ahead, the HAR-RV-NG-SV-TY for the 44-days ahead and the HAR-RV-EC-TY for the 66-days ahead.

Table 8: MCS p-values for 1-day to 66-days ahead: During crisis period

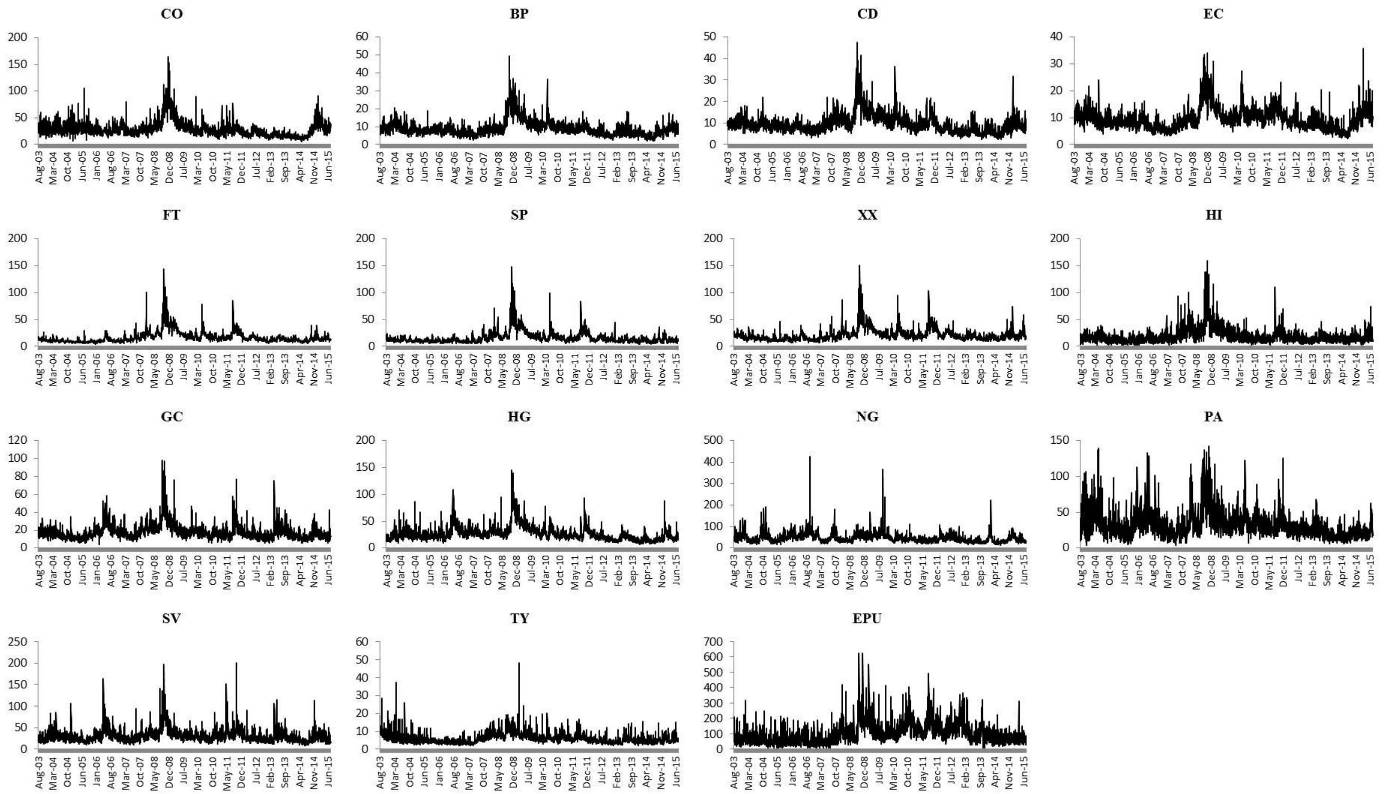
Loss function: MAE	Forecasting Horizon					
	1-day	5-days	10-days	22-days	44-days	66-days
	p-values are reported					
RW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR(1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV	0.0251	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV-STOCKS	**0.9047**	**0.5657**	0.0587	0.0001	0.0046	0.0000
HAR-RV-FOREX	**0.9047**	**0.9642**	**0.9222**	**0.4875**	0.0000	0.0000
HAR-RV-COMMODITIES	**0.9047**	**0.9642**	**0.9900**	**1.0000**	0.0001	0.0000
HAR-RV-MACRO	**0.9047**	**0.5657**	**0.9900**	**0.4875**	0.0000	0.0000
HAR-RV-COMBINED	**1.0000**	**1.0000**	**0.9222**	**0.4875**	**1.0000**	**1.0000**
HAR-RV-AVERAGE	**0.9047**	**0.9682**	**1.0000**	**0.4875**	0.0000	0.0000

Loss function: MSE	Forecasting Horizon					
	1-day	5-days	10-days	22-days	44-days	66-days
	p-values are reported					
RW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR(1)	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV	**0.1656**	0.0000	0.0000	0.0000	0.0000	0.0000
HAR-RV-STOCKS	**0.4984**	**0.2302**	0.0080	0.0003	0.0000	0.0000
HAR-RV-FOREX	**1.0000**	**1.0000**	**0.9863**	**0.1536**	0.0000	0.0000
HAR-RV-COMMODITIES	**0.9983**	**0.6918**	**0.9425**	**0.1752**	0.0000	0.0000
HAR-RV-MACRO	**0.9983**	**0.8484**	**0.9820**	**0.1752**	0.0000	0.0000
HAR-RV-COMBINED	**0.8862**	**0.9820**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
HAR-RV-AVERAGE	**0.9983**	**0.9609**	**0.9863**	**0.1752**	0.0000	0.0000

Note: figures in bold denote that the model belongs to the confidence set of the best performing models. The interpretation of the MCS p-value is analogous to that of a classical p-value; a $(1-a)$ confidence interval that contains the 'true' parameter with a probability no less than $(1-a)$.

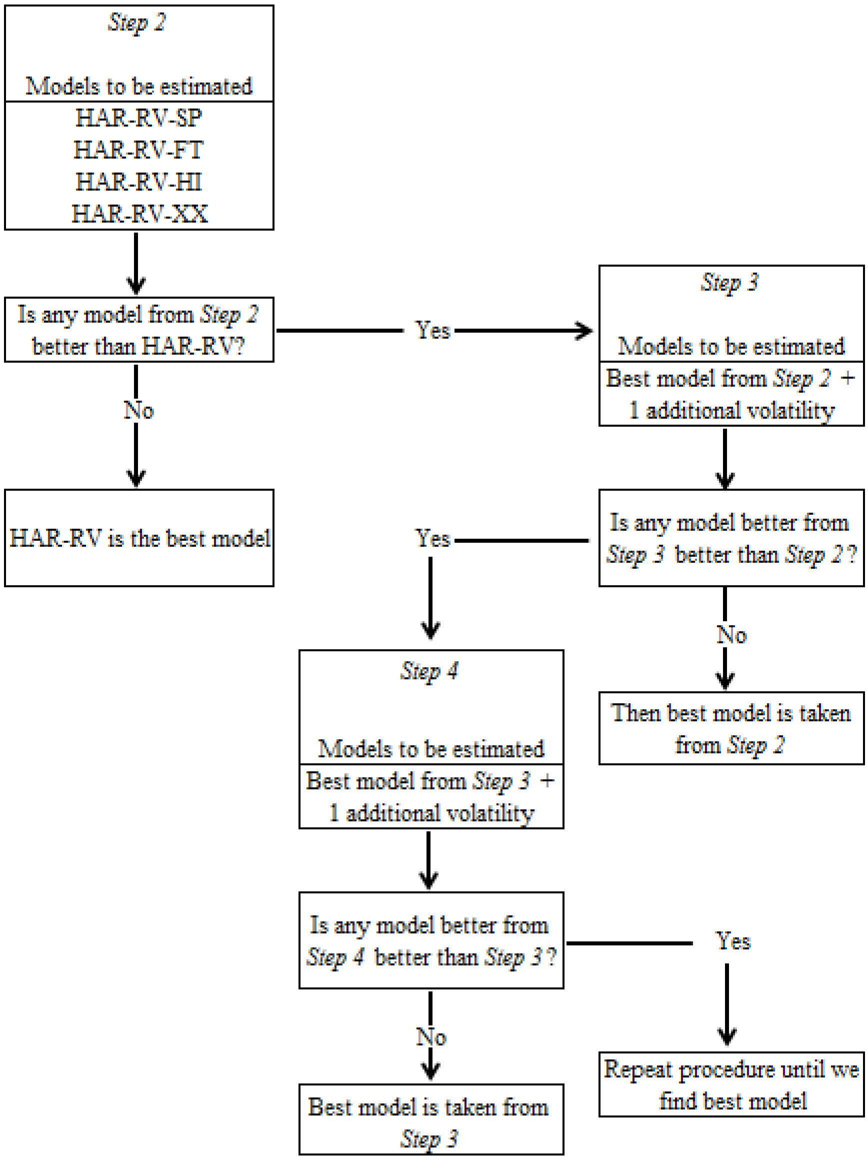
Figures

Figure 1: Realized volatility plots



Note: From the top left to the bottom right the variables are as follows: CO=Brent Crude oil, BP=GBP/USD, CD=CAD/USD, EC=EUR/USD, FT=FTSE100, SP=S&P500, XX=Euro Stoxx 50, HI=Hang Seng, GC=Gold, HG=Copper, NG=Natural gas, PA=Palladium, SV=Silver, TY=US 10yr T-Bills, EPU=Economic Policy Uncertainty. All value in the y-axis refers to percentages.

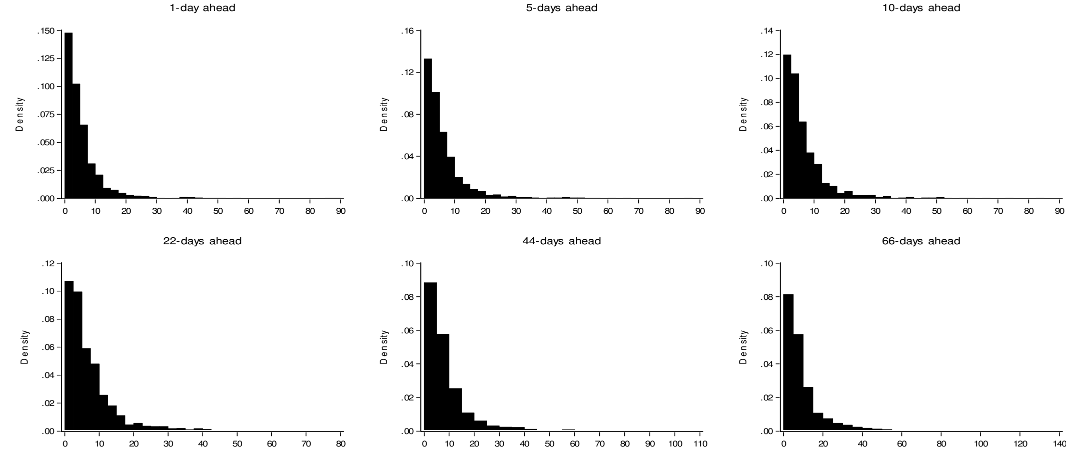
Figure 2: Forecasting strategy flow chart



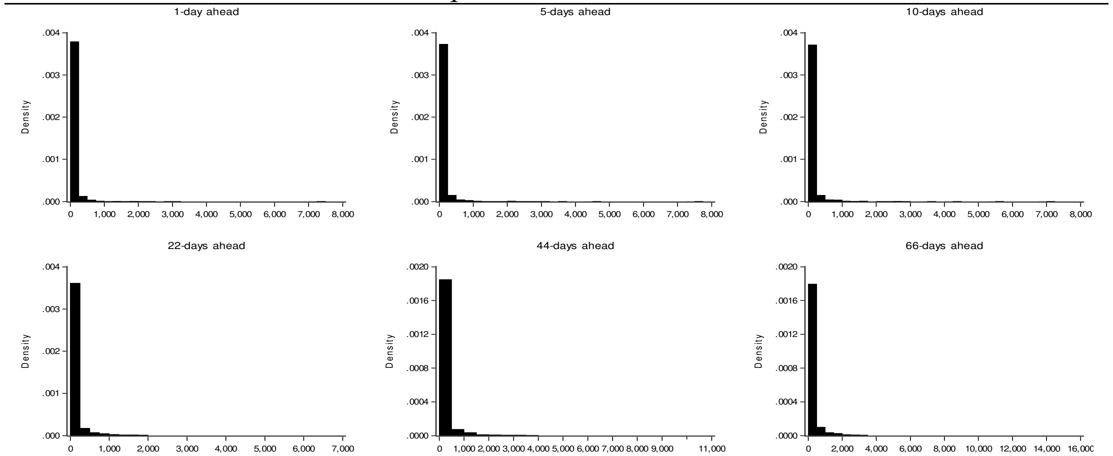
Note: This flow chart indicatively shows the forecasting strategy for the *Stocks* asset class. The same procedure is followed for all remaining three flow charts. The best model for the *Stocks* is denoted as HAR-RV-STOCKS.

Figure 3: Distribution of the absolute and squared deviations between the actual oil realized volatility and the predicted values.

Absolute Deviations

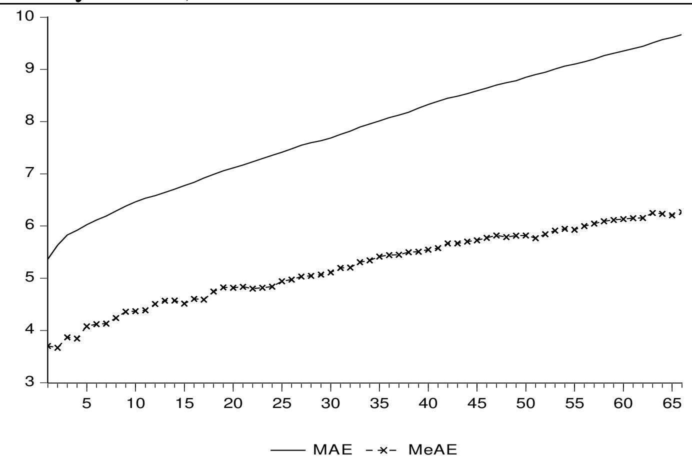


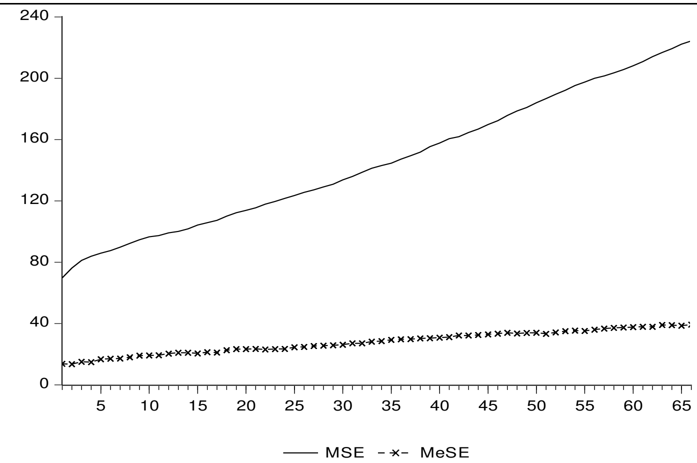
Squared Deviations



Note: These figures correspond to the HAR-RV-COMBINED model.

Figure 4: Mean and median values of the squared difference and the absolute difference between the oil realized volatility forecast $\left(RV^{(\tau)}_{(HL),OIL,t+s|t}\right)$ and the realized volatility $\left(RV^{(\tau)}_{(HL),OIL,t+s}\right)$ across the forecasting horizons (1-day to 66-days ahead).





Note: The values presented in this figure correspond to the HAR-RV-COMBINED model.