International Review of Economics and Finance 71 (2021) 811–829

Contents lists available at ScienceDirect

International Review of Economics and Finance

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

Modeling realized volatility of the EUR/USD exchange rate: Does implied volatility really matter?

Tomáš Plíhal [a,*], Štefan Lyócsa [a,b]

[a] *Institute of Financial Complex Systems, Masaryk University, Brno, Czech Republic*
[b] *Faculty of Management, University of Presov, Presov, Slovakia*

ARTICLE INFO

JEL classification:
C5
C53
G17

Keywords:
High-frequency data
Implied volatility
Realized volatility
Forecasting
Options

ABSTRACT

We model future EUR/USD exchange rate realized volatility (RV) within a class of heterogeneous autoregressive (HAR) models augmented by implied volatilities (IVs). The existing literature has almost unanimously employed IVs from options with one-month maturities; however, our in-sample analysis shows that using IVs from options with a shorter maturity (of one day and one week) might be more relevant when explaining the volatility of the next day and week. In general, IVs are more useful in predicting future RV than past RVs (daily, weekly and monthly averages). At the same time, RVs seem to contain only small incremental predictive power compared to IVs. The out-of-sample results strengthen our in-sample results, as they show the increased predictive power of the models with implied volatility up to 17.3% for one-day-ahead, 42.1% for one-week-ahead, and 22.8% for one-month-ahead forecasts. Additionally, the superior set of models contains only volatility model specifications with IVs. Our results hold not only for individual forecast models but also for combinations of volatility forecasts. We show that increased forecasting accuracy is stable across time and that it is achieved during periods of high market volatility. Our study also provides new evidence that implied volatility from short-lived options as a serious contender for modeling realized volatility.

1. Introduction

For all assets where future returns are uncertain, return volatility plays a central role. Understanding and forecasting volatility helps us i) to understand how prices behave, ii) to allocate assets more efficiently, iii) to better manage financial risks, and iv) to value financial derivatives. It is therefore not surprising that since the time variation in volatility has been broadly accepted as a fact, estimation, modeling, and forecasting volatility have received considerable attention in the literature (e.g. Andersen & Bollerslev, 1998; Bollerslev, 1986; Corsi, 2009; Engle, 1982).

Volatility estimation concentrates around historical realized volatility and forward-looking implied volatility. Several studies have investigated the predictive power of realized and implied volatility with respect to future price variation, but the results are mixed, particularly about the predictive power of implied volatility. First, Canina and Figlewski (1993), Becker et al. (2006, 2007), and Becker and Clements (2008) concluded that in stock markets, implied volatility contains very little incremental information beyond historical realized volatility. Second, Day and Lewis (1992), Christensen and Prabhala (1998), Blair et al. (2001), Busch et al. (2011), Han and Park (2013), Kambouroudis et al. (2016), Wang and Wang (2016), and Kourtis et al. (2016) concluded that implied volatility contains

* Corresponding author.
E-mail address: tomas.plihal@econ.muni.cz (T. Plíhal).

https://doi.org/10.1016/j.iref.2020.10.001
Received 5 April 2019; Received in revised form 20 July 2020; Accepted 1 October 2020
Available online 7 October 2020

additional information on future levels of stock market volatility beyond what is captured by historical volatility. Specifically, Kourtis et al. (2016) found that implied volatility from options that mature in the next 30 days tends to be useful in predicting month-ahead realized volatility on multiple stock markets around the world but is less helpful with a shorter forecast horizon. Seo and Kim (2015) also found that the forecasting power of volatility models based on implied volatility varies with the investor sentiment. Third, Christensen and Prabhala (1998) and Blair et al. (2001) argued that implied volatility is not only useful for forecasting purposes but that all important information is already included in implied volatility, and in fact, realized volatility does not contain any useful information for forecasting volatility. These are three different views on the role of implied volatility. The literature is scarce for other asset classes. For example, the bond market was studied in Amin and Morton (1994) and Busch et al. (2011). Amin and Morton (1994) found that implied volatility has some merits when pricing and trading bond derivatives. Based on the earlier works on model-free implied volatility (e.g., Bakshi et al. (2003); Jiang and Tian (2005)), Chatrath et al. (2015) showed that compared to historical realized volatility, model-free implied volatility (from options with 30 days and longer maturity) and risk-neutral moments have superior forecasting performance with respect to future realized volatility.

The foreign exchange market has also been of interest in this strand of the literature. Concerning volatility modeling, Jorion (1995) and Xinzhong and Taylor (1995) found that implied volatility is superior to realized volatility. However, within a class of autoregressive fractionally integrated moving average models, Pong et al. (2004) claimed the exact opposite and argued that for up to 10-day forecasts, using historical volatility led to more accurate forecasts because of the existence of high-frequency (high-precision) volatility estimators, which are not available for implied volatility. For longer forecast horizons, implied volatility was still not better than historical volatility. Covrig and Low (2003) and Busch et al. (2011) also found that implied volatility contains incremental information about the realized volatility on the foreign exchange market. However, Covrig and Low (2003) used only a larger forecast horizon starting from 1 to 6 months and found that for one-month-ahead forecasts, implied volatility shows very accurate forecasts (superior to realized volatility), but an increasing forecast horizon leads to negligible differences between the use of the two volatility measures.

The standard in the existing volatility modeling literature is to rely on implied volatilities derived from options that mature in more than 30 days (e.g. Becker & Clements, 2008; Becker et al., 2007, 2006; Blair et al., 2001; Busch et al., 2011; Christensen & Prabhala, 1998; Han & Park, 2013; Kambouroudis et al., 2016; Wang & Wang, 2016). The role of implied volatility from options with a shorter maturity has not been investigated. This is surprising, as the very existence of a market for options with short-term maturity (less than 30 days) means that there is a demand for insurance against undesirable price movements in the near future. On the foreign exchange market, such price movements might impend from scheduled macroeconomic news announcements, which is a heavily investigated topic in the literature (e.g. Andersen & Bollerslev, 1998; Andersen, Bollerslev, Diebold, et al., 2007, 2003; Fatum et al., 2012; Faust et al., 2007; Gau & Wu, 2017; Kocenda & Moravcova, 2016). Overall, the existing evidence suggests that implied volatility increases several days before the scheduled news announcement and decreases after the news release and thus carries a volatility risk premium. Short-lived maturity options therefore provide an exceptional opportunity to capture anticipated price shifts in the near future that would otherwise be blurred by the noise from options with a longer maturity. It is thus surprising that the existing literature on volatility modeling is missing the implied volatility from short-lived options.

Our study fills this gap in the literature as it is the first that aims to assess the role of short-lived options in predicting EUR/USD currency market realized volatility within a common framework of the popular heterogeneous autoregressive volatility models of Corsi (2009).

We contribute to the literature in three ways. First, in the current literature, most authors employ implied volatilities from option prices with maturity usually set to one month. This approach is very convenient as volatility indices are readily available (e.g., the widely used VIX index of the CBOE). Some empirical studies that have in some way used implied volatility with a shorter maturity are Day and Lewis (1992), Canina and Figlewski (1993), Jorion (1995), and Xinzhong and Taylor (1995). However, these studies are currently outdated and do not use high-frequency data. Moreover, there seems to be a mismatch in the literature between the volatility forecast horizon and the days to maturity used to calculate implied volatility. In our study, we focus on the foreign exchange market and on the very liquid EUR/USD exchange rate. Therefore, our first contribution is that instead of using implied volatilities with only one-month maturity, we explore implied volatilities from options with three maturities, namely, one-day, one-week and one-month maturities, to model and forecast the day-, week- and month-ahead volatility on the EUR/USD currency market. Compared to existing studies that employ only a one-month long maturity, our approach is more general with respect to the potential predictive power of implied volatilities.

Second, the existing literature seems to suggest that the role of implied and historical volatility differs concerning the forecast horizon. However, the recent evidence is based mostly on at least one-month-ahead forecasts. Earlier research (e.g. Xinzhong & Taylor, 1995) studied shorter forecast horizons but did not rely on high-frequency data. We fill this gap in the literature, as our setup allows us to study how the predictive power of the implied volatility of different maturities differs with respect to forecast horizons.

Third, we employ several volatility model specifications and combination forecasts to address the possibility that the findings from previous studies might have been subject to volatility model choice uncertainty. More specifically, we explore whether our results are driven by one specific volatility model. We compare individual forecasting models that i) separate volatility into positive and negative semi-volatilities, ii) separate the continuous volatility component and the jump component, and iii) weight volatility based on the estimated variance of the measurement error. Next, as it is well known that combination forecasts in the volatility literature tend to improve forecasts from individual models, we also address the potential issue that through model combination, all potential benefits of using implied volatility diminish. Within the combination forecasts, we compare the combined forecasting accuracy of the models that ignore information from implied volatility with the models that include information from implied volatility.

By using an augmented heterogeneous autoregressive model (HAR), we provide new empirical evidence that implied volatilities (IVs) from options with a one-day and one-week maturity might be more relevant when explaining the volatility of the next one day and one week on the EUR/USD foreign exchange market. Overall, the in-sample and out-of-sample results show that IVs are more useful in predicting future realized volatility (RV) than historical RVs (daily, weekly and monthly averages). Adding RVs to the model that contains IVs led only to small increases in predictive power. At the same time, although we are using different popular specifications of the baseline HAR model with only realized volatilities, significant improvements in predicting and forecasting volatility occur when IVs are also added. We also observe that forecast accuracy improved differently with respect to the forecast horizon. For the in-sample analysis, employing implied volatility leads to larger improvements for shorter forecast horizons and decreases with a twenty-two-day-ahead forecasting horizon. However, the out-of-sample analysis shows the largest improvements for one-week-ahead forecasts. We also decompose forecast improvement to find that most of the improved forecasting performance comes during periods of high market volatility. An analysis of the time-varying cumulative forecast improvements shows that models with IVs tend to constantly outperform the standard HAR model across the time period of our sample. Interestingly, large forecast improvements provided by implied volatility are dated to the second half of 2016 around the Brexit referendum (a weakening of the EUR) and U.S. presidential elections in November (a strengthening of the USD). Finally, we also perform a sensitivity check to determine whether our results hold for simplified HAR model specifications and within a different class of volatility model, specifically, an autoregressive fractionally integrated model with exogenous variables (implied volatilities), where we allow errors to follow the standard GARCH model of Bollerslev (1986). We denote this to be the ARFIX-GARCH model. The model is estimated in a day-ahead setting, and we find that the inclusion of implied volatilities improves forecasting accuracy.

The rest of this paper is organized as follows. Section 2 describes the data used. Section 3 discusses our methodology. In Section 4, we first present the results from the in-sample analysis that show the relative importance of IVs being calculated from options with shorter maturities followed by an out-of-sample study where we compare the overall forecasting accuracy of the models that rely on implied volatilities. In this section, we also perform an ARFIX-GARCH analysis. Section 5 concludes.

2. Data

In our analysis, we focus on one of the most liquid markets in the world, the EUR/USD exchange rate market. We study twelve years from 2006 to 01-01 to 2017-12-31. Our dataset consists of two parts. The first part contains the high-frequency data required for the calculation of realized volatility, and the second part comprises data related to implied volatility. The choice of the EUR/USD exchange rate is motivated by the notion that if short-lived options contain useful information in predicting future realized volatility, we should find that the effect on a market with high liquidity, where possible lack of trading activity, larger spreads, and market micro-structure noise might influence our results to a lesser extent. The price data correspond to over-the-counter (OTC) quoted spot prices[1] collected from the DukasCopy.[2] The following sections describe the data in more detail.

2.1. Realized volatility

2.1.1. Trading period

The realized volatility estimation is based on observed historical price movements, while implied volatility is derived from the market prices of financial derivatives. Before high-frequency data were available, daily realized volatility was estimated through daily squared returns, and currently, with the increasing availability of high-frequency data, volatility is estimated by using intra-day returns.

In our analysis, we use the realized volatility of the EUR/USD calculated from high-frequency data. The EUR/USD trade volume changes dramatically with the time of day and day of the week. According to Aloud et al. (2013), this intra-day and intra-week seasonality depends on the opening hours of the relevant markets. Concerning the foreign exchange market where trading is possible over the entire week, we remove weekends, but in contrast to previous studies, we calculate the realized volatility over the whole day, *starting at* 16:00 GMT and *ending at* the 16:00 GMT the *next trading day*, which allows us to capture all relevant price movements, not just the price movements during a chosen 'trading window' that is usually selected based on the most liquid and active trading times. Our motivation is related to the topic of our research in two ways. First, in our dataset, implied volatilities are reported as of 16:00 GMT. Second, implied volatilities should entail expectations over the price variation in the next periods regardless of an arbitrary 'trading window'. For example, if the calculation of realized volatility excludes an arbitrary period in a given day, one could argue that the different information content of implied volatility is due to the exclusion of this specific period. With this difference in mind, weekends are still excluded for the days when no implied volatilities are reported.

2.1.2. Data cleaning and estimation procedures

To estimate the realized volatility, we undergo four steps. First, we use the tick-level data of the quoted bid-ask prices and apply an appropriate cleaning procedure to eliminate possible errors in the dataset. Second, we create an evenly spaced 5-min series of prices, where the price representing each 5-min window is the most recent mid-price from a given 5-min window (i.e., closing prices). Third, we calculate the intra-day continuous returns and subsequently realized volatility (RV). A more detailed description follows.

As our quoted prices are the tick-level data, we follow the standard in the literature, and in the *first step*, we apply an appropriate

[1] We choose the analysis of spot prices as they drive prices on derivative markets as well. Our choice for spot prices is also motivated by the literature (Andersen, Bollerslev, & Diebold, 2007; Busch et al., 2011). Alternatively, our analysis could be based on future prices, which is left for future research.

[2] The data are publicly available at https://www.dukascopy.com/swiss/english/marketwatch/historical/.

cleaning procedure to eliminate possible data errors. To filter the high-frequency data, we use the cleaning procedure suggested by Barndorff-Nielsen et al. (2009). This procedure consists of applying the following rules.

1. Delete entries with zero quotes.
2. Delete entries with a negative spread.
3. Replace multiple quotes that have the same time stamp and a single entry with the median price.
4. Delete entries for which the spread is more than 50 times the median spread on that day.
5. Delete entries for which the mid-quote deviates by more than ten mean absolute deviations from a centered mean (excluding the observation under consideration) of 25 observations before and 25 observations after.

In the *second step*, we choose a 5-min calendar sampling frequency and calculate the mid-price from the latest bid-ask prices in the given 5-min sampling window. The 5-min sampling frequency should be sufficiently high to produce a volatility estimate with negligible sampling variation but sufficiently low to mitigate micro-structure noise. Moreover, the 5-min sampling frequency is usually recommended as a rule of thumb for very liquid equity and foreign exchange markets. Liu et al. (2015) considered over 400 different estimators of realized volatility and used 11 years of data on different asset classes. The authors concluded that significantly outperforming the 5-min RV sampling frequency is difficult in terms of estimation accuracy.

After applying the data cleaning procedure and sampling, we calculate the daily realized volatility (RV) in the *third step*. More specifically,

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2 \tag{1}$$

$$r_{t,i} = \ln c_{t,i} - \ln c_{t,i-1} \tag{2}$$

where $r_{t,i}$ represents the intra-day return at day t in the i^{th} sampling window. The returns for each sampling window are calculated as the log-return from close ($c_{t,i}$) to close ($c_{t,i-1}$) mid-prices.

2.2. Implied volatility

Derivatives are priced on the market based on the supply and demand of the market participants, who given an asset pricing model, enter contracts based on their beliefs about future volatility. Therefore, by observing option prices and assuming a particular asset pricing model, we can calculate their expected volatility, namely, the implied volatility. Implied volatility therefore represents the market expectations about future volatility over the remaining life of the given option. If the markets are efficient, implied volatility should contain all available information and expectations about future price movements. Consequently, implied volatility should be an efficient forecast of future levels of price variation. However, in practice, implied volatility may provide biased expectations, as underlying option pricing models (and assumptions) might differ or volatility risk might be priced or because at the daily data frequency, implied volatilities are only a snapshot at a particular time of the day. Accordingly, implied volatilities derived from options with shorter maturities might be noisier and are rarely used in the empirical literature.[3]

The trading of options on the OTC foreign exchange market is realized by quoting implied volatility directly. If the parties agree on the implied volatility, the remaining contract terms are specified. In practice, every volatility quote set by an options trader is converted into monetary units by using an option-pricing model, which is usually the Garman-Kohlhagen model (Garman & Kohlhagen, 1983). This is possible as volatility is the only unobserved variable, and the remaining variables are given by option characteristics (e.g., maturity), are publicly known (e.g., interest rates) or are to be agreed on in the contract (size of the position). We use the data on implied volatility from Bloomberg, which aggregates data on a daily basis from the brokers and dealers of large banks and insurance companies.[4] This contribution of the Bloomberg database reduces the idiosyncratic effect that is specific to individual market participants who provide quotes. We refer to this approach of estimating implied volatility as the hard data approach.

The EUR/USD currency pair represents approximately 23% of the total foreign exchange market turnover.[5] The EUR/USD market is the most liquid market in the world, which also minimizes the possible data issues with the quality of quotes. We use implied volatility quotes from Bloomberg for various maturities. All of our quotes are from at-the-money options with the same delta. The reason for using at-the-money options to calculate implied volatility (IV) is based on theoretical foundations (e.g. Hull & White, 1987; Jorion, 1995). At-the-money options should produce the smallest bias in the case of non-constant volatility. They are also considered to be the least affected by the non-normal distribution of returns and non-simultaneity problems. The earlier studies of Beckers (1981) and Canina and Figlewski (1993) suggest that the IVs from near the money options tend to predict volatility better than the IVs from deep out/in money options.

Compared to the existing studies, this research differs in that our volatility models explore the role of IVs for options with a one day, one week, and one month time to maturity. We selected these maturities to match the forecast horizon used in the volatility model of

[3] For a discussion concerning equity options, see Baruník and Hlínková (2016).
[4] The number of institutions varies but is usually around 80, e.g., GFI Group, HSBC, Raiffeisen, Societe Generale, InvestAZ, etc.
[5] Bank of International Settlements report, available at www.bis.org/publ/rpfx16fx.pdf.

Corsi (2009). Even when modeling one- or five-day-ahead volatility, the existing studies rely on volatility indices or on other implied volatilities that are derived from options with a maturity of around one month or 30 days. We argue that if the day-ahead volatility forecast is of interest, one- or five-day maturity option volatilities should contain expectations that are more relevant for the upcoming days.

3. Methodology

In this section, we present our methodology. First, we describe the volatility models that are based on the HAR model by Corsi (2009). The HAR model is widely used for volatility modeling and forecasting as it well describes high volatility persistence and a long memory of volatility while providing very good forecasting performance. As discussed in Corsi (2009), the standard GARCH model of Bollerslev (1986) and stochastic volatility models have difficulties in explaining these volatility features. The autoregressive fractionally integrated realized volatility (ARFI-RV) model is designed to capture the long-memory property, but fractional integration lacks economic interpretation. The estimation of these models is also much more complicated. In contrast, compared to the simple autoregressive (AR) model, the HAR model can be understood as an extension of an AR(p) process with less parameters, which might prove to be useful for forecasting purposes. Recently, Kourtis et al. (2016) and Horpestad et al. (2019) compared the performance of the GARCH and HAR class models to find that the HAR model performed much better. Furthermore, HAR models are easily estimated, and model specifications can be straightforwardly modified to account for various volatility features. HAR models are therefore well suited for our purpose of exploring how implied volatilities, which are calculated for various expiration periods, influence future volatility. Our specific models are described in more detail in the following subsection. Second, we specify the design of the out-of-sample study. Third, we describe the unconditional forecast combinations. Finally, we outline the evaluation of volatility forecasts.

3.1. Volatility models

We consider several alternative HAR model specifications. The idea is to rule out that in- and out-of-sample forecast improvements are random artifacts of the baseline HAR model. In all of these specifications, we report the results when only the lagged realized volatilities are decomposed. In this way, we follow the work of Sevi (2014) and Patton and Sheppard (2015).[6]

3.1.1. HAR and IV-HAR

The first reference model (benchmark) is the original HAR model developed by Corsi (2009). The future levels of volatility are predicted through three historical volatility components that should capture short- and long-term volatility behavior or the behavior of heterogeneous market agents that have different investment horizons (e.g. Muller et al., 1997). This model is denoted simply as HAR:

$$RV_{t+1}^{H} = \beta_1 + \beta_2 RV_t^{D} + \beta_3 RV_t^{W} + \beta_4 RV_t^{M} + \varepsilon_t \tag{3}$$

where RV_{t+1}^{H} is the average daily realized volatility over the next H days, i.e., the forecast horizon. In this study, we consider H = 1, 5, and 22, i.e., forecasts of intra-day volatility over the next one, five and twenty-two trading days, respectively. RV_t^D, RV_t^W, and RV_t^M are the averages of realized volatility over the previous one, five and twenty-two trading days, respectively. These volatility components should mimic daily (D), weekly (W), and monthly (M) volatility trends. The idea is that an investor with monthly investment horizons and longer should not be influenced by short-term price movements as much as a day-trader, who might be influenced by not only short- but also long-term price variations.

Our first alternative specification mimics the reference HAR model in that instead of realized volatilities, the right-hand side variables are implied volatilities derived from options with corresponding one- (IV_t^D), five- (IV_t^W), and twenty-two- (IV_t^M) day maturities. We refer to this model as full implied volatility HAR (FIV-HAR):

$$RV_{t+1}^{H} = \beta_1 + \beta_2 IV_t^{D} + \beta_3 IV_t^{W} + \beta_4 IV_t^{M} + \varepsilon_t \tag{4}$$

Here and in the remaining implied volatility models, we use implied volatilities with the same three maturities, regardless of the forecasting horizon H. We expected that for day-ahead volatility forecasts, short-term implied volatilities should be most informative. Therefore, an alternative is to use specifications with only one implied volatility component that matches the forecasting horizon H.[7] However, the current specification with all three implied volatility components allows us to compare the relative merits of each implied volatility component.

Our second alternative specification allows us to inspect the incremental information content of the implied volatility, as the plain HAR model is augmented by implied volatility variables and is referred to as the IV-HAR model:

$$RV_{t+1}^{H} = \beta_1 + \beta_2 RV_t^{D} + \beta_3 RV_t^{W} + \beta_4 RV_t^{M} +$$

[6] We also consider a case where weekly and monthly realized volatilities were decomposed for weekly and monthly volatility forecasting models. These results led to qualitatively similar results as the findings reported in the main text of this paper.

[7] The in- and out-of-sample results from these specifications are available upon request. These results support our conclusions drawn from our specifications, while implied volatility models with only one implied volatility component tend to lead to similar or less accurate out-of-sample predictions.

$$\beta_5 IV_t^D + \beta_6 IV_t^W + \beta_7 IV_t^M + \varepsilon_t \quad (5)$$

Thus, we can directly observe whether implied volatilities with different maturities up to one month offer relevant predictive information beyond what is incorporated in the corresponding realized volatility. To rule out that the possible fit and forecast improvements are just artifacts of the baseline HAR model, we also augment several popular HAR model specifications.[8] In the following subsection, we describe these models in more detail.

3.1.2. CJ-HAR and IV-CJ-HAR

In the presence of discontinuous price changes, the realized volatility is a non-consistent estimator of price variation. Moreover, price discontinuities (jumps) might distort the autoregressive structure of the underlying continuous volatility component. Andersen, Bollerslev, and Diebold (2007) argues that jump and continuous volatility components behave differently and might have a different impact on the future levels of volatility. Regardless of the underlying asset, the existing empirical studies seem to confirm these observations (e.g. Lyócsa et al., 2017; Lyócsa & Molnár, 2016; Sevi, 2014). We follow the work of Andersen et al. (2012) and decompose the realized volatility on the continuous (CC) and jump (JC) components through the estimation of the median realized volatility (MRV_t):

$$JC_t = max\{0, (RV_t - MRV_t) I[JT_t > \varphi_{1-\alpha}]\} \quad (6)$$

$$CC_t = MRV_t I[JT_t > \varphi_{1-\alpha}] + RVI[JT_t \leq \varphi_{1-\alpha}] \quad (7)$$

where MRV_t is the median realized volatility, namely, the jump-robust estimator of integrated volatility, and $I[.]$ is an indicator function that returns 1 if the test statistic JT_t (related to the jump component) exceeds the 5% critical value $\varphi_{1-\alpha}$ (see Eq. (6) in Andersen et al. (2012) for further details). The MRQ_t represents the median realized quarticity. More specifically,

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

$$JT_t = \sqrt{f} \frac{(RV_t - MRV_r)RV_t^{-1}}{\left(0.96 max\left\{1, MRQ_t / RV_t^2\right\}\right)^{1/2}} \quad (9)$$

$$MRQ_{t,f} = \frac{3\pi f}{9\pi + 72 - 52\sqrt{3}} \left(\frac{f}{f-2}\right) \sum_{j=2}^{f-1} \left[med\left|r_{t,j-1}\right|, \left|r_{t,j}\right|, \left|r_{t,j+1}\right|\right]^4 \quad (10)$$

The increased number of parameters does not guarantee improved out-of-sample forecast accuracy. Therefore, in our analysis, we use a simplified version of the HAR model with continuous and jump components in that only the previous day's volatility is decomposed into its continuous and potential jump components.[9] The CJ-HAR model is as follows:

$$RV_{t+1}^H = \beta_1 + \beta_2 CC_t^D + \beta_3 JC_t^D + \beta_4 RV_t^W + \beta_5 RV_t^M + \varepsilon_t \quad (11)$$

where CC^D and JC^D represent the continuous component and jump component of realized volatility, respectively, for the previous day.

We are interested not only in comparing the forecasting accuracy of the specification defined above with the plain HAR model but also in how well the following augmented specification IV-CJ-HAR will compare with the CJ-HAR model:

$$RV_{t+1}^H = \beta_1 + \beta_2 CC_t^D + \beta_3 JC_t^D + \beta_4 RV_t^W + \beta_5 RV_t^M +$$

$$\beta_6 IV_t^D + \beta_7 IV_t^W + \beta_8 IV_t^M + \varepsilon_t \quad (12)$$

Variables IV^D, IV^W, and IV^M stand for the implied volatility derived from options with one-day, one-week and one-month maturity, respectively.

3.1.3. Q-HAR and IV-Q-HAR

Most of the empirical studies on volatility modeling silently assume that the measurement error of the realized volatility estimator is constant over time. However, Bollerslev et al. (2016) argues that this is not likely and shows how ignoring measurement error leads to an attenuation bias, i.e., the latent integrated volatility is more persistent than the observed realized volatility. Bollerslev et al. (2016) propose weighting each volatility with the variance of the measurement error. During days when the variance of measurement error is low, the realized volatility should be more indicative of future volatility than during days with a higher variance of measurement error. In this way, the time variation in the measurement error is considered. As the measurement error is estimated through the MRQ_t defined before, we can directly specify the Q-HAR model as follows:

[8] For recent applications of these HAR model extensions, see Çelik and Ergin (2014) and Ma et al. (2017).

[9] A similar approach is not uncommon in the volatility literature; see Patton and Sheppard (2015) or Bollerslev et al. (2016).

$$RV_{t+1}^{H} = \beta_1 + \left(\beta_2 + \beta_{2Q}\sqrt{MRQ_t^D}\right)RV_t^D + \beta_3 RV_t^W + \beta_4 RV_t^M + \varepsilon_t \tag{13}$$

As before, the alternative specification is the one augmented with implied volatilities and denoted as IV-Q-HAR as follows:

$$RV_{t+1}^{H} = \beta_1 + \left(\beta_2 + \beta_{2Q}\sqrt{MRQ_t^D}\right)RV_t^D + \beta_3 RV_t^W + \beta_4 RV_t^M +$$

$$\beta_5 IV_t^D + \beta_6 IV_t^W + \beta_7 IV_t^M + \varepsilon_t \tag{14}$$

3.1.4. SV-HAR and IV-SV-HAR

The asymmetric volatility effect is a well-documented feature of financial markets and states that during periods of market downturn, volatility tends to be larger than during periods when the market thrives. With positive and negative semi-volatilities, it is straightforward to incorporate this idea into volatility modeling through the HAR models (e.g., Barndorff-Neilsen et al., 2010; Patton & Sheppard, 2015; Sevi, 2014). Positive (RVP_t) and negative (RVN_t) semi-volatilities are calculated as squared sums of positive and negative intra-day returns ($r_{t,i}$) as follows:

$$\text{RVP}_t = \sum_{i=2} r_{t,i}^2 I[r_{t,i} > 0] \tag{15}$$

$$\text{RVN}_t = \sum_{i=2} r_{t,i}^2 I[r_{t,i} < 0] \tag{16}$$

where I represents the indicator function that returns one if the condition in square brackets holds and is zero otherwise.

If asymmetric volatility holds, it might be that negative semi-volatilities are more relevant for forecasting the next period's volatility than positive semi-volatilities, and they therefore deserve their persistence parameter. Consequently, the SV-HAR model is defined as follows:

$$RV_{t+1}^{H} = \beta_1 + \beta_2 RVP_t^D + \beta_3 RVN_t^D + \beta_4 RV_t^W + \beta_5 RV_t^M + \varepsilon_t \tag{17}$$

Regressors RVP_t^D and RVN_t^D denote realized positive and negative semi-volatilities, respectively, for the previous day and replace RV_t^D in the model specification. The augmented version IV-SV-HAR contains implied volatilities with maturities of one day, one week and one month (IV^D, IV^W, IV^M) and is specified as follows:

$$RV_{t+1}^{H} = \beta_1 + \beta_2 RVP_t^D + \beta_3 RVN_t^D + \beta_4 RV_t^W + \beta_5 RV_t^M +$$

$$\beta_6 IV_t^D + \beta_7 IV_t^W + \beta_8 IV_t^M + \varepsilon_t \tag{18}$$

We decompose only the one-day lagged realized volatility (RV^D) into positive and negative semivariances, following the studies of Sevi (2014) and Patton and Sheppard (2015). This idea is similar to the case of the HAR-CJ model.

3.2. Forecasting procedure

Volatility tends to cluster into periods of excess volatility. Although volatility shocks are difficult to forecast, they are of high interest to market participants as during such times, accurate volatility forecasts are most important. In the context of volatility models, sudden spikes in volatility often lead to the heteroskedasticity of forecast errors, and the estimated autoregressive parameters are less reliable. Patton and Sheppard (2015) therefore recommend estimating HAR models through weighted least squares (WLS), where weights are inverse fitted values from the initial ordinary least squares (OLS) estimation. Thus, extreme volatility receives a lower weight. All in- and out-of-sample regressions were estimated through the WLS. The out-of-sample study is performed in an expanding window framework where the initial estimation window is set to $1,000$ observations. After forecasting one-, five-, and twenty-two-day-ahead forecasts, additional observation is added, and all models are re-estimated, after which the next forecast is calculated. The first forecast used for evaluation purposes is the $1,001$ forecast. As a result, we obtained $2,083$ volatility forecasts.

3.3. Combination forecasts

Forecast combination is a popular approach to improve the forecasting performance from individual models. The idea was introduced by Bates and Granger (1969) for airline passenger data. Since then, this idea has been investigated by many authors, for example, Clemen (1989), Timmermann (2006) and Newbold and Harvey (2007). In general, different model specifications could comprise different information from the underlying data, and forecast combination could lead to better performance, especially when individual forecasts are unbiased and not highly correlated.

The idea of using combination forecasts in this study is motivated by the fact that when benefits from combination forecasts are utilized, adding implied volatility might not be useful at all. For example, we might combine all forecasts that exclude information from implied volatility. Such a forecast might not be worse than a combination forecast from implied volatility models (or individual implied volatility models). If this is true, there is little benefit from including implied volatility into forecasting realized volatility.

A number of different approaches to combining individual forecasts exist, e.g., Genre et al. (2013) compare many combination methods with the simplest method of a simple average. The authors conclude that a simple average is not the best combination method all the time, but it is very difficult to pick a single preferred method. We therefore use the simple average of individual forecasting models as an unconditional approach to combining forecasts. Regarding volatility forecasting, such a naïve approach was challenged by Wang and Wang (2016) in favor of Bayesian model averaging and dynamic model averaging by using a sample of S&P 500 index returns. In contrast, in Santos and Ziegelmann (2014), Lyócsa and Molnár (2016), and Lyócsa et al. (2017), simple averaging was not worse than individual forecasting models. We follow the latter approach and take the simple average of individual forecasts across two groups of models as follows:

- Across all individual volatility forecast models that do not include information from implied volatility.
- Across all individual volatility forecast models that include information from implied volatility.

3.4. Forecast evaluation

First, we compare the average forecast accuracy by using two loss functions. Next, a statistical test compares a set of competing volatility forecasting models with a given confidence level. This iterative testing procedure leads to set of superior forecasting models.

3.4.1. Loss function

Forecasts are evaluated with loss functions that provide a consistent ranking of forecasts, even if the proxy of the underlying latent volatility is measured with noise (Patton, 2011), the MSE and QLIKE, which are defined as follows:

$$MSE_i = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{RV}_{i,t} - RV_t\right)^2 \tag{19}$$

$$QLIKE_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{RV_t}{\widehat{RV}_{i,t}} - \ln\frac{RV_t}{\widehat{RV}_{i,t}} - 1\right) \tag{20}$$

where $\widehat{RV}_{i,t}$ is the forecasted realized variance at time t of model i, and RV_t is the proxy of the true volatility, specifically, the 5 min realized variance.

The choice between the two loss functions depends on the application at hand. Although MSE weights volatility over- and under-estimation equally and extreme forecast errors are increasingly penalized, the QLIKE weights forecast under-estimation more severely than they forecast over-estimation and are less sensitive to extremes.

3.4.2. Model confidence set

We follow the sequential testing procedure of Hansen et al. (2011) that leads to the set of superior forecasting models. This procedure is a top-down approach in which at each step, a null hypothesis of the equal predictive ability of competing models is tested. If the hypothesis is rejected, the model with the worst relative loss is excluded, and the testing procedure is repeated until the null of equal predictive ability cannot be rejected at given confidence $1-\alpha$ (we set $\alpha = 0.05$). Let the loss function for the i^{th} model be defined as follows:

$$L(i,t) = L\left(\widehat{RV}_{i,t}, RV_t\right) \tag{21}$$

When two competing models i and j are compared, a loss differential $D(i,j,t)$ is evaluated as follows:

$$D(i,j,t) = L(i,t) - L(j,t) \tag{22}$$

Let m denote the number of competing models and $D(i,.,t)$ denote the average loss of model i at time t with respect to the remaining models as follows:

$$D(i,.,t) = \frac{1}{m-1}\sum_{j,j\neq i} D(i,j,t) \tag{23}$$

3.5. The hypothesis of interest is defined as

$$H_0 : E[D(i,.)] = 0, \text{ for all}, \ i = 1,\dots,m \tag{24}$$

$$H_1 : E[D(i,.)] \neq 0, \text{ for some}, \ i = 1,\dots,m \tag{25}$$

Given that $\overline{d}_{i,.}$ is an average of the average loss differential of the i^{th} model with respect to the other competing models, the statistic is defined as follows:

$$t_{i,.} = \frac{\overline{d}_{i,.}}{\sqrt{\widehat{var}(\overline{d}_{i,.})}} \quad (26)$$

The denominator is a block bootstrap estimator of the variance of $\overline{d}_{i,.}$. Finally, the test statistic is defined as follows:

$$T_{max} = max_{(i)} t_{i,.} \quad (27)$$

The distribution under the null hypothesis of the test statistics T_{max} is bootstrapped as was the variance before. If the null hypothesis is rejected, then the model with the highest $t_{i,.}$ is eliminated, and the testing procedure is repeated until the null is not rejected.

4. Results

This section describes our results from the in-sample and out-of-sample analysis. First, we perform preliminary data analysis, and then, we move to the model estimation and forecasting.

4.1. Preliminary data analysis

The descriptive statistics of our volatility measures presented in Table 1 show that the annualized average volatility was 95.705, which corresponds to 9.783% when square root is taken and is a similar number to 9.10% found in Bollerslev et al. (2018, Table B1). The results also seem to confirm some stylized facts known in the literature. For example, the distribution of the volatility is skewed to the right and shows long memory properties as even at lag 100, the autocorrelation is a considerable 0.221.[10]

The jump and continuous volatility components show very distinct statistical properties. The jumps are considerably less persistent, while the continuous volatility components have persistence above the realized volatility. The decomposition of the realized volatility into its continuous and jump components appears to be helpful. Andersen, Bollerslev, and Diebold (2007) argued that because the two components are so different, they may contain distinct predictive power of future volatility. Similarly, negative semi-volatilities have higher persistence than positive semi-volatilities, which also suggests that the two components might have different roles in forecasting future volatility.

Finally, we turn our attention to implied volatilities. The annualized implied volatility calculated from the options that expire the next day is much higher than the corresponding RV. The differences between the two volatility measures can be attributed to the scaling.[11] However, as we use IV_t as the independent variables, the different scaling is of little importance to our empirical findings. The persistence of the two volatility measures is also different. At the first lag, the results are similar, but IV^D has an even higher dependence at the 10^{th} lag than at the first lag. IV decays much slower than RV. These results suggest that implied volatility is distinct from realized volatility and might contain different information.

The time series plots in Fig. 1 show that the daily, weekly and monthly implied volatilities peaked at the end of October 2008, around the period when Lehman Brothers filed for bankruptcy and when the U.S. presidential election occurred in the United States. RV, MRQ, RVP, and RVN reached maximal values on 18 March 2015, the day in which the Federal Open Market Committee (FOMC) meeting signaled possible rate hikes.

Table 1
Descriptive statistics of volatility measures.

	Mean	$\sqrt{\text{Mean}}$	St.Dev.	Skew.	Kurt.	Min.	Median	Max.	$\rho(1)$	$\rho(10)$	$\rho(100)$	$\rho(250)$
RV	95.705	9.783	105.620	5.032	43.995	0.011	67.536	1671.232	0.610	0.472	0.221	0.014
MRQ	56.779	7.535	509.870	29.796	1052.850	0.000	8.083	20858.825	0.035	0.030	0.022	−0.007
JC	12.883	3.589	24.356	6.828	69.336	0.000	6.358	398.680	0.071	0.094	0.054	−0.014
CC	82.822	9.101	93.150	5.129	45.648	0.011	56.562	1536.965	0.657	0.512	0.231	0.019
RVP	48.396	6.957	58.038	5.893	59.402	0.005	32.859	1014.250	0.520	0.376	0.186	0.013
RVN	47.308	6.878	53.733	5.000	45.184	0.006	31.885	889.019	0.582	0.475	0.216	0.009
IV^D	134.632	11.603	150.174	4.301	30.063	5.954	92.930	2172.865	0.662	0.715	0.332	0.092
IV^W	112.772	10.619	104.253	3.533	18.872	13.988	84.318	1127.952	0.961	0.807	0.311	0.015
IV^M	109.930	10.485	90.059	3.014	12.697	17.347	86.211	834.343	0.987	0.910	0.380	0.071

Note: RV = Realized Volatility; MRQ = Median Realized Quarticity; JC = Jump Component of RV; CC = Continuous Component of RV; RVP = Realized Semi-Volatility Positive; RVN = Realized Semi-Volatility Negative; and IV^D, IV^W, IV^M = Implied volatilities from options with one-day, one-week, and one-month maturity, respectively. $\rho(.)$ is the value of the auto-correlation coefficient at the given lag by using the Lin and McLeod (2006) test. All auto-correlation coefficients are significant at least at the 0.05 significance level.

[10] Again, this result is similar to the finding in Bollerslev et al. (2018) of 0.22.

[11] For example, VIX is calculated over 30 days (including weekends) and then scaled to annual values for 12 months.

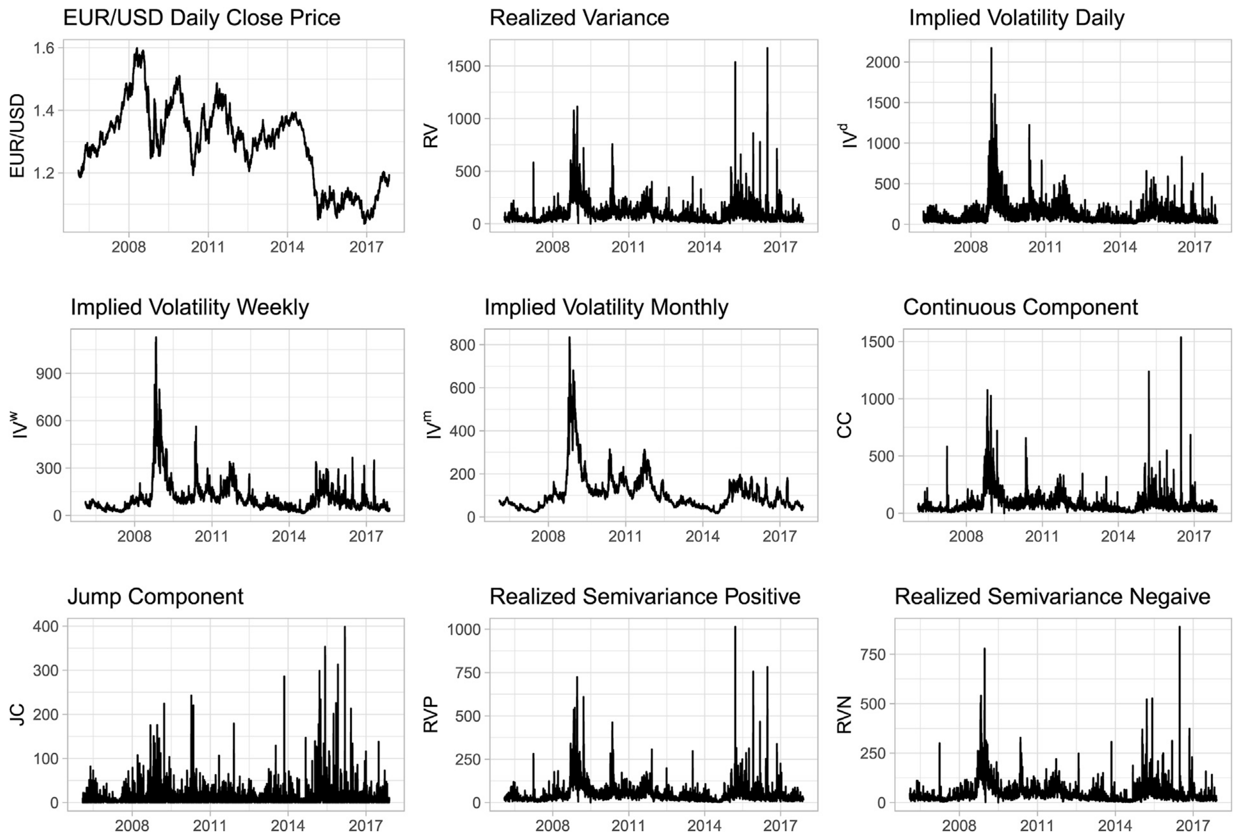

Fig. 1. Time series plots of volatility measures.

4.2. *In-sample results*

The in-sample results of our models are summarized in Tables 2–4. The coefficients in bold are statistically significant at $\alpha = 5\%$. Under each coefficient, we report the standard errors in brackets, and the model fit is reported at the bottom of each table.

When we examine the statistical significance of individual regressors, we notice that at least one of three coefficients loaded on implied volatility becomes significant in all models. This result is robust across different predictive horizons. The demand and supply factors that drive pricing might be noisy from a short-term perspective; therefore, the IVs from options with a shorter maturity also might be expected to be too noisy to contain valuable information concerning future realized volatility. Moreover, the IVs should represent the overall expectations of the market participants about the future volatility of the spot market, but it is not guaranteed as arbitrage trading might either not be that profitable or be too risky. However, our results show that even one-day and one-week-ahead market expectations are actually useful with respect to future realized volatility. More specifically, for the one-day and one-week-ahead regressions, the most important regressor appears to be daily and weekly implied volatilities (IV^D, IV^W), and for one-month-ahead forecasts, all implied volatilities are of statistical importance. Implied volatility therefore seems to have merit in predictive volatility regressions. Interestingly, for shorter forecast horizons, the monthly implied volatility, which is the implied volatility used in most existing empirical studies, seems to matter the least.

Comparing individual models with their non-implied volatility versions shows that the inclusion of implied volatility tends to decrease the size of the coefficients that correspond to realized volatilities. For example, in Table 2, the coefficient of the lagged daily realized volatility of 0.253 in the plain RV-HAR model decreases to 0.131 for the IV-HAR model. This behavior is also observed for other model specifications and predictive forecast horizons. However, this effect also seems stronger with longer forecast horizons. Especially in the case of one-week and one-month-ahead volatility predictions, the inclusion of implied volatility causes several regressors of realized volatility to be insignificant.

The results reported in Tables 2–4 show that the IVs seem to almost subsume the information content of RVs, and compared to the RV-HAR model, the FIV-HAR model leads to a higher mode fit. At the same time, including past RVs into the FIV-HAR model, which produces the IV-HAR model, increases the model fit only slightly, which suggests that past RVs contain only little incremental information that is useful for predicting future RV. The general pattern that we observe is that the R^2 in the individual models with IV increased (also for the adjusted R^2) considerably by approximately 0.10 for one-day, by 0.15 for one-week and by 0.06 for one-month-ahead volatility predictions. These fit improvements are non-trivial, as given the baseline R^2, they correspond to an increase of the fit of the model by approximately 29%, 27%, and 9% for one-day, one-week and one-month-ahead volatility predictions, respectively. In

Table 2
One day ahead variance predictive regressions.

	RV-HAR	FIV-HAR	IV-HAR	CJ-HAR	IV-CJ-HAR	Q-HAR	IV-Q-HAR	PS-HAR	IV-PS-HAR
Intercept	7.235***	8.313***	6.850***	8.659***	7.596***	6.219***	6.359***	7.226***	6.876***
	(1.736)	(1.620)	(1.380)	(1.693)	(1.369)	(1.668)	(1.344)	(1.731)	(1.374)
RV^D	0.253***		0.131***			0.331***	0.167***		
	(0.045)		(0.031)			(0.051)	(0.032)		
RV^W	0.313***		0.115**	0.265***	0.096*	0.288***	0.108*	0.313***	0.116**
	(0.072)		(0.043)	(0.074)	(0.044)	(0.071)	(0.042)	(0.072)	(0.042)
RV^M	0.358***		− 0.045	0.327***	− 0.043	0.337***	− 0.047	0.360***	− 0.046
	(0.051)		(0.048)	(0.048)	(0.048)	(0.053)	(0.049)	(0.052)	(0.048)
IV^D		0.057*	0.067**		0.067**		0.069**		0.066*
		(0.024)	(0.026)		(0.025)		(0.025)		(0.026)
IV^W		0.735***	0.607***		0.590***		0.595***		0.608***
		(0.068)	(0.071)		(0.071)		(0.069)		(0.070)
IV^M		− 0.029	− 0.071		− 0.078		− 0.074		− 0.068
		(0.058)	(0.047)		(0.047)		(0.046)		(0.047)
CC^D				0.390***	0.203***				
				(0.054)	(0.041)				
JC^D				− 0.152*	− 0.061				
				(0.065)	(0.048)				
$RV^D \times MRQ^D$						− 0.002***	− 0.001*		
						(0.000)	(0.000)		
RVP^D								0.269***	0.178***
								(0.077)	(0.048)
RVN^D								0.233***	0.076*
								(0.063)	(0.038)
R^2	0.361	0.465	0.472	0.373	0.476	0.365	0.473	0.362	0.474
Adj. R^2	0.360	0.464	0.471	0.372	0.475	0.364	0.472	0.361	0.473

Note: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Significances are based on the estimation of the variance-covariance matrix of coefficients by using the automatic bandwidth selection as in Whitney et al., (1994) with a quadratic spectral weighting scheme. We use the implementation procedures provided in Zeileis (2004). The variance inflation factors are all below 10. Model variables: RV^D, RV^W, RV^M = Realized Volatility for the previous day and average for the previous week and month, respectively; IV^D, IV^W, IV^M = Implied Volatility from options with one-day, one-week, and one-month maturity, respectively; CC^D = Continuous Component from the previous day; JC^D = Jump Component for the previous day; MRQ^D = Median Realized Quarticity; RVP^D = Realized Semi-Volatility Positive for the previous day; and RVN^D = Realized Semi-Volatility Negative for the previous day.

contrast to these improvements, a comparison of the RV-HAR with the CJ-HAR, Q-HAR and PS-HAR models (i.e., the models without implied volatility) shows that different model specifications cause only mild improvements to the fit of the model.

Accordingly, the results in Tables 2–4 show not only that implied volatility is superior to the model with only realized volatility but also that realized volatility provides new incremental information for volatility models, as the (parameter-adjusted) fit of the IV-HAR model (which includes RV and IV) is slightly above that of the FIV-HAR model (with only IV information) across all forecast horizons.

4.3. *Out of sample results*

The in-sample results do not necessarily manifest into the same out-of-sample results, which are usually of higher importance for market participants. For one-day, one-week and one-month-ahead volatility forecasts, the results are presented in Table 5. For each model, we report the average values of two loss functions and compare the percentage improvement (negative value) with respect to the RV-HAR model. Panel A includes the results for the individual models, and panel B shows the combinations of the model forecasts. In Table 5, we compare the performance of the forecasts based on the model confidence set of Hansen et al. (2011). In Table 6, we perform a bi-variate comparison of models, i.e., within each class of HAR models, we compare the performance of the models with and without implied volatility.

4.3.1. *Evidence from individual forecasts*

Before we proceed to interpret the key results, we establish several characteristics of our findings. In Fig. 2, we plot the realized volatility and corresponding forecasts from selected forecasting models across the out-of-sample period. The forecasting models evidently do not catch all 'spikes' in volatility, which implies heteroskedastic forecast errors. Such errors are often observed in the volatility literature (e.g. Lyócsa & Molnár, 2016). With respect to the benchmark plain RV-HAR model, almost all individual forecasts fare better. The benchmark model is outperformed more for the one-week-ahead forecasts. In addition, as we expected, the forecast combinations are better than the plain RV-HAR model. Interestingly, a comparison among the different model specifications shows that regardless of the extensions of the standard RV-HAR model, all models lead to similar forecast improvements. In addition, the forecast improvements are somewhat higher for the QLIKE loss function. Compared to the plain RV-HAR model, the forecast improvements

Table 3
One week ahead variance predictive regressions.

	RV-HAR	FIV-HAR	IV-HAR	CJ-HAR	IV-CJ-HAR	Q-HAR	IV-Q-HAR	PS-HAR	IV-PS-HAR
Intercept	9.309***	10.480***	8.790***	10.008***	8.863***	8.392**	8.515***	9.354***	8.794***
	(2.682)	(1.725)	(1.795)	(2.607)	(1.795)	(2.596)	(1.866)	(2.660)	(1.795)
RV^D	0.094***		0.005			0.164***	0.027		
	(0.019)		(0.012)			(0.027)	(0.018)		
RV^W	0.276***		0.069	0.254**	0.067	0.251**	0.064	0.273***	0.068
	(0.080)		(0.050)	(0.079)	(0.050)	(0.078)	(0.047)	(0.079)	(0.049)
RV^M	0.532***		0.098	0.519***	0.098	0.511***	0.095	0.530***	0.098
	(0.088)		(0.072)	(0.087)	(0.072)	(0.085)	(0.066)	(0.088)	(0.072)
IV^D		0.040**	0.040**		0.040**		0.040**		0.040**
		(0.015)	(0.014)		(0.014)		(0.014)		(0.014)
IV^W		0.685***	0.642***		0.641***		0.637***		0.642***
		(0.068)	(0.074)		(0.074)		(0.072)		(0.074)
IV^M		0.024	− 0.067		− 0.067		− 0.069		− 0.068
		(0.062)	(0.053)		(0.053)		(0.052)		(0.053)
CC^D				0.155***	0.011				
				(0.025)	(0.016)				
JC^D				− 0.089*	− 0.011				
				(0.035)	(0.026)				
$RV^D \times MRQ^D$						− 0.001***	− 0.000		
						(0.000)	(0.000)		
RVP^D								0.028	− 0.013
								(0.044)	(0.020)
RVN^D								0.169**	0.028
								(0.058)	(0.030)
R^2	0.560	0.711	0.717	0.562	0.717	0.567	0.718	0.560	0.717
Adj. R^2	0.559	0.711	0.716	0.562	0.716	0.566	0.718	0.560	0.716

Note: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Significances are based on the estimation of the variance-covariance matrix of coefficients by using the automatic bandwidth selection as in Whitney et al., (1994) with a quadratic spectral weighting scheme. We use the implementation procedures provided in Zeileis (2004). The variance inflation factors are all below 10. Model variables: RV^D, RV^W, RV^M = Realized Volatility for the previous day and average for the previous week and month, respectively; IV^D, IV^W, IV^M = Implied Volatility from options with one-day, one-week, and one-month maturity, respectively; CC^D = Continuous Component from the previous day; JC^D = Jump Component for the previous day; MRQ^D = Median Realized Quarticity; RVP^D = Realized Semi-Volatility Positive for the previous day; and RVN^D = Realized Semi-Volatility Negative for the previous day.

(negative values) for the one-day-ahead, one-week-ahead, and one-month-ahead forecasts range from 0.2%, − 0.5%, and 0.2% to − 13.2%, − 38.9%, and − 22.8%, respectively, for the MSE and from 0.3%, 0.0%, and 0.1% to − 17.3%, − 42.1%, and − 20.8%, respectively, for the QLIKE.

Our key finding is that the inclusion of implied volatilities systematically improves the forecast accuracy for both loss functions across all models and forecast horizons. Comparing the loss functions within each class of volatility model shows that the inclusion of implied volatilities always leads to improved forecasts (see Table 6). The improvements for the MSE range from − 11.36% to − 38.60%, with stronger improvements for one-week-ahead forecasts. The forecast improvements based on QLIKE are even larger as they range from −14.76% to − 41.95%.

Comparing the forecasting accuracy of the RV-HAR, FIV-HAR and IV-HAR models shows that the implied volatility FIV-HAR model outperforms the RV-HAR model, while the IV-HAR model leads to further improvements. These results therefore confirm that the in-sample results reported earlier also manifest into improved out-of-sample forecasts. As IVs appear to be superior to past RVs, it is therefore not surprising to see that different HAR model specifications (a decomposition of RVs into continuous and jump components, the use of semi-volatilities, or the estimation of measurement error) lead to much smaller benefits compared to the simple addition of implied volatilities, which seems to be the preferred choice when forecast accuracy is of interest.

4.3.2. Evidence from combination forecasts

An important question is whether the improved forecasting accuracy of the individual forecasting models that *include* implied volatility is matched by the improved forecasting accuracy from the combination forecasts that do *not include* implied volatility. If this were true, then for forecasting purposes, the use of implied volatility is limited as it would be sufficient to take the combination forecast of the models that ignore the information contained in implied volatilities. However, as shown in Tables 5 and 6, this is not the case, as the combination forecasts that do not contain information from implied volatilities are clearly outperformed by the individual and combination forecasts that use the information contained in implied volatilities.

4.3.3. Statistical evaluation of forecast improvements

In Table 5, the symbol ‡ denotes whether a given row model was included in the superior set of models as indicated by the model confidence set approach of Hansen et al. (2011). All models in this superior set of models are assumed to have similar predictive power.

Table 4
One month ahead variance predictive regressions.

	RV-HAR	FIV-HAR	IV-HAR	CJ-HAR	IV-CJ-HAR	Q-HAR	IV-Q-HAR	PS-HAR	IV-PS-HAR
Intercept	13.979***	14.712***	12.176***	14.335***	12.085***	13.143***	11.805***	14.032***	12.194***
	(3.720)	(2.604)	(2.875)	(3.635)	(2.834)	(2.840)	(2.324)	(3.700)	(2.873)
RV^D	0.051***		0.004			0.109***	0.032		
	(0.013)		(0.009)			(0.024)	(0.018)		
RV^W	0.201**		0.084	0.190**	0.086	0.182**	0.078	0.198**	0.083
	(0.063)		(0.049)	(0.062)	(0.049)	(0.065)	(0.050)	(0.063)	(0.049)
RV^M	0.601***		0.203*	0.594***	0.203*	0.582***	0.200*	0.598***	0.203*
	(0.080)		(0.101)	(0.080)	(0.101)	(0.076)	(0.089)	(0.080)	(0.101)
IV^D		0.051*	0.051**		0.051**		0.051**		0.051**
		(0.021)	(0.019)		(0.019)		(0.016)		(0.019)
IV^W		0.231**	0.180**		0.181**		0.175**		0.178**
		(0.086)	(0.068)		(0.068)		(0.059)		(0.068)
IV^M		0.437***	0.258*		0.259*		0.256**		0.258*
		(0.089)	(0.118)		(0.119)		(0.095)		(0.119)
CC^D				0.080***	− 0.003				
				(0.023)	(0.017)				
JC^D				− 0.032	0.025				
				(0.045)	(0.040)				
$RV^D \times MRQ^D$						− 0.001***	− 0.000		
						(0.000)	(0.000)		
RVP^D								− 0.009	− 0.015
								(0.033)	(0.024)
RVN^D								0.122*	0.030
								(0.051)	(0.037)
R^2	0.618	0.668	0.681	0.618	0.681	0.623	0.683	0.619	0.681
Adj. R^2	0.618	0.668	0.681	0.618	0.681	0.623	0.682	0.619	0.681

Note: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Significances are based on the estimation of the variance-covariance matrix of coefficients by using the automatic bandwidth selection as in Whitney et al., (1994) with a quadratic spectral weighting scheme. We use the implementation procedures provided in Zeileis (2004). The highest variance inflation factor is 10.18 for IV_t^M, but the corresponding coefficients are significant in all models. Model variables: RV^D, RV^W, RV^M = Realized Volatility for the previous day and average for the previous week and month, respectively; IV^D, IV^W, IV^M = Implied Volatility from options with one-day, one-week, and one-month maturity, respectively; CC^D = Continuous Component from the previous day; JC^D = Jump Component for the previous day; MRQ^D = Median Realized Quarticity; RVP^D = Realized Semi-Volatility Positive for the previous day; and RVN^D = Realized Semi-Volatility Negative for the previous day.

The initial set consists of all models with and without implied volatilities alike (for a given loss function and forecast horizon). However, the superior set for one-day, one-week, and one-month-ahead forecasts consists only of the models that include implied volatility. This result provides substantial statistical evidence that implied volatility matters for forecasting purposes. For one-day and one-week-ahead forecasts, the superior set of models according to QLIKE contains only models with both realized and implied volatilities, which in contrast suggests that RVs contain incremental information with respect to IVs, particularly for shorter time horizons.

Table 6 compares the models in a bi-variate setting.[12] The results confirm statistically significant differences in the forecasting accuracy of the models and combination forecasts with and without implied volatility included. The model specifications with implied volatility always outperform their model versions without implied volatility in a statistically significant manner. The same results apply to the combinations of forecasts. This feature is also robust for all forecasting horizons, but implied volatility helps most for one-week forecasting horizons.

4.3.4. Dissecting the superior performance of IV models

Accurate forecasts are most valued during uncertain times. We inspect when forecast improvements are generated. We therefore stratify and average the forecast losses with respect to the realized volatility. In Table 7, we compare the average loss values for the days after a day with an RV that belongs to the top 10% highest realized volatilities or lower. When the ratio is lower, the forecast improvement of the competing model is larger than the benchmark model.

The emerging pattern is that forecast improvements are the largest for the days that follow periods of high market uncertainty. This effect declines with increased forecast horizons. For example, for one-day-ahead forecasts, the improvements are up to 35%, while for one-month-ahead forecasts, the improvements after high vs. low volatile periods tend to be very similar but are still larger after more uncertain days.

Finally, in Fig. 3, we plot the percentage difference between the cumulative loss function of the benchmark HAR model and the best performing HAR model with IVs. Across all forecast horizons, average forecast improvements appear to stabilize after one year of data. The day-ahead volatility forecast improvements are stable across the entire time period. Five-day-ahead and one-month-ahead volatility

[12] This comparison uses the model confidence set of Hansen et al. (2011) with an initial set of two models only.

Table 5
One day ahead, one week ahead, and one month ahead forecasts evaluation.

Panel A: Individual model forecasts	One day						One week						One month					
	MSE			QLIKE			MSE			QLIKE			MSE			QLIKE		
RV-HAR	5992			0.165			2260			0.088			1402			0.067		
FIV-HAR	5382	−10.2%		0.143	−13.8%		1396	−38.3%	‡	0.053	−39.3%		1140	−18.7%	‡	0.055	−16.7%	
IV-HAR	5227	−12.8%	‡	0.138	−16.6%		1380	−38.9%	‡	0.051	−41.5%	‡	1088	−22.4%	‡	0.053	−19.8%	‡
CJ-HAR	5915	−1.3%		0.161	−3.0%		2249	−0.5%		0.086	−1.5%		1405	0.2%		0.066	−0.1%	
IV-CJ-HAR	5200	−13.2%	‡	0.137	−17.3%	‡	1382	−38.9%	‡	0.051	−41.4%	‡	1090	−22.2%	‡	0.054	−19.5%	‡
Q-HAR	5875	−1.9%		0.166	0.3%		2230	−1.4%		0.087	−0.2%		1389	−0.9%		0.066	−0.4%	
IV-Q-HAR	5208	−13.1%	‡	0.138	−16.4%		1408	−37.7%	‡	0.051	−42.1%	‡	1101	−21.4%	‡	0.053	−20.8%	‡
SV-HAR	6003	0.2%		0.166	0.2%		2249	−0.5%		0.088	0.0%		1396	−0.4%		0.067	0.1%	
IV-SV-HAR	5231	−12.7%	‡	0.138	−16.6%		1381	−38.9%	‡	0.051	−41.5%	‡	1089	−22.4%	‡	0.054	−19.5%	‡
Panel B: Unconditional combination model forecasts																		
IV models	5209	−13.1%	‡	0.137	−16.9%	‡	1379	−39.0%	‡	0.051	−41.5%	‡	1082	−22.8%	‡	0.053	−20.5%	‡
non-IV models	5921	−1.2%		0.164	−0.8%		2237	−1.0%		0.087	−0.6%		1392	−0.7%		0.066	−0.4%	
MCS p-value	0.154			0.152			0.545			0.247			0.618			0.637		

Models: RV-HAR = benchmark model that uses only Realized Volatilities; FIV-HAR = alternative to RV-HAR that contains only implied volatilities; IV-HAR = combination of RV-HAR and IV-HAR; CJ-HAR = HAR specification that decomposes RV^D to continuous and jump components; IV-CJ-HAR = CJ-HAR with added implied volatilities; Q-HAR = HAR specification that controls for measurement error; IV-Q-HAR = Q-HAR with added implied volatilities; SV-HAR = HAR specification that decomposes RV^D to positive and negative semi-volatilities; and IV-SV-HAR = SV-HAR with added implied volatilities. Apart from the average value of the loss function, we report in [%] the improvement compared to the benchmark HAR model. The symbol ‡ denotes that the given model (in the row) belongs to the set of superior models. In the row MCS p-value, we report the p-value of the test of equal predictive ability of the models that belong to the set of superior models. Therefore, a higher level of the p-value suggests that at the given significance level, the remaining models have the same predictive ability.

Table 6
Comparing IV and non-IV models forecasting accuracy.

	One day		One week		One month	
	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE
Panel A: Comparing individual models						
HAR vs. IV-HAR	– 12.8%***	– 16.6%****	– 38.9%***	– 41.5%****	– 22.4%***	– 19.8%***
CJ-HAR vs. IV-CJ-HAR	– 12.1%**	– 14.8%****	– 38.6%***	– 40.6%****	– 22.4%***	– 19.4%***
Q-HAR vs. IV-Q-HAR	– 11.4%***	– 16.7%****	– 36.8%***	– 42.0%****	– 20.7%***	– 20.5%****
SV-HAR vs. IV-SV-HAR	– 12.9%***	– 16.8%****	– 38.6%***	– 41.5%****	– 22.0%***	– 19.6%***
Panel B: Comparing combination forecasts; non IV models to IV models						
Mean	– 12.0%***	– 16.3%****	– 38.4%***	– 41.2%****	– 22.3%***	– 20.2%****

Note: Symbols"*","**","***", and"****" denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively, of the test of the equal predictive power of the MCS test applied only on two models. Statistical significance suggests that the model with the lower loss value outperforms the model with a higher loss value. Models: CJ-HAR = HAR specification that decomposes RV^D to continuous and jump components; IV-CJ-HAR = CJ-HAR with added implied volatilities; Q-HAR = HAR specification that controls for measurement error; IV-Q-HAR = Q-HAR with added implied volatilities; SV-HAR = HAR specification that decomposes RV^D to positive and negative semi-volatilities; and IV-SV-HAR = SV-HAR with added implied volatilities.

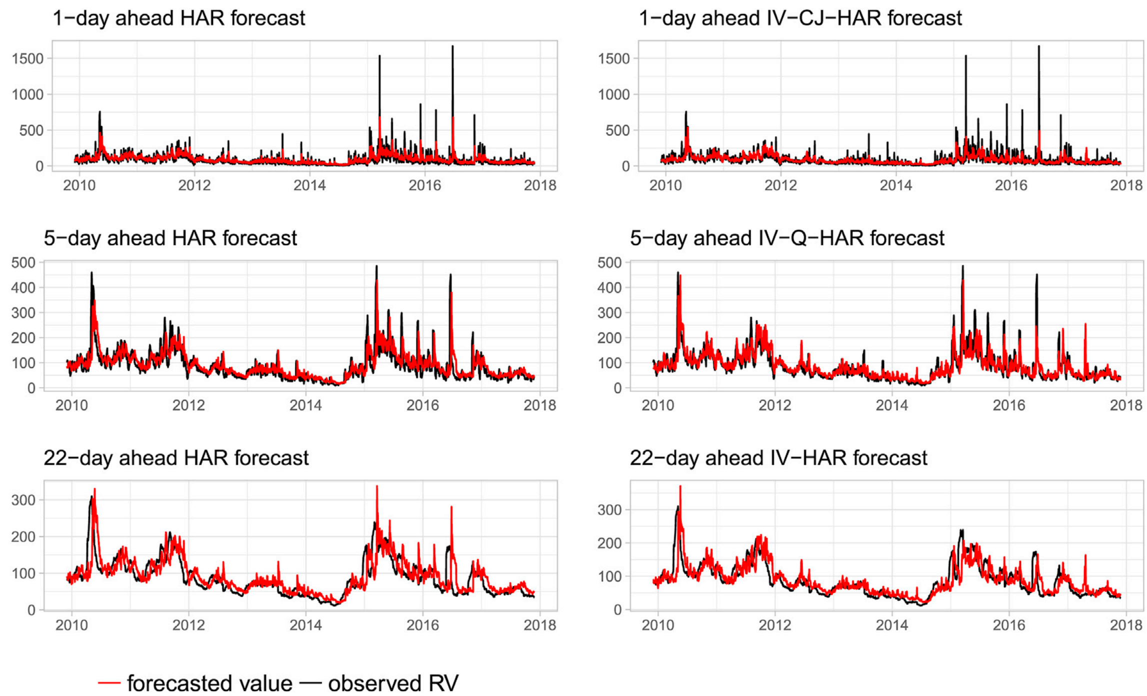

Fig. 2. Time series plots of observed volatility and forecasts from the best model.

forecasts have improved since half of the year into 2016 until the year ends. These improvements in forecasting accuracy seem to coincide with the time around the Brexit referendum in June 2016 that led to a sharp decline in the EUR against the USD and the time of the November U.S. presidential elections that led to the strengthening of the USD over the EUR and other currencies.[13] In general, the results in Fig. 3 show that forecast improvements are pervasive across the time.

4.4. ARFIX-GARCH approach

As we argue in Section 3, the HAR model is currently considered to be one of the best performing volatility models, which is easy to adjust for exogenous variables and easy to implement and interpret. To observe whether our results are an artifact of the HAR models, we also estimate the following ARFIX-GARCH model: $(1-L)^d(RV_{t+1} - \beta_0 - \beta_1 IV_t^D - \beta_2 IV_t^W - \beta_3 IV_t^M) = \varepsilon_{t+1}$, where L is the lag

[13] This strengthening occurred due to expectations of a fiscal stimulus from the new White House cabinet and the subsequent rise in inflation.

Table 7
Stratified out-of-sample forecast losses.

MSE				QLIKE	
Benchmark model	Competing IV-HAR model	top 10% RV	bottom 90% RV	top 10% RV	bottom 90% RV
Panel A: One day ahead forecasts					
RV-HAR	FIV-HAR	0.70	0.95	0.82	0.87
RV-HAR	IV-HAR	0.65	0.93	0.70	0.85
CJ-HAR	IV-CJ-HAR	0.66	0.94	0.74	0.86
Q-HAR	IV-Q-HAR	0.70	0.93	0.72	0.85
SV-HAR	IV-SV-HAR	0.65	0.93	0.69	0.85
Cmb-HAR	Cmb-IV-HAR	0.66	0.94	0.70	0.85
Panel B: One week ahead forecasts					
RV-HAR	FIV-HAR	0.56	0.64	0.51	0.62
RV-HAR	IV-HAR	0.55	0.64	0.48	0.60
CJ-HAR	IV-CJ-HAR	0.55	0.64	0.49	0.61
Q-HAR	IV-Q-HAR	0.62	0.64	0.51	0.59
SV-HAR	IV-SV-HAR	0.57	0.64	0.49	0.60
Cmb-HAR	Cmb-IV-HAR	0.57	0.64	0.49	0.60
Panel C: One month ahead forecasts					
RV-HAR	FIV-HAR	0.69	0.85	0.69	0.85
RV-HAR	IV-HAR	0.63	0.83	0.66	0.82
CJ-HAR	IV-CJ-HAR	0.64	0.82	0.67	0.82
Q-HAR	IV-Q-HAR	0.67	0.83	0.68	0.81
SV-HAR	IV-SV-HAR	0.64	0.83	0.67	0.82
Cmb-HAR	Cmb-IV-HAR	0.63	0.83	0.65	0.82

Note: The values in the table are the loss ratios of the competing HAR model (with implied volatilities) to the benchmark HAR model (without implied volatilities). Values below 1 indicate that the competing model outperformed its benchmark model. Models: RV-HAR = benchmark model that uses only Realized Volatilities; FIV-HAR = alternative to RV-HAR that contains only implied volatilities; IV-HAR = combination of RV-HAR and IV-HAR; CJ-HAR = HAR specification that decomposes RV^D to continuous and jump components; IV-CJ-HAR = CJ-HAR with added implied volatilities; Q-HAR = HAR specification that controls for measurement error; IV-Q-HAR = Q-HAR with added implied volatilities; SV-HAR = HAR specification that decomposes RV^D to positive and negative semi-volatilities; IV-SV-HAR = SV-HAR with added implied volatilities; Cmb-HAR = a combination of forecasts for the models without implied volatility; and Cmb-IV-HAR = a combination of forecasts for the models with implied volatility.

operator, and d the fractional difference parameter, $\varepsilon_{t+1} = \sigma_{t+1}\nu_{t+1}$, $\nu_{t+1} \tilde{} N(0,1)$, and $\sigma^2_{t+1} = \gamma_0 + \gamma_1\varepsilon^2_{t+1} + \gamma_2\sigma^2_{t+1}$. In the in-sample framework, all implied volatility coefficients were significant, but the fractional difference parameter was practically 0 and insignificant, while the monthly IV^M_t was negative. We also run an out-of-sample study by using the same design as the design described in Section 3.2. Given the mean square error loss function (see Section 3.4), the forecasting accuracy of the ARFIX-GARCH model with implied volatilities increased by 17% compared to the standard ARFI-GARCH model estimated without exogenous variables in the mean equation. Given the QLIKE loss function, the predictions had similar accuracy.

5. Conclusion

We study the role of implied volatility calculated from options with different maturities in forecasting the future realized volatility on the highly liquid EUR/USD currency market. We use implied volatility from options with expiration in a day, week, or month. In the modeling framework, we rely on different extensions of the HAR model of Corsi (2009). Except for the original HAR model, we use extensions that account for continuous and jump components and the time-varying measurement error of realized volatility and extensions that separate positive and negative semi-volatilities.

In the in-sample analysis, we compare the relative importance of different IVs, while in the out-of-sample analysis, we study the overall usefulness of IVs with respect to future levels of volatility. To some extent, our results complement earlier findings, most notably from Covrig and Low (2003); Busch et al. (2011) in that we find that implied volatility outperforms the historical realized volatility in in-sample and out-of-sample frameworks. We extend the existing literature in several ways.

First, we provide evidence that even IVs from options with short-term maturity seem to contain non-trivial predictive power. More specifically, in the case of one-day and one-week volatility forecast horizons, the IVs from options with short-term maturity are significant, while the IVs from options with a longer one-month maturity are not significant. Even when modeling one-month-ahead volatility, all IVs seem to be relevant, not just the IV from the option with a one-month maturity. The general pattern that we observed is that the inclusion of the implied volatility in the volatility models decreased the size of the coefficients of realized volatility while the fit of the models increased (up to 29%, 27% and 9% for the one-day, one-week, and one-month-ahead forecasts, respectively).

Second, our results also suggest that utilizing RVs is still beneficial, as the models that include both RVs and IVs show an even greater fit or forecast improvement, but the improvements achieved by including RVs are much smaller than the improvements achieved by utilizing IVs.

Third, we show that the role of implied volatility changes concerning the volatility forecast horizon, as with lower forecast horizons, implied volatilities tend to matter more.

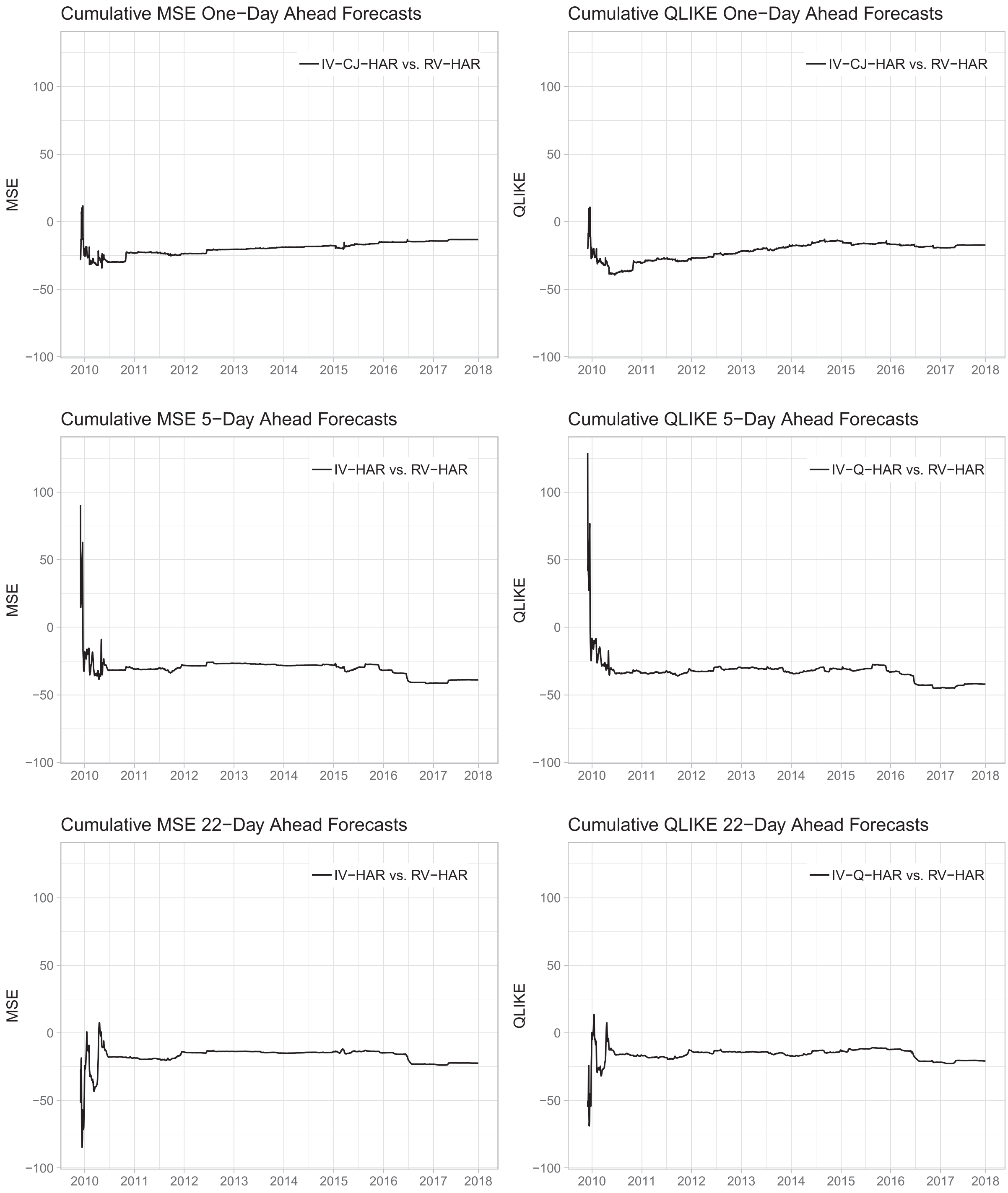

Fig. 3. Percentage difference of cumulative loss functions. Note: The value on the y-axis is the percentage difference between the cumulative loss function of the benchmark HAR model and the best performing HAR model with IVs.

Fourth, our out-of-sample analysis for all time horizons showed that the model confidence set approach resulted in a superior set of models that consisted solely of the models that included implied volatility. These results were supported by estimated improvements in forecasting accuracy as measured through MSE and QLIKE loss functions and regardless of the HAR model employed or whether we used unconditional combination volatility forecasts. For one-day-ahead, one-week-ahead, and one-month-ahead forecasts, the improvements in forecasting accuracy were up to 17.3%, 42.1%, and 22.8%, respectively.

Overall, our results provide substantial evidence that implied volatility from options with a short-term maturity also has predictive power with respect to future levels of volatility on the EUR/USD exchange rate market. Furthermore, different specifications of the HAR models and combination forecasts do not help to improve the predictive power of the volatility models to the same extent as including implied volatilities as additional regressors. However, RV still seems to contain some valuable information because our model with only IV (FIV-HAR) does not dominate the other model specifications with both RV and IV and is also excluded from the superior set of models for one-day and one-month-ahead forecasts according to the QLIKE loss function. Forecast improvements tend to follow highly volatile periods (i.e., when accurate forecasts matter the most and are pervasive across time) and are stable across time.

CRediT authorship contribution statement

Tomáš Plíhal: Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Štefan Lyócsa:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing.

Acknowledgements

We would like to thank the participants of the INFINITI Conference on International Finance 2018 in Poznań and participants at the ECE 2018 in Brno, particularly Katarina Lucivjanska and Mariya Gubareva for their constructive comments. We would like to thank Oleg Deev for his research assistance and useful comments.This research was supported by the Czech Science Foundation (GACR), nr. 18-05829S.

References

Aloud, M., Fasli, M., Tsang, E., Dupuis, R., & Olsen, R. (2013). Stylized facts of trading activity in the high frequency fx market: An empirical study. *Journal of Finance and Investment Analysis, 2*, 145–183.

Amin, K. I., & Morton, A. J. (1994). Implied volatility functions in arbitrage-free term structure models. *Journal of Financial Economics, 35*, 141–180.

Andersen, T. G., & Bollerslev, T. (1998). Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance, 219*.

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

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *The American Economic Review, 38*.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2007b). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics, 73*, 251–277.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics, 169*, 75–93.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies, 16*, 101–143.

Barndorff-Neilsen, O. E., Kinnebrouk, S., & Shephard, N. (2010). In t. bollerslev, j. russell, & m. watson (Eds.), *Measuring downside risk: Realised semivariance* (pp. 117–136). Oxford University Press. chapter Measuring downside risk: realised semivariance.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal, 12*.

Baruník, J., & Hlínková, M. (2016). Revisiting the long memory dynamics of the implied–realized volatility relationship: New evidence from the wavelet regression. *Economic Modelling, 54*, 503–514.

Bates, J. M., & Granger, C. W. J. (1969). *The combination of forecasts* (p. 451). OR.

Becker, R., & Clements, A. E. (2008). Are combination forecasts of s&p 500 volatility statistically superior? *International Journal of Forecasting, 24*, 122–133.

Becker, R., Clements, A. E., & White, S. I. (2006). On the informational efficiency of s&p500 implied volatility. *The North American Journal of Economics and Finance, 17*, 139–153.

Becker, R., Clements, A. E., & White, S. I. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking & Finance, 31*, 2535–2549.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking & Finance, 5*, 363–381.

Blair, B. J., Poon, S.h., & Taylor, S. J. (2001). Forecasting s&p 100 volatility : The incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics, 105*, 5–26.

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

Bollerslev, T., Hood, B., Huss, J., & Pedersen, L. H. (2018). Risk everywhere: Modeling and managing volatility. *Review of Financial Studies, 31*, 2729–2773.

Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics, 192*, 1–18.

Busch, T., Christensen, B. J., & 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.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *Review of Financial Studies, 6*, 659.

Çelik, S., & Ergin, H. (2014). Volatility forecasting using high frequency data: Evidence from stock markets. *Economic Modelling, 36*, 176–190.

Chatrath, A., Miao, H., Ramchander, S., & Wang, T. (2015). The forecasting efficacy of risk-neutral moments for crude oil volatility. *Journal of Forecasting, 34*, 177–190.

Christensen, B., & Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics, 50*, 125–150.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559–583.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics, 7*, 174.

Covrig, V., & Low, B. S. (2003). The quality of volatility traded on the over-the-counter currency market: A multiple horizons study. *Journal of Futures Markets, 23*, 261–284.

Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics, 52*, 267–287.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 987*.

Fatum, R., Hutchison, M., & Wu, T. (2012). Asymmetries and state dependence: The impact of macro surprises on intraday exchange rates. *Journal of the Japanese and International Economies, 26*, 542–560.

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. *Journal of Monetary Economics, 54*, 1051–1068.

Garman, M., & Kohlhagen, S. (1983). Foreign currency option values. *Journal of International Money and Finance, 2*, 231–237.

Gau, Y. F., & Wu, Z. X. (2017). Macroeconomic announcements and price discovery in the foreign exchange market. *Journal of International Money and Finance, 79*, 232–254.

Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting, 29*, 108–121.
Han, H., & Park, M. D. (2013). Comparison of realized measure and implied volatility in forecasting volatility. *Journal of Forecasting, 32*, 522–533.
Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica, 79*, 453–497.
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.
Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance, 281*.
Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *Review of Financial Studies, 18*, 1305–1342.
Jorion, P. (1995). Predicting volatility in the foreign exchange market. *The Journal of Finance, 50*, 507–528.
Kambouroudis, D. S., McMillan, D. G., & Tsakou, K. (2016). Forecasting stock return volatility: A comparison of garch, implied volatility, and realized volatility models. *Journal of Futures Markets, 36*, 1127–1163.
Kocenda, E., & Moravcova, M. (2016). Ntraday effect of news on emerging European forex markets: An event study analysis. In *I Working papers IES 2016/20*. Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.
Kourtis, A., Markellos, R. N., & Symeonidis, L. (2016). An international comparison of implied, realized, and garch volatility forecasts. *Journal of Futures Markets, 36*, 1164–1193.
Lin, J. W., & McLeod, A. I. (2006). Improved peňa–rodriguez portmanteau test. *Computational Statistics & Data Analysis, 51*, 1731–1738.
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*, 293–311.
Lyócsa, S., & Molnár, P. (2016). Volatility forecasting of strategically linked commodity etfs: Gold-silver. *Quantitative Finance, 16*, 1809–1822.
Lyócsa, S., 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.
Ma, F., Liu, J., Huang, D., & Chen, W. (2017). Forecasting the oil futures price volatility: A new approach. *Economic Modelling, 64*, 560–566.
Muller, U. A., Dacorogna, M. M., Dave, R. D., Olsen, R. B., Pictet, O. V., & von Weizsacker, J. E. (1997). Volatilities of different time resolutions — analyzing the dynamics of market components. *Journal of Empirical Finance, 4*, 213–239. High Frequency Data in Finance, Part 1.
Newbold, P., & Harvey, D. (2007). *Forecast combination and encompassing*. wiley, University of Nottingham.
Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics, 160*, 246–256.
Patton, A. J., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *The Review of Economics and Statistics, 97*, 683–697.
Pong, S., Shackleton, M. B., Taylor, S. J., & Xu, X. (2004). Forecasting currency volatility : A comparison of implied volatilities and ar[fi]ma models. *Journal of Banking & Finance, 28*, 2541–2563.
Santos, D. G., & Ziegelmann, F. A. (2014). Volatility forecasting via midas, har and their combination: An empirical comparative study for ibovespa. *Journal of Forecasting, 33*, 284–299.
Seo, S. W., & Kim, J. S. (2015). The information content of option-implied information for volatility forecasting with investor sentiment. *Journal of Banking & Finance, 50*, 106–120.
Sevi, B. (2014). Stochastics and statistics: Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research, 235*, 643–659.
Timmermann, A. (2006). Chapter 4 forecast combinations. In G. Elliott, C. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting, 1* pp. 135–196). Elsevier.
Wang, Y. H., & Wang, Y. Y. (2016). The information content of intraday implied volatility for volatility forecasting. *Journal of Forecasting, 35*, 167–178.
Whitney, K., Newey, a, & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 631.
Xinzhong, X., & Taylor, S. J. (1995). Conditional volatility and the informational efficiency of the phlx currency options market. *Journal of Banking & Finance, 19*, 803–821.
Zeileis, A. (2004). Econometric computing with hc and hac covariance matrix estimators. *Journal of Statistical Software, 11*, 1–17.