Data Science in Finance and Economics

https://www.aimspress.com/journal/dsfe

DSFE, 4(4): 514–530.
DOI:10.3934/DSFE.2024021
Received: 21 April 2024
Revised: 15 September 2024
Accepted: 28 October 2024
Published: 30 October 2024

Research article

On the relevance of realized quarticity for exchange rate volatility forecasts

Morten Risstad[1]***and Mathias Holand**

Norwegian University of Science and Technology, Dept. of Industrial Economics and Technology Management, Norway

* **Correspondence:** Email: morten.risstad@ntnu.no.

Abstract: High-frequency tick data have proved helpful for forecasting volatility across asset classes. In the finite samples typically faced by practitioners, however, noise inherent in tick-level prices creates inaccuracies in model parameter estimates and resulting forecasts. A remedy proposed to alleviate these measurement errors is to include higher-order moments, more specifically the realized quarticity, in volatility prediction models. In this paper, we investigate the relevance of this approach in foreign exchange markets, as represented by EURUSD and USDJPY data from 2010 to 2022. Using well-established realized volatility models, we find that including realized quarticity leads to higher precision in daily, weekly, and monthly out-of-sample forecasts. These results are robust across estimation windows, evaluation metrics, and model specifications.

Keywords: foreign exchange; volatility; high-frequency data; realized quarticity; multi-period forecasts

JEL Codes: F31, C58, D53, G17

1. Introduction

Estimating foreign exchange rate (FX) volatility is a core risk management activity for financial institutions, corporates and regulators. The subject has been extensively investigated among both practitioners and scientific researchers, and several alternative models exist. Among the most prominent are the models belonging to GARCH and stochastic volatility classes. However, the true value of volatility cannot be directly observed. Hence, volatility must be estimated, inevitably with error. This constitutes a fundamental problem in implementing parametric models, especially in the context of high-frequency data. Andersen and Bollerslev (1998) proposed using realized volatility, as derived from high-frequency data, to accurately measure the true latent integrated volatility. This approach has gained attention for volatility modeling in markets where tick-level data is available (Andersen et al., 2013). Andersen et al. (2003) suggest fractionally integrated ARFIMA models in this context. Still, the long-memory HAR (heterogeneous autoregressive) model of Corsi (2009) is arguably the most widely used to capture the high persistence typically observed in realized volatility of financial prices. The HAR model is relatively simple and easy to estimate. In empirical applications, the model tends to perform better than GARCH and stochastic volatility models possibly due to the sensitivity of tightly parameterized volatility models to minor model misspecifications (Sizova., 2011). Although realized volatility (RV) is a consistent estimator of the true latent volatility, it is subject to measurement error in empirical finite samples. Hence, RV will not only reflect the true latent integrated volatility (IV), but also additional measurement errors. Bollerslev et al. (2016) propose utilizing higher-order realized moments of the realized distribution to approximate these measurement errors. More specifically, Bollerslev et al. (2016) propose the HARQ-model, which augments the HAR model with realized quarticity as an additional covariate.

The empirical performance of the HARQ and related extensions has been extensively studied. The focus has predominantly been on equity markets. A majority of the studies analyze U.S. data; see Bollerslev et al. (2016); Clements, A. and Preve, D. (2021); Pascalau and Poirier (2023); Andersen et al. (2023) and others. Liu et al. (2018) and Wang et al. (2020) investigate Chinese equity markets, whereas Liang et al. (2022); Ma et al. (2019) analyse international data. Bitcoin and electricity markets have attracted some attention; see, for instance, Shen et al. (2020); Qieu et al. (2021), and Qu et al. (2018).

Empirical applications of the HARQ model in the context of foreign exchange rate risk are sparse. Lyócsa et al. (2016) find that the standard HAR model rarely is outperformed by less parsimonious specifications on CZKEUR, PLZEUR, and HUFEUR data. Plíhal et al. (2021) and Rokicka and Kudła (2020) estimate the HARQ model on EURUSD and EURGBP data, respectively. Their focus is different from ours, as they investigate the incremental predictive power of implied volatility for a broad class of HAR models. In a similar vein, Götz (2023) and Lyócsa et al. (2024) utilize the HARQ model for the purpose of estimating foreign exchange rate tail risk.

Using updated tick-level data from two major currency pairs, EURUSD and USDJPY, this paper documents the relevance of realized quarticity for improving volatility estimates across varying forecasting horizons. These results are robust across estimation windows, evaluation metrics, and model specifications.

2. Materials and method

2.1. Data

We use high-frequency intraday ticklevel spot data, publicly available at DukasCopy.* The sample period is 1. January 2010 to 31. December 2022. Liu et al. (2015) investigate the optimal intraday sampling frequency across a significant number of asset classes and find that 5-min intervals usually outperform others. Hence, as common in the literature, we estimate the realized volatility from 5-minute returns.

To filter tick-level data, we follow a two-step cleaning procedure based on the recommendations by Barndorff-Nielsen et al. (2009). Initially, we eliminate data entries that exhibit any of the following issues: (i) absence of quotes, (ii) a negative bid-ask spread, (iii) a bid-ask spread exceeding 50 times the median spread of the day, or (iv) a mid-quote deviation beyond ten mean absolute deviations from a centered mean (computed excluding the current observation from a window of 25 observations before and after). Following this, we calculate the mid-quotes as the average of the bid and ask quotes and then resample the data at 5-minute intervals.

*This data source is also used by Plíhal et al. (2021), Risstad et al. (2023) and Lyócsa et al. (2024), among others.

We compute the consistent estimator of the true latent time-t variance from

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

where $M = 1/\Delta$, and the Δ-period intraday return is $r_{t,i} \equiv \log(S_{t-1+i\times\Delta}) - \log(S_{t-1+(i-1)\times\Delta})$, where S is the spot exchange rate. Analogously, the multi(h)-period realized variance estimator is

$$RV_{t-1,t-h}^2 = \frac{1}{h}\sum_{i=1}^{h} RV_{t-h}^2. \tag{2}$$

Setting $h = 5$ and $h = 22$ yields weekly and monthly estimates, respectively.

Table 1 displays descriptive statistics for daily realized variances, as computed from (1).

Table 1. Realized Variance (daily).

	Min	Mean	Median	Max	ρ_1
EURUSD	0.1746	3.0606	2.2832	59.4513	0.5529
USDJPY	0.1018	3.2460	2.0096	168.0264	0.2860

The table contains summary statistics for the daily RV s for EURUSD and USDJPY. ρ_1 is the standard first order autocorrelation coefficient. Sample period: 1. January 2010 to 31. December 2022.

2.2. The HARQ models

To represent the long-memory dynamic dependencies in volatility, Corsi (2009) proposed using daily, weekly, and monthly lags of realized volatility as covariates. The original HAR model is defined as

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t, \tag{3}$$

where RV is computed from (1) and (2). If the variables in (2.2) contain measurement errors, the beta coefficients will be affected. Bollerslev et al. (2016) suggests two measures to alleviate this. First, they include a proxy for measurement error as an additional explanatory variable. Furthermore, they directly adjust the coefficients in proportion to the magnitude of the measurement errors:

$$\begin{aligned} RV_t =&\beta_0 + \underbrace{\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right)}_{\beta_{1,t}} RV_{t-1} \\ &+ \underbrace{\left(\beta_2 + \beta_{2Q} RQ_{t-1|t-5}^{1/2}\right)}_{\beta_{2,t}} RV_{t-1|t-5} \\ &+ \underbrace{\left(\beta_3 + \beta_{3Q} RQ_{t-1|t-22}^{1/2}\right)}_{\beta_{3,t}} RV_{t-1|t-22} + u_t, \end{aligned}$$

where realized quarticity RQ is defined as

$$RQ_t \equiv \frac{M}{3}\sum_{i=1}^{M} r_{t,i}^4 \tag{4}$$

The *full HARQ* model in (2.2) adjusts the coefficients on all lags of RV. A reasonable conjecture is that measurement errors in realized volatilities tend to diminish at longer forecast horizons, as these errors are diversified over time. This suggests that measurement errors in daily lagged realized volatilities are likely to be relatively more important. Motivated by this Bollerslev et al. (2016) specify the *HARQ* model as

$$\begin{aligned} RV_t =&\beta_0 + \underbrace{\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right)}_{\beta_{1,t}} RV_{t-1} + \beta_2 RV_{t-1|t-5} \\ &+ \beta_3 RV_{t-1|t-22} + u_t. \end{aligned} \tag{5}$$

Although there is no reason to expect that autoregressive models of order one will be able to accurately capture long memory in realized volatility, we estimate AR(1) models as a point of reference. The *AR* and *ARQ* models are defined as

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + u_t. \tag{6}$$

and

$$RV_t = \beta_0 + \underbrace{\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right)}_{\beta_{1,t}} RV_{t-1} + u_t. \tag{7}$$

in equations (6) and (7), respectively.

3. Results and discussion

Due to noisy data and related estimation errors, forecasts from realized volatility models might occasionally appear as unreasonably high or low. Thus, in line with Swanson et al. (1997) and Bollerslev et al. (2016), we filter forecasts from all models so that any forecast outside the empirical distribution of the estimation sample is replaced by the sample mean.

3.1. In-sample estimation results

Table 2 reports in-sample parameter estimates for the ARQ, HARQ, and HARQ-F models, along with the benchmark AR and ARQ models, for one-day ahead EURUSD (upper panel) and USDJPY (lower panel) volatility forecasts. Robust standard errors (s.e.) are computed as proposed by White (1980). R^2, MSE, and $QLIKE$ are displayed at the bottom of each panel.

Table 2. In-sample estimation results, one-day-ahead volatility forecasts.

EURUSD	AR	HAR	ARQ	HARQ	HARQ-F
β_0	1.3663*	0.3961*	0.7428*	0.2785*	−0.0651
s.e.	0.1843	0.0598	0.0969	0.0586	0.0685
β_1	0.5530*	0.2364*	0.7903*	0.4349*	0.3740*
s.e.	0.0653	0.0730	0.0388	0.0754	0.0792
β_2		0.3767*		0.3072*	0.4613*
s.e.		0.0717		0.0697	0.1031
β_3		0.2572*		0.1850*	0.2398*
s.e.		0.0532		0.0515	0.0822
β_{1Q}			−2.4914*	−1.3708*	−0.9710*
s.e.			0.3377	0.1939	0.2266
β_{2Q}					−1.7578***
s.e.					0.8706
β_{3Q}					−3.9819*
s.e.					1.1618
R^2	0.3058	0.3956	0.3685	0.4101	0.4166
MSE	6.3005	5.4852	5.7315	5.3538	5.2950
$QLIKE$	0.1647	0.1230	0.1540	0.1217	0.1199

USDJPY	AR	HAR	ARQ	HARQ	HARQ-F
β_0	2.3073*	1.0682*	1.3537*	0.7811*	0.5218*
s.e.	0.2362	0.1381	0.2207	0.1429	0.1328
β_1	0.2854*	0.1819*′	0.6180*	0.5177***	0.4416*
s.e.	0.0804	0.0806	0.0853	0.1106	0.1260
β_2		0.1441**		0.0542	0.2345***
s.e.		0.0585		0.0543	0.1072
β_3		0.3443*		0.2188*	0.2228*
s.e.		0.0499		0.0493	0.0658
β_{1Q}			−0.2295*	−0.1967*	−0.1526*
s.e.			0.0318	0.0386	0.0476
β_{2Q}					−0.2296**
s.e.					0.0849
β_{3Q}					−0.3573*
s.e.					0.1142
R^2	0.0814	0.1154	0.1489	0.1581	0.1642
MSE	33.6096	32.3668	31.1409	30.8063	30.5818
$QLIKE$	0.3214	0.2561	0.2663	0.2377	0.2242

Note: The table contains in-sample parameter estimates and corresponding standard errors (White, 1980), together with R^2. MSE and $QLIKE$ computed from (12) and (13). Superscripts *, **, and *** represent statistical significance in a two-sided t-test at 1%, 5% and 10% levels, respectively.

The coefficients β_{1Q} are negative and exhibit strong statistical significance, aligning with the hypothesis that RQ represents time-varying measurement error. When comparing the autoregressive

(AR) coefficient of the AR model to the autoregressive parameters in the ARQ model, the AR coefficient is markedly lower, reflecting the difference in in persistence between the models.

In the comparative analysis of the HAR and HARQ models applied to both currency pairs, the HAR model assigns more emphasis to the weekly and monthly lags, which are generally less sensitive to measurement errors. In contrast, the HARQ model typically assigns a higher weight to the daily lag. However, when measurement errors are substantial, the HARQ model reduces the weight on the daily lag to accommodate the time-varying nature of the measurement errors in the daily realized volatility (RV). The flexible version of this model, the HARQ-F, allows for variability in the weekly and monthly lags, resulting in slightly altered parameters compared to the standard HARQ model. Notably, the coefficients β_{2Q} and β_{3Q} in the HARQ-F model are statistically significant, and this model demonstrates a modest enhancement in in-sample fit relative to the HARQ model.

3.2. Out-of-sample forecasting results

To further assess the out-of-sample performance of the HARQ model, we consider three alternative HAR type specifications. More specifically, we include both the HAR-with-Jumps (HAR-J) and the Continuous-HAR (CHAR) proposed by Andersen et al. (2007), as well as the SHAR model proposed by Patton and Sheppard (2015), in the forecasting comparisons. Based on the Bi-Power Variation (BPV) measure of Barndorff-Nielsen and Shephard (2004), HAR-J and CHAR decompose the total variation into a continuous and a discontinuous (jump) part.

The HAR-J model augments the standard HAR model with a measure of the jump variation;

$$\begin{aligned} RV_t =& \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} \\ &+ \beta_J J_{t-1} + u_t, \end{aligned} \tag{8}$$

where $J_t \equiv \max\left[RV_t - BPV_t, 0\right]$, and the BPV measure is defined as,

$$BPV_t \equiv \mu_1^{-2} \sum_{i=1}^{M-1} \left|r_{t,i}\right| \left|r_{t,i+1}\right|, \tag{9}$$

with $\mu_1 = \sqrt{2/\pi} = \mathbb{E}(|Z|)$, and Z is a standard normal random variable.

The CHAR model includes measures of the continuous component of the total variation as covariates;

$$RV_t = \beta_0 + \beta_1 BPV_{t-1} + \beta_2 BPV_{t-1|t-5} + \beta_3 BPV_{t-1|t-22} + u_t. \tag{10}$$

Inspired by the semivariation measures of Barndorff-Nielsen et al. (2008), Patton and Sheppard (2015) propose the SHAR model, which, in contrast to the HAR model, effectively allows for asymmetric responses in volatility forecasts from negative and positive intraday returns. More specifically, when $RV_t^- \equiv \sum_{i=1}^{M} r_{t,i}^2 \mathbb{I}_{\{r_{t,i}<0\}}$ and $RV_t^+ \equiv \sum_{i=1}^{M} r_{t,i}^2 \mathbb{I}_{\{r_{t,i}>0\}}$, the SHAR model is defined as:

$$\begin{aligned} RV_t =& \beta_0 + \beta_1^+ RV_{t-1}^+ + \beta_1^- RV_{t-1}^- + \beta_2 RV_{t-1|t-5} \\ &+ \beta_3 RV_{t-1|t-22} + u_t. \end{aligned} \tag{11}$$

To evaluate model performance, we consider the mean squared error (MSE) and the QLIKE loss, which, according to Patton (2011), both are robust to noise. MSE is defined as

$$MSE(RV_t, F_t) \equiv (RV_t - F_t)^2, \tag{12}$$

where F_t refers to the one-period direct forecast. QLIKE is defined as

$$QLIKE(RV_t, F_t) \equiv \frac{RV_t}{F_t} - \ln\left(\frac{RV_t}{F_t}\right) - 1. \tag{13}$$

3.2.1. Daily forecasting horizon

Table 3. Out-of-sample forecast losses, one-day-ahead volatility forecasts.

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.1483	1.0000	1.0088	0.9945	1.0080	1.0311	0.9759	**0.9655***
MSE-EW	1.1619	1.0000	0.9984	0.9908	1.0050	1.02660	0.9742	**0.9720***
QLIKE-RW	1.3153	1.0000	0.9907	0.9813	1.0078	1.1575	0.9767	**0.9582***
QLIKE-EW	1.3915	1.0000	0.9907	0.9944	1.0052	1.1927	0.9952	**0.9721****

USDJPY	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.0502	1.0000	1.0053	0.9979	1.0238	0.8907	0.8885	**0.8832***
MSE-EW	1.0475	1.0000	1.0243	1.0133	1.0515	0.9558	0.9446	**0.9376***
QLIKE-RW	1.2320	1.0000	1.0748	0.9944	0.9811	0.9482	0.8824	**0.8667***
QLIKE-EW	1.3066	1.0000	1.0023	0.9800	0.9941	1.0039	0.8949	**0.8519***

Note: Model performance, expressed as model loss normalized by the loss of the HAR model. Each row reflects a combination of estimation window and loss function. Ratio for the best performing model on each row in bold.Corresponding asterix * and ** denote 1% and 5% confidence levels from Diebold-Mariano test for one-sided tests of superior performance of the best performing model compared to the HAR model.

Table 3 contains one-day-ahead forecasts for EURUSD and USDJPY. The table reports model performance, expressed as model loss normalized by the loss of the HAR model. Each row reflects a combination of estimation window and loss function. The lowest ratio on each row, highlighting the best, performing model, is in bold. We evaluate the models using both a rolling window (RW) and an expanding window (EW). In both cases, forecasts are derived from model parameters re-estimated each day with a fixed length RW comprised of the previous 1000 days, as well as an EW using all of the available observations. The sample sizes for EW thus range from 1000 to 3201 days. The results are consistent in that the HARQ-F model is the best performer for both currency pairs and across loss functions and estimation windows. The HARQ model is closest to HARQ-F. Neither HAR-J, CHAR, nor SHAR appear to consistently improve upon the standard HAR model.

Table 4. Stratified one-day-ahead out-of-sample forecast losses.

(**a**) Bottom 95% RQ

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.1156	1.0000	0.9937	0.9907	1.0021	1.0636	0.9925	**0.9794**
MSE-IW	1.1175	1.0000	0.9887	0.9885	1.0020	1.0711	0.9967	**0.9866**
QLIKE-RW	1.3299	1.0000	0.9975	0.9855	1.0071	1.1598	0.9745	**0.9555**
QLIKE-IW	1.4108	1.0000	0.9956	0.9980	1.0055	1.1995	0.9944	**0.9720**

height**USDJPY**	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.0330	1.0000	1.0146	0.9984	0.9940	0.9592	0.9526	**0.949**5
MSE-IW	1.0590	1.0000	0.9962	0.9925	1.0001	0.9849	0.9681	**0.9601**
QLIKE-RW	1.2507	1.0000	1.1353	0.9877	0.9829	0.9542	0.8797	**0.8450**
QLIKE-IW	1.3266	1.0000	0.9883	0.9734	0.9993	1.0100	0.8887	**0.8434**

(**b**) Top 5% RQ

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.2276	1.0000	1.0453	1.0036	1.0225	0.9523	0.9355	**0.9316**
MSE-IW	1.2642	1.0000	1.0206	0.9960	1.0121	**0.9218**	0.9224	0.9382
QLIKE-RW	1.0876	1.0000	**0.8851**	0.9152	1.0186	1.1223	1.0116	0.9996
QLIKE-IW	1.0902	1.0000	**0.9141**	0.9389	1.0006	1.0856	1.0081	0.9745

USDJPY	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.0674	1.0000	1.0025	0.9974	1.0535	0.9425	0.8700	**0.8518**
MSE-IW	1.0347	1.0000	1.0566	1.0365	1.1090	0.9246	0.9183	**0.9126**
QLIKE-RW	1.0202	1.0000	1.5755	1.0697	0.9601	**0.8803**	0.9135	0.9999
QLIKE-IW	1.0544	1.0000	1.1789	1.0628	0.9279	**0.9278**	0.9730	0.9588

Note: The table segments the results in Table 3 according to *RQ*. The bottom panel shows the ratios for days following a value of *RQ* in the top 5%. The top panel shows the results for the remaining 95% of sample. Ratio for the best performing model on each row in bold.

Judging from Table 3, it is beneficial to include RQ as an explanatory variable when RV is measured inaccurately. However, precise measurement of RV becomes more difficult when RV is high, inducing a positive correlation between RV and RQ. At the same time, high *RV* often coincides with jumps. To clarify whether the performance of RQ-based models is due to jump dynamics, Table 4 further segments the results in Table 3 into forecasts for days when the previous day's *RQ* was very high (Top 5% RQ, Table 4b) and the remaining sample (Bottom 95% RQ, Table 4a). As this breakdown shows, the RQ-based models perform relatively well also during periods of non-extreme heteroscedasticity of RQ.

3.2.2. Longer forecast horizons

In practitioner applications, longer forecasts than one day are often of interest. We now extend our analysis to weekly and monthly horizons, using direct forecasts. The daily forecast analysis in subsubsection 3.2.1 indicates the lag order of RQ plays an important role in forecast accuracy. Hence, following Bollerslev et al. (2016), we consider the HARQ-h model, and adjust the lag corresponding to the specific forecast horizon only. Specifically, for the weekly and monthly forecasts analysed here, the relevant HARQ-h specifications become

$$\begin{aligned} RV_{t+4|t} =&\beta_0 + \beta_1 RV_{t-1} + \underbrace{\left(\beta_2 + \beta_{2Q} RQ^{1/2}_{t-1|t-5}\right)}_{\beta_{2,t}} RV_{t-1|t-5} \\ &+ \beta_3 RV_{t-1|t-22} + u_t \end{aligned} \tag{14}$$

and

$$\begin{aligned} RV_{t+21|t} =&\beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + u_t, \\ &+ \underbrace{\left(\beta_3 + \beta_{3Q} RQ^{1/2}_{t-1|t-22}\right)}_{\beta_{3,t}} RV_{t-1|t-22} + u_t, \end{aligned} \tag{15}$$

respectively.

Table 5. In-sample weekly and monthly model estimates.

(a) EURUSD

	Weekly						Monthly					
	AR	ARQ	HAR	HARQ	HARQ-F	HARQ-h	AR	ARQ	HAR	HARQ	HARQ-F	HARQ-h
β_0	0.8646*	0.2634*	0.5680*	0.4758*	-0.0250	0.2275**	1.6388*	0.9642*	0.9269*	0.8452*	0.2328	0.2153
s.e.	0.1345	0.0927	0.0997	0.0882	0.0895	0.0861	0.1806	0.1840	0.2246	0.2099	0.2080	0.2093
β_1	0.7168*	0.9620*	0.1194*	0.2752*	0.1836*	0.1181*	0.4616*	0.7373*	0.0717*	0.2097*	0.1131*	0.0646*
s.e.	0.0480	0.0400	0.0264	0.0395	0.0269	0.0214	0.0564	0.0616	0.0205	0.0401	0.0248	0.0185
β_2			0.3938*	0.3395*	0.5777*	0.7635*			0.2091*	0.1606*	0.3706*	0.2176*
s.e.			0.0887	0.0881	0.1282	0.1139			0.0587	0.0554	0.0962	0.0563
β_3			0.3008*	0.2440*	0.3131**	0.0876			0.4163*	0.3661*	0.5153*	0.7179*
s.e.			0.0880	0.0817	0.1275	0.0940			0.1186	0.1174	0.1498	0.1106
β_{1Q}		−5.4876*		−1.0749*	−0.4728*			−6.1534*		−0.9499*	−0.3246*	
s.e.		0.4817		0.1377	0.1005			0.9900		0.1815	0.0846	
β_{2Q}					−2.7357*	−4.9739*					−2.3111*	
s.e.					0.9302	0.7181					0.8020	
β_{3Q}					−5.6441*						−7.8467*	−10.9979*
s.e.					1.4540						2.1071	1.9082
R^2	0.5138	0.5642	0.5453	0.5604	0.5843	0.5756	0.4297	0.5191	0.5072	0.5237	0.5678	0.5568
MSE	2.6073	2.3370	2.4385	2.3576	2.2292	2.2759	2.1913	1.8477	1.8932	1.8299	1.6606	1.7027
$QLIKE$	0.0862	0.0731	0.0752	0.0735	0.0679	0.0704	0.1073	0.0804	0.0839	0.0012	0.0760	0.0788

(b) USDJPY

	Weekly						Monthly					
	AR	ARQ	HAR	HARQ	HARQ-F	HARQ-h	AR	ARQ	HAR	HARQ	HARQ-F	HARQ-h
β_0	2.0305*	1.1976*	1.3310*	1.1591*	0.8708*	0.9646*	2.5786*	2.2815*	1.7358*	1.6356*	1.3894*	1.3900*
s.e.	0.2484	0.1550	0.1967	0.1646	0.1701	0.1564	0.1928	0.2245	0.2792	0.2678	0.3106	0.3151
β_1	0.3709*	0.6801*	0.0687*	0.2722*	0.1650*	0.0668*	0.2011*	0.3121*	0.0286**	0.1460*	0.0829*	0.0283*
s.e.	0.0717	0.0512	0.0266	0.0500	0.0373	0.0207	0.0363	0.0566	0.0119	0.0258	0.0166	0.0113
β_2			0.1294	0.0742	0.3558*	0.4971*			0.0865*	0.0541	0.1886*	0.0923*
s.e.			0.0700	0.0609	0.0787	0.0790			0.0389	0.0333	0.0487	0.0376
β_3			0.3910*	0.3147*	0.2622*	0.1829*			0.3460*	0.3030*	0.3340*	0.4811*
s.e.			0.0703	0.0621	0.0959	0.0693			0.0916	0.0883	0.1346	0.1220
β_{1Q}		−0.6085*		−0.1190*	−0.0571*			−0.2167**		−0.0678*	−0.0318*	
s.e.		0.0534		0.0173	0.0141			0.0832		0.0093	0.0068	
β_{2Q}					−0.3653*	−0.5357*					−0.1659*	
s.e.					0.0704	0.0648					0.0465	
β_{3Q}					-0.2750						-0.3946	−0.7392**
s.e.					0.2010						0.2942	0.2900
R^2	0.1367	0.2323	0.1848	0.2270	0.2557	0.2475	0.1414	0.2106	.2205	0.2496	0.2761	0.2542
MSE	11.6923	10.3980	11.0412	10.4701	10.0811	10.1919	5.4365	4.9983	4.9351	4.7513	4.58326	4.7220
$QLIKE$	0.2361	0.4197	0.2057	0.1937	0.4076	0.1405	0.2143	0.1973	0.1801	0.1734	0.1634	0.1680

Note: In-sample parameter estimates for weekly (h = 5) and monthly (h = 22) forecasting models. EURUSD in upper panel (Table 5a) and USDJPY in lower panel (Table 5b). Robust standard errors (s.e.) using Newey and West (1987) accommodate autocorrelation up to order 10 (h = 5), and 44 (h = 22), respectively. Superscripts *, ** and *** represent statistical significance in a two-sided t-test at 1%, 5%, and 10% levels.

Table 5 presents in-sample parameter estimates across model specifications. The patterns observed here closely resemble those of the daily estimates detailed in Table 2. All coefficients on RQ (β_{1Q},β_{2Q},β_{3Q}) are negative, except for the ($h = 22$) lag statistically significant. This indicates that capturing measurement errors is relevant also for forecast horizons beyond one day. The HARQ model consistently allocates greater weight to the daily lag compared to the standard HAR model. Similarly, the HARQ-h model predominantly allocates its weight towards the time-varying lag. The weights of the HARQ-F model on the different lags are relatively more stable when compared to the HARQ-h model.

Table 6. Weekly out-of-sample forecast losses.

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F	HARQ-h
MSE-RW	1.3063	1.0000	0.9636	0.9884	1.0017	1.1459	0.9677	**0.9024***	0.9205
MSE-EW	1.2702	1.0000	0.9433	0.9559	0.9997	1.1288	0.9501	**0.8996***	0.9117
QLIKE-RW	1.5923	1.0000	0.9819	0.9840	0.9995	1.3558	0.9932	**0.8701**	0.9283
QLIKE-EW	1.7682	1.0000	0.9874	1.0031	1.0033	1.4134	0.9648	**0.8832***	0.9297
USDJPY	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F	HARQ-h
MSE-RW	1.0618	1.0000	0.9464	0.9509	0.9965	0.9064	0.8971	**0.8393***	0.8443
MSE-EW	1.1707	1.0000	1.0148	1.0021	1.0336	1.0194	0.9388	0.8993	**0.8976***
QLIKE-RW	1.3119	1.0000	1.0057	0.9910	0.9740	1.0493	0.9099	**0.8246***	0.8359
QLIKE-EW	1.3847	1.0000	0.9918	0.9768	1.0002	1.1391	0.9179	**0.8350***	0.8463

Note: Model performance, expressed as model loss normalized by the loss of the HAR model. Each row reflects a combination of estimation window and loss function. Ratio for the best-performing model on each row in bold. Corresponding asterix * and ** denote 1% and 5% confidence levels from Diebold-Mariano test for one-sided tests of superior performance of the best performing model compared to the HAR model.

Table 7. Monthly out-of-sample forecast losses.

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F	HARQ-h
MSE-RW	1.3289	1.0000	0.9876	0.9952	1.0003	1.1876	0.9625	**0.8803***	0.9004
MSE-IW	1.3265	1.0000	0.9759	1.0010	1.0044	1.1707	0.9537	**0.8723**	0.9070
QLIKE-RW	1.4301	1.0000	0.9945	0.9950	0.9982	1.2380	0.9622	**0.9215***	0.9279
QLIKE-IW	1.5155	1.0000	0.9951	1.0051	1.0011	1.2596	0.9599	**0.9333**	0.9784
USDJPY	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F	HARQ-h
MSE-RW	1.2529	1.0000	1.0215	1.0086	0.9893	1.5820	1.0500	1.0070	**0.9621***
MSE-IW	1.2547	1.0000	1.0073	1.0029	1.0119	1.1181	0.9620	**0.9495***	0.9780
QLIKE-RW	1.1937	1.0000	1.0023	0.9963	0.9893	1.0313	**0.9307**	0.9454	1.0318
QLIKE-IW	1.2894	1.0000	0.9959	0.9909	1.0000	1.1453	0.9452	**0.8932***	1.0143

Note: Model performance, expressed as model loss normalized by the loss of the HAR model. Each row reflects a combination of estimation window and loss function. Ratio for the best performing model on each row in bold. Corresponding asterix * and ** denote 1% and 5% confidence levels from Diebold-Mariano test for one-sided tests of superior performance of the best performing model compared to the HAR model.

Table 6 and Table 7 detail the out-of-sample performance for weekly and monthly forecasts, respectively. Notably, the HAR-J, CHAR, and SHAR models generally fail to demonstrate consistent improvements over the basic HAR model. This is a sharp contrast to the RQ-augmented models. The

HARQ-F model outperforms the HAR model both for EURUSD and USDJPY for nearly all instances. Also, HARQ-h delivers forecasts that are relatively consistent with the HAR model. Judging from both weekly and monthly results, the inherent flexibility of the HARQ-F is beneficial also for longer-term forecasts. We note that, at the monthly forecasting horizon for USDJPY, there is some variability as to preferred Q-specifications. Also, in some monthly instances, the Diebold-Mariano null hypothesis of equal predictability cannot be rejected. This is not unreasonable, since the number of independent monthly observations naturally becomes lower than for corresponding shorter forecasting horizons, leading to higher parameter uncertainty and related noise in volatility estimates.

3.3. Robustness

3.3.1. Alternative HARQ Specifications

The intention of the HARQ model is to capture the heteroskedastic measurement error of realized variance. The HARQ model in (5) approximates this through the square root of RQ. Bollerslev et al. (2016) argues that this encounters possible issues with numerical stability. Still, this specification is somewhat ad-hoc and a number of reasonable alternatives exist. To clarify whether the performance of the HARQ model is sensitive to the definition of RQ, we follow Bollerslev et al. (2016) and substitute $RQ, RQ^{-1/2}, RQ^{-1}$, and $\log(RQ)$ in place of $RQ^{1/2}$. Furthermore, we augment the standard HAR and HARQ models with $RQ^{1/2}$ as an additional explanatory variable, which allows the HAR(Q) model intercept to be time-varying.

Table 8 reports the out-of-sample forecast results from the alternative HARQ specifications. We normalize all losses by those of the HARQ model based on $RQ^{1/2}$.

Table 8. Alternative HARQ Specifications

	Alternative RQ transformations					*Adding $RQ^{1/2}$*	
EURUSD	RQ	$RQ^{1/2}$	$RQ^{-1/2}$	RQ^{-1}	$\log(RQ)$	HAR	HARQ
MSE-RW	1.0023	**1.0000**	1.0263	1.0246	1.0092	1.0309	1.0052
MSE-IW	1.0016	**1.0000**	1.0274	1.0265	1.0069	1.0292	1.0086
QLIKE-RW	1.0042	**1.0000**	1.0326	1.0304	1.0007	1.0250	1.0067
QLIKE-IW	1.0014	1.0000	1.0064	1.0254	**0.9937**	1.0044	1.0164
USDJPY	RQ	$RQ^{1/2}$	$RQ^{-1/2}$	RQ^{-1}	$\log(RQ)$	HAR	HARQ
MSE-RW	1.0001	**1.0000**	1.1345	1.1225	1.0516	1.1202	1.0118
MSE-IW	1.0049	1.0000	1.0606	1.0543	**0.9931**	1.0512	1.0186
QLIKE-RW	1.0097	1.0000	1.1439	1.1067	**0.9794**	1.0731	1.0455
QLIKE-IW	1.0188	1.0000	1.1105	1.0841	**0.9322**	1.0358	0.9989

Note: Model performance, expressed as model loss normalized by the loss of the HARQ model, relies on $RQ^{1/2}$. Each row reflects a combination of estimation window and loss function. Ratio for the best, performing model on each row in bold. The left panel reports the results based on alternative RQ interaction terms. The right panel reports the results from including $RQ^{1/2}$ as an explanatory variable.

The two rightmost columns of Table 8 reveal that including $RQ^{1/2}$ as an explanatory variable in the HAR and HARQ models does not lead to improved forecasts. Similarly, applying alternative RQ transformations does not appear to be helpful. Overall, we conclude that the HARQ model demonstrates greater stability and is generally favored over the alternative specifications.

3.3.2. Alternative Q-Models

HARQ is essentially an expansion of the HAR model. In a similar vein, the other benchmark volatility models can be extended accordingly. Following Bollerslev et al. (2016), from the HAR-J model defined in (3.2), we construct the HARQ-J model;

$$\begin{aligned} RV_t = &\beta_0 + \left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right) RV_{t-1} + \beta_2 RV_{t-1|t-5} \\ &+ \beta_3 RV_{t-1|t-22} + \beta_J J_{t-1} + u_t. \end{aligned} \tag{16}$$

Furthermore, from the CHAR model defined in (3.2), we construct the CHARQ model;

$$\begin{aligned} RV_t = &\beta_0 + \left(\beta_1 + \beta_{1Q} TPQ_{t-1}^{1/2}\right) BPV_{t-1} + \beta_2 BPV_{t-1|t-5} \\ &+ \beta_3 BPV_{t-1|t-22} + u_t. \end{aligned} \tag{17}$$

Lastly, from the SHAR model defined in (3.2), we construct the SHARQ model;

$$\begin{aligned} RV_t = &\beta_0 + \left(\beta_1^+ + \beta_{1Q}^+ RQ_{t-1}^{1/2}\right) RV_{t-1}^+ + \left(\beta_1^- + \beta_{1Q}^- RQ_{t-1}^{1/2}\right) RV_{t-1}^- \\ &+ \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t. \end{aligned} \tag{18}$$

Table 9 compares out-of-sample forecast results from each of the alternative Q-models (HARQ-J, CHARQ, and SHARQ), to their non-Q adjusted baseline specification. We also include the HARQ model. For both currencies, the enhancements seen in the HARQ-J and CHARQ models align with those observed in the basic HARQ model. This is in contrast to the SHARQ model, which is outperformed by SHAR. Bollerslev et al. (2016) report similar results.

Table 9. Out-of-sample forecast losses for alternative Q-models.

EURUSD	HARQ	HARQ-J	CHARQ	SHARQ
MSE-RW	0.9759	0.9693	0.9749	1.0613
MSE-IW	0.9742	0.9563	0.9567	1.0315
QLIKE-RW	0.9767	0.9845	0.9750	1.1473
QLIKE-IW	0.9952	0.9960	0.9893	0.9987
USDJPY	HARQ	HARQ-J	CHARQ	SHARQ
MSE-RW	0.8885	0.8916	0.8914	1.0953
MSE-IW	0.9446	0.9322	0.9389	0.8965
QLIKE-RW	0.8824	0.8471	0.9040	1.3887
QLIKE-IW	0.8949	0.8942	0.9178	0.8974

Note: Model performance, expressed as model loss normalized by the loss of the relevant baseline models without the Q-adjustment terms. Each row reflects a combination of estimation window and loss function. Ratio for the best performing model on each row in bold.

3.3.3. Subsample analysis

Recent history contains two independent events that separately have induced turbulence in the global macroeconomy and financial markets. One is the outbreak of COVID-19 in March 2020; another is the Russian invasion of Ukraine in the second half of 2022, as illustrated in Figure 1.

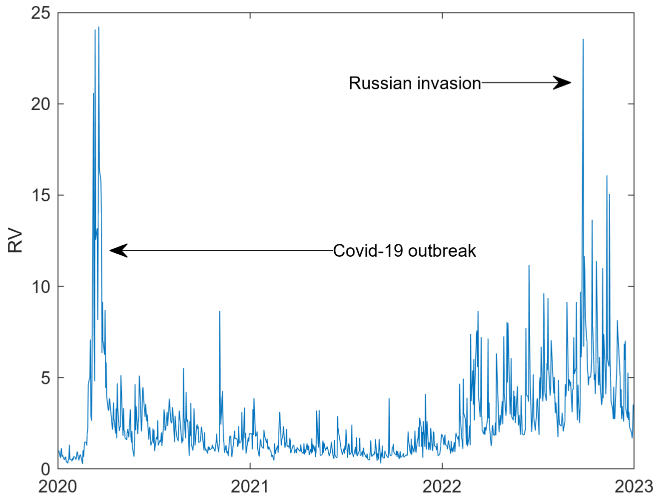

Figure 1. EURUSD realized variance

To analyze this period of extreme market conditions in isolation, we perform a sub-sample analysis covering 2020–2022. Table 10 contains out-of-sample results for day-ahead volatility forecasts. Reassuringly, the overall results remain intact, in that the HARQ-F model is the best performing model also when this extreme period is considered in isolation.

Table 10. Day ahead out-of-sample forecast losses, 2020–2022 subsample.

EURUSD	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.2522	1.0000	0.9781	0.9745	1.0041	1.0425	0.9517	**0.9304**
MSE-IW	1.2068	1.0000	0.9813	0.9764	0.9979	1.0976	0.9806	**0.9677**
QLIKE-RW	1.3216	1.0000	1.0169	0.9829	1.0093	1.1370	0.9446	**0.9065**
QLIKE-IW	1.5585	1.0000	1.0085	1.0119	1.0059	1.2338	0.9725	**0.9701**

USDJPY	AR	HAR	HAR-J	CHAR	SHAR	ARQ	HARQ	HARQ-F
MSE-RW	1.0930	1.0000	1.0555	0.9909	0.9822	0.9895	0.9564	**0.9348**
MSE-IW	1.1099	1.0000	0.9958	0.9850	1.0112	1.0523	1.0071	**0.9827**
QLIKE-RW	1.3404	1.0000	1.2635	1.0136	0.9845	0.9509	**0.8611**	0.8677
QLIKE-IW	1.4766	1.0000	0.9939	0.9808	1.0108	1.0231	0.8453	**0.7868**

Note: Model performance, expressed as model loss normalized by the loss of the HAR model. Each row reflects a combination of estimation window and loss function. Ratio for the best performing model on each row in bold.

4. Conclusions

This study uses updated tick-level data from two major currency pairs, EURUSD and USDJPY, covering January 2010 to December 2022, to investigate the relevance of realized quarticity for out-of-sample volatility forecasts. We find that realized quarticity effectively captures noise caused by measurement errors, as evidenced by increased precision in daily, weekly, and monthly volatility estimates from models augmented with realized quarticity as an additionally explanatory variable. These results are robust across estimation windows, evaluation metrics, and model specifications. As such, the results conform to comparable studies from other markets, predominantly on equity indices and single stocks. This paper also complements the relatively scarce body of literature on foreign exchange markets in this context.

A myriad of volatility models based on the HAR framework have been proposed. Still, simple linear HAR specifications have proven remarkably difficult to beat, as shown by Audrino et al. (2024) and Branco et al. (2024). In a recent survey, Gunnarsson et al. (2024) report promising results for machine learning models and volatility forecasting across asset classes. The FX implied volatility surface contains a rich set of relevant predictive information across forecasting horizons and quantiles (de Lange et al., 2022). Thus, combining implied volatilities and high-frequency data using machine learning models, along the lines of Blom et al. (2023), appears as an interesting avenue for future research.

Rarely, one single model dominates others in terms of statistical and economic criteria. To this end, investigating ensemble models where high-frequency models are combined with other volatility model classes, such as time series models and stochastic volatility models-possibly including jump-processes, should be of interest. The recently developed rough-path volatility models based on fractional Brownian motion (Salmon and SenGuptz, 2021; Bayer et al., 2023) appear particularly relevant in this context.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Author contributions

M.R.: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft, Writing - Review & Editing.

M.H.: Data Curation, Writing - Original Draft, Writing - Review & Editing.

Acknowledgments

We would like to thank Andrew Patton for making the Matlab code from Bollerslev et al. (2016) available at https://public.econ.duke.edu/ ap172/. Furthermore, we are grateful for insightful comments from the Editor and two anonymous reviewers, which helped us improve the paper.

Conflict of interest

The authors declare no conflicts of interest.

References

Andersen TG, Bollerslev T (1998) Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Int Econ Rev* 9: 885–905. https://doi.org/10.2307/2527343

Andersen TG, Bollerslev T, Diebold FX (2007) Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Rev Econ Stat* 4: 701–720. https://doi.org/10.1162/rest.89.4.701

Andersen TG, Bollerslev T, Diebold FX, et al. (2003) Modeling and forecasting realized volatility. *Econometrica* 2: 579–625. https://doi.org/10.1111/1468-0262.00418

Andersen TG, Bollerslev T, Christoffersen PF, et al. (2013) Financial risk measurement for financial risk management. *Handbook Econ Financ* 2: 1127–1220. https://doi.org/10.1016/B978-0-44-459406-8.00017-2

Andersen TG, Li Y, Todorov V, et al. (2023) Volatility measurement with pockets of extreme return persistence. *J Econometrics* 2: 105048. https://doi.org/10.1016/j.jeconom.2020.11.005

Audrino F, Chassot J (2024) HARd to Beat: The Overlooked Impact of Rolling Windows in the Era of Machine Learning. *arXiv Preprint* arXiv:2406.08041. https://doi.org/10.48550/arXiv.2406.08041

Barndorff-Nielsen OE, Hansen PR, Lunde A, et al. (2009) Realized kernels in practice: Trades and quotes. *Economet J*. https://doi.org/10.1111/j.1368-423X.2008.00275.x

Barndorff-Nielsen OE, Kinnebrok S, Lunde A, et al. (2009) Measuring downside risk-realised semivariance. *CREATES Res Pap*. 2008–42. http://dx.doi.org/10.2139/ssrn.1262194

Barndorff-Nielsen OE, Shephard N (2004) Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72: 885–925. https://doi.org/10.1111/j.1468-0262.2004.00515.x

Bayer Christian, Friz PK, Fukasawa M, et al. (2023) Rough volatility. SIAM

Blom HM, de Lange PE, Risstad M (2023) Estimating Value-at-Risk in the EURUSD Currency Cross from Implied Volatilities Using Machine Learning Methods and Quantile Regression. *J Risk Financ Manag* 16: 312. https://doi.org/10.3390/jrfm16070312

Bollerslev T, Patton AJ, Quaedvlieg R (2016) Exploiting the errors: A simple approach for improved volatility forecasting. *J Econom* 192: 1–18. https://doi.org/10.1016/j.jeconom.2015.10.007

Branco R, Rubesam A, Zevallos M (2024) Forecasting realized volatility: Does anything beat linear models? *J Empir Financ* 2024: 101524. https://doi.org/10.1016/j.jempfin.2024.101524

Clements A, Preve D (2021) A practical guide to harnessing the HAR volatility model. *J Bank Financ* 133: 106285. https://doi.org/10.1016/j.jbankfin.2021.106285

Corsi F (2009) A simple approximate long-memory model of realized volatility. *J Financ Economet* 7: 174–196. https://doi.org/10.1093/jjfinec/nbp001

Götz P (2023) Realized quantity extended conditional autoregressive Value-at-Risk models. *J Risk* 26.

Liang C, Li Y, Ma F, et al. (2022) Forecasting international equity market volatility: A new approach. *J Forecast* 41: 1433–1457.

Gunnarsson ES, Isern HR, Kaloudis A, et al. (2024) Prediction of realized volatility and implied volatility indices using AI and machine learning: A review *Int Rev Financ Anal* 2024: 103221 https://doi.org/10.1016/j.irfa.2024.103221

de Lange PE, Risstad M, Westgaard S (2022) Estimating value-at-risk using quantile regression and implied volatilities. *J Risk Model Validat*.

Liu LY, Patton AJ, Sheppard K (2015) Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *J Econom* 187: 293–311. https://doi.org/10.1016/j.jeconom.2015.02.008

Liu G, Wei Y, Chen Y, et al. (2018) Forecasting the value-at-risk of Chinese stock market using the HARQ model and extreme value theory. *Physica A* 499: 288–297. https://doi.org/10.1016/j.physa.2018.02.033

Lyócsa S, Molnár P, Fedorko I (2016) Forecasting exchange rate volatility: The case of the Czech Republic, Hungary and Poland. *Financ Uver* 66: 453–463.

Lyócsa S, Plíhal T, Vỳrost T (2024) Forecasting day-ahead expected shortfall on the EUR/USD exchange rate: The (I) relevance of implied volatility. *Int J Forecast*. https://doi.org/10.1016/j.ijforecast.2023.11.003

Ma F, Wahab M, Zhang Y (2019) Forecasting the US stock volatility: An aligned jump index from G7 stock markets. *Pac-Basin Financ J* 54: 132–146. https://doi.org/10.1016/j.pacfin.2019.02.006

Newey WK, West KD (1987) Hypothesis testing with efficient method of moments estimation. *Int Econom Rev* 54: 777–787. https://doi.org/10.2307/2526578

Pascalau R, Poirier R (2023) Increasing the information content of realized volatility forecasts. *J Financ Econom* 21: 1064–1098. https://doi.org/10.1093/jjfinec/nbab028

Patton AJ (2011) Volatility forecast comparison using imperfect volatility proxies. *J Econom* 160: 246–256. https://doi.org/10.1016/j.jeconom.2010.03.034

Patton AJ, Sheppard K (2015) Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Rev Econom Stat* 97: 683–697. https://doi.org/10.1162/REST$_{a0}$0503

Plíhal Tomáš, Lyócsa Štefan (2021) Modeling realized volatility of the EUR/USD exchange rate: Does implied volatility really matter? *Int Rev Econom Financ* 71: 811–829. https://doi.org/10.1016/j.iref.2020.10.001

Qiu Y, Wang Z, Xie T, et al. (2021) Forecasting Bitcoin realized volatility by exploiting measurement error under model uncertainty. *J Empir Financ* 62: 179–201. https://doi.org/10.1016/j.jempfin.2021.03.003

Qu H, Duan Q, Niu M (2018) Modeling the volatility of realized volatility to improve volatility forecasts in electricity markets. *Energ Econ* 74: 767–776. https://doi.org/10.1016/j.eneco.2018.07.033

Rokicka A, Kudła J (2020) Modeling Realized Volatility with Implied Volatility for the EUR/GBP Exchange Rate markets. *J Risk* 23.

Risstad M, Thodesen A, Thune KA, et al. (2023) On the Exchange Rate Dynamics of the Norwegian Krone. *J Risk Financ Manag* 16: 308. https://doi.org/10.3390/jrfm16070308

Shen D, Urquhart A, Wang P (2020) Forecasting the volatility of Bitcoin: The importance of jumps and structural breaks. *Eur Financ Manag* 26: 1294–1323. https://doi.org/10.1111/eufm.12254

Salmon N, SenGupta I (2021) Fractional Barndorff-Nielsen and Shephard model: Applications in variance and volatility swaps, and hedging. *Ann Financ* 17: 529–558. https://doi.org/10.1007/s10436-021-00394-4

Sizova N (2011) Integrated variance forecasting: Model based vs. reduced form. *J Econom* 162: 294–311. https://doi.org/10.1016/j.jeconom.2011.02.004

Swanson NR, White H (1997) Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. *Int J Forecast* 13: 439–461. https://doi.org/10.1016/S0169-2070(97)00030-7

Wang Y, Liang F, Wang T, et al. (2020) Does measurement error matter in volatility forecasting? Empirical evidence from the Chinese stock market. *Econ Model* 87: 148–157. https://doi.org/10.1016/j.econmod.2019.07.014

White H (1980) DA heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 1980: 817–838. https://doi.org/10.2307/1912934

AIMS Press