International Journal of Forecasting 39 (2023) 1221–1237

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Improving variance forecasts: The role of Realized Variance features[☆]

Ioannis Papantonis [a,b,*], Leonidas Rompolis [b], Elias Tzavalis [b]

[a] *Bank of England, United Kingdom*
[b] *Athens University of Economics & Business, Greece*

ARTICLE INFO

Keywords:
Realized Variance
Forecasting
EGARCH
Heterogeneous Autoregressive
HAR
MLE

ABSTRACT

In this paper, we effectively extend the Realized-EGARCH (R-EGARCH) framework by allowing the conditional variance process to incorporate exogenous variates related to different observable features of Realized Variance (RV). The choice of these features is well motivated by recent studies on the Heterogeneous Autoregressive (HAR) class of models. We examine several specifications nested within our augmented R-EGARCH representation, and we find that they perform significantly better than the standard R-EGARCH model. These specifications incorporate realized semi-variances, heterogeneous long-memory effects of RV, and jump variation. We also show that the performance of our framework further improves if we allow for skewness and excess kurtosis for asset return innovations, instead of assuming normality. This can better filter the true distribution of the return innovations, and thus can more accurately estimate their effects on the variance process. This is also supported by a Monte Carlo simulation exercise executed in the paper.

1. Introduction

Modeling and forecasting volatility have always been at the core of modern financial economics and econometrics. The first generation of volatility models developed, known since Engle (1982) and Bollerslev (1986) as ARCH/GARCH models, treated variance as an unobservable (latent) variable that can be filtered through the conditional distribution of asset returns.

Recently, there has been growing research interest in extending the aforementioned class of models by using as exogenous regressors measures of Realized Variance (RV) constructed from high-frequency (intra-day) data. Pioneering works of this second generation of models include those of Engle (2002) and Engle and Gallo (2006), who use RV within a multiplicative error model (MEM) context, as well as the high-frequency volatility (HEAVY) model of Shephard and Sheppard (2010) for jointly describing latent and observable variance dynamics.[1] The statistical and economic gains from incorporating RV variates in ARCH/GARCH models are emphasized in numerous studies.[2]

☆ We would like to thank the editor, Dick van Dijk, as well as the associate editor and the two anonymous reviewers for providing useful comments and suggestions on an earlier version of the paper. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee, Financial Policy Committee or Prudential Regulation Committee.

* Corresponding author at: Bank of England, United Kingdom.
E-mail addresses: Ioannis.Papantonis@bankofengland.co.uk (I. Papantonis), rompolis@aueb.gr (L. Rompolis), e.tzavalis@aueb.gr (E. Tzavalis).

1 As Cipollini et al. (2012) discuss, the HEAVY model can also be a special case of a bivariate vector MEM. The framework that they suggest is more general than HEAVY in that it introduces the possibility of testing for dynamic feedback effects.

2 Several authors report significant gains for option valuation after incorporating information from realized variance; see, for instance, Christoffersen et al. (2014, 2010), Corsi et al. (2013), Feunou and Okou

https://doi.org/10.1016/j.ijforecast.2022.05.006

An appropriate econometric framework for studying these extensions to the ARCH/GARCH class was suggested by Hansen et al. (2012), known as the Realized-GARCH (R-GARCH). This framework provides a joint structure for modeling the dynamics of returns and conditional (or realized) variance, by introducing a measurement equation for variance in a system directly linking the observable RV dynamics to the latent conditional volatility.[3] The R-GARCH framework preserves some of the interesting properties of the standard GARCH and, therefore, can be applied for option pricing and Value-at-Risk, inter alia. Furthermore, it can explicitly model the impact of return news on volatility, the so-called leverage effect. At the same time, it allows for evaluating the ability of conditional volatility models to predict realized variance.

In this paper, we extend the R-GARCH framework to incorporate different variants of RV in the variance process. These include the following: (i) realized semi-variances capturing asymmetries in variance arising from negative skewness in asset returns due to upside/downside daily variations, (ii) moving-average terms of RV accounting for a long-memory pattern in the conditional variance process (as also suggested in Huang et al. (2016)), and (iii) a measure of jump variation capturing the discontinuous component of quadratic variation. The choice of these exogenous variables can be well justified from the literature on heterogeneous autoregressive (HAR) models, popularized by Corsi (2009), showing that their incorporation in a HAR-type regression can enhance not only the in-sample explanatory power but also the out-of-sample forecasting performance of this class of models.

We employ a flexible exponential model class to encapsulate the information from the above variables in the conditional variance function. This nests the standard realized EGARCH (R-EGARCH) model of Hansen and Huang (2016) as a special case. Working with an exponential representation for the conditional variance is beneficial and provides a more flexible framework for modeling variance dynamics, as it allows for distinguishing between the sign and magnitude effects of asset return innovations and incorporating exogenous variates in the variance process without imposing any restrictions on the parameters of the model. These properties are also discussed in Hansen and Huang (2016) and Borup and Jakobsen (2019).

In our analysis, we also account for asymmetric/fat-tailed distributions of asset return innovations, which can be shown to more accurately filter the return innovation terms and help to better identify their effects on the variance process. This finding is further supported by our empirical results, as well as by a Monte Carlo simulation exercise carried out in the paper. Our suggested framework encompasses the basic HAR model and its various extensions suggested in the literature. This enables us not only to appraise the importance of jointly modeling the dynamics of asset returns and variance, which is behind the motivation of the standard R-EGARCH framework, but also to examine the impact of the different distributional assumptions for the return innovations on realized variance. We empirically compare a wide variety of model specifications nested within our augmented R-EGARCH framework. Our results lead to several interesting observations. Specifically, we uncover two separate channels of asymmetry in the conditional variance function that are both found to be significant in modeling the variance process: one originates from the responses of variance to lagged asset return innovations, and the other is due to the asymmetric effects arising from the dispersion between the daily realized semi-variances. The heterogeneous (moving-average) RV terms and the jump variation components that we consider further improve the performance of the R-EGARCH model. We find that the heterogeneous terms, together with the autoregressive component of the conditional variance, can better capture the long-term properties of the variance process. On the contrary, the jump variation component has a very short-lived effect on variance.

Regarding the performance of our framework compared to HAR-type model extensions, our results emphasize that incorporating lagged values of the conditional variance function, as well as filtered return innovations (instead of lagged returns, as in Corsi and Renò (2012)), in order to capture leverage effects can considerably improve the ability of the models to forecast realized variance. This adds support to modeling asset returns and variance dynamics, jointly. The out-of-sample results presented in the paper provide further support for these findings.

The remainder of the paper is organized as follows. We start by introducing the generic framework for the joint dynamics of returns and variance, followed by a description of the exponential variance model class and the several specifications nested within. We continue by presenting details on the joint estimation approach that we follow, along with results from a Monte Carlo simulation exercise examining the bias and efficiency of the estimated volatility model parameters under different distributional assumptions for the asset returns. Next, we provide a thorough comparative discussion of both the in-sample fit and the out-of-sample forecasting performance of the models. Finally, we conclude the paper by summarizing our findings.

2. The generalized R-EGARCH-SHARJ framework

Let us write the continuously compounded daily returns as $R_t = p_t - p_{t-1}$, where $p_t = \log(P_t)$ is the log price of an observed price process $\{P_t\}$ of a financial asset. We denote the conditional mean and variance of returns as $m_t \equiv E[R_t|\mathcal{F}_{t-1}]$ and $\upsilon_t \equiv Var[R_t|\mathcal{F}_{t-1}]$, respectively, where $\mathcal{F}_{t-1}$ defines a σ-algebra formed by the information set on the observed variables up to time $t-1$. We can also think of the conditional variance function as $\upsilon_t = E\left[\langle p, p\rangle_{t-1,t} |\mathcal{F}_{t-1}\right]$, where $\langle p, p\rangle_{t-1,t}$ is the quadratic variation of the log-price process $\{p_t\}$ defined on a daily fixed time interval $(t-1, t)$. As follows, the log-quadratic variation $\log\langle p, p\rangle_{t-1,t}$ has a conditional mean

(2019), and Huang et al. (2017), among others. Other academics and practitioners demonstrate that realized variance can lead to improved Value-at-Risk estimates; see Louzis et al. (2014) and the references therein.

[3] Takahashi et al. (2009) used this framework in the context of stochastic volatility modeling.

$E\left[\log\langle p,p\rangle_{t-1,t}|\mathcal{F}_{t-1}\right] = \log E\left[\langle p,p\rangle_{t-1,t}|\mathcal{F}_{t-1}\right] + o \simeq \log v_t \equiv h_t$, where the convexity term o is negligible. Since $\log\langle p,p\rangle_{t-1,t}$ is unobservable, a natural measure used in the literature to proxy it is the logarithm of realized variance RV_t, i.e. $\log RV_t$, which is a consistent estimator of $\log\langle p,p\rangle_{t-1,t}$; see Barndorff-Nielsen and Shephard (2002a, 2002b, 2004a, 2004b, 2006).

We consider a flexible exponential model class for describing the joint dynamics of returns and log variance, which in a generic form can be written as follows:

$$R_t = m_t + \sqrt{v_t}z_t,\ \text{ where } v_t = \exp(h_t) \tag{1a}$$

$$\log RV_t = h_t + \sigma u_t, \tag{1b}$$
$$\text{with } h_t = \beta_0 + \beta_h h_{t-1} + \tau(t-1) + x(t-1).$$

In the above, z_t and u_t are jointly $\sim i.i.d.(0,1)$; i.e. there are no cross-correlations or serial correlations. In the conditional log-variance function, the term $\tau(t)$ represents the news impact function (NIF), and $x(t)$ captures the response of log variance to exogenous realized variates, as defined below. Note that the conditional variance of the log-quadratic variation $\sigma^2 = Var\left[\log\langle p,p\rangle_{t-1,t}|\mathcal{F}_{t-1}\right]$ is considered constant.[4] We intentionally leave the densities governing z_t and u_t unspecified for now, in order to maintain generality. However, we return to this point in the following sections, as we intend to highlight the clear benefits of a more flexible density specification, especially for volatility forecasting.

The exponential class given by Eqs. (1a) and (1b) constitutes a framework that assumes a one-to-one relation between $\log RV_t$ and h_t. This parsimonious relation can ensure robustness and accuracy of the estimation procedure, as it helps the parameter identification by avoiding over-parameterization and is a very common choice among econometricians (see also Bekierman and Gribisch (2016), Hansen and Huang (2016), Koopman and Scharth (2012), Takahashi et al. (2009), among others). For the conditional mean of returns, we follow the standard approach in the literature (see Engle et al. (1987)) and we assume that $m_t = r_f + \mu\sqrt{v_t}$, which incorporates the risk-free rate r_f, as well as a volatility-in-mean effect capturing a time-varying risk premium.

To complete the specification of the model, we assume that the NIF $\tau(t)$ has an equivalent form to that of the standard EGARCH (Nelson, 1991). That is,

$$\tau(t) = \beta_z z_t + \beta_{|z|}(|z_t| - E[|z_t|]), \tag{1c}$$

where the two components of $\tau(t)$ allow variance to respond asymmetrically to lagged return innovations, a stylized fact also known as the leverage effect.[5] For the component $x(t)$, which captures the response of log variance to exogenous realized variates, we assume that it is driven by a set of variables employed in the ongoing literature studying HAR-type RV regressions:

$$x(t) = \gamma_D^+\log RV_t^+ + \gamma_D^-\log RV_t^- + \gamma_W\log RV_t^{[5]}$$
$$+ \gamma_M\log RV_t^{[20]} + \gamma_J\log RJ_t. \tag{1d}$$

The definition of these variables, and the justification for their inclusion in $x(t)$, is based on the following intuition.[6]

Firstly, the signed variables $\log RV_t^+$ and $\log RV_t^-$ respectively denote the upside and downside log realized semi-variances. The differences between $\log RV_t^+$ and $\log RV_t^-$ are directly related to asymmetries/skewness in intra-day return distributions. These asymmetries can be attributed primarily to the presence of jumps in the process of returns and/or variance; among other sources, see Audrino and Hu (2016). Note that in the absence of skewness in returns, the two semi-variances are almost identical (i.e. $\log RV_t^+ = \log RV_t^-$), which implies that the so-called signed jump variation arising from negative/positive jumps (in other studies defined as the semi-variance differential $RV_t^+ - RV_t^-$) shrinks to zero. If this is the case, then the variance Eq. (1d) can be simplified to consider only $\log RV_t$, with a corresponding coefficient γ_D. Using $\log RV_t^+$ and $\log RV_t^-$ in the R-EGARCH framework (instead of relying solely on $\log RV_t$) can be beneficial for the variance process, as $\log RV_t$ alone may omit important information by blending the two semi-variance components. Furthermore, incorporating the semi-variances in $x(t)$ can show whether these two variables have additional/different information with respect to NIF $\tau(t)$ regarding the asymmetric responses of variance to return innovations. On statistical grounds, the inclusion of $\log RV_t^+$ and $\log RV_t^-$ in $x(t)$ can be justified by the work of Patton and Sheppard (2015) on the HAR framework, providing clear evidence that using these variables instead of $\log RV_t$ is beneficial for forecasting realized variance.

Secondly, the variables $\log RV^{[5]}$ and $\log RV^{[20]}$ denote the 5- and 20-day moving-average terms of $\log RV$.[7] These terms have been extensively used in the HAR regression framework to approximate in a parsimonious way the long-memory dependence structure of realized variance, which would otherwise require non-trivial parametric fractionally integrated processes. Their inclusion in $x(t)$ can help to mimic this structure exogenously (as shown in Huang et al. (2016)). In a Realized GARCH context, Huang et al. (2016) were the first to use these variables to improve upon the forecasting ability of the GARCH. A more recent study by Baillie et al. (2019) provides a detailed discussion of the long-memory property of realized variance and its connection with the heterogeneous terms.

Finally, the variable $\log RJ_t$ is a measure of relative jump variation, capturing effects from the discontinuous

4 As in other studies, we can think that $\log\langle p,p\rangle_{t-1,t}$ follows a normal distribution, so that the convexity term o equals $-\frac{1}{2}\sigma^2$; unreported evidence indicates that the empirical results remain robust to its inclusion in the conditional mean of log variance. We also find the time variation of the variance of log variance, i.e. σ^2, to be negligible. Hence, accounting for time variation in σ is unnecessary and beyond the scope of this paper.

5 Regardless of the distribution of z_t, the expectation $E[|z_t|]$ will be (or converge to) a constant. Therefore, we ignore it in the estimation procedure, as this is absorbed by the constant coefficient ω.

6 A more detailed description of the statistical background of the realized estimators and the way they are constructed from high-frequency data can be found in Appendix.

7 The κ-day moving average terms are obtained as $\log RV_t^{[\kappa]} \equiv \log RV_{(t+1-\kappa:t)} = \kappa^{-1}\sum_{j=1}^{\kappa}\log RV_{t+1-j}$.

part of the log-quadratic variation $\log \langle p, p\rangle_{t-1,t}$. This can be attributed to jumps in either the returns or the underlying variance process. As shown by the estimates of the HAR regressions provided by Andersen et al. (2007) and Patton and Sheppard (2015), the inclusion of a jump measure may lead to improvements in short-term forecasting of future levels of variance, given that it represents a less persistent component of $\log \langle p, p\rangle_{t-1,t}$ (see also Bollerslev et al. (2009)). These improvements can be important for pricing options at short maturities. In the context of our suggested framework given by Eqs. (1a)–(1d), the inclusion of $\log RJ_t$ in $x(t)$ can reveal whether the discontinuous component of the variation has a significant impact on h_t, beyond the asymmetric effects captured by the semi-variances $\log RV_t^+$ and $\log RV_t^-$ and/or the NIF components in $\tau(t)$.[8]

2.1. Nested specifications

The general framework given by the system of Eqs. (1a)–(1d) nests several specifications that are widely used for modeling conditional and/or realized variance dynamics. In fact, it encompasses the standard HAR model for log realized variance, as well as its various extensions previously suggested in the literature. All of the specifications within the (1a)–(1d) framework can be evaluated by comparing the values of their joint likelihood functions or other goodness-of-fit and/or prediction performance metrics.

For $b_h = b_z = b_{|z|} = \gamma_D^+ = \gamma_D^- = \gamma_J = 0$, we obtain the HAR model of Corsi (2009) for log RV. Recall that if $RV_t^+ = RV_t^-$, then it is natural to replace the semi-variances with $\log RV_t$ in the $x(t)$ equation; i.e. we consider $\gamma_D^+ = \gamma_D^- = 0$ and $\gamma_D \neq 0$ for identification purposes. Also, for $\gamma_D^+ = \gamma_D^- = \gamma_W = \gamma_M = \gamma_J = 0$ we obtain the Realized-EGARCH (R-EGARCH) model described in Hansen and Huang (2016). The extensions of the R-EGARCH that include the heterogeneous components $RV^{[5]}$ and $RV^{[20]}$, can be obtained for $\gamma_D^+ = \gamma_D^- = \gamma_J = 0$ (denoted as R-EGARCH-HAR), while those that include the semi-variances RV_t^+ and RV_t^- can be obtained for $\gamma_J = 0$ (denoted as R-EGARCH-SHAR). Finally, we denote as R-GARCH-SHARJ our "full" specification, which also incorporates the jump variation measure $\log RJ_t$ in (1d). We consider this specification our benchmark for future model comparisons. All of the models discussed here are summarized in Table 1 (Panel A).

The HAR class of models can be obtained from the R-GARCH-SHARJ by setting restrictions on its beta slope coefficients. In particular, for $\beta_h = \beta_z = \beta_{|z|} = 0$ and $\gamma_J = 0$, we obtain the HAR extension with the semi-variances $\log RV_t^+$ and $\log RV_t^-$ (known as SHAR since Patton and Sheppard (2015)), while for $\beta_h = \beta_z = \beta_{|z|} = 0$ and $\gamma_D^+ = \gamma_D^- = 0$, we obtain the extension of the HAR with jumps (denoted as HARJ). Another interesting extension of the HAR model studied in the literature is the one by Bollerslev et al. (2009) and Corsi and Renò (2012) allowing for leverage effects. In the standard HAR literature, leverage can be captured by including lagged (signed) returns, i.e. by using R_{t-1} and/or $|R_{t-1}|$ in the realized variance equation, similarly applied by Audrino and Hu (2016). However, in our R-EGARCH-SHARJ framework, these effects can be captured by the NIF $\tau(t)$ through the return innovations z_t and $|z_t|$, instead of the raw returns R_t and $|R_t|$. Capturing the leverage effects through the filtered innovation terms, rather than the returns themselves, can lead to a better identification of the true effects of leverage on volatility; i.e. the net of any variance-in-mean or other latent premia effects that might be embedded in the returns process. As follows, a natural extension of the HAR model towards the above directions is to consider conditional information from the NIF (denoted as SHARJL), which may lead to important gains in forecasting realized variance. Again, refer to Table 1 (Panel B) for a summary of the nested HAR-type model extensions.

3. Empirical analysis

In this section, we first present the econometric method to estimate the R-EGARCH-SHARJ model and the different specifications nested within, as discussed in Section 2. Next, we present our dataset and discuss the estimation results. We provide a thorough in-sample and out-of-sample analysis. To evaluate the performance of the models, we use a number of goodness-of-fit and prediction performance metrics for the log realized variance.

3.1. Joint maximum likelihood estimation

Let $\vartheta = (\beta_0, \beta_h, \beta_z, \beta_{|z|}, \gamma_D^+, \gamma_D^-, \gamma_W, \gamma_M, \gamma_J)'$ denote the parameter vector of the R-EGARCH-SHARJ model, described by Eqs. (1a)–(1d). To estimate ϑ, we employ a standard joint maximum likelihood estimation (MLE) approach. This can be applied under different assumptions for the probability density functions (PDFs) of the returns R_t and log realized variance $\log RV_t$, which (conditional on $\mathcal{F}_{t-1}$) can be denoted $f_t^R(R_t|\mathcal{F}_{t-1};\vartheta)$ and $f_t^V(\log RV_t|\mathcal{F}_{t-1};\vartheta)$, respectively. The joint-MLE method is a very straightforward and robust way to obtain parameter estimates and has been used in numerous similar applications (see Hansen et al. (2012), Hansen and Huang (2016), Papantonis (2016), and the references therein). The assumption that the two error terms z_t and u_t are jointly i.i.d. implies that the joint log-likelihood function of R_t and $\log RV_t$ conditional on $\mathcal{F}_{t-1}$ can be written as

$$\begin{aligned} &\mathcal{L}(R_t, \log RV_t|\mathcal{F}_{t-1};\vartheta) \qquad (2)\\ &\equiv \mathcal{L}^R(R_t|\mathcal{F}_{t-1};\vartheta) + \mathcal{L}^V(\log RV_t|\mathcal{F}_{t-1};\vartheta)\\ &= \sum_{t=1}^{T} \log f_t^R(R_t|\mathcal{F}_{t-1};\vartheta) + \sum_{t=1}^{T} \log f_t^V(\log RV_t|\mathcal{F}_{t-1};\vartheta) \end{aligned}$$

This means that the total log-likelihood $\mathcal{L}(R_t, \log RV_t|\mathcal{F}_{t-1};\vartheta)$, simply abbreviated by $\mathcal{L}$, is the sum of two independent log-likelihood components for the returns R_t and log variance $\log RV_t$: $\mathcal{L}^R$ and $\mathcal{L}^V$, respectively. The parameter vector ϑ can be estimated by a numerical optimization of the non-linear maximization problem:

$$\hat{\vartheta} = \arg\max_{\vartheta} \mathcal{L}$$

[8] Huang and Tauchen (2005) demonstrate for the S&P 500 index that the relative jump variation can account for approximately 7% of the total variation of the index returns.

Table 1
Nested models within the R-EGARCH-SHARJ class.

R-EGARCH-SHARJ	$h_t = \beta_0 + \beta_h h_{t-1} + \beta_z z_t + \beta_{\|z\|}(\|z_t\| - E[\|z_t\|]) + x(t)$	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$
Panel A: R-EGARCH extensions		
R-EGARCH	#	$x(t) = \gamma_D \log RV_t$, if $\log RV_t^+ = \log RV_t^-$
R-EGARCH-S	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^-$
R-EGARCH-HAR	#	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
R-EGARCH-SHAR	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
R-EGARCH-HARJ	#	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$
Panel B: HAR extensions		
HAR	$h_t = \beta_0 + x(t)$	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
HARJ	#	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$
SHAR	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
SHARJ	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$
HARL	$h_t = \beta_0 + \beta_z z_t + \beta_{\|z\|}(\|z_t\| - E[\|z_t\|]) + x(t)$	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
SHARL	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]}$
HARJL	#	$x(t) = \gamma_D \log RV_t + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$
SHARJL	#	$x(t) = \gamma_D^+ \log RV_t^+ + \gamma_D^- \log RV_t^- + \gamma_W \log RV_t^{[5]} + \gamma_M \log RV_t^{[20]} + \gamma_J \log RJ_t$

Notes: This table shows the various models used in the paper. In Panel A, we present extensions to the R-EGARCH of Hansen and Huang (2016) nested within the R-EGARCH-SHARJ framework. In Panel B, we show several extensions to the HAR model class that can also be represented as nested models within the R-EGARCH-SHARJ framework. The first column of the table reports the abbreviation of each model used in the text. The second column presents the specification of log conditional variance h_t, where z_t is the innovation term of the mean equation following an *i.i.d.*(0, 1) process. The third column shows the specification of x_t, which captures the response of h_t to a set of exogenous realized measures. In particular, $\log RV_t$ is the daily log realized variance, $\log RV_t^+$ and $\log RV_t^-$ denote daily upside and downside log realized semi-variance, respectively, $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$ are the 5- and 20-day moving average terms of $\log RV$, respectively, and $\log RJ_t$ denotes the discontinuous component of $\log RV$. Appendix A presents the definitions of these measures and their estimators relying on high-frequency data.

We report the values of both $\mathcal{L}^R$ and $\mathcal{L}^V$ components in order to get a more complete picture of how each of them contributes to the total likelihood. Additionally, we estimate the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) as alternative measures of assessing the goodness of fit of the different models to the actual data.[9]

3.1.1. Misspecification of the return distribution and the impact on variance

Efficient estimation of vector ϑ by the MLE approach requires correctly specifying the PDFs for returns and log variance, $f_t^R(R_t|\mathcal{F}_{t-1};\vartheta)$ and $f_t^V(\log RV_t|\mathcal{F}_{t-1};\vartheta)$. Specifically, ignoring the prevalent asymmetries and fat-tails in the return distribution may lead to inaccurate or inefficient ϑ estimates, therefore influencing the conditional variance process h_t. Since there is an interaction between the conditional variance h_t and the return innovations z_t, this can lead to poor filtering of z_t, which in turn can result in an inaccurate estimation of the NIF function $\tau(t)$ governing the conditional variance dynamics.

Responding to pronounced evidence that asset returns are characterized by high levels of skewness and excess kurtosis (leading us to reject the normality assumption for z_t), we consider that $f_t^R(R_t|\mathcal{F}_{t-1};\vartheta)$ has a normal-inverse Gaussian (NIG) density. The NIG was popularized in the financial econometrics literature by Barndorff-Nielsen (1978), and since then it has been found in many studies to outperform other parametric densities when fitting asset returns.[10] One of the most interesting properties of the NIG has to do with its ability to span a significantly wider domain of theoretically feasible skewness–kurtosis combinations, compared to other parametric densities (Jondeau & Rockinger, 2003). Furthermore, it nests a wide range of popular parametric PDFs (including the normal PDF) as special cases. For the log realized variance, we consider that $f_t^V(\log RV_t|\mathcal{F}_{t-1};\vartheta)$ follows a normal density with constant variance. This assumption has been followed in a few other studies, including Andersen et al. (2001, 2003), and can be justified by the empirical properties of our actual log realized variance time series. As noted by Bollerslev et al. (2009), the logarithmic function of realized variance absorbs almost any heteroscedasticity in the error term u_t of the variance equation. This issue is further explored in a recent study by Cipollini et al. (2021).

To investigate the effects of a possibly misspecified return PDF on the parameter vector of h_t and, in particular, on the NIF, which depends on the innovation terms z_t, we carried out a Monte Carlo (MC) simulation study. We generated data for a standard EGARCH model

9 The standard errors of the parameter estimator $\hat{\vartheta}$ are calculated from the inverse of the information matrix. The variance–covariance matrix of the MLE estimator can be written as follows:

$$\mathrm{var}(\hat{\vartheta}) = \left[\mathbf{I}(\hat{\vartheta})\right]^{-1} = \left(-E\left[\mathbf{H}(\hat{\vartheta})\right]\right)^{-1} = \left(-E\left[\frac{\partial^2 \mathcal{L}(\hat{\vartheta})}{\partial\hat{\vartheta}\partial\hat{\vartheta}'}\right]\right)^{-1}$$

where the Hessian matrix $\mathbf{H}$ is obtained numerically. As follows, the standard errors of $\hat{\vartheta}$ are simply obtained as the square root of the diagonal elements of the above covariance matrix.

10 The NIG has been shown to perform better than conventional parametric densities in the context of stochastic volatility models (Andersson, 2001; Barndorff-Nielsen, 1997) as well as discrete-time conditional volatility models (Forsberg & Bollerslev, 2002; Jensen & Lunde, 2001). It has also been used to produce more reliable VaR estimates (Venter & de Jongh, 2002; Wilhelmsson, 2009) and to boost the accuracy of option pricing models (Chorro et al., 2012; Eriksson et al., 2009; Ghysels & Wang, 2014; Stentoft, 2008).

under different distributional assumptions for z_t, to produce different combinations of skewness (Sk) and kurtosis (Ku) levels, and we estimated the model based on the MLE method under different densities for z_t. Specifically, we considered that z_t is randomly drawn from either an NIG or a normal distribution.[11] The values that we chose for the true parameter vector ϑ of our simulations were based on estimates of the EGARCH model on actual data. Similarly, for the levels of the skewness and kurtosis coefficients we selected values very close to those found in our empirical analysis section; precisely, we used $Sk = -0.32$ and $Ku = 3.20$. For comparison and robustness purposes, we also examined a theoretical case with more pronounced levels of skewness and kurtosis, assuming $Sk = -0.60$ and $Ku = 3.80$. The simulation results along with more details about the design of our MC study can be found in Table 2. The table presents the standard deviation (std), root mean squared error (RMSE), and bias of the estimated conditional log-variance function coefficients (namely for β_0, β_h, β_z, and $\beta_{|z|}$) resulting from all MC iterations. These were estimated under both NIG and normal assumptions for the innovation term z_t. For the NIG, we also present the results for the Sk and Ku coefficients, calculated based on the estimates of the distributional shape parameters.

A few interesting conclusions can be drawn from analyzing the MC results in Table 2. First, we observe that when the underlying true distribution of innovations z_t is the NIG, relying on the normal assumption for the estimation will increase the bias and inefficiency of the MLE parameters of the conditional variance process—especially those of the NIF $\tau(t)$, as they are the ones reflecting the sign and magnitude effects of z_t on h_t (through the slope coefficients β_z and $\beta_{|z|}$). Note that the size of the bias and inefficiency of the parameters will increase for more pronounced skewness/kurtosis levels, i.e. for $Sk = -0.60$ and $Ku = 3.80$. Also, the negative and positive values for the bias of β_z and $\beta_{|z|}$, respectively, indicate that for negative values of z_t, the estimates of $\tau(t)$ will tend to overstate the true effects of z_t and $|z_t|$ on h_t.

This intuition is graphically confirmed by the plots in Fig. 1, which presents the true NIF $\tau(t)$ (for $Sk = -0.32$ and $Ku = 3.20$) against those estimated under the NIG and the normal assumptions. These graphs reveal that (for $z_t \in (-\infty, 0)$) the MLE estimates of $\tau(t)$ under the normal assumption provide an upward-shifted NIF, compared to the true one generated under the NIG. On the other hand, the estimated $\tau(t)$ under the NIG tracks the true NIF very closely, across the whole spectrum of z_t. The overestimation of $\tau(t)$ can be clearly attributed to the prevalent skewness and kurtosis in z_t, which is omitted by the normal distribution. This will tend to magnify the effects of z_t on h_t, by boosting the values of β_z and $\beta_{|z|}$.

Another interesting conclusion that can be drawn from our simulation results is that the MLE approach with an NIG assumption for z_t can efficiently estimate the true parameter vector of the model, even if the true underlying distribution of z_t is normal (refer to the case of $Sk = 0$ and $Ku = 3.0$ in the table). The levels of bias, standard deviation, and RMSE reported in the table for this particular simulation scenario are very close, irrespective of which distribution is used in the MLE method. This can be attributed to the fact that the normal distribution is nested in the NIG. It also highlights the flexibility of the NIG to cover a richer pattern of distributional characteristics of the return innovations z_t, without incurring any significant estimation efficiency loss.

3.2. Data

We implemented our generalized R-EGARCH-SHARJ framework on the S&P 500 index.[12] We obtained ultra-high-frequency data (UHF) from the Trade and Quote (TAQ) database. This allowed us to monitor all the recorded tick-level data on the S&P 500 index at millisecond-level precision. Our dataset covered the period from 1995–2016. Given the UHF data on the index, we followed the usual procedures to construct evenly spaced observations at 5-minute intervals for the trading hours between 9:30 and 16:00 (EST) by taking the last price that was recorded within the previous 5-minute period.[13] This ultimately gave us 78 intra-day observations per day, which we used to calculate the (aforementioned) realized measures. A detailed presentation of realized measures estimated from UHF data can be found in Appendix.

3.3. Estimates of the R-EGARCH-SHARJ model

We start by presenting the estimates of the R-EGARCH-SHARJ framework and its nested specifications, under the assumption of an NIG distribution for returns. The results can be found in Table 3. To assess the performance of the alternative model specifications considered, we also display values of several in-sample goodness-of-fit metrics (namely, the total log-likelihood $\mathcal{L}$ and its two components $\mathcal{L}^R$ and $\mathcal{L}^V$, along with the AIC and BIC), as well as some widely used prediction performance metrics (or loss functions) indicating the ability of the conditional

[11] In an earlier version of the paper, we also considered the SGED distribution of Theodossiou (1998), which is often used in practice to capture skewness and excess kurtosis in asset returns; see e.g. Dendramis et al. (2014), Theodossiou and Savva (2016) and Feunou et al. (2016). However, our conclusions remained unchanged. Hence, the SGED results were excluded from this version of the paper in order to adhere to space limitations.

[12] The S&P 500 data provide an attractive ground for an analysis of this type. The U.S. stock market is very liquid and, therefore, one can easily obtain a rich dataset consisting of equity index options as well as ultra-high-frequency intra-day tick data. Also, there is an abundance of studies documenting the interesting stylized facts of the S&P 500 dataset, asymmetric volatility/feedback effects, negative variance risk-premia, significant deviations from normality of both physical and risk-neutral densities, etc.

[13] In our analysis we used 5-minute sampling for the intra-day data, as it has been shown that this sampling frequency offers an ideal balance between estimation accuracy and robustness to micro-structural noise. We followed the standard data-cleaning procedures for our high-frequency data, as thoroughly presented in Falkenberry (2002) and Brownlees and Gallo (2006) (we would like to thank the anonymous reviewer for pointing this out). In cases where there was no trading in a specific interval, the corresponding return was set to zero.

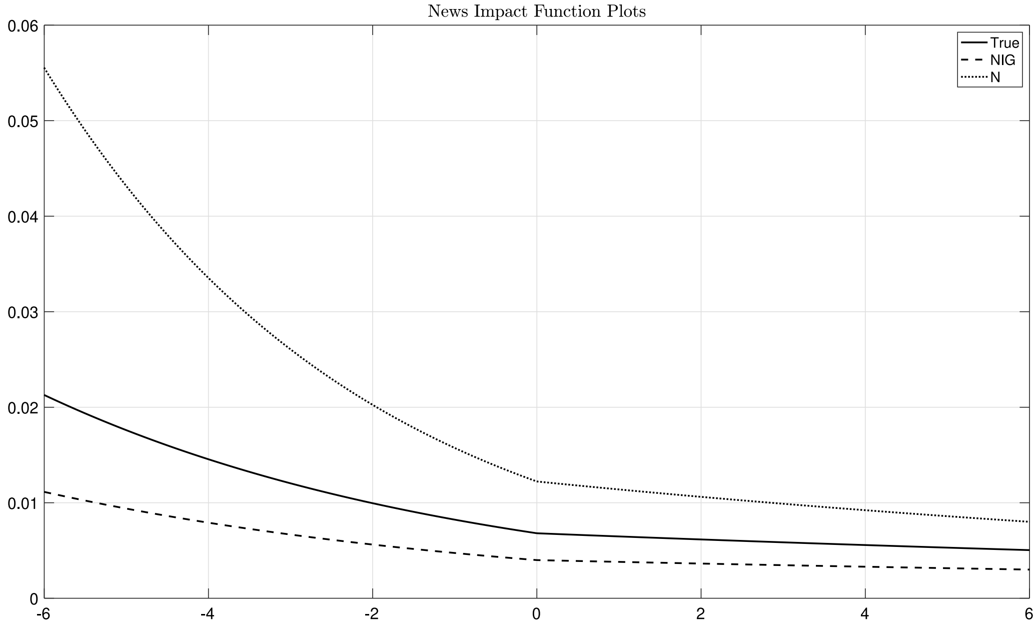

Fig. 1. This figure plots the "true NIF", where we consider that return innovations follow the NIG with $Sk = -0.32$ and $Ku = 3.2$, against those produced when estimating the model for the simulated paths under the NIG and the standard normal distribution.

Table 2
Monte Carlo simulation results.

	true	NIG-MLE			Normal-MLE		
	ϑ^*	Std	RMSE	Bias	Std	RMSE	Bias
Panel A: Return innovations follow the standard normal distribution							
β_0	−0.32	0.0520	0.0522	−0.0049	0.0506	0.0507	−0.0016
β_h	0.95	0.0086	0.0086	−0.0009	0.0084	0.0085	−0.0009
β_z	−0.12	0.0132	0.0134	0.0026	0.0120	0.0121	−0.0011
$\beta_{\|z\|}$	0.07	0.0175	0.0186	−0.0061	0.0177	0.0182	−0.0042
Sk	0.00	0.0491	0.0491	0.0011			
Ku	3.00	0.0530	0.0665	0.0404			
Panel B: Return innovations follow the NIG with $Sk = -0.32$ and $Ku = 3.2$							
β_0	−0.32	0.0488	0.0519	−0.0177	0.0887	0.0919	0.0241
β_h	0.95	0.0068	0.0069	−0.0011	0.0193	0.0193	−0.0016
β_z	−0.12	0.0320	0.0337	0.0108	0.0141	0.0437	−0.0414
$\beta_{\|z\|}$	0.07	0.0241	0.0255	−0.0084	0.0241	0.0319	0.0208
Sk	−0.32	0.0515	0.0519	−0.0062			
Ku	3.20	0.1184	0.1237	0.0357			
Panel C: Return innovations follow the NIG with $Sk = -0.60$ and $Ku = 3.8$							
β_0	−0.32	0.0498	0.0513	−0.0123	0.0519	0.0536	0.0136
β_h	0.95	0.0071	0.0072	−0.0011	0.0089	0.0090	−0.0014
β_z	−0.12	0.0221	0.0230	0.0063	0.0154	0.0285	−0.0240
$\beta_{\|z\|}$	0.07	0.0197	0.0207	−0.0064	0.0237	0.0259	0.0104
Sk	−0.60	0.0641	0.0641	−0.0001			
Ku	3.80	0.2196	0.2196	−0.0008			

Notes: This table shows how the distributional assumption for the return innovations z_t influences the bias and efficiency of the estimator of the conditional variance model parameters. We simulate and examine three different scenarios. In Panel A we assume that the true distribution of return innovations is the standard normal. In Panels B and C we consider that the underlying distribution is the NIG with $Sk = -0.32$ and $Ku = 3.2$, and $Sk = -0.60$ and $Ku = 3.8$, respectively. The column "true ϑ^*" shows the true parameters of the variance model and the distribution of the return innovations. The estimated parameter vector $\hat{\vartheta}$ of each iteration is obtained by direct MLE on the simulated path of returns, i.e. $\hat{\vartheta} = \arg\max_{\vartheta} \mathcal{L}^R$, where $R_t = \sqrt{\exp(h_t)}z_t$ with $h_t = \beta_0 + \beta_h h_{t-1} + \beta_z z_{t-1} + \beta_{|z|}|z_{t-1}|$. We simulate 3000 random paths of 3000 daily observations each, also allowing for a burn-in sample of 500 points.

log variance h_t to forecast future levels of log realized variance $\log RV_t$. We selected the following loss functions:

$$L_t^{MSE} = [\log RV_t - h_t]^2,\ L_t^{MAE} = |\log RV_t - h_t|, \quad (3)$$

$$L_t^{HMSE} = \left[\frac{RV_t - v_t}{RV_t}\right]^2,\ L_t^{MAPE} = \left|\frac{RV_t - v_t}{RV_t}\right|,$$

$$\text{and } L_t^{QLike} = h_t + \frac{RV_t}{v_t}$$

The first two loss metrics, i.e. mean squared error (MSE) and mean absolute error (MAE), focus on the ability of the conditional log variance h_t of the models to predict $\log RV$. As expected, the squared errors place more weight (penalizing) to periods with larger deviations, whereas the absolute errors provide a more equally weighted penalization scheme. The latter two metrics aim at comparing the conditional variance level $v_t \equiv \exp h_t$ to the realized variance RV_t. However, due to the highly heteroscedastic nature of variance when working with levels, we consider the percentage errors to tackle this issue; hence, we obtain the so-called heteroscedasticity-adjusted mean squared error (HMSE) and the mean absolute percentage error (MAPE) that are robust to this problem (and so are the MSE and MAE metrics for log variance). In addition to these metrics, we also consider the QLike measure introduced in Patton (2011), which is robust to the presence of noise in the measurement of variance proxies (see Table 3).

The main conclusion that can be drawn from the results in Table 3 is that the full model specification of the R-EGARCH-SHARJ considerably improves upon the performance of the standard R-EGARCH suggested by Hansen and Huang (2016). This can be confirmed by all goodness-of-fit and prediction performance metrics reported in the table. Note that the total log-likelihood $\mathcal{L}$ and the HMSE metric illustrate an improvement of approximately 13% for the R-EGARCH-SHARJ. Comparing the alternative specifications nested within this framework clearly indicates that all of them constitute valuable extensions to the R-EGARCH, since they contribute to improving the fit and prediction performance of the R-EGARCH. The standard errors reported below the parameter estimates provide further evidence in support of this finding. The estimated parameters are significant at conventional levels (5% and below); this result is consistent across all specifications. Let us also note that the residual term of the log-variance equation of the R-EGARCH-SHARJ model does not demonstrate any serial correlation, which implies that there are no indications of misspecification in the dynamics of the conditional log-variance function h_t.[14]

Regarding the separate impact of the exogenous variables included in $x(t)$ on the conditional variance h_t, the results of the table lead to a number of interesting observations. Firstly, the different values of the γ_D^+ and γ_D^- coefficient estimates reveal prevalent asymmetries in the responses of h_t to the lagged semi-variances $\log RV_{t-1}^+$ and $\log RV_{t-1}^-$. The downside semi-variance seems to have a stronger effect on h_t (almost twice as large, according to the estimates of γ_D^+ and γ_D^-), reflecting a higher response to variations in negative returns rather than variations in positive returns. Our results suggest that the inclusion of $\log RV_{t-1}^+$ and $\log RV_{t-1}^-$ in $x(t)$, instead of $\log RV_{t-1}$, considerably improves the fit and prediction performance of R-EGARCH. This can be seen by comparing the goodness-of-fit and loss-function results of the R-EGARCH and R-EGARCH-S models, and this is robust to other extensions of the R-EGARCH model. This result is also consistent with the HAR regression results discussed by Patton and Sheppard (2015). As noted in Section 2, the ability of $\log RV_t^+$ and $\log RV_t^-$ to better forecast future levels of h_t can be attributed to the asymmetries in the intra-day distributions of returns R_t that can only be captured when decomposing the (upside/downside) semi-variances. Also, note that in the presence of $\log RV_t^+$ and $\log RV_t^-$, the persistency in h_t (captured by the autoregressive coefficient β_h) remains unaffected. This implies that the effects of $\log RV_t^+$ and $\log RV_t^-$ may be considered more short-lived compared to those described by h_{t-1}.

Secondly, incorporating $\log RV_t^+$ and $\log RV_t^-$ in the R-EGARCH model does not absorb the asymmetric effects that arise in the variance process through leverage, captured by the NIF $\tau(t) = \beta_z z_t + \beta_{|z|}(|z_t| - E[|z_t|])$. The estimates of the slope coefficients β_z and $\beta_{|z|}$ of the NIF hardly change after the inclusion of $\log RV_t^+$ and $\log RV_t^-$ into the model, while they also maintain the correct sign; the estimated values of β_z are negative while those of $\beta_{|z|}$ are positive, as theory would predict. These results imply that we have two separate or different channels of volatility asymmetry, both of which appear to be statistically significant. The one originates from the asymmetric responses of variance to intra-day signed variations of $\log RV_t^+$ and $\log RV_t^-$, whereas the other is driven by the lagged innovations z_t throughout $\tau(t)$. The two effects seem to complement rather than compete with each other.

Thirdly, the positive sign of the slope coefficients of the heterogeneous terms $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$ (namely γ_W and γ_M, respectively) are indicative of important long-memory patterns in the volatility process h_t, which implies a slow decay rate for h_t; this result is in line with Huang et al. (2016). As expected, including lagged moving-average values of $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$ in the R-EGARCH model reduces the degree of persistency in h_t captured by the slope of its autoregressive term h_{t-1}, β_h. For instance, a comparison of the specifications R-EGARCH and R-EGARCH-HAR shows that the estimate of the autoregressive coefficient β_h drops from 0.72 (under R-EGARCH) to 0.59 (under R-EGARCH-HAR). However, the h_{t-1} term still appears to capture a substantial part of the variance persistency, beyond that captured by $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$. These results imply that we need both the

[14] To examine whether there is evidence of serial correlation in u_t, we carried out a series of Ljung–Box Q-tests. Below, we summarize the values of this test statistic for different lags, along with p-values in parentheses. These values clearly indicate that we cannot reject the null hypothesis of no serial correlation in u_t at the 5% or 10% level for all lags considered. This result also holds if we extend the number of lags to a much higher order.

lags	1	2	3	4	5
Q	2.54	2.91	2.92	3.09	3.95
p-values	(0.11)	(0.40)	(0.40)	(0.54)	(0.55)

Table 3
In-sample MLE results for the augmented R-EGARCH class of models (assuming NIG-distributed returns).

	R-EGARCH	R-EGARCH-S	R-EGARCH-HAR	R-EGARCH-SHAR	R-EGARCH-HARJ	R-EGARCH-SHARJ
μ	0.0360 [0.0170]	0.0371 [0.0168]	0.0365 [0.0170]	0.0375 [0.0169]	0.0360 [0.0170]	0.0370 [0.0169]
β_0	−0.3605 [0.0267]	−0.1815 [0.0270]	−0.3713 [0.0359]	−0.1809 [0.0362]	−0.3237 [0.0383]	−0.1732 [0.0394]
β_h	0.7291 [0.0116]	0.7289 [0.0117]	0.5928 [0.0237]	0.6294 [0.0235]	0.5510 [0.0248]	0.5739 [0.0252]
β_z	−0.1089 [0.0040]	−0.0811 [0.0052]	−0.1399 [0.0053]	−0.1101 [0.0067]	−0.1352 [0.0052]	−0.1174 [0.0070]
$\beta_{\|z\|}$	0.0230 [0.0070]	0.0476 [0.0070]	0.0289 [0.0079]	0.0499 [0.0076]	0.0300 [0.0080]	0.0518 [0.0079]
γ_D	0.2365 [0.0106]	–	0.2121 [0.0130]	–	0.2312 [0.0135]	–
γ_D^+	–	0.0557 [0.0104]	–	0.0705 [0.0123]	–	0.0848 [0.0132]
γ_D^-	–	0.1829 [0.0103]	–	0.1602 [0.0115]	–	0.1445 [0.0122]
γ_W	–	–	0.0550 [0.0217]	0.0180 [0.0207]	0.0709 [0.0217]	0.0607 [0.0213]
γ_M	–	–	0.1053 [0.0098]	0.0904 [0.0096]	0.1118 [0.0103]	0.1021 [0.0104]
γ_J	–	–	–	–	−0.2427 [0.0223]	−0.1682 [0.0222]
η	50.8663 [4.8938]	51.4899 [5.8701]	50.0695 [5.2237]	50.2103 [4.9380]	48.0058 [4.2153]	49.0001 [4.6577]
λ	−0.6867 [0.0109]	−0.6840 [0.0110]	−0.6867 [0.0109]	−0.6849 [0.0110]	−0.6844 [0.0110]	−0.6838 [0.0110]
$\mathcal{L}^R$	18095.15	18109.02	18095.46	18104.85	18108.75	18110.99
$\mathcal{L}^V$	−4612.67	−4543.69	−4517.88	−4473.60	−4458.44	−4445.82
$\mathcal{L}$	13482.48	13565.33	13577.58	13631.25	13650.32	13665.17
AIC	−26948.96	−27112.67	−27135.16	−27240.50	−27278.64	−27306.34
BIC	−26896.01	−27053.09	−27068.97	−27167.68	−27205.82	−27226.91
MSE	0.3097	0.3021	0.2993	0.2946	0.2930	0.2916
MAE	0.4297	0.4248	0.4228	0.4195	0.4175	0.4168
HMSE	0.5484	0.5047	0.5070	0.4814	0.4750	0.4693
MAPE	0.4738	0.4642	0.4615	0.4555	0.4522	0.4509
QLike	0.5488	0.5459	0.5447	0.5428	0.5431	0.5424

Notes: This table presents maximum likelihood estimation (MLE) results for the augmented R-EGARCH class of models defined in Table 1 (Panel A) assuming that the return innovations follow the NIG distribution. $\mathcal{L}^R$ denotes the maximum log-likelihood value of the return component, while $\mathcal{L}^V$ is the maximum log-likelihood value of the log realized variance component. $\mathcal{L}$ is the sum of the two components. AIC and BIC denote the Akaike and Bayesian Information Criterion, respectively. MSE is the mean-squared error, MAE denotes the mean absolute error, HMSE is the heteroskedasticity-adjusted squared error, MAPE denotes the mean absolute percentage error and QLike is the quasi-likelihood measure of Patton (2011). Standard errors are shown in brackets.

autoregressive term h_{t-1} and the lagged heterogeneous components $\log RV_{t-1}^{[5]}$ and $\log RV_{t-1}^{[20]}$ in order to better capture the latent long-term dependence structure embedded in the variance process.

Fourthly, the extension of R-EGARCH to incorporate the (relative) realized jump variation measure $\log RJ_t$ also brings important information for modeling h_t and forecasting RV_t. This is easily seen by comparing the fit and prediction performance metrics between the full specification R-EGARCH-SHARJ and its nested specification R-EGARCH-SHAR. The negative estimate of γ_J, i.e. the slope coefficient of $\log RJ_t$, reveals that volatility declines after days with higher activity in terms of jump variation. This implies that the effect on h_t coming from the jump variation has an opposite direction to that of $\log RV_t^+$ and $\log RV_t^-$ and/or $\tau(t)$. As argued by Patton and Sheppard (2015), this can be attributed to the very high mean reversion of the jump variation measure, which can be also justified by the fact that $\log RJ_t$ was found to demonstrate negligible levels of autocorrelation. The fact that $\log RJ_t$ captures a different effect on h_t from that of the semi-variances $\log RV_t^+$ and $\log RV_t^-$ or the NIF $\tau(t)$ possibly relates to the rare extreme events in stock index returns that can only be characterized by $\log RJ_t$.

To sum up, the estimation results of this section provide clear evidence that extending R-EGARCH to incorporate (i) the two semi-variances $\log RV_t^+$ and $\log RV_t^-$, (ii) the heterogeneous terms $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$, and (iii) the relative jump variation measure $\log RJ_t$ considerably improves the goodness-of-fit and prediction performance of the model. The added information can capture asymmetries in the intra-day return distributions and/or jumps (extremely rare events), as well as a richer time-dependence structure characterizing the variance process. Our results highlight that it is crucial for future academic research on parametric models of conditional volatility to account for all of the above effects jointly.

3.3.1. Non-normal asset returns and the impact on the R-EGARCH-SHARJ model

Employing a more flexible NIG distribution for describing asset returns (in other words, allowing for skewness and excess kurtosis) can also have a beneficial impact on modeling the variance process h_t. In order to empirically assess this assertion, we compare the results of Table 3 to those of Table 4, where we report estimates of the R-EGARCH-SHARJ framework (and its several nested specifications) under the assumption that returns

Table 4
In-sample MLE results for the augmented R-EGARCH class of models (assuming normality for return innovations).

	R-EGARCH	R-EGARCH-S	R-EGARCH-HAR	R-EGARCH-SHAR	R-EGARCH-HARJ	R-EGARCH-SHARJ
μ	0.0342 [0.0134]	0.0354 [0.0134]	0.0348 [0.0134]	0.0356 [0.0134]	0.0344 [0.0134]	0.0352 [0.0134]
β_0	−0.3273 [0.0251]	−0.1495 [0.0257]	−0.3392 [0.0343]	−0.1511 [0.0345]	−0.2898 [0.0364]	−0.1407 [0.0373]
β_h	0.7298 [0.0111]	0.7302 [0.0112]	0.5897 [0.0225]	0.6309 [0.0223]	0.5554 [0.0233]	0.5821 [0.0236]
β_z	−0.1158 [0.0041]	−0.0864 [0.0053]	−0.1487 [0.0054]	−0.1156 [0.0069]	−0.1434 [0.0053]	−0.1228 [0.0070]
$\beta_{\lvert z\rvert}$	0.0133 [0.0070]	0.0389 [0.0070]	0.0183 [0.0080]	0.0407 [0.0077]	0.0196 [0.0082]	0.0427 [0.0080]
γ_D	0.2361 [0.0102]	–	0.2110 [0.0127]	–	0.2293 [0.0132]	–
γ_D^+	–	0.0532 [0.0100]	–	0.0660 [0.0120]	–	0.0792 [0.0128]
γ_D^-	–	0.1843 [0.0100]	–	0.1633 [0.0112]	–	0.1480 [0.0119]
γ_W	–	–	0.0641 [0.0207]	0.0246 [0.0197]	0.0752 [0.0205]	0.0630 [0.0201]
γ_M	–	–	0.0992 [0.0095]	0.0829 [0.0092]	0.1040 [0.0099]	0.0931 [0.0099]
γ_J	–	–	–	–	−0.2401 [0.0215]	−0.1642 [0.0213]
$\mathcal{L}^R$	17785.30	17804.62	17780.90	17794.12	17795.84	17800.77
$\mathcal{L}^V$	−4672.70	−4601.72	−4576.94	−4531.67	−4515.97	−4503.31
$\mathcal{L}$	13112.60	13202.91	13203.97	13262.45	13279.86	13297.46
AIC	−26213.19	−26391.81	−26391.93	−26506.90	−26541.73	−26574.92
BIC	−26173.48	−26345.48	−26338.98	−26447.33	−26482.16	−26508.73
MSE	0.3165	0.3085	0.3057	0.3008	0.2991	0.2977
MAE	0.4376	0.4330	0.4307	0.4277	0.4255	0.4247
HMSE	0.6764	0.6203	0.6243	0.5915	0.5837	0.5763
MAPE	0.5201	0.5087	0.5060	0.4990	0.4952	0.4935
QLike	0.5387	0.5363	0.5350	0.5334	0.5337	0.5331

Notes: This table presents maximum likelihood estimation (MLE) results for the augmented R-EGARCH class of models defined in Table 1 (Panel A) assuming that the return innovations follow the standard normal distribution. $\mathcal{L}^R$ denotes the maximum log-likelihood value of the return component, while $\mathcal{L}^V$ is the maximum log-likelihood value of the log realized variance component. $\mathcal{L}$ is the sum of the two components. AIC and BIC denote the Akaike and Bayesian information criteria, respectively. MSE is the mean squared error, MAE denotes the mean absolute error, HMSE is the heteroskedasticity-adjusted squared error, MAPE denotes the mean absolute percentage error and QLike is the quasi-likelihood measure of Patton (2011). Standard errors are shown in brackets.

follow a normal density, i.e. $z_t \sim N(0, 1)$. The results of the two tables indicate that the assumption of z_t being NIG-distributed (instead of normally distributed) leads to significantly better fit and prediction performance; this holds across all the specifications within the R-EGARCH-SHARJ framework. The estimates of the shape parameters η and λ of the NIG reported in Table 3 imply a substantial level of negative skewness in the distribution of returns (varying around −0.29 across models) that should not be neglected, as well as a mild level of excess kurtosis (approximately 3.17).

We also observe differences to the estimated parameters of the variance process, induced by the parametric assumption governing the return distribution. Under the normality assumption, the estimates of the slope coefficients tend to overstate the effects of almost all the explanatory variables in the R-EGARCH-SHARJ model and its various specifications. This result also holds for the NIF $\tau(t)$ component of h_t, which is consistent with the findings from our MC simulation analysis. As noted in the simulation section, the normality assumption leads to poor filtering of the true distribution of the return innovations z_t, due to its inability to capture the prevalent skewness and kurtosis in the data. Consequently, the R-EGARCH-SHARJ model parameter estimates compensate by overstating the magnitude of the slope coefficient estimates—particularly those of $\tau(t)$, which directly depend on z_t and $|z_t|$.

Our empirical results, together with the outcome of our MC study, highlight the utmost importance of using a more flexible distribution (such as the NIG) when filtering return innovations, as this can influence the parameter estimates of a volatility model that aims to capture leverage effects from shocks in returns.

3.3.2. Extensions to the HAR class: Using return innovations to capture leverage

As noted in Section 2, the R-EGRACH-SHARJ framework can encompass the HAR model and many of its variations in the literature. In this subsection, we compare the R-EGRACH-SHARJ to extensions of the HAR model suggested in the literature. There is a key benefit to relying on the R-EGRACH-SHARJ framework for these comparisons. Doing so allows us to examine whether using filtered return innovations z_t can better capture leverage effects and hence improve the performance of a HAR-type model, compared to using raw lagged returns as often done in practice (e.g. Audrino and Hu (2016), Bollerslev et al. (2009), Corsi and Renò (2012)). Thus, we compare some alternative HAR-type specifications, namely the HAR, SHAR, SHARJ, and SHARJL models (as listed in Panel B of Table 1). For the richer specification, i.e. for the SHARJL model, we distinguish two different cases of modeling the NIF $\tau(t)$: the first one assumes that $\tau(t) = \beta_z z_t + \beta_{|z|}(|z_t| - E[|z_t|])$, and the second replaces z_t with the raw returns R_t (i.e. $\tau(t) =$

$\beta_R R_t + \beta_{|R|}|R_t|$). In order to ensure clarity, we henceforth denote the two cases of the SHARJL model as SHARJL$_z$ and SHARJL$_R$, respectively. The results are reported in Table 5.

Two interesting conclusions can be drawn from the results of the table. First, the best specification in the HAR class of models is the SHARJL$_z$, which uses filtered return innovations to capture the NIF. This can be justified by the results of both the fit and prediction metrics reported in the table. This finding is attributed to the fact that return innovations z_t better capture the leverage effects on variance than returns R_t. As we mentioned above (see Section 2), this happens because the filtered innovations z_t are not contaminated by conditional variance and/or risk (e.g. variance-in-mean) effects.

A second interesting conclusion that can be drawn from the results of Table 5 concerns the performance of the SHARJL$_z$ model compared to the R-EGARCH-SHARJ framework. Recall that the last model also includes the autoregressive term h_{t-1} in the conditional log-variance function. The fit and prediction performance metrics reported in the tables clearly indicate that R-EGARCH-SHARJ constitutes a better specification than SHARJL$_z$. Further support for this can be obtained by the Ljung–Box Q-test for serial correlation in the error term u_t of the realized variance Eq. (1b). We find that the Ljung–Box Q-test rejects the null hypothesis of no serial correlation in u_t for SHARJL$_z$, in contrast to R-EGARCH-SHARJ.[15] As noted above, the importance of the h_{t-1} term is attributed to the fact that it appears to be a necessary component of the dynamic structure of the conditional variance. This term (endogenously) captures a smoother and longer-term component of the variance process. The terms $\log RV_{t-1}^{[5]}$ and $\log RV_{t-1}^{[20]}$, and the semi-variances $\log RV_{t-1}^{+}$ and $\log RV_{t-1}^{-}$, cannot represent this structure alone. Summing up, these results also add support to the case for jointly modeling the returns and variance dynamics.

3.4. Out-of-sample results

In order to assess the out-of-sample (OOS) prediction performance of the R-EGARCH-SHARJ model and the several models nested within, we split the available sample into two parts; the first 3000 observations (4/01/1995–30/11/2006) were used to initialize fitting the models (also known as the training sample) and the remaining 2538 observations (1/12/2006–30/12/2016) were used to test the actual performance of the models (the testing sample). In this way, we maintain a good balance between having a significantly large sample to accurately estimate the model parameters while also keeping a considerably large period for our OOS testing (Diebold, 2015). Note that our OOS period covers the major financial crisis of 2007–2008 and other short-lived shocks such as the 2010 flash crash; this enables us to examine the robustness of our models against those extreme events as well.

In order to provide a complete performance evaluation of the R-EGARCH-SHARJ framework, we carried out both one- and multi-period-ahead prediction exercises. We can think of the one-period-ahead performance evaluation as an extension to the in-sample analysis, which can reveal the best model specification fitting the actual data. Yet, another reason we focus on evaluating the one-period-ahead variance predictions has to do with the statistical properties of the tests aiming to assess the equivalent prediction accuracy in pair-wise model comparisons. It has been shown in numerous theoretical econometric studies that these tests have better size and power properties when used for one-period-ahead predictions.[16] On the other hand, the multi-period-ahead predictions can be less efficient, but they are useful for revealing which of the alternative specifications of R-EGARCH-SHARJ (i.e. which extension of R-EGARCH) can provide long-horizon forecast gains. These results may lead to improvements when considered in an asset pricing context.

3.4.1. One-step-ahead predictions

First, we consider the one-day-ahead prediction of log realized variance $\log RV_t$, obtained as $E[\log RV_{t+1}|\mathcal{F}_t] = h_{t+1}$. We apply a daily rolling re-estimation of the models, always keeping the window size fixed at 3000 observations.[17] Note that the R-EGARCH-based framework is by construction optimized to provide one-period-ahead conditional variance forecasts at a daily frequency. We ran a series of comparative tests which helped us to assess the model performance and determine whether our findings from the in-sample analysis also hold for OOS predictions. In order to facilitate the comparisons, we used the same prediction loss functions as the ones used during the in-sample analysis. Additionally, we employed the reality check (RC) test introduced by White (2000) to test for equal predictive power between model pairs. This test is asymptotically valid for both non-nested and nested models, when the OOS estimates are obtained using a fixed rolling-window. The RC is a common choice for studies of this type (see for instance Hansen and Lunde (2005) as well as Bollerslev et al. (2016)), since it provides a straightforward framework for comparing whether the loss of a model A is (statistically) significantly lower than that of a benchmark model B.[18] The statistical hypotheses for the RC test can be simply formulated as follows:

$$H_0 \;:\; E[\hat{g}_{t+1}] \leq 0 \quad \text{versus}$$
$$H_1 \;:\; E[\hat{g}_{t+1}] > 0$$

15 The p-values of the Ljung–Box Q-test reported below clearly reject the null hypothesis of no serial correlation in u_t for the SHARJL$_z$ model.

lags	1	2	3	4	5
Q	12.60	36.15	37.32	38.56	47.07
p-values	(3.0E−4)	(1.4E−08)	(3.9E−08)	(8.4E−08)	(5.5E−09)

16 See West (2006) and Clark and McCracken (2013) for a thorough discussion. Hansen and Lunde (2005) (and similar studies) also focus on one-step-ahead forecasts when comparing volatility models.

17 Obtaining forecasts based on rolling re-estimation of model parameters is also robust to the occurrence of possibly unaccounted structural breaks; this problem is discussed in Giraitis et al. (2015, 2013).

18 A notable advantage of White's RC when used for multiple model comparisons is that it does not rely on the use of probability inequalities and typically gives quite conservative tests (in contrast to Bonferroni bound tests). See also Hansen (2005).

Table 5
In-sample MLE results for the HAR class of models (assuming NIG-distributed returns).

	HAR	HARJ	SHAR	SHARJ	SHARJL$_z$	SHARJL$_R$
μ	0.0392 [0.0173]	0.0381 [0.0172]	0.0401 [0.0172]	0.0393 [0.0172]	0.0366 [0.0155]	0.0364 [0.0172]
β_0	−0.5418 [0.0812]	−0.4054 [0.0800]	−0.2780 [0.0807]	−0.1825 [0.0805]	−0.3258 [0.0810]	−0.5386 [0.0949]
β_h	–	–	–	–	–	–
β_z	–	–	–	–	−0.1117 [0.0084]	−8.6768 [0.7558]
$\beta_{\|z\|}$	–	–	–	–	0.0255 [0.0089]	5.5671 [0.9465]
γ_D	0.3237 [0.0147]	0.3786 [0.0147]	–	–	–	–
γ_D^+	–	–	0.0346 [0.0134]	0.0684 [0.0136]	0.1627 [0.0159]	0.1225 [0.0148]
γ_D^-	–	–	0.3018 [0.0109]	0.2852 [0.0108]	0.1304 [0.0153]	0.1890 [0.0134]
γ_W	0.4112 [0.0238]	0.3674 [0.0235]	0.4227 [0.0237]	0.4107 [0.0236]	0.4364 [0.0234]	0.4122 [0.0233]
γ_M	0.2100 [0.0196]	0.2021 [0.0192]	0.1858 [0.0192]	0.1824 [0.0191]	0.2102 [0.0190]	0.1943 [0.0188]
γ_J	–	−0.4969 [0.0277]	–	−0.3165 [0.0275]	−0.3272 [0.0275]	−0.3339 [0.0273]
η	54.3235 [5.3014]	49.1036 [5.3308]	55.9353 [5.3559]	51.9697 [4.5848]	50.0760 [5.1978]	49.2296 [5.0291]
λ	−0.6986 [0.0104]	−0.6943 [0.0106]	−0.6942 [0.0106]	−0.6917 [0.0107]	−0.6913 [0.0107]	−0.6924 [0.0107]
$\mathcal{L}^R$	18027.92	18054.17	18050.96	18066.50	18069.62	18064.41
$\mathcal{L}^V$	−4862.33	−4732.82	−4711.63	−4661.97	−4555.97	−4569.30
$\mathcal{L}$	13165.59	13321.35	13339.34	13404.52	13513.65	13495.11
AIC	−26317.19	−26626.70	−26662.67	−26791.05	−27005.30	−26968.21
BIC	−26270.85	−26573.74	−26609.72	−26731.48	−26932.48	−26895.40
MSE	0.3390	0.3235	0.3210	0.3153	0.3035	0.3049
MAE	0.4506	0.4410	0.4386	0.4342	0.4255	0.4267
HMSE	0.6099	0.5258	0.5392	0.5130	0.4862	0.4936
MAPE	0.4996	0.4813	0.4798	0.4726	0.4611	0.4633
QLike	0.5708	0.5640	0.5614	0.5592	0.5510	0.5513

Notes: This table presents maximum likelihood estimation (MLE) results for the HAR-type class of models defined in Table 1 (Panel B) assuming that the return innovations follow the NIG distribution. $\mathcal{L}^R$ denotes the maximum log-likelihood value of the return component, while $\mathcal{L}^V$ is the maximum log-likelihood value of the log realized variance component. $\mathcal{L}$ is the sum of the two components. AIC and BIC denote the Akaike and Bayesian Information Criterion, respectively. MSE is the mean squared error, MAE denotes the mean absolute error, HMSE is the heteroskedasticity-adjusted squared error, MAPE denotes the mean absolute percentage error and QLike is the quasi-likelihood measure of Patton (2011). Standard errors are shown in brackets.

with $\hat{g}_{t+1} \equiv L_B(\log RV_{t+1}, h^B_{t+1}|\hat{\vartheta}^B_t) - L_A(\log RV_{t+1}, h^A_{t+1}|\hat{\vartheta}^A_t)$, where $L(.)$ can be any of the variance loss functions presented in (3). Here, model A stands for the alternative specification, which in our case is the full R-EGARCH-SHARJ model. Model B corresponds to a selected benchmark or restricted model, for which we (sequentially) consider all the nested specifications (since they constitute restricted cases of the full model). Note that in implementing the RC, we employ the stationary bootstrap technique of Politis and Romano (1994) to obtain the asymptotic variance of the loss-differential distribution. For the selection of the block length of the bootstrap, we follow the approach proposed by Politis and White (2004), taking into account the correction discussed by Patton et al. (2009).

In Table 6 we collect all the prediction loss-function results for our one-period-ahead OOS forecasting exercise. In brackets, we present bootstrapped p-values of the RC tests. The RC p-values in Panel A represent pair-wise comparisons between the R-EGARCH-SHARJ model and the several nested specifications, assuming an NIG distribution for R_t. In Panel B, the p-values focus on comparisons of identical model configurations, when the benchmark is estimated under the normal assumption, instead of the NIG assumed for the alternative model.

A general conclusion that can be drawn from our OOS results is that the performance of the R-EGARCH-SHARJ model is in line with our findings from the in-sample analysis. This framework considerably improves upon the OOS prediction performance of the R-EGARCH and the various extensions that we consider in the paper. This can be justified by the loss-function prediction metrics reported in the table, as well as by the RC test p-values. In almost all cases, the results indicate that the improvements in prediction performance for the R-EGARCH-SHARJ over its nested specifications are statistically significant at conventional levels (5% or less). A closer look to the magnitude of these improvements reveals that the R-EGARCH-SHARJ model under the NIG reduces the MSE (MAE) from 0.3713 (0.4718) found for the R-EGARCH to just 0.3486 (0.4530). The improvement seems even more pronounced if we use as a prediction evaluation criterion the squared (or absolute) percentage error metric HMSE (MAPE): the forecast gains reach a level of almost 20% for the HMSE. The conclusions that can be drawn for the normal assumption case are qualitatively very similar.

When comparing identical model configurations under different distributional assumptions for the returns,

Table 6
Out-of-sample loss-function results and reality check test p-values.

	Panel A: NIG					Panel B: Normal				
	MSE	MAE	HMSE	MAPE	QLike	MSE	MAE	HMSE	MAPE	QLike
R-EGARCH-SHARJ	0.3486	0.4530	0.5307	0.4750	0.6358	0.3482	0.4553	0.6220	0.5060	0.6228
	[–]	[–]	[–]	[–]	[–]	[1.0000]	[0.0288]	[0.0000]	[0.0000]	[1.0000]
R-EGARCH-HARJ	0.3497	0.4530	0.5338	0.4754	0.6367	0.3490	0.4553	0.6255	0.5063	0.6237
	[0.1321]	[1.0000]	[0.1906]	[0.3651]	[0.1229]	[1.0000]	[0.0282]	[0.0001]	[0.0000]	[1.0000]
R-EGARCH-SHAR	0.3510	0.4581	0.5808	0.4967	0.6296	0.3542	0.4629	0.6822	0.5318	0.6197
	[0.1158]	[0.0007]	[0.0000]	[0.0000]	[1.0000]	[0.0174]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
R-EGARCH-HAR	0.3594	0.4635	0.6404	0.5125	0.6317	0.3642	0.4693	0.7565	0.5505	0.6231
	[0.0014]	[0.0000]	[0.0002]	[0.0000]	[1.0000]	[0.0015]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
R-EGARCH-S	0.3590	0.4640	0.5990	0.5050	0.6318	0.3615	0.4681	0.7008	0.5397	0.6205
	[0.0006]	[0.0000]	[0.0023]	[0.0000]	[1.0000]	[0.0533]	[0.0009]	[0.0000]	[0.0000]	[1.0000]
R-EGARCH	0.3713	0.4718	0.6968	0.5285	0.6339	0.3761	0.4772	0.8206	0.5675	0.6239
	[0.0000]	[0.0000]	[0.0021]	[0.0000]	[1.0000]	[0.0011]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
SHARJLz	0.3629	0.4622	0.5699	0.4917	0.6459	0.3640	0.4664	0.6762	0.5286	0.6320
	[0.0000]	[0.0000]	[0.0076]	[0.0000]	[0.0134]	[0.2556]	[0.0011]	[0.0000]	[0.0000]	[1.0000]
SHARJ	0.3725	0.4686	0.5954	0.5020	0.6458	0.3732	0.4719	0.7061	0.5388	0.6363
	[0.0000]	[0.0000]	[0.0013]	[0.0000]	[0.0167]	[0.3486]	[0.0045]	[0.0000]	[0.0000]	[1.0000]
HARJ	0.3806	0.4756	0.5948	0.5068	0.6461	0.3811	0.4781	0.7130	0.5453	0.6408
	[0.0000]	[0.0000]	[0.0007]	[0.0000]	[0.0352]	[0.4089]	[0.0411]	[0.0000]	[0.0000]	[1.0000]
SHAR	0.3826	0.4776	0.6871	0.5297	0.6471	0.3871	0.4832	0.8203	0.5720	0.6361
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0368]	[0.0045]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
HAR	0.4118	0.4967	0.8311	0.5660	0.6513	0.4181	0.5033	1.0048	0.6154	0.6488
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0190]	[0.0008]	[0.0000]	[0.0000]	[0.0000]	[1.0000]

Notes: This table presents out-of-sample results for one-step-ahead predictions of $\log RV$ generated by the models shown in the first column of the table. It reports the mean squared error (MSE), mean absolute error (MAE), heteroskedasticity-adjusted mean squared error (HMSE), mean absolute percentage error (MAPE), and the quasi-likelihood measure of Patton (2011) (QLike). We apply a daily rolling re-estimation of the models with a fixed window size of 3000 observations. Panels A and B present the results assuming that return innovations follow the NIG and standard normal distribution, respectively. In Panel A, the reality check test p-values in square brackets stand for comparisons of the nested (benchmark) models to the alternative R-EGARCH-SHARJ model. In Panel B, the reality check test p-values compare each one of the identical model configurations under the normal (benchmark) and NIG (alternative) assumptions.

we found significant gains in predicting realized variance when assuming that return innovations are NIG- instead of normally distributed (refer to Panel B of the Table). This holds not only for the R-EGARCH-SHARJ model but also for all the nested model specifications, including R-EGARCH. As before, the most significant prediction gains from employing a more flexible NIG density are found for the HMSE and MAPE metrics. For the full R-EGARCH-SHARJ specification, the reported value for the HMSE (MAPE) loss metric declines from 0.65 (0.51) to 0.55 (0.48) when using the NIG. These differences are significant at 5% (or lower) levels, according to the RC test p-values reported in the table. Note that the only metric that does not seem to favor the NIG over the normality assumption is the QLike, the difference being statistically not significant.[19]

Finally, for the HAR class of models, the OOS results are consistent with those of the in-sample analysis. The values of the loss-function metrics and the p-values of the RC test reported in the table clearly indicate that the extension to the HAR that allows for semi-variance, jumps, and leverage effects through filtered innovations z_t (i.e. SHARJL$_z$) performs significantly better than the other specifications nested within it. However, this model does not outperform the full R-EGARCH-SHARJ, incorporating the latent/conditional dynamic term h_{t-1}, thus implying the importance of this term in predicting realized variance OOS.

3.4.2. Multi-period-ahead predictions

To obtain the multi-period-ahead predictions we rely on a direct forecasting approach (see for instance Clark and McCracken (2005, 2013), Marcellino et al. (2006), and Ghysels et al. (2019)), and we employ the following horizon-specific regression model implied by the R-EGARCH-SHARJ framework:

$$\begin{aligned} \log RV_{t+s} =& \beta_{0,s} + \beta_{h,s}h_t + \beta_{z,s}z_t + \beta_{|z|,s}(|z_t| - E[|z_t|]) \\ &+ x_s(t) + \sigma_s u_{t+s} \quad \text{with} \\ x_s(t) =& \gamma^{+}_{D,s}\log RV^{+}_t + \gamma^{-}_{D,s}\log RV^{-}_t + \gamma_{W,s}\log RV^{[5]}_t \\ &+ \gamma_{M,s}\log RV^{[20]}_t + \gamma_{J,s}\log RJ_t \end{aligned} \quad (4)$$

In this case, our aim is to forecast the value of dependent variable $\log RV_{t+s}$ at s steps ahead, but conditional on the information set up to time t, i.e. $\mathcal{F}_t$.[20] This model is estimated by rolling over our OOS daily observations, and based on these estimates, we obtain predictions of $\log RV_{t+s}$ by calculating $\mathbb{E}[\log RV_{t+s}|\mathcal{F}_t]$ at each point of time t.

In Table 7 we present estimates of the prediction loss functions used in our analysis to evaluate the performance of the alternative specifications of (4) considered. We ran the regressions for $s = \{5, 10, 20\}$ days ahead. The results

[19] The reason behind this finding is twofold. First, under the normality assumption, the models overestimate variance, as seen from the NIF plots in Fig. 1. Second, the QLike loss function is asymmetric, hence penalizing more in cases when variance is underestimated (see also Figure 1 in Patton (2011)).

[20] Several studies utilize a similar framework for forecasting volatility over long horizons; see e.g. Andersen et al. (2003), as well as Ghysels et al. (2019) for a recent survey.

Table 7
Out-of-sample loss-function results for multi-step-ahead predictions.

	5-days ahead					10-days ahead					20-days ahead				
	MSE	MAE	HMSE	MAPE	QLike	MSE	MAE	HMSE	MAPE	QLike	MSE	MAE	HMSE	MAPE	QLike
R-EGARCH-SHARJ	0.6257 [−]	0.6104 [−]	1.8488 [−]	0.7532 [−]	0.8476 [−]	0.7664 [−]	0.6770 [−]	2.4442 [−]	0.8589 [−]	1.0004 [−]	0.8981 [−]	0.7390 [−]	2.6865 [−]	0.9348 [−]	1.1264 [−]
R-EGARCH-HARJ	0.6265 [0.0533]	0.6103 [0.2511]	1.8485 [0.2973]	0.7525 [0.3827]	0.8553 [0.0825]	0.7646 [0.2399]	0.6762 [0.1604]	2.4284 [1.0000]	0.8567 [1.0000]	0.9956 [0.3741]	0.8979 [0.1185]	0.7390 [0.2140]	2.6974 [0.0614]	0.9351 [0.1318]	1.1274 [0.1040]
R-EGARCH-SHAR	0.6231 [0.4311]	0.6103 [0.1727]	1.7782 [1.0000]	0.7504 [0.4780]	0.8423 [1.0000]	0.7614 [1.0000]	0.6767 [0.2307]	2.3376 [1.0000]	0.8559 [0.4309]	0.9815 [1.0000]	0.8920 [1.0000]	0.7372 [0.2624]	2.6388 [1.0000]	0.9308 [0.3947]	1.1180 [1.0000]
R-EGARCH-HAR	0.6261 [0.1100]	0.6117 [0.0369]	1.7853 [1.0000]	0.7520 [0.2639]	0.8499 [0.0796]	0.7575 [1.0000]	0.6751 [0.2465]	2.2968 [1.0000]	0.8524 [0.4835]	0.9709 [1.0000]	0.8917 [0.3576]	0.7372 [0.2192]	2.6349 [1.0000]	0.9305 [0.2853]	1.1210 [0.4057]
R-EGARCH-S	0.6341 [0.0026]	0.6170 [0.0003]	1.9648 [0.0535]	0.7688 [0.0014]	0.8367 [1.0000]	0.7734 [0.0066]	0.6813 [0.0069]	2.4924 [0.0310]	0.8697 [0.0066]	0.9832 [1.0000]	0.9038 [0.0137]	0.7432 [0.0085]	2.7202 [0.0588]	0.9433 [0.0226]	1.1241 [0.3471]
R-EGARCH	0.6400 [0.0010]	0.6197 [0.0000]	2.0414 [0.0414]	0.7752 [0.0017]	0.8445 [1.0000]	0.7705 [0.0090]	0.6798 [0.0146]	2.4615 [0.0534]	0.8676 [0.0109]	0.9708 [1.0000]	0.9055 [0.0070]	0.7447 [0.0041]	2.7241 [0.0508]	0.9455 [0.0150]	1.1272 [0.1949]
SHARJLz	0.6348 [0.0255]	0.6114 [0.3004]	1.9682 [0.0306]	0.7592 [0.1673]	0.8544 [0.1356]	0.7690 [0.1717]	0.6763 [0.1751]	2.8043 [0.1765]	0.8614 [0.2268]	1.0204 [0.1087]	0.8962 [0.3976]	0.7361 [1.0000]	2.7089 [0.2648]	0.9303 [0.4547]	1.1274 [0.4197]
SHARJ	0.6373 [0.0062]	0.6130 [0.0487]	1.9725 [0.1016]	0.7606 [0.0908]	0.8503 [0.3400]	0.7734 [0.0754]	0.6784 [0.1014]	2.8107 [0.2471]	0.8612 [0.3733]	1.0312 [0.0585]	0.8948 [0.3018]	0.7356 [0.3788]	2.6559 [1.0000]	0.9275 [0.4572]	1.1242 [1.0000]
HARJ	0.6352 [0.0035]	0.6104 [0.0516]	2.0260 [0.0863]	0.7573 [0.0991]	0.8582 [0.0026]	0.7685 [0.0638]	0.6766 [0.0641]	2.8085 [0.2496]	0.8564 [0.3651]	1.0145 [0.0453]	0.8929 [0.0590]	0.7362 [0.0693]	2.6491 [0.3752]	0.9284 [0.0684]	1.1222 [0.2419]
SHAR	0.6373 [0.0012]	0.6147 [0.0065]	1.9102 [0.2976]	0.7609 [0.0409]	0.8453 [0.4388]	0.7717 [0.0262]	0.6805 [0.0184]	2.6713 [0.2353]	0.8625 [0.1277]	1.0013 [1.0000]	0.8910 [0.1710]	0.7347 [0.1394]	2.6182 [1.0000]	0.9247 [0.2422]	1.1202 [0.3713]
HAR	0.6413 [0.0000]	0.6171 [0.0002]	1.9672 [0.1492]	0.7653 [0.0149]	0.8550 [0.0246]	0.7683 [0.0102]	0.6793 [0.0031]	2.5549 [0.2438]	0.8573 [0.0638]	0.9776 [1.0000]	0.8901 [0.0520]	0.7352 [0.0342]	2.5725 [0.4225]	0.9252 [0.0767]	1.1215 [0.1578]

Notes: This table presents out-of-sample results for 5-, 10-, and 20-day-ahead predictions of $\log RV$ generated by the models shown in the first column of the table assuming that return innovations follow the NIG distribution. It reports the mean squared error (MSE), mean absolute error (MAE), heteroskedasticity-adjusted mean squared error (HMSE), mean absolute percentage error (MAPE,) and the quasi-likelihood measure of Patton (2011) (QLike). The reality check test p-values in square brackets stand for comparisons of the nested (benchmark) models to the alternative R-EGARCH-SHARJ model.

of the table indicate that the full specification of the R-EGARCH-SHARJ model preserves its prediction superiority that we found for one period ahead only for a horizon up to five days ahead. If we increase the prediction horizon to $s = \{10, 20\}$ days, the R-EGARCH-SHAR and R-EGARCH-HAR specifications appear to provide marginally better forecasts than the full-model, with R-EGARCH-HAR being the superior one between these two. The above results hold for the majority of the performance metrics reported in the table and across the different prediction horizons.

Summing up, these results highlight the importance of considering $\log RV_t^+$ and $\log RV_t^-$, and/or $\log RV_t^{[5]}$ and $\log RV_t^{[20]}$, for providing short- and medium-term predictions of the realized variance of asset returns. The coefficient estimates of the model obtained by the rolling regression procedure (not reported for reasons of space) also indicate that $\gamma_{D,s}^+$ and $\gamma_{D,s}^-$ are significant for $s = \{5, 10\}$, as are $\gamma_{W,s}$ and $\gamma_{M,s}$ for all horizons, i.e. $s = \{5, 10, 20\}$. On the contrary, the effects of the slope coefficients of the NIF and the jump variation measure (i.e. $\beta_{z,s}$, $\beta_{|z|,s}$, and $\gamma_{J,s}$) tend to be informative only for short-run predictions; these die out for prediction horizons of roughly $s = \{10, 20\}$ days.

4. Summary & conclusions

In this paper, we effectively built on the Realized-EGARCH model and we suggested extensions of it that include a richer set of realized measures to improve the fit and prediction of the variance process. The choice of these variables was well motivated by the ongoing literature on HAR models for realized variance forecasting. We constructed an exponential class of models for the variance process, and we analyzed both the in- and out-of-sample forecasting performance of the several specifications nested within the suggested framework. Additionally, we examined the performance of the suggested models under different assumptions for the distribution of asset return innovations. In particular, we considered not only the normal distribution (often used in practice) but also the NIG. These were shown to perform very well at capturing the distributional asymmetries of financial data and fat tails.

We found that our augmented R-EGARCH framework improves considerably upon the goodness of fit and prediction performance of the standard R-EGARCH. We demonstrated that the improvement of the model originates from three sources: (i) the decomposition of the RV into upside/downside semi-variances, which better capture asymmetric behaviors in variance related to negative skewness in asset return distributions; (ii) the addition of heterogeneous components of the realized variance approximating a long-memory effect in variance process; and (iii) the inclusion of the jump (discontinuous) component of the quadric variation, reflecting jumps in asset prices and/or the variance process itself.

Interestingly, we showed that our suggested framework performs better than the HAR class of models for realized variance and its extensions suggested in the literature. Our framework encompasses this class of models. We showed that its benefits come mainly from two sources: (i) the use of asset return innovations (instead of returns) to capture the leverage effects on variance; and (ii) the incorporation of the autoregressive component of the conditional variance, filtered throughout the R-EGARCH model. The importance of the above two sources in forecasting realized variance emphasizes the need to jointly model the dynamics of asset returns and variance.

Finally, regarding the performance of our framework under alternative parametric density assumptions of asset return innovations, our results demonstrated that the NIG constitutes a useful extension. This can be attributed to fact that, by allowing for skewness and excess kurtosis in asset return innovations, this distribution can better capture the probability density function of the innovation terms and, hence, estimate the parameters of the model more efficiently, especially those of the news impact function (NIF). Our MC simulation exercise provided further evidence in favor of this intuition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Realized variance estimators

In this appendix, we explain how the realized estimators that we use as exogenous drivers of variance are constructed based on high-frequency data. Let us start by representing the log-price process $\{p_t\}$ as the sum of a continuous and a pure-jump component:

$$p_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + J_t \ ,$$

where J_t stands for the jump process, μ_s is a deterministic drift process, and σ_s is a strictly positive càdlàg instantaneous volatility process. We define the ith intraday log-return at day t as $r_{t,i} = p_{t,\frac{i}{n}} - p_{t,\frac{i-1}{n}}$, with $i = 1, 2, \ldots, n$, so that $p_{t,0}$ and $p_{t,1}$ are the opening and closing log prices, respectively.[21] Recall that the true quadratic variation $\langle p, p\rangle_{t-1,t}$ is unobservable. However, it is well known that under mild conditions it can be estimated by the realized variance RV_t from intra-day log returns; see Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002a, 2002b, 2004a, 2004b, 2006), among others. It can be shown that for n equally spaced intra-day observations in $(t-1, t)$, RV_t converges in probability to the quadratic variation as the time interval between two consecutive observations shrinks. That is,

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2 \xrightarrow{p} \int_0^t \sigma_s^2 ds + \sum_{0<s\le t} (\Delta p_s)^2 \equiv \langle p, p\rangle_{t-1,t} \quad \text{as } n \to \infty,$$

which is the sum of integrated variance and jump variation, with $\Delta p_s = p_s - p_{s-}$ representing the jumps. Following Barndorff-Nielsen et al. (2008), RV_t can be further decomposed into upside and downside semi-variances (RV_t^+ and RV_t^-, respectively), which are defined as

$$RV_t^+ = \sum_{i=1}^{n} r_{t,i}^2 \mathbb{I}_{\{r_{t,i}>0\}} \xrightarrow{p} \frac{1}{2}\int_0^t \sigma_s^2 ds + \sum_{0<s\le t} (\Delta p_s)^2 \mathbb{I}_{\{\Delta p_s>0\}} \quad \text{and}$$

$$RV_t^- = \sum_{i=1}^{n} r_{t,i}^2 \mathbb{I}_{\{r_{t,i}\le 0\}} \xrightarrow{p} \frac{1}{2}\int_0^t \sigma_s^2 ds + \sum_{0<s\le t} (\Delta p_s)^2 \mathbb{I}_{\{\Delta p_s\le 0\}} \quad \text{as } n \to \infty,$$

which basically implies that $RV_t = RV_t^+ + RV_t^-$.[22]

Finally, we also construct a measure of "relative" jump variation, which we denote as RJ_t. In doing that, we first need to isolate the continuous from the discontinuous part of the variation, i.e. the integrated variance from the jump variation. In the limit, the part of the variation that arises purely due to the continuous price movements, also known as integrated variance, can be estimated by the so-called bi-power variation (Barndorff-Nielsen & Shephard, 2004b) as

$$BV_t = \frac{\pi}{2}\sum_{i=2}^{n} |r_{t,i}||r_{t,i-1}| \xrightarrow{p} \int_{t-1}^t \sigma_s^2 ds \quad \text{as } n \to \infty.$$

As follows, the difference between the realized and the bi-power variation can be thought of as a non-parametric estimator of the jump variation. As in Bollerslev et al. (2009), we employ a relative jump variation measure, as it has been shown to be more robust; Huang and Tauchen (2005) provide an exhaustive simulation analysis. The approximate logarithmic version of the relative jump variation can be expressed as

$$\log RJ_t \equiv \log RV_t - \log BV_t \ ,$$

which we can directly use in our econometric framework for volatility forecasting purposes.

References

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135–167.

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

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

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Andersson, J. (2001). On the normal inverse Gaussian stochastic volatility model. *Journal of Business & Economic Statistics*, *19*(1), 44–54.

Audrino, F., & Hu, Y. (2016). Volatility forecasting: Downside risk, jumps and leverage effect. *Econometrics*, *4*(1), 8.

21 Not to be confused with the R_t notation used in the main document to denote the close-to-close daily returns.

22 Patton and Sheppard (2015) use the two semi-variances to introduce a *signed jump variation* measure as

$$SJV_t \equiv RV_t^+ - RV_t^- \xrightarrow{p} \sum_{0<s\le t} (\Delta p_s)^2 \mathbb{I}_{\{\Delta p_s>0\}} - \sum_{0<s\le t} (\Delta p_s)^2 \mathbb{I}_{\{\Delta p_s\le 0\}} \quad \text{as } n \to \infty \ ,$$

which we can think of as an alternative *realized skewness* proxy, since it asymptotically converges to the differential between the upside and downside variation of the latent jumps. We found that this metric strongly correlates with various realized skewness proxies suggested in other empirical studies on high-frequency econometrics; see, for instance, Amaya et al. (2015) and Feunou et al. (2016).

Baillie, R. T., Calonaci, F., Cho, D., & Rho, S. (2019). Long memory, realized volatility and heterogeneous autoregressive models. *Journal of Time Series Analysis*, *40*(4), 609–628.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, 151–157.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, *24*(1), 1–13.

Barndorff-Nielsen, O. E., Kinnebrock, S., & Shephard, N. (2008). Measuring downside risk-realised semivariance. *CREATES Research Paper*, (2008–42).

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *64*(2), 253–280.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, *17*(5), 457–477.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, *72*(3), 885–925.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Barndorff-Nielsen, O. E., & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1–30.

Bekierman, J., & Gribisch, B. (2016). Estimating stochastic volatility models using realized measures. *Studies in Nonlinear Dynamics & Econometrics*, *20*(3), 279–300.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bollerslev, T., Kretschmer, U., Pigorsch, C., & Tauchen, G. (2009). A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics*, *150*(2), 151–166.

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

Borup, D., & Jakobsen, J. S. (2019). Capturing volatility persistence: a dynamically complete realized EGARCH-MIDAS model. *Quantitative Finance*, *19*(11), 1839–1855.

Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis*, *51*(4), 2232–2245.

Chorro, C., Guégan, D., & Ielpo, F. (2012). Option pricing for GARCH-type models with generalized hyperbolic innovations. *Quantitative Finance*, *12*(7), 1079–1094.

Christoffersen, P., Feunou, B., Jacobs, K., & Meddahi, N. (2014). The economic value of realized volatility: Using high-frequency returns for option valuation. *Journal of Financial and Quantitative Analysis*, *49*(3), 663–697.

Christoffersen, P., Jacobs, K., & Mimouni, K. (2010). Volatility dynamics for the S&P500: Evidence from realized volatility, daily returns, and option prices. *The Review of Financial Studies*, *23*(8), 3141–3189.

Cipollini, F., Engle, R. F., & Gallo, G. M. (2012). Semiparametric vector MEM. *Journal of Applied Econometrics*, *28*(7), 1067–1086.

Cipollini, F., Gallo, G. M., & Otranto, E. (2021). Realized volatility forecasting: Robustness to measurement errors. *International Journal of Forecasting*, *37*(1), 44–57.

Clark, T. E., & McCracken, M. W. (2005). Evaluating direct multistep forecasts. *Econometric Reviews*, *24*(4), 369–404.

Clark, T. E., & McCracken, M. (2013). Advances in forecast evaluation. In *Handbook of economic forecasting, Vol. 2* (pp. 1107–1201). Elsevier.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Corsi, F., Fusari, N., & La Vecchia, D. (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics*, *107*(2), 284–304.

Corsi, F., & Renò, R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, *30*(3), 368–380.

Dendramis, Y., Spungin, G. E., & Tzavalis, E. (2014). Forecasting VaR models under different volatility processes and distributions of return innovations. *Journal of Forecasting*, *33*(7), 515–531.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, *33*(1), 1.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 987–1007.

Engle, R. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, *17*(5), 425–446.

Engle, R., & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, *131*(1–2), 3–27.

Engle, R., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-m model. *Econometrica*, 391–407.

Eriksson, A., Ghysels, E., & Wang, F. (2009). The normal inverse Gaussian distribution and the pricing of derivatives. *The Journal of Derivatives*, *16*(3), 23–37.

Falkenberry, T. N. (2002). High frequency data filtering. *White Paper, Tick Data, Inc*.

Feunou, B., Jahan-Parvar, M. R., & Tédongap, R. (2016). Which parametric model for conditional skewness? *The European Journal of Finance*, *22*(13), 1237–1271.

Feunou, B., & Okou, C. (2019). Good volatility, bad volatility, and option pricing. *Journal of Financial and Quantitative Analysis*, *54*(2), 695–727.

Forsberg, L., & Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model. *Journal of Applied Econometrics*, *17*(5), 535–548.

Ghysels, E., Plazzi, A., Valkanov, R., Rubia, A., & Dossani, A. (2019). Direct versus iterated multiperiod volatility forecasts. *Annual Review of Financial Economics*, *11*, 173–195.

Ghysels, E., & Wang, F. (2014). Moment-implied densities: properties and applications. *Journal of Business & Economic Statistics*, *32*(1), 88–111.

Giraitis, L., Kapetanios, G., Mansur, M., & Price, S. (2015). Forecasting under structural change. In *Empirical economic and financial research* (pp. 401–419). Springer.

Giraitis, L., Kapetanios, G., & Price, S. (2013). Adaptive forecasting in the presence of recent and ongoing structural change. *Journal of Econometrics*, *177*(2), 153–170.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, *23*(4), 365–380.

Hansen, P. R., & Huang, Z. (2016). Exponential GARCH modeling with realized measures of volatility. *Journal of Business & Economic Statistics*, *34*(2), 269–287.

Hansen, P. R., Huang, Z., & Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, *27*(6), 877–906.

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

Huang, Z., Liu, H., & Wang, T. (2016). Modeling long memory volatility using realized measures of volatility: A realized HAR GARCH model. *Economic Modelling*, *52*, 812–821.

Huang, X., & Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, *3*(4), 456–499.

Huang, Z., Wang, T., & Hansen, P. R. (2017). Option pricing with the realized GARCH model: An analytical approximation approach. *Journal of Futures Markets*, *37*(4), 328–358.

Jensen, M. B., & Lunde, A. (2001). The NIG-S&ARCH model: A fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model. *The Econometrics Journal*, *4*(2), 319–342.

Jondeau, E., & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics & Control*, *27*(10), 1699–1737.

Koopman, S. J., & Scharth, M. (2012). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics*, *11*(1), 76–115.

Louzis, D. P., Xanthopoulos-Sisinis, S., & Refenes, A. P. (2014). Realized volatility models and alternative value-at-risk prediction strategies. *Economic Modelling*, *40*, 101–116.

Marcellino, M., Stock, J. H., & Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *135*(1–2), 499–526.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 347–370.

Papantonis, I. (2016). Volatility risk premium implications of GARCH option pricing models. *Economic Modelling*, *58*, 104–115.

Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246–256.

Patton, A., Politis, D. N., & White, H. (2009). Correction to "automatic block-length selection for the dependent bootstrap" by D. Politis and H. White. *Econometric Reviews*, *28*(4), 372–375.

Patton, A., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *The Review of Economics and Statistics*, *97*(3), 683–697.

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303–1313.

Politis, D. N., & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, *23*(1), 53–70.

Shephard, N., & Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, *25*(2), 197–231.

Stentoft, L. (2008). American option pricing using GARCH models and the normal inverse Gaussian distribution. *Journal of Financial Econometrics*, *6*(4), 540–582.

Takahashi, M., Omori, Y., & Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis*, *53*(6), 2404–2426.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, *44*(12-part-1), 1650–1661.

Theodossiou, P., & Savva, C. S. (2016). Skewness and the relation between risk and return. *Management Science*, *62*(6), 1598–1609.

Venter, J. H., & de Jongh, P. J. (2002). Risk estimation using the normal inverse Gaussian distribution. *Journal of Risk*, *4*, 1–24.

West, K. D. (2006). Forecast evaluation. *Handbook of Economic Forecasting*, *1*, 99–134.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*(5), 1097–1126.

Wilhelmsson, A. (2009). Value at risk with time varying variance, skewness and kurtosis—the NIG-ACD model. *The Econometrics Journal*, *12*(1), 82–104.