

Ioannis Papantonis*, Leonidas S. Rompolis, Elias Tzavalis and Orestis Agapitos

Augmenting the Realized-GARCH: the role of signed-jumps, attenuation-biases and long-memory effects

<https://doi.org/10.1515/snde-2020-0131>

Received December 11, 2020; accepted April 7, 2022; published online August 10, 2022

Abstract: This paper extends the Realized-GARCH framework, by allowing the conditional variance equation to incorporate exogenous variables related to intra-day realized measures. The choice of these measures is motivated by the so-called heterogeneous auto-regressive (HAR) class of models. Our augmented model is found to outperform both the Realized-GARCH and the various HAR models in terms of in-sample fitting and out-of-sample forecasting accuracy. The new model specification is examined under alternative parametric density assumptions for the return innovations. Non-normality seems to be very important for filtering the return innovations to which variance responds and helps significantly upon the prediction performance of the suggested model.

Keywords: HAR; heterogeneous auto-regressive; NIG; realized variance; Realized-GARCH; SGED.

1 Introduction

Modelling and accurately predicting the volatility of asset returns is a key issue behind many financial applications, such as portfolio allocation, risk management and option pricing. Since Engle (1982) and Bollerslev (1986), a first “generation” of volatility models evolved and has been extensively used; the well-documented generalized autoregressive conditional heteroscedasticity (GARCH) class of models. The simple autoregressive structure of these models provided an easy way to capture clusters in the latent volatility dynamics of returns. Numerous different parametric GARCH-type specifications have been suggested to extend the GARCH models, leading to more sophisticated volatility models which incorporate leverage/feedback effects, long-memory, regime-switches, among other stylized properties of volatility.¹

Normally, standard GARCH models (and their extensions) use daily returns to extract information about the current level of volatility, treating volatility itself as unobservable. However, more recent advances in the high-frequency financial econometrics literature utilized intra-day data to construct observable proxies of variance, the so-called as realized variance (RV) estimators. These observable variance measures gave rise to a new econometric modelling framework, aiming to either incorporate this information into GARCH models

¹ Andersen et al. (2009), Bauwens, Hafner, and Laurent (2012) as well as Francq and Zakoian (2019) provide excellent surveys on GARCH-type volatility models.

*Corresponding author: Ioannis Papantonis, Department of Economics, Athens University of Economics and Business, 10434 Athens, Greece, E-mail: papantonis@aueb.gr

Leonidas S. Rompolis, Department of Accounting and Finance, Athens University of Economics and Business, 10434 Athens, Greece

Elias Tzavalis and Orestis Agapitos, Department of Economics, Athens University of Economics and Business, 10434 Athens, Greece

or directly model the realized variance process. Engle (2002) and Engle and Gallo (2006), first, used RV in a Multiplicative Error Model (MEM) context, while Shephard and Sheppard (2010) presented a so-called HEAVY (high-frequency volatility) model for jointly modelling latent and observable variance dynamics. The most popular and easily applicable models of this research-strand were the HAR-type realized variance regressions popularized by Corsi (2009) and the Realized-GARCH framework of Hansen, Huang, and Shek (2012). These models, which we can think of as a second generation of volatility models, have been shown to boost the forecasting performance of the variance process, especially when a realized variance proxy is included in the GARCH.²

One of the most important benefits of GARCH over HAR-type models is that they can jointly model the dynamics of returns and variance.³ Hence, in this paper we employ and effectively build on the Realized-GARCH framework of Hansen, Huang, and Shek (2012). We first extend the Realized-GARCH model to incorporate additional intra-day realized measures capturing different volatility asymmetry sources, i.e. asymmetries originating from both responses of variance to lagged return innovations and/or short-lived effects from realized upside and downside intra-day variations. In addition to these measures, we also consider extensions including a volatility-of-volatility measure, which can correct for attenuation-biases in measuring realized variance, as well as heterogeneous terms of realized variance, which approximate long-memory properties of variance. These extensions are well-justified by the literature on HAR models. We show that the inclusion of these metrics in parametric GARCH specifications may improve the efficiency of the conditional variance estimates. We therefore refer to these extensions of the Realized-GARCH model as a third generation of volatility models.

We empirically assess the ability of the above new extensions to improve upon the performance of the standard Realized-GARCH models to fit the data and predict realized volatility. We use intra-day data of the S & P500 data and we allow return innovations to follow alternative parametric probability density functions (PDFs), apart from the Normal (N) that is often used in practice. We consider the cases of the skewed-GED (SGED) and normal-inverse-Gaussian (NIG) PDFs that have been shown to outperform other parametric densities in fitting financial return data.⁴ These PDFs are more flexible and can account for more pronounced levels of skewness and excess kurtosis in asset returns. Therefore, they can help to better filter the return innovations, and hence to improve the model-inference for both returns and variance.

Two main conclusions can be drawn from the results of our empirical analysis. First, the model-extensions that we suggest seem to improve both the in- and out-of-sample performance of the model to predict realized volatility compared to the standard Realized-GARCH. This finding is justified by a number of goodness-of-fit and prediction-accuracy metrics reported, as well as by a series of equal-prediction performance tests evaluating the relative out-of-sample forecasting performance of the models. Our results show that the significant improvement in the prediction performance of the model comes primarily from the inclusion of the upside and downside intra-day movements of realized variance. Second, allowing for asymmetric/fat-tailed return distributions also plays a crucial role in the accurate filtering of the innovations, and consistently

² There are several papers documenting the importance of RV for several financial applications. For instance Christoffersen et al. (2014), Hansen, Huang, and Wang (2015), Feunou, Jahan-Parvar, and Okou (2017), Feunou and Okou (2019), among others, demonstrate that incorporating RV for option pricing is beneficial. Also, Louzis, Xanthopoulos-Sisinis, and Refenes (2014), Wang, Chen, and Gerlach (2019), and Chen, Watanabe, and Lin (2021) highlight the impact of RV on accurately forecasting tail-risk at multiple horizons ahead.

³ Kourtis, Markellos, and Symeonidis (2016) provide a comparison of the volatility forecasting performance of HAR and Realized-GARCH models across different markets.

⁴ The SGED has been used since Theodossiou (1998) in several financial applications; for instance, it has been used in capturing higher-order risk-premia (Theodossiou and Savva 2016), in option-pricing theory (Theodossiou 2015), in conditional density modelling (Feunou, Jahan-Parvar, and Tédongap 2016) and value-at-risk (VaR) forecasting (Dendramis, Spungin, and Tzavalis 2014). Similarly, the NIG has been shown to perform better than conventional parametric densities in the context of stochastic volatility models (Andersson 2001; Barndorff-Nielsen 1997), discrete-time conditional volatility models (Forsberg and Bollerslev 2002; Jensen and Lunde 2001), VaR forecasting (Venter and de Jongh 2002; Wilhelmsson 2009), as well as option-pricing (Chorro, Guégan, and Ielpo 2012; Eriksson, Ghysels, and Wang 2009; Ghysels and Wang 2014; Stentoft 2008).

helps the identification of the parameters of the volatility process. This boosts the prediction performance of all the augmented GARCH-type specifications that we consider in our analysis.

The remainder of the paper is organized as follows. Section 2 presents a generic framework for the joint dynamics of returns and variance and specifies the conditional variance model suggested by the paper. Section 3 provides details on the joint estimation procedure that we follow. Section 4 provides a detailed discussion on both the in-sample fit and the out-of-sample forecasting performance of the models. Section 5 concludes the paper.

2 The model

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. The conditional PDF of r_t is described by $f_r(r_t | \mathcal{F}_{t-1})$, with $\mathcal{F}_{t-1}$ being the σ -algebra formed by the information-set on the observed variables up to time $t-1$. We denote the conditional mean and variance of returns as $m_t \equiv E[r_t | \mathcal{F}_{t-1}]$ and $h_t \equiv \text{var}[r_t | \mathcal{F}_{t-1}]$, which are linked to the conditional *location* and *scale* parameters of $f_r(r_t | \mathcal{F}_{t-1})$, respectively. The conditional variance function h_t can also be written as $h_t = E[\langle p, p \rangle_{t-1,t} | \mathcal{F}_{t-1}]$, 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 can be denoted as $v_t \equiv \log \langle p, p \rangle_{t-1,t}$ and has a conditional PDF $f_v(v_t | \mathcal{F}_{t-1})$, with conditional mean $E[v_t | \mathcal{F}_{t-1}] = E[\log \langle p, p \rangle_{t-1,t} | \mathcal{F}_{t-1}] = \log E[\langle p, p \rangle_{t-1,t} | \mathcal{F}_{t-1}] + o \simeq \log h_t$, where the convexity term o is negligible. Since v_t is unobservable, a natural measure used in the literature to proxy v_t is the logarithm of RV, where $\log(RV_t)$ is a consistent estimator of $\log \langle p, p \rangle_{t-1,t}$; see for example Barndorff-Nielsen and Shephard (2002a), Barndorff-Nielsen and Shephard (2002b), Barndorff-Nielsen and Shephard (2004a), Barndorff-Nielsen and Shephard (2004b), and Barndorff-Nielsen and Shephard (2006).⁵

Under any (generic) parametric specification for the returns r_t and log-variance v_t dynamics, we can write the *joint* conditional PDF as $f(r_t, v_t | (\mathcal{F}_{t-1}; \theta))$, where θ is a one-dimensional real-valued vector of model parameters, i.e. $\theta \in \mathbb{R}^{k \times 1}$. Given the above definitions, the generic framework of the joint dynamics of r_t and v_t can be written as:

$$r_t = m_t + \sqrt{h_t} z_t \quad (1a)$$

$$v_t = \log h_t + \sigma_v u_t \quad (1b)$$

where $z_t = (r_t - m_t) / \sqrt{h_t}$ and $u_t = (v_t - \log h_t) / \sigma_v$ being jointly $\sim \text{i.i.d.}(0, 1)$; i.e. there are no cross- or serial-correlations between z_t and u_t . Note that σ_v^2 is the conditional variance $\text{Var}[v_t | \mathcal{F}_{t-1}]$, which is assumed constant.⁶ The above framework assumes there is a one-to-one relation between v_t and $\log h_t$. This parsimonious relation can ensure robustness and accuracy of the estimation procedure, as it helps the parameter identification by avoiding over-parametrization and is a very common choice among econometricians (see Bekierman and Gribisch (2016), Hansen and Huang (2016), Koopman and Scharth (2012), and Takahashi, Omori, and Watanabe (2009) among others). For the conditional mean of r_t we follow the standard approach in the literature (see Engle, Lilien, and Robins (1987)), and we assume that $m_t = r_f + \mu \sqrt{h_t}$, which incorporates the risk-free rate r_f , as well as a volatility-in-mean effect capturing a time-varying risk-premium embodied in r_t .

⁵ Details on the estimation of RV and other realized measures can be found in Section 3.

⁶ In the remainder of the paper we will assume that v_t follows the normal distribution so that convexity term o of the mean equation of v_t equals $-\frac{1}{2}\sigma_v^2$. Unreported evidence indicate that the empirical results remain robust to its inclusion in the mean equation of v_t . We also found the time-variation of the variance of v_t to be negligible; hence beyond the scope of this paper.

2.1 Conditional variance function specifications

In this section we introduce a new conditional variance framework which nests the standard GJR threshold GARCH threshold specification (Glosten, Jagannathan, and Runkle 1993) and the HAR family of RV models. We can write the variance function as follows:

$$h_t = b_0 + b_1 h_{t-1} + \tau(t-1) + x(t-1), \quad (2)$$

where

$$\tau(t) = b_2 \epsilon_t^2 + b_3 \mathbb{1}_{\{z_t < 0\}} \epsilon_t^2, \quad \epsilon_t = \sqrt{h_t} z_t, \quad (3)$$

is the news impact function (NIF) of variance (with b_3 capturing the variance asymmetry induced by the well-documented leverage effect),⁷ and

$$x(t) = c_0 RV_t + c_1 RV_t^+ + c_2 RV_t^- + c_3 RV_t^{[5]} + c_4 RV_t^{[20]} + c_5 RQ_t^{1/2} \quad (4)$$

is the component of h_t driven by exogenous variables related to intra-day realized measures. Note that this NIF specification $\tau(t)$ is more flexible than the one in Hansen's Realized GARCH, as it can also accommodate for asymmetric volatility effects.

The rationale behind the choice of the realized measures included in $x(t)$ (Eq. (4)) is motivated by recent empirical findings, and primarily from the literature of HAR models aiming to predict the dynamics of RV. We first include RV as in Hansen, Huang, and Shek (2012), which has been shown to enhance the forecasting performance of conditional/realized variance models. Next, we consider the realized upside and downside semi-variances, denoted as RV^+ and RV^- , respectively. Patton and Sheppard (2015) provide clear evidence supporting the decomposition of RV into its two semi-variance components, as it can be beneficial for volatility forecasting when there are prevalent asymmetries in high-frequency data. They use an HAR-type regression framework to show that the response of RV to the two semi-variances can differ, with the negative realized semi-variance being more informative for predicting future RV. At this point, we should highlight that semi-variances have never been used before into a GARCH-based framework, and therefore it would be interesting to examine (see Section 4) whether the parametric NIF component of the model can absorb similar information to the two semi-variances regarding the asymmetric responses of variance.

Moreover, we incorporate in $x(t)$ the 5- and 20-day (weekly and monthly) moving-average terms of RV, i.e. RV Andersen et al. (2009) and RV Bollerslev et al. (2009), which have been used in the HAR literature to approximate the long-memory behaviour of RV.⁸ Interestingly, the inclusion of these so-called “heterogeneous” terms in $x(t)$ and, hence, in the conditional variance model (2) allows capturing long-memory patterns of h_t exogenously, without affecting the parsimonious parametric structure of the model (see also Huang, Liu, and Wang (2016)). Baillie et al. (2019) provide a very detailed discussion on the long-memory property of RV and its connection with the heterogeneous terms.

Last, the term $RQ_t^{1/2}$ that we include in $x(t)$ constitutes a measure of realized quarticity, which effectively reflects the variance-of-variance. Despite the fact that RV can be a consistent estimator of quadratic variation, it is still subject to measurement error due to finite sampling. It has been shown that the variance of this error is proportional to RQ (see Andersen, Bollerslev, and Meddahi (2005)). Therefore, taking $RQ_t^{1/2}$ into account can absorb this error and correct for attenuation-biases in estimating/forecasting variance. Bollerslev, Patton, and Quaadvlieg (2016) and Cipollini, Gallo, and Otranto (2020), again within the context of an HAR model, exploit RQ to attenuate the bias of the estimated model parameters driven by the measurement error of RV.

⁷ Alternative GARCH-type models may offer different specifications for the NIF specification of variance. However, there is enough consensus that the simple threshold representation of the GJR model performs extremely well, while at the same time being parsimonious and easy to interpret. In Appendix A.3, we consider and estimate alternative specifications for the NIF in order to examine if they can capture the leverage effect better than the GJR model. See also the discussion in Section 4.1.

⁸ The m -day moving average terms are simply obtained as $RV_t^{[m]} \equiv RV_{(t+1-m:t)} = m^{-1} \sum_{j=1}^m RV_{t+1-j}$.

The above general framework, to which we henceforth refer as GARCH-HAR-X class and is described by Eqs. (2)–(4), nests a number of specifications suggested in the literature for modelling conditional and/or realized variance dynamics. These nested specifications include several HAR-type models, as well as the Realized-GARCH and its extensions produced by the additional exogenous realized measures; i.e. from the factors included in $x(t)$. More specifically, for $b_1 = b_2 = b_3 = c_1 = c_2 = c_5 = 0$ we obtain the standard HAR model of Corsi (2009). Similarly, for $b_1 = b_2 = b_3 = c_0 = c_5 = 0$, we get the more flexible SHAR model of Patton and Sheppard (2015), where the capital letter “S” indicates the inclusion of semi-variances in (4). Finally, for $c_1 = c_2 = c_3 = c_4 = c_5 = 0$, the model reduces to the Realized-GARCH, which for simplicity in this paper we denote as GARCH-R. Again, note that our GARCH-R is more generic than the standard Realized-GARCH as it allows for volatility asymmetries through the b_3 term in the NIF (3).

We further utilize this framework to study several extensions to the aforementioned existing models. First, augmenting the standard HAR model to include information regarding the realized variance-of-variance (setting $b_1 = b_2 = b_3 = c_1 = c_2 = 0$) yields a representation that we denote as HARQ, where the capital letter “Q” reflects the realized quarticity term in (3).⁹ This can be combined with the two semi-variances (when $b_1 = b_2 = b_3 = c_0 = 0$) to generate the SHARQ model.

Next, if we allow for $b_2 \neq 0$ and (possibly) $b_3 \neq 0$, the above specifications of the HAR model (namely, the HAR, SHAR, HARQ and SHARQ) are extended to include information from the filtered return innovations. This extension allows capturing magnitude and leverage effects on RV , through out return innovations functions ϵ_t^2 and $I_{\{\epsilon_t < 0\}} \epsilon_t^2$, respectively. We will henceforth refer to these models, simply, as HARz, SHARz, HARQz, SHARQz (where the “z” suffix stands for the filtered return innovation).

Furthermore, we present extensions of the GARCH-R model nested in the above framework. For $c_0 = c_3 = c_4 = c_5 = 0$ we obtain the GARCH-S model which augments the GARCH-R by allowing for upside and downside semi-variances. When $c_1 = c_2 = c_5 = 0$ and $c_0 = c_5 = 0$, we obtain the GARCH-HAR and GARCH-SHAR models, respectively. These two specifications include the heterogeneous (moving-average) terms of RV , as in the HAR-type models literature. Finally, for $c_1 = c_2 = 0$ and $c_0 = 0$, we get the GARCH-HARQ and GARCH-SHARQ models, which incorporate realized quarticity information in the GARCH-HAR and GARCH-SHAR models. Note that the GARCH-SHARQ model constitutes what we call the “full” specification produced by Eqs. (2)–(4), and it encompasses all the specifications considered. We will use GARCH-SHARQ as benchmark for future model comparisons, especially vis-a-vis the GARCH-R and the augmented HAR-type representations using the same exogenous variables. The parameter restrictions for the nested specifications that we discussed are also summarized in Table 1.

3 Data & estimation

3.1 Data

We implement our generalized Realized-GARCH-HAR framework on the S&P500 index.¹⁰ We obtain ultra-high-frequency data (UHF) from the trade and quote (TAQ) Database. This allows us to monitor all the recorded tick-level data on the S&P500 index, at millisecond-level precision. Our dataset covers the period of 1995–2016. Given the UHF data on the index, we construct evenly-spaced observations of 5 min intervals for

⁹ Note that our HARQ representation is slightly different from that in Bollerslev, Patton, and Quaadvlieg (2016); in their paper they only consider the interaction between RQ and RV , instead of using RQ as explanatory variable to predict RV . However, for simplicity we preserve the same name for the model.

¹⁰ The S&P500 data provide an attractive ground for an analysis of this type. The US 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 this market; asymmetric volatility/feedback effects, negative variance risk-premia, significant deviations from normality of both physical and risk-neutral densities, etc.

Table 1: Nested models within the GARCH-HAR-X class.

	h_t	ϵ_t^2	$\epsilon_t^2 \mathbb{I}_{\{z_t < 0\}}$	RV_t	RV_t^+	RV_t^-	$RV_t^{[5]}$	$RV_t^{[20]}$	$RQ_t^{1/2}$
HAR	—	—	—	X	—	—	X	X	—
SHAR	—	—	—	—	X	X	X	X	—
HARz	—	X	X	X	—	—	X	X	—
SHARz	—	X	X	—	X	X	X	X	—
HARQ	—	—	—	X	—	—	X	X	X
SHARQ	—	—	—	—	X	X	X	X	X
HARQz	—	X	X	X	—	—	X	X	X
SHARQz	—	X	X	—	X	X	X	X	X
GARCH-R	X	X	X	X	—	—	—	—	—
GARCH-S	X	X	X	—	X	X	—	—	—
GARCH-HAR	X	X	X	X	—	—	X	X	—
GARCH-SHAR	X	X	X	—	X	X	X	X	—
GARCH-HARQ	X	X	X	X	—	—	X	X	X
GARCH-SHARQ	X	X	X	—	X	X	X	X	X

This table presents the variables included in the different models nested within the GARCH-HAR-X class. The first column corresponds to the model abbreviations used throughout the paper. Recall that $\epsilon_t = \sqrt{h_t} z_t$, where h_t denotes the conditional variance and z_t the return innovation term. RV_t stands for the daily realized variance, while RV_t^+ and RV_t^- denote daily upside and downside realized semi-variances, respectively. The $RV_t^{[5]}$ and $RV_t^{[20]}$ variables are the 5- and 20-day moving average terms of RV , and RQ_t denotes the so-called realized quarticity.

the trading hours between 9:30 and 16:00 (EST) by taking the last price that was recorded within the previous 5 min period. This gives us 78 intra-day observations per day, which we use to calculate the realized measures. In case there is no trading in a specific interval, then the corresponding return is set to zero.

To calculate the realized measure of variance RV_t , included in relation (4), we rely on the heavily-used estimator (see Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002a), Barndorff-Nielsen and Shephard (2002b), Barndorff-Nielsen and Shephard (2004a), Barndorff-Nielsen and Shephard (2004b), and Barndorff-Nielsen and Shephard (2006) among others):

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

where $r_{t,i}^2$ is the i th intra-day log-return at day t defined as $r_{t,i} = p_{t,\frac{i}{n}} - p_{t,\frac{i-1}{n}}$ (with $i = 1, 2, \dots, n$), where p stands for the log-price. This converges in probability to the quadratic variation $\langle p, p \rangle_{t-1,t}$ as $n \rightarrow \infty$, i.e. as the time-interval between two consecutive observations shrinks. For the upside and downside semi-variances we follow Barndorff-Nielsen, Kinnebrock, and Shephard (2008) and Patton and Sheppard (2015). They show that RV can be further decomposed into upside and downside semi-variances, defined as:

$$RV_t^+ = \sum_{i=1}^n r_{t,i}^2 \mathbb{I}_{\{r_{t,i} > 0\}} \quad \text{and} \quad RV_t^- = \sum_{i=1}^n r_{t,i}^2 \mathbb{I}_{\{r_{t,i} \leq 0\}},$$

respectively. The RV decomposition simply implies that $RV_t = RV_t^+ + RV_t^-$. Finally, as an estimator for the realized quarticity we use

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

which, in the absence of jumps, it can be shown to provide a measure proportional to the asymptotic variance of RV ; see Barndorff-Nielsen and Shephard (2002a) and Andersen, Bollerslev, and Diebold (2007).

3.2 A joint maximum likelihood estimation approach

We apply a joint maximum likelihood estimation (MLE) of returns r_t and log-quadratic variation v_t on the system of equations defined by (1a) and (1b), where v_t is approximated by $\log(RV_t)$ and conditional variance h_t is described by (2)–(4). Responding to heavily-documented evidence that return distributions exhibit pronounced skewness and/or kurtosis levels, we consider different parametric PDFs for the return innovations z_t . Beyond the Normal (N) distribution, we consider the cases of the skewed-GED (SGED), as well as the very flexible normal-inverse-Gaussian (denoted as NIG).¹¹

Both the SGED and the (standardized) NIG are two-parameter distributions that have been shown to outperform other parametric distributions in fitting financial return data. This is because they are able to span a significantly wider domain of theoretically feasible skewness-kurtosis combinations, compared to other conventional parametric densities; see discussion in Jondeau and Rockinger (2003).¹² They also nest a range of popular parametric densities (including the Normal) as limiting cases. Capturing asymmetry and fat-tails in returns helps to better filter the return innovations and hence to improve the model-inference for both returns and variance. Ignoring these features might induce inefficiency and bias in the parameter estimates which would impact both the in-sample and out-of-sample performance of the system of Eqs. (1a) and (1b); see discussion in Papantonis, Rompolis, and Tzavalis (2021).

On the other hand, evidence suggests that the distribution of $\log(RV)$ is approximately Normal; see Andersen et al. (2001) and Andersen et al. (2003). Therefore, it follows naturally to assume that the innovation term u in Eq. (1b) is Normally distributed, i.e. $u \sim N(0, 1)$. Also, note that working with $\log(RV)$, instead of RV , absorbs almost entirely the heteroscedasticity in the error term of Eq. (1b); see Bollerslev et al. (2009). Simulation results in Papantonis (2016) and Papantonis, Rompolis, and Tzavalis (2021) show significant improvements in terms of bias and efficiency for the parameter estimates of volatility models when a similar joint MLE framework is considered.

The joint MLE provides a very straightforward and robust method of obtaining parameter estimates and has been used in numerous similar applications (see Hansen, Huang, and Shek (2012), Hansen and Huang (2016), Papantonis (2016), Papantonis, Rompolis, and Tzavalis (2021) and references therein). The joint log-likelihood function that we maximize can be simply expressed as follows:

$$\begin{aligned}\mathcal{L}(r_t, v_t | \mathcal{F}_{t-1}; \vartheta) &\equiv \mathcal{L}^R(r_t | \mathcal{F}_{t-1}; \vartheta) + \mathcal{L}^V(v_t | \mathcal{F}_{t-1}; \vartheta) \\ &= \sum_{t=1}^T \log f_r(r_t | \mathcal{F}_{t-1}; \vartheta) + \sum_{t=1}^T \log f_v(v_t | \mathcal{F}_{t-1}; \vartheta),\end{aligned}\quad (5)$$

where T is the sample size. The total log-likelihood $\mathcal{L}(r_t, v_t | \mathcal{F}_{t-1}; \vartheta)$ is simply obtained as the sum of two independent log-likelihood components for the returns $\mathcal{L}^R(r_t | \mathcal{F}_{t-1}; \vartheta)$ and log-realized variance $\mathcal{L}^V(v_t | \mathcal{F}_{t-1}; \vartheta)$. The parameter vector ϑ can be estimated by solving the following numerical non-linear maximization problem:

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

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:

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

¹¹ A more detailed description of the standardized NIG and SGED is provided in Appendix A.

¹² Specifically, the SGED behaves similarly to the skewed-student- t distribution, therefore has been used in several financial risk applications focusing on capturing fat-tails in returns, for instance when modelling value-at-risk or expected shortfall; see Dendramis, Spungin, and Tzavalis (2014).

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

4 Estimation results

4.1 In-sample estimates

The in-sample MLE parameter estimates, along with their standard errors in parenthesis, are presented in Tables 2–4. Each table presents results under the three alternative density specifications for the return innovation term z_t that we consider; i.e. for the Normal (Table 2), SGED (Table 3) and NIG (Table 4) distributions. Note that the tables include not only the results of the full model, but also those of the several nested specifications discussed in Section 2. For all the specifications considered, the tables report the values of both components of $\mathcal{L}(r_t, v_t | (F_{t-1}; \vartheta))$, i.e. $\mathcal{L}^R(r_t | F_{t-1}; \vartheta)$ and $\mathcal{L}^V(v_t | F_{t-1}; \vartheta)$, in order to provide a more complete picture of the marginal contribution of each component to the total likelihood value; for brevity we denote the likelihood components also as $\mathcal{L}$, $\mathcal{L}^R$ and $\mathcal{L}^V$, respectively. Additionally, we provide the Akaike information criterion (AIC) and Bayesian information criterion (BIC) in order to demonstrate how well the alternative model and/or density specifications fit the data. Finally, to assess the prediction performance of the models, in Table 5 we report results for several widely-used prediction performance metrics calculated using the full sample estimates of conditional variance. These metrics include the mean squared and absolute errors (denoted as MSE and MAE), as well as the heteroscedasticity-adjusted mean squared (HMSE) and mean absolute percentage (MAPE) error metrics, which are more robust to possible heteroscedastic patterns in the forecast errors. The same metrics are also used in the next subsection to evaluate the out-of-sample accuracy of variance predictions. These loss-functions are calculated as

$$L_t^{\text{MSE}} = [\log RV_t - \log h_t]^2, \quad L_t^{\text{MAE}} = |\log RV_t - \log h_t|, \quad L_t^{\text{HMSE}} = \left[\frac{RV_t - h_t}{RV_t} \right]^2 \quad \text{and} \quad L_t^{\text{MAPE}} = \left| \frac{RV_t - h_t}{RV_t} \right| \quad (6)$$

Several interesting conclusions can be drawn from the results shown in Tables 2–5. First, they clearly demonstrate that the full GARCH-SHARQ specification considerably improves upon the in-sample fit and predictive performance of realized variance compared to the HAR model of Corsi (2009) or the GARCH-R model of Hansen, Huang, and Shek (2012). This result can be clearly justified by all the prediction and fit performance metrics reported in the tables. Recall that the GARCH-R and HAR-type models are the key reference models for our comparisons, since the GARCH-SHARQ follows as direct extension of these.

The enhanced performance of the GARCH-SHARQ model with respect to GARCH-R comes primarily from the decomposition of RV into upside and downside realized semi-variances (RV^+ and RV^-) therefore introducing a source of asymmetry (this also holds for the GARCH-S/SHAR/SHARQ models), as well as from the inclusion of realized quarticity RQ in $x(t)$, which helps controlling for attenuation biases in realized variance measurement (refer to the GARCH-HARQ/SHARQ). It becomes obvious that this particular extension of the GARCH-R model increases the variance prediction performance significantly (e.g. the HMSE metric drops from 0.6336 to 0.5056 under the NIG distribution). A similar conclusion can be drawn from the values of the log-likelihood function $\mathcal{L}$ and its two component terms $\mathcal{L}^R$ and $\mathcal{L}^V$ reported in the tables. Note that the values of $\mathcal{L}^R$ and $\mathcal{L}^V$ indicate that the gains in the log-likelihood of the model come primarily from improving its fit into the realized measure of log-variance v_t . Additionally, we find that the improved performance of the GARCH-SHARQ model to some extent is also due to the 5-day and 20-day (heterogeneous) moving-average terms of RV (RV (Andersen et al. 2009) and RV (Bollerslev et al. 2009)), which help approximating the long-memory property of variance (see also Huang, Liu, and Wang (2016)). The standard errors reported below the parameter estimates indicate that all the exogenous variables are significant at conventional levels (5% and below) and, according to AIC and BIC, including them in the equation significantly improves the ability of the model to fit the data.

Table 2: Full sample maximum likelihood parameter estimates for linear GARCH-HAR-X(N) models.

	HAR	SHAR	HARz	SHARz	HARQ	SHARQ	HARQz	SHARQz
μ	0.0371 [0.0134]	0.0373 [0.0134]	0.0408 [0.0134]	0.0394 [0.0134]	0.0370 [0.0134]	0.0375 [0.0134]	0.0411 [0.0134]	0.0410 [0.0134]
b_0	3.7334 [0.3908]	3.9286 [0.3766]	4.0149 [0.3763]	3.9895 [0.3747]	3.5804 [0.3685]	3.6333 [0.3655]	3.8294 [0.3614]	3.8179 [0.3614]
b_1	—	—	—	—	—	—	—	—
b_2	—	—	−0.0400 [0.0036]	−0.0124 [0.0054]	—	—	−0.0223 [0.0037]	−0.0195 [0.0044]
b_3	—	—	0.1216 [0.0072]	0.0606 [0.0108]	—	—	0.0936 [0.0070]	0.0873 [0.0088]
c_0	0.3407 [0.0172]	—	0.2584 [0.0177]	—	0.7518 [0.0237]	—	0.6207 [0.0244]	—
c_1	—	−0.0556 [0.0223]	—	0.0289 [0.0334]	—	0.5189 [0.0350]	—	0.5964 [0.0323]
c_2	—	0.6692 [0.0261]	—	0.5045 [0.0387]	—	0.7964 [0.0239]	—	0.6349 [0.0272]
c_3	0.3964 [0.0234]	0.4469 [0.0225]	0.4487 [0.0229]	0.4558 [0.0226]	0.3043 [0.0210]	0.3414 [0.0212]	0.3544 [0.0211]	0.3561 [0.0212]
c_4	0.1487 [0.0162]	0.1356 [0.0154]	0.1463 [0.0156]	0.1395 [0.0154]	0.1277 [0.0148]	0.1219 [0.0145]	0.1272 [0.0144]	0.1264 [0.0144]
c_5	—	—	—	—	−0.1775 [0.0055]	−0.1400 [0.0068]	−0.1508 [0.0054]	−0.1478 [0.0060]
η	—	—	—	—	—	—	—	—
λ	—	—	—	—	—	—	—	—
$\mathcal{L}^R$	17,696.22	17,718.79	17,704.42	17,709.31	17,732.42	17,740.41	17,732.05	17,732.45
$\mathcal{L}^V$	−5019.80	−4871.16	−4863.74	−4841.68	−4768.40	−4725.40	−4663.16	−4662.86
$\mathcal{L}$	12,676.42	12,847.63	12,840.68	12,867.64	12,964.02	13,015.02	13,068.89	13,069.59
AIC	−25,342.84	−25,683.27	−25,667.36	−25,719.27	−25,916.03	−26,016.03	−26,121.78	−26,121.18
BIC	−25,309.75	−25,643.55	−25,621.03	−25,666.32	−25,876.31	−25,969.69	−26,068.82	−26,061.60
K	5	6	7	8	6	7	8	9
	GARCH-R	GARCH-S	GARCH-HAR	GARCH-SHAR	GARCH-HARQ	GARCH-SHARQ		
μ	0.0430 [0.0134]	0.0419 [0.0134]	0.0435 [0.0134]	0.0427 [0.0134]	0.0413 [0.0134]	0.0416 [0.0134]		
b_0	1.8323 [0.1371]	1.8050 [0.1350]	2.0252 [0.1831]	1.9573 [0.1748]	2.1203 [0.2034]	2.0054 [0.1969]		
b_1	0.6829 [0.0117]	0.6816 [0.0114]	0.5681 [0.0243]	0.5730 [0.0220]	0.4814 [0.0256]	0.5005 [0.0252]		
b_2	−0.0339 [0.0029]	−0.0092 [0.0037]	−0.0390 [0.0031]	−0.0106 [0.0040]	−0.0321 [0.0031]	−0.0220 [0.0039]		
b_3	0.1138 [0.0050]	0.0612 [0.0067]	0.1398 [0.0066]	0.0781 [0.0080]	0.1214 [0.0062]	0.0989 [0.0078]		
c_0	0.2409 [0.0113]	—	0.2248 [0.0137]	—	0.4524 [0.0218]	—		
c_1	—	0.0430 [0.0173]	—	−0.0030 [0.0215]	—	0.3563 [0.0295]		
c_2	—	0.4519 [0.0242]	—	0.4630 [0.0266]	—	0.4999 [0.0242]		
c_3	—	—	0.0425 [0.0208]	0.0460 [0.0187]	0.0615 [0.0180]	0.0529 [0.0175]		
c_4	—	—	0.0652 [0.0080]	0.0597 [0.0076]	0.0670 [0.0084]	0.0636 [0.0081]		
c_5	—	—	—	—	−0.0984 [0.0056]	−0.0853 [0.0060]		
η	—	—	—	—	—	—		
λ	—	—	—	—	—	—		
$\mathcal{L}^R$	17,758.16	17,775.49	17,748.91	17,765.45	17,774.01	17,777.40		
$\mathcal{L}^V$	−4803.51	−4755.93	−4739.31	−4691.87	−4576.14	−4568.83		
$\mathcal{L}$	12,954.64	13,019.56	13,009.61	13,073.58	13,197.88	13,208.57		
AIC	−25,897.29	−26,025.11	−26,003.21	−26,129.15	−26,377.75	−26,397.13		
BIC	−25,857.57	−25,978.78	−25,950.26	−26,069.58	−26,318.18	−26,330.94		
K	6	7	8	9	9	10		

This table presents maximum likelihood estimation (MLE) results for the GARCH-HAR-X class of models defined in Table 1 assuming that the return innovations follow the 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. K is the number of model parameters. Standard errors are shown in brackets.

Table 3: Full sample maximum likelihood parameter estimates for linear GARCH-HAR-X(SGED) models.

	HAR	SHAR	HARz	SHARz	HARQ	SHARQ	HARQz	SHARQz
μ	0.0420 [0.0149]	0.0419 [0.0150]	0.0482 [0.0148]	0.0459 [0.0149]	0.0409 [0.0150]	0.0413 [0.0149]	0.0473 [0.0150]	0.0472 [0.0150]
b_0	3.3310 [0.3936]	3.5336 [0.3791]	3.6432 [0.3785]	3.6084 [0.3767]	3.1686 [0.3702]	3.2321 [0.3673]	3.4414 [0.3626]	3.4315 [0.3631]
b_1	—	—	—	—	—	—	—	—
b_2	—	—	−0.0375 [0.0037]	−0.0105 [0.0054]	—	—	−0.0201 [0.0038]	−0.0178 [0.0045]
b_3	—	—	0.1236 [0.0073]	0.0638 [0.0109]	—	—	0.0950 [0.0071]	0.0896 [0.0089]
c_0	0.3372 [0.0174]	—	0.2504 [0.0179]	—	0.7493 [0.0240]	—	0.6135 [0.0247]	—
c_1	—	−0.0523 [0.0223]	—	0.0286 [0.0332]	—	0.5188 [0.0352]	—	0.5929 [0.0325]
c_2	—	0.6635 [0.0263]	—	0.4909 [0.0388]	—	0.7932 [0.0242]	—	0.6255 [0.0275]
c_3	0.3888 [0.0234]	0.4359 [0.0226]	0.4378 [0.0230]	0.4448 [0.0227]	0.2953 [0.0211]	0.3314 [0.0212]	0.3434 [0.0212]	0.3449 [0.0212]
c_4	0.1453 [0.0162]	0.1343 [0.0155]	0.1455 [0.0156]	0.1388 [0.0155]	0.1255 [0.0149]	0.1201 [0.0146]	0.1263 [0.0145]	0.1256 [0.0145]
c_5	—	—	—	—	−0.1763 [0.0056]	−0.1388 [0.0068]	−0.1489 [0.0054]	−0.1463 [0.0060]
η	1.4332 [0.0288]	1.4376 [0.0289]	1.4293 [0.0287]	1.4315 [0.0288]	1.4431 [0.0290]	1.4452 [0.0291]	1.4382 [0.0289]	1.4384 [0.0289]
λ	−0.0713 [0.0138]	−0.0714 [0.0137]	−0.0710 [0.0132]	−0.0709 [0.0136]	−0.0735 [0.0138]	−0.0727 [0.0137]	−0.0720 [0.0138]	−0.0720 [0.0139]
Skew	−0.1817	−0.1810	−0.1816	−0.1809	−0.1854	−0.1831	−0.1825	−0.1823
Kurt	3.9491	3.9375	3.9595	3.9536	3.9242	3.9183	3.9362	3.9356
$\mathcal{L}^R$	17,828.06	17,849.71	17,836.58	17,840.82	17,861.87	17,868.51	17,861.68	17,861.88
$\mathcal{L}^V$	−4988.08	−4841.32	−4830.21	−4808.78	−4739.42	−4696.98	−4632.22	−4631.92
$\mathcal{L}$	12,839.97	13,008.39	13,006.37	13,032.04	13,122.45	13,171.54	13,229.46	13,229.97
AIC	−25,665.95	−26,000.77	−25,994.74	−26,044.08	−26,228.91	−26,325.07	−26,438.92	−26,437.93
BIC	−25,619.61	−25,947.82	−25,935.16	−25,977.88	−26,175.95	−26,265.50	−26,372.73	−26,365.12
K	7	8	9	10	8	9	10	11
	GARCH-R	GARCH-S	GARCH-HAR	GARCH-SHAR	GARCH-HARQ	GARCH-SHARQ		
μ	0.0486 [0.0151]	0.0462 [0.0151]	0.0497 [0.0151]	0.0474 [0.0151]	0.0459 [0.0151]	0.0459 [0.0150]		
b_0	1.7720 [0.1404]	1.7491 [0.1382]	1.8945 [0.1843]	1.8344 [0.1765]	1.9665 [0.2062]	1.8633 [0.1995]		
b_1	0.6746 [0.0124]	0.6737 [0.0120]	0.5593 [0.0252]	0.5647 [0.0229]	0.4666 [0.0261]	0.4861 [0.0256]		
b_2	−0.0306 [0.0030]	−0.0062 [0.0038]	−0.0357 [0.0032]	−0.0078 [0.0041]	−0.0289 [0.0032]	−0.0191 [0.0040]		
b_3	0.1132 [0.0051]	0.0612 [0.0068]	0.1393 [0.0067]	0.0787 [0.0081]	0.1211 [0.0063]	0.0994 [0.0080]		
c_0	0.2395 [0.0117]	—	0.2229 [0.0139]	—	0.4551 [0.0222]	—		
c_1	—	0.0448 [0.0177]	—	0.0004 [0.0219]	—	0.3629 [0.0300]		
c_2	—	0.4470 [0.0247]	—	0.4559 [0.0270]	—	0.5006 [0.0246]		
c_3	—	—	0.0386 [0.0211]	0.0419 [0.0191]	0.0599 [0.0182]	0.0514 [0.0177]		
c_4	—	—	0.0686 [0.0081]	0.0632 [0.0077]	0.0710 [0.0086]	0.0675 [0.0083]		
c_5	—	—	—	—	−0.0996 [0.0057]	−0.0870 [0.0061]		
η	1.4502 [0.0294]	1.4574 [0.0296]	1.4465 [0.0292]	1.4533 [0.0294]	1.4553 [0.0293]	1.4566 [0.0294]		
λ	−0.0718 [0.0142]	−0.0719 [0.0144]	−0.0714 [0.0142]	−0.0711 [0.0142]	−0.0728 [0.0139]	−0.0723 [0.0139]		
Skew	−0.1800	−0.1790	−0.1795	−0.1777	−0.1815	−0.1799		
Kurt	3.9047	3.8863	3.9141	3.8962	3.8921	3.8885		
$\mathcal{L}^R$	17,875.72	17,889.54	17,870.04	17,882.95	17,893.30	17,895.70		
$\mathcal{L}^V$	−4771.74	−4725.34	−4707.22	−4660.95	−4546.11	−4538.94		
$\mathcal{L}$	13,103.98	13,164.20	13,162.83	13,222.00	13,347.19	13,356.76		

Table 3: (continued)

	GARCH-R	GARCH-S	GARCH-HAR	GARCH-SHAR	GARCH-HARQ	GARCH-SHARQ
AIC	−26,191.96	−26,310.40	−26,305.65	−26,422.00	−26,672.39	−26,689.52
BIC	−26,139.01	−26,250.83	−26,239.46	−26,349.19	−26,599.57	−26,610.08
K	8	9	10	11	11	12

This table presents maximum likelihood estimation (MLE) results for the GARCH-HAR-X class of models defined in Table 1 assuming that the return innovations follow the SGED 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. K is the number of model parameters. Standard errors are shown in brackets.

Table 4: Full sample maximum likelihood parameter estimates for linear GARCH-HAR-X(NIG) models.

	HAR	SHAR	HARz	SHARz	HARQ	SHARQ	HARQz	SHARQz
μ	0.0392 [0.0173]	0.0393 [0.0172]	0.0471 [0.0172]	0.0440 [0.0172]	0.0390 [0.0172]	0.0394 [0.0171]	0.0468 [0.0171]	0.0466 [0.0172]
b_0	3.2043 [0.3659]	3.3970 [0.3534]	3.5057 [0.3535]	3.4730 [0.3519]	3.0709 [0.3466]	3.1268 [0.3441]	3.3332 [0.3406]	3.3255 [0.3404]
b_1	—	—	—	—	—	—	—	—
b_2	—	—	−0.0335 [0.0035]	−0.0092 [0.0051]	—	—	−0.0176 [0.0036]	−0.0159 [0.0043]
b_3	—	—	0.1153 [0.0068]	0.0615 [0.0103]	—	—	0.0892 [0.0067]	0.0852 [0.0084]
c_0	0.3120 [0.0162]	—	0.2301 [0.0167]	—	0.6971 [0.0225]	—	0.5705 [0.0232]	—
c_1	—	−0.0447 [0.0209]	—	0.0307 [0.0311]	—	0.4869 [0.0329]	—	0.5552 [0.0304]
c_2	—	0.6126 [0.0246]	—	0.4464 [0.0362]	—	0.7381 [0.0227]	—	0.5795 [0.0258]
c_3	0.3631 [0.0218]	0.4070 [0.0211]	0.4109 [0.0215]	0.4166 [0.0213]	0.2784 [0.0198]	0.3117 [0.0199]	0.3236 [0.0199]	0.3248 [0.0199]
c_4	0.1382 [0.0151]	0.1293 [0.0145]	0.1387 [0.0146]	0.1332 [0.0145]	0.1209 [0.0139]	0.1162 [0.0137]	0.1219 [0.0136]	0.1214 [0.0136]
c_5	—	—	—	—	−0.1642 [0.0052]	−0.1302 [0.0064]	−0.1391 [0.0051]	−0.1372 [0.0056]
η	55.3203 [5.3562]	56.8253 [4.3716]	53.1010 [4.5842]	54.9382 [5.1462]	51.6319 [4.5846]	53.7877 [5.9446]	51.1707 [4.6025]	51.3957 [4.7125]
λ	−0.6989 [0.0104]	−0.6954 [0.0105]	−0.6963 [0.0105]	−0.6961 [0.0105]	−0.6930 [0.0107]	−0.6919 [0.0107]	−0.6919 [0.0107]	−0.6919 [0.0107]
Skew	−0.2819	−0.2768	−0.2867	−0.2818	−0.2893	−0.2830	−0.2902	−0.2895
Kurt	3.1602	3.1549	3.1661	3.1605	3.1697	3.1626	3.1709	3.1701
$\mathcal{L}^R$	18,026.91	18,045.10	18,035.71	18,038.15	18,060.83	18,065.43	18,061.25	18,061.25
$\mathcal{L}^V$	−4953.26	−4808.52	−4796.64	−4775.34	−4708.11	−4666.18	−4601.30	−4600.98
$\mathcal{L}$	13,073.66	13,236.58	13,239.07	13,262.81	13,352.72	13,399.25	13,459.96	13,460.27
AIC	−26,133.31	−26,457.16	−26,460.13	−26,505.62	−26,689.45	−26,780.50	−26,899.92	−26,898.55
BIC	−26,086.97	−26,404.21	−26,400.56	−26,439.43	−26,636.49	−26,720.92	−26,833.72	−26,825.74
K	7	8	9	10	8	9	10	11
	GARCH-R	GARCH-S	GARCH-HAR	GARCH-SHAR	GARCH-HARQ	GARCH-SHARQ		
μ	0.0514 [0.0171]	0.0486 [0.0170]	0.0520 [0.0171]	0.0494 [0.0171]	0.0475 [0.0170]	0.0476 [0.0170]		
b_0	1.6796 [0.1313]	1.6554 [0.1290]	1.8057 [0.1733]	1.7489 [0.1658]	1.8992 [0.1950]	1.8060 [0.1890]		
b_1	0.6770 [0.0123]	0.6770 [0.0119]	0.5624 [0.0253]	0.5687 [0.0231]	0.4681 [0.0264]	0.4870 [0.0260]		
b_2	−0.0271 [0.0028]	−0.0051 [0.0036]	−0.0320 [0.0030]	−0.0069 [0.0039]	−0.0263 [0.0031]	−0.0178 [0.0038]		
b_3	0.1051 [0.0048]	0.0583 [0.0064]	0.1298 [0.0063]	0.0754 [0.0076]	0.1139 [0.0060]	0.0950 [0.0075]		
c_0	0.2215 [0.0109]	—	0.2065 [0.0130]	—	0.4229 [0.0209]	—		
c_1	—	0.0454 [0.0166]	—	0.0060 [0.0206]	—	0.3419 [0.0284]		
c_2	—	0.4082 [0.0228]	—	0.4163 [0.0251]	—	0.4625 [0.0231]		
c_3	—	—	0.0344 [0.0199]	0.0369 [0.0181]	0.0562 [0.0174]	0.0486 [0.0169]		

Table 4: (continued)

	GARCH-R	GARCH-S	GARCH-HAR	GARCH-SHAR	GARCH-HARQ	GARCH-SHARQ
c_4	—	—	0.0655 [0.0077]	0.0607 [0.0073]	0.0687 [0.0081]	0.0655 [0.0079]
c_5	—	—	—	—	−0.0926 [0.0054]	−0.0816 [0.0058]
η	50.1807 [4.4828]	52.4872 [3.9657]	49.3305 [4.6749]	52.3148 [4.4833]	49.8684 [4.6023]	51.1580 [3.8995]
λ	−0.6891 [0.0108]	−0.6872 [0.0109]	−0.6897 [0.0108]	−0.6881 [0.0108]	−0.6863 [0.0109]	−0.6859 [0.0109]
Skew	−0.2918	−0.2846	−0.2946	−0.2854	−0.2916	−0.2877
Kurt	3.1733	3.1651	3.1765	3.1659	3.1735	3.1690
$\mathcal{L}^R$	18,073.95	18,084.96	18,070.49	18,080.10	18,091.53	18,092.74
$\mathcal{L}^V$	−4739.17	−4693.48	−4675.14	−4629.49	−4516.30	−4509.27
$\mathcal{L}$	13,334.78	13,391.48	13,395.35	13,450.61	13,575.24	13,583.46
AIC	−26,653.57	−26,764.96	−26,770.70	−26,879.23	−27,128.48	−27,142.93
BIC	−26,600.61	−26,705.38	−26,704.51	−26,806.41	−27,055.66	−27,063.49
K	8	9	10	11	11	12

This table presents maximum likelihood estimation (MLE) results for the GARCH-HAR-X class of models defined in Table 1 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 Criteria, respectively. K is the number of model parameters. Standard errors are shown in brackets.

Second, our results clearly indicate that this generalized/augmented GARCH-HAR-X framework also provides the means for useful extensions to the standard HAR-type model-variants suggested in the literature to predict RV . As can be seen from Tables 2–4, comparing the standard HAR-type models to their corresponding GARCH-HAR-X counterparts shows that all the GARCH-model features have significant impact on RV_t , even in the presence of the other HAR-model terms. Their sign and magnitude is in line with other studies in the standard GARCH model literature, and they seem to improve both the joint fit to returns and variance and the performance of the HAR model to predict the realized variance measure v_t . The values of the loss-metrics reported in Table 5 indicate a sizeable performance improvement (for instance the HMSE drops from 0.6647 to 0.5750 when comparing the SHAR of Patton and Sheppard (2015) to the corresponding GARCH-SHAR), while those of $\mathcal{L}$, $\mathcal{L}^R$ and $\mathcal{L}^V$ again reflect that most of the gains come from improving its fit into the realized measure of log-variance v_t .

The improved performance of the augmented HAR-type models that include the GARCH terms stems firstly from the autoregressive conditional variance component h_{t-1} , differentiating the impact from that of lagged RV . As seen by the estimates of the models, the inclusion of h_{t-1} in the HAR model does not interact with the lagged realized measures, which means that it reflects different dynamic effects of conditional variance. A second source of performance improvement over the standard HAR models comes from adding the return innovation functions, i.e. through the NIF $\tau(t)$; refer to the HARz, SHARz, HARQz, SHARQz models. Doing so allows us to capture asymmetries in realized variance in a similar way as in GARCH models. Other studies have tried to capture leverage effects in HAR models by simply using lagged returns (Audrino and Hu 2016; Corsi and Renò 2012). Even though this approach might look very similar conceptually, we have found that using lagged filtered innovations z_t instead (as in a GARCH model) performs much better in fitting the data. This is an interesting finding and can be justified by the fact that filtered innovations provide a “cleaner” source of news information and have time-invariant distributional properties (at least for the first two moments of the distribution) since they are not contaminated by conditional variance or variance-in-mean effects, unlike raw returns.

A third conclusion that can be drawn from our results is related to the importance of the NIF component $\tau(t)$. It appears that including in h_t the exogenous realized variables through $x(t)$ cannot fully absorb the

Table 5: Full sample variance loss-function results for GARCH-HAR-X models.

	MSE	MAE	HMSE	MAPE
Panel A: GARCH-HAR-X(N)				
HAR	0.3587	0.4649	0.8972	0.5752
SHAR	0.3400	0.4520	0.8343	0.5526
HARz	0.3390	0.4514	0.8211	0.5518
SHARz	0.3364	0.4495	0.8124	0.5483
HARQ	0.3276	0.4459	0.6897	0.5313
SHARQ	0.3225	0.4423	0.6880	0.5267
HARQz	0.3154	0.4364	0.6659	0.5177
SHARQz	0.3153	0.4364	0.6664	0.5178
GARCH-R	0.3318	0.4458	0.7897	0.5428
GARCH-S	0.3261	0.4427	0.7478	0.5356
GARCH-HAR	0.3242	0.4417	0.7455	0.5337
GARCH-SHAR	0.3187	0.4382	0.7146	0.5268
GARCH-HARQ	0.3056	0.4292	0.6258	0.5054
GARCH-SHARQ	0.3048	0.4288	0.6252	0.5050
Panel B: GARCH-HAR-X(SGED)				
HAR	0.3547	0.4618	0.8389	0.5578
SHAR	0.3364	0.4489	0.7824	0.5366
HARz	0.3350	0.4481	0.7665	0.5351
SHARz	0.3324	0.4463	0.7591	0.5319
HARQ	0.3242	0.4430	0.6449	0.5162
SHARQ	0.3193	0.4394	0.6443	0.5119
HARQz	0.3119	0.4331	0.6217	0.5026
SHARQz	0.3119	0.4332	0.6222	0.5027
GARCH-R	0.3280	0.4427	0.7402	0.5276
GARCH-S	0.3226	0.4397	0.7024	0.5210

parameters b_2 and b_3 of $\tau(t)$, which control for asymmetric effects of lagged innovations on h_t . This can be seen by directly comparing the estimates of the HAR-type models (HAR, SHAR, HARQ and SHARQ) to those of their corresponding NIF-augmented representations (HARz, SHARz, HARQz and SHARQz), which have been intentionally designed to include the $\tau(t)$ terms. The parameters of $\tau(t)$ are statistically significant even in the presence of the two semi-variances RV^+ and RV^- .¹³ This result implies that the realized semi-variances bring different information to the conditional variance equation. Almost certainly, this source of information relates mostly to short-term effects from signed-jump variation. Moreover, the different values of the coefficient estimates for RV^+ and RV^- indicate that there are prevalent asymmetries in the responses of the conditional variance to positive and negative jumps. These asymmetric responses are directly related to realized skewness in intra-day data. In the absence of realized skewness the two semi-variances would be almost identical, hence the signed jump variation would collapse to zero, therefore having no contribution in forecasting conditional variance. This finding is in line with the HAR-type regression results discussed in Patton and Sheppard (2015). Summing up, the above results highlight that there are two channels of volatility asymmetry; the one originates

¹³ Let us highlight here that this might as well depend on the parametric choice for the NIF. Therefore, in order to examine if the GJR model provides a good specification for the leverage effect, we also consider some alternative and more generic parametrizations of the $\tau(t)$ function, such as the smooth-transition GARCH of González-Rivera (1998) as well as a generalized version of the GJR which allows the threshold parameter to be different from zero (i.e. the threshold is treated as an unknown parameter that can be estimated from the data, as in Zakoian (1994)). For reasons of space, in Appendix A.3 we present results only for the full GARCH-SHARQ specification. These results suggest that the GJR provides a robust specification for $\tau(t)$. This is true across the three distributions considered.

Table 5: (continued)

	MSE	MAE	HMSE	MAPE
Panel B: GARCH-HAR-X(SGED)				
GARCH-HAR	0.3205	0.4387	0.6975	0.5188
GARCH-SHAR	0.3152	0.4352	0.6701	0.5123
GARCH-HARQ	0.3024	0.4262	0.5856	0.4916
GARCH-SHARQ	0.3016	0.4259	0.5854	0.4914
Panel C: GARCH-HAR-X(NIG)				
HAR	0.3503	0.4577	0.7069	0.5183
SHAR	0.3324	0.4452	0.6647	0.5007
HARz	0.3310	0.4441	0.6522	0.4991
SHARz	0.3285	0.4425	0.6457	0.4964
HARQ	0.3206	0.4384	0.5505	0.4819
SHARQ	0.3158	0.4345	0.5514	0.4781
HARQz	0.3085	0.4281	0.5338	0.4698
SHARQz	0.3084	0.4282	0.5342	0.4699
GARCH-R	0.3242	0.4388	0.6336	0.4932
GARCH-S	0.3189	0.4360	0.6022	0.4878
GARCH-HAR	0.3168	0.4348	0.5978	0.4854
GARCH-SHAR	0.3116	0.4313	0.5750	0.4798
GARCH-HARQ	0.2991	0.4216	0.5057	0.4608
GARCH-SHARQ	0.2984	0.4214	0.5056	0.4608

This table presents in-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), and mean absolute percentage error (MAPE). Panel A reports the results assuming that return innovations follow the standard Normal distribution. In Panel B and C return innovations follow the SGED and NIG distributions, respectively.

from the asymmetric responses to intra-day realized upside/downside variation, while the other is driven from non-linearities with respect to lagged return innovations. Both of these channels are found to carry statistically significant information for volatility forecasting.

Fourth, the impact of realized quarticity RQ on conditional variance is found to be negative and significant, with a sizeable contribution to the likelihood. This indicates that beyond the skewness effect on conditional variance that comes from signed-jump variations, there is also a material volatility-of-volatility effect related to intra-day realized kurtosis. Bollerslev, Patton, and Quaedvlieg (2016) and Cipollini, Gallo, and Otranto (2020), in a slightly different context, use the RQ in order to indirectly account for time-variation in the HAR parameters and they also report results similar in direction and magnitude; they attribute this effect to attenuation-biases in measuring realized variance.

Finally, regarding the alternative densities that we used to model the returns innovations z_t , our results show that the NIG distribution provides much better fit to the actual data, followed by the SGED. This improvement is significant not only in terms of joint log-likelihood and information criteria, but also in terms of in-sample variance loss-functions. This result is robust across all the restricted specifications nested within our generalized GARCH-HAR-X framework. In particular, our results indicate that the NIG distribution is able to capture a much higher level of negative skewness in the data (varying around -0.29 across models), while producing less pronounced kurtosis (approximately 3.17). The SGED distribution on the other hand generates lower values of negative skewness (approximately -0.19), while generating higher values for kurtosis (approximately 3.86). It is evident that the fit improves significantly when accounting for higher negative skewness levels. The semi-heavy behaviour of the NIG tails seems to be more flexible in capturing the skewness of the actual data, hence improving the filtering of z_t and the identification of the parameters of h_t , especially for the NIF component.

4.2 Out-of-sample results

In order to assess the out-of-sample (OOS) performance of our models, we split the sample in two parts; we use the first 3000 observations (4/01/1995 to 30/11/2006) to initialize fitting the models (“training” sample) and the remaining approximately 2500 observations (that is from 1/12/2006 onwards) to test the OOS performance. In this way, we maintain a good balance between having a significantly large sample that allows us to accurately estimate model-parameters and also keeping a considerably large period for our OOS testing (Diebold 2015). Note that our OOS period covers the major financial crisis of 2007–08 and other short-lived shocks such as the flash-crash of 2010; this will enable us to examine the robustness of our models against those severe events as well.

We carry out a daily rolling re-estimation of the models to obtain one-step ahead predictions, always keeping the window-size fixed at 3000 observations.¹⁴ Despite the fact that the GARCH-based framework is by construction optimized to provide one-step ahead conditional variance forecasts at a daily frequency, another reason we focus on evaluating the one-step ahead variance predictions is because it has been shown in numerous theoretical econometric studies that the test statistics used in the assessment of equivalent prediction accuracy in pair-wise model-comparisons have good size and power properties primarily for one-step ahead forecasts.¹⁵

We carry out a series of comparative tests; these will help us shed more light to the model-performance and ensure that our findings from the full-sample analysis also hold OOS. In order to facilitate the comparisons we use the same variance loss-functions as in the in-sample analysis and we also employ the reality check (RC) introduced in White (2000). This test is asymptotically valid for both non-nested and nested model-comparisons when the OOS estimates are obtained using a rolling-window. RC is a very common choice among similar studies (see for instance Hansen and Lunde (2005) as well as Bollerslev, Patton, and Quaadvlieg (2016)), since it provides a very straightforward framework for comparing whether the loss of a model A is (statistically) significantly lower than that of a benchmark model B.¹⁶ The statistical hypotheses for the RC test can be simply formulated as

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

$$H_1 : E[\hat{g}_{t+1}] > 0$$

$$\text{with } \hat{g}_{t+1} \equiv L_B(RV_{t+1}, h_{t+1}^B | \hat{\theta}_t^B) - L_A(RV_{t+1}, h_{t+1}^A | \hat{\theta}_t^A)$$

where $L(RV_{t+1}, h_{t+1} | \hat{\theta}_t)$ can be any of the loss functions in (6). In our pair-wise comparisons model A will always be the full GARCH-SHARQ model whereas we (sequentially) consider all the restricted specifications as benchmark models B. Note that during the implementation of the RCs we use 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 in Politis and White (2004), taking into account the correction discussed in Patton, Politis, and White (2009).

We compare the prediction performance of the full GARCH-SHARQ model relative to all the restricted cases. As before, we are primarily interested in the comparisons vis-a-vis the standard GARCH-R model and the nested HAR-type variants. In Tables 6 and 7 we show the OOS variance loss-functions and RC p -values, respectively, based on the MSE, MAE, HMSE and MAPE metrics. To save some space we report the results for the NIG distribution, since it was found to outperform both the Normal and SGED.

¹⁴ Obtaining forecasts based on rolling re-estimation of model parameters is also robust to the occurrence of possibly unaccounted structural breaks; this problem is also discussed in Giraitis et al. (2015) and Giraitis, Kapetanios, and Price (2013).

¹⁵ See West (2006) and Clark and McCracken (2013) for a more thorough discussion. Hansen and Lunde (2005) (and similar studies) also focus on one-step ahead forecasts when comparing volatility models.

¹⁶ A notable advantage of White’s RC when used for multiple model comparison 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 6: Out-of-sample variance loss-function results for GARCH-HAR-X(NIG) models.

	MSE	MAE	HMSE	MAPE
GARCH-SHARQ	0.3613	0.4627	0.6325	0.5062
GARCH-HARQ	0.3603	0.4616	0.6274	0.5027
GARCH-SHAR	0.3807	0.4808	0.8222	0.5715
GARCH-HAR	0.3893	0.4863	0.8735	0.5815
GARCH-S	0.3869	0.4841	0.8461	0.5762
GARCH-R	0.3957	0.4884	0.9150	0.5858
SHARQz	0.3725	0.4697	0.7074	0.5253
HARQz	0.3717	0.4690	0.7049	0.5242
SHARQ	0.3791	0.4744	0.7240	0.5333
HARQ	0.3797	0.4743	0.7036	0.5269
SHARz	0.4081	0.4948	0.9918	0.5977
HARz	0.4130	0.4979	1.0059	0.6032
SHAR	0.4082	0.4946	1.0076	0.5985
HAR	0.4314	0.5100	1.0892	0.6241

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 assuming that return innovations follow the NIG distribution. It reports the mean squared error (MAE), mean absolute error (MAE), heteroskedasticity-adjusted mean squared error (HMSE), and mean absolute percentage error (MAPE). We apply a daily rolling re-estimation of the models with a fixed window size of 3000 observations.

Table 7: Reality check tests results comparing restricted models within the GARCH-HAR-X(NIG) to the full GARCH-SHARQ specification.

	MSE	MAE	HMSE	MAPE
GARCH-HARQ	1.0000	1.0000	1.0000	1.0000
GARCH-SHAR	0.0001	0.0000	0.0000	0.0000
GARCH-HAR	0.0000	0.0000	0.0000	0.0000
GARCH-S	0.0000	0.0000	0.0000	0.0000
GARCH-R	0.0000	0.0000	0.0000	0.0000
SHARQz	0.0136	0.0074	0.0014	0.0001
HARQz	0.0175	0.0168	0.0023	0.0001
SHARQ	0.0021	0.0005	0.0002	0.0000
HARQ	0.0006	0.0003	0.0010	0.0000
SHARz	0.0000	0.0000	0.0000	0.0000
HARz	0.0000	0.0000	0.0000	0.0000
SHAR	0.0000	0.0000	0.0000	0.0000
HAR	0.0000	0.0000	0.0000	0.0000

This table reports the p -values of Reality Check tests for the GARCH-HAR-X class of models with NIG-distributed return innovations. All (nested) models, shown in the first column of the table, are compared to the full GARCH-SHARQ specification.

Table 6 results indicate that our extended specifications outperform the benchmark GARCH-R and HAR-type models, corroborating the findings of the in-sample analysis. Overall, the GARCH-HARQ model was found to yield the best OOS performance, followed closely by the GARCH-SHARQ. For example, the GARCH-HARQ representation reduces the MSE from 0.3957 under the GARCH-R or 0.3797 under the HARQ model to 0.3603. Results for the variance percentage errors (i.e. for the HMSE metric) are also very interesting. Again, if we compare the GARCH-R and HARQ models to our augmented GARCH-HARQ specification we see that HMSE drops significantly from 0.9150 and 0.7036 to 0.6274. These findings reveal that the GARCH-HARQ (and the full GARCH-SHARQ) model can not only shrink the magnitude of the average errors, but also better capture extreme errors relative to the actual realized variance levels. The above arguments can be further

justified by the p -values of the RC tests that we have performed. Results in Table 7 suggest that all of the aforementioned loss-function improvements are statistically significant at conventional levels of statistical significance (bootstrapped p -values are below 5%, in most cases close to zero).

Turning again the discussion to the relative contribution of the additional variables that extend the GARCH-R and HAR-type models to the GARCH-SHARQ framework, we observe that the results of the OOS exercise are consistent with those of the in-sample (full-sample) analysis. In particular, the performance metrics indicate that the GARCH-S improves upon the standard GARCH-R, highlighting the importance of the RV^+ and RV^- semi-variances. Similar conclusions can be drawn for the heterogeneous components RV (Andersen et al. 2009) and RV (Bollerslev et al. 2009) (e.g. by comparing the GARCH-HAR to the GARCH-R) and the realized quarticity measure RQ (by comparing the GARCH-SHARQ to the GARCH-SHAR). The above results become more clear when looking at the HMSE or MAPE metrics. As can be seen by the estimates of these metrics the majority of the prediction gains come firstly from correcting for attenuation-biases through RQ and secondly through decomposing RV into its two upside/downside semi-variance components RV^+ and RV^- . The relative contribution of the heterogeneous terms is less pronounced but still very significant, accommodating for long-memory patterns beyond the persistency inherent in the parametric structure of the GARCH-based framework.

Looking into the HAR-type model extensions, we conclude again that the OOS results confirm the in-sample ones. The reported prediction performance metrics indicate that the OOS gains come primarily from the extension of the HAR to include information from filtered return innovations (e.g. compare SHARQ to SHARQz), followed by the extensions also incorporating filtered conditional variance h_{t-1} (compare for instance the SHARQz to the GARCH-SHARQ).

Finally, results for the OOS exercise evaluating the impact of the distributional assumption for the return innovations z_t on the prediction performance of the models are presented in Tables 8 and 9. In particular, Table 8 presents the variance loss-function results under all three parametric densities that we employ for the returns distribution and, similarly, Table 9 presents the RC test results for the models using NIG innovations for z_t versus those assuming SGED or Normal. Note that the models with SGED and/or Normal are the benchmark models when compared to the NIG, since the NIG has been generally found to perform more flexibly. The table results clearly indicate that, for all loss/prediction metrics, the NIG distribution enhances the performance of variance forecasts compared to both SGED and Normal distribution, and this holds across all model specifications. The superiority of the NIG distribution can be statistically confirmed by the p -values of the RC tests. These are almost zero, thus clearly rejecting the null hypothesis suggesting that the SGED (or Normal) distribution is equivalent to the NIG in terms of prediction performance.

4.3 Robustness of OOS results to longer horizons

The OOS results of the previous section are focused on one-period ahead, where inference procedures on the prediction performance have better size and power properties as noted before. In this section, we examine the OOS forecasting performance of the models for longer horizons ahead to see if our results remain robust. To obtain the multi-period ahead predictions, we rely on the “direct” forecasting approach (see for instance Clark and McCracken (2005), Marcellino, Stock, and Watson (2006), Clark and McCracken (2013), and Ghysels et al. (2019)) and we employ the following horizon-specific regression model as implied by the full GARCH-SHARQ specification:

$$\widetilde{RV}_{t+s} = \beta_{0,s} + \beta_{h,s} \widetilde{h}_{t+1} + \beta'_{x,s} \widetilde{x}(t) + \sigma_{u,s} u_{t+s} \quad (7)$$

where we forecast the log-realized variance $\widetilde{RV}_{t+s}$ at multiple periods ahead, while also using the log-transformations of the exogenous variables in $\widetilde{x}(t)$ as “control” variables (since these are also observable in $\mathcal{F}_t$). In the equation above (Eq. (7)) we have $\widetilde{x}'(t) = \left\{ \widetilde{RV}_t^+, \widetilde{RV}_t^-, \widetilde{RV}_t^{[5]}, \widetilde{RV}_t^{[20]}, \widetilde{RQ}_t \right\}'$ (with corresponding vector of regression coefficients $\beta'_{x,s}$), where the tilde notation has been used to denote the log-transformations

Table 8: Out-of-sample variance loss-function results for GARCH-HAR-X models under different distributional assumptions for the return innovations.

	MSE	MAE	HMSE	MAPE
GARCH-SHARQ(NIG)	0.3613	0.4627	0.6325	0.5062
GARCH-SHARQ(SGED)	0.3689	0.4660	0.7088	0.5314
GARCH-SHARQ(N)	0.3660	0.4675	0.7452	0.5425
GARCH-HAR(NIG)	0.3893	0.4863	0.8735	0.5815
GARCH-HAR(SGED)	0.3974	0.4926	0.9946	0.6165
GARCH-HAR(N)	0.4023	0.4959	1.0471	0.6308
GARCH-SHAR(NIG)	0.3807	0.4808	0.8222	0.5715
GARCH-SHAR(SGED)	0.3885	0.4874	0.9346	0.6057
GARCH-SHAR(N)	0.3932	0.4906	0.9814	0.6192
GARCH-S(NIG)	0.3869	0.4841	0.8461	0.5762
GARCH-S(SGED)	0.3942	0.4900	0.9588	0.6101
GARCH-S(N)	0.3991	0.4935	1.0093	0.6242
GARCH-R(NIG)	0.3957	0.4884	0.9150	0.5858
GARCH-R(SGED)	0.4033	0.4945	1.0397	0.6212
GARCH-R(N)	0.4084	0.4980	1.0962	0.6361
SHAR(NIG)	0.4082	0.4946	1.0076	0.5985
SHAR(SGED)	0.4163	0.5009	1.1500	0.6359
SHAR(N)	0.4211	0.5043	1.2106	0.6515
HAR(NIG)	0.4314	0.5100	1.0892	0.6241
HAR(SGED)	0.4403	0.5160	1.2526	0.6648
HAR(N)	0.4456	0.5196	1.3217	0.6820

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 assuming that return innovations follow the standard Normal, SGED, and NIG distribution. It reports the mean squared error (MSE), mean absolute error (MAE), heteroskedasticity-adjusted mean squared error (HMSE), and mean absolute percentage error (MAPE). The definitions of these metrics are given in (6). We apply a daily rolling re-estimation of the models with a fixed window size of 3000 observations.

of the variables in $\tilde{\mathbf{x}}$.¹⁷ The conditional variance dynamics which we incorporate in the regression model through $\tilde{h}_{t+1}$ (known at time t) are obtained from the previous one-step ahead forecasting exercise, i.e. by daily re-estimating the model using a rolling-window approach. Using these estimates, alongside with the observable exogenous variables, allows us to obtain multi-step OOS predictions of the log-realized variance $\tilde{RV}_{t+s}$ by calculating the conditional expectation $E[\tilde{RV}_{t+s}|\mathcal{F}_t]$, at each point in time t .

There are several reasons behind our decision to run the OOS predictive regressions using log-transformations. Firstly, we want to maintain consistency with log-variance Eq. (1b), as this is the one we actually estimate in the rolling-sample MLE. Secondly, predicting the log-transformation of variance (at any horizon) is the best way to ensure that we will obtain non-negative variance estimates OOS, without having to impose any parameter restrictions. Additionally, as we mentioned in our empirical estimation section, the log-realized variance approximately follows a Normal distribution, which is a desired property for our multi-step OOS regression framework, as the OLS estimator will demonstrate lower bias and improved efficiency (see also Papantonis, Rompolis, and Tzavalis (2021)). Effectively, what we do here is very similar to a log-HAR regression specification for predicting variance at multiple-horizons ahead; the only difference being the addition of the log-conditional variance in the regression, which encapsulates in $\tilde{h}_{t+1}$ the endogenous conditional variance dynamics, as well as the variance response to lagged return innovations.

¹⁷ This regression-based framework has been used extensively in the literature to examine the forecasting performance of variance over long-horizons; see, e.g. Corsi (2009), Andersen, Bollerslev, and Diebold (2007) as well as Ghysels et al. (2019) for a recent survey.

Table 9: Reality Check tests results comparing different distributions for the return innovations.

	MSE	MAE	HMSE	MAPE
Panel A: NIG vs SGED				
GARCH-SHARQ	0.1229	0.0092	0.0000	0.0000
GARCH-SHAR	0.0000	0.0000	0.0000	0.0000
GARCH-HAR	0.0000	0.0000	0.0000	0.0000
GARCH-S	0.0000	0.0000	0.0000	0.0000
GARCH-R	0.0000	0.0000	0.0000	0.0000
SHAR	0.0000	0.0000	0.0000	0.0000
HAR	0.0000	0.0000	0.0000	0.0000
Panel B: NIG vs N				
GARCH-SHARQ	0.0187	0.0000	0.0000	0.0000
GARCH-SHAR	0.0000	0.0000	0.0000	0.0000
GARCH-HAR	0.0000	0.0000	0.0000	0.0000
GARCH-S	0.0000	0.0000	0.0000	0.0000
GARCH-R	0.0000	0.0000	0.0000	0.0000
SHAR	0.0000	0.0000	0.0000	0.0000
HAR	0.0000	0.0000	0.0000	0.0000

This table reports the p -values of reality check tests comparing similar models (reported in the first column) under different distributions for the return innovations. Panel A compares the NIG versus SGED, while panel B compares the NIG versus the standard Normal distribution.

In Table 10, we present estimates of the OOS prediction loss-metrics for all the alternative specifications nested with the GARCH-SHARQ. This is done for $s = \{5, 10, 20\}$ days ahead. For reasons of space, we present results for the full set of models only for the NIG distribution, since it has been found to perform better under all circumstances.¹⁸ The results of the table are consistent with those for the case of one-period ahead for all models, and indicate that the full GARCH-SHARQ model provides the best OOS performance compared to all the other nested specifications. This result is true for all the different horizons ahead considered and under all the forecasting performance metrics reported in the table. Especially when comparing with the GARCH-R model, the improvements in variance forecasts appear to be very substantial. As can be seen from the table, these improvements are mainly due to the inclusion of the $RV_t^{[5]}$ and $RV_t^{[20]}$ terms approximating the long-memory property of variance process, and the two semi-variances RV_t^+ and RV_t^- . As expected, the forecasting performance of all the alternative models deteriorates (in terms of the prediction loss-metrics reported in the table) as we projection horizon increases.¹⁹

¹⁸ We replicate the multi-step ahead OOS forecasting exercise under the assumption of SGED- and Normally-distributed return innovations. We find that assuming an NIG distribution for returns leads to more accurate variance forecasts at all different horizons. This is in line not only with our full-sample results, but also with the one-step ahead predictions discussed in the previous sections.

¹⁹ The coefficient estimates of the regression model in Eq. (6) – not reported here for reasons of space – indicate that the model keeps its prediction ability even for $h = 20$ days ahead.

Table 10: Out-of-sample variance loss-function results for multi-step ahead predictions.

	Horizon $s = 5$				Horizon $s = 10$				Horizon $s = 20$			
	MSE	MAE	HMSE	MAPE	MSE	MAE	HMSE	MAPE	MSE	MAE	HMSE	MAPE
GARCH-SHARQ	0.6088	0.6013	1.6880	0.7307	0.7386	0.6657	2.1816	0.8239	0.8699	0.7265	2.4475	0.9066
GARCH-HARQ	0.6134	0.6027	1.7160	0.7323	0.7399	0.6666	2.2092	0.8254	0.8709	0.7273	2.4589	0.9083
GARCH-SHAR	0.6119	0.6037	1.6152	0.7314	0.7421	0.6676	2.0901	0.8262	0.8738	0.7284	2.4402	0.9097
GARCH-HAR	0.6240	0.6099	1.6885	0.7422	0.7455	0.6702	2.0859	0.8303	0.8768	0.7300	2.4529	0.9142
GARCH-S	0.6324	0.6173	1.9495	0.7697	0.7697	0.6791	2.5397	0.8683	0.8982	0.7411	2.6443	0.9391
GARCH-R	0.6441	0.6235	2.1038	0.7823	0.7693	0.6792	2.4811	0.8673	0.8987	0.7429	2.6430	0.9425
SHARQz	0.6159	0.6029	1.7055	0.7331	0.7431	0.6676	2.1937	0.8258	0.8729	0.7271	2.4710	0.9074
SHARQ	0.6187	0.6045	1.6795	0.7336	0.7452	0.6681	2.2252	0.8266	0.8731	0.7271	2.4718	0.9073
HARQ	0.6198	0.6037	1.7009	0.7327	0.7449	0.6684	2.2332	0.8271	0.8737	0.7276	2.4798	0.9086
SHAR	0.6227	0.6069	1.6693	0.7369	0.7485	0.6702	2.1848	0.8295	0.8761	0.7287	2.4682	0.9103
HAR	0.6312	0.6113	1.7184	0.7449	0.7525	0.6728	2.1832	0.8336	0.8798	0.7304	2.4820	0.9144

This table presents out-of-sample results for 5, 10, and 20 days 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. The multi-period ahead predictions of $\log RV$ is estimated using the OOS regression model (6). It reports the mean squared error (MAE), mean absolute error (MAE), heteroskedasticity-adjusted mean squared error (HMSE), and mean absolute percentage error (MAPE).

5 Summary & conclusions

In this paper, we extend the Realized-GARCH model (or GARCH-R under our notation) with exogenous variates related to measurable features of realized variance, the selection of which is motivated by the ongoing literature on HAR models for realized variance. The suggested GARCH-HAR-X framework can be also seen as an a dual extension of the HAR model to include GARCH terms. We analyse both the in-sample and out-of-sample prediction performance of the generalized model, as well as the several specifications nested within, and we do that under different assumptions for the probability density function of the return innovations. We consider not only the Normal, often used in practice, but also the SGED and the NIG distribution which have been shown to perform very well in capturing distributional asymmetries of financial data.

We find our “full” GARCH-SHARQ specification to perform significantly better than the standard GARCH-R in fitting the data and forecasting realized variance. We show that its enhanced performance originates from the following three sources/extensions to the conditional variance function: (i) the decomposition of realized variance into upside downside semi-variances, which allows to better capture short-term asymmetric behaviours in variance due to the signed-jump variation; (ii) the addition of the heterogeneous components, which effectively approximate a long-memory effect in conditional variance dynamics and, most importantly, (iii) the inclusion of a variance-of-variance (quarticity) proxy responsible for capturing attenuation-biases in variance forecasts. The in-sample and out-of-sample performance of the models is assessed based on a number of different fit and prediction-accuracy metrics; these clearly indicate that the inclusion of upside/downside realized semi-variances, as well as realized quarticity, are the main drivers for the significant gains in the prediction performance of the augmented GARCH-SHARQ model.

Furthermore, our GARCH-SHARQ model has also been found to perform better than the several nested HAR-type representations. We show that this is primarily due to the filtered return innovations when included in conditional variance through the news-impact function. This allows for two separate channels of leverage-effect; through asymmetric responses to both return innovations and semi-variances. Finally, regarding the performance of the alternative model specifications under different parametric density assumptions, our analysis shows that the NIG distribution enhances the in-sample fit and out-of-sample forecasting performance compared to both the SGED and the Normal distribution, and this holds across all model specifications. We show that this may be attributed to the enhanced flexibility of the NIG in capturing asymmetries in the data.

Acknowledgement: We would like to thank the editor Bruce Mizrach, as well as the two anonymous reviewers for providing useful comments and suggestions on an earlier version of the paper. We also acknowledge financial support from the Operational Programme “*Human Resources Development, Education and Lifelong Learning 2014–2020*” in the context of the project MIS-5049538, co-financed by Greece and the European Union (European Social Fund, ESF).

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This research is co-financed by Greece and the European Union (European Social Fund, ESF) through the Operational Programme “*Human Resources Development, Education and Lifelong Learning 2014–2020*” in the context of the project MIS-5,049,538.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

Appendix A: Distributions

A.1 The standardized normal-inverse-Gaussian distribution

A.1.1 The normal-inverse-Gaussian (NIG) distribution

The normal-inverse-Gaussian, henceforth NIG, results as a special case of the generalized hyperbolic (GH) family of (mean-variance) mixture distributions popularized by Barndorff-Nielsen (1978), with the mixing distribution in this case being the inverse-Gaussian (IG). A very detailed description of this family of distributions can be found in McNeil, Frey, and Embrechts (2005). The location-scale invariant parametrization of the PDF of a random variable X that follows an NIG distribution, i.e. $x \sim \text{NIG}(\bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta})$, is defined by:

$$f_{\text{NIG}}(x; \bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta}) = \frac{\bar{\alpha}}{\pi \bar{\delta}} \exp \left[\sqrt{\bar{\alpha}^2 - \bar{\beta}^2} + \bar{\beta} \frac{x - \bar{\mu}}{\bar{\delta}} \right] q \left(\frac{x - \bar{\mu}}{\bar{\delta}} \right)^{-1} \mathbb{K}_1 \left(\bar{\alpha} q \left(\frac{x - \bar{\mu}}{\bar{\delta}} \right) \right),$$

where $\mathbb{K}_1(\cdot)$ is the modified Bessel function of third order and index 1, and $q(x) = \sqrt{1 + x^2}$. All distributional parameters must be real-valued while satisfying $0 \leq |\bar{\beta}| < \bar{\alpha}$ and $\bar{\delta} > 0$. The interpretation of the parameters is quite straightforward: $\bar{\alpha}$ and $\bar{\beta}$ are shape parameters, driving the steepness and the asymmetry of the density, respectively, while $\bar{\mu}$ and $\bar{\delta}$ correspond to the location and the scale of the distribution. In the limiting case of $\bar{\beta} = 0$ the resulting density is symmetric, and also as $\bar{\alpha} \rightarrow \infty$ then the density converges to the Normal. It can be shown that (under the location-scale invariant parametrization) it holds that:

$$X \sim \text{NIG}(\bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta}) \Leftrightarrow \frac{X - \bar{\mu}}{\bar{\delta}} \sim \text{NIG}(\bar{\alpha}, \bar{\beta}, 0, 1),$$

while the moment-generating function (MGF) can be expressed as:

$$\bar{M}(\tau; \bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta}) = \exp \left[\bar{\alpha} \left(\sqrt{1 - \left(\frac{\bar{\beta}}{\bar{\alpha}} \right)^2} - \sqrt{1 - \left(\frac{\bar{\beta}}{\bar{\alpha}} + \frac{\bar{\delta}}{\bar{\alpha}} \tau \right)^2} \right) + \tau \bar{\mu} \right].$$

To get the first four central moments we solve the above MGF:

$$\begin{aligned} E[X] &= \bar{\mu} + \frac{\lambda \bar{\delta}}{\sqrt{1 - \lambda^2}} \\ \text{Var}[X] &= \frac{\bar{\delta}^2}{\bar{\alpha} \left(\sqrt{1 - \lambda^2} \right)^{3/2}} \\ \text{Skew}[X] &= 3 \left(\frac{\lambda}{\left(\bar{\alpha} \sqrt{1 - \lambda^2} \right)^{1/2}} \right) \\ \text{Kurt}[X] &= 3 \left(1 + \frac{4\lambda^2 + 1}{\bar{\alpha} \sqrt{1 - \lambda^2}} \right), \end{aligned}$$

where $\lambda = \bar{\beta}/\bar{\alpha}$.

A.1.2 The zero-mean & unit-variance NIG specification

We want to ensure that the innovation z follows a zero-mean and unit-variance NIG distribution. Using the MGF of the NIG to solve for the first central moments and equating the first two moments of r_{t+1} , we can show that the returns follow a conditional NIG distribution as:

$$r_{t+1} \sim \text{NIG}(\bar{\alpha}_{t+1}, \bar{\beta}_{t+1}, \bar{\mu}_{t+1}, \bar{\delta}_{t+1} | \mathcal{F}_t),$$

with

$$\begin{aligned}\bar{\mu}_{t+1} &= E_t[r_{t+1}] - \lambda_{t+1} \bar{\delta}_{t+1} \left(\sqrt{1 - \lambda_{t+1}^2} \right)^{-1} \\ \bar{\delta}_{t+1} &= \text{Var}_t[r_{t+1}]^{1/2} \left[\bar{\alpha}_{t+1} \left(\sqrt{1 - \lambda_{t+1}^2} \right)^{3/2} \right]^{1/2},\end{aligned}$$

i.e. it is entirely driven by the constant values $\bar{\alpha}$ and $\bar{\beta}$. Given $0 \leq |\bar{\beta}| < \bar{\alpha}$ we define

$$\lambda = \bar{\beta}/\bar{\alpha} \in (-1, 1) \quad \text{and} \quad \eta = \sqrt{\bar{\alpha}^2 - \bar{\beta}^2} > 0, \quad (\text{A.1})$$

which we can think of as two quantities driving the “asymmetry” and “steepness” of the density. Finally, we solve the system of equations implied in (A.1) for $\bar{\alpha}$ and $\bar{\beta}$ to get the parameters

$$\bar{\beta} = \eta \lambda (1 - \lambda^2)^{-1/2} \quad \text{and} \quad \bar{\alpha} = \sqrt{\bar{\beta}^2 + \eta^2},$$

which we substitute in the density function above to get the log-likelihood of returns. Using the parametrization above, we can think of the standardized innovation z as having a standardized NIG density purely defined by two parameters, i.e. $z \sim \text{NIG}(\eta, \lambda)$.

A.2 The SGED (skewed generalized error) distribution

A zero-mean and unit-variance random variable Z is SGED-distributed as $z \sim \text{SGED}(\eta, \lambda)$ (Theodossiou 1998) when it adheres to the PDF:

$$f_{\text{SGED}}(z; \eta, \lambda) = C \exp \left(- \frac{|z + \delta|^\eta}{[1 + \text{sign}(z + \delta)\lambda]\theta^\eta} \right),$$

where the following parametrizations hold:

$$\begin{aligned}C &= (\eta/2\theta) \Gamma(1/\eta)^{-1} \\ \theta &= \Gamma(1/\eta)^{1/2} \Gamma(3/\eta)^{-1/2} S^{-1} \\ \delta &= 2\lambda A S^{-1} \\ S &= \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2} \\ A &= \Gamma(2/\eta) \Gamma(1/\eta)^{-1/2} \Gamma(3/\eta)^{-1/2},\end{aligned}$$

with Γ being the *gamma*-function $\Gamma(q) = \int_0^\infty x^{q-1} e^{-x} dx$. The density is defined iff

$$\eta > 0 \quad \text{and} \quad -1 < \lambda < 1.$$

This density nests several conventional densities as well. For instance, for $\lambda = 0$ the density reduces to the GED distribution of Nelson (1991), which is symmetric but fat-tailed. When $\lambda = 0$ and $\eta = 2$, the density collapses to the standard Normal. For $\lambda = 0$ and $\eta = 1$ we get the double-exponential distribution and finally for $\lambda = 0$ and $\eta = +\infty$ we get a uniform in the $[-\sqrt{3}, +\sqrt{3}]$ interval. Hence, it is easy to interpret η as the degrees of freedom and λ as the asymmetry parameter of the density.

Regarding the higher moments of an SGED-distributed random variable, the following mapping holds:

$$\begin{aligned}\text{skew}[Z] &= E[z^3] = A_3 - 3\delta - \delta^3 \\ \text{kurt}[Z] &= E[z^4] = A_4 - 4A_3\delta + 6\delta^2 + 3\delta^4,\end{aligned}$$

with

$$\begin{aligned} A_3 &= 4\lambda(1 + \lambda^2)\Gamma(4/\eta)\Gamma(1/\eta)^{-1}\theta^3 \\ A_4 &= (1 + 10\lambda^2 + 5\lambda^4)\Gamma(5/\eta)\Gamma(1/\eta)^{-1}\theta^4. \end{aligned}$$

A.3 Robustness of the GARCH-SHARQ model to alternative specifications of the NIF

The NIF given by Eq. (3) assumes abrupt shifts in variance h_{t+1} driven by the sign of ϵ_t . Alternative specifications of this function considered smooth transition shifts have been also suggested in the literature (see Franses and Van Dijk (2000), for a survey). In this direction, the logistic smooth transition model of González-Rivera (1998) is one popular alternative. This model assumes that $\tau(t)$ is given as

$$\tau(t) = b_2\epsilon_t^2 + b_3\epsilon_t^2 g(z_t; \gamma), \quad (\text{A.2})$$

where $g(z_t; \gamma) = [1 + \exp(-\gamma z_t)]^{-1} \in (0, 1)$ is the logistic function governing the transition between the two regimes assumed by the NIF. The value of parameter γ , known as the speed-of-transition parameter, determines the smoothness of the transition between the two regimes implied by $g(z_t)$. For $\gamma \rightarrow \infty$, $g(z_t; \gamma)$ tends to the indicator function implied by Eq. (3); i.e. equation (A.2) reduces to (3) and the two NIF functions are equivalent. On the other hand, when $\gamma \rightarrow 0$, then $g(z_t; \gamma) \rightarrow \frac{1}{2}$ which means that the smooth-transition model reduces to the standard GARCH model, also nested within the GJR.

Another direction towards which Eq. (3) can be extended is to allow for an unknown threshold parameter, instead of assuming $\delta = 0$ as in (3), i.e.,

$$\tau(t) = b_2\epsilon_t^2 + b_3\mathbb{I}(z_t < \delta)\epsilon_t^2. \quad (\text{A.3})$$

The threshold parameter δ can be estimated endogenously from the data, using a grid-search approach.²⁰

In Table 11, we present estimates of the fully specified GARCH-SHARQ model under the three distributions considered. The results show that the estimates of γ are quite large (implying almost abrupt shifts between the two regimes of the NIF) and the threshold parameter δ is very close to zero; the bootstrapped standard errors reported in square brackets indicate that the null hypothesis of $\delta = 0$ cannot be rejected. These results hold for all distributions considered. This allows us to conclude that the NIF function given by (3) provides a correct specification of the data. Further support of (3) can be obtained by the values of the likelihood components $\mathcal{L}$, $\mathcal{L}^R$ and $\mathcal{L}^V$ (or the information criteria AIC and BIC), as well the prediction performance metrics reported in the table. The use of the NIFs (A.2) and (A.3) does not improve upon the performance of (3).

²⁰ Note that both (A.2) and (A.3) can be nested in a more general specification of the logistic transition function given as $g(z_t; \gamma, \delta) = [1 + \exp(-\gamma(z_t - \delta))]^{-1}$. However, estimation of the model with this transition function often leads to identification problems of γ and δ . Thus, estimation of (A.2) and (A.3), separately, often constitutes a preferable approach to obtain robust estimates of γ and δ .

Table 11: Full sample maximum likelihood parameter estimates for linear GARCH-HAR-X models under different NIF specifications.

	NIG			SGED			Normal		
μ	0.0476 [0.0170]	0.0476 [0.0171]	0.0472 [0.0171]	0.0459 [0.0150]	0.0459 [0.0150]	0.0456 [0.0150]	0.0416 [0.0134]	0.0413 [0.0135]	
b_0	1.8060 [0.1891]	1.8060 [0.1891]	1.8060 [0.1891]	1.8633 [0.1996]	1.8633 [0.1995]	1.8633 [0.1969]	2.0055 [0.1969]	2.0055 [0.1969]	
b_1	0.4870 [0.0260]	0.4870 [0.0260]	0.4869 [0.0260]	0.4861 [0.0256]	0.4860 [0.0256]	0.4859 [0.0252]	0.5004 [0.0252]	0.5003 [0.0252]	
b_2	0.0773 [0.0063]	0.0773 [0.0063]	0.0773 [0.0063]	0.0803 [0.0067]	0.0803 [0.0067]	0.0803 [0.0067]	0.0769 [0.0066]	0.0770 [0.0066]	
b_3	-0.0178 [0.0038]	-0.0178 [0.0038]	-0.0178 [0.0038]	-0.0191 [0.0040]	-0.0191 [0.0040]	-0.0191 [0.0039]	-0.0220 [0.0039]	-0.0220 [0.0039]	
c_1	0.3419 [0.0284]	0.3419 [0.0284]	0.3420 [0.0284]	0.3629 [0.0300]	0.3629 [0.0300]	0.3629 [0.0300]	0.3563 [0.0295]	0.3564 [0.0295]	
c_2	0.4625 [0.0231]	0.4626 [0.0231]	0.4626 [0.0231]	0.5006 [0.0246]	0.5006 [0.0246]	0.5006 [0.0246]	0.4999 [0.0242]	0.4999 [0.0242]	
c_3	0.0486 [0.0169]	0.0486 [0.0169]	0.0487 [0.0169]	0.0514 [0.0177]	0.0514 [0.0177]	0.0514 [0.0175]	0.0529 [0.0175]	0.0530 [0.0175]	
c_4	0.0655 [0.0079]	0.0655 [0.0079]	0.0656 [0.0079]	0.0675 [0.0083]	0.0675 [0.0083]	0.0675 [0.0083]	0.0637 [0.0081]	0.0637 [0.0081]	
c_5	-0.0816 [0.0058]	-0.0816 [0.0058]	-0.0817 [0.0058]	-0.0870 [0.0061]	-0.0870 [0.0061]	-0.0870 [0.0061]	-0.0853 [0.0060]	-0.0853 [0.0060]	
γ	-	17.6271 [7.2796]	-	-	17.3428 [6.8258]	-	17.7180 [5.9556]	-	
δ	-	-	0.1442 [0.1120]	-	-	0.1429 [0.1083]	-	0.1460 [0.1040]	
η	51.1581 [5.1227]	51.1610 [5.2285]	51.1699 [4.1611]	1.4566 [0.0294]	1.4566 [0.0294]	1.4567 [0.0294]	-	-	
λ	-0.6859 [0.0109]	-0.6859 [0.0109]	-0.6860 [0.0109]	-0.0723 [0.0339]	-0.0723 [0.0339]	-0.0724 [0.0341]	-	-	
$\mathcal{L}^R$	18,092.74	18,092.74	18,092.76	17,895.70	17,895.70	17,895.73	17,777.39	17,777.42	
$\mathcal{L}^V$	-4509.27	-4509.27	-4509.23	-4538.94	-4538.94	-4538.90	-4568.82	-4568.77	
$\mathcal{L}$	13,583.46	13,583.47	13,583.53	13,356.76	13,356.76	13,356.83	13,208.57	13,208.65	
AIC	-27,142.93	-27,140.93	-27,141.05	-26,689.52	-26,687.53	-26,687.65	-26,395.14	-26,395.29	
BIC	-27,063.49	-27,054.88	-27,055.00	-26,610.08	-26,601.47	-26,601.60	-26,322.33	-26,322.48	
MSE	0.2984	0.2984	0.2984	0.3016	0.3016	0.3016	0.3048	0.3048	
MAE	0.4214	0.4214	0.4214	0.4259	0.4259	0.4259	0.4288	0.4288	
HMSE	0.5056	0.5056	0.5056	0.5854	0.5854	0.5854	0.6252	0.6252	
MAPE	0.4608	0.4608	0.4608	0.4914	0.4914	0.4914	0.5050	0.5050	

This table presents maximum likelihood estimation (MLE) results for the GARCH-SARQ model under different NIF specifications assuming that the return innovations follow the standard Normal, SGED, and NIG distribution. The NIF specifications considered are the GJR, LST-GARCH, and T-GARCH. $\mathcal{L}^R$ 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 denotes 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 and MAPE denotes the mean absolute percentage errors. The definitions of these metrics are given in (6). Standard errors are shown in brackets.

References

- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2001. "The Distribution of Realized Exchange Rate Volatility." *Journal of the American Statistical Association* 96 (453): 42–55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. "Modeling and Forecasting Realized Volatility." *Econometrica* 71 (2): 579–625.
- Andersen, T. G., T. Bollerslev, and N. Meddahi. 2005. "Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities." *Econometrica* 73 (1): 279–96.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold. 2007. "Roughing it up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility." *The Review of Economics and Statistics* 89 (4): 701–20.
- Andersen, T. G., R. A. Davis, J.-P. Kreiß, and T. V. Mikosch. 2009. *Handbook of Financial Time-Series*. New York: Springer.
- Andersson, J. 2001. "On the Normal Inverse Gaussian Stochastic Volatility Model." *Journal of Business & Economic Statistics* 19 (1): 44–54.
- Audrino, F., and Y. Hu. 2016. "Volatility Forecasting: Downside Risk, Jumps and Leverage Effect." *Econometrics* 4 (1): 8.
- Baillie, R. T., F. Calonaci, D. Cho, and S. Rho. 2019. "Long Memory, Realized Volatility and Heterogeneous Autoregressive Models." *Journal of Time Series Analysis* 40 (4): 609–28.
- Barndorff-Nielsen, O. E. 1978. "Hyperbolic Distributions and Distributions on Hyperbolae." *Scandinavian Journal of Statistics* 5 (3): 151–7.
- 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., and N. Shephard. 2002a. "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): 253–80.
- Barndorff-Nielsen, O. E., and N. Shephard. 2002b. "Estimating Quadratic Variation Using Realized Variance." *Journal of Applied Econometrics* 17 (5): 457–77.
- Barndorff-Nielsen, O. E., and N. Shephard. 2004a. "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics." *Econometrica* 72 (3): 885–925.
- Barndorff-Nielsen, O. E., and N. Shephard. 2004b. "Power and Bipower Variation with Stochastic Volatility and Jumps." *Journal of Financial Econometrics* 2 (1): 1–37.
- Barndorff-Nielsen, O. E., and N. Shephard. 2006. "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation." *Journal of Financial Econometrics* 4 (1): 1–30.
- Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard. 2008. "Measuring Downside Risk-Realised Semivariance." In *CREATES Research Paper*, (2008-42).
- Bauwens, L., C. M. Hafner, and S. Laurent. 2012. *Handbook of Volatility Models and their Applications*, 3. Chichester: John Wiley & Sons.
- Bekierman, J., and B. Gribisch. 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): 307–27.
- Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen. 2009. "A Discrete-Time Model for Daily S & P500 Returns and Realized Variations: Jumps and Leverage Effects." *Journal of Econometrics* 150 (2): 151–66.
- Bollerslev, T., A. Patton, and R. Quaedvlieg. 2016. "Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting." *Journal of Econometrics* 192 (1): 1–18.
- Chen, C. W., T. Watanabe, and E. M. Lin. 2021. "Bayesian Estimation of Realized Garch-Type Models with Application to Financial Tail Risk Management." *Econometrics and Statistics*.
- Chorro, C., D. Guégan, and F. Ielpo. 2012. "Option Pricing for GARCH-Type Models with Generalized Hyperbolic Innovations." *Quantitative Finance* 12 (7): 1079–94.
- Christoffersen, P., B. Feunou, K. Jacobs, and N. Meddahi. 2014. "The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation." *Journal of Financial and Quantitative Analysis* 49 (3): 663–97.
- Cipollini, F., G. M. Gallo, and E. Otranto. 2020. "Realized Volatility Forecasting: Robustness to Measurement Errors." *International Journal of Forecasting* 37 (1): 44–57.
- Clark, T. E., and M. W. McCracken. 2005. "Evaluating Direct Multistep Forecasts." *Econometric Reviews* 24 (4): 369–404.
- Clark, T. E., and M. McCracken. 2013. "Advances in Forecast Evaluation." In *Handbook of Economic Forecasting*, 2, 1107–201. Amsterdam: Elsevier.
- Corsi, F. 2009. "A Simple Approximate Long-Memory Model of Realized Volatility." *Journal of Financial Econometrics* 7 (2): 174–96.
- Corsi, F., and R. Renò. 2012. "Discrete-Time Volatility Forecasting with Persistent Leverage Effect and the Link with Continuous-Time Volatility Modeling." *Journal of Business & Economic Statistics* 30 (3): 368–80.
- Dendramis, Y., G. E. Spungin, and E. Tzavalis. 2014. "Forecasting Var Models under Different Volatility Processes and Distributions of Return Innovations." *Journal of Forecasting* 33 (7): 515–31.

- 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* 50 (4): 987–1007.
- Engle, R. 2002. "New Frontiers for Arch Models." *Journal of Applied Econometrics* 17 (5): 425–46.
- Engle, R., and G. M. Gallo. 2006. "A Multiple Indicators Model for Volatility Using Intra-Daily Data." *Journal of Econometrics* 131 (1–2): 3–27.
- Engle, R., D. M. Lilien, and R. P. Robins. 1987. "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model." *Econometrica* 55 (2): 391–407.
- Eriksson, A., E. Ghysels, and F. Wang. 2009. "The Normal Inverse Gaussian Distribution and the Pricing of Derivatives." *The Journal of Derivatives* 16 (3): 23–37.
- Feunou, B., and C. Okou. 2019. "Good Volatility, Bad Volatility, and Option Pricing." *Journal of Financial and Quantitative Analysis* 54 (2): 695–727.
- Feunou, B., M. R. Jahan-Parvar, and R. Tédongap. 2016. "Which Parametric Model for Conditional Skewness?" *The European Journal of Finance* 22 (13): 1237–71.
- Feunou, B., M. R. Jahan-Parvar, and C. Okou. 2017. "Downside Variance Risk Premium." *Journal of Financial Econometrics* 16 (3): 341–83.
- Forsberg, L., and T. Bollerslev. 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): 535–48.
- Franco, C., and J.-M. Zakoian. 2019. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: John Wiley & Sons.
- Franses, P. H., and D. Van Dijk. 2000. *Non-linear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press.
- Ghysels, E., and F. Wang. 2014. "Moment-implied Densities: Properties and Applications." *Journal of Business & Economic Statistics* 32 (1): 88–11.
- Ghysels, E., A. Plazzi, R. Valkanov, A. Rubia, and A. Dossani. 2019. "Direct versus Iterated Multiperiod Volatility Forecasts." *Annual Review of Financial Economics* 11: 173–95.
- Giraitis, L., G. Kapetanios, and S. Price. 2013. "Adaptive Forecasting in the Presence of Recent and Ongoing Structural Change." *Journal of Econometrics* 177 (2): 153–70.
- Giraitis, L., G. Kapetanios, M. Mansur, and S. Price. 2015. "Forecasting under Structural Change." In *Empirical Economic and Financial Research*, 401–19. New York: Springer.
- Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *The Journal of Finance* 48 (5): 1779–801.
- González-Rivera, G. 1998. "Smooth-Transition Garch Models." *Studies in Nonlinear Dynamics & Econometrics* 3 (2).
- Hansen, P. R. 2005. "A Test for Superior Predictive Ability." *Journal of Business & Economic Statistics* 23 (4): 365–80.
- Hansen, P. R., and Z. Huang. 2016. "Exponential GARCH Modeling with Realized Measures of Volatility." *Journal of Business & Economic Statistics* 34 (2): 269–87.
- Hansen, P. R., and A. Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1, 1)?" *Journal of Applied Econometrics* 20 (7): 873–89.
- Hansen, P. R., Z. Huang, and H. H. Shek. 2012. "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility." *Journal of Applied Econometrics* 27 (6): 877–906.
- Hansen, P. R., Z. Huang, and T. Wang. 2015. "Realized EGARCH, Cboe Vix and Variance Risk Premium." In *Working Paper*.
- Huang, Z., H. Liu, and T. Wang. 2016. "Modeling Long Memory Volatility Using Realized Measures of Volatility: A Realized HAR GARCH Model." *Economic Modelling* 52: 812–21.
- Jensen, M. B., and A. Lunde. 2001. "The Nig-S&arch Model: A Fat-Tailed, Stochastic, and Autoregressive Conditional Heteroskedastic Volatility Model." *The Econometrics Journal* 4 (2): 319–42.
- Jondeau, E., and M. Rockinger. 2003. "Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements." *Journal of Economic Dynamics and Control* 27 (10): 1699–737.
- Koopman, S. J., and M. Scharth. 2012. "The Analysis of Stochastic Volatility in the Presence of Daily Realized Measures." *Journal of Financial Econometrics* 11 (1): 76–115.
- Kourtis, A., R. N. Markellos, and L. Symeonidis. 2016. "An International Comparison of Implied, Realized, and GARCH Volatility Forecasts." *Journal of Futures Markets* 36 (12): 1164–93.
- Louzis, D. P., S. Xanthopoulos-Sisinis, and A. P. Refenes. 2014. "Realized Volatility Models and Alternative Value-At-Risk Prediction Strategies." *Economic Modelling* 40: 101–16.
- Marcellino, M., J. H. Stock, and M. W. Watson. 2006. "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series." *Journal of Econometrics* 135 (1–2): 499–526.
- McNeil, A. J., R. Frey, P. Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*, Vol. 3. Princeton: Princeton University Press.

- Nelson, D. B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59 (2): 347–70.
- Papantonis, I. 2016. "Volatility Risk Premium Implications of GARCH Option Pricing Models." *Economic Modelling* 58: 104–15.
- Papantonis, I., L. Rompolis, and E. Tzavalis. 2021. *Improving Variance Forecasts: The Role of Realized Variance Features*. Athens: Athens University of Economics and Business, Mimeo.
- Patton, A., and K. Sheppard. 2015. "Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility." *Review of Economics and Statistics* 97 (3): 683–97.
- Patton, A., D. N. Politis, and H. White. 2009. "Correction to "Automatic Block-Length Selection for the Dependent Bootstrap" by D. Politis and H. White." *Econometric Reviews* 28 (4): 372–5.
- Politis, D. N., and J. P. Romano. 1994. "The Stationary Bootstrap." *Journal of the American Statistical Association* 89 (428): 1303–13.
- Politis, D. N., and H. White. 2004. "Automatic Block-Length Selection for the Dependent Bootstrap." *Econometric Reviews* 23 (1): 53–70.
- Sheppard, N., and K. Sheppard. 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): 540–82.
- Takahashi, M., Y. Omori, and T. Watanabe. 2009. "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously." *Computational Statistics & Data Analysis* 53 (6): 2404–26.
- Theodossiou, P. 1998. "Financial Data and the Skewed Generalized T Distribution." *Management Science* 44 (12-part-1): 1650–61.
- Theodossiou, P. 2015. "Skewed Generalized Error Distribution of Financial Assets and Option Pricing." *Multinational Finance Journal* 19 (4): 223–66.
- Theodossiou, P., and C. S. Savva. 2016. "Skewness and the Relation between Risk and Return." *Management Science* 62 (6): 1598–609.
- Venter, J. H., and P. J. de Jongh. 2002. "Risk Estimation Using the Normal Inverse Gaussian Distribution." *Journal of Risk* 4: 1–24.
- Wang, C., Q. Chen, and R. Gerlach. 2019. "Bayesian Realized-Garch Models for Financial Tail Risk Forecasting Incorporating the Two-Sided Weibull Distribution." *Quantitative Finance* 19 (6): 1017–42.
- West, K. D. 2006. "Forecast Evaluation." In *Handbook of Economic Forecasting*, 1, 99–134. Amsterdam: Elsevier.
- White, H. 2000. "A Reality Check for Data Snooping." *Econometrica* 68 (5): 1097–126.
- Wilhelmsson, A. 2009. "Value at Risk with Time Varying Variance, Skewness and Kurtosis—The Nig-Acd Model." *The Econometrics Journal* 12 (1): 82–104.
- Zakoian, J.-M. 1994. "Threshold Heteroskedastic Models." *Journal of Economic Dynamics and Control* 18 (5): 931–55.

Supplementary Material: The online version of this article offers supplementary material (<https://org/10.1515/snde-2020-0131>).