Time Varying Heteroskedastic Realized GARCH models for tracking measurement error bias in volatility forecasting

Gerlach, Richard and Naimoli, Antonio and Storti, Giuseppe

University of Sydney, University of Salerno, University of Salerno

8 January 2018

Time-Varying Heteroskedastic Realized GARCH models for tracking measurement error bias in volatility forecasting

Richard Gerlach[1], *Antonio Naimoli*[2] and *Giuseppe Storti*[2] *

1. The University of Sydney – Business School

2. University of Salerno – Department of Economics and Statistics

*Corresponding author. Address for correspondence: Dipartimento di Scienze Economiche e Statistiche, Università degli studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy. Email: storti@unisa.it. Office phone: 0039 089 962212.

Abstract

This paper proposes generalisations of the Realized GARCH model by Hansen et al. (2012), in three different directions. First, heteroskedasticity of the noise term in the measurement equation is taken into account. Namely, the variance of the measurement error is assumed to be time-varying as a function of an estimator of the Integrated Quarticity obtained from intra-daily returns. Second, in order to account for attenuation bias effects, the volatility dynamics are allowed to depend on the accuracy of the realized measure. This is achieved by letting the response coefficient of the lagged realized measure be a function of the time-varying variance of the volatility measurement error. This feature allows the model to assign more weight to lagged volatilities when they are more accurately measured. Finally, a further extension is proposed by introducing an additional explanatory variable into the measurement equation, aiming to quantify the bias due to the effect of jumps.

JEL Codes: C58, C22 ,C53.

Keywords: Realized GARCH, Realized Volatility, Realized Quarticity, Jumps.

1 Introduction

It is widely acknowledged that the use of realized volatility measures (Hansen and Lunde, 2011) can be beneficial for improving the accuracy of volatility forecasts on a daily scale. This is typically done by choosing one of the following approaches.

First, dynamic models can be directly fitted to time series of realized measures. Examples include the Heterogeneous AutoRegressive (HAR) (Corsi, 2009) and the class of Multiplicative Error Models (MEM) (Engle, 2002; Engle and Gallo, 2006). A drawback of this approach is that the focus is on the estimation of the expected level of the realized measure, rather than on the estimation of the conditional variance of returns. As it will be clarified in the next section, realized measures are designed to consistently estimate the integrated variance which is related to but different from the conditional variance. Namely, in the absence of microstructure noise and jumps, the integrated variance can be interpreted as an unbiased estimator of the conditional variance of returns.

The second approach makes use of time series models for daily returns, e.g. GARCH-type models, where the conditional variance is driven by one or more realized measures. The main idea is to replace a noisy volatility proxy, such as the squared daily returns used in standard GARCH models, with a more efficient realized measure. Differently from the above-mentioned approach, in this case, both low (daily returns) and high (realized measures) frequency information are employed in the model. Examples of models falling within this class include the HEAVY model of Shephard and Sheppard (2010) and the Realized GARCH model of Hansen et al. (2012). These two models are closely related but, nevertheless, they are characterised by some distinctive features. HEAVY models are designed for the generation of multi-step ahead forecasts, guaranteed by the inclusion of a dynamic updating equation for the conditional expectation of the chosen realized measure. On the other hand, Realized GARCH models include a measurement equation allowing to gain, in a fully data-driven fashion, deeper insight on the statistical properties of the realized measure and its relationship with the latent volatility.

A complication arising with both approaches is that realized measures are noisy estimates of the underlying integrated variance, generating a classical errors-in-variables problem. This typically leads to the rise of what is often called attenuation bias. More precisely, the estimated response of the conditional variance to the past realized measure will be negatively biased, compared to what we would have found replacing the realized measure by the latent integrated variance. Although it is evident that correcting for this attenuation bias can potentially lead to improved volatility forecasts, this issue has not received much attention in the literature. Recently, Bollerslev et al. (2016) found that, in a HAR model, letting the volatility persistence depend on the estimated degree of measurement error leads to remarkable improvements in the model's predictive performance. In the same vein, Shephard and Xiu (2016) found evidence that, in a GARCH-X model, the magnitude of the response coefficients associated with different realized volatility measures is related to the quality of the measure itself. Finally, Hansen and Huang (2016) observe that the response of the current conditional variance to past unexpected volatility shocks is negatively correlated with the accuracy of the associated realized volatility measure.

In this paper we develop a novel modelling approach that accounts for the attenuation bias effect in a natural and fully data-driven way. To this purpose, we first extend the standard Realized GARCH model by letting the variability of the measurement error vary over time as a function of an estimator of the integrated quarticity of intra-daily returns. In this way, we obtain a model-based time-varying estimate of the accuracy of the realized measure used. Consequently, we adjust the volatility dynamics for attenuation bias effects by allowing the response coefficient of the lagged realized volatility to depend on this quantity. In particular, the model is designed so that more weight is given to lagged volatilities when these are more accurately measured. Finally, the proposed modelling approach is further extended to explicitly model the impact of jumps on the predicted conditional variance of returns. This is achieved by introducing into the measurement equation an additional component that controls for the amount of bias generated by the occurrence of jumps. A notable feature of the proposed solution is that the jump correction only occurs on days in which jumps are effectively observed, while resorting to the use of more efficient standard measures, such as realized variances and kernels, in jumps-free periods.

The paper is organised as follows. In Section 2 the basic theoretical framework behind the computation of realized measures is reviewed, while Section 3 discusses the Realized GARCH model of Hansen et al. (2012). Section 4 illustrates the proposed time-varying parameter heteroskedastic Realized GARCH model. A jumps-free setting is considered first, then a modification of the proposed model, explicitly taking into account the impact of jumps, is introduced. QML estimation of the proposed models is discussed at the end of the same section. Sections 5 to 7 are dedicated to the empirical application. Section 5 presents the main features of the data used for the analysis; Section 6 focuses on the in-sample performance of the proposed models, taking the standard Realized GARCH model as a benchmark, whereas the out-of-sample forecasting performance is analysed in Section 7. Finally, Section 8 concludes.

2 Realized measures: a short review

In recent years, the availability of high-frequency financial market data has enabled researchers to build reliable measures of the latent daily volatility, based on the use of intra-daily returns. In the econometric and financial literature, these are widely known as realized volatility measures. The theoretical background to these measures is given by the dynamic specification of the price process in continuous time. Formally, let the logarithmic price p_t of a financial asset be determined by the stochastic differential process

$$dp_t = \mu_t dt + \sigma_t dW_t + dJ_t \qquad 0 \le t \le T \ , \tag{1}$$

where μ_t and σ_t are the drift and instantaneous volatility processes, respectively, whilst W_t is a standard Brownian motion; σ_t is assumed to be independent of W_t and J_t is a finite activity jump process. Under assumption of jump absence ($dJ_t = 0$) and a frictionless market, the logarithmic price p_t follows a semi-martingale process.

In that case, the Quadratic Variation (QV) of log-returns $r_t = p_t - p_{t-1}$ coincides with the Integrated Variance (IV), given by

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds \ . \tag{2}$$

In the absence of jumps, microstructure noise and measurement error, Barndorff-Nielsen and Shephard (2002) show that IV is consistently estimated by Realized Volatility (RV)

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2 \ , \tag{3}$$

where

$$r_{t,i} = p_{t-1+i\Delta} - p_{t-1+(i-1)\Delta}$$

is the i-th Δ-period intraday return, $M = 1/\Delta$. Although IV and the conditional variance of returns do not coincide, there is a precise relationship between these two quantities: under standard integrability conditions (Andersen et al., 2001) it can be shown that

$$E(IV_t|\mathscr{F}_{t-1}) = var(r_t|\mathscr{F}_{t-1}) \ ,$$

where $\mathscr{F}_{t-1}$ denotes the information set at time $(t-1)$. In other words, the optimal forecast of IV can be interpreted as the conditional variance of returns and the difference between these two quantities is given by a zero mean error.

Barndorff-Nielsen and Shephard (2002) show that RV consistently estimates the true latent volatility, when $\Delta \longrightarrow 0$, but in practice, due to data limitations, the following results hold

$$RV_t = IV_t + \varepsilon_t \tag{4}$$

and

$$\varepsilon_t \sim N(0, 2\Delta IQ_t), \tag{5}$$

where $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ is the Integrated Quarticity (IQ). This, in turn, can be consistently estimated as

$$RQ_t = \frac{M}{3} \sum_{i=1}^{M} r_{t,i}^4 \ . \tag{6}$$

On the other hand, if jumps are present, QV will differ from IV, with the difference given by the accumulated squared jumps. Formally, let

$$dJ_t = k_t dq_t \ ,$$

where $k_t = p_t - p_{t-}$ is the size of the jump in the log-price p_t and q_t is a counting process, with possibly time-varying intensity λ_t, such that

$$P(dq_t = 1) = \lambda_t dt \ .$$

Then, under the assumptions in Andersen et al. (2007)

$$RV_t \xrightarrow[p]{} QV_t = IV_t + \sum_{t-1 \le s \le t} k^2(s) \, .$$

Hence, RV is a consistent estimator of QV, but not of IV. An alternative here is to use jump-robust estimators, such as the *Bipower* and *Tripower Variation* (Barndorff-Nielsen and Shephard, 2004), *minRV* or *medRV* (Andersen et al., 2012), that are consistent for IV even in the presence of jumps. In the empirical applications carried out in this work, among the different proposals arising in the literature, focus here is put on the *medRV* estimator, mainly for theoretical reasons. Specifically, Andersen et al. (2012) show that in the jump-free case *"the medRV estimator has better theoretical efficiency properties than the tripower variation measure and displays better finite-sample robustness to both jumps and the occurrence of "zero" returns in the sample"*. In addition, unlike the *Bipower Variation* measure, for the *medRV* estimator an asymptotic limit theory in the presence of jumps is available.

The *medRV* estimator proposed by Andersen et al. (2012) is

$$medRV_t = \frac{\pi}{6 - 4\sqrt{3}\pi} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} med \left(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}| \right)^2 \, . \tag{7}$$

Nevertheless, in the jump-free case, these jump-robust estimators are substantially less efficient than the simple RV estimator: i.e. *Bipower* and *Tripower Variation*, *medRV* and *minRV* are all asymptotically normal, with asymptotic variance proportional (up to different scale factors) to the IQ (Andersen et al., 2012). Further, in presence of jumps, this quantity will be not consistently estimated by RQ; thus, some alternative jump-robust estimator will be needed. For the same reasons discussed above, focus here is on the *medRQ* estimator proposed by Andersen et al. (2012)

$$medRQ_t = \frac{3\pi M}{9\pi + 72 - 52\sqrt{3}} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} med \left(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}| \right)^4 \, . \tag{8}$$

A further issue is how to consistently estimate QV in the presence of market microstructure frictions. In this direction, several estimators are proposed in the literature to mitigate the influence of market microstructure noise, such as the Two Time Scales approach of Zhang et al. (2005), the Realized Kernel of Barndorff-Nielsen et al. (2008) and the pre-averaged RV of Jacod et al. (2009), among others. In this paper, the Realized Kernel (RK) is employed, specified as

$$RK = \sum_{h=-H}^{H} K\left(\frac{h}{H+1} \right) \zeta_H, \qquad \zeta_H = \sum_{j=|h|+1}^{M} r_{t,i} r_{t,i-|h|} \, , \tag{9}$$

where $K(\cdot)$ is a kernel weight function and H a bandwidth parameter[†].

3 Realized GARCH models

The Realized GARCH (RGARCH), introduced by Hansen et al. (2012), extends the class of GARCH models by first replacing squared returns, as the driver of the volatility dynamics, with a more efficient proxy, such as a RV measure. With this change alone, the resulting

[†] For details about the optimal choice of the kernel type and the bandwidth selection see Barndorff-Nielsen et al. (2009).

specification can be seen as a GARCH-X model, where the realized measure is used as an explanatory variable. A second extension is that the Realized GARCH "completes" the GARCH-X, by adding a measurement equation that explicitly models the contemporaneous relationship between the realized measure and the latent conditional variance.

Formally, let $\{r_t\}$ be a time series of stock returns and $\{x_t\}$ be a time series of realized measures of volatility. Focus here is on the logarithmic RGARCH model, defined via

$$r_t = \mu_t + \sqrt{h_t}\, z_t \,, \tag{10}$$

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma log(x_{t-1}) \,, \tag{11}$$

$$log(x_t) = \xi + \varphi\, log(h_t) + \tau(z_t) + u_t \,, \tag{12}$$

where $h_t = var(r_t|\mathscr{F}_{t-1})$ is the conditional variance and $\mathscr{F}_{t-1}$ the historical information set at time $t-1$. To simplify the exposition, in the reminder, it is assumed that the conditional mean $\mu_t = E(r_t|\mathscr{F}_{t-1}) = 0$. The innovations z_t and u_t are assumed to be mutually independent, with $z_t \overset{iid}{\sim} (0,1)$ and $u_t \overset{iid}{\sim} (0,\sigma_u^2)$.

The function $\tau(z_t)$ can accommodate leverage effects, since it captures the dependence between returns and future volatility. A common choice (see e.g. Hansen et al. (2012)), found to be empirically satisfactory, is

$$\tau(z_t) = \tau_1\, z_t + \tau_2(z_t^2 - 1) \,.$$

Substituting the measurement equation into the volatility equation, the model implies an AR(1) representation for $log(h_t)$

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma) log(h_{t-1}) + \gamma w_{t-1} \,, \tag{13}$$

where $w_t = \tau(z_t) + u_t$ and $E(w_t) = 0$. The coefficient $(\beta + \varphi\gamma)$ reflects the persistence in (the logarithm of) volatility, whereas γ represents the impact of both the lagged return and realized measure on future (log-)volatility. To ensure the volatility process h_t is stationary the required restriction is $\beta + \varphi\gamma < 1$.

Compared to the linear RGARCH, the log-linear specification has two main advantages: first, it is more flexible, since no constraints on the parameters are required in order to ensure positivity of the conditional variance, which holds automatically by construction; and second, the logarithmic transformation substantially reduces, but does not eliminate, the heteroskedasticity of the measurement equation error term. For these reasons, this paper exclusively focuses on the log-linear specification of the Realized GARCH model.

4 Time Varying Coefficient Heteroskedastic Realized GARCH models with dynamic attenuation bias

In this section a generalisation of the basic Realized GARCH specification is proposed that accounts and allows for the natural heteroskedasticity of the measurement error u_t, as well as for dynamic attenuation bias.

In a jump-free world, any consistent estimator of the *IV* can be written as the sum of the conditional variance plus a random innovation. Since the variance of this innovation term is function of the *IQ*, it seems natural to model the variance of the noise u_t in equation (12) as function of the *RQ*. Thus, it is assumed that the measurement noise variance is time-varying, i.e. $u_t \overset{iid}{\sim} (0,\sigma_{u,t}^2)$. In order to model the time-varying variance of the measurement noise, the specification

$$\sigma_{u,t}^2 = exp\left\{\delta_0 + \delta_1 log\left(\sqrt{RQ_t}\right)\right\} \tag{14}$$

is considered, where the exponential formulation guarantees the positivity of the estimated variance, without imposing constraints on the parameters δ_0 and δ_1. The resulting model is denoted the *Heteroskedastic* Realized GARCH (HRGARCH). It is easy to see that the homoskedastic Realized GARCH is nested within this class, by setting $\delta_1 = 0$, and that this restriction can be tested by means of a simple Wald statistic.

In order to account for dynamic attenuation effects in the volatility persistence, in the sense of Bollerslev et al. (2016), the basic HRGARCH specification is further extended, allowing for time-varying persistence in the volatility equation. This is achieved by letting γ, the impact coefficient of the lagged realized measure, depend on the time-varying variance of the measurement noise u_t. In line with Bollerslev et al. (2016), the impact of past realized measures on current volatility is expected to be down-weighted in periods in which the efficiency of the realized measure is low. The resulting model is called the *Time Varying Heteroskedastic* Realized GARCH (TV-HRGARCH).

The volatility updating equation of the TV-HRGARCH is given by

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma_t\, log(x_{t-1}), \tag{15}$$

where

$$\gamma_t = \gamma_0 + \gamma_1\, \sigma^2_{u,t-1} \tag{16}$$

and $\sigma^2_{u,t}$ follows the specification in (14). Accordingly, as its fixed coefficients counterpart, the TV-HRGARCH can be represented in terms of a time-varying coefficients AR(1) model for $log(h_t)$

$$log(h_t) = (\omega + \xi\gamma_t) + (\beta + \varphi\gamma_t) log(h_{t-1}) + \gamma_t\, w_{t-1}. \tag{17}$$

Also in this case, the simple Realized GARCH can be obtained by imposing the restriction $\delta_1 = \gamma_1 = 0$.

The setting which has been considered so far does not allow for the occurrence of jumps. Consideration is now given to a variant of the proposed modelling approach that, in order to capture this additional source of bias, features a jumps component as an additional variable in the measurement equation. This is achieved by adding the log-ratio between a non jump-robust realized measure x_t, such as the standard *RV* estimator or the *RK*, and a jump-robust realized measure x_t^J, as an explanatory variable. In the empirical application, as anticipated in Section 2, the *medRV* estimator proposed by Andersen et al. (2012) is employed.

Generally, let $C_t = x_t/x_t^J$. In the limit, this ratio will converge in probability to the ratio between QV and IV. Values of $C_t > 1$ are interpreted as providing evidence of jumps occurring at time t, while the discrepancy between the two measures is expected to disappear in absence of jumps, leading to values of $C_t \approx 1$. Naturally, sampling variability will play a role here and values $C_t < 1$ will be possible, in a small proportion of cases. This is compatible with the fact that the observed C_t is given by the combination of a latent signal $\bar{C}_t \geq 1$ and a measurement error, thus explaining observed values of C_t below the threshold 1. A simple way to avoid observations below 1 is to truncate the distribution of C_t at this threshold, setting all the values below the truncation point equal to 1 (see e.g. Andersen et al. (2007) for a similar approach). However, this does not ensure consistent filtering of the measurement error (the truncation on the left tail is somewhat arbitrary and the right tail would be untouched) with the potential drawback of introducing an additional source of bias into the analysis. Therefore, taking into account the limited empirical incidence of values of $C_t < 1$, it is decided to work with uncensored values of C_t.

After adding the bias correction variable C_t, the proposed modified measurement equation is

$$log(x_t) = \xi + \varphi log(h_t) + \eta\, log(C_t) + \tau(z_t) + u_t^* \tag{18}$$

or equivalently

$$log(x_t^*) = \xi + \varphi log(h_t) + \tau(z_t) + u_t^* \,, \tag{19}$$

where $log(x_t^*) = log(x_t/C_t^{\eta})$.

Considering the chosen RGARCH and the AR(1) representation for the log-conditional variance, it follows that

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma) log(h_{t-1}) + \gamma w_{t-1}^* \,, \tag{20}$$

where

$$w_t^* = \tau(z_t) + u_t^*$$

and

$$u_t^* = log(x_t^*) - \xi - \varphi log(h_t) - \tau(z_t). \tag{21}$$

By substituting equation (21) in (20), the log-conditional variance can be alternatively written as

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma log(x_{t-1}) - \gamma\eta\, log(C_{t-1}) \tag{22}$$

or equivalently

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma log(x_{t-1}^*). \tag{23}$$

In this modified framework, looking at equation (22), it then turns out that the log-conditional variance $log(h_t)$ is driven not only by past values of the realized measure but also, with opposite sign, by past values of the associated bias. The additional parameter η allows to adjust the contribution of C_{t-1}. From a different point of view, equations (19) and (23) suggest that the volatility updating equation can be rewritten in a form similar to that of the standard RGARCH model, with the substantial difference that volatility changes are driven instead by the bias-corrected measure $log\left(x_{t-1}/C_{t-1}^{\eta}\right) = log\left(x_{t-1}^*\right)$; the amount of correction is determined by the estimated scaling parameter η. This specification, of course, extends to the HRGARCH and TV-HRGARCH models.

By simple algebra it is easy to show that

$$\begin{aligned} log(x_t^*) &= log(x_t) - \eta log(C_t) \\ &= log(x_t) - \eta[log(x_t) - log(x_t^J)] \\ &= (1-\eta) log(x_t) + \eta log(x_t^J). \end{aligned}$$

If $0 < \eta < 1$, as it results from our empirical analysis, $log(x_t^*)$ is a weighted average of $log(x_t)$ and $log(x_t^J)$ and the parameter η can be seen as the weight to be assigned to the jump-robust log-transformed realized measure $log(x_t^J)$, whereas $(1-\eta)$ is the weight corresponding to the non-jump robust realized measure $log(x_t)$. If $\eta = 0$, $log(x_t^*)$ coincides with the non-robust realized measure $log(x_t)$, while for $\eta = 1$, $log(x_t^*)$ reduces to the jump robust log-transformed realized measure $log(x_t^J)$.

In the remainder, to distinguish models incorporating the bias correction variable C_t in the measurement equation, these models will be denoted by addition of the superscript " $*$ ", namely: RGARCH*, HRGARCH* and TV-HRGARCH*.

The model parameters can be estimated by standard Quasi Maximum Likelihood (QML) techniques. Let Y_t indicate any additional explanatory variable eventually included in the measurement equation. Following Hansen et al. (2012), the quasi log-likelihood function, conditionally on past information $\mathscr{F}_{t-1}$ and Y_t, is given by

$$\mathscr{L}(\mathbf{r}, \mathbf{x}; \theta) = \sum_{t=1}^{T} log\, f(r_t, x_t | \mathscr{F}_{t-1}, Y_t) \,,$$

where $\theta = (\theta_h', \theta_x', \theta_\sigma')'$ with θ_h, θ_x and θ_σ respectively being the vectors of parameters appearing in the volatility equation (θ_h), in the level of the measurement equation (θ_x) and in the noise variance specification (θ_σ).

An attractive feature of the Realized GARCH structure is that the conditional density $f(r_t, x_t|\mathscr{F}_{t-1}, Y_t)$ can be easily decomposed as

$$f(r_t, x_t|\mathscr{F}_{t-1}, Y_t) = f(r_t|\mathscr{F}_{t-1}) f(x_t|r_t; \mathscr{F}_{t-1}, Y_t).$$

Assuming a Gaussian specification for z_t and u_t, such as $z_t \overset{iid}{\sim} N(0,1)$ and $u_t \overset{iid}{\sim} N(0, \sigma_u^2)$, the quasi log-likelihood function is

$$\mathscr{L}(\mathbf{r}, \mathbf{x}; \theta) = \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(h_t) + \frac{r_t^2}{h_t}}_{\ell(\mathbf{r})} + \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(\sigma_u^2) + \frac{u_t^2}{\sigma_u^2}}_{\ell(\mathbf{x}|\mathbf{r})}. \quad (24)$$

Since standard GARCH models do not include an equation for x_t, the overall maximised log-likelihood values given by RGARCH models are not comparable to those returned from the estimation of standard GARCH-type models; the former will tend to be larger. Nevertheless, the partial log-likelihood value of the returns component, $\ell(\mathbf{r}) = \sum_{t=1}^{T} log\, f(r_t|\mathscr{F}_{t-1})$, can be still meaningfully compared to the maximised log-likelihood value achieved for a standard GARCH type model.

5 The Data

To assess the performance of the proposed models, an empirical application to four stocks traded on the Xetra Market in the German Stock Exchange has been performed. This section presents the salient features of the data analysed. In particular, the following assets are considered: Allianz (ALV), a financial services company dealing mainly with insurance and asset management; Bayerische Motoren Werke (BMW), a company engaged in vehicle and engine manufacturing; Metro Group (MEO), a cash and carry group and RWE (RWE), a company providing electric utilities.

The original dataset included tick-by-tick data on transactions (trades only) in the period 02/01/2002 to 27/12/2012. The raw data are cleaned, using the procedure described in Brownlees and Gallo (2006), then converted to an equally spaced series of five-minute log-returns, which are aggregated on a daily basis to: compute a time series of 2791 daily log-returns; two different realized volatility measures, *RV* and *RK*; the jump-robust *medRV* estimator and two realized quarticity measures, *RQ* and *medRQ*. Only continuous trading transactions during the regular market hours 9:00 am - 5:30 pm are considered.

Table 1 reports some descriptive statistics for daily log-returns (r_t), *RV*, *RK* and *medRV*; as well as for the bias correction variables related to RV_t and RK_t, denoted by C_t^{RV} and C_t^{RK}, respectively. For ease of presentation, the values associated with *RV* and *RK* are multiplied by 100. The daily returns have standard deviation typically around 0.020 and are slightly skewed, negatively so for ALV, BMW and MEO, but positively for RWE. Furthermore, the high kurtosis values indicate much heavier tails than the normal distribution, as expected. Looking at the realized measures, all *RV* and *RK* series present very strong positive skew; *medRV* has smaller standard deviations than *RV* and *RK* as it could be expected given its jump-robustness. The bias correction variables C_t^{RV} and C_t^{RK} have mean slightly above one and positive skewness. Their minimums are ≈ 0.75 and maximums $\in [2.66, 3.70]$. This preliminary analysis suggests that the impact of jumps ($C_t > 1$) is more important and

Table 1: Summary statistics

		Min.	1Qu.	Med.	Mean	3Qu.	Max.	S.dev.	Skew.	Kurt.
r_t	ALV	-0.147	-0.010	0.000	-0.001	0.008	0.135	0.021	-0.066	8.402
	BMW	-0.135	-0.010	0.000	0.000	0.010	0.153	0.020	-0.039	7.497
	MEO	-0.150	-0.010	-0.001	-0.001	0.009	0.122	0.019	-0.377	8.900
	RWE	-0.108	-0.009	0.000	-0.001	0.008	0.097	0.016	0.065	7.415
$RV_t \times 100$	ALV	0.002	0.011	0.021	0.050	0.047	1.732	0.089	6.999	93.681
	BMW	0.004	0.016	0.028	0.045	0.051	0.842	0.057	5.254	49.634
	MEO	0.004	0.016	0.026	0.045	0.048	1.047	0.060	5.030	48.277
	RWE	0.003	0.012	0.020	0.034	0.035	1.011	0.046	6.635	92.977
$RK_t \times 100$	ALV	0.002	0.011	0.021	0.049	0.047	1.730	0.089	7.012	94.289
	BMW	0.004	0.016	0.028	0.045	0.050	0.842	0.056	5.289	50.367
	MEO	0.004	0.015	0.025	0.044	0.047	1.041	0.059	5.077	49.355
	RWE	0.003	0.012	0.019	0.033	0.035	1.009	0.045	6.703	96.421
$medRV_t \times 100$	ALV	0.002	0.010	0.019	0.045	0.043	1.606	0.083	7.241	99.614
	BMW	0.003	0.014	0.025	0.041	0.046	0.773	0.052	4.999	44.906
	MEO	0.002	0.013	0.023	0.040	0.042	1.032	0.053	5.395	61.048
	RWE	0.002	0.011	0.018	0.031	0.032	0.876	0.042	5.991	75.030
$C_t^{RV} = \frac{RV_t}{medRV_t}$	ALV	0.760	1.000	1.100	1.134	1.219	2.655	0.195	1.536	7.596
	BMW	0.738	0.996	1.093	1.126	1.213	3.131	0.191	1.867	11.612
	MEO	0.744	1.015	1.120	1.162	1.263	3.704	0.223	2.355	17.469
	RWE	0.742	1.003	1.092	1.127	1.213	2.751	0.187	1.635	8.842
$C_t^{RK} = \frac{RK_t}{medRV_t}$	ALV	0.747	0.989	1.090	1.123	1.208	2.654	0.193	1.567	7.835
	BMW	0.736	0.984	1.080	1.114	1.196	3.126	0.190	1.889	11.761
	MEO	0.741	1.005	1.105	1.148	1.243	3.670	0.221	2.399	17.998
	RWE	0.742	0.988	1.076	1.111	1.195	2.747	0.185	1.684	9.111

Summary statistics of daily log-returns r_t, daily Realized Variance RV_t* (*: $\times 100$), daily Realized Kernel RK_t* (*: $\times 100$), daily $medRV_t*$ (*: $\times 100$), bias correction variable C_t^{RV} for RV_t and bias correction variable C_t^{RK} for RK_t . Sample period: January 2002 – December 2012. Min.: Minimum; 1Qu.: First Quartile; Med.: Median; Mean; 3Qu.: Third Quartile; Max.: Maximum; S.dev.: Standard deviation; Skew.: Skewness; Kurt.: Kurtosis.

prevalent compared to the measurement error bias ($C_t < 1$). These aspects are also confirmed by the distributional information on C_t presented in Table 2. Only approximately 5% of observations have C_t below 0.90.

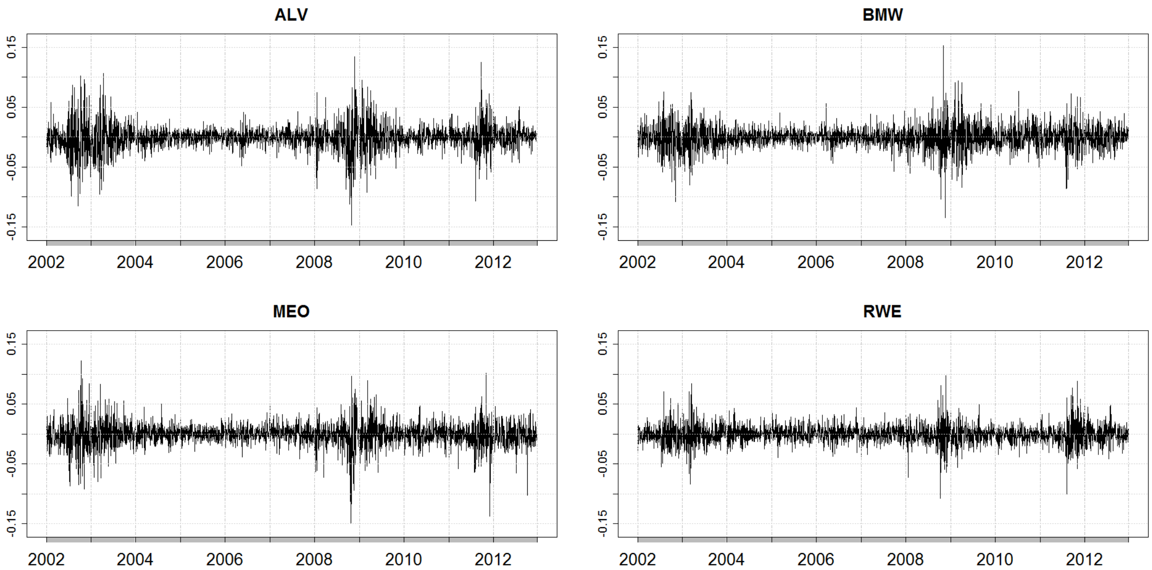
Figure 1 displays the daily returns for the four analysed stocks. These reveal three periods of high volatility common to all assets: the first relates to the dot com bubble in 2002; the second is the financial crisis starting in mid 2007 and peaking in 2008; the crisis in Europe then progressed from the banking system to a sovereign debt crisis, with the highest turmoil level in the late 2011, the 3rd period. These are clearly evident in Figure 2, reporting the time plots of the daily 5-minute RV series.

Finally, Figure 3 shows the evolution of the bias correction variables C_t over time. This fluctuates approximately around a base level ≈ 1, with an evident positive skewness due to the the upward peaks (jumps), while downward variations due to measurement noise appear to be much less pronounced and negligible.

Table 2: C_t distribution for RV and RK

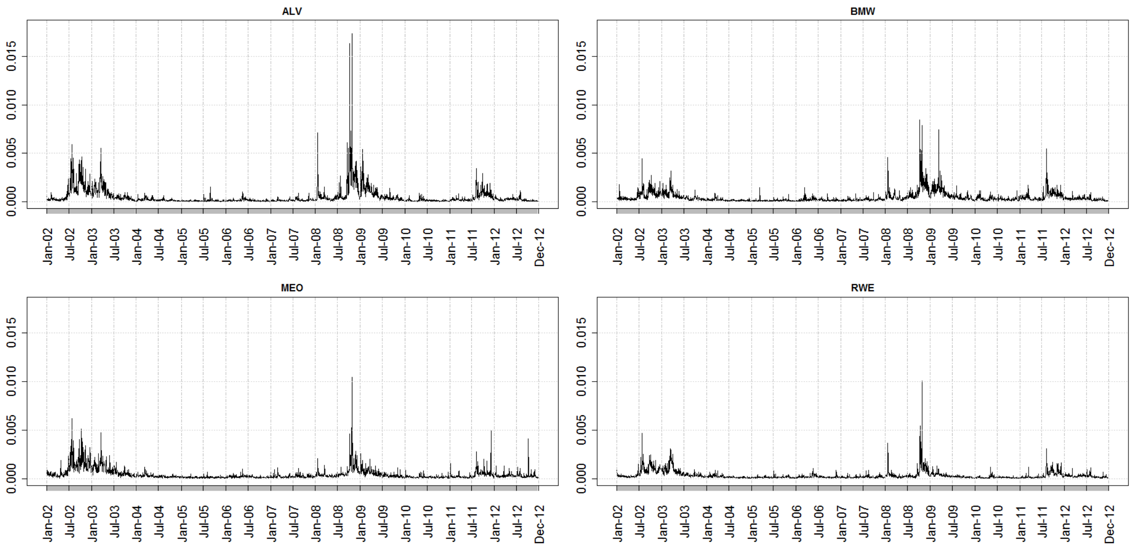
	Distribution of C_t^{RV}				Distribution of C_t^{RK}			
	ALV	BMW	MEO	RWE	ALV	BMW	MEO	RWE
0%	0.760	0.738	0.744	0.742	0.747	0.736	0.741	0.742
5%	0.894	0.893	0.900	0.895	0.887	0.883	0.892	0.883
10%	0.930	0.930	0.941	0.931	0.923	0.919	0.929	0.919
25%	1.000	0.996	1.015	1.003	0.989	0.984	1.005	0.988
50%	1.100	1.093	1.120	1.092	1.090	1.080	1.105	1.076
75%	1.219	1.213	1.263	1.213	1.208	1.196	1.243	1.195
90%	1.375	1.354	1.428	1.359	1.362	1.342	1.410	1.343
95%	1.492	1.475	1.565	1.477	1.472	1.463	1.542	1.453
100%	2.655	3.131	3.704	2.751	2.654	3.126	3.670	2.747

Figure 1: Time series of daily log-returns



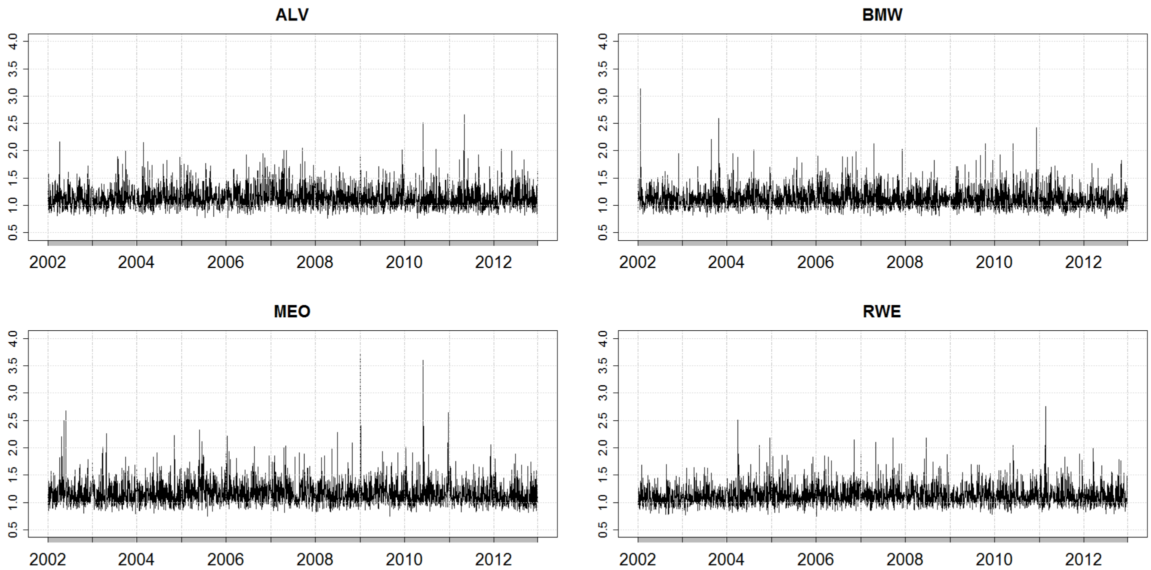
Daily log-returns for the stocks ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right) for the full sample period 02/01/2002 – 27/12/2012.

Figure 2: Daily Realized Volatility



Daily Realized Volatility computed using a sampling frequency of 5 minutes. ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right). Full sample period 02/01/2002 – 27/12/2012.

Figure 3: Time series of daily bias correction variable C_t^{RV}



Daily bias correction variable $C_t^{RV} = RV_t/medRV_t$ for the stocks ALV (top-left), BMW (top-right), MEO (bottom-left) and RWE (bottom-right) for the full sample period 02/01/2002 – 27/12/2012.

6 In-sample analysis

This section discusses the in-sample performance of the proposed models. The full data is employed here and focus is on the log-linear specification. It is worth noting that in presence of jumps, RQ_t could itself be affected by bias. So as a robustness check, in addition to the jump free models (RGARCH, HRGARCH and TV-HRGARCH) and the models in which the impact of jumps is considered in the measurement equation (RGARCH*, HRGARCH* and TV-HRGARCH*), we also consider a variant of the latter class of models where, in the specification of $\sigma^2_{u,t}$, RQ_t is replaced by the jump-robust estimator $medRQ$. In order to distinguish this class of models the subscript "MRQ" has been used.

6.1 In-sample estimation results

Estimation results using the 5-min RV as realized measure are reported in Table 3, showing parameter estimates and robust standard errors (in small font underneath), together with values of the log-likelihood $\mathscr{L}(\mathbf{r},\mathbf{x})$, partial log-likelihood $\ell(\mathbf{r})$ and Bayesian Information Criterion (BIC), for the four analysed stocks [‡].

The parameter ω is in most cases not significant, except for the TV-HRGARCH specifications, where its value is considerably greater than other models, since it is influenced by the dynamics of the time-varying coefficient γ_t. The parameter β is between 0.565 and 0.752, always taking the highest value for HRGARCH and the lowest for TV-HRGARCH, whereas φ takes values close to one with a limited variability across different models. This suggests that the log-transformed realized measure $log(x_t)$ is roughly proportional to the log-conditional variance. These results are in line with the findings in Hansen et al. (2012). The parameters of the leverage function $\tau(z)$ are always significant (except for τ_1 using the RGARCH* for MEO stock), with τ_1 negative and τ_2 positive, as expected.

The parameter δ_1 is always positive and statistically significant at the 0.05 level, thus giving empirical confirmation to the intuition that the variance of the measurement error $\sigma^2_{u,t}$ is time-varying and, in accordance with the asymptotic theory, suggesting that this is positively related to the IQ. This also implies that $\sigma^2_{u,t}$ tends to take on higher values in periods of turmoil and lower values when volatility tends to stay low, as it can be easily seen in Figure 4 which compares the constant variance σ^2_u estimated by RGARCH with the time-varying variance $\sigma^2_{u,t}$ given by the HRGARCH model. For the four analysed stocks, the trend of $\sigma^2_{u,t}$ follows the dynamics of the realized measure, being higher in turbulent periods and lower in calm periods, while the constant variance σ^2_u estimated within the RGARCH (red line in the plots) is approximately equal to the average level of the time-varying variance of the measurement noise.

Another key parameter is the coefficient γ, which summarises the impact of the realized measure on future volatility. It ranges from 0.240 to 0.402 for both RGARCH and HRGARCH and its value is generally increased by the introduction of the bias correction variable C_t in the measurement equation; this provides additional evidence supporting the idea that accounting for jumps further reduces the attenuation bias effect on γ. For the TV-HRGARCH this effect is explained, in an adaptive fashion, by the time-varying coefficient γ_t, depending on the past noise variance $\sigma^2_{u,t-1}$ through the slope coefficient γ_1. This is always positive and statistically significant at the 0.05 level, except for the TV-HRGARCH$^*_{MRQ}$ model fitted to RWE, with an associated p-value of 0.107. Interestingly, in the TV-HRGARCH* and TV-

[‡]Differently from Hansen et al. (2012) positive values are obtained for the log-likelihood. This is mainly due to the fact that they use percentage log-returns, which approximately fall in the range (-30, 30). It follows that the conditional variances are often above 1, returning positive log-variances that multiplied by -1 in the log-likelihood, explaining the comparatively large negative log-likelihoods that they typically get.

Table 3: In-Sample Estimation Results using 5-minutes Realized Volatility

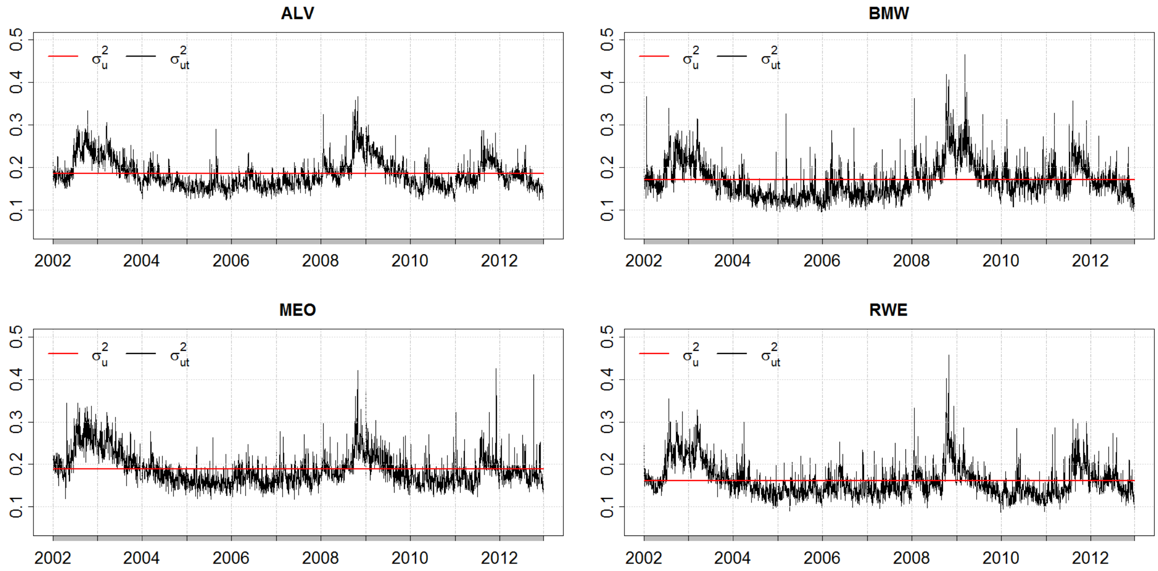
		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathscr{L}(\mathbf{r},\mathbf{x})$	*BIC*
ALV	RGARCH	0.008	0.402	-	-	0.598	-0.399	0.953	-0.069	0.108	-	0.185	-	-	7609.434	6006.082	-4.281
		(0.102)	0.032			0.031	(0.217)	0.027	0.008	0.007		0.008					
	HRGARCH	0.009	0.381	-	-	0.620	-0.473	0.948	-0.069	0.111	-	-	-0.405	0.162	7609.002	6032.686	-4.297
		(0.097)	0.030			0.029	0.217	0.027	0.007	0.007			(0.286)	0.034			
	TV-HRGARCH	1.995	-	0.489	1.060	0.565	-0.462	0.949	-0.069	0.111	-	-	-0.394	0.166	7611.877	6063.920	-4.317
		0.292		0.049	0.206	0.032	0.203	0.025	0.007	0.006			(0.228)	0.027			
	RGARCH*	0.002	0.408	-	-	0.590	-0.382	0.960	-0.068	0.107	0.398	0.180	-	-	7609.635	6042.332	-4.304
		(0.099)	0.031			0.031	(0.207)	0.026	0.008	0.006	0.045	0.008					
	HRGARCH*	0.010	0.392	-	-	0.607	-0.471	0.953	-0.068	0.110	0.388	-	-0.493	0.154	7609.412	6066.969	-4.319
		(0.085)	0.030			0.030	0.179	0.022	0.007	0.006	0.043		(0.296)	0.035			
	TV-HRGARCH*	1.265	-	0.446	0.747	0.575	-0.466	0.952	-0.069	0.110	0.285	-	-0.530	0.151	7610.934	6082.354	-4.327
		0.336		0.049	0.206	0.033	0.189	0.023	0.007	0.006	0.049		(0.284)	0.034			
	HRGARCH$^*_{MRQ}$	0.010	0.394	-	-	0.605	-0.467	0.954	-0.069	0.110	0.403	-	-0.539	0.145	7609.429	6065.011	-4.318
		(0.094)	0.030			0.030	0.202	0.025	0.007	0.006	0.043		(0.299)	0.035			
	TV-HRGARCH$^*_{MRQ}$	1.534	-	0.449	0.885	0.579	-0.469	0.954	-0.069	0.110	0.466	-	-0.499	0.151	7613.060	6083.501	-4.328
		0.381		0.053	0.211	0.032	0.182	0.023	0.007	0.006	0.043		(0.291)	0.034			
BMW	RGARCH	0.044	0.304	-	-	0.705	-0.520	0.927	-0.029	0.082	-	0.171	-	-	7470.471	5976.587	-4.260
		(0.111)	0.027			0.025	(0.325)	0.040	0.008	0.006		0.007					
	HRGARCH	0.036	0.280	-	-	0.727	-0.591	0.924	-0.032	0.086	-	-	0.321	0.270	7470.853	6026.783	-4.293
		(0.100)	0.025			0.023	(0.314)	0.039	0.008	0.006			(0.309)	0.039			
	TV-HRGARCH	1.614	-	0.410	0.617	0.692	-0.577	0.926	-0.032	0.085	-	-	0.147	0.250	7469.434	6053.708	-4.310
		0.316		0.039	0.143	0.025	0.234	0.029	0.008	0.006			(0.218)	0.027			
	RGARCH*	0.047	0.309	-	-	0.699	-0.526	0.930	-0.029	0.081	0.277	0.169	-	-	7471.600	5993.826	-4.270
		(0.110)	0.027			0.026	(0.314)	0.039	0.008	0.006	0.057	0.006					
	HRGARCH*	0.039	0.285	-	-	0.722	-0.588	0.927	-0.031	0.085	0.223	-	0.201	0.256	7471.749	6037.920	-4.298
		(0.087)	0.025			0.023	0.259	0.032	0.008	0.006	0.053		(0.310)	0.039			
	TV-HRGARCH*	1.358	-	0.389	0.524	0.696	-0.588	0.926	-0.032	0.085	0.138	-	0.115	0.247	7470.247	6058.335	-4.310
		0.324		0.040	0.135	0.025	0.258	0.032	0.008	0.006	0.055		(0.248)	0.031			
	HRGARCH$^*_{MRQ}$	0.041	0.287	-	-	0.720	-0.574	0.928	-0.031	0.084	0.243	-	-0.097	0.212	7471.793	6024.967	-4.289
		(0.107)	0.026			0.024	(0.330)	0.041	0.008	0.006	0.054		(0.359)	0.044			
	TV-HRGARCH$^*_{MRQ}$	1.067	-	0.349	0.469	0.707	-0.522	0.935	-0.031	0.083	0.309	-	-0.117	0.211	7469.812	6035.934	-4.294
		0.259		0.034	0.125	0.024	0.263	0.033	0.008	0.005	0.053		(0.300)	0.037			

Table 3 continued

		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathscr{L}(\mathbf{r},\mathbf{x})$	*BIC*
MEO	RGARCH	-0.080	0.252	-	-	0.740	-0.082	0.985	-0.018	0.082	-	0.189	-	-	7463.731	5828.595	-4.154
		(0.097)	0.025			0.024	(0.364)	0.043	0.009	0.007		0.007					
	HRGARCH	-0.077	0.240	-	-	0.752	-0.157	0.982	-0.023	0.090	-	-	-0.045	0.209	7463.768	5860.955	-4.174
		(0.095)	0.023			0.023	(0.371)	0.044	0.008	0.006			(0.294)	0.036			
	TV-HRGARCH	1.433	-	0.339	0.699	0.712	-0.178	0.978	-0.024	0.090	-	-	-0.225	0.188	7469.276	5890.380	-4.193
		0.275		0.041	0.152	0.026	(0.345)	0.041	0.008	0.006			(0.242)	0.030			
	RGARCH*	-0.072	0.257	-	-	0.734	-0.117	0.987	-0.016	0.081	0.362	0.185	-	-	7464.330	5860.417	-4.174
		(0.119)	0.025			0.024	(0.443)	0.053	(0.009)	0.007	0.054	0.007					
	HRGARCH*	-0.069	0.246	-	-	0.745	-0.181	0.984	-0.022	0.089	0.341	-	-0.151	0.198	7464.372	5888.727	-4.191
		(0.088)	0.023			0.023	(0.333)	0.039	0.008	0.006	0.052		(0.286)	0.035			
	TV-HRGARCH*	0.945	-	0.307	0.487	0.722	-0.191	0.980	-0.023	0.089	0.253	-	-0.278	0.183	7467.683	5905.626	-4.201
		(0.252)		0.035	0.145	0.025	(0.361)	0.042	0.008	0.006	0.055		(0.270)	0.034			
	HRGARCH$^*_{MRQ}$	-0.066	0.248	-	-	0.744	-0.193	0.982	-0.020	0.088	0.362	-	-0.253	0.180	7464.360	5883.987	-4.188
		(0.117)	0.025			0.023	(0.456)	0.054	0.008	0.007	0.052		(0.313)	0.038			
	TV-HRGARCH$^*_{MRQ}$	1.222	-	0.314	0.633	0.721	-0.211	0.980	-0.020	0.088	0.430	-	-0.308	0.174	7470.239	5903.041	-4.199
		0.324		0.042	0.175	0.025	(0.338)	0.040	0.008	0.007	0.051		(0.287)	0.035			
RWE	RGARCH	-0.406	0.315	-	-	0.644	0.736	1.067	-0.041	0.082	-	0.162	-	-	7991.189	6569.620	-4.685
		0.115	0.031			0.029	0.373	0.043	0.008	0.006		0.007					
	HRGARCH	-0.366	0.292	-	-	0.671	0.606	1.058	-0.040	0.083	-	-	0.293	0.262	7989.793	6614.132	-4.714
		0.151	0.030			0.024	(0.551)	0.063	0.007	0.005			0.371	0.044			
	TV-HRGARCH	0.933	-	0.389	0.613	0.634	0.660	1.063	-0.041	0.083	-	-	-0.128	0.212	7992.232	6631.893	-4.724
		0.319		0.043	0.195	0.028	(0.520)	0.059	0.007	0.005			(0.321)	0.038			
	RGARCH*	-0.410	0.317	-	-	0.640	0.743	1.072	-0.041	0.080	0.284	0.160	-	-	7991.585	6588.595	-4.696
		0.153	0.034			0.029	(0.509)	0.058	0.008	0.006	0.057	0.007					
	HRGARCH*	-0.377	0.294	-	-	0.666	0.639	1.064	-0.040	0.082	0.245	-	0.179	0.250	7990.196	6628.284	-4.721
		0.157	0.031			0.025	(0.570)	0.065	0.007	0.005	0.053		(0.404)	0.048			
	TV-HRGARCH*	0.458	-	0.357	0.374	0.646	0.675	1.067	-0.040	0.082	0.180	-	-0.046	0.223	7991.355	6638.507	-4.726
		(0.320)		0.040	0.159	0.027	(0.544)	0.062	0.007	0.005	0.054		(0.344)	0.041			
	HRGARCH$^*_{MRQ}$	-0.379	0.295	-	-	0.665	0.649	1.065	-0.041	0.081	0.271	-	0.115	0.236	7990.271	6622.699	-4.717
		0.122	0.028			0.025	(0.430)	0.049	0.007	0.005	0.053		(0.402)	0.047			
	TV-HRGARCH$^*_{MRQ}$	0.284	-	0.338	0.303	0.654	0.677	1.069	-0.041	0.081	0.308	-	-0.119	0.209	7990.301	6627.033	-4.718
		(0.316)		0.033	(0.188)	0.025	(0.409)	0.047	0.007	0.005	0.055		(0.440)	0.052			

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012. $\ell(\mathbf{r})$: partial log-likelihood. $\mathscr{L}(\mathbf{r},\mathbf{x})$: log-likelihood. BIC: Bayesian Information Criterion. Standard errors are reported in small font under the parameters value: in parenthesis parameter not significant at 5%.

Figure 4: Constant versus time-varying variance of the noise u_t of the HRGARCH fitted using the 5 minute *RV*



The Figure shows the constant variance σ^2_u (red-line) estimated with RGARCH together with the time-varying variance $\sigma^2_{u,t}$ (black-line) estimated with HRGARCH. Both models have been fitted taking the 5-minutes *RV* as volatility proxy. Full sample period 02 January 2002 - 27 December 2012.

HRGARCH$^*_{MRQ}$ models, the introduction of the bias correction variable in the measurement equation has the effect of reducing the value of the fitted γ_1, compared to what we find for the TV-HRGARCH.

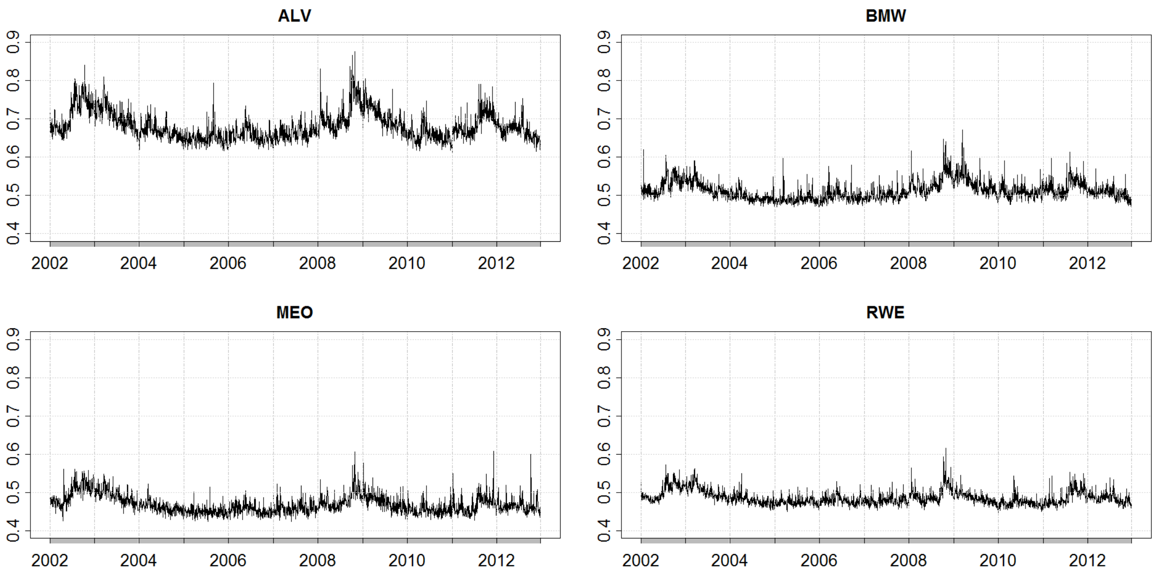
The value of γ_1 determines the amount by which the response to past volatility is corrected to account for attenuation bias effects. Since γ_1 is always positive and $log(x_t)$ is negative, when the lagged variance of the error term of the realized measure $\sigma^2_{u,t-1}$ is high, the impact of the lagged log-transformed realized measure $log(x_{t-1})$ on $log(h_t)$ will be negative and lower than what would have been implied by the same value of $log(x_{t-1})$ in correspondence of a lower value of $\sigma^2_{u,t-1}$. Said differently, the impact of x_{t-1} on h_t will be down-scaled towards zero when $\sigma^2_{u,t-1}$ increases. Equivalently, variations in h_t ($\nabla h_t = h_t - h_{t-1}$) will be negatively correlated with the values of γ_t and $\sigma^2_{u,t-1}$. These results are in line with the recent findings of Bollerslev et al. (2016).

Figure 5 displays the time plot of the γ_t coefficient for the four considered stocks. It is evident that when the variance of the measurement error is high, γ_t is also high, leading to a less substantial increase of h_t compared to days in which, ceteris paribus, $\sigma^2_{u,t}$ is low and the realized measure provides a stronger more reliable signal. Further, the value of γ_t tends to be higher than the value of the time invariant γ estimated within the RGARCH and HRGARCH models.

Finally, it is interesting to note that, for each series, the variance $\sigma^2_{u^*}$ of the measurement equation error u^*_t of the RGARCH* model is slightly lower than what observed for the RGARCH model, providing evidence of an improved goodness of fit in the modified measurement equation and efficiency of the bias corrected realized measure x^*_t. In this context, an important role is played by the smoothing parameter η, determining the amount of jump-implied bias correction associated to the ratio C_t. Overall, the variability of the estimated η coefficients across assets makes evident the flexibility of our modelling approach: the amount of smoothing in $log(x^*_t)$ is not arbitrarily chosen, but data driven through the estimated parameter η.

Reminding the interpretation of η discussed in Section 4, since the estimated η results lower than 0.5 for all the analysed assets, the impact of the realized measure $log(x_t)$ is, on average, greater than the impact of the jump-robust realized measure $log(x_t^J)$. Furthermore, as a general trend we observe that, for all stocks, the fitted value of η tends to be higher for heteroskedastic models relying on the jump robust quarticity estimator *medRQ*, (TV-)HRGARCH$^*_{MRQ}$, rather than for models based on *RQ*, (TV-)HRGARCH*.

Figure 5: Time-varying coefficient γ_t given by the TV-HRGARCH model



The Figure shows the time-varying coefficient $\gamma_t = \gamma_0 + \gamma_1 \sigma^2_{u,t-1}$ for the full sample period 02 January 2002 - 27 December 2012.

As a further robustness check, all the models are re-estimated using the 5-minute Realized Kernel as volatility proxy. The results and conclusions are very similar to those obtained using the 5-minute *RV*. The estimated parameters and standard errors obtained using *RK* as a volatility proxy have been reported in Table 8 in the Empirical Appendix. In the next subsection we provide a detailed analysis of the goodness of fit achieved through the models considered.

6.2 Log-likelihood analysis

To evaluate the performance of the estimated models we first refer to the commonly used Bayesian Information Criterion (BIC).

From the top panel of Table 4 it clearly emerges that the specifications accounting for both heteroskedasticity and attenuation bias effects, within the class of models corrected by the C_t variable, tend to minimise the BIC. In particular, the TV-HRGARCH* features the lowest value of the considered information criterion in three cases out of four, whereas for the stock ALV the TV-HRGARCH$^*_{MRQ}$ prevails, closely followed by the TV-HRGARCH* model. Interestingly, in the latter case, the parameter η given by the use of $medRQ_t$ instead of RQ_t as state variable of $\sigma^2_{u,t}$ into TV-HRGARCH* specification, takes on the highest estimated values within the class of models built using information on C_t (see Table 3), meaning that the jump-robust realized measure *medRV* has an higher impact in determining volatility dynamics for ALV compared to the other stocks. The bottom panel of Table 4 shows that similar results

Table 4: BIC in-sample comparison

		ALV	BMW	MEO	RWE
RV	RGARCH	-4.281	-4.260	-4.154	-4.685
	HRGARCH	-4.297	-4.293	-4.174	-4.714
	TV-HRGARCH	-4.317	-4.310	-4.193	-4.724
	RGARCH*	-4.304	-4.270	-4.174	-4.696
	HRGARCH*	-4.319	-4.298	-4.191	-4.721
	TV-HRGARCH*	-4.327	**-4.310**	**-4.201**	**-4.726**
	HRGARCH$^*_{MRQ}$	-4.318	-4.289	-4.188	-4.717
	TV-HRGARCH$^*_{MRQ}$	**-4.328**	-4.294	-4.199	-4.718
RK	RGARCH	-4.271	-4.250	-4.137	-4.674
	HRGARCH	-4.288	-4.283	-4.157	-4.704
	TV-HRGARCH	-4.306	-4.297	-4.171	-4.711
	RGARCH*	-4.299	-4.262	-4.163	-4.688
	HRGARCH*	-4.314	-4.290	-4.180	-4.714
	TV-HRGARCH*	-4.320	**-4.300**	-4.186	**-4.716**
	HRGARCH$^*_{MRQ}$	-4.312	-4.281	-4.177	-4.710
	TV-HRGARCH$^*_{MRQ}$	**-4.323**	-4.286	**-4.187**	-4.710

BIC values for the analysed models using the 5-min *RV* (first panel) and the 5-min *RK* (second panel). Best models are reported in **bold**.

Table 5: Likelihood Ratio statistics computed using the RGARCH as benchmark model

		ALV	BMW	MEO	RWE	Average
RV	HRGARCH	53.208	100.392	64.72	89.024	76.836
	TV-HRGARCH	115.676	154.242	123.57	124.546	129.508
	RGARCH*	72.500	34.478	63.644	37.950	52.143
	HRGARCH*	121.774	122.666	120.264	117.328	120.508
	TV-HRGARCH*	152.544	163.496	154.062	137.774	151.969
	HRGARCH$^*_{MRQ}$	117.858	96.76	110.784	106.158	107.890
	TV-HRGARCH$^*_{MRQ}$	154.838	118.694	148.892	114.826	134.312
RK	HRGARCH	54.932	98.828	64.582	89.456	76.949
	TV-HRGARCH	112.462	146.648	112.002	117.262	122.093
	RGARCH*	84.232	42.500	81.208	46.534	63.619
	HRGARCH*	134.232	127.994	136.640	125.086	130.988
	TV-HRGARCH*	160.270	162.366	161.556	139.658	155.962
	HRGARCH$^*_{MRQ}$	130.434	102.370	127.622	114.492	118.730
	TV-HRGARCH$^*_{MRQ}$	167.740	124.328	164.730	122.190	144.747

apply when the 5-min Realized Kernel is used as realized measure in the measurement equation x_t. In this case, the BIC is minimised by TV-HRGARCH* for BMW and RWE and by TV-HRGARCH$^*_{MRQ}$, again closely followed by TV-HRGARCH*, for ALV and MEO.

The BIC analysis also suggests that the standard Realized GARCH provides the highest BIC values, but the simple inclusion of the C_t variable (RGARCH*) remarkably improves the fit highlighting the importance of modelling the bias due to the occurrence of jumps.

Information criteria, such as BIC, can be used to compare competing models, but they do not allow to assess the statistical significance of differences in goodness of fit among candidate models. Since the Realized GARCH can be obtained as a special case of each of the models estimated in the previous section, we refer to the Likelihood Ratio Test (LRT) to perform pairwise comparisons of the performance of the proposed models with that of the simple RGARCH, taken as a benchmark. The LRT statistic for the i-th model is then given by:

$$LRT_i = -2\left[\mathscr{L}(\mathbf{r},\mathbf{x})_{RGARCH} - \mathscr{L}(\mathbf{r},\mathbf{x})_i\right].$$

It is worth remarking that, in a ML context, the LRT statistic is asymptotically distributed as a χ^2 random variable with degrees of freedom given by the number of constrained parameters under the null. However, if the estimation is carried out through QML, the usual χ^2 value corresponding to the desired statistical significance should be considered as approximate, since the statistical test behaves asymptotically as a weighted sum of independent chi-squares (it is necessary to compute critical values by numerical simulation). In our analysis we compute p-values referring to the χ^2 distribution and hence these are not exact but should be considered only "indicative of significance" (see also (Hansen et al. (2012))).

Table 5 presents the values of LRT statistic both for RV (top panel) and RK (bottom panel). In all the cases, the benchmark model is rejected. Even in the case of the simple HRGARCH model, the LRT statistics are always remarkably larger than the indicative χ^2_1 critical value, providing evidence in favour of the heteroskedastic nature of the measurement error u_t, in agreement with our findings in subsection 6.1.

Summarising: the in-sample results show that the introduction of heteroskedasticity and time-varying persistence, as well as the bias correction for jumps and measurement errors, has positive effects on the accuracy of the estimated volatility. Consequently, in sample, the proposed specifications show notable improvements over the standard RGARCH in terms of goodness of fit. Overall, the models incorporating heteroskedasticity, time-varying volatility response coefficient and jumps correction are those returning the best performances.

7 Out-of-sample Analysis

In this section, the out-of-sample predictive ability of the models estimated in Section 6 is assessed via a rolling window forecasting exercise, using an estimation window of 1500 days. The out-of-sample period starts on 26 November 2007 and includes 1270 daily observations, covering the credit crisis and the turbulent period from November 2011 to the beginning of 2012. For the sake of brevity, we only report results obtained for models fitted using the 5-min RV as volatility proxy. However, very similar performances are obtained when the 5-min RV is replaced by the 5-min RK estimator (the results are reported in the Empirical Appendix).

In order to assess the forecasting performance of the proposed models, the predictive (quasi) log-likelihood and the QLIKE loss function (Patton, 2011) are employed. Furthermore, the Model Confidence Set (MCS) of Hansen et al. (2011) is used to evaluate the comparative predictive ability of all the models, considering a confidence level of 75%. In particular, the Semi-Quadratic statistic, based on a block-bootstrap procedure with 5000 re-samples, is employed to sequentially test the hypothesis of equal predictive ability, where the optimal block length has been chosen through the method described in Patton et al. (2009). It is worth noting that the QLIKE loss function specifically measures the ability in forecasting the conditional variance of returns, while the predictive log-likelihood assesses the ability of a given model to predict the conditional distribution of $(r_t, x_t|\mathscr{F}_{t-1}, Y_t)$, thus considering both components of the Realized GARCH quasi likelihood.

Table 6: Predictive log-likelihood using 5-min *RV* as volatility proxy (top): in **bold** the preferred model according to predictive log-likelihood. MCS p-values of predictive log-likelihood (bottom): in box models $\in$ 75% MCS. The p-values refer to the negative predictive log-likelihoods.

	ALV	BMW	MEO	RWE
RGARCH	2617.961	2478.886	2498.497	2938.425
HRGARCH	2635.038	2500.582	2542.889	2974.401
TV-HRGARCH	2654.591	2515.229	2560.055	2982.102
RGARCH*	2634.241	2488.724	2523.193	2958.191
HRGARCH*	2651.794	2507.150	2560.524	2987.250
TV-HRGARCH*	**2660.343**	**2517.435**	2569.528	**2990.083**
TV-HRGARCH$_{MRQ}$	2633.243	2495.127	2535.734	2964.466
HRGARCH$^*_{MRQ}$	2651.363	2502.983	2555.388	2983.233
TV-HRGARCH$^*_{MRQ}$	2659.316	2507.181	**2571.215**	2982.434
	MCS p-values			
RGARCH	0.0006	0.0022	0.0000	0.0026
HRGARCH	0.0060	0.0322	0.0012	0.0578
TV-HRGARCH	$\boxed{0.3128}$	0.1182	0.0886	0.2200
RGARCH*	0.0830	0.0322	0.0008	0.0468
HRGARCH*	0.1086	0.1182	0.1858	$\boxed{0.5568}$
TV-HRGARCH*	$\boxed{1.0000}$	$\boxed{1.0000}$	$\boxed{0.8324}$	$\boxed{1.0000}$
HRGARCH$^*_{MRQ}$	0.0856	0.0730	0.0386	$\boxed{0.2692}$
TV-HRGARCH$^*_{MRQ}$	$\boxed{0.8762}$	0.1182	$\boxed{1.0000}$	0.2200

The first criterion we use for assessing predictive accuracy is, as in Hansen et al. (2012), the predictive log-likelihood, given for time $t+1$ by:

$$\hat{\mathscr{L}}(\mathbf{r},\mathbf{x})_{t+1} = -\frac{1}{2}\left[log(2\pi)+log(\hat{h}_{t+1})+\frac{r^2_{t+1}}{\hat{h}_{t+1}}\right] - \frac{1}{2}\left[log(2\pi)+log(\hat{\sigma}^2_u)+\frac{u^2_{t+1}}{\hat{\sigma}^2_u}\right]. \quad (25)$$

Subsequently, the aggregated predictive log-likelihood is computed by summing the density estimates for each day in the forecast period. Table 6 shows the values of the predictive log-likelihood corresponding to all models employed. The models including the bias correction variable C_t in the measurement equation maximise the predictive log-likelihood in all cases. In particular, for ALV, BMW and RWE, the specification that allows for heteroskedasticity and time-varying persistence, TV-HRGARCH*, returns the highest values of the predictive log-likelihood, whereas for MEO, the TV-HRGARCH$^*_{MRQ}$, replacing RQ_t by $medRQ_t$ in the specification of $\sigma^2_{u,t}$, prevails.

The bottom panel of Table 6 shows the MCS p-values associated to the predictive log-likelihood (multiplied by -1). Interestingly, the only model always coming into the MCS at the 75% confidence level is the TV-HRGARCH*, while the standard Realized GARCH never enters the set of superior models. For the asset ALV, all the time-varying specifications enter the MCS at the considered confidence level, while for BMW, the TV-HRGARCH* uniquely enters the set. On the other hand, for MEO, both the TV-HRGARCH* and TV-HRGARCH$^*_{MRQ}$ are included into the MCS. Finally, the HRGARCH* and the HRGARCH$^*_{MRQ}$ belong to the set of superior models only in the analysis of the RWE stock. As expected, we obtain analogous

Table 7: Average values of QLIKE loss using 5-min *RV* as volatility proxy (top) and MCS p-values (bottom). For each stock in **bold** minimum loss and in box models $\in$ 75% MCS.

	QLIKE			
	ALV	BMW	MEO	RWE
RGARCH	-6.9723	-6.6765	-6.9460	-7.2618
HRGARCH	-6.9721	-6.6755	-6.9464	-7.2606
TV-HRGARCH	**-6.9755**	**-6.6788**	**-6.9529**	**-7.2652**
RGARCH*	-6.9728	-6.6775	-6.9480	-7.2625
HRGARCH*	-6.9728	-6.6763	-6.9482	-7.2612
TV-HRGARCH*	-6.9746	-6.6788	-6.9525	-7.2647
HRGARCH^*_{MRQ}	-6.9728	-6.6766	-6.9482	-7.2614
$\text{TV-HRGARCH}^*_{MRQ}$	-6.9754	-6.6776	-6.9525	-7.2620

	MCS p-values			
RGARCH	0.0766	0.1616	0.0286	0.0294
HRGARCH	0.0772	0.0778	0.0714	0.0230
TV-HRGARCH	1.0000	1.0000	1.0000	1.0000
RGARCH*	0.0772	0.4294	0.1336	0.0310
HRGARCH*	0.0772	0.1534	0.2692	0.0280
TV-HRGARCH*	0.4234	0.7902	0.8038	0.0758
HRGARCH^*_{MRQ}	0.0772	0.2548	0.2174	0.0280
$\text{TV-HRGARCH}^*_{MRQ}$	0.9418	0.4294	0.8196	0.0294

results when the 5-min *RK* is used (Table 9 in the Empirical Appendix).

As a further criterion for assessing and comparing the forecasting accuracy of the fitted models, the QLIKE loss function is employed. This choice is motivated by two considerations. First, the QLIKE is robust to noisy volatility proxies (Patton, 2011). Second, compared to other robust alternatives, this loss function has been found to be more powerful in rejecting poorly performing predictors (Liu et al., 2015). The QLIKE loss has been computed according to the formula

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T} log(\hat{h}_t) + \frac{x_t}{\hat{h}_t}, \tag{26}$$

where the 5-minute *RV* and *RK* are chosen as alternative volatility proxies. In this section, we only report results obtained for the 5-minute *RV*. Again, very similar results are obtained when using the 5-min *RK* (Table 10 in the Empirical Appendix).

Observing the average values of the QLIKE loss function (top panel of Table 7), we find that the lowest value is always obtained by using the TV-HRGARCH model. The MCS p-values, in the bottom panel of Table 7, show that the TV-HRGARCH is always included in the MCS and, for RWE, it is the only specification appearing in the MCS. Within the class of models which include a jump correction component in the measurement equation, the RGARCH* is included into the MCS for BMW, whereas the HRGARCH* falls into the MCS only for MEO. The TV-HRGARCH* is always included in the set of superior models, except for RWE. This result also applies to the $\text{TV-HRGARCH}^*_{MRQ}$, while the simplified heteroskedastic specification HRGARCH^*_{MRQ} enters the MCS just for BMW. On the other hand, the standard RGARCH never comes into the set of superior models, nor does the HRGARCH.

8 Conclusion

We propose a generalisation of the class of Realized GARCH models that accounts for heteroskedasticity in measurement error and which explicitly reduces the magnitude of the attenuation bias through the temporal variation of the parameters driving the volatility persistence. Furthermore, in order to deal with the presence of jumps in stock prices, we further extend the proposed modelling approach introducing a bias correction variable that allows to control the impact of jumps on the predicted volatility in a fully data driven fashion.

The empirical analysis points out that our modelling approach outperforms the standard Realized GARCH both in fitting and forecasting volatility. In particular we find evidence in favour of the use of models accounting for heteroskedasticity, time-varying attenuation bias as well as the presence of jumps. More in detail, focusing on the out-of-sample predictive performance, we find that, when the predictive log-likelihood is used as a measure of accuracy, the TV-HRGARCH* model, incorporating all the mentioned features, is the only model always entering the 75% MCS, while the TV-HRGARCH model, not incorporating the jump correction term, only enters the MCS for the stock ALV. Moving to consider the QLIKE as a measure of predictive accuracy, the performance of the TV-HRGARCH* model is still remarkably good since this model enters the MCS in three cases out of four. However, it is slightly outperformed by the TV-HRGARCH model entering the MCS for all the stocks considered. Furthermore, the use of different realized volatility, *RV* and *RK*, and quarticity measures, *RQ* and *medRQ*, substantially confirms the robustness of our findings.

Empirical Appendix

Table 8: In-Sample Estimation Results using 5-minutes Realized Kernel

		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathscr{L}(\mathbf{r},\mathbf{x})$	BIC
ALV	RGARCH	0.012	0.399	-	-	0.602	-0.409	0.953	-0.067	0.109	-	0.186	-	-	7608.548	5992.439	-4.271
		(0.105)	0.032			0.031	(0.229)	0.028	0.008	0.007		0.008					
	HRGARCH	0.014	0.379	-	-	0.622	-0.489	0.947	-0.068	0.112	-	-	-0.375	0.165	7608.163	6019.905	-4.288
		(0.090)	0.030			0.029	0.196	0.024	0.007	0.007			(0.299)	0.036			
	TV-HRGARCH	1.930	-	0.485	1.004	0.568	-0.476	0.949	-0.068	0.112	-	-	-0.378	0.167	7610.464	6048.670	-4.306
		0.309		0.050	0.203	0.032	0.206	0.025	0.007	0.006			(0.234)	0.028			
	RGARCH*	0.003	0.406	-	-	0.592	-0.385	0.961	-0.067	0.108	0.429	0.181	-	-	7609.049	6034.555	-4.299
		(0.091)	0.031			0.031	0.188	0.023	0.008	0.006	0.044	0.008					
	HRGARCH*	0.013	0.392	-	-	0.607	-0.480	0.953	-0.068	0.110	0.417	-	-0.480	0.155	7608.862	6059.555	-4.314
		(0.094)	0.030			0.030	0.202	0.025	0.007	0.006	0.043		(0.299)	0.036			
	TV-HRGARCH*	1.118	-	0.440	0.648	0.579	-0.474	0.953	-0.068	0.111	0.327	-	-0.516	0.152	7609.911	6072.574	-4.320
		0.279		0.044	0.174	0.032	0.173	0.021	0.007	0.006	0.048		0.284	0.034			
	HRGARCH$^*_{MRQ}$	0.013	0.393	-	-	0.605	-0.475	0.954	-0.068	0.111	0.432	-	-0.526	0.146	7608.885	6057.656	-4.312
		(0.086)	0.030			0.030	0.180	0.022	0.007	0.006	0.043		(0.319)	0.037			
	TV-HRGARCH$^*_{MRQ}$	1.573	-	0.451	0.891	0.578	-0.477	0.954	-0.068	0.110	0.493	-	-0.483	0.153	7612.481	6076.309	-4.323
		0.461		0.057	0.226	0.032	0.198	0.024	0.007	0.006	0.043		(0.286)	0.033			
BMW	RGARCH	0.042	0.300	-	-	0.707	-0.513	0.929	-0.028	0.082	-	0.172	-	-	7470.179	5962.526	-4.250
		(0.115)	0.027			0.025	(0.341)	0.042	0.008	0.006		0.007					
	HRGARCH	0.035	0.277	-	-	0.729	-0.585	0.926	-0.032	0.086	-	-	0.312	0.268	7470.559	6011.940	-4.283
		(0.086)	0.024			0.023	0.265	0.033	0.008	0.006			(0.325)	0.041			
	TV-HRGARCH	1.508	-	0.397	0.570	0.698	-0.571	0.928	-0.032	0.086	-	-	0.155	0.250	7469.159	6035.850	-4.297
		0.300		0.038	0.127	0.024	(0.361)	0.044	0.008	0.005			(0.237)	0.030			
	RGARCH*	0.046	0.307	-	-	0.701	-0.523	0.931	-0.029	0.081	0.308	0.170	-	-	7471.498	5983.776	-4.262
		(0.092)	0.027			0.025	0.260	0.032	0.008	0.006	0.056	0.007					
	HRGARCH*	0.039	0.283	-	-	0.723	-0.589	0.928	-0.032	0.085	0.256	-	0.179	0.252	7471.632	6026.523	-4.290
		(0.109)	0.025			0.023	(0.341)	0.042	0.008	0.006	0.051		(0.325)	0.041			
	TV-HRGARCH*	1.211	-	0.373	0.465	0.702	-0.585	0.928	-0.032	0.085	0.182	-	0.107	0.245	7470.231	6043.709	-4.300
		0.321		0.040	0.123	0.024	(0.330)	0.041	0.008	0.006	0.053		(0.267)	0.034			
	HRGARCH$^*_{MRQ}$	0.040	0.285	-	-	0.721	-0.570	0.930	-0.031	0.084	0.276	-	-0.128	0.207	7471.682	6013.711	-4.281
		(0.111)	0.026			0.023	(0.345)	0.043	0.008	0.006	0.052		(0.341)	0.042			
	TV-HRGARCH$^*_{MRQ}$	1.069	-	0.346	0.475	0.708	-0.517	0.937	-0.031	0.084	0.341	-	-0.139	0.207	7469.674	6024.690	-4.286
		0.308		0.038	0.129	0.024	0.257	0.032	0.008	0.005	0.053		(0.325)	0.040			

Table 8 continued

		ω	γ	γ_0	γ_1	β	ξ	φ	τ_1	τ_2	η	σ_u^2	δ_0	δ_1	$\ell(\mathbf{r})$	$\mathscr{L}(\mathbf{r},\mathbf{x})$	*BIC*
MEO	RGARCH	-0.077	0.250	-	-	0.742	-0.097	0.985	-0.018	0.082	-	0.192	-	-	7463.020	5804.847	-4.137
		(0.097)	0.025			0.025	(0.363)	0.043	0.009	0.007		0.007					
	HRGARCH	-0.074	0.239	-	-	0.753	-0.171	0.981	-0.024	0.091	-	-	-0.032	0.209	7463.082	5837.138	-4.157
		(0.101)	0.024			0.024	(0.403)	0.048	0.008	0.006			(0.290)	0.036			
	TV-HRGARCH	1.262	-	0.324	0.609	0.719	-0.181	0.979	-0.024	0.090	-	-	-0.215	0.187	7467.454	5860.848	-4.171
		0.291		0.040	0.159	0.027	(0.335)	0.040	0.008	0.006			(0.249)	0.031			
	RGARCH*	-0.069	0.256	-	-	0.735	-0.130	0.987	-0.016	0.081	0.412	0.187	-	-	7463.913	5845.451	-4.163
		(0.091)	0.024			0.024	(0.328)	0.039	(0.009)	0.007	0.053	0.007					
	HRGARCH*	-0.064	0.245	-	-	0.746	-0.199	0.983	-0.021	0.089	0.392	-	-0.165	0.195	7463.935	5873.167	-4.180
		(0.094)	0.023			0.023	(0.358)	0.042	0.008	0.006	0.051		(0.293)	0.036			
	TV-HRGARCH*	0.755	-	0.293	0.389	0.729	-0.200	0.982	-0.022	0.090	0.323	-	-0.269	0.183	7466.313	5885.625	-4.186
		0.278		0.036	0.128	0.025	(0.447)	0.053	0.008	0.006	0.053		(0.283)	0.035			
	HRGARCH$^*_{MRQ}$	-0.064	0.247	-	-	0.744	-0.204	0.983	-0.020	0.088	0.412	-	-0.256	0.178	7463.961	5868.658	-4.177
		(0.117)	0.025			0.023	(0.452)	0.054	0.008	0.007	0.051		(0.307)	0.037			
	TV-HRGARCH$^*_{MRQ}$	1.229	-	0.312	0.637	0.721	-0.220	0.981	-0.020	0.088	0.477	-	-0.320	0.172	7469.650	5887.212	-4.187
		0.288		0.040	0.183	0.025	(0.348)	0.041	0.008	0.007	0.049		(0.283)	0.034			
RWE	RGARCH	-0.389	0.315	-	-	0.645	0.680	1.062	-0.040	0.081	-	0.164	-	-	7991.294	6554.930	-4.674
		0.138	0.033			0.029	(0.460)	0.053	0.008	0.006		0.007					
	HRGARCH	-0.358	0.292	-	-	0.671	0.575	1.056	-0.040	0.083	-	-	0.306	0.263	7989.920	6599.658	-4.704
		0.155	0.030			0.024	(0.560)	0.064	0.007	0.005			(0.400)	0.048			
	TV-HRGARCH	0.809	-	0.380	0.527	0.639	0.600	1.058	-0.040	0.082	-	-	-0.085	0.216	7992.073	6613.561	-4.711
		0.355		0.042	(0.205)	0.028	0.482	0.055	0.007	0.005			(0.345)	0.041			
	RGARCH*	-0.398	0.317	-	-	0.641	0.705	1.069	-0.040	0.080	0.316	0.161	-	-	7991.651	6578.197	-4.688
		0.125	0.032			0.029	(0.410)	0.047	0.008	0.006	0.057	0.007					
	HRGARCH*	-0.367	0.294	-	-	0.667	0.605	1.062	-0.040	0.081	0.276	-	0.178	0.248	7990.251	6617.473	-4.714
		0.128	0.028			0.025	(0.449)	0.051	0.007	0.005	0.053		(0.381)	0.045			
	TV-HRGARCH*	0.412	-	0.352	0.350	0.647	0.615	1.062	-0.040	0.081	0.219	-	-0.081	0.217	7991.414	6624.759	-4.716
		(0.317)		0.037	0.178	0.027	(0.416)	0.047	0.007	0.005	0.054		(0.386)	0.046			
	HRGARCH$^*_{MRQ}$	-0.369	0.295	-	-	0.666	0.613	1.063	-0.040	0.081	0.301	-	0.122	0.236	7990.312	6612.176	-4.710
		0.115	(0.027)			0.025	(0.401)	0.046	0.007	0.005	0.053		(0.402)	0.047			
	TV-HRGARCH$^*_{MRQ}$	0.269	-	0.337	0.287	0.655	0.641	1.066	-0.041	0.080	0.337	-	-0.107	0.209	7990.403	6616.025	-4.710
		(0.360)		0.037	(0.190)	0.025	(0.533)	0.061	0.007	0.005	0.056		(0.432)	0.051			

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012. $\ell(\mathbf{r})$: partial log-likelihood. $\mathscr{L}(\mathbf{r},\mathbf{x})$: log-likelihood. BIC: Bayesian Information Criterion. Standard errors are reported in small font under the parameters value: in parenthesis parameter not significant at 5%.

Table 9: Predictive log-likelihood using 5-min *RK* as volatility proxy (top): in **bold** the preferred model according to predictive log-likelihood. MCS p-values of predictive log-likelihood (bottom): in box models ∈ 75% MCS. The p-values refer to the negative predictive log-likelihoods.

	ALV	BMW	MEO	RWE
RGARCH	2614.147	2476.051	2492.735	2933.013
HRGARCH	2631.776	2497.190	2536.516	2968.781
TV-HRGARCH	2650.184	2511.197	2552.354	2975.101
RGARCH*	2631.724	2487.072	2521.690	2955.767
HRGARCH*	2649.767	2505.092	2557.727	2984.190
TV-HRGARCH*	2657.061	**2514.476**	2565.812	**2986.220**
HRGARCH$^*_{MRQ}$	2649.392	2501.026	2552.834	2980.199
TV-HRGARCH$^*_{MRQ}$	**2657.407**	2505.357	**2568.678**	2979.078
MCS p-values				
RGARCH	0.0010	0.0032	0.0002	0.0012
HRGARCH	0.0052	0.0292	0.0004	0.0308
TV-HRGARCH	0.2184	0.1406	0.0442	0.1372
RGARCH*	0.1016	0.0292	0.0032	0.0308
HRGARCH*	0.1102	0.1406	0.1712	0.6244
TV-HRGARCH*	0.9554	1.0000	0.6822	1.0000
HRGARCH$^*_{MRQ}$	0.1016	0.0810	0.0276	0.2848
TV-HRGARCH$^*_{MRQ}$	1.0000	0.1406	1.0000	0.2298

Table 10: Average values of QLIKE loss using 5-min *RK* as volatility proxy (top) and MCS p-values (bottom). For each stock in **bold** minimum loss and in box models $\in$ 75% MCS.

QLIKE				
	ALV	BMW	MEO	RWE
RGARCH	-6.9809	-6.6857	-6.9577	-7.2790
HRGARCH	-6.9807	-6.6847	-6.9582	-7.2778
TV-HRGARCH	-6.9839	**-6.6880**	**-6.9645**	**-7.2816**
RGARCH*	-6.9814	-6.6866	-6.9596	-7.2797
HRGARCH*	-6.9813	-6.6855	-6.9598	-7.2784
TV-HRGARCH*	-6.9829	-6.6878	-6.9635	-7.2808
HRGARCH^*_{MRQ}	-6.9814	-6.6858	-6.9599	-7.2786
$\text{TV-HRGARCH}^*_{MRQ}$	**-6.9841**	-6.6870	-6.9644	-7.2790

MCS p-values				
RGARCH	0.0776	0.1458	0.0262	0.0866
HRGARCH	0.0776	0.0702	0.0596	0.0282
TV-HRGARCH	$\boxed{0.8484}$	$\boxed{1.0000}$	$\boxed{1.0000}$	$\boxed{1.0000}$
RGARCH*	0.0776	$\boxed{0.3940}$	0.0870	0.0952
HRGARCH*	0.0776	0.1314	0.1118	0.0386
TV-HRGARCH*	$\boxed{0.3062}$	$\boxed{0.5222}$	$\boxed{0.4572}$	0.1192
HRGARCH^*_{MRQ}	0.0776	0.2162	0.1564	0.0444
$\text{TV-HRGARCH}^*_{MRQ}$	$\boxed{1.0000}$	$\boxed{0.4242}$	$\boxed{0.9192}$	0.0586

References

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), 701–720.

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., D. Dobrev, and E. Schaumburg (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics 169*(1), 75–93.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal 12*(3).

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

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

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

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

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

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

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

Hansen, P. and A. Lunde (2011). Forecasting volatility using high frequency data. *The Oxford Handbook of Economic Forecasting, Oxford: Blackwell*, 525–556.

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

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., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications 119*(7), 2249–2276.

Liu, L. Y., A. J. Patton, and K. Sheppard (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics 187*(1), 293–311.

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

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), 372–375.

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

Shephard, N. and D. Xiu (2016). Econometric analysis of multivariate realised qml: efficient positive semi-definite estimators of the covariation of equity prices.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.