Cardiff Economics Working Papers

Working Paper No. E2022/5

Exponential High-Frequency-Based-Volatility (EHEAVY) Models

Yogndeng Xu

March 2022

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Exponential High-Frequency-Based-Volatility (EHEAVY) Models

Yongdeng Xu (Cardiff University)*

February 2022

Abstract

This paper proposes an Exponential HEAVY (EHEAVY) model. The model specifies the dynamics of returns and realized measures of volatility in an exponential form, which guarantees the positivity of volatility without restrictions on parameters and naturally allows the asymmetric effects. It provides a more flexible modelling of the volatility than the HEAVY models. A joint quasi-maximum likelihood estimation and closed form multi-step ahead forecasting is derived. The model is applied to 31 assets extracted from the Oxford-Man Institute's realized library. The empirical results show that the dynamic of return volatility is driven by the realized measure, while the asymmetric effect is captured by the return shock (not by the realized return shock). Hence, both return and realized measure are included in the return volatility equation. Out-of-sample forecast and portfolio exercise further shows the superior forecasting performance of the EHEAVY model, in both statistical and economic sense.

Keywords: HEAVY model, High-frequency data, Asymmetric effects, Realized variance, Portfolio

JEL Classification: C32, C53, G11, G17

*Corresponding author. Cardiff Business School, Cardiff University, CF103EU, UK; Email: xuy16@cardiff.ac.uk

1 Introduction

Modelling and forecasting the return volatility has many implications in asset pricing, portfolio selection and risk management practices. Several studies have introduced non-parametric estimators of realized volatility using intra-day data (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002; Barndorff-Nielsen et al., 2008, 2009) and voluminous empirical evidence on modelling and forecasting the realized measure of volatility is developed, for example, the ARFIMA model in the original or logarithmic form (see Andersen et al., 2003; Chiriac and Voev, 2011; Koopman et al., 2005; Asai et al., 2012; Allen et al., 2014) and the Heterogeneous Autoregressive (HAR-RV) model by Corsi (2009).

Recently, the intra-daily estimators of volatility - known as realized measures — have been used to improve the volatility of return models. One of the popular models is the so-called "High-rEquency-bAsed VolatilitY"(HEAVY) model, initially proposed by Shephard and Sheppard (2010). The two-equation system, the HEAVY-r and the HEAVY-RM, which jointly estimates conditional variances of return and the realized measures of volatility based on daily and intra-daily data. The HEAVY model adopts to information arrival more rapidly than the classic daily GARCH process and hence it provides more reliable forecasts. Various extensions of HEAVE models have been developed. Hansen et al. (2012) introduced the Realized GARCH model that corresponds most closely to the HEAVY framework. The realized GARCH model is based on measurement equations that tie the realized measure to the latent conditional variance of return. An exponential type of realized GARCH model is developed by Hansen and Huang (2016). Cipollini et al. (2013) refer to the HEAVY model by simply restricting the bivariate vector multiplicative error representation for squared returns and realized variance. Borovkova and Mahakena (2015) apply the HEAVY models with different error distributions (student-t and skewed-t). They also extend the HEAVY-r equation with a news sentiment proxy and a time to maturity variable alternatively. Keranasos and Yfanti (2020) enrich the HEAVY with the long memory features and asymmetric effects. Yfanti et al. (2020) add a range based Garman–Klass volatility into the HEAVY framework. The multivariate specification of HEAVY model is also developed by Noureldin et al. (2012) and extended by Opschoor et al. (2017), Creal et al. (2013), Sheppard and Xu (2019) and Bauwen and Xu (2021).

The HEAVY models proposed so far are linear models. To guarantee the positivity of volatility, the parameters in HEAVY-r and HEAVY-RM equations are constrained to be positive, which is too restrictive. In addition, the asymmetric effect, in which the variances respond asymmetrically to positive and negative shocks, is not well addressed. In the HEAVY models, the past shock can either be represented by return shock, or by realized return shock. It is unclear which one should be used for the asymmetric effects. In the HEAVY model of Shephard and Sheppard (2010), the dynamics of volatility is only driven by lagged realized measure. It is straightforward to use lagged realized measures to capture the asymmetric effect. However, in the Realized GARCH model of Hansen et al. (2012, 2016),the asymmetric effects in both the return and realized measure equation are captured by the return shocks. In Keranasos and Yfanti (2020)'s "double asymmetric effects" HEAVY model, both return shocks and realized measures are used in modelling the asymmetric effects.

In this paper, we extend the HEAVY model to an EHEAVY model, in which both the return and realized measure equation takes an exponential form. The EHEAVY maintains the advantages of the EGARCH model. The conditional variance of return and realized measure is guaranteed to be positive without restrictions on the parameter set. In addition, both types of asymmetric effects are included as additional exogenous variables. Within the EHEAVY framework, we empirically study the asymmetric effects, without any positivity restrictions on the parameters. In our full EHEAVY model, we have both types of asymmetric. We estimate the EHEAVY model of 31 stocks from the Oxford man institution of realized lib. Indeed, we find that both two types of asymmetric effect are significant if only one is included. However, if both are included, only the standardized return is significant, while the standardized realized return is insignificant. The log-likelihood and BIC comparison confirm the finding. This indicates that the asymmetric effects in the conditional variance of return equation are captured by the return shock, not the realized return shock. Interestingly, this finding also holds for the realized measure of volatility. That is, the asymmetric effect in realized measure equation is also captured by return shock, not the realized measure.[1]

We then conduct an out-of-sample forecasts exercise and compare the EHEAVY model with HEAVY, asymmetric HEAVY (AHEAVY) of Shephard and Sheppard (2010) and Realized EGARCH model of Hansen and Huang (2016) at the daily, weekly and monthly horizon. The

[1]For a robustness check, we also estimate an extended Asymmetric HEAVY (AHEAVY) model of Shepard and Shepperd (2010). Similar results hold. That is, the asymmetric effect in the return and realized measure equation is captured by return shock, not the realized measure. The results are available upon request.

results suggest that EHEAVY model outperforms the benchmark HEAVY, AHEAVY model when forecasting the conditional variance of return and realized measure. It has a similar performance with Realized EGARCH model when forecasting the conditional variance of return. But it performs better than the realized EGARCH model when forecasting the realized measure. The super forecasting performance is further illustrated in a portfolio exercise, where the EHEAVY strategy results in a portfolio with higher Certain Equivalent Return and Sharpe's Ratio. The empirical evidence indicates the gain of using the EHEAVY model, in both statistical and economic sense.

It is notable that our EHEAVY model is closely related to the realized EGARCH model of Hansen and Huang (2016), but it differs with Hansen and Huang (2016) with several aspects. First, we adopt an EGARCH specification and the absolute standardized return is excluded in the return equation, which implies that the informative absolute return about future volatility is small. In the EGARCH model of Hansen and Huang (2016), the absolute standardized return [2] is included. Secondly, Hansen and Huang use a measurement equation, where the log of realized volatility is a function of conditional volatility of return in the same period. In our model, we adopt a HEAVY-RM type of structure, where realized volatility is a lagged function of return. The empirical results suggest that our model produces better out-of-sample forecasts of realized volatility than the realized EGARCH model. Third, we derive a joint Quasi maximum likelihood estimation approach for the EHEAVY model and a closed form multi-step ahead forecasts procedure.

The remainder of the paper is organized as follows. Section 2 introduces the EHEAVY models. Section 3 proposes Quasi maximum likelihood estimation and the multi-step ahead forecasts procedure. Section 4 is the empirical application. Section 5 concludes. A supplementary appendix (SA) includes additional empirical results.

2 The Exponential HEAVY Models

The benchmark HEAVY specification of Shephard and Sheppard (2010) use two variables: daily financial returns (r_t) and a corresponding sequence of daily realized measures of volatility, RM_t. Realised measures are theoretically high-frequency, nonparametric-based estimators of

[2] Hansen and Huang use a squared standardized return. Actually, the effect between the absolute and squared standardized return is rather close. We adopt the absolute standardized return in line with the EGARCH model.

the variation of open-to-close returns. We form the signed square rooted realized measures as follows: $\widetilde{RM_t} = sign(r_t)\sqrt{RM_t}$, where the $sign(r_t) = 1$, if $r_t \geqslant 0$ and $sign(r_t) = -1$, if $r_t < 0$. $\widetilde{RM_t}$ is also known as realized return. Then, the return and realized measure are characterized by the following relation:

$$\begin{array}{rcl} r_t^2 = h_t \varepsilon_{rt} & \text{or} & r_t = \sqrt{h_t} e_{rt} \\ RM_t = m_t \varepsilon_{Rt} & & \widetilde{RM_t} = \sqrt{m_t} e_{Rt} \end{array} \tag{1}$$

The first representation is multiplicative error specification, where the stochastic term ε_{it} ($i = r, R$) is independent and identically distributed, which is positively defined and has a unit mean. This implies that $\mathbb{E}(r_t^2 | \mathcal{F}_{t-1}) = h_t$. The second representation is a GARCH type model, where e_{it} is independent and identically distributed, which has zero mean and unit variance. This implies that $\mathbb{V}ar(r_t | \mathcal{F}_{t-1}) = h_t$. In other words, the GARCH model for the conditional variance of the returns (or the realized returns), is similar to the multiplicative error model (MEM)[3] for the conditional mean of the squared returns (or the realized measures).

We firstly present the full EHEAVY model, which consists of the following two equations:

$$\begin{aligned} \log h_t &= \omega_r + \beta_r \log h_{t-1} + \alpha_{rr}|e_{rt-1}| + \gamma_{rr} e_{rt-1} \\ &\quad + \alpha_{rR}|e_{Rt-1}| + \gamma_{rR} e_{Rt-1}, \\ \log m_t &= \omega_R + \beta_R \log m_{t-1} + \alpha_{RR}|e_{Rt-1}| + \gamma_{RR} e_{Rt-1} \\ &\quad + \alpha_{Rr}|e_{rt-1}| + \gamma_{rR} e_{rt-1}. \end{aligned} \tag{2}$$

where $corr(e_{rt}, e_{Rt}) = \rho$.

The first equation is the EHEAVY-r equation and the second equation is EHEAVY-RM equation. In the EHEAVY-r equation, the parameter β_r summarizes the persistence of volatility, whereas α_{rR} represents how informative the realized measures are about future volatility of return. The asymmetric effect is represented by $\gamma_{rr} e_{rt-1}$ and $\gamma_{rR} e_{Rt-1}$. In the EHEAVY-RM equation, the parameter β_R summarizes the persistence of realized measure volatility and the asymmetric effect is represented by $\gamma_{Rr} e_{rt-1}$ and $\gamma_{RR} e_{Rt-1}$. EHEAVY model is stationary if $\beta_r < 1$ and $\beta_R < 1$. One advantage of the EHEAVY versus HEAVY model is that the positivity

[3] Engle (2002) first proposed the MEM model using the various GARCH family specifications to estimate the volatility, which is a non-negative process.

of variance is guaranteed without any restrictions on the parameter set.

In the empirical applications, we find that α_{rr} and γ_{rR} are often insignificant in EHEAVY-r equation, and α_{Rr} and γ_{RR} are insignificant in EHEAVY-RM equation. Take the EURO50 index close-to-close return for example, the estimated full EHEAVY model is given by

$$\begin{aligned}\log h_t &= \underset{(0.028)}{-0.193} + \underset{(0.003)}{0.970}\log h_{t-1} + \underset{(0.018)}{0.001}|e_{rt-1}| - \underset{(0.017)}{0.151}e_{rt-1} \\ &\quad + \underset{(0.042)}{0.305}|e_{Rt-1}| - \underset{(0.018)}{0.005}e_{Rt-1} \\ \log m_t &= \underset{(0.002)}{-0.225} + \underset{(0.004)}{0.967}\log m_{t-1} + \underset{(0.020)}{0.002}|e_{rt-1}| - \underset{(0.017)}{0.146}e_{rt-1} \\ &\quad + \underset{(0.047)}{0.339}|e_{Rt-1}| - \underset{(0.019)}{0.012}e_{Rt-1}\end{aligned}$$

with $\rho = \underset{(0.003)}{0.841}$, The numbers in parentheses are the robust standard errors for each of the point estimates. More details of estimation are shown in the empirical application Section 4. Excluding the insignificant terms, we have the chosen EHEAVY model.

The chosen EHEAVY model is

$$\begin{aligned}\log h_t &= \omega_r + \beta_r \log h_{t-1} + \alpha_{rR}|e_{Rt-1}| + \gamma_{rr}e_{rt-1}, \\ \log m_t &= \omega_R + \beta_R \log m_{t-1} + \alpha_{RR}|e_{Rt-1}| + \gamma_{rR}e_{rt-1}\end{aligned} \tag{3}$$

Without a further illustration, the EHAVY model thereafter is the chosen model defined in (3). It is notable that the realized EGARCH model of Hansen and Huang (2016) also includes $\alpha_{rr}|e_{rt-1}|$ in the HEAVY-r equation[4]. Their representation is more like an EGARCH-X model, where both $\alpha_{rr}|e_{rt-1}|$ and $\alpha_{rR}|e_{Rt-1}|$ are included. Consistent with the evidence in HEAVY, we found that the estimated α_{rr} is very small or insignificant, which implies that the informative the absolute (or squared) return about future volatility is small. In the EHEAVY model, only $\alpha_{rr}|e_{rt-1}|$ is excluded. The EHEAVY-r equation has the same number of parameters as the EGARCH model.

The EHEAVY-RM equation is closer to the HEAVY-RM of Shepard and Shepperd (2010) with the exponential representation. Shepard and Shepperd (2010) suggested an asymmetric HEAVY model, where the asymmetric effect is captured by the binary lagged realized measure in the HEAVY-r and HEAVY-RM equation. Our empirical evidence shows that the asymmetric

[4] Hansen and Huang (2016) use a quadratic form $\alpha_{rr}(e_{rt-1})^2$

effects are mostly captured by the return shock, not the realized measure. So, the EHEAVY model includes $\gamma_{rr} e_{rt-1}$ and $\gamma_{rR} e_{rt-1}$ terms to capture the asymmetric effects. This is the main feature of the EHEAVY model. The asymmetric effects in both return and realized measure equation is modelled by the return shock, not by the realized return shock.

The multiple-step ahead forecasts from the EHEAVY model only relies on the HEAVY-r equation. The information from realized measures is not required. Hence, even if the dynamics of realized measure is misspecified, the effect to the conditional variances of the return is small. Shepard and Shepperd (2010) find that the HEAVY model performs not as good as the GARCH model for multi-step ahead (e.g., 22 step ahead) forecasts, which is also confirmed by Bauwens and Xu (2021). The EHEAVY model is expected to have better multi-step ahead forecasts than the HEAVY model.

To better understand the dynamics, we express the EHEAVY models in a vector form. Define $x_t = [r_t^2, RM_t]'$, $\widetilde{x}_t = [r_t, \widetilde{RM_t}]'$, $\mu_t = [h_t, m_t]'$ and $e_t = [e_{rt}, e_{Rt}]'$, the vector multiplicative representation of EHEAVY model is

$$\widetilde{x}_t = \sqrt{\mu_t} \odot e_t, \quad e_t|\mathcal{F}_{t-1} \sim D(0, P) \tag{4}$$

$$\log \mu_t = \omega + B \log \mu_{t-1} + A e_{t-1} + \Gamma |e_{t-1}|$$

where e_t are a sequence of independent and identically distributed variables with mean 0 and time-invariant positive definite covariance matrix P with ones on the main diagonal so that $E(x_t|\mathcal{F}_{t-1}) = \mu_t$, and

$$\omega = \begin{bmatrix} \omega_r \\ \omega_R \end{bmatrix}, \; A = \begin{bmatrix} 0 & \alpha_{rR} \\ 0 & \alpha_{RR} \end{bmatrix}, \; \Gamma = \begin{bmatrix} \gamma_{rr} & 0 \\ \gamma_{Rr} & 0 \end{bmatrix}, \; B = \begin{bmatrix} \beta_r & \\ & \beta_R \end{bmatrix}. \tag{5}$$

It is notable that if

$$A = \begin{bmatrix} \alpha_{rr} & \alpha_{rR} \\ \alpha_{Rr} & \alpha_{RR} \end{bmatrix}, \; \Gamma = \begin{bmatrix} \gamma_{rr} & \gamma_{rR} \\ \gamma_{Rr} & \gamma_{RR} \end{bmatrix}, \tag{6}$$

(4) becomes the full EHEAVY model defines in (2). If

$$A = \begin{bmatrix} \alpha_{rr} & 0 \\ 0 & \alpha_{RR} \end{bmatrix}, \; \Gamma = \begin{bmatrix} \gamma_{rr} & 0 \\ 0 & \gamma_{RR} \end{bmatrix}, \tag{7}$$

the top part of (4) becomes the EGARCH model.

3 Estimation and Forecasting

This section presents a quasi-maximum likelihood estimation approach and a multi-step ahead forecasting procedure.

3.1 Quasi-maximum Likelihood Estimation (QMLE)

The parameters in return and realized measure equation are not variation free in the EHEAVY models, hence a joint estimation method is required. Below we derive a Quasi-maximum likelihood estimation approach.

The vector representation of the EHEAVY model is

$$\begin{aligned} \widetilde{x}_t &= \sqrt{\mu_t} \odot e_t, \quad e_t|\mathcal{F}_{t-1} \sim D(0, P) \\ \log \mu_t &= \omega + B \log \mu_{t-1} + A e_{t-1} + \Gamma |e_{t-1}|. \end{aligned} \tag{8}$$

More general, let $\widetilde{x}_t$ as $k-$dimension process and let $\theta' = [\theta'_1, \theta'_2]$, where $\theta'_1 = vech(P)$, operator *vech* stacks the lower triangular elements of an symmetric $(k \times k)$ matrix into a $k \times (k+1)/2$ vector and θ'_2 contains the parameters in μ_t. Assuming e_t follows a multivariate normal distribution $e_t|\mathcal{F}_{t-1} \sim N(0, P)$, the likelihood function is equivalent to the one in the Constant Conditional Correlation (CCC)-GARCH model (Bollerslev,1990; Jeantheau,1998). The log-likelihood function for the observation t is given by

$$\begin{aligned} l_t(\theta) &= -\frac{k}{2}\log(2\pi) - \frac{1}{2}\log|M_t P M_t| - \frac{1}{2}\widetilde{x}_t{}' \left(M_t P M_t\right)^{-1} \widetilde{x}_t \\ &= -\frac{k}{2}\log(2\pi) - \log|M_t| - \frac{1}{2}\log|P| - \frac{1}{2}\widetilde{x}_t{}' M_t^{-1} P^{-1} M_t^{-1} \widetilde{x}_t \end{aligned} \tag{9}$$

where $M_t = diag(\sqrt{\mu_{1t}}, \sqrt{\mu_{2t}}, ..., \sqrt{\mu_{kt}})$.

Let $l(\theta) = \sum_{t=1}^{T} l_t(\theta)$, the QMLE for $\hat{\theta}$ equals

$$\hat{\theta} = \arg\max_{\theta} l(\theta)$$

Explicit expressions for the score vector and the Hessian matrix of the log-likelihood function can be derived following the CCC-GARCH literature; see Nakatani and Teräsvirta (2009) lemma 3.1 and 3.2 for example.

The detailed asymptotic distribution theory is not yet available. The asymptotic analysis of the EHEAVY model is similarly complicated as EGARCH model [5], so it is beyond the scope of this article to fully establish the asymptotic theory for the estimators. We leave this for future research.

3.2 Multi-Step Ahead Forecasts

The HEAVY and EHEAVY model can be used to predict both the conditional variance of return and the realized measure of volatility. The latter has been the subject of very active literature (see, for example, Andersen et al., 2001, 2003 ; Corsi, 2009 ; Bollerslev et al., 2016; Taylor, 2017).

Suppose the forecaster models x_t and obtains s-step-ahead forecasts given by the conditional mean of x_t; that is $E(x_{t+s}|\mathcal{F}_t)$, where $\mathcal{F}_t$ is the forecaster's information set. Let $\mu_{t+s|t} = E(x_{t+s}|\mathcal{F}_t)$. Now let's move steps ahead, x_{t+s}, $s > 0$ is not known and needs to be substituted with its corresponding conditional expectation μ_{t+s}. The multi-step ahead forecasts of the EHEAVY model are not straightforward, as the conditional expectation of log function is not equal to the log function of the conditional expectation. To do so, we denote $\phi_t = \log(\mu_t)$ and

$$\begin{aligned}
\phi_{t+1|t} &= \omega + B\phi_t + \Gamma e_t + A|e_t|, \\
\phi_{t+2|t} &= \omega + A\bar{e} + B\phi_{t+1|t} \\
&= \bar{\omega} + B\phi_{t+1|t}, \qquad (10)
\end{aligned}$$

where $\bar{e} = E(|e_t|)$ and $\bar{\omega} = \omega + A\bar{e}$. If e_t is symmetric normally distributed, $E(|e_t|) = \sqrt{2/\pi}$. More wisely, $E(|e_t|)$ can be estimated by the unconditional mean of $|e_t|$

And then, for $s > 2$,

$$\phi_{t+s|t} = \bar{\omega} + B\phi_{t+s-1|t}, \qquad (11)$$

[5] The consistency and asymptotic properties $\widehat{\boldsymbol{\theta}}$ for the multivariate EGARCH or EHEAVY model are not available under general conditions (see for example, Nakatani and Teräsvirta, 2009 and Francq et al., 2013). A limitation in the development of the asymptotic properties for the (multivariate) EGARCH is the lack of an invertibility condition (See Wintenberger, 2013 and Martinet and McAleer, 2018 for a discussion). More recently, Xu and Keranosas (2021) shows that the QML estimates of multivariate EGARCH model are unbiased and normal distributed, when the sample size is relatively large (i.e., sample size ≥ 2500) by Monte Carlo simulation.

which can be solved recursively for any horizon s. A closed form forecasts for $\phi_{t+s|t}$ can also be derived as:

$$\phi_{t+s|t} = \tilde{\omega} + B^{s-1}\phi_{t+1|t} \tag{12}$$

where $\tilde{\omega} = \frac{(1-B^{s-1})\bar{\omega}}{1-B}$.

We then derive a formula for $\mu_{t+s|t} = E(x_{t+s}|\mathcal{F}_t)$. With the log specification one would have to account for distributional aspects of $\log(\mu_{t+s|t})$ in order to produce an unbiased forecast of $\mu_{t+s|t}$. Using the second-order approximation [6]

$$\mu_{t+s|t} \approx \exp(\phi_{t+s|t})(1 + \frac{\sigma^2_{\phi,t+s|t}}{2}) \tag{13}$$

where $\sigma^2_{\phi,t+s|t}$ is is the s-step-ahead conditional second moment $\phi_{t+s|t}$. The conditional second moments are estimated using their unconditional sample counterparts.

The EHEAVY model s-step ahead forecasts $\mu_{t+s|t}$ are derived by setting A, B, Γ to the matrices defines in (5). Then, the s-step ahead forecast of the conditional variance of return ($h_{t+s|t}$) corresponds to the first element of $\mu_{t+s|t}$ and the s-step ahead forecast of the realized measure of volatility ($m_{t+s|t}$) corresponds to the second element of $\mu_{t+s|t}$.

4 Estimation Application

4.1 Data

We use daily data for 31 assets extracted from the Oxford-Man Institute's (OMI) realized library. Our sample covers the period from 03/01/2000 to 31/5/2021. The OMI's realized library includes daily stock market returns and several realized volatility measures calculated on high-frequency data from the Reuters DataScope Tick History database. The data are first cleaned and then used in the realized measures calculations. According to the library's documentation, the data cleaning consists of deleting records outside the time interval that the stock exchange is open. Some minor manual changes are also needed, when results are ineligible due to the rebasing of indices. The library's realized measures are calculated in the

[6] The full approximation is given by $\exp\left(\phi_{t+s|t}\right)\left(1 + \sum_{k=1}^{\infty} \frac{1}{k!}\phi_{k,t+s|t}\right)$ where $\phi_{k,t+s|t}$ is is the s-step-ahead kth conditional moment about the conditional mean. See Taylor (2017) for the details of a full approximation.

way described in Shephard and Sheppard (2010). We adopt the realized kernel as the realized measure, using Parzen kernal function. This estimator is similar to the well-known realized variance, but is robust to market microstructure noise and is a more accurate estimator of the quadratic variation. The realized kernel is calculated as follows: $RK_t = \sum_{k=-H}^{H} k(h/(H+1))\gamma_h$, where $k(x)$ is the Parzen kernel function with $\gamma_h = \sum_{j=|h|+1}^{n} x_{jt}x_{j-|h|,t}$; $x_{jt} = X_{t_{j,t}} - X_{t_{j-1,t}}$ are the 5-minute intra-daily returns where $X_{t_{j,t}}$ are the intra-daily prices and $t_{j,t}$ are the times of trades on the t-th day. Shephard and Sheppard (2010) declare that they select the bandwidth of H as in Barndorff-Nielsen et al. (2009).

The realized measure is directly related to the volatility of open-to-close returns, but only captures a fraction of the volatility of close-to-close returns. In the estimation, we'll use both open-to-close returns and close-to-close returns.

Table 1 presents the 31 assets extracted from the database and provides volatility estimations for each one's squared returns and realized kernels time series for the respective sample period. We calculate the mean and standard deviation (StDev) of the annualized volatility. Annualized volatility is the square root of 252 times the squared return or the realized kernel. The mean figures show that the assets have the annualized volatilities of the realized measure between 9% and 30%, with the corresponding results for the squared close-to-close returns between 14% and 40%. On average, the realized measure is about 63% of squared return. The realized kernel missed out on the overnight return, which accounts for their lower level. On the other hand side, the annualized volatility of open-to-close return is similar to the annualized volatility of realized measure. It is typically a little higher than the realized measure, but the difference is very small. The StDev figures show much higher standard deviations for the squared return than the realized measure. The standard deviations of squared close-to-close return are usually twice as higher as the standard deviations of realized measure. The squared open-to-close returns also have much higher standard deviation than the realized measure. It turns out that the realized measure is a more stable measurement of volatility than the squared returns.

4.2 Estimation results

We estimate the EHEAVY model using both the open-to-close returns and close-to-close returns. First, we compare the performance of the EHEAVY model with the EGARCH and the full EHEAVY model. We also estimate an alternative EHEAVY (AEHEAVY) model, where the asymmetric is only from the realized return shock (see (21) in Appendix B for the model specification). The estimated parameters are summarized in table 2. Results based on open-to-close returns are presented in the left panel and the analogous results for close-to-close returns are presented in the right panel. The detailed estimates for each of the assets are presented in Appendix A table A2 and A3.

Based on open-to-close returns estimation, the empirical results are summarized as follows.

- The EGARCH estimates are as expected. The persistence parameter β is higher and close to one. The leverage parameter γ is negative and significant.

- In the full EHEAVY model, the estimated α_{rr} is significant 13 out of 31 cases. The size is small, with a median value of 0.02 ranging from -0.032 to 0.142. The estimated α_{rR} is significant in all 31 cases, with the median value of 0.366 ranging from 0.211 to 0.602. The size of α_{rR} is much large than that of α_{rr}, which is consistent with the findings in the HEAVY literature. It shows that the future volatility of return is mainly driven by the information from the realized measure.

- The estimated γ_{rr} is relatively large and significant in 28 out of 31 cases. The estimated γ_{rR} is significant only in 4 out of 31 cases with a much smaller size. It indicates that the leverage effects in the return variance equation are mainly driven by the return shock, not by the realized return shock.

- The estimated α_{Rr} is small and only significant in 15 out of 31 cases. The estimated α_{RR} is large and significant in all 31 cases, which is consistent with HEAVY literature. The estimated γ_{rr} is relatively large and significant in 27 out of 31 cases. The estimated γ_{RR} is small and significant only in 5 out of 31 cases. It shows that the leverage effect in the realized measure equation is also driven by the return shock, not by the realized measure.

- In the EHEAVY model, all coefficients are significant in almost all 31 cases. In particular, the coefficients for asymmetric effect γ_{rr} and γ_{Rr} are negative and significant, showing that the asymmetric effect is a common stylized fact in volatility modelling.

- In the alternative EHEAVY model, we use the realized measure only to capture asymmetric effects. Interestingly, γ_{Rr} and γ_{RR} are also negative and significant. The size of

γ_{Rr} and γ_{RR} are marginally smaller than γ_{rr} and γ_{Rr} in the EHEAVY model. If we only estimated the alternative EHEAVY model, we may mis-conclude that the asymmetric effect can be modelled by the realized return shock.

- The estimates of ρ is about 0.8 and very similar across the assets, showing a high correlation between return and realized return. This is evidence of joint estimation of return and realized measure, as proposed in section 3.

The analogous results for close-to-close returns are similar. To summarize the main findings, negative news increase volatility more than positive news. However, the size of the increments is measured by previous periods' return shocks, not by the previous period's realized return shock. This is true for both conditional variance of return and realized volatility modelling.[7]

Further insight can be gained from the in-sample partial log-likelihood value and Bayesian information criteria (BIC). To have a measure of fit that can be compared with conventional EGARCH and HEAVY model, we follow Hansen et al., (2012) and define the partial log-likelihood function for the time series of returns

$$\ell_P(r;\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log(2\pi) + \log(h_t) + (r_t^2/h_t)\right] \tag{14}$$

This quantity is the Kullback–Leibler measure associated with the conditional distribution of returns.

We define the partial log-likelihood function for the time series of realized measure

$$\ell_P(RM;\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log(2\pi) + \log(m_t) + (rm_t/m_t)\right] \tag{15}$$

The comparison is reported in table 3. The full EHEAVY model has the highest log-likelihood values. However, The log-likelihood gain of full EHEAVY over EHEAVY is equal to 17 for the 4 additional parameters, hence it appears to be minor. On the contrary, the gains of full EHEAVY and EHEAVY over EGARCH are substantial (131 and 124 respectively). Further, the EHEAVY model has a much higher log-likelihood value than the alternative EHEAVY model, and the log-likelihood gain is 44 which shows that the EHEVY model has a better fits to

[7]It also raises an interesting question about realized volatility modelling and forecasting (i.e., ARFIMA or HAR model). When leverage effect is considered, whether the lagged absolute (or squared) return, or lagged realized measure should be included in the model. We leave this question for future research.

the data than the alternative EHEAVY model. In addition, there are 19 out of 31 cases where full EHEAVY has the highest log-likelihood values, and 11 out of 31 cases where EHEAVY has the highest log-likelihood values. When comparing the BIC criteria, EHEAVY has the smallest value among all three models. There are 21 out of 31 cases, where the EHEAVY model achieves the best BIC criteria. Notice that EGARCH, EHEAVY and alternative EHEAVY are not nested, but they have the same number of parameters, so choosing between them using their log-likelihood values is equivalent to a choice based on model choice criteria.

In brief, the EHEAVY model strongly dominates the conventional EGARCH models that rely exclusively on daily returns, and the alternative EHEAVY model that uses realized measure for asymmetric effect. These results are consistent with the findings in table 2, suggesting that the return volatility dynamic is mainly captured by the lagged realized measure and leverage effects is better modelled by the return shock.

4.3 News impact curve

Additional insight about the value of the EHEAVY structure is evident from the news impact curve. This curve was introduced by Engle and Ng (1993), and is used to illustrate the impact that return shocks has on volatility. News impact curve is the impact that e_{rt} has on h_{t+1} measured in percentages, as defined by $\mathrm{E}(\log h_{t+1} \mid e_{rt} = e_r) - \mathrm{E}(\log h_{t+1}))$. We plot the impact curve of the EHEAVY model and the EGARCH, Realized EGARCH model. As the return shocks are contemporaneously correlated with realized return shocks, one unit return shocks will also incur $\hat{\rho}$ unit realized return shock. The news impact curve for EHEAVY model is given by $\alpha_{rR}|e_{Rt-1}| + \gamma_{rr}e_{rt-1} = \rho\alpha_{rR}|e_{rt-1}| + \gamma_{rr}e_{rt-1}$. For the EGARCH and realized EGARCH model, the news impact curve is simply given by $\alpha_{rr}|e_{rt-1}| + \gamma_{rr}e_{rt-1}$.

Taking EURO50 close-to-close return for example, the news impact curve of EGARCH, realized EGARCH and EHEAVY model is plotted in Figure 1. As is evident from Figure 1, the generalized structure of the EHEAVY model has more a profound effect on the news impact curve than the EGARCH and realized EGARCH model. The news impact curve of realized EGARCH model is very close to that of the EGARCH model, and it does not show an increasing news impact curve when news is positive. The EHEAVY model allows good news and bad news to have a different impact on volatility. It allows big news to have a greater impact on volatility than the EGARCH and realized EGARCH model in both directions.

Figure 2 gives another example: SPX close-to-close return news impact curve. Again, the HEAVY model has the highest variation in both directions among the three completion models. It allows big news to have a much greater impact on volatility than the other models.

4.4 Out of sample forecasting comparison

Next, we conduct an out-of-sample forecasting comparison. In this application, we consider forecasting volatility of close-to-close return only, which is more appropriate for most applications in portfolio allocation or risk management.

The EHEAVY model is compared with the following three models: 1) benchmark HEAVY model; 2) the asymmetric HEAVY model of Shepard and Shepperd (2010); 3) the realized EGARCH of Hansen and Huang (2016).[8] The out-of-sample period comprises the last 1000 observations of the full-sample period for each asset. The four models are re-estimated every observation based on an rolling sample windows of sample size $T-1000$. As shown in Appendix table A1, the full sample size is around $T = 5000$ for most of the assets. That leaves the estimated sample around 4000 observations. We report $s = 1$, 5 and 22 for horizons of 1-day, 5-day and 22-days ahead out-of-sample forecasts.

We use the following two loss functions for the volatility of the close-to-close return

$$
\begin{aligned}
MSE(r_{t+s}^2, h_{t+s|t}) &= \sum_{t=T-1000+s}^{T} (r_{t+s}^2 - h_{t+s|t})^2 \\
QMLIK(r_{t+s}^2, h_{t+s|t}) &= \sum_{t=T-1000+s}^{T} \left(\frac{r_{t+s}^2}{h_{t+s|t}} - \log\left(\frac{r_{t+s}^2}{h_{t+s}}\right) - 1 \right).
\end{aligned}
$$

And the corresponding loss functions for the realized measure of volatility are

$$
\begin{aligned}
MSE(RM_{t+s}, m_{t+s|t}) &= \sum_{t=T-1000+s}^{T} (RM_{t+s} - m_{t+s|t})^2 \\
QMLIK(RM_{t+s}, m_{t+s|t}) &= \sum_{t=T-1000+s}^{T} \left(\frac{RM_{t+s}}{m_{t+s|t}} - \log\left(\frac{RM_{t+s}}{m_{t+s}}\right) - 1 \right).
\end{aligned}
$$

Table 4 shows the comparisons of the out-of-sample forecasts by reporting the ratio of the

[8]See appendix B for the specification of the two models.

losses for the different models relative to the losses of the benchmark HEAVY model. The average ratios across the 31 assets are reported. The upper panel is the forecasting of the conditional variance of return. It can be seen that the average losses between EHEAVY and realized EGARCH model are very close, and both are smaller than the average losses of the two HEAVY models at all forecasting horizons. This is not surprising, as the specification of the return equation between the EHEAVY and realized EGARCH are very close. The asymmetric HEAVY model is slightly better than the HEAVY model. The lower panel is the forecasting of the realized measure of volatility. The EHEAVY model has the smallest losses among the four competition models at the three forecasting horizons. Apparently, the EHEAVY model has a better forecasting of realized measures of volatility than the HEAVY and realized EGARCH model. The realized EGARCH model is the second best model.

In order to formally determine whether the quality of the forecasts differ significantly across the different models, we apply the Model Confidence Set (MCS) of Hansen, Lunde, and Nason (2011). This approach identifies the (sub)set of models that contain the best forecasting model with 90% confidence. For each of the two loss functions and three forecast horizon we determine the subset of models that comprise the MCS.

Table 5 summarizes the MCS of the out-of-sample forecasts by reporting the numbers of the asset in which each model is part of the 90% MCS. For the conditional variance of return, EHEAVY and realized EGARCH model belong to the 90% MCS of both MSE and QLIK loss at the forecasting horizons in almost all assets. There are only a few number of assets that the HEAVY model is in the 90% MCS of QML loss. The asymmetric HEAVY model has slightly larger numbers than the HEAVY model. For the realized measure of volatility, the EHEAVY model is always in the 90% MCS of MSE loss. HEAVY, asymmetric HEAVY and realized EGARCH model are mostly in the 90% MCS of MSE loss. However, when the QLIK loss function is used, only EHEAVY model is in the 90% MCS. Realized EGARCH model is in the 90% MCS in a few assets. HEAVY and asymmetric HEAVY model are not in the 90% MCS of QLIK in almost most all assets at the three horizons.

To summarize, the MCS test suggests that the EHEAVY model performs similarly well with the REGARC model when forecasting the conditional variance of return. However, the EHEAVY model outperforms the realized EGARCH and other HEAVY models when forecasting the realized measure of volatility throughout the forecast horizons.

4.5 Portfolio exercise

Compared with the statistical gains of volatility predictability, market investors care more about economic significance. Specifically, they are interested in how well these volatility forecasts do in asset allocation. To evaluate the economic value of volatility forecasts, we consider a mean–variance utility investor who allocates the assets between stock index and risk-free asset following the literature (see, e.g., g Campbell and Thompson, 2008; Ferreira and Santa-Clara, 2011; Neely et al., 2014; Rapach et al., 2010,Wang, et al., 2016). The utility from investing in this portfolio is:

$$U_t\left(r_t\right) = E_t\left(w_t r_t + r_{t.f}\right) - \frac{1}{2}\gamma \operatorname{Var}_t\left(w_t r_t + r_{t,f}\right) \tag{16}$$

where w_t is the weight of stock in this portfolio, r_t is the stock return in excess of risk-free rate, $r_{t,f}$ is the risk-free rate and c is the risk aversion coefficient. $E_t(.)$ and $Var_t(.)$ denote conditional mean and variance given information at time t. Maximizing $U_t(r_t)$ respect to w_t yield the ex-ante optimal weight of stock index at day $t+1$

$$w_t^* = \frac{1}{\gamma}\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}\right) \tag{17}$$

where $\hat{r}_{t+1}$ and $\hat{\sigma}_{t+1}^2$ are the mean and volatility forecasts of asset excess returns, respectively. We follow the literature by restricting the optimal weight between 0 and 1.5 (i.e., $0 < w_t^* <= 1.5$) to preclude short sales and preventing more than 50% leverage (Rapach et al., 2010; Neely et al., 2014).

In this way, the portfolio return at day $t+h$ is given by:

$$R_{t+h} = w_t^* r_{t+h} + r_{t+h,f} \tag{18}$$

We employ two popular criteria to evaluate the performance of a portfolio constructed based on return and volatility forecasts. The first is the Sharpe ratio:

$$\mathrm{SR_{t+h}} = \frac{\bar{\mu}_{p,t+h}}{\bar{\sigma}_{p,t+h}} \tag{19}$$

where $\bar{\mu}_{p,t+h}$ and $\bar{\sigma}_{p,t+h}$ are the mean and standard deviation of portfolio excess returns over the out-of-sample period, respectively. The second criterion for evaluating portfolio performance is the certainty equivalent return (CER):

$$\text{CER}_{p,t+h} = \hat{\mu}_{p,t+h} - \frac{\gamma}{2}\hat{\sigma}^2_{p,t+h} \tag{20}$$

where $\hat{\mu}_{p,t+h}$ and $\hat{\sigma}^2_{p,t+h}$ are the mean and variance of portfolio returns over the out-of-sample period, respectively.

We use value 6 for γ [9]. For the risk-free rate $r_{t+1,f}$, we use the 3-month Treasury bill rate. Actually, the daily risk-free rate is rather close to zero. For the mean forecasts, we use the popular historical average(HA) forecasts. HA is generally accepted as the benchmark model in forecasting stock return (see, e.g., Rapach et al., 2010; Neely et al., 2014). Goyal and Welch (2008) find that it is difficult to beat this benchmark in forecasting stock returns out-of-sample. In this way, the optimal weights of the stock index are only determined by the volatility forecasts because different strategies share the same mean forecasts of returns when γ is fixed.

Table 6 shows the Return, Sharpe ratio and CER of portfolios formed by the conditional variance of return forecasts by the four models. We report the average (Mean), minimum (Min), maximum (Max) of Return, Sharpe ratio and CER across the 31 assets. From the Mean, we find that the EHEAVY strategy results in the portfolio with the Return of about 8%, the Sharpe ratios of about 58% and the CER about 105% higher than the HEAVY strategy. The asymmetric HEAVY and realized GARCH portfolio has a Return, Sharpe Ratio and CER of that are close to the HEAVY portfolio. The Min and Max columns show similar results to the Mean column. Overall, the portfolio exercise shows the greater performances of EHEAVY in the economic sense. In indicates that the EHEAVY model can improve the economic value of volatility forecasts significantly.

5 Conclusions

This paper proposes an EHEAVY model. It extends the HEAVY model of Shephard and Sheppard (2010) in several ways. First, the EHEAVY model guarantees the positivity of vari-

[9] We also check the values of 3, and 9 for γ, the results in table 6 does not change significantly

ance without restrictions on the parameter set, hence it is more flexible. Second, the asymmetric effects, which is an important feature for volatility modelling, are included. Third, a joint quasi-maximum likelihood estimation and closed form multi-step ahead forecast procedure is derived. The empirical results show that the dynamic of return volatility is driven by the realized measure. Further, the asymmetric effects in both HEAVY-r and HEAVY-RM equation are from the return shock, not from the realized measure. The Out-of-sample forecasting comparison shows that the EHEAVY model forecasts the conditional volatility of return and realized measure volatility better than the competition models. This result is also confirmed by a portfolio exercise, implying the greater performances of EHEAVY in both statistical and economic sense.

The EHEAVY model has a simple structure, a straightforward estimation and inference procedure. Thus, empirical researchers and practitioners can readily use the model to forecast volatility. For future research, the EHEAVY model can be extended by adding other realized measures, additional exogenous variables, jump components in return or realized measure equations.

References

[1] Allen, D.E., McAleer, M., Scharth, M., 2014. Asymmetric Realized Volatility Risk. Journal of Risk and Financial Management 7, 80-109.

[2] Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885-905.

[3] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of exchange rate volatility. Journal of the American Statistical Association 96, 42-55.

[4] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and Forecasting Realized Volatility. Econometrica 71, 579-626.

[5] Asai, M., McAleer, M., Medeiros, M.C., 2012. Asymmetry and long memory in volatility modeling. Journal of Financial Econometrics 10, 495-512.

[6] Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47-78.

[7] Bai, J., Perron, P., 2003a. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1-22.

[8] Bai, J., Perron, P., 2003b. Critical values for multiple structural change tests. Econometrics Journal 6, 72-78.

[9] Baillie, R. T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3-30.

[10] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76, 1481-1536.

[11] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: trades and quotes. Econometrics Journal 12, C1-C32.

[12] Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society, Series B 64, 253-280.

[13] Bauwens, L. and Xu, Y., 2020. DCC-HEAVY: A multivariate GARCH model based on realized variances and correlations (No. E2019/5). Cardiff University, Cardiff Business School, Economics Section.

[14] Bollerslev, T., Patton, A. J., and Quaedvlieg, R. 2016. Exploiting the errors: A simple approach for improved volatility forecasting. Journal of econometrics, 192(1), 1-18.

[15] Borovkova, S., Mahakena, D., 2015. News, volatility and jumps: the case of natural gas futures. Quantitative Finance 15(7), 1217-1242.

[16] Chiriac, R., Voev, V., 2011. Modelling and forecasting multivariate realized volatility. Journal of Applied Econometrics 26, 922-947.

[17] Chortareas, G., Jiang, Y., Nankervis, J.C., 2011. Forecasting exchange rate volatility using high-frequency data: Is the euro different? International Journal of Forecasting 27, 1089-1107.

[18] Chou, R.Y., 2005. Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. Journal of Money, Credit and Banking 37, 561-582.

[19] Cipollini, F., Engle, R.F., Gallo, G.M., 2009. A Model for Multivariate Non-Negative Valued Processes in Financial Econometrics. Available at SSRN: https://ssrn.com/abstract=1333869 or http://dx.doi.org/10.2139/ssrn.1333869.

[20] Cipollini, F., Engle, R.F., Gallo, G.M., 2013. Semiparametric vector MEM. Journal of Applied Econometrics 28, 1067-1086.

[21] Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174-196.

[22] Corsi, F., Mittnik, S., Pigorsch, C., Pigorsch, U., 2008. The volatility of realized volatility. Econometric Reviews 27, 46-78.

[23] Davidson, J., 2004. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. Journal of Business and Economic Statistics 22, 16-29.

[24] Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83-106.

[25] Engle, R.F., 2002. New frontiers for ARCH models. Journal of Applied Econometrics 17, 425-446.

[26] Engle, R.F., Gallo, G.M., 2006. A multiple indicator model for volatility using intra-daily data. Journal of Econometrics 131, 3-27.

[27] Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: A new model for irregularly spaced transaction data. Econometrica 66, 1127-1162.

[28] Francq, C., Wintenberger, O., and Zakoïan, J.-M. 2013. GARCH models without positivity constraints: Exponential or Log GARCH? Journal of econometrics, 177(1), 34-46.

[29] Gallo, G.M., Otranto, E., 2015. Forecasting realized volatility with changing average levels. International Journal of Forecasting 31(3), 620-634.

[30] Glosten, L.R., Jagannathan R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48(5), 1779-1801.

[31] Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

[32] Hansen, P. R. and Huang, Z., 2016. Exponential GARCH modeling with realized measures of volatility. Journal of Business & Economic Statistics 34(2): 269-287.

[33] Hansen, P.R., Huang, Z., Shek, H., 2012. Realized GARCH: A joint model for returns and realized measures of volatility. Journal of Applied Econometrics 27, 877-906.

[34] Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise (with discussion). Journal of Business and Economic Statistics 24, 127-218.

[35] Hansen, P. R., Lunde, A. and Nason, J.M., 2011. The model confidence set. Econometrica, 79, 453-497.

[36] Karanasos, M., Yfanti, S. and Christopoulos, A., 2020. The long memory HEAVY process: modeling and forecasting financial volatility. Annals of Operations Research, pp.1-20.

[37] Lanne, M., 2006. A mixture multiplicative error model for realized volatility. Journal of Financial Econometrics 4, 594-616.

[38] Martinet, G. G., and McAleer, M., 2018. On the Invertibility of EGARCH (p, q). Econometric Reviews, 37(8), 824-849.

[39] Nakatani, T., and Teräsvirta, T., 2009. Testing for volatility interactions in the Constant Conditional Correlation GARCH model. The Econometrics Journal, 12(1), 147-163.

[40] Noureldin, D., Shephard, N., Sheppard, K., 2012. Multivariate high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics 27, 907-933.

[41] Opschoor, A., Janus, P., Lucas, A. and Van Dijk, D., (2018). New HEAVY models for fat-tailed realized covariances and returns. Journal of Business and Economic Statistics, 1-15.

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

[43] Sheppard, K. and Xu, W. (2019). Factor High-Frequency-Based Volatility (HEAVY) Models. Journal of Financial Econometrics, 17, 33-65.

[44] Tse, Y.K., 1998. The conditional heteroscedasticity of the yen–dollar exchange rate. Journal of Applied Econometrics 13, 49-55.

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

[46] Taylor, N. 2017. Realised variance forecasting under Box-Cox transformations. International Journal of Forecasting, 33(4), 770-785.

[47] Yfanti, S., Chortareas, G., Karanasos, M. and Noikokyris, E., 2020. A three-dimensional asymmetric power HEAVY model. International Journal of Finance and Economics.

Table 1 – Data descriptive statistics

	$r^2_{t,oc}$		$r^2_{t,cc}$		rk_t	
Symbol	Mean	StDev	Mean	StDev	Mean	StDev
Symbol	Mean	StDev	Mean	StDev	Mean	StDev
AEX	19.14	58.66	30.20	90.48	18.94	41.00
AORD	11.37	33.18	14.18	40.78	11.43	33.62
BFX	16.91	53.84	24.97	82.66	15.56	33.11
BSESN	26.52	73.62	34.11	113.67	21.74	55.93
BVLG	10.33	27.44	20.37	63.63	9.61	18.94
BVSP	32.46	84.07	52.22	154.24	30.25	59.08
DJI	19.23	62.21	22.57	87.40	17.04	40.12
FCHI	20.99	55.16	31.85	89.56	21.01	40.85
FTMIB	21.79	66.42	39.34	140.98	18.84	28.39
FTSE	21.33	64.90	21.73	65.81	20.27	51.21
GDAXI	24.64	66.65	33.99	98.28	24.44	48.37
GSPTSE	15.27	59.91	17.63	76.25	12.04	42.55
HSI	17.00	60.45	33.57	104.67	15.53	32.19
IBEX	23.02	65.02	33.73	101.83	23.17	41.90
IXIC	27.46	84.76	40.05	118.11	23.86	57.06
KS11	20.55	62.85	34.64	106.81	17.60	33.17
KSE	24.59	58.84	29.44	73.34	19.17	33.58
MXX	23.59	64.79	25.83	70.04	17.69	32.61
N225	19.72	71.59	35.04	100.69	17.60	49.25
NSEI	22.07	70.98	33.46	115.24	18.65	52.81
OMXC20	21.14	63.40	26.12	74.09	19.12	48.68
OMXHPI	21.90	61.17	28.30	77.33	20.08	44.57
OMXSPI	21.60	67.17	27.65	85.35	19.03	53.95
OSEAX	27.16	77.85	30.13	89.10	21.63	60.62
RUT	23.10	68.98	38.58	117.01	18.65	45.35
SMSI	22.77	69.16	32.52	107.83	21.71	47.22
SPX	19.90	64.16	24.46	86.94	16.68	40.08
SSEC	30.76	74.33	37.52	99.60	26.41	43.12
SSMI	14.76	52.42	21.43	68.79	13.74	35.12
STI	10.50	24.38	18.40	78.33	9.05	15.51
STOXX50E	26.55	75.92	32.56	90.84	25.59	57.26
Average	21.23	62.72	29.89	92.57	18.91	42.49

This table provides descriptive statistics for the dataset of 31 assets. Columns 2-3 report the corresponding time-series averages (Mean) and standard deviations (StDev) of the squared open-to-close returns ($r^2_{t,oc}$). Columns 4-5 report the corresponding Mean and StDev of the squared close-to-close returns ($r^2_{t,cc}$). Columns 6-7 report the corresponding Mean and StDev of the realized kernel (rk_t).

Table 2 – EHEAVY models estimated parameters

	Open-Close Returns				Close-Close Returns			
Penal A: EGARCH model estimation results								
	Min	Med	Max	No Sig.	Min	Med	Max	No Sig.
w_r	-0.111	-0.059	-0.022	**28**	-0.087	-0.043	0.005	**24**
α_{rr}	0.118	0.160	0.227	**31**	0.103	0.158	0.214	**31**
β_r	0.960	0.978	0.988	**31**	0.956	0.976	0.985	**31**
γ_{rr}	-0.143	-0.09	-0.023	**31**	-0.152	-0.106	-0.025	**31**
Penal B: Full EHEAVY model (2) estimation results								
	Open-Close Returns				Close-Close Returns			
	Min	Med	Max	No Sig.	Min	Med	Max	No Sig.
w_r	-0.414	-0.273	-0.115	**30**	-0.468	-0.316	-0.038	**30**
α_{rr}	-0.032	0.021	0.142	**13**	-0.013	0.044	0.205	**23**
α_{rR}	0.221	0.366	0.602	**31**	0.051	0.195	0.619	**30**
β_r	0.919	0.965	0.980	**31**	0.939	0.949	0.998	**31**
γ_{rr}	-0.158	-0.106	0.001	**28**	-0.169	-0.138	-0.013	**29**
γ_{rR}	-0.072	0.009	0.059	**4**	-0.068	-0.028	0.087	**2**
w_R	-0.441	-0.305	-0.158	**31**	-0.515	-0.376	-0.162	**31**
α_{rR}	-0.018	0.040	0.201	**15**	-0.011	0.026	0.220	**20**
α_{RR}	0.232	0.389	0.573	**31**	0.231	0.315	0.640	**31**
β_R	0.888	0.959	0.978	**31**	0.920	0.946	0.998	**31**
γ_{Rr}	-0.160	-0.107	-0.009	**27**	-0.175	-0.134	-0.031	**30**
γ_{RR}	-0.067	0.010	0.084	**5**	-0.031	-0.015	0.091	**1**
ρ	0.819	0.845	0.890	**31**	0.780	0.816	0.888	**31**
Penal C: EHEAVY model (3) estimation results								
	Open-Close Returns				Close-Close Returns			
w_r	-0.432	-0.282	-0.158	**31**	-0.478	-0.268	-0.104	**31**
α_{rR}	0.275	0.406	0.588	**31**	0.277	0.394	0.675	**31**
β_r	0.921	0.966	0.983	**31**	0.935	0.967	0.990	**30**
γ_{rr}	-0.150	-0.093	-0.023	**30**	-0.154	-0.108	-0.052	**31**
w_R	-0.504	-0.327	-0.204	**31**	-0.523	-0.305	-0.186	**31**
α_{RR}	0.301	0.448	0.696	**31**	0.298	0.430	0.685	**31**
β_R	0.887	0.962	0.981	**31**	0.919	0.962	0.989	**31**
γ_{Rr}	-0.149	-0.093	-0.026	**30**	-0.146	-0.105	-0.033	**31**
ρ	0.818	0.845	0.890	**31**	0.773	0.815	0.886	**31**
Penal D: Alternative EHEAVY model (21) estimation results								
	Open-Close Returns				Close-Close Returns			
w_r	-0.455	-0.329	-0.171	**31**	-0.566	-0.329	-0.113	**31**
α_{rR}	0.292	0.437	0.613	**31**	0.289	0.433	0.704	**31**
β_r	0.927	0.967	0.982	**31**	0.932	0.967	0.986	**31**
γ_{rR}	-0.133	-0.077	-0.008	**28**	-0.142	-0.092	-0.020	**30**
w_R	-0.525	-0.360	-0.224	**31**	-0.585	-0.351	-0.207	**31**
α_{RR}	0.319	0.481	0.720	**31**	0.320	0.471	0.698	**31**
β_R	0.894	0.963	0.980	**31**	0.916	0.967	0.985	**31**
γ_{RR}	-0.125	-0.073	-0.014	**29**	-0.127	-0.085	0.008	**30**
ρ	0.818	0.845	0.890	**31**	0.771	0.815	0.887	**31**

Minimum (Min), median (Med), and maximum (Max) are the summary statistics of the estimates for the 31 assets. All estimates are provided the supplementary appendix. No Sig. denotes the numbers of significance for the corresponding parameters. The total number of assets is 31. In the alternative EHEAVY model, the asymmetric effect is modelled by realized return shock only.

Table 3 – In sample statistics comparison

	Close-Close Returns				Open-Close Returns			
	LL		BIC		LL		BIC	
	Mean	No.	Mean	No.	Mean	No.	Mean	No.
EGARCH	-19352	1	38738	1	-17834	0	35701	0
FEHEAVY	-19221	19	38494	8	-17672	20	35394	6
EHEAVY	**-19238**	**10**	**38510**	**21**	**-17674**	**10**	**35381**	**24**
AEHEAVY	-19282	1	38596	1	-17696	1	35426	1

LL denotes partial log-likelihood value defined in (14) and (15). BIC denotes the Bayesian information criteria. No. is numbers of assets whereby each model achieves the smallest criteria. The total number of assets is 31. FEHEAVY is the full EHEAVY model defined in (3). AEHEAVY is the alternative EHEAVY model defined in (21), where the asymmetric effect is modelled by realized return shock only.

Table 4 – Out of sample forecasts

Conditional Variance of Returns						
	s=1		s=5		s=22	
(Average)	MSE	QLIK	MSE	QLIK	MSE	QLIK
HEAVY	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
AHEAVY	0.9651	0.8811	0.9916	0.9813	1.0003	0.9857
REGARCH	0.9440	0.8249	0.9635	0.9288	0.9620	0.9182
EHEAVY	**0.9486**	**0.8266**	**0.9660**	**0.9359**	**0.9630**	**0.9118**

Realized Measures of Volatility						
	s=1		s=5		s=22	
(Average)	MSE	QLIK	MSE	QLIK	MSE	QLIK
HEAVY	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
AHEAVY	0.8983	0.9009	0.9915	0.9429	1.0142	0.9866
REGARCH	0.7642	0.6364	0.9145	0.8931	0.9450	0.9432
EHEAVY	**0.7230**	**0.5878**	**0.8984**	**0.8049**	**0.9244**	**0.8658**

This table reports the ratio of the losses for the different models relative to the losses of the benchmark HEAVY model. The average ratios across the 31 assets are reported. AHEAVY denotes the asymmetric HEAVY model. REGARCH denotes the realized GARCH model.

Table 5 – Out of sample forecasts - MCS test

	Conditional Variance of Returns					
	s=1		s=5		s=22	
MSC	MSE	QLIK	MSE	QLIK	MSE	QLIK
HEAVY	31	2	30	10	28	5
AHEAVY	30	8	31	13	26	11
REGARCH	31	29	31	30	31	25
EHEAVY	**31**	**28**	**31**	**30**	**31**	**27**
	Realized Measures of Volatility					
	s=1		s=5		s=22	
	MSE	QLIK	MSE	QLIK	MSE	QLIK
HEAVY	29	0	30	7	28	0
AHEAVY	27	1	31	11	29	0
REGARCH	30	12	31	15	30	1
EHEAVY	**31**	**30**	**31**	**31**	**31**	**31**

This table reports the numbers whereby each model is part of the 90% model confidence set. The total number of assets is 31. AHEAVY denotes the asymmetric HEAVY model. REGARCH denotes the realized GARCH model.

Table 6 – Performances of portfolios formed by volatility forecasts

	Mean			Min			Max		
	R	SR	CER	R	SR	CER	R	SR	CER
HEAVY	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
AHEAVY	0.997	0.978	1.030	0.989	0.928	0.637	1.006	1.054	2.063
RGARCH	1.011	1.102	1.009	0.976	0.909	0.631	1.052	1.396	1.369
EHEAVY	**1.081**	**1.579**	**2.052**	**1.058**	**1.160**	**1.853**	**1.109**	**1.800**	**2.883**

This table shows the performances of portfolios formed by the different volatility forecasts. It gives the mean excess return (R), Sharpe ratio and certainty equivalent return (CER) of each portfolio. The ratios of the performance for the different models relative to the that of HEAVY model are calculated. The Mean, Minimum (Min), Maximum (Max) of the ratios across the 31 assets are reported in the table.

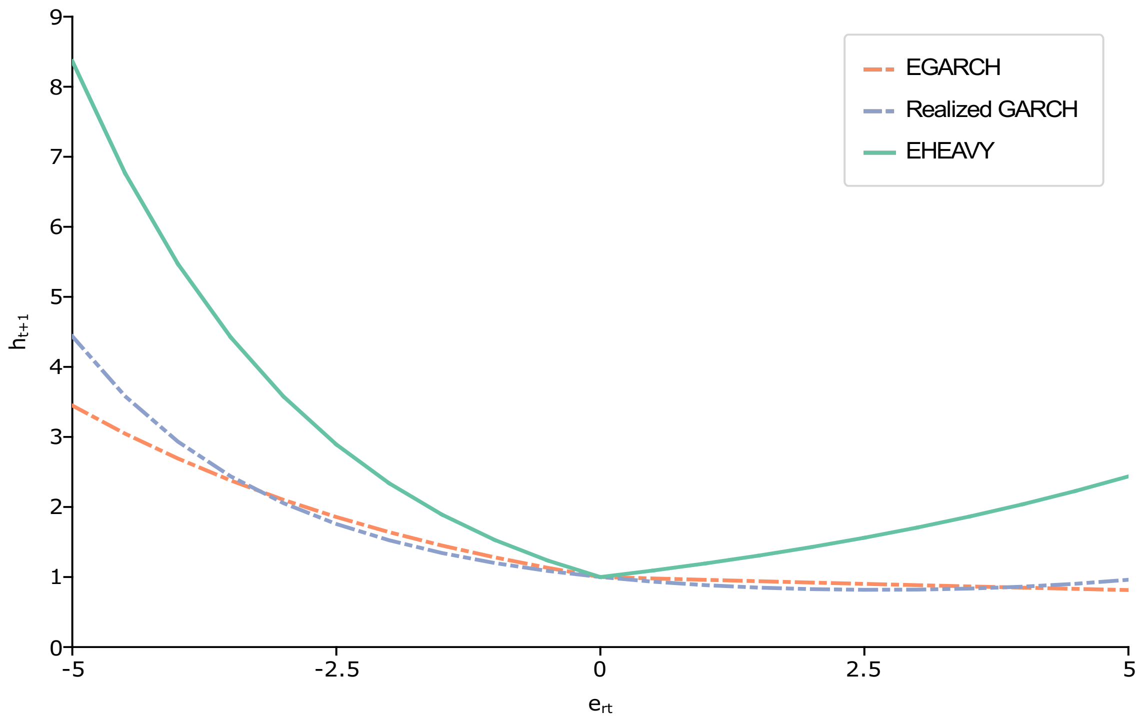

Figure 1 – News impact curve for the EGARCH, Realized EGARCH and EHEAVY model: EURO50 close-to-close return.

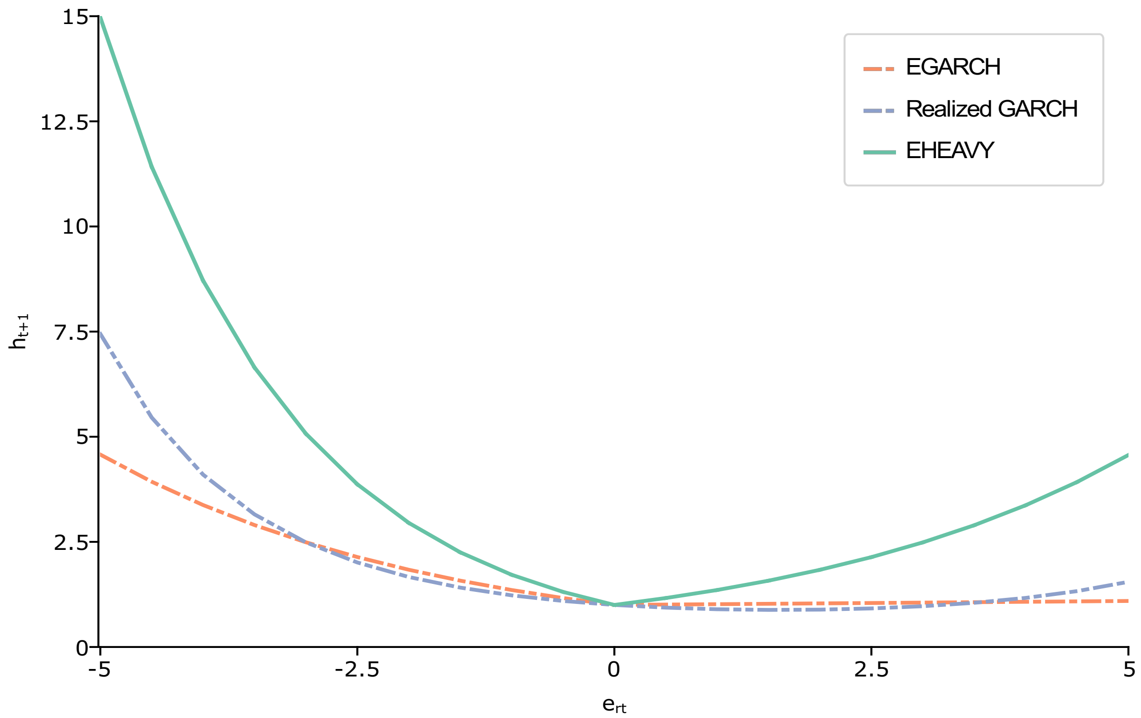

Figure 2 – News impact curve for the EGARCH, EHEAVY and Realized EGARCH model: SPX close-to-close return

Supplementary Appendix: Additional Tables

Table A1: Symbol Names

Symbol	Name	Start Date	Obs.(T)
.AEX	AEX	January 03, 2000	5473
.AORD	All Ordinaries	January 04, 2000	5422
.BFX	Bell 20	January 03, 2000	5471
.BSESN	S&P BSE Sensex	January 03, 2000	5322
.BVLG	PSI All-Share	October 15, 2012	2207
.BVSP	BVSP BOVESPA	January 03, 2000	5282
.DJI	Dow Jones Industrials	January 03, 2000	5380
.FCHI	CAC 40	January 03, 2000	5475
.FTMIB	FTSE MIB	June 01, 2009	3056
.FTSE	FTSE 100	January 04, 2000	5414
.GDAXI	DAX	January 03, 2000	5440
.GSPTSE	S&P/TSX Composite	May 02, 2002	4786
.HSI	HANG SENG	January 03, 2000	5258
.IBEX	IBEX 35	January 03, 2000	5440
.IXIC	Nasdaq 100	January 03, 2000	5384
.KS11	Korea Composite	January 04, 2000	5283
.KSE	Karachi SE 100	January 03, 2000	5229
.MXX	IPC Mexico	January 03, 2000	5384
.N225	Nikkei 225	February 02, 2000	5219
.NSEI	NIFTY 50	January 03, 2000	5313
.OMXC20	OMX Copenhagen 20	October 03, 2005	3902
.OMXHPI	OMX Helsinki All Share	October 03, 2005	3943
.OMXSPI	OMX Stockholm All Share	October 03, 2005	3943
.OSEAX	Oslo Exchange All-share	September 03, 2001	4931
.RUT	Russel 2000	January 03, 2000	5381
.SMSI	Madrid General	July 04, 2005	4069
.SPX	S&P 500	January 03, 2000	5383
.SSEC	Shanghai Composite	January 04, 2000	5184
.SSMI	Swiss Stock Market	January 04, 2000	5377
.STI	Straits Times	January 03, 2000	3439
.STOXX50E	EURO STOXX 50	January 03, 2000	5472

This table provides descriptive statistics for the dataset of 31 assets. The sample ending date is June 24, 2021. Column 2 show the asset names. Column 3 show the sample starting date. Column 4 is in the number of observations (or the sample size T) of the asset.

Table A2: EHEAVY models Estimated parameters: Open-Close Returns

	HEAVY-r: Stock Returns					HEAVY-RM: Realized Measure					
Symbol	w	α_{RR}	β_R	γ_{Rr}	LL	w	α_{RR}	β_R	γ_{Rr}	ρ	LL
AEX	-0.287	0.378	0.971	-0.120	-18627	-0.310	0.412	0.967	-0.120	0.833	-18636
AORD	-0.187	0.275	0.970	-0.104	-16490	-0.208	0.301	0.968	-0.110	0.890	-16233
BFX	-0.323	0.432	0.964	-0.107	-18377	-0.337	0.445	0.964	-0.101	0.837	-17974
BSESN	-0.422	0.588	0.952	-0.064	-20360	-0.504	0.696	0.941	-0.059	0.841	-19099
BVLG	-0.407	0.497	0.970	-0.064	-6636	-0.424	0.519	0.967	-0.064	0.845	-6511
BVSP	-0.236	0.384	0.962	-0.071	-22071	-0.294	0.473	0.953	-0.080	0.864	-21470
DJI	-0.328	0.463	0.957	-0.146	-18132	-0.329	0.462	0.956	-0.137	0.833	-17603
FCHI	-0.245	0.352	0.967	-0.115	-19767	-0.248	0.360	0.966	-0.112	0.838	-19774
FTMIB	-0.358	0.516	0.951	-0.115	-11535	-0.354	0.501	0.952	-0.111	0.818	-11081
FTSE	-0.300	0.405	0.969	-0.093	-19310	-0.341	0.456	0.966	-0.081	0.836	-19093
GDAXI	-0.234	0.326	0.972	-0.109	-20444	-0.252	0.346	0.971	-0.114	0.828	-20174
GSPTSE	-0.260	0.337	0.975	-0.096	-14759	-0.264	0.344	0.972	-0.108	0.859	-13711
HSI	-0.244	0.338	0.971	-0.023	-18324	-0.299	0.420	0.961	-0.025	0.846	-17965
IBEX	-0.312	0.411	0.971	-0.088	-20435	-0.339	0.448	0.969	-0.083	0.831	-20496
IXIC	-0.327	0.447	0.965	-0.109	-20077	-0.362	0.488	0.962	-0.110	0.836	-19240
KS11	-0.326	0.393	0.983	-0.047	-18679	-0.341	0.412	0.981	-0.049	0.846	-18044
KSE	-0.158	0.424	0.921	-0.066	-20175	-0.212	0.560	0.887	-0.063	0.866	-18901
MXX	-0.252	0.410	0.955	-0.078	-20552	-0.327	0.512	0.942	-0.077	0.837	-19042
N225	-0.269	0.401	0.963	-0.057	-18818	-0.293	0.436	0.958	-0.070	0.863	-18157
NSEI	-0.432	0.551	0.966	-0.073	-19105	-0.464	0.598	0.959	-0.070	0.834	-18192
OMXC20	-0.262	0.437	0.948	-0.076	-14503	-0.355	0.565	0.935	-0.079	0.846	-13957
OMXHPI	-0.231	0.335	0.970	-0.100	-14382	-0.282	0.408	0.962	-0.112	0.863	-13890
OMXSPI	-0.273	0.376	0.972	-0.135	-13966	-0.293	0.402	0.968	-0.149	0.864	-13343
OSEAX	-0.237	0.364	0.966	-0.093	-19049	-0.274	0.407	0.962	-0.100	0.857	-17775
RUT	-0.192	0.345	0.954	-0.116	-20173	-0.204	0.344	0.955	-0.120	0.855	-18928
SMSI	-0.321	0.431	0.971	-0.092	-15299	-0.349	0.465	0.968	-0.093	0.855	-15042
SPX	-0.282	0.406	0.960	-0.150	-18218	-0.283	0.404	0.959	-0.146	0.840	-17358
SSEC	-0.387	0.511	0.972	-0.040	-21147	-0.466	0.612	0.964	-0.043	0.836	-20196
SSMI	-0.422	0.559	0.950	-0.087	-17316	-0.441	0.565	0.954	-0.072	0.825	-16734
STI	-0.406	0.513	0.962	-0.047	-10360	-0.438	0.550	0.958	-0.043	0.846	-9884
EURO50	-0.207	0.320	0.967	-0.138	-20799	-0.223	0.344	0.964	-0.133	0.847	-20619
Mean	-0.294	0.417	0.963	-0.091	-17674	-0.326	0.460	0.958	-0.091	0.846	-17068

This table provides QML estimates of EHEAVY model for the dataset of 31 assets. All parameters are statistically significant at 5% level.

Table A3: EHEAVY models Estimated parameters: Close-Close Returns

	HEAVY-r: Stock Returns					HEAVY-RM: Realized Measure					
Symbol	w	α_{RR}	β_R	γ_{Rr}	LL	w	α_{RR}	β_R	γ_{Rr}	ρ	LL
AEX	-0.235	0.341	0.970	-0.154	-20928	-0.259	0.356	0.969	-0.137	0.807	-18617
AORD	-0.181	0.278	0.970	-0.107	-17547	-0.200	0.298	0.966	-0.113	0.886	-16222
BFX	-0.338	0.450	0.970	-0.121	-20266	-0.337	0.430	0.972	-0.105	0.815	-17962
BSESN	-0.362	0.536	0.955	-0.075	-21524	-0.451	0.643	0.943	-0.064	0.804	-19085
BVLG	-0.366	0.526	0.954	-0.106	-8136	-0.388	0.517	0.952	-0.099	0.809	-6501
BVSP	-0.130	0.296	0.961	-0.091	-24674	-0.268	0.439	0.955	-0.080	0.815	-21449
DJI	-0.343	0.478	0.959	-0.132	-18813	-0.312	0.428	0.963	-0.133	0.824	-17586
FCHI	-0.268	0.394	0.967	-0.141	-21842	-0.290	0.401	0.968	-0.113	0.811	-19761
FTMIB	-0.389	0.626	0.938	-0.140	-13167	-0.416	0.599	0.942	-0.109	0.779	-11070
FTSE	-0.278	0.371	0.973	-0.103	-19409	-0.332	0.436	0.969	-0.091	0.835	-19079
GDAXI	-0.240	0.335	0.975	-0.129	-22124	-0.272	0.354	0.977	-0.118	0.810	-20152
GSPTSE	-0.214	0.307	0.970	-0.124	-15426	-0.223	0.305	0.970	-0.113	0.850	-13693
HSI	-0.189	0.295	0.972	-0.074	-21655	-0.222	0.319	0.969	-0.057	0.785	-17947
IBEX	-0.317	0.458	0.963	-0.108	-22331	-0.368	0.500	0.962	-0.086	0.806	-20478
IXIC	-0.299	0.427	0.967	-0.128	-21883	-0.305	0.418	0.967	-0.123	0.797	-19215
KS11	-0.308	0.412	0.975	-0.082	-21169	-0.326	0.420	0.973	-0.064	0.787	-18026
KSE	-0.104	0.277	0.954	-0.056	-21268	-0.186	0.445	0.919	-0.068	0.841	-18884
MXX	-0.231	0.383	0.958	-0.081	-20975	-0.323	0.506	0.943	-0.070	0.831	-19037
N225	-0.124	0.362	0.937	-0.104	-22153	-0.228	0.420	0.939	-0.098	0.824	-18140
NSEI	-0.271	0.327	0.990	-0.078	-22000	-0.422	0.483	0.989	-0.063	0.773	-18200
OMXC20	-0.230	0.459	0.935	-0.093	-15350	-0.289	0.495	0.936	-0.092	0.834	-13948
OMXHPI	-0.177	0.302	0.967	-0.115	-15525	-0.268	0.404	0.961	-0.109	0.838	-13882
OMXSPI	-0.260	0.395	0.965	-0.151	-15115	-0.338	0.472	0.962	-0.146	0.835	-13334
OSEAX	-0.244	0.360	0.971	-0.087	-19406	-0.276	0.397	0.966	-0.098	0.853	-17771
RUT	-0.172	0.295	0.968	-0.119	-22618	-0.190	0.308	0.962	-0.128	0.810	-18899
SMSI	-0.335	0.454	0.972	-0.118	-16590	-0.401	0.519	0.970	-0.107	0.815	-15026
SPX	-0.284	0.422	0.959	-0.141	-19172	-0.257	0.373	0.961	-0.145	0.830	-17342
SSEC	-0.410	0.571	0.962	-0.052	-22109	-0.478	0.651	0.955	-0.054	0.813	-20200
SSMI	-0.375	0.539	0.947	-0.140	-19147	-0.408	0.536	0.952	-0.111	0.812	-16706
STI	-0.478	0.675	0.940	-0.071	-12173	-0.523	0.685	0.938	-0.033	0.792	-9890
EURO50	-0.192	0.301	0.970	-0.147	-21887	-0.224	0.337	0.967	-0.136	0.841	-20602
Mean	-0.269	0.408	0.962	-0.109	-19238	-0.316	0.448	0.959	-0.099	0.818	-17055

This table provides QML estimates of EHEAVY model for the dataset of 31 assets. All parameters are statistically significant at 5% level.

Appendix B: AHEAVY and Realized GARCH model specification

In this appendix, we present the specification of alternative EHEAVY model, asymmetric HEAVY of Shepard and Shepperd (2010) and realized EGARCH Hansen and Huang (2016) model.

The alternative EHEAVY model, where the asymmetric is modelled by realized return shock only, is given by

$$\log h_t = \omega_r + \beta_r \log h_{t-1} + \alpha_{rR}|e_{Rt-1}| + \gamma_{rR}e_{Rt-1},$$
$$\log m_t = \omega_R + \beta_R \log m_{t-1} + \alpha_{RR}|e_{Rt-1}| + \gamma_{RR}e_{Rt-1} \quad (21)$$

The AHEAVY model of Shepard and Shepperd (2010) is

$$h_t = \omega_r + (\alpha_{rR} + \gamma_{rR} s_{t-1}) RM_{t-1} + \beta_r h_{t-1}, \tag{22}$$
$$m_t = \omega_R + (\alpha_{RR} + \gamma_{RR} s_{t-1}) RM_{t-1} + \beta_R m_{t-1}.$$

where $s_t = 0.5[1 - sign(r_t)]$, that is, $s_t = 1$ if $r_t < 0$ and 0 otherwise; γ_{ii}, γ_{ij} $(i \neq j)$ are the own and cross leverage parameters, respectively[10]; positive γ_{ii}, γ_{ij} means a larger contribution of negative 'shocks' in the volatility process.

The realized EGARCH model of Hansen and Huang (2016) is

$$\log h_t = \omega_r + \beta_r \log h_{t-1} + \alpha_{rr} e_{rt-1}^2 + \gamma_{rr} e_{rt-1} + \alpha_{rR} u_{Rt-1}, \tag{23}$$
$$\log RM_t = \omega_R + \beta_R \log h_t + \alpha_{Rr} e_{rt-1}^2 + \gamma_{Rr} e_{Rt-1} + u_{Rt}.$$

where $u_{Rt} \sim N(0, \sigma_u)$. The first equation is referred to GARCH equation, and the second one is referred to measurement equation. It can be seen that the GARCH equation is close to our EHEAVY-r equation, where the measurement equation has a different specification as EHEAVY-RM equation.

[10]This type of asymmetry was introduced by Glosten et. al., (1993).