S.I.: REGRESSION METHODS BASED ON OR TECHNIQUES

The long memory HEAVY process: modeling and forecasting financial volatility

M. Karanasos[1] · **S. Yfanti[2]** · **A. Christopoulos[3]**

Abstract
This paper studies the bivariate HEAVY system of volatility regression equations and its various extensions that are directly applicable to the day-to-day business treasury operations of trading in foreign exchange and commodities, investing in bond and stock markets, hedging out market risk, and capital budgeting. We enrich the HEAVY framework with powers, asymmetries, and long memory that improve its forecasting accuracy significantly. Other findings are as follows. First, hyperbolic memory fits the realized measure better, whereas fractional integration is more suitable for the powered returns. Second, the structural breaks applied to the bivariate system capture the time-varying behavior of the parameters, in particular during and after the global financial crisis of 2007/2008.

Keywords Asymmetries · Financial crisis · Forecasting · HEAVY model · High-frequency data · Long memory · Power transformations · Realized variance · Risk management · Structural breaks

JEL Classification C22 · C52 · C58 · G01 · G15

1 Introduction

The volatility of financial returns constitutes a pivotal part of the empirical finance and econometrics research, with crucial implications in portfolio optimization and risk management practices. Robust modeling and reliable forecasting of the volatility trajectory of financial instruments has been the main task and objective of financial economics applications for business operations, given that volatility constitutes one of the fundamental input variables in estimations and decision processes of any non-financial corporation on derivatives pricing, portfolio immunization, investment diversification, firm valuation and funding choices. There are several studies introducing non-parametric estimators of realized volatility

✉ M. Karanasos
    menelaos.karanasos@brunel.ac.uk

[1] Economics and Finance, Brunel University London, London UB8 3PH, UK

[2] Loughborough University, Loughborough, UK

[3] National and Kapodistrian University of Athens, Athens, Greece

using high-frequency market data. Andersen and Bollerslev (1998), Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) were the first that econometrically formalized the realized variance with quadratic variation-like measures, while Barndorff-Nielsen et al. (2008, 2009) focused on the realized kernel estimation as a realized measure which is more robust to noise.

A large body of empirical research focuses on modeling and forecasting the realized volatility. Various studies combine it with the conditional variance of returns. Engle (2002b) proposed the GARCH-X process, where the former is included as an exogenous variable in the equation of the latter. Corsi et al. (2008) suggested the HAR-GARCH formulation for modeling the volatility of realized volatility.

Hansen et al. (2012) introduced the Realized GARCH model that corresponds most closely to the HEAVY framework of Shephard and Sheppard (2010), which jointly estimates conditional variances based on both daily (squared returns) and intra-daily (it uses the realized kernel as a measure of ex-post volatility) data, so that the system of equations adopts to information arrival more rapidly than the classic daily GARCH process. One of its advantages is the robustness to certain forms of structural breaks, especially during crisis periods, since the mean reversion and short-run momentum effects result in higher quality performance in volatility level shifts and more reliable forecasts. Borovkova and Mahakena (2015) employed a HEAVY specification with a skewed-t error distribution, while Huang et al. (2016) incorporated the HAR structure of the realized measure in the GARCH conditional variance specification in order to capture the long memory of the volatility dynamics.

This paper examines the HEAVY model by enriching the bivariate system with asymmetries and power transformations, through the structure of Ding et al. (1993). Among others, Pérez et al. (2009, see the references therein for more details) show that the presence of an asymmetric response of volatility to positive and negative returns shows up in non-zero cross-correlations between original returns and future powers of absolute returns. One of our main findings is that each of the two powered conditional variances is significantly affected by the first lags of both power transformed variables, that is, squared negative returns, and realized kernel (or, for the latter, its negative signed values).

We analyze the various specifications in depth and we investigate their performance over six stock indices. We also take into account long memory (either fractionally integrated or hyperbolic), by employing the framework of Davidson (2004) (see, Schoffer 2003 and Dark 2005, 2010, as well). We find that a fractionally integrated formulation better fits the squared returns, whereas a hyperbolic type of memory is more suitable for the realized measure. The long memory feature reinforces our main argument that the lagged values of the power transformations of both aforementioned variables move the dynamics of the two powered conditional variances. The fractionally integrated (asymmetric power) model for the returns equation pools information across both low-frequency and high-frequency based volatility indicators. Similarly, the more richly parametrized hyperbolic process for the realized kernel equation is bolstered with low-frequency information as well since the lagged value of the powered squared negative returns improves the forecasting performance of the model. Finally, in the presence of structural breaks, which are apparent in the two power transformed volatility measures, we re-estimate the bivariate system including dummy variables, and we present the time-varying behavior of the parameters. Focusing on the recent global financial crisis, we observe that their values increase after the crisis.

With the advent of the crisis, when volatilities increased sharply and persistently with crucial systemic risk externalities, we witnessed a resurgence of regulators' and academics' interest in meaningful volatility estimates, while, at the same time, practitioners remained alert to improving the relevant volatility frameworks on a day-to-day basis. Financial eco-

nomics scholars focused on volatility as a potent catalyst of systemic risk build-up, which policymakers tried to limit. We demarcate this study from the extant finance bibliography by extending the benchmark HEAVY model with asymmetries, power transformations and long memory providing a well-defined framework that adequately fits the volatility process. The bivariate system of the two volatility regression equations, we establish, is ready-to-use, not only on stock markets returns, but also on multiple financial economics applications of business operations research, such as bonds investing, foreign exchange trading and commodities hedging, core daily functions in the treasuries of most non-financial corporations, besides the financial services sector.

Overall, our proposed volatility modeling framework improves the HEAVY model, with major implications for market practitioners and policymakers on forecasting the financial returns second moment. Volatility modeling and forecasting is essential for asset allocation, pricing and risk management hedging strategies. A reliable volatility forecast, exploiting in full the high-frequency domain, is the input variable of paramount importance for the processes of derivatives pricing, effective cross-hedging, Value-at-Risk measurement, investment allocation and portfolio optimization with different asset classes and financial instruments. Moreover, the robust volatility modeling approach we introduce provides a tool of utmost significance not only for market players but also for policymakers. Policymaking includes continuous oversight duties and prudential regulation practices. In this vein, it is imperative for the authorities to account for the volatility of financial markets across every aspect of the financial system's policy responses, both post-crisis through stabilization policy reactions and pre-crisis through proactive assessment of financial risks.

The remainder of the paper is structured as follows. In Sect. 2, we detail the HEAVY formulation and our first extension, which allows for asymmetries and power transformations. Section 3 describes the data and Sect. 4 presents the results for the asymmetric power specification. The next section studies the long memory process and discusses the relevant empirical findings. In Sect. 6, we estimate multiple-step-ahead forecasts to measure the out-of-sample performance of the various specifications. The following section takes into consideration the presence of structural breaks. In Sect. 8, we discuss the dynamic conditional correlations estimation procedure for the extended HEAVY models. Finally, Sect. 9 concludes the analysis.

2 The framework

The benchmark HEAVY model of Shephard and Sheppard (2010) can be extended in many directions. We allow for power transformations, leverage effects and long memory (see Sects. 4 and 5 below) in the conditional variance process. We run the estimated benchmark specification of Shephard and Sheppard (2010), enriched with the three key features to improve the volatility modeling and forecasting further.

2.1 The HEAVY model

The HEAVY model uses two variables: the close-to-close stock returns (r_t) and the realized measure of variation based on high frequency data, RM_t. We first form the signed square rooted (SSR) realized measure as follows: $\widetilde{RM_t} = sign(r_t)\sqrt{RM_t}$, where $sign(r_t) = 1$, if $r_t \geq 0$ and $sign(r_t) = -1$, if $r_t < 0$.

We assume that the returns and the SSR realized measure are characterized by the following relations:

$$r_t = e_{rt}\sigma_{rt}, \quad \widetilde{RM_t} = e_{Rt}\sigma_{Rt}, \tag{1}$$

where the stochastic term e_{it} is independent and identically distributed (*i.i.d*), $i = r, R$; σ_{it} is positive with probability one for all t and it is a measurable function of $\mathcal{F}_{t-1}^{(XF)}$, which is the filtration generated by all available information through time $t - 1$. We will use $\mathcal{F}_{t-1}^{(HF)}$ ($X = H$) for the high frequency past data, i.e., for the case of the realized measure, or $\mathcal{F}_{t-1}^{(LoF)}$ ($X = Lo$) for the low frequency past data, i.e., for the case of the close-to-close returns. Hereafter, for notational convenience we will drop the superscript XF.

In the HEAVY/GARCH model e_{it} has zero mean and unit variance. Therefore, the two series have zero conditional means, and their conditional variances are given by

$$\mathbb{E}\left(r_t^2 \,|\mathcal{F}_{t-1}\right) = \sigma_{rt}^2, \text{ and } \mathbb{E}\left(\widetilde{RM_t}^2 \,|\mathcal{F}_{t-1}\right) = \mathbb{E}(RM_t \,|\mathcal{F}_{t-1}) = \sigma_{Rt}^2, \tag{2}$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator.

2.2 Asymmetric power specification

The asymmetric power (AP) specification for the HEAVY(1, 1) model consists of the following equations (in what follows for notational simplicity we will drop the order of the model if it is (1, 1)):

$$(1 - \beta_i L)\left(\sigma_{it}^2\right)^{\frac{\delta_{ii}}{2}} = \omega_i + (\alpha_{ir} + \gamma_{ir}s_{t-1})L\left(r_t^2\right)^{\frac{\delta_{ir}}{2}}$$
$$+ (\alpha_{iR} + \gamma_{iR}s_{t-1})L(RM_t)^{\frac{\delta_{iR}}{2}}, \quad \text{for } i = r, R \tag{3}$$

where L is the lag operator, $\delta_{ij} \in \mathbb{R}_{>0}$ (the set of the positive real numbers) are the power parameters, and $s_t = 0.5[1 - sign(r_t)]$, 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[1]; positive γ_{ii}, γ_{ij} means larger contribution of negative 'shocks' in the volatility process (in our long memory AP specification we will replace $\alpha_{ii} + \gamma_{ii}s_{t-1}$ by $\alpha_{ii}(1 + \gamma_{ii}s_{t-1})$; see Sect. 5 below, and, in particular, Eq. (4)). In this specification the powered conditional variance, $(\sigma_{it}^2)^{\delta_{ii}/2}$, is a linear function of the lagged values of the powered transformed squared returns and realized measure.

We will distinguish between three different asymmetric cases: the double one (DA: $\gamma_{ij} \neq 0$) and two more: own asymmetry (OA: $\gamma_{ii} \neq 0$, $\gamma_{ij} = 0$ for $i \neq j$) and cross asymmetry (CA: $\gamma_{ii} = 0$, $\gamma_{ij} \neq 0$ for $i \neq j$).

The α_{iR} and γ_{iR} are called the (four) Heavy parameters (own when $i = R$ and cross when $i \neq R$). These parameters capture the impact of the realized measure on the two conditional variances. Similarly, the α_{ir} and γ_{ir} (four in total) are called the Garch parameters (own when $i = r$ and cross for $i \neq r$). They depict the influence of the squared returns on the two conditional variances. When all four Heavy parameters are zero the AP-HEAVY model is reduced to a bivariate AP-GARCH process (see, for example, Conrad and Karanasos 2010). If, on the other hand, all four Garch parameters are zero, then we have the AP-HEAVY specification where the only unconditional regressor is the first lag of the powered RM_t. Finally, we should mention that all the parameters in this bivariate system should take non-negative values (see, for example, Conrad and Karanasos 2010).

---

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

To sum up, the benchmark HEAVY(1, 1) model, to be extended in this study, is characterized by two conditional variance equations:

$$(1 - \beta_r L)\sigma_{rt}^2 = \omega_r + \alpha_{rR} L(RM_t),$$
$$(1 - \beta_R L)\sigma_{Rt}^2 = \omega_R + \alpha_{RR} L(RM_t)$$

Equation (3) gives the general formulation of our first HEAVY extension for both variables, where we add asymmetries and power transformations. We will first estimate both conditional variance equations in the general form with all Heavy, Garch and Asymmetry parameters given by Eq. (3) and in case a parameter is insignificant, we will exclude it and this will result in a reduced form that is statistically preferred for each volatility process. For example, in the returns and realized conditional variances estimation, the own and cross Garch parameters (α_{rr} and α_{Rr} respectively) are proved insignificant and, therefore, excluded (see Sect. 4, Table 2, Panels A and B) since this is the way to reach the returns and realized measure formulations that are statistically preferred.

We conclude this section by highlighting the fact that the estimation (see Sect. 4) of our extended asymmetric power specification proves that both powered conditional variances receive the notable impact from the first lag of the power transformed negative returns. Therefore, the conclusion that the realized measure of variation does all the work at moving around the conditional variance of stock returns does not hold for the more richly parametrized model.

3 Data description

The HEAVY framework is estimated for six stock indices returns and realized volatilities. According to the analysis in Shephard and Sheppard (2010), the HEAVY formulation considerably improves the volatility modeling by allowing momentum and mean reversion effects and adjusting quickly to the structural breaks in volatility. We extend the benchmark specification in Shephard and Sheppard (2010), by adding the features of power transformed conditional variances, leverage effects and long memory (see Sect. 5 below) in the volatility process. Moreover, in order to identify the possible recent global financial crisis effects on the volatility process and to take into account the structural breaks in the two powered series (squared returns and realized measure), in Sect. 7 we incorporate dummies in our empirical investigation. The analysis with the structural breaks can be considered as an alternative to the long memory investigation.

3.1 Oxford-Man Institute's library

We use daily data for six market indices extracted from the Oxford-Man Institute's (OMI) realized library version 0.2 of Heber et al. (2009): S&P 500 from the US (SP), Nikkei 225 from Japan (NIKKEI), TSE from Canada, FTSE 100 from the UK (FTSE), DAX from Germany and Eustoxx 50 from the Eurozone (EUSTOXX). Our sample covers the period from 03/01/2000 to 30/11/2017 for most indices. For the Canadian stock market index TSE the data begin from 2002. 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

Table 1 Data description

Index	Sample period			r_t^2		RK_t	
	Start date	End date	Obs.	Avol	SD	Avol	SD
SP	03/01/2000	30/11/2017	4479	0.192	0.047	0.165	0.025
NIKKEI	04/01/2000	30/11/2017	4343	0.243	0.067	0.167	0.019
TSE	02/05/2002	30/11/2017	3894	0.165	0.035	0.117	0.014
FTSE	04/01/2000	30/11/2017	4502	0.183	0.039	0.144	0.015
DAX	03/01/2000	30/11/2017	4535	0.236	0.062	0.208	0.031
EUSTOXX	03/01/2000	30/11/2017	4540	0.233	0.056	0.201	0.032

Avol, annualized volatility; SD, standard deviation

changes are also needed, when results are ineligible due to the rebasing of indices. We use the daily closing prices, P_t^C, to form the daily returns as follows: $r_t = \ln(P_t^C) - \ln(P_{t-1}^C)$, and two realized measures as drawn from the library: the realized kernel and the 5-min realized variance. The estimation results using the two alternative measures are very similar, so we present only the ones with the realized kernels (the results for the realized variances are available upon request).

3.2 Realized measures

The library's realized measures are calculated in the way described in Shephard and Sheppard (2010). The realized kernel, which we present in our analysis here, is chosen as a measure more robust to noise, where the exact calculation with a Parzen weight function is described as follows: $RK_t = \sum_{k=-H}^{H} k(h/(H+1))\gamma_h$, where $k(x)$ is the Parzen kernel function with $\gamma_h = \sum_{j=|h|+1}^{n} x_{j,t} x_{j-|h|,t}$; $x_{jt} = X_{t_{j,t}} - X_{t_{j-1,t}}$ are the 5-min 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) declared that they selected the bandwidth of H as in Barndorff-Nielsen et al. (2009).

The 5-min realized variance, RV_t, which we also employ as an alternative realized measure, is calculated with the formula: $RV_t = \sum x_{j,t}^2$ (results are not reported but they are available upon request). Heber et al. (2009) additionally implement a subsampling procedure from the data to the most feasible level in order to eliminate the stock market noise effects. The subsampling involves averaging across many realized variance estimations from different data subsets (see also the references in Shephard and Sheppard 2010 for realized measures surveys, noise effects and subsampling procedures).

Table 1 presents the main six stock indices extracted from the database and provides volatility estimations for each one's squared returns and realized kernels time series for the respective sample period (see also the SP series graphs in the "Appendix", Figs. 3, 4 and 5). We calculate the standard deviation of the series and the annualized volatility. Annualized volatility is the square rooted mean of 252 times the squared return or the realized kernel. The standard deviations are always lower than the annualized volatilities. The realized kernels have lower annualized volatilities and standard deviations than the squared returns since they ignore the overnight effects and are affected by less noise. The returns represent the close-to-close yield and the realized kernels the open-to-close variation. The annualized volatility of the realized measure is between 12 and 21%, while the squared returns show figures from 17 to 24%.

4 Estimation results

Building upon the introduction of the GARCH-X process by Engle (2002b) to include realized measures as exogenous regressors in the conditional variance equation, Han and Kristensen (2014) and Han (2015) studied the asymptotic properties of this new specification with a fractionally integrated (nonstationary) process included as covariate. Moreover, Pedersen and Rahbek (2019) developed likelihood-ratio tests on the significance of the nonstationary covariate in the above mentioned model, while Halunga and Orme (2009) provided some asymmetry and nonlinearity tests. Lastly, Nakatani and Teräsvirta (2009) and Pedersen (2017) focused on the multivariate case, the so called extended constant conditional correlation, which allows for volatility spillovers and they developed inference and testing for the quasi-maximum likelihood estimator (QMLE) parameters (see also Ling and McAleer 2003 for the asymptotic theory of vector ARMA-GARCH processes). Within the HEAVY framework we first estimate the benchmark formulation as in Shephard and Sheppard (2010), that is, without asymmetries and power transformations, obtaining very similar results (available upon request).

Table 2 presents the estimation results for the chosen asymmetric power specifications. Wald and t tests are used to test the significance of the Heavy and Garch parameters, rejecting the null hypothesis at 10% in all cases. We should highlight the fact that since all the parameters take non-negative values, we use one-sided tests. Following Pedersen and Rahbek (2019), we first test for arch effects and after rejecting the conditional homoscedasticity hypothesis we apply one-sided significance tests of the covariates added to the estimated GARCH-X processes.

For the returns, we statistically prefer the double asymmetric power (DAP) specification since the own estimated power term is $1.18 \le \delta_{rr} \le 1.59$ in all cases (see Table 2, Panel A; see also the Wald tests for the power terms, Table 8, Panel A, in the "Appendix", where the hypotheses of $\delta_{rr} = 1$ and $\delta_{rr} = 2$ are rejected in most cases).

Since when we tried to estimate both power terms simultaneously there was no convergence, we estimated them separately with a two-stage procedure, as follows: We firstly estimate for the realized measure univariate asymmetric power specifications, that is, without the effect from the returns, and, then, in the second stage we use the estimated powers in the returns equations. The sequential procedure produces power term values close to the ones estimated for the realized measure in the respective DAP specification (compare the values of δ_{rR} with those of δ_{RR} in Table 2, Panels A and B).

The Heavy asymmetry parameter, γ_{rR}, is significant and around 0.05 (min. value) to 0.16 (max. value).[2] Although α_{rr} is insignificant and excluded in all cases, the own asymmetry parameter (γ_{rr}) is significant with $\gamma_{rr} \in [0.07, 0.14]$. In other words, the lagged values of both powered variables, that is, the negative signed realized measure and the squared negative returns drive the model of the power transformed conditional variance of the returns. Moreover, the momentum parameter, β_r, is estimated to be around 0.87–0.91. All six indices generated very similar DAP specifications.

Similarly, for the realized measure the most preferred specification is the DAP one, as the estimated power is $\delta_{RR} \in [1.05, 1.28]$ in all cases (see Table 2, Panel B). The Wald tests of the power terms (see Panel B in Table 8) mostly reject the hypotheses of $\delta_{RR} = 1$ and $\delta_{RR} = 2$. The δ_{Rr} is the power term of the returns estimated separately using the univariate specification,

---

[2] The DAP-HEAVY-r equation for returns is also estimated with the direct Heavy effect from the power transformed realized measure, α_{rR}, instead of the Heavy asymmetry, γ_{rR} (these results are available upon request).

Table 2 The DAP-HEAVY model. Estimation results

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX

Panel A: Stock returns

$$(1 - \beta_r L)(\sigma_{rt}^2)^{\frac{\delta_{rr}}{2}} = \omega_r + \gamma_{rr} s_{t-1} L(r_t^2)^{\frac{\delta_{rr}}{2}} + \gamma_{rR} s_{t-1} L(RM_t)^{\frac{\delta_{rR}}{2}}$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_r	**0.87** (73.46)***	**0.87** (51.51)***	<u>0.91</u> (72.00)***	0.88 (50.10)***	0.90 (73.10)***	0.88 (66.48)***
γ_{rr}	0.08 (2.24)***	**0.07** (1.47)*	0.08 (1.51)*	<u>0.14</u> (1.80)**	0.10 (2.61)***	0.11 (2.68)***
γ_{rR}	0.14 (2.72)***	0.07 (1.46)*	**0.05** (1.48)*	0.06 (1.93)***	<u>0.16</u> (2.28)***	**0.05** (1.75)**
δ_{rr}	1.39 (14.71)***	<u>1.59</u> (8.46)***	1.19 (7.37)***	1.46 (11.54)***	**1.18** (11.46)***	<u>1.59</u> (13.91)***
δ_{rR}	1.25	1.26	1.08	<u>1.33</u>	1.32	1.29

Panel B: Realized measure

$$(1 - \beta_R L)(\sigma_{Rt}^2)^{\frac{\delta_{RR}}{2}} = \omega_R + (\alpha_{RR} + \gamma_{RR} s_{t-1}) L(RM_t)^{\frac{\delta_{RR}}{2}} + \gamma_{Rr} s_{t-1} L(r_t^2)^{\frac{\delta_{Rr}}{2}}$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_R	0.66 (71.64)***	0.67 (14.49)***	0.66 (33.41)***	**0.64** (19.79)***	0.67 (21.05)***	<u>0.69</u> (16.31)***
α_{RR}	0.25 (18.83)***	0.26 (6.47)***	0.28 (14.47)***	<u>0.31</u> (9.33)***	0.26 (8.35)***	**0.24** (4.61)***
γ_{RR}	<u>0.07</u> (4.26)***	0.02 (1.49)*	0.02 (1.89)**	**0.001** (1.53)*	0.02 (1.99)***	0.03 (1.51)*
γ_{Rr}	<u>0.33</u> (6.23)***	0.23 (1.52)*	0.11 (7.12)***	**0.10** (1.86)**	0.11 (2.26)***	0.32 (1.72)**
δ_{RR}	1.16 (56.23)***	1.21 (5.48)***	**1.05** (67.37)***	<u>1.28</u> (6.09)***	1.22 (6.01)***	1.25 (7.60)***
δ_{Rr}	1.45	1.56	1.22	1.44	1.27	<u>1.64</u>

The numbers in parentheses are t-statistics
Bold (underlined) numbers indicate minimum (maximum) values across the six indices
***, **, *Significance at the 0.05, 0.10, 0.15 level respectively

as discussed previously. Both Heavy parameters, α_{RR} and γ_{RR}, are significant: α_{RR} is around 0.24 (min. value) to 0.31 (max. value), while γ_{RR}, is between 0.001 and 0.07. Moreover, the cross asymmetry Garch parameter is always significant with $\gamma_{Rr} \in [0.10, 0.33]$. This means that the power transformed conditional variance of $\widetilde{RM}_t$ is significantly affected by the lagged values of both powered variables: squared negative returns and realized measure. Lastly, the momentum parameter, β_R, is estimated to be around 0.64–0.69.

Overall, our results show strong Heavy effects (captured by the γ_{rR}, α_{RR} and γ_{RR} parameters), as well as asymmetric Garch influences (as the estimated γ_{rr} and γ_{Rr} are significant).

5 Long memory

In this section we extend the DAP-HEAVY framework by incorporating long memory.

5.1 Hyperbolic specification

First, we present the most general hyperbolic (HY) specification (see, for example, in the context of a univariate GARCH model Davidson 2004; Dark 2005, 2010 and Schoffer 2003):

$$(1 - \beta_r L)\left[(\sigma_{rt}^2)^{\frac{\delta_{rr}}{2}} - \omega_r\right] = A_r(L)(1 + \gamma_{rr} s_t)\left(r_t^2\right)^{\frac{\delta_{rr}}{2}} + (\alpha_{rR} + \gamma_{rR} s_{t-1})L(RM_t)^{\frac{\delta_{rR}}{2}},$$

$$(1 - \beta_R L)\left[(\sigma_{Rt}^2)^{\frac{\delta_{RR}}{2}} - \omega_R\right] = A_R(L)(1 + \gamma_{RR} s_t)(RM_t)^{\frac{\delta_{RR}}{2}} + (\alpha_{Rr} + \gamma_{Rr} s_{t-1})L\left(r_t^2\right)^{\frac{\delta_{Rr}}{2}},$$

$$(4)$$

with

$$A_i(L) = (1 - \beta_i L) - (1 - \phi_i L)[(1 - \zeta_i) + \zeta_i(1 - L)^{d_i}], \quad i = r, R,$$

where $|\phi_i| < 1$, d_i, is the fractional differencing parameter with $0 \leq d_i \leq 1$, and ζ_i, is the amplitude or hyperbolic parameter with $0 \leq \zeta_i \leq 1$. In other words, we have three long memory parameters, ϕ_i, ζ_i, and d_i. So now the Heavy parameters are six in total. Similarly, the Garch parameters are six.

If $\zeta_i = 0$ and $\phi_i - \beta_i = \alpha_{ii}$, the HYDAP specifications reduce to the DAP ones (see Eq. (3)), since in this case $A_i(L) = \alpha_{ii} L$.

The HY specification also nests the fractional integrated (FI) one (see, for example, Baillie et al. 1996; Tse 1998; Karanasos et al. 2004, and Conrad and Karanasos 2006) by imposing the restriction $\zeta_i = 1$. In this case $A_i(L)$, in Eq. (4) becomes

$$A_i(L) = (1 - \beta_i L) - (1 - \phi_i L)(1 - L)^{d_i}. \tag{5}$$

Finally, note that the sufficient conditions of Dark (2005, 2010) for the non-negativity of the conditional variance of a HYAPARCH $(1, d_i, 1)$ specification are: $\omega_i > 0$, $\beta_i - \zeta_i d_i \leq \phi_i \leq \frac{2-d_i}{3}$ and $\zeta_i d_i(\phi_i - \frac{1-d_i}{2}) \leq \beta_i(\phi_i - \beta_i + \zeta_i d_i)$, $i = r, R$ (see also Conrad 2010). When $\zeta_i = 1$ they reduce to the ones for the FIGARCH $(1, d_i, 1)$ model (see Bollerslev and Mikkelsen 1996).

5.2 Estimated model

We further extend the HEAVY framework by incorporating long memory. For the returns the chosen specification is the FIOAP, whereas for the realized measure we select the HYDAP one. In all cases the power terms of the covariates are presented as fixed parameters since they are estimated separately using univariate models.

In the FIOAP specification for the returns (see Table 3, Panel A) δ_{rr} is around 1.27–1.57 and d_r close to 0.50 (around 0.42–0.49). In most cases the Wald tests (see Table 9) reject the null hypotheses of $d_r = 0$ or 1 and $\delta_{rr} = 1$ or 2. The other two long memory parameters, ϕ_r and the hyperbolic one, ζ_r, were insignificant and, therefore, they were excluded. The own asymmetry parameter, γ_{rr}, is significant and around 0.26–0.60. The Heavy parameter, α_{rR}, is significant as well and with estimated values between 0.04 and 0.15. In other words, the lagged values of both powered variables (the squared returns and the realized measure) drive the model of the power transformed conditional variance of returns.

In the HYDAP specification for the realized measure (see Table 3, Panel B) δ_{RR} is around 1.13–1.43. There is also strong evidence of hyperbolic memory as not only d_R but also ζ_R is significant, with estimated values 0.57–0.68 and 0.86–0.92, respectively, while the Wald tests always reject the null of either a FIDAP ($H_0 : \zeta_R = 1$) or a DAP formulation ($H_0 : \zeta_R = 0$). The own and the cross (Garch) asymmetric parameters, $\gamma_{RR} \in [0.06, 0.24]$ and $\gamma_{Rr} \in [0.08, 0.22]$, are also significant. This means that the power transformed conditional variance of $\widetilde{RM_t}$ is significantly affected by the lagged values of both powered variables: realized measure and squared negative returns.

Table 3 The HYDAP-HEAVY model. Estimation results

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX

Panel A: Stock returns, FIOAP specification

$$(1 - \beta_r L)[(\sigma_{rt}^2)^{\frac{\delta_{rr}}{2}} - \omega_r] = \left[(1 - \beta_r L) - (1 - L)^{d_r}\right](1 + \gamma_{rr} s_t)(r_t^2)^{\frac{\delta_{rr}}{2}} + \alpha_{rR} L (RM_t)^{\frac{\delta_{rR}}{2}}$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_r	0.43 (3.02)***	0.43 (2.11)***	0.38 (2.77)***	**0.35** (3.98)***	0.39 (3.14)***	0.38 (4.15)***
d_r	0.48 (3.34)***	0.49 (2.20)***	**0.42** (3.11)***	**0.42** (6.20)***	0.46 (6.53)***	0.43 (5.46)***
γ_{rr}	0.60 (1.67)**	**0.26** (1.46)*	0.39 (3.04)***	0.56 (3.22)***	0.46 (4.71)***	0.54 (2.36)***
α_{rR}	0.05 (2.22)***	**0.04** (2.36)***	0.07 (2.83)***	0.07 (3.01)***	0.14 (1.90)**	0.15 (1.51)*
δ_{rr}	1.36 (2.17)***	1.57 (2.67)***	**1.27** (7.69)***	1.41 (6.63)***	1.41 (10.17)***	1.46 (5.64)***
δ_{rR}	1.24	1.21	1.13	1.25	1.44	1.50

Panel B: Realized measure, HYDAP specification

$$(1 - \beta_R L)[(\sigma_{Rt}^2)^{\frac{\delta_{RR}}{2}} - \omega_R] = \gamma_{Rr} s_{t-1} L (r_t^2)^{\frac{\delta_{Rr}}{2}} + (1 - \phi_R L)[(1 - \zeta_R) +$$
$$\zeta_R (1 - L)^{d_R}](1 + \gamma_{RR} s_t)(RM_t)^{\frac{\delta_{RR}}{2}}$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_R	0.58 (15.08)***	**0.54** (7.55)***	0.56 (11.65)***	**0.54** (10.76)***	0.68 (16.38)***	0.68 (13.18)***
ϕ_R	0.30 (6.65)***	0.27 (5.93)***	0.32 (7.20)***	**0.24** (8.97)***	0.37 (7.60)***	0.35 (7.26)***
ζ_R	**0.86** (77.36)***	0.90 (63.01)***	0.87 (70.16)***	0.92 (90.13)***	0.89 (88.20)***	0.87 (54.09)***
d_R	0.62 (17.01)***	0.63 (15.76)***	**0.57** (19.92)***	0.65 (23.37)***	0.68 (19.33)***	0.68 (16.95)***
γ_{RR}	0.24 (5.02)***	0.08 (1.83)**	0.12 (2.45)***	**0.06** (3.21)***	0.12 (2.84)***	0.13 (1.95)**
γ_{Rr}	0.22 (11.05)***	0.19 (6.93)***	0.11 (8.65)***	0.11 (9.66)***	**0.08** (10.36)***	0.10 (4.36)***
δ_{RR}	**1.13** (7.31)***	1.24 (7.04)***	1.17 (11.24)***	1.29 (3.88)***	1.43 (7.25)***	1.43 (8.21)***
δ_{Rr}	1.33	1.49	1.25	1.42	1.38	1.39

See notes in Table 2

All in all, our long memory extension of the asymmetric power specification proves once more that both powered conditional variances receive the notable impact from the first lags of both power transformed variables. Intriguingly, this result stands in sharp contrast to the benchmark HEAVY model, where the intradaily realized measure is not affected by squared daily returns and the daily returns conditional variance is only determined by the lagged realized measure and the lagged returns variance since the asymmetries from negative returns are completely neglected.

6 Forecasting performance

Following the estimation of all possible extensions to the HEAVY framework of equations, we perform multistep-ahead out-of-sample forecasting in order to compare the forecasting

Table 4 Root mean square error (RMSE) of m-step ahead forecasts for SP as a Ratio of the benchmark model

Specifications ↓	Restrictions ↓ m-steps → :	1	5	20	120	231
Panel A: Stock returns						
Benchmark HEAVY	$\alpha_{rr} = \gamma_{rr} = \gamma_{rR} = 0$ $\delta_{rj}=2(\,j=r,R)$, in Eq. (3)	1.000	1.000	1.000	1.000	1.000
DAP-HEAVY	$\alpha_{ri} = 0$ $(i=r,R)$, in Eq. (3)	0.907	**0.861**	**0.878**	**0.979**	**0.989**
FIOAP-HEAVY	$\phi_r = \varsigma_r = \gamma_{rR} = 0$ in Eq. (4)	**0.626**	1.256	1.619	1.456	1.626
Panel B: Realized measure						
Benchmark HEAVY	$\alpha_{Rr} = \gamma_{Rr} = \gamma_{RR} = 0$ $\delta_{RR}=2$, in Eq. (3)	1.000	1.000	1.000	1.000	1.000
DAP-HEAVY	$\alpha_{Rr} = 0$ in Eq. (3)	**0.884**	0.612	0.386	0.156	0.126
HYDAP-HEAVY	$\alpha_{Rr} = 0$ in Eq. (4)	0.893	**0.396**	**0.245**	**0.086**	**0.070**

Bold numbers indicate minimum values across the different specifications

accuracy of the different specifications proposed in this study. We re-estimate the benchmark model, the DAP process and its long memory extension, for the shortened sample from 3/1/2000 up to 30/12/2016 (4,248 observations: in-sample estimation) and keep the remaining 231 observations from 3/1/2017 to 30/11/2017 for out-of-sample comparison purposes. With the shortened sample, for each specification we estimate the 1-, 5-, 20-, 120- and 231-step-ahead forecasted (power transformed) conditional variances and calculate the standard measure of forecasting performance, that is the Root Mean Square Error (RMSE) based on the out-of-sample observations up to 30/11/2017.

The results, presented in Table 4 for the SP index (similar forecasting results for the other five indices available upon request), clearly show the preference for our extensions over the benchmark models across all time horizons. Regarding the returns equations (see Panel A), the FIOAP is the best performing specification in the 1-day forecasted variance, while for the one-week and up to the eleven-month-forecasts the DAP formulation dominates the alternative extensions with the lowest RMSE. In the realized measure equation, we get the best 1-step-ahead forecasting performance from the DAP specification (see Panel B). For the remaining time horizons the preferred formulation is the Hyperbolic one.

Overall, the extensions of the HEAVY specification proposed in our study perform significantly better than the benchmark one in the short- and the long-term horizons.

Value at risk application

The forecasting performance of the proposed models can be further examined in a real-world business operation. Value-at-Risk (VaR) is a daily metric for market risk measurement, defined as the potential loss in the value of a portfolio, over a pre-defined holding period, for a given confidence level. The most important input in the VaR calculation is the 1-day volatility forecast of the risk factor relevant to the trading portfolio under scope. We directly apply our conditional variance forecasts in a long position to a portfolio of 10,000 S&P 500 index contracts starting from 30/12/2016. We calculate 231 daily VaR values from 3/1/2017 to 30/11/2017 using the 1-day conditional variance forecasts of each model for returns and realized measure (6 models in total). Given that the conditional mean return is zero and the returns follow the normal distribution, we first calculate the 1-day VaR with a 99% and 95%

Table 5 VaR backtesting for the SP portfolio. No. of exceptions of the 99% and 95% VaR

Specifications	99% VaR	95% VaR
Panel A: Stock returns		
Benchmark HEAVY	2	6
DAP-HEAVY	2	3
FIOAP-HEAVY	1	3
Panel B: Realized measure		
Benchmark HEAVY	2	5
DAP-HEAVY	1	2
HYDAP-HEAVY	1	2

confidence level. According to the parametric approach to VaR calculation, we multiply the daily portfolio value by the 1-day conditional volatility forecast (equal to the square root of the conditional variance forecast) and the left quantile at the respective confidence level of the normal distribution (the z-scores for 99% and 95% confidence level are 2.326 and 1.645, respectively). Secondly, we calculate the daily realized return of the portfolio (gains and losses) and, thirdly, we perform the backtesting exercise comparing the 231 realized returns to the respective 1-day VaR for the 99% and 95% confidence levels. In the cases where the realized loss exceeds the respective day's VaR value, we record it as an exception in the backtesting procedure, meaning that the VaR metric fails to cover the loss of the specific day's portfolio value.

Table 5 reports the backtesting results that prove the superiority of the long memory asymmetric power HEAVY models in forecasting the 1-day-ahead volatility since they give fewer exceptions than the VaR estimates of the benchmark models for both confidence levels applied. The robustness check of the VaR model through the backtesting exercise results in the critical number of exceptions which should be, first, in line with the selected confidence level (the 99% and 95% confidence levels allow for 1 and 5 exceptions, respectively, every 100 days) and, second, the lowest possible. The higher number of exceptions means higher market risk capital requirements for financial institutions. This is the way regulators heavily penalize banks' internal VaR models that do not cover trading losses.

7 Structural breaks

In this section, we investigate the impact of structural changes (detected in the two time series used) on the Heavy and Garch estimated parameters. The time-varying behavior of these parameters can be significant, specifically around a financial crisis break, indicative of the crisis effects on the volatility pattern. As an alternative to the long memory specification, we incorporate structural break dummies in the DAP specification. We identify the structural breaks in the two volatility series for SP, focusing mainly on the recent global financial crisis, and study their impact on the HEAVY framework. We test for structural breaks by employing the methodology in Bai and Perron (1998, 2003a, b), who address the problem of testing for multiple structural changes in a least squares context and under very general conditions on the data and the errors. In addition to testing for the presence of breaks, these statistics identify the number and location of multiple breaks. So, we identify the structural breaks in the two powered series (power transformations of squared returns and realized measure) with the Bai and Perron methodology (see Table 6 and Figs. 1, 2). We use the breaks of the two series in

Table 6 The break dates for SP

	1st break	2nd break	3rd break
r	**02/04/2003**	**31/10/2007**	**30/11/2011**
R	11/04/2003	23/07/2007	13/09/2010

Bai and Perron breaks identification: results selected from the repartition procedure for 1% significance level with 5 maximum number of breaks and 0.15 trimming parameter

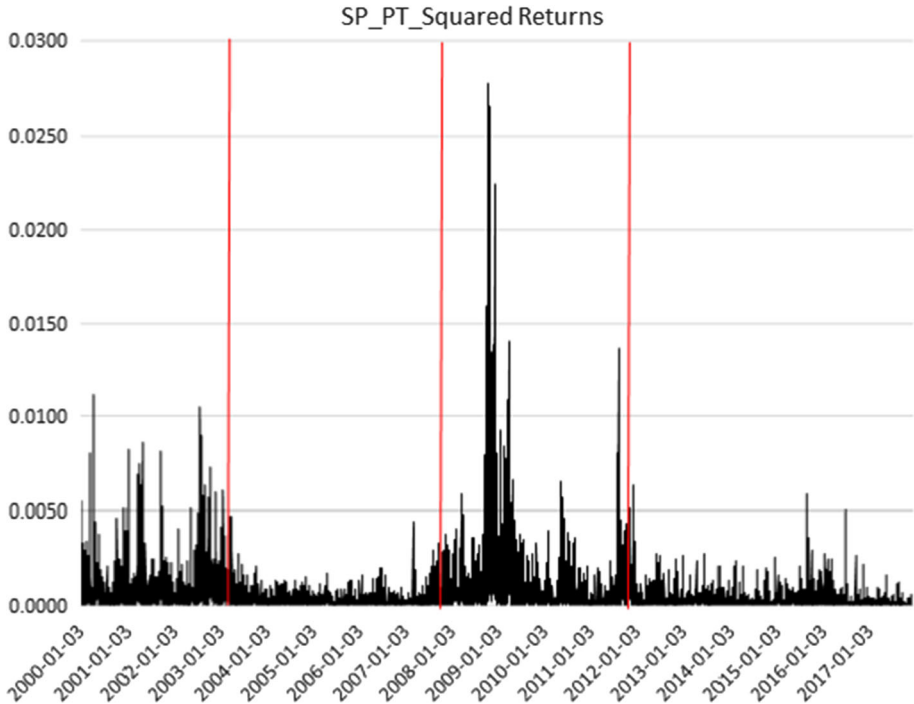Dates in bold indicate that the corresponding dummy coefficient is used in the DAP-HEAVY model

Fig. 1 SP power transformed squared returns with breaks

order to build the slope dummies for the various parameters. We observe that a break date for the recent financial crisis of 2007/2008 is detected, so that we can focus on the crisis effect. We also detect one break date before and one after the crisis.

We present the estimation results for the SP index in Table 7 (similar results for the other five indices available upon request), where we choose to use the 3 breaks of the power transformed squared returns series: (1) 02/04/2003: pre-crisis break, (2) 31/10/2007: crisis break and (3) 30/11/2011: post-crisis break. The three dummy variables multiplied by the respective Heavy and Garch parameters (to construct the slope dummies) are defined as follows: $D_{i,t} = 0$, if $t < T_i$ and $D_{i,t} = 1$, if $t \geq T_i$, $i = (1), (2), (3)$ the three break dates. In the returns equation, the own asymmetry (Garch) parameter, γ_{rr}, receives a decreasing impact (-0.04) from the pre-crisis break, while the cross asymmetry (Heavy) parameter is increased by the crisis dummy $(+0.05)$ and decreased by the post-crisis dummy (-0.04). Regarding the realized measure equation, the Heavy impact, as captured by the own asymmetry γ_{RR},
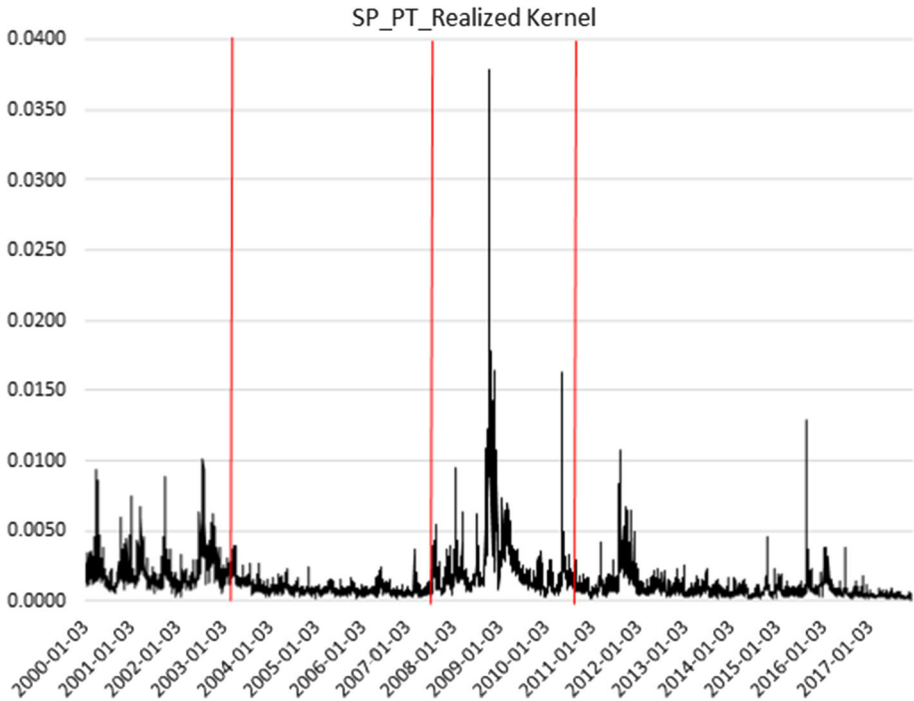
Fig. 2 SP power transformed realized kernel with breaks

Table 7 The DAP-HEAVY model with structural breaks for SP. Estimation results

Panel A: Stock returns

$$(1 - \beta_r L)(\sigma_{rt}^2)^{\frac{\delta_{rr}}{2}} = \omega_r + (\gamma_{rr} + \gamma_{rr}^{(1)} D_{1,t-1}) s_{t-1} L(r_t^2)^{\frac{\delta_{rr}}{2}} + (\gamma_{rR} + \gamma_{rR}^{(2)} D_{2,t-1}$$
$$+\gamma_{rR}^{(3)} D_{3,t-1}) s_{t-1} L(RM_t)^{\frac{\delta_{rR}}{2}}$$

β_r	γ_{rr}	$\gamma_{rr}^{(1)}$	γ_{rR}	$\gamma_{rR}^{(2)}$	$\gamma_{rR}^{(3)}$	δ_{rr}	δ_{rR}
0.86	0.10	−0.04	0.13	0.05	−0.04	1.39	1.25
(70.92)***	(2.23)***	(−1.83)**	(2.27)***	(2.52)***	(−2.46)***	(12.57)***	

Panel B: Realized measure

$$(1 - \beta_R L)(\sigma_{Rt}^2)^{\frac{\delta_{RR}}{2}} = \omega_R + [\alpha_{RR} + \alpha_{RR}^{(3)} D_{3,t-1} + (\gamma_{RR} + \gamma_{RR}^{(2)} D_{2,t-1}) s_{t-1}] L(RM_t)^{\frac{\delta_{RR}}{2}}$$
$$+(\gamma_{Rr} + \gamma_{Rr}^{(1)} D_{1,t-1}) s_{t-1} L(r_t^2)^{\frac{\delta_{Rr}}{2}}$$

β_R	α_{RR}	$\alpha_{RR}^{(3)}$	γ_{RR}	$\gamma_{RR}^{(2)}$	γ_{Rr}	$\gamma_{Rr}^{(1)}$	δ_{RR}	δ_{Rr}
0.65	0.25	−0.06	0.04	0.06	0.36	−0.08	1.19	1.45
(26.94)***	(10.91)***	(−2.43)***	(1.97)***	(2.17)***	(2.66)***	(−1.70)**	(12.09)***	

See notes in Table 2. Superscripts in parentheses indicate the break date

rises with the crisis break, and the Heavy parameter α_{RR} falls post-crisis, whereas the Garch asymmetric influence (captured by γ_{Rr}) falls after the pre-crisis break.

Overall, our finding is that the dummy parameters corresponding to the 2003 and 2011 breaks are negative, whereas the ones for the 2007/2008 crisis are positive.

8 Dynamic correlations

Lastly, we estimated the bivariate system of the extended HEAVY models (Eqs. (3)–(4)) with four alternative correlation models: the Constant Conditional Correlations (CCC) model of Bollerslev (1990), the Dynamic Conditional Correlations (DCC) specification of Engle (2002a), the Asymmetric DCC process (see Cappiello et al. 2006) and the Dynamic Equicorrelations (DECO) model of Engle and Kelly (2012).

The conditional covariance matrix for the 2-dimensional vector $\mathbf{r}_t = [r_t, \widetilde{RM}_t]'$, $\mathbb{V}ar(\mathbf{r}_t \,|F_{t-1}) = \mathbf{H}_t$, is given by:

$$\mathbf{H}_t = \begin{bmatrix} \sigma_{rt}^2 & \sigma_{rR,t} \\ \sigma_{rR,t} & \sigma_{Rt}^2 \end{bmatrix},$$

where σ_{it}^2, $i = r, R$, are the conditional variances (see Eq. (2)), and $\sigma_{rR,t}$ denotes the conditional covariance, that is $\sigma_{rR,t} = Cov(r_t, \widetilde{RM}_t\,[F_{t-1}])$. $\mathbf{H}_t$ can be written as:

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t,$$

while $\mathbf{D}_t$ is a diagonal matrix of standard deviations, that is $\mathbf{D}_t = dg(\mathbf{H}_t^{\frac{1}{2}})$, and $\mathbf{R}_t$ is the conditional correlation matrix with unit diagonal elements and off-diagonal elements given by $\rho_{rR,t} = \sigma_{rR,t}/\sigma_{rt}\sigma_{Rt}$. In our HEAVY model in Eqs. (3)–(4) we, initially, assumed that the conditional covariances and dynamic correlations are zero: $\rho_{rR,t} = \sigma_{rR,t} = 0$ for all t. This implies that, $\mathbf{R}_t = \mathbf{I}_2$ (the identity matrix) and $\mathbf{H}_t$ is a diagonal matrix ($\mathbf{H}_t = \mathbf{D}_t^2$). Allowing for non-zero conditional correlations does not alter our results because the estimation of various non-zero correlation models-the four alternative specifications, namely the CCC, DCC, ADCC, and DECO-is a two-step procedure, where in the first step the parameters in the $\mathbf{D}_t$ matrix are estimated (using Eqs. (3)–(4)), while the second step consists of estimating the (off-diagonal) parameters in $\mathbf{R}_t$. To see this more explicitly, we present the quasi-likelihood function (QL). But first, note that $\mathbf{r}_t$ can be written as: $\mathbf{r}_t = \mathbf{D}_t\mathbf{e}_t$ or equivalently, $\mathbf{e}_t = \mathbf{D}_t^{-1}\mathbf{r}_t$, where $\mathbf{e}_t = [e_{rt}, e_{Rt}]'$. Then QL is given by

$$QL = QL_1 + QL_2$$
$$= -\underbrace{\sum_{t=1}^{T}(n\log(2\pi) + 2\log|\mathbf{D}_t| + \mathbf{r}_t'\mathbf{D}_t^{-2}\mathbf{r}_t)}_{QL_1} - \underbrace{\sum_{t=1}^{T}(\log|\mathbf{D}_t| + \mathbf{e}_t'\mathbf{R}_t^{-1}\mathbf{e}_t + \mathbf{e}_t'\mathbf{e}_t)}_{QL_2}.$$

Thus in the first-step the parameters of the bivariate (HY)DAP-HEAVY process, see again Eqs. (3)–(4), are estimated using QL_1, and in the second-step we estimate the off-diagonal element in $\mathbf{R}_t$ using the standardized residuals: $\widehat{\mathbf{e}}_t = \widehat{\mathbf{D}}_t^{-1}\mathbf{r}_t$ in QL_2. In all cases, the three alternative dynamic models (DCC, ADCC and DECO) estimate the average conditional correlations for the two volatility measures around 0.80–0.90 similar to the constant correlation values that we get from the CCC model.

All in all, the conditional correlations extension provides identical results for the conditional variance equations and estimates similar correlation levels for all indices formulation (results not reported but are available upon request).

9 Conclusions

Our study examined the HEAVY model and extended it by taking into consideration leverage, power transformations and long memory characteristics. For the realized measure our empirical results favor the most general hyperbolic asymmetric power specification, where the lags of both powered variables—squared negative returns, and realized kernel—move the dynamics of the power transformed conditional variance of the latter. Similarly, modeling the returns with an asymmetric power process, we found that not only the powered realized measure, but the power transformed squared negative returns, as well, help in forecasting the conditional variance of the latter.

The long memory (hyperbolic or fractionally integrated) of volatility, its asymmetric response to negative and positive shocks and its power transformations ensures the superiority of our contribution, which can be implemented in the areas of asset allocation and portfolio selection, as well as in several risk management practices. Further, for the U.S. stock index we proved the forecasting superiority of our extensions over the benchmark HEAVY model through the out-of-sample forecasting across multiple short- and long-term horizons. Finally, the detection of structural breaks and the inclusion of break dummies in the asymmetric power formulation capture the time-varying pattern of the parameters, as the break corresponding to the financial crisis of 2007/2008, in particular, increases the values of the parameters.

Our empirical findings on the nexus between low-frequency daily squared returns and high-frequency intradaily realized measures provide a volatility forecasting framework with major implications for policymakers and market practitioners, from investors, risk and portfolio managers up to financial chiefs, leaving ample room for future research on further HEAVY model extensions. Thereupon, policymakers and market players should use our HEAVY framework to closely watch and forecast financial volatility patterns in the process of devising drastic policies, enforcing the financial system's regulations, deciding on asset allocation, hedging strategies and investment projects. Future research should focus on extending the multivariate HEAVY formulation of Noureldin et al. (2012) with long memory, asymmetries and power transformations, as in the recent study of Dark (2018), who uses a long memory multivariate GARCH model, or Opschoor et al. (2018) within the Generalised Autoregressive Score (GAS) process of Creal et al. (2013).

A Appendix

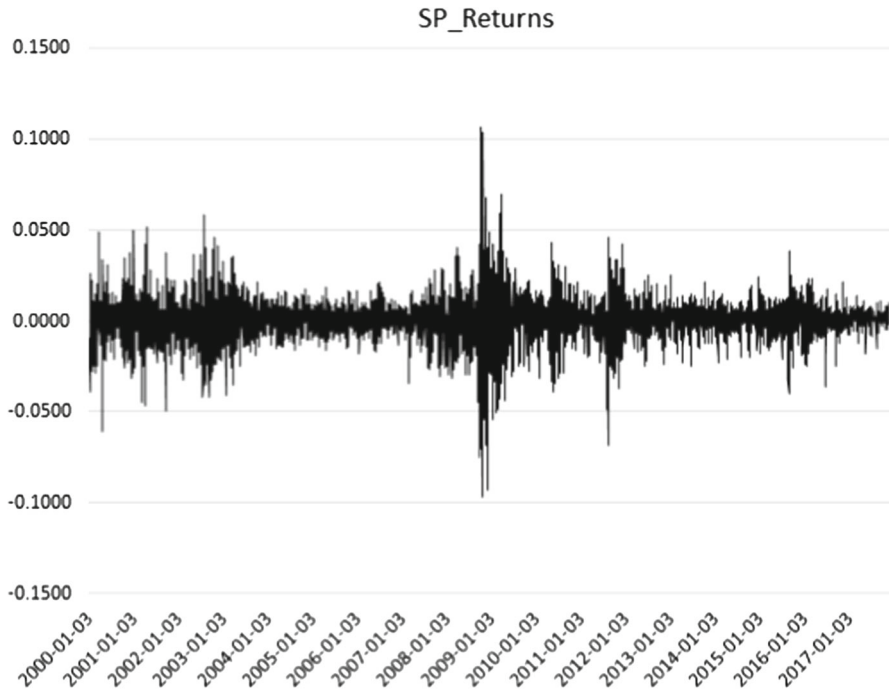### A.1 S&P 500 Graphs

See Figs. 3, 4 and 5

Fig. 3 S&P 500 close-to-close returns
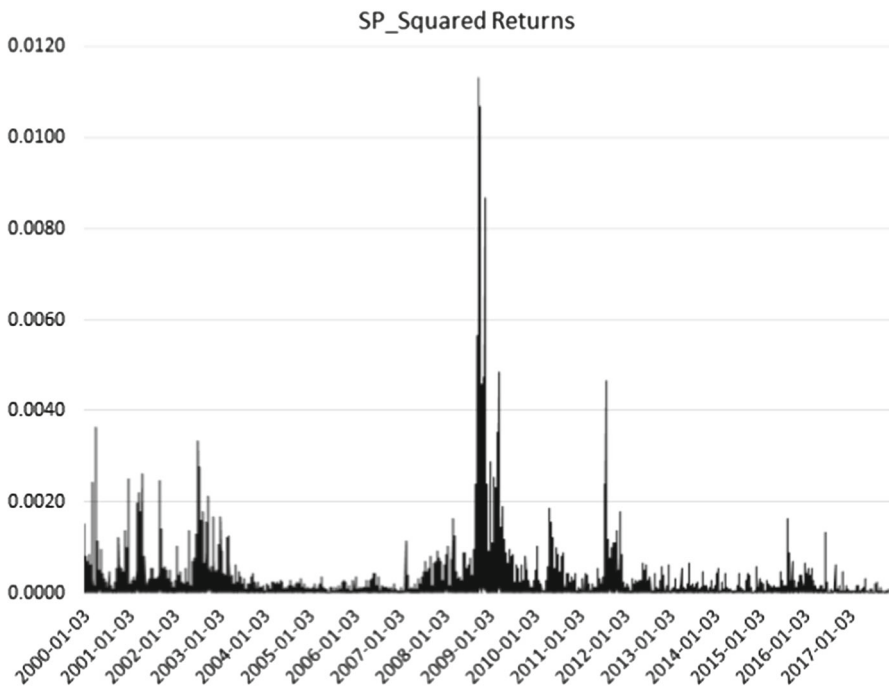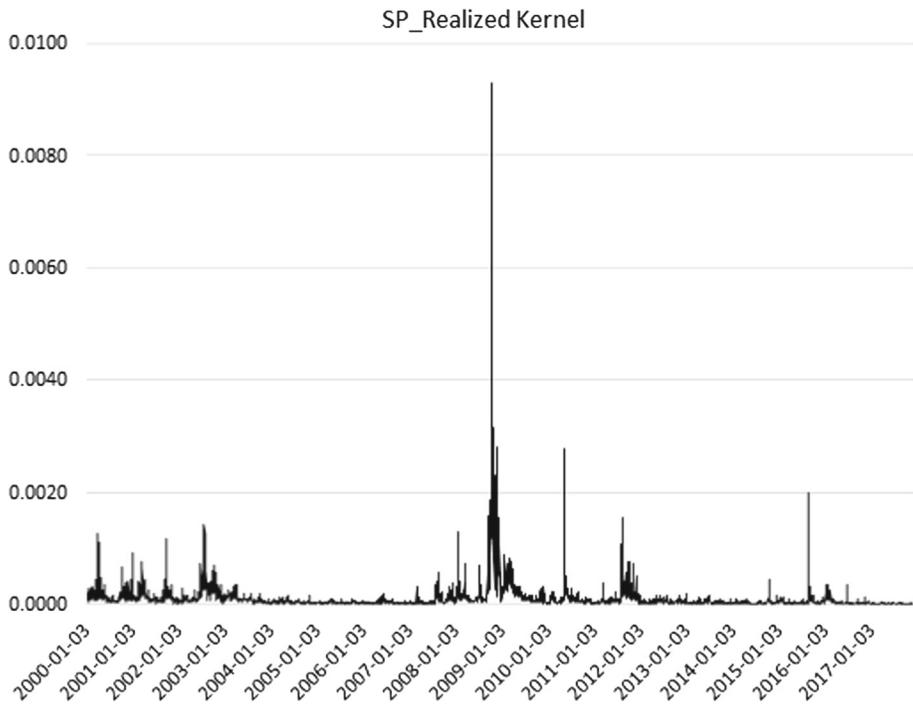

Fig. 4 S&P 500 squared returns

Fig. 5 S&P 500 realized kernel

A.2 Wald tests

See Tables 8, 9.

Table 8 Wald tests for the DAP-HEAVY specification

Restrictions↓	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
Panel A: Stock returns						
$\delta_{rr} = 1$	16.81 [0.00]	9.82 [0.00]	1.40 [0.24]	13.08 [0.00]	2.92 [0.09]	26.90 [0.00]
$\delta_{rr} = 2$	42.43 [0.00]	4.83 [0.03]	25.03 [0.00]	18.54 [0.00]	64.77 [0.00]	12.55 [0.00]
Panel B: Realized measure						
$\delta_{RR} = 1$	60.16 [0.00]	0.91 [0.34]	11.03 [0.00]	1.79 [0.18]	1.19 [0.28]	2.34 [0.13]
$\delta_{RR} = 2$	165.8 [0.00]	12.72 [0.00]	36.88 [0.00]	11.66 [0.00]	14.67 [0.00]	20.64 [0.00]

The Table reports values for ChiSq(1) tests
The numbers in square brackets are *p* values

Table 9 Wald tests for the HYDAP-HEAVY specification

Restrictions↓	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
Panel A: Stock returns						
$\delta_{rr} = 1$	0.33 [0.56]	0.95 [0.33]	2.70 [0.10]	3.65 [0.06]	8.63 [0.00]	3.18 [0.07]
$\delta_{rr} = 2$	10.18 [0.00]	0.53 [0.47]	19.41 [0.00]	7.88 [0.01]	18.46 [0.00]	4.29 [0.04]
$d_r = 0$	11.18 [0.00]	4.85 [0.03]	9.65 [0.00]	38.42 [0.00]	42.62 [0.00]	29.78 [0.00]
$d_r = 1$	12.97 [0.00]	5.62 [0.02]	19.14 [0.00]	75.61 [0.00]	57.94 [0.00]	53.86 [0.00]
Panel B: Realized measure						
$\delta_{RR} = 1$	3.02 [0.08]	0.44 [0.51]	7.97 [0.00]	4.92 [0.03]	4.59 [0.03]	5.00 [0.03]
$\delta_{RR} = 2$	179.2 [0.00]	15.23 [0.00]	21.71 [0.00]	4.93 [0.03]	6.25 [0.01]	6.04 [0.01]
$d_R = 0$	422.8 [0.00]	292.4 [0.00]	306.1 [0.00]	89.92 [0.00]	202.6 [0.00]	140.1 [0.00]
$d_R = 1$	262.8 [0.00]	183.9 [0.00]	143.9 [0.00]	40.03 [0.00]	126.6 [0.00]	80.65 [0.00]
$\zeta_R = 0$	511.9 [0.00]	402.2 [0.00]	457.9 [0.00]	613.7 [0.00]	592.7 [0.00]	241.1 [0.00]
$\zeta_R = 1$	15.25 [0.00]	14.35 [0.00]	19.90 [0.00]	3.33 [0.07]	5.04 [0.03]	3.61 [0.06]

See notes in Table 8

References

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*, 885–905.

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.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*, 47–78.

Bai, J., & Perron, P. (2003a). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*, 1–22.

Bai, J., & Perron, P. (2003b). Critical values for multiple structural change tests. *Econometrics Journal*, *6*, 72–78.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*, 3–30.

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.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, *12*, C1–C32.

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, *72*, 498–505.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modelling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*, 151–184.

Borovkova, S., & Mahakena, D. (2015). News, volatility and jumps: The case of natural gas futures. *Quantitative Finance*, *15*, 1217–1242.

Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, *4*, 537–572.

Conrad, C. (2010). Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics*, *157*, 441–457.

Conrad, C., & Karanasos, M. (2006). The impulse response function of the long memory GARCH model. *Economics Letters*, *90*, 34–41.

Conrad, C., & Karanasos, M. (2010). Negative volatility spillovers in the unrestricted ECCC-GARCH model. *Econometric Theory*, *26*, 838–862.

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, *27*, 46–78.

Creal, D. D., Koopman, S. J., & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, *28*, 777–795.

Dark, J. G. (2005). *Modelling the conditional density using a hyperbolic asymmetric power ARCH model*. Mimeo: Monash University.

Dark, J. G. (2010). Estimation of time varying skewness and kurtosis with an application to value at risk. *Studies in Nonlinear Dynamics and Econometrics, 14*, Article 3.

Dark, J. G. (2018). Multivariate models with long memory dependence in conditional correlation and volatility. *Journal of Empirical Finance*, *48*, 162–180.

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.

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.

Engle, R. F. (2002a). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*, *20*, 339–350.

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

Engle, R. F., & Kelly, B. T. (2012). Dynamic equicorrelation. *Journal of Business and Economic Statistics*, *30*, 212–228.

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*, 1779–1801.

Halunga, A. G., & Orme, C. D. (2009). First-order asymptotic theory for parametric misspecification tests of GARCH models. *Econometric Theory*, *25*, 364–410.

Han, H. (2015). Asymptotic properties of GARCH-X processes. *Journal of Financial Econometrics*, *13*, 188–221.

Han, H., & Kristensen, D. (2014). Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *Journal of Business and Economic Statistics*, *32*, 416–429.

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.

Heber, G., Lunde, A., Shephard, N., & Sheppard, K. (2009). *Oxford-Man Institute's (OMI's) realized library, version 0.2*. Oxford: Oxford-Man Institute, University of Oxford.

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

Karanasos, M., Psaradakis, Z., & Sola, M. (2004). On the autocorrelation properties of long-memory GARCH processes. *Journal of Time Series Analysis*, *25*, 265–281.

Ling, S., & McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory*, *19*, 280–310.

Nakatani, T., & Teräsvirta, T. (2009). Testing for volatility interactions in the constant conditional correlation GARCH model. *Econometrics Journal*, *12*, 147–163.

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

Opschoor, A., Janus, P., Lucas, A., & Van Dijk, D. (2018). New HEAVY models for fat-tailed realized covariances and returns. *Journal of Business and Economic Statistics*, *36*, 643–657.

Pedersen, R. S. (2017). Inference and testing on the boundary in extended constant conditional correlation GARCH models. *Journal of Econometrics*, *196*, 23–36.

Pedersen, R. S., & Rahbek, A. (2019). Testing GARCH-X type models. *Econometric Theory*, *35*, 1012–1047.

Pérez, A., Ruiz, E., & Veiga, H. (2009). A note on the properties of power-transformed returns in long-memory stochastic volatility models with leverage effect. *Computational Statistics and Data Analysis*, *53*, 3593–3600.

Schoffer, O. (2003). *HY-A-PARCH: A stationary A-PARCH model with long memory*. Mimeo: University of Dortmund.

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

Tse, Y. K. (1998). The conditional heteroscedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics*, *13*, 49–55.