Stylized Facts for Extended HEAVY/GARCH models and MEM: the importance of asymmetries, power transformations, long memory, structural breaks and spillovers

M. Karanasos[†,*], Y. Xu[‡], S. Yfanti[§]

[†]*Brunel University, London, UK;* [‡] Cardiff University; [§]Lancaster University

This draft (**Incomplete not to be quoted**): December 2017

Abstract

This paper studies and extends the HEAVY model. Our main contribution is the enrichment of the model with asymmetries, power transformations and long memory -fractionally integrated or hyperbolic. The conclusion that the lagged realized measure does all the work at moving around the conditional variance of stock returns, while it holds in the benchmark specification, it does not hold once we allow for asymmetric, power and long memory effects, since we find that the two power transformed conditional variances are significantly affected by the lagged *power transformed* squares of *negative* returns.

Other findings are as follows. First, hyperbolic memory fits the model of the realized measure better, whereas fractional integration is more suitable for modelling the conditional variance of the returns. Second, the augmentation of the HEAVY framework with the Garman-Klass range-based volatility estimator further improves the forecasting accuracy of the volatility process. Third, the structural breaks applied to the trivariate system capture the time-varying behavior of the parameters, in particular during and after the global financial crisis of 2008.

Keywords: Asymmetries, HEAVY and GARCH models, financial crisis, high-frequency data, hyperbolic long memory, MEM, power transformations, realized variance, structural breaks.

JEL Classification Codes: C32; C58; G15; F3

[*]Address for correspondence: Menelaos Karanasos, Economics and Finance, Brunel University London, UB8 3PH, UK; email: menelaos.karanasos@brunel.ac.uk, tel: +44(0)1895265284, fax: +44 (0)1895269770.

1 Introduction

The asset return volatility has attracted major interest of the financial econometrics research, with crucial implications in asset pricing, portfolio selection and risk management practices. Several studies have introduced non-parametric estimators of realized volatility using high-frequency market data, trying to overcome the market microstructure noise contained in the datasets. Andersen and Bollerslev (1998), Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) were the first studies that formalized econometrically the realized variance with quadratic variation-like measures. Hansen and Lunde (2006) studied, amongst others, the effect of market frictions on the measurement of realized volatility and proved the superiority of kernel-based estimators, while Barndorff-Nielsen et al. (2008, 2009) focused on the realized kernel estimation as the realized measure the more robust to noise.

Voluminous empirical evidence on modelling and forecasting the realized volatility is developed. A popular approach broadly used is the ARFIMA time series model of realized variance in its 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) propose a fractionally integrated model with asymmetries named Dually Asymmetric Realized Volatility model (DARV-FI), where the ARFIMA model incorporates leverage effect parameters to measure the higher volatility risk in periods of negative returns. Another popular approach to model the temporal aggregation of realized volatility is the Heterogeneous Autoregressive (HAR-RV) model introduced by Corsi (2009).

Within the GARCH framework, researchers combine the realized volatility with the conditional variance of asset returns. Engle (2002) introduced the GARCHX model of daily returns, where the realized volatility is included as exogenous variable in the conditional variance equation. Corsi et al. (2008) extended the HAR model of realized volatility with a GARCH error process (HAR-GARCH) to model the volatility of realized volatility, in order to account for the time-varying conditional heteroscedasticity of the normally distributed HAR errors and improve its predictive power. Hansen et al. (2012) introduce the Realized GARCH model, which is the most close specification to the HEAVY model, we study in this paper.

Moreover, several studies use directly high frequency returns in the GARCH models instead of incorporating daily realized volatility measures in the daily returns GARCH equation. For example, Chortareas et al. (2011) calculate the 15min returns of Euro exchange rates, estimate intra-daily GARCH and FIGARCH processes and compare them to daily returns GARCH and FIGARCH, as well as to the daily realized volatility ARFIMA model. The intra-daily GARCH models and the ARFIMA realized volatility model perform better than the daily processes. Engle (2000) moves the attention from high-frequency data to ultra-high-frequency (UHF) data, that are irregularly spaced in time, introducing the UHF-

GARCH, where the conditional duration from the Autoregressive Conditional Duration (ACD) model (Engle and Russell, 1998) is incorporated into the GARCH specification. Li et al. (2016) develop the Markov-Switching Autoregressive Conditional Intensity (MS-ACI) model on intraday Trade and Quote (TAQ) stock data by extending and improving the original Autoregressive Conditional Intensity (ACI) model of Russell (1999).

Furthermore, Engle (2002) first introduced the MEM specification for the conditional expectation of non-negative valued time series. MEM nests the GARCH structure with the squared returns series being replaced by any non-negative process. The MEM structure also nests several GARCH-type models for positive valued processes like the ACD model of Engle and Russell (1998) for durations, the Conditional Autoregressive Range (CARR) of Chou (2005) for the price range and the Autoregressive Conditional Volume (ACV) of Manganelli (2005) for the transaction volume. Engle and Gallo (2006) estimate a trivariate MEM for three non-negative series: the squared returns, the high-low range and the realized variance, including cross effects and asymmetries. Cipollini et al. (2009, 2013) estimate multivariate MEMs allowing for interdependence across the terms of the vector representation of the model and formalize the joint probability density function of the vector error term with a copula approach (Cipollini et al., 2009) and a semiparametric approach (Cipollini et al., 2013). Brownlees et al. (2011) propose a further MEM extension, the Component MEM, which incorporates both daily and intra-daily components in the non-negative process modelling. Finally, Gallo and Otranto (2015), following Lanne (2006), focus on the time-varying and asymmetric behaviour of the MEM's parameters in the realized volatility modelling and propose Markov Switching and Smooth Transition parameters in order to capture the volatility regimes with different dynamics.

Following the MEM framework, Shephard and Sheppard (2010) model the realized volatility with a MEM$(1,1)$ equation, the HEAVY-RM. They model also the returns with a GARCH$(1,0)$-X process, the HEAVY-r equation, where the ARCH term is replaced by the lagged realized volatility. The two-equation system, the HEAVY-r and the HEAVY-RM, defines the HEAVY model, which is extended to its multivariate specification by Noureldin et al. (2012). Cipollini et al. (2013) refer to the HEAVY model by simply restricting the bivariate Vector MEM representation for squared returns and realized variance. Lastly, Borovkova and Mahakena (2015) are the first to apply the HEAVY models with different error distributions (student-t and skewed-t). They also extend the HEAVY-r equation with a leverage term, a news sentiment proxy and a time to maturity variable alternatively.

This paper studies and extends the HEAVY framework of Shephard and Sheppard (2010), which jointly models financial volatility based on both daily (squared returns) and intra-daily (realized measure) data, so that the estimated system of equations adopts to information arrival more rapidly than the classic daily GARCH process. This new class of models uses recently developed estimators of *ex-post* volatility of daily stock returns (i.e., the realized variance or the realized kernel). 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 to higher quality performance in volatility level shifts and more reliable forecasts.

The purpose of this paper is to analyze in depth various HEAVY specifications looking at their performance over six stock index returns. Our first contribution is the enrichment of the model with asymmetries and power transformations, through the Asymmetric Power ARCH framework of Ding et al. (1993). We find that both the power transformed lagged squared negative returns and negative signed realized measure affect the power transformed conditional variance of the returns. Similarly, for the realized measure the extended HEAVY with a double asymmetric impact is also the chosen model. In other words, the benchmark formulation of Shephard and Sheppard (2010) ignores both power transformations and asymmetric effects, which are found to be significant.

Moreover, we take into account long memory (either fractionally integrated or hyperbolic), by employing the hyperbolic framework (see, for example, the HYAPARCH model in Dark, 2005, 2010 and Schoffer, 2003, and the HYGARCH in Davidson, 2004). We find that a fractionally integrated own asymmetric power specification better fits the power transformed conditional mean of the squared returns, whereas a hyperbolic type of memory with double asymmetries is more suitable for the power transformed conditional mean of the realized measure. Once the long memory feature is taken into consideration along with asymmetries and power transformations, both the lagged power transformed realized measure and the lagged values of the power transformed squared returns move the dynamics of the power transformed conditional variance of the returns. In other words, the fractionally integrated asymmetric power process for the preceding conditional variance pools information across both low-frequency and high-frequency based volatility indicators. Similarly, the more richly parameterized hyperbolic process for the power transformed conditional variance of the signed square rooted realized measure is bolstered with low-frequency information as well, since the lagged power transformed squares of the negative returns together with the lagged values of the power transformed realized measure help forecast the aforementioned conditional variance.

We further augment the bivariate model with a third volatility variable, namely, the range-based Garman-Klass volatility building a trivariate system. We observe that the power transformed Garman-Klass volatility affects the power transformed conditional variances of both the returns and the signed square rooted realized measure, whereas the Garman-Klass volatility equation is estimated with significant effects from the lagged power transformed realized measure and squared negative returns. Furthermore, we re-estimate the trivariate system, allowing for volatility spillovers. Finally, in the presence of structural breaks, which are apparent in the three volatility measures, we re-estimate the augmented trivariate system including dummy variables, and we evaluate the time-varying behavior of the parameters. Focusing on the recent global financial crisis, we observe that the values of the parameters increase after the crisis.

The remainder of the paper is structured as follows. In Section 2 we detail the benchmark HEAVY formulation and our first extension, which allows for asymmetries and power transformations. Section 3 describes the data and Section 4 presents the results for the asymmetric power specification. In the following Section we study the long memory process and discuss the relevant empirical findings. Section 6 explores some further HEAVY extensions. In Section 6.1 we estimate the augmented (by the Garman-Klass volatility) model, in Section 6.2 we take into consideration the presence of structural breaks and in Section 6.3 we present the results from the trivariate system allowing for volatility spillovers. In Section 7, within the context of an N-dimensional vector asymmetric power HEAVY process, we derive explicit formulas for the optimal predictors of the three power transformed conditional variances and their second moment structure. Finally, Section 8 concludes the analysis.

2 The HEAVY/GARCH Framework

The benchmark HEAVY specification of Shephard and Sheppard (2010) can be extended in many directions. We allow for power transformations of the volatilities, leverage effects and long memory (see Section 5 below) in the conditional variance process. We re-run the estimated benchmark specification of Shephard and Sheppard (2010), enriched with the three key features to improve further the HEAVY/GARCH volatility modelling.

2.1 Asymmetric Power Specifications

The HEAVY/GARCH type of models use two variables: the close-to-close stock returns (r_t) and the open-to-close variation proxied by the realized measure, 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 $r_t \geqslant 0$ and $sign(r_t) = -1$, if $r_t < 0$. Hereafter, we will denote the returns and the SSR realized measure by ε_{rt} and ε_{Rt}, respectively. We assume that they are charcterized by the following relation:

$$\varepsilon_{it} = e_{it}\sigma_{it}$$

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 $\mathcal{F}_{t-1}^{(XF)}$, that 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 GARCH model e_{it} has zero mean and unit variance. Therefore, ε_{it} has zero conditional mean, and its conditional variance is given by $\mathbb{E}(\varepsilon_{it}^2 | \mathcal{F}_{t-1}) = \sigma_{it}^2$. In the MEM $e_{it} > 0$, with unit mean and variance equal to $q_i > 0$. This implies that $\mathbb{E}(\varepsilon_{it} | \mathcal{F}_{t-1}) = \sigma_{it}$ and $\mathbb{V}ar(\varepsilon_{it} | \mathcal{F}_{t-1}) = q_i \sigma_{it}^2$. In other words, the GARCH model for the conditional variance of the returns (or the SSR realized measure), is similar to the MEM for the conditional mean of the squared returns (or the realized measure).[1]

The double asymmetric power (DAP) specification for the HEAVY$(1,1)$-i -or GARCH$(1,1)$- model, $i = r, R$, consists of the following equations (in what follows for notational simplicity we will drop the order of the model if it is $(1,1)$):

$$(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} = \omega_r + (\alpha_{rr} + \gamma_{rr} s_{t-1}) L (\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{rR} + \gamma_{rR} s_{t-1}) L (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}, \tag{1}$$

$$(1-\beta_R L)(\sigma_{Rt}^2)^{\frac{\delta_R}{2}} = \omega_R + (\alpha_{RR} + \gamma_{RR} s_{t-1}) L (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}} + (\alpha_{Rr} + \gamma_{Rr} s_{t-1}) L (\varepsilon_{rt}^2)^{\frac{\delta_r}{2}}, \tag{2}$$

where L is the lag operator, δ_r, $\delta_R \in \mathbb{R}^+$ (the set of the positive real numbers) are the two 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[2]; 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 Section 5 below, and, in particular, eq. (4)). In this model/specification the power transformed (hereafter PT) conditional variance, $(\sigma_{it}^2)^{\delta_i/2}$, is a linear function of the lagged PT squared returns, $(\varepsilon_{r,t-1}^2)^{\frac{\delta_r}{2}}$, and realized measure, $(\varepsilon_{R,t-1}^2)^{\frac{\delta_R}{2}}$.

We will distinguish between three different asymmetric cases. The double one (A: γ_{ii}, $\gamma_{ij} \neq 0$) and two more: own asymmetry (OA: $\gamma_{ii} \neq 0$, $\gamma_{ij} = 0$) and cross asymmetry (CA: $\gamma_{ii} = 0$, $\gamma_{ij} \neq 0$). For notational convenience we will drop the letter D when we have a double asymmetry.

It will be convenient to have labels for the six models/specifications that we estimate (see also Table 1 below). For example, the abbreviation AP HEAVY-E-r (or GARCH-X-r) stands for the model for stock returns in eq. (1) with $\delta_r, \alpha_{rr}, \gamma_{rr}, \alpha_{rR}, \gamma_{rR} \neq 0$. That is, in this model both the lagged PT squared returns and realized measure move the dynamics of the PT conditional variance of the returns.

The simple OAP GARCH process is the one with $\alpha_{rR} = \gamma_{rR} = 0$:

$$(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} = \omega_r + (\alpha_{rr} + \gamma_{rr} s_{t-1}) L (\varepsilon_{rt}^2)^{\frac{\delta_r}{2}},$$

while the CAP HEAVY (or GARCH $(1,0)$-X) process is the one with $\alpha_{rr} = \gamma_{rr} = 0$:

$$(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} = \omega_r + (\alpha_{rR} + \gamma_{rR} s_{t-1})(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}.$$

[1] Engle (2002) first proposed the MEM model using the various GARCH family specifications to estimate the volatility of volatility, which is a non-negative process. He uses the AP-MEM model in his Volatility Laboratory (V-Lab) amongst other processes for real-time financial volatility modelling.

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

The α_{rR} and γ_{rR} parameters will be called the Heavy or ARCH-X parameters, while α_{rr} and γ_{rr} are the ARCH or Heavy-E parameters.

The general model in eq. (1) can be thought of as an extended AP HEAVY-r process with the lagged PT squared returns included as an additional regressor. The name suggests that it is the lagged PT realized measure which does almost all the work at moving around the PT conditional variance of the returns (see Shephard and Sheppard, 2010).

Alternatively, it can be considered as an AP GARCH-X-r process, that is the lagged PT realized measure is used as an additional regressor in the AP GARCH process (see also Engle, 2002). As pointed out by Shephard and Sheppard (2010), the GARCH-X terminology suggests that it is the lagged PT squared returns which drive the model.

Similarly, in the AP HEAVY-E-R (or GARCH-X-R) model (see eq. (2)) for the realized measure the α_{RR} and γ_{RR} parameters will be called the Heavy (or ARCH) parameters, while α_{Rr} and γ_{Rr} are the Heavy-E or ARCH-X parameters. The Table below summarizes the six different models.

Table 1: Six different AP HEAVY/GARCH/MEM models/specifications

PT Squared Returns, $(r_t^2)^{\frac{\delta_r}{2}}$:

$$(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} = \omega_r + (\alpha_{rr} + \gamma_{rr} s_{t-1}) L (\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{rR} + \gamma_{rR} s_{t-1}) L (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}$$

$$\underbrace{\alpha_{rr}, \gamma_{rr}, \alpha_{rR}, \gamma_{rR} \neq 0}_{\substack{\text{DAP HEAVY-E} \\ \text{or GARCH-X} \\ \text{(or MEM-X)}}} \qquad \underbrace{\alpha_{rR} = \gamma_{rR} = 0}_{\substack{\text{OAP GARCH} \\ \text{(or MEM)}}} \qquad \underbrace{\alpha_{rr} = \gamma_{rr} = 0}_{\substack{\text{CAP HEAVY} \\ \text{or GARCH(1,0)-X} \\ \text{[or MEM(1,0)-X]}}}$$

PT Realized Measure, $RM_t^{\frac{\delta_R}{2}}$:

$$(1-\beta_R L)(\sigma_{Rt}^2)^{\frac{\delta_R}{2}} = \omega_r + (\alpha_{Rr} + \gamma_{Rr} s_{t-1}) L (\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{RR} + \gamma_{RR} s_{t-1}) L (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}$$

$$\underbrace{\alpha_{Rr}, \gamma_{Rr}, \alpha_{RR}, \gamma_{RR} \neq 0}_{\substack{\text{DAP HEAVY-E} \\ \text{or GARCH-X} \\ \text{or MEM-X}}} \qquad \underbrace{\alpha_{Rr} = \gamma_{Rr} = 0}_{\substack{\text{OAP HEAVY} \\ \text{or GARCH} \\ \text{(or MEM)}}} \qquad \underbrace{\alpha_{RR} = \gamma_{RR} = 0}_{\substack{\text{CAP GARCH(1,0)-X} \\ \text{[or MEM(1,0)-X]}}}$$

2.1.1 Bivariate Representation

The two AP HEAVY-E (or GARCH-X) models/specifications in eqs. (1) and (2), can be expressed/interpreted as a bivariate system with shocks or unconditional spillovers (in order to distinguish matrices (vectors) from scalars, the former are denoted by upper (lower)-case boldface symbols):

$$(\mathbf{I}_2 - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + (\mathbf{A} + \boldsymbol{\Gamma} s_{t-1}) L \boldsymbol{\varepsilon}_t, \tag{3}$$

where $\mathbf{I}_2$ is the identity matrix and $\mathbf{B}$ is a diagonal matrix of order 2 with nonzero elements β_r, β_R; $\boldsymbol{\sigma}_t = [\ (\sigma_{rt}^2)^{\frac{\delta_r}{2}}, (\sigma_{Rt}^2)^{\frac{\delta_R}{2}}\ ]'$; $\boldsymbol{\omega} = [\omega_r, \omega_R]'$; $\boldsymbol{\varepsilon}_t = [(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}}, (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}\ ]'$; $\mathbf{A}$ and $\boldsymbol{\Gamma}$ are full matrices:

$$\mathbf{A} = \begin{bmatrix} \alpha_{rr} & \alpha_{rR} \\ \alpha_{Rr} & \alpha_{RR} \end{bmatrix}, \ \boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{rr} & \gamma_{rR} \\ \gamma_{Rr} & \gamma_{RR} \end{bmatrix}.$$

In other words, the diagonal elements are the ARCH parameters (α_{rr}, γ_{rr}, α_{RR}, γ_{RR}), whereas the cross diagonal elements (α_{rR}, γ_{rR}, α_{Rr}, γ_{Rr}) are the ARCH-X parameters.

Observe that if $\mathbf{\Gamma}$ is a diagonal matrix, then we only have own asymmetries, while if it is cross diagonal, then we only have cross asymmetries. If $\mathbf{A}$ and $\mathbf{\Gamma}$ are diagonal matrices and $\delta_r, \delta_R = 2$, then the bivariate system reduces to two univariate asymmetric GARCH processes: one for the returns and one for the SSR realized measure. Finally, if $\mathbf{\Gamma} = \mathbf{0}$ and $\delta_r, \delta_R = 2$, then we have the benchmark specification of Shephard and Sheppard (2010).

Next let $\boldsymbol{\alpha}_r(\boldsymbol{\gamma}_r)$ and $\boldsymbol{\alpha}_R(\boldsymbol{\gamma}_R)$ denote the two columns of $\mathbf{A}$ ($\mathbf{\Gamma}$):

$$\boldsymbol{\alpha}_r = \begin{bmatrix} \alpha_{rr} \\ \alpha_{Rr} \end{bmatrix}, \quad \boldsymbol{\alpha}_R = \begin{bmatrix} \alpha_{rR} \\ \alpha_{RR} \end{bmatrix}, \quad \boldsymbol{\gamma}_r = \begin{bmatrix} \gamma_{rr} \\ \gamma_{Rr} \end{bmatrix}, \quad \boldsymbol{\gamma}_R = \begin{bmatrix} \gamma_{rR} \\ \gamma_{RR} \end{bmatrix}.$$

That is, $\boldsymbol{\alpha}_R + \boldsymbol{\gamma}_R$ is the vector with the four Heavy parameters, whereas $\boldsymbol{\alpha}_r + \boldsymbol{\gamma}_r$ is the vector with the four Heavy-E parameters. Thus the bivariate system in eq. (3) can be written as

$$(\mathbf{I}_2 - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + (\alpha_r + \gamma_r s_{t-1})L(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\boldsymbol{\alpha}_R + \boldsymbol{\gamma}_R s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}.$$

If $\boldsymbol{\alpha}_r = \boldsymbol{\gamma}_r = \mathbf{0}$ then we have the simple bivariate AP HEAVY system (with cross asymmetries for the returns and own asymmetries for the SSR realized measure). Later on, we will extend the model by allowing for volatility spillovers (see Section 6.3 below), that is we will examine the case where the $\mathbf{B}$ matrix is full. In Section 7, within the context of an N-dimensional vector AP HEAVY/GARCH/MEM model, we will derive closed form expressions for the optimal predictors of the PT variables and their conditional variances, as well as their unconditional moment structure.

3 Data Description

The various HEAVY/GARCH/MEM models are estimated for six stock indices returns and realized volatilities. According to the analysis in Shephard and Sheppard (2010), the HEAVY formulations improve considerably the volatility modelling by allowing momentum and mean reversion effects and adjusting quickly to the structural breaks in volatility. We first run the benchmark specifications, as in Shephard and Sheppard (2010), for the six indices and then we extend them by adding the features of power transformation of the conditional variances, leverage effects and long memory (see Sections 4 and 5 below) in the volatility process.

We will also extend the bivariate model to a trivariate system by including the Garman-Klass (GK) volatility measure as an additional variable (see Section 6.1). 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 three series (squared returns, realized measure and GK volatility), in Section 6.2 we will incorporate dummies in our empirical investigation. Finally, we will take into consideration volatility spillovers (see Section 6.3).

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, Nikkei 225 from Japan, TSE from Canada, FTSE 100 from the UK, DAX from Germany and Eustoxx 50 from the Eurozone. Our sample covers the period from 03/01/2000 to 01/03/2013 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 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-minute 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 5-minute realized variances are available upon request).

3.1 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 $\gamma_h = \sum_{j=-|h|+1}^{n} x_{jt}x_{j-|h|,t}$; $x_{jt} = X_{t_{j,t}} - X_{t_{j-1,t}}$ 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 5-minute realized variance, RV_t, which we also employ as an alternative realized measure, is calculated with the formula: $RV_t = \sum x_{j,t}^2$. Heber et al. (2009) implement additionally 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 A.1 in the supplementary Appendix 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. 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 13% and 23%, while the squared returns show figures from 18% to 25%.

4 Asymmetric Power Specifications (Stylized Facts)

After running the benchmark HEAVY models/specifications[3], we add asymmetries and power transformations to enrich our HEAVY/GARCH volatility modelling. From the estimated results we choose to present in Table 2, we conclude to the following stylized facts for the asymmetric power specifications.

For the PT squared returns, we statistically prefer the HEAVY-E (or GARCH-X) model with the A[P] specification[4] since the power term is $1.40 \leq \delta_r \leq 1.70$ in all cases (see also the Wald tests of the power terms, in the supplementary Appendix, where the hypotheses of $\delta_r = 1$ and $\delta_r = 2$ are rejected for all six indices); the asymmetric Heavy (or ARCH) parameter, γ_{rR}, is significant and around 0.03 (min. value) to 0.13 (max. value). Although α_{rr} is insignificant and excluded in all cases, the own asymmetry parameter (γ_{rr}) is significant with $\gamma_{rr} \in [0.06, 0.09]$. In other words, not only the PT lagged negative signed realized measure, but also the lagged PT squared negative returns drive the model of the PT conditional variance of returns. Moreover, the momentum parameter, β_r, is estimated to be around 0.87 to 0.91. All six indices generated very similar A[P] specifications.

Similarly, for the realized measure the most preferred model/specification is the A[P] HEAVY-E (or GARCH-X), where we model the PT realized measure, as the estimated power is $\delta_R \in [1.20, 1.50]$ in all cases. The Wald tests of the power terms (see the supplementary Appendix) reject the hypotheses of $\delta_R = 1$ and $\delta_R = 2$. The Heavy parameter, α_{RR}, is significant and around 0.19 (min. value) to 0.29 (max. value), while the asymmetric Heavy-E (or ARCH-X) parameter, γ_{Rr}, is between 0.07 and 0.13. This means that both the PT lagged realized measure and squared negative returns affect significantly the PT conditional variance of $\widetilde{RM}_t$. Lastly, the own asymmetry, γ_{RR}, is significant and around 0.02 to 0.06, while the momentum parameter, β_R, is estimated to be around 0.64 to 0.71.

[3] The benchmark HEAVY models estimated without asymmetries and power transformations result to the HEAVY-r and the HEAVY-R as the models that best describe the two volatility processes (results are available upon request). These are exactly the two models proposed also by Shephard and Sheppard, 2010, to constitute the bivariate HEAVY system.

[4] [P] means that the power parameters are estimated for one equation and for the second are fixed to the values of the first equation.

To sum up, in our first HEAVY/GARCH extension with the inclusion of power transformations and asymmetries, we estimate the HEAVY-E (or GARCH-X) models with $\delta_i \neq 2$ and $\gamma_{ij} \neq 0$, $i = r, R$. The A[P] HEAVY-E is the chosen model/specification for the PT conditional variance of the returns, since the estimated power is significantly different from either 1 or 2, and both the asymmetric Heavy (γ_{rR}) and ARCH (γ_{rr}) parameters are significant. So, it appears that in the original HEAVY-r model of Shephard and Sheppard (2010) the squared returns had no effect on the conditional variance of the returns, because power transformations and the own asymmetric influence were ignored.

Regarding the realized measure, the A[P] HEAVY-E is also the chosen model/specification, since the estimated power is again signifcantly different from either 1 or 2, and both Heavy (or ARCH: α_{RR}, γ_{RR}) parameters are significant as well as the asymmetric Heavy-E or ARCH-X (γ_{Rr}) one. Thus, it is optimal i) to model the PT conditional variance (and not the variance as in the original HEAVY-R model) of the SSR realized measure, which is significantly affected not only by the lagged PT realized measure, but by the lagged PT squared negative returns as well, and ii) to include own asymmetries.

Table 2: A[P] HEAVY-E (or GARCH-X) Specifications

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
			Panel A: Stock Returns			
			$(1-\beta_r L)(\sigma^2_{rt})^{\frac{\delta_r}{2}} = \omega_r + \gamma_{rr} s_{t-1} L(\varepsilon^2_{rt})^{\frac{\delta_r}{2}} + \gamma_{rR} s_{t-1} L(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}}$			
β_r	0.88 (63.35)***	**0.87** (47.33)***	<u>0.91</u> (69.68)***	0.89 (74.02)***	0.88 (55.82)***	0.88 (63.48)***
γ_{rr}	0.07 (4.47)***	<u>0.09</u> (4.68)***	0.07 (5.87)***	<u>0.09</u> (6.97)***	**0.06** (4.05)***	0.08 (5.05)***
γ_{rR}	0.09 (4.70)***	0.09 (3.87)***	0.04 (3.25)***	**0.03** (4.43)***	<u>0.13</u> (4.13)***	0.06 (4.81)***
			Panel B: Realized Measure			
			$(1-\beta_R L)(\sigma^2_{Rt})^{\frac{\delta_R}{2}} = \omega_R + \alpha_{RR}(1+\gamma_{RR} s_{t-1}) L(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}} + \gamma_{Rr} s_{t-1} L(\varepsilon^2_{rt})^{\frac{\delta_r}{2}}$			
β_R	<u>0.71</u> (31.01)***	**0.64** (20.16)***	<u>0.71</u> (26.25)***	0.65 (18.37)***	0.68 (22.07)***	<u>0.71</u> (27.83)***
α_{RR}	**0.19** (9.04)***	0.28 (10.53)***	0.23 (9.39)***	<u>0.29</u> (8.54)***	0.24 (8.23)***	0.20 (8.87)***
γ_{RR}	<u>0.06</u> (4.58)***	0.04 (2.56)***	0.03 (2.48)***	<u>0.06</u> (4.89)***	**0.02** (1.88)**	0.03 (3.34)***
γ_{Rr}	0.13 (10.17)***	0.10 (5.96)***	0.10 (6.27)***	**0.07** (3.50)***	0.09 (10.54)***	<u>0.19</u> (11.03)***
			Powers δ_i			
δ_r	1.50	<u>1.70</u>	**1.40**	1.50	**1.40**	1.50
δ_R	1.40	<u>1.50</u>	**1.20**	**1.20**	1.40	1.30

Notes: The numbers in parentheses are t-statistics.

***, **, * denote significance at the 0.05, 0.10, 0.15 level respectively.

Bold (underlined) numbers indicate minimum (maximum) values across the six indices.

5 Long Memory Extension

After adding asymmetries and power transformations to enrich our HEAVY volatility modelling, we further extend the HEAVY framework with long memory. In this Section we present the most general hyperbolic (HY) specification, that is the HYAP HEAVY-E (or GARCH-X) model (see, for example, the HYAPARCH framework in Dark, 2005, 2010, and Schoffer, 2003):

$$\begin{aligned}(1-\beta_r L)[(\sigma^2_{rt})^{\frac{\delta_r}{2}} - \omega_r] &= A_r(L)(1+\gamma_{rr} s_t)(\varepsilon^2_{rt})^{\frac{\delta_r}{2}} + (\alpha_{rR} + \gamma_{rR} s_{t-1}) L(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}}, \\ (1-\beta_R L)[(\sigma^2_{Rt})^{\frac{\delta_R}{2}} - \omega_R] &= A_R(L)(1+\gamma_{RR} s_t)(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}} + (\alpha_{Rr} + \gamma_{Rr} s_{t-1}) L(\varepsilon^2_{rt})^{\frac{\delta_r}{2}},\end{aligned} \tag{4}$$

with

$$A_r(L) = (1-\beta_r L) - (1-\phi_r L)[(1-\zeta_r) + \zeta_r(1-L)^{d_r}],$$
$$A_R(L) = (1-\beta_R L) - (1-\phi_R L)[(1-\zeta_R) + \zeta_R(1-L)^{d_R}],$$

where $|\phi_r|$, $|\phi_R| < 1$; d_r, d_R are the two long memory parameters: $0 \leq d_r$, $d_R \leq 1$ and ζ_r, ζ_R are the two amplitude parameters: $0 \leq \zeta_r$, $\zeta_R \leq 1$. For example, in the HEAVY-E-R model the HYAP specification has six Heavy parameters: β_R, δ_R, ϕ_R, ζ_R, d_R, and γ_{RR}.

If $\zeta_i = 0$ and $\phi_i - \beta_i = \alpha_{ii}$, $i = r$, R, the HYAP specifications reduce to the AP ones (see eqs. (1)-(2) and footnote 2), since in this case $A_i(L) = \alpha_{ii}L$.

The HY specification also nests the fractional integrated (FI) one (see, for example, the FIAPARCH formulation of Tse, 1998 and the FIGARCH of Baillie et al., 1996) by imposing the restriction $\zeta_i = 1$. In this case $A_i(L)$, $i = r$, R in eq. (4) become

$$A_r(L) = (1-\beta_r L) - (1-\phi_r L)(1-L)^{d_r}, \tag{5}$$
$$A_R(L) = (1-\beta_R L) - (1-\phi_R L)(1-L)^{d_R}.$$

It also nests the symmetric specification by imposing the restriction $\gamma_{ii} = \gamma_{ij} = 0$ which if, in addition $\delta_i = 2$, $i = r, R$, it reduces to the one without power transformations.

Overall for each of the six HEAVY/GARCH/MEM models we estimate nine specifications (see Table 3 below).[5]

Table 3: The HYAP HEAVY specification and eight alternative restricted ones

$$(1-\beta_i L)[(\sigma_{it}^2)^{\frac{\delta_i}{2}} - \omega_i] = A_i(L)(1+\gamma_{ii}s_t)(\varepsilon_{it}^2)^{\frac{\delta_i}{2}} + (\alpha_{ij} + \gamma_{ij}s_{t-1})L(\varepsilon_{jt}^2)^{\frac{\delta_j}{2}},$$
$$A_i(L) = (1-\beta_i L) - (1-\phi_i L)[(1-\zeta_i) + \zeta_i(1-L)^{d_i}],\ i{,}j = r{,}R,\ i \neq j.$$

Restrictions↓→:	$\zeta_i = 0$	FI: $\zeta_i = 1$	HY: $\zeta_i \in (0,1)$
$\delta_i = 2$ and $\gamma_{ij} = 0\ \forall\ i,j$	HEAVY	FI HEAVY	HY HEAVY
P and $\gamma_{ij} = 0\ \forall\ i,j$	P HEAVY	FIP HEAVY	HYP HEAVY
AP: no restrictions (on δ_i, γ_{ij})	AP HEAVY	FIAP HEAVY	HYAP HEAVY

Notes: For the HEAVY-r and GARCH$(1,0)$-X-R models
we estimate only the three power specifications since $A_i(L) = 0$.

The power transformations (captured by δ_i, $i = r, R$), leverage effects (captured by γ_{ij}, i, $j = r$, R)

[5] Shephard and Sheppard (2010) proposed as an extension of the HEAVY-R model the fractional process with leverage effects or Corsi's (2009) long memory HAR structure. They also suggested the use of realized semivariances in the HEAVY formulations, to capture leverage effects or the inclusion of a leverage parameter multiplied with the realized measure as in Engle and Gallo (2006).

and long memory (captured by d_i, ζ_i) are our main contribution to the HEAVY-E model of Shephard and Sheppard (2010), as well as to the GARCH-X model and the MEM of Engle (2002).

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).

5.1 *Bivariate System*

Next, we will present the two expressions in eq. (4) as a bivariate system.

First, define the two matrix polynomials

$$\mathbf{A}(L)=\begin{bmatrix} A_r(L) & \alpha_{rR}L \\ \alpha_{Rr}L & A_R(L) \end{bmatrix}, \ \mathbf{\Gamma}(L)=\begin{bmatrix} A_r(L)\gamma_{rr} & \gamma_{rR}L \\ \gamma_{Rr}L & A_R(L)\gamma_{RR} \end{bmatrix},$$

where $A_i(L)$ are given in eq. (4).

The two HYAP HEAVY-E or GARCH-X processes in eq. (4) can be written in a matrix form as

$$(\mathbf{I} - \mathbf{B}L)[\boldsymbol{\sigma}_t - \boldsymbol{\omega}] = [\mathbf{A}(L) + \mathbf{\Gamma}(L)s_t]\boldsymbol{\varepsilon}_t,$$

where $\mathbf{B}$, $\boldsymbol{\sigma}_t$, $\boldsymbol{\omega}$, s_t and $\boldsymbol{\varepsilon}_t$ are as in eq. (3). As noted earlier, when the HY parameters, ζ_i, are equal to zero and $\phi_i - \beta_i = \alpha_{ii}$, for $i = r, R$, then $A_i(L) = \alpha_{ii}L$, and the above bivariate system reduces to our second AP formulation (see eq. (3) and footnote 2).

Let also $\boldsymbol{\alpha}_i(L)$ and $\boldsymbol{\gamma}_i(L)$, $i = r, R$, denote the two columns of the matrix polynomials $\mathbf{A}(L)$ and $\mathbf{\Gamma}(L)$, respectively. That is, the second and first columns contain the Heavy and Heavy-E lag polynomials, respectively.

Then the bivariate HYAP HEAVY/GARCH system can be written as:

$$(\mathbf{I} - \mathbf{B}L)[\boldsymbol{\sigma}_t - \boldsymbol{\omega}] = [\boldsymbol{\alpha}_r(L) + \boldsymbol{\gamma}_r(L)s_t](\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + [\boldsymbol{\alpha}_R(L) + \boldsymbol{\gamma}_R(L)s_t](\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}.$$

Clearly, if $\boldsymbol{\alpha}_r(L) = \boldsymbol{\gamma}_r(L) = \mathbf{0}$ then the HEAVY-E bivariate model reduces to the simple HEAVY one.

5.2 Stylized Facts

We further extend the asymmetric power formulation by incorporating long memory through the HY framework and present the preferred model for each volatility process. For the squared returns, the chosen specification is the FIOAP one and for the realized measure, we select the HYAP one. In both cases the HEAVY-E (or GARCH-X) model is the preferred one.

In the FIOA[P] specification for the squared returns (see Table 4) δ_r is around 1.40 to 1.70 and d_r close to 0.50 (around 0.41 to 0.52). In most cases the Wald tests (not reported) reject the null hypotheses of $d_r = 0$ or 1 and $\delta_r = 1$ or 2. The Heavy coefficient, α_{rR}, is significant and around 0.03 to 0.07. In other words, both the lagged PT realized measure and lagged values of the PT squared returns drive the model of the PT conditional variance of returns. Furthermore, the own asymmetry coefficient (γ_{rr}) is significant and around 0.24 to 0.52, while the ϕ_r and γ_{rR} coefficients were insignificant in all six cases and, therefore, they were excluded.

Table 4: FIOA[P] HEAVY-E-r Specification ($\zeta_r = 1$, $\gamma_{rR} = 0$)

$$(1-\beta_r L)[(\sigma^2_{rt})^{\frac{\delta_r}{2}} - \omega_r] = A_r(L)(1+\gamma_{rr}s_t)(\varepsilon^2_{rt})^{\frac{\delta_r}{2}} + \alpha_{rR}(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}},$$
$$A_r(L) = (1-\beta_r L) - (1-L)^{d_r} \ (\varphi_r = 0)$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_r	0.48 (7.42)***	0.48 (5.45)***	$\underline{0.50}$ (8.11)***	**0.37** (2.52)***	0.45 (4.61)***	0.43 (3.77)***
d_r	0.51 (7.86)***	$\underline{0.52}$ (6.25)***	0.52 (8.66)***	**0.41** (4.97)***	0.49 (6.18)***	0.52 (4.97)***
γ_{rr}	$\underline{0.52}$ (6.70)***	**0.24** (4.00)***	0.28 (3.96)***	0.48 (8.12)***	0.50 (6.82)***	0.33 (8.79)***
α_{rR}	$\underline{0.07}$ (2.03)***	0.05 (2.74)***	0.07 (2.12)***	**0.03** (1.78)**	0.07 (2.04)***	0.05 (2.48)***
δ_r^*	**1.40**	1.60	1.40	1.50	1.40	$\underline{1.70}$
δ_R	$\underline{1.40}$	1.30	1.30	**1.20**	1.40	1.40

Notes: See Notes in Table 2.

*The powers used are approximations of the estimated powers from the non-linear models (available upon request).

In the HYA[P] specification for the realized measure (see Table 5) δ_R is around 1.20 to 1.40. The Wald tests (not reported) do not reject the null of $\delta_R = 1$ at 5% significance level for two out of six cases. There is also strong evidence of hyperbolic memory as ζ_R and d_R are around $0.85 - 0.90$ and $0.54 - 0.71$, respectively, with the Wald tests always rejecting the null of either a FIAP ($H_0 : \zeta_R = 1$) or an AP specification ($H_0 : \zeta_R = 0$). We further include the two Heavy-E (or ARCH-X) parameters, α_{Rr} and γ_{Rr}. The former is always insignificant and excluded and the latter, which captures the cross asymmetries, is significant and around $0.07 - 0.13$ in all but one case. So, both the lagged values of the PT realized measure and the lagged PT squared negative returns affect significantly the PT conditional variance of the SSR realized measure. The own asymmetry (γ_{RR}) is significant in all cases and around 0.08 to 0.25 and the other two Heavy parameters, β_R and ϕ_R, are around $0.50 - 0.70$ and $0.22 - 0.36$, respectively. Note that for five out of the six cases (the only exception is the TSE index) the HYA[P] specification of the HEAVY-E-R model with $0.69 \leq d_R \leq 0.71$, $\zeta_R \in [0.85, 0.90]$ and $1.20 \leq \delta_R \leq 1.40$ is the preferred one.

Table 5: HYA[P] HEAVY-E-R Specification

$$(1-\beta_R L)[(\sigma^2_{rt})^{\frac{\delta_R}{2}}-\omega_R]=A_R(L)(1+\gamma_{RR}s_t)(\varepsilon^2_{Rt})^{\frac{\delta_R}{2}}+\gamma_{Rr}s_{t-1}(\varepsilon^2_{rt})^{\frac{\delta_r}{2}},$$
$$A_R(L)=(1-\beta_R L)-(1-\phi_R L)[(1-\zeta_R)+\zeta_R(1-L)^{d_R}]$$

	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
β_R	0.63 (16.94)***	0.58 (7.53)***	**0.50** (7.08)***	0.67 (12.67)***	<u>0.70</u> (16.33)***	0.65 (14.52)***
ϕ_R	**0.22** (8.87)***	0.31 (4.51)***	0.30 (5.04)***	0.35 (7.11)***	<u>0.36</u> (7.62)***	0.30 (7.88)***
ζ_R	**0.85** (64.07)***	0.89 (51.19)***	0.87 (50.01)***	<u>0.90</u> (66.52)***	0.89 (80.48)***	0.85 (61.10)***
d_R	0.70 (17.96)***	0.70 (12.57)***	**0.54** (16.02)***	<u>0.71</u> (16.50)***	0.71 (18.07)***	0.69 (17.23)***
γ_{RR}	**0.08** (4.70)***	0.24 (6.51)***	0.10 (1.77)**	<u>0.25</u> (4.33)***	0.11 (2.32)***	0.24 (3.75)***
γ_{Rr}	0.10 (10.35)***	**0.07** (4.05)***	0.11 (7.97)***	0.13 (6.22)***	0.09 (9.68)***	<u>0.31</u> (10.18)***
δ_R	<u>1.40</u>	1.30	1.30	**1.20**	1.40	1.40
δ_r	**1.40**	1.60	1.40	1.50	1.40	<u>1.70</u>

Notes: See Notes in Table 4.

6 Further Extensions

6.1 The Garman-Klass Volatility

Following the HEAVY extensions with asymmetries, power transformations and long memory, in this Section we test the inclusion of an alternative measure of volatility (apart from the squared returns and the realized measure) to the HEAVY/GARCH framework already analyzed. Using data on the daily high, low, opening and closing prices of each index in the OMI's realized library we generate an additional daily measure of price volatility. To avoid the microstructure biases introduced by high-frequency data and based on the conclusion of Chen et al. (2006), that the range-based and high-frequency integrated volatility provide essentially equivalent results, we employ the classic range-based estimator of Garman and Klass (1980) to construct the daily GK volatility as follows:

$$GK_t=\frac{1}{2}u_t^2-(2\ln 2-1)c_t^2,$$

where u_t and c_t are the differences in the natural logarithms (as of time t) of the high and low and of the closing and opening prices, respectively. The Garman-Klass is an open-to-close range-based volatility estimator that is proved to be a more precise volatility proxy, with superior empirical performance in the

GARCH/MEM framework. Recently, Molnar (2016) proved that the inclusion of the Parkinson and GK estimators in the Range-(R)GARCH(1,1) model he proposes, outperforms the standard squared returns GARCH(1,1), and RGARCH performs particularly better in situations, where volatility level changes rapidly. Several studies also discuss the improvement of the GARCH/MEM framework through the open-to-close range-based volatility proxies, regarded as more accurate than the close-to-close squared returns: they exclude the noise from the dynamics of the opening jumps and they ensure greater accuracy in volatility forecasting through the range information they provide (see Chou et al. 2010 and 2015, Molnar, 2012 and the references therein). Therefore, we incorporate the GK variable in our system of HEAVY equations, in order to improve the forecasting performnace of the HEAVY framework.

We further form the SSR GK volatility ($\widetilde{GK}_t = sign(r_t)\sqrt{GK_t}$), which we will denote by ε_{gt}. As with the stock returns and the realized measure, we assume that $\varepsilon_{gt} = e_{gt}\sigma_{gt}$, where in the GARCH model the stochastic term e_{gt} is *i.i.d* with zero mean and unit variance. Therefore, $\mathbb{E}(\varepsilon_{gt}^2 | \mathcal{F}_{t-1}) = \sigma_{gt}^2$.

By using either the PT conditional variance of $\widetilde{GK}_t$ as dependent variable or the PT GK volatility as a regressor in the other two equations, we augment the AP (hereafter, we will use the abbreviation: Au-AP) HEAVY-E or GARCH-X models as follows:

$$\begin{aligned}
(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} &= \omega_r + (\alpha_{rr} + \gamma_{rr}s_{t-1})L(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{rR} + \gamma_{rR}s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}} + (\alpha_{rg} + \gamma_{rg}s_{t-1})L(\varepsilon_{gt}^2)^{\frac{\delta_g}{2}}, \\
(1-\beta_R L)(\sigma_{Rt}^2)^{\frac{\delta_R}{2}} &= \omega_R + (\alpha_{Rr} + \gamma_{Rr}s_{t-1})L(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{RR} + \gamma_{RR}s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}} + (\alpha_{Rg} + \gamma_{Rg}s_{t-1})L(\varepsilon_{gt}^2)^{\frac{\delta_g}{2}}, \\
(1-\beta_g L)(\sigma_{gt}^2)^{\frac{\delta_g}{2}} &= \omega_g + (\alpha_{gr} + \gamma_{gr}s_{t-1})L(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{gR} + \gamma_{gR}s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}} + (\alpha_{gg} + \gamma_{gg}s_{t-1})L(\varepsilon_{gt}^2)^{\frac{\delta_g}{2}}.
\end{aligned} \tag{6}$$

For example, the Au-AP HEAVY models for the stock returns and GK volatility are obtained from the above equations by setting $\alpha_{rr} = \gamma_{rr} = 0$ and $\alpha_{gg} = \gamma_{gg} = 0$:

$$\begin{aligned}
(1-\beta_r L)(\sigma_{rt}^2)^{\frac{\delta_r}{2}} &= \omega_r + (\alpha_{rR} + \gamma_{rR}s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}} + (\alpha_{rg} + \gamma_{rg}s_{t-1})L(\varepsilon_{gt}^2)^{\frac{\delta_g}{2}}, \\
(1-\beta_g L)(\sigma_{gt}^2)^{\frac{\delta_g}{2}} &= \omega_g + (\alpha_{gr} + \gamma_{gr}s_{t-1})L(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}} + (\alpha_{gR} + \gamma_{gR}s_{t-1})L(\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}},
\end{aligned} \tag{7}$$

which reduce to the HEAVY-$r(g)$ models if the Au-Heavy coefficients α_{rg}, $\gamma_{rg}(\alpha_{gr}, \gamma_{gr})$ are zero, and to the GARCH$(1,0)$-X-$r(g)$ models when the Heavy coefficients α_{rR}, $\gamma_{rR}(\alpha_{gR}, \gamma_{gR})$ are zero.

As with the bivariate case (see eq. (3)) we can express the three expressions of the trivariate system, eq. (6) in a matrix form:

$$(\mathbf{I}_3 - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + (\mathbf{A} + \boldsymbol{\Gamma}s_{t-1})L\boldsymbol{\varepsilon}_t, \tag{8}$$

where $\mathbf{B}$ is a diagonal matrix of order 3, with nonzero elements β_r, β_R and β_g; $\boldsymbol{\sigma}_t = [(\sigma_{rt}^2)^{\frac{\delta_r}{2}}, (\sigma_{Rt}^2)^{\frac{\delta_R}{2}}, (\sigma_{gt}^2)^{\frac{\delta_g}{2}}]'$, $\boldsymbol{\omega} = [\omega_r, \omega_R, \omega_g]'$, $\boldsymbol{\varepsilon}_t = [(\varepsilon_{rt}^2)^{\frac{\delta_r}{2}}, (\varepsilon_{Rt}^2)^{\frac{\delta_R}{2}}, (\varepsilon_{gt}^2)^{\frac{\delta_g}{2}}]'$, and s_t is defined in eq. (1); $\mathbf{A}$ and $\boldsymbol{\Gamma}$ are full matrices:

$$\mathbf{A}=\begin{bmatrix} \alpha_{rr} & \alpha_{rR} & \alpha_{rg} \\ \alpha_{Rr} & \alpha_{RR} & \alpha_{Rg} \\ \alpha_{gr} & \alpha_{gR} & \alpha_{gg} \end{bmatrix},\ \mathbf{\Gamma}=\begin{bmatrix} \gamma_{rr} & \gamma_{rR} & \gamma_{rg} \\ \gamma_{Rr} & \gamma_{RR} & \gamma_{Rg} \\ \gamma_{gr} & \gamma_{gR} & \gamma_{gg} \end{bmatrix}.$$

Table 6 presents the results for the chosen Au-AP HEAVY/GARCH models/specifications.

Regarding the stock returns, the A[P] HEAVY-E (or GARCH-X) model augmentd by the α_{rg} parameter is the statistically chosen model (see Panel A). That is, even when we take into account the presence of GK volatility the chosen model is the HEAVY–E one, since: i) the asymmetric Heavy parameter, γ_{rR}, ii) the own asymmetry parameter, γ_{rr}, and iii) the Au parameter, α_{rg}, are significant in all six stock indices.

Similarly, for the SSR realized measure the Au-A[P] HEAVY-E process is the preferred one in four out of the six cases (see Panel B). That is, for the SP, NIKKEI, DAX and EUSTOXX indices the two Heavy (or ARCH) parameters, α_{RR}, γ_{RR}, the asymmetric Heavy-E parameter, γ_{Rr}, and the Au parameter, α_{Rg}, are all significant. For the TSE index, the Au effect (captured by α_{Rg}) dominates the HEAVY-E impact, since for this case γ_{Rr} becomes insignifant, whereas for the FTSE index, the Heavy-E parameter, α_{Rr}, replaces the equivalent asymmetric one.

Finally, regarding the conditional variance of the SSR GK volatility the Au-A[P] HEAVY-E (or GARCH-X) model is again the chosen one (see panel C). In particular, the power term is $1.00 \leq \delta_g \leq 1.30$ in all cases. In addition, the Heavy parameter (α_{gR}), the own asymmetry parameter, γ_{gg}, and the asymmetric Au parameter, γ_{gr}, are significant in all cases. That is, not only the PT lagged realized measure, but also the lagged PT negative signed GK volatility and squared negative returns drive the model of the PT $\widetilde{GK}_t$.

Overall, our results show strong HEAVY effects (captured by the γ_{rR}, α_{RR}, γ_{RR} and α_{gR} parameters), asymmetric HEAVY-E influences (as the estimated γ_{rr}, γ_{Rr} and γ_{gr} are significant), as well as Au-HEAVY-E impacts (captured by the α_{rg} and α_{Rg} parameters).

Table 6: Au-A[P] HEAVY-E (or GARCH-X) Specifications (δ_i is fixed)

	SP	NIKKEI	TSE	FTSE*	DAX	EUSTOXX
			Panel A: Stock Returns			
β_r	0.84 (39.03)***	0.84 (32.30)***	0.86 (28.22)***	0.86 (37.26)***	**0.81** (25.60)***	0.84 (29.44)***
α_{rg}	0.02 (2.89)***	0.02 (4.04)***	0.03 (2.33)***	0.02 (2.28)***	0.02 (3.02)***	**0.01** (1.96)***
γ_{rr}	**0.05** (2.65)***	0.10 (4.64)***	0.07 (4.14)***	0.10 (6.61)***	**0.05** (2.52)***	0.07 (3.98)***
γ_{rR}	0.10 (4.13)***	0.04 (1.69)**	0.03 (1.84)**	**0.02** (3.18)***	0.14 (3.96)***	0.06 (4.25)***
			Panel B: Realized Measure*			
β_R	0.73 (33.03)***	0.66 (21.50)***	0.65 (22.59)***	**0.64** (20.64)***	0.70 (23.45)***	0.71 (29.20)***
α_{RR}	**0.10** (3.89)***	0.17 (5.49)***	0.26 (8.78)***	0.25 (8.43)***	0.17 (6.03)***	0.12 (5.07)***
α_{Rg}	0.04 (5.23)***	0.03 (5.21)***	0.02 (1.61)*	0.04 (1.68)**	**0.01** (4.22)***	0.04 (5.09)***
γ_{RR}	0.05 (4.18)***	0.03 (2.23)***	0.07 (7.55)***	0.07 (6.43)***	**0.02** (1.80)**	0.03 (3.28)***
γ_{Rr}	0.14 (10.88)***	0.10 (6.42)***			**0.09** (10.65)***	0.19 (11.25)***
			Panel C: GK volatility			
β_g	0.80 (30.31)***	0.68 (12.95)***	**0.62** (15.04)***	0.77 (28.11)***	0.78 (25.78)***	0.80 (31.04)***
α_{gR}	0.18 (4.31)***	0.61 (5.72)***	0.32 (7.61)***	**0.10** (6.50)***	0.53 (5.45)***	0.22 (5.21)***
γ_{gg}	0.06 (5.67)***	0.04 (2.83)***	**0.02** (1.48)*	**0.02** (1.50)*	0.05 (5.47)***	0.04 (4.41)***
γ_{gr}	0.21 (7.34)***	0.24 (4.11)***	0.20 (6.40)***	**0.17** (8.27)***	0.24 (6.96)***	0.31 (9.37)***
			Powers δ_i			
δ_r	1.50	1.70	**1.40**	1.50	**1.40**	1.50
δ_R	1.40	1.50	**1.20**	**1.20**	1.40	1.30
δ_g	1.20	1.20	**1.15**	1.30	1.00	1.10

Notes: See notes in Table 2.

FTSE: Realized measure equation: $\alpha_{Rr} = \underset{(5.34)^{***}}{0.11}$

6.2 Structural Breaks

After augmenting the AP HEAVY models with the GK volatility measure, in this Section we identify the structural breaks in the three volatility series for SP, focusing mainly on the recent global financial crisis, and study their impact on the Au-AP HEAVY models. 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, for each index we identify the structural breaks in the three series (PT squared returns, PT realized measure and PT GK volatility) with the Bai and Perron methodology (see Table 7). We use the breaks of the three series in order to build the slope dummies for the various coefficients in the Au-AP HEAVY-E models. We observe that a break date for the recent financial crisis of 2007-08 is detected, so that we can focus on the crisis effect. We also detect one break date before and one after the crisis.

Table 7: The break dates for SP

	1^{st} Break	2^{nd} Break	3^{rd} Break
r	**28/04/2003**	**31/10/2007**	**30/10/2009**
R	11/04/2003	06/11/2007	02/11/2009
g	06/08/2003	23/07/2007	15/07/2009

Notes: Bai & 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 Au-AP HEAVY-E models.

We present the estimation results for the SP index in Table 8, where we choose to use the 3 breaks of the PT squared returns series: (1) 28/04/2003: pre-crisis break, (2) 31/10/2007: crisis break and (3) 30/10/2009: post-crisis break. In the returns equation, the coefficient of the lagged GK volatility measure, α_{rg}, receives an impact from the crisis and the post-crisis break .The GK estimate is increased by the crisis dummy ($+0.01$) and decreased by the post-crisis dummy (-0.01). Regarding the realized measure equation, the ARCH/Heavy effect, α_{RR}, rises with the crisis break, while the lagged GK, α_{Rg}, and the cross effect from the returns, γ_{Rr}, parameters fall after the post- and pre-crisis breaks, respectively. Finally, in the GK equation, the heavy coefficient, α_{gR}, and the returns asymmetry, γ_{gr}, receive a negative effect from the 1st and the 3rd break dummies, respectively, but the own asymmetry coefficient, γ_{gg}, increases with the crisis dummy.

Overall, our finding is that the dummy coefficients corresponding to the 2003 and 2009 breaks are negative, whereas the one for the 2007 crisis is always positive and gives an increment to the coefficient it refers to.

Table 8: Au-A[P] HEAVY-E Specifications for SP with breaks

Panel A: Stock Returns							
β_r	α_{rg}	$\alpha_{rg}^{(2)}$	$\alpha_{rg}^{(3)}$	γ_{rr}	γ_{rR}		
0.83 $(37.70)^{***}$	0.02 $(2.53)^{***}$	0.01 $(2.44)^{***}$	-0.01 $(-1.67)^{**}$	0.05 $(2.35)^{***}$	0.10 $(4.43)^{***}$		
Panel B: Realized Measure							
β_R	α_{RR}	$\alpha_{RR}^{(2)}$	α_{Rg}	$\alpha_{Rg}^{(3)}$	γ_{RR}	γ_{Rr}	$\gamma_{Rr}^{(1)}$
0.72 $(32.47)^{***}$	0.09 $(3.43)^{***}$	0.03 $(3.78)^{***}$	0.04 $(5.23)^{***}$	-0.01 $(-2.72)^{***}$	0.06 $(4.35)^{***}$	0.17 $(9.89)^{***}$	-0.05 $(-2.94)^{***}$
Panel C: GK volatility							
β_g	α_{gR}	$\alpha_{gR}^{(1)}$	γ_{gg}	$\gamma_{gg}^{(2)}$	γ_{gr}	$\gamma_{gr}^{(3)}$	
0.79 $(29.16)^{***}$	0.21 $(4.45)^{***}$	-0.06 $(-3.90)^{***}$	0.04 $(3.47)^{***}$	0.05 $(4.30)^{***}$	0.23 $(8.10)^{***}$	-0.05 $(-1.86)^{**}$	
Powers δ_i							
δ_r	δ_R	δ_g					
1.50	1.40	1.20					

Notes: See notes in Table 2.

6.3 Volatility Spillovers (Incomplete Section)

7 Theoretical Results

7.1 N-dimensional Process

In this Section we consider the N-dimensional AP HEAVY/GARCH/MEM system. First, we will introduce some further notation.

Let $\boldsymbol{\varepsilon}_t = [|\varepsilon_{it}|^{\delta_i}]_{i=1,\ldots,N}$ (hereafter for typographical convenience we will drop the subscript), where $\delta_i \in \mathbb{R}^+$. We assume that the vector $\boldsymbol{\varepsilon}_t$ is characterized by the relation

$$\boldsymbol{\varepsilon}_t = \mathbf{Z}_t \boldsymbol{\sigma}_t, \tag{9}$$

where $\mathbf{Z}_t = diag[\mathbf{z}_t]-\ diag[\mathbf{y}] = diag\{y_1, \ldots, y_N\}$ refers to a diagonal matrix- with $\mathbf{z}_t = [|e_{it}|^{\delta_i}]$, and $\boldsymbol{\sigma}_t$ is $\mathcal{F}_{t-1}$ measurable with $\mathbf{F}_{t-1} = \sigma(\boldsymbol{\varepsilon}_{t-1}, \boldsymbol{\varepsilon}_{t-2}, \ldots)$. That is, $\boldsymbol{\varepsilon}_t = [|e_{it}|^{\delta_i}\, \sigma_{it}^{\delta_i}]$.

Let also $\widetilde{\boldsymbol{\sigma}}_t = [\sigma_{it}]$, in other words, $\widetilde{\boldsymbol{\sigma}}_t$ is equal to $\boldsymbol{\sigma}_t$ when $\delta_i = 1$ for all i. Further, let $\mathbf{E}_t = diag[\mathbf{e}_t]$ where the stochastic vector $\mathbf{e}_t = [e_{it}]$ is independent and identically distributed ($i.i.d$); notice that the ith element of $\mathbf{e}_t$ is equal to the corresponding element of $\mathbf{z}_t$, when $\delta_i = 1$ for all i, multiplied by $sign(e_{it})$. In addition, let $\widetilde{\boldsymbol{\varepsilon}}_t = \mathbf{E}_t \widetilde{\boldsymbol{\sigma}}_t = [e_{it}\sigma_{it}]$.

In the N-dimensional GARCH model $\mathbf{e}_t$ has zero mean, unit variance, and positive definite time invariant correlation matrix $\mathbf{R} = [\rho_{ij}]$ with $\rho_{ii} = 1$ therefore, $\widetilde{\boldsymbol{\varepsilon}}_t$ is a vector with zero conditional mean: $\mathbf{E}(\widetilde{\boldsymbol{\varepsilon}}_t | \mathbf{F}_{t-1}) = \mathbf{0}$; $\mathbf{E}(\cdot)$ refers to the elementwise expectation operator. The conditional covariance matrix of $\widetilde{\boldsymbol{\varepsilon}}_t$ is given by $\boldsymbol{\Sigma}_t = \mathbf{E}(\widetilde{\boldsymbol{\varepsilon}}_t \widetilde{\boldsymbol{\varepsilon}}_t' | \mathbf{F}_{t-1}) = diag[\widetilde{\boldsymbol{\sigma}}_t]\mathbf{R}diag[\widetilde{\boldsymbol{\sigma}}_t]$.

In the N-dimensional MEM $\mathbf{e}_t > \mathbf{0}$, with $\mathbf{E}(\mathbf{e}_t) = \mathbf{j}$ (where $\mathbf{j}$ is the unit vector), and positive definite covariance matrix $\mathbf{Q} = [q_{ij}]$, with $\mathbf{q} = diag[\mathbf{Q}]$. That is, $\mathbf{E}(\widetilde{\boldsymbol{\varepsilon}}_t | \mathbf{F}_{t-1}) = \widetilde{\boldsymbol{\sigma}}_t$. In this case $\boldsymbol{\Sigma}_t = \mathbf{E}(\widetilde{\boldsymbol{\varepsilon}}_t \widetilde{\boldsymbol{\varepsilon}}_t' | \mathbf{F}_{t-1}) = diag[\widetilde{\boldsymbol{\sigma}}_t]\mathbf{Q}diag[\widetilde{\boldsymbol{\sigma}}_t]$.

The N-dimensional semi unrestricted (SUE) AP model[6] of order $(1,1)$ -in what follows for notational simplicity we will drop the order of the model if it is $(1,1)$- consists of the following equations:

$$\sigma_{it}^{\delta_i} = \omega_i + \sum\nolimits_{j=1}^{N} (\alpha_{ij} + \gamma_{ij} s_{t-1}) \left|\varepsilon_{j,t-1}\right|^{\delta_j} + \sum\nolimits_{j=1}^{N} \beta_{ij} \sigma_{j,t-1}^{\delta_j},$$

where we recall that $s_t = 0.5[1 - sign(r_t)]$.

This can be either a multivariate HEAVY/GARCH model or a MEM. For example, in our trivariate context the three GARCH variables, ε_{it}, $i = 1, 2$, are the stock returns and the signed square rooted (SSR) realized measure and GK volatility, whereas $\sigma_{it}^2 = \mathbb{E}(\varepsilon_{it}^2 | \mathcal{F}_{t-1})$ are their conditional variances. The HEAVY formulation parallels the GARCH one. It is also very similar to the bivariate MEM. In the latter model the three variables (ε_{it}) are the squared returns, the realized measure and the GK volatility, whereas $\sigma_{it} = \mathbb{E}(\varepsilon_{it} | \mathcal{F}_{t-1})$ are their conditional means. Therefore, as noted earlier, we will use the three terms, GARCH, HEAVY, MEM, interchangeably.

The SUE-AP model can be expressed/interpreted as an N-dimensional system with shock (unconditional) and conditional spillovers:

$$(\mathbf{I} - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + L\mathbf{A}(t)\boldsymbol{\varepsilon}_t, \tag{10}$$

where $\mathbf{B} = [\beta_{ij}]$ is a full matrix (of order N), that its cross diagonal elements capture the conditional spillovers; $\boldsymbol{\omega} = [\omega_i]$ is a vector that contains the drifts; $\mathbf{A}(t) = \mathbf{A} + \boldsymbol{\Gamma} s_t$, where $\mathbf{A} = [\alpha_{ij}]$ and $\boldsymbol{\Gamma} = [\gamma_{ij}]$ are full matrices as well. The cross diagonal elements of $(\boldsymbol{\Gamma})\mathbf{A}$ capture the (asymmetric) shock (or unconditional) spillovers.

[6] It is termed semi unrestricted extended, because the three matrices are full (extended), and although some of the elements of the $\mathbf{B}$ matrix are allowed to take negative values, the $\mathbf{A}$ and $\boldsymbol{\Gamma}$ matrices should be non-negative (semi-unrestricted); see Karanasos and Hu (2017).

7.2 Optimal Predictors

In order to derive the optimal predictors, we need to obtain the ARMA representation of the SUE-AP model in eqs. (9)-(10).

ARMA Representation and General Solution

First, let $\mathbf{Z} = \mathbf{E}(\mathbf{Z}_t) < \infty$ (the inequality sign refers to element-by-element inequality). We also define the serially uncorrelated vector (with zero mean): $\mathbf{v}_t = \boldsymbol{\varepsilon}_t - \mathbf{E}(\boldsymbol{\varepsilon}_t \,|\mathbf{F}_{t-1}\,)$, where $\mathbf{E}(\boldsymbol{\varepsilon}_t \,|\mathbf{F}_{t-1}\,) = \mathbf{Z}\boldsymbol{\sigma}_t$.

Corollary 1 *The ARMA representation of the N-dimensional SUE-AP process in eqs. (9) and (10) is given by*

$$[\mathbf{I} - L\mathbf{C}(t)]\boldsymbol{\sigma}_t = \boldsymbol{\omega} + L\mathbf{A}(t)\mathbf{v}_t, \tag{11}$$

where $\mathbf{C}(t) = \mathbf{B} + \mathbf{A}(t)\mathbf{Z}$.

The proof is trivial: we add and subtract $\mathbf{A}(t)\mathbf{Z}\boldsymbol{\sigma}_\mathbf{t}$ in the right-hand side of eq. (10).

Next, we will present the general solution, which generates all the main time series properties of the SUE-AP model.

But, first we define

$$\mathbf{D}_{t,k} = \prod\nolimits_{r=0}^{k-1} \mathbf{C}(t-r-1), \tag{12}$$

coupled with the initial value $\mathbf{D}_{t,0} = \mathbf{I}$, where $k \in Z^*$ (Z^* is the set of non-negative integers).

Theorem 1 *The general solution of eq. (11) with initial condition value $\mathbf{c}_{t-k} = \boldsymbol{\sigma}_{t-k}$, is given by*

$$\boldsymbol{\sigma}_t = \underbrace{\sum\nolimits_{r=0}^{k-1} \mathbf{D}_{t,r}[\boldsymbol{\omega} + \mathbf{A}(t-r-1)\mathbf{v}_{t-r-1}]}_{(Particular\ Solution)} + \underbrace{\mathbf{D}_{t,k}\mathbf{c}_{t-k}}_{(Homoneneous\ Solution)}. \tag{13}$$

The proof is trivial. It is obtained by using repeated substitution in eq. (11).

In the above Theorem $\boldsymbol{\sigma}_{t,k}$ is decomposed in two parts: the homogeneous part consists of the initial condition $\mathbf{c}_{t-k}$; the particular part contains the drift ($\boldsymbol{\omega}$) and the lags of $\mathbf{v}_t$ from time $t-k$ to time $t-1$. Notice that the 'matrix coefficients' or weights are the terms in the generating sequence $\{\mathbf{D}_{t,r}\}_{0 \le r \le k-1}$. Moreover, for '$k = 0$' (for $i > j$ we will use the convention $\sum_{r=i}^{j}(\cdot) = 0$), since $\mathbf{D}_{t,0} = \mathbf{I}$ (see eq.(12)), eq. (13) becomes an 'identity': $\boldsymbol{\sigma}_t = \mathbf{c}_t = \boldsymbol{\sigma}_t$. Similarly, when $k = 1$ eq. (13), since $\mathbf{D}_{t,1} = \mathbf{C}(t-1)$, reduces to 'eq. (11)' with initial condition value $\mathbf{c}_{t-1} = \boldsymbol{\sigma}_{t-1}$:

$$\boldsymbol{\sigma}_t = \boldsymbol{\omega} + \mathbf{C}(t-1)\mathbf{c}_{t-1} + \mathbf{A}(t-1)\mathbf{v}_{t-1} = \boldsymbol{\omega} + \mathbf{C}(t-1)\boldsymbol{\sigma}_{t-1} + \mathbf{A}(t-1)\mathbf{v}_{t-1}.$$

In what follows, we will obtain the linear predictor of the SUE-AP model.

First, we will introduce some additional notation. Let $\mathbf{C} = \mathbf{E}[\mathbf{C}(t)]$ (where $\mathbf{C}(t)$ has been given in eq. (11)). Thus,

$$\mathbf{C} = \mathbf{E}[\mathbf{C}(t)] = \mathbf{B} + (\mathbf{A} + \mathbf{\Gamma}\frac{\mathbf{1}}{\mathbf{2}})\mathbf{Z}, \tag{14}$$

since $\mathbf{E}[diag[\mathbf{s}_t]] = \mathbf{E}[diag[\mathbf{s}_t^{\wedge 2}]] = 1/2\mathbf{I}$ ($^\wedge$ denotes the elementwise exponentiation) and, therefore, $\mathbf{E}[\mathbf{A}(t)] = \mathbf{A} + \mathbf{\Gamma}\frac{\mathbf{1}}{\mathbf{2}}$**,** which implies that $\mathbf{E}(\mathbf{D}_{t,k}) = \mathbf{C}^k$ ($\mathbf{Y}^k$ denotes the matrix $\mathbf{Y}$ raised to the power of k).

Taking the conditional expectation of eq. (13) with respect to the σ field $\mathbf{F}_{t-k-1}$ yields the following Proposition.

Proposition 1 *The k-step-ahead optimal (in L_2 sense) linear predictor of $\boldsymbol{\sigma}_t$, $\mathbf{E}(\boldsymbol{\sigma}_t\,|\mathbf{F}_{t-k-1})$, is readily seen to be*

$$\mathbf{E}(\boldsymbol{\sigma}_t\,|\mathbf{F}_{t-k-1}) = \left(\sum\nolimits_{r=0}^{k-1}\mathbf{C}^r\right)\boldsymbol{\omega} + \mathbf{C}^k\mathbf{c}_{t-k}, \tag{15}$$

Further, $\mathbf{C}^k$ can be expressed as

$$\mathbf{C}^k = \widetilde{\mathbf{C}}diag[\boldsymbol{\phi}^{\wedge k}]\widetilde{\mathbf{C}}^{-1}$$

(see, for example, Hamilton, 1994), where $\widetilde{\mathbf{C}} = [\widetilde{c}_{ij}]$ is the matrix with the N eigenvectors of $\mathbf{C}$, and $\boldsymbol{\phi} = [\phi_i]$ is the vector of the N eigenvalues. Denote the ijth element of $\widetilde{\mathbf{C}}^{-1}$by $\widetilde{c}_{ij}^{*}$ and define $\boldsymbol{\omega}_k = \sum_{r=0}^{k-1}\mathbf{C}^r\boldsymbol{\omega}$ $=[\omega_i^{(k)}]$ (the superscript in parenthesis denotes an index). Then the ith element of eq. (15) is given by $\underline{^{gal}}$

$$\mathbb{E}(\sigma_{it}^{\delta_i}\,|\mathcal{F}_{t-k}) = \omega_i^{(k)} + \sum\nolimits_{m=1}^{N}\sum\nolimits_{l=1}^{N}\widetilde{c}_{il}\widetilde{c}_{lm}^{*}\phi_l^k\sigma_{m,t-k}^{\delta_m}$$

(results for the associated forecast error and its variance are available upon request).

First-order Moment

Next, we will obtain the first unconditional moment of SUE-AP model.

First, let $\lambda(\mathbf{Y})$ refer to the modulus of the largest eigenvalue of $\mathbf{Y}$. Also, let $adj[\mathbf{Y}]$ denote the adjoint of matrix $\mathbf{Y}$.

Assumption 1. $\lambda(\mathbf{C}) < 1$.

Corollary 2 *Under Assumption A1, the first-order moment vector $\boldsymbol{\sigma} = \mathbf{E}(\boldsymbol{\sigma}_t) = \lim_{k\to\infty}\mathbf{E}(\boldsymbol{\sigma}_t\,|\mathbf{F}_{t-k-1})$, if and only if $adj(\mathbf{I}-\mathbf{C})\boldsymbol{\omega} > \mathbf{0}$, is given by*

$$\boldsymbol{\sigma} = (\mathbf{I}-\mathbf{C})^{-1}\boldsymbol{\omega}. \tag{16}$$

Notice that eq. (16) imposes an additional matrix inequality constraint on the parameter space, that is $adj(\mathbf{I}-\mathbf{C})\boldsymbol{\omega} > \mathbf{0}$. Finally, the following corollary gives the optimal linear predictors of $\boldsymbol{\varepsilon}_t$, and its first unconditional moment as well. The proof follows from Proposition 1 and Corollary 2, and it is trivial.

Corollary 3 *The k-step-ahead optimal (in L_2 sense) linear predictor of $\boldsymbol{\varepsilon}_t$ is given by*

$$\mathbf{E}(\boldsymbol{\varepsilon}_t \,|\mathbf{F}_{t-k-1}) = \mathbf{ZE}(\boldsymbol{\sigma}_t \,|\mathbf{F}_{t-k-1}),$$

where $\mathbf{E}(\boldsymbol{\sigma}_t \,|\mathbf{F}_{t-k-1})$ is given in eq. (15). Under assumption A1, and if and only if $adj[\mathbf{I}-\mathbf{C}]\boldsymbol{\omega} > \mathbf{0}$, then the first unconditional moment vector, $\boldsymbol{\varepsilon} = \mathbf{E}(\boldsymbol{\varepsilon}_t) = \lim_{k\to\infty} \mathbf{E}(\boldsymbol{\varepsilon}_t \,|\mathbf{F}_{t-k-1})$, is given by

$$\boldsymbol{\varepsilon} = \mathbf{Z}(\mathbf{I}-\mathbf{C})^{-1}\boldsymbol{\omega}.$$

Verification of the above corollary is straightforward and hence its proof is omitted.

7.3 Second Moment Structure

Now that we have derived the optimal predictors and the first unconditional moment of the SUE-AP model, we will examine its second moments. But first, we will introduce some further notation.

NOTATION

Let $\boldsymbol{\Gamma}(l) = [\gamma_{ij}(l)]$, $l \in \mathbb{Z}^*$ ($\mathbb{Z}^*$ is the set of non negative integers), denote the multidimensional covariance function of $\{\boldsymbol{\sigma}_t\}$, that is

$$\boldsymbol{\Gamma}(l) = \mathbf{E}[(\boldsymbol{\sigma}_{t-l}-\boldsymbol{\sigma})(\boldsymbol{\sigma}_t-\boldsymbol{\sigma})'], \tag{17}$$

or

$$\boldsymbol{\Gamma}(l) = \boldsymbol{\Sigma}(l) - \boldsymbol{\sigma}\boldsymbol{\sigma}',$$

where $\boldsymbol{\Sigma}(l) = \mathbb{E}(\boldsymbol{\sigma}_{t-l}\boldsymbol{\sigma}_t')$. In addition, let the vec forms of $\boldsymbol{\Sigma}(l)$ and $\boldsymbol{\Gamma}(l)$ be denoted by $\mathbf{s}(l)$ and $\boldsymbol{\gamma}(l)$, respectively. Explicit solutions for the $\boldsymbol{\Gamma}(l)$ and conditions for its existence will be presented below.

Further, let

$$\mathbf{D} = diag[\sqrt{\gamma_{11}(0)}, \ldots, \sqrt{\gamma_{NN}(0)}], \tag{18}$$

where $\gamma_{ii}(0)$ is the ith diagonal element of $\boldsymbol{\Gamma}(0)$. To further fix notation, write the lth-order, for $l \geq 1$, autocorrelation matrix of $\{\boldsymbol{\sigma}_t\}$ as

$$\mathbf{R}(l) = \mathbf{D}^{-1}\boldsymbol{\Gamma}(l)\mathbf{D}^{-1}. \tag{19}$$

Clearly, the lth-order autocorrelation matrix $\mathbf{R}(l)$ has the stacked form: $vec[\mathbf{R}(l)] = (\mathbf{D}^{-1})^{\otimes 2}\boldsymbol{\gamma}(l)$, where $\mathbf{X}^{\otimes 2} = \mathbf{X} \otimes \mathbf{X}$, and $\otimes$ is the Kronecker product.

Kronecker Products

Next, we will introduce some additional notation, which involves various Kronecker products. Specifically, let

$$\begin{aligned} \mathbf{Z}^{\otimes 2} &= \mathbf{Z} \otimes \mathbf{Z},\ \mathbf{Z}_*^{\otimes 2} = \mathbb{E}(\mathbf{Z}_t \otimes \mathbf{Z}_t),\ \widetilde{\mathbf{Z}} = \mathbf{Z}_*^{\otimes 2} - \mathbf{Z}^{\otimes 2}, \\ \mathbf{C}^{\otimes 2} &= \mathbf{C} \otimes \mathbf{C},\ \mathbf{C}^{\otimes I} = \mathbf{C} \otimes \mathbf{I},\ \mathbf{I}^{\otimes C} = \mathbf{I} \otimes \mathbf{C}, \\ \widetilde{\mathbf{A}} &= \mathbf{E}[\mathbf{A}(t) \otimes \mathbf{A}(t)]. \end{aligned} \tag{20}$$

We will assume that $\mathbf{Z}_*^{\otimes 2} < \infty$, that is $\mathbb{E}(|e_{it}|^{\delta_i} |e_{jt}|^{\delta_j}) < \infty$, for all i and j. Notice that $\widetilde{\mathbf{Z}}$ in eq. (20) is a diagonal matrix (of order N^2), and its rth element, with $r = [(i-1)N + j]$, where for each $i = 1, \dots, N$, $j = 1, \dots, N$, is given by

$$z_{(i-1)N+j} = \mathbb{E}(|e_{it}|^{\delta_i} |e_{jt}|^{\delta_j}) - \mathbb{E}(|e_{it}|^{\delta_i})\mathbb{E}(|e_{jt}|^{\delta_j}).$$

Therefore, $\widetilde{\mathbf{Z}}\mathbf{j}_{N^2}$ is a vector of order N^2 with $[(i-1)N + j]$th element $z_{(i-1)N+j}$.

Denote also

$$\widetilde{\mathbf{C}} = \mathbf{I}_{N^2} - \mathbf{I}^{\otimes \mathbf{C}}\mathbf{C}^{\otimes \mathbf{I}} - \widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}. \tag{21}$$

Assumption 2. $\lambda(\widetilde{\mathbf{C}}) < 1$.

Theorem 2 *Under Assumption A2 the vec form of* $\mathbf{\Gamma}(0)$*, if and only if* $\mathbf{D} > \mathbf{0}$*, is given by*

$$\boldsymbol{\gamma}(0) = \widetilde{\mathbf{C}}^{-1}\widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}\boldsymbol{\sigma}^{\otimes 2}. \tag{22}$$

Further, the vec form of the covariance function, for lag $l \geq 1$*,* $\boldsymbol{\gamma}(l)$*, is given by*

$$\boldsymbol{\gamma}(l) = \mathbf{I}^{\otimes \mathbf{C}}\boldsymbol{\gamma}(0).$$

Notice that eq. (22) imposes an additional matrix inequality constraint on the parameter space, that is $\mathbf{D} > \mathbf{0}$ (18).

8 Conclusions

Our study extended the HEAVY models by taking into consideration leverage, power transformations and long memory characteristics. For the realized measure our empirical results favour the most general hyperbolic asymmetric power specification, where both the power transformed lagged squares of the negative returns and lagged values of the signed square rooted realized measure move the dynamics of the power transformed conditional variance of the latter. Similarly, modelling the returns with a fractionally integrated process, we found that both the power transformed lagged negative signed realized measure and lagged values of the returns help forecasting the power transformed 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 on the areas of asset allocation and portfolio selection, as well as on several risk management practices.

Moreover, we augmented the HEAVY model with a third variable, the Garman-Klass range-based volatility, creating a trivariate system to achieve greater accuracy in volatility forecasting. Further, the detection of structural breaks and the inclusion of break dummies in the augmented formulation capture the time-varying pattern of the parameters, as the break corresponding to the financial crisis of 2008, in particular, increases the values of the parameters.

Future research should focus on applying the long memory asymmetric power extensions on the multivariate HEAVY model of Noureldin et al. (2012), and employ different error distributions like the Skewed-t.

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] Bollerslev, T., Mikkelsen, H.O., 1996. Modelling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151-184.

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

[15] Brownlees, C.T., Cipollini, F., Gallo, G.M., 2011. Intra-daily volume modeling and prediction for algorithmic trading. Journal of Financial Econometrics 9, 489-518.

[16] Chen, Z., Daigler, R., Parhizgari, A., 2006. Persistence of volatility in future markets. Journal of Futures Markets 26, 571-594.

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

[18] 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.

[19] 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.

[20] Chou, R. Y., Chou, H., and Liu, N., 2010. Range Volatility Models and Their Applications in Finance. In Lee, C.-F., and Lee, J.: Handbook of Quantitative Finance and Risk Management, 1273-1282. New York, NY: Springer.

[21] Chou, R. Y., Chou, H., and Liu, N., 2015. Range Volatility: A Review of Models and Empirical Studies. In Lee, C.-F., and Lee, J.: Handbook of Financial Econometrics and Statistics, 2029-2050. New York, NY: Springer.

[22] 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.

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

[24] Conrad, C., 2010. Non-negativity conditions for the hyperbolic GARCH model. Journal of Econometrics 157, 441-457.

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

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

[27] Dark, J.G., 2005. Modelling the conditional density using a hyperbolic asymmetric power ARCH model. Mimeo, Monash University.

[28] 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(2), Article 3.

[29] 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.

[30] Davidson, J., 2013. Time Series Modelling 4.37. http://www.timeseriesmodelling.com.

[31] 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.

[32] Engle, R.F., 2000. The econometrics of ultra-high-frequency data. Econometrica 68, 1-22.

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

[34] Engle, R.F., 2009. The Volatility Laboratory (V-Lab), Version 2.4.9, New York University Stern, The Volatility Institute.

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

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

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

[38] Garman, M., Klass, M., 1980. On the estimation of security price volatilities from historical data. Journal of Business 53, 67-78.

[39] 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.

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

[41] 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.

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

[43] Heber, G., Lunde, A., Shephard, N., Sheppard, K., 2009. Oxford-Man Institute's (OMI's) realized library, Version 0.2. Oxford-Man Institute: University of Oxford.

[44] Karanasos, M., Xu, Y., 2017. Matrix Inequality Constraints for Vector (Asymmetric Power) GARCH/HEAVY Models and MEM with spillovers: some New (Mixture) Formulations. Working Paper, Brunel University.

[45] Koopman, S.J., Jungbacker, B., Hol, E., 2005. Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. Journal of Empirical Finance 12, 445-475.

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

[47] Li, Y., Nolte, I., Nolte, S., 2016. High-Frequency Volatility Modelling: A Markov-Switching Autoregressive Conditional Intensity Model. Working Paper, Lancaster University.

[48] Manganelli, S., 2005. Duration, volume and volatility impact of trades. Journal of Financial Markets 8, 377-399.

[49] Molnár, P., 2012. Properties of range-based volatility estimators. International Review of Financial Analysis 23, 20-29.

[50] Molnár, P., 2016. High-low range in GARCH models of stock return volatility, Applied Economics, 48:51, 4977-4991.

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

[52] Russell, J. R., 1999. Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data. Working Paper, University of Chicago.

[53] Schoffer, O., 2003. HY-A-PARCH: A stationary A-PARCH model with long memory. Mimeo, University of Dortmund.

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

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

A APPENDIX

Unconditional Moments

Proof. (of Theorem 2) Rewrite the ARMA representation, eq. (11) as

$$\boldsymbol{\sigma}_t = \boldsymbol{\omega} + \mathbf{C}(t-1)\boldsymbol{\sigma}_{t-1} + \mathbf{A}(t-1)\mathbf{v}_{t-1}, \tag{A.1a}$$

or

$$\boldsymbol{\sigma}'_t = \boldsymbol{\omega}' + \boldsymbol{\sigma}'_{t-1}\mathbf{C}(t-1)' + \mathbf{v}'_{t-1}\mathbf{A}(t-1)'. \tag{A.1b}$$

Right-multiplying eq. (A.1a) by $\boldsymbol{\sigma}'_t$ yields

$$\boldsymbol{\sigma}_t\boldsymbol{\sigma}'_t = \boldsymbol{\omega}\boldsymbol{\sigma}'_t + \mathbf{C}(t-1)\boldsymbol{\sigma}_{t-1}\boldsymbol{\sigma}'_t + \mathbf{A}(t-1)\mathbf{v}_{t-1}\boldsymbol{\sigma}'_t. \tag{A.2}$$

Under Assumption A2, taking expectations on both sides of (A.2) yields

$$\boldsymbol{\Sigma}(0) = \boldsymbol{\omega}\boldsymbol{\sigma}' + \mathbf{C}\boldsymbol{\Sigma}(1) + \mathbf{E}[\mathbf{A}(t-1)\mathbf{v}_{t-1}\boldsymbol{\sigma}'_t] \tag{A.2a}$$

(we recall that $\boldsymbol{\Sigma}(0)$ has been defined in eq. (17), $\mathbf{C}$ and $\boldsymbol{\sigma}$ and have been given in Proposition 1 and Corollary 2, respectively).

On account of eq. (A.1b) the last term in the right hand-side of eq. (A.2a) is given by

$$\mathbf{E}[\mathbf{A}(t-1)\mathbf{v}_{t-1}\boldsymbol{\sigma}_t'] = \mathbf{E}[\mathbf{A}(t-1)\mathbf{v}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)']. \tag{A.2b}$$

Using $\boldsymbol{\omega}$ =$(\mathbf{I}-\mathbf{C})\boldsymbol{\sigma}$ (see Corollary 2) and eq. (A.2b), eq. (A.2a) produces

$$\boldsymbol{\Sigma}(0) = (\mathbf{I}-\mathbf{C})\boldsymbol{\sigma}\boldsymbol{\sigma}' + \mathbf{C}\boldsymbol{\Sigma}(1) + \mathbf{E}[\mathbf{A}(t-1)\mathbf{v}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)'] \ ,$$

or

$$\boldsymbol{\Gamma}(0) = \mathbf{C}\boldsymbol{\Gamma}(1) + \mathbf{E}[\mathbf{A}(t-1)\mathbf{v}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)'] \tag{A.3}$$

($\boldsymbol{\Gamma}(l)$ has been defined in eq. (17)).

Taking the vec form of (A.3) and using

$$\mathbf{E}[vec[\mathbf{A}(t-1)\mathbf{v}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)']] = \widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}\mathbf{s}(0) = \widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}[\boldsymbol{\gamma}(0) + \boldsymbol{\sigma}^{\otimes 2}],$$

where $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{Z}}$ have been defined in eq. (20), yields

$$\boldsymbol{\gamma}(0) = \mathbf{I}^{\otimes C}\boldsymbol{\gamma}(1) + \widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}[\boldsymbol{\gamma}(0) + \boldsymbol{\sigma}^{\otimes 2}]. \tag{A.4}$$

Next, left-multiplying eq. (A.1b) by $\boldsymbol{\sigma}_{t-1}$ yields

$$\boldsymbol{\sigma}_{t-1}\boldsymbol{\sigma}_t' = \boldsymbol{\sigma}_{t-1}\boldsymbol{\omega}' + \boldsymbol{\sigma}_{t-1}\boldsymbol{\sigma}_{t-1}'\mathbf{C}(t-1)' + \boldsymbol{\sigma}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)'. \tag{A.5}$$

Under Assumption A2, taking expectations on both sides of (A.5) yields

$$\boldsymbol{\Sigma}(1) = \boldsymbol{\sigma}\boldsymbol{\omega}' + \boldsymbol{\Sigma}(0)\mathbf{C}', \tag{A.6}$$

(since $\mathbf{E}[\boldsymbol{\sigma}_{t-1}\mathbf{v}_{t-1}'\mathbf{A}(t-1)'] = 0$). Taking the vec form of the above equation, and using the fact that $\boldsymbol{\omega}$ =$(\mathbf{I}-\mathbf{C})\boldsymbol{\sigma}$ (see eq. (16)), eq. (A.6) yields

$$\boldsymbol{\gamma}(1) = \mathbf{C}^{\otimes I}\boldsymbol{\gamma}(0). \tag{A.7}$$

Substituting eq. (A.7) into eq. (A.4) and solving for $\boldsymbol{\gamma}(0)$ gives

$$\boldsymbol{\gamma}(0) = \widetilde{\mathbf{C}}^{-1}\widetilde{\mathbf{A}}\widetilde{\mathbf{Z}}\boldsymbol{\sigma}^{\otimes 2}$$

(we recall that $\widetilde{\mathbf{C}}$ has been defined in eq. (21)), as claimed.

Next, rewrite the general solution eq. (13) as

$$\boldsymbol{\sigma}_t' = \sum\nolimits_{r=0}^{l-1}[\boldsymbol{\omega}' + \mathbf{v}_{t-r-1}'\mathbf{A}(t-r-1)']\mathbf{D}_{t,r}' + \boldsymbol{\sigma}_{t-l}'\mathbf{D}_{t,l}'. \tag{A.8}$$

Left-multiplying eq. (A.8) by $\boldsymbol{\sigma}_{t-l}$, taking expectations on both sides under Assumption A2, and using $\mathbf{E}(\mathbf{D}_{t,l}) = \mathbf{C}^l$ (see eq. 14), yields

$$\boldsymbol{\Sigma}(l) = \boldsymbol{\sigma}\boldsymbol{\omega}'[(\mathbf{I}-\mathbf{C})^{-1}]'[\mathbf{I} - (\mathbf{C}^l)'] + \boldsymbol{\Sigma}(0)(\mathbf{C}^l)'. \tag{A.9}$$

On account of $\boldsymbol{\omega} = (\mathbf{I} - \mathbf{C})\boldsymbol{\sigma}$, it follows that

$$\boldsymbol{\Gamma}(l) = \boldsymbol{\Gamma}(0)(\mathbf{C}^l)'. \tag{A.10}$$

Taking the vec form of eq. (A.10) yields

$$\boldsymbol{\gamma}(l) = (\mathbf{C}^l)^{\otimes I}\boldsymbol{\gamma}(0),$$

as claimed. ■

B SUPPLEMENTARY APPENDIX

Table A.1: Data Description

		Sample period			r_t^2		RK_t	
Index symbol	Index name (Country)	Strart date	End date	Obs.	Avol	sd	Avol	sd
SP	S&P 500 (US)	03/01/2000	01/03/2013	3281	0.212	0.054	0.183	0.029
NIKKEI	NIKKEI 225 (Japan)	04/01/2000	01/03/2013	3184	0.250	0.073	0.174	0.020
TSE	S&P/TSX Comp. Index (Canada)	02/05/2002	01/03/2013	2701	0.183	0.041	0.128	0.016
FTSE	FTSE 100 (UK)	04/01/2000	01/03/2013	3301	0.197	0.044	0.159	0.017
DAX	DAX (Germany)	03/01/2000	01/03/2013	3333	0.253	0.070	0.227	0.035
EUSTOXX	EUROSTOXX 50 (Eurozone)	03/01/2000	01/03/2013	3325	0.248	0.062	0.216	0.035

Notes: Avol is the annualised volatility and sd is the standard deviation.

Table A.2: Wald Tests for the AP HEAVY-E-r Specification

Restrictions↓	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
$\delta_r = 1$	14.45 [0.00]	11.52 [0.00]	9.68 [0.00]	11.14 [0.00]	17.07 [0.00]	25.62 [0.00]
$\delta_r = 2$	31.03 [0.00]	28.77 [0.00]	20.72 [0.00]	21.67 [0.00]	13.92 [0.00]	11.12 [0.00]

Notes: The Table report values for ChiSq(1) tests.

The numbers in square brackets are p-values.

Table A.3: Wald Tests for the AP HEAVY-E-R Specification

Restrictions↓	SP	NIKKEI	TSE	FTSE	DAX	EUSTOXX
$\delta_R = 1$	4.27 [0.04]	4.20 [0.04]	73.02 [0.00]	40.32 [0.00]	17.03 [0.00]	30.13 [0.00]
$\delta_R = 2$	59.00 [0.00]	15.47 [0.00]	21.87 [0.00]	12.65 [0.00]	55.09 [0.00]	6.09 [0.01]

Notes: See notes in Table A.2.