University of Messina

DEPARTMENT OF ECONOMICS

Ph.D. Programme in Economics, Management and Statistics

XXXVI Cycle

Modelling the Covariance Dynamics of Multivariate Financial Time Series

SECS-P/05 - Econometrics

Candidate:
Emilija Dzuverovic

Supervisor:
Prof. Edoardo Otranto

A.Y. 2022/2023

Knowledge grows exponentially. The more we know, the greater our ability to learn, and the faster we expand our knowledge base.

- Dan Brown

Preface

Writing this thesis provided me the chance to meet many exceptional people, travel the world, and, most importantly, learn lots of things. In this regard, there are many people whom I need to acknowledge for their assistance and support with this journey.

First and foremost, I am extremely grateful to my supervisor, Prof. Edoardo Otranto, for his invaluable advice and patience during my Ph.D. studies. His immense knowledge and plentiful experience have encouraged me all the time in my research. I still remember Prof. Edoardo going through the R code on simulating the benchmark HAR model with me patiently and explaining everything clearly a couple of times. I would like to take this opportunity to additionally thank him for opening so many doors for me.

Next, I would like to thank my hosts at Université catholique de Louvain and University of Bologna, Prof. Luc Bauwens, Prof. Christian Hafner, and Prof. Matteo Barigozzi. I immensely enjoyed my visits to Louvain-la-Neuve and Bologna, and I am looking forward to our continued collaboration. Their support has helped me to grow as a researcher, learn and develop new skills, and, importantly, gain confidence in our field.

In particular, the knowledge of Prof. Luc about almost everything has taught me how little I still know. His ability to grasp things I have been looking at for weeks in about two minutes is inspiring. I hope to obtain his level of intuition. A special thanks to Prof. Matteo for providing me with heaps of invaluable advice on my career.

I am incredibly grateful to my grandmother, Milica, who taught me to embrace quantitative problems and challenge my beliefs. I extend my deepest gratitude to Giovanni for his incredible patience and love and to my sister Masa for being a pillar

of strength through all the highs and lows. Finally, I dedicate this thesis to my mother, Ljiljana, who has always been there to give me every opportunity to get the most out of my education and every other aspect of life. Without them, I would not have come this far.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Over the last three decades, the notion of the covariance matrix of asset returns has pervaded almost every financial aspect, and covariance forecasts are a central input to various financial applications such as asset pricing, portfolio selection, and risk management.

Nevertheless, covariances are not directly observable in the market. In this regard, a large body of literature studies the different estimators for the latent covariance matrix. Correspondingly, another stream exploits the distinct modelling techniques to effectively forecast the covariance matrix.

With the increased availability of high-frequency (HF) data for many financial assets, researchers have focused on investigating the rich inherent information to estimate the covariances. In particular, Andersen and Bollerslev (1998) introduce the realized volatility as an accurate proxy of latent volatility. Due to the non-latent feature, it can be observed and modelled directly via the time-series models. The corresponding extension to the multivariate setting, i.e., the realized covariance (RC) matrix, has been initially proposed by Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004).

In contrast to the univariate case, modelling and forecasting the covariances introduce two major technical challenges. The first is the so-called 'curse of dimensionality' issue. Specifically, most of the models have a number of parameters that scale with the number of considered assets, thus affecting the statistical efficiency of the estimation or even making the latter infeasible. The second challenge is the necessity of guaranteeing the positive-definiteness (PD-ness) of covariance matrices.

Given the increasing interconnectedness among financial markets and the rapidly accelerating degree of globalization, among others, this thesis aims at establishing and verifying new methods to model and forecast the covariance matrices, thus allowing for better analyses of financial problems and improved decision-making. The thesis consists of three studies.

In the first study, Chapter 2, we analyze empirically the asymptotic properties of the LS estimator with respect to widely utilized multivariate HAR models. Due to the emphasized 'curse of dimensionality' issue, an easy-to-implement framework contains the adoption of scalar versions of HAR-based models. This implies a strongly binding assumption of the analogous dynamics of all variances and covariances, or correlations. As such, Bauwens and Otranto (2020, 2023) propose HE–HAR models that allow for asset pair-specific and time-varying impact parameters using the element-by-element Hadamard exponential (HE) function of the matrix, thus adding great flexibility to scalar HAR dynamics via a unique additional parameter.

We initially verify asymptotically unbiased, efficient, and normal OLS estimates for benchmark HAR models, i.e., the vech-HAR model of Chiriac and Voev (2011) and the HAR-DRD model of Oh and Patton (2016), under three different distributions. Correspondingly, the results of the extensive Monte Carlo (MC) simulation experiment confirm both consistency and asymptotic normality for regular HAR parameters in HE–vech-HAR models. The nonlinear coefficient of the HE parameterization shows a degree of bias depending on the degree of nonlinearity of the distribution generat-

ing the data, i.e., multivariate Normal, Wishart, and Matrix F distributions. Such a problem further amplifies for HE–HAR-DRD variants, where even some conventional HAR coefficients appear biased and non-normal. In fact, using a more fat-tailed and asymmetric Matrix F data generation process (DGP) compared to Wishart implies a relatively small loss of satisfaction of asymptotic properties.

While distinct models have revealed that HF data provides important additional information for modelling and forecasting covariance matrices, Chapter 3 documents the importance of capturing the asymmetric responses of the RC to shocks, i.e., the 'leverage effect'. The first method of accounting for the 'asymmetry' effect is based on the decomposition of covariances via signs of daily returns and implies that, e.g., the conditional variance of an asset today is higher if the corresponding lagged daily return is negative than if it is positive. The second way is based on the estimates of realized variances and covariances defined via signs of underlying HF returns, i.e., measures known as realized semi-variances and semi-covariances proposed by Shephard and Sheppard (2010) and Bollerslev et al. (2020a), respectively.

To perform empirical evaluations of the introduced models, a HF dataset for the main S&P500 ETF and five large US stocks of the banking sector have been employed. Both the in-sample and forecasting results show that the asymmetric models significantly outperform the symmetric benchmark specification, i.e., the conditional autoregressive Wishart (CAW) model of Golosnoy et al. (2012), in all cases. Furthermore, the results strongly indicate that the 'daily' asymmetric models not only result in significantly better fitting but also tend to generate more accurate predictions than the models built upon the semi-covariance decomposition. Such a finding is not surprising since the close-to-close return incorporates the information of the overnight period, whereas the intra-daily returns do not. To corroborate this finding, we further confirm that the asymmetric models using the decomposition of the RC matrix based on the signs of daily *open-to-close* returns fit less well and have worse

forecasting performance than the *close-to-close* returns-based ones.

Factor models have been widely used both theoretically and empirically in finance to deal with the dimensionality problem. In this regard, the number of parameters in covariance matrix estimation is significantly reduced by assuming that a few factors can completely capture the cross-sectional risks.

Chapter 4 presents the third original contribution of the thesis, where we introduce a class of models for high-dimensional covariance matrices, i.e., HD DCC-HEAVY, by combining the hierarchical factor modelling approach of Hansen et al. (2014) and recent dynamic conditional correlation formulation of a HEAVY model (Noureldin et al. (2012)) introduced by Bauwens and Xu (2023). Such a framework is independent of the cross-sectional dimension of the assets under consideration, allowing for relatively straightforward estimation and forecasting schemes.

Given that there exists no evidence on the forecasting ability of the hierarchical-type factor models, we assess the statistical and economic performance of the distinct variants of our model in terms of the set of factors and attribution of asymmetric dynamics, comparing them with the cDCC model (Aielli (2013)) considered a benchmark, the Realized Beta GARCH model (Hansen et al. (2014)), and its 3-Fama-French (FF) extension (Archakov et al. (2020)).

An illustrative empirical study for the S&P500 constituents over the period from January 1962 until January 2023 shows that the model's forecasts consistently beat both the existing hierarchical models of Hansen et al. (2014) and Archakov et al. (2020), as well as the benchmark cDCC model. We confirm the robustness of our findings under changing market conditions.

Chapter 2

Nonlinear HAR Models and Nonlinear Least Squares: Asymptotic Properties

The contents of this chapter are the result of joint work with Prof. Edoardo Otranto (University of Messina). The paper has been accepted for publication in Advanced Methods in Statistics, Data Science, and Related Applications.

2.1 Introduction

Forecasting the covariance matrix of asset returns has paramount importance in asset pricing, portfolio allocation, and risk management. As such, the majority of introduced econometric models assume that variances and covariances are either assessable conditional upon past daily information, i.e., multivariate GARCH (MGARCH) models (Bauwens et al. (2006)), or latent, i.e., multivariate stochastic volatility (MSV) models (Asai et al. (2006)).

In contrast, a recent preferable realized covariance (RC) method allows treating the volatility as an 'observable' quantity via the high-frequency (HF) data (e.g., Andersen et al. (2003); Barndorff-Nielsen and Shephard (2004); Barndorff-Nielsen et al. (2008, 2011)). However, RC models generally face several challenges, and the most prominent include ensuring the positive definite (PD)-ness of forecasts coupled with the dimensionality issue. The latter implies the rapid increase in the number of parameters with the cross-sectional dimension that has a potential pronounced impact on the statistical efficiency of the estimation and often even renders the model estimation unfeasible. Hence, the RC modelling framework calls for parsimonious yet flexible parameterizations, capable of capturing complex serial dependencies observed in realized variances and covariances.

Among the HF data-based models, the multivariate extensions of the Heterogeneous Auto-Regressive (HAR) model (Corsi (2009)) are widely employed, i.e., vech-HAR model (Chiriac and Voev (2011)), HAR-DRD model (Oh and Patton (2016)). This class of models, originally inspired by the Heterogenous Market Hypothesis of Müller et al. (1993) and asymmetric propagation of volatility between long and short horizons, appears capable of capturing a commonly recognized long memory of RC via simple linear regression structures.

Nevertheless, in face of the emphasized 'curse of dimensionality' issue, an easy-to-implement framework contains the adoption of scalar versions of HAR-based models. This implies that all variances and covariances, or correlations, obey the same dynamics, an assumption that is obviously highly restrictive. In this regard, using the element-by-element Hadamard exponential (HE) function of the matrix, the asset pair-specific and time-varying impact coefficients of lagged realized variances and covariances, or correlations, could be defined, adding great flexibility in dynamics via a single parameter (Bauwens and Otranto (2020, 2023)).

A practical advantage of the HAR family of models is the straightforward estimation via OLS. By incorporating the HE extension, the nonlinear LS (NLS) could be adopted. Thus, great computational advantages in terms of simplicity, stability, and cost potentially arise on behalf of both flexible and parsimonious HE–HAR models compared to all other maximum likelihood (ML)-based RC models. It has therefore been natural to ask if the convincing NLS method would generate, at least asymptotically, unbiased, efficient, and normally distributed estimates. The approach taken to provide a reliable and empirically valid answer to this question contains the extensive Monte Carlo (MC) simulation experiment.

As the LS method has been extensively adopted within the nonlinear regression framework, the asymptotic properties of the estimator have been subject to notable research interest. In general, the consistency is theoretically established by Wu (1981) and Lai (1994), among others, whereas, e.g., Wooldridge (1986) and Pollard and Radchenko (2006) prove the asymptotic normality under certain regularity conditions.

Recently, Wang (2021) extrapolates analogous theoretical results to nonstationary and heteroscedastic models, which obviously contain a broad class of commonly used volatility models such as GARCH (Bollerslev et al. (1988)), time-varying GARCH (Rohan and Ramanathan (2013)), and nonlinear GARCH (Lanne and Saikkonen (2005)). Correspondingly, Dias (2013) provides the empirical evidence via MC study that the proposed nonlinear iterative LS (NL-ILS) estimator is consistent and even outperforms the ML benchmark for many GARCH-type models. Not to mention, Hamadeh and Zakoïan (2011) derive the proof of both consistent and asymptotically normal LS estimates of the power-transformed ARCH model (Hwang and Kim (2004)) under analogous assumptions as for standard ARCH models (Bose and Mukherjee (2003)).

Earlier, the rate of the NLS convergence and asymptotic distribution have been proved and empirically supported by Ciuperca (2011) for the long memory auto-regressive fractionally integrated moving-average (ARFIMA) model (Granger and Joyeux (1980); Hosking (1981)), widely advocated within the volatility modelling framework. In addition, Chang and Park (2003) prove the consistency of the NLS estimator for smooth transition regressions, often useful to capture the asymmetry in volatility dynamics. Nevertheless, some components have a non-Gaussian distribution that causes usual statistical methods to yield inefficient estimates and/or invalid tests.

Andreou (2016) compares the common LS regression models with corresponding mixed data sampling (MIDAS) regressions (Ghysels et al. (2005, 2006)), assuming the underlying MIDAS data generating process (DGP). The analytical and numerical results are presented for the asymptotic bias and relative efficiency of the slope estimator in these two classes for a number of alternative HF volatilities. They show that the LS estimator is asymptotically biased for all cases considered. In contrast, the MIDAS-NLS slope estimator turns out to be not only unbiased but also relatively more efficient. The analogous conclusions are achieved earlier by Andreou et al. (2010), where asymptotic properties of the FLAT-LS and MIDAS-NLS are tested in models that contain either the i.i.d., ARCH(1), or AR(1) HF regressors.

Chen et al. (2018) show that NLS estimates for the logarithmic auto-regressive conditional duration (log-ACD) models utilized for the HF data analyses are consistent and asymptotically normal, as well as strongly competitive with respect to the ones generated via the quasi-ML method (QMLE).

In this work, we investigate asymptotic properties of the NLS for the multivariate HE–HAR family of models (Bauwens and Otranto (2023)) via MC experiments. A decision to examine NLS estimates for the HE–HAR class can be justified by the possibility of having a very flexible but parsimonious model while maintaining the computational advantages of the HAR family. Initially, we adopt the vech-HAR

model of Chiriac and Voev (2011) in order to verify whether our exercise confirms the theoretically established asymptotic properties of linear OLS estimates (e.g., McAleer and Medeiros (2008); Hwang and Shin (2014)). Then, we consider the two HE-based specifications introduced by Bauwens and Otranto (2023), extending the vech-HAR model of Chiriac and Voev (2011) and HAR-DRD model of Oh and Patton (2016), under different distributional assumptions, i.e., multivariate Normal (applied to the log transformation of RC matrices), Wishart, and Matrix F distribution.

Our results show that the convergence is achieved for conventional HAR parameters, while there is a small-medium bias for the HE coefficient. Asymptotic normality seems satisfied only for the HE–vech-HAR specification with log-transformed RC matrices under Normal distribution.

The rest of the chapter is organized as follows. Section 2.2 introduces the HAR-type framework adopted for MC experiments; Section 2.3 presents the design of implemented simulations and provides a discussion on empirical results; Section 2.4 gives some concluding comments.

2.2 Modelling Framework

One of the most widespread models in the econometric literature for modelling the realized volatility is the HAR model of Corsi (2009); its linear auto-regressive form has a great advantage to be estimated via OLS. To account for not only serial but also cross-correlation dynamics across the elements of covariance matrices, due to the common movements among financial markets and the increasing degree of globalization in the economy, among others, recent literature has attempted to extend the most utilized univariate volatility models to the multivariate framework.

Let us denote the $m \times m$ RC matrix of day t as $\boldsymbol{C}_t$, the information set available at time $t-1$ as $\mathcal{F}_{t-1} = \{\boldsymbol{C}_{t-1}, \boldsymbol{C}_{t-2}, ...\}$, and the conditional expectation of $\boldsymbol{C}_t$, i.e., $\mathbf{E}(\boldsymbol{C}_t|\mathcal{F}_{t-1})$, as $\boldsymbol{S}_t$.

2.2.1 Linear vech-HAR Model

A natural extension of the HAR model to the multivariate setting was proposed by Chiriac and Voev (2011), i.e., the vech-HAR model:

$$\begin{aligned} \operatorname{vech}(\boldsymbol{C}_t) = \; & (1-\alpha_D-\alpha_W-\alpha_M)\overline{\boldsymbol{C}} + \alpha_D \operatorname{vech}(\boldsymbol{C}_{t-1}) + \alpha_W \operatorname{vech}(\overline{\boldsymbol{C}}_{t-1:t-5}) \\ & + \alpha_M \operatorname{vech}(\overline{\boldsymbol{C}}_{t-1:t-22}) + \boldsymbol{\varepsilon}_t, \end{aligned} \tag{2.1}$$

where $\operatorname{vech}(\cdot)$ is the operator that stacks the lower triangular portion of its matrix argument as a vector, $\overline{\boldsymbol{C}}$ is the sample mean of RC matrices, regressors $\overline{\boldsymbol{C}}_{t-1:t-h}$ are the averages of past RC matrices between $t-1$ and $t-h$ ($h = 5, 22$), and α's are scalar parameters. Finally, $\boldsymbol{\varepsilon}_t$ is a vector of disturbances.
The use of the so-called targeting, i.e., $(1-\alpha_D-\alpha_W-\alpha_M)\overline{\boldsymbol{C}}$, as a constant, guarantees that the estimated matrix $\boldsymbol{S}_t$ is positive-definite (PD), while at the same time reduces significantly the number of parameters to be estimated.

We consider three alternative specifications for the distribution of $\boldsymbol{\varepsilon}_t$, ensuring the PD-ness of the RC series:

- Multivariate Normal, which implies the log transformation of RC (Bauer and Vorkink (2011)), so that all variables in (2.1) are expressed in logarithms. We call this model "log vech-HAR";

- Wishart distribution (e.g., Golosnoy et al. (2012); Bonato et al. (2012); Bauwens et al. (2012)), with $\boldsymbol{C}_t|\mathcal{F}_{t-1} \sim W_m(v, \boldsymbol{S}_t/v)$, where $W_m(v, \boldsymbol{S}_t/v)$ denotes the m-dimensional central Wishart distribution with $v \geq m$ degrees of freedom and $m \times m$ PD scale matrix $\boldsymbol{S}_t/v$, i.e., $\mathbf{E}(\boldsymbol{C}_t|\mathcal{F}_{t-1})) = \boldsymbol{S}_t$. We call this model

"vech-HAR Wishart";

- Matrix F distribution (e.g., Opschoor et al. (2018); Vassallo et al. (2021)) due to the sudden large movements in RC that might potentially impair estimated dynamics of the series if a 'thin-tailed' distribution, such as Wishart, is adopted (e.g., Creal et al. (2011); Harvey (2013)). In this regard, we suppose $\boldsymbol{C}_t|\mathcal{F}_{t-1} \sim F_m(v_1, v_2, ((v_2 - m - 1)/v_1)\boldsymbol{S}_t)$, where $v_1, v_2 \geq m + 1$ are degrees of freedom. When v_2 is finite, the Matrix F distribution exhibits a leptokurtic behavior, typical of realized measures of variance and covariance. As $v_2 \to \infty$, the convergence to the Wishart distribution prevails. We call this model "vech-HAR Matrix F".

2.2.2 Nonlinear HE–vech-HAR Model

The adoption of the scalar vech-HAR model requires imposing a very strong assumption, i.e., all the variances and covariances obey the same dynamics. Bauwens and Otranto (2023) suggest a way of adding great flexibility in model dynamics while preserving the parsimonious parameterization, i.e., via a single additional parameter, using the HE function of the matrix.

In practical terms, let us consider again the simple scalar vech-HAR model (2.1), where each daily RC matrix is specified as a linear function of lagged daily, weekly, and monthly covariances. Considering the HE operator's algebraic properties, Bauwens and Otranto (2023) propose replacing the scalar impact coefficient α_D of the lagged daily RC on the current RC matrix by a PD matrix of time-varying coefficients determined via a unique new parameter.

The new HE–vech-HAR model is given by:

$$\begin{aligned}\operatorname{vech}(\boldsymbol{C}_t) = \; & (1-\overline{\alpha}_{D,t}-\alpha_W-\alpha_M)\overline{\boldsymbol{C}}+\alpha_D \exp\odot(\phi \operatorname{vech}(\boldsymbol{P}_{t-1}-\mathbf{J}_m))\odot \operatorname{vech}(\boldsymbol{C}_{t-1}) \\ & +\alpha_W \operatorname{vech}(\overline{\boldsymbol{C}}_{t-1:t-5})+\alpha_M \operatorname{vech}(\overline{\boldsymbol{C}}_{t-1:t-22})+\boldsymbol{\varepsilon}_t,\end{aligned} \tag{2.2}$$

where $\odot$ denotes the Hadamard product and $\exp\odot$ the HE operator, i.e., element-by-element exponential; ϕ is the new parameter and $\overline{\alpha}_{D,t}$ is the average of the entries of $\alpha_D \exp\odot(\phi \operatorname{vech}(\boldsymbol{P}_{t-1}-\mathbf{J}_m))$, where $\boldsymbol{P}_t$ is the realized correlation matrix and $\mathbf{J}_m$ is a $m\times m$ matrix of ones. Thus, if ϕ differs from zero, the impact coefficient becomes both asset pair-specific and time-varying. Indeed, the HE operator applied to a PD matrix guarantees the PD-ness of the resulting matrix, whereas the Hadamard product of two PD matrices is again PD.

The adoption of lagged correlations in the impact parameter can be justified by the strong evidence that volatilities and correlations move together (e.g., Andersen et al. (2001)). As such, we expect the transfer of the volatility clustering phenomenon to correlations as well. During turbulent periods, the correlations increase, but the level of persistence of such an increase can be distinct across asset pairs. Hence, by adding a dependence of the impact coefficient on lagged correlations, we are able to account for the volatility clustering effect on RC that differs across asset pairs and changes through time.

The vech-HAR model with the HE extension could be estimated by NLS. In this regard, we investigate the asymptotic properties of HE–vech-HAR NLS estimates generated via the Gauss-Newton algorithm. Again, we consider RC in log form, thus following the Normal distribution, conditional Wishart, as well as conditional Matrix F distributed ones. The resulting models are denoted by "log HE–vech-HAR", "HE–vech-HAR Wishart", and "HE–vech-HAR Matrix F", respectively.

2.2.3 Nonlinear HE–HAR-DRD Model

Inspired by the seminal DCC model of Engle (2002), Oh and Patton (2016) introduce the widely utilized HAR-DRD. We begin by decomposing the RC matrix as:

$$\boldsymbol{C}_t = \boldsymbol{V}_t^{1/2} \boldsymbol{P}_t \boldsymbol{V}_t^{1/2}, \tag{2.3}$$

where $\boldsymbol{V}_t$ is the diagonal matrix of realized variances (RV) and $\boldsymbol{P}_t$ is the realized correlation matrix, which allows for estimating the full model in two steps (see Engle (2002)).

In particular, the first step consists of modelling m univariate volatilities, i.e., $\boldsymbol{V}_{ii,t}$ $(i = 1, \ldots, m)$, or their logs to ensure the positiveness, via the benchmark HAR model (Corsi (2009)). Hence, the OLS estimator applies to:

$$\boldsymbol{V}_{ii,t} = c_i + \alpha_{i,D} \boldsymbol{V}_{ii,t-1} + \alpha_{i,W} \overline{\boldsymbol{V}}_{ii,t-1:t-5} + \alpha_{i,M} \overline{\boldsymbol{V}}_{ii,t-1:t-22} + \varepsilon_{i,t}. \tag{2.4}$$

Subsequently the benchmark HAR-type structure is adopted to model realized correlations, i.e.,

$$\begin{aligned} \operatorname{vech}(\boldsymbol{P}_t) = (1 - \alpha_D - \alpha_W - \alpha_M)\overline{\boldsymbol{P}} + \alpha_D \operatorname{vech}(\boldsymbol{P}_{t-1}) + \alpha_W \operatorname{vech}(\overline{\boldsymbol{P}}_{t-1:t-5}) \\ + \alpha_M \operatorname{vech}(\overline{\boldsymbol{P}}_{t-1:t-22}) + \boldsymbol{\varepsilon}_t, \end{aligned} \tag{2.5}$$

where, similarly to (2.1), regressors are the averages of past realized correlation matrices, $\overline{\boldsymbol{P}}$ is the sample mean of the realized correlation series, and α_i $(i = D, W, M)$ are scalar parameters to be estimated by OLS.

To allow for distinct dynamics of realized correlations, Bauwens and Otranto (2023) introduce the HE parameterization to (2.5), i.e.,

$$\begin{aligned}\operatorname{vech}(\boldsymbol{P}_t) = (1 - \overline{\alpha}_{D,t} - \alpha_W - \alpha_M)\overline{\boldsymbol{P}} + \alpha_D \exp \odot(\phi \operatorname{vech}(\boldsymbol{P}_{t-1} - \mathbf{J}_m)) \odot \operatorname{vech}(\boldsymbol{P}_{t-1}) \\ + \alpha_W \operatorname{vech}(\overline{\boldsymbol{P}}_{t-1:t-5}) + \alpha_M \operatorname{vech}(\overline{\boldsymbol{P}}_{t-1:t-22}) + \boldsymbol{\varepsilon}_t.\end{aligned} \tag{2.6}$$

Similarly to HE–vech-HAR, HAR-DRD model with the HE extension could be estimated by NLS. As follows, we examine the asymptotic properties of HE–HAR-DRD NLS estimates generated via the Gauss-Newton algorithm for "log HE–HAR-DRD", "HE–HAR-DRD Wishart", and "HE–HAR-DRD Matrix F" specifications.

2.3 Monte Carlo Experiments

For the subsequent empirical analyses, we employ the three HAR models under three distributional assumptions, i.e., a total of 9 specifications. Specifically, 3 linear vech-HAR (2.1), 3 nonlinear HE–vech-HAR (2.2), and 3 nonlinear HE–HAR-DRD specifications (2.3, 2.4, 2.6), distinguished by the distribution, i.e., Normal for logs, Wishart, and Matrix F.

To evaluate the asymptotic properties of OLS/NLS, we generate data from each specification referring to a 3–variate time series ($m = 3$) of increasing length $T = 1000, 3000, 5000, 10000, 50000$, with the corresponding parameters fixed in the range of estimated values on real data for similar models.[1] In particular, for vech-HAR, we consider the range of values presented in Tables 1-3 in Buccheri and Corsi (2021); for HE–vech-HAR, the values are taken from Table 14 of Bauwens and Otranto (2023), but fixed $\phi = 0.15$ to ensure PD matrices;[2] for HE–HAR-DRD, we consider Table 2 of Bollerslev et al. (2018) for the HAR part and keep $\phi = 0.15$. The sample covariance

[1] We perform empirical exercises via the R software.

[2] In the paper of Bauwens and Otranto (2023), this coefficient is negative.

matrix for each model is randomly generated from the underlying distribution. For Wishart, we set $v = 4$ degrees of freedom, coupled with $v_1 = 22$ and $v_2 = 35$ for Matrix F. We replicate each experiment $N = 1000$ times.

Concerning the RC in log form, we initially generate a disturbance term $\boldsymbol{\varepsilon}_t$ from a standardized Normal distribution, with the series of $\boldsymbol{C}_t$ obtained iteratively from a model considered, i.e., vech-HAR (2.1), HE–vech-HAR (2.2), HE–HAR-DRD (2.6). Instead, disturbances do not enter models under Wishart and Matrix F distributions, where RC matrices are generated by the conditional distribution of $\boldsymbol{C}_t$. More in detail, for both cases, we first generate the unconditional $\overline{\boldsymbol{C}}$ discussed above, with the expectation equal to the identity matrix.

Then, for the Wishart case, we simulate $\boldsymbol{S}_t$ from one of the three models, such that the RC matrix $\boldsymbol{C}_t$ at time t is generated from $W_3(4, \boldsymbol{S}_t/4)$.

Analogously, for the Matrix F case, we simulate $\boldsymbol{S}_t$ from one of the three models. Then, $\boldsymbol{C}_t$ follows from $F_3(v_1, v_2, ((v_2 - 3 - 1)/v_1)\boldsymbol{S}_t)$, with $v_1 = 22$ and $v_2 = 35$. To guarantee both symmetry and PD-ness of the generated RC, we adopt the subsequent decomposition property (Konno (1991)):

$$\boldsymbol{C}_t = ((v_2 - m - 1)/v_1)\boldsymbol{S}_t^{1/2}\boldsymbol{M}_t^{1/2}\boldsymbol{L}_t^{-1}\boldsymbol{M}_t^{1/2}\boldsymbol{S}_t^{1/2}, \tag{2.7}$$

where $\boldsymbol{M}_t$ and $\boldsymbol{L}_t$ are independent Wishart matrices, i.e., $\boldsymbol{M}_t \sim W_m(v_1, \mathrm{I}_m)$, $\boldsymbol{L}_t \sim W_m(v_2, \mathrm{I}_m)$.

For HE–HAR–DRD specifications, we do not consider the part relative to the RV, so that we generate directly realized correlations $\boldsymbol{P}_t$ from the corresponding conditional distribution, putting $\boldsymbol{S}_t = \boldsymbol{R}_t$, and implementing the same steps previously described.

In the following tables, reporting results of our experiments, the second column shows the 'true' value of each parameter. To evaluate the unbiasedness and efficiency, we report the average and standard deviation of N estimates for each parameter and T, as well as the degree of biasedness. Ultimately, we show the p-value of the Jarque-Bera test (Jarque and Bera (1980)) on the set of N estimates for each coefficient to evaluate the normality.

In what follows, we comment results for each DGP.

2.3.1 vech-HAR Model

We run experiments on the discussed linear vech-HAR variants to confirm the theoretical asymptotic properties under the Normal DGP and verify their validity under the other two distributional assumptions.

Table 2.1 shows the corresponding results. For the log and Wishart cases, we confirm the unbiasedness of α_D and α_W already at small sample sizes, whereas α_M converges to the true value at $T = 3000$ for the Wishart case and $T = 5000$ for the log case. The decrease in standard deviations is consistent with the efficiency of the OLS estimator, and the normality of estimates is also satisfied in small samples. A different behavior characterizes the Matrix F case with some puzzling results. The convergence seems to be achieved only asymptotically, and the normality of estimators is ambiguous, in particular for the parameter α_M. In practice, the strong asymmetric and leptokurtic form of the Matrix F distribution may affect the goodness of the OLS asymptotic performance.

2.3.2 HE–vech-HAR Model

The NLS estimates for HE–vech-HAR models are obtained by utilizing the Gauss-Newton method, which is based on the Taylor series approximation to obtain the linearized model and solve it via OLS (Greene (2018)). Such a choice might be sup-

ported not only because this algorithm avoids computing the second-order derivatives but also due to the absence of potential issues arising from initial starting values provided for estimations given the specified simulation parameters.

As indicated in Table 2.2, all NLS estimates of regular HAR parameters within HE–vech-HAR models reach the convergence for the log and Wishart cases. The full convergence of all coefficients prevails for the log case, i.e., at $T = 5000$, coupled with the continuous reduction of respective standard deviations. Correspondingly, the estimate of the HE term coefficient ϕ is featured by a relatively small but persistent negative bias around 2%.

Concerning the Matrix F case, the partial convergence of conventional HAR parameters is attained by $T = 5000$, while the monthly coefficient α_M subsequently fully converges as well. At the same time, the bias does not exceed 1.3%, being less than 0.5% in the majority of cases. Similar comments with respect to the two previous cases hold for regular HAR coefficients, while the bias of ϕ has an opposite sign but higher maximum magnitude, i.e., around 6%.

Due to the absence of the overall consistency evidence for NLS estimates of HE–vech-HAR variants and in line with related literature, the underlying distribution of estimated coefficients often diverges from Normal. On the other hand, we confirm the normality hypothesis for all parameter estimates, even at $T = 5000$, in log model.

2.3.3 HE–HAR-DRD Model

The MC results for HE–HAR-DRD variants reported in Table 2.3 suggest a relatively inferior NLS performance compared to the estimation of HE–vech-HAR models.

The regular HAR coefficients only occasionally hit the 'true' value under Wishart and Matrix F DGP. Considering the Normal, the coefficient α_D exhibits bias for all T. The HE term estimator is generally characterized by not only protracted but also notable bias, close to 10% at $T = 50000$ for the Matrix-F case and more than the

Table 2.1: vech-HAR Model: $N = 1000$ MC results in correspondence to three different distributions

DGP	Parameter/Sample Size	$T = 1000$	$T = 3000$	$T = 5000$	$T = 10000$	$T = 50000$
	$\alpha_D =$ **0.45**					
	Sample Mean	**0.45**	**0.45**	**0.45**	**0.45**	**0.45**
	SD	(0.00114)	(0.00065)	(0.00050)	(0.00035)	(0.00016)
	Bias	-0.239%	-0.142%	-0.110%	-0.112%	-0.013%
	Jarque-Bera test (p-value)	**0.8037**	**0.5434**	**0.6151**	**0.3466**	**0.3127**
	$\alpha_W =$ **0.25**					
	Sample Mean	0.24	**0.25**	**0.25**	**0.25**	**0.25**
Normal	SD	(0.00173)	(0.00101)	(0.00079)	(0.00055)	(0.00025)
	Bias	-0.560%	-0.207%	-0.079%	-0.020%	-0.079%
	Jarque-Bera test (p-value)	**0.1051**	**0.9129**	**0.2670**	**0.9632**	**0.1026**
	$\alpha_M =$ **0.15**					
	Sample Mean	0.14	0.14	**0.15**	**0.15**	**0.15**
	SD	(0.00174)	(0.00103)	(0.00080)	(0.00055)	(0.00024)
	Bias	-1.452%	-0.519%	-0.347%	-0.177%	-0.030%
	Jarque-Bera test (p-value)	**0.9688**	**0.8918**	**0.8063**	**0.4520**	**0.8113**
	$\alpha_D =$ **0.45**					
	Sample Mean	**0.45**	**0.45**	**0.45**	**0.45**	**0.45**
	SD	(0.00115)	(0.00065)	(0.00050)	(0.00034)	(0.00016)
	Bias	-0.198%	-0.159%	-0.140%	-0.068%	-0.047%
	Jarque-Bera test (p-value)	**0.4210**	**0.8894**	**0.2636**	**0.8559**	**0.1866**
	$\alpha_W =$ **0.25**					
	Sample Mean	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**
Wishart	SD	(0.00176)	(0.00102)	(0.00080)	(0.00055)	(0.00025)
	Bias	-0.126%	-0.181%	-0.100%	-0.122%	-0.030%
	Jarque-Bera test (p-value)	**0.9815**	**0.6958**	**0.7642**	**0.3186**	**0.6312**
	$\alpha_M =$ **0.15**					
	Sample Mean	0.13	**0.15**	**0.15**	**0.15**	**0.15**
	SD	(0.00180)	(0.00105)	(0.00080)	(0.00055)	(0.00024)
	Bias	-1.547%	-0.272%	-0.147%	-0.054%	-0.023%
	Jarque-Bera test (p-value)	**0.1738**	**0.5051**	**0.8959**	0.0413	**0.4031**
	$\alpha_D =$ **0.45**					
	Sample Mean	0.49	0.47	**0.45**	0.46	0.44
	SD	(0.00069)	(0.00041)	(0.00037)	(0.00026)	(0.00018)
	Bias	4.357%	1.616%	0.296%	0.622%	-0.560%
	Jarque-Bera test (p-value)	0.0008	**0.1099**	**0.9438**	0.0209	0.0025
	$\alpha_W =$ **0.25**					
	Sample Mean	0.23	0.24	**0.25**	0.24	**0.25**
Matrix F	SD	(0.00055)	(0.00042)	(0.00031)	(0.00021)	(0.00015)
	Bias	-2.474%	-0.686%	-0.254%	-1.409%	-0.185%
	Jarque-Bera test (p-value)	0.0222	**0.5801**	0.0322	0.0011	0.0000
	$\alpha_M =$ **0.15**					
	Sample Mean	0.13	0.13	0.14	**0.15**	0.14
	SD	(0.00056)	(0.00030)	(0.00020)	(0.00019)	(0.00016)
	Bias	-2.097%	-2.183%	-0.986%	-0.406%	-1.085%
	Jarque-Bera test (p-value)	**0.3034**	0.0000	0.0084	0.0004	0.0000

Notes: The table reports the sample mean of 1000 OLS estimates of each vech-HAR parameter with the corresponding standard deviation in parentheses (at sample sizes $T = 1000, 3000, 5000, 10000, 50000$). The bias is calculated as the difference between the sample mean and the true parameter value, multiplied by 100. We also show the p-value of the Jarque-Bera test for each coefficient and respective sample size, where the entries in boldface indicate the failure to reject the null hypothesis of normality at the 5% significance level.

negative 8% for the log case. The estimates do not follow Normal distribution, also asymptotically, excluding α_W and α_M in "log HE–HAR-DRD".

2.4 Final Remarks

To accurately capture the cross-correlation dynamics among assets and volatility spillovers in financial markets via the HF data, the multivariate HAR-type models have been widely adopted. This class of models appears particularly attractive, being featured by a solid, empirically proved forecasting performance (e.g., Taylor (2015); Oh and Patton (2016); Cubadda et al. (2017); Vassallo et al. (2021)), coupled with a rather simple LS estimation method.

As such, this paper presents the empirical evidence on the asymptotic properties of the LS estimator, i.e., consistency and normality, with respect to both linear and nonlinear HAR-type framework. To the best of our knowledge, this is the first study that not only verifies empirically consistent and asymptotically Normal OLS estimates for benchmark HAR models but also investigates the properties of the NLS for flexible yet parsimonious nonlinear HAR extensions. In particular, we consider adding the HE extension (Bauwens and Otranto (2020, 2023)) to HAR benchmarks, which implies the asset pair-specific and time-varying impact parameters of the lagged RC, or correlations, adding great flexibility to scalar HAR dynamics via a unique additional parameter. Thus, it enables to adequately counteract the well-known 'curse of dimensionality' issue of the RC modelling framework. Not to mention great computational advantages in terms of the simplicity, stability, and cost on behalf of HE–HAR models with respect to all ML-based RC models.

Based on $N = 1000$ MC replication outcomes, we aim at verifying consistent, efficient, and asymptotically Normal OLS/NLS estimates for benchmark multivariate HAR specifications (with covariance targeting), i.e., vech-HAR models, as well as their HE version, under the three different distributions. Moreover, we consider the HE–HAR-DRD specification under analogous distributional assumptions, as it provides the possibility of a practically convenient 2-step estimation, separating the estimation of the RV from the estimation of the correlation part. All nonlinear HE

extensions are estimated by NLS, utilizing the Gauss-Newton algorithm.

Our experiments confirm both consistency and asymptotic normality of the OLS for vech-HAR variants, except for the Matrix F DGP. In fact, using a stronger leptokurtic and asymmetric distribution than Wishart to generate data, we notice a (small) loss of satisfaction of asymptotic properties. Similar results are obtained for regular HAR parameters in HE–vech-HAR models. The nonlinear coefficient of the HE parameterization shows a degree of biasedness, which increases with the degree of nonlinearity of the distribution generating data. This problem further amplifies for HE–HAR-DRD specifications, where even some conventional HAR estimates exhibit biasedness and non-normality.

Given the breadth of the HAR modelling framework and their widespread and straightforward adoption in the context of the accurate RC matrix forecasting, these results provide a valuable extension to existing ones in the literature and immediate practical implications. In general, we believe that it is important to highlight advantages and problems in the performance of HAR estimators and to draw the researchers' attention to potential pitfalls in using them in empirical studies. Thus, our findings suggest that researchers and practitioners can confidently adopt linear HAR models estimated by OLS to obtain the RC estimates. In addition, they might implement the NLS for the HE–HAR class with caution.

However, an avenue we have not explored in this study is building up alternative flexible but still parsimonious multivariate HAR extensions subject to the NLS estimation. E.g., incorporating the smooth transition structure within benchmark HAR to explicitly account for the asymmetric behavior in RV and RC (see McAleer and Medeiros (2008); Qu et al. (2016)). Hence, it remains completely an open empirical question whether the consistency and asymptotic normality of the NLS could be confirmed for upcoming nonlinear HAR models for the RC matrix, which arguably reflect interesting topics for further research.

Table 2.2: HE–vech-HAR Model: $N = 1000$ MC results in correspondence to three different distributions

DGP	Parameter/Sample Size	$T = 1000$	$T = 3000$	$T = 5000$	$T = 10000$	$T = 50000$
	$\alpha_D =$ **0.21**					
	Sample Mean	**0.21**	**0.21**	**0.21**	**0.21**	**0.21**
	SD	(0.00060)	(0.00034)	(0.00026)	(0.00019)	(0.00009)
	Bias	-0.274%	-0.335%	-0.301%	-0.275%	-0.266%
	Jarque-Bera test (p-value)	**0.0579**	**0.4740**	**0.1157**	0.0050	**0.1787**
	$\alpha_W =$ **0.33**					
	Sample Mean	**0.33**	**0.33**	**0.33**	**0.33**	**0.33**
	SD	(0.00099)	(0.00057)	(0.00043)	(0.00031)	(0.00014)
	Bias	-0.041%	0.047%	0.015%	-0.005%	-0.004%
Normal	Jarque-Bera test (p-value)	**0.8834**	**0.9225**	**0.7298**	**0.3233**	**0.1471**
	$\alpha_M =$ **0.22**					
	Sample Mean	0.21	**0.22**	**0.22**	**0.22**	**0.22**
	SD	(0.00115)	(0.00062)	(0.00049)	(0.00035)	(0.00016)
	Bias	-1.327%	-0.431%	-0.249%	-0.089%v	-0.007%
	Jarque-Bera test (p-value)	**0.8069**	**0.1961**	**0.3642**	**0.8022**	**0.2799**
	$\phi =$ **0.15**					
	Sample Mean	0.14	0.13	0.13	0.13	0.13
	SD	(0.00327)	(0.00185)	(0.00143)	(0.00100)	(0.00047)
	Bias	-1.005%	-1.621%	-1.757%	-1.836%	-2.082%
	Jarque-Bera test (p-value)	0.0000	0.0379	**0.5345**	**0.7134**	**0.7335**
	$\alpha_D =$ **0.21**					
	Sample Mean	**0.21**	0.18	0.20	**0.21**	0.19
	SD	(0.00068)	(0.00056)	(0.00043)	(0.00041)	(0.00028)
	Bias	0.493%	-2.737%	-1.099%	0.202%	-1.525%
	Jarque-Bera test (p-value)	0.0000	0.0113	0.0001	0.0000	0.0000
	$\alpha_W =$ **0.33**					
	Sample Mean	0.31	**0.33**	0.34	**0.33**	0.34
	SD	(0.00096)	(0.00081)	(0.00049)	(0.00059)	(0.00035)
	Bias	-2.208%	0.068%	0.891%	0.360%	0.770%
Wishart	Jarque-Bera test (p-value)	0.0008	0.0000	0.0114	0.0055	0.0000
	$\alpha_M =$ **0.22**					
	Sample Mean	0.19	0.19	0.19	0.21	**0.22**
	SD	(0.00173)	(0.00117)	(0.00088)	(0.00073)	(0.00017)
	Bias	-3.121%	-2.985%	-3.143%	-1.482%	0.032%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\phi =$ **0.15**					
	Sample Mean	0.11	0.14	0.12	0.13	0.12
	SD	(0.00399)	(0.00342)	(0.00218)	(0.00145)	(0.00114)
	Bias	-3.634%	-1.405%	-2.519%	-2.396%	-3.215%
	Jarque-Bera test (p-value)	**0.3721**	0.0000	0.0000	0.0000	0.0016
	$\alpha_D =$ **0.21**					
	Sample Mean	0.24	0.22	**0.21**	0.22	**0.21**
	SD	(0.00031)	(0.00034)	(0.00032)	(0.00020)	(0.00016)
	Bias	3.269%	0.936%	0.436%	1.250%	-0.288%
	Jarque-Bera test (p-value)	0.0491	0.0000	**0.1033**	0.0000	0.0000
	$\alpha_W =$ **0.33**					
	Sample Mean	**0.33**	0.34	0.34	0.32	**0.33**
	SD	(0.00056)	(0.00041)	(0.00029)	(0.00021)	(0.00012)
	Bias	0.325%	0.985%	0.505%	-0.601%	0.052%
Matrix F	Jarque-Bera test (p-value)	0.0000	0.0182	0.0153	**0.2053**	0.0000
	$\alpha_M =$ **0.22**					
	Sample Mean	0.20	0.20	**0.22**	**0.22**	**0.22**
	SD	(0.00073)	(0.00052)	(0.00032)	(0.00027)	(0.00017)
	Bias	-2.456%	-1.755%	-0.043%	0.052%	-0.413%
	Jarque-Bera test (p-value)	0.0000	0.0000	**0.1033**	0.0000	0.0038
	$\phi =$ **0.15**					
	Sample Mean	0.21	0.16	0.17	0.21	0.14
	SD	(0.00209)	(0.00173)	(0.00128)	(0.00080)	(0.00084)
	Bias	5.799%	1.183%	2.042%	6.319%	-1.020%
	Jarque-Bera test (p-value)	**0.1970**	0.0000	0.0000	0.0000	0.0000

Notes: The table reports the sample mean of 1000 NLS estimates of each HE–vech-HAR parameter with the corresponding standard deviation in parentheses (at sample sizes $T = 1000, 3000, 5000, 10000, 50000$). The bias is calculated as the difference between the sample mean and the true parameter value, multiplied by 100. We also show the p-value of the Jarque-Bera test for each coefficient and respective sample size, where the entries in boldface indicate the failure to reject the null hypothesis of normality at the 5% significance level.

Table 2.3: HE–HAR-DRD Model: $N = 1000$ MC results in correspondence to three different distributions

DGP	Parameter/Sample Size	$T = 1000$	$T = 3000$	$T = 5000$	$T = 10000$	$T = 50000$
	$\alpha_D =$ **0.05**					
	Sample Mean	0.09	0.07	0.07	0.06	0.06
	SD	(0.01248)	(0.00265)	(0.00185)	(0.00086)	(0.00034)
	Bias	4.130%	2.370%	2.028%	1.002%	0.517%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_W =$ **0.16**					
	Sample Mean	**0.16**	**0.16**	**0.16**	**0.16**	**0.16**
	SD	(0.00185)	(0.00088)	(0.00068)	(0.00047)	(0.00021)
	Bias	0.273%	0.106%	0.149%	-0.029%	-0.005%
Normal	Jarque-Bera test (p-value)	**0.5544**	**0.9242**	**0.3741**	**0.9167**	**0.5518**
	$\alpha_M =$ **0.56**					
	Sample Mean	0.55	0.55	0.55	**0.56**	**0.56**
	SD	(0.00199)	(0.00101)	(0.00080)	(0.00057)	(0.00026)
	Bias	-1.056%	-0.927%	-0.604%	-0.303%	-0.032%
	Jarque-Bera test (p-value)	**0.5414**	**0.9225**	**0.7957**	**0.2093**	**0.6409**
	$\phi =$ **0.15**					
	Sample Mean	0.06	0.02	0.12	0.08	0.07
	SD	(0.04148)	(0.02834)	(0.02192)	(0.01367)	(0.00599)
	Bias	-8.599%	-12.743%	-3.256%	-7.105%	-8.307%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	**0.1189**	**0.5491**
	$\alpha_D =$ **0.05**					
	Sample Mean	0.15	0.03	**0.05**	0.04	**0.05**
	SD	(0.00645)	(0.00076)	(0.00066)	(0.00033)	(0.00011)
	Bias	10.211%	-2.272%	0.039%	-1.166%	-0.290%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_W =$ **0.16**					
	Sample Mean	0.23	0.17	**0.16**	0.17	0.15
	SD	(0.00131)	(0.00129)	(0.00085)	(0.00086)	(0.00027)
	Bias	7.146%	1.422%	0.049%	0.971%	-0.780%
Wishart	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_M =$ **0.56**					
	Sample Mean	0.45	0.52	0.54	0.55	**0.56**
	SD	(0.00061)	(0.00011)	(0.00006)	(0.00046)	(0.00075)
	Bias	-11.009%	-4.354%	-2.055%	-1.070%	0.190%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\phi =$ **0.15**					
	Sample Mean	1.27	-0.47	0.04	0.14	0.10
	SD	(0.05209)	(0.01711)	(0.00742)	(0.00286)	(0.00110)
	Bias	111.601%	-62.171%	-11.331%	-0.750%	-4.610%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_D =$ **0.05**					
	Sample Mean	0.08	0.06	0.07	0.07	0.06
	SD	(0.00030)	(0.00093)	(0.00028)	(0.00036)	(0.00024)
	Bias	3.285%	1.128%	1.945%	1.829%	1.078%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_W =$ **0.16**					
	Sample Mean	0.17	0.15	0.15	**0.16**	0.17
	SD	(0.00005)	(0.00002)	(0.00008)	(0.00014)	(0.00022)
	Bias	1.497%	-0.578%	-1.488%	0.471%	1.088%
Matrix F	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\alpha_M =$ **0.56**					
	Sample Mean	0.51	0.52	**0.56**	0.55	0.54
	SD	(0.00193)	(0.00089)	(0.00081)	(0.00068)	(0.00056)
	Bias	-5.058%	-3.662%	0.055%	-0.634%	-1.558%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
	$\phi =$ **0.15**					
	Sample Mean	0.51	-0.05	0.18	0.39	0.25
	SD	(0.00273)	(0.01592)	(0.00925)	(0.00036)	(0.00041)
	Bias	36.419%	-19.616%	3.304%	24.420%	9.630%
	Jarque-Bera test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: The table reports the sample mean of 1000 NLS estimates of each HE–HAR-DRD parameter with the corresponding standard deviation in parentheses (at sample sizes $T = 1000, 3000, 5000, 10000, 50000$). The bias is calculated as the difference between the sample mean and the true parameter value, multiplied by 100. We also show the p-value of the Jarque-Bera test for each coefficient and respective sample size, where the entries in boldface indicate the failure to reject the null hypothesis of normality at the 5% significance level.

Chapter 3

Asymmetric Models for Realized Covariances

The contents of this chapter are the result of joint work with Prof. Luc Bauwens (Université catholique de Louvain) and Prof. Christian Hafner (Université catholique de Louvain).

3.1 Introduction

Forecasts of the covariance matrix of asset returns are a central input to asset pricing, portfolio allocation, and risk management decisions. Such forecasts can be computed using a multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) model – see Bauwens et al. (2006) for a survey – that specify the unobserved covariance matrix as a function of past (usually daily) returns. Forecasts can also be based on models for realized covariance (RC) matrices, which are 'observable' measures of variances and covariances based on high-frequency (intraday) returns, see e.g., Andersen et al. (2003), Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen et al. (2008), Barndorff-Nielsen et al. (2011). Several types of RC matrix (RCM) models have been introduced in the literature, each facing the need to ensure

the positive definite(PD)-ness of the RCM forecasts and to avoid parameter proliferation. In particular, the conditional autoregressive Wishart (CAW) class of models specifies a probability distribution for the RCM, such that its conditional expectation is a parametric function of past RC matrices (Gouriéroux et al. (2009)). The parameterization of this function is broadly similar to that of MGARCH models, in particular the BEKK (Baba-Engle-Kent-Kroner)-type (Engle and Kroner (1995); Golosnoy et al. (2012)) and DCC (dynamic conditional correlation)-type specifications (Engle (2002); Bauwens et al. (2012)).[1]

Both RCM and MGARCH models are designed to capture the main properties of the time series of covariance matrices of asset returns, corresponding to the clustering and the persistence of the volatilities of financial returns. Another stylized fact, specific to stock returns, is the negative correlation between returns and volatilities, initially expounded by Black (1976) and developed by Christie (1982). Based on the Modigliani-Miller framework, they explain that an unexpected stock price drop raises the debt-to-equity ratio, i.e., leverage, of a firm, which implies increased riskiness and higher volatility. The alternative interpretation, commonly referred to as the volatility feedback effect, was proposed by French et al. (1987) – see also Campbell and Hentschel (1992) and Wu (2001). It is based on the evidence of a positive correlation between future volatility and market risk premium. I.e., in the occurrence of an expected volatility increase, the risk premium increases as well, such that the risk adverse investors sell the stock, putting a downward pressure on its price. Whatever interpretation is favoured, it implies that the volatility increases more strongly after a negative unexpected return than after a positive one of the same magnitude, what has been named 'asymmetry' in volatility or (with some misuse of language) the 'leverage effect'.

[1]Most other RCM models are multivariate generalizations of the heterogenous autoregressive (HAR) model of Corsi (2009), such as the vech-HAR model derived from Chiriac and Voev (2011) and the HAR-DRD model of Oh and Patton (2016).

In this regard, while this type of asymmetry has been widely and diversely incorporated in the specification of GARCH models, such as the widely used GJR-GARCH model of Glosten et al. (1993), this is less the case for RCM models.[2] Our research contribution consists in developing a class of models for RC matrices based on the BEKK-type CAW model of Golosnoy et al. (2012), to capture the asymmetric responses of the elements of the RCM to shocks, and empirically assessing and comparing these models.

We introduce the 'asymmetry' effect in RC models in two ways. The first one consists of adding terms to the benchmark specification, which are active if the daily returns at $t-1$ are negative. This is designed in such a way that the conditional variance of an asset at date t is higher if the daily return of the asset is negative at date $t-1$ than if it is positive. Likewise, the covariance at t between two assets is higher if the returns are both negative at date $t-1$ than if at least one of them is positive. The conditional threshold autoregressive Wishart (CTAW) model of Anatolyev and Kobotaev (2018) is of this type. We propose a more flexible model in Section 3.2.1. The signs of returns to specify an 'asymmetry' effect have also been used, with some noticeable differences, in MGARCH (Kroner and Ng (1998); De Goeij and Marquering (2004); Cappiello et al. (2006); Audrino and Trojani (2011); Francq and Zakoïan (2012)), HAR (Qu and Zhang (2022)), and HEAVY models (Bauwens and Xu (2023)).

The second way to introduce the 'asymmetry' effect is based on the estimates of realized variances and covariances via the signs of high-frequency returns, i.e., measures known as realized semi-variances and semi-covariances proposed by Bollerslev et al. (2020a) and subsequently used by Bollerslev et al. (2020b) to develop asymmetric MGARCH and realized GARCH models. These authors show that the RCM can be decomposed into the sum of three terms, i.e., the positive semi-covariance term, the negative one, and relevant only for covariances, the mixed one. Instead of

[2]Univariate realized variance (RV) models that include an asymmetric effect have been developed by Corsi and Renò (2012), McAleer and Medeiros (2008), Patton and Sheppard (2015).

assuming that the conditional variance (i.e., the conditional mean of the RCM) of an asset at date t depends on the realized variance of the same asset at date $t-1$, we assume that it depends additively on the realized positive semi-variance (at $t-1$), with a specific coefficient, and on the realized negative semi-variance, with another coefficient. If the latter coefficient is larger than the former, an 'asymmetry' effect is present, corresponding to the leverage effect described previously, whereas if the coefficients are equal, there is no asymmetric effect. A similar specification can be used for a realized covariance, by assuming that it depends linearly on the three semi-covariances (positive, negative, mixed) with different coefficients, instead of the realized covariance (when the three coefficients are equal).[3] The detailed definitions and specifications are provided in Section 3.2.2.

To perform empirical evaluations of the models, we have built a time series of daily returns and RC matrices based on a high-frequency data set for five stocks (of the banking sector) and an exchange traded fund (ETF) tracking the S&P500 market index. Statistical evaluation criteria consist of the in-sample fit and out-of-sample forecast loss functions, i.e., mean squared error and quasi-likelihood. In order to formally determine whether the quality of the forecasts differs significantly across the models, we apply the model confidence set (MCS) procedure of Hansen et al. (2011), which allows us to identify the subset of models that contains the best forecasting model given a pre-specified level of confidence. We also compare the model performances from a portfolio allocation perspective, using as loss functions the global minimum variance portfolio (GMVP) and the mean-variance portfolio (MVP).

Both the in-sample and forecasting results essentially underscore that the asymmetric models significantly outperform the symmetric benchmark specification in all cases. Such results underline the importance of accounting for the asymmetries in

[3] In Bollerslev et al. (2020b), the same specification is used for the conditional mean of the outer product of the daily return vector, as in the traditional MGARCH and realized GARCH models, instead of being used for the conditional mean of the RCM, as we do.

modelling, estimating, and forecasting RC matrices. Furthermore, the results strongly indicate that the models that rely on the sign of underlying daily returns to capture asymmetry not only result in significantly better fitting but also tend to generate more accurate predictions than the models built upon the semi-covariance decomposition. This, perhaps surprisingly, suggests that the daily returns provide a more useful information than the intra-daily returns to account for the 'asymmetry' effect in RC matrices.

The rest of the chapter is organized as follows. Section 3.2 introduces the asymmetric extensions of the benchmark 'symmetric' BEKK-CAW model of Golosnoy et al. (2012). Section 3.3 explains the estimation method. Section 3.4 provides information on the data used to obtain the empirical results presented in Section 3.5. Section 3.6 concludes.

3.2 Introducing 'asymmetry' effects in the BEKK model

Let us consider the daily RCM C_t, defined as the sum of m outer-products of the intraday return vectors over the day t (Barndorff-Nielsen and Shephard (2004)), i.e.,

$$C_t = \sum_{j=1}^{m} r_{j,t} r'_{j,t}, \tag{3.1}$$

where $r_{j,t}$ is a $n \times 1$ vector of returns for the j-th time interval of day t.

To capture the temporal and contemporaneous dependences of the elements in C_t, the BEKK parameterization adopted by Golosnoy et al. (2012) and inspired by the BEKK-MGARCH model of Engle and Kroner (1995), is used for the conditional covariance matrix S_t, defined as the conditional expectation $\mathbf{E}(C_t|\mathcal{F}_{t-1})$ based on the filtration $\mathcal{F}_{t-1} = \{C_{t-1}, C_{t-2}, ...\}$.

The BEKK equation, with one lag of S_t and of C_t, is

$$S_t = CC' + AC_{t-1}A' + BS_{t-1}B', \tag{3.2}$$

where C is a $n \times n$ full rank lower-triangular matrix and A and B are $n \times n$ parameter matrices, for a total of $n(n+1)/2+2n^2$ parameters to be estimated. Given a positive-semidefinite (PSD) S_0, the symmetry and PD-ness of S_t are automatically guaranteed. Following Engle and Kroner (1995), sufficient conditions to identify the benchmark BEKK-CAW model are positive diagonal elements of the matrix C, and that the first diagonal elements of A and of B are positive.

The number of parameters is reduced by imposing the matrices A and B to be diagonal, which leads to a total of $n(n+1)/2 + 2n$ parameters. In this case, each conditional covariance only depends on the corresponding lagged realized and conditional covariances. Finally, the scalar version, i.e., $A = aI_n$ and $B = bI_n$, requires the estimation of only $n(n+1)/2 + 2$ parameters.

3.2.1 'Asymmetry' terms based on the signs of daily returns

To introduce 'asymmetry' terms in the BEKK-CAW model (3.2) that we consider a benchmark, we decompose a RCM additively into several parts based on the signs of daily returns.

Decomposition of a RCM based on the signs of daily returns

We denote by $r_{t,i}$ the daily return of asset i on day t, by $I_t^- = [1_{\{r_{t,1}\leq 0\}}, \ldots, 1_{\{r_{t,n}\leq 0\}}]'$ the indicator vector of the negative daily returns, and by $I_t^+ = [1_{\{r_{t,1}>0\}}, \ldots, 1_{\{r_{t,n}>0\}}]'$ the indicator vector of the positive daily returns. The decomposition of C_t as defined in (3.1) into positive ($C_{P,t}$), negative ($C_{N,t}$), and mixed ($C_{M,t}$) parts based on the

signs of daily returns is then

$$\begin{aligned} C_t &= C_{P,t} + C_{N,t} + C_{M,t}, \text{ where} \\ C_{P,t} &= C_t \odot I_t^+ I_t^{+\prime}, \; C_{N,t} = C_t \odot I_t^- I_t^{-\prime}, \\ C_{M,t} &= C_t \odot (I_t^+ I_t^{-\prime} + I_t^- I_t^{+\prime}), \end{aligned} \tag{3.3}$$

with $\odot$ denoting the Hadamard (element-wise) product of matrices. The decomposition holds because the elements of the matrix $I_t^+ I_t^{+\prime} + I_t^- I_t^{-\prime} + I_t^+ I_t^{-\prime} + I_t^- I_t^{+\prime}$ are all equal to 1. The positive and negative parts, i.e., $C_{P,t}$ and $C_{N,t}$, are PSD; the qualifier positive (negative) is a shortcut for 'positively (negatively) signed'. So, it does not indicate that the off-diagonal elements of $C_{P,t}$ ($C_{N,t}$) are positive (negative). The diagonal elements of $I_t^+ I_t^{-\prime}$ and $I_t^- I_t^{+\prime}$, and therefore of $C_{M,t}$, are always equal to zero, i.e., the mixed part $C_{M,t}$ is necessarily indefinite. We denote by $c_{\bullet,ij,t}$ the (i,j)-th entry of $C_{\bullet,t}$ (where $\bullet$ stands for P, N, or M).

Models using the decomposition based on the signs of daily returns

To include an 'asymmetry' effect in RC matrices based on the signs of daily returns, we add to (3.2) a term that uses the negative component $C_{N,t}$. The conditional covariance matrix dynamic equation is then

$$S_t = CC' + AC_{t-1}A' + \tilde{A}_N C_{N,t-1} \tilde{A}_N' + BS_{t-1}B', \tag{3.4}$$

where $\tilde{A}_N$ is a $n \times n$ parameter matrix. This specification is the same as the conditional threshold autoregressive Wishart (CTAW) model of Anatolyev and Kobotaev (2018) and in the context of MGARCH models, it corresponds to the multivariate version of the GJR univariate GARCH model of Glosten et al. (1993). Using the decomposition

of C_t in (3.3), we parameterize the previous equation equivalently as

$$S_t = CC' + A_P(C_{P,t-1} + C_{M,t-1})A'_P + A_N C_{N,t-1} A'_N + BS_{t-1}B', \tag{3.5}$$

where A_P and A_N are equal to A and $A + \tilde{A}_N$ of (3.4), respectively. Note that if $A_P = A_N := A$ in (3.5), this model is equivalent to the benchmark symmetric model (3.2).

To illustrate this model very simply, we consider two assets ($n = 2$), assuming that $A_P = \text{diag}(a_{P11}, a_{P22})$, $A_N = \text{diag}(a_{N11}, a_{N22})$, and $B = C = 0$, because the corresponding parameters are irrelevant for the asymmetry terms. Then, the dynamic equations for the conditional variances $s_{11,t}$, $s_{22,t}$, and the conditional covariance $s_{12,t}$ are

$$s_{ii,t} = a^2_{Pii} c_{P,ii,t-1} + a^2_{Nii} c_{N,ii,t-1}, \text{ for } i = 1, 2; \tag{3.6}$$

$$s_{12,t} = a_{P11} a_{P22} (c_{P,12,t-1} + c_{M,12,t-1}) + a_{N11} a_{N22} c_{N,12,t-1}. \tag{3.7}$$

The leverage effect in the variances corresponds to $a^2_{Nii} > a^2_{Pii}$ ($i = 1, 2$), i.e., this implies that the conditional variance ($s_{ii,t}$) increases more if the lagged return ($r_{t-1,i}$) is negative than if it is positive, for a given value of the lagged realized variance ($c_{ii,t-1}$), since $c_{P,ii,t-1} = c_{ii,t-1} 1_{\{r_{t-1,i} > 0\}}$ and $c_{N,ii,t-1} = c_{ii,t-1} 1_{\{r_{t-1,i} \leq 0\}}$. For the covariance, the leverage effect corresponds to $a_{N11} a_{N22} > a_{P11} a_{P22}$, which is surely true if the effect holds for both variances, but may be true also if it holds only for one variance. Indeed, for a given level of the lagged realized covariance ($c_{12,t-1}$), the conditional covariance ($s_{12,t}$) increases more if the lagged returns ($r_{t-1,i}$ and $r_{t-1,j}$) are negative than if they are positive or of opposite signs.

A more flexible model is obtained by removing the constraint of equal coefficients of $c_{P,12,t-1}$ and $c_{M,12,t-1}$ in (3.7), resulting in the covariance equation $s_{12,t} = a_{P11} a_{P22} c_{P,12,t-1} + a_{M11} a_{M22} c_{M,12,t-1} + a_{N11} a_{N22} c_{N,12,t-1}$. The corresponding more

flexible version of (3.5) is

$$S_t = CC' + A_P C_{P,t-1} A_P' + A_N C_{N,t-1} A_N' + A_M C_{M,t-1} A_M' + B S_{t-1} B', \tag{3.8}$$

where $A_M = (a_{Mij})$ is a $n \times n$ parameter matrix. The scalar version of this formulation is proposed by Bollerslev et al. (2020b) in the context of realized GARCH models. If $A_P = A_N = A_M := A$ in (3.8), this model is equivalent to the benchmark symmetric model (3.2).

Since $C_{M,t} = C_t \odot (I_t^+ I_t^{-'} + I_t^- I_t^{+'})$, one can split the term $A_M C_{M,t-1} A_M'$ into two terms with different parameter matrices, A_M^+ and A_M^-, creating a more flexible model. Extending the scalar version of Bollerslev et al. (2020b), this is done as follows:

$$S_t = CC' + A_P C_{P,t-1} A_P' + A_N' C_{N,t-1} A_N' + A_M^+ C_{M,t-1}^+ A_M^{+'} + A_M^- C_{M,t-1}^- A_M^{-'} + B S_{t-1} B', \tag{3.9}$$

where $C_{M,t}^+ = C_t \odot \tau(I_t^+ I_t^{-'})$, $C_{M,t}^- = C_t \odot \tau(I_t^- I_t^{+'})$, and the operator $\tau(\cdot)$ sets the lower triangular part of the matrix argument equal to the upper triangular part.[4] The $\tau(\cdot)$ operator is needed to obtain the symmetry of the $C_{M,t}^+$ and $C_{M,t}^-$ matrices.

We refer to (3.5) as the tr-BEKK-CAW model ('tr' for 'threshold'), to (3.8) as the trPNM-BEKK-CAW model, and to (3.9) as the trPNτM-BEKK-CAW, omitting BEKK-CAW when it is clear that we refer to this class of models. Diagonal and scalar versions are obtained by restricting the parameter matrices in the same way as for the model (3.2).

[4]Following He and Teräsvirta (2002), $\tau(M) = \text{ivech}(\text{vech}(M'))$, where $\text{vech}(\cdot)$ stacks the lower triangle of a $n \times n$ matrix into a $n(n+1)/2 \times 1$ vector and $\text{ivech}(\cdot)$ is its inverse, thus generating a symmetric matrix. E.g., for two assets, if $I_t^+ = (1\,0)'$ and $I_t^- = (0\,1)'$ (the first return is positive, the second is negative), $I_t^+ I_t^{-'} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\tau(I_t{}^+ I_t^{-'}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

3.2.2 'Asymmetry' terms based on the signs of intra-daily returns

To enable refined intraday asymmetric RC dynamics (Bollerslev et al. (2020b)), we exploit the semi-covariance decomposition of the RCM into several components.

Semi-covariance decomposition of a RCM

Bollerslev et al. (2020a) provide a decomposition of C_t, as defined in (3.1), into positive (P_t), negative (N_t), and two mixed (M_t^+, M_t^-) realized semi-covariance matrices defined by using the signs of the underlying intraday returns, extending the idea of Barndorff-Nielsen et al. (2010) to the multivariate setting. The semi-covariance matrices are defined as

$$\begin{aligned} P_t &= \sum_{j=1}^{m} r_{j,t}^{+} r_{j,t}^{+\prime};\ N_t = \sum_{j=1}^{m} r_{j,t}^{-} r_{j,t}^{-\prime}; \\ M_t^{+} &= \sum_{j=1}^{m} r_{j,t}^{+} r_{j,t}^{-\prime};\ M_t^{-} = \sum_{j=1}^{m} r_{j,t}^{-} r_{j,t}^{+\prime}, \end{aligned} \tag{3.10}$$

where $r_{j,t}^{+} = r_{j,t} \odot \mathrm{I}_{j,t}^{+}$ and $r_{j,t}^{-} = r_{j,t} \odot \mathrm{I}_{j,t}^{-}$ denote the vectors of positive and negative intra-daily returns, with $I_{j,t}^{+} = [1_{\{r_{j,t,1}>0\}}, \ldots, 1_{\{r_{j,t,n}>0\}}]'$ and $I_{j,t}^{-} = [1_{\{r_{j,t,1}\leq 0\}}, \ldots, 1_{\{r_{j,t,n}\leq 0\}}]'$ denoting the corresponding indicator vectors of the signs of intraday returns.

The positive and negative semi-covariance matrices, i.e., P_t and N_t, are PSD; the qualifiers positive and negative do not imply that the off-diagonal elements of these matrices are positive or negative. The mixed components M_t^+ and M_t^- have zero diagonal elements and, thus, are indefinite, with off-diagonal elements that are necessarily negative. Obviously,

$$C_t = P_t + N_t + M_t, \text{ where } M_t = M_t^{+} + M_t^{-}, \tag{3.11}$$

given that $M_t^{+} = M_t^{-\prime}$.

The corresponding positive, negative and mixed terms of the decompositions (3.3) and (3.11) generally differ. They may be equal under specific conditions that are unlikely to hold in practice. E.g., if $r_{j,t} > 0 \; \forall j$, i.e., the intraday returns (of all stocks) are positive during day t, and the daily return vector r_t used to define $C_{P,t}$ is chosen to be the return over the trading period, i.e., $\sum_{j=1}^{m} r_{j,t}$, then $P_t = C_{P,t} = C_t$. Conversely, if the daily return r_t used to extract the positive daily return indicator vector I_t^+ is the close-to-close return and thus differs from $\sum_{j=1}^{m} r_{j,t}$ (unless the overnight return is equal to zero), then P_t may differ from $C_{P,t}$. A simple example is developed in Appendix A.1.

Models using the semi-covariance decomposition

We exploit the realized semi-covariance decomposition of a RCM to define a semi-BEKK-CAW model for the conditional covariance matrix S_t corresponding to the RCM C_t:

$$S_t = CC' + A_P P_{t-1} A_P' + A_N N_{t-1} A_N' + A_M M_{t-1} A_M' + B S_{t-1} B', \tag{3.12}$$

where A_P, A_N, and A_M are $n \times n$ parameter matrices, while P_t, N_t, and M_t denote the positive, negative, and mixed semi-covariance matrices, respectively.

By construction, M_t has zero diagonal elements, implying that the term $A_M M_{t-1} A_M'$ is irrelevant for the conditional variances. If $A_P = A_N = A_M := A$, the semi-CAW model is equivalent to the benchmark symmetric model (3.2).

To illustrate the terms of this model, we write its equations of the bivariate version, eliminating the constant and $BS_{t-1}B'$ terms, and assuming that A_P, A_N, and A_M are lower triangular (LT), i.e., $a_{P12} = a_{N12} = a_{M12} = 0$. The triangularity assumption could be relevant if the first asset is a market index and the second one is a particular stock so that the market may have an impact on the stock but no impact of the stock

on the market is allowed:

$$
\begin{aligned}
s_{11,t} &= a_{P11}^2 p_{11,t-1} + a_{N11}^2 n_{11,t-1};\\
s_{22,t} &= a_{P22}^2 p_{22,t-1} + a_{N22}^2 n_{22,t-1}\\
&\quad + a_{P21}^2 p_{11,t-1} + a_{N21}^2 n_{11,t-1}\\
&\quad + 2a_{P21}a_{P22}p_{12,t-1} + 2a_{N21}a_{N22}n_{12,t-1} + 2a_{M21}a_{M22}m_{12,t-1};\\
s_{12,t} &= a_{P11}a_{P22}p_{12,t-1} + a_{M11}a_{M22}m_{12,t-1} + a_{N11}a_{N22}n_{12,t-1}\\
&\quad + a_{P11}a_{P21}p_{11,t-1} + a_{N11}a_{N21}n_{11,t-1}.
\end{aligned}
\tag{3.13}
$$

In each equation, the first line corresponds to the diagonal model (i.e., when both off-diagonal elements of A_P, A_N, and A_M are set to zero). It is clear that in the diagonal model, each conditional variance only depends on the corresponding lagged positive and negative semi-variances. A larger coefficient a_{N11}^2 (a_{N22}^2) of the negative semi-variance than of the positive one a_{P11}^2 (a_{P22}^2) for the first (second) asset corresponds to the 'leverage' effect. In the LT model, several other terms appear in the particular asset conditional variance, i.e., two terms that correspond to the effect of the market positive and negative semi-variances, with the possibility of a cross-leverage effect (if $a_{N21}^2 > a_{P21}^2$), and three terms that correspond to the impacts of the three semi-covariances, with coefficients that can be of any sign.

The conditional covariance equation has three terms that capture the effect of the lagged positive, negative, and mixed semi-covariances ($p_{12,t-1}$, $n_{12,t-1}$, $m_{12,t-1}$). It is possible that the coefficient ($a_{N11}a_{N22} > 0$) of the negative semi-covariance ($n_{12,t-1}$) is larger than that ($a_{P11}a_{P22} > 0$) of the positive one ($p_{12,t-1}$), which can be interpreted as a 'leverage' effect on the conditional covariance. The coefficient $a_{M11}a_{M22}$ of the mixed semi-covariance can be of any sign, since a_{M11} and a_{M22} do not appear squared in the variance equations. The two additional terms of the LT version (in the third line of the covariance equation) correspond to cross-effects of the market semi-variances,

with a possibility of a 'cross-leverage' effect if $a_{N11}a_{N21} > a_{P11}a_{P21}$; these coefficients can be of any sign.

Like for the trPNM-CAW model, the 'semi' model can be made more flexible by splitting the mixed semi-covariance matrix M_t into its two components M_t^+ and M_t^- (the positive and negative mixed semi-covariance matrices, respectively, such that $M_t = M_t^+ + M_t^-$) and applying the τ transformation to each component (Bollerslev et al. (2020a)), i.e.,

$$\begin{aligned} S_t = CC' + BS_{t-1}B' + A_P P_{t-1}A_P' + A_N N_{t-1}A_N' \\ + A_M^+ \tau(M_{t-1}^+)A_M^{+'} + A_M^- \tau(M_{t-1}^-)A_M^{-'}, \end{aligned} \tag{3.14}$$

where A_M^+ and A_M^- are $n \times n$ are parameter matrices. We refer to this specification as the semi-τ-BEKK-CAW model. We further discuss the use of the mixed semi-covariance decomposition in Appendix A.2.

3.3 Estimation

The estimation of the parameters of the models presented in Section 3.2 is carried out by maximizing a log-likelihood function (LLF). As in Golosnoy et al. (2012), the latter is based on the assumption that the probability density function of the RC matrices C_t, conditional on the filtration $\mathcal{F}_{t-1} = \{C_{t-1}, C_{t-2}, ...\}$, is Wishart, i.e.,

$$C_t|\mathcal{F}_{t-1} \sim W_n(v, S_t(\theta)/v), \tag{3.15}$$

where $W_n(v, S_t(\theta)/v)$ denotes the n-dimensional central Wishart distribution with v degrees of freedom, with $v \geq n$, and PD $n \times n$ scale matrix $S_t(\theta)/v$, implying $\mathbf{E}(C_t|\mathcal{F}_{t-1}) = S_t(\theta)$; θ is the vector of parameters appearing in the equation defining S_t. E.g., for equation (3.2), θ consists of the elements of C, A, and B.

The LLF for T observations is

$$LLF(\theta|C_1,\ldots,C_T) = -\frac{\nu}{2}\sum_{t=1}^{T}\left\{\log|S_t(\theta)| + \text{trace}\left[S_t(\theta)^{-1}C_t\right]\right\}. \tag{3.16}$$

Bauwens et al. (2012) show that the parameter v can be treated as nuisance parameter, meaning that it can be fixed to an arbitrary value (in practice, 1) to estimate θ. They also show that the Wishart-based LLF provides a quasi-maximum likelihood (QML) estimator for the parameters θ, under suitable conditions, so that the QML estimator is consistent.

The maximization of the LLF is typically difficult due to the dimension of θ, denoted by d_θ, which is of order n^2. E.g., in the case of (3.14), $d_\theta = n(n+1)/2+5n^2(=$ 201 if $n = 6$); in the scalar version of the same model $d_\theta = n(n+1)/2+5(= 26)$ and in the diagonal version, $d_\theta = n(n+1)/2+5n(= 51)$. To get rid of the $n(n+1)/2$ parameters of the C matrix in the maximization of the LLF, it is possible to estimate CC' consistently in a first step. In a second step, the remaining parameters are estimated by QML, conditional on the first step estimates. This procedure reduces the number of parameters by $n(n+1)/2$ in the second step. The estimation of C in the first step is called 'covariance targeting'. It is based on writing the constant term (CC') of each dynamic equation for S_t as a function of $\mathbf{E}(C_t)$, and replacing the latter by the sample mean of the C_t matrices. The covariance targeting parameterizations of the models are defined in Appendix A.3. When targeting is used, it is understood that θ in (3.16) does not include the elements of C.

3.4 Data construction and description

For the subsequent empirical analyses, we have constructed the time series of daily RC matrices with the corresponding decompositions (3.3) and (3.11) into positive, negative, and mixed matrices, based on a high-frequency data set for the SPDR S&P500 (SPY or Spyder), an exchange traded fund that tracks the S&P500 index, and five stocks of the banking sector, i.e., Bank of America Corp. (BAC), Citigroup Inc. (C), Goldman Sachs Group Inc. (GS), JPMorgan Chase & Co. (JPM), and Wells Fargo & Co. (WFC).[5]

To avoid the measurement drawbacks due to microstructure effects when sampling returns at very high frequencies, we compute each daily realized (semi-)covariance matrix as the sum of the outer products of the five-minute log-return vectors of the trading period of the day. Given the high liquidity of all the stocks, the effect of non-synchronicity is rather negligible at the chosen frequency; the synchronization was done globally for all the stocks, using the closest prior price. The sample period is January 3, 2012 - December 31, 2021, resulting in 2517 observations.

Table 3.1 reports, for each asset, the time series means and standard deviations of the realized variances (annualized in percentage, i.e., multiplied by 252 and by 100), and of their 'positive' and 'negative' components used in the two broad classes of asymmetric models. The same statistics for the squared close-to-close and open-to-close log-returns of each asset are shown in the first two rows.

[5]The data provider is the AlgoSeek company (30 Wall Street, New York, NY, 10005, USA). The data provided to us by Algoseek are the prices of the assets, observed every minute during the trading period (9:30-16:00), compiled from the trades that occurred in sixteen US exchanges and marketplaces.

Table 3.1: Time series means and standard deviations (between parentheses) of realized variances, their positive and negative decompositions, and squared daily returns

Asset	SPY	BAC	C	GS	JPM	WFC
r^2_{cc}	2.63 (11.98)	9.87 (32.75)	10.44 (40.77)	7.71 (27.07)	7.02 (28.64)	7.92 (30.00)
r^2_{oc}	1.31 (3.75)	5.51 (13.07)	5.47 (13.66)	4.45 (10.60)	3.54 (7.91)	4.23 (11.74)
RV	5.16 (26.73)	5.45 (9.70)	5.78 (14.12)	4.62 (8.35)	4.06 (9.76)	4.63 (11.08)
P	2.55 (13.16)	2.77 (5.36)	2.89 (7.30)	2.36 (4.55)	2.08 (5.25)	2.34 (6.14)
N	2.61 (13.61)	2.68 (4.74)	2.89 (7.30)	2.26 (4.16)	1.99 (4.73)	2.29 (5.41)
C_P-cc	2.43 (17.36)	2.70 (7.18)	2.68 (7.89)	2.29 (6.31)	1.93 (6.32)	2.28 (8.66)
C_N-cc	2.73 (20.65)	2.75 (7.57)	3.10 (12.39)	2.33 (6.37)	2.13 (7.97)	2.35 (7.66)
C_P-oc	2.71 (18.99)	2.74 (7.86)	2.73 (8.74)	2.37 (6.62)	2.00 (6.88)	2.35 (9.02)
C_N-oc	2.45 (19.15)	2.71 (6.86)	3.05 (11.81)	2.25 (6.05)	2.06 (7.49)	2.28 (7.22)

r^2_{cc}: squared close-to-close daily return; r^2_{oc}: squared open-to-close daily return; RV: realized variance; P: positive semi-variance; N: negative semi-variance; C_P: RV if daily return is positive, 0 if negative; C_N: RV if daily return is negative, 0 if positive; the suffixes -cc and -oc indicate that the different terms of the decomposition (3.3) are based on the signed daily *close-to-close* and *open-to-close* returns, respectively.

Regarding the statistics reported in Table 3.1, several comments are worth making:

1. In each row, the time series means are comparable between the six assets, except for SPY squared returns. There is more heterogeneity in the standard deviations. I.e., due to more extreme values, those for SPY are larger than for the banks.

2. The average positive semi-variance (P) of each asset is a bit larger than the average positive component (C_P-cc), and the average N is smaller than the corresponding C_N-cc (since $P+N = C_P+C_N = C$). Correspondingly, the ratio P/C is slightly over 50% (except for SPY), and the ratio C_P-cc/C under 50%.

3. The average standard deviations of C_P-cc exceed those of P (by 51% on average over the six assets), and likewise, but less strongly, for C_N-cc with respect to N (5% on average).

4. C_P-cc is smaller than C_P-oc (by 10% for SPY, and 3% for the other assets), hence C_N-cc is larger than C_N-oc.

5. Except for SPY, each average realized variance (covering the open-to-close trading period), is only a fraction of the corresponding average squared close-to-close returns (between 55 and 60%), but it is much closer to the average squared open-to-close returns. Also visible are the larger standard deviations of the time series of squared close-to-close returns compared to those of the realized variances.

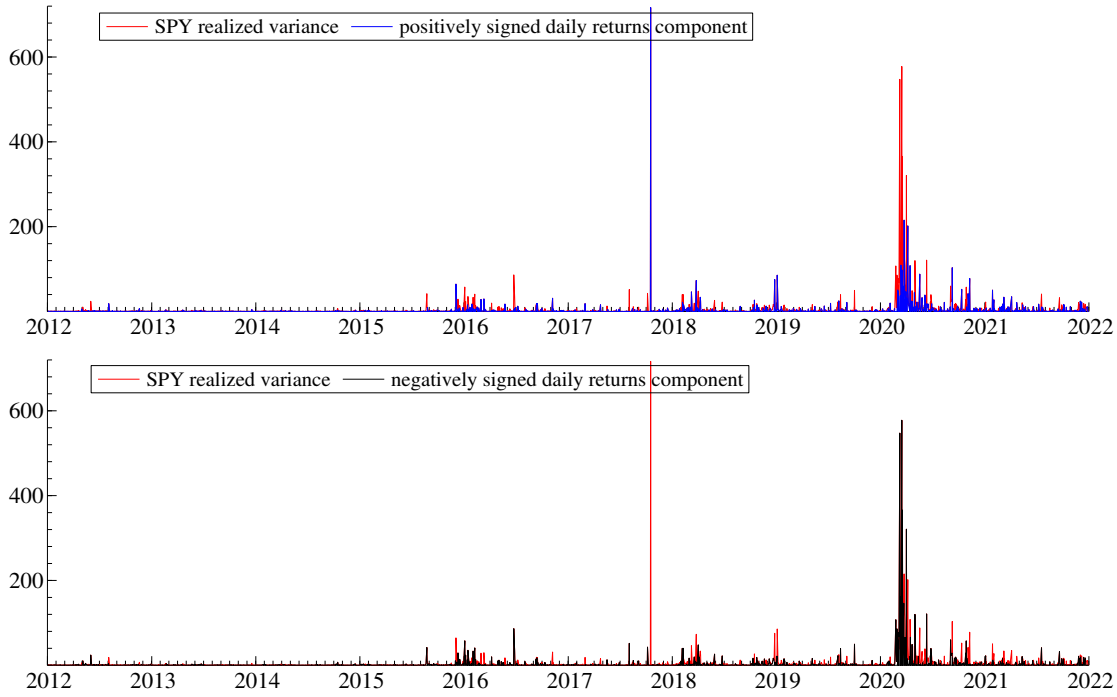


Figure 3.1: Annualized realized variances of SPY and the terms of their decomposition (3.3) using the signed daily close-to-close returns

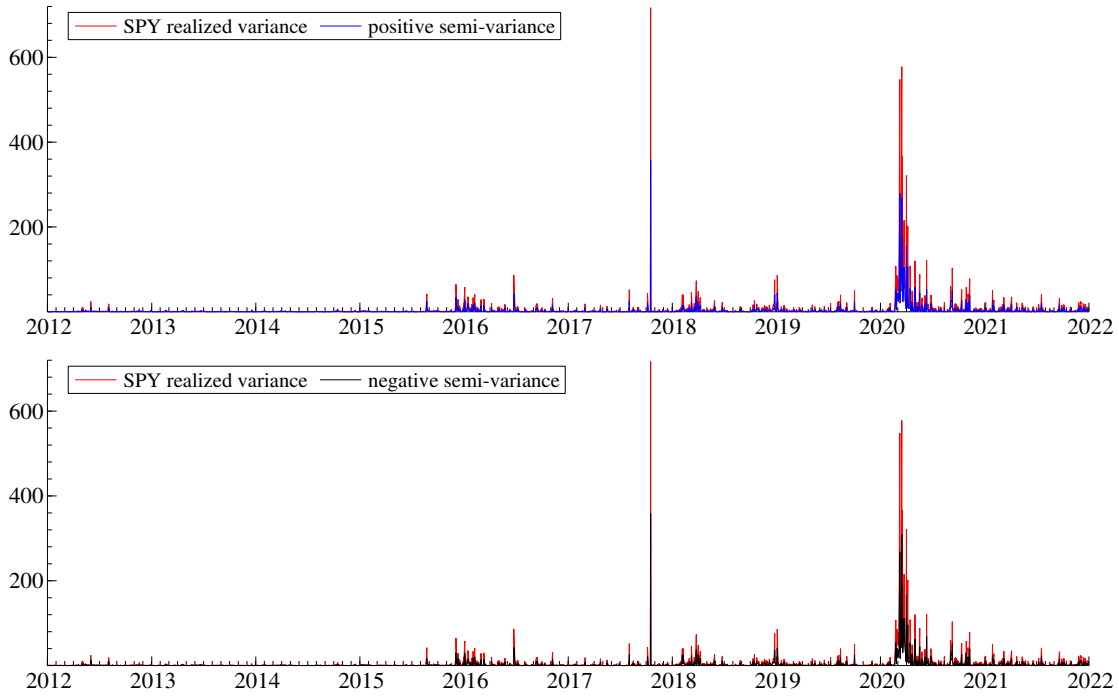


Figure 3.2: Annualized realized variances of SPY and the terms of their decomposition (3.11)

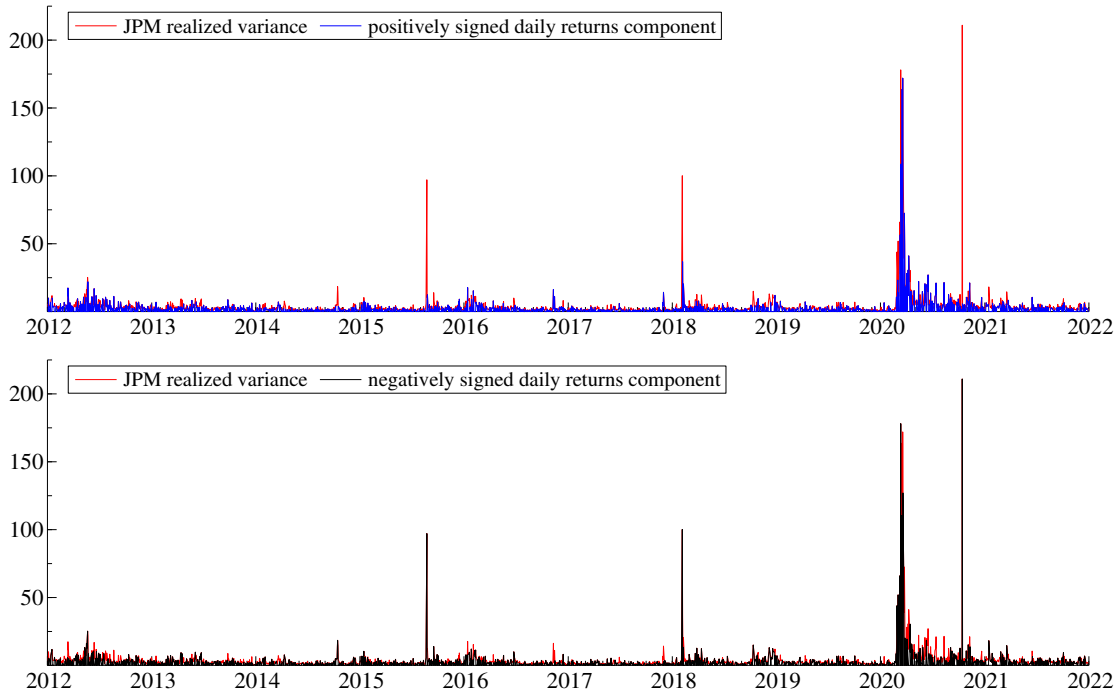


Figure 3.3: Annualized realized variances of JPM and the terms of their decomposition (3.3) using the signed daily close-to-close returns

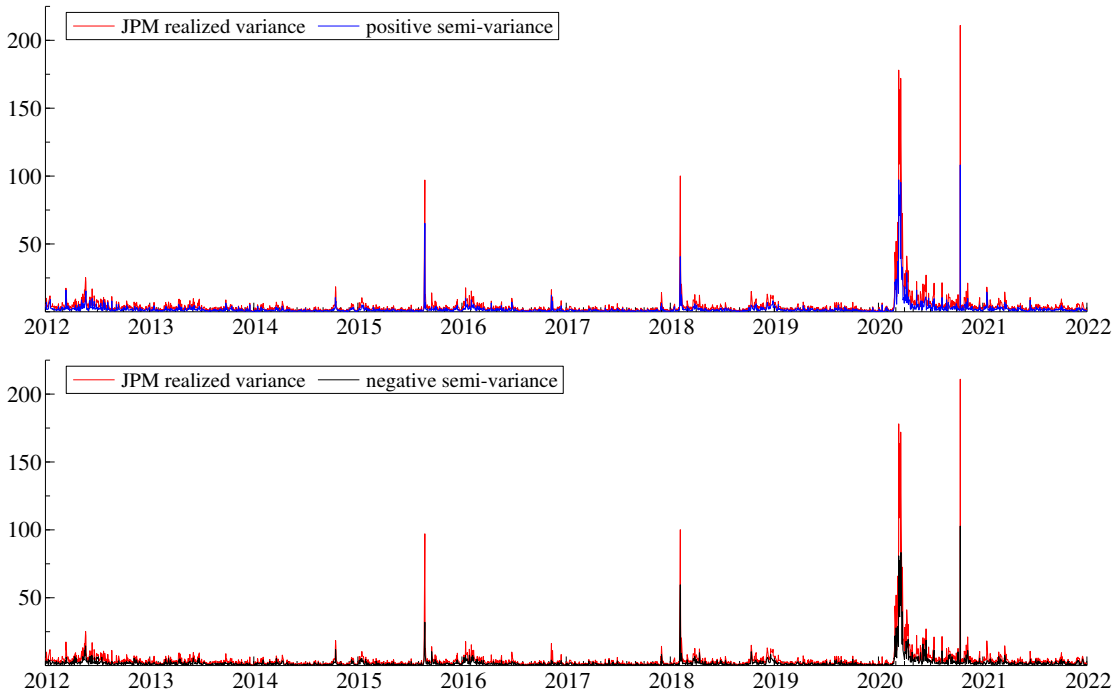


Figure 3.4: Annualized realized variances of JPM and the terms of their decomposition (3.11)

Figures 3.1-3.4 show the time series of the realized variances of SPY and JPM, and the components of their decompositions defined by (3.3), i.e., using close-to-close returns, and (3.11). They illustrate the occurrence of a few extreme values, either isolated (e.g., for JPM, on October 10, 2020) or clustered (mainly in March 2020). The extreme value of the SPY realized variance (on October 17, 2017) is fully attributed to its positive component in the decomposition (3.3) due to a positive return on the day, and in almost equal proportions to the positive and negative semi-variances in the decomposition (3.11). More generally, the profiles of the series on the two figures relative to the same asset illustrate that the two decompositions differ.

This is more visible in Figure 3.5, which shows a zoom of the realized volatility of SPY and its decompositions during the year 2020, with the very high volatility period starting around the middle of February. One can see that in the decomposition using the signed daily returns (right graphs), the realized volatility (in red on each graph) of each day is fully attributed either to the positively signed component (in blue, top

right graph) or to the negatively signed one (in black, bottom right graph). In the decomposition into semi-variances, each realized volatility (in red) is split into a part attributed to the positive semi-variance (in blue, top left graph) and the other to the negative semi-variance (in black, bottom left graph).

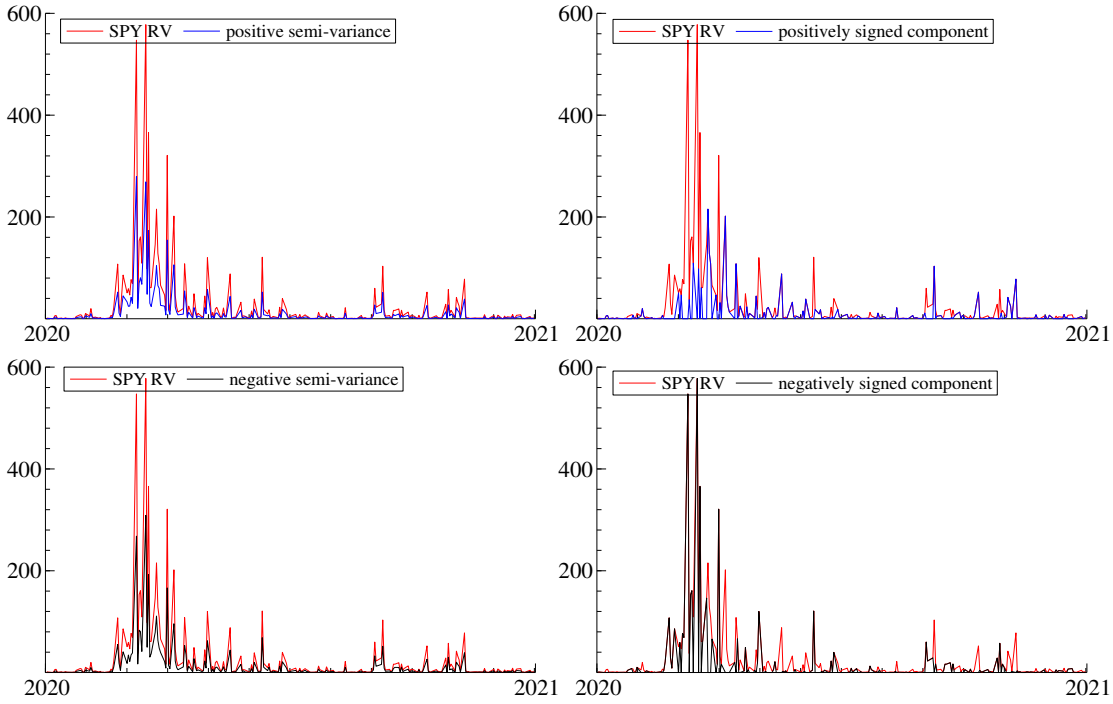


Figure 3.5: Annualized realized variances of SPY and their decompositions: zoom on the year 2020

In Appendix A.5, Table A5.1 shows the time series means and standard deviations of the realized covariances between the six assets and their decomposition into semi-covariances, Table A5.2 shows the analogous statistics with respect to the decomposition using the signed daily close-to-close returns, and Table A5.3 the same information when the signed open-to-close returns are used. Table 3.2 contains representative extracts from the tables in the Appendix, for the asset pairs SPY-BAC and BAC-JPM.

Several comments can be made:

1. The average realized covariances (C) and their average components (in the other columns) are smaller for SPY-BAC than for BAC-JPM. This holds also for the standard deviations. The results in the full tables (in Appendix A.5) show that the statistics for the five pairs with SPY as the first asset and a bank as the second asset show a relative homogeneity, and likewise for the fifteen pairs involving two banks, while the levels of the statistics in the two groups are different.

2. A noticeable difference is the opposite signs of M and C_M. Each average mixed covariance (M) is negative because the mixed covariances of each day are negative by construction. The fact that each average mixed covariance term (C_M) is positive is not a necessity, but a feature of the data, i.e., the sample correlations (and covariances) between the assets are positive, which is not a surprise since five assets are in the same sector and the first one is tracking a market index. Thus, we observe that the inequalities $M < 0 < C_M$ are confirmed for each asset pair. Hence, $P + N$, which is positive, must be larger than $C_P + C_N$, which is also positive. This is even holding term by term, as stated below.

3. Each average positive semi-covariance (P) is larger than the corresponding average positive component (C_P), and each average N is larger than the corresponding C_N.

4. The two parts M^+ and M^- of M are close; likewise for the two parts C_M^+ and C_M^- of C_M.

5. Comparing the averages of the terms of decompositions using the close-to-close and open-to-close returns reveals that C_M is larger for open-to-close than for close-to-close, and correspondingly, $C_P + C_N$, but mainly C_N, is smaller. This holds for all pairs, as can be seen in the full tables of Appendix A.5.

Table 3.2: Time series means and standard deviations (between parentheses) of realized covariances and the terms of three decompositions

Asset Pair	C	P	N	M	M^+	M^-
SPY-BAC	1.58 (4.29)	1.00 (2.91)	0.98 (2.51)	-0.41 (1.77)	-0.20 (0.82)	-0.21 (0.99)
BAC-JPM	3.44 (7.27)	1.84 (4.05)	1.78 (3.73)	-0.18 (0.36)	-0.09 (0.21)	-0.09 (0.18)
Asset Pair	C	C_P-*cc*	C_N-*cc*	C_M-*cc*	C_M^+-*cc*	C_M^--*cc*
SPY-BAC	1.58 (4.29)	0.60 (3.21)	0.69 (3.00)	0.29 (0.83)	0.14 (0.56)	0.15 (0.65)
BAC-JPM	3.44 (7.27)	1.45 (5.07)	1.57 (5.70)	0.42 (1.30)	0.19 (0.91)	0.22 (0.97)
Asset Pair	C	C_P-*oc*	C_N-*oc*	C_M-*oc*	C_M^+-*oc*	C_M^--*oc*
SPY-BAC	1.58 (4.29)	0.61 (3.41)	0.63 (2.49)	0.34 (1.46)	0.16 (0.82)	0.18 (1.24)
BAC-JPM	3.44 (7.27)	1.42 (5.35)	1.45 (5.00)	0.57 (2.59)	0.28 (1.71)	0.29 (1.99)

C: realized covariance. Corresponding to decomposition (3.11): P: positive semi-covariance; N: negative semi-covariance; M: total mixed semi-covariance; M^+ and M^-: positive and negative mixed semi-covariances. Corresponding to decomposition (3.3): C_P: positive part; C_N: negative part; C_M : total mixed part; C_M^+ and C_M^-: positive and negative mixed parts. The suffixes -*cc* and -*oc* in the last two panels indicate that the different terms of the decomposition (3.3) are based on the signed daily *close-to-close* and *open-to-close* returns, respectively.

Figures 3.6 and 3.7 show the time series of the realized covariances between SPY and BAC and the components of their decompositions (3.3), i.e., using the close-to-close daily returns, and (3.11). The peaks of the covariances occur at the same periods as those of the variances. The covariances are almost always positive, but a close look at the (identical) top left graphs reveals a few isolated and slightly negative covariances. In the decomposition based on the signed daily returns (Figure 3.6), these negative values are attributed to one of the three components (as is the case of any positive value). In the decomposition based on the intra-daily returns (Figure 3.7), positive and negative semi-covariances are positive, and the mixed one is negative (by definition). The two figures illustrate the differences between the two decompositions, in particular in their mixed components.

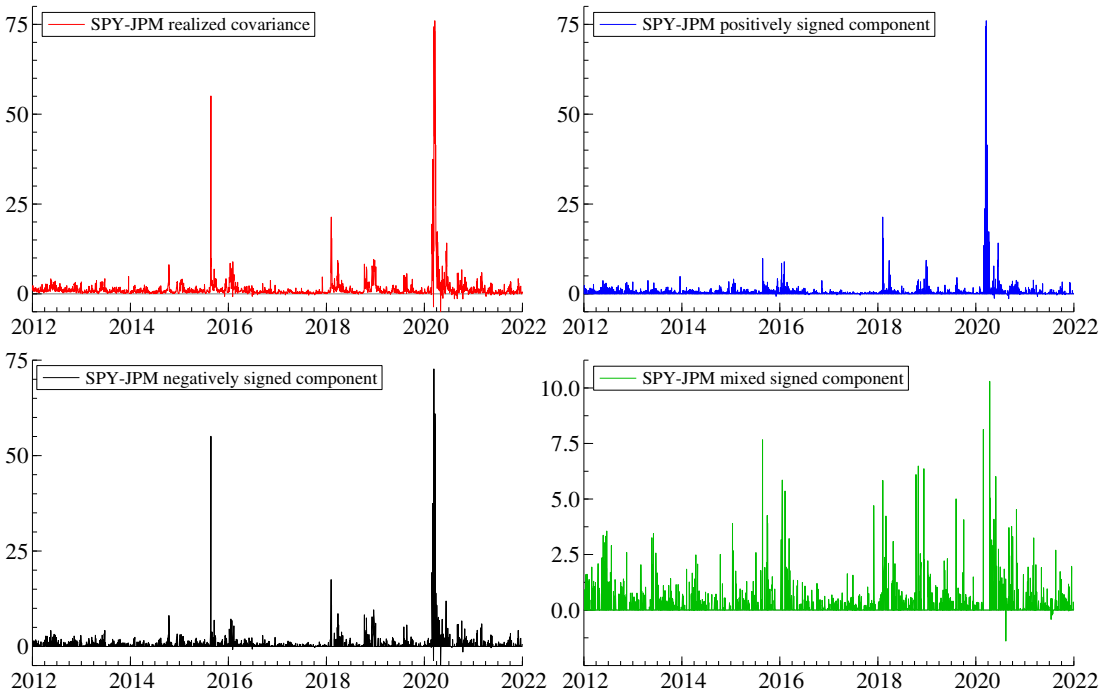


Figure 3.6: Annualized realized covariances of SPY and JPM and the components of their decomposition (3.3) using the signed daily close-to-close returns

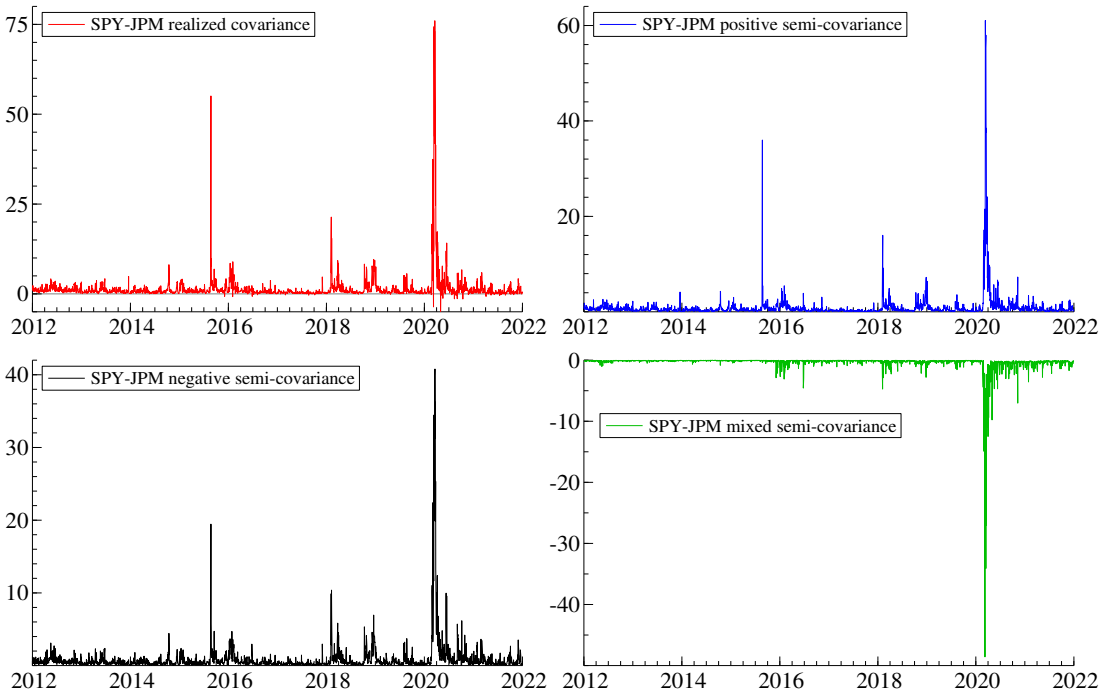


Figure 3.7: Annualized realized covariances of SPY and JPM and the semi-covariance components of their decomposition (3.11)

Table 3.3 shows the means of the indicators of the signed daily close-to-close and open-to-close returns; the corresponding statistics of the two types of returns are close. The daily close-to-close returns of SPY are positive for 56% of the days of the sample period, and 51 or 52% for the other stocks (column 2); negative returns occur in the corresponding complementary percentages (column 3). In column 4, the value 0.406 for SPY is indicating that in 40.6% of the days, both SPY and at least one of the five other stocks have positive returns; the other values in the same columns are close to 41%. In column 5, the values fluctuate between 0.333 and 0.38; the latter (for stock C) means that returns of C and at least one of the other assets were simultaneously negative. The values in the last column are the proportions of days when the return of the stock (in the first column) has a different sign than at least one of the other stocks.

Table 3.3: Time series means of indicators of signed daily returns

Ticker	diag. $I^+I^{+'}$	diag. $I^-I^{-'}$	off-diag. $I^+I^{+'}$	off-diag. $I^-I^{-'}$	$I^+I^{-'}+ I^-I^{+'}$
			Close-to-close returns		
SPY	0.557	0.443	0.406	0.333	0.261
BAC	0.521	0.479	0.418	0.373	0.209
C	0.512	0.488	0.418	0.380	0.202
GS	0.521	0.479	0.412	0.368	0.220
JPM	0.515	0.485	0.416	0.376	0.208
WFC	0.512	0.488	0.402	0.365	0.233
			Open-to-close returns		
SPY	0.548	0.452	0.384	0.319	0.297
BAC	0.510	0.490	0.397	0.364	0.239
C	0.514	0.486	0.399	0.362	0.239
GS	0.519	0.481	0.393	0.352	0.255
JPM	0.522	0.478	0.403	0.359	0.238
WFC	0.514	0.486	0.385	0.348	0.267

diag. $I^+I^{+'}$: indicator of positive returns; diag. $I^-I^{-'}$: indicator of negative returns; off-diag. $I^+I^{+'}$ and off-diag. $I^-I^{-'}$: for the asset indicated in the table row, the average of the five time series means of the off-diagonal elements of the indicated matrix. See comments in the text.

3.5 Empirical results

3.5.1 Estimation results

In this part of the empirical results, we present and interpret the estimates of the symmetric and five asymmetric BEKK-CAW models defined in Section 3.2. Each model is estimated in three versions, i.e., scalar, diagonal, and partly lower triangular, which are explained below. The models are estimated for six assets on the dataset of 2517 observations described in Section 3.4, corresponding to the period from January 3, 2012 to December 31, 2021. The constant terms are estimated by covariance targeting and the remaining parameters by maximizing the Wishart quasi-likelihood function.

The covariance targeting formulas of the models presented in their general version in Section 3.2 are derived in Appendix A.3. In the scalar versions, the parameter matrices, except the constant terms, are restricted to be scalar (see Appendix A.4 for the covariance targeting formulas and the PD-ness conditions). In the diagonal versions, the parameter matrices, except the constant terms, are restricted to be diagonal, so that, like in the scalar versions, there is no influence of the SPY asset on the five banking stocks. To include such effects, we modify each diagonal parameter matrix by adding five parameters in the first column. E.g., the matrix A_P in (3.12) is parameterized as

$$A_P = \begin{pmatrix} a_{P11} & 0 & 0 & 0 & 0 & 0 \\ a_{P21} & a_{P22} & 0 & 0 & 0 & 0 \\ a_{P31} & 0 & a_{P33} & 0 & 0 & 0 \\ a_{P41} & 0 & 0 & a_{P44} & 0 & 0 \\ a_{P51} & 0 & 0 & 0 & a_{P55} & 0 \\ a_{P61} & 0 & 0 & 0 & 0 & a_{P66} \end{pmatrix}, \tag{3.17}$$

and thus appears as partly lower triangular (PLT). The same extensions are introduced in the matrices A, A_N, A_M, A_M^+, A_M^-, but not B that remains diagonal. The variance equation of asset 1 (SPY) is then like written in the first equation of (3.13), and the variance equation of each other asset is as written in the second equation of (3.13), replacing every index 2 by another value (3, 4, 5, 6) for another asset. The covariance equations between asset 1 and the other assets are as written in the third equation of (3.13) for the pair 1-2 (replacing 2 by 3, 4, 5, 6, for assets 3 to 6). Each of these five covariance equations includes two terms that correspond to the impact of the semi-variances of asset 1, due to the introduction of the parameters $a_{\bullet j1}$. The ten covariance equations between the assets i and j (from 2 to 6) do not include these two terms and are like

$$s_{ij,t} = a_{Pii}a_{Pjj}p_{ij,t-1} + a_{Mii}a_{Mjj}m_{ij,t-1} + a_{Nii}a_{Njj}n_{ij,t-1}. \tag{3.18}$$

For the other models, the variables that multiply the coefficients of the variance and covariance equations are changed according to the specification of each model presented in Section 3.2.

The full sets of estimates of each model are reported in Tables A6.1-A6.3 of Appendix A.6. In these tables, the estimates of the models **tr**, **trPNM**, and **trPNτM** are obtained with the data based on the decomposition (3.3) of C_t that uses the *close-to-close* returns. The estimates of the same models with the data based on the decomposition that uses the *open-to-close* returns are reported in Tables A7.1-A7.3 of Appendix A.7. These models are designated by $\mathbf{tr}^{oc}$, $\mathbf{trPNM}^{oc}$, and $\mathbf{trPN\tau M}^{oc}$ in the sequel.

In-sample fit comparisons and some hypothesis tests

Table 3.4 collects three in-sample fit criteria for each model: the maximized log-likelihood function (LLF), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). The LLF values are not always comparable (see further down about this). In view of the information criteria, several conclusions can be drawn:

1. The **tr**, **trPNM**, and **trPNτM** models have smaller AIC and BIC than the corresponding models **tr**^{oc}, **trPNM**^{oc}, and **trPNτM**^{oc}. Using the close-to-close returns, instead of the open-to-close ones, improves the in-sample fit of each model.

2. According to AIC, each asymmetric model has a better fit than the symmetric benchmark model. According to BIC, this is also the case for the comparisons that do not involve the *oc* models, where the exceptions are the following. For the scalar versions, **sym** has a lower BIC than **semi** and **semi-τ**, and for the PLT versions, **sym** has a lower BIC than **trPNτM** and **semi-τ**.

3. Among the asymmetric models, for each version and both information criteria, the best fitting model is among **tr**, **trPNM**, and **trPNτM**.

4. The extensions of **semi** to **semi-τ** do not improve the information criteria, i.e., the likelihood improvements are insufficient to counter the increases in the number of parameters. For the extensions of **trPNM** to **trPNτM**, the AIC values indicate a better fit of the latter in the scalar and PLT versions.

Table 3.4: Maximum log-likelihood function (LLF), AIC, and BIC values of estimated BEKK-CAW models

	sym	**tr**	**trPNM**	**trPNτM**	**semi**	**semi-τ**	**tr^{oc}**	**trPNM^{oc}**	**trPNτM^{oc}**
LLF-scalar	-940.88	-919.30	-901.22	**-897.39**	-934.15	-932.78	-938.76	-934.44	-924.82
LLF-diagonal	-929.52	-887.79	-835.76	**-834.41**	-856.09	-852.68	-922.08	-913.90	-900.21
LLF-PLT	-926.78	-842.03	-822.42	**-810.29**	-837.69	-833.30	-879.97	-864.14	-858.24
AIC-scalar	9.490	9.473	9.460	**9.458**	9.486	9.486	9.489	9.486	9.479
AIC-diagonal	9.489	9.460	**9.424**	9.427	9.440	9.442	9.488	9.486	9.480
AIC-PLT	9.491	9.432	9.425	**9.424**	9.437	9.442	9.462	9.458	9.462
BIC-scalar	9.494	9.480	**9.469**	**9.469**	9.495	9.497	9.496	9.496	9.491
BIC-diagonal	9.517	9.502	**9.479**	9.497	9.496	9.511	9.529	9.541	9.549
BIC-PLT	9.530	**9.497**	9.515	9.540	9.528	9.558	9.527	9.549	9.578

LLF values have been shifted by adding 11,000. The values in bold correspond to the best model of each row.

Table 3.5 reports the likelihood ratio (LR) statistics for testing the null hypothesis of a simpler version against a more complex one that nests it, not considering the *oc* models. Each LR statistics is assumed to be asymptotically chi-squared, the degrees of freedom are indicated below each LR value, and the corresponding p-value in the next row. Based on the p-values, each diagonal version significantly improves the corresponding scalar version, and each PLT version improves the corresponding diagonal, except for **sym**, and scalar versions. Note that for the **sym** model, the p-values are higher than for the asymmetric ones.

Table 3.5: Likelihood ratio tests of scalar versus diagonal, diagonal versus PLT, and scalar versus PLT

	sym	**tr**	**trPNM**	**trPNτM**	**semi**	**semi-τ**
LR s/d	22.72	63.02	130.92	125.96	156.12	160.20
df s/d	10	15	20	25	20	25
p s/d	0.012	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
LR d/plt	5.48	91.52	26.68	48.24	36.80	38.76
df d/plt	5	10	15	20	15	20
p d/plt	0.36	< 0.01	0.032	< 0.01	< 0.01	< 0.01
LR s/plt	28.2	154.54	157.6	174.2	192.92	198.96
df s/plt	15	25	35	45	35	45
p s/plt	0.02	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

LR s/d, df s/d, p s/d: respectively, likelihood ratio statistics, degrees of freedom, p-value for testing the scalar model against the diagonal one; s/plt (d/plt) is for testing the scalar (diagonal) model against the partly lower triangular (plt) one.

Table 3.6 reports the LR statistics and the associated degrees of freedom of the possible nesting tests in each version of the models, not considering the *oc* models. Each null hypothesis cannot be rejected at the 1% level of significance. Note that some pairs of models are not nested, e.g., **trPNM** and **trPNτM**, **semi** and **semi-τ**, any model in the set (**tr**, **trPNM**, **trPNτM**) and any model in the set (**semi**, **semi-τ**) since the two sets use different data. However, even if this is the case, we can compare, equivalently to using an information criterion, the maximized log-likelihood values of models that have the same number of parameters since the penalty term of a given information criterion does not differ. E.g., in the scalar versions, **trPNτM** has 35 likelihood points more than **semi-τ**, and **trPNM** has 33 more than **semi**.

Table 3.6: Likelihood ratio statistics and their degrees of freedom for hypotheses making some models restricted cases of the other

model			version		
nested	nesting	H_0	scalar	diagonal	PLT
sym	**tr**	$A_P = A_N$ in (2.7)	43.2 (1)	83.5 (6)	169.5 (11)
sym	**trPNM**	$A_P = A_N = A_M$ in (3.8)	79.3 (2)	188 (12)	209 (22)
sym	**semi**	$A_P = A_N = A_M$ in (3.12)	13.5 (2)	146 (12)	178 (22)
tr	**trPNM**	$A_P = A_M$ in (3.8)	36.2 (1)	104 (6)	39.2 (11)

The p-value of each test is below 0.01, assuming a chi-squared distribution, with the degrees of freedom reported in parentheses beside the corresponding LR statistics.

Interpretations of the estimates

The estimates reported in the tables of Appendices A.6 and A.7 for the diagonal and PLT versions of different models are difficult to interpret, because the coefficients of the variables that appear in the conditional variance and covariance equations are nonlinear functions of the parameters that are estimated, as illustrated, e.g., by the equations (3.13) and (3.18) for the **semi** model. In Tables 3.7-3.10, we report the estimates in an interpretable way, without considering the *oc* models.

In Table 3.7, we show the estimated coefficients, except the constant terms, that appear in the variance equations of all models. Since the diagonal and PLT versions of the different models have six variance equations with different parameters, we report the estimates of the variance equation for SPY (asset 1), and only the averages of the estimates for the five banking assets (numbered 2 to 6). Indeed, the estimates for the banks are close to each other, but rather different from those of SPY. For the latter, the diagonal element of B is higher than for the banking stocks, and correspondingly the parameters of the A_P and A_N matrices are smaller. These differences reflect the fact that the conditional volatility of SPY is smoother than for the banks, while the bank conditional volatilities are more sensitive to the lagged realized volatility changes.

By definition, the symmetric (**sym**) model has the same parameters for the lagged $c_{P,ii,t-1}$ and $c_{N,ii,t-1}$ variables, i.e., its variance equations can be written as (3.6), where $a^2_{Pii} = a^2_{Nii}$. The estimates reported in Table 3.7 show that for the asymmetric models, these coefficients differ, with a^2_{Nii} in most cases much larger than a^2_{Pii}. There are minor exceptions to this for assets 2, 3, and 4 in the diagonal and PLT versions of both **semi** models, as can be seen in the tables of Appendix A.6. Hence, the leverage effect is essentially confirmed in the variance equations.

Table 3.7: Variance equations of scalar, diagonal, and PLT BEKK-CAW models: average coefficients of variables present in all models

		Coeff:	b^2_{ii}	a^2_{Pii}	a^2_{Nii}
Model	Version	Asset	$s_{ii,t-1}$	$c_{P,ii,t-1}$	$c_{N,ii,t-1}$
	scalar	1-6	0.78	0.20	0.20
	diagonal	1	0.75	0.02	0.02
sym	diagonal	2-6	0.42	0.12	0.12
	PLT	1	0.84	0.16	0.16
	PLT	2-6	0.66	0.32	0.32
	scalar	1-6	0.79	0.17	0.22
	diagonal	1	0.70	0.004	0.07
tr	diagonal	2-6	0.54	0.04	0.10
	PLT	1	0.86	0.08	0.22
	PLT	2-6	0.71	0.26	0.30
	scalar	1-6	0.79	0.14	0.23
	diagonal	1	0.76	0.004	0.05
trPNM	diagonal	2-6	0.51	0.06	0.11
	PLT	1	0.89	0.05	0.21
	PLT	2-6	0.71	0.24	0.32
	scalar	1-6	0.83	0.11	0.20
	diagonal	1	0.76	0.004	0.05
trPNτM	diagonal	2-6	0.50	0.06	0.11
	PLT	1	0.88	0.05	0.25
	PLT	2-6	0.70	0.24	0.34
			$s_{ii,t-1}$	$p_{ii,t-1}$	$n_{ii,t-1}$
	scalar	1-6	0.78	0.14	0.26
	diagonal	1	0.66	0.03	0.10
semi	diagonal	2-6	0.38	0.13	0.24
	PLT	1	0.82	0.13	0.33
	PLT	2-6	0.62	0.33	0.45
	scalar	1-6	0.78	0.14	0.26
	diagonal	1	0.66	0.02	0.10
semi-τ	diagonal	2-6	0.38	0.13	0.24
	PLT	1	0.82	0.13	0.33
	PLT	2-6	0.61	0.35	0.45

Each row corresponds to the conditional variance equation of the model identified in the first two columns. For a scalar model, the coefficients are the same for assets 1 to 6; for a diagonal and PLT model, the coefficients are reported for assets 2 to 6 as the means of the coefficients for these 5 assets. The coefficients, as defined in the first row, are computed using the estimates reported in Table A6.1 for the scalar models, A6.2 for the diagonal models, and A6.3 for the PLT models. Each coefficient multiplies the variable written in row 2 below it for the models of the first four blocks and the variable shown above the 'semi' model for the last two blocks. For PLT models and assets 2-6, the additional terms of the variance equations are reported in Table 3.8.

Table 3.8: Variance equations of PLT BEKK-CAW models: average coefficients of additional terms for assets 2-6

Coefficient	Variable	**sym**	**tr**	**trPNM**	**trPNτM**	Variable	**semi**	**semi-τ**
a^2_{Pi1}	$c_{P,11,t-1}$	0.00084	0.0012	0.0031	0.0035	$p_{11,t-1}$	0.0034	0.0059
a^2_{Ni1}	$c_{N,11,t-1}$	0.00084	0.00034	0.0029	0.00075	$n_{11,t-1}$	0.00089	0.0013
$2a_{Pi1}a_{Pii}$	$c_{P,1i,t-1}$	-0.025	0.022	0.035	0.0035	$p_{1i,t-1}$	-0.0022	-0.059
$2a_{Ni1}a_{Nii}$	$c_{N,1i,t-1,}$	-0.025	0.017	-0.018	0.030	$n_{1i,t-1}$	-0.0025	0.0022
$2a_{Mi1}a_{Mii}$	$c_{M,1I,t-1,}$	-0.025	0.022	0.11	-	$m_{1i,t-1}$	-0.15	-
$2a^+_{Mi1}a^+_{Mii}$	$c^+_{M,1i,t-1}$	-	-	-	0.26	$\tau(m^+_{1i,t-1})$	-	-0.15
$2a^-_{Mi1}a^-_{Mii}$	$c^-_{M,1i,t-1}$	-	-	-	-0.06	$\tau(m^-_{1i,t-1})$	-	-0.21

The coefficient values for each model are computed using the estimates reported in Table A6.3 and the coefficients as defined in the first column, corresponding to the variables in the second column for the models in columns 3 to 6, and the variables in column 7 for the last two models.

In Table 3.8, we report the average estimated coefficients of the additional terms that appear only in the variance equations of the PLT models of the five banking stocks, due to the introduction of the off-diagonal parameters in the first column of the A matrices, as shown in (3.17). The coefficients (a^2_{Pi1} and a^2_{Ni1}), measuring the impact of the positive and negative components of the lagged realized variance of SPY on the other asset variances, are very small and do not reveal a cross-leverage effect. The coefficients in the last five rows of the table correspond to the impacts of components of covariance terms, being negligible only for the **sym** and **tr** models.

In Tables 3.9 and 3.10, we report the average estimated coefficients of the covariance equations. For the diagonal and PLT versions, we report separately the averages for the five covariance equations between asset 1 and the five banking stocks, and the averages for the ten equations between the five banking stocks. We find an asymmetric effect in the covariance equations, i.e., $a_{Nii}a_{Njj}$ is always larger than $a_{Pii}a_{Pjj}$ and $a_{Mii}a_{Mjj}$, with a single and minor exception for the latter. Moreover, the conditional covariances between the five banks are less smooth and more reactive to lagged realized covariance terms than the covariances between the banks and SPY. The additional terms in these five covariances of the PLT models, i.e., last two columns of Table 3.10, have a very small impact.

Table 3.9: Covariance equations of scalar, diagonal, and PLT models, part 1

Model	Version	Coeff: Asset	$a_{Pii}a_{Pjj}$ $c_{P,ij,t-1}$	$a_{Nii}a_{Njj}$ $c_{N,ij,t-1}$	$a_{Mii}a_{Mjj}$ $c_{M,ij,t-1}$	$a^+_{Mii}a^+_{Mjj}$ $c^+_{M,ij,t-1}$	$a^-_{Mii}a^-_{Mjj}$ $c^-_{M,ij,t-1}$
	scalar	1-6	0.20	0.20	0.20	-	
	diagonal	1-other	0.05	0.05	0.05	-	
sym	diagonal	2-6	0.12	0.12	0.12	-	
	PLT	1-other	0.22	0.22	0.22	-	
	PLT	2-6	0.32	0.32	0.32	-	
	scalar	1-6	0.17	0.22	0.17	-	
	diagonal	1-other	0.01	0.08	0.01	-	
tr	diagonal	2-6	0.04	0.09	0.04	-	
	PLT	1-other	0.15	0.26	0.15	-	
	PLT	2-6	0.26	0.30	0.26	-	
	scalar	1-6	0.14	0.23	0.18	-	
	diagonal	1-other	0.02	0.07	0.04	-	
trPNM	diagonal	2-6	0.06	0.11	0.08	-	
	PLT	1-other	0.11	0.26	0.26	-	
	PLT	2-6	0.24	0.32	0.27	-	
	scalar	1-6	0.11	0.20	-	0.15	0.17
	diagonal	1-other	0.02	0.07	-	0.05	0.04
trPNτM	diagonal	2-6	0.06	0.11	-	0.08	0.08
	PLT	1-other	0.11	0.29	-	0.30	0.19
	PLT	2-6	0.24	0.33	-	0.28	0.28
			$p_{ij,t-1}$	$n_{ij,t-1}$	$m_{ij,t-1}$	$\tau(m^+_{ij,t-1})$	$\tau(m^-_{ij,t-1})$
	scalar	1-6	0.14	0.26	0.17	-	
	diagonal	1-other	0.06	0.14	0.02	-	
semi	diagonal	2-6	0.12	0.22	0.002	-	
	PLT	1-other	0.20	0.38	0.05	-	
	PLT	2-6	0.33	0.43	0.12	-	
	scalar	1-other	0.14	0.26	-	0.21	0.15
	diagonal	1-other	0.05	0.15	-	0.02	0.50
semi-τ	diagonal	2-6	0.12	0.22	-	0.00	0.37
	PLT	1-other	0.21	0.38	-	0.03	-0.01
	PLT	2-6	0.34	0.44	-	0.13	0.00

Each row corresponds to the conditional covariance equation of the model identified in the first two columns. For a scalar model, the coefficients are the same for all covariance equations. For a diagonal and PLT '1-other' model, the reported coefficient values are the means of the coefficients of the 5 covariance equations between asset 1 and assets 2 to 6. For a diagonal and PLT '2-6' model, the reported coefficient values are the means of the coefficients of the 10 covariance equations between the assets 2 to 6. The coefficients, as defined in the first row, are computed using the estimates reported in Table A6.1 for the scalar models, A6.2 for the diagonal models, and A6.3 for the PLT models. Each coefficient multiplies the variable written in row 2 below it for the models of the first four blocks and the variable shown above the 'semi' model for the last two blocks.

Table 3.10: Covariance equations of scalar, diagonal, and PLT models, part 2

		Coeff:	$b_{ii}b_{jj}$	$a_{P11}a_{Pi1}$	$a_{N11}a_{Ni1}$
Model	Version	Asset	$s_{ij,t-1}$	$c_{P,11,t-1}$	$c_{N,11,t-1}$
	scalar	1-6	0.78	-	-
	diagonal	1-other	0.55	-	-
sym	diagonal	2-6	0.41	-	-
	PLT	1-other	0.75	-0.0088	-0.0088
	PLT	2-6	0.66	-	-
	scalar	1-6	0.79	-	-
	diagonal	1-other	0.61	-	-
tr	diagonal	2-6	0.55	-	-
	PLT	1-other	0.78	0.0061	0.0076
	PLT	2-6	0.71	-	-
	scalar	1-6	0.79	-	-
	diagonal	1-other	0.62	-	-
trPNM	diagonal	2-6	0.50	-	-
	PLT	1-other	0.79	0.0077	-0.0073
	PLT	2-6	0.71	-	-
	scalar	1-6	0.83	-	-
	diagonal	1-other	0.62	-	-
trPNτM	diagonal	2-6	0.50	-	-
	PLT	1-other	0.78	0.0099	0.0130
	PLT	2-6	0.70	-	-
	Variable		$m_{ij,t-1}$	$p_{11,t-1}$	$n_{11,t-1}$
	scalar	1-6	0.78	-	-
	diagonal	1-other	0.50	-	-
semi	diagonal	2-6	0.38	-	-
	PLT	1-other	0.71	-0.00078	0.0000
	PLT	2-6	0.61	-	-
	scalar	1-6	0.78	-	-
	diagonal	1-other	0.50	-	-
semi-τ	diagonal	2-6	0.38	-	-
	PLT	1-other	0.71	-0.017	0.010
	PLT	2-6	0.61	-	-

See note below Table 3.9.

Appendix A.8 illustrates graphically the estimated variance equations of SPY and JPM and the covariance equations between them for the two models.

3.5.2 Forecast comparisons using statistical loss functions

To compare the out-of-sample forecasting accuracy of the symmetric and asymmetric BEKK-CAW models, we use the so-called James-Stein loss (James and Stein (1976)), which is usually referred to as the multivariate Quasi-Likelihood (QLIK) loss function, i.e.,

$$QLIK_t(S_t^{(i)}, C_t) = \ln |S_t^{(i)}| + \text{trace}([S_t^{(i)}]^{-1} C_t), \tag{3.19}$$

where $S_t^{(i)}$ is a forecast by model i of the observed RC matrix C_t that is used as a proxy for the unobservable covariance matrix Σ_t. This loss function satisfies the conditions for producing a consistent ranking (Hansen and Lunde (2006); Laurent et al. (2013)).[6]

To jointly compare the forecasts of a set of models, we rely on the model confidence set (MCS) of Hansen et al. (2003, 2011). The procedure does not necessarily select a single best model, allowing for the possibility of equal forecasting ability. Hence, a model is removed from the MCS only if it is significantly inferior to the other models.

In particular, starting with a set of candidate models $\mathcal{M}_0$, given a loss function, the loss difference between each pair of models in the set is computed at every time point $t = 1, \ldots, T$, so that for models i and j, we get $d_{t,ij} = L_{t,i} - L_{t,j}$, where $L_{t,(\cdot)} = L_t(S_t^{(\cdot)}, C_t)$ is (3.19). At each step of the procedure, the null hypothesis of equal predictive accuracy, i.e., $H_0 : \mathbf{E}[d_{t,ij}] = 0$, is tested for $\forall i > j \in \mathcal{M}$, a subset of models $\mathcal{M} \subset \mathcal{M}_0$, with $\mathcal{M} = \mathcal{M}_0$ at the initial step. If H_0 is rejected at a chosen significance level α, the worst performing model is removed. This process continues until a set of models remains that includes no model that can be rejected at the level α.

[6] We have also used the squared Frobenius norm loss but do not report the results since this loss function does not help to discriminate clearly between the models.

We adopt the range statistics of Hansen et al. (2011) to test H_0, i.e.,

$$T_{R,\mathcal{M}} = \max_{i,j \in \mathcal{M}} \frac{| \overline{d}_{ij} |}{\sqrt{\widehat{\mathrm{Var}}(\overline{d}_{ij})}}, \tag{3.20}$$

where $\overline{d}_{ij} = \frac{1}{T} \sum_{t=1}^{T} d_{t,ij}$, and $\widehat{\mathrm{Var}}(\overline{d}_{ij})$ is obtained by a circular block bootstrap approach (Hansen et al. (2003)), which we implement with 10,000 replications and the varying block length to verify the robustness of the results.

To compute the forecasts, each model is estimated five times using a rolling window scheme. We start with the fitting period from January 3, 2012 to June 30, 2020 ($T = 2137$) and compute 76 one-step-ahead forecasts after the last in-sample date, assuming that the information related to returns is updated each day. Then, we re-estimate the parameters using the next window of 2137 observations obtained by removing the first 76 observations of the previous window and adding 76 new observations to it, and forecast again the next 76 observations following the end of the estimation sample. This procedure is continued until the end of the sample, resulting in 380 out-of-sample forecasts.

Table 3.11 reports the model confidence sets at the 90% confidence level obtained for several starting model sets, in each case using the QLIK loss function. The values of the loss function are reported for each model. The **sym** model is always excluded from the reported model confidence sets. For the PLT models, all asymmetric models are included in the MCS, except the *oc* models. For the diagonal models, compared to the PLT models, the **tr** model is not in the MCS. For the scalar models, only **trPN**τ**M** and **trPN**τ**M**^{oc} are included, the latter being the single case of inclusion of an *oc* model.

When all twenty-seven models are compared together, the composition of the MCS is the same as when only the PLT models are compared, i.e., all scalar and diagonal models are excluded, as well as the *oc* models and the symmetric model.

Table 3.11: Model confidence sets at 90% level of BEKK-CAW models, with QLIK loss function

Version	Model	QLIK	MCS by version	MCS all
	sym	54.216	0.0000	0.0000
	tr	54.262	0.0000	0.0000
	trPNM	54.299	0.0000	0.0000
	trPNτM	54.087	**0.4297**	0.0000
Scalar	**tr**^{oc}	54.215	0.0001	0.0000
	trPNM^{oc}	54.241	0.0000	0.0000
	trPNτM^{oc}	**54.062**	**1.0000**	0.0000
	semi	54.203	0.0001	0.0000
	semi-τ	54.128	0.0002	0.0000
	sym	54.246	0.0000	0.0000
	tr	53.939	0.0058	0.0000
	trPNM	53.823	**0.5500**	0.0000
	trPNτM	53.825	**0.2660**	0.0000
Diagonal	**tr**^{oc}	54.136	0.0000	0.0000
	trPNM^{oc}	54.144	0.0000	0.0000
	trPNτM^{oc}	54.085	0.0000	0.0000
	semi	**53.796**	**1.0000**	0.0000
	semi-τ	53.865	**0.2660**	0.0000
	sym	54.111	0.0000	0.0000
	tr	53.756	**0.4367**	**0.4262**
	trPNM	53.788	**0.3469**	**0.3521**
	trPNτM	53.769	**0.4367**	**0.4262**
PLT	**tr**^{oc}	53.825	0.0000	0.0000
	trPNM^{oc}	53.936	0.0000	0.0000
	trPNτM^{oc}	53.834	0.0000	0.0000
	semi	53.691	**0.4367**	**0.4262**
	semi-τ	**53.676**	**1.0000**	**1.0000**

'QLIK' column: average value of QLIK losses over the forecast period; bold values identify the minimum loss over the nine models of each version.
'MCS by version' column: p-values of the tests of the MCS procedure when the starting set consists of the nine models of a given version; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).
'MCS all' column: same as previous column when the starting set consists of the twenty-seven models of the table.

3.5.3 Forecast comparisons using economic loss functions

Given that the statistical superiority of one model over another does not automatically translate into a more meaningful impact on investment decisions that depend on the covariance matrix forecasts (Fleming et al. (2003)), we further evaluate the out-of-sample forecasting performance of all the considered models from an economic perspective. We consider two loss functions, i.e., the standard deviation of the out-of-sample global minimum variance portfolio (GMVP) and the standard deviation of the out-of-sample mean-variance portfolio (MVP).

We start by performing a GMVP optimization, where the investor is focused exclusively to reduce the portfolio volatility and ignores the use of the mean of returns. Hence, the optimal portfolio weights are independent of the expected returns, thus allowing to exclusively evaluate the covariance matrix forecasts.

The GMVP optimization problem can be expressed as:

$$\min_{w_t} w_t' \Sigma_t w_t \quad \text{s. t.} \quad w_t' \mathbf{1} = 1, w_t \geq 0, \tag{3.21}$$

where Σ_t is the unobservable covariance matrix of returns for time t, $\mathbf{1}$ is a $n \times 1$ vector of ones, and w_t represents a vector of non-negative portfolio weights, i.e., short-selling is not allowed. In practice, Σ_t is replaced by the forecast $S_t^{(i)}$ of model i. Given that the short-selling restrictions prevent an analytical solution for the optimal weights, numerical optimization is used, for which we rely on the MATLAB Financial Toolbox.

The computed weights (Appendix A.9) are applied to the observed returns of the forecasting period, resulting in 380 optimal GMVP returns. The standard deviation of these returns is computed and serves as a loss function (e.g., Engle and Kelly (2012), Bauwens and Xu (2023)). The best model minimizes the portfolio standard deviation.

The MCS procedure is implemented to evaluate the performance of different starting sets of models, with the results presented in Table 3.12. The MCS of the PLT models includes seven models, only two *oc* models being excluded. The MCS of the diagonal models includes only the **tr** and **trPNτM** models, while the MCS of the scalar models excludes the **semi** models. The MCS of the twenty-seven models coincides with the MCS of the scalar models. Thus, it even includes the symmetric model and the *oc* models.

The second loss function is based on the classical mean-variance portfolio (MVP) optimization, which adds to the GMVP optimization problem a constraint that the targeted portfolio return is larger than a pre-set threshold that we set at 3.5% per year. Once the optimal weights are computed, the procedure is the same as for the GMVP loss function. Table 3.13 provides the results. In the comparisons of the scalar models, and of all models, each MCS consists only of the scalar **trPNτM** and **trPNτM**^{oc} models.

As follows, in contrast to the statistical performance where the PLT asymmetric models based on the signs of close-to-close daily or intra-daily returns are superior, only the scalar models using the decomposition based on daily (*cc* and *oc*) returns appear preferable from an economic perspective.

Table 3.12: Model confidence sets at 90% level of BEKK-CAW models, with GMVP loss function

Version	Model	Loss	MCS by version	MCS all
	sym	1.578	**0.8167**	**0.8153**
	tr	1.577	**0.9200**	**0.9206**
	trPNM	**1.576**	**1.0000**	**1.0000**
	trPNτM	1.577	**0.9200**	**0.9206**
Scalar	**tr**^{oc}	1.578	**0.8167**	**0.8153**
	trPNM^{oc}	1.578	**0.8167**	**0.8153**
	trPNτM^{oc}	1.579	**0.7031**	**0.3555**
	semi	1.581	0.0006	0.0028
	semi-τ	1.580	0.0243	0.0224
	sym	1.634	0.0007	0.0003
	tr	**1.600**	**1.0000**	0.0028
	trPNM	1.603	0.0013	0.0028
	trPNτM	1.602	**0.5094**	0.0028
Diagonal	**tr**^{oc}	1.642	0.0007	0.0001
	trPNM^{oc}	1.609	0.0013	0.0028
	trPNτM^{oc}	1.618	0.0013	0.0028
	semi	1.633	0.0013	0.0003
	semi-τ	1.612	0.0013	0.0028
	sym	1.628	**0.7733**	0.0003
	tr	1.623	**0.8073**	0.0028
	trPNM	1.619	**0.8073**	0.0028
	trPNτM	**1.614**	**1.0000**	0.0028
PLT	**tr**^{oc}	1.624	**0.8073**	0.0028
	trPNM^{oc}	1.651	0.0034	0.0000
	trPNτM^{oc}	1.645	0.0045	0.0000
	semi	1.643	**0.1048**	0.0000
	semi-τ	1.639	**0.2452**	0.0003

'Loss' column: standard deviation of GMVP returns over the forecast period; bold values identify the minimum loss over the nine models of each version.
'MCS by version' column: p-values of the tests of the MCS procedure when the starting set consists of the nine models of a given version; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).
'MCS all' column: same as previous column when the starting set consists of the twenty-seven models of the table.

Table 3.13: Model confidence sets at 90% level of BEKK-CAW models, with MVP loss function

Version	Model	Loss	MCS by version	MCS all
	sym	1.683	0.0001	0.0004
	tr	1.683	0.0001	0.0004
	trPNM	1.681	0.0001	0.0004
	trPNτM	**1.677**	**1.0000**	**1.0000**
Scalar	**tr**^{oc}	1.683	0.0001	0.0004
	trPNM^{oc}	1.685	0.0001	0.0004
	trPNτM^{oc}	1.681	**0.3469**	**0.3457**
	semi	1.684	0.0001	0.0004
	semi-τ	1.683	0.0001	0.0004
	sym	1.738	0.0001	0.0000
	tr	1.709	0.0103	0.0004
	trPNM	1.695	**0.1068**	0.0004
	trPNτM	**1.694**	**1.0000**	0.0004
Diagonal	**tr**^{oc}	1.734	0.0001	0.0000
	trPNM^{oc}	1.696	**0.1068**	0.0004
	trPNτM^{oc}	1.712	0.0066	0.0004
	semi	1.724	0.0001	0.0001
	semi-τ	1.703	0.0103	0.0004
	sym	1.723	0.0038	0.0001
	tr	1.708	**0.5565**	0.0004
	trPNM	1.711	**0.4396**	0.0004
	trPNτM	**1.698**	**1.0000**	0.0004
PLT	**tr**^{oc}	1.706	**0.5565**	0.0004
	trPNM^{oc}	1.720	0.0038	0.0004
	trPNτM^{oc}	1.719	0.0038	0.0004
	semi	1.728	0.0001	0.0001
	semi-τ	1.730	0.0001	0.0001

'Loss' column: standard deviation of MVP returns over the forecast period; bold values identify the minimum loss over the nine models of each version.
'MCS by version' column: p-values of the tests of the MCS procedure when the starting set consists of the nine models of a given version; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).
'MCS all' column: same as previous column when the starting set consists of the twenty-seven models of the table.

3.6 Conclusions

This paper introduces and compares empirically BEKK-CAW models that account for asymmetric dynamics in RC matrices, based on a high-frequency dataset for the main S&P500 ETF and five large US banks. While distinct RC models have revealed that high-frequency data provides important additional information for modelling and forecasting covariance matrices, our study documents the importance of capturing distinct responses of the conditional variances and covariances to lagged realized (co)variances decomposed additively into components weighted either by signed daily returns or by signed intra-daily returns.

The proposed asymmetric BEKK-CAW models show a better in-sample fit and out-of-sample forecasting performance than the benchmark symmetric model. We find that the forecasts of the more flexible (PLT) asymmetric models based on either signed daily close-to-close or intra-daily returns statistically dominate the forecasts of their scalar and diagonal versions. Conversely, in terms of portfolio optimization, the scalar models that capture the asymmetry via the signs of the daily close-to-close returns are superior to the (scalar or more general) models that attempt the same via the signs of the intra-daily returns.

The finding that the asymmetric models using the decomposition of the RCM based on the signs of the daily close-to-close returns have better forecasting performances than the models that use the decomposition based on intra-daily returns is not surprising, since the close-to-close return incorporates the information of the overnight period, whereas the intra-daily returns do not.

In general, we also find that the asymmetric models (**tr**, **trPNM**, **trPNτM)** using the decomposition of the RCM based on the signs of the daily *open-to-close* returns fit less well and have worse forecasting performances than those using the decomposition based on the signs of the *close-to-close* returns. Hence, we conclude that it seems preferable to use close-to-close returns in the decomposition based on

daily returns, and to use the latter rather than the decomposition based on intra-daily returns.

This conclusion differs from that of Bollerslev et al. (2020b), who conclude in favour of the models using the decomposition based on intra-daily returns. There are several explanations to this difference: i) the data sets are different in terms of asset number, composition, and sample period; ii) Bollerslev et al. (2020b) model the daily covariance matrix, i.e., the covariance matrix of the daily returns, as a function of the decomposed RCM of the trading period, which does not take into account the volatility during the non-trading (overnight) period, whereas we model the RCM of the trading period.

Several research tracks are open: i) developing asymmetric dynamic conditional correlation (DCC)-type models based on the decompositions of the RCM; ii) adding HAR-type dynamics to asymmetric BEKK-CAW models to explicitly account for the possible long memory feature of volatility and see how this impacts the forecasting performance; iii) using the maximum likelihood estimation of the asymmetric models assuming a Matrix-F conditional distribution instead of a Wishart (e.g., Zhou et al. (2022); Opschoor et al. (2018)).

A Appendix of Chapter 3

A.1 Impact of the definition of daily return on decompositions of C_t

Let $n = 2$, $r_{j,t} = (x, y)'/\sqrt{m}$, $\forall j = 1, 2, \ldots, m$, with $x, y > 0$.

Then, $r^+_{j,t} = (x, y)'/\sqrt{m}$, $r^-_{j,t} = (0, 0)'$, and

$$C_t = \sum_{j=1}^{m} r_{j,t} r'_{j,t} = \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}, \; P_t = C_t, \; N_t = M_t = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

If the daily return r_t is the open-to-close return $\sum_{j=1}^{m} r_{j,t} = \sqrt{m}(x, y)'$, then $I^+_t = (1, 1)'$, $I^-_t = (0, 0)'$, and

$$C_{P,t} = P_t, \; C_{N,t} = N_t, \; C_{M,t} = M_t.$$

If the daily close-to-close return r_t is different from the open-to-close one, e.g., $r_t = (z, 0)'$ with $z > 0$, then $I^+_t = (1, 0)'$, $I^-_t = (0, 1)'$, and

$$C_{P,t} = \begin{pmatrix} x^2 & 0 \\ 0 & 0 \end{pmatrix} \neq P_t, \; C_{N,t} = \begin{pmatrix} 0 & 0 \\ 0 & y^2 \end{pmatrix} \neq N_t, \; C_{M,t} = \begin{pmatrix} 0 & xy \\ xy & 0 \end{pmatrix} \neq M_t.$$

A.2 Use of the decomposition of M_t

The adoption of the semi-covariance decomposition in the modelling framework with or without separating the components of the mixed matrix M_t depends on the application context. In a bivariate model of the volatility of a specific asset and of a market index, separating the effect of the two realized semi-covariance matrices M^+_t and M^-_t might be relevant. The off-diagonal elements of the M^+_t and M^-_t matrices are necessarily negative.

The covariance equation of the model (3.14) for 2 assets (asset 1 is the market portfolio, asset 2 is a stock), assuming that all the A matrices are lower triangular, e.g.,

$$A_M^- = \begin{pmatrix} a_{M11}^- & 0 \\ a_{M21}^- & a_{M22}^- \end{pmatrix}$$

is

$$\begin{aligned} s_{12,t} = {} & a_{P11}a_{P22}p_{12,t-1} + a_{N11}a_{N22}n_{12,t-1} \\ & + a_{M11}^+ a_{M22}^+ [\tau(M_{t-1}^+)]_{12} + a_{M11}^- a_{M22}^- [\tau(M_{t-1}^-)]_{12} \\ & + a_{P11}a_{P21}p_{11,t-1} + a_{N11}a_{N21}n_{11,t-1}, \end{aligned}$$

where $[\tau(M_t^+)]_{12}$ is the (1,2) element of the matrix $\tau(M_t^+)$.

Given the mixed matrix $\tau(M^-)$ implied via a negative market return and a positive stock return, a negative value of the coefficient $a_{M11}^- a_{M22}^-$ implies that the covariance between the asset and the market increases, which is consistent with the stylized fact that in a 'bear' market period, the covariances tend to increase. The opposite holds for a 'bull' period, i.e., given the mixed matrix $\tau(M^+)$ implied via a positive market return and a negative stock return, a positive $a_{M11}^+ a_{M22}^+$ coefficient implies that the covariance declines.

The above arguments readily extend to a partially lower triangular model of higher dimension, such as estimated in the empirical application (see Section 3.5.1).

A.3 Covariance targeting parameterizations of BEKK-CAW models

To define their targeting parameterizations, the models presented in Section 3.2 are transformed in vector form, i.e., for $s_t = \text{vech}(S_t)$, where $\text{vech}(\cdot)$ is the operator that stacks the lower triangular part of a symmetric $n \times n$ matrix argument into a $n(n+1)/2 \times 1$ vector. Matrices appearing in the equations specifying S_t, such as C_t, $C_{i,t}$ for $i = P, N, M, ...$, and P_t, N_t, M_t, ..., are transformed in the same way and the corresponding vectors are denoted by lower-case letters (e.g., c_t, $c_{N,t}$, p_t, ...). Adjustment matrices K_i, of order $n(n+1)/2$, for $i = P, N, M, ...$, as above, are introduced in the targeting terms to account for the difference between the unconditional levels of c_t and the vectors corresponding to the covariances of (signed) daily/intraday returns. In each adjustment matrix, the matrix $\bar{C} = \sum_{t=1}^{T} C_t/T$ appears.

The parameter matrices of the vectorized models are obtained as $\tilde{M} = L_n(M \otimes M)D_n$, where M is a parameter square matrix of order n (e.g., A, A_N, A_P, ...) of the model in matrix format, and L_n and D_n denote the $n(n+1)/2 \times n^2$ elimination and $n^2 \times n(n+1)/2$ duplication matrices, respectively.[7] Hence $\tilde{M}$ is in each case square and of order $n(n+1)/2$. We refer to Noureldin et al. (2012) for details. Each targeting term below is the product of a matrix of order $n(n+1)/2$ and the vector $\text{vech}(\bar{C})$.

Symmetric model (**sym**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A} - \tilde{B})\bar{c} + \tilde{A}c_{t-1} + \tilde{B}s_{t-1}. \tag{A3.1}$$

Noureldin et al. (2012) prove that the unconditional mean of c_t exists, corresponding to the condition derived in Engle and Kroner (1995), if the eigenvalues of the matrix

[7] L_n is defined such that for any $n \times n$ matrix Q, $\text{vech}(Q) = L_n \,\text{vec}(Q)$, and D_n such that for any symmetric matrix R, $\text{vec}(R) = D_n \,\text{vech}(R)$ (see e.g., Lütkepohl (1996)), with $\text{vec}(\cdot)$ denoting the operator that stacks the columns of a $n \times n$ matrix into a $n^2 \times 1$ vector.

$\tilde{A}+\tilde{B}$ are less than one in modulus.

As a result, $E(c_t) = (\mathrm{I}_{n(n+1)/2} - (\tilde{A}+\tilde{B}))^{-1}c$, and c can be estimated by $\bar{c}$.

Threshold model (**tr**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A}^* - \tilde{B})\bar{c} + \tilde{A}_P(c_{P,t-1} + c_{M,t-1}) + \tilde{A}_N c_{N,t-1} + \tilde{B}s_{t-1}, \qquad \text{(A3.2)}$$

where $c_{P,t} = \mathrm{vech}(C_t \odot I_t^+ I_t^{+'})$, $c_{N,t} = \mathrm{vech}(C_t \odot I_t^- I_t^{-'})$, and $c_{M,t} = \mathrm{vech}(C_t \odot (I_t^+ I_t^{-'} + I_t^- I_t^{+'}))$; $\tilde{A}^* = \sum_{i=1}^{2} \tilde{A}_i K_i$, with $\tilde{A}_i = L_n(A_i \otimes A_i)D_n$, for $i = P, N$,
$K_P = L_n[\big(\overline{C}_P\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_P\big)^{1/2}\overline{C}^{-1/2} + \big(\overline{C}_M\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_M\big)^{1/2}\overline{C}^{-1/2}]D_n$,
with $\overline{C}_P = 1/T\sum_{t=1}^{T} C_t \odot I_t^+ I_t^{+'}$, $\overline{C}_M = 1/T\sum_{t=1}^{T} C_t \odot (I_t^+ I_t^{-'} + I_t^- I_t^{+'})$, and
$K_N = L_n[\big(\overline{C}_N\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_N\big)^{1/2}\overline{C}^{-1/2}]D_n$, with $\overline{C}_N = 1/T\sum_{t=1}^{T} C_t \odot I_t^- I_t^{-'}$.

Threshold model with PNM terms (**trPNM**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A}^* - \tilde{B})\bar{c} + \tilde{A}_P c_{P,t-1} + \tilde{A}_N c_{N,t-1} + \tilde{A}_M c_{M,t-1} + \tilde{B}s_{t-1}, \qquad \text{(A3.3)}$$

where $c_{P,t}$, $c_{N,t}$, and $c_{M,t}$ are defined under (A3.2); $\tilde{A}^* = \sum_{i=1}^{3} \tilde{A}_i K_i$, with $\tilde{A}_i = L_n(A_i \otimes A_i)D_n$ and $K_i = L_n[\big(\overline{C}_i\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_i\big)^{1/2}\overline{C}^{-1/2}]D_n$, with $\overline{C}_i$ as under (A3.2), for $i = P, N, M$.

Threshold model with PNτ(M) terms (**trPNτM**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A}^* - \tilde{B})\bar{c} + \tilde{A}_P c_{P,t-1} + \tilde{A}_N c_{N,t-1} + \tilde{A}_M^+ c_{M,t-1}^+ + \tilde{A}_M^- c_{M,t-1}^- + \tilde{B}s_{t-1}, \qquad \text{(A3.4)}$$

where $c_{M,t}^+ = \mathrm{vech}(C_{M,t}^+)$ with $C_{M,t}^+ = C_t \odot \tau(I_t^+ I_t^{-'})$, and $c_{M,t}^- = \mathrm{vech}(C_{M,t}^-)$ with $C_{M,t}^- = C_t \odot \tau(I_t^- I_t^{+'})$; $\tilde{A}^* = \sum_{i=1}^{2} \tilde{A}_i K_i + \tilde{A}_M^+ K_M^+ + \tilde{A}_M^- K_M^-$, with $\tilde{A}_i = L_n(A_i \otimes A_i)D_n$ and $K_i = L_n[\big(\overline{C}_i\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_i\big)^{1/2}\overline{C}^{-1/2}]D_n$, for $i = P, N$,
$\tilde{A}_M^+ = L_n(A_M^+ \otimes A_M^+)D_n$, $K_M^+ = L_n[\big(\overline{C}_M^+\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_M^+\big)^{1/2}\overline{C}^{-1/2}]D_n$,
$\tilde{A}_M^- = L_n(A_M^- \otimes A_M^-)D_n$, $K_M^- = L_n[\big(\overline{C}_M^-\big)^{1/2}\overline{C}^{-1/2} \otimes \big(\overline{C}_M^-\big)^{1/2}\overline{C}^{-1/2}]D_n$, with
$\overline{C}_M^+ = 1/T\sum_{t=1}^{T} C_{M,t}^+$ and $\overline{C}_M^- = 1/T\sum_{t=1}^{T} C_{M,t}^-$.

Semi-covariance model (**semi**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A}^* - \tilde{B})\overline{c} + \tilde{A}_P p_{t-1} + \tilde{A}_N n_{t-1} + \tilde{A}_M m_{t-1} + \tilde{B} s_{t-1}, \qquad \text{(A3.5)}$$

where $p_t = \mathrm{vech}(P_t)$, $n_t = \mathrm{vech}(N_t)$, and $m_t = \mathrm{vech}(M_t)$; $\tilde{A}^* = \sum_{i=1}^{3} \tilde{A}_i K_i$, with $\tilde{A}_i = L_n(A_i \otimes A_i)D_n$ and $K_i = L_n[\left(\bar{i}\right)^{1/2}\overline{C}^{-1/2} \otimes \left(\bar{i}\right)^{1/2}\overline{C}^{-1/2}]D_n$, for $i = P, N, M$, with $\overline{P} = 1/T \sum_{t=1}^{T} P_t$, $\overline{N} = 1/T \sum_{t=1}^{T} N_t$, and $\overline{M} = 1/T \sum_{t=1}^{T} M_t$.

Semi-covariance model (**semi-τ**):

$$s_t = (\mathrm{I}_{n(n+1)/2} - \tilde{A}^* - \tilde{B})\overline{c} + \tilde{A}_P p_{t-1} + \tilde{A}_N n_{t-1} + \tilde{A}_M^+ m_{t-1}^+ + \tilde{A}_M^- m_{t-1}^- + \tilde{B} s_{t-1}, \quad \text{(A3.6)}$$

where $m_t^+ = \mathrm{vech}(\tau(M_t^+))$ and $m_t^- = \mathrm{vech}(\tau(M_t^-))$;

$\tilde{A}^* = \sum_{i=1}^{2} \tilde{A}_i K_i + \tilde{A}_M^+ K_M^+ + \tilde{A}_M^- K_M^-$, K_i is defined as under (A3.5), for $i = P, N$,

$K_M^+ = L_n[\left(\overline{M}^+\right)^{1/2}\overline{C}^{-1/2} \otimes \left(\overline{M}^+\right)^{1/2}\overline{C}^{-1/2}]D_n$, with $\overline{M}^+ = 1/T \sum_{t=1}^{T} \tau(M_t^+)$, and

$K_M^- = L_n[\left(\overline{M}^-\right)^{1/2}\overline{C}^{-1/2} \otimes \left(\overline{M}^-\right)^{1/2}\overline{C}^{-1/2}]D_n$, with $\overline{M}^- = 1/T \sum_{t=1}^{T} \tau(M_t^-)$.

A.4 Scalar BEKK-CAW models with covariance targeting

With scalar parameter matrices, it is convenient to write the equations using the matrix format. The equations below are obtained as particular cases of the corresponding equations of Appendix A.3, when the parameter matrices A, A_N, ..., are scalar, i.e., $A = aI_n$, $A_N = a_N I_n$, ... The largest eigenvalue of a matrix M is denoted by $\rho(M)$; $\rho(M) < 1$ means that the largest eigenvalue is smaller than 1 in modulus.

Symmetric model (**sym**):

$$S_t = (1 - a^2 - b^2)\overline{C} + a^2 C_{t-1} + b^2 S_{t-1}, \qquad \text{(A4.1)}$$

$\overline{C} = (1/T) \sum_{t=1}^{T} C_t$ (PD);

$a^2 + b^2 < 1$ (covariance stationarity of S_t and PD target).

Threshold model (**tr**):

$$S_t = (1-b^2)\overline{C} - A^*\overline{C} + a^2 C_{t-1} + a_N^2 (I_{t-1}^- I_{t-1}^{-'}) \odot C_{t-1} + b^2 S_{t-1}, \tag{A4.2}$$

$A^* = a^2\,\mathrm{I}_n + (a_N^2 \times \overline{C}_N \times \overline{C}^{-1})$; $\overline{C}_N = (1/T)\sum_{t=1}^{T} C_t \odot I_t^- I_t^{-'}$ (PSD);
$\rho(A^* + b^2\,\mathrm{I}_n) < 1$ (covariance stationarity of S_t and PD target).

For the models listed below, it seems impossible to derive sufficient conditions to guarantee the PD-ness of S_t. Consequently, we impose no a priori restrictions on the parameters. However, during the estimation, we require that the coefficients jointly behave in such a way that S_t is PD $\forall t$.

Threshold model with PNM terms (**trPNM**):

$$\begin{aligned} S_t =& (1-b^2)\overline{C} - A^*\overline{C} + (a_P^2 I_{t-1}^+ I_{t-1}^{+'} + a_N^2 I_{t-1}^- I_{t-1}^{-'} \\ &+ a_M^2 (I_{t-1}^+ I_{t-1}^{-'} + I_{t-1}^- I_{t-1}^{+'})) \odot C_{t-1} + b^2 S_{t-1}, \end{aligned} \tag{A4.3}$$

$A^* = (a_P^2 \overline{C}_P + a_N^2 \overline{C}_N + a_M^2 \overline{C}_M) \times \overline{C}^{-1}$; $\overline{C}_P = (1/T)\sum_{t=1}^{T} C_t \odot I_t^+ I_t^{+'}$ (PSD), with $\overline{C}_N$ (PSD) and $\overline{C}_M$ (indefinite) defined analogously;
$\rho(A^* + b^2\,\mathrm{I}_n) < 1$ (covariance stationarity of S_t).

Threshold model with PNτ(M) terms (**trPNτM**):

$$\begin{aligned} S_t =& (1-b^2)\overline{C} - A^*\overline{C} + (a_P^2 I_{t-1}^+ I_{t-1}^{+'} + a_N^2 I_{t-1}^- I_{t-1}^{-'} \\ &+ (a_M^+)^2 \tau(I_{t-1}^+ I_{t-1}^{-'}) + (a_M^-)^2 \tau(I_{t-1}^- I_{t-1}^{+'})) \odot C_{t-1} + b^2 S_{t-1}, \end{aligned} \tag{A4.4}$$

$A^* = (a_P^2 \overline{C}_P + a_N^2 \overline{C}_N + (a_M^+)^2 \overline{C}_M^+ + (a_M^-)^2 \overline{C}_M^-) \times \overline{C}^{-1}$; $\overline{C}_M^+ = (1/T)\sum_{t=1}^{T} C_t \odot \tau(I_t^+ I_t^{-'})$ (indefinite), with $\overline{C}_M^-$ (indefinite) defined analogously;
$\rho(A^* + b^2\,\mathrm{I}_n) < 1$ (covariance stationarity of S_t).

Semi-covariance model (**semi**):

$$S_t = (1-b^2)\overline{C} - A^*\overline{C} + a_P^2 P_{t-1} + a_N^2 N_{t-1} + a_M^2 M_{t-1} + b^2 S_{t-1}, \tag{A4.5}$$

$A^* = (a_P^2\overline{P} + a_N^2\overline{N} + a_M^2\overline{M}) \times \overline{C}^{-1}$; $\overline{P} = (1/T)\sum_{t=1}^{T} P_t$ (PSD), with $\overline{N}$ (PSD) and $\overline{M}$ (indefinite) defined analogously;

$\rho(A^* + b^2\,\mathrm{I}_n) < 1$ (covariance stationarity of S_t).

Semi-covariance model (**semi-τ**):

$$\begin{aligned} S_t =& (1-b^2)\overline{C} - A^*\overline{C} + a_P^2 P_{t-1} + a_N^2 N_{t-1} \\ &+ (a_M^+)^2 \tau(M_{t-1}^+) + (a_M^-)^2 \tau(M_{t-1}^-) + b^2 S_{t-1}, \end{aligned} \tag{A4.6}$$

$A^* = (a_P^2\overline{P} + a_N^2\overline{N} + (a_M^+)^2\overline{M}^+ + (a_M^-)^2\overline{M}^-) \times \overline{C}^{-1}$; $\overline{M}^+ = (1/T)\sum_{t=1}^{T} \tau(M_t^+)$ and $\overline{M}^- = (1/T)\sum_{t=1}^{T} \tau(M_t^-)$;

$\rho(A^* + b^2\,\mathrm{I}_n) < 1$ (covariance stationarity of S_t).

A.5 Complete tables of the statistics of covariance decompositions

Table A5.1: Time series means and standard deviations (between parentheses) of realized covariances and their decomposition into semi-covariances

Asset Pair	C	P	N	M	M^+	M^-
SPY-BAC	1.58 (4.29)	1.00 (2.91)	0.98 (2.51)	-0.41 (1.77)	-0.20 (0.82)	-0.21 (0.99)
SPY-C	1.64 (4.85)	1.03 (3.21)	1.03 (2.92)	-0.42 (1.99)	-0.21 (1.06)	-0.22 (1.00)
SPY-GS	1.48 (4.08)	0.94 (2.75)	0.91 (2.35)	-0.37 (1.65)	-0.18 (0.70)	-0.20 (0.99)
SPY-JPM	1.42 (4.22)	0.90 (2.94)	0.87 (2.32)	-0.35 (1.59)	-0.17 (0.79)	-0.18 (0.86)
SPY-WFC	1.35 (4.16)	0.89 (2.93)	0.88 (2.59)	-0.42 (1.85)	-0.20 (0.81)	-0.22 (1.08)
BAC-C	4.10 (8.38)	2.18 (4.96)	2.14 (4.59)	-0.22 (0.61)	-0.10 (0.25)	-0.11 (0.38)
BAC-GS	3.20 (6.35)	1.78 (3.64)	1.72 (3.40)	-0.31 (0.53)	-0.15 (0.27)	-0.15 (0.37)
BAC-JPM	3.44 (7.27)	1.84 (4.05)	1.78 (3.73)	-0.18 (0.36)	-0.09 (0.21)	-0.09 (0.18)
BAC-WFC	3.19 (7.05)	1.76 (3.97)	1.71 (3.76)	-0.27 (0.51)	-0.13 (0.25)	-0.14 (0.33)
C-BS	3.37 (7.74)	1.84 (4.27)	1.82 (4.14)	-0.29 (0.51)	-0.15 (0.27)	-0.14 (0.29)
C-JPM	3.54 (8.28)	1.88 (4.63)	1.84 (4.33)	-0.18 (0.49)	-0.09 (0.33)	-0.09 (0.22)
C-WFC	3.30 (8.07)	1.81 (4.51)	1.80 (4.53)	-0.30 (0.91)	-0.14 (0.34)	-0.16 (0.63)
GS-JPM	2.88 (6.47)	1.58 (3.65)	1.53 (3.37)	-0.23 (0.42)	-0.12 (0.26)	-0.12 (0.24)
GS-WFC	2.62 (6.35)	1.51 (3.80)	1.47 (3.36)	-0.35 (0.65)	-0.17 (0.34)	-0.18 (0.42)
JPM-WFC	2.83 (6.55)	1.56 (3.93)	1.50 (3.38)	-0.23 (0.64)	-0.11 (0.27)	-0.12 (0.44)

C: realized covariance; P: positive semi-covariance; N: negative semi-covariance; M: total mixed semi-covariance; M^+ and M^-: positive and negative mixed semi-covariances. See Section 3.2.2 for definitions.

Table A5.2: Time series means and standard deviations (between parentheses) of realized covariances and their decomposition into parts via daily *close-to-close* returns

Asset Pair	C	C_P	C_N	C_M	C_M^+	C_M^-
SPY-BAC	1.58 (4.29)	0.60 (3.21)	0.69 (3.00)	0.29 (0.83)	0.14 (0.56)	0.15 (0.65)
SPY-C	1.64 (4.85)	0.59 (2.90)	0.74 (3.63)	0.31 (1.90)	0.13 (0.54)	0.19 (1.84)
SPY-GS	1.48 (4.08)	0.57 (2.86)	0.68 (3.05)	0.23 (0.70)	0.11 (0.46)	0.13 (0.55)
SPY-JPM	1.42 (4.22)	0.55 (3.08)	0.64 (3.01)	0.24 (0.71)	0.11 (0.47)	0.13 (0.55)
SPY-WFC	1.35 (4.16)	0.52 (3.05)	0.60 (2.93)	0.23 (0.77)	0.10 (0.48)	0.13 (0.63)
BAC-C	4.10 (8.38)	1.70 (5.02)	1.91 (6.82)	0.49 (2.94)	0.20 (0.97)	0.29 (2.79)
BAC-GS	3.20 (6.35)	1.30 (4.52)	1.41 (4.93)	0.49 (1.36)	0.23 (0.99)	0.26 (1.00)
BAC-JPM	3.44 (7.27)	1.45 (5.07)	1.57 (5.70)	0.42 (1.30)	0.19 (0.91)	0.22 (0.97)
BAC-WFC	3.19 (7.05)	1.34 (5.35)	1.39 (5.03)	0.47 (1.40)	0.22 (0.92)	0.25 (1.11)
C-BS	3.37 (7.74)	1.34 (4.57)	1.54 (6.24)	0.49 (2.64)	0.27 (2.50)	0.22 (0.93)
C-JPM	3.54 (8.28)	1.40 (4.67)	1.63 (6.35)	0.50 (2.96)	0.28 (2.81)	0.23 (1.01)
C-WFC	3.30 (8.07)	1.30 (4.70)	1.47 (6.24)	0.53 (3.28)	0.29 (3.09)	0.24 (1.17)
GS-JPM	2.88 (6.47)	1.18 (4.42)	1.31 (5.11)	0.39 (1.08)	0.19 (0.79)	0.19 (0.79)
GS-WFC	2.62 (6.35)	1.05 (4.89)	1.11 (4.36)	0.46 (1.30)	0.22 (0.86)	0.24 (1.03)
JPM-WFC	2.83 (6.55)	1.14 (4.68)	1.23 (4.91)	0.45 (1.37)	0.22 (0.90)	0.23 (1.09)

C: realized covariance; C_P: positive part; C_N: negative part; C_M: total mixed part; C_M^+ and C_M^-: positive and negative mixed parts. See Section 3.2.1 for definitions.

Table A5.3: Time series means and standard deviations (between parentheses) of realized covariances and their decomposition into parts via daily *open-to-close* returns

Asset Pair	C	C_P	C_N	C_M	C_M^+	C_M^-
SPY-BAC	1.58 (4.29)	0.61 (3.41)	0.63 (2.49)	0.34 (1.46)	0.16 (0.82)	0.18 (1.24)
SPY-C	1.64 (4.85)	0.58 (2.79)	0.67 (3.21)	0.39 (2.67)	0.15 (0.85)	0.24 (1.24)
SPY-GS	1.48 (4.08)	0.57 (2.84)	0.57 (2.47)	0.34 (1.97)	0.16 (0.86)	0.18 (1.79)
SPY-JPM	1.42 (4.22)	0.56 (3.24)	0.56 (2.55)	0.30 (1.44)	0.14 (1.00)	0.16 (1.06)
SPY-WFC	1.35 (4.16)	0.50 (2.87)	0.52 (2.49)	0.33 (2.01)	0.15 (1.10)	0.18 (1.70)
BAC-C	4.10 (8.38)	1.67 (5.23)	1.81 (6.29)	0.62 (3.69)	0.28 (1.34)	0.34 (3.47)
BAC-GS	3.20 (6.35)	1.29 (4.79)	1.30 (4.39)	0.61 (2.13)	0.31 (1.32)	0.30 (1.72)
BAC-JPM	3.44 (7.27)	1.42 (5.35)	1.45 (5.00)	0.57 (2.59)	0.28 (1.71)	0.29 (1.99)
BAC-WFC	3.19 (7.05)	1.29 (5.44)	1.27 (4.67)	0.63 (2.18)	0.31 (1.27)	0.32 (1.83)
C-GS	3.37 (7.74)	1.33 (5.03)	1.42 (5.84)	0.62 (2.76)	0.35 (2.59)	0.27 (1.05)
C-JPM	3.54 (8.28)	1.39 (5.03)	1.51 (6.22)	0.64 (3.53)	0.35 (3.22)	0.29 (1.51)
C-WFC	3.30 (8.07)	1.24 (4.71)	1.37 (6.14)	0.69 (3.50)	0.37 (3.18)	0.32 (1.54)
GS-JPM	2.88 (6.47)	1.15 (4.70)	1.17 (4.53)	0.56 (2.13)	0.27 (1.68)	0.29 (1.36)
GS-WFC	2.62 (6.35)	1.07 (4.99)	1.04 (4.17)	0.51 (1.53)	0.24 (0.92)	0.27 (1.27)
JPM-WFC	2.83 (6.55)	1.13 (4.79)	1.15 (4.56)	0.55 (2.07)	0.28 (1.24)	0.27 (1.70)

C: realized covariance; C_P: positive part; C_N: negative part; C_M: total mixed part; C_M^+ and C_M^-: positive and negative mixed parts. See Section 3.2.1 for definitions.

A.6 Estimation results of BEKK-CAW models

Table A6.1: Scalar BEKK-CAW model QML estimates with robust standard errors in parentheses

	sym (A4.1)	**tr** (A4.2)	**trPNM** (A4.3)	**trPNτM** (A4.4)	**semi** (A4.5)	**semi-τ** (A4.6)
b^2	0.782	0.787	0.793	0.828	0.784	0.784
	(0.057)	(0.057)	(0.057)	(0.035)	(0.057)	(0.057)
a^2	0.199					
	(0.049)					
a_P^2		0.166	0.141	0.113	0.139	0.138
		(0.052)	(0.052)	(0.034)	(0.046)	(0.045)
a_N^2		0.217	0.232	0.201	0.258	0.259
		(0.048)	(0.050)	(0.031)	(0.059)	(0.059)
a_M^2			0.184		0.173	
			(0.050)		(0.037)	
$(a_M^+)^2$				0.145		0.209
				(0.030)		(0.048)
$(a_M^-)^2$				0.170		0.149
				(0.035)		(0.036)
LLF	-11940.88	-11919.30	-11901.22	**-11897.39**	-11934.15	-11932.78
AIC	9.490	9.473	9.460	**9.458**	9.486	9.486
BIC	9.494	9.480	**9.469**	**9.469**	9.495	9.497

Each column corresponds to a model; the models are defined in Appendix A.4, corresponding to the headers in row 1; row 2 refers to the equation numbers in the appendix. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

Table A6.2: Diagonal BEKK-CAW model QML estimates

	sym (A3.1)	**tr** (A3.2)	**trPNM** (A3.3)	**trPNτM** (A3.4)	**semi** (A3.5)	**semi-τ** (A3.6)
A	0.134*					
	0.294*					
	0.314*					
	0.284*					
	0.387*					
	0.432*					
A_P		0.062	0.065	0.064	0.158*	0.151*
		0.187*	0.261*	0.263*	0.455*	0.434*
		0.189*	0.247*	0.251*	0.420*	0.403*
		0.175*	0.229*	0.229*	0.285*	0.341*
		0.228*	0.281*	0.279*	0.416*	0.424*
		0.259*	0.228*	0.231*	0.177*	0.154*
A_N		0.259*	0.222*	0.224*	0.308*	0.314*
		0.267*	0.335*	0.330*	0.429*	0.424*
		0.273*	0.319*	0.315*	0.404*	0.397*
		0.240*	0.281*	0.281*	0.308*	0.328*
		0.342*	0.368*	0.373*	0.485*	0.515*
		0.413*	0.335*	0.334*	0.723*	0.705*
A_M			0.149*		0.487*	
			0.298*		0.032	
			0.283*		0.045	
			0.249*		0.082	
			0.339*		0.031	
			0.239*		0.041	
A_M^+				0.174*		1.024
				0.297*		0.010
				0.264*		0.021
				0.243*		0.019
				0.334*		0.005
				0.253*		0.038
A_M^-				0.123*		0.015
				0.294*		0.092
				0.305*		0.085
				0.254*		0.865
				0.351*		0.023
				0.229*		0.015
B	0.866*	0.835*	0.873*	0.873*	0.812*	0.814*
	0.706*	0.768*	0.701*	0.703*	0.589*	0.600*
	0.686*	0.756*	0.711*	0.711*	0.613*	0.622*
	0.716*	0.787*	0.742*	0.742*	0.720*	0.678*
	0.613*	0.710*	0.673*	0.671*	0.580*	0.561*
	0.474*	0.662*	0.723*	0.721*	0.583*	0.602*
LLF	-11929.52	-11887.79	-11835.76	**-11834.41**	-11856.09	-11852.68
AIC	9.489	9.460	**9.424**	9.427	9.440	9.442
BIC	9.517	9.502	**9.479**	9.497	9.496	9.511

* denotes statistical significance at the 5% level. In each cell, the first value is the estimate for the market index (SPY), the next ones are for the banking stocks (ordered as BAC, C, GS, JPM, WFC). Each column corresponds to a model; the models are defined in Appendix A.3, corresponding to the headers in row 1; row 2 refers to the equation numbers in the appendix. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

Table A6.3: Partly lower triangular BEKK-CAW model QML estimates

	sym (A3.1)	**tr** (A3.2)	**trPNM** (A3.3)	**trPNτM** (A3.4)	**semi** (A3.5)	**semi-τ** (A3.6)
A	0.394* -0.029 0.520* -0.025 0.535* -0.026 0.514* 0.012 0.617* -0.044* 0.644*					
A_P		0.292* -0.002 0.530* 0.007 0.516* 0.006 0.487* 0.075 0.537* 0.018 0.482*	0.219 0.008 0.517* 0.029 0.494* 0.011 0.473* 0.119 0.508* 0.009 0.456*	0.228 0.016 0.519* 0.041 0.506* 0.018 0.479* 0.121 0.508* 0.021 0.455*	0.357* -0.033 0.658* -0.003 0.635* -0.038 0.548* 0.111 0.556* -0.048 0.472*	0.356* -0.100 0.661* -0.037 0.638* -0.096 0.602* 0.061 0.553* -0.073 0.483*
A_N		0.474* 0.030 0.552* 0.022 0.543* 0.013 0.513* 0.004 0.581* 0.011 0.571*	0.461* -0.007 0.581* -0.020 0.566* -0.012 0.527* -0.024 0.601* -0.016 0.561*	0.495* 0.038* 0.591* 0.029 0.573* 0.024* 0.539* 0.017 0.624* 0.024 0.563*	0.572* 0.035 0.613* -0.006 0.598* 0.025 0.557* -0.051 0.695* -0.003 0.838*	0.578* 0.066 0.615* 0.015 0.596* 0.020 0.584* -0.031 0.714* 0.020 0.821*
A_M			0.495* 0.076* 0.551* 0.088* 0.523* 0.093* 0.486* 0.162* 0.565* 0.099* 0.467*		0.151 -0.215 0.414* -0.125 0.269 -0.263* 0.573* -0.300 0.357* 0.002 0.136	
A_M^+				0.573* 0.227* 0.557* 0.209* 0.522* 0.201* 0.495* 0.322* 0.586* 0.251* 0.476*		0.077 -0.170 0.536 -0.300 0.445 0.398 0.157 -0.299 0.287 -0.300 0.383
A_M^-				0.360* -0.093 0.556* -0.055 0.549* -0.040 0.493* -0.058 0.577* -0.027 0.455*		0.123 1.000 -0.201 0.666 -0.093 0.772 -0.135 0.974 -0.300 0.123 0.044
B	0.919* 0.854* 0.845* 0.858* 0.787* 0.720*	0.930* 0.836* 0.841* 0.862* 0.814* 0.848*	0.943* 0.834* 0.843* 0.863* 0.814* 0.861*	0.937* 0.829* 0.835* 0.857* 0.806* 0.858*	0.905* 0.777* 0.796* 0.829* 0.769* 0.756*	0.905* 0.772* 0.792* 0.820* 0.762* 0.755*
LLF	-11926.78	-11842.03	-11822.42	**-11810.29**	-11837.69	-11833.30
AIC	9.491	9.432	9.425	**9.424**	9.437	9.442
BIC	9.530	**9.497**	9.515	9.540	9.528	9.558

* denotes statistical significance at the 5% level. Each column corresponds to a model; the models are defined in Appendix A.3, corresponding to the headers in row 1; row 2 refers to the equation numbers in the appendix. For each parameter matrix A, the first column of coefficients gives the impacts of SPY on the banking stocks (ordered as BAC, C, GS, JPM, WFC), and the second column reports the diagonal parameters, with SPY first. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

A.7 Estimation results of BEKK-CAW models based on daily open-to-close returns

Table A7.1: Scalar OC BEKK-CAW model QML estimates

	tr (A4.2)	**trPNM** (A4.3)	**trPNτM** (A4.4)
b^2	0.782	0.783	0.784
	(0.055)	(0.052)	(0.052)
a_P^2	0.192	0.180	0.182
	(0.041)	(0.030)	(0.029)
a_N^2	0.205	0.217	0.220
	(0.054)	(0.063)	(0.066)
a_M^2		0.200	
		(0.047)	
$(a_M^+)^2$			0.195
			(0.046)
$(a_M^-)^2$			0.241
			(0.065)
LLF	-11938.76	-11934.44	**-11924.82**
AIC	9.489	9.486	**9.479**
BIC	9.496	9.496	**9.491**

Each column corresponds to a model; the models are defined in Appendix A.4, corresponding to the headers in row 1, with indicator vectors in each specification defined via OC returns; row 2 refers to the equation numbers in the appendix. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

Table A7.2: Diagonal OC BEKK-CAW model QML estimates

	tr (A3.2)	**trPNM** (A3.3)	**trPN$_7$M** (A3.4)
A_P	0.149*	0.185*	0.191
	0.265*	0.305*	0.286*
	0.262*	0.295*	0.281
	0.245*	0.270	0.259*
	0.301*	0.329*	0.308
	0.282*	0.253	0.367
A_N	0.177*	0.174*	0.176*
	0.369*	0.357*	0.349*
	0.357*	0.319*	0.329*
	0.312*	0.301*	0.300*
	0.416*	0.419*	0.408
	0.415*	0.331	0.448
A_M		0.104*	
		0.316*	
		0.315	
		0.266*	
		0.387*	
		0.255	
A_M^+			0.090*
			0.311*
			0.284
			0.261*
			0.376
			0.418
A_M^-			0.127*
			0.303*
			0.310*
			0.274*
			0.383
			0.390
B	0.841*	0.832*	0.828*
	0.664*	0.653*	0.673*
	0.660*	0.647*	0.681*
	0.709*	0.702*	0.712*
	0.626*	0.608*	0.623
	0.536*	0.703*	0.576
LLF	-11922.08	-11913.90	**-11900.21**
AIC	9.488	9.486	**9.480**
BIC	**9.529**	9.541	9.549

* denotes statistical significance at the 5% level. In each cell, the first value is the estimate for the market index (SPY), the next ones are for the banking stocks (ordered as BAC, C, GS, JPM, WFC). Each column corresponds to a model; the models are defined in Appendix A.3, corresponding to the headers in row 1, with indicator vectors in each specification defined via OC returns; row 2 refers to the equation numbers in the appendix. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

Table A7.3: PLT OC BEKK-CAW model QML estimates

	tr (A3.2)		**trPNM** (A3.3)		**trPNτM** (A3.4)	
A_P		0.435*		0.450*		0.461
	0.015	0.561*	0.041	0.549*	0.026	0.539*
	0.029	0.551*	0.056	0.539*	0.041	0.536*
	0.011	0.529*	0.028	0.527*	0.017	0.517*
	0.038	0.572*	0.071	0.569*	0.054	0.555*
	0.013	0.544	0.036	0.530*	0.027	0.523*
A_N		0.429*		0.456*		0.458*
	0.002	0.579*	-0.011	0.604*	0.017	0.611*
	-0.009	0.568*	-0.022	0.596*	0.006	0.598*
	-0.005	0.542*	-0.009	0.565*	0.012	0.570*
	-0.007	0.608*	-0.024	0.667*	0.003	0.674*
	-0.002	0.600	-0.015	0.611*	0.007	0.619*
A_M				0.503*		
			0.078*	0.574*		
			0.094*	0.564*		
			0.078*	0.538*		
			0.088*	0.645*		
			0.041	0.541*		
A_M^+						0.478*
					0.216*	0.577*
					0.214*	0.558*
					0.186*	0.532*
					0.223*	0.638*
					0.114*	0.522*
A_M^-						0.559*
					-0.032	0.569*
					0.007	0.569*
					-0.008	0.537*
					-0.031	0.645*
					-0.004	0.558*
B	0.917*		0.913*		0.911*	
	0.816*		0.813*		0.814*	
	0.821*		0.816*		0.816*	
	0.841*		0.834*		0.836*	
	0.797*		0.777*		0.780*	
	0.820		0.823*		0.822*	
LLF	-11879.97		-11864.14		**-11858.24**	
AIC	9.462		**9.458**		9.462	
BIC	**9.527**		9.549		9.578	

* denotes statistical significance at the 5% level. Each column corresponds to a model; the models are defined in Appendix A.3, corresponding to the headers in row 1, with indicator vectors in each specification defined via OC returns; row 2 refers to the equation numbers in the appendix. For each parameter matrix A, the first column of coefficients gives the impacts of SPY on the banking stocks (ordered as BAC, C, GS, JPM, WFC), and the second column reports the diagonal parameters, with SPY first. The last lines report the obtained maximum value of the log-likelihood function (LLF) and the corresponding Akaike (AIC) and Bayesian information criteria (BIC) values. The models are estimated using the dataset of 2517 observations described in Section 3.4.

A.8 Graphical illustrations of (co)variance equations

Figures A8.1 and A8.2 show the fitted conditional variances of SPY over the three years of different levels of volatility, i.e., 2014 (low level), 2020 (extreme level, in March and April), 2021 (medium level). Each graph also shows the contribution of each term of the right-hand side of the conditional variance equation of the chosen model, which is **semi-τ** for Figure A8.1 and **trPNτM** for Figure A8.2, both in their PLT version. The equations for SPY (asset 1) are

$$\begin{aligned}
\textbf{semi-}\tau\textbf{:} \quad s_{11,t} &= a_{P11}^2 p_{11,t-1} + a_{N11}^2 n_{11,t-1} + b_{11}^2 s_{11,t-1} + \text{constant};\\
\textbf{trPN}\tau\textbf{M:} \quad s_{11,t} &= a_{P11}^2 C_{P11,t-1} + a_{N11}^2 C_{N11,t-1} + b_{11}^2 s_{11,t-1} + \text{constant}.
\end{aligned}$$

On each graph, the red line corresponds to the first term on the right-hand side, the blue line to the sum of the first two terms, and the green line to the fitted conditional variance. Hence, the spread between the green and blue lines is the cumulative contribution of the last two terms term (where the constant is relatively small).

For the same models, Figures A8.3 and A8.4 report the fitted conditional variances of JPM and the terms of the corresponding variance equations, only for the year 2021. There are six more terms in the variance equations of JPM than of SPY, due to the spillover effects introduced by the PLT specification. The variance equations for JPM (asset 5) are

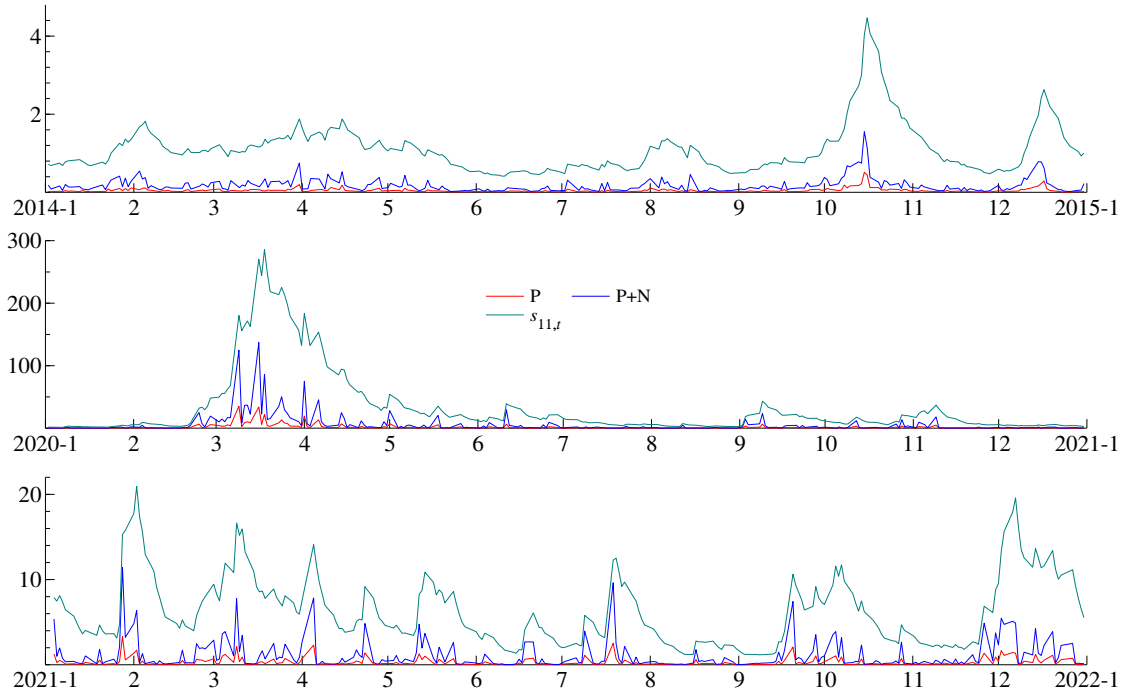
$$\begin{aligned}
\textbf{semi-}\tau\textbf{:} \quad s_{55,t} &= a_{P55}^2 p_{55,t-1} + a_{N55}^2 n_{55,t-1} + a_{P51}^2 p_{11,t-1} + a_{N51}^2 n_{11,t-1}\\
&+ b_{55}^2 s_{55,t-1} + 2a_{P51}a_{P55}p_{51,t-1} + 2a_{N51}a_{N55}n_{51,t-1} + \text{constant}\\
&+ 2a_{M51}^{+}a_{M55}^{+}[\tau(M_{t-1}^{+})]_{51} + 2a_{M51}^{-}a_{M55}^{-}[\tau(M_{t-1}^{-})]_{51};\\
\textbf{trPN}\tau\textbf{M:} \quad s_{55,t} &= a_{P55}^2 C_{P55,t-1} + a_{N55}^2 C_{N55,t-1} + a_{P51}^2 C_{P11,t-1} + a_{N51}^2 C_{N11,t-1}\\
&+ b_{55}^2 s_{55,t-1} + 2a_{P51}a_{P55}C_{P51,t-1} + 2a_{N51}a_{N55}C_{N51,t-1} + \text{constant}\\
&+ 2a_{M51}^{+}a_{M55}^{+}C_{M51,t-1}^{+} + 2a_{M51}^{-}a_{M55}^{-}C_{M51,t-1}^{-}.
\end{aligned}$$

The correspondence between the lines on the different graphs and the terms on the right-hand side of the equations is explained below each figure.

Figures A8.5 and A8.6 report the fitted conditional covariances of the pair SPY-JPM and the terms of the corresponding covariance equations, only for the year 2021. The equations are

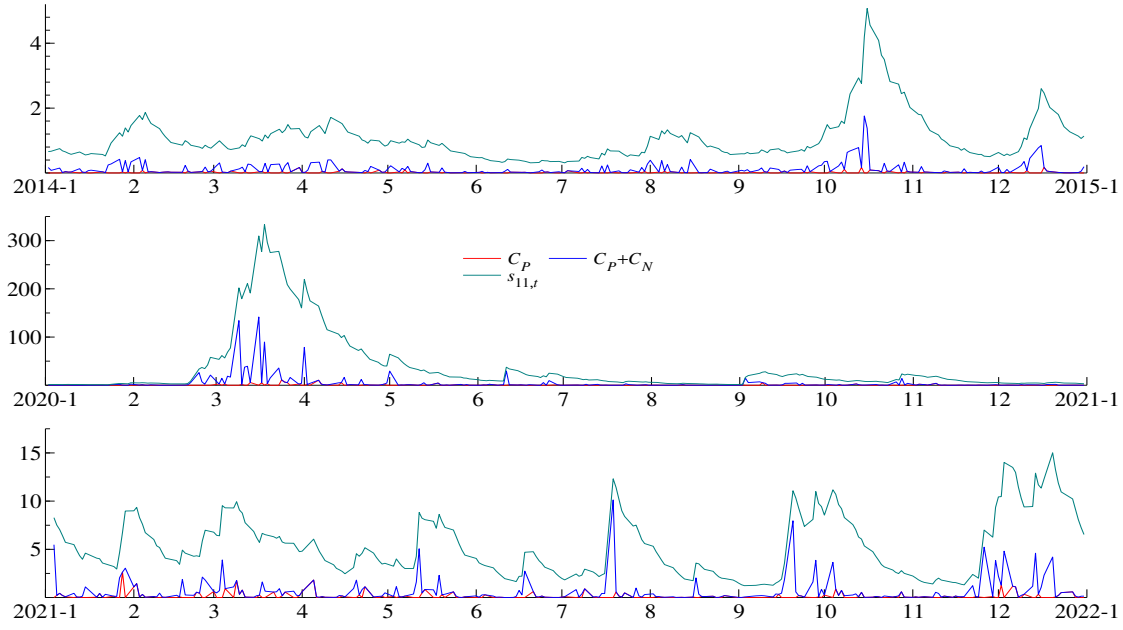
$$\begin{aligned}
\textbf{semi-}\tau\textbf{:} \quad s_{51,t} &= b_{11}b_{55}s_{51,t-1} + a_{P11}a_{P51}p_{11,t-1} + a_{N11}a_{N51}n_{11,t-1} + \text{constant} \\
&+ a_{P11}a_{P55}p_{51,t-1} + a_{N11}a_{N55}n_{51,t-1} \\
&+ a^{+}_{M11}a^{+}_{M55}[\tau(M^{+}_{t-1})]_{51} + a^{-}_{M11}a^{-}_{M55}[\tau(M^{-}_{t-1})]_{51}; \\
\textbf{trPN}\tau\textbf{M:} \quad s_{51,t} &= b_{11}b_{55}s_{51,t-1} + a_{P11}a_{P51}C_{P11,t-1} + a_{N11}a_{N51}C_{N11,t-1} + \text{constant} \\
&+ a_{P11}a_{P55}C_{P51,t-1} + a_{N11}a_{N55}C_{N51,t-1} \\
&+ a^{+}_{M11}a^{+}_{M55}C^{+}_{M51,t-1} + a^{-}_{M11}a^{-}_{M55}C^{-}_{M51,t-1}.
\end{aligned}$$

Figure A8.1: Terms of the SPY variance equation of the **semi-τ** model during the years 2014, 2020, and 2021



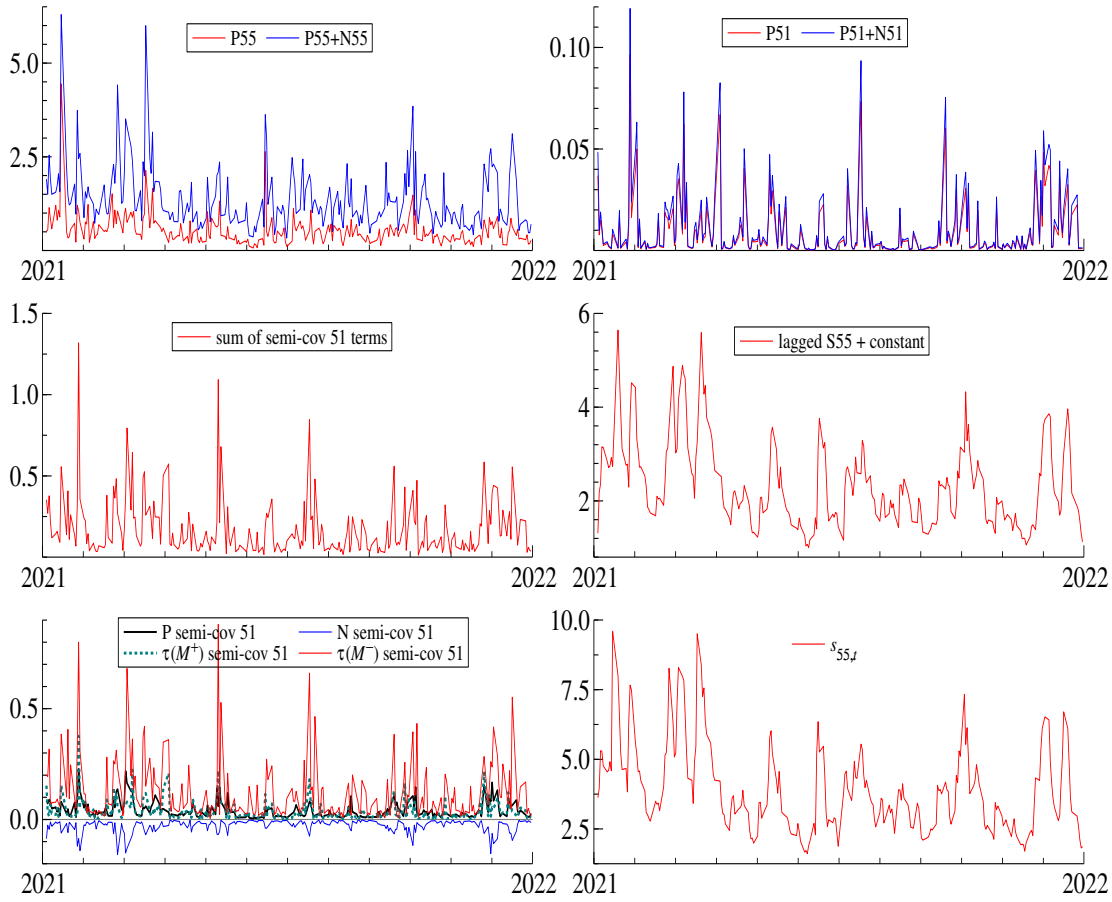
P: $a^2_{P11}p_{11,t-1}$; P+N: P+$a^2_{N11}n_{11,t-1}$; $s_{11,t}$: P+N+$b^2_{11}s_{11,t-1}$+constant, i.e., fitted conditional variance. Values on the vertical axes are annualized.

Figure A8.2: Terms of the SPY variance equation of the **trPNτM** model during the years 2014, 2020, and 2021



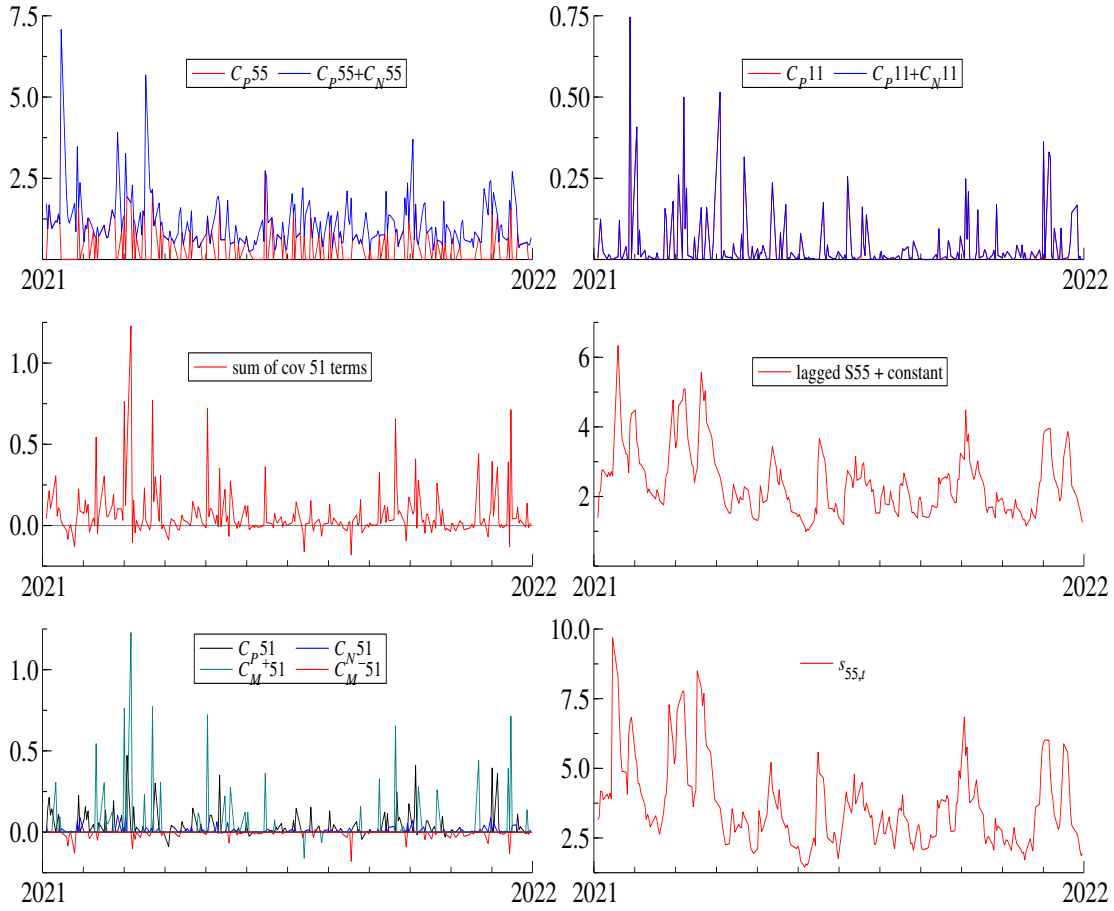
P: $a^2_{P11}C_{P11,t-1}$; P+N: P+$a^2_{N11}C_{N11,t-1}$; $s_{11,t}$: P+N+$b^2_{11}s_{11,t-1}$+constant, i.e., fitted conditional variance. Values on the vertical axes are annualized.

Figure A8.3: Terms of the JPM variance equation of the **semi-τ** model during the year 2021



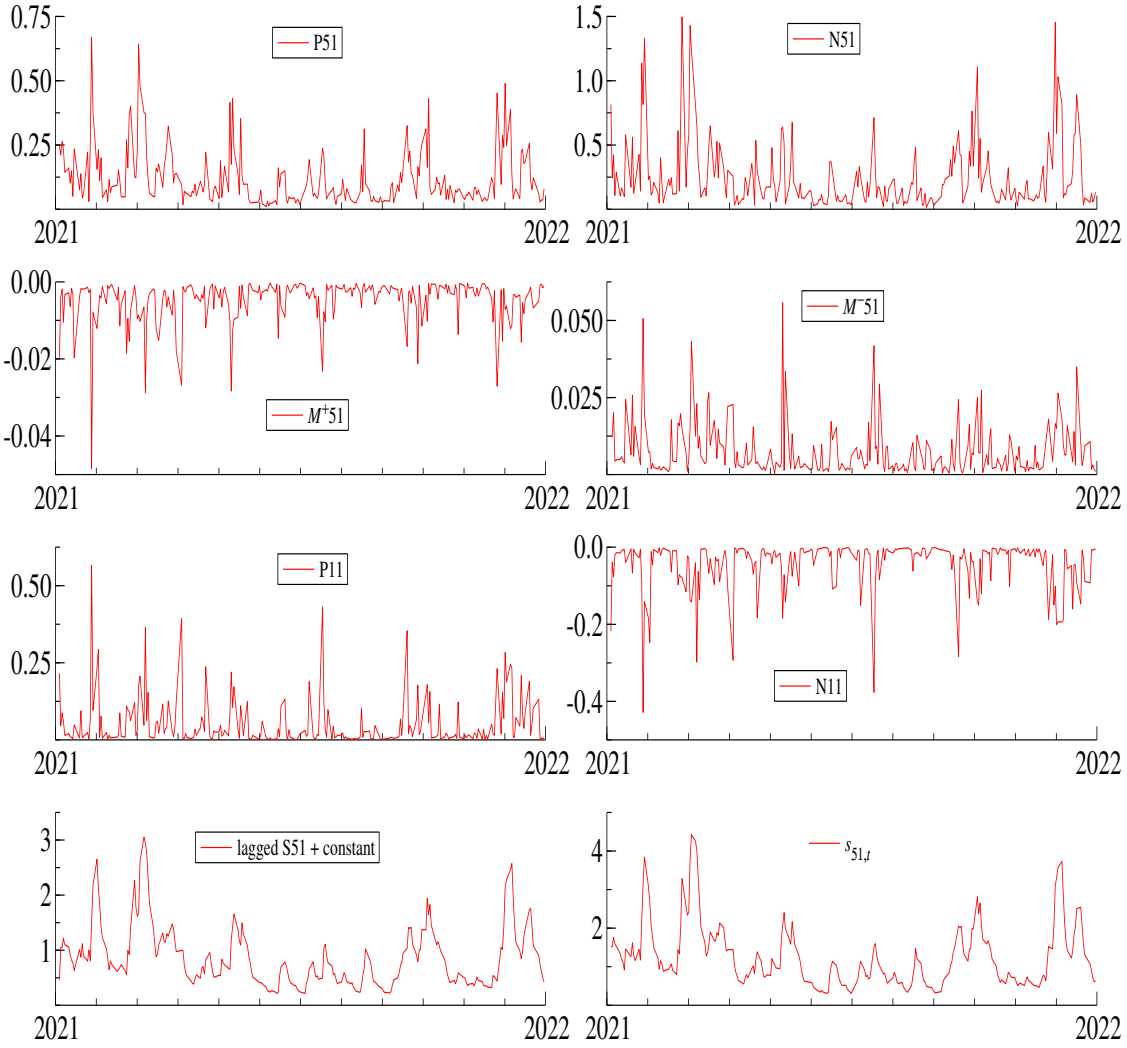
P55: $a^2_{P55}p_{55,t-1}$; P55+N55: P55+$a^2_{N55}n_{55,t-1}$; P51: $a^2_{P51}p_{11,t-1}$; P51+N51: P51+$a^2_{N51}n_{11,t-1}$;
sum of semi-cov 51 terms: sum of the four terms plotted in the bottom left graph;
P semi-cov 51: $2a_{P51}a_{P55}p_{51,t-1}$; N semi-cov 51: $2a_{N51}a_{N55}n_{51,t-1}$;
$\tau(M^+)$ semi-cov 51: $2a^+_{M51}a^+_{M55}[\tau(M^+_{t-1})]_{51}$; $\tau(M^-)$ semi-cov 51: $2a^-_{M51}a^-_{M55}[\tau(M^-_{t-1})]_{51}$;
lagged S55+constant: $b^2_{55}s_{55,t-1}$+constant;
$s_{55,t}$: P55+N55+P51+N51+sum of 4 semi-cov 51 terms+lagged S55+constant, i.e., fitted conditional variance. Values on the vertical axes are annualized.

Figure A8.4: Terms of the JPM variance equation of the **trPNτM** model during the year 2021



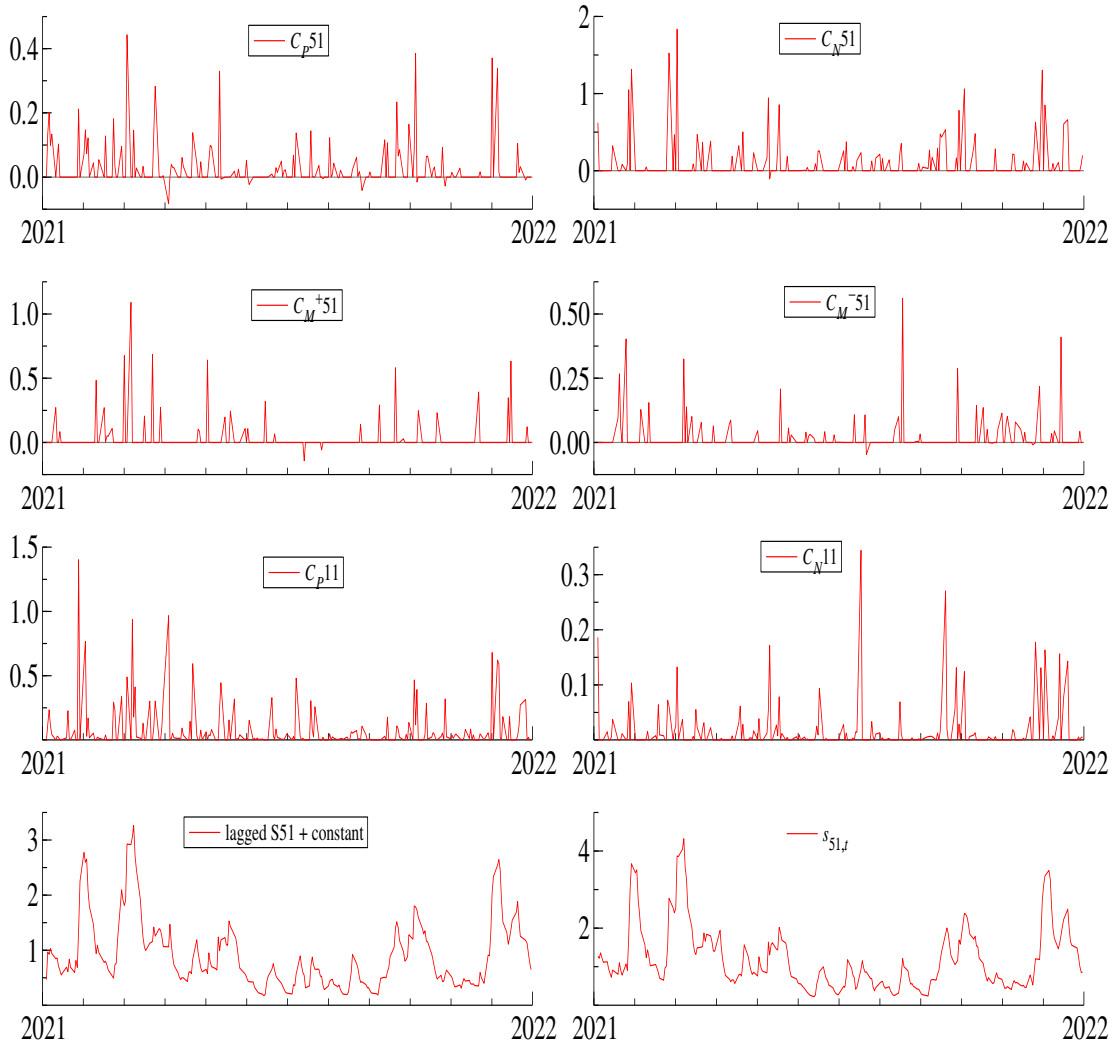
C_P55: $a^2_{P55}C_{P55,t-1}$; C_P55+C_N55: $C_P55+a^2_{N55}C_{N55,t-1}$;
C_P11: $a^2_{P51}C_{P11,t-1}$; C_P11+C_N11: $C_P11+a^2_{N51}C_{N11,t-1}$;
sum of cov 51 terms: sum of the four terms plotted in the bottom left graph;
C_P51: $2a_{P51}a_{P55}C_{P51,t-1}$; C_N51: $2a_{N51}a_{N55}C_{N51,t-1}$; C^+_M51: $2a^+_{M51}a^+_{M55}C^+_{M51,t-1}$;
C^-_M51: $2a^-_{M51}a^-_{M55}C^-_{M51,t-1}$; lagged S55+constant: $b^2_{55}s_{55,t-1}$+constant;
$s_{55,t}$: $C_P55+C_N55+C_P11+C_N11$+sum of 4 $C_\bullet 51$ terms+lagged S55+constant, i.e., fitted conditional variance. Values on the vertical axes are annualized.

Figure A8.5: Terms of the SPY-JPM covariance equation of the **semi-τ** model during the year 2021



P51: $a_{P11}a_{P55}p_{51,t-1}$; N51: $a_{N11}a_{N55}n_{51,t-1}$; M^+51: $a^+_{M11}a^+_{M55}[\tau(M^+_{t-1})]_{51}$; M^-51: $a^-_{M11}a^-_{M55}[\tau(M^-_{t-1})]_{51}$,;
P11: $a_{P11}a_{P51}p_{11,t-1}$; P11+N11: P11+$a_{N11}a_{N51}n_{11,t-1}$; lagged S51+constant: $b_{11}b_{55}s_{51,t-1}$+constant;
$s_{51,t}$: sum of 7 terms, i.e., fitted conditional covariance. Values on the vertical axes are annualized.

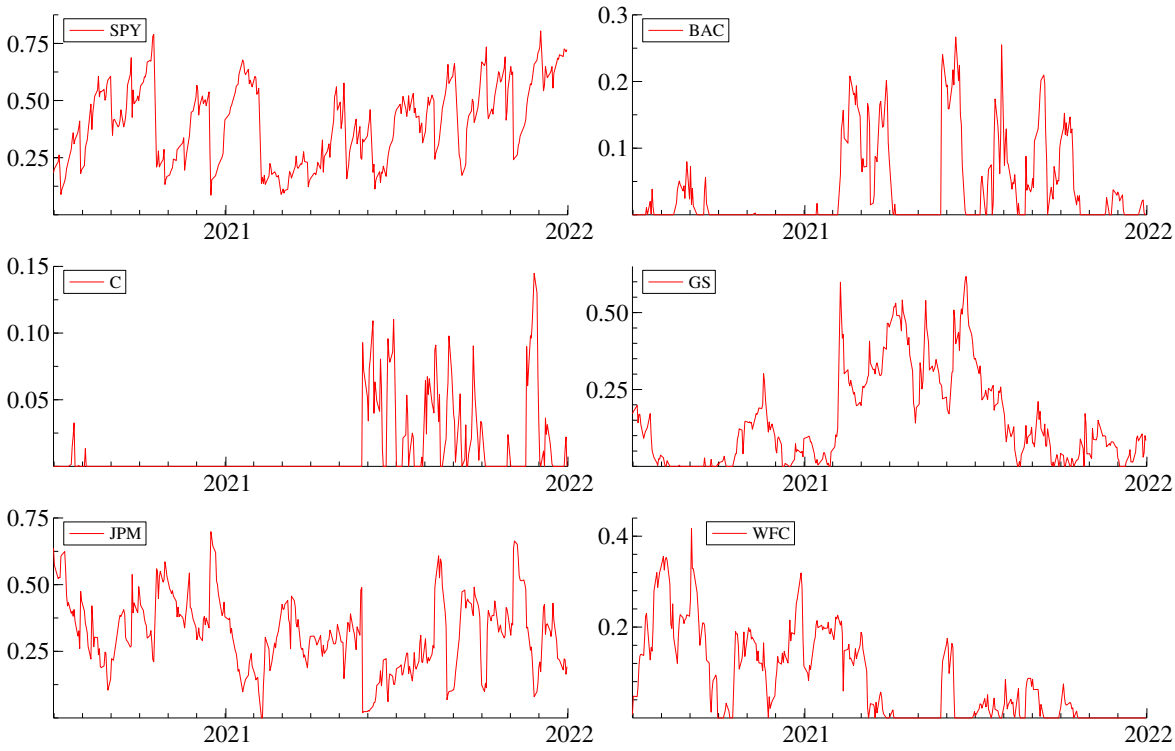
Figure A8.6: Terms of the SPY-JPM covariance equation of the **trPNτM** model during the year 2021



C_P51: $a_{P11}a_{P55}C_{P51,t-1}$; C_N51: $a_{N11}a_{N55}C_{N51,t-1}$; C_M^+51: $a^+_{M11}a^+_{M55}C^+_{M51,t-1}$; C_M^-51: $a^-_{M11}a^-_{M55}C^-_{M51,t-1}$;
C_P11: $a_{P11}a_{P51}C_{P11,t-1}$; C_N11: $a_{N11}a_{N51}C_{N11,t-1}$; lagged S51+constant: $b_{11}b_{55}s_{51,t-1}$+constant;
$s_{51,t}$: sum of 7 terms, i.e., fitted conditional covariance. Values on the vertical axes are annualized.

A.9 Optimal GMVP weights

Figure A9.1: Optimal GMVP weights



GMVP weights over the forecasting period (2020/07/01-2021/12/31)

Chapter 4

Hierarchical DCC-HEAVY Model for High-Dimensional Covariance Matrices

The contents of this chapter are the result of joint work with Prof. Matteo Barigozzi (University of Bologna). The paper is available at `2305.08488[econ.EM]`.

4.1 Introduction

In this paper, we develop a flexible framework that accurately captures the latent covariance structure of the high-dimensional asset returns and allows for sophisticated asymmetric dynamics in the covariances while at the same time keeping the estimation and forecasting straightforward and independent of the cross-sectional dimension of the assets under consideration.

Our methodology relates to the Realized Beta GARCH model of Hansen et al. (2014) and the corresponding extension of Archakov et al. (2020) that introduce the hierarchical-type factor framework based on the realized GARCH model (Hansen et al. (2012)), taking realized measures as direct inputs. In contrast, we model the

dynamics of both conditional and realized covariances (RC) in a GJR-type spirit (Glosten et al. (1993)). In addition, they focus on modelling the dynamics of daily returns and adopt intra-daily realized measures, leaving the dynamics of the residuals unspecified. Instead, we use monthly returns and construct realized measures via daily data. As such, we estimate and test our model that defines the conditional covariance matrices completely for much longer sample periods.

Given that no prior study investigates the forecasting ability of the hierarchical-type factor models, we assess both the statistical and economic performance of the distinct versions of our model in terms of the factor set and asymmetric dynamics, comparing them with the benchmark cDCC model (Aielli (2013)), the Realized Beta GARCH model (Hansen et al. (2014)), and its 3-Fama-French (FF) extension (Archakov et al. (2020)).

To perform empirical evaluations of the models, we utilize the data from a Kenneth French library on the three FF factors (Fama and French (1993)), i.e., market risk, size, and value, together with the momentum factor (Carhart (1997)), coupled with Yahoo Finance time series of the daily and monthly adjusted prices for a selected cross-section of individual assets, including all the stocks that belong to the S&P500 Index during the entire sample period from January 1962 until January 2023, i.e., $T = 732$.

Statistical evaluation criteria consist of the in-sample fit and out-of-sample forecast loss functions, i.e., the Euclidean distance (ED) and Frobenius norm (FN). From the economic point of view, we focus on the global minimum variance portfolio (GMVP) optimization, as the corresponding weights are determined solely by forecasts of the conditional covariance matrices over the given investment horizon. In this regard, the models are evaluated in terms of the forecasted conditional portfolio volatility. In order to formally determine whether the quality of the forecasts differs significantly across the models, we apply the model confidence set (MCS) procedure of Hansen

et al. (2011), which allows us to identify the subset of models that contains the best forecasting model given a pre-specified level of confidence. We also consider some typical features of the implied portfolio allocations, such as portfolio turnover rates and the short-selling proportion. Finally, we examine the economic significance of differences in portfolio volatility via a utility-based framework of Fleming et al. (2001, 2003).

Both the in-sample and forecasting results imply that our High-Dimensional (HD) DCC-HEAVY class of models significantly outperforms the existing hierarchical models of Hansen et al. (2014) and Archakov et al. (2020), as well as the benchmark cDCC model (Aielli (2013)). With regard to the latter, we prove the benefits of employing the higher-frequency data to model conditional covariances of lower-frequency returns. Conversely, the importance of specifying the RC dynamics could explain the poor performance of Realized GARCH-based models (Hansen et al. (2014); Archakov et al. (2020)). We confirm the robustness of our findings under changing market conditions.

The rest of the chapter is organized as follows. Section 4.2 introduces the hierarchical HD DCC-HEAVY models. Section 4.3 expounds on the estimation scheme, while the forecast formulas are provided in Section 4.4. Section 4.5 describes the empirical methodology, details the data used in the paper, and presents the in- and out-of-sample results of empirical exercises. Section 4.6 concludes.

4.2 Modelling Framework

Let us define a $K \times 1$ vector of returns related to the set of factors on month t as r_t^c and the corresponding RC matrix as RC_t^c. In addition, for $i = 1, \ldots, N$, we consider an individual asset return $r_{i,t}$ and associated realized measure between an individual asset and the set of factors $RC_{i,t}^c$.

In this regard, we observe the two types of information sets. $\mathcal{F}_t^c$, composed of the variables related to the set of factors, and $\mathcal{F}_t^{c,i}$, which further incorporates the observable information on an individual asset (for $i = 1, \ldots, N$).

We consider the factor model for an individual asset return:

$$r_{i,t} = \alpha_{i,t} + (\beta_{i,t})' r_t^c + \varepsilon_{i,t}; \tag{4.1}$$

$$\begin{aligned} \beta_{i,t} &= \operatorname{Var}(r_t^c | \mathcal{F}_{t-1}^c)^{-1} \operatorname{Cov}(r_{i,t}, r_t^c | \mathcal{F}_{t-1}^{c,i}) \\ &= (\operatorname{diag}(H_t^c)^{1/2} R_t^c \operatorname{diag}(H_t^c)^{1/2})^{-1} \operatorname{diag}(H_t^c)^{1/2} \rho_{i,t} (h_{i,t})^{1/2} \\ &= (\operatorname{diag}(H_t^c)^{1/2})^{-1} (R_t^c)^{-1} \rho_{i,t} (h_{i,t})^{1/2}; \\ \alpha_{i,t} &= \mu_i - (\beta_{i,t})' \mu^c, \end{aligned} \tag{4.2}$$

where $r_{i,t}$ is a close-to-close return of an individual asset on month t, r_t^c is a $K \times 1$ vector of returns of K factors, $\alpha_{i,t}$ and $\varepsilon_{i,t}$ are the intercept and idiosyncratic return component related to $r_{i,t}$, respectively, and $\beta_{i,t}$ is a $K \times 1$ vector of asset betas; $\operatorname{diag}(H_t^c)$ is a $K \times K$ diagonal matrix composed of the conditional variances of factors on month t, R_t^c is the corresponding $K \times K$ conditional correlation matrix, while $h_{i,t}$ denotes the conditional variance of an asset i and $\rho_{i,t}$ a $K \times 1$ vector of conditional correlations between an asset and the factors.

Similarly, the factor model for N individual asset returns:

$$r_t = \alpha_t + B_t r_t^c + \varepsilon_t, \tag{4.3}$$

where r_t is a $N \times 1$ vector of returns of individual assets on month t, α_t and ε_t are the corresponding $N \times 1$ vectors of intercepts and idiosyncratic return components, respectively, and B_t is a $N \times K$ matrix of asset betas. It follows readily:

$$\operatorname{Var}(r_t | \mathcal{F}_{t-1}^{c,i}) = B_t \operatorname{Var}(r_t^c | \mathcal{F}_{t-1}^c)(B_t)' + \Sigma_t, \tag{4.4}$$

with $\Sigma_t = \mathrm{E}(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}^{c,i})$.

To model (4.1)-(4.4), we primarily rely on a hierarchical method introduced by Hansen et al. (2014). In particular, $\mathcal{F}_t^c$ is adopted to build up the model for the dynamics of the set of factors. Subsequently, conditional on former estimates, we set up the framework for the dynamics between each individual asset and the factors by utilizing $\mathcal{F}_t^{c,i}$. Ultimately, the nonlinear shrinkage method (Ledoit and Wolf (2017)) is employed to define the covariances between idiosyncratic return components of the individual assets.

4.2.1 Marginal Model for a Set of Factors

We initially specify the marginal model for a set of factors by extending the recent DCC-HEAVY model (Bauwens and Xu (2023)) to allow for sophisticated asymmetric dynamics in the covariance matrices.

In this regard, we decompose a $K \times K$ conditional covariance matrix of K factors, i.e., $\mathrm{E}(r_t^c r_t^{c'} | \mathcal{F}_{t-1}^c) = H_t^c$, as

$$H_t^c = \mathrm{diag}(h_t^c)^{1/2} R_t^c \, \mathrm{diag}(h_t^c)^{1/2}, \tag{4.5}$$

where h_t^c is a $K \times 1$ vector of the conditional variances of factors on month t and R_t^c is the corresponding $K \times K$ conditional correlation matrix, given $\mathrm{E}(\mathrm{diag}(r_t^c r_t^{c'}) | \mathcal{F}_{t-1}^c) = \mathrm{diag}(h_t^c)$ and $\mathrm{E}(u_t^c u_t^{c'} | \mathcal{F}_{t-1}^c) = R_t^c$, with $u_t^c = r_t^c \odot (h_t^c)^{-1/2}$.
The dynamics of the conditional variances and correlations, allowing for asymmetric effects, are specified as

$$h_t^c = w_h + A_h^+ v_{t-1}^c \odot \mathrm{I}_{t-1}^+ + A_h^- v_{t-1}^c \odot \mathrm{I}_{t-1}^- + B_h h_{t-1}^c, \tag{4.6}$$

where v_t^c is a $K \times 1$ vector of the realized variances of factors on month t, w_h is a $K \times 1$ positive vector, and A_h^+, A_h^-, and B_h are the $K \times K$ diagonal matrices of

coefficients with positive diagonal entries less than 1, with $\odot$ denoting the Hadamard (element-wise) product of matrices, $\mathrm{I}_t^+ = [1_{\{r_{1,t}^c>0\}}, \ldots, 1_{\{r_{K,t}^c>0\}}]'$ the indicator vector of the positive monthly returns, and $\mathrm{I}_t^- = [1_{\{r_{1,t}^c\leq 0\}}, \ldots, 1_{\{r_{K,t}^c\leq 0\}}]'$ the indicator vector of the negative monthly returns.

Correspondingly,

$$R_t^c = \tilde{R} + \alpha_R RL_{t-1}^c + \beta_R R_{t-1}^c, \tag{4.7}$$

where RL_t^c is a $K \times K$ realized correlation matrix of the factors on month t, and α_R and β_R are non-negative scalar parameters, i.e., $\beta_R = 0$ if $\alpha_R = 0$ and $\beta_R < 1$, with $\tilde{R} = (1 - \beta_R)\overline{R} - \alpha_R\overline{P}$, i.e., $\mathrm{E}(u_t^c u_t^{c'}) = \overline{R}$ and $\mathrm{E}(RL_t^c) = \overline{P}$ set to the empirical counterparts.

Analogously, we decompose a $K \times K$ conditional mean of the RC matrix of K factors, i.e., $\mathrm{E}(RC_t^c|\mathcal{F}_{t-1}^c) = M_t^c$, as

$$M_t^c = \mathrm{diag}(m_t^c)^{1/2} P_t^c \,\mathrm{diag}(m_t^c)^{1/2}, \tag{4.8}$$

where m_t^c is a $K \times 1$ vector of the conditional means of realized variances of factors on month t and P_t^c is the corresponding $K \times K$ conditional mean of realized correlations, i.e., $\mathrm{E}(RL_t^c|\mathcal{F}_{t-1}^c) = P_t^c$.

The dynamics of the realized variances and correlations, allowing for daily asymmetric effects, are specified as

$$m_t^c = w_m + A_m^+ v_{t-1}^{c+} + A_m^- v_{t-1}^{c-} + B_m m_{t-1}^c, \tag{4.9}$$

where v_t^{c+} and v_t^{c-} are the $K \times 1$ vectors of the positive and negative realized semi-variances (Shephard and Sheppard (2010)) of factors, respectively, w_m is a $K \times 1$ positive vector, and A_m^+, A_m^-, and B_m are the $K \times K$ diagonal matrices of coefficients with positive diagonal entries below 1.

Specifically, for $i = 1, \ldots, K$ and $j = 1, \ldots, m$, $v_{i,t}^{c+} = \sum_{j=1}^{m} (r_{i,j,t}^{c+})^2$ and $v_{i,t}^{c-} = \sum_{j=1}^{m} (r_{i,j,t}^{c-})^2$, where $r_{i,j,t}^{c+} = r_{i,j,t}^{c} \times 1_{\{r_{i,j,t}^{c} > 0\}}$ and $r_{i,j,t}^{c-} = r_{i,j,t}^{c} \times 1_{\{r_{i,j,t}^{c} \leq 0\}}$ denote the positive and negative daily returns, respectively.

Correspondingly,

$$P_t^c = (1 - \alpha_P - \beta_P)\overline{P} + \alpha_P RL_{t-1}^c + \beta_R P_{t-1}^c, \tag{4.10}$$

where α_P and β_P are non-negative scalar parameters, i.e., $\beta_P = 0$ if $\alpha_P = 0$ and $\alpha_P + \beta_P < 1$, with $\mathrm{E}(RL_t^c) = \overline{P}$ set to the empirical counterpart.

4.2.2 Model for Individual Asset Returns

By assuming that the conditional distribution of individual asset returns depends on the factors but not vice versa (Hansen et al. (2014)), the standardized return of each asset is conditionally jointly distributed with 'degarched' factors, i.e.,

$$\begin{pmatrix} u_t^c \\ u_{i,t} \end{pmatrix} | \mathcal{F}_{t-1}^{c,i} \sim N \left(\begin{pmatrix} 0_{K\times 1} \\ 0_{1\times 1} \end{pmatrix}, R_{i,t}^c \right), \tag{4.11}$$

where a $(K+1) \times (K+1)$ joint conditional correlation matrix $R_{i,t}^c$ is given by:

$$R_{i,t}^c = \begin{pmatrix} R_t^c & \rho_{i,t} \\ (\rho_{i,t})' & 1 \end{pmatrix}, \tag{4.12}$$

where R_t^c and $\rho_{i,t}$ denote the $K \times K$ conditional correlation matrix of factors filtered from a marginal model and $K \times 1$ vector of correlations between an individual asset and the factors on month t, respectively.

In accordance to the framework for a set of factors, the dynamics of the conditional and realized variance of an individual asset, allowing for corresponding asymmetric

effects are specified as

$$h_{i,t} = c_{i,h} + a^{+}_{i,h} v_{i,t-1} 1_{[r_{i,t-1}>0]} + a^{-}_{i,h} v_{i,t-1} 1_{[r_{i,t-1}\leq 0]} + b_{i,h} h_{i,t-1}, \tag{4.13}$$

where $h_{i,t}$ and $v_{i,t}$ denote the conditional and realized variance of an asset i on month t, respectively, and $c_{i,h}, a^{+}_{i,h}, a^{-}_{i,h}$, and $b_{i,h}$ are non-negative scalar coefficients;

$$m_{i,t} = c_{i,m} + a^{+}_{i,m} v^{+}_{i,t-1} + a^{-}_{i,m} v^{-}_{i,t-1} + b_{i,m} m_{i,t-1}, \tag{4.14}$$

where $m_{i,t}$, $v^{+}_{i,t}$, and $v^{-}_{i,t}$ denote the conditional mean of the realized variance, positive and negative semi-variance of an asset i on month t, respectively, and $c_{i,m}, a^{+}_{i,m}, a^{-}_{i,m}$, and $b_{i,m}$ are non-negative scalar coefficients.

Finally, to model the vectors of correlations between the returns of an individual asset and the set of factors, we utilize the Fisher transformation, i.e., $\mathbb{F}(\cdot)$, to map each element from a closed interval $(-1, 1)$ into $\mathbb{R}$ within the typical HEAVY-type recursions (Noureldin et al. (2012); Bauwens and Xu (2023)):

$$\mathbb{F}(\rho_{i,t}) = \phi_{i,R} + \alpha_{i,R} \mathbb{F}(rl_{i,t-1}) + \beta_{i,R} \mathbb{F}(\rho_{i,t-1}), \tag{4.15}$$

where $\rho_{i,t}$ and $rl_{i,t}$ denote the $K \times 1$ vectors of conditional and realized correlations of an asset i with factors on month t, respectively, and $\phi_{i,R}$, $\alpha_{i,R}$, and $\beta_{i,R}$ are non-negative scalar parameters;

$$\mathbb{F}(p_{i,t}) = \phi_{i,P} + \alpha_{i,P} \mathbb{F}(rl_{i,t-1}) + \beta_{i,P} \mathbb{F}(p_{i,t-1}), \tag{4.16}$$

where $p_{i,t}$ denotes a $K \times 1$ vector of the conditional means of realized correlations of an asset i with factors on month t, and $\phi_{i,P}$, $\alpha_{i,P}$, and $\beta_{i,P}$ are non-negative scalar parameters.

4.2.3 Idiosyncratic Dynamics

Based on formulas (4.1)–(4.4), to fully specify the conditional covariance matrices of individual asset returns, we should define the dynamics of the residuals, i.e., $\mathrm{E}(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}^{c,i})$.

In line with most of the literature, we treat the assumption of an exact factor model as strict. As such, for the underlying approximate factor model, we propose applying the nonlinear (NL) shrinkage method of Ledoit and Wolf (2017) to the sample covariance matrix of the residuals, which has been proved preferable with respect to both the linear shrinkage of Ledoit and Wolf (2004) (Ledoit and Wolf (2017)) and thresholding schemes (De Nard et al. (2021)).[1] This methodology implies shifting the eigenvalues of the empirical covariance matrix via the out-of-sample optimization of the minimum variance loss function subject to a required return constraint (Engle and Colacito (2006)).

It follows directly:

$$\hat{\beta}_{i,t} = (\mathrm{diag}(\hat{H}_t^c)^{1/2})^{-1} (\hat{R}_t^c)^{-1} \hat{\rho}_{i,t} (\hat{h}_{i,t})^{1/2} \tag{4.17}$$

and

$$\hat{\mathrm{Var}}(r_t | \mathcal{F}_{t-1}^{c,i}) = \hat{B}_t \hat{\mathrm{Var}}(r_t^c | \mathcal{F}_{t-1}^c)(\hat{B}_t)' + \hat{\Sigma}_{\hat{\varepsilon}} = \hat{B}_t \hat{H}_t^c (\hat{B}_t)' + \hat{\Sigma}_{\hat{\varepsilon}}, \tag{4.18}$$

where matrices $\hat{H}_t^c$ and $\hat{R}_t^c$ are filtered from the core model, i.e., (4.5)-(4.10), whereas each conditional variance $\hat{h}_{i,t}$ and the corresponding correlation vector $\hat{\rho}_{i,t}$ are extracted from the individual factor model related to an asset i, i.e., (4.13)-(4.16), for $i = 1, ..., N$. The NL shrinkage method (Ledoit and Wolf (2017)) delivers $\hat{\Sigma}_{\hat{\varepsilon}}$.

[1] Alternatively, the dynamic Σ_t could be defined via the benchmark dynamic conditional correlation (DCC) model (Engle (2002)) for the cross-section of $N \leq 100$ assets. Conversely, when the number of individual assets is large, the DCC-NL model introduced by Engle et al. (2019) might be adopted. In each case, the estimation of the additional $3N + 2$ parameters is required.

4.3 Estimation

The hierarchical structure of the introduced model suggests a convenient step-by-step estimation procedure independent of the cross-sectional dimension of the assets under consideration. As follows, we discuss the quasi-maximum likelihood (QML) estimation scheme and define the corresponding log-likelihood functions (LLF).

Initially, to estimate the core model for a set of factors, we essentially follow the approach of Bauwens and Xu (2023), by partitioning the parameters of both conditional and realized covariances into the coefficients of the corresponding variance and correlation equations.[2]

In particular, let us define the two parameter sets θ_H^c and θ_M^c for the conditional and realized covariances of factors, respectively.

Given the hypothesis that the distribution of the 'degarched' monthly return vector is multivariate Gaussian (4.11), the first step consists of estimating the parameters of the conditional variances (4.6), i.e., θ_{H1}^c, and correlations (4.7), i.e., θ_{H2}^c, for the set of factors by maximizing the following QML functions:

$$\begin{aligned} LLF_{H1}^c(\theta_{H1}^c|\mathcal{F}_{t-1}^c) &= -\frac{1}{2}\sum_{t=1}^{T}\left\{2\log\left|\mathrm{diag}(h_t^c)^{1/2}\right| + u_t^{c'}u_t^c\right\}; \\ LLF_{H2}^c(\theta_{H2}^c|\hat{\theta}_{H1}^c;\mathcal{F}_{t-1}^c) &= -\frac{1}{2}\sum_{t=1}^{T}\left\{\log|R_t^c| + \hat{u}_t^{c'}(R_t^c)^{-1}\hat{u}_t^c\right\}, \end{aligned} \tag{4.19}$$

where $\hat{u}_t^c = r_t^c \odot (\hat{h}_t^c)^{-1/2}$, with $\hat{h}_t^c$ defined via $\hat{\theta}_{H1}^c$.

Bauwens and Xu (2023) show that the estimated parameters for conditional correlations (4.7), i.e., (α_R, β_R), do not automatically guarantee the PD-ness of R_t^c. As such, we proceed by checking the condition during the numerical maximization of LLF_{H2}^c.

To specify the dynamics of realized measures, we assume that the probability density function of RC matrices RC_t^c, conditional on the filtration $\mathcal{F}_{t-1}^c$, is Wishart,

[2]The parameter sets can be alternatively estimated without splitting by maximizing the corresponding full LLFs (see Bauwens and Xu (2023)).

i.e.,

$$RC_t^c|\mathcal{F}_{t-1}^c \sim W_K(\nu, M_t^c(\theta_M^c)/\nu), \tag{4.20}$$

where $W_K(\nu, M_t^c(\theta_M^c)/\nu)$ denotes the K-dimensional central Wishart distribution with $\nu \geq K$ degrees of freedom and PD $K \times K$ scale matrix $M_t^c(\theta_M^c)/\nu$, implying $\mathrm{E}(RC_t^c|\mathcal{F}_{t-1}^c) = M_t^c(\theta_M^c)$.

Correspondingly, we split θ_M^c into the parameters for realized variances (4.9), i.e., θ_{M1}^c, and realized correlations (4.10), i.e., θ_{M2}^c. The second-step objective functions for T observations are given by:

$$\begin{aligned}
LLF_{M1}^c(\theta_{M1}^c|\mathcal{F}_{t-1}^c) &= -\frac{\nu}{2}\sum_{t=1}^{T}\left\{2\log|L_t^c| + \mathrm{trace}\left[(L_t^c)^{-1}RC_t^c(L_t^c)^{-1}\right]\right\}; \\
LLF_{M2}^c(\theta_{M2}^c|\hat{\theta}_{M1}^c;\mathcal{F}_{t-1}^c) &= -\frac{\nu}{2}\sum_{t=1}^{T}\left\{\log|P_t^c| + \mathrm{trace}\left[((P_t^c)^{-1} - \mathrm{I}_K)(\hat{L}_t^c)^{-1}RC_t^c(\hat{L}_t^c)^{-1}\right]\right\},
\end{aligned} \tag{4.21}$$

where I_K denotes the identity matrix of order K, $L_t^c = \mathrm{diag}(m_t^c)^{1/2}$, with $\hat{L}_t^c$ defined via $\hat{\theta}_{M1}^c$, and the parameter ν set equal to 1.[3]

Next, we consider the likelihood contributions for the conditional model of each individual asset return. It follows from the assumptions (4.11) and (4.12), the conditional distribution of the standardized monthly asset return:

$$u_{i,t}|u_t^c \sim N\left((\rho_{i,t})'(R_t^c)^{-1}u_t^c, 1 - (\rho_{i,t})'(R_t^c)^{-1}\rho_{i,t}\right). \tag{4.22}$$

As such, the underlying LLF with regard to the conditional covariances of an asset i

$$LLF_{H_i}^{c,i}(\theta_{H_i}|\mathcal{F}_{t-1}^{c,i}) = -\frac{1}{2}\sum_{t=1}^{T}\left\{\log\left(h_{i,t}\left(1 - (\rho_{i,t})'(R_t^c)^{-1}\rho_{i,t}\right)\right) + \frac{(u_{i,t} - (\rho_{i,t})'(R_t^c)^{-1}u_t^c)^2}{(1 - (\rho_{i,t})'(R_t^c)^{-1}\rho_{i,t})}\right\}, \tag{4.23}$$

[3]The score for θ_M^c is proportional to ν (Bauwens et al. (2012)).

directly follows from:

$$
\begin{aligned}
\mathrm{Cov}(r_{i,t}, r_t^c|\mathcal{F}_{t-1}^{c,i}) &= \mathrm{diag}(h_t^c)^{1/2}\rho_{i,t}(h_{i,t})^{1/2};\\
\mathrm{Var}(r_{i,t}|r_t^c, \mathcal{F}_{t-1}^{c,i}) &= h_{i,t} - \frac{(\mathrm{diag}(h_t^c)^{1/2}\rho_{i,t}(h_{i,t})^{1/2})'(\mathrm{diag}(h_t^c)^{1/2}\rho_{i,t}(h_{i,t})^{1/2})}{\mathrm{diag}(h_t^c)^{1/2}R_t^c\,\mathrm{diag}(h_t^c)^{1/2}}\\
&= h_{i,t}\left(1 - (\rho_{i,t})'(R_t^c)^{-1}\rho_{i,t}\right);\\
\mathrm{E}(r_{i,t}|r_t^c, \mathcal{F}_{t-1}^{c,i}) &= \mu_i + \frac{(\mathrm{diag}(h_t^c)^{1/2}\rho_{i,t}(h_{i,t})^{1/2})'}{\mathrm{diag}(h_t^c)^{1/2}R_t^c\,\mathrm{diag}(h_t^c)^{1/2}}(r_t^c - \mu^c) = \mu_i + (h_{i,t})^{1/2}(\rho_{i,t})'(R_t^c)^{-1}u_t^c.
\end{aligned}
\tag{4.24}
$$

To ensure the positivity of the joint conditional correlation matrix $R_{i,t}^c$ (4.12), we must ensure $(\rho_{i,t})'(R_t^c)^{-1}\rho_{i,t} < 1$ for each $t = 1, ..., T$, during the estimation (Archakov et al. (2020)).

In analogous fashion as for the conditional correlations (4.12), we use a partitioning of the realized measures so that a $(K+1)\times(K+1)$ joint conditional mean of the realized correlation matrix, i.e., $P_{i,t}^c$, is given by:

$$
P_{i,t}^c = \begin{pmatrix} P_t^c & p_{i,t} \\ (p_{i,t})' & 1 \end{pmatrix}, \tag{4.25}
$$

where P_t^c and $p_{i,t}$ denote the $K \times K$ conditional mean of the realized correlation matrix of factors filtered from a marginal model and $K \times 1$ vector of the conditional expectations of correlations between an individual asset and the factors on month t, respectively.

In this regard, the QML function reads as

$$
LLF_{M_i}^{c,i}(\theta_{M_i}|\mathcal{F}_{t-1}^{c,i}) = -\frac{\nu}{2}\sum_{t=1}^{T}\left\{\log\left(m_{i,t}(1-p_{i|c,t})\right) + \frac{v_{i,t} - (rc_{i,t})'(RC_t^c)^{-1}rc_{i,t}}{m_{i,t}(1-p_{i|c,t})}\right\}, \tag{4.26}
$$

where $m_{i,t}$ denotes the conditional mean of the realized variance of an asset i, i.e., $v_{i,t}$, $rc_{i,t}$ is a $K \times 1$ vector of RC between an asset i and K factors, and

$p_{i|c,t} = (p_{i,t})'(P_t^c)^{-1}p_{i,t}$. Analogously, we set ν equal to 1.

In order to estimate the model for the cross-section of N assets, we initially estimate the marginal model for a set of factors followed by the separate estimations of individual models for $i = 1, ..., N$, conditional on variables obtained via the core model. Finally, we apply the NL shrinkage of Ledoit and Wolf (2017) to obtain the conditional covariances of derived residuals, i.e., $\hat{\varepsilon}_t = r_t - \hat{\alpha}_t - \hat{B}_t r_t^c$.

Considering the estimation of the core model, the total number of parameters with respect to K factors is $8K + 4$. Given the assumption of the diagonal matrices of coefficients for the variance equations, we split the estimation of $8K$ parameters for the variances into K univariate HEAVY models (Shephard and Sheppard (2010)). Conversely, the model for each individual asset requires the specification of 14 additional parameters. As follows, a total of $8K + 4 + 14N$ coefficients is generated for the cross-sectional dimension of N assets.

4.4 Forecasting

Forecasting the covariance matrices of asset returns is paramount in derivative pricing, asset allocation, and risk management decisions.
In this regard, in our experiments, we focus on the 1-step-ahead predictions of the conditional covariances of monthly returns for the selected cross-section of N individual assets, i.e., $\text{Var}(r_{t+1}|\mathcal{F}_t^{c,i})$, directly computable via

$$\hat{\text{Var}}(r_{t+1}|\mathcal{F}_t^{c,i}) = \hat{B}_{t+1}\hat{\text{Var}}(r_{t+1}^c|\mathcal{F}_t^c)(\hat{B}_{t+1})' + \hat{\Sigma}_{\hat{\varepsilon}} = \hat{B}_{t+1}\hat{H}_{t+1}^c(\hat{B}_{t+1})' + \hat{\Sigma}_{\hat{\varepsilon}}, \qquad (4.27)$$

where $\hat{H}_{t+1}^c$ is a $K \times K$ predicted conditional covariance matrix of factors for the month $t + 1$, $\hat{B}_{t+1}$ is a $N \times K$ matrix of predicted asset betas, and $\hat{\Sigma}_{\hat{\varepsilon}}$ is a $N \times N$ conditional covariance matrix of the forecasted residuals.

In particular, for each asset i and time $t+1$:

$$\hat{\beta}_{i,t+1} = (\text{diag}(\hat{H}^c_{t+1})^{1/2})^{-1}(\hat{R}^c_{t+1})^{-1}\hat{\rho}_{i,t+1}(\hat{h}_{i,t+1})^{1/2}, \tag{4.28}$$

where $\text{diag}(\hat{H}^c_{t+1})$ is a $K \times K$ diagonal matrix composed of the conditional variances of factors for the month $t+1$, $\hat{R}^c_{t+1}$ is the corresponding $K \times K$ conditional correlation matrix, $\hat{h}_{i,t+1}$ denotes the predicted conditional variance of an asset i, and $\hat{\rho}_{i,t+1}$ a $K \times 1$ vector of the forecasted conditional correlations between an asset i and the factors.

4.5 Empirical Application

4.5.1 Data Construction and Description

For the subsequent empirical analyses, we use monthly returns on factors and assets, and construct realized measures of variances and covariances using daily returns observed within each month. In particular, we compute monthly covariance matrices with the corresponding realized analogues with respect to the three FF factors (Fama and French (1993)), i.e., market risk, size, and value, together with the momentum factor (Carhart (1997)), based on the data obtained from a Kenneth French library.

Our model is tested for a selected cross-section of individual assets, consisting of all the stocks that belong to the S&P500 Index during the entire sample period from January 1962 until January 2023, i.e., $N = 20$ and $T = 732$.

The stock names and tickers are: American Electric Power Company Inc. (AEP), The Boeing Company (BA), Caterpillar Inc. (CAT), Chevron Corporation (CVX), DTE Energy Company (DTE), Consolidated Edison Inc. (ED), General Dynamics Corporation (GD), General Electric Company (GE), Honeywell International Inc. (HON), International Business Machines Corporation (IBM), International Paper Company

(IP), The Coca-Cola Company (KO), The Kroger Co. (KR), 3M Company (MMM), Altria Group Inc. (MO), Merck & Co. Inc. (MRK), Marathon Oil Corporation (MRO), Motorola Solutions Inc. (MSI), The Procter & Gamble Company (PG), and Exxon Mobil Corporation (XOM).

We build the corresponding time series of the monthly and close-to-close daily returns for each asset based on the prices adjusted for dividends and splits available on Yahoo Finance.

As a result, the empirical application at the monthly frequency with realized measures built upon daily data allows for estimating and testing the models for a long sample period.[4]

Table 4.1 reports, for each factor, the time series means and standard deviations of the realized variances, their 'positive' and 'negative' components used to specify the asymmetric dynamics, and squared monthly returns.[5] The last row indicates the average of the time series means and standard deviations of realized correlations between the factors. The same statistics for the individual assets are shown in the Appendix B.1, i.e., Table B.1.1.

Table 4.1: Summary statistics for the 3 FF and MOM factors

Factor	MKT	SMB	HML	MOM
r_{cc}^2	2.51 (5.07)	1.09 (2.78)	1.05 (2.21)	2.25 (9.37)
RV	2.65 (5.75)	0.73 (1.34)	0.83 (1.66)	1.49 (3.40)
P	1.25 (2.30)	0.34 (0.47)	0.44 (0.93)	0.65 (1.19)
N	1.40 (3.73)	0.39 (0.98)	0.39 (0.82)	0.84 (2.40)
GJR_P	1.12 (2.15)	0.31 (0.49)	0.46 (1.30)	0.66 (1.60)
GJR_N	1.53 (5.65)	0.42 (1.35)	0.37 (1.19)	0.83 (3.17)
RL	–0.10 (0.50)	–0.03 (0.42)	–0.14 (0.44)	0.01 (0.50)

r_{cc}^2: squared close-to-close monthly return; RV: realized variance; P: positive semi-variance; N: negative semi-variance; GJR_P: RV if monthly return is positive, 0 if negative; GJR_N: RV if monthly return is negative, 0 if positive; RL: realized correlation, the average of the 3 time series means and sd-s of realized correlations with the other 3 factors.

[4]N.B. In order to estimate and forecast daily conditional covariances within the current framework, the accurate replication of the factors intra-daily requires high-frequency (HF) data access with respect to the entire universe of stocks listed on the NYSE, NASDAQ, and AMEX (see Aït-Sahalia et al. (2020)).

[5]The data are annualized in percentage, i.e., multiplied by 1200.

Considering the statistics reported in Table 4.1, the market and momentum factors appear more volatile compared to the size and value factors. Except for the market factor, each average realized variance is only a fraction of the corresponding average squared close-to-close return. The average negative semi-variance (N) of each factor, except HML, is larger than the average positive component (P). The same applies for the portions of the variances with respect to the signs of monthly returns, i.e., GJR_P and GJR_N. Besides MOM, all the factors have a negative average realized correlation with respect to the others.

The analogous summary measures for the individual assets, i.e., Table B.1.1, generally suggest that the average realized variance exceeds the corresponding average squared close-to-close return. It might not be suprising, given the realized measures obtained via daily returns that account for the overnight information. In contrast to the set of factors, the average positive semi-variance (P) is larger than the negative component (N). Conversely, the portions of the variances with respect to the negative monthly returns, i.e., GJR_N, exceed the GJR_P. Ultimately, the average realized correlations of all the assets with factors lie in a narrow interval, ranging from 0.28 to 0.36, with rather similar standard deviations.

Figures 4.1-4.2 show the time series of the realized variances of the market and momentum factors and the components of their semi-variance decompositions. They illustrate the occurrence of a few clustered extreme values, consistent with periods of financial turbulence. In both cases, the extreme volatility is largely attributed to the negative semi-variance due to the prevailing negative daily returns during the crises.

Figure 4.3 illustrates the time series of the realized correlations between BA and each factor. The patterns of correlations with the three FF factors are comparable, with BA being correlated the most with the market factor. On the other hand, the correlations with MOM are more dispersed and volatile.

Figure 4.1: Annualized realized variances of the market factor and the corresponding semi-variance decomposition

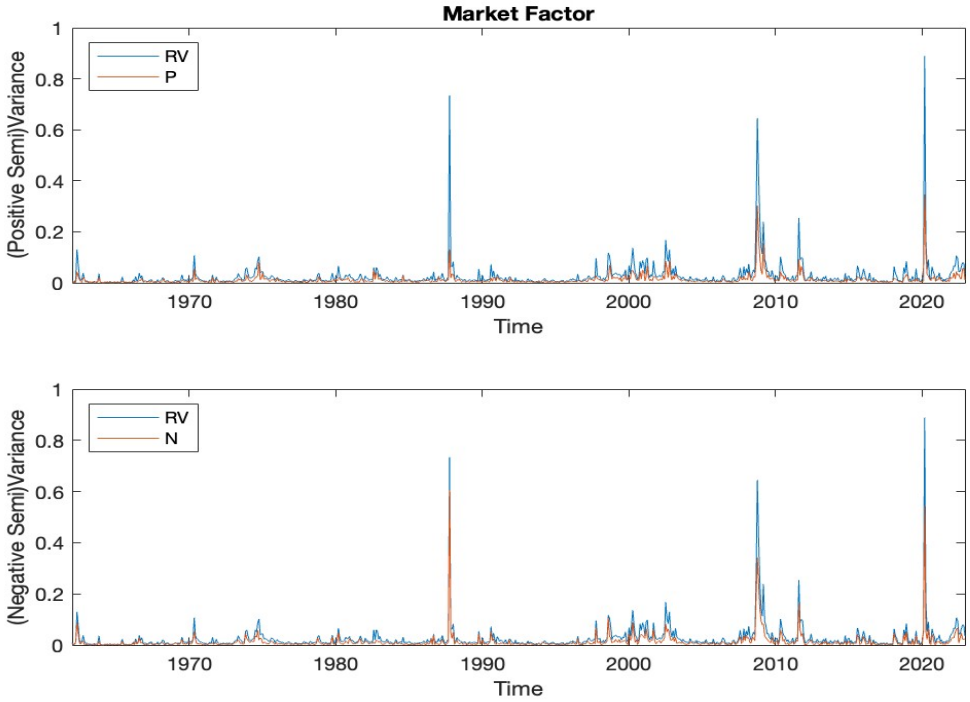


Figure 4.2: Annualized realized variances of the momentum factor and the corresponding semi-variance decomposition

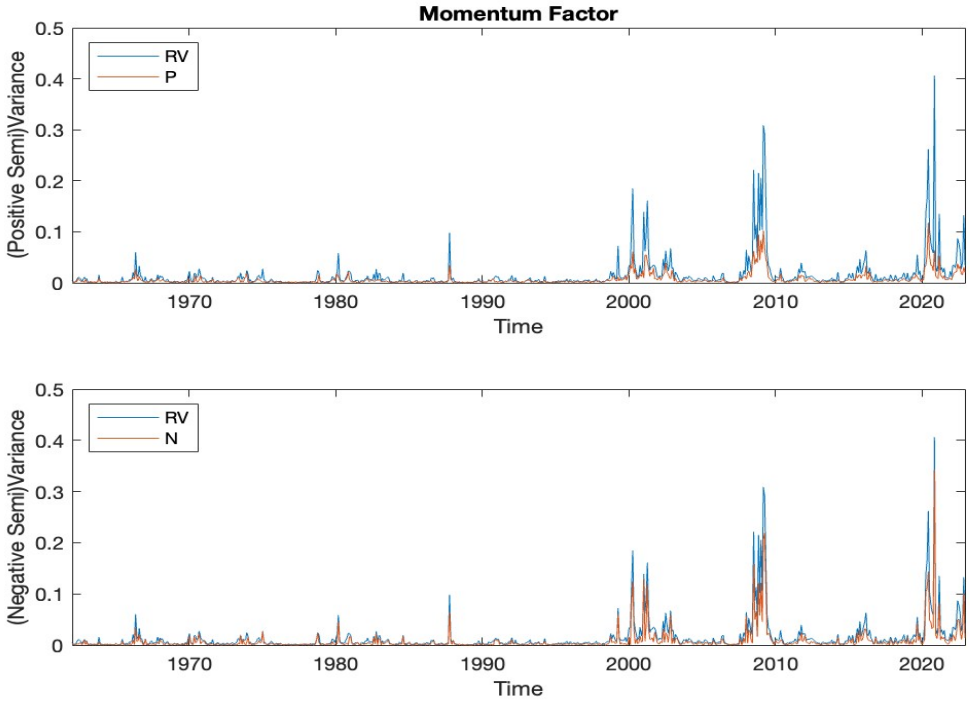
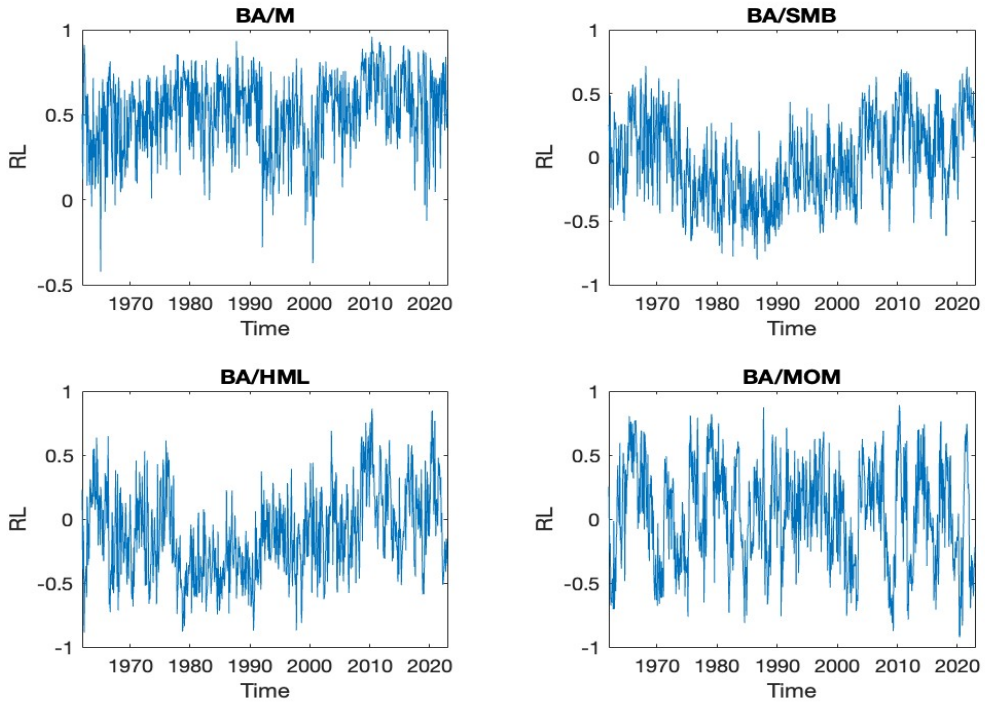


Figure 4.3: Realized correlations of BA with the three FF factors and MOM



4.5.2 In-Sample Fit

To evaluate the in-sample fit of the benchmark HD DCC-HEAVY model, i.e., "4F-HD DCC-HEAVY" with $K = 4$ (3 FF and MOM factors), we additionally consider restricted versions with respect to the set of factors and asymmetric effects.

In particular, we estimate the variants by assuming that equity returns are either explained via the 3 FF factors ("FF-HD DCC-HEAVY") or market factor only ("M-HD DCC-HEAVY").

In addition, we examine whether allowing for asymmetries in the covariance dynamics allows for improving the fit by specifying the modelling equations of the benchmark model without accounting for the signs of underlying returns ("sym-HD DCC-HEAVY").

The corresponding variance equations for "sym-HD DCC-HEAVY" are given by:

$$h_t^c = w_h + A_h v_{t-1}^c + B_h h_{t-1}^c, \tag{4.29}$$

where v_t^c is a $K \times 1$ vector of the realized variances of factors on month t, w_h is a $K \times 1$ positive vector, and A_h and B_h are the $K \times K$ diagonal matrices of coefficients;

$$m_t^c = w_m + A_m v_{t-1}^c + B_m m_{t-1}^c, \tag{4.30}$$

where m_t^c is a $K \times 1$ vector of the conditional means of realized variances of factors on month t, w_m is a $K \times 1$ positive vector, and A_m and B_m are the $K \times K$ diagonal matrices of coefficients;

$$h_{i,t} = c_{i,h} + a_{i,h} v_{i,t-1} + b_{i,h} h_{i,t-1}, \tag{4.31}$$

where $h_{i,t}$ and $v_{i,t}$ denote the conditional and realized variance of an asset i on month t, respectively, and $c_{i,h}, a_{i,h}$, and $b_{i,h}$ are non-negative scalar coefficients;

$$m_{i,t} = c_{i,m} + a_{i,m} v_{i,t-1} + b_{i,m} m_{i,t-1}, \tag{4.32}$$

where $m_{i,t}$ denotes the conditional mean of the realized variance of an asset i on month t, and $c_{i,m}, a_{i,m}$, and $b_{i,m}$ are non-negative scalar coefficients.

Correspondingly, to capture potential additional information provided by the 'high-frequency (HF)' daily data, we estimate cDCC model (Aielli (2013)) built exclusively upon monthly data, which has been widely applied to capture the dynamics of time-varying betas (e.g., Engle and Kelly (2012); Bali et al. (2017)).

In cDCC model (Aielli (2013)), the variance of the market and individual asset is modelled as the benchmark univariate GARCH (Bollerslev (1986)), while the dynamics of the conditional correlations are specified as:

$$\begin{aligned} R_t &= \operatorname{diag}(Q_t)^{-1/2} Q_t \operatorname{diag}(Q_t)^{-1/2}, \\ Q_t &= (1-\alpha-\beta)S + \alpha[\operatorname{diag}(Q_{t-1})^{1/2} u_{t-1} u_{t-1}' \operatorname{diag}(Q_{t-1})^{1/2}] + \beta Q_{t-1}, \end{aligned} \tag{4.33}$$

where R_t denotes the conditional correlation matrix on month t, S is a symmetric matrix with unit diagonal elements, $u_t = r_t \odot (h_t)^{-1/2}$, with the vector of returns of the market and individual asset $r_t = (r_{m,t}, r_{i,t})'$, coupled with conditional variances $h_t = (h_{m,t}, h_{i,t})'$, and α and β are scalar coefficients.

Ultimately, we consider the existing hierarchical factor models, including the benchmark Realized Beta GARCH model (Hansen et al. (2014)) and the extended version introduced by Archakov et al. (2020) ("Multivariate Realized Beta GARCH").

We estimate each model for a cross-section of N selected assets.[5] The in-sample fit of the seven models has been assessed using the three criteria, i.e., the value of the maximized LLF, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). Given all the models assume that monthly returns are conditionally normal, the LLFs evaluated using only monthly data are directly comparable, i.e., the highest value indicates a superior in-sample fit. Conversely, the lower AIC/BIC values are better. For all models, we present the average value of each criterion with respect to the N assets.[6]

[5]Given all the competing models leave the conditional covariance matrices of idiosyncratic return components unspecified, we do not account for them in the in- and out-of-sample comparisons.

[6]The full set of results is available upon request.

Table 4.2 collects the three in-sample fit criteria for each model and comparison. In view of the results obtained, several conclusions can be drawn:

1. The "**4F-HD DCC-HEAVY**" model has a larger LLF value and correspondingly smaller AIC and BIC than the symmetric version "**sym-HD DCC-HEAVY**" with respect to both core and conditional models for individual assets. As follows, allowing for the asymmetric dynamics in the covariances of factors, as well as an individual asset vs. the set of factors, based on the signs of underlying daily/monthly returns, improves the in-sample fit of the model.

2. Among the market factor-based models, considering the total LLF values and both information criteria evaluated at the monthly data, the best fitting model is "**M-HD DCC-HEAVY**". The relative superiority of our model suggests the benefits of adopting the higher-frequency data to model conditional covariances of lower-frequency returns as opposed to **cDCC** model. Furthermore, specifying the dynamics of the RC is important as "**M-HD DCC-HEAVY**" readily outperforms **Realized Beta GARCH** model of Hansen et al. (2014). The latter provides for a better fit with respect to each criterion compared to the low-frequency data-based **cDCC** model.

3. The "**FF-HD DCC-HEAVY**" model outperforms the scalar version of the competing "**Multivariate Realized Beta GARCH**" of Archakov et al. (2020) in terms of a possible comparison of the conditional LLF for individual assets vs. factors evaluated at the monthly data, thus confirming the advantages of explicitly modelling the dynamics of realized measures.

Table 4.2: Maximized log-likelihood function (LLF), AIC, and BIC values of estimated models

	4F-HD DCC-HEAVY	**sym-HD DCC-HEAVY**	
LLF^{c}	**-17912.52**	-18469.61	
AIC	**49.040**	50.551	
BIC	**49.266**	50.752	
$\text{LLF}^{c,i}$	**-3509.52**	-3548.47	
AIC	**9.627**	9.728	
BIC	**9.715**	9.803	
$\text{LLF}^{c} + \text{LLF}^{c,i}$	**-21422.04**	-22018.08	
AIC	**58.667**	60.279	
BIC	**58.981**	60.555	
	M-HD DCC-HEAVY	**Real. Beta GARCH**	**cDCC**
$\text{LLF}^{c}_{H} + \text{LLF}^{c,i}_{H_i}$	**-3672.63**	-3991.02	-4111.86
AIC	**10.065**	10.943	11.256
BIC	**10.134**	11.031	11.307
	FF-HD DCC-HEAVY	**Mult. Real. Beta GARCH**	
$\text{LLF}^{c,i}_{H_i}$	**-1678.84**	-1962.60	
AIC	**4.606**	5.384	
BIC	**4.650**	5.434	

LLF^{c}: total LLF for the core model;
$\text{LLF}^{c,i}$: average (across N assets) total LLF for the conditional model for individual assets;
$\text{LLF}^{c}_{H} + \text{LLF}^{c,i}_{H_i}$: average (across N assets) total LLF evaluated at the monthly data;
$\text{LLF}^{c,i}_{H_i}$: average (across N assets) LLF for the conditional model for individual assets evaluated at the monthly data;
For each maximum value of the log-likelihood function (LLF), we report the corresponding Akaike (AIC) and Bayesian information criteria (BIC). The values in bold correspond to the best model of each row. The models are estimated using the dataset of 732 observations described in Section 4.5.1.

The estimates of the parameters of the core model for each asymmetric HD DCC-HEAVY version are reported in Table 4.3. The results demonstrate that the coefficients in columns III-V noticeably differ for the three models, implying distinct dynamics of the variances of factors.

In each case, the average estimate of the b_h parameter is much smaller compared to standard GARCH models, while the average estimates of the a_h^+ and a_h^- parameters are much larger compared to conventional ARCH terms. In line with the findings of Shephard and Sheppard (2010), Noureldin et al. (2012), and Bauwens and Xu (2023), these results suggest that the dynamics of conditional variances are better captured

by realized variances than by squared returns. Columns VI-VII present the parameter estimates of the correlations, implying rather responsive series.[7]

Table 4.3: Parameter estimates of the core model for HD DCC-HEAVY variants

Coeff.	w_h	a_h^+	a_h^-	b_h	α_R	β_R
Model						
4F-HD DCC-HEAVY	0.000	0.699	0.519	0.495	0.272	0.669
FF-HD DCC-HEAVY	0.000	0.418	0.543	0.518	0.229	0.729
M-HD DCC-HEAVY	0.000	0.286	0.826	0.487	-	-
Coeff.	w_m	a_m^+	a_m^-	b_m	α_P	β_P
Model						
4F-HD DCC-HEAVY	0.000	0.140	0.103	0.752	0.249	0.739
FF-HD DCC-HEAVY	0.000	0.061	0.115	0.819	0.221	0.639
M-HD DCC-HEAVY	0.000	0.000	0.135	0.860	-	-

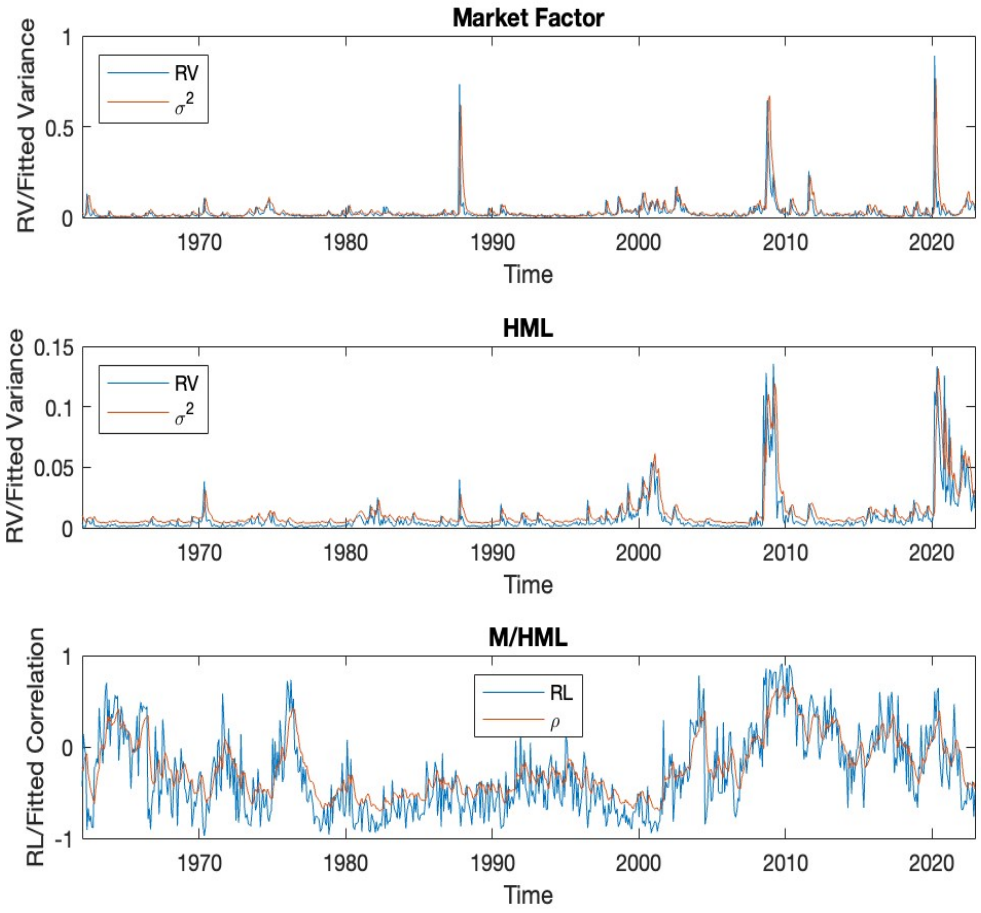
Presented are the estimates of the parameters that appear in the HD DCC-HEAVY equations of the core model for the conditional variances and correlations (upper panel) and the corresponding realized analogues (lower panel). Columns II-V provide the (average of the) estimates of univariate models for the variance of each factor. Columns VI-VII provide the estimates of the parameters of correlations. The estimation period is January 1962 - December 2022, i.e., $T = 732$.

All three FF factors exhibit a significant asymmetry effect with respect to underlying monthly and daily returns, i.e., the greater a_h^- and a_m^- parameters compared to a_h^+ and a_m^+, respectively. Thus, one of the main stylized facts of the financial time series, i.e., the stronger impact of negative returns on the volatility, seems incorporated in the dynamics of the FF portfolio returns. The same conclusion no longer holds with the addition of a MOM factor.

In Figure 4.4, we plot realized variances and correlations for the market and HML factors against fitted conditional variances and correlations via the benchmark "4F-HD DCC-HEAVY" model. Clearly, conditional variances track the corresponding realized series closely. In addition, Figure 4.4 demonstrates a significant temporal variation in the correlation dynamics of selected factors, suggesting the potential importance of accounting for time-varying factor covariances.

[7]The conditional models for individual assets (e.g., Table 4.4) would suggest very persistent correlations typically found in the literature. However, these estimates cannot be directly associated with the correlations between the individual assets and factors because we model the dynamics of the (Fisher transformed) vectors of correlations and not the correlation elements directly.

Figure 4.4: Annualized realized and fitted conditional variances of the market and HML factors and the corresponding correlation series



In our empirical analyses, we estimate conditional models for the cross-section of $N = 20$ individual assets (see Section 4.5.1). The corresponding estimation results for the "FF-HD DCC-HEAVY" model are reported in Table 4.4.

The effects of the lagged realized variances on the current conditional variances are high, on average. Thus, we further confirm realized measures as more informative about volatility than squared returns. Correspondingly, the average $a_{i,h}^{-}$ exceeds $a_{i,h}^{+}$, indicating the presence of the asymmetry effect. Ultimately, the coefficients associated with the dynamics of the realized variances of individual stocks are relatively dispersed, implying distinct dynamics of corresponding series.

Table 4.4: "FF-HD DCC-HEAVY" parameter estimates of the conditional models for individual assets

Coeff.	$c_{i,h}$	$a^+_{i,h}$	$a^-_{i,h}$	$b_{i,h}$	$\phi_{i,R}$	$\alpha_{i,R}$	$\beta_{i,R}$
mean	0.000	0.159	0.222	0.755	0.017	0.022	0.944
min	0.000	0.026	0.082	0.545	0.001	0.001	0.631
1	0.000	0.031	0.085	0.561	0.001	0.001	0.675
5	0.000	0.048	0.097	0.626	0.001	0.001	0.849
10	0.000	0.052	0.099	0.661	0.002	0.002	0.892
25	0.000	0.093	0.144	0.692	0.002	0.002	0.922
50	0.000	0.145	0.213	0.752	0.003	0.002	0.978
75	0.000	0.198	0.278	0.827	0.014	0.014	0.984
90	0.000	0.293	0.305	0.872	0.022	0.026	0.987
95	0.000	0.313	0.442	0.886	0.056	0.082	0.989
99	0.000	0.412	0.458	0.889	0.149	0.236	0.991
max	0.000	0.436	0.462	0.890	0.168	0.275	0.991
Coeff.	$c_{i,m}$	$a^+_{i,m}$	$a^-_{i,m}$	$b_{i,m}$	$\phi_{i,P}$	$\alpha_{i,P}$	$\beta_{i,P}$
mean	0.000	0.134	0.110	0.751	0.000	0.028	0.819
min	0.000	0.042	0.024	0.603	0.000	0.000	0.611
1	0.000	0.043	0.027	0.611	0.000	0.000	0.616
5	0.000	0.046	0.037	0.632	0.000	0.000	0.636
10	0.000	0.082	0.043	0.653	0.000	0.000	0.689
25	0.000	0.121	0.080	0.687	0.000	0.000	0.745
50	0.000	0.133	0.096	0.748	0.000	0.000	0.844
75	0.000	0.147	0.131	0.803	0.000	0.055	0.905
90	0.000	0.181	0.201	0.853	0.000	0.069	0.931
95	0.000	0.221	0.205	0.869	0.000	0.134	0.940
99	0.000	0.225	0.239	0.914	0.000	0.151	0.941
max	0.000	0.226	0.248	0.925	0.000	0.155	0.941

Presented are the estimates of the parameters that appear in the "FF-HD DCC-HEAVY" equations of the conditional model for an individual asset, i.e., conditional variances and conditional correlation vectors (upper panel), and the corresponding realized analogues (lower panel). Estimation period is January 1962 - December 2022, i.e., $T = 732$.

For HD DCC-HEAVY models, we implicitly assume that the correlations across the selected cross-section of asset returns are explained via either, a single, three, or four sources of the systematic risk, i.e., market, size, value, and momentum. In this regard, the vector of model-implied betas for each asset given by (4.2) is obtained by accounting for the information from higher- and lower-frequency data.

To present the rich dynamics of estimated betas, we graphically illustrate the "4F-HD DCC-HEAVY" fitted measures for IP in Figure 4.5. The average market beta is close to 1, implying the IP closely tracks the S&P500 dynamics. Conversely, the means of the value and MOM factors lie in the interval 0.5-0.8, while the average SMB beta is around -0.1. The exposure to the size risk factor varies the most. All the betas hit a range of extreme values during the financial crisis episode. The corresponding summary statistics are given in Table 4.5.

Figure 4.5: Fitted betas for IP for the period 1962–2022

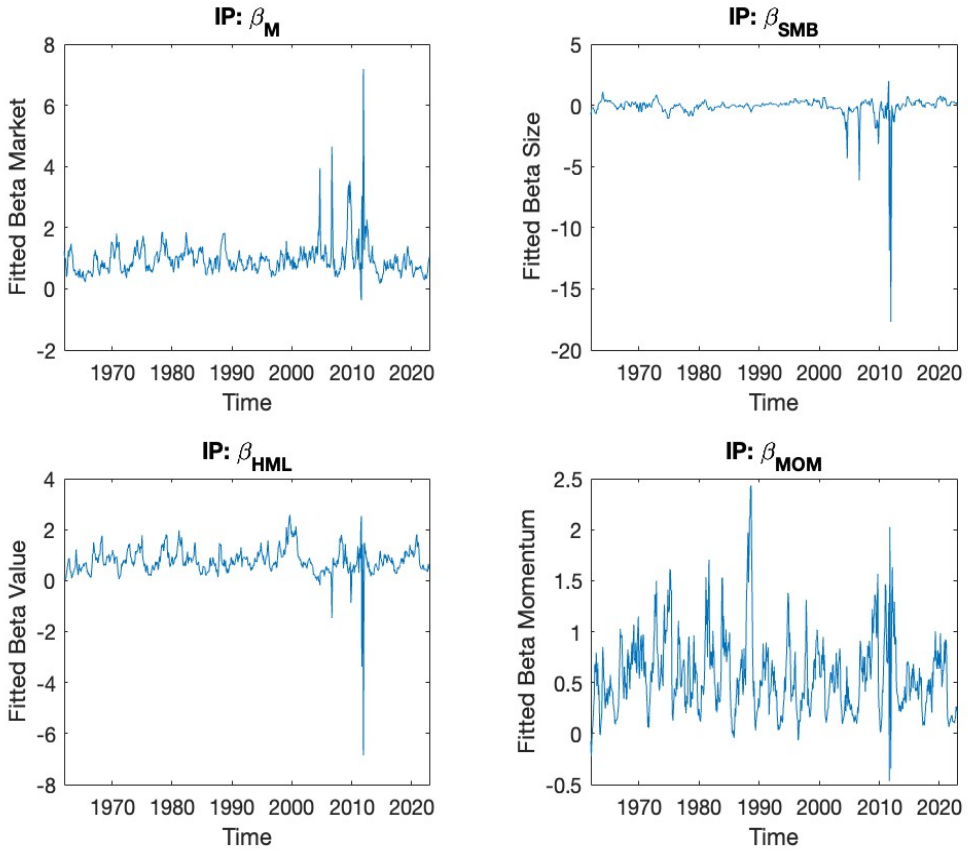


Table 4.5: "4F-HD DCC-HEAVY" beta estimates for IP

	$\hat{\beta}_M$	$\hat{\beta}_{SMB}$	$\hat{\beta}_{HML}$	$\hat{\beta}_{MOM}$	$\hat{\beta}_M$	$\hat{\beta}_{SMB}$	$\hat{\beta}_{HML}$	$\hat{\beta}_{MOM}$
mean	0.959	-0.116	0.763	0.567	1.433	-0.858	0.628	0.794
sd.	0.541	0.990	0.552	0.378	1.157	2.704	1.181	0.450
min	-0.369	-17.697	-6.852	-0.466	-0.369	-17.697	-6.852	-0.466
1	0.273	-2.549	-0.051	-0.015	-0.251	-13.556	-4.387	-0.378
5	0.414	-0.844	0.195	0.100	0.404	-2.654	-0.648	0.062
10	0.499	-0.541	0.302	0.145	0.473	-1.858	0.302	0.241
25	0.645	-0.221	0.491	0.299	0.678	-0.928	0.507	0.535
50	0.886	0.003	0.717	0.514	1.091	-0.193	0.690	0.764
75	1.122	0.203	0.983	0.763	1.805	0.203	1.066	1.113
90	1.426	0.384	1.384	1.008	3.006	0.363	1.483	1.340
95	1.705	0.514	1.621	1.296	3.364	0.450	1.671	1.446
99	3.190	0.753	2.106	1.755	5.141	1.634	2.273	1.745
max	7.185	2.006	2.582	2.433	7.185	2.006	2.536	2.026

Presented are the "4F-HD DCC-HEAVY" estimates of betas for IP for the full sample period, i.e., 1962-2022 (left panel), and the financial crisis turbulence, i.e., 2007-2012 (right panel).

4.5.3 Out-of-Sample Forecasting

We compute the out-of-sample forecasts discussed in Section 4.4 with regard to all the asymmetric hierarchical-type factor models, which fit the data better compared to the cDCC model, i.e., "4F-HD DCC-HEAVY", "FF-HD DCC-HEAVY","M-HD DCC-HEAVY", Realized Beta GARCH, and "Multivariate Realized Beta GARCH". Starting from the fitting period from January 1962 to December 2016 ($T_e = 660$), we generate the forecasts by re-estimating the models every year on a rolling window with T_e monthly observations and then producing a sequence of the twelve 1-step-ahead predictions based on the updated parameter estimates. We consider the two out-of-sample forecasting periods.[8] The first, characterized by the relatively low volatility of returns, includes the years 2017-2019. The second period lasts until the end of 2022, with volatility at a relatively high level, triggered by the COVID pandemic.

[8]The results for a full out-of-sample period are available in Appendix B.2.

Statistical Accuracy

In order to assess the statistical accuracy of all models, we adopt the two loss functions that produce the consistent ranking (Patton (2011); Laurent et al. (2013)), i.e., the Euclidean distance (ED) and squared Frobenius norm (FN).

The first is based on the vech$(\cdot)$[9] transformation of the forecast error matrix, where the prediction errors on variances and covariances are equally weighted:

$$ED_t(C_{t+1}, \hat{H}_{t+1}) = \text{vech}(C_{t+1} - \hat{H}_{t+1})' \, \mathrm{I}_{N*} \, \text{vech}(C_{t+1} - \hat{H}_{t+1}), \tag{4.34}$$

where $\hat{H}_{t+1}$ is the conditional forecast of the covariances of r_{t+1}, C_{t+1} is a proxy for the unobserved covariance matrix at time $t+1$, and I_{N*} is the identity matrix of order $N(N+1)/2$. Indeed, the natural proxy for latent covariances is given by $r_t r_t'$, although others, such as the RC, can be used.[10]

The second loss function is the matrix equivalent of the MSE loss function, where the weights on the covariance forecast errors are doubled compared to the ones on variances:

$$FN_t(C_{t+1}, \hat{H}_{t+1}) = \text{trace}[(C_{t+1} - \hat{H}_{t+1})'(C_{t+1} - \hat{H}_{t+1})] = \sum_{i,j} (c_{ij,t+1} - \hat{h}_{ij,t+1})^2. \tag{4.35}$$

For assessing the significance of differences in the ED and FN losses across the five models, we rely on the model confidence set (MCS) approach of Hansen et al. (2011). The MCS identifies the model or subset of models with the best forecasting performance, given the pre-specified confidence level. It is computed at the 10% significance level using a block bootstrap (Hansen et al. (2003)) with 10,000 replications and the varying block length to verify the robustness of the results.

[9] The operator that stacks the lower triangular part of a symmetric $N \times N$ matrix argument into a $N(N+1)/2 \times 1$ vector.

[10] The adoption of $r_t r_t'$ appears more suitable when forecasting the covariances over the entire month.

Table 4.6: Model confidence sets at 90% level of hierarchical factor models, with ED and FN loss functions

Model	ED	MCS 2017-2019	ED	MCS 2020-2022
4F-HD DCC-HEAVY	0.080	0.006	0.828	0.002
FF-HD DCC-HEAVY	**0.075**	**1.000**	**0.796**	**1.000**
M-HD DCC-HEAVY	0.075	**0.885**	0.879	0.002
Realized Beta GARCH	0.078	0.091	0.951	0.002
Multivariate Realized Beta GARCH	0.093	0.000	0.954	0.002
Model	FN	MCS 2017-2019	FN	MCS 2020-2022
4F-HD DCC-HEAVY	0.138	0.002	1.340	0.002
FF-HD DCC-HEAVY	0.132	**0.320**	**1.276**	**1.000**
M-HD DCC-HEAVY	**0.129**	**1.000**	1.399	0.002
Realized Beta GARCH	0.135	0.008	1.514	0.002
Multivariate Realized Beta GARCH	0.164	0.000	1.527	0.002

'ED/FN' columns: the average annualized value of ED/FN losses over the corresponding forecast period; bold values identify the minimum loss over the five models.
'MCS 2017-2019' columns: p-values of the MCS tests over the out-of-sample period including the years 2017-2019; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).
'MCS 2020-2022' columns: the analogous results for the period 2020-2022.

Table 4.6 reports the model confidence sets at the 90% confidence level using the ED and FN loss functions. The hierarchical models of Hansen et al. (2014) and Archakov et al. (2020) are always excluded from the reported model confidence sets. The "FF-HD DCC-HEAVY" significantly outperforms all the other models during financial turbulence, while during calm times the MCS also incorporates the "M-HD DCC-HEAVY".

As follows, when all the hierarchical factor models are compared in statistical terms, the new HD DCC-HEAVY models are superior compared to the models built upon the Realized GARCH framework in all cases. Considering the full out-of-sample period, only the "FF-HD DCC-HEAVY" model enters the MCS in terms of both ED and FN losses (Appendix B.2, Table B.2.1).

Economic Performance

In order to perform the economic evaluation of the forecasting performance, we rely on the global minimum variance portfolio (GMVP) optimization (e.g., Engle and Kelly (2012); Bauwens and Xu (2023)) since it does not require the estimation of expected returns, providing an essentially clean framework for assessing the merits of distinct covariance forecasting models.

Given a covariance matrix forecast $\hat{H}_{t+1}$, the portfolio weights $\hat{\omega}_{t+1}$ are obtained by solving the minimization problem:

$$\min_{\omega_{t+1}} \omega'_{t+1} \hat{H}_{t+1} \omega_{t+1} \quad \text{s. t.} \quad \omega'_{t+1} \mathbf{1} = 1, \tag{4.36}$$

where $\mathbf{1}$ is a $N \times 1$ vector of ones.

It follows readily that the optimal GMVP weights are given by:

$$\hat{\omega}_{t+1} = \frac{\hat{H}_{t+1}^{-1} \mathbf{1}}{\mathbf{1}' \hat{H}_{t+1}^{-1} \mathbf{1}}. \tag{4.37}$$

In addition, we consider the optimization under a short-selling restriction and compute the weights via numerical optimization, i.e., MATLAB Financial Toolbox, given the absence of a closed-form analytical solution. The results are available in Appendix B.2 (Table B.2.3). Given the main aim of assessing the accuracy of distinct covariance matrix estimators, our performance measures do not account for transaction costs.

Initially, we adopt the MCS to select the best-performing models that minimize the standard deviation (SD) of the portfolios obtained by applying the computed weights to the observed returns.

The results presented in Table 4.7 show that the "M-HD DCC-HEAVY" model provides for the lowest out-of-sample SD during the calm periods, whereas only the "4F-HD DCC-HEAVY" enters the MCS when the volatility is at a relatively high level. Considering the entire out-of-sample period, the MCS includes only the "4F-HD DCC-HEAVY" model (Appendix B.2, Table B.2.2), while the analogous conclusion applies for long-only portfolios (Appendix B.2, Table B.2.3). Therefore, in contrast to the statistical performance where the "FF-HD DCC-HEAVY" model is superior, the "4F-HD DCC-HEAVY" appears preferable from a variance minimization perspective.

In general, the "M-HD DCC-HEAVY" model outperforms the competing market factor-based Realized Beta GARCH of Hansen et al. (2014) in all cases. The same applies for a corresponding comparison between the three-factor "FF-HD DCC-HEAVY" and "Multivariate Realized Beta GARCH" (Archakov et al. (2020)) models (see Tables 4.7, B.2.2, and B.2.3).

Table 4.7: Model confidence sets at 90% level of hierarchical factor models, with GMVP loss function

Model	SD	MCS 2017-2019	SD	MCS 2020-2022
4F-HD DCC-HEAVY	0.653	0.098	**0.571**	**1.000**
FF-HD DCC-HEAVY	0.715	0.000	0.709	0.088
M-HD DCC-HEAVY	**0.565**	**1.000**	0.952	0.000
Realized Beta GARCH	0.643	0.098	0.878	0.000
Multivariate Realized Beta GARCH	0.723	0.000	0.718	0.067

'SD' columns: the average annualized standard deviation of GMVP returns over the corresponding forecast period; bold values identify the minimum loss over the five models.
'MCS 2017-2019' column: p-values of the MCS tests over the out-of-sample period including the years 2017-2019; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).
'MCS 2020-2022' column: the analogous results for the period 2020-2022.

In addition, we examine some basic features of the portfolios, including the Average Return (AR), i.e., the average of out-of-sample returns for the corresponding period; the Information Ratio (IR), i.e., the ratio AR/SD; portfolio turnover rates

(TO); and the proportion of short positions (SP).[11]

The latter are specified as follows:

$$TO_t = \sum_{i}^{N} \left| \hat{w}_{i,t} - \hat{w}_{i,t-1} \frac{1 + r_{t-1}^{i}}{1 + r_{t-1}^{p}} \right|; \tag{4.38}$$

$$SP_t = \sum_{i}^{N} \mathbb{1}_{\{\hat{w}_{i,t}<0\}}, \tag{4.39}$$

where r_t^p is the total return of the portfolio for the month t, $\hat{w}_{i,t}$ and r_t^i are the weight and return of stock i, respectively, and $\mathbb{1}_{\{.\}}$ denotes the indicator function.[12]

The results reported in Table B.2.4 confirm that hierarchical HD DCC-HEAVY models consistently and notably outperform Realized GARCH variants. In particular, the "M-HD DCC-HEAVY" features the highest IR during turbulent periods and overall. On the other hand, the findings summarized in Table B.2.5 suggest that the propensity of models with respect to short positions is very similar and, in general, moderately increases for HD DCC-HEAVY models during turmoils. The increasing trend of the average monthly turnover rates for all models is also visible.

Given that the GMVPs aim at minimizing the variance, and thus the SD, rather than maximizing the expected returns or the IR, the most important performance measure is the out-of-sample SD. In this regard, the out-of-sample returns and IR are also beneficial but should be considered of secondary importance.

Finally, to assess the economic gains of utilizing distinct HD DCC-HEAVY covariance matrix estimators, following Fleming et al. (2001, 2003), we determine the maximum performance fee a risk-averse investor would be willing to pay to switch from using one model to another.

[11] The resulting AR and IR are computed with respect to estimated non-negative weights since short-selling is difficult to implement, thus it is not generally the common practice for most investors.

[12] We do not set constraints on the turnover rate and leverage proportion in the optimization.

Accordingly, we assume that the investor has quadratic preferences of the form:

$$U(r_t^p) = 1 + r_t^p - \frac{\gamma}{2(1+\gamma)}(1 + r_t^p)^2, \tag{4.40}$$

where r_t^p is the portfolio return and γ is the investor's relative risk aversion, taking values 1 and 10 (Fleming et al. (2003)). As follows, we determine a fee Δ_γ by equating the average realized utilities from two alternative portfolios:

$$\sum_{t=1}^{T} U(r_t^{p_1}) = \sum_{t=1}^{T} U(r_t^{p_2} - \Delta_\gamma), \tag{4.41}$$

where $r_t^{p_1}$ and $r_t^{p_2}$ are the portfolio returns related to competing HD DCC-HEAVY forecasting strategies.

Major observations based on results in Table 4.8 are as follows. First, by utilizing the "4F-HD DCC-HEAVY" covariance forecasts, a risk-averse investor can achieve notable economic gains that become pronounced during the crisis period. Overall, an investor with low (high) risk aversion would be willing to pay on average 27 (38) bps to switch from the "FF-HD DCC-HEAVY" strategy to the "4F-HD DCC-HEAVY" and 15 (35) bps to switch from the "M-HD DCC-HEAVY". These results provide further support that the "4F-HD DCC-HEAVY" might be a preferable hierarchical factor model from the investor's point of view.

Table 4.8: BPS fees for switching from simpler HD DCC-HEAVY to the "4F-HD DCC-HEAVY" covariance matrix forecasts

Period	2017-2019		2020-2022		2017-2022	
Model	Δ_1	Δ_{10}	Δ_1	Δ_{10}	Δ_1	Δ_{10}
FF-HD DCC-HEAVY	-6.92	-6.92	61.52	83.84	27.30	38.46
M-HD DCC-HEAVY	28.74	28.74	0.46	40.64	14.60	34.69

'Δ_γ' columns: the basis points fee an investor with quadratic utility and relative risk aversion γ would pay to switch from the covariance matrix estimator indicated in column 1 to the "4F-HD DCC-HEAVY" model over the period indicated in row 1.

4.6 Conclusion

In this paper, we introduce a class of models for high-dimensional covariance matrices by combining the hierarchical approach of Hansen et al. (2014) and dynamic conditional correlation formulation of a HEAVY model (Noureldin et al. (2012)) recently proposed by Bauwens and Xu (2023). In this regard, we rely on the evidence to adopt the higher-frequency data to model more accurate realized measures of covariances and employ them to forecast the conditional covariance matrix of lower-frequency returns (i.e., Noureldin et al. (2012); Gorgi et al. (2019); Bauwens and Xu (2023)).

An illustrative empirical study for the S&P500 constituents over the period from January 1962 until January 2023, i.e., $N = 20$ and $T = 732$, shows that our method always significantly outperforms the benchmark and existing hierarchical factor models in statistical and economic terms. The findings are robustified under distinct market conditions.

The avenues for future research are twofold. First, a promising feature of the framework is the ability to readily extract inherently time-varying factor loadings for a given asset or portfolio, thus conforming to the extensive literature that proves the dynamic nature of betas (e.g., Bollerslev et al. (1988); Jagannathan and Wang (1996); etc.) but also potentially improving the commonly adopted rolling regression approach for their estimation. Second, to verify the relevance of adopted factors and thus employ the optimal HD DCC-HEAVY model, the asymptotic theory on estimated loadings and corresponding testing procedures should be derived.

B Appendix of Chapter 4

B.1 Data cont'd

Table B.1.1: Summary statistics for the individual assets

	r_{cc}^2	RV	P	N	GJR_P	GJR_N	RL
AEP	3.68 (6.54)	4.34 (9.99)	2.21 (4.69)	2.13 (5.55)	1.97 (3.47)	2.37 (9.85)	0.30 (0.21)
BA	11.32 (26.15)	11.54 (19.78)	6.04 (8.81)	5.50 (12.40)	5.69 (9.58)	5.85 (19.13)	0.34 (0.22)
CAT	8.53 (17.46)	8.76 (10.13)	4.43 (4.75)	4.33 (6.71)	4.61 (7.31)	4.15 (9.35)	0.35 (0.23)
CVX	5.27 (10.15)	6.63 (12.97)	3.38 (5.72)	3.25 (8.01)	3.18 (5.45)	3.45 (12.67)	0.34 (0.23)
DTE	3.26 (8.77)	3.96 (7.89)	2.04 (3.97)	1.92 (4.55)	1.97 (4.17)	1.99 (7.26)	0.31 (0.21)
ED	4.40 (26.05)	4.03 (11.62)	1.93 (3.92)	2.10 (9.68)	1.92 (4.17)	2.11 (11.22)	0.30 (0.22)
GD	9.45 (16.93)	9.23 (9.68)	4.94 (5.89)	4.29 (5.42)	4.99 (8.33)	4.24 (8.16)	0.34 (0.22)
GE	6.62 (13.82)	7.52 (12.06)	3.83 (5.71)	3.69 (7.23)	3.35 (6.91)	4.17 (11.21)	0.33 (0.24)
HON	7.53 (18.62)	8.55 (14.84)	4.36 (6.91)	4.19 (9.34)	4.15 (7.29)	4.40 (14.27)	0.36 (0.23)
IBM	5.77 (11.11)	6.39 (8.69)	3.24 (4.12)	3.15 (5.86)	2.92 (4.98)	3.47 (8.43)	0.30 (0.24)
IP	8.44 (22.52)	9.31 (16.22)	4.71 (8.55)	4.60 (9.46)	4.58 (11.59)	4.73 (13.12)	0.35 (0.23)
KO	4.33 (9.19)	5.37 (9.22)	2.77 (4.01)	2.60 (5.74)	2.66 (4.05)	2.71 (9.11)	0.28 (0.23)
KR	7.33 (14.14)	8.42 (10.08)	4.26 (5.36)	4.16 (6.96)	4.34 (7.21)	4.08 (9.22)	0.31 (0.20)
MMM	4.51 (8.47)	5.38 (6.69)	2.72 (3.13)	2.66 (4.52)	2.53 (6.11)	2.85 (4.68)	0.32 (0.24)
MO	6.29 (11.74)	6.63 (8.01)	3.35 (3.95)	3.28 (5.59)	3.43 (5.08)	3.20 (7.76)	0.29 (0.23)
MRK	5.59 (9.53)	6.26 (7.79)	3.17 (3.58)	3.09 (5.79)	3.21 (4.97)	3.05 (7.45)	0.28 (0.22)
MRO	12.67 (46.27)	12.92 (32.01)	6.27 (10.28)	6.65 (24.93)	5.89 (11.51)	7.03 (31.22)	0.35 (0.23)
MSI	11.54 (20.10)	13.72 (17.49)	6.87 (8.08)	6.85 (11.45)	6.56 (10.94)	7.16 (16.74)	0.32 (0.23)
PG	3.73 (10.02)	4.70 (11.17)	2.32 (3.71)	2.38 (8.44)	2.14 (3.52)	2.56 (11.11)	0.28 (0.22)
XOM	3.63 (7.75)	5.28 (10.27)	2.71 (4.72)	2.57 (6.06)	2.51 (4.27)	2.77 (10.05)	0.33 (0.24)

r_{cc}^2: squared close-to-close monthly return; RV: realized variance; P: positive semi-variance; N: negative semi-variance; GJR_P: RV if monthly return is positive, 0 if negative; GJR_N: RV if monthly return is negative, 0 if positive; RL: realized correlation, the average of the 4 time series means and sd-s of realized correlations with the set of factors.

B.2 Out-of-Sample Performance cont'd

Table B.2.1: Model confidence sets at 90% level of hierarchical factor models, with ED and FN loss functions over the full out-of-sample period

Model	ED	MCS 2017-2022	FN	MCS 2017-2022
4F-HD DCC-HEAVY	0.454	0.005	0.739	0.003
FF-HD DCC-HEAVY	**0.435**	**1.000**	**0.704**	**1.000**
M-HD DCC-HEAVY	0.477	0.005	0.764	0.003
Realized Beta GARCH	0.515	0.005	0.824	0.003
Multivariate Realized Beta GARCH	0.524	0.005	0.845	0.003

'ED/FN' column: the average annualized value of ED/FN losses over the full forecast period; bold values identify the minimum loss over the five models.
'MCS 2017-2022' columns: p-values of the MCS tests over the out-of-sample period including the years 2017-2022; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).

Table B.2.2: Model confidence set at 90% level of hierarchical factor models, with GMVP loss function over the full out-of-sample period

Model	SD	MCS 2017-2022
4F-HD DCC-HEAVY	**0.612**	**1.000**
FF-HD DCC-HEAVY	0.712	0.027
M-HD DCC-HEAVY	0.758	0.005
Realized Beta GARCH	0.761	0.005
Multivariate Realized Beta GARCH	0.720	0.005

'SD' column: the average annualized standard deviation of GMVP returns over the full forecast period; bold values identify the minimum loss over the five models.
'MCS 2017-2022' column: p-values of the MCS tests over the out-of-sample period including the years 2017-2022; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).

Table B.2.3: Model confidence set at 90% level of hierarchical factor models, with GMVP loss function under long-only portfolios

Period	2017-2019		2020-2022		2017-2022	
Model	SD	MCS	SD	MCS	SD	MCS
4F-HD DCC-HEAVY	0.105	0.079	**0.137**	**1.000**	**0.121**	**1.000**
FF-HD DCC-HEAVY	0.115	0.000	0.144	**0.659**	0.130	0.000
M-HD DCC-HEAVY	0.112	0.079	0.153	**0.659**	0.132	0.000
Realized Beta GARCH	0.157	0.000	0.242	0.000	0.199	0.000
Multivariate Realized Beta GARCH	**0.080**	**1.000**	0.199	0.091	0.140	0.000

'SD' columns: the average annualized standard deviation of GMVP returns with short sale restrictions over the forecast period indicated in row 1; bold values identify the minimum loss over the five models.
'MCS' columns: p-values of the MCS tests over the out-of-sample period indicated in row 1; bold values identify the models included in the MCS at the 90% confidence level (i.e., p-values larger than 0.10).

Table B.2.4: Alternative economic performance measures for hierarchical factor models based on GMVP optimization

Period	2017-2019		2020-2022		2017-2022	
Model	AR	IR	AR	IR	AR	IR
4F-HD DCC-HEAVY	0.099	0.937	0.145	1.066	0.122	1.010
FF-HD DCC-HEAVY	0.121	1.051	0.086	0.596	0.104	0.798
M-HD DCC-HEAVY	0.105	0.935	0.221	1.449	0.163	1.231
Realized Beta GARCH	0.027	0.173	0.005	0.020	0.016	0.080
Multivariate Realized Beta GARCH	0.005	0.059	0.116	0.579	0.060	0.431

'AR' columns: the average annualized GMVP return over the forecast period indicated in row 1.
'IR' columns: the average annualized AR/SD ratio over the forecast period indicated in row 1.

Table B.2.5: Alternative economic performance measures for hierarchical factor models based on GMVP optimization cont'd

Period	2017-2019		2020-2022		2017-2022	
Model	TO	SP	TO	SP	TO	SP
4F-HD DCC-HEAVY	1.227	0.471	1.441	0.478	1.334	0.474
FF-HD DCC-HEAVY	0.828	0.463	1.089	0.476	0.959	0.469
M-HD DCC-HEAVY	1.023	0.469	1.258	0.478	1.141	0.474
Realized Beta GARCH	1.085	0.476	1.139	0.468	1.112	0.472
Multivariate Realized Beta GARCH	0.671	0.479	1.247	0.475	0.960	0.477

'TO' columns: the average portfolio turnover over the forecast period indicated in row 1.
'SP' columns: the average leverage proportion over the forecast period indicated in row 1.

Chapter 5

Conclusions

Considering the importance of identifying the advantages and potential pitfalls in the performance of distinct estimation techniques adopted for covariance models, in Chapter 2, we empirically investigate the asymptotic properties of the NLS for the HE–HAR family of models (Bauwens and Otranto (2023)). The latter allows for the possibility of having a very flexible but parsimonious model, thus adequately offsetting the well-known 'curse of dimensionality' issue while maintaining the computational advantages of the HAR family, i.e., simplicity, stability, and cost with respect to all ML-based RC models. Not to mention, the modelling framework automatically guarantees the PD-ness of the covariances under the proper assumption for error terms, e.g., multivariate Normal, Wishart, or Matrix F distribution.

To the best of our knowledge, this is the first study that not only empirically verifies consistent and asymptotically Normal OLS estimates for benchmark HAR models but also investigates the properties of the NLS for flexible nonlinear HAR extensions. Our results show that conventional HAR parameters are consistent (and Normal under Gaussian errors), while the HE coefficient features a small-medium bias conditional upon the underlying distributional assumption. In this regard, we argue for confidently adopting linear HAR models estimated by OLS to obtain RC forecasts.

In addition, the NLS for the HE–HAR class might be employed with caution.

Regardless of the assumption on the assessment of the latent covariances, all models aim to contain the main stylized facts of the financial time series, such as persistence, clustering, etc. On the other hand, while the 'leverage effect' specific to stock returns has been incorporated in the dynamics of realized variances (RV) by, e.g., Corsi and Renò (2012), McAleer and Medeiros (2008), Patton and Sheppard (2015), etc., the corresponding extension to the multivariate setting is rather rare.

In this regard, via Chapter 3, we contribute to the literature by developing a class of asymmetric models for RC matrices based on the CAW model of Golosnoy et al. (2012) that guarantees the PD-ness of covariances by construction and empirically evaluating the statistical and economic performance of these models.

The asymmetric dynamics in RC are captured by utilizing the signs of either lagged daily or intra-daily returns. Both the in-sample and forecasting results demonstrate that the asymmetric models always significantly outperform the symmetric benchmark specification. In addition, we show that the models that rely on the signs of underlying daily returns to capture asymmetry not only better fit the data but also provide for more statistically accurate forecasts and notably lower out-of-sample portfolio volatility than the models built upon the semi-covariance decomposition (Bollerslev et al. (2020a)). As such, we underline the importance of accounting for the asymmetries in modelling and forecasting RC matrices. To do so, we recommend using the RC decomposition via daily close-to-close rather than HF returns.

In Chapter 4, we propose a class of models for high-dimensional covariance matrices by utilizing the hierarchical modelling method of Hansen et al. (2014) and dynamic conditional correlation (DCC)-HEAVY model of Bauwens and Xu (2023). An assumption of a factor structure for cross-correlation dynamics, together with the hierarchical approach, makes the estimation and forecasting independent of the cross-sectional dimension of the assets under consideration. In particular, the estimation

procedure relies on two main steps. In the first step, we estimate the marginal model for the set of factors; in the second step, we estimate the models for individual asset returns conditionally on the variables filtered from the core part.

Empirical results of the out-of-sample exercises imply that our High-Dimensional (HD) DCC-HEAVY class of models brings statistically more accurate forecasts with significant economic gains with respect to the existing hierarchical-type models (Hansen et al. (2014), Archakov et al. (2020)) and the benchmark cDCC model (Aielli (2013)). The robustness of the findings is verified under distinct market conditions.

The topics analyzed in this thesis are clearly relevant outside of academia. I.e., the presented methods and conclusions might be of interest to, e.g., asset managers, banks, pension funds, etc. All these institutions face precisely the same problem of selecting the estimator and forecasting methodology for covariance matrices of asset returns in order to decide on their operations, which have a substantial impact on their clients' investments, retirement savings, etc.

Bibliography

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Aït-Sahalia, Y., Kalnina, I., and Xiu, D. (2020). High-frequency factor models and regressions. *Journal of Econometrics*, 216(1):86–105.

Anatolyev, S. and Kobotaev, N. (2018). Modeling and forecasting realized covariance matrices with accounting for leverage. *Econometric Reviews*, 37(2):114–139.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andreou, E. (2016). On the use of high frequency measures of volatility in MIDAS regressions. *Journal of Econometrics*, 193(2):367–389.

Andreou, E., Ghysels, E., and Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, 158(2):246–261.

Archakov, I., Hansen, P. R., and Lunde, A. (2020). A Multivariate Realized GARCH Model [Unpublished manuscript]. Cornell University.

Asai, M., McAleer, M., and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews*, 25(2-3):145–175.

Audrino, F. and Trojani, F. (2011). A general multivariate threshold GARCH model with dynamic conditional correlations. *Journal of Business & Economic Statistics*, 29(1):138–149.

Bali, T. G., Engle, R. F., and Tang, Y. (2017). Dynamic conditional beta is alive and well in the cross section of daily stock returns. *Management Science*, 63(11):3760–3779.

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

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.

Barndorff-Nielsen, O. E., Kinnebrock, S., and Shephard, N. (2010). Measuring downside risk: Realised semivariance. In "Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle"(Edited by T. Bollerslev, J. Russell and M. Watson).

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

Bauer, G. H. and Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160(1):93–101.

Bauwens, L., Laurent, S., and Rombouts, J. V. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21(1):79–109.

Bauwens, L. and Otranto, E. (2020). Nonlinearities and regimes in conditional correlations with different dynamics. *Journal of Econometrics*, 217(2):496–522.

Bauwens, L. and Otranto, E. (2023). Modeling Realized Covariance Matrices: A Class of Hadamard Exponential Models. *Journal of Financial Econometrics*, 21(4):1376–1401.

Bauwens, L., Storti, G., and Violante, F. (2012). Dynamic conditional correlation models for realized covariance matrices. CORE Discussion Paper 2012/60, Université catholique de Louvain, Louvain La Neuve.

Bauwens, L. and Xu, Y. (2023). DCC- and DECO-HEAVY: Multivariate GARCH models based on realized variances and correlations. *International Journal of Forecasting*, 39(2):938–955.

Black, F. (1976). Studies of stock market volatility changes. In *Proceedings of 1976 Meeting of the American Statistical Association, Business and Economic Statistics Section*.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96(1):116–131.

Bollerslev, T., Li, J., Patton, A. J., and Quaedvlieg, R. (2020a). Realized semicovariances. *Econometrica*, 88(4):1515–1551.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2018). Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. *Journal of Econometrics*, 207(1):71–91.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2020b). Multivariate leverage effects and realized semicovariance GARCH models. *Journal of Econometrics*, 217(2):411–430.

Bonato, M., Caporin, M., and Ranaldo, A. (2012). A forecast-based comparison of restricted Wishart autoregressive models for realized covariance matrices. *The European Journal of Finance*, 18(9):761–774.

Bose, A. and Mukherjee, K. (2003). Estimating the ARCH parameters by solving linear equations. *Journal of Time Series Analysis*, 24(2):127–136.

Buccheri, G. and Corsi, F. (2021). HARK the SHARK: Realized volatility modeling with measurement errors and nonlinear dependencies. *Journal of Financial Econometrics*, 19(4):614–649.

Campbell, J. Y. and Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31(3):281–318.

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

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Chang, Y. and Park, J. Y. (2003). Index models with integrated time series. *Journal of Econometrics*, 114(1):73–106.

Chen, Z., Liu, W., Wang, C. D., Wu, W.-q., and Wu, Y.-h. (2018). Nonlinear Least Squares Estimation of Log-ACD Models. *Acta Mathematicae Applicatae Sinica*, English Series, 34(3):516–533.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26(6):922–947.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, 10(4):407–432.

Ciuperca, G. (2011). Asymptotic behaviour of the LS estimator in a nonlinear model with long memory. *Journal of the Korean Statistical Society*, 40(2):193–203.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Corsi, F. and Renò, R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, 30(3):368–380.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563.

Cubadda, G., Guardabascio, B., and Hecq, A. (2017). A vector heterogeneous autoregressive index model for realized volatility measures. *International Journal of Forecasting*, 33(2):337–344.

De Goeij, P. C. and Marquering, W. (2004). Modeling the conditional covariance between stock and bond returns: A multivariate GARCH approach. *Journal of Financial Econometrics*, 2(4):531–564.

De Nard, G., Ledoit, O., and Wolf, M. (2021). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*, 19(2):236–257.

Dias, G. F. (2013). The Nonlinear Iterative Least Squares (NL-ILS) Estimator: An Application to Volatility Models. DP, School of Economics and Finance, Queen Mary University.

Engle, R. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business & Economic Statistics*, 20:339–350.

Engle, R. and Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2):238–253.

Engle, R. F. and Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11(1):122–150.

Engle, R. F., Ledoit, O., and Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, 37(2):363–375.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1):329–352.

Fleming, J., Kirby, C., and Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics*, 67(3):473–509.

Francq, C. and Zakoïan, J.-M. (2012). QML estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory*, 28(1):179–206.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95.

Glosten, L. R., Jagannathan, R., and 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.

Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167(1):211–223.

Gorgi, P., Hansen, P. R., Janus, P., and Koopman, S. J. (2019). Realized Wishart-GARCH: A score-driven multi-asset volatility model. *Journal of Financial Econometrics*, 17(1):1–32.

Gouriéroux, C., Jasiak, J., and Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150(2):167–181.

Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

Greene, W. H. (2018). *Econometric Analysis*. Pearson Education Limited, London.

Hamadeh, T. and Zakoïan, J.-M. (2011). Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH processes. *Journal of Statistical Planning and Inference*, 141(1):488–507.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R. and Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, 131(1-2):97–121.

Hansen, P. R., Lunde, A., and Nason, J. M. (2003). Choosing the best volatility models: The Model Confidence Set approach. *Oxford Bulletin of Economics and Statistics*, 65(s1):839–861.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79(2):453–497.

Hansen, P. R., Lunde, A., and Voev, V. (2014). Realized beta GARCH: A multivariate GARCH model with realized measures of volatility. *Journal of Applied Econometrics*, 29(5):774–799.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: With applications to financial and economic time series*. Cambridge University Press, Cambridge.

He, C. and Teräsvirta, T. (2002). An application of the analogy between vector ARCH and vector random coefficient autoregressive models. SSE/EFI Working Paper Series in Economics and Finance, No. 516.

Hosking, J. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.

Hwang, E. and Shin, D. W. (2014). Infinite-order, long-memory heterogeneous autoregressive models. *Computational Statistics & Data Analysis*, 76:339–358.

Hwang, S. and Kim, T. Y. (2004). Power transformation and threshold modeling for ARCH innovations with applications to tests for ARCH structure. *Stochastic Processes and their Applications*, 110(2):295–314.

Jagannathan, R. and Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *The Journal of Finance*, 51(1):3–53.

James, W. and Stein, C. (1976). Estimation with quadratic loss. In *Volume 1 of Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability: Contributions to the Theory of Statistics*, pages 361–379.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Konno, Y. (1991). A note on estimating eigenvalues of scale matrix of the multivariate F-distribution. *Annals of the Institute of Statistical Mathematics*, 43(1):157–165.

Kroner, K. F. and Ng, V. K. (1998). Modelling asymmetric comovements of asset returns. *The Review of Financial Studies*, 11(4):817–844.

Lai, T. L. (1994). Asymptotic properties of nonlinear least squares estimates in stochastic regression models. *Annals of Statistics*, 22(4):1917–1930.

Lanne, M. and Saikkonen, P. (2005). Non-linear GARCH models for highly persistent volatility. *Econometrics Journal*, 8(2):251–276.

Laurent, S., Rombouts, J. V., and Violante, F. (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics*, 173(1):1–10.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *The Review of Financial Studies*, 30(12):4349–4388.

Lütkepohl, H. (1996). *Handbook of Matrices*. Wiley, Chichester.

McAleer, M. and Medeiros, M. C. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics*, 147(1):104–119.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B., and Ward, J. R. (1993). Fractals and intrinsic time: A challenge to econometricians [Unpublished manuscript]. Olsen & Associates, Zürich.

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27(6):907–933.

Oh, D. and Patton, A. (2016). High-dimensional copula-based distributions with mixed frequency data. *Journal of Econometrics*, 193:349–366.

Opschoor, A., Janus, P., Lucas, A., and Van Dijk, D. (2018). New HEAVY models for fat-tailed realized covariances and returns. *Journal of Business & Economic Statistics*, 36(4):643–657.

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

Patton, A. J. and Sheppard, K. (2015). Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. *Review of Economics and Statistics*, 97(3):683–697.

Pollard, D. and Radchenko, P. (2006). Nonlinear least-squares estimation. *Journal of Multivariate Analysis*, 97(2):548–562.

Qu, H., Chen, W., Niu, M., and Li, X. (2016). Forecasting realized volatility in electricity markets using logistic smooth transition heterogeneous autoregressive models. *Energy Economics*, 54:68–76.

Qu, H. and Zhang, Y. (2022). Asymmetric multivariate HAR models for realized covariance matrix: A study based on volatility timing strategies. *Economic Modelling*, 106:105699.

Rohan, N. and Ramanathan, T. (2013). Nonparametric estimation of a time-varying GARCH model. *Journal of Nonparametric Statistics*, 25(1):33–52.

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

Taylor, N. (2015). Realized volatility forecasting in an international context. *Applied Economics Letters*, 22(6):503–509.

Vassallo, D., Buccheri, G., and Corsi, F. (2021). A DCC-type approach for realized covariance modeling with score-driven dynamics. *International Journal of Forecasting*, 37(2):569–586.

Wang, Q. (2021). Least squares estimation for nonlinear regression models with heteroscedasticity. *Econometric Theory*, 37(6):1267–1289.

Wooldridge, J. M. (1986). Estimation and inference for dependent processes. In *Handbook of Econometrics*. Elsevier.

Wu, C.-F. (1981). Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics*, 9(3):501–513.

Wu, G. (2001). The determinants of asymmetric volatility. *The Review of Financial Studies*, 14(3):837–859.

Zhou, J., Jiang, F., Zhu, K., and Li, W. K. (2022). Time series models for realized covariance matrices based on the matrix-F distribution. *Statistica Sinica*, 32:755–786.