This paper introduces a novel multi-variate volatility model that can accommodate appropriately defined network structures based on low-frequency and high-frequency data. The model offers substantial reductions in the number of unknown parameters and computational complexity. The model formulation, along with iterative multi-step-ahead forecasting and targeting parameterization are discussed. Quasi-likelihood functions for parameter estimation are proposed and their asymptotic properties are established. A series of simulation studies are carried out to assess the performance of parameter estimation in finite samples. Furthermore, a real data analysis demonstrates that the proposed model outperforms the existing volatility models in prediction of future variances of daily return and realized measures.

Received 7 April 2022; Accepted 19 October 2023

Keywords: High-frequency data; low-frequency data; network structure; quasi-maximum likelihood estimators; volatility prediction power.

MOS subject classification: 62M10; 62M20; 62F12.

1. INTRODUCTION

Volatility analysis is an important issue in modern financial markets. As such, a great many researchers have focused on developing and evaluating volatility models. The natural information source of the volatility is the historical data of the security, which can be further divided into low-frequency and high-frequency historical data. The low-frequency historical data refer to the observed prices of the security daily or at longer time horizons. The autoregressive conditional heteroskedasticity (ARCH) in Engle (1982) and the generalized autoregressive conditional heteroskedasticity (GARCH) models in Bollerslev (1986) are the most famous models for the analysis of low-frequency data. The high-frequency data refer to the intra-day prices of a security, such as tick-by-tick, 1-second, 5-minute data, and etc. The researchers often model the high-frequency historical data by continuous-time Itô processes and develop realized volatility estimators. These estimators include realized volatility (RV) (Anderson et al., 2001; Barndorff-Nielsen, 2002), two-time scale realized volatility (Zhang et al., 2005), multi-scale realized volatility (Zhang, 2006), kernel realized volatility (Barndorff-Nielsen et al., 2008), pre-averaging realized volatility (Jacod et al., 2009) and quasi-maximum likelihood estimator (Xiu, 2010).

However, despite that the high-frequency and low-frequency data must be interrelated at different time scales, the corresponding models were developed quite separately in the literature. There are some attempts to bridge the gap between high-frequency and low-frequency data. Wang (2002) studied the statistical relationship between the GARCH and diffusion model; Engle and Gallo (2006) proposed the multiplicative error
model; Ghysels et al. (2006) studied the Mixed Data Sampling model (MIDAS); Corsi (2009) provided Heterogeneous Autoregressive model for Realized Volatility (HAR) model; Shephard and Sheppard (2010) proposed the high-frequency-based volatility models, they call ‘HEAVY’ models (High-frequency-bAsed Volatility models) for simplicity; Hansen et al. (2012) studied volatilities analysis by combining the realized GARCH model and the high-frequency volatility model; Kim and Wang (2016) proposed GARCH-Itô model for merging low-frequency data and high-frequency data. Song et al. (2021) and Yuan et al. (2022) extended GARCH-Itô model to Realized GARCH-Itô model and QGARCH-Itô model respectively.

All the findings show that the volatility models combining low-frequency and high-frequency data have stronger forecasting power than those based on one information source. The GARCH-Itô-type models bridge the dynamic high-frequency structures and the low-frequency GARCH model in a complicated way. On the other hand, the realized volatility-based models, such as HAR, HEAVY, MIDAS and realized GARCH models can directly forecast realized volatilities from high-frequency data. Among them, the HEAVY model stands out as a direct approach for predicting daily asset return volatility, utilizing high-frequency constructed realized measures. The model structure is very simple, and allows us to clearly understand the prediction for the variance of daily return and the realized measure, which is out of the intellectual insights of the GARCH model. Empirical studies show that the HEAVY model’s in- and out-of-sample performance are both better than the GARCH model. Moreover, portfolio optimization and risk management in the financial markets often require cross-sections of hundreds of different stocks. This led to the development of multi-variate high-frequency-based volatility (multi-variate HEAVY) model (Noureldin et al., 2012). The empirical results suggest that the multi-variate HEAVY model outperforms the multi-variate GARCH model. Jin and Maheu (2013) provided a Bayesian estimation method to improve the density forecasts of multi-variate daily returns, however, the multi-variate HEAVY model is much easier to estimate and allows for straightforward out-of-sample model evaluation due to its closed-form forecasting formulae.

Nevertheless, owing to the structure of the multi-variate HEAVY model, it suffers from the ‘curse of dimensionality’, not only in terms of the number of parameters in the model but also in the dimension of the realized measure required to drive the dynamics. Therefore, many efforts have been taken to solve the issue by dimension reduction. In particular, dimension reduction by factor modeling has been proved very useful. For example, Kim and Fan (2019) and Kim et al. (2022) developed the factor-based diffusion process to account for the low-frequency large volatility dynamics. Shin et al. (2021) proposed the factor and idiosyncratic VHAR-Itô volatility models to account for the factor and idiosyncratic dynamics. Sheppard and Xu (2019) proposed the factor HEAVY models by resembling the $\beta$-GARCH model. The factor HEAVY model requires dynamic factor variance, dynamic factor loading, and dynamic idiosyncratic variance. To our knowledge, none of the methods have taken the observed network structures into consideration for the volatility models combining high-frequency and low-frequency data. This is the key contribution we intend to make in this work.

In real world, many systems can be described by complex networks (Newman, 2003; Reka & Barabasi, 2002). In essence, a financial market can be represented as a network, where nodes represent financial entities such as stocks, and the edges connecting them represent the correlations between their returns (Huang et al., 2009; Mantegna, 1999; Tumminello et al., 2010). Some researchers studied the network structures between the stocks by the minimal spanning tree method (Bonanno et al., 2001; Brida et al., 2016). Empirical data analysis shows that model performance can be improved significantly by incorporating network structure information (Nitzan & Libai, 2011; Goel & Goldstein, 2014; Wei et al., 2016). Moreover, to incorporate the network information among individuals, Zhu et al. (2017) developed a network vector autoregression (NAR) model. The response of each individual can be explained by its lagged values, the average of its neighbors, and a set of node-specific covariates. Inspired by Zhu et al. (2017)’ paper, Zhou et al. (2020) proposed a network GARCH model that used information derived from an appropriately defined network structure. This reduces the number of parameters and the computational complexity of the network GARCH model.

In this article, we develop a new method for multi-variate HEAVY model with the observed network structure by combining high-frequency and low-frequency data. We call the proposed model the network HEAVY model.
the regime, each asset has several different network structures, which can be represented by the adjacency matrix
\[ A^{(i)} = (a^{(i)}_{ij}), \ i, j = 1, \ldots, N. \] As mentioned in Section 2, the network structure can be characterized as the industry sectors’ network structure, that is, \( a_{ij} = 1 \), if the \( i \)th stock and the \( j \)th stock belong to the same industry sector, and \( a_{ij} = 0 \) otherwise. We can also choose the stock returns’ network structure represented by the sample correlation matrix between the stocks’ returns. In contrast to the traditional multi-variate HEAVY model, the network HEAVY model uses information of the stocks by defining the appropriate network structures, which can overcome the ‘curse of dimensionality’ of the multi-variate HEAVY model. The parameters of the network HEAVY model can be estimated by the quasi-maximum likelihood functions, and the computational complexity drops from \( \mathcal{O}(N^2) \) to \( \mathcal{O}(N) \). Compared with the multi-variate HEAVY model and the other volatility models such as network GARCH, multi-variate GARCH, and HAR model, real data analysis of the constitute stocks from S&P 500 illustrates that the network HEAVY model has stronger volatility forecasting power in the variance of the daily return and the realized measure.

This article is organized as follows. Section 2 introduces the network HEAVY model, as well as the model parameterization, iterative multi-step-ahead forecasts, and targeting parameterization. Section 3 derives the asymptotic properties. Section 4 conducts a series of simulation studies to check the finite in-of-sample performance and out-of-sample prediction of the proposed methodology. Section 5 carries out a real data analysis with 12 constituent stocks from the S&P 500 to demonstrate the usefulness of the proposed model in the variances of daily returns and daily realized measures forecasting. Section 6 concludes the article. All the proofs are relegated to Appendix A.

2. NETWORK HEAVY MODEL

The HEAVY models are direct models of daily asset return volatility based on realized measures constructed from high-frequency data. These models allow for both mean reversion and momentum,

\[ \text{var}(r_t | P^{HF}_{t-1}) = h_t = \omega + a \text{RM}_{t-1} + \beta h_{t-1}, \quad \omega, \alpha \geq 0, \quad \beta \in [0, 1), \]  

\[ \text{E}(\text{RM}_t | P^{HF}_{t-1}) = \mu_t = \omega_R + \alpha_R \text{RM}_{t-1} + \beta_R \mu_{t-1}, \quad \omega_R, \alpha_R, \beta_R \geq 0, \quad \alpha_R + \beta_R \in [0, 1), \]  

where \( \{r_t\} \) is the daily financial return sequence and \( \{\text{RM}_t\} \) is a corresponding sequence of daily realized measures. \( P^{HF}_{t-1} \) denotes the past information of \( r_t \) and \( \text{RM}_t \), and the superscript \( HF \) denotes the high-frequency dataset. Particularly, (2.1) is called the HEAVY-\( r \) model and (2.2) the HEAVY-RM model. It is noted that the HEAVY model is non-nested.

Supposed that there are \( N \) stocks, and let \( r_{it} \) denote the daily return, \( \text{RM}_{it} \) denote the corresponding daily realized measures of \( i \)th stock at \( t \)th day, where \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \). To model the network structures among stocks, we propose the network HEAVY model, abbreviated as NHEAVY model for simplicity,

\[ \text{var}(r_{it} | P^{HF}_{t-1}) = h_{it} = \omega + a \text{RM}_{t-1} + \sum_{\ell=1}^{L} \lambda^{(\ell)} d^{(\ell)} a^{(\ell)} \sum_{j \neq i} a^{(\ell)}_{ij} \text{RM}_{jt-1} + \beta h_{t-1}, \]  

\[ \text{E}(\text{RM}_{it} | P^{HF}_{t-1}) = \mu_{it} = \omega_R + \alpha_R \text{RM}_{t-1} + \sum_{\ell=1}^{L} \lambda^{(\ell)} d^{(\ell)} a^{(\ell)} \sum_{j \neq i} a^{(\ell)}_{ij} \text{RM}_{jt-1} + \beta_R \mu_{t-1}, \]  

where \( a^{(\ell)}_{ij} \) represents the edge between stock \( i \) and \( j \) in the \( \ell \)th network. And \( \sum_{j \neq i} \) denotes \( \sum_{j=1, j \neq i}^{N} \). Specifically, \( a^{(\ell)}_{ij} \) can be constructed into the adjacency matrices \( A^{(\ell)} = (a^{(\ell)}_{ij}) \in \mathbb{R}^{N \times N} \), \( \ell = 1, \ldots, L \), which can portray different networks, such as the overall market index, industry sector, the stocks’ return and some other
network. We follow the convention and do not allow any stock (node) to be self-related, so that $a_{ij}^{(r)} = 0$ for any $1 \leq i \leq N$ (Zhu et al., 2017; Zhou et al., 2020). If $i \neq j$, we can take the following two examples for $a_{ij}^{(r)}$:

**Example 1.** $a_{ij}^{(r)} = a_{ji}^{(r)} = 1$, which represents the $i$th stock and $j$th stocks are connected, if they belong to the same industry sector;

**Example 2.** $a_{ij}^{(r)} = \rho_{ij}$, and $\rho_{ij}$ is the sample correlation coefficient of the returns between any two stock $i$ and $j$ with

$$\rho_{ij} = \frac{\sum_{t=1}^{T} (r_{ij} - \bar{r}_i) (r_{ij} - \bar{r}_j)}{\sqrt{\sum_{t=1}^{T} (r_{ij} - \bar{r}_i)^2} \sqrt{\sum_{t=1}^{T} (r_{ij} - \bar{r}_j)^2}}$$

where $\bar{r}_i = \frac{1}{T} \sum_{t=1}^{T} r_{ij}$ is the sample mean for the $i$th stock. It is clear that $\rho_{ij} = \rho_{ji}$.

In the meanwhile, $d_{ij}^{(r)} = \sqrt{\sum_{j=1}^{N} a_{ij}^{(r)} \sum_{j=1}^{N} a_{ij}^{(r)}}$ represents the sum of weights of the connected stocks (neighbors) of the $i$th stock, which is similar to the out-degree. In the case, $d_{ij}^{(r)} = \sum_{j \neq i} a_{ij}^{(r)}$ is the standardized weight (effect) of the $j$th stock for the $i$th stock. If $d_{ij}^{(r)} = 0$ for some $i$, the $i$th stock is regarded as isolated in the $r$th network. Following the convention, we set $d_{ij}^{(r)} = \sum_{j \neq i} a_{ij}^{(r)}$ for any network.

In contrast to the HEAVY models, the $i$th stock is assumed to be only affected by its directly connected neighbors in the determined networks, which is particularly reasonable in practice. We call (2.3) NHEAVY-$\tau$ model and (2.4) NHEAVY-RM model respectively. Here, $\lambda$ and $\lambda_R$ are used to capture the influence of other stocks on the $i$th stock over the network structure. It is also note that the NHEAVY model merely requires $O(N)$ parameters to model the dependency structure among $N$ stocks. Similar discussions have been introduced in the network vector autoregression (NAR) model (Zhu et al., 2017) and the NARQ model (Zhou et al., 2020).

Then we discuss the stationary condition for the proposed NHEAVY model. Define $r_t^2 = (r_{1t}^2, \ldots, r_{Nt}^2)^T$, $RM_t = (RM_{1t}, \ldots, RM_{Nt})^T$, $h_t = (h_{1t}, \ldots, h_{Nt})^T$, $\mu_t = (\mu_{1t}, \ldots, \mu_{Nt})^T$, $\mathbf{D}^{(r)} = \text{diag}(d_{1t}^{(r)}, \ldots, d_{Nt}^{(r)})$, $\omega = \omega \mathbf{1}_N$, $\alpha_R = \omega_R \mathbf{1}_N$, $\beta = \beta \mathbf{1}_N$, $\beta_R = \beta_R \mathbf{1}_N$, $\mathbf{A}(\cdot) = \alpha(\cdot) \mathbf{1}_N$, $\lambda(\cdot) = \lambda_R(\cdot) \mathbf{1}_N$, and the $r$th adjacency matrix $\mathbf{A}^{(r)}(\cdot) = (a_{ij}^{(r)}) \in \mathbb{R}^{N \times N}$, where $\mathbf{1}_N = (1, \ldots, 1)^T$ and $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. Following the NHEAVY representation in (2.3) and (2.4), the dynamic structure of the bivariate model can be gleaned from rewriting

$$
\begin{pmatrix}
\mathbf{h}_t \\
\mu_t
\end{pmatrix} = \mathbf{w} + \mathbf{B} \begin{pmatrix}
\mathbf{h}_{t-1} \\
\mu_{t-1}
\end{pmatrix} + \begin{pmatrix}
\alpha + \sum_{\ell=1}^{L} \lambda(\cdot) \mathbf{D}^{(r)-1} \mathbf{A} \\
\alpha_R + \sum_{\ell=1}^{L} \lambda_R(\cdot) \mathbf{D}^{(r)-1} \mathbf{A}^{(r)}
\end{pmatrix} \left(\mathbf{RM}_{t-1} - \mu_{t-1}\right),
$$

where $\mathbf{w} = (\omega, \omega_R)^T \in \mathbb{R}^{2N \times 1}$, $\mathbf{B} = \begin{pmatrix}
\beta & \alpha + \sum_{\ell=1}^{L} \lambda(\cdot) \mathbf{D}^{(r)-1} \mathbf{A} \\
0 & \alpha_R + \sum_{\ell=1}^{L} \lambda_R(\cdot) \mathbf{D}^{(r)-1} \mathbf{A}^{(r)} + \beta_R
\end{pmatrix} \in \mathbb{R}^{2N \times 2N}$.

Let $\mathbf{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})^T$, $\varepsilon_{it}$ is i.i.d with $\mathbb{E}\{\varepsilon_{it} | \mathbf{F}_{t-1}^{\text{HF}}\} = 1$ and $\mathbb{E}\{\varepsilon_{it}^2 | \mathbf{F}_{t-1}^{\text{HF}}\} = \kappa_{\varepsilon_t}^2$. Similarly, $\mathbf{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})^T$, $\varepsilon_{it}$ is i.i.d with $\mathbb{E}\{\varepsilon_{it} | \mathbf{F}_{t-1}^{\text{HF}}\} = 1$ and $\mathbb{E}\{\varepsilon_{it}^2 | \mathbf{F}_{t-1}^{\text{HF}}\} = \kappa_{\varepsilon_t}^2$. Furthermore, $\mathbb{E}\{\varepsilon_{it} \varepsilon_{jt} | \mathbf{F}_{t-1}^{\text{HF}}\} = \kappa_{\varepsilon_t}^{\varepsilon_j}$, $\mathbb{E}\{\varepsilon_{it} \varepsilon_{jt} | \mathbf{F}_{t-1}^{\text{HF}}\} = 0, i \neq j$. 

---

1 We thank the two referees. One referee suggests that the sample correlation matrix is reasonable to form a network, and another suggests that we can construct several different networks such as the overall market index, industry sector, or some other networks. In addition, the number of network $L$ can be chosen empirically in practice.
Denote \(r^2 = \varepsilon_i h_i\) and \(\text{RM}_i = \varepsilon_i \mu_i\), \(\varepsilon_i, h_i = (\varepsilon_{i,1} h_{i,1}, \ldots, \varepsilon_{i,N} h_{i,N})^\top\), \(\varepsilon_i, \mu_i = (\varepsilon_{i,1} \mu_{i,1}, \ldots, \varepsilon_{i,N} \mu_{i,N})^\top\), then the vector multiplicative representation of HEAVY models in (2.3) and (2.4) can be rewritten as,

\[
\begin{pmatrix}
  r^2_i \\
  \text{RM}_i
\end{pmatrix} = \begin{pmatrix}
  \varepsilon_i, h_i \\
  \varepsilon_i, \mu_i
\end{pmatrix} = \begin{pmatrix}
  h_i \\
  \mu_i
\end{pmatrix} + \begin{pmatrix}
  (\varepsilon_i - 1_N) \\
  \mu_i(\varepsilon_i - 1_N)
\end{pmatrix},
\]

where \((\varepsilon_i - 1_N)^\top, (\varepsilon_i - 1_N)^\top\)^\top can be seen as a vector martingale difference sequence with respect to \(\mathcal{F}_{\text{HF}}^t\) (see Engle and Gallo, 2006; Shephard and Sheppard, 2010). Let \(u_t = (u_{1t}, \ldots, u_{Nt})^\top\), \(u_{R_t} = (u_{1Rt}, \ldots, u_{NRt})^\top\), with \(u_{it} = r^2 - h_{it} = (\varepsilon_{i,1} - 1)h_{it}, u_{it} = \text{RM}_{it} = \mu_{it} = (\varepsilon_{i,1} - 1)\mu_{it}\). Then, both \(\{u_{it}\}\) and \(\{u_{Rt}\}\) can be viewed as i.i.d. ‘noise’ with zero mean, and finite variance.

Combining (2.5) and (2.6), a bi-product of the processes \(r^2_i\) and \(\text{RM}_i\) is the VARMA(1,1) representation,

\[
\begin{pmatrix}
  r^2_i \\
  \text{RM}_i
\end{pmatrix} = w + B \begin{pmatrix}
  r^2_{i-1} \\
  \text{RM}_{i-1}
\end{pmatrix} + \begin{pmatrix}
  (1_N - \beta_1 1_L)u_t \\
  (1_N - \beta_R 1_L)u_{R_t}
\end{pmatrix},
\]

where \(1_N = (1, \ldots, 1)^\top\), and \(L\) is the lag operator.

By the VARMA(1,1) representation in (2.7), we immediately have the necessary and sufficient condition for the existence of a unique strictly stationary solution to the NHEAVY model. Strictly speaking, the eigenvalues of \(B\) must be less than one in modulus. Since \(B\) is block triangular, its eigenvalues are members of the multi-set of the eigenvalues of \(\beta\) and \(\alpha_R + \sum_{l=1}^L \lambda^{(e)}_R \mathbf{A}^{(e)} + \beta_R\). Suppose that \(\lambda^{(e)}_d\) is any arbitrary eigenvalue of \(\mathbf{A}^{(e)}\). Since \(\mathbf{D}^{(e)}\mathbf{A}^{(e)} = (\alpha^{(e)}_0 + \lambda^{(e)}_d)^\top\), by the Gershgorin circle theorem on the eigenvalues of matrices and by the definition of \(d_i\), we can get that \(|\lambda^{(e)}_d| \leq 1\). Thus, the largest eigenvalue in modular of \(\alpha_R + \sum_{l=1}^L \lambda^{(e)}_R \mathbf{A}^{(e)} + \beta_R\) is smaller than \(\alpha_R + \sum_{l=1}^L \lambda^{(e)}_R + \beta_R\), i.e. \(\lambda^{(e)}_c(\alpha_R + \sum_{l=1}^L \lambda^{(e)}_R \mathbf{A}^{(e)} + \beta_R) \leq \alpha_R + \sum_{l=1}^L \lambda^{(e)}_R + \beta_R < 1\). The largest eigenvalue of \(\beta\), i.e., \(\rho(B) = \beta < 1\).

Multi-step-ahead forecasts of volatility are very important for asset allocation and risk assessment. Since these tasks are usually carried out over multiple days, in contrast to one-step-ahead forecasts of volatility that only require the NHEAVY-r model, both NHEAVY-r and NHEAVY-RM models play a central role in multi-step-ahead forecasts.

**Remark 1.** Let \(1_N\) denote an \(N \times N\) identity matrix, for \(s \geq 0\), from the dynamic structure representation (2.5), we have the multi-step-ahead forecasts as follows,

\[
\begin{pmatrix}
  h_{i+s|s-1} \\
  \mu_{i+s|s-1}
\end{pmatrix} = (1_N + B + \cdots + B^s)w + B^{s+1} \begin{pmatrix}
  h_{i-1} \\
  \mu_{i-1}
\end{pmatrix},
\]

Write \(v_1 = \alpha + \sum_{e=1}^L \lambda^{(e)} \mathbf{D}^{(e)} \mathbf{A}^{(e)}\) and \(v_2 = \alpha + \sum_{e=1}^L \lambda^{(e)} \mathbf{D}^{(e)} \mathbf{A}^{(e)} + \beta_R\), then for \(J = 1, 2, \ldots\),

\[
B^J = \begin{pmatrix}
  \beta^J_1 & v_1^{J-1} + v_2^{J-2} \beta + \cdots + \beta^{J-1} \\
  0 & v_2^J
\end{pmatrix}.
\]

Similar to the HEAVY model, there are benefits to re-parameterize the NHEAVY model, as it allows the the intercepts to be explicitly related to the unconditional mean of squared returns and realized measures. Let \(\mu = (\mu_1, \ldots, \mu_N)^\top\) and \(\mu_R = (\mu_{R1}, \ldots, \mu_{RN})^\top\), where \(\mu_i = E \left( r^2_{it} \right)\) and \(\mu_{Rt} = E \left( \text{RM}_{it} \right)\).
Remark 2. The ‘targeting parameterization’ for the NHEAVY model is provided as follows,

\[ h_i = \left( I_N - \left( \alpha + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} \right) \kappa - \beta \right) \mu + \left( \alpha + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} \right) \mathbf{R}_{M_{i-1}} + \beta \mathbf{h}_{i-1}, \]  
\[ \mu_i = \left( I_N - \mathbf{a}_R - \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} - \beta_R \right) \mu_R + \left( \alpha_R + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} \right) \mathbf{R}_{M_{i-1}} + \beta_R \mu_{i-1}, \]

where \( \kappa = \text{diag}(\kappa_1, \ldots, \kappa_N)^T \), \( \kappa_i = \mu_R \mu_i^{-1} \).

Using (2.9) and (2.10), it is easier to impose the condition: the eigenvalues of \( (\alpha + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} ) \kappa + \beta \) must be less than one in modulus. By the arguments of block triangular and Gershon circle theorem, \( \rho(\alpha + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} ) \kappa + \beta < 1 \), where \( \kappa_{(i)} \) is the maximum of \( \kappa_i \) for \( i = 1, \ldots, N \).

It is worth mentioning that the ‘targeting parameterization’ can make use of the estimations of \( \mu_R, \mu, \) and \( \kappa \), where

\[ \hat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} \frac{r_{it}^2}{R_{M_{it}}}, \quad \hat{\mu}_R = \frac{1}{T} \sum_{i=1}^{T} \mathbf{R}_{M_{it}}, \quad \hat{\kappa}_i = \hat{\mu}_R \hat{\mu}_{i}^{-1}. \]

When these estimators are plugged into the quasi-likelihood functions, the optimization tasks can be substantially simplified, but it does alter the resulting asymptotic standard errors.

3. QUASI-MAXIMUM LIKELIHOOD ESTIMATION

We assume that returns \( \{r_{it}^2\} \) and the related realized measures \( \{\mathbf{RM}_{it}\} \) are from model (2.7). Let \( \lambda_0 = \left( \lambda_{0(1)}, \ldots, \lambda_{0(L)} \right)^T, \lambda_{R0} = \left( \lambda_{R0(1)}, \ldots, \lambda_{R0(L)} \right)^T, \mathbf{\Phi}_0 = \left( \alpha_0, \alpha^T, \lambda_{R0}^T, \beta_0 \right)^T, \) and \( \mathbf{\Phi}_{R0} = \left( \omega_{R0}, \alpha_{R0}, \lambda_{R0}^T, \beta_{R0} \right)^T. \)

Denote \( \theta_0 = (\mathbf{\phi}_0^T, \mathbf{\Phi}_{R0}^T)^T \in \mathbb{R}^{2L+6} \) to be the true values and \( \theta = (\mathbf{\phi}_i^T, \mathbf{\Phi}_{Ri}^T)^T \in \mathbb{R}^{2L+6} \) to be the parameters. When modeling inference, we will regard the parameters as having no link between the NHEAVY-r and NHEAVY-RM models, i.e. \( \mathbf{\phi}_i \) and \( \mathbf{\Phi}_{Ri} \) are variation free, which is important in the one-step estimation for the NHEAVY model.

The quasi-log-likelihood function (ignoring the constants) of returns \( \{r_{it}^2\} \) is given by

\[ \hat{\mathcal{L}}(\mathbf{\phi}) = \frac{1}{T} \sum_{t=2}^{T} \hat{\mathcal{L}}_t(\mathbf{\phi}) \text{ with } \hat{\mathcal{L}}_t(\mathbf{\phi}) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \log \hat{\lambda}_i(\mathbf{\phi}) + \frac{r_{it}^2}{\hat{h}_i(\mathbf{\phi})} \right\}, \]

where \( \hat{\lambda}_i(\mathbf{\phi}) \) is defined recursively for \( t > 1 \) by

\[ \hat{\lambda}_i(\mathbf{\phi}) = \omega + a \mathbf{R}_{M_{i-1}} + \sum_{\ell=1}^{L} \lambda^{(\ell)} \left( D^{(\ell)} \right)^{-1} \mathbf{A}^{(\ell)} \mathbf{R}_{M_{i-1}} + \beta \hat{\lambda}_{i-1}(\mathbf{\phi}), \]

with: \( \hat{\lambda}_{i1}(\mathbf{\phi}) = T^{-1/2} \sum_{j=1}^{[T(i-1)/T]} r_{it}^2 \).

The quasi-log-likelihood function (ignoring the constants) of realized measures \( \{\mathbf{RM}_{it}\} \) is given by

\[ \hat{\mathcal{L}}(\mathbf{RM} \mathbf{\phi}_i) = \frac{1}{T} \sum_{t=2}^{T} \hat{\mathcal{L}}_{RM}(\mathbf{RM} \mathbf{\phi}_i) \text{ with } \hat{\mathcal{L}}_{RM}(\mathbf{RM} \mathbf{\phi}_i) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \log \hat{\lambda}_i(\mathbf{RM} \mathbf{\phi}_i) + \frac{\mathbf{RM} \mathbf{\phi}_i}{\hat{h}_i(\mathbf{\phi})} \hat{\mathbf{RM}} \mathbf{\phi}_i \right\}, \]
where \( \hat{\mu}_i(\phi_R) \) is defined recursively for \( t > 1 \) by

\[
\hat{\mu}_i(\phi_R) = \omega_R + \alpha_R RM_{it-1} + \sum_{\ell=1}^{L} \alpha^{(\ell)}_{it} \sum_{j \neq i} a^{(\ell)}_{ij} RM_{jt-1} + \beta_R \hat{\mu}_{i-1}(\phi_R),
\]

with \( \hat{\mu}_1(\phi_R) = T^{-1/2} \sum_{t=1}^{\lfloor T/\gamma \rfloor} RM_t \).

Then the QMLE of \( \phi \) is defined as

\[
\hat{\phi} = (\hat{\omega}, \hat{\alpha}, \hat{\lambda}, \hat{\beta})^T = \arg \min_{\phi \in \Phi} L'(\phi),
\]

where \( \Phi \) is the parameter space of \( \phi \) and the QMLE of \( \phi_R \) is defined as

\[
\hat{\phi}_R = (\hat{\omega}_R, \hat{\alpha}_R, \hat{\lambda}_R, \hat{\beta}_R)^T = \arg \min_{\phi_R \in \Phi_R} L_{RM}(\phi_R),
\]

where \( \Phi_R \) is the parameter space of \( \phi_R \).

We define \( \hat{\theta} = (\hat{\phi}^T, \hat{\phi}_R^T)^T \). To discuss the asymptotic properties of \( \hat{\theta} \), we first study those of \( \hat{\phi} \) and \( \hat{\phi}_R \). It is convenient to approximate the sequences \( \{ \hat{h}_i(\phi) \} \) and \( \{ \hat{\mu}_i(\phi_R) \} \) by the ergodic stationary sequences \( \{ \hat{h}_i(\phi) \} \) and \( \{ \hat{\mu}_i(\phi_R) \} \) as follows,

\[
h_i(\phi) = \omega + \alpha RM_{it-1} + \sum_{\ell=1}^{L} \alpha^{(\ell)}_{it} \sum_{j \neq i} a^{(\ell)}_{ij} RM_{jt-1} + \beta h_{it-1}(\phi), \]

\[
\mu_i(\phi_R) = \omega_R + \alpha_R RM_{it-1} + \sum_{\ell=1}^{L} \alpha^{(\ell)}_{it} \sum_{j \neq i} a^{(\ell)}_{ij} RM_{jt-1} + \beta_R \mu_{i-1}(\phi_R),
\]

for any \( t \) and each \( i \). Similarly to the definition of \( \hat{L}'(\phi), \hat{\mu}'(\phi), \hat{L}_{RM}(\phi_R), \) and \( \hat{\mu}_{RM}(\phi_R) \), we can define

\[
L'(\phi) = \frac{1}{T} \sum_{t=2}^{T} \epsilon'_i(\phi) \quad \text{with} \quad \epsilon'_i(\phi) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \log h_i(\phi) + \frac{r^2}{h_{ii}(\phi)} \right\}, \tag{3.3}
\]

\[
L_{RM}(\phi_R) = \frac{1}{T} \sum_{t=2}^{T} \epsilon'_{RM}(\phi_R) \quad \text{with} \quad \epsilon'_{RM}(\phi_R) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \log \mu_i(\phi_R) + \frac{RM_{ii}}{\mu_{ii}(\phi_R)} \right\}. \tag{3.4}
\]

Before stating our main results, we give the following assumptions that are standard in quasi-maximum likelihood estimation.

**Assumption 1.** The parameter space \( \Theta \) is a compact subset of \( \{ \theta : \omega > 0, \alpha > 0, \lambda > 0, \alpha_R > 0, \lambda_R > 0, \beta_R > 0, 0 < \beta < 1, \alpha + \sum_{\ell=1}^{L} \lambda^{(\ell)}_{N(\ell)} + \beta < 1, \alpha_R + \sum_{\ell=1}^{L} \lambda^{(\ell)}_{R(\ell)} + \beta_R < 1 \} \) and \( \theta_0 \in \Theta \).

**Assumption 2.** Both \( \{ u_i \} \) and \( \{ u_{Ri} \} \) in model (2.7), are i.i.d. across \( i \) and \( t \) with zero mean and finite variance, and they are Gaussian errors. Moreover, \( \{ u_i \} \) and \( \{ u_{Ri} \} \) are non-degenerate.

The strong consistency and asymptotic normality of the QMLE \( \hat{\theta} \) for the NHEAVY model is stated in the Theorem 1.
Theorem 1. If Assumption 1 and 2 hold, then \( \hat{\theta} \to \theta_0 \) almost surely as \( T \to \infty \). Furthermore, if \( \theta_0 \) is an interior point of \( \Theta \), we have

\[
\sqrt{NT} \left( \hat{\theta} - \theta_0 \right) \xrightarrow{d} N \left( 0, I^{-1} \mathbf{J} (I^{-1})^\top \right), \tag{3.5}
\]

where

\[
I = \begin{pmatrix} \Sigma_r & 0 \\ 0 & \Sigma_r \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} (\kappa^2_r - 1) \Sigma_r & (\kappa^2_r - 1) \Sigma_r^R \\ (\kappa^2_r - 1) \Sigma_r^R & (\kappa^2_r - 1) \Sigma_r \end{pmatrix}.
\]

where

\[
\Sigma_r = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{\nabla_i^2 (\rho_i) \partial \theta_i} \frac{\partial \theta_i}{\partial \theta_i} \right), \quad \Sigma_r^R = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{\nabla_i^2 (\rho_i) \partial \theta_i} \frac{\partial \theta_i}{\partial \theta_i^R} \right), \quad \Sigma_r^R = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{\nabla_i^2 (\rho_i) \partial \theta_i} \frac{\partial \theta_i}{\partial \theta_i^R} \right), \quad \kappa^2_r = E(\epsilon_i^2), \quad \kappa^2_R = E(\epsilon_i^2), \quad \text{and} \quad \kappa^2_r = E(\epsilon_i^2).
\]

The block diagonality of (3.5) is due to the variation-free property of parameters, which is a straightforward application of quasi-likelihood theory and can be viewed as an extension of Bollerslev and Wooldridge (2007), and is also discussed extensively in Cipollini et al. (2007).

Then, we discuss the two-step estimation for the NHEAVY model. With the help of Remark 2, the targeting parameterization can be written in an alternative form,

\[
\begin{align*}
\hat{h}_i &= (1 - \alpha \kappa_i - \beta) \mu_i - \sum_{\ell=1}^{1} \lambda^{(c)} d^{(c)}_i \sum_{j \neq i} a_j^{(c)} \mu_i \mu_j + \alpha \text{RM}_{it-1} \\
&\quad + \sum_{\ell=1}^{1} \lambda^{(c)} d^{(c)}_i \sum_{j \neq i} a_j^{(c)} \text{RM}_{it-1} + \beta h_{it-1}, \\
\hat{\mu}_i &= (1 - \alpha_R - \beta_R) \mu_i - \sum_{\ell=1}^{1} \lambda^{(c)} d^{(c)}_i \sum_{j \neq i} a_j^{(c)} \mu_i \mu_j + \alpha \text{RM}_{it-1} \\
&\quad + \sum_{\ell=1}^{1} \lambda^{(c)} d^{(c)}_i \sum_{j \neq i} a_j^{(c)} \text{RM}_{it-1} + \beta_R h_{it-1}, \tag{3.6}
\end{align*}
\]

From model (3.6), a convenient two-step approach can be constructed. The unconditional moments, \( \mu_i \) and \( \mu_{Ri} \), are estimated in the first step by

\[
\hat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} r_{it}^2, \quad \hat{\mu}_{Ri} = \frac{1}{T} \sum_{t=1}^{T} \text{RM}_{it}.
\]

Define \( \hat{L}(\mu, \mu_{R}, \hat{\phi}) \) and \( \hat{L}^\text{RM}(\mu_{R}, \hat{\phi}_R) \) to be the observed log-likelihood function for the covariance targeting HEAVY model. Then \( \hat{\phi} \) and \( \hat{\phi}_R \) can be estimated by QMLE in the second step.

\[
\hat{\phi} = \arg \min_{\theta \in \Theta} \hat{L}(\mu, \mu_{R}, \phi) \quad \text{and} \quad \hat{\phi}_R = \arg \min_{\theta_{R} \in \Theta_{R}} \hat{L}^\text{RM}(\mu_{R}, \phi_R).
\]

With ‘targeting parameterization’ in Remark 2, the variation free property between the parameters of the NHEAVY-r and NHEAVY-RM no longer holds since \( \kappa \) depends on \( \mu_R \). And the asymptotic properties of the estimators can refer to in the HAC estimator (Noureldin et al., 2012).
4. SIMULATION STUDIES

Two simulation experiments are conducted in this section: the first is to evaluate the finite-sample performance of the quasi-maximum likelihood estimation for the proposed NHEAVY model in Section 3, and the second one is to show the forecasting power of the estimation.

In the first experiment, the log prices \( X_t \) are generated from the following model,

\[
dX_t = \mu_t ds + \sigma_t^T dB_t,
\]

where \( s = t - 1 + \ell / m \) for \( t = 1, \ldots, T, \ell = 0, \ldots, m, \mu_t \in \mathbb{R}^N \) is the drift term, \( \sigma_t \in \mathbb{R}^{N \times N} \) is the volatility matrix of \( X_t \) which satisfies \( \gamma(s) = \sigma_s^T \sigma_s \), and \( B_t \in \mathbb{R}^N \) is a standard \( N \)-dimensional Brownian motion. For convenience, we let \( \mu_t = 0 \) and the volatility \( \sigma_t \) is a Cholesky decomposition of \( \gamma(s) \) with

\[
\gamma(s) = (\gamma(s)), \quad \gamma(s) = \sqrt{\tau_i \tau_j k^{i-j}},
\]

where \( \{ \tau_i, i = 1, \ldots, N \} \) are independently generated from a uniform distribution on \( [0, 1] \), and \( \kappa \) is set at 0.5. We set the time interval to \( \Delta = 1/78 \) during discretion and fix the initial log prices \( X_0 = \log(50) \) for all \( 1 \leq i \leq N \).

The model parameters are set to \( \phi_0 = (0.01, 0.06, 0.015, 0.015, 0.7), \phi_{R0} = (0.05, 0.1, 0.05, 0.05, 0.65) \), and \( \theta_0 = (\phi_0^T, \phi_{R0}^T)^T \).

Moreover, we let the number of networks to be \( L = 2 \), described by adjacency matrices \( A^{(1)} \) and \( A^{(2)} \) respectively. The first adjacency matrix \( A^{(1)} \) describes the network structure between the industry sectors. In the existing literature, a power-law distribution reflects a popular network phenomenon, that is, the majority of nodes have very few edges but a small amount has a huge number of edges. To mimic this phenomenon, we simulate \( A^{(1)} \) as follows. First, for each node, its in-degree satisfies \( P(d_i = k) = c k^{-\alpha} \) for a normalizing constant \( c \) and exponent parameter \( \alpha \in 1, 2, 3 \). A smaller \( \alpha \) value implies a heavier tail. Next, for the \( i \)-th node, we randomly select \( d_i \) nodes to be its followers. Another adjacency matrix \( A^{(2)} \) describes the network structures between the stocks’ return, see Example 2 in Section 2. We consider two numbers of assets \( N = 5 \) or \( N = 20 \) and three different sample sizes, \( T = 100j \) with \( j = 1, 2, 6 \), and there are \( P = 1000 \) replications for each setting. Let \( \theta^{(p)} = (\theta^{(p)}_k)^T = \left( (\phi^{(p)}_k, \phi^{(p)}_{R_k})^T \right)^T \) be the estimators obtained in the \( p \)-th \( (1 \leq p \leq P) \) replication. The RV is used to study the finite-sample performance of the parameter estimators. First, for a given parameter \( \theta_k \), with \( 1 \leq k \leq 10 \), the mean square error is evaluated by MSE_k = \( P^{-1} \sum_{p=1}^{P} (\hat{\theta}^{(p)}_k - \theta_k)^2 \). Second, for each \( 1 \leq k \leq 10 \), a 95\% confidence interval is constructed for \( \hat{\theta}_k \) as CI_k = \( (\hat{\theta}^{(p)}_k - z_{0.975} \bar{\theta}^{(p)}_k, \hat{\theta}^{(p)}_k + z_{0.975} \bar{\theta}^{(p)}_k) \), where \( \bar{\theta}^{(p)}_k \) is the square root of the \( j \)-th diagonal element of \( \bar{\theta}_k \), and \( z_{0.975} \) is the \( 0.975 \)-quantile of a standard normal distribution. Then, the coverage probability is computed as CP_k = \( P^{-1} \sum_{p=1}^{P} 1 \left( \theta_k \in \text{CI}_k^{(p)} \right) \), where IC(\cdot) is the indicator function. The results are summarized in Table I.

From Table I, all estimated parameters are consistent, and MSEs decease toward zero as \( T \to \infty \). This may be due to that, when \( T \) is larger, more sampled time intervals are available to fit the model. Moreover, the reported coverage probabilities (i.e., CP) for each parameter \( \theta_k \) are all fairly close to the nominal level of 95\%. This suggests that the estimated standard error (i.e., \( \bar{\theta}^{(p)}_k \)) approximates the true SE well. These findings confirm that the proposed estimator \( \hat{\theta} \) is consistent and asymptotically normal.

The second experiment is to evaluate the forecasting power of the NHEAVY model, and the original HEAVY model with the multi-variate setting (MHEAVY) is seen as a competitor. A sample is generated using the data-generating process in the first experiment with \( (N, T) = (10, 101) \), and the first 100 days are used for in-sample data. Meanwhile, the 101st day is the target for the out-of-sample forecast of daily volatility. The one-step predictors are denoted as \( \hat{\theta}_{101|100}(\hat{\theta}) \) and \( \mu_{101|100}(\hat{\theta}) \). Note that \( \hat{h}_{101|100}(\hat{\theta}) \) and \( \mu_{101|100}(\hat{\theta}) \) are the variance of daily return and the daily realized measures respectively. The 1-norm, 2-norm, and ∞-norm of the averaged
Table I. Simulation results for the NHEAVY model with 1000 replications

|   | \( N = 5 \) | \( N = 20 \) | \( N = 5 \) | \( N = 20 \) | \( N = 5 \) | \( N = 20 \) |
|---|---|---|---|---|---|---|
| \( \omega \) | 0.26 (94.3) | 0.03 (95.0) | 0.14 (95.0) | 0.02 (94.9) | 0.09 (94.6) | 0.02 (94.9) |
| \( \alpha \) | 8 (93.9) | 2.12 (92.4) | 5.05 (93.8) | 1.31 (93.0) | 2.11 (93.6) | 0.66 (93.5) |
| \( \lambda_1 \) | 0.85 (94.5) | 0.06 (92.6) | 0.47 (92.7) | 0.04 (92.5) | 0.15 (92.6) | 0.03 (92.4) |
| \( \lambda_2 \) | 2.98 (93.4) | 0.03 (93.7) | 1.58 (92.6) | 0.03 (92.9) | 0.63 (92.4) | 0.03 (92.3) |
| \( \beta \) | 1.48 (95.0) | 1.03 (95.0) | 0.47 (92.7) | 0.04 (92.5) | 0.15 (92.6) | 0.03 (92.4) |
| \( \omega_k \) | 0.09 (95.0) | 0.08 (95.0) | 0.04 (95.0) | 0.02 (95.0) | 0.02 (95.0) | 0.02 (95.0) |
| \( \alpha_k \) | 1.67 (92.3) | 1.22 (92.0) | 1.20 (92.9) | 1.07 (93.2) | 1.1 (92.5) | 1.01 (94.5) |
| \( \lambda_{1k} \) | 0.26 (92.5) | 0.26 (92.6) | 0.26 (93.8) | 0.25 (92.5) | 0.25 (92.5) | 0.25 (92.9) |
| \( \lambda_{2k} \) | 0.30 (92.6) | 0.25 (92.5) | 0.26 (94.3) | 0.25 (92.4) | 0.15 (92.6) | 0.23 (92.3) |
| \( \beta_k \) | 5.54 (95.2) | 3.10 (95.6) | 1.28 (95.7) | 0.04 (95.4) | 0.06 (95.6) | 0.01 (95.7) |

Notes: The MSEs \((\times 10^{-2})\) are reported for each estimator. The corresponding CPs (in %) are given in parentheses.

Table II. The 1-norm, 2-norm and \( \infty \)-norm \((\times 10^{-2})\) of the averaged forehead one step predictors for the NHEAVY and MHEAVY models

| One-step predictors | NHEAVY | MHEAVY |
|---------------------|--------|--------|
| \( h_{101|100}(\hat{\theta}) - h_{101|100}(\theta_0) \) | 2.938 | 7.773 |
| \( \mu_{101|100}(\hat{\theta}) - \mu_{101|100}(\theta_0) \) | 0.245 | 0.323 |

\( h_{101|100}(\hat{\theta}) - h_{101|100}(\theta_0) \) and \( \mu_{101|100}(\hat{\theta}) - \mu_{101|100}(\theta_0) \) are used to evaluate the forecast power of the NHEAVY and MHEAVY models. All the results are summarized in Table II. Table II shows that our proposed NHEAVY model has stronger forecast power.

5. REAL DATA ANALYSIS

This section analyzes the high-frequency trading data for the constituent stocks of S&P 500 index from 2 January 2011 to 30 December 2013. Specifically, we consider \( N = 12 \) stocks with largest trading volumes on 2 January 2013, and the data in 2013 are used to evaluated the out-of-sample performance. As a result, there are \( T = 497 \) days for estimation and \( M = 249 \) days for prediction.

The daily trading data from 9:30 am to 4:00 pm are downloaded from the Wharton Research Data Services, and we use the most common 5-minute returns, corresponding to the the number of intraday observations \( m = 78 \), in the literature (Andersen & Bollerslev, 1998; Zhang, 2011; Liu et al., 2015). Overnight returns are excluded since they usually have jumps influenced by external factors. RV is used to estimate the integrated volatility.

We consider empirically two network structures, i.e., \( A^{(1)} = (\alpha_{ij}^{(1)}), i = 1, \ldots, 12, \text{ and } j = 1, \ldots, 12 \) depicting the industry section network. Another network structure \( A^{(2)} = (\alpha_{ij}^{(2)}), i = 1, \ldots, 12, \text{ and } j = 1, \ldots, 12 \), describes the correlations between the stocks’ return. All the chosen stocks and the industry classification are in Table III.

The proposed NHEAVY model is compared to four competitors: the vector HAR (VHAR) model, the original HEAVY model, the GARCH model with their multi-variate settings, named MHEAVY and MGARCH.
Table III. Stocks’ name and their industry classification (divided by the global industry classification standard)

| Industry classification | Stocks name |
|-------------------------|-------------|
| Information technology  | AAPL, AMD, HPQ, INTC, MSFT, MU, ORCL |
| Consumer discretionary   | EBAY, F     |
| Financials              | AIG, MET, PGR |

respectively, and the network GARCH (NGARCH) model. We also construct another model with different factors as follows

\[
\text{var}(r_{it} | F_{t-1}^{HF}) = h_{it} = \omega + \alpha \text{RM}_{it-1} + \beta h_{it-1},
\]

where \(F_{t-1}^{HF}\) represents the \(\ell\)-th factors, which can be represented by the linear combination of the realized measures in different stocks. We named (5.1) and (5.2) the revised factor HEAVY (RFHEAVY) model.

A rolling forecast procedure is employed to evaluate the out-of-sample performance of the six models: the ending point of historical data iterates in the out-of-sample period of 2013 with the window size being fixed at \(T = 497\), and then one-step ahead prediction is conducted for each iteration. We adopt the most commonly used empirical quasi-likelihood (QLIKE) in the literature to evaluate the forecasting accuracy,

\[
\text{QLIKE}_i = \frac{1}{M} \sum_{t=1}^{M} \left( \frac{y_{it}}{\hat{y}_{it}} - \log \left( \frac{y_{it}}{\hat{y}_{it}} \right) - 1 \right) \quad \text{with} \quad 1 \leq i \leq N,
\]

where \(\hat{y}_{it}\) and \(y_{it}\) are the predicted and calculated the variances of daily returns/realized measures for the \(i\)th asset at the \(t\)th trading day in 2013 respectively, and \(M = 249\) is the number of trading days in 2013; see Patton and Sheppard and Patton (2011a, b). Figure 1 gives the boxplots of QLIKEs from NHEAVY, MHEAVY, NGARCH, MGARCH and RFHEAVY, and VHAR models with \(N = 12\) stocks, and we have four findings below. (i) The NHEAVY model has much better forecasting accuracy than the other competitors, indicating the importance of exploring network structures in the HEAVY model. (ii) Although the RFHEAVY model shows better performance than MHEAVY, it does not defeat the NHEAVY model. (iii) The HEAVY-type model, including NHEAVY, RFHEAVY, and MHEAVY models, show better performance than NGARCH and MGARCH model in forecasting the variance of daily return. (iv) The VHAR model is close to the NHEAVY model in forecasting future realized measures, but it cannot forecast the variances of the daily returns.

To further compare the prediction accuracy among the six models, we conduct the Diebold-Mariano (DM) test (Diebold & Mariano, 1995) as follows. We first calculate the residuals for each model:

\[
e_i = y_i - \hat{y}_i, \quad 1 \leq i \leq N,
\]

where \(y_i\) is the calculated variances of daily returns/realized measures for the \(i\)th stock with \(y_i = (y_{i,1}, \ldots, y_{i,T})^T\) and \(\hat{y}_i\) is the predicted variances of daily returns/realized measures for the \(i\)th stock with \(\hat{y}_i = (\hat{y}_{i,1}, \ldots, \hat{y}_{i,T})^T\).

---

2 To check the factors’ contribution to the dimension reduction of the HEAVY model, we consider the revised factor HEAVY in the same data-generating setting as the NHEAVY model.
Figure 1. Boxplots of QLIKEs from the NHEAVY, MHEAVY, NGARCH, MGARCH, RFHEAVY, and VHAR models for the out-of-sample variances of daily returns and daily realized measures. The left panel is for the variances of daily returns, and the right panel is for daily realized measures.

Table IV. The p-values of the less DM tests for the NHEAVY model vs. one of the MHEAVY, RFHEAVY, NGARCH, MGARCH, and VHAR models for the stock AAPL and AMD

| Model    | Variance of daily return | Daily realized measure |
|----------|--------------------------|------------------------|
| AAPL     | AMD                      |                         |
| MHEAVY   | 0.04                     | <2.20e−16              |
| RFHEAVY  | 9.99e−06                 | 0.02                   |
| NGARCH   | <2.20e−16                | 0.02                   |
| MGARCH   | <2.20e−16                | 1.89e−09               |
| VHAR     | —                        | —                      |

Then, we define

\[
d_i = e_{i}^{*2} - e_i^2 = \left( e_{i,j}^{*2} - e_{i,j}^2 \right),
\]

where \(e_{i,j}^{*}\) is the residuals from the NHEAVY and \(e_{i,j}\) is the residuals from one of MHEAVY, NGARCH, MGARCH, RFHEAVY and VHAR model. Then, we conduct hypothesis tests for

\[
H'_0 : \mathbb{E}[d_i] = 0 \quad \text{vs.} \quad H_1 : \mathbb{E}[d_i] < 0 \quad \text{(or} \ \mathbb{E}[d_i] > 0) \quad .
\]

The first alternative statement (\(\mathbb{E}[d_i] < 0\)) is to test whether the NHEAVY model is better, while the second alternative statement (\(\mathbb{E}[d_i] > 0\)) is to test whether another model is better than the NHEAVY model. We call them ‘less’ and ‘greater’ tests respectively. The p-values of the less and greater DM tests for the NHEAVY model vs. one of the MHEAVY, NGARCH, MGARCH, RFHEAVY, and VHAR models for the stock AAPL, AMD are presented in Tables IV and V respectively. The other stocks’ DM test results are shown in Tables B1–B3 in Appendix B. For the variance of the daily return, the less tests show that the p-values of the MHEAVY, NGARCH,
Table V. The $p$-values of the greater DM tests for the NHEAVY model vs. one of the MHEAVY, RFHEAVY, NGARCH, MGARCH and VHAR models for the stock AAPL and AMD

|          | AAPL   | AMD   | AAPL   | AMD   |
|----------|--------|-------|--------|-------|
| MHEAVY   | 0.96   | 1     | 1      | 1     |
| RFHEAVY  | 1      | 0.98  | 1      | 1     |
| NGARCH   | 1      | 1     | —      | —     |
| MGARCH   | 1      | 1     | —      | —     |
| VHAR     | —      | —     | 0.88   | 0.84  |

Figure 2. ACF plots for the residuals of out-of-sample variances of daily returns from the NHEAVY model. The left panel is for AAPL, and the right panel is for AMD

Figure 3. ACF plots for the residuals of out-of-sample daily realized measures from the NHEAVY model. The left panel is for AAPL, and the right panel is for AMD

MGARCH, and RFHEAVY models are very significant, and the greater tests indicate that none of the competitors has significant $p$-values. For the daily realized measures, the MHEAVY and RFHEAVY model have the significant $p$-values in the less tests, and have no significant $p$-values in the greater tests; The VHAR model has the significant $p$-values in the less tests except for the stock PGR, and has no significant $p$-values in the greater tests. From the results, we can conclude that the proposed NHEAVY model shows better prediction accuracy than the five competitors.

To check the volatility model persistency of the NHEAVY model, we study the residuals between the predicted and calculated variances of daily returns/realized measures. Their ACF plots can be drawn. The results of the stock AAPL and AMD can be presented in Figures 2 and 3 respectively. The ACF plots for the other stocks are shown in Figures C1 and C2 in Appendix C. All the ACF plots show that the residuals have no significant time series structures, indicating that the proposed NHEAVY model can explain the market dynamics.
6. CONCLUSION

This article introduces a novel network NHEAVY model, which takes the observed network structures into consideration by combining high-frequency and low-frequency data. The parameters in network HEAVY models are estimated by the quasi-likelihood function. The proposed model reduces the computational complexity substantially from $O(N^2)$ to $O(N)$. All statistical properties of estimators are presented in Section 3, and the findings are confirmed by simulation studies. We further illustrate the usefulness of our models using a real data set from the S&P 500. It shows that the network HEAVY model outperforms other volatility models such as the network GARCH, original GARCH, and HEAVY with the multivariate setting, HAR, and the revised factor HEAVY.

The proposed network HEAVY models can be extended in further directions. First, it would be interesting to add asymmetric terms to network HEAVY models to explicitly capture the leverage effects and this may improve forecast performance further. It might also be beneficial to use a long-run/short-run component model in the dynamic equations to separate our transitory movements in volatility.

ACKNOWLEDGEMENTS

We are deeply grateful to the editor, an associate editor and two anonymous referees for their valuable comments that led to the substantial improvement of the manuscript. Yuan’s research is sponsored by Shanghai Pujiang Program (No. 23PJ1402400), a Project funded by China Postdoctoral Science Foundation (No. 2023M741190), and State Key Program of National Natural Science Foundation of China (No. 71931004). Li’s research is supported in part by GRF grant 17306121 from the Hong Kong Research Grant Council and a Key Program grant 72033002 from National Natural Science Foundation of China. Wang’s research is supported in part by HK RGC Grants GRF-11301521, GRF-11311022, GRF-14306523 and CUHK Startup Grant 4937091.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Wharton Research Data Services at https://wrds-www.wharton.upenn.edu. The data that support the findings of this study are available from Wharton Research Data Services. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the author(s) with the permission of Wharton Research Data Services.

REFERENCES

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

Anderson TG, Bollerslev T, Diebold F, Labys P. 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96:42–55.

Barndorff-Nielsen OE. 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal Royal Statistical Society (B) 64:253–280.

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. 2008. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica 76:1481–1536.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31:307–327.

Bollerslev T, Wooldridge J. 2007. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11:143–172.

Bonanno G, Lillo F, Mantegna R. 2001. High-frequency cross-correlation in a set of stocks. Quantitative Finance 1:96–104.

Brida J, Matesanz D, Seihas M. 2016. Network analysis for returns and volume trading in stock markets: The euro Stoxx case. Physica A: Statistical Mechanics and its Applications 444:751–764.

Cipollini, F., Engle, R., & Gallo, G. (2007). A model for multivariate non-negative valued processes in financial econometrics. Working Papers Archive No, wp2007-16.

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

Diebold F, Mariano R. 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13:253–263.
Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* 50:987–1007.

Engle RF, Gallo GM. 2006. A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* 131:3–27.

Glyns E, Santa-Clara P, Valkanov R. 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131:59–95.

Goel S, Goldstein DG. 2014. Predicting individual behavior with social networks. *Marketing Science* 33:82–93.

Hansen PR, Huang Z, Shek HH. 2012. Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics* 27:877–906.

Huang WQ, Zhang XT, Yao S. 2009. A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and its Applications* 388:2956–2964.

Jacod J, Li Y, Mykland PA, Podolskij M, Vetter M. 2009. Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications* 119:2249–2276.

Jin X, Maheu J. 2013. Modeling realized covariances and returns. *Journal of Financial Econometrics* 11:335–369.

Kim D, Fan J. 2019. Factor GARCH-Itô models for high-frequency data with application to large volatility matrix prediction. *Journal of Econometrics* 208:395–417.

Kim D, Wang Y. 2016. Unified discrete-time and continuous-time models and statistical inferences for merged low-frequency and high-frequency financial data. *Journal of Econometrics* 194:220–230.

Kim D, Song X, Wand Y. 2022. Unified discrete-time factor stochastic volatility and continuous-time Itô models for combining inference based on low-frequency and high-frequency. *Journal of Multivariate Analysis* 192:105091.

Liu LY, Patton AJ, Sheppard K. 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics* 187:293–311.

Mantegna RN. 1999. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems* 11:193–197.

Newman MEJ. 2003. The structure and function of complex networks. *SIAM Review* 45:167–256.

Nitzan I, Libai B. 2011. Social effects on customer retention. *Journal of Marketing* 75:24–38.

Noureldin D, Shephard N, Sheppard K. 2012. Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* 119:907–933.

Patton AJ. 2011a. Data-based ranking of realised volatility estimators. *Journal of Econometrics* 161:284–303.

Patton AJ. 2011b. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160:246–256.

Reka A, Barabasi AL. 2002. Statistical mechanics of complex network. *Review of Modern Physics* 74:47–97.

Shephard N, Sheppard K. 2010. Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* 119:907–933.

Sheppard K, Xu W. 2019. Factor high-frequency-based volatility(HEAVY) models. *Journal of Financial Econometrics* 17:33–65.

Shin, M., Kim, D., Wang, Y., & Fan, J. (2021). Factor and idiosyncratic VAR-Itô volatility models for heavy-tailed high-frequency financial data. *arXiv preprint arXiv:2109.05227*.

Song X, Kim D, Yuan H, Cui X, Lu Z, Zhou Y, Wang Y. 2021. Volatility analysis with realized GARCH-Itô models. *Journal of Econometrics* 222:393–410.

Tumminello M, Lillo F, Mantegna RN. 2010. Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization* 75:40–58.

Wang Y. 2002. Asymptotic nonequivalence of GARCH models and diffusions. *The Annals of Statistics* 30:754–783.

Wei Y, Yildirim P, Christophe VDB, Chrysanthos D. 2016. Credit scoring with social network data. *Marketing Science* 35:201–340.

Xiu D. 2010. Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics* 159:235–250.

Yuan H, Sun Y, Xu L, Zhou Y, Cui X. 2022. A new volatility: GQARCH-Itô model. *Journal of Time Series Analysis* 43:345–370.

Zhang L. 2006. Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli* 12:1019–1043.

Zhang L. 2011. Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics* 160:33–47.

Zhang L, Mykland PA, Aït-Sahalia Y. 2005. A tale of two time scales; determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100:1394–1411.

Zhou J, Li D, Pan R, Wang H. 2020. Network GARCH model. *Statistica Sinica* 30:1723–1740.

Zhu X, Pan R, Li G, Liu Y, Wang H. 2017. Network vector autoregression. *The Annals of Statistics* 45:1096–1123.
A1. Proof of Theorem 1

Proof of Theorem 1. Let \( \lambda = (\lambda^{(1)}, \ldots, \lambda^{(L)})^T \) and \( \lambda_R = (\lambda^{(1)}_R, \ldots, \lambda^{(L)}_R)^T \). We first prove that the parameters \( \phi = (\omega, \alpha, \lambda^T, \beta)^T \) and \( \phi_R = (\omega_R, \alpha_R, \lambda_R^T, \beta_R)^T \) satisfy the consistency and asymptotics respectively. Then the statistical properties of the parameters \( \theta = (\phi^T, \phi_R^T)^T \) are intuitive.

We give some notations. Let \( K \) and \( \rho \) be generic constants taking many different values with \( K > 0 \) and \( \rho \in (0, 1) \).

By the expressions of \( \hat{h}_n(\phi), h_n(\phi), \hat{\mu}_n(\phi) \) and \( \mu_n(\phi_R) \) in Section 2, we have the following facts

\[
\hat{h}_n(\phi) = \sum_{k=1}^{n} 2^{-1} \left\{ \omega + a \mathbf{R} \mathbf{M}_{1,n-k} + \sum_{t=1}^{L} \lambda_t^{(e)} d_t^{(e)-1} \sum_{j \neq i} \lambda_t^{(e)} d_t^{(e)} \mathbf{R} \mathbf{M}_{j,n-k} \right\} + 2^{-1} \hat{h}_n(\phi), \quad t \geq 2,
\]

\[
h_n(\phi) = \sum_{k=1}^{n} 2^{-1} \left\{ \omega + a \mathbf{R} \mathbf{M}_{1,n-k} + \sum_{t=1}^{L} \lambda_t^{(e)} d_t^{(e)-1} \sum_{j \neq i} \lambda_t^{(e)} d_t^{(e)} \mathbf{R} \mathbf{M}_{j,n-k} \right\} + 2^{-1} h_n(\phi), \quad t \geq 2,
\]

\[
\hat{\mu}_n(\phi_R) = \sum_{k=1}^{n} 2^{-1} \left\{ \omega + a \mathbf{R} \mathbf{M}_{1,n-k} + \sum_{t=1}^{L} \lambda_t^{(e)} d_t^{(e)-1} \sum_{j \neq i} \lambda_t^{(e)} d_t^{(e)} \mathbf{R} \mathbf{M}_{j,n-k} \right\} + 2^{-1} \hat{\mu}_n(\phi_R), \quad t \geq 2,
\]

\[
\mu_n(\phi_R) = \sum_{k=1}^{n} 2^{-1} \left\{ \omega + a \mathbf{R} \mathbf{M}_{1,n-k} + \sum_{t=1}^{L} \lambda_t^{(e)} d_t^{(e)-1} \sum_{j \neq i} \lambda_t^{(e)} d_t^{(e)} \mathbf{R} \mathbf{M}_{j,n-k} \right\} + 2^{-1} \mu_n(\phi_R), \quad t \geq 2.
\]

Then, we have

\[
h_n(\phi) = \hat{h}_n(\phi) - \beta^{t-1} \hat{h}_n(\phi) + \beta^{t-1} h_n(\phi), \quad (A1)
\]

\[
\mu_n(\phi_R) = \hat{\mu}_n(\phi_R) - \beta^{t-1} \hat{\mu}_n(\phi_R) + \beta^{t-1} \mu_n(\phi_R), \quad (A2)
\]

due to \( \hat{h}_n(\phi) = T^{-1/2} \sum_{t=1}^{[T]^{1/2}} R_{n,t}^2 \), and \( \hat{\mu}_n(\phi_R) = T^{-1/2} \sum_{t=1}^{[T]^{1/2}} \mathbf{R} \mathbf{M}_{n,t} \), then

\[
\frac{\partial h_n(\phi)}{\partial \phi} = \frac{\partial \hat{h}_n(\phi)}{\partial \phi} + (t-1)\beta^{t-2}(h_n(\phi) - \hat{h}_n(\phi))e + \beta^{t-1}\frac{\partial \hat{h}_n(\phi)}{\partial \phi},
\]

\[
\frac{\partial^2 h_n(\phi)}{\partial \phi \partial \Phi^T} = \frac{\partial^2 \hat{h}_n(\phi)}{\partial \phi \partial \Phi^T} + (t-1)(t-2)\beta^{t-3}(h_n(\phi) - \hat{h}_n(\phi))ee^T + 2(t-1)\beta^{t-2}\frac{\partial \hat{h}_n(\phi)}{\partial \phi} e^T + \beta^{t-1}\frac{\partial^2 \hat{h}_n(\phi)}{\partial \phi \partial \Phi^T},
\]

\[
\frac{\partial \mu_n(\phi_R)}{\partial \phi_R} = \frac{\partial \hat{\mu}_n(\phi_R)}{\partial \phi_R} + (t-1)\beta^{t-2}(\mu_n(\phi_R) - \hat{\mu}_n(\phi_R))e + \beta^{t-1}\frac{\partial \hat{\mu}_n(\phi_R)}{\partial \phi_R},
\]

\[
\frac{\partial^2 \mu_n(\phi_R)}{\partial \phi_R \partial \Phi_R^T} = \frac{\partial^2 \hat{\mu}_n(\phi_R)}{\partial \phi_R \partial \Phi_R^T} + (t-1)(t-2)\beta^{t-3}(\mu_n(\phi_R) - \hat{\mu}_n(\phi_R))ee^T + 2(t-1)\beta^{t-2}\frac{\partial \hat{\mu}_n(\phi_R)}{\partial \phi_R} e^T + \beta^{t-1}\frac{\partial^2 \hat{\mu}_n(\phi_R)}{\partial \phi_R \partial \Phi_R^T}, \quad (A3)
\]
where \( e = (0, \ldots, 0, 1)^T \in \mathbb{R}^{L+1} \), and

\[
\begin{align*}
\sup_{\phi \in \Phi} \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \phi_k} & \leq K, \\
\sup_{\phi \in \Phi} \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \beta} & \leq K \sum_{j=1}^{\infty} \rho^j \left( \sup_{\phi \in \Phi} h_{i,j-1}(\phi) \right)^s,
\end{align*}
\tag{A4}
\]

where \( \phi_k \in \{ \omega, \alpha, \lambda^T \} \), by the inequality \( x/(1+x) \leq x' \) for \( x > 0 \) and \( s \in (0, 1] \). After the similar calculating, we have

\[
\begin{align*}
\sup_{\phi \in \Phi} \frac{1}{\mu_i(\phi_R)} \frac{\partial \mu_i(\phi_R)}{\partial \phi_k} & \leq K, \\
\sup_{\phi \in \Phi} \frac{1}{\mu_i(\phi_R)} \frac{\partial \mu_i(\phi_R)}{\partial \beta} & \leq K \sum_{j=1}^{\infty} \rho^j \left( \sup_{\phi \in \Phi} \mu_{i,j-1}(\phi_R) \right)^s,
\end{align*}
\tag{A5}
\]

where \( \phi_{R_k} \in \{ \omega_R, \alpha_R, \lambda^T_R \} \).

**Proof of Theorem 1** By a standard compactness argument, using Lemma 1 and Lemma 2, we have

\[
0 = \frac{\partial L' (\hat{\phi})}{\partial \phi} = \frac{\partial \hat{L}' (\hat{\phi}_0)}{\partial \phi} + \frac{\partial \hat{L}' (\phi^*)}{\partial \phi} (\hat{\phi} - \phi_0),
\]

where \( \phi^* \) lies in \( \hat{\phi} \) and \( \phi_0 \), thus

\[
\sqrt{NT} (\hat{\phi} - \phi_0) = -\left( \frac{\partial \hat{L}' (\phi^*)}{\partial \phi} \right)^{-1} \sqrt{NT} \frac{\partial \hat{L}' (\phi_0)}{\partial \phi} + o_p(1) \xrightarrow{d} \mathcal{N} \left( 0, (\kappa_2^r - 1) \Sigma^{-1}_r \right).
\]

Similarly, using Lemma 3 and Lemma 4, we have

\[
0 = \frac{\partial \hat{L}' (\hat{\phi}_R)}{\partial \phi_R} = \frac{\partial \hat{L}' (\phi_{R_0})}{\partial \phi_R} + \frac{\partial \hat{L}' (\phi^*)_R}{\partial \phi_R} (\hat{\phi}_R - \phi_{R_0}),
\]

where \( \phi^*_R \) lies in \( \hat{\phi}_R \) and \( \phi_{R_0} \), thus

\[
\sqrt{NT} (\hat{\phi}_R - \phi_{R_0}) = -\left( \frac{\partial \hat{L}' (\phi^*_R)}{\partial \phi_R} \right)^{-1} \sqrt{NT} \frac{\partial \hat{L}' (\phi_{R_0})}{\partial \phi_R} + o_p(1) \xrightarrow{d} \mathcal{N} \left( 0, (\kappa_2^R - 1) \Sigma^{-1}_R \right).
\]
Since $\theta = (\phi^T, \phi_R^T)^T$, by a straightforward application of quasi-likelihood theory (Cipollini et al., 2007), we have

$$\sqrt{NT}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I^{-1}(I^{-1})^T),$$

where

$$I = \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma_R \end{pmatrix}, \quad J = \begin{pmatrix} (\kappa^2_2 - 1) \Sigma & (\kappa^2_2 - 1) \Sigma_{R,R} \\ (\kappa^2_2 - 1) \Sigma_{R,T} & (\kappa^2_2 - 1) \Sigma_R \end{pmatrix}. \quad \square$$

### A2. Four useful lemmas

We provide four useful lemmas.

**Lemma 1.** If Assumptions 1 and 2 hold, then

(i) $\lim_{T \to \infty} \sup_{\phi \in \Phi} \left| L'(\phi) - \hat{L}'(\hat{\theta}) \right| = 0$ a.s.;

(ii) $E \left| \epsilon_i' \right| < \infty$, and $E \epsilon_i' \phi) \geq E \epsilon_i' \phi_0)$, where the equality holds if and only if $\phi = \phi_0$.

(iii) Any $\phi \neq \phi_0$ has a neighborhood $V(\phi)$ such that

$$\liminf_{T \to \infty} \inf_{\phi \in V(\phi)} \hat{L}'(\phi) > E \epsilon_i' \phi_0).$$

**Proof.** The proof is similar to that of theorem 1 in Zhou et al. (2020), and thus it is omitted. \quad \square

**Lemma 2.** If the conditions in Theorem 1 hold, then

(i) $\sqrt{NT} \left\| \frac{\partial L'(\phi_0)}{\partial \phi} - \frac{\partial L'(\phi)}{\partial \phi} \right\| = o(1)$ a.s.;

(ii) $\sup_{\phi \neq \phi_0} \left\| \frac{\partial^2 L'(\phi)}{\partial \phi \partial \phi} \right\| = O_p(\eta)$;

(iii) $\frac{\partial^2 L'(\phi)}{\partial \phi \partial \phi} \to \Sigma'$, where $\Sigma' = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{h_i^2(\phi_0)} \frac{\partial h_i(\phi_0)}{\partial \phi} \frac{\partial h_i(\phi_0)}{\partial \phi} \right)$;

(iv) $\sqrt{NT} \frac{\partial L'(\phi)}{\partial \phi} \xrightarrow{d} N(0, (\kappa^2_2 - 1) \Sigma')$, where $\Sigma' = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{h_i^2(\phi_0)} \frac{\partial h_i(\phi_0)}{\partial \phi} \frac{\partial h_i(\phi_0)}{\partial \phi} \right)$, and $\kappa^2_2 = E(\epsilon_i^2)$.

**Proof.** (i) By simply calculating the first derivative of $\hat{L}(\phi)$ and $L(\phi)$, we have

$$\sqrt{NT} \left\| \frac{\partial L'(\phi)}{\partial \phi} - \frac{\partial L'(\phi)}{\partial \phi} \right\| \leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \frac{1}{\hat{h}_t(\phi)} \frac{\partial \hat{h}_t(\phi)}{\partial \phi} - \frac{1}{h_t(\phi)} \frac{\partial h_t(\phi)}{\partial \phi} \right\|^2 + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \frac{r_t^2}{\hat{h}_t^2(\phi)} \frac{\partial \hat{h}_t(\phi)}{\partial \phi} - \frac{r_t^2}{h_t^2(\phi)} \frac{\partial h_t(\phi)}{\partial \phi} \right\|^2. \quad (A6)$$

The first term of the right hand in (A6) can have the following results,

$$\begin{align*}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \frac{1}{\hat{h}_t(\phi)} \frac{\partial \hat{h}_t(\phi)}{\partial \phi} - \frac{1}{h_t(\phi)} \frac{\partial h_t(\phi)}{\partial \phi} \right\|^2 & \leq \frac{K}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \frac{\partial \hat{h}_t(\phi)}{\partial \phi} - \frac{\partial h_t(\phi)}{\partial \phi} \right\|^2 \\
& + \frac{K}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \hat{h}_t(\phi) - h_t(\phi) \right\| \frac{1}{h_t(\phi)} \frac{\partial h_t(\phi)}{\partial \phi}. \quad (A7)
\end{align*}$$

wileyonlinelibrary.com/journal/jtsa © 2023 The Authors. J. Time Ser. Anal. 45: 533–557 (2024) Journal of Time Series Analysis published by John Wiley & Sons Ltd. DOI: 10.1111/jtsa.12726
By the facts in (A1) and (A2), we have
\[
\begin{align*}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left\| \frac{\partial \hat{h}_i(\phi)}{\partial \phi} - \frac{\partial h_i(\phi)}{\partial \phi} \right\| \\
\leq \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sup_{\phi \in \Phi} h_i(\phi) \left( \frac{1}{\sqrt{T}} \sum_{t=2}^{T} (t-1)\beta^{-2} \right) \\
+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sup_{\phi \in \Phi} \left\| \frac{\partial h_i(\phi)}{\partial \phi} \right\| \left( \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \beta^{-1} \right) \\
\to 0 \quad \text{a.s.}
\end{align*}
\]

Due to (A4) and Markov’s inequality, we can have
\[
\sum_{t=2}^{\infty} P \left( \beta^{-1} \sup_{\phi \in \Phi} \left\| \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \phi} \right\| > \epsilon \right) \leq \sum_{t=2}^{\infty} \frac{\beta^{-1}}{\epsilon} \sup_{\phi \in \Phi} \left\| \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \phi} \right\| < \infty,
\]

Then, the following results can be easily derived,
\[
\begin{align*}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left\| \left[ h_i(\phi) - \hat{h}_i(\phi) \right] \right\| \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \phi} \\
\leq \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sup_{h_i(\phi) - \hat{h}_i(\phi)} \left( \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \beta^{-1} \sup_{\phi \in \Phi} \left\| \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \phi} \right\| \right) \\
\to 0 \quad \text{a.s.}
\end{align*}
\]

Thus, the first term of the right hand in (A6) convergence 0 a.s. Similarly, we can prove the second term of the right hand in (A6) convergence 0 a.s. Therefore, the proof of (i) is completed.

(ii) We first show
\[
\sup_{\| \phi - \phi_0 \| < \eta} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} - \frac{\partial^2 L(\phi_0)}{\partial \beta^2} \right| = O_p(\eta).
\]

Since
\[
\sup_{\| \phi - \phi_0 \| < \eta} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{L}(\phi_0)}{\partial \beta^2} \right| \leq \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{L}(\phi_0)}{\partial \beta^2} \right| \\
+ \frac{1}{T} \sum_{t=2}^{T} \sup_{\| \phi - \phi_0 \| < \eta} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{L}(\phi_0)}{\partial \beta^2} \right|.
\]

It is note that
\[
\frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} = \frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{1}{h_i(\phi)} \frac{\partial^2 h_i(\phi)}{\partial \beta^2} - \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \beta} \frac{\partial^2 \hat{h}_i(\phi)}{\partial \beta^2} \right\} - \frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{r_i^2}{h_i(\phi)} \frac{\partial^2 h_i(\phi)}{\partial \beta^2} - 2 \frac{r_i^2}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \beta} \frac{\partial \hat{h}_i(\phi)}{\partial \beta} \right\}.
\]
Then, we have

\[
\frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{\partial^2 \hat{\epsilon}_t^2(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{\epsilon}_t^2(\phi)}{\partial \beta^2} \right| \leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{1}{h_i(\phi)} \frac{\partial^2 h_i(\phi)}{\partial \beta^2} - \frac{1}{\hat{h}_i(\phi)} \frac{\partial^2 \hat{h}_i(\phi)}{\partial \beta^2} \right|
\]

\[+ \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi) \partial h_i(\phi)}{\partial \beta} - \frac{1}{\hat{h}_i(\phi)} \frac{\partial \hat{h}_i(\phi) \partial \hat{h}_i(\phi)}{\partial \beta} \right|
\]

\[+ \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{1}{h_i(\phi)} \frac{\partial^2 h_i(\phi)}{\partial \beta^2} - \frac{1}{\hat{h}_i(\phi)} \frac{\partial^2 \hat{h}_i(\phi)}{\partial \beta^2} \right|
\]

\[+ \frac{2}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{r_{it}^2}{h_i(\phi)} \frac{\partial h_i(\phi) \partial h_i(\phi)}{\partial \beta} - \frac{r_{it}^2}{\hat{h}_i(\phi)} \frac{\partial \hat{h}_i(\phi) \partial \hat{h}_i(\phi)}{\partial \beta} \right| = I + II + III + IV.
\]

For the term I, by the facts in (A3), we have

\[
I \leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{h_i(\phi) - \hat{h}_i(\phi)}{\hat{h}_i(\phi)h_i(\phi)} \right| \frac{\partial^2 h_i(\phi)}{\partial \beta^2}
\]

\[+ \frac{K}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{\partial^2 h_i(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{h}_i(\phi)}{\partial \beta^2} \right|
\]

\[
\leq \frac{1}{N} \sum_{i=1}^{N} \sup_{\phi \in \Phi} \left| h_i(\phi) - \hat{h}_i(\phi) \right| \frac{1}{T} \sum_{t=2}^{T} \beta^{-1} \sup_{\phi \in \Phi} \left| \frac{\partial^2 h_i(\phi)}{\partial \beta^2} \right|
\]

\[+ \frac{1}{N} \sum_{i=1}^{N} \sup_{\phi \in \Phi} h_i(\phi) \left( \frac{1}{T} \sum_{t=2}^{T} (t-1)(t-2) \beta^{-3} \right)
\]

\[+ \frac{2}{N} \sum_{i=1}^{N} \sup_{\phi \in \Phi} \left| \frac{\partial h_i(\phi)}{\partial \beta} \right| \left( \frac{1}{T} \sum_{t=1}^{T} (t-1) \beta^{-2} \right)
\]

\[+ \frac{1}{N} \sum_{i=1}^{N} \sup_{\phi \in \Phi} \left| \frac{\partial^2 h_i(\phi)}{\partial \beta^2} \right| \left( \frac{1}{T} \sum_{t=2}^{T} \beta^{-1} \right) \rightarrow 0 \quad \text{as } T \rightarrow \infty. \quad (A7)
\]

Similarly, we can prove that II, III, and IV converge to 0 a.s. Thus,

\[
\frac{1}{T} \sum_{t=2}^{T} \sup_{\phi \in \Phi} \left| \frac{\partial^2 \hat{\epsilon}_t^2(\phi)}{\partial \beta^2} - \frac{\partial^2 \hat{\epsilon}_t^2(\phi)}{\partial \beta^2} \right| \rightarrow 0 \quad \text{as } T \rightarrow \infty. \quad (A8)
\]
On the other hand, by the Taylor expansion, we have
\[
E \sup_{\|\phi - \phi_0\| \leq \eta} \left| \frac{\partial^2 \epsilon_i(\phi)}{\partial \beta^2} - \frac{\partial^2 \epsilon_i(\phi_0)}{\partial \beta^2} \right| \leq \eta \sup_{\|\phi - \phi_0\| \leq \eta} \left| \frac{\partial^3 \epsilon_i(\phi)}{\partial \beta^3} \right|.
\]

By a simple calculation, it follows that
\[
\frac{\partial^3 \epsilon_i(\phi)}{\partial \beta^3} = \frac{1}{N} \sum_{j=1}^{N} \left\{ 2 - \frac{6\epsilon_{it}^2}{h_i(\phi)} \left[ \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \beta} \right]^3 \right. \\
+ \frac{1}{N} \sum_{j=1}^{N} \left\{ 1 - \frac{r_{it}^2}{h_i(\phi)} \right\} \frac{1}{h_i(\phi)} \frac{\partial^3 h_i(\phi)}{\partial \beta^3} \\
+ \frac{1}{N} \sum_{j=1}^{N} \left\{ \frac{6\epsilon_{it}^2}{h_i(\phi)} - 3 \right\} \frac{1}{h_i(\phi)} \frac{\partial h_i(\phi)}{\partial \beta} \frac{1}{h_i(\phi)} \frac{\partial^2 h_i(\phi)}{\partial \beta^2}.
\]

Similar to the proof of lemma 2 in Zhou et al. (2020), it is not hard to show that \(E \sup_{\|\phi - \phi_0\| \leq \eta} \left| \frac{\partial^3 \epsilon_i(\phi)}{\partial \beta^3} \right| = O(1)\), which together with (A8), we can prove
\[
\sup_{\|\phi - \phi_0\| \leq \eta} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \beta^2} - \frac{\partial^2 L(\phi_0)}{\partial \beta^2} \right| = O_p(\eta).
\]

Similarly, we can show that
\[
\sup_{\|\phi - \phi_0\| \leq \eta} \left| \frac{\partial^2 \hat{L}(\phi)}{\partial \phi_i \partial \phi_j} - \frac{\partial^2 L(\phi_0)}{\partial \phi_i \partial \phi_j} \right| = O_p(\eta)
\]
where \(\phi_i, \phi_j \in \{\omega, \alpha, \lambda^T\}\). Thus, (ii) holds.

(iii) By the simple calculation, we have
\[
\frac{\partial^2 \epsilon_i(\phi_0)}{\partial \phi_i \partial \phi_j} = \frac{1}{N} \sum_{j=1}^{N} \left\{ (1 - \epsilon_{it}) \frac{1}{h_i(\phi_0)} \frac{\partial^2 h_i(\phi_0)}{\partial \phi_i \partial \phi_j} \\
+ (2\epsilon_{it} - 1) \frac{1}{h_i(\phi_0)} \frac{\partial h_i(\phi_0)}{\partial \phi_i} \frac{\partial h_i(\phi_0)}{\partial \phi_j} \right\}.
\]

Since \(\epsilon_{it}\) is i.i.d with \(E\{\epsilon_{it}\} = 1\) and \(\text{var}\{\epsilon_{it}\} = \kappa_2^r\), by the strong law of large numbers, (iii) holds.

(iv) Note that
\[
\sqrt{NT} \frac{\partial L'(\phi_0)}{\partial \phi} = \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \sum_{i=1}^{N} (1 - \epsilon_{it}) \frac{1}{h_i(\phi_0)} \frac{\partial h_i(\phi_0)}{\partial \phi}.
\]

By the martingale central limit theorem, we have
\[
\sqrt{NT} \frac{\partial L'(\phi_0)}{\partial \phi} \overset{d}{\rightarrow} \mathcal{N} \left( 0, \kappa_2^r \Sigma' \right),
\]
where \(\Sigma' = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{h_i^2(\phi_0)} \frac{\partial h_i(\phi_0)}{\partial \phi_i} \frac{\partial h_i(\phi_0)}{\partial \phi^T} \right)\), and \(\kappa_2^r = E(\epsilon_{it}^2)\).
Then, we provide Lemma 3 and 4 for the quasi-likelihood functions $\hat{L}^\text{RM}$ and $L^\text{RM}$.

**Lemma 3.** If Assumptions 1 and 2 hold, then

(i) $\lim_{T \to \infty} \sup_{\phi_R \in \Phi_R} |L^\text{RM}(\phi_R) - \hat{L}^\text{RM}(\phi_R)| = 0$ a.s.

(ii) $E|\varepsilon_t^\text{RM}(\phi_R)| < \infty$, and $E\varepsilon_t^\text{RM}(\phi_R) \geq E\varepsilon_t^\text{RM}(\phi_{R0})$, where the equality holds if and only if $\phi_R = \phi_{R0}$.

(iii) Any $\phi_R \neq \phi_{R0}$ has a neighborhood $V(\phi_R)$ such that

$$\liminf_{T \to \infty} \inf_{\phi^*_R \in V(\phi_R)} \hat{L}^\text{RM}(\phi^*_R) > E\varepsilon_t^\text{RM}(\phi_{R0}).$$

**Proof.** The proof is similar to that of Lemma 1.

**Lemma 4.** If the conditions in Theorem 1 hold, then

(i) $\sqrt{NT} \left\| \frac{\partial^2 L(\phi_{R0})}{\partial \phi} - \frac{\partial^2 L(\phi_{R0})}{\partial \phi} \right\| = o(1)$ a.s.;

(ii) $\sup_{\phi_R \neq \phi_{R0}} \left\| \frac{\partial^2 L(\phi_{R0})}{\partial \phi} \right\| = O_p(\eta)$;

(iii) $\frac{\partial^2 L(\phi_{R0})}{\partial \phi} \rightarrow \Sigma^R$, where $\Sigma^R = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{\xi_i^2(\phi_{R0})} \frac{\partial \mu_i(\phi_{R0})}{\partial \phi} \frac{\partial \mu_i(\phi_{R0})}{\partial \phi} \right)$;

(iv) $\sqrt{NT} \frac{\partial^2 L(\phi_{R0})}{\partial \phi} \rightarrow N \left( 0, \left( \kappa^R - 1 \right) \Sigma^R \right)$ where $\kappa^R = E(\xi^2)$, and $\Sigma^R = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{1}{\xi_i^2(\phi_{R0})} \frac{\partial \mu_i(\phi_{R0})}{\partial \phi} \frac{\partial \mu_i(\phi_{R0})}{\partial \phi} \right)$.

**Proof.** The proof is similar in Lemma 2.

---

**APPENDIX B: THE DM TESTS FOR THE OTHER STOCKS**

Table B1. The $p$-values of the less DM tests for the NHEAVY model vs. one of the MHEAVY, RFHEAVY, NGARCH, MGARCH and VHAR models for the variances of daily returns from the stock HPQ, INTC, MSFT, MU, ORCL, AIG, MET, PGR, EBAY and F

| Stock | NHEAVY | RFHEAVY | MGARCH | NGARCH |
|-------|--------|---------|--------|--------|
| HPQ   | <2.20e−16 | 2.20e−06 | 2.11e−08 | 4.83e−07 |
| INTC  | <2.20e−16 | 2.68e−09 | <2.20e−16 | <2.20e−16 |
| MSFT  | <2.20e−16 | 2.38e−08 | <2.20e−16 | <2.20e−16 |
| MU    | <2.20e−16 | 8.76e−07 | <2.20e−16 | <2.20e−16 |
| ORCL  | <2.20e−16 | 1.05e−07 | <2.20e−16 | <2.20e−16 |
| AIG   | <2.20e−16 | 3.56e−08 | <2.20e−16 | <2.20e−16 |
| MET   | <2.20e−16 | 1.57e−08 | <2.20e−16 | <2.20e−16 |
| PGR   | <2.20e−16 | 5.46e−10 | <2.20e−16 | <2.20e−16 |
| EBAY  | <2.20e−16 | 5.62e−10 | <2.20e−16 | <2.20e−16 |
| F     | <2.20e−16 | 1.85e−08 | <2.20e−16 | <2.20e−16 |
Table B2. The *p*-values of the greater DM tests for the NHEAVY model vs. one of the MHEAVY, RFHEAVY, NGARCH, MGARCH and VHAR models for the variances of daily returns from the stock HPQ, INTC, MSFT, MU, ORCL, AIG, MET, PGR, EBAY and F.

|       | MHEAVY | RFHEAVY | MGARCH | NGARCH |
|-------|--------|---------|--------|--------|
| HPQ   | 1      | 1       | 1      | 1      |
| INTC  | 1      | 1       | 1      | 1      |
| MSFT  | 0.91   | 1       | 1      | 1      |
| MU    | 1      | 1       | 1      | 1      |
| ORCL  | 1      | 1       | 1      | 1      |
| AIG   | 1      | 1       | 1      | 1      |
| MET   | 1      | 1       | 1      | 1      |
| PGR   | 1      | 1       | 1      | 1      |
| EBAY  | 1      | 1       | 1      | 1      |
| F     | 1      | 1       | 1      | 1      |

Table B3. The *p*-values of the DM tests for the NHEAVY model vs. one of the MHEAVY, RFHEAVY, NGARCH, MGARCH and VHAR models for the daily realized measures from the stock HPQ, INTC, MSFT, MU, ORCL, AIG, MET, PGR, EBAY and F.

|       | 'less test' | 'greater test' |
|-------|-------------|----------------|
|       | MHEAVY      | RFHEAVY        | VHAR |
| HPQ   | <2.20e−16   | 6.75e−08       | 0.02 | 1    | 1    | 0.98 |
| INTC  | <2.20e−16   | 1.72e−08       | 3.06e−05 | 1 | 1 | 1 |
| MSFT  | <2.20e−16   | 6.55e−09       | 1.86e−06 | 1 | 1 | 1 |
| MU    | <2.20e−16   | 1.15e−08       | 0.02 | 1 | 1 | 0.98 |
| ORCL  | <2.20e−16   | 2.19e−07       | 0.01 | 1 | 1 | 0.99 |
| AIG   | <2.20e−16   | 1.52e−08       | 5.44e−05 | 1 | 1 | 0.99 |
| MET   | <2.20e−16   | 8.83e−09       | 6.13e−02 | 1 | 1 | 0.94 |
| PGR   | <2.20e−16   | 0.04           | 0.16 | 1 | 1 | 0.84 |
| EBAY  | <2.20e−16   | 8.57e−09       | 0.01 | 1 | 1 | 0.99 |
| F     | <2.20e−16   | 8.36e−09       | 5.38e−07 | 1 | 1 | 1 |
APPENDIX C: THE ACF PLOTS FOR THE OTHER STOCKS

Figure C1. ACF plots for the residuals of out-of-sample variances of daily return from the proposed NHEAVY model. The stocks are HPQ, INTC, MSFT, MU, ORCL, AIG, MET, PGR, EBAY and F.
Figure C2. ACF plots for the residuals of out-of-sample daily realized measures from the proposed NHEAVY model. The stocks are HPQ, INTC, MSFT, MU, ORCL, AIG, MET, PGR, EBAY and F.