

Article

A Note on the Asymptotic Normality Theory of the Least Squares Estimates in Multivariate HAR-RV Models

Won-Tak Hong ¹, Jiwon Lee ² and Eunju Hwang ^{2,*}

¹ College of International Studies, KyungHee University, Yongin 446-701, Korea; wontak@khu.ac.kr

² Department of Applied Statistics, Gachon University, Seongnam 13120, Korea; zzwon1212@gmail.com

* Correspondence: ehwang@gachon.ac.kr; Tel.: +82-31-750-5373

Received: 28 September 2020; Accepted: 18 November 2020; Published: 22 November 2020



Abstract: In this work, multivariate heterogeneous autoregressive-realized volatility (HAR-RV) models are discussed with their least squares estimations. We consider multivariate HAR models of order p with q multiple assets to explore the relationships between two or more assets' volatility. The strictly stationary solution of the HAR(p, q) model is investigated as well as the asymptotic normality theories of the least squares estimates are established in the cases of i.i.d. and correlated errors. In addition, an exponentially weighted multivariate HAR model with a common decay rate on the coefficients is discussed together with the common rate estimation. A Monte Carlo simulation is conducted to validate the estimations: sample mean and standard error of the estimates as well as empirical coverage and average length of confidence intervals are calculated. Lastly, real data of volatility of Gold spot price and S&P index are applied to the model and it is shown that the bivariate HAR model fitted by selected optimal lags and estimated coefficients is well matched with the volatility of the financial data.

Keywords: multivariate HAR models; least squares estimation; asymptotic normality; exponentially weighted HAR models

1. Introduction

Volatility in financial asset returns is one of the most important components in the financial market for optimization decisions such as portfolio selection, risk management and asset pricing. Over recent decades, extensive research works by statisticians and econometricians have analyzed volatility alongside time series modelings. In particular, because multiple assets are correlated with each other in the financial markets, the cross-correlation of two or more asset returns and the spillover effect of the volatility have been represented in multivariate time series models rather than the univariate models. Probabilistic properties such as long-memory of the volatility, which means that historical volatility has a persistent impact on future volatility, have been imposed efficiently to time series models through various statistical techniques.

In order to capture the long-memory persistence of the volatility, Corsi [1] suggested a simple but remarkably efficient time series model with additive cascades defined on different time periods by incorporating the daily, weekly and monthly volatility components. Corsi's model [1] called a heterogeneous autoregressive-realized volatility model of order 3, (HAR-RV(3) or simply HAR(3)), is a type of linear autoregression with heterogeneous regressors, which are previous lag, lagged weekly moving average, and lagged monthly moving average. Since its introduction, various adaptive versions of the HAR model are used to analyze the volatility along with empirical data analysis. Here, we refer to [2–5] for univariate data and [6–11] for multivariate data.

In particular, Bubak et al. [6] used a multivariate extension of the HAR model to uncover volatility transmission between Central European currencies and the EUR/USA foreign exchange rate, whereas Soucek and Todorova [8] employ a bivariate HAR model to explore the relationship between equity and oil market volatility. Cubadda et al. [9] proposed a vector HAR index model for detecting the presence of co-movements and analyzing the joint behavior in a set of daily realized volatility measures. Cech and Barunik [10] proposed a generalized HAR model for dynamic covariance matrix modelling and forecasting.

All of the references above empirically only dealt with the multivariate HAR models, but they have not presented theoretical results of the HAR model. Indeed, there is only a few of theoretical analysis of the HAR model, even in the univariate case. Furthermore, the least squares estimate (LSE) method has been popularly used as a tool of estimation in time series data analysis. The LSEs are used often in the literature to analyze the multivariate data in the multivariate HAR models and good performances on the volatility forecasting are obtained with forecast accuracy, for instance, see [9,12]. However, in the multivariate HAR models, rigorous formulation of asymptotic multivariate normality theory of the LSE has not yet been established in spite of the frequent uses of the LSEs. The earlier studies motivate us to consider multivariate HAR models of order p with q multiple assets, and look into theoretical analysis of the LSEs. In this work we provide, as a main result, strictly stationarity of the HAR(p, q) model and asymptotic normality theories of LSEs for the parameters in the multivariate HAR models. The multivariate normalities of the LSEs are established in two cases of i.i.d. and correlated errors of the models. In addition, we consider an exponentially weighted multivariate HAR model and discuss the estimation of a common decay rate on coefficients in the exponentially weighted HAR model.

A Monte Carlo experiment is conducted to illustrate performance of the estimations. Sample mean and standard error of the LSEs as well as empirical coverage and average length of confidence intervals are calculated. In addition, we model volatility of Gold spot price and S&P500 index using a bivariate HAR model and estimate the coefficient parameters. The parameter estimates and confidence intervals are computed according to the asymptotic normality we derived. Finally, the bivariate HAR fitting is demonstrated to the market data with optimally selected lags under some criterion rules.

The remainder of the paper is organized as follows: In Section 2, the multivariate HAR model is presented with discussion on strictly stationary solution and in Section 3 the LSEs and their multivariate normalities are derived. Section 4 deals with the exponentially weighted multivariate HAR model. A Monte-Carlo experiment is performed and an application with market data is given in Sections 5 and 6, respectively. We conclude in Section 7, where the main conclusion and discussion on future works are stated.

2. Multivariate HAR(p, q) Models

We consider the following HAR-RV model $\{Y_{1,t} : t \in \mathbb{Z}\}$ with multivariate data $\{Y_{j,t} : t \in \mathbb{Z}, j = 2, \dots, q\}$ which has been discussed in [8] as an extension of the HAR-RV(3) model in [1].

$$Y_{1,t} \equiv Y_{1,t}^{(d)} = \beta_0 + \sum_{i=d,w,m} \beta_1^{(i)} Y_{1,t-1}^{(i)} + \sum_{i=d,w,m} \beta_2^{(i)} Y_{2,t-1}^{(i)} + \dots + \sum_{i=d,w,m} \beta_q^{(i)} Y_{q,t-1}^{(i)} + \epsilon_{1,t} \quad (1)$$

where $\{\epsilon_{1,t}, t \in \mathbb{Z}\}$ are random variables for error terms and $\{\beta_0, \beta_j^{(i)}, j = 1, \dots, q; i = 1, \dots, p\}$ are coefficient parameters. For each j th-asset, $j = 1, 2, \dots, q$, three volatility components $Y_{j,t-1}^{(i)}, i = d, w, m$, corresponding to time horizons of one day (d), one week (w), one month (m) are given by the realized volatility $Y_{j,t-1}$ on day $t - 1$ if $i = d$, and

$$Y_{j,t-1}^{(i)} = \frac{1}{h_i} \left(Y_{j,t-1}^{(d)} + \dots + Y_{j,t-h_i}^{(d)} \right)$$

weakly and monthly averages of realized volatility if $i = w, m; j = 1, \dots, q$; and $h_d = 1, h_w = 5, h_m = 22$. Soucek and Todorova [8] examined the volatility linkages of oil with three considered equity markets in three separate bivariate models in the bivariate case of assets $j = 1, 2$, with $q = 2$ in model (1).

Motivated from the above, we consider a general multivariate HAR-RV(p, q) model $\{Y_{j,t} : t \in \mathbb{Z}, j = 1, \dots, q\}$:

$$Y_{j,t}^{(1)} = \beta_{j0} + \sum_{i=1}^p \beta_{j1}^{(i)} Y_{1,t-1}^{(i)} + \sum_{i=1}^p \beta_{j2}^{(i)} Y_{2,t-1}^{(i)} + \dots + \sum_{i=1}^p \beta_{jq}^{(i)} Y_{q,t-1}^{(i)} + \epsilon_{j,t} \tag{2}$$

where $\{\epsilon_{j,t}, t \in \mathbb{Z}, j = 1, \dots, q\}$ are random variables for errors with mean zero and finite variance, $\{h_i, i = 1, 2, \dots, p\}$ are assumed to satisfy $1 = h_1 < h_2 < \dots < h_p < \infty$, and $\{\beta_0, \beta_{jk}^{(i)}, j, k \in \{1, \dots, q\}; i \in \{1, \dots, p\}\}$ are parameters to be estimated. Volatility components are given in the same as above. For $i = 1, \dots, p$ and $j = 1, \dots, q, Y_{j,t-1}^{(i)} = \frac{1}{h_i} (Y_{j,t-1}^{(1)} + \dots + Y_{j,t-h_i}^{(1)})$.

Model (2) is more general than model (1) and thus we consider model (2) in this work. As main results we investigate the strict stationarity of the model and develop the estimation of parameters.

Let $\mathbb{Y}_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{q,t})^\top$ with $Y_{j,t} \equiv Y_{j,t}^{(1)}$, for $j = 1, \dots, q$. Model (2) can be written as follows:

$$Y_{j,t} = \beta_{j0} + \sum_{k=1}^{h_p} \phi_k^{(j,1)} Y_{1,t-k} + \sum_{k=1}^{h_p} \phi_k^{(j,2)} Y_{2,t-k} + \dots + \sum_{k=1}^{h_p} \phi_k^{(j,q)} Y_{q,t-k} + \epsilon_{j,t}$$

where, for $j, \ell \in \{1, 2, \dots, q\}$, (with $h_1 = 1$), $\phi_1^{(j,\ell)} = \sum_{i=1}^p \beta_{j\ell}^{(i)} / h_i, \phi_{h_1+1}^{(j,\ell)} = \dots = \phi_{h_2}^{(j,\ell)} = \sum_{i=2}^p \beta_{j\ell}^{(i)} / h_i, \phi_{h_2+1}^{(j,\ell)} = \dots = \phi_{h_3}^{(j,\ell)} = \sum_{i=3}^p \beta_{j\ell}^{(i)} / h_i, \dots, \phi_{h_{p-1}+1}^{(j,\ell)} = \dots = \phi_{h_p}^{(j,\ell)} = \beta_{j\ell}^{(p)} / h_p$. We write, for each $j = 1, 2, \dots, q, Y_{j,t} = \beta_{j0} + \sum_{k=1}^{h_p} \phi_{j,k}^\top \mathbb{Y}_{t-k} + \epsilon_{j,t}$ where $\phi_{j,k} = (\phi_k^{(j,1)}, \phi_k^{(j,2)}, \dots, \phi_k^{(j,q)})^\top$ and thus

$$\mathbb{Y}_t = \mathcal{B}_0 + \mathcal{H}_{t-1}^* + \mathcal{E}_t = \mathcal{B}_0 + \mathcal{H}\mathbb{Y}_{t-1}^* + \mathcal{E}_t$$

where $\mathcal{B}_0 = (\beta_{10}, \beta_{20}, \dots, \beta_{q0})^\top, \mathcal{E}_t = (\epsilon_{1,t}, \epsilon_{2,t}, \dots, \epsilon_{q,t})^\top$ and

$$\mathcal{H}_{t-1}^* = \begin{bmatrix} \sum_{k=1}^{h_p} \phi_{1,k}^\top \mathbb{Y}_{t-k} \\ \sum_{k=1}^{h_p} \phi_{2,k}^\top \mathbb{Y}_{t-k} \\ \vdots \\ \sum_{k=1}^{h_p} \phi_{q,k}^\top \mathbb{Y}_{t-k} \end{bmatrix} = \begin{bmatrix} \phi_{1,1}^\top & \phi_{1,2}^\top & \dots & \phi_{1,h_p}^\top \\ \phi_{2,1}^\top & \phi_{2,2}^\top & \dots & \phi_{2,h_p}^\top \\ \vdots & \vdots & & \vdots \\ \phi_{q,1}^\top & \phi_{q,2}^\top & \dots & \phi_{q,h_p}^\top \end{bmatrix} \begin{bmatrix} \mathbb{Y}_{t-1} \\ \mathbb{Y}_{t-2} \\ \vdots \\ \mathbb{Y}_{t-h_p} \end{bmatrix} =: \mathcal{H}\mathbb{Y}_{t-1}^*$$

Note that $\mathcal{H} = (\phi_{j,k}^\top), j = 1, \dots, q; k = 1, 2, \dots, h_p$, is a matrix in $\mathbb{R}^{q \times q h_p}$ and \mathbb{Y}_{t-1}^* is a column vector in $\mathbb{R}^{q h_p}$. Furthermore note that $\mathcal{H}\mathbb{Y}_{t-1}^*$ is written as

$$\mathcal{H}\mathbb{Y}_{t-1}^* = \Phi_1 \mathbb{Y}_{t-1} + \Phi_2 \mathbb{Y}_{t-2} + \dots + \Phi_{h_p} \mathbb{Y}_{t-h_p}$$

where, for $k = 1, 2, \dots, h_p$,

$$\Phi_k = \begin{bmatrix} \phi_{1,k}^\top \\ \phi_{2,k}^\top \\ \vdots \\ \phi_{q,k}^\top \end{bmatrix} = \begin{bmatrix} \phi_k^{(1,1)} & \phi_k^{(1,2)} & \dots & \phi_k^{(1,q)} \\ \phi_k^{(2,1)} & \phi_k^{(2,2)} & \dots & \phi_k^{(2,q)} \\ \vdots & \vdots & & \vdots \\ \phi_k^{(q,1)} & \phi_k^{(q,2)} & \dots & \phi_k^{(q,q)} \end{bmatrix} \in \mathbb{R}^{q \times q}$$

Therefore we have the following vector-autoregression of order h_p :

$$\mathbb{Y}_t = \mathcal{B}_0 + \Phi_1 \mathbb{Y}_{t-1} + \Phi_2 \mathbb{Y}_{t-2} + \dots + \Phi_{h_p} \mathbb{Y}_{t-h_p} + \mathcal{E}_t \tag{3}$$

which is written as

$$(I_q - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_{h_p} L^{h_p}) \mathbb{Y}_t = \mathcal{B}_0 + \mathcal{E}_t$$

with lag operator L and the identity matrix $I_q \in \mathbb{R}^{q \times q}$.

We make the following assumption:

(A1) The roots of $\det(I_q - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_{h_p} z^{h_p}) = 0$ lie outside the complex unit circle (have modulus greater than one), or equivalently, the eigenvalues of the companion matrix (of dimension $qh_p \times qh_p$)

$$\mathcal{F} = \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{h_p-1} & \Phi_{h_p} \\ I_q & O & \dots & O & O \\ O & I_q & & & O \\ \vdots & & \ddots & & \vdots \\ O & & & I_q & O \end{bmatrix}$$

have modulus less than one.

Denote $[A]_{qq}$ by the $q \times q$ submatrix of a square matrix A , consisting of the first q rows and the first q columns of A , that is $[A]_{qq} = \mathcal{I}_q A \mathcal{I}_q^\top$ where $\mathcal{I}_q = (I_q, O) \in \mathbb{R}^{q \times qh_p}$. I is the identity matrix with compatible dimension.

Theorem 1. Assume (A1) holds. Then there exists a unique strictly stationary solution with a finite first-order moment to the model and the solution has the form of

$$\mathbb{Y}_t = [(I - \mathcal{F})^{-1}]_{qq} \mathcal{B}_0 + \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} \mathcal{E}_{t-j}. \tag{4}$$

Proof of Theorem 1. From (3) we write

$$\begin{bmatrix} \mathbb{Y}_t \\ \mathbb{Y}_{t-1} \\ \mathbb{Y}_{t-2} \\ \vdots \\ \mathbb{Y}_{t-h_p-1} \end{bmatrix} = \begin{bmatrix} \mathcal{B}_0 \\ O \\ O \\ \vdots \\ O \end{bmatrix} + \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{h_p-1} & \Phi_{h_p} \\ I_q & O & \dots & O & O \\ O & I_q & & & O \\ \vdots & & \ddots & & \vdots \\ O & & & I_q & O \end{bmatrix} \begin{bmatrix} \mathbb{Y}_{t-1} \\ \mathbb{Y}_{t-2} \\ \vdots \\ \mathbb{Y}_{t-h_p-1} \\ \mathbb{Y}_{t-h_p} \end{bmatrix} + \begin{bmatrix} \mathcal{E}_t \\ O \\ O \\ \vdots \\ O \end{bmatrix}$$

which is denoted by

$$\mathbb{Y}_t^* = \mathcal{B}_0^* + \mathcal{F} \mathbb{Y}_{t-1}^* + \mathcal{E}_t^*. \tag{5}$$

Let $\lambda_i(\mathcal{F})$ be the i th eigenvalue of the matrix \mathcal{F} and $\rho := \max_i |\lambda_i(\mathcal{F})|$. Under (A1), we have $\rho < 1$. Furthermore it follows that $\lim_{m \rightarrow \infty} \sum_{j=1}^m \mathcal{F}^j \mathcal{E}_{t-j}^*$ exists and $\mathbb{Y}_t^* = (I - \mathcal{F})^{-1} \mathcal{B}_0^* + \sum_{j=0}^{\infty} \mathcal{F}^j \mathcal{E}_{t-j}^*$ exists. Thus $\{\mathbb{Y}_t^*\}$ is strictly stationary and satisfies (5). It can be given straightforwardly.

Let $[A]_q$ denote $q \times 1$ column vector consisting of the first q components of a column vector A . Since $[\mathbb{Y}_t^*]_q = \mathbb{Y}_t$, $[\mathcal{B}_0^*]_q = \mathcal{B}_0$ and $[\mathcal{E}_t^*]_q = \mathcal{E}_t$, $\{\mathbb{Y}_t\}$ is also strictly stationary and

$$\mathbb{Y}_t = [\mathbb{Y}_t^*]_q = [(I - \mathcal{F})^{-1} \mathcal{B}_0^*]_q + \left[\sum_{j=0}^{\infty} \mathcal{F}^j \mathcal{E}_{t-j}^* \right]_q = [(I - \mathcal{F})^{-1}]_{qq} \mathcal{B}_0 + \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} \mathcal{E}_{t-j}$$

that is the desired result in (4). Now we show the uniqueness of the strictly stationary solution (4). Assume that $\{\check{\mathbb{Y}}_t\}$ is another strictly stationary solution of the vector autoregression (3) with $E \|\check{\mathbb{Y}}_t\| < \infty$. Then we have $\check{\mathbb{Y}}_t^* = \mathcal{B}_0^* + \mathcal{F} \check{\mathbb{Y}}_{t-1}^* + \mathcal{E}_t^*$ in the same way as above in (5). For any positive integer m we have $\check{\mathbb{Y}}_t^* = \sum_{j=0}^{m-1} \mathcal{F}^j (\mathcal{B}_0^* + \mathcal{E}_{t-j}^*) + \mathcal{F}^m \check{\mathbb{Y}}_{t-m}^*$, and we observe

$$E \|\mathbb{Y}_t^* - \check{\mathbb{Y}}_t^*\| = E \left\| \sum_{j=m}^{\infty} \mathcal{F}^j (\mathcal{B}_0^* + \mathcal{E}_{t-j}^*) - \mathcal{F}^m \check{\mathbb{Y}}_{t-m}^* \right\| \leq C \rho^m$$

where C is a constant independent of t and m . Since m is arbitrary, we see $E\|\mathbb{Y}_t - \check{\mathbb{Y}}_t\| \leq E\|\mathbb{Y}_t^* - \check{\mathbb{Y}}_t^*\| = 0$ and thus $\mathbb{Y}_t = \check{\mathbb{Y}}_t$ almost surely. This completes the uniqueness. \square

3. The Least Squares Estimation

We consider two cases on the error processes to discuss the least squares estimates : In the first case, we assume that $\{\epsilon_{j,t}\}$ are i.i.d. random variables with mean zero and variance σ^2 for all $j = 1, 2, \dots, q$, that is, $E[\mathcal{E}_t \mathcal{E}_t^\top] = \sigma^2 I_q$. As a second but more general case, we adopt correlated error processes with $E[\mathcal{E}_t \mathcal{E}_t^\top] = (\sigma_{jl}) \in \mathbb{R}^{q \times q}$ where $\sigma_{jl} = Cov(\epsilon_{j,t}, \epsilon_{l,t}) \neq 0$.

First, the case of i.i.d. error processes are stated on assumption as follows and the ordinary least squares estimator (OLSE) of the parameters in model (2) is investigated.

(A2) $\{\epsilon_{j,t}\}$ are i.i.d. random variables with mean zero and variance σ^2 for all $j = 1, 2, \dots, q$, that is, $E[\mathcal{E}_t \mathcal{E}_t^\top] = \sigma^2 I_q$. Under (A1) and (A2), let $\mu = (\mu_1, \dots, \mu_q)^\top = E[\mathbb{Y}_t]$ with $\mu_j = E[Y_{j,t}]$ and $\Gamma(\ell) = Cov(\mathbb{Y}_t, \mathbb{Y}_{t-\ell})$, lag- ℓ autocovariance matrix function. Note that, by using (4),

$$\mu = [(I - \mathcal{F})^{-1}]_{qq} \mathcal{B}_0 = (I_q - \Phi_1 - \Phi_2 - \dots - \Phi_{h_p})^{-1} \mathcal{B}_0, \quad \Gamma(0) = \sigma^2 \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} [\mathcal{F}^j]_{qq}^\top$$

$$\Gamma(\ell) = \sigma^2 \sum_{j=0}^{\infty} [\mathcal{F}^{j+\ell}]_{qq} [\mathcal{F}^j]_{qq}^\top \text{ for } \ell > 0, \quad \Gamma(\ell) = \sigma^2 \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} [\mathcal{F}^{j-\ell}]_{qq}^\top \text{ for } \ell < 0.$$

Suppose that data $\{\mathbb{Y}_{-h_p+1}, \dots, \mathbb{Y}_{-1}, \mathbb{Y}_0, \mathbb{Y}_1, \dots, \mathbb{Y}_n\}$ are observed where $\mathbb{Y}_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{q,t})^\top$ with $Y_{j,t} \equiv Y_{j,t}^{(1)}$ for $j = 1, \dots, q$.

Let

$$\beta_j = (\beta_{j0}, \beta_{j1}^{(1)}, \dots, \beta_{j1}^{(p)}, \beta_{j2}^{(1)}, \dots, \beta_{j2}^{(p)}, \dots, \beta_{jq}^{(1)}, \dots, \beta_{jq}^{(p)})^\top \in \mathbb{R}^{(1+pq)}$$

$$\mathbb{X}_{t-1} = (1, Y_{1,t-1}^{(1)}, \dots, Y_{1,t-1}^{(p)}, Y_{2,t-1}^{(1)}, \dots, Y_{2,t-1}^{(p)}, \dots, Y_{q,t-1}^{(1)}, \dots, Y_{q,t-1}^{(p)})^\top \in \mathbb{R}^{(1+pq)}$$

$$= (1, \bar{\mathbb{Y}}_{1,t-1}^\top, \bar{\mathbb{Y}}_{2,t-1}^\top, \dots, \bar{\mathbb{Y}}_{q,t-1}^\top)^\top \text{ with } \bar{\mathbb{Y}}_{j,t-1} = (Y_{j,t-1}^{(1)}, \dots, Y_{j,t-1}^{(p)})^\top.$$

Then Equation (2) is, for each j , given by

$$Y_{j,t} = \beta_j^\top \mathbb{X}_{t-1} + \epsilon_{j,t} = \mathbb{X}_{t-1}^\top \beta_j + \epsilon_{j,t} \tag{6}$$

and its matrix form is

$$\mathbb{Y}_t = \mathcal{B} \mathbb{X}_{t-1} + \mathcal{E}_t, \quad \mathbb{Y}_t^\top = \mathbb{X}_{t-1}^\top \mathcal{B}^\top + \mathcal{E}_t^\top, \tag{7}$$

where $\mathcal{B} = (\beta_1, \beta_2, \dots, \beta_q)^\top$, $q \times (1 + pq)$ matrix, and $\mathcal{E}_t = (\epsilon_{1,t}, \epsilon_{2,t}, \dots, \epsilon_{q,t})^\top$. The matrix form with $t = 1, 2, \dots, n$ is given by

$$\mathbb{Y} = \mathbb{X} \mathcal{B}^\top + \mathcal{E}$$

where $\mathbb{Y} = (\mathbb{Y}_1, \dots, \mathbb{Y}_n)^\top \in \mathbb{R}^{n \times q}$, $\mathbb{X} = (\mathbb{X}_0, \mathbb{X}_1, \dots, \mathbb{X}_{n-1})^\top \in \mathbb{R}^{n \times (1+pq)}$ and $\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_n)^\top \in \mathbb{R}^{n \times q}$. The ordinary least squares estimator (OLSE) of \mathcal{B} is obtained by

$$\hat{\mathcal{B}} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_q)^\top = \arg \min_{(\beta_1, \beta_2, \dots, \beta_q)} \sum_{j=1}^q \sum_{t=1}^n \epsilon_{j,t}^2.$$

Its transpose is given by

$$\hat{\mathcal{B}}^\top = (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \mathbb{Y}) = \left(\sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top \right)^{-1} \left(\sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{Y}_t \right).$$

Equivalently, from (6) with $Y_{j,t} = \mathbb{X}_{t-1}^\top \beta_j + \epsilon_{j,t}$, whose matrix form with $t = 1, 2, \dots, n$, is given by

$$\begin{aligned} \tilde{Y}_j &:= (Y_{j,1}, Y_{j,2}, \dots, Y_{j,n})^\top = (\mathbb{X}_0, \mathbb{X}_1, \dots, \mathbb{X}_{n-1})^\top \beta_j + \tilde{\epsilon}_j \\ \tilde{Y}_j &= \mathbb{X} \beta_j + \tilde{\epsilon}_j \end{aligned}$$

where $\tilde{\epsilon}_j = (\epsilon_{j,1}, \epsilon_{j,2}, \dots, \epsilon_{j,n})^\top$, we obtain OLSE of β_j for each j as follows.

$$\hat{\beta}_j = (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \tilde{Y}_j) = \left(\sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top \right)^{-1} \left(\sum_{t=1}^n \mathbb{X}_{t-1} Y_{j,t} \right).$$

The following theorem states the asymptotic normality of the OLSE. We need the notations for the theorem:

Under the strictly stationary condition in (A1), for $j, l \in \{1, \dots, q\}$, let $\mu_j = E[Y_{j,t}]$ and

$$\gamma_{jl}(k_1, k_2) = Cov(Y_{j,t-k_1}, Y_{l,t-k_2}), \quad \tilde{\gamma}_{jl}(i_1, i_2) = \frac{1}{h_{i_1} h_{i_2}} \sum_{k_1=1}^{h_{i_1}} \sum_{k_2=1}^{h_{i_2}} \gamma_{jl}(k_1, k_2).$$

Furthermore let J_p be $p \times p$ matrix with all components as ones and $vec(\cdot)$ the usual operator, stacking the columns of a given matrix.

Theorem 2. Assume that (A1) and (A2) holds, and the matrix $\mathbb{X}^\top \mathbb{X}$ is of full rank. (a) For each $j = 1, 2, \dots, q$, as $n \rightarrow \infty$ we have

$$\sqrt{n}(\hat{\beta}_j - \beta_j) \xrightarrow{d} N(0, \sigma^2 \Sigma^{-1})$$

where Σ is given by

$$\Sigma = E[\mathbb{X}_0 \mathbb{X}_0^\top] = \begin{bmatrix} 1 & S_{01} & \cdots & S_{0q} \\ S_{10} & S_{11} & \cdots & S_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ S_{q0} & S_{q1} & \cdots & S_{qq} \end{bmatrix} \in \mathbb{R}^{(1+pq) \times (1+pq)}$$

with $S_{j0} = (\mu_1, \dots, \mu_l)^\top = S_{0j}^\top \in \mathbb{R}^p$ and

$$S_{jl} = \begin{bmatrix} \tilde{\gamma}_{jl}(1, 1) & \cdots & \tilde{\gamma}_{jl}(1, p) \\ \vdots & \ddots & \vdots \\ \tilde{\gamma}_{jl}(p, 1) & \cdots & \tilde{\gamma}_{jl}(p, p) \end{bmatrix} + \mu_j \mu_l J_p \in \mathbb{R}^{p \times p}.$$

(b) As $n \rightarrow \infty$ we have

$$\frac{\sqrt{n}}{\sigma} [vec(\hat{\mathcal{B}}) - vec(\mathcal{B})] \xrightarrow{d} N(0, I_q \otimes \Sigma^{-1})$$

where \otimes is the Kronecker product.

Proof of Theorem 2. For (a), $\hat{\beta}_j$ is written as

$$\hat{\beta}_j = \beta_j + \left(\sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top \right)^{-1} \left(\sum_{t=1}^n \mathbb{X}_{t-1} \epsilon_{j,t} \right), \quad \sqrt{n}(\hat{\beta}_j - \beta_j) = \sqrt{n} \hat{\Sigma}_x^{-1} \hat{\Sigma}_{xe,j}$$

where

$$\hat{\Sigma}_x = \frac{1}{n} \sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top, \quad \hat{\Sigma}_{x_e, j} = \frac{1}{n} \sum_{t=1}^n \mathbb{X}_{t-1} \epsilon_{j,t}. \tag{8}$$

For the desired result, it suffices to show that as $n \rightarrow \infty$,

$$\hat{\Sigma}_x \xrightarrow{P} \Sigma \quad \text{and} \quad \sqrt{n} \hat{\Sigma}_{x_e, j} \xrightarrow{d} N(0, \sigma^2 \Sigma).$$

First, we prove the convergence of $\hat{\Sigma}_x$.

$$\hat{\Sigma}_x = \frac{1}{n} \sum_{t=1}^n \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top = \frac{1}{n} \sum_{t=1}^n \begin{bmatrix} 1 & \hat{\Sigma}_{12,t} \\ \hat{\Sigma}_{21,t} & \hat{\Sigma}_{22,t} \end{bmatrix} \tag{9}$$

where

$$\hat{\Sigma}_{12,t} = (\bar{Y}_{1,t-1}^\top, \bar{Y}_{2,t-1}^\top, \dots, \bar{Y}_{q,t-1}^\top) = \hat{\Sigma}_{21,t}^\top, \quad 1 \times pq \text{ row vectors}$$

$$\hat{\Sigma}_{22,t} = \hat{\Sigma}_{21,t} \hat{\Sigma}_{12,t}, \quad pq \times pq \text{ matrix.}$$

It can be shown straightforwardly that the right-hand side of (9) converges to Σ as $n \rightarrow \infty$ by the WLLN of the stationary sequences.

Second, we verify the convergence of $\sqrt{n} \hat{\Sigma}_{x_e, j} \left(= \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbb{X}_{t-1} \epsilon_{j,t} \right)$ in distribution. For this purpose, it is enough to show that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \eta^\top \mathbb{X}_{t-1} \epsilon_{j,t} \xrightarrow{d} N(0, \sigma^2 \eta^\top \Sigma \eta) \tag{10}$$

for any $\eta \in \mathbb{R}^{1+pq}$. Fix j and let $\zeta_t = \frac{1}{\sqrt{n}} \eta^\top \mathbb{X}_{t-1} \epsilon_{j,t}$ and $\mathcal{F}_t^* = \sigma\{\epsilon_{j,s}, -\infty < s \leq t\}$. Let $\mathbb{M}_t = \sum_{s=1}^t \zeta_s$. Note that $E[\mathbb{M}_t | \mathcal{F}_{t-1}^*] = E[\sum_{s=1}^t \zeta_s | \mathcal{F}_{t-1}^*] = E[\zeta_t + \sum_{s=1}^{t-1} \zeta_s | \mathcal{F}_{t-1}^*] = \mathbb{M}_{t-1}$ and thus $\{\mathbb{M}_t\}$ is a martingale sequence with respect to $\{\mathcal{F}_t^*\}$.

Similar to the proof of the convergence of $\hat{\Sigma}_x$, we can show that $E[(\eta^\top \mathbb{X}_{t-1})(\mathbb{X}_{t-1}^\top \eta)]^2 < \infty$, which implies that for any $\delta > 0$

$$\sum_{t=1}^n E[\zeta_t^2 \mathbb{I}(|\zeta_t| > \delta) | \mathcal{F}_{t-1}^*] \leq \frac{1}{\delta^2} \sum_{t=1}^n E[\zeta_t^4 | \mathcal{F}_{t-1}^*] \leq \frac{C}{n^2 \delta^2} \sum_{t=1}^n (\eta^\top \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top \eta)^2 \xrightarrow{P} 0$$

for some constant $C > 0$. Furthermore $\sum_{t=1}^n E[\zeta_t^2 | \mathcal{F}_{t-1}^*] = \frac{\sigma^2}{n} \sum_{t=1}^n \eta^\top \mathbb{X}_{t-1} \mathbb{X}_{t-1}^\top \eta \xrightarrow{P} \sigma^2 \eta^\top \Sigma \eta$ as $n \rightarrow \infty$. Hence, by the central limit theorem for martingale difference sequences (see [13]), the desired convergence of \mathbb{M}_n to $N(0, \sigma^2 \eta^\top \Sigma \eta)$ in (10) in distribution holds. We complete the proof of (a).

For (b), we observe covariance matrix $Cov(\sqrt{n}(\hat{\beta}_j - \beta_j), \sqrt{n}(\hat{\beta}_l - \beta_l))$ for $j \neq l$, which is equal to

$$nE \left[(\hat{\beta}_j - \beta_j)(\hat{\beta}_l - \beta_l)^\top \right] = nE \left[E \left[(\hat{\beta}_j - \beta_j)(\hat{\beta}_l - \beta_l)^\top | \mathbb{Y}_t, t = -h_p + 1, \dots, n \right] \right]$$

By (8), $nE \left[(\hat{\beta}_j - \beta_j)(\hat{\beta}_l - \beta_l)^\top | \mathbb{Y}_t, t = -h_p + 1, \dots, n \right] =$

$$nE \left[\hat{\Sigma}_x^{-1} \hat{\Sigma}_{x_e, j} (\hat{\Sigma}_x^{-1} \hat{\Sigma}_{x_e, l})^\top | \mathbb{Y}_t, t = -h_p + 1, \dots, n \right] = \hat{\Sigma}_x^{-1} \left(\frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n \mathbb{X}_{t-1} E[\epsilon_{j,t} \epsilon_{l,s}] \mathbb{X}_{s-1}^\top \right) \hat{\Sigma}_x^{-1}$$

which is zero because $\{\epsilon_{j,t}, j = 1, 2, \dots, q\}$ are independent. Thus the covariance matrix is zero, and therefore the desired asymptotic multivariate normality of $\text{vec}(\hat{\beta})$ is obtained with the covariance matrix $I_q \otimes \Sigma^{-1}$. \square

Now we adopt correlated error processes by assuming the following:

(A3) $\{\epsilon_{j,t}, j = 1, 2, \dots, q, t \in \mathbb{Z}\}$ are correlated random variables with mean zero and $Cov(\epsilon_{j,t}, \epsilon_{l,t}) = \sigma_{jl} \neq 0$, that is, the covariance matrix Ψ is assumed to be $\Psi := E[\mathcal{E}_t \mathcal{E}_t^\top] = (\sigma_{jl}) \in \mathbb{R}^{q \times q}$.

The covariance matrix Ψ is nonsingular and positive definite so there exists a nonsingular symmetric matrix $\Psi^{\frac{1}{2}} \in \mathbb{R}^{q \times q}$ and $\Psi^{\frac{1}{2}} \Psi^{\frac{1}{2}} = \Psi$.

Under (A1) and (A3), note that lag- ℓ autocovariance matrix functions $\Gamma(\ell)$ of \mathbb{Y}_t are given by $\Gamma(0) = \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} \Psi [\mathcal{F}^j]_{qq}^\top$ and

$$\Gamma(\ell) = \sum_{j=0}^{\infty} [\mathcal{F}^{j+\ell}]_{qq} \Psi [\mathcal{F}^j]_{qq}^\top \text{ for } \ell > 0, \quad \Gamma(\ell) = \sum_{j=0}^{\infty} [\mathcal{F}^j]_{qq} \Psi [\mathcal{F}^{j-\ell}]_{qq}^\top \text{ for } \ell < 0.$$

In this case the generalized least squares estimator (GLSE) is computed by minimizing the sum of squared standardized errors. The GLSE $\hat{\mathcal{B}}_{GLS} = (\hat{\beta}_{1, GLS}, \dots, \hat{\beta}_{q, GLS})^\top$ of \mathcal{B} is given as follows: From (7) we have

$$\Psi^{-\frac{1}{2}} \mathbb{Y}_t = \Psi^{-\frac{1}{2}} \mathcal{B} \mathbb{X}_{t-1} + \Psi^{-\frac{1}{2}} \mathcal{E}_t, \quad \mathbb{Y}_t^\top \Psi^{-\frac{1}{2}} = \mathbb{X}_{t-1}^\top \mathcal{B}^\top \Psi^{-\frac{1}{2}} + \mathcal{E}_t^\top \Psi^{-\frac{1}{2}}.$$

Let $\mathbb{Y}_{t, \Psi} = \Psi^{-\frac{1}{2}} \mathbb{Y}_t$, $\mathcal{B}_\Psi = \Psi^{-\frac{1}{2}} \mathcal{B}$ and $\mathcal{U}_t = \Psi^{-\frac{1}{2}} \mathcal{E}_t$. Note that $E[\mathcal{U}_t \mathcal{U}_t^\top] = I_q$ and we have

$$\mathbb{Y}_{t, \Psi} = \mathcal{B}_\Psi \mathbb{X}_{t-1} + \mathcal{U}_t, \quad \mathbb{Y}_{t, \Psi}^\top = \mathbb{X}_{t-1}^\top \mathcal{B}_\Psi^\top + \mathcal{U}_t^\top. \tag{11}$$

The GLSE is obtained by

$$\hat{\mathcal{B}}_{GLS} = \arg \min_{\mathcal{B}} \sum_{t=1}^n \mathcal{U}_t^\top \mathcal{U}_t = \arg \min_{\mathcal{B}} \sum_{t=1}^n (\mathbb{Y}_t - \mathcal{B} \mathbb{X}_{t-1})^\top \Psi^{-1} (\mathbb{Y}_t - \mathcal{B} \mathbb{X}_{t-1}).$$

The matrix form of (11) with $t = 1, 2, \dots, n$ is given as $\mathbb{Y}_\Psi = \mathbb{X} \mathcal{B}_\Psi^\top + \mathcal{U}$ where

$$\mathbb{Y}_\Psi = (\mathbb{Y}_{1, \Psi}, \dots, \mathbb{Y}_{n, \Psi})^\top \in \mathbb{R}^{n \times q}, \quad \mathbb{X} = (\mathbb{X}_0, \mathbb{X}_1, \dots, \mathbb{X}_{n-1})^\top \in \mathbb{R}^{n \times (1+pq)}, \quad \mathcal{U} = (\mathcal{U}_1, \dots, \mathcal{U}_n)^\top \in \mathbb{R}^{n \times q}.$$

The least squares estimator of \mathcal{B}_Ψ is of the form $\hat{\mathcal{B}}_\Psi^\top = (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \mathbb{Y}_\Psi)$, which is the estimator of $(\Psi^{-\frac{1}{2}} \mathcal{B})^\top = \mathcal{B}^\top \Psi^{-\frac{1}{2}}$. Furthermore note that $\mathbb{Y}_\Psi = \mathbb{Y} \Psi^{-\frac{1}{2}}$. Hence the GLSE of \mathcal{B} is given as

$$\hat{\mathcal{B}}_{GLS}^\top = (\mathbb{X}^\top \mathbb{X})^{-1} (\mathbb{X}^\top \mathbb{Y})$$

that is of the same form as the OLSE. The following theorem presents the asymptotic normality of the GLSE.

Theorem 3. Assume that (A1) and (A3) hold, and the matrix $\mathbb{X}^\top \mathbb{X}$ is of full rank.

(a) For each $j = 1, 2, \dots, q$, as $n \rightarrow \infty$ we have

$$\sqrt{n}(\hat{\beta}_{j, GLS} - \beta_j) \xrightarrow{d} N(0, \sigma_j^2 \Sigma^{-1})$$

where Σ is given as in Theorem 2 and $\sigma_j^2 = \sigma_{jj}$.

(b) As $n \rightarrow \infty$ we have

$$\sqrt{n}[\text{vec}(\hat{\mathcal{B}}_{GLS}) - \text{vec}(\mathcal{B})] \xrightarrow{d} N(0, \Psi \otimes \Sigma^{-1}).$$

Theorem 3 includes the following case of the uncorrelated but heterogeneous variance of the error processes with $Cov(\epsilon_{j,t}, \epsilon_{l,t}) = \sigma_j^2$ if $j = l$, and zero if $j \neq l$.

Corollary 1. Assume (A1) and if $E[\mathcal{E}_t] = O$ and $E[\mathcal{E}_t \mathcal{E}_t^\top] = \text{diag}(\sigma_j^2) \in \mathbb{R}^{q \times q}$, then as $n \rightarrow \infty$, $\sqrt{n}(\hat{\beta}_{j,GLS} - \beta_j) \xrightarrow{d} N(0, \sigma_j^2 \Sigma^{-1})$ and $\sqrt{n}[\text{vec}(\hat{\mathcal{B}}_{GLS}) - \text{vec}(\mathcal{B})] \xrightarrow{d} N(0, \Sigma_{diag})$ where Σ_{diag} is given by

$$\Sigma_{diag} = \begin{bmatrix} \sigma_1^2 \Sigma^{-1} & O & \dots & O \\ O & \sigma_2^2 \Sigma^{-1} & \dots & O \\ \vdots & & \ddots & \\ O & \dots & O & \sigma_q^2 \Sigma^{-1} \end{bmatrix}.$$

Proof of Theorem 3. The proof is similar to that of the Theorem 2, except for the limit of the covariance matrix $Cov(\sqrt{n}(\hat{\beta}_{j,GLS} - \beta_j), \sqrt{n}(\hat{\beta}_{l,GLS} - \beta_l))$ for $j \neq l$. Following what we've done in the proof of Theorem 2, we get $nE[(\hat{\beta}_{j,GLS} - \beta_j)(\hat{\beta}_{l,GLS} - \beta_l)^\top | \mathbb{Y}_t, t = -h_p + 1, \dots, n] =$

$$nE \left[\hat{\Sigma}_x^{-1} \hat{\Sigma}_{xe,j} (\hat{\Sigma}_x^{-1} \hat{\Sigma}_{xe,l})^\top | \mathbb{Y}_t, t = -h_p + 1, \dots, n \right] = \hat{\Sigma}_x^{-1} \left(\frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n \mathbb{X}_{t-1} E[\epsilon_{j,t} \epsilon_{l,s}] \mathbb{X}_{s-1}^\top \right) \hat{\Sigma}_x^{-1}$$

which is equal to $\hat{\Sigma}_x^{-1} \left(\frac{1}{n} \sum_{t=1}^n \mathbb{X}_{t-1} \sigma_{jl} \mathbb{X}_{t-1}^\top \right) \hat{\Sigma}_x^{-1} = \sigma_{jl} \hat{\Sigma}_x^{-1} \xrightarrow{P} \sigma_{jl} \Sigma^{-1}$. Thus the desired asymptotic multivariate normality of $\text{vec}(\hat{\mathcal{B}}_{GLS})$ is completed with the covariance matrix $\Psi \otimes \Sigma^{-1}$. \square

We have discussed parameter estimation by means of the least squared method and have established the multivariate normalities in multivariate HAR(p, q) models. For a univariate HAR model, Hwang and Shin [3] proposed infinite-order, long-memory HAR model to capture the long-memory property and studied the asymptotic theories of the LSE. In the work of [3,14], it was assumed that HAR coefficients decrease exponentially, and it was shown that, under the exponential decay condition, the autocorrelation function of the HAR model is algebraically decreasing and thus the model is of long-memory. For this reason, additionally we consider the exponentially weighted multivariate HAR model with exponential decay rate, and develop its asymptotic normality of the rate estimator. As a simple case we assume that multiple assets have a common rate on coefficients, and the rate estimation problem is investigated in the following section. A general case of the exponentially weighted multivariate HAR model will be dealt in the future study.

4. Exponentially Weighted Multivariate HAR Model

In this section, as a special case of model (2), we consider an exponentially weighted multivariate HAR model with common rate on the parameters of common regressors as follows: In model (2), for $j, k \in \{1, 2, \dots, q\}$, and for $i \in \{1, 2, \dots, p\}$, $\beta_{jk}^{(i)} = c_{jk} \lambda^{i-1}$ for some c_{jk} and $|\lambda| < 1$. The parameters c_{jk} and λ are estimated using the LSE $\hat{\beta}_{jk}^{(i)}$ in Section 3 as follows:

$$\hat{c}_{jk} = \hat{\beta}_{jk}^{(1)} \quad \text{and} \quad \hat{\lambda}_n = \frac{\sum_{k=1}^q \sum_{j=1}^q \hat{\beta}_{jk}^{(2)}}{\sum_{k=1}^q \sum_{j=1}^q \hat{\beta}_{jk}^{(1)}}.$$

We write $\Sigma^{-1} = (q_{l_1, l_2}^{(-1)})$, which is given in the asymptotic variances in Theorems 2 and 3, for $l_1, l_2 \in \{1, 2, \dots, 1 + pq\}$, that is, $q_{l_1, l_2}^{(-1)}$ is (l_1, l_2) -component of Σ^{-1} . By Theorem 2 for each i, j, k , we have $\sqrt{n}(\hat{\beta}_{jk}^{(i)} - \beta_{jk}^{(i)}) \xrightarrow{d} N(0, q_{jk}^{(i)})$ for some $q_{jk}^{(i)}$ which is the corresponding component of the matrix $\sigma^2 \Sigma^{-1}$ in Theorem 2. Indeed, $q_{jk}^{(i)}$ is the same variance for all j , say, $q_k^{(i)}$, and we can easily represent $q_k^{(i)} = \sigma^2 q_{l,l}^{(-1)}$ with $l = (k - 1)p + 1 + i$. The following theorem states the asymptotic normality of the estimates for the common rate.

Theorem 4. Assume (A1) and (A2) hold. In model (2) with $\beta_{jk}^{(i)} = c_{jk}\lambda^{i-1}$, as $n \rightarrow \infty$, we have

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow{d} N(0, v^2)$$

where $v^2 = \text{Var}(\sum_{k=1}^q \sum_{j=1}^q Z_{jk}^*) / C^2$ with $C = \sum_{k=1}^q \sum_{j=1}^q c_{jk}$ and Z_{jk}^* following normal distribution with mean zero and variance

$$v_{jk}^* = \sigma^2 \left[\varrho_{(k-1)p+3, (k-1)p+3}^{(-1)} + \lambda^2 \varrho_{(k-1)p+2, (k-1)p+2}^{(-1)} - 2\lambda \varrho_{(k-1)p+2, (k-1)p+3}^{(-1)} \right].$$

Proof of Theorem 4. By Theorem 2 for each i, j, k , we have $\sqrt{n}(\hat{\beta}_{jk}^{(i)} - \beta_{jk}^{(i)}) \xrightarrow{d} N(0, \varrho_{jk}^{(i)})$ for some $\varrho_{jk}^{(i)}$ which is the corresponding component of the matrix $\sigma^2 \Sigma^{-1}$ in Theorem 2, where $\varrho_{jk}^{(i)} = \varrho_k^{(i)} = \sigma^2 \varrho_{(k-1)p+1+i, (k-1)p+1+i}^{(-1)}$. We may write

$$\sum_{k=1}^q \sum_{j=1}^q \hat{\beta}_{jk}^{(i)} = C\lambda^{i-1} + \frac{1}{\sqrt{n}} \sum_{k=1}^q \sum_{j=1}^q Z_{jk}^{(i)} + o_p(1/\sqrt{n})$$

where $Z_{jk}^{(i)}$ s are normal random variables with mean zero and variance $\varrho_k^{(i)}$ for each j and we have

$$\hat{\lambda}_n = \frac{\lambda + \frac{1}{C\sqrt{n}} \sum_{k=1}^q \sum_{j=1}^q Z_{jk}^{(2)} + o_p(1/\sqrt{n})}{1 + \frac{1}{C\sqrt{n}} \sum_{k=1}^q \sum_{j=1}^q Z_{jk}^{(1)} + o_p(1/\sqrt{n})},$$

$$\sqrt{n}(\hat{\lambda}_n - \lambda) = \frac{1}{C} \sum_{k=1}^q \sum_{j=1}^q (Z_{jk}^{(2)} - \lambda Z_{jk}^{(1)}) + O_p(1/\sqrt{n}).$$

Note that $Z_{jk}^{(2)} - \lambda Z_{jk}^{(1)}$ has asymptotically normal distribution with mean zero and variance $\varrho_k^{(2)} + \lambda^2 \varrho_k^{(1)} - 2\lambda \text{Cov}(Z_{jk}^{(2)}, Z_{jk}^{(1)}) = \sigma^2 \left[\varrho_{(k-1)p+3, (k-1)p+3}^{(-1)} + \lambda^2 \varrho_{(k-1)p+2, (k-1)p+2}^{(-1)} - 2\lambda \varrho_{(k-1)p+2, (k-1)p+3}^{(-1)} \right]$. Therefore, the desired asymptotic normality of $\sqrt{n}(\hat{\lambda}_n - \lambda)$ is obtained. \square

Remark 1. As pointed out by [3], the exponential decay condition $\beta_{jk}^{(i)} = c_{jk}\lambda^{i-1}$ is a condition for the long-memory property of HAR models. Ref. [3] discussed the HAR model of infinity order and its approximation of finite orders, where it has been shown that the exponential decay condition is equivalent to algebraically decay autocorrelation functions along with a mild lag condition. We are interested in testing whether or not the model is the exponentially weighted multivariate HAR model with common rate λ . For example, we construct the following hypothesis and test statistic: the null hypothesis $H_0 : \beta_{jk}^{(i)} = c_{jk}\lambda^{i-1}$ for some c_{jk} and common rate $|\lambda| < 1$, for each j, k , versus the alternative hypothesis H_A : the model does not have the common rate nor the exponentially weighted multivariate HAR model. For $i = 1, 2, \dots, p - 1$, let

$$\hat{\lambda}_{(i),n} = \frac{\sum_{k=1}^q \sum_{j=1}^q \hat{\beta}_{jk}^{(i+1)}}{\sum_{k=1}^q \sum_{j=1}^q \hat{\beta}_{jk}^{(i)}}$$

and note that asymptotic normality of $\hat{\lambda}_{(i),n}$ holds like Theorem 4 by that of the OLSEs. A collection of the differences $\{\hat{\lambda}_{(i_1),n} - \hat{\lambda}_{(i_2),n} : \text{for all pairs } (i_1, i_2), i_1 \neq i_2\}$ is considered. In the collection, all distinct elements in absolute values are at most $(p - 1)(p - 2) / 2 (= p^*)$. These are relabelled as $\{\Lambda_\ell : \ell = 1, 2, \dots, p^*\}$. Let $\bar{\Lambda}_n = \sum_{\ell=1}^{p^*} \Lambda_\ell / p^*$ or some related statistics of $\{\Lambda_\ell\}$. Similarly to Theorem 4, we can find the limiting distribution of $\bar{\Lambda}_n$ under the null hypothesis H_0 . The null H_0 might be rejected if $|\bar{\Lambda}_n|$ is large or if $\max_\ell \Lambda_\ell$ is

large. In a future work, a specific test statistic related to Λ_ℓ will be constructed and the limiting distribution of the test will be investigated.

5. A Monte-Carlo Study

We present simulation results for model (2) with $p = 3$ and $q = 2$, which is a bivariate case of two assets, using parameters $\beta_1 = (\beta_{10}, \beta_{11}^{(d)}, \beta_{11}^{(w)}, \beta_{11}^{(m)}, \beta_{12}^{(d)}, \beta_{12}^{(w)}, \beta_{12}^{(m)})^\top = (0.3, 0, 2, 0.1, 0.05, 0.1, 0.05, 0.02)^\top$, $\beta_2 = (\beta_{20}, \beta_{21}^{(d)}, \beta_{21}^{(w)}, \beta_{21}^{(m)}, \beta_{22}^{(d)}, \beta_{22}^{(w)}, \beta_{22}^{(m)})^\top = (-0.3, 0.07, 0.04, 0.01, 0.25, 0.07, 0.1)^\top$. We consider both independent errors and correlated ones for $\{\epsilon_{j,t}\}$, $j = 1, 2$: (i) i.i.d. normal distribution $N(0, 1)$ with $E[\mathcal{E}_t \mathcal{E}_t^\top] = I_2$ and (ii) correlated normal distribution with $E[\mathcal{E}_t \mathcal{E}_t^\top] = (\sigma_{jl}) \in \mathbb{R}^{2 \times 2}$ with $\sigma_{11} = \sigma_1^2 = 1, \sigma_{22} = \sigma_2^2 = 1.21, \sigma_{12} = \rho\sigma_1\sigma_2$ with $\rho = 0.5$. The bivariate HAR(3,2) processes are generated according to (2) with sample size $n = 1000$.

We compute the LSEs for the parameters with replication numbers 500 and give the sample means and standard errors of the 500 estimates in Tables 1 and 2 with $n = 300, 600, 1000$. It is reported in Tables that sample means are closer to the real values of parameters as well as standard errors decreases as n increases. Additionally, Figures 1 and 2 illustrate plots of sample means (a), (b) and standard errors (c), (d) of the LSEs of the parameters as n increases on the horizontal axis: Figures 1a,c and 2a,c for the estimates of β_1 and Figures 1b,d and 2b,d for the estimates of β_2 .

Confidence intervals using the normal approximations in Theorems 2 and 3 with confidence level 95% are constructed: $0.95 = P(\hat{\beta} - z_{0.975}\hat{se} \leq \beta \leq \hat{\beta} + z_{0.975}\hat{se})$, for each parameter component β , with its LSE $\hat{\beta}$ and standard-error estimate $\hat{se} = \hat{\sigma} / \sqrt{n}$, where $\hat{\sigma}^2$ is the corresponding (diagonal) component of estimate of asymptotic covariance matrix in Theorems 2 and 3. The confidence intervals of 500 samples as well as the empirical coverage probabilities and average lengths in the i.i.d. error case by the normal approximations are demonstrated in Figure 3, where sample size $n = 1000$ and replication number 500 are used. To illustrate the multivariate asymptotic normality, plots of the normal approximations for estimates of some parameters are depicted in Figure 4, and the bivariate normalities of some pairs of two chosen estimates also can be seen in Figure 5. The three figures support normality results established in the theory.

Table 1. Sample mean and standard error of 500 estimates in i.i.d. case.

Parameter		Sample Mean (Standard Error)			
		$n = 300$	$n = 600$	$n = 1000$	
β_1	β_{10}	0.30	0.36 (0.17)	0.33 (0.12)	0.32 (0.08)
	$\beta_{11}^{(d)}$	0.20	0.19 (0.07)	0.20 (0.04)	0.20 (0.03)
	$\beta_{11}^{(w)}$	0.10	0.08 (0.14)	0.09 (0.10)	0.10 (0.07)
	$\beta_{11}^{(m)}$	0.05	-0.05 (0.26)	0.00 (0.18)	0.02 (0.13)
	$\beta_{12}^{(d)}$	0.10	0.10 (0.07)	0.10 (0.04)	0.10 (0.03)
	$\beta_{12}^{(w)}$	0.05	0.06 (0.13)	0.06 (0.09)	0.06 (0.07)
	$\beta_{12}^{(m)}$	0.02	0.04 (0.25)	0.02 (0.17)	0.02 (0.12)
	β_2	β_{20}	-0.30	-0.37 (0.16)	-0.34 (0.12)
$\beta_{21}^{(d)}$		0.07	0.06 (0.07)	0.07 (0.05)	0.07 (0.04)
$\beta_{21}^{(w)}$		0.04	0.06 (0.14)	0.05 (0.09)	0.05 (0.07)
$\beta_{21}^{(m)}$		0.01	0.01 (0.27)	0.01 (0.18)	0.01 (0.14)
$\beta_{22}^{(d)}$		0.25	0.24 (0.06)	0.25 (0.05)	0.25 (0.04)
$\beta_{22}^{(w)}$		0.07	0.05 (0.14)	0.06 (0.10)	0.06 (0.07)
$\beta_{22}^{(m)}$		0.10	0.00 (0.24)	0.05 (0.16)	0.08 (0.12)

Table 2. Sample mean and standard error of 500 estimates in correlated case.

Parameter	Sample Mean (Standard Error)			
	$n = 300$	$n = 600$	$n = 1000$	
β_{10}	0.30	0.36 (0.30)	0.33 (0.20)	0.32 (0.15)
$\beta_{11}^{(d)}$	0.20	0.18 (0.09)	0.19 (0.07)	0.19 (0.05)
$\beta_{11}^{(w)}$	0.10	0.07 (0.22)	0.09 (0.15)	0.10 (0.12)
$\beta_{11}^{(m)}$	0.05	−0.04 (0.44)	0.00 (0.29)	0.01 (0.22)
$\beta_{12}^{(d)}$	0.10	0.11 (0.09)	0.11 (0.07)	0.11 (0.05)
$\beta_{12}^{(w)}$	0.05	0.05 (0.21)	0.05 (0.14)	0.04 (0.11)
$\beta_{12}^{(m)}$	0.02	0.03 (0.44)	0.03 (0.28)	0.04 (0.21)
β_{20}	−0.30	−0.38 (0.33)	−0.33 (0.21)	−0.31 (0.15)
$\beta_{21}^{(d)}$	0.07	0.06 (0.11)	0.06 (0.07)	0.06 (0.06)
$\beta_{21}^{(w)}$	0.04	0.04 (0.25)	0.05 (0.16)	0.05 (0.12)
$\beta_{21}^{(m)}$	0.01	0.06 (0.48)	0.02 (0.31)	0.01 (0.22)
$\beta_{22}^{(d)}$	0.25	0.25 (0.10)	0.25 (0.07)	0.26 (0.06)
$\beta_{22}^{(w)}$	0.07	0.05 (0.23)	0.06 (0.15)	0.06 (0.11)
$\beta_{22}^{(m)}$	0.10	−0.03 (0.47)	0.05 (0.30)	0.08 (0.21)

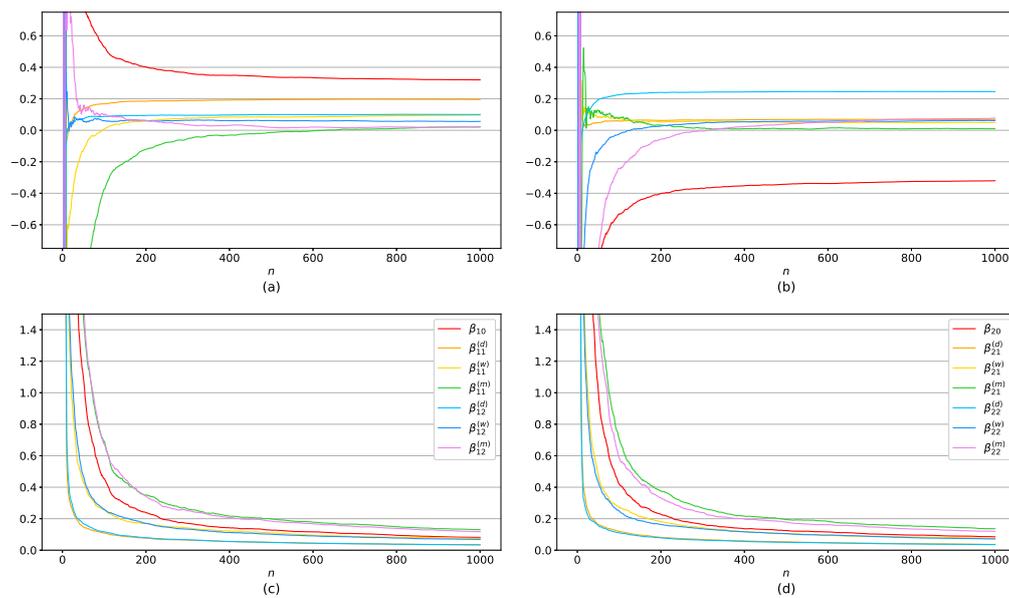


Figure 1. Sample mean (a,b) and standard error (c,d) of the 500 estimates in i.i.d. case.

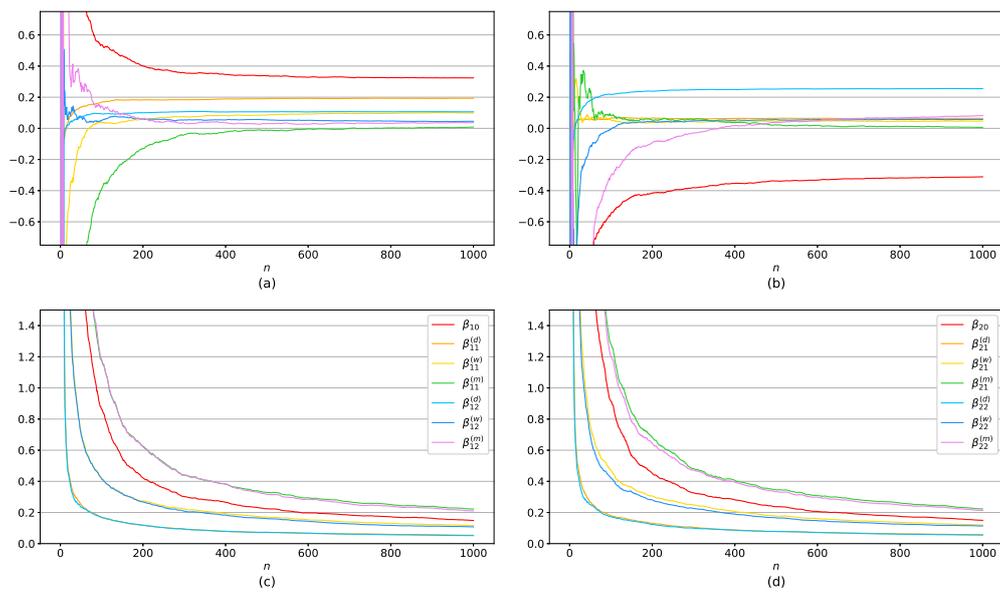


Figure 2. Sample mean (a,b) and standard error (c,d) of the 500 estimates in correlated case.

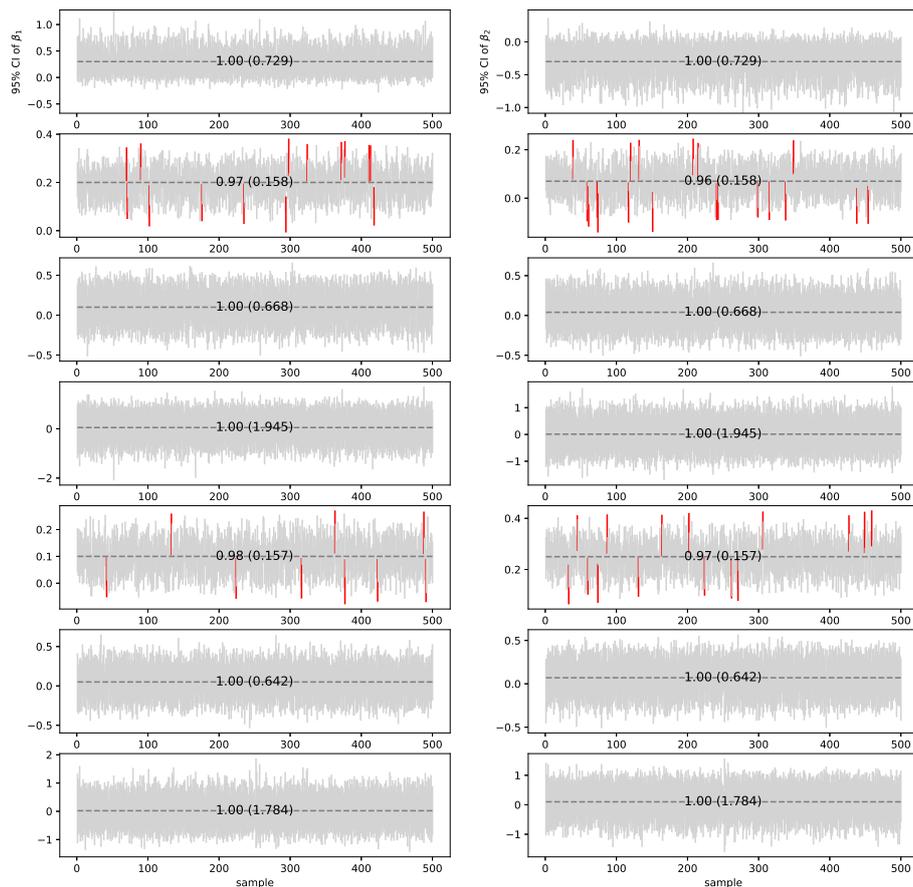


Figure 3. Empirical coverage and average length (in parenthesis) of 95% confidence intervals for seven components of β_1 (left column) and β_2 (right column) in i.i.d. error case. The horizontal dotted line indicates the true value of the parameter in each plot. Confidence intervals that do not contain the parameter are depicted as red color.

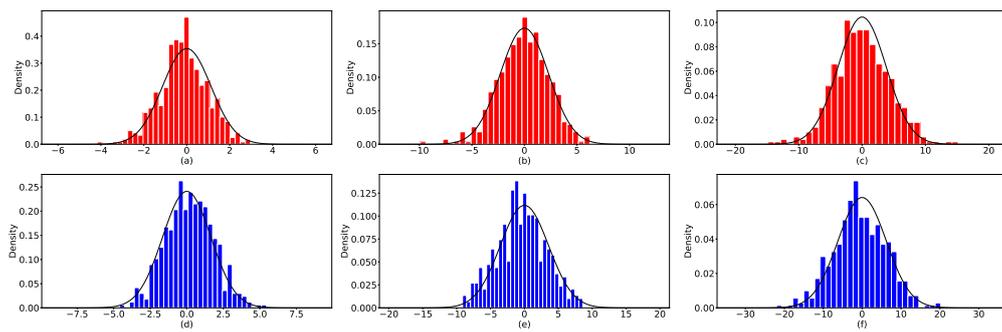


Figure 4. Normal approximation of estimates $\hat{\beta}_{11}^{(d)}, \hat{\beta}_{11}^{(w)}, \hat{\beta}_{11}^{(m)}$ in i.i.d.case (a–c) and correlated case (d–f).

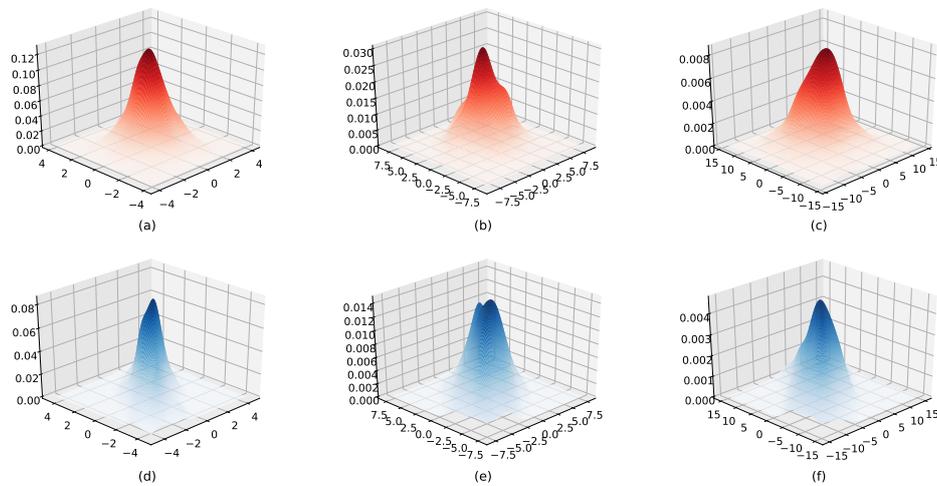


Figure 5. Bivariate normal approximation of some pairs of two chosen estimates $(\hat{\beta}_{11}^{(d)}, \hat{\beta}_{21}^{(d)})$, $(\hat{\beta}_{11}^{(w)}, \hat{\beta}_{21}^{(w)})$, $(\hat{\beta}_{11}^{(m)}, \hat{\beta}_{21}^{(m)})$ in i.i.d.case (a–c) and correlated case (d–f).

As for the exponentially weighted multivariate HAR(3,2) model with $\beta_{jk}^{(i)} = c_{jk}\lambda^{i-1}$ in Section 4, the estimates for the common rate are given in Table 3, from which we see that sample means of estimated values are close to the true ones with reasonable standard errors.

Table 3. Sample mean and standard error of 500 estimates in exponentially weighted HAR model.

Parameter		Sample Mean (Standard Error)		
		$n = 300$	$n = 600$	$n = 1000$
β_{10}	0.10	0.11 (0.09)	0.10 (0.06)	0.10 (0.05)
c_{11}	0.30	0.29 (0.05)	0.29 (0.04)	0.30 (0.03)
c_{12}	0.20	0.20 (0.05)	0.20 (0.04)	0.20 (0.03)
β_{20}	−0.30	−0.33 (0.09)	−0.32 (0.07)	−0.31 (0.05)
c_{21}	0.10	0.10 (0.05)	0.10 (0.04)	0.10 (0.03)
c_{22}	0.25	0.24 (0.05)	0.25 (0.04)	0.25 (0.03)
λ	0.50	0.49 (0.23)	0.52 (0.19)	0.50 (0.15)

6. Application

This section addresses empirical data analysis on the Gold spot price and S&P500 index. Their volatility is modeled by bivariate HAR model, using three years of daily closing price, from 18 September 2017 to 17 September 2020, of Gold and S&P500. The three years of Gold price

and S&P500 index movement and their log return are shown in Figure 6 while their volatility and autocorrelation coefficients functions (ACFs) in Figure 7.

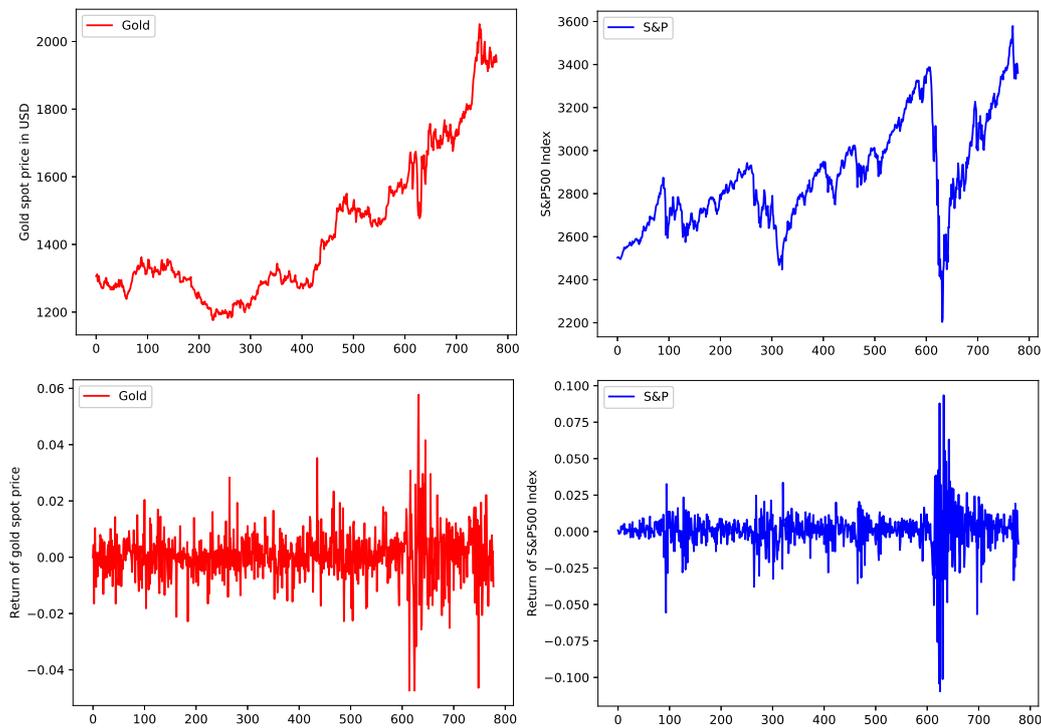


Figure 6. Gold spot price in USD and its return (in red) and S&P500 index and its return (in blue) against number of days starting from 18 September 2017 to 17 September 2020.

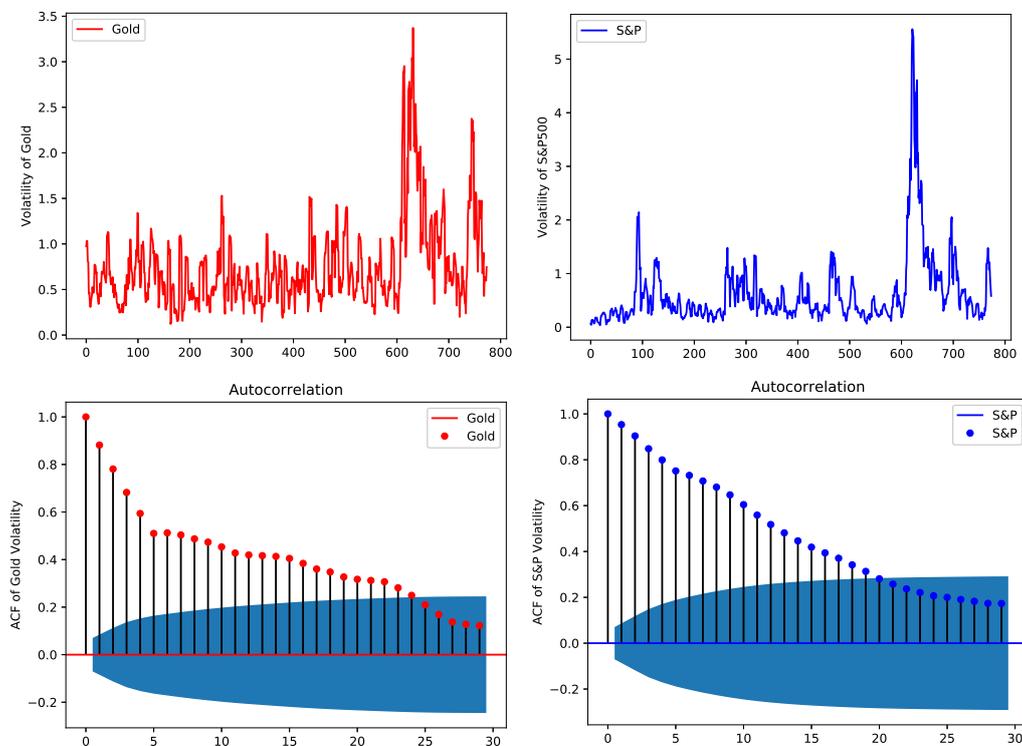


Figure 7. Volatility and ACF for Gold (in red) and S&P 500 (in blue).

We list some criteria of MSE, R^2 , AIC, BIC in Table 4 for the OLSEs (ordinary least squares estimators) of the bivariate HAR model of order $p = 3, 4$ by examining the volatility of Gold and S&P500. Conventionally lags of order $p = 3, 4$ are used for a day ($h_1 = 1$), a week ($h_2 = 5$), a month ($h_3 = 22$) and a quarter ($h_4 = 66$) in a HAR model. Nevertheless, in this work, we consider optimal lags in the sense of minimizing MSE of OLSEs for Gold and S&P500 volatility simultaneously. For order $p = 3, 4$, we choose lags h_2, \dots, h_p to satisfy the condition, $1 = h_1 < h_2 < \dots < h_p < \bar{h}$ for some large \bar{h} , so that the sum of the two mean squared residuals of the OLSEs for the volatility of Gold and S&P500 is minimized. As a result, we found optimal lags $(h_1, h_2, h_3) = (1, 5, 6)$ for $p = 3$ with $(MSE_{Gold}, MSE_{S\&P}) = (0.0461, 0.0387)$ and $(h_1, h_2, h_3, h_4) = (1, 3, 6, 7)$ for $p = 4$ with $(MSE_{Gold}, MSE_{S\&P}) = (0.0488, 0.0410)$. In Table 4, we compare our optimal lags with the conventional lags in different criteria. In all cases considered, our selection of optimal lags $(h_1, h_2, h_3) = (1, 5, 6)$ with $p = 3$ turns out to be the best.

Table 4. OLSEs in the bivariate HAR($p, 2$) model on volatility of Gold and S&P500.

		Gold				S&P500			
		MSE	R ²	AIC	BIC	MSE	R ²	AIC	BIC
$p = 3$	(1, 5, 22)	0.0492	0.790	−117.0	−84.68	0.0420	0.910	−236.6	−204.2
	(1, 5, 6)	0.0461 *	0.801 *	−169.8 *	−137.3 *	0.0387 *	0.916 *	−303.2 *	−270.7 *
$p = 4$	(1, 5, 22, 66)	0.0519	0.789	−73.30	−32.24	0.0448	0.909	−180.3	−139.3
	(1, 3, 6, 7)	0.0488	0.791	−128.1	−86.28	0.0410	0.912	−264.4	−222.6

* Denotes the best.

Therefore we report estimation results of the HAR(3,2) model with $(h_1, h_2, h_3) = (1, 5, 6)$ for the volatilities of Gold and S&P500 in Table 5, where coefficients estimates, standard errors and 95% confidence intervals are provided. In Figure 8, two plots of the bivariate HAR(3,2) fitted model by the OLSEs for this optimal case on the datasets of Gold and S&P volatilities are depicted along with residuals. We see that the bivariate HAR(3,2) fittings are similar to the real volatility plots of both datasets as reported upon the criterions.

Table 5. Estimation results for Gold and S&P500 using HAR(3,2) with $(h_1, h_2, h_3) = (1, 5, 6)$.

Parameter	Gold ($j = 1$)		S&P ($j = 2$)	
	Coeff. est.(s.e.)	(95% C.I.)	Coeff. est.(s.e.)	(95% C.I.)
β_{j0}	0.0790 (0.016)	(0.047, 0.111)	−0.0088 (0.015)	(−0.038, 0.021)
$\beta_{j1}^{(1)}$	0.9751 (0.038)	(0.901, 1.049)	0.0534 (0.034)	(−0.014, 0.121)
$\beta_{j1}^{(2)}$	−1.0822 (0.187)	(−1.450, −0.714)	−0.4278 (0.172)	(−0.765, −0.091)
$\beta_{j1}^{(3)}$	0.9301 (0.175)	(0.586, 1.274)	0.4458 (0.161)	(0.130, 0.761)
$\beta_{j2}^{(1)}$	0.1478 (0.039)	(0.071, 0.225)	1.1153 (0.036)	(1.045, 1.186)
$\beta_{j2}^{(2)}$	−0.6484 (0.198)	(−1.036, −0.261)	−1.1196 (0.181)	(−1.475, −0.764)
$\beta_{j2}^{(3)}$	0.5879 (0.179)	(0.237, 0.939)	0.9338 (0.164)	(0.612, 1.256)

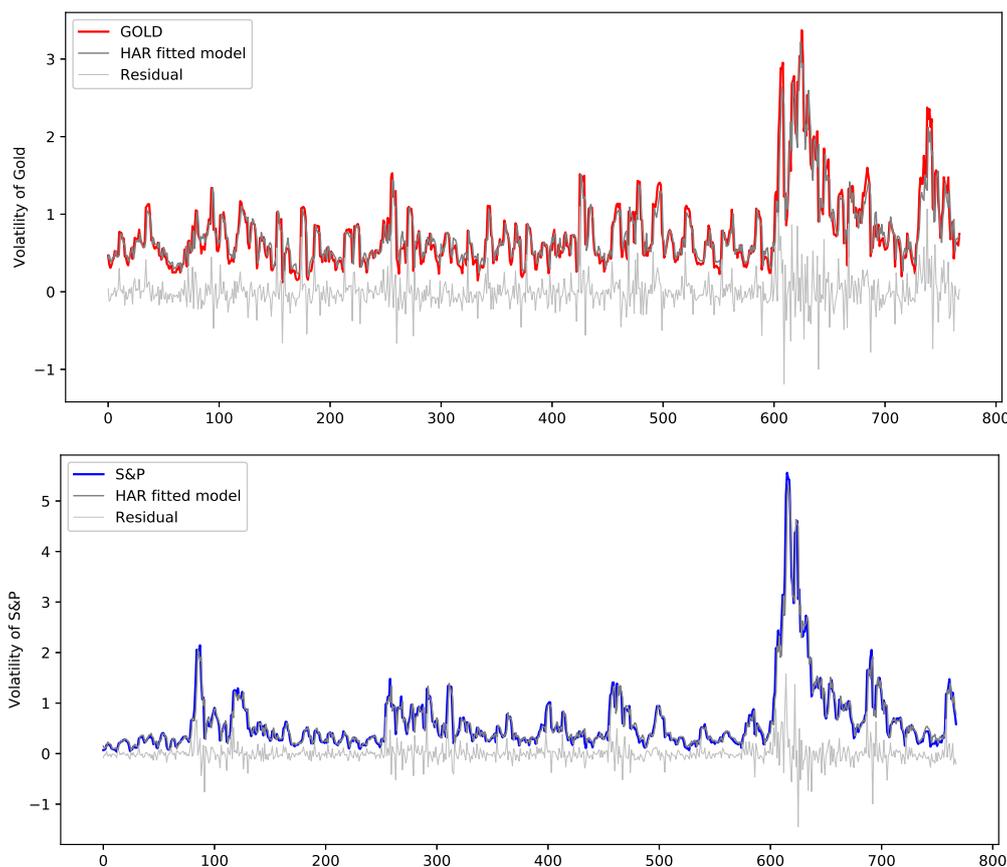


Figure 8. Fitted HAR model and residual using OLS for the volatility of Gold and S&P500.

7. Conclusions

Due to the cross-correlation of multiple assets and spillover effect of volatility in the financial market, a multivariate heterogeneous autoregressive-realized volatility (HAR-RV) model has attained much attention recently. In the multivariate HAR model, its stationarity is discussed and estimation problems are studied. We first investigate the strictly stationarity solution of the multivariate HAR model and second develop the asymptotic normality theory for the least squares estimates (LSEs) with i.i.d. and correlated errors, respectively. Third, we propose an exponentially weighted multivariate HAR model and estimate its common exponential decay rate. In a Monte-Carlo experiment, performances of the LSEs are numerically illustrated with sample mean and standard error of the estimates as well as empirical coverage and average length of confidence intervals by using the normal approximation. In addition, as a real data example, volatilities of Gold spot price and S&P500 index during recent three years are used to analyze in a bivariate HAR model. The coefficient estimates and confidence intervals are found in the bivariate HAR model of volatility of Gold and S&P500, along with choosing optimal lags, and it is shown that the bivariate HAR model with the proposed optimal lags is well matched with the volatility of the financial data.

We suggest some problems on the multivariate HAR model. As we proposed before, the exponentially weighted HAR models with decay rates are of interest owing to reduced numbers of parameters as well as the long-memory property. In modeling the multivariate HAR model, testing whether the HAR coefficients have exponentially weighted decay rates or not and furthermore whether multiple assets have a common decay rate or not might provide statistically useful tools to analyze the time series model. In follow-up studies we will deal with the hypotheses tests by constructing the test statistics and establishing the null limiting distribution. Finally, we mention that in a multivariate HAR model with heteroscedasticity errors, asymptotic properties of the estimates differ from the existing

results, and thus we will derive the asymptotic theory on the HAR models in the presence of dynamic heteroscedasticity. In this case, financial market data can be represented more remarkably and hence forecasting volatility with forecast accuracy will be carried out in the further research.

Author Contributions: Validation, W.-T.H., J.L. and E.H.; writing–review and editing, W.-T.H, J.L and E.H; funding acquisition, E.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Research Fund of Gachon University (GCU-2019-0299).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Corsi, F. A simple approximate long-memory model of realized volatility. *J. Financ. Econom.* **2009**, *7*, 174–196. [[CrossRef](#)]
2. Chen, Y.; Härdle, W.K.; Pigorsch, U. Localized realized volatility modeling. *J. Am. Stat. Assoc.* **2010**, *105*, 1376–1393. [[CrossRef](#)]
3. Hwang, E.; Shin, D.W. Infinite-order, long-memory heterogeneous autoregressive models. *Comput. Stat. Data Anal.* **2014**, *76*, 339–358. [[CrossRef](#)]
4. Qu, H.; Ji, P. Modeling realized volatility dynamics with a genetic algorithm. *J. Forecast.* **2016**, *35*, 434–444. [[CrossRef](#)]
5. Audrino, F.; Huang, C.; Okhrin, O. Flexible HAR model for realized volatility. *Stud. Nonlinear Dyn. Econom.* **2018**, *23*. [[CrossRef](#)]
6. Bubák, V.; Kočenda, E.; Žikeš, F. Volatility transmission in emerging European foreign exchange markets. *J. Bank. Financ.* **2011**, *35*, 2829–2841. [[CrossRef](#)]
7. Chiriac, R.; Voev, V. Modelling and forecasting multivariate realized volatility. *J. Appl. Econom.* **2011**, *26*, 922–947. [[CrossRef](#)]
8. Souček, M.; Todorova, N. Realized volatility transmission between crude oil and equity futures markets: A multivariate HAR approach. *Energy Econ.* **2013**, *40*, 586–597. [[CrossRef](#)]
9. Cubadda, G.; Guardabascio, B.; Hecq, A. A vector heterogeneous autoregressive index model for realized volatility measures. *Int. J. Forecast.* **2017**, *33*, 337–344. [[CrossRef](#)]
10. Čech, F.; Baruník, J. On the modelling and forecasting of multivariate realized volatility: Generalized heterogeneous autoregressive (GHAR) model. *J. Forecast.* **2017**, *36*, 181–206. [[CrossRef](#)]
11. Luo, J.; Chen, L. Realized volatility forecast with the Bayesian random compressed multivariate HAR model. *Int. J. Forecast.* **2020**, *36*, 781–799. [[CrossRef](#)]
12. Taylor, N. Realized volatility forecasting in an international context. *Appl. Econom. Lett.* **2015**, *22*, 503–509. [[CrossRef](#)]
13. Hall, P.; Heyde, C.C. *Martingale Limit Theory and Its Application*; Academic Press: Cambridge, MA, USA, 2014.
14. Hwang, E.; Shin, D.W. A CUSUM test for a long memory heterogeneous autoregressive model. *Econ. Lett.* **2013**, *121*, 379–383. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).