

CIRJE-F-1158

**Dynamic Factor, Leverage and Realized Covariances  
in Multivariate Stochastic Volatility**

Yuta Yamauchi

Graduate School of Economics, The University of Tokyo

Yasuhiro Omori  
The University of Tokyo

November 2020

CIRJE Discussion Papers can be downloaded without charge from:

<http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html>

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

# Dynamic factor, leverage and realized covariances in multivariate stochastic volatility

Yuta Yamauchi\* and Yasuhiro Omori<sup>†</sup>

November 13, 2020

## Abstract

In the stochastic volatility models for multivariate daily stock returns, it has been found that the estimates of parameters become unstable as the dimension of returns increases. To solve this problem, we focus on the factor structure of multiple returns and consider two additional sources of information: first, the realized stock index associated with the market factor, and second, the realized covariance matrix calculated from high frequency data. The proposed dynamic factor model with the leverage effect and realized measures is applied to ten of the top stocks composing the exchange traded fund linked with the investment return of the S&P500 index and the model is shown to have a stable advantage in portfolio performance.

**JEL classification:** C15, C32, C38, C58, G11

**Keywords:** Dynamic factor, leverage, Markov chain Monte Carlo, portfolio performance, realized covariance matrix, stochastic volatility, stock returns

---

\*Graduate School of Economics, The University of Tokyo, Tokyo, Japan. E-mail:mchyuta@gmail.com.

<sup>†</sup>Faculty of Economics, The University of Tokyo, Tokyo, Japan. E-mail:omori@e.u-tokyo.ac.jp. Phone: +81-3-5841-5516. Fax: +81-3-5841-5521.

Dynamic factor, leverage and realized covariances  
in multivariate stochastic volatility: Supplementary material

Yuta Yamauchi and Yasuhiro Omori

November 13, 2020

## A MCMC algorithm for sampling other parameters

### A.1 Joint posterior density

The logarithm of the joint posterior density of  $\boldsymbol{\theta}$  is given by

$$\begin{aligned}
& \log \pi(\boldsymbol{\theta} | \mathbf{x}, \mathbf{y}) \\
&= \text{const} + \sum_{t=1}^T \frac{s_0}{2} \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} \sum_{t=1}^T \text{tr} \{ (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}) \mathbf{W}_t^{-1} \} \\
&\quad - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{V}_{1t}| - \frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t)' \mathbf{V}_{1t}^{-1} (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t) \\
&\quad - \frac{T-1}{2} \log |\boldsymbol{\Sigma}_{\eta\eta}| - \frac{1}{2} \sum_{t=1}^{T-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t \}' \boldsymbol{\Sigma}_{\eta\eta}^{-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t \} \\
&\quad - \frac{1}{2} \log |\boldsymbol{\Sigma}_{h,0}| - \frac{1}{2} (\mathbf{h}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}_{h,0} (\mathbf{h}_1 - \boldsymbol{\mu}) \\
&\quad - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{V}_{2t|h}| - \frac{1}{2} \sum_{t=1}^T (\mathbf{f}_t - \boldsymbol{\gamma} - \boldsymbol{\Psi}(\mathbf{f}_{t-1} - \boldsymbol{\gamma}) - \mathbf{c}_t)' \mathbf{V}_{2t|h}^{-1} (\mathbf{f}_t - \boldsymbol{\gamma} - \boldsymbol{\Psi}(\mathbf{f}_{t-1} - \boldsymbol{\gamma}) - \mathbf{c}_t) \\
&\quad - \frac{1}{2} \sum_{t=1}^T \log |\boldsymbol{\Sigma}_\nu| - \frac{1}{2} \sum_{t=1}^T (\mathbf{x}_t - \mathbf{A} \mathbf{f}_t)' \boldsymbol{\Sigma}_\nu^{-1} (\mathbf{x}_t - \mathbf{A} \mathbf{f}_t) + \log \pi(\boldsymbol{\theta}),
\end{aligned}$$

where  $\pi(\boldsymbol{\theta})$  denotes a prior density of  $\boldsymbol{\theta}$ .

## A.2 Generation of $\mathbf{f}_t$

To sample  $(\mathbf{f}_1, \dots, \mathbf{f}_T)$  simultaneously, we consider the following state space equations:

$$\begin{pmatrix} \mathbf{x}_t \\ \mathbf{y}_t \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \mathbf{f}_t + \begin{pmatrix} \boldsymbol{\nu}_t \\ \mathbf{V}_{1t}^{1/2} \boldsymbol{\epsilon}_{1t} \end{pmatrix},$$

$$\mathbf{f}_t = (\mathbf{I}_q - \boldsymbol{\Psi}) \boldsymbol{\gamma} + \mathbf{c}_t + \boldsymbol{\Psi} \mathbf{f}_{t-1} + \mathbf{V}_{2t|h}^{1/2} \boldsymbol{\epsilon}_{2t}, \quad \boldsymbol{\Psi} = \text{diag}(\boldsymbol{\psi}),$$

where  $\mathbf{c}_t$  is defined in Equation (32). We implement the simulation smoother for the linear and Gaussian state space model.

## A.3 Generation of $\boldsymbol{\alpha}$

The conditional posterior density of  $\boldsymbol{\alpha}$  is given by

$$\begin{aligned} \pi(\boldsymbol{\alpha} | \cdot) &\propto \exp \left[ -\frac{1}{2} \sum_{j=2}^q (\boldsymbol{\alpha}_j - \mathbf{m}_{\alpha_j})' \mathbf{S}_{\alpha_j}^{-1} (\boldsymbol{\alpha}_j - \mathbf{m}_{\alpha_j}) - \frac{1}{2} \sum_{t=1}^T (\mathbf{A} \mathbf{f}_t - \mathbf{x}_t)' \boldsymbol{\Sigma}_{\nu}^{-1} (\mathbf{A} \mathbf{f}_t - \mathbf{x}_t) \right] \\ &\propto \prod_{j=2}^q \exp \left[ -\frac{1}{2} (\boldsymbol{\alpha}_j - \mathbf{m}_{\alpha_j})' \mathbf{S}_{\alpha_j}^{-1} (\boldsymbol{\alpha}_j - \mathbf{m}_{\alpha_j}) - \frac{1}{2\sigma_{\nu,j}^2} \sum_{t=1}^T \{\boldsymbol{\alpha}_j' \mathbf{f}_{jt}^* - (x_{jt} - f_{jt})\}^2 \right] \end{aligned} \quad (46)$$

where  $\mathbf{f}_{jt}^* = (f_{1t}, f_{2t}, \dots, f_{j-1,t})'$ . Noting that  $\boldsymbol{\alpha}_j$ s are conditionally independent, we generate  $\boldsymbol{\alpha}_j$  ( $j = 2, \dots, q$ ) from  $N(\hat{\mathbf{m}}_{\alpha_j}, \hat{\boldsymbol{\Sigma}}_{\alpha_j})$  where

$$\hat{\mathbf{m}}_{\alpha_j} = \hat{\boldsymbol{\Sigma}}_{\alpha_j} \left\{ \sum_{t=1}^T \left( \mathbf{f}_{jt}^* \sigma_{\nu,j}^{-2} (x_{jt} - f_{jt}) \right) + \mathbf{S}_{\alpha_j}^{-1} \mathbf{m}_{\alpha_j} \right\}, \quad \hat{\boldsymbol{\Sigma}}_{\alpha_j}^{-1} = \sum_{t=1}^T \sigma_{\nu,j}^{-2} \mathbf{f}_{jt}^* \mathbf{f}_{jt}^{*'} + \mathbf{S}_{\alpha_j}^{-1}.$$

## A.4 Generation of $\boldsymbol{\mu}$

The conditional posterior density of  $\boldsymbol{\mu}$  is given by

$$\begin{aligned} \pi(\boldsymbol{\mu} | \cdot) &\propto \exp \left[ -\frac{1}{2} (\boldsymbol{\mu} - \mathbf{m}_{\boldsymbol{\mu}})' \mathbf{S}_{\boldsymbol{\mu}}^{-1} (\boldsymbol{\mu} - \mathbf{m}_{\boldsymbol{\mu}}) - \frac{1}{2} (\mathbf{h}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}_{h,0}^{-1} (\mathbf{h}_1 - \boldsymbol{\mu}) \right] \\ &\times \exp \left[ -\frac{1}{2} \sum_{t=1}^{T-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t - \mathbf{c}_t^* \}' \boldsymbol{\Sigma}_{\eta|\epsilon}^{-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t - \mathbf{c}_t^* \}' \right], \end{aligned} \quad (47)$$

where  $\mathbf{c}_t^* = (c_{1t}^*, \dots, c_{p+q,t}^*)'$  (see A.6 for the definition of  $c_{it}^*$ ). We generate  $\boldsymbol{\mu}$  from  $N(\hat{\mathbf{m}}_\mu, \hat{\boldsymbol{\Sigma}}_\mu)$  where

$$\begin{aligned}\hat{\mathbf{m}}_\mu &= \hat{\boldsymbol{\Sigma}}_\mu \times \left[ \mathbf{S}_\mu^{-1} \mathbf{m}_\mu + \boldsymbol{\Sigma}_{h,0}^{-1} \mathbf{h}_1 + (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\Sigma}_{\eta|\epsilon}^{-1} \sum_{t=1}^{T-1} \{\mathbf{h}_{t+1} - \boldsymbol{\Phi} \mathbf{h}_t - \mathbf{c}_t^*\} \right] \\ \hat{\boldsymbol{\Sigma}}_\mu^{-1} &= \mathbf{S}_\mu^{-1} + \boldsymbol{\Sigma}_{h,0}^{-1} + (T-1)(\mathbf{I}_{p+q} - \boldsymbol{\Phi}) \boldsymbol{\Sigma}_{\eta|\epsilon}^{-1} (\mathbf{I}_{p+q} - \boldsymbol{\Phi}), \\ \boldsymbol{\Sigma}_{\eta|\epsilon} &= \text{diag}(\sigma_{\eta,1}^2, \dots, \sigma_{\eta,p}^2, \sigma_{\eta,p+1}^2(1 - \rho_{p+1}^2), \dots, \sigma_{\eta,p+q}^2(1 - \rho_{p+q}^2)).\end{aligned}$$

## A.5 Generation of $\boldsymbol{\gamma}$

The conditional posterior density of  $\boldsymbol{\gamma}$  is given by

$$\begin{aligned}\pi(\boldsymbol{\gamma}|\cdot) &\propto \exp\left[-\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{m}_\gamma)' \mathbf{S}_\gamma^{-1} (\boldsymbol{\gamma} - \mathbf{m}_\gamma)\right] \\ &\times \exp\left[-\frac{1}{2} \sum_{t=2}^T \{\mathbf{f}_t - (\mathbf{I} - \boldsymbol{\Psi}) \boldsymbol{\gamma} - \boldsymbol{\Psi} \mathbf{f}_{t-1} - \mathbf{c}_t\}' \mathbf{V}_{2t|h}^{-1} \{\mathbf{f}_t - (\mathbf{I} - \boldsymbol{\Psi}) \boldsymbol{\gamma} - \boldsymbol{\Psi} \mathbf{f}_{t-1} - \mathbf{c}_t\}\right],\end{aligned}\tag{48}$$

where  $\mathbf{c}_t$  is defined in Equation (32). We generate  $\boldsymbol{\gamma}$  from  $N(\hat{\mathbf{m}}_\gamma, \hat{\boldsymbol{\Sigma}}_\gamma)$  where

$$\begin{aligned}\hat{\mathbf{m}}_\gamma &= \hat{\boldsymbol{\Sigma}}_\gamma \times \left[ \mathbf{S}_\gamma^{-1} \mathbf{m}_\gamma + (\mathbf{I}_q - \boldsymbol{\Psi}) \sum_{t=2}^T \mathbf{V}_{2t|h}^{-1} \{\mathbf{f}_t - \boldsymbol{\Psi} \mathbf{f}_{t-1} - \mathbf{c}_t\} \right] \\ \hat{\boldsymbol{\Sigma}}_\gamma^{-1} &= \mathbf{S}_\gamma^{-1} + \sum_{t=2}^T (\mathbf{I}_q - \boldsymbol{\Psi}) \mathbf{V}_{2t|h}^{-1} (\mathbf{I}_q - \boldsymbol{\Psi}).\end{aligned}$$

## A.6 Generation of $\boldsymbol{\phi}$

The posterior distributions of  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{p+q})'$  are conditionally independent and the conditional posterior density of  $\phi_i$  is given by

$$\pi(\phi_i|\cdot) \propto k(\phi_i) \times \exp\left[-\frac{1}{2\sigma_{\eta,i}^2(1 - \rho_i^2)} \sum_{t=1}^{T-1} \{h_{i,t+1} - \mu_i - \phi_i(h_{it} - \mu_i) - c_{it}^*\}^2\right],\tag{49}$$

where  $\rho_1 = \dots = \rho_p = 0$ ,  $c_{1t}^* = \dots = c_{pt}^* = 0$  and

$$\begin{aligned}k(\phi_i) &= \sqrt{1 - \phi_i^2} \left(\frac{1 + \phi_i}{2}\right)^{a_\phi - 1} \left(\frac{1 - \phi_i}{2}\right)^{b_\phi - 1}, \quad i = 1, \dots, p+q, \\ c_{p+i,t}^* &= \rho_{p+i} \sigma_{\eta,p+i} \exp(-h_{p+i,t}/2) \{f_{it} - \gamma_i - \psi_i(f_{i,t-1} - \gamma_i)\}, \quad i = 1, \dots, q,\end{aligned}$$

Let  $\phi_i^o$  denote a current sample of  $\phi_i$ . Then we generate  $\phi_i^n$  from  $N(m_{\phi,i}, \sigma_{\phi,i}^2)$ , where

$$m_{\phi,i} = \frac{\sum_{t=1}^{T-1} (h_{i,t+1} - \mu_i - c_{it}^*)(h_{it} - \mu_i)}{\sum_{t=2}^{T-1} (h_{it} - \mu_i)^2}, \quad \sigma_{\phi,i}^2 = \frac{\sigma_{\eta,i}^2(1 - \rho_i^2)}{\sum_{t=2}^{T-1} (h_{it} - \mu_i)^2},$$

and accept it with probability  $\min\{1, k(\phi_i^n)/k(\phi_i^o)\}$ .

## A.7 Generation of $\psi$

The posterior distributions of  $\boldsymbol{\psi} = (\psi_1, \dots, \psi_q)'$  are conditionally independent and the conditional posterior density of  $\psi_i$  is given by

$$\pi(\psi_i|\cdot) \propto k(\psi_i) \times \exp \left[ -\frac{1}{2} \sum_{t=2}^T \frac{\{f_{it} - \gamma_i - \psi_i(f_{i,t-1} - \gamma_i) - c_{it}\}^2}{\{1 - \rho_{p+i}^2 I(t < T)\} \exp(h_{p+i,t})} \right] \quad (50)$$

where  $c_{it}$  is defined in Equation (32), and

$$k(\psi_i) = \left( \frac{1 + \psi_i}{2} \right)^{a_\psi - 1} \left( \frac{1 - \psi_i}{2} \right)^{b_\psi - 1}, \quad i = 1, \dots, q.$$

Let  $\psi_i^o$  denote a current sample of  $\psi_i$ . Then we generate  $\psi_i^n$  from  $N(m_{\psi,i}, \sigma_{\psi,i}^2)$ , where

$$m_{\psi,i} = \sigma_{\psi,i}^2 \sum_{t=2}^T \frac{(f_{i,t} - \gamma_i - c_{it})(f_{i,t-1} - \gamma_i)}{\{1 - \rho_{p+i}^2 I(t < T)\} \exp(h_{p+i,t})}, \quad \sigma_{\psi,i}^{-2} = \sum_{t=2}^T \frac{(f_{i,t-1} - \gamma_i)^2}{\{1 - \rho_{p+i}^2 I(t < T)\} \exp(h_{p+i,t})},$$

and accept it with probability  $\min\{1, k(\psi_i^n)/k(\psi_i^o)\}$ .

## A.8 Generation of $\rho$

The conditional posterior density of  $\rho_k$  ( $k = p + 1, \dots, p + q$ ) is given by

$$\begin{aligned} \pi(\rho_k|\cdot) &\propto \left( \frac{1 + \rho_k}{2} \right)^{a_\rho - 1} \left( \frac{1 - \rho_k}{2} \right)^{b_\rho - 1} (1 - \rho_k^2)^{-(T-1)/2} \\ &\times \exp \left[ -\frac{1}{2\sigma_{\eta,k}^2 (1 - \rho_k^2)} \sum_{t=1}^{T-1} \{h_{k,t+1} - \mu_k - \phi_k(h_{k,t} - \mu_k) - c_{kt}^*\}^2 \right], \end{aligned}$$

where  $c_{kt}^*$  is defined in A.6. In order to sample  $\rho_k$ , we consider Fisher transformation,  $g_k = \log(1 + \rho_k) - \log(1 - \rho_k)$  (or  $\rho_k = \{\exp(g_k) - 1\}/\{\exp(g_k) + 1\}$ ). Let  $l(g_k)$  denote the logarithm of the posterior density of  $g_k$  given by

$$l(g_k) \equiv \log \pi(\rho_k|\cdot) + \log(J(g_k)), \quad k = p + 1, \dots, p + q,$$

where  $J(g_k)$  is Jacobian such that

$$J(g_k) = \frac{2 \exp(g_k)}{(\exp(g_k) + 1)^2}.$$

First, we compute the mode  $\hat{g}_k$  of  $l(g_k)$  numerically. Then we consider Taylor expansion of  $l(g_k)$  around  $\hat{g}_k$ :

$$l(g_k) \approx l(\hat{g}_k) + m_{g_k}(g_k - \hat{g}_k) - \frac{(g_k - \hat{g}_k)^2}{2S_{g_k}} \equiv l^*(g_k),$$

where

$$m_{g_k} = \left. \frac{\partial l(g_k)}{\partial g_k} \right|_{g_k = \hat{g}_k}, \quad S_{g_k}^{-1} = - \left. \frac{\partial^2 l(g_k)}{\partial g_k^2} \right|_{g_k = \hat{g}_k}.$$

To conduct MH algorithm, we propose a candidate  $g_k^n$  from  $N(\hat{g}_k + S_{g_k} m_{g_k}, S_{g_k})$  and accept it with probability  $\{1, \exp(l(g_k^n) - l(g_k^o) - l^*(g_k^n) + l^*(g_k^o))\}$  where  $g_k^o$  is a current sample.

### A.9 Generation of $\sigma_\eta$

First we define  $\tilde{n}_{\eta,i} = n_\eta + T$  and

$$\tilde{d}_{\eta,i} = d_\eta + \sum_{t=1}^{T-1} \{h_{i,t+1} - (1 - \phi_i)\mu_i - \phi h_{i,t}\}^2 + (1 - \phi_i^2)(h_{i1} - \mu_i)^2,$$

for  $i = 1, \dots, p + q$ .

*Generation of  $\sigma_{\eta,1}^2, \dots, \sigma_{\eta,p}^2$ .* We generate  $\sigma_{\eta,i}^2$  from the conditional posterior distribution,  $\sigma_{\eta,i}^2 \sim \text{IG}(\tilde{n}_{\eta,i}/2, \tilde{d}_{\eta,i}/2)$ ,  $i = 1, \dots, p$ .

*Generation of  $\sigma_{\eta,p+1}^2, \dots, \sigma_{\eta,p+q}^2$ .* The conditional posterior density for  $\sigma_{\eta,p+i}^2$  is given by

$$\begin{aligned} \pi(\sigma_{\eta,p+i}^2 | \cdot) &\propto (\sigma_{\eta,p+i}^2)^{-\frac{\tilde{n}_{\eta,i}}{2}-1} \exp \left[ -\frac{\tilde{d}_{\eta,i}}{2\sigma_{\eta,p+i}^2} + k(\sigma_{\eta,p+i}^2) \right], \quad i = 1, \dots, q, \quad (51) \\ k(\sigma_{\eta,p+i}^2) &= -\frac{1}{2} \sum_{t=1}^{T-1} \frac{(f_{it} - \gamma - \psi_i(f_{i,t-1} - \gamma) - c_{it})^2}{(1 - \rho_{p+i}^2) \exp(h_{p+i,t})}, \end{aligned}$$

where  $c_{it}$  is defined in Equation (32). We conduct MH algorithm to sample from its conditional posterior distribution. Let  $\sigma_{\eta,p+i}^{2,o}$  denote a current sample, and we propose a candidate  $\sigma_{\eta,p+i}^{2,n} \sim \text{IG}(\tilde{n}_{\eta,i}/2, \tilde{d}_{\eta,i}/2)$ . Accept it with probability  $\min\{1, \exp(k(\sigma_{\eta,p+i}^{2,n}) - k(\sigma_{\eta,p+i}^{2,o}))\}$ .

### A.10 Generation of $\sigma_\nu$

Let  $\tilde{x}_{jt}$  denote the  $j$ -th element of  $\mathbf{x}_t - \mathbf{A}\mathbf{f}_t$ . We generate  $\sigma_{\nu,j}^2$  from the conditional posterior distribution,  $\sigma_{\nu,j}^2 \sim \text{IG}(\tilde{n}_{\nu,j}/2, \tilde{d}_{\nu,j}/2)$ ,  $j = 1, \dots, q$  where  $\tilde{n}_{\nu,j} = n_\nu + T$  and  $\tilde{d}_{\nu,j} = d_\nu + \sum_{t=1}^T \tilde{x}_{jt}^2$ .

### A.11 Generation of $\delta$

The conditional posterior density of  $\delta$  is

$$\begin{aligned} \pi(\delta|\cdot) &= \frac{1}{2^{\frac{(\delta+p+3)p}{2}} \Gamma_p\left(\frac{\delta+p+3}{2}\right)} |(\delta+2)(\mathbf{B}\mathbf{V}_{2t}\mathbf{B}' + \mathbf{V}_{1t})|^{\frac{\delta+p+3}{2}} \times |\mathbf{W}_t|^{-\frac{\delta+2p+4}{2}} \\ &\times \exp\left[-\frac{1}{2} \text{tr}\left\{(\delta+2)(\mathbf{B}\mathbf{V}_{2t}\mathbf{B}' + \mathbf{V}_{1t})\mathbf{W}_t^{-1}\right\}\right]. \end{aligned}$$

We generate  $\log \delta$  using random-walk MH algorithm. Given the current value of  $\log \delta^o$ , we generate  $\log \delta^n \sim N(\log \delta^o, \sigma_\delta^2)$ . The  $\sigma_\delta$  is a tuning parameter to optimize the sampling efficiency. In our empirical study, we set  $\sigma_\delta = 0.001$  and the acceptance rate was 0.898.

## B Figures

### B.1 The effect of the realized measures on the estimated dynamic correlations

Figures 1 and 2 show the estimated dynamic correlation coefficients between AAPL and MSFT. There are rises in the correlation coefficients among stocks around the period of the global financial crisis.

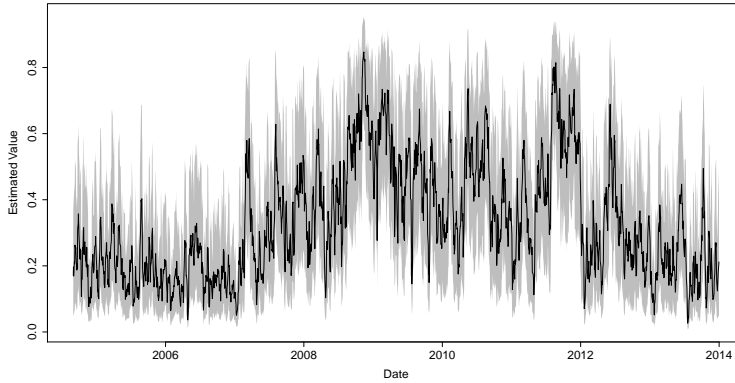


Figure 1: FMSV model. The posterior mean (solid line) and 95% credible interval (shaded area) of estimated dynamic correlation coefficients between AAPL and MSFT.



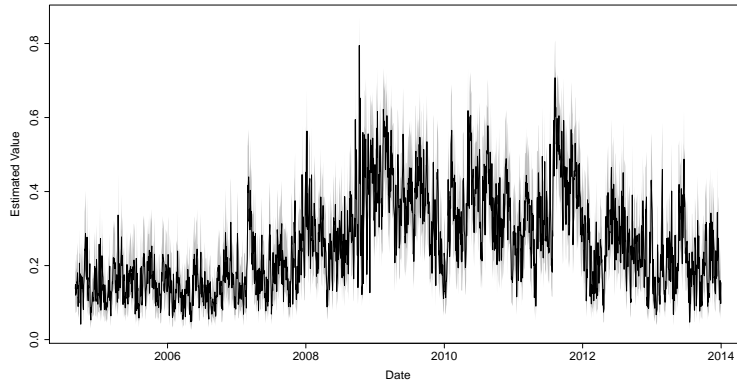


Figure 2: FMRSV model. The posterior mean (solid line) and 95% credible interval (shaded area) of estimated dynamic correlation coefficients between AAPL and MSFT.

Figure 3 shows the correlation coefficients in the FMSV and FMRSV models simultaneously. The estimates of the FMSV model are sometimes larger than those of the FMRSV model.

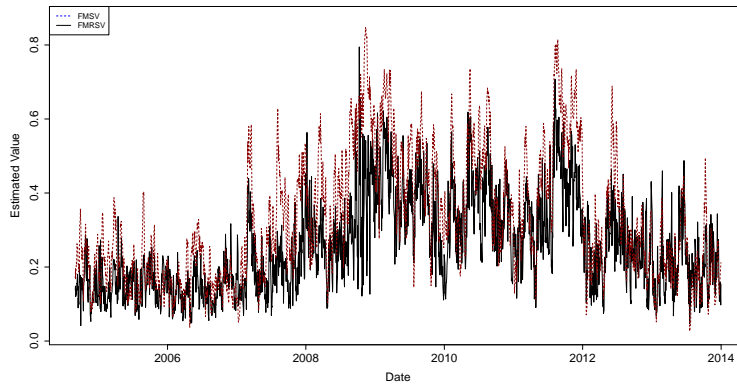


Figure 3: FMSV (black solid line) and FMRSV (red dashed line) models. The posterior mean of estimated dynamic correlation coefficients between AAPL and MSFT.

Further, Figure 4 shows the boxplots of the difference between the estimated dynamic correlations in the FMSV and FMRSV models. The estimates of the FMSV model tend to be larger than those of the FMRSV model.

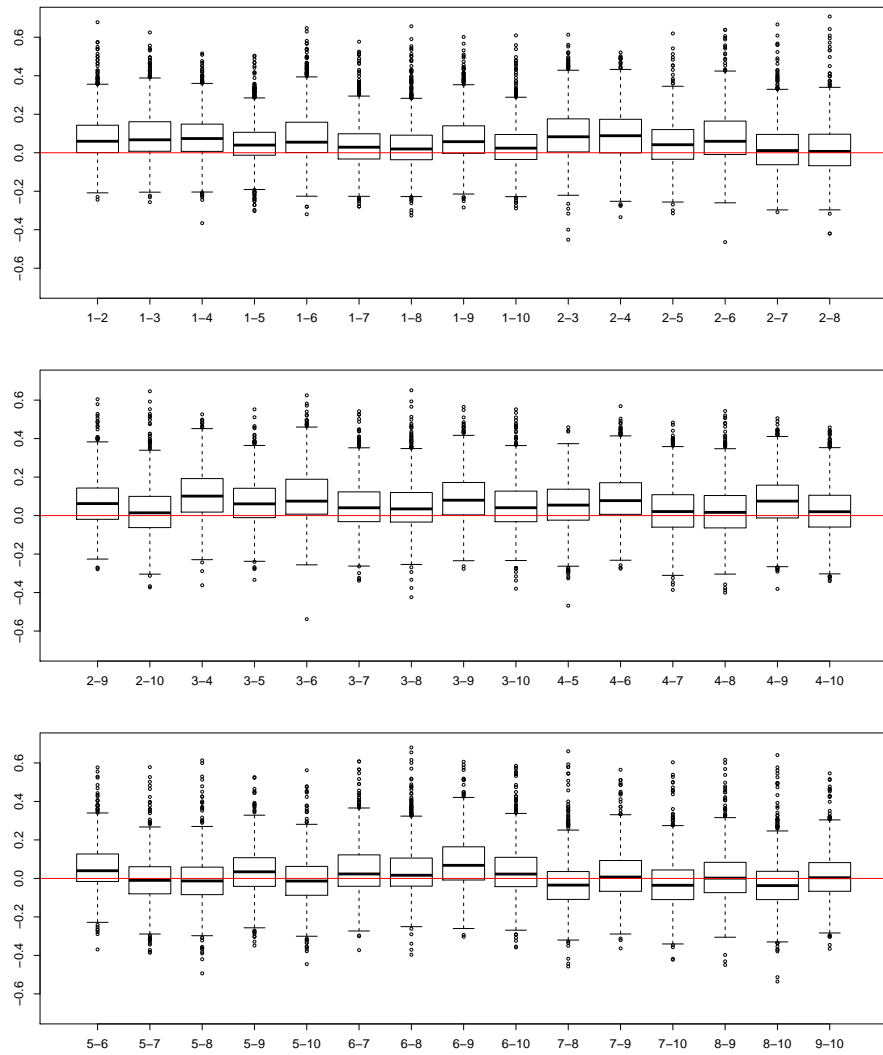


Figure 4: Boxplots of the difference between dynamic correlation estimates (the estimate of FMSV model minus the estimate of FMRSV model).

# 1 Introduction

Portfolio management is one of the major purposes of constructing statistical models that estimate and predict time-varying variances and covariances of asset returns in financial econometrics. The weights on the financial assets are chosen to optimize the objective function for the portfolio based on the estimated models. The univariate stochastic volatility (SV) model is a popular statistical model that can well describe the dynamic stochastic process of the time-varying volatility of an asset return and there are various extensions to the multivariate SV model. As is often the case with the multivariate volatility model, the number of model parameters and dynamic latent variables increases as the number of assets increases, which leads to unstable and unreliable estimation results. Since a small number of market factors are often found to exist in empirical studies, this paper focuses on the factor multivariate stochastic volatility with the leverage effect (FMSV) model to overcome this problem. The FMSV model describes the factor structure of the dynamic covariance matrices among asset returns, and has been investigated in the past literature (e.g. Pitt and Shephard (1999), Aguilar and West (2000), Chib et al. (2006), Lopes and Carvalho (2007)), and extended to incorporate the leverage effect (Ishihara and Omori (2017)) which implies a decrease in the stock return followed by the increase in its volatility (see e.g. Yu (2005)).

Even in such a parsimonious model, the information from daily returns is often insufficient to obtain accurate and stable estimates of the parameters and dynamic latent variables for multivariate SV models. A common way to incorporate additional information on the volatilities and covolatilities of asset returns is to use high-frequency data that include information on intraday asset trades. The realized stochastic volatility (RSV) models, for example, are this type of extension and are known to outperform models without realized measures at estimating model parameters, forecasting volatilities and portfolio performance (Takahashi et al. (2009), Hansen et al. (2012), Koopman and Scharth (2013), Takahashi et al. (2016), Shirota et al. (2017), Kurose and Omori (2019), Yamauchi and Omori (2019)). However, it is not straightforward to extend the FMSV model using the realized covariance matrices in a similar manner, since the realized covariance matrices do not directly correspond to the factor loading matrix and the idiosyncratic volatilities that are not explained by the factors. This paper constructs the appropriate relationship between the realized covariance matrix and the true covariance matrix in the FMSV model, and proposes a dynamic factor multivariate realized stochastic volatility with the leverage effect (FMRSV) model as an extension

of both FMSV and RSV models.

It should be mentioned that the estimate of the true covariance matrix using the realized volatilities and realized covolatilities is known to be biased due to market microstructure noise, nontrading hours, nonsynchronous trading and so forth. Although the biases can be removed by introducing the corresponding adjustment terms to the measurement equation in some RSV models, it turns out to be difficult in the multivariate SV model because of its factor structure. We instead show how to adjust the biases in the preprocessing step using the information of daily returns, thus extending the method of Hansen and Lunde (2005) for the univariate SV model.

Furthermore, we describe a novel method to estimate the relative weight of additional information from realized measures through the precision parameter. As we shall see in our empirical studies, the weight of the realized measure equation is found to be large; and consequently, the FMRSV model estimates the leverage effect of latent factors and the correlation coefficients between assets with smaller absolute values than the FMSV model. In other words, without the additional information from realized covariances, we tend to overestimate the leverage effect and the strength of the linear relationship between asset returns.

The rest of this paper is organized as follows. Section 2 introduces the factor multivariate SV model with daily stock returns, realized factors and realized covariance matrices. Section 3 describes the estimation method using the Markov chain Monte Carlo simulation. In Section 4, we applied the proposed model to ten U.S. stock returns data.

## 2 Factor multivariate stochastic volatility with realized measures

### 2.1 Dynamic factor and stochastic volatility

First, we describe the factor multivariate stochastic volatility (SV) model. Let  $\mathbf{y}_t = (y_{1t}, \dots, y_{pt})'$  and  $\mathbf{f}_t = (f_{1t}, \dots, f_{qt})'$  denote a  $p \times 1$  stock return vector and a  $q \times 1$  latent factor vector. As we shall see in our empirical studies, it is often the case that there is co-movement among stock returns (see Figure 1). To model the co-movement, we assume that the return is the sum of the factor and idiosyncratic components as in Chib et al. (2002) who considered the

static factor without the leverage effect:

$$\mathbf{y}_t = \mathbf{B} \mathbf{f}_t + \mathbf{V}_{1t}^{1/2} \boldsymbol{\epsilon}_{1t}, \quad t = 1, \dots, T, \quad (1)$$

$$\boldsymbol{\epsilon}_{1t} \sim N(\mathbf{0}, \mathbf{I}_p), \quad (2)$$

$$\mathbf{V}_{1t} = \text{diag}(\exp(h_{1t}), \dots, \exp(h_{pt})), \quad (3)$$

where  $\mathbf{B}$  is the  $p \times q$  coefficient matrix of the factor, and  $\mathbf{I}_p$  denotes the  $p \times p$  identity matrix.

Let us denote

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\beta}'_1 \\ \vdots \\ \boldsymbol{\beta}'_p \end{bmatrix}, \quad \boldsymbol{\beta} \equiv \text{vec}(\mathbf{B}') = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_p \end{bmatrix}.$$

Further, we consider a dynamic process for the latent factor  $\mathbf{f}_t$  (e.g. Han (2006)). Assume that it follows the first order stationary autoregressive process:

$$\mathbf{f}_t = \boldsymbol{\gamma} + \boldsymbol{\psi} \odot (\mathbf{f}_{t-1} - \boldsymbol{\gamma}) + \mathbf{V}_{2t}^{1/2} \boldsymbol{\epsilon}_{2t}, \quad t = 1, \dots, T, \quad (4)$$

$$\boldsymbol{\epsilon}_{2t} \sim N(\mathbf{0}, \mathbf{I}_q), \quad \mathbf{f}_0 \equiv \boldsymbol{\gamma}, \quad (5)$$

$$\mathbf{V}_{2t} = \text{diag}(\exp(h_{p+1,t}), \dots, \exp(h_{p+q,t})), \quad (6)$$

where  $\odot$  denotes the Hadamard product,  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)'$  is a mean vector of  $\mathbf{f}_t$  and  $\boldsymbol{\psi} = (\psi_1, \dots, \psi_q)'$  ( $|\psi_i| < 1, i = 1, \dots, q$ ) is an autoregressive coefficient vector. For the initial factor  $\mathbf{f}_1$ , we assume  $\mathbf{f}_0 \equiv \boldsymbol{\gamma}$  in Equation (4) for simplicity. The log volatility  $\mathbf{h}_t = (h_{1t}, \dots, h_{p+q,t})'$  is assumed to follow the first order stationary autoregressive process

$$\mathbf{h}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\phi} \odot (\mathbf{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\eta\eta}), \quad t = 1, \dots, T-1, \quad (7)$$

$$\boldsymbol{\Sigma}_{\eta\eta} = \text{diag}(\boldsymbol{\sigma}_\eta), \quad \boldsymbol{\sigma}_\eta = (\sigma_{\eta,1}^2, \dots, \sigma_{\eta,p+q}^2), \quad (8)$$

$$\mathbf{h}_1 = \boldsymbol{\mu} + \boldsymbol{\eta}_0, \quad \boldsymbol{\eta}_0 \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{h,0}), \quad \boldsymbol{\Sigma}_{h,0} = \text{diag}\left(\frac{\sigma_{\eta,1}^2}{1 - \phi_1^2}, \dots, \frac{\sigma_{\eta,p+q}^2}{1 - \phi_{p+q}^2}\right), \quad (9)$$

where  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_{p+q})'$  is a mean vector of  $\mathbf{h}_t$  and  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{p+q})'$  ( $|\phi_i| < 1, i = 1, \dots, p+q$ ) is an autoregressive coefficient vector. To incorporate the leverage effects, we model the joint distribution of the error terms as follows:

$$\begin{pmatrix} \boldsymbol{\epsilon}_{1t} \\ \boldsymbol{\epsilon}_{2t} \\ \boldsymbol{\eta}_t \end{pmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \mathbf{I}_{p+q} & \boldsymbol{\Sigma}_{\epsilon\eta} \\ \boldsymbol{\Sigma}_{\eta\epsilon} & \boldsymbol{\Sigma}_{\eta\eta} \end{bmatrix}\right), \quad \boldsymbol{\Sigma}_{\epsilon\eta} = \text{diag}(\rho_1 \sigma_{\eta,1}, \dots, \rho_{p+q} \sigma_{\eta,p+q}), \quad (10)$$

where  $\rho_i < 0$  implies that there is a leverage effect between  $\epsilon_{it}$  and  $h_{i,t+1}$ . In empirical studies, the leverage effect is often found to exist only for the factor process, especially the first factor that represents the market factor (e.g. Ishihara and Omori (2017), Yamauchi and Omori (2019)). Thus, we assume that there is no leverage effect for those idiosyncratic components, i.e.,  $\rho_1 = \dots = \rho_p = 0$ ,

$$\Sigma_{\epsilon\eta} = \text{diag}(0, \dots, 0, \rho_{p+1}\sigma_{\eta,p+1}, \dots, \rho_{p+q}\sigma_{\eta,p+q}), \quad (11)$$

and denote  $\boldsymbol{\rho} \equiv (\rho_{p+1}, \dots, \rho_{p+q})'$ . We define the FMSV model by Equations (1)–(11).

## 2.2 Realized factor and realized covariance matrix

When there are many parameters and latent variables in the multivariate SV model, the parameter estimates of interest often become unstable and inaccurate. We overcome these difficulties by introducing additional measurement equations based on the realized measures.

*Realized factor.* First, we introduce the realized measure for factor  $\mathbf{f}_t$ . In stock markets, it is usual that some major indices represent the market factor dynamics, such as the S&P500 index return and sector index returns. We call them realized factor series in this paper and denote them by a  $q \times 1$  observed realized factor vector  $\mathbf{x}_t$ . In practice,  $q$  is expected to be small. Since the realized factors are considered to be correlated, we assume

$$\mathbf{x}_t = \mathbf{A} \mathbf{f}_t + \boldsymbol{\nu}_t \quad \boldsymbol{\nu}_t \sim \text{N}(\mathbf{0}, \Sigma_{\nu}), \quad (12)$$

$$\Sigma_{\nu} = \text{diag}(\boldsymbol{\sigma}_{\nu}), \quad \boldsymbol{\sigma}_{\nu} = (\sigma_{\nu 1}^2, \dots, \sigma_{\nu q}^2), \quad t = 1, \dots, T, \quad (13)$$

where a loading matrix  $\mathbf{A}$  is a  $q \times q$  lower triangular matrix such that

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ a_{21} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{q1} & \dots & a_{q,q-1} & 1 \end{pmatrix}, \quad (14)$$

for identification reasons, and we let the lower triangular parameters of  $\mathbf{A}$  be  $\boldsymbol{\alpha} \equiv (\boldsymbol{\alpha}'_2, \dots, \boldsymbol{\alpha}'_q)'$ , where  $\boldsymbol{\alpha}_j = (\alpha_{j1}, \dots, \alpha_{j,j-1})'$  is a  $(j-1) \times 1$  vector. The first element of the realized factor  $\mathbf{x}_t$  is chosen to represent the overall dynamics of the market such as the S&P500.

*Realized covariance matrix.* Next, we consider the realized measures for  $\mathbf{h}_t$ . The high-frequency data are used to compute the realized covariance matrix  $\mathbf{W}_t$  which is assumed to follow

$$\mathbf{W}_t \sim \text{IW}(s_0, \{k_0 \text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta})\}^{-1}), \quad \text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta}) = \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}, \quad (15)$$

where  $s_0$  and  $k_0$  are the constant hyperparameters, and the probability density function of  $\mathbf{W}_t$  given the parameters and latent variables is

$$f(\mathbf{W}_t | \cdot) \propto |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|^{\frac{s_0}{2}} \times |\mathbf{W}_t|^{-\frac{s_0+p+1}{2}} \times \exp \left[ -\frac{1}{2} \text{tr} \{k_0(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})\} \mathbf{W}_t^{-1} \right]. \quad (16)$$

How we set the hyperparameters  $s_0$  and  $k_0$  affects to what extent we incorporate the information of the realized covariance matrix. Thus, we introduce a new model parameter  $\delta$  where  $s_0 = \delta + p + 3$  and  $k_0 = \delta + 2$ . The expected values and covariances of  $\mathbf{W}_t$  are given by

$$\text{E}(\mathbf{W}_t) = \frac{k_0 \text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta})}{s_0 - (p + 1)} = \text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta}), \quad (17)$$

$$\text{Var}(w_{ii,t}) = \frac{2}{\delta} \sigma_{ii,t}^2, \quad \text{Cov}(w_{ij,t}, w_{kl,t}) = \frac{2\sigma_{ij,t}\sigma_{kl,t} + (\delta + 2)(\sigma_{ik,t}\sigma_{jl,t} + \sigma_{il,t}\sigma_{jk,t})}{\delta(\delta + 3)}, \quad (18)$$

where  $w_{ij,t}$  and  $\sigma_{ij,t}$  denote the  $(i, j)$ -th element of  $\mathbf{W}_t$  and  $\text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta})$ , respectively. While keeping  $\mathbf{W}_t$  the unbiased estimator of the true covariance matrix  $\text{Cov}(\mathbf{y}_t | \mathbf{h}_t, \boldsymbol{\theta})$ , we control its variance using the precision parameter  $\delta$ . A large  $\delta$  implies a small variance for the realized covariance matrix, and consequently we place more weight on the information of the realized covariance matrix. We define the FMRSV model by Equations (1)–(18).

### 2.3 Bias correction of realized volatilities and correlations using daily returns

The realized covariance matrices are known to have estimation biases due to market microstructure noise, nontrading hours, nonsynchronous trades and so forth. In the preprocessing step, we use the information of daily returns to correct these biases for volatilities and correlation matrices, respectively.

*Bias correction of realized volatilities.* We first correct the bias of the variance following Hansen and Lunde (2005). Let  $s_i^2$  and  $\tilde{w}_{ii,t}$  respectively denote the sample variance of the

daily return and the realized volatility at time  $t$  for the  $i$ -th stock return. Compute a constant  $c_i$  such that

$$c_i = \frac{s_i^2}{\frac{1}{T} \sum_{t=1}^T \tilde{w}_{ii,t}}, \quad i = 1, \dots, p, \quad t = 1, \dots, T.$$

Define the bias-corrected realized volatility  $w_{ii,t}$  as

$$w_{ii,t} = c_i \tilde{w}_{ii,t}, \quad i = 1, \dots, p, \quad t = 1, \dots, T,$$

so that the average of bias-corrected realized volatilities is equal to the sample variance of daily returns. For example, if we ignore the overnight returns and compute the open-to-close realized volatilities, we tend to underestimate the true volatilities. By using the close-to-close daily returns, we can correct the biases as above. In Section 4, we illustrate examples where we found most of the values of  $c_i$ 's are greater than one.

*Bias correction of realized correlation matrices.* Let  $\mathbf{R}$  and  $\tilde{\mathbf{R}}_t$  respectively denote the sample correlation matrix using daily returns and the sample correlation matrix using intraday returns at time  $t$ . We compute the bias-corrected realized correlation matrix  $\mathbf{R}_t$  where it is guaranteed to be positive definite:

1. Compute the spectral decompositions of  $\mathbf{R}$  and  $\tilde{\mathbf{R}}_t$  as

$$\mathbf{R} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}', \quad \tilde{\mathbf{R}}_t = \mathbf{P}_t\mathbf{\Lambda}_t\mathbf{P}_t',$$

where the  $i$ -th column of  $\mathbf{P}$  ( $\mathbf{P}_t$ ) is the eigenvector of  $\mathbf{R}$  ( $\tilde{\mathbf{R}}_t$ ), and  $\mathbf{\Lambda}$  ( $\mathbf{\Lambda}_t$ ) is the diagonal matrix whose  $i$ -th diagonal element is the eigenvalue corresponding to the  $i$ -th columns of  $\mathbf{P}$  ( $\mathbf{P}_t$ ). We define the logarithms of  $\mathbf{R}$  and  $\tilde{\mathbf{R}}_t$  as

$$\mathbf{L}\mathbf{R} = \mathbf{P}(\log \mathbf{\Lambda})\mathbf{P}', \quad \tilde{\mathbf{L}}\mathbf{R}_t = \mathbf{P}_t(\log \mathbf{\Lambda}_t)\mathbf{P}_t',$$

where  $\log \mathbf{\Lambda}$  ( $\log \mathbf{\Lambda}_t$ ) denotes a diagonal matrix whose  $i$ -th diagonal element is a logarithm of the  $i$ -th diagonal element of  $\mathbf{\Lambda}$  ( $\mathbf{\Lambda}_t$ ).

2. Compute a constant matrix  $\mathbf{C}$  such that

$$\mathbf{C} = \mathbf{L}\mathbf{R} - \frac{1}{T} \sum_{t=1}^T \tilde{\mathbf{L}}\mathbf{R}_t,$$

and define the bias-corrected log correlation matrix  $\mathbf{L}\mathbf{R}_t$  as

$$\mathbf{L}\mathbf{R}_t = \tilde{\mathbf{L}}\mathbf{R}_t + \mathbf{C}.$$



3. Compute the bias-corrected correlation matrix  $\mathbf{R}_t$  from  $\mathbf{LR}_t$  as follows. The convergence of the algorithm is usually very fast as discussed in Archakov and Hansen (2018).

- (a) Set  $k = 0$ .
- (b) Let  $\mathbf{x}_k$  denote the vector whose  $i$ -th element is the  $i$ -th diagonal element of  $\mathbf{LR}_t$ .
- (c) Compute the spectral decomposition of  $\mathbf{LR}_t$  such that

$$\mathbf{LR}_t = \mathbf{Q}_t \mathbf{D}_t \mathbf{Q}_t',$$

where the  $i$ -th column of  $\mathbf{Q}_t$  is the eigenvector of  $\mathbf{LR}_t$ , and  $\mathbf{D}_t$  is the diagonal matrix whose  $i$ -th diagonal element is the eigenvalue corresponding to the  $i$ -th columns of  $\mathbf{Q}_t$ . Then  $\exp(\mathbf{LR}_t)$  is

$$\exp(\mathbf{LR}_t) = \mathbf{Q}_t \exp(\mathbf{D}_t) \mathbf{Q}_t',$$

Let  $\mathbf{\Delta}$  denote the vector whose  $i$ -th element is the logarithm of the  $i$ -th diagonal element of  $\exp(\mathbf{LR}_t)$ .

- (d) Update  $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{\Delta}$ . Replace the diagonal elements ( $\mathbf{x}_k$ ) of  $\mathbf{LR}_t$  with  $\mathbf{x}_{k+1}$ .
- (e) Set  $k \leftarrow k + 1$  and return to (b). Repeat until convergence is achieved.
- (f) Replace the nondiagonal elements of the identity matrix by those of  $\exp(\mathbf{LR}_t)$  and save the resulting matrix as  $\mathbf{R}_t$ .

4. Compute  $\mathbf{W}_t$  using the bias-corrected realized volatilities ( $w_{ii,t}$ 's) and the bias-corrected realized correlation matrices ( $\mathbf{R}_t$ 's).

In empirical studies, it is often pointed out that the realized correlations depend on the data sampling frequency. That is, the correlations computed from the high frequency data tend to be smaller due to the market microstructure noise than those of the daily returns. This is well known to exist in the stock market and is known as the Epps effect (Epps (1979), Yamauchi and Omori (2019)). In fact, in Figure 2 of Section 4, some of the realized correlations are shown to have such biases.

### 3 Markov chain Monte Carlo estimation

#### 3.1 Prior distributions for parameters

Since there are many parameters and latent variables in our FMRSV model, it is difficult to evaluate the likelihood and to implement the maximum likelihood estimation. Thus taking

a Bayesian approach, we estimate the parameters and conduct statistical inference using a Markov chain Monte Carlo simulation.

First, we set the prior distributions of parameters  $\boldsymbol{\theta} \equiv (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\rho}, \boldsymbol{\sigma}_\eta, \boldsymbol{\sigma}_\nu, \delta)$ . We assume the prior distributions of  $\boldsymbol{\alpha}_j$ ,  $\boldsymbol{\beta}_i$ ,  $\boldsymbol{\mu}$ , and  $\boldsymbol{\gamma}$  to follow multivariate normal distributions. The prior distributions of  $\boldsymbol{\phi}$ ,  $\boldsymbol{\psi}$  and  $\boldsymbol{\rho}$  are assumed to follow beta distributions. We assume that  $\boldsymbol{\sigma}_\eta$  and  $\boldsymbol{\sigma}_\nu$  follow independent inverse gamma distributions. We assume  $\delta$  follows a noninformative improper prior distribution. In summary, the following prior distributions are assumed:

$$\begin{aligned} \boldsymbol{\mu} &\sim \text{N}(\mathbf{m}_\mu, \mathbf{S}_\mu), \quad \boldsymbol{\gamma} \sim \text{N}(\mathbf{m}_\gamma, \mathbf{S}_\gamma), \\ \boldsymbol{\beta}_i &\sim \text{N}(\mathbf{m}_{\beta_i}, \mathbf{S}_{\beta_i}), \quad i = 1, \dots, p, \quad \boldsymbol{\alpha}_j \sim \text{N}(\mathbf{m}_{\alpha_j}, \mathbf{S}_{\alpha_j}), \quad j = 2, \dots, q, \\ \frac{1 + \phi_i}{2} &\sim \text{Beta}(a_\phi, b_\phi), \quad \sigma_{\eta,i}^2 \sim \text{IG}\left(\frac{n_\eta}{2}, \frac{d_\eta}{2}\right), \quad i = 1, \dots, p + q, \\ \frac{1 + \psi_j}{2} &\sim \text{Beta}(a_\psi, b_\psi), \quad \frac{1 + \rho_{p+j}}{2} \sim \text{Beta}(a_\rho, b_\rho), \quad \sigma_{\nu,j}^2 \sim \text{IG}\left(\frac{n_\nu}{2}, \frac{d_\nu}{2}\right), \quad j = 1, \dots, q, \\ \pi(\delta) &\propto I(\delta > 0). \end{aligned}$$

### 3.2 Markov chain Monte Carlo algorithm

Let  $\mathbf{f} = (\mathbf{f}'_1, \dots, \mathbf{f}'_T)'$ ,  $\mathbf{h} = (\mathbf{h}'_1, \dots, \mathbf{h}'_T)'$ ,  $\mathbf{x} = (\mathbf{x}'_1, \dots, \mathbf{x}'_T)'$ ,  $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_T)'$  and  $\mathbf{W} = \{\mathbf{W}_t\}_{t=1}^T$ . Furthermore, let  $\boldsymbol{\theta}_{\setminus \boldsymbol{\alpha}}$  denote  $\boldsymbol{\theta}$  excluding  $\boldsymbol{\alpha}$ . We implement the Markov chain Monte Carlo simulation as follows:

1. Initialize  $\mathbf{h}$ ,  $\mathbf{f}$  and  $\boldsymbol{\theta}$ .
2. Generate  $\mathbf{h} \mid \boldsymbol{\theta}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
3. Generate  $\mathbf{f} \mid \boldsymbol{\theta}, \mathbf{h}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
4. Generate  $\boldsymbol{\alpha} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\alpha}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
5. Generate  $\boldsymbol{\beta} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\beta}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
6. Generate  $\boldsymbol{\mu} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\mu}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
7. Generate  $\boldsymbol{\gamma} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\gamma}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
8. Generate  $\boldsymbol{\phi} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\phi}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
9. Generate  $\boldsymbol{\psi} \mid \boldsymbol{\theta}_{\setminus \boldsymbol{\psi}}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .

10. Generate  $\boldsymbol{\rho} \mid \boldsymbol{\theta} \setminus \boldsymbol{\rho}, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
11. Generate  $\boldsymbol{\sigma}_\eta \mid \boldsymbol{\theta} \setminus \boldsymbol{\sigma}_\eta, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
12. Generate  $\boldsymbol{\sigma}_\nu \mid \boldsymbol{\theta} \setminus \boldsymbol{\sigma}_\nu, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
13. Generate  $\delta \mid \boldsymbol{\theta} \setminus \delta, \mathbf{h}, \mathbf{f}, \mathbf{x}, \mathbf{y}, \mathbf{W}$ .
14. Return to Step 2.

Let  $\mathbf{h}_t^{(1)} = (h_{1t}, \dots, h_{pt})'$  and  $\mathbf{h}_t^{(2)} = (h_{p+1,t}, \dots, h_{p+q,t})'$ , and let  $\mathbf{h}^{(1)} = \{\mathbf{h}_t^{(1)}\}_{t=1}^T$  and  $\mathbf{h}^{(2)} = \{\mathbf{h}_t^{(2)}\}_{t=1}^T$ . We describe the generations of  $\boldsymbol{\beta}$  and  $\mathbf{h}^{(1)}$  below. See Appendix A for the generation of  $\mathbf{h}^{(2)}$ , and the supplementary material for other steps.

### 3.2.1 Generation of $\boldsymbol{\beta}$

The logarithm of the conditional posterior density of  $\boldsymbol{\beta}$  given the other parameters and latent variables is

$$\begin{aligned} & \log \pi(\boldsymbol{\beta} \mid \cdot) \\ &= \text{const} + \frac{s_0}{2} \sum_{t=1}^T g_t(\boldsymbol{\beta}_i) - \frac{k_0}{2} \text{tr} \left( \sum_{t=1}^T \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' \mathbf{W}_t^{-1} \right) - \frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t)' \mathbf{V}_{1t}^{-1} (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t). \end{aligned} \tag{19}$$

where  $g_t(\boldsymbol{\beta}_i) \equiv \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|$ . Since the logarithm of the determinant component,  $g_t(\boldsymbol{\beta}_i)$ , cannot be transformed to some well-known density form, we can construct a proposal distribution for the Metropolis-Hastings (MH) algorithm without this term and adjust it by the acceptance probability in the MH algorithm. However, it results in inefficient sampling and we need to approximate  $g_t(\boldsymbol{\beta}_i)$  using some known density to improve the sampling efficiency. In order to approximate it using the normal density, we consider Taylor expansion around  $\hat{\boldsymbol{\beta}}_i$ , the mode of the conditional posterior density,

$$g_t(\boldsymbol{\beta}_i) \approx g_t(\hat{\boldsymbol{\beta}}_i) + \mathbf{g}'_t(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) - \frac{1}{2} (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)' \mathbf{G}_t^{-1} (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i).$$

where

$$\mathbf{g}_t = \left[ \frac{\partial g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} \right]_{\boldsymbol{\beta}_i = \hat{\boldsymbol{\beta}}_i}, \quad \mathbf{G}_t^{-1} = - \left[ \frac{\partial^2 g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i'} \right]_{\boldsymbol{\beta}_i = \hat{\boldsymbol{\beta}}_i}.$$

It can be shown that

$$\begin{aligned}\frac{\partial g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} &= 2 \mathbf{V}_{2t} \mathbf{B}' (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{e}_i, \\ \frac{\partial^2 g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i'} &= 2d_{ii} \{ \mathbf{V}_{2t} - \mathbf{V}_{2t} \mathbf{B}' (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{B} \mathbf{V}_{2t} \} - \frac{1}{2} \left[ \frac{\partial g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} \right] \left[ \frac{\partial g_t(\boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} \right]'.\end{aligned}$$

where  $\mathbf{e}_i$  denotes a  $p \times 1$  vector with the  $i$ -th element equal to one and zero otherwise, and  $d_{ii}$  is the  $(i, i)$ -th element of  $(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1}$  (the proof is given by Proposition 1 of Appendix B.1). Further, let  $w_t^{ij}$  denote the  $(i, j)$ -th element of  $\mathbf{W}_t^{-1}$ . Noting that

$$\begin{aligned}\text{tr}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' \mathbf{W}_t^{-1}) &= \text{vec}(\mathbf{B}')' (\mathbf{W}_t^{-1} \otimes \mathbf{V}_{2t}) \text{vec}(\mathbf{B}'), \\ &= \text{const} + \boldsymbol{\beta}_i' (w_t^{ii} \mathbf{V}_{2t}) \boldsymbol{\beta}_i + 2 \boldsymbol{\beta}_i' (\mathbf{V}_{2t} \sum_{i \neq j} w_t^{ij} \boldsymbol{\beta}_j),\end{aligned}\quad (20)$$

$$\begin{aligned}\text{tr}(\mathbf{f}_t' \mathbf{B}' \mathbf{V}_{1t}^{-1} \mathbf{B} \mathbf{f}_t) &= \text{tr}(\mathbf{B} \mathbf{f}_t \mathbf{f}_t' \mathbf{B}' \mathbf{V}_{1t}^{-1}) \\ &= \text{const} + \boldsymbol{\beta}_i' \{ \exp(-h_{it}) \mathbf{f}_t \mathbf{f}_t' \} \boldsymbol{\beta}_i,\end{aligned}\quad (21)$$

$$\begin{aligned}\text{tr}\{\mathbf{y}_t' \mathbf{V}_{1t}^{-1} \mathbf{B} \mathbf{f}_t\} &= \text{tr}\{\mathbf{f}_t \mathbf{y}_t' \mathbf{V}_{1t}^{-1} \mathbf{B}\} = \text{vec}(\mathbf{B}')' \text{vec}\{\mathbf{f}_t \mathbf{y}_t' \mathbf{V}_{1t}^{-1}\}, \\ &= \text{const} + \boldsymbol{\beta}_i' \{ y_{it} \exp(-h_{it}) \mathbf{f}_t \},\end{aligned}\quad (22)$$

we obtain the normal approximation for the conditional posterior density

$$\log \pi(\boldsymbol{\beta} | \cdot) \approx \text{const} - \frac{1}{2} (\boldsymbol{\beta} - \hat{\mathbf{m}}_{\beta_i})' \hat{\boldsymbol{\Sigma}}_{\beta_i} (\boldsymbol{\beta} - \hat{\mathbf{m}}_{\beta_i}) + r(\boldsymbol{\beta}_i), \quad (23)$$

where

$$\begin{aligned}\hat{\mathbf{m}}_{\beta_i} &= \hat{\boldsymbol{\Sigma}}_{\beta_i} \left[ \sum_{t=1}^T \frac{s_0}{2} \mathbf{G}_t^{-1} (\hat{\boldsymbol{\beta}}_i + \mathbf{G}_t \mathbf{g}_t) - k_0 \sum_{t=1}^T \mathbf{V}_{2t} \left( \sum_{i \neq j} w_t^{ij} \boldsymbol{\beta}_j \right) + \sum_{t=1}^T y_{it} \exp(-h_{it}) \mathbf{f}_t + \mathbf{S}_{\beta_i}^{-1} \mathbf{m}_{\beta_i} \right] \\ \hat{\boldsymbol{\Sigma}}_{\beta_i} &= \left[ \frac{s_0}{2} \sum_{t=1}^T \mathbf{G}_t^{-1} + k_0 \sum_{t=1}^T w_t^{ii} \mathbf{V}_{2t} + \sum_{t=1}^T \exp(-h_{it}) \mathbf{f}_t \mathbf{f}_t' + \mathbf{S}_{\beta_i}^{-1} \right]^{-1}, \\ r(\boldsymbol{\beta}_i) &= \frac{s_0}{2} \sum_{t=1}^T \left\{ g_t(\boldsymbol{\beta}_i) - \mathbf{g}_t'(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) + \frac{1}{2} (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)' \mathbf{G}_t^{-1} (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) \right\}.\end{aligned}$$

When the current value is  $\boldsymbol{\beta}_i^o$ , we generate  $\boldsymbol{\beta}_i^n$  from  $\text{N}(\hat{\mathbf{m}}_{\beta_i}, \hat{\boldsymbol{\Sigma}}_{\beta_i})$  and accept  $\boldsymbol{\beta}_i^n$  with probability  $\min\{1, \exp(r(\boldsymbol{\beta}_i^n) - r(\boldsymbol{\beta}_i^o))\}$ .

### 3.2.2 Generation of $\mathbf{h}^{(1)}$

The log conditional posterior density of  $\mathbf{h}^{(1)}$  given other latent variables and parameters is

$$\begin{aligned}
& \log \pi(\mathbf{h}^{(1)} | \cdot) \\
&= \text{const} + \frac{s_0}{2} \sum_{t=1}^T \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} \sum_{t=1}^T \text{tr}(\mathbf{V}_{1t} \mathbf{W}_t^{-1}) \\
&\quad - \frac{1}{2} \sum_{t=1}^T \{ \log |\mathbf{V}_{1t}| + (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t)' \mathbf{V}_{1t}^{-1} (\mathbf{y}_t - \mathbf{B} \mathbf{f}_t) \} - \frac{1}{2} \{ \mathbf{h}_1 - \boldsymbol{\mu} \}' \boldsymbol{\Sigma}_{h,0}^{-1} (\mathbf{h}_1 - \boldsymbol{\mu}) \\
&\quad - \frac{1}{2} \sum_{t=1}^{T-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t \}' \boldsymbol{\Sigma}_{\eta}^{-1} \{ \mathbf{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi}) \boldsymbol{\mu} - \boldsymbol{\Phi} \mathbf{h}_t \} \\
&= \text{const} - \sum_{i=1}^p \frac{(1 - \phi_i^2)(h_{i1} - \mu_i)^2}{2\sigma_{\eta,i}^2} - \sum_{t=1}^{T-1} \sum_{i=1}^p \frac{\{h_{i,t+1} - h_{it} - (1 - \phi_i)\mu_i\}^2}{2\sigma_{\eta,i}^2} + \sum_{t=1}^T l_t, \quad (24)
\end{aligned}$$

where

$$l_t = \frac{s_0}{2} \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} w_t^{ii} \exp(h_{it}) - \frac{1}{2} \sum_{i=1}^p \{ h_{it} + (y_{it} - \beta_i' \mathbf{f}_t)^2 \exp(-h_{it}) \},$$

and  $w_t^{ii}$  denotes the  $(i, i)$ -th element of  $\mathbf{W}_t^{-1}$ . Although it is simple and easy to implement a single move sampler that generates a single state variable  $h_{it}$  ( $i = 1, \dots, p$ ,  $t = 1, \dots, T$ ) at a time, it would result in inefficient sampling. That is, it is well known to generate highly autocorrelated samples when state variables are highly correlated as in stochastic volatility models. A multi-move sampler, which generates a block of state variables using stochastic knots (e.g. Shephard and Pitt (1997), Watanabe and Omori (2004)), is one of the most efficient ways to generate latent state variables with high autocorrelations. In the multi-move sampler, first, we divide  $(h_{i,1}, \dots, h_{i,T})$  into  $K$  blocks  $(h_{i,s_k}, \dots, h_{i,s_{k+1}-1})$ , where  $k = 1, \dots, K$  with  $1 = s_1 < s_2 < \dots < s_{K+1} = T + 1$ . Then we approximate the nonlinear Gaussian state space model using the linear Gaussian state space model to sample from the conditional posterior distribution of the state variables for each block. To sample the state variables  $(h_{i,s}, \dots, h_{i,s+m})$  from their conditional posterior distribution efficiently, we consider sampling the corresponding disturbances  $(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$ . The log posterior density of  $(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$ , ( $i = 1, \dots, p$ ) is

$$\begin{aligned}
\log f(\eta_{i,s-1}, \dots, \eta_{i,s+m-1} | \cdot) &= \text{const} - \frac{1}{2} \sum_{t=s-1}^{t=s+m-1} \eta_{it}^2 + \sum_{t=s}^{s+m} l_{it} \\
&\quad - \frac{\{h_{i,s+m+1} - \phi_i h_{i,s+m} - (1 - \phi_i)\mu_i\}^2}{2\sigma_{\eta,i}^2} I(s+m < T), \quad (25)
\end{aligned}$$

where

$$l_{it} = \frac{s_0}{2} \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} w_t^{ii} \exp(h_{it}) - \frac{1}{2} \{h_{it} + (y_{it} - \boldsymbol{\beta}'_i \mathbf{f}_t)^2 \exp(-h_{it})\},$$

and  $I(A)$  is an indicator function such that  $I(A) = 1$  if  $A$  is true and 0 otherwise. By Taylor expansion of  $l_{it}$  around the conditional mode,  $\hat{h}_{it}$ , the approximated conditional posterior density  $f^*(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$  of the block is given by

$$\begin{aligned} \log f^*(\eta_{i,s-1}, \dots, \eta_{i,s+m-1}) = & \text{const} - \frac{1}{2} \sum_{t=s-1}^{t=s+m-1} \eta_{it}^2 + \sum_{t=s}^{t=s+m} \left[ \hat{l}_{it} + (h_{it} - \hat{h}_{it}) \hat{l}'_{it} + \frac{1}{2} (h_{it} - \hat{h}_{it})^2 \hat{l}''_{it} \right], \\ & - \frac{\{h_{i,s+m+1} - \phi_i h_{i,s+m} - (1 - \phi_i) \mu_i\}^2}{2\sigma_{\eta,i}^2} I(s+m < T). \end{aligned} \quad (26)$$

where  $\hat{l}'_{it}$  and  $\hat{l}''_{it}$  are

$$\begin{aligned} \hat{l}'_{it} &= \frac{s_0}{2} d_{ii} \exp(h_{it}) - \frac{k_0}{2} w_t^{ii} \exp(h_{it}) + \frac{1}{2} \{-1 + (y_{it} - \boldsymbol{\beta}'_i \mathbf{f}_t)^2 \exp(-h_{it})\}, \\ \hat{l}''_{it} &= \frac{s_0}{2} d_{ii} \exp(h_{it}) - \frac{s_0}{2} d_{ii}^2 \exp(2h_{it}) - \frac{k_0}{2} w_t^{ii} \exp(h_{it}) - \frac{1}{2} (y_{it} - \boldsymbol{\beta}'_i \mathbf{f}_t)^2 \exp(-h_{it}), \end{aligned}$$

respectively, evaluated at  $h_{it} = \hat{h}_{it}$  using

$$\frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{it}} = d_{ii} \exp(h_{it}), \quad \frac{\partial^2 \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{it}^2} = d_{ii} \exp(h_{it}) - d_{ii}^2 \exp(2h_{it}),$$

for  $i = 1, \dots, p$  and  $t = 1, \dots, T$  (the proof is given by Proposition 2 of Appendix B.2 ). To construct the approximated linear Gaussian state space model from which we sample a proposal of  $(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$ , we define the auxiliary variables  $\hat{y}_{it}$  and  $v_{it}$  as follows. For  $t = s, \dots, s+m-1$  or  $t = s+m = T$ ,

$$\hat{y}_{it} = \hat{h}_{it} + v_{it} \hat{l}'_{it}, \quad v_{it} = -\hat{l}''_{it}^{-1}, \quad (27)$$

and, for  $t = s+m < T$ ,

$$\hat{y}_{it} = v_{it} \left[ \left\{ \hat{l}'_{it} - \hat{l}''_{it} \hat{h}_{it} \right\} + \phi_i \sigma_{\eta,i}^{-2} \{h_{t+1,i} - (1 - \phi_i) \mu_i\} \right], \quad v_{it} = \left( \phi_i^2 \sigma_{\eta,i}^{-2} - \hat{l}''_{it} \right)^{-1}. \quad (28)$$

Then, consider the following linear Gaussian state space model,

$$\hat{y}_{it} = h_{it} + \epsilon_t, \quad \epsilon_t \sim \text{N}(0, v_{it}), \quad (29)$$

$$h_{i,t+1} = (1 - \phi_i) \mu_i + \phi_i h_{it} + \eta_{it}, \quad \eta_{it} \sim \text{N}(0, \sigma_{\eta,i}). \quad (30)$$

Given  $h_{i,s-1}, (\hat{y}_{is}, \dots, \hat{y}_{i,s+m})$  and other parameters, we can generate the candidate of  $(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$  from the approximated density  $f^*$  using a simulation smoother (e.g.

de Jong and Shephard (1995), Durbin and Koopman (2002)), and conduct the MH algorithm. That is, when we have the current samples  $(\eta_{i,s-1}^o, \dots, \eta_{i,s+m-1}^o)$ , we accept the new samples  $(\eta_{i,s-1}^n, \dots, \eta_{i,s+m-1}^n)$  generated from  $f^*(\eta_{i,s-1}, \dots, \eta_{i,s+m-1})$  with probability

$$\min \left\{ 1, \frac{f(\eta_{i,s-1}^n, \dots, \eta_{i,s+m-1}^n) f^*(\eta_{i,s-1}^o, \dots, \eta_{i,s+m-1}^o)}{f(\eta_{i,s-1}^o, \dots, \eta_{i,s+m-1}^o) f^*(\eta_{i,s-1}^n, \dots, \eta_{i,s+m-1}^n)} \right\}.$$

To obtain the conditional mode  $(\hat{h}_{i,s}, \dots, \hat{h}_{i,s+m})$ , we select some initial mode values such as the current state vector of  $(h_{i,s}, \dots, h_{i,s+m})$  and repeat the disturbance smoother several times (see e.g. Shephard and Pitt (1997), Watanabe and Omori (2004)).

Remark. Parameter  $K$  is chosen to obtain stable and efficient estimation results. In our empirical study, we used  $K = 470$ , but, in general, a smaller  $K$  could be used.

## 4 Empirical studies

### 4.1 Data

*Data.* In this section, we apply our proposed model to the daily returns of ten U.S. stocks with the bias-corrected realized covariance matrices. The ten series of stock returns are chosen from top stocks composing the exchange traded fund (ETF) that seeks to track the performance of a benchmark index that measures the investment return of the S&P500 index<sup>1</sup>. They are 1: Apple Inc. (AAPL), 2: Microsoft Corp. (MSFT), 3: Amazon.com Inc. (AMZN), 4: JPMorgan Chase & Co. (JPM), 5: Berkshire Hathaway Inc. Class B (BRKB), 6: Alphabet Inc. Class A (GOOGL), 7: Johnson & Johnson (JNJ), 8: Proctor & Gamble Co. (PG), 9: Exxon Mobil Corp (XOM), and 10: AT&T (T). The federal funds rate is used as a risk-free asset. The daily returns for the  $i$ -th stocks are defined as  $y_{it} = 100 \times (\log p_{it} - \log p_{i,t-1})$ , where  $p_{it}$  is the closing price of the  $i$ -th asset at time  $t$ . The (open-to-close) realized covariance matrices are computed as  $\sum_{s=1}^{78} \mathbf{r}_{st} \mathbf{r}'_{st}$ , where  $\mathbf{r}_{st}$  is the  $s$ -th return vector during day  $t$  at intervals of 5 minutes from 9:35 to 16:00<sup>2</sup>. The sample period is from September 1, 2004 to December 31, 2013, and the number of observations is  $T = 2350$ .

*Realized factor.* The time series plots of  $y_{it}$ 's are shown in Figure 1, which shows that there is a very high volatility period in 2008 ( the time of the global financial crisis when Lehman

<sup>1</sup>The ETF is Vanguard S&P 500 ETF (VOO).

<sup>2</sup>The intraday price data was obtained from Tick Data (<http://www.tickdata.com>).

Brothers filed for Chapter 11 bankruptcy protection). We can see the co-movement among the ten stock returns and the S&P500 index. Since the stock index is considered to represent the overall movement in the stock market, we use the S&P500 index as the realized factor  $x_t$  which corresponds to the U.S. stock market factor; and we set  $q = 1$  and  $\mathbf{A} = \mathbf{1}$  in this analysis.

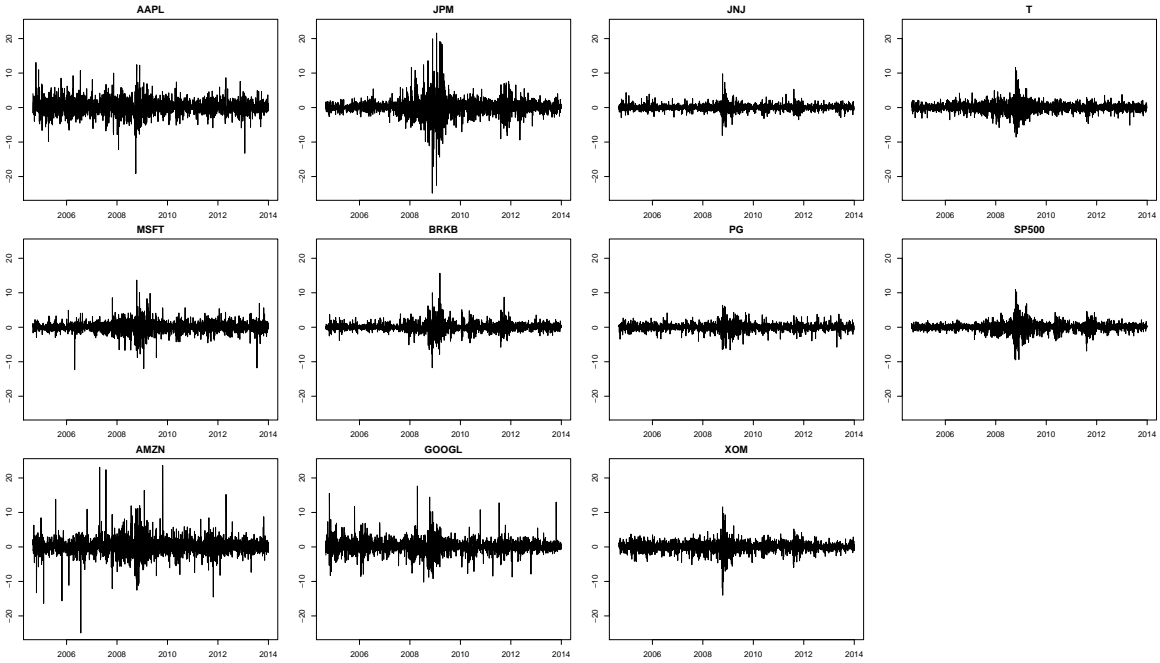


Figure 1: Time series plots of ten U.S. stock (close-to-close) returns and S&P 500 return.

*Bias correction of the realized volatilities and correlations.* The bias correction vector  $\mathbf{c}$  for the realized volatilities is obtained as

$$\mathbf{c} = (1.53, 1.39, 1.55, 1.37, 1.02, 1.50, 1.04, 1.06, 1.21, 0.99)',$$

and the realized volatilities (except for T) appear to underestimate the true volatility due to ignoring the overnight returns. The bias correction matrix  $\mathbf{C}$  for the log correlation matrices is omitted since it is difficult to interpret intuitively; instead, we show the boxplots of the differences between the bias-corrected realized correlation and the raw realized correlation in Figure 2. Most of the bias-corrected realized correlations are found to be larger (except those for AAPL-AMZN (1-3) and AMZN-GOOGL (3-6)) than the realized correlations, indicating the existence of the Epps effect.



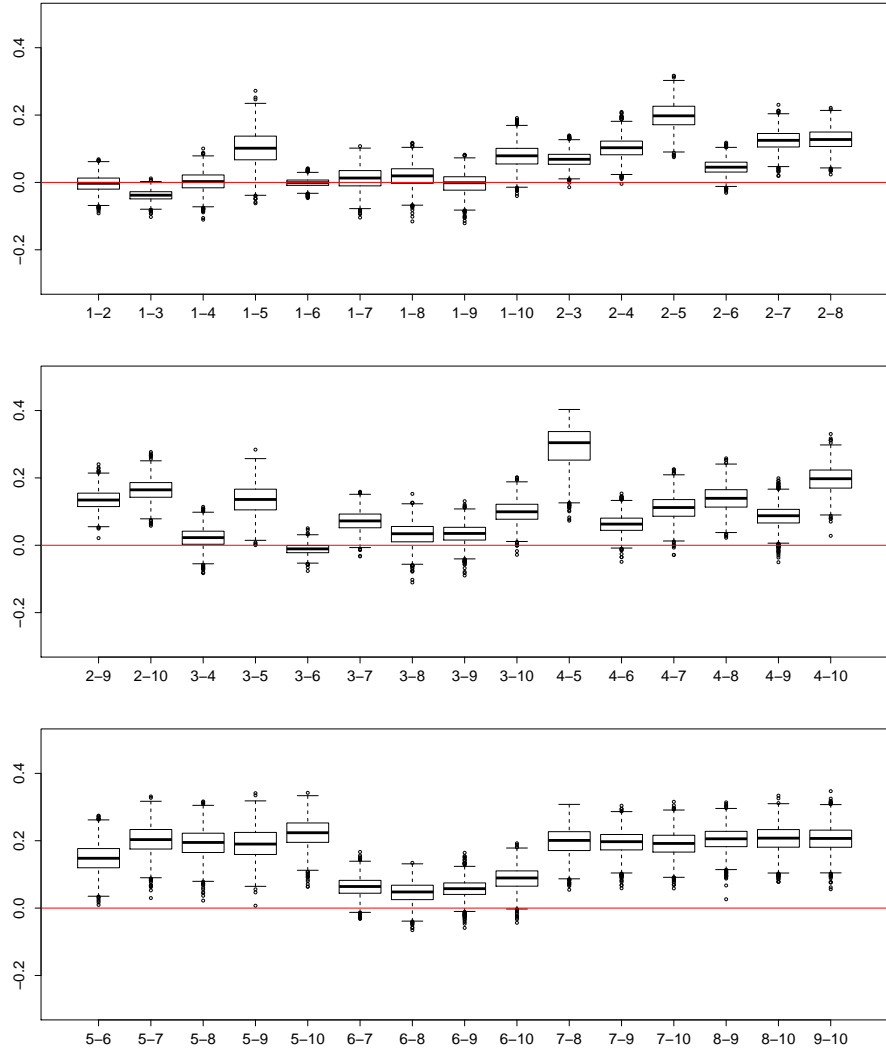


Figure 2: Boxplot of  $\mathbf{R}_t - \tilde{\mathbf{R}}_t$ . Differences between the bias-corrected and raw realized correlations.

The prior distributions are assumed as follows:

$$\begin{aligned}
 \mu_i &\sim N(0, 4), \quad \gamma_j \sim N(0, 1), \quad \beta_i \sim N(0, 1), \\
 \frac{1 + \phi_i}{2} &\sim \text{Beta}(20, 1.5), \quad \sigma_{\eta,i}^2 \sim \text{IG}\left(\frac{0.1}{2}, \frac{0.1}{2}\right), \\
 \frac{1 + \psi_j}{2} &\sim \text{Beta}(1, 1), \quad \frac{1 + \rho_{p+j}}{2} \sim \text{Beta}(1, 1), \quad \sigma_{\nu,j}^2 \sim \text{IG}\left(\frac{0.1}{2}, \frac{0.1}{2}\right), \\
 \pi(\delta) &\propto I(\delta > 0).
 \end{aligned}$$

## 4.2 Estimation results

The MCMC simulation is iterated to obtain 20,000 posterior samples after discarding 10,000 samples as the burn-in period for the FMSV model, and 10,000 posterior samples after discarding 2,000 samples as the burn-in period for the FMRSV model. Tables 1 and 2 show the posterior estimation results of the parameters for the FMSV model (without realized covariances) and the FMRSV model (with realized covariances), respectively. The inefficiency factors<sup>3</sup> range from 1 to 114 (the effective sample sizes range from 175 to 20,000) for the FMSV model, and from 1 to 94 (the effective sample sizes are from 106 to 10,000) for the FMRSV model, which implies that our sampling algorithm works quite efficiently.

The conditional means of the log volatilities ( $\mu_j$ ) vary from  $-1.105$  (JNJ) to  $0.804$  (AAPL) in the FMSV model and from  $-1.047$  (JNJ) to  $1.240$  (AMZN) in the FMRSV model. Overall, these posterior means of the FMRSV model are larger than those of the FMSV model. In both models,  $\mu_1$  (AAPL) and  $\mu_3$  (AMZN) are the largest two, while  $\mu_7$  (JNJ) is the smallest, as expected from Figure 1.

The persistences of the log volatilities ( $\phi_j$ ) are high from  $0.619$  (AMZN) to  $0.982$  (factor) in the FMSV model and from  $0.640$  (AMZN) to  $0.927$  (XOM) in the FMRSV model, where the estimates of the FMRSV model are relatively lower than those of the FMSV model. The persistences ( $\phi_{11}$ ) of the log volatilities of the dynamic factor are higher than those for most of the stock returns. The factor loadings ( $\beta_j$ ) are all positive, ranging from  $0.470$  (JNJ) to  $1.260$  (JPM) in the FMSV model and from  $0.521$  (JNJ) to  $1.126$  (JPM) in the FMRSV model, suggesting co-movement between stock returns. Among the factor loadings, the estimates of AMZN and JPM are found to be the largest. The posterior probability that the factor mean ( $\gamma_1$ ) is positive is greater than  $0.975$  since the 95% credible interval is above 0. This implies that the expected market return is positive during this sample period. The autoregressive coefficient ( $\psi_1$ ) is estimated to be negative but close to zero.

The leverage effect of the factor, denoted by  $\rho$ , is estimated to be  $-0.609$  in the FMSV model, and  $-0.234$  in the FMRSV model. It is expected to be negative with the posterior probability greater than  $0.975$ , suggesting the existence of the leverage effect. The absolute

---

<sup>3</sup>The inefficiency factor is defined as  $1 + 2 \sum_{g=1}^{\infty} \rho(g)$ , where  $\rho(g)$  is the sample autocorrelation at lag  $g$ . This is interpreted as the ratio of the numerical variance of the posterior mean from the chain to the variance of the posterior mean from hypothetical uncorrelated draws. The smaller the inefficiency factor is, the closer the MCMC sampling is to the uncorrelated sampling. The effective sample size is obtained as the posterior sample size divided by the inefficiency factor.

value of the posterior estimate is found to be much smaller in the FMRSV model.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
$\mu_1$	0.804	[0.653, 0.960]	6	$\phi_1$	0.823	[0.759, 0.879]	68
$\mu_2$	-0.281	[-0.433, -0.124]	5	$\phi_2$	0.799	[0.729, 0.858]	71
$\mu_3$	0.767	[0.645, 0.893]	8	$\phi_3$	0.619	[0.503, 0.708]	70
$\mu_4$	0.067	[-0.453, 0.575]	1	$\phi_4$	0.976	[0.960, 0.989]	64
$\mu_5$	-0.478	[-0.856, -0.103]	1	$\phi_5$	0.971	[0.954, 0.985]	59
$\mu_6$	0.292	[0.117, 0.467]	4	$\phi_6$	0.832	[0.775, 0.880]	67
$\mu_7$	-1.105	[-1.272, -0.936]	5	$\phi_7$	0.864	[0.811, 0.908]	65
$\mu_8$	-0.777	[-0.927, -0.622]	5	$\phi_8$	0.858	[0.801, 0.904]	79
$\mu_9$	-0.398	[-0.761, -0.028]	2	$\phi_9$	0.977	[0.961, 0.989]	62
$\mu_{10}$	-0.423	[-0.623, -0.220]	3	$\phi_{10}$	0.932	[0.893, 0.961]	87
$\mu_{11}^\dagger$	-0.382	[-0.867, 0.042]	2	$\phi_{11}^\dagger$	0.982	[0.972, 0.990]	41
$\beta_1$	0.934	[0.880, 0.990]	3	$\sigma_{\eta,1}$	0.572	[0.470, 0.684]	87
$\beta_2$	0.835	[0.798, 0.873]	6	$\sigma_{\eta,2}$	0.656	[0.546, 0.778]	90
$\beta_3$	1.146	[1.088, 1.205]	5	$\sigma_{\eta,3}$	0.896	[0.791, 1.024]	81
$\beta_4$	1.260	[1.209, 1.311]	5	$\sigma_{\eta,4}$	0.277	[0.216, 0.350]	110
$\beta_5$	0.687	[0.646, 0.727]	6	$\sigma_{\eta,5}$	0.243	[0.189, 0.305]	101
$\beta_6$	0.865	[0.823, 0.907]	5	$\sigma_{\eta,6}$	0.649	[0.548, 0.756]	88
$\beta_7$	0.470	[0.445, 0.495]	4	$\sigma_{\eta,7}$	0.497	[0.407, 0.594]	89
$\beta_8$	0.516	[0.486, 0.546]	4	$\sigma_{\eta,8}$	0.472	[0.386, 0.569]	102
$\beta_9$	0.860	[0.828, 0.891]	3	$\sigma_{\eta,9}$	0.186	[0.142, 0.241]	109
$\beta_{10}$	0.678	[0.645, 0.711]	4	$\sigma_{\eta,10}$	0.301	[0.234, 0.389]	114
$\psi_1^\dagger$	-0.059	[-0.102, -0.016]	4	$\sigma_{\eta,11}^\dagger$	0.194	[0.159, 0.234]	88
$\gamma_1^\dagger$	0.051	[0.022, 0.081]	4	$\sigma_{\nu,1}^\dagger$	0.136	[0.110, 0.161]	78
				$\rho_1^\dagger$	-0.609	[-0.705, -0.489]	44

Table 1: FMSV model. The posterior means, 95 % credible intervals and inefficiency factors (IFs) of the parameters for ten U.S. stocks returns.

†: parameters related to the market factor.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
$\mu_1$	0.898	[0.794, 1.002]	1	$\phi_1$	0.781	[0.752, 0.808]	3
$\mu_2$	0.113	[0.0368, 0.190]	1	$\phi_2$	0.769	[0.736, 0.801]	6
$\mu_3$	1.240	[1.178, 1.303]	2	$\phi_3$	0.640	[0.603, 0.676]	6
$\mu_4$	0.355	[0.219, 0.489]	1	$\phi_4$	0.815	[0.791, 0.840]	3
$\mu_5$	-0.643	[-0.794, -0.489]	1	$\phi_5$	0.849	[0.824, 0.872]	4
$\mu_6$	0.615	[0.503, 0.729]	1	$\phi_6$	0.820	[0.794, 0.846]	4
$\mu_7$	-1.047	[-1.134, -0.960]	1	$\phi_7$	0.803	[0.773, 0.834]	6
$\mu_8$	-0.804	[-0.890, -0.718]	1	$\phi_8$	0.799	[0.768, 0.829]	5
$\mu_9$	-0.280	[-0.451, -0.106]	1	$\phi_9$	0.927	[0.909, 0.945]	6
$\mu_{10}$	-0.551	[-0.661, -0.443]	1	$\phi_{10}$	0.845	[0.819, 0.870]	5
$\mu_{11}^\dagger$	-0.278	[-0.433, -0.127]	4	$\phi_{11}^\dagger$	0.900	[0.881, 0.919]	6
$\beta_1$	0.787	[0.770, 0.804]	20	$\sigma_{\eta,1}$	0.557	[0.535, 0.579]	6
$\beta_2$	0.858	[0.843, 0.875]	34	$\sigma_{\eta,2}$	0.420	[0.399, 0.443]	11
$\beta_3$	1.100	[1.078, 1.123]	24	$\sigma_{\eta,3}$	0.533	[0.513, 0.553]	12
$\beta_4$	1.126	[1.105, 1.147]	33	$\sigma_{\eta,4}$	0.613	[0.592, 0.636]	8
$\beta_5$	0.606	[0.595, 0.618]	36	$\sigma_{\eta,5}$	0.566	[0.542, 0.592]	12
$\beta_6$	0.808	[0.792, 0.825]	26	$\sigma_{\eta,6}$	0.499	[0.478, 0.521]	10
$\beta_7$	0.521	[0.512, 0.532]	36	$\sigma_{\eta,7}$	0.417	[0.394, 0.441]	13
$\beta_8$	0.557	[0.546, 0.568]	35	$\sigma_{\eta,8}$	0.416	[0.395, 0.438]	11
$\beta_9$	0.796	[0.783, 0.810]	37	$\sigma_{\eta,9}$	0.310	[0.290, 0.331]	16
$\beta_{10}$	0.694	[0.682, 0.706]	35	$\sigma_{\eta,10}$	0.415	[0.393, 0.437]	10
$\psi_1^\dagger$	-0.0677	[-0.107, -0.0281]	4	$\sigma_{\eta,11}^\dagger$	0.362	[0.343, 0.382]	16
$\gamma_1^\dagger$	0.0902	[0.0604, 0.121]	2	$\sigma_{\nu,1}^\dagger$	0.132	[0.108, 0.159]	58
$\delta$	16.478	[16.243, 16.692]	94	$\rho_1^\dagger$	-0.234	[-0.288, -0.181]	8

Table 2: FMRSV model. The posterior means, 95 % credible intervals and inefficiency factors (IFs) of the parameters for ten U.S. stocks returns.

†: parameters related to the market factor.

The posterior mean of the precision parameter  $\delta$  is estimated to be large at approximately 16.5 in the FMRSV model. This result suggests that the distribution of the realized covariances is concentrated around the expected value of the true covariance matrix (under our

factor model), and that the model fit is good for the measurement equation of the realized covariance matrix.

Figures 3 and 4 are the time series plots of the log idiosyncratic volatilities in the two models. The credible intervals are narrower and more stable in the FMRSV model than those in the FMSV model. Similar results are found for the dynamic correlations (see the supplementary material). This finding implies that the additional information of the realized covariances enables us to estimate the true volatilities and correlations more accurately. Moreover, Figure 5 shows the posterior means of the estimated volatilities for ten stock returns. For example, the estimates of the return volatilities of AMZN in the FMSV model are overall smooth with many large irregular jumps throughout the sample period, while those in the FMRSV model are large only around the time of the global financial crisis with small irregular jumps. These results also confirm the usefulness of the additional information of realized covariances.

In Figure 6, we compare boxplots of the correlation coefficients between AAPL and MSFT during the three periods: (1) period 1: from September 1, 2004, to October 5, 2007, (2) period 2: from October 8, 2007, to November 9, 2010, and (3) period 3: from November 10, 2010, to December 31, 2013. The estimated correlations are high for both models during period 2, which includes the global financial crisis. Overall, the correlations in the FMSV model are estimated to be larger and more dispersed than those of the FMRSV models (the boxplots of the difference between the correlation estimates of the FMSV and FMRSV models are shown in the supplementary material). This indicates that we overestimate the correlations among stock returns, especially during volatile markets when we do not use the information of the realized covariances.

The heatmaps of the posterior means of all correlations between the ten stocks for the above three periods are also shown in Figure 7. All posterior means of the correlations are found to be positive and suggest the existence of a common market factor. For the FMSV model, in period 1, MSFT and JPM have larger correlations with other stock returns; while in periods 2 and 3, XOM seems to have the largest correlations. However, for the FMRSV model, we do not observe such significant differences among correlations. As is also indicated by the box plots of the correlations between AMZN and MSFT in the three periods, most correlations are larger in the FMSV model than in the FMRSV model for all periods. During period 2, which includes the global financial crisis, all correlations are the highest, suggesting the co-movement of all stock returns through the market factor.

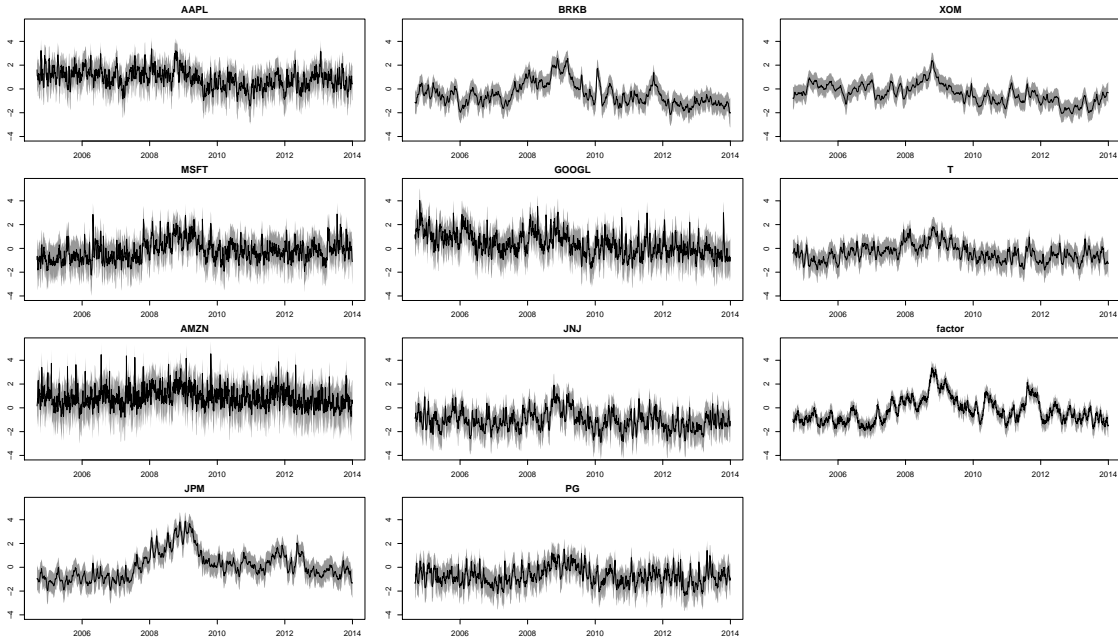


Figure 3: The posterior means (solid lines) and 95% credible intervals (shaded areas) of each estimated log volatility of the stocks and the latent factor in the FMSV model.

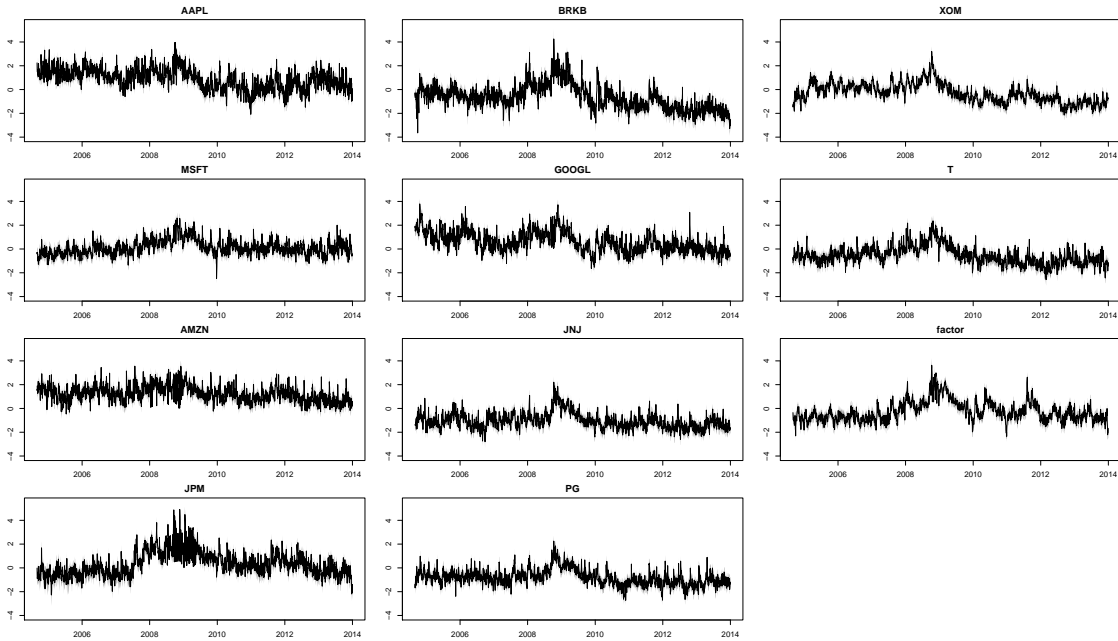


Figure 4: The posterior means (solid lines) and 95% credible intervals (shaded areas) of the estimated idiosyncratic log volatilities of the stocks and the latent factor in the FMRSV model.

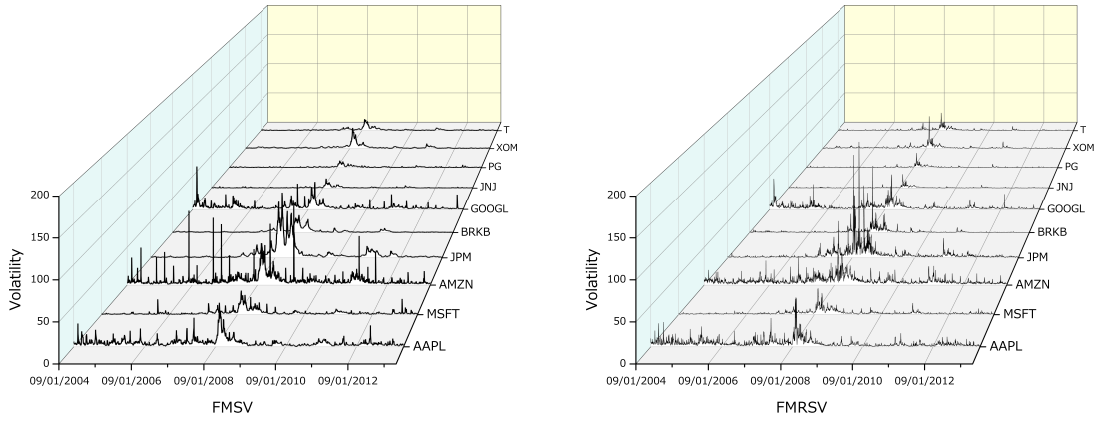


Figure 5: The posterior means of the estimated volatilities.  
 Left: FMSV model. Right: FMRSV model.

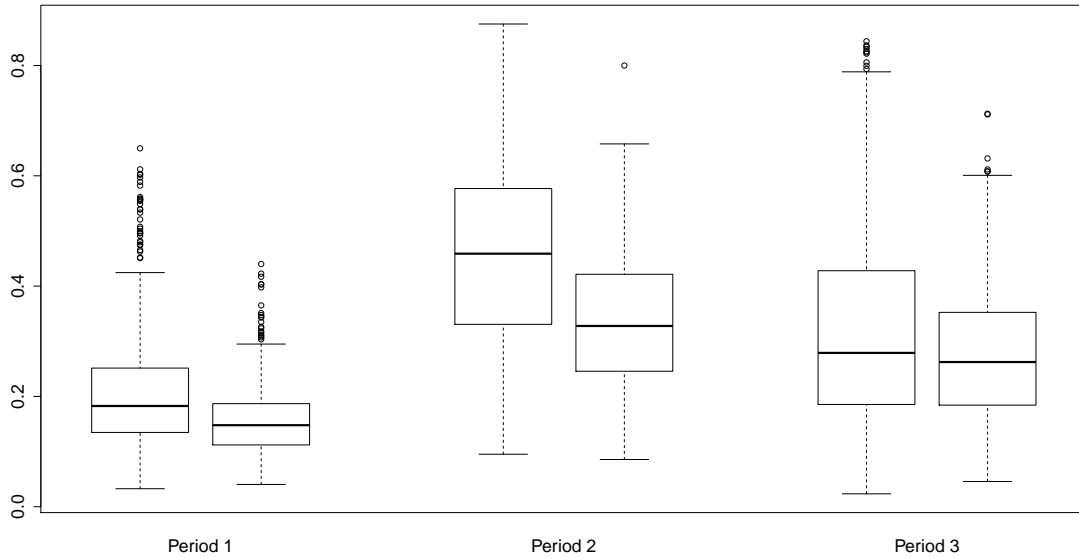


Figure 6: Boxplots of the posterior means of the correlation between AAPL and MSFT.  
 1: 9/1/2004–10/5/2007. 2: 10/8/2007–11/9/2010. 3: 11/10/2010–12/31/2013.  
 Left: FMSV model. Right: FMRSV model.

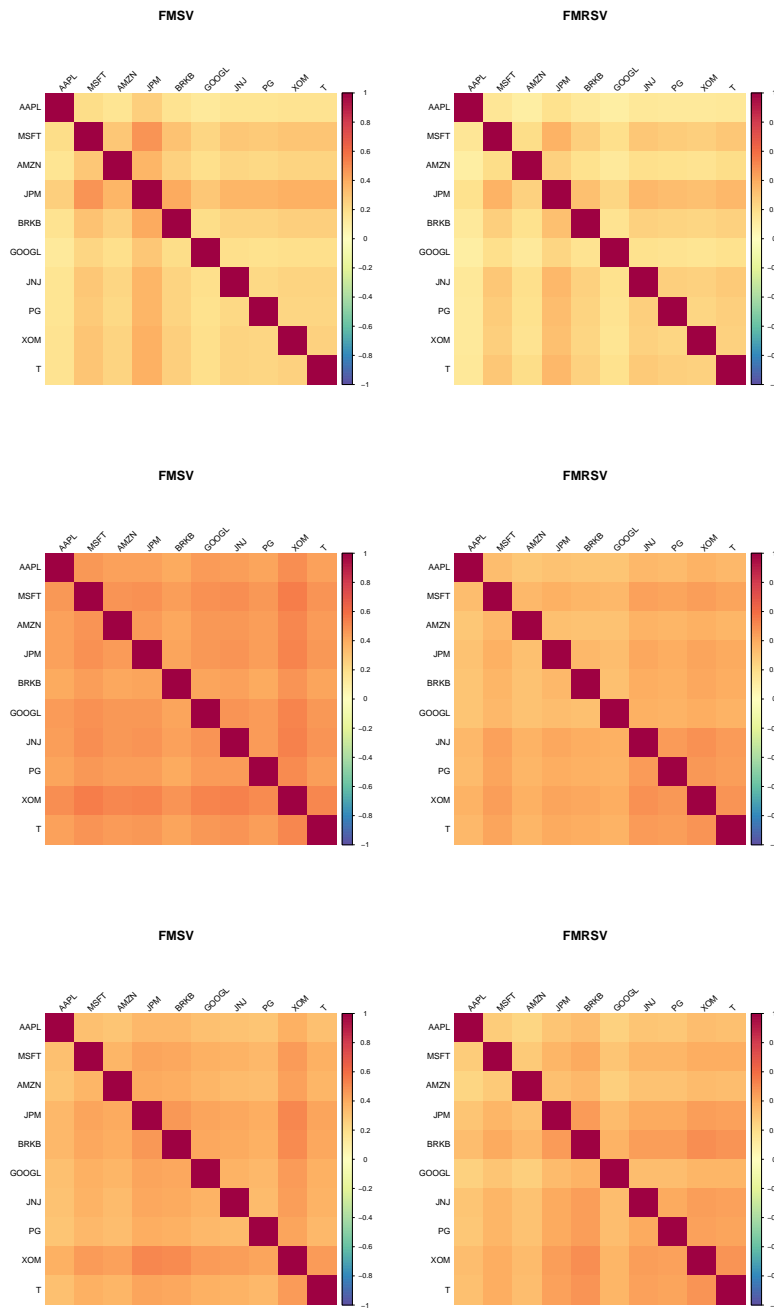


Figure 7: Heatmaps of the posterior means of the correlation coefficients.

Top: 9/1/2004–10/5/2007. Middle: 10/8/2007–11/9/2010.

Bottom: 11/10/2010–12/31/2013. Left: FMSV model. Right: FMRSV model.



### 4.3 Comparison of portfolio performance

In addition to the estimation results, we also compare the portfolio performance for the FMSV and FMRSV models. The portfolio return at time  $t + 1$  is defined as

$$r_{p,t+1} = \mathbf{w}'_t \mathbf{y}_{t+1} + (1 - \mathbf{w}'_t \mathbf{1}_p) r_f, \quad (31)$$

where  $\mathbf{w}_t$  is a  $p \times 1$  portfolio weight vector for the stock return  $\mathbf{y}_{t+1}$ ,  $\mathbf{1}_p$  denotes a  $p \times 1$  vector with all elements equal to one, and  $r_f$  is the risk-free asset return. The weight  $\mathbf{w}_t$  is unrestricted (allowing short selling) and chosen to optimize the objective function based on the portfolio strategy, as follows.

The conditional mean and conditional variance of  $r_{p,t+1}$  given the information set  $\mathcal{F}_t$  are

$$\begin{aligned} \mu_{p,t+1} &\equiv \mathbb{E}[r_{p,t+1} | \mathcal{F}_t] = \mathbf{w}'_t \mathbf{m}_{t+1|t} + (1 - \mathbf{w}'_t \mathbf{1}_p) r_f, \\ \sigma_{p,t+1}^2 &\equiv \text{Var}[r_{p,t+1} | \mathcal{F}_t] = \mathbf{w}'_t \boldsymbol{\Sigma}_{t+1|t} \mathbf{w}_t, \end{aligned}$$

where

$$\begin{aligned} \mathbf{m}_{t+1|t} &\equiv \mathbb{E}[\mathbf{y}_{t+1} | \mathcal{F}_t] = \mathbf{B}\{\boldsymbol{\gamma} + \boldsymbol{\psi} \odot (\mathbf{f}_t - \boldsymbol{\gamma})\}, \\ \boldsymbol{\Sigma}_{t+1|t} &\equiv \text{Var}[\mathbf{y}_{t+1} | \mathcal{F}_t] = \mathbf{B}\mathbf{V}_{2,t+1}\mathbf{B}' + \mathbf{V}_{1,t+1}. \end{aligned}$$

This paper considers the portfolio strategy to minimize the conditional expected variance  $\sigma_{p,t+1}^2$  given the target conditional expected return  $\mu_{p,t+1} = \mu_p^*$ . The solution of the weight is given by

$$\hat{\mathbf{w}}_t = \boldsymbol{\Sigma}_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}_p) \frac{\mu_p^* - r_f}{(\mathbf{m}_{t+1|t} - r_f \mathbf{1}_p)' \boldsymbol{\Sigma}_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}_p)}$$

where we obtain the estimates of  $\mathbf{m}_{t+1|t}$  and  $\boldsymbol{\Sigma}_{t+1|t}$  via a one-step ahead forecast using a rolling estimation.

To investigate the effect of including the realized covariances as the additional information and the leverage effect, we compare the portfolio performance using the following three models.

1. FMSV model: Factor multivariate stochastic volatility model with the leverage effect, but without realized covariances.
2. FMRSV-NL model: Factor multivariate stochastic volatility model without the leverage effect, but with realized covariances.

3. FMRSV model: Factor multivariate stochastic volatility model, with the leverage effect and realized covariances.

Two different forecast periods with 100 one-day ahead forecasts are considered using the rolling estimation with the number of observations equal to 2250:

- Period I. From August 9, 2013, to December 31, 2013.
- Period II. From August 9, 2019, to December 31, 2019.

Period I includes the time of the global financial crisis in the estimation period. The rolling forecast and estimation are implemented as follows.

Step 1. First, we estimate parameters using the first 2250 observations from September 1, 2004, to August 8, 2013. and forecast the mean, the volatility and the correlation of the multiple stock returns for August 9, 2013. Use them to obtain the optimal weights of the assets for the above portfolio strategies where the federal funds rate is used as the risk-free asset return  $r_f$ .

Step 2. Next, we drop the first observation (September 1, 2004) from the sample period and add a new observation (August 9, 2013). The new sample period is from September 2, 2004, to August 9, 2013. We estimate the parameters using these observations and forecast the mean, the volatility and the correlation for August 10, 2013. Then, we use them to obtain the optimal weights in a similar manner.

Step 3. We iterate these rolling forecasts until December 31, 2013, to obtain the 100 one-day ahead forecasts and corresponding weights.

Table 3 displays the cumulative realized variances of three models in the two periods. The FMRSV-NL and FMRSV models show the best performance in periods I and II (except for  $\mu_p^* = 0.007$  in Period I), respectively among these models, suggesting that the introduction of realized covariances improves the prediction of the conditional means and covariances of stock returns. On the other hand, the effect of introducing the leverage depends on the forecasting period. The performance of the FMRSV-NL model is the best in period I, but the worst in period II. The performance of the FMRSV model is overall good and stable in both periods.

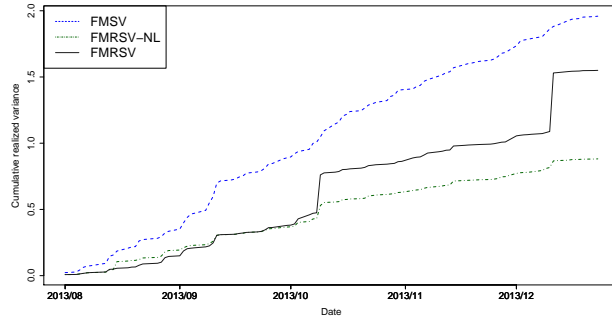
	$\mu_p^* = 0.004$	$\mu_p^* = 0.01$	$\mu_p^* = 0.02$
FMSV	0.291	1.959	8.027
FMRSV-NL	<b>0.131</b>	<b>0.882</b>	<b>3.616</b>
FMRSV	0.229	1.550	6.362

Period I (8/9/2013–12/31/2013)

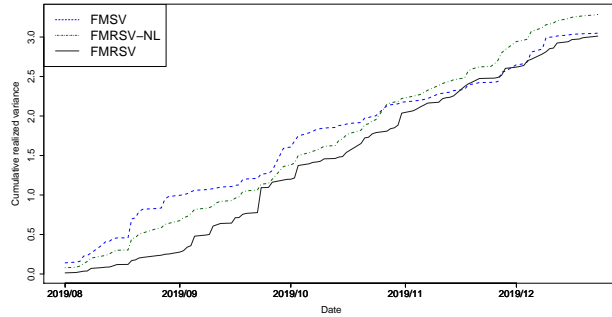
	$\mu_p^* = 0.007$	$\mu_p^* = 0.015$	$\mu_p^* = 0.03$
FMSV	<b>0.115</b>	3.049	19.481
FMRSV-NL	0.135	3.285	20.722
FMRSV	0.126	<b>3.011</b>	<b>18.918</b>

Period II (8/9/2019–12/31/2019)

Table 3: The cumulative realized variances computed as  $\sum_{t=T-100}^{T-1} \hat{w}_t' \Sigma_{t+1} \hat{w}_t$  where  $\Sigma_{t+1}$  is evaluated using the bias-corrected realized covariances at time  $t + 1$ .



Period I (8/9/2013–12/31/2013)



Period II (8/9/2019–12/31/2019)

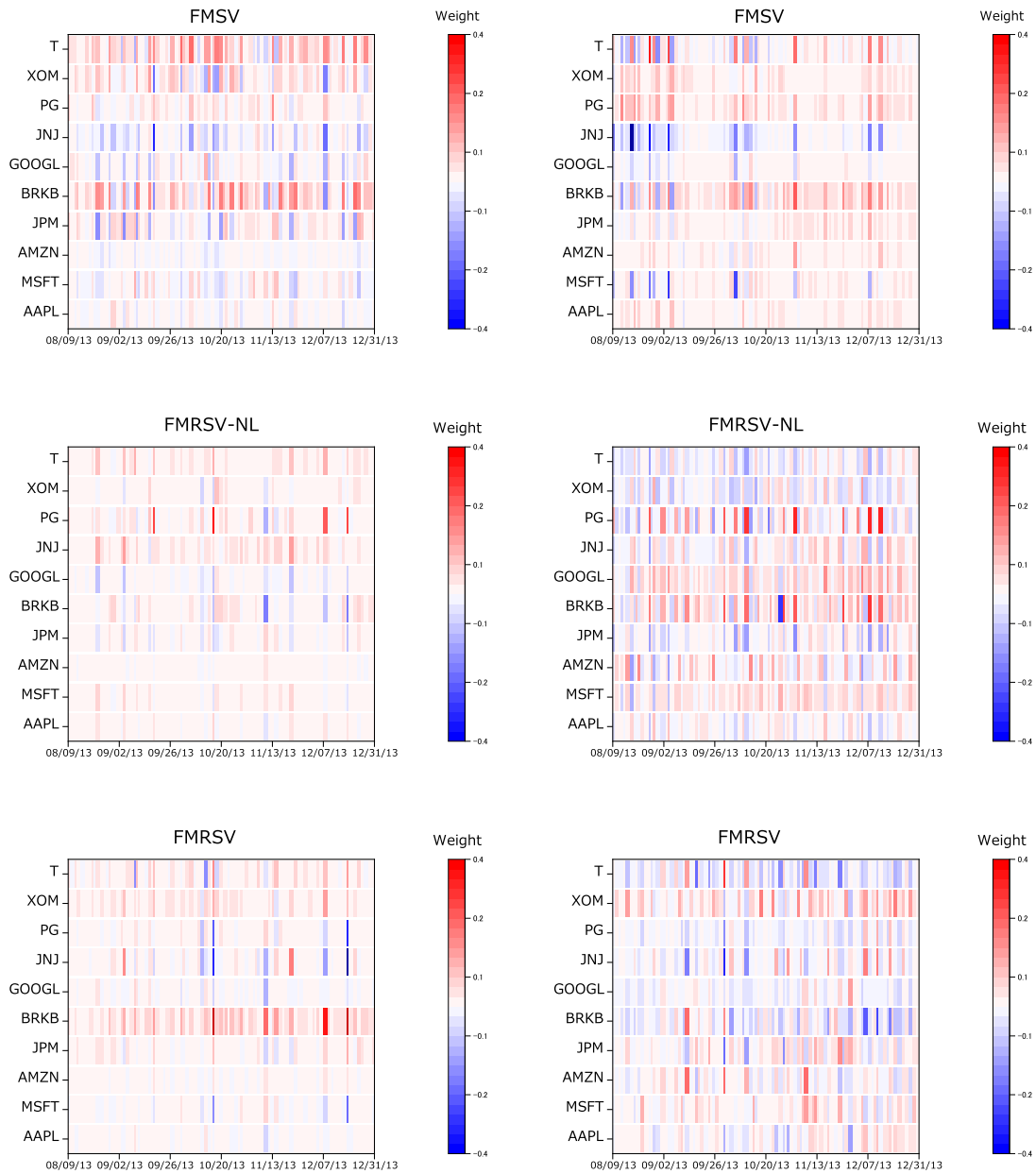
Figure 8: The cumulative realized variances with  $\mu_p^* = 0.01$  (Period I) and  $\mu_p^* = 0.015$  (Period II). FMRSV: solid black. FMRSV-NL: dashed green. FMSV: dotted blue.

Figure 8 shows the time series plots of the cumulative realized variances for the three models. The FMRSV-L and FMRSV models outperform the FMSV model in periods I and II respectively, implying that the information of the realized covariances improves the portfolio performance consistently. The performance of the FMRSV-NL model seems to depend on the forecast period. For example, it outperforms the FMRSV model after November 2013, while it underperforms the FMSV and FMRSV models after November 2019.

Figure 9 displays the split heatmaps of the portfolio weights for the three models during the two periods. We first compare three models in period I. In the FMSV model, the weights for T and BRKB tend to be positive and much larger than those of other stock returns, while the weights of JNJ are often found to be negative. Those weights are relatively unstable and sometimes become negative even for BRKB and T. For the FMRSV-NL and FMRSV models, most weights are positive for all ten stock returns and stable throughout period I. However, the FMRSV-NL model places relatively large weights on PG and JNJ, while the FMRSV model places heavy weights on BRKB. In period II, the FMSV model places large positive weights on BRKB and negative weights on JNJ as in period I, but the weights for T are smaller and sometimes negative. They are large for BRKB and PG in the FMRSV-NL model, while they are more stable and larger for AMZN and XOM in the FMRSV model. The weights for T and BRKB seem unstable but become very small during the latter part of the forecasting period in the FMRSV model.

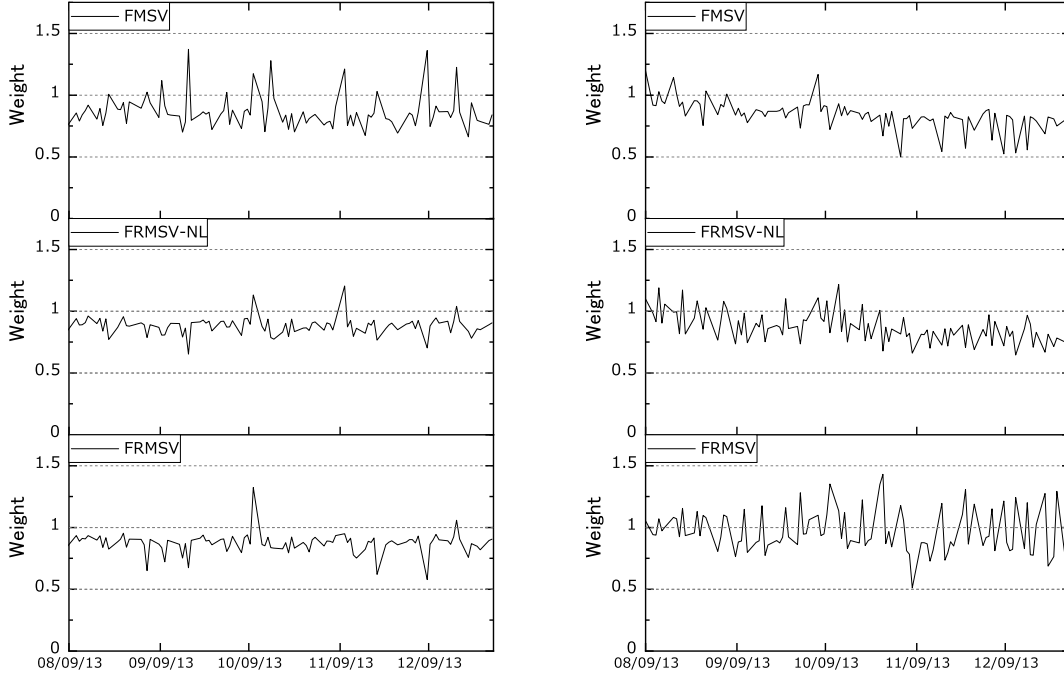
On the other hand, Figure 10 shows the time series plot of the weights of the risk-free asset. In period I, the weights for the federal funds rate are volatile in the FMSV model. They sometimes increase and decrease from over 1.3 to approximately 0.7. In the FMRSV-NL and FMRSV models, they are basically stable at approximately 0.9 except a couple of days. In period II, the weights gradually decrease from approximately 1.0 to around 0.7 in the FMSV and FMRSV-NL models, while they are very volatile around approximately 1.0 in the FMRSV model.

In summary, the weights of the asset returns change over time and lead to better portfolio performance when we use the realized covariances, suggesting that including such an additional information is very effective. The importance of the leverage effect depends on the period, but the overall performance is more stable and better for the model with leverage. Using the information from both daily returns and realized covariances gives us more accurate and stable results in the estimation and in the portfolio performance based on the forecast.



Left: Period I (8/9/2013–12/31/2013). Right: Period II (8/9/2019–12/31/2019).

Figure 9: Split heatmaps of the portfolio weight  $w_t$  for the ten stocks with  $\mu_p^* = 0.01$  (Period I) and  $\mu_p^* = 0.015$  (Period II).



Left: Period I (8/9/2013–12/31/2013). Right: Period II (8/9/2019–12/31/2019).

Figure 10: Time series plots of the portfolio weight  $1 - \mathbf{w}'_t \mathbf{1}_p$  for the federal funds rate with  $\mu_p^* = 0.01$  (Period I) and  $\mu_p^* = 0.015$  (Period II).

## 5 Conclusion

We propose a multivariate SV model with a dynamic factor structure and leverage effect, incorporating the realized measures of latent covariance and latent factors. Using the information of realized measures in addition to daily stock returns, we are able to estimate the model parameters and latent variables more accurately, and give more stable one-step ahead forecasts of the covariance matrices, which improves the portfolio performance as illustrated in our empirical studies. Taking account of the leverage effect from the information of daily returns is found to be important to obtain stable portfolio performance and becomes critical depending on the forecast period.

## Acknowledgements

The computational results were obtained by using Ox version 7 (Doornik (2007)). This work was supported by JSPS KAKENHI Grant Numbers 19H00588, 20H00073.

## References

- Aguilar, O. and M. West (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics* 18(3), 338–357.
- Archakov, I. and P. Hansen (2018). A new parametrization of correlation matrices. Technical report, Working Paper.
- Chib, S., F. Nardari, and N. Shephard (2002). Markov chain monte carlo methods for stochastic volatility models. *Journal of Econometrics* 108-2, 281–316.
- Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134(2), 341–371.
- de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika* 82, 339–350.
- Doornik, J. (2007). *Object-Oriented Matrix Programming Using Ox, 3rd ed.* London: Timberlake Consultants Press and Oxford: www.doornik.com.
- Durbin, J. and S. J. Koopman (2002). Simple and efficient simulation smoother for state space time series analysis. *Biometrika* 89, 603–616.
- Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association* 74, 291–298.
- Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies* 19(1998), 237–271.
- Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics* 27, 877–906.
- Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: does anything beat a garch (1, 1)? *Journal of applied econometrics* 20(7), 873–889.
- Ishihara, T. and Y. Omori (2017). Portfolio optimization using dynamic factor and stochastic volatility: evidence on fat-tailed error and leverage. *Japanese Economic Review* 68-1, 63–94.

- Koopman, S. J. and M. Scharth (2013). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics* 11, 76–115.
- Kurose, Y. and Y. Omori (2019). Multiple-block dynamic equicorrelations with realized measures, leverage and endogeneity. *Econometrics and Statistics*. in press.
- Lopes, H. F. and C. M. Carvalho (2007). Factor stochastic volatility with time varying loadings and Markov switching regimes. *Journal of Statistical Planning and Inference* 137(10), 3082–3091.
- Omori, Y. and T. Watanabe (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis* 52(6), 2892–2910.
- Pitt, M. and N. Shephard (1999). Time varying covariances: a factor stochastic volatility approach. In A. D. J.M. Bernardo, J.O. Berger and A. Smith (Eds.), *Bayesian Statistics*, Volume 6, pp. 547–570. Oxford University Press.
- Seber, G. A. F. (2008). *A Matrix Handbook for Statisticians*. John Wiley & Sons.
- Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika* 84(3), 653–667.
- Shirota, S., Y. Omori, H. F. Lopes, and H. Piao (2017). Cholesky realized stochastic volatility model. *Econometrics and Statistics* 3, 34–59.
- Takahashi, M., Y. Omori, and T. Watanabe (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis* 53-6, 2404–2426.
- Takahashi, M., T. Watanabe, and Y. Omori (2016). Volatility and quantile forecasts by realized stochastic volatility models with generalized hyperbolic distribution. *International Journal of Forecasting* 32(2), 437–457.
- Watanabe, T. and Y. Omori (2004). A multi-move sampler for estimating non-gaussian time series models: Comments on shephard & pitt (1997). *Biometrika*, 246–248.
- Yamauchi, Y. and Y. Omori (2019). Multivariate stochastic volatility model with realized volatilities and pairwise realized correlations. *Journal of Business & Economic Statistics*. in press.
- Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics* 127(2), 165–178.



## Appendix

### A MCMC algorithm

Generation of  $\mathbf{h}^{(2)}$ . First we define  $\mathbf{c}_t$  and  $\mathbf{V}_{2t|h}$  such that

$$\mathbf{c}_t = I(t < T) \begin{pmatrix} \frac{\rho_{p+1} \exp(h_{p+1,t}/2)}{\sigma_{p+1,\eta}} \{h_{p+1,t+1} - \mu_{p+1} - \phi_{p+1}(h_{p+1,t} - \mu_{p+1})\} \\ \vdots \\ \frac{\rho_{p+q} \exp(h_{p+q,t}/2)}{\sigma_{p+q,\eta}} \{h_{p+q,t+1} - \mu_{p+q} - \phi_{p+q}(h_{p+q,t} - \mu_{p+q})\} \end{pmatrix}, \quad (32)$$

$$\mathbf{V}_{2t|h} = \{\mathbf{I} - I(t < T) \text{diag}(\rho_{p+1}^2, \dots, \rho_{p+q}^2)\} \mathbf{V}_{2t}. \quad (33)$$

Then, the log conditional posterior density of  $\mathbf{h}^{(2)}$  given the other latent variables and parameters is

$$\begin{aligned} & \log \pi(\mathbf{h}^{(2)} | \cdot) \\ &= \text{const} + \frac{s_0}{2} \sum_{t=1}^T \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} \text{tr} \left( \sum_{t=1}^T \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' \mathbf{W}_t^{-1} \right) \\ & \quad - \frac{1}{2} \sum_{t=1}^T \left\{ \log |\mathbf{V}_{2t|h}| + (\mathbf{f}_t - \gamma - \Psi(\mathbf{f}_{t-1} - \gamma) - \mathbf{c}_t)' \mathbf{V}_{2t|h}^{-1} (\mathbf{f}_t - \gamma - \Psi(\mathbf{f}_{t-1} - \gamma) - \mathbf{c}_t) \right\} \\ & \quad - \frac{1}{2} \sum_{t=1}^{T-1} \left\{ \mathbf{h}_{t+1} - (\mathbf{I} - \Phi) \boldsymbol{\mu} - \Phi \mathbf{h}_t \right\}' \boldsymbol{\Sigma}_{\eta\eta}^{-1} \left\{ \mathbf{h}_{t+1} - (\mathbf{I} - \Phi) \boldsymbol{\mu} - \Phi \mathbf{h}_t \right\} - \frac{1}{2} \left\{ \mathbf{h}_1 - \boldsymbol{\mu} \right\}' \boldsymbol{\Sigma}_{h,0}^{-1} (\mathbf{h}_1 - \boldsymbol{\mu}) \right\} \\ &= \text{const} - \sum_{t=1}^{T-1} \sum_{i=1}^q \frac{\{h_{p+i,t+1} - h_{p+i,t} - (1 - \phi_{p+i})\mu_{p+i}\}^2}{2\sigma_{\eta,p+i}^2} - \sum_{i=1}^q \frac{(1 - \phi_{p+i}^2)(h_{p+i,1} - \mu_{p+i})^2}{2\sigma_{\eta,p+i}^2} \\ & \quad + \sum_{t=1}^T l_t^*, \end{aligned} \quad (34)$$

where

$$\begin{aligned} l_t^* &= \frac{s_0}{2} \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} \text{tr} (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' \mathbf{W}_t^{-1}) - \frac{1}{2} \sum_{i=1}^q \left\{ h_{p+i,t} + (f_{it} - \mu_{f,it})^2 \sigma_{f,it}^{-2} \right\}, \\ \mu_{f,it} &= \gamma_i + \psi_i(f_{i,t-1} - \gamma_i) + \frac{\rho_{p+i} \exp(h_{p+i,t}/2)}{\sigma_{p+i,\eta}} \{h_{p+i,t+1} - \mu_{p+i} - \phi_{p+i}(h_{p+i,t} - \mu_{p+i})\} I(t < T), \\ \sigma_{f,it}^2 &= \{1 - I(t < T)\rho_{p+i}^2\} \exp(h_{p+i,t}). \end{aligned}$$

When there is a correlation between  $f_{it}$  and  $h_{p+i,t+1}$ , the approximated conditional distribution of  $(h_{p+i,s}, \dots, h_{p+i,s+m})$  does not have a diagonal covariance matrix, and the sampling algorithm for  $\mathbf{h}^{(1)}$  does not apply to that for  $\mathbf{h}^{(2)}$ . Thus we take an alternative sampling algorithm based on Omori and Watanabe (2008) to approximate the nonlinear Gaussian state space model using the linear Gaussian state space model. As in sampling  $\mathbf{h}^{(1)}$ , we consider sampling the disturbances  $(\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1})$  instead of the state variables  $(h_{p+i,s}, \dots, h_{p+i,s+m})$  given the other state variables and parameters. The logarithm of the conditional posterior density of  $(\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1})$

is given by

$$\log f(\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1} | \cdot) = \text{const} - \frac{1}{2} \sum_{t=s-1}^{t=s+m-1} \eta_{p+i,t}^2 + L, \quad (35)$$

where

$$L = \sum_{t=s}^{t=s+m} l_{it}^* - \frac{\{h_{p+i,s+m+1} - (1 - \phi_{p+i})\mu_{p+i} - \phi_{p+i}h_{p+i,s+m}\}^2}{2\sigma_{\eta,p+i}^2} I(s+m < T),$$

$$l_{it}^* = \frac{s_0}{2} \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}| - \frac{k_0}{2} \text{tr}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' \mathbf{W}_t^{-1}) - \frac{1}{2} \left\{ h_{p+i,t} + (f_{it} - \mu_{f,it})^2 \sigma_{f,it}^{-2} \right\}.$$

In order to approximate  $L$  using a logarithm of a normal probability density, we define  $d_t, A_t, B_t$  and  $\mathbf{Q}$  as follows.

$$d_t = \frac{\partial L}{\partial h_{p+i,t}}, \quad A_t = -E \left( \frac{\partial^2 L}{\partial h_{p+i,t}^2} \right), \quad B_t = -E \left( \frac{\partial^2 L}{\partial h_{p+i,t} \partial h_{p+i,t-1}} \right),$$

$$\mathbf{Q} = \begin{bmatrix} A_s & B_{s+1} & 0 & \cdots & 0 \\ B_{s+1} & A_{s+1} & B_{s+2} & \cdots & 0 \\ 0 & B_{s+2} & A_{s+2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & B_{s+m} \\ 0 & \cdots & 0 & B_{s+m} & A_{s+m} \end{bmatrix},$$

where the expected value is taken with respect to  $f_{it}$ . Let  $\mathbf{b}_i$  denote the  $i$ -th column vector ( $i = 1, \dots, q$ ) of  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_q)$ . Using

$$\frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{p+i,t}} = \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \exp(h_{p+i,t}),$$

(the proof is given by Proposition 3 in Appendix B.3), it can be shown that

$$d_t = \frac{s_0}{2} \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \exp(h_{p+i,t}) - \frac{k_0}{2} \mathbf{b}'_i \mathbf{W}_t^{-1} \mathbf{b}_i \exp(h_{p+i,t})$$

$$- \frac{1}{2} + \frac{(f_{it} - \mu_{f,it})^2}{2\sigma_{f,it}^2} + \frac{(f_{it} - \mu_{f,it})}{\sigma_{f,it}^2} \frac{\partial \mu_{f,t,i}}{\partial h_{p+i,t}} + \frac{(f_{i,t-1} - \mu_{f,i,t-1})}{\sigma_{f,i,t-1}^2} \frac{\partial \mu_{f,t-1,i}}{\partial h_{p+i,t}}$$

$$+ \frac{\phi_{p+i}(h_{p+i,t+1} - (1 - \phi_{p+i})\mu_{p+i} - \phi_{p+i}h_{p+i,t})}{\sigma_{p+i,\eta}^2} I(t = s+m < T),$$

where

$$\frac{\partial \mu_{f,t,i}}{\partial h_{p+i,t}} = \frac{\rho_{p+i}}{\sigma_{p+i,\eta}} \left\{ -\phi_{p+i} + \frac{h_{p+i,t+1} - (1 - \phi_{p+i})\mu_{p+i} - \phi_{p+i}h_{p+i,t}}{2} \right\} \exp\left(\frac{h_{p+i,t}}{2}\right) I(t < T),$$

$$\frac{\partial \mu_{f,i,t-1}}{\partial h_{p+i,t}} = \frac{\rho_{p+i}}{\sigma_{p+i,\eta}} \exp\left(\frac{h_{p+i,t-1}}{2}\right) I(t > 1).$$

Further,

$$\begin{aligned}
A_t &= \frac{s_0}{2} \{ \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \}^2 \exp(2h_{2t,i}) \\
&\quad + \frac{1}{2} + \sigma_{f,it}^{-2} \left( \frac{\partial \mu_{f,it}}{\partial h_{p+i,t}} \right)^2 + \sigma_{f,i,t-1}^{-2} \left( \frac{\partial \mu_{f,i,t-1}}{\partial h_{p+i,t}} \right)^2 + I(t = s + m < T) \phi_{p+i}^2 \sigma_{\eta,p+i}^{-2}, \\
B_t &= \sigma_{f,t-1}^{-2} \frac{\partial \mu_{f,i,t-1}}{\partial h_{p+i,t-1}} \frac{\partial \mu_{f,i,t-1}}{\partial h_{p+i,t}}, \quad t = s + 1, \dots, s + m, \quad B_s = 0,
\end{aligned}$$

using

$$\begin{aligned}
&\frac{\partial^2 \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{p+i,t}^2} \\
&= \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \exp(h_{2t,i}) - \{ \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \}^2 \exp(2h_{2t,i}),
\end{aligned}$$

(the proof is given by Proposition 3 in Appendix B.3) and

$$E(\mathbf{W}_t^{-1}) = \frac{s_0}{k_0} (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1}.$$

Let  $\mathbf{h}_{p+i}$  denote  $(h_{p+i,s}, \dots, h_{p+i,s+m})$ . Then, via the Taylor expansion of  $L$  around the conditional mode,  $\hat{\mathbf{h}}_{p+i}$ , we obtain

$$\begin{aligned}
\log f(\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1} | \cdot) &= \text{const} - \frac{1}{2} \sum_{t=s-1}^{t=s+m-1} \eta_{p+i,t}^2 + L \\
&\approx \text{const} - \frac{1}{2} \sum_{t=s-1}^{t=s+m-1} \eta_{p+i,t}^2 + \hat{L} + \hat{\mathbf{d}}' (\mathbf{h}_{p+i} - \hat{\mathbf{h}}_{p+i}) - \frac{1}{2} (\mathbf{h}_{p+i} - \hat{\mathbf{h}}_{p+i})' \hat{\mathbf{Q}} (\mathbf{h}_{p+i} - \hat{\mathbf{h}}_{p+i}) \\
&\equiv \text{const} + \log f^*(\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1} | \cdot).
\end{aligned}$$

where  $\mathbf{d} = (d_s, \dots, d_{s+m})'$  and  $\hat{L}$ ,  $\hat{\mathbf{d}}$ , and  $\hat{\mathbf{Q}}$  are the values of  $L$ ,  $\mathbf{d}$ , and  $\mathbf{Q}$  evaluated at  $\mathbf{h}_{p+i} = \hat{\mathbf{h}}_{p+i}$ .  $f^*$  is the posterior density of the disturbances for the linear Gaussian state space model in (36), (37) and (38). We sample  $\eta_{p+i,s-1}, \dots, \eta_{p+i,s+m-1}$  as follows:

1. Set some initial value of  $\hat{\mathbf{h}}_{p+i}$ .
2. Compute  $\hat{d}_t$ ,  $\hat{A}_t$ , and  $\hat{B}_t$  at  $\hat{\mathbf{h}}_{p+i}$  for  $t = s, \dots, s + m$ .
3. Initialize  $D_s = \hat{A}_s$ ,  $J_s = 0$ , and  $b_s = \hat{d}_s$  and derive  $D_t$ ,  $J_t$ , and  $b_t$  recursively for  $t = s + 1, \dots, s + m$ :

$$D_t = \hat{A}_t - \hat{B}_t D_{t-1}^{-1} \hat{B}_t, \quad J_t = \hat{B}_t K_{t-1}^{-1}, \quad b_t = \hat{d}_t - J_t K_{t-1}^{-1} b_{t-1},$$

where  $K_t = \sqrt{D_t}$  and set  $J_{s+m+1} = 0$ .

4. Define  $\hat{y}_t = \hat{\gamma}_t + D_t^{-1} b_t$ , where

$$\hat{\gamma}_t = \hat{h}_{p+i,t} + K_t^{-1} J_{t+1} \hat{h}_{p+i,t+1} \quad t = s, \dots, s + m.$$

5. Construct the approximated linear Gaussian state space model given by

$$\hat{y}_{i,t} = z_t h_{p+i,t} + \mathbf{G}_t \boldsymbol{\xi}_t, \quad t = s, \dots, s+m, \quad (36)$$

$$h_{p+i,t+1} = (1 - \phi_{p+i})\mu_{p+i} + \phi_{p+i}h_{p+i,t} + \mathbf{H}_t \boldsymbol{\xi}_t, \quad t = s, \dots, s+m-1, \quad (37)$$

$$\boldsymbol{\xi}_t = (\xi_{1t}, \xi_{2t})' \sim N(\mathbf{0}, \mathbf{I}_2),$$

$$h_{p+i,s} = \begin{cases} (1 - \phi_{p+i})\mu_{p+i} + \phi_{p+i}h_{p+i,s-1} + \mathbf{H}_{s-1}\boldsymbol{\xi}_{s-1}, & s > 1, \\ \mu_{p+i} + \mathbf{H}_0\boldsymbol{\xi}_0, & s = 1, \end{cases} \quad (38)$$

where

$$z_t = 1 + \phi_{p+i}K_t^{-1}J_{t+1}, \quad \mathbf{G}_t = K_t^{-1}[1, \sigma_{\eta,p+i}J_{t+1}], \quad t = 1, \dots, T,$$

$$\mathbf{H}_t = \begin{cases} (0, \sigma_{\eta,p+i}), & t = 1, \dots, T-1, \\ \left(0, \frac{\sigma_{\eta,p+i}}{\sqrt{1-\phi_{p+i}^2}}\right), & t = 0. \end{cases}$$

In order to find the posterior mode, we implement the Kalman filter and a disturbance smoother for (36) – (38) and update  $\hat{\mathbf{h}}_{p+i}$ . We repeat Steps 2-5 several times or until some convergence criterion is met. Otherwise, we go to Step 6.

6. Generate a candidate  $(\eta_{p+i,s-1}^n, \dots, \eta_{p+i,s+m-1}^n)$  via a simulation smoother using models (36) – (38) given  $\hat{\mathbf{h}}_{p+i}$ . These samples are generated from  $f^*$ .
7. Conduct the MH algorithm where we accept a candidate  $(\eta_{p+i,s-1}^n, \dots, \eta_{p+i,s+m-1}^n)$  with probability

$$\min \left\{ 1, \frac{f(\eta_{p+i,s-1}^n, \dots, \eta_{p+i,s+m-1}^n | \cdot) f^*(\eta_{p+i,s-1}^o, \dots, \eta_{p+i,s+m-1}^o | \cdot)}{f(\eta_{p+i,s-1}^o, \dots, \eta_{p+i,s+m-1}^o | \cdot) f^*(\eta_{p+i,s-1}^n, \dots, \eta_{p+i,s+m-1}^n | \cdot)} \right\},$$

where  $(\eta_{p+i,s-1}^o, \dots, \eta_{p+i,s+m-1}^o)$  is a current sample.

## B Propositions

### B.1 Proposition 1

(i)

$$\frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial \beta_i} = 2 \mathbf{V}_{2t} \mathbf{B}' (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{e}_i, \quad (39)$$

where  $\mathbf{e}_i$  denotes a  $p \times 1$  vector with the  $i$ -th element equal to one and zero elements.

(ii)

$$\begin{aligned} \frac{\partial^2 \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial \beta_i \partial \beta_i'} &= 2d_{ii} \{ \mathbf{V}_{2t} - \mathbf{V}_{2t} \mathbf{B}' (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{B} \mathbf{V}_{2t} \} \\ &\quad - \frac{1}{2} \left[ \frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial \beta_i} \right] \left[ \frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial \beta_i} \right]', \end{aligned} \quad (40)$$

where  $d_{ii}$  is the  $(i, i)$ -th element of  $(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1}$ .

**Proof:**

(i) Let  $\mathbf{X}_t = \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}$ . Then, using a chain rule and  $\partial \log |\mathbf{X}| / \partial \text{vec}(\mathbf{X}) = \text{vec}(\mathbf{X}^{-1})'$  (see, e.g., 17.29 and 17.52 of Seber (2008)), we obtain the first derivative

$$\begin{aligned} \frac{\partial \log |\mathbf{X}_t|}{\partial \beta'_i} &= \frac{\partial \log |\mathbf{X}_t|}{\partial \text{vec}(\mathbf{X}_t)'} \times \frac{\partial \text{vec}(\mathbf{X}_t)}{\partial \text{vec}(\mathbf{B}')'} \times \frac{\partial \text{vec}(\mathbf{B}')}{\partial \beta'_i} \\ &= \text{vec}(\mathbf{X}_t^{-1})' \times \frac{\partial \text{vec}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}')}{\partial \text{vec}(\mathbf{B}')'} \times (\mathbf{e}_i \otimes \mathbf{I}_q) \\ &= \text{vec}(\mathbf{X}_t^{-1})' \times \{(\mathbf{I}_{p^2} + \mathbf{K}_{pp})(\mathbf{I}_p \otimes \mathbf{B} \mathbf{V}_{2t})\} \times (\mathbf{e}_i \otimes \mathbf{I}_q) \\ &= 2 \text{vec}(\mathbf{X}_t^{-1})' (\mathbf{I}_p \otimes \mathbf{B} \mathbf{V}_{2t}) (\mathbf{e}_i \otimes \mathbf{I}_q) \\ &= 2 \text{vec}(\mathbf{X}_t^{-1})' (\mathbf{e}_i \otimes \mathbf{B} \mathbf{V}_{2t}). \end{aligned}$$

where  $\mathbf{K}_{mn}$  is a vec-permutation matrix such that  $\text{vec}(\mathbf{A}') = \mathbf{K}_{mn} \text{vec}(\mathbf{A})$  for an  $m \times n$  matrix  $\mathbf{A}$ . The third equality follows from  $\partial \text{vec}(\mathbf{X}' \mathbf{A} \mathbf{X}) / \partial \text{vec}(\mathbf{X})' = (\mathbf{I}_{n^2} + \mathbf{K}_{nn})(\mathbf{I}_n \otimes \mathbf{X}' \mathbf{A})$  for an  $m \times n$  matrix  $\mathbf{X}$  and an  $m \times m$  symmetric matrix  $\mathbf{A}$  (e.g., 17.30(f) of Seber (2008)). Therefore,

$$\frac{\partial \log |\mathbf{X}_t|}{\partial \beta'_i} = 2(\mathbf{e}'_i \otimes \mathbf{V}_{2t} \mathbf{B}') \text{vec}(\mathbf{X}_t^{-1}) = 2 \mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{e}_i,$$

using  $\text{vec}(\mathbf{A} \mathbf{X} \mathbf{C}) = (\mathbf{C}' \otimes \mathbf{A}) \text{vec}(\mathbf{X})$  (e.g., 11.16(b) of Seber (2008)).

(ii) From (i) and a product rule (e.g., 17.30(h) Seber (2008)),

$$\begin{aligned} \frac{\partial^2 \log |\mathbf{X}_t|}{\partial \beta'_i \partial \beta'_i} &= 2 \times \frac{\partial \mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{e}_i}{\partial \beta'_i} \\ &= 2 \times \left[ \{(\mathbf{e}'_i \mathbf{X}_t^{-1}) \otimes \mathbf{I}_q\} \frac{\partial \text{vec}(\mathbf{V}_{2t} \mathbf{B}')}{\partial \beta'_i} + (\mathbf{1} \otimes \mathbf{V}_{2t} \mathbf{B}') \frac{\partial \text{vec}(\mathbf{X}_t^{-1} \mathbf{e}_i)}{\partial \beta'_i} \right]. \end{aligned}$$

Since

$$\frac{\partial \text{vec}(\mathbf{V}_{2t} \mathbf{B}')}{\partial \beta'_i} = \frac{\partial \text{vec}(\mathbf{V}_{2t} \mathbf{B}')}{\partial \text{vec}(\mathbf{B}')'} \times \frac{\partial \text{vec}(\mathbf{B}')}{\partial \beta'_i} = (\mathbf{I}_p \otimes \mathbf{V}_{2t}) (\mathbf{e}_i \otimes \mathbf{I}_q) = \mathbf{e}_i \otimes \mathbf{V}_{2t},$$

using  $\partial \text{vec}(\mathbf{A} \mathbf{X} \mathbf{B}) / \partial \text{vec}(\mathbf{X})' = \mathbf{B}' \otimes \mathbf{A}$  (e.g. 17.30(b) of Seber (2008)), the first term in the bracket is

$$\{(\mathbf{e}'_i \mathbf{X}_t^{-1}) \otimes \mathbf{I}_q\} \frac{\partial \text{vec}(\mathbf{V}_{2t} \mathbf{B}')}{\partial \beta'_i} = (\mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{e}_i) \otimes \mathbf{V}_{2t} = d_{ii} \mathbf{V}_{2t}.$$

For the second term in the bracket, via a product rule and a chain rule,

$$\begin{aligned} (\mathbf{1} \otimes \mathbf{V}_{2t} \mathbf{B}') \frac{\partial \text{vec}(\mathbf{X}_t^{-1} \mathbf{e}_i)}{\partial \beta'_i} &= \mathbf{V}_{2t} \mathbf{B}' \left\{ (\mathbf{e}'_i \otimes \mathbf{I}_p) \frac{\partial \text{vec}(\mathbf{X}_t^{-1})}{\partial \beta'_i} \right\} \\ &= \mathbf{V}_{2t} \mathbf{B}' (\mathbf{e}'_i \otimes \mathbf{I}_p) \times \frac{\partial \text{vec}(\mathbf{X}_t^{-1})}{\partial \text{vec}(\mathbf{X}_t)'} \frac{\partial \text{vec}(\mathbf{X}_t)}{\partial \beta'_i} \\ &= \mathbf{V}_{2t} \mathbf{B}' (\mathbf{e}'_i \otimes \mathbf{I}_p) \left[ \{-(\mathbf{X}_t^{-1})' \otimes \mathbf{X}_t^{-1}\} \frac{\partial \text{vec}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}')}{\beta'_i} \right]. \quad (41) \end{aligned}$$

Noting that  $\mathbf{K}_{pm}(\mathbf{b} \otimes \mathbf{A}) = \mathbf{A} \otimes \mathbf{b}$  (e.g. 11.19(c)(ii) of Seber (2008)), we substitute

$$\begin{aligned} \frac{\text{vec}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}')}{\beta'_i} &= (\mathbf{I}_{p^2} + \mathbf{K}_{pp})(\mathbf{I}_p \otimes \mathbf{B} \mathbf{V}_{2t})(\mathbf{e}_i \otimes \mathbf{I}_q) = (\mathbf{I}_{p^2} + \mathbf{K}_{pp})(\mathbf{e}_i \otimes \mathbf{B} \mathbf{V}_{2t}) \\ &= \mathbf{e}_i \otimes \mathbf{B} \mathbf{V}_{2t} + \mathbf{B} \mathbf{V}_{2t} \otimes \mathbf{e}_i, \end{aligned}$$

into Equation (41). Then (41) reduces to

$$\begin{aligned} & -\mathbf{V}_{2t} \mathbf{B}' \{(\mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{e}_i) \otimes \mathbf{X}_t^{-1} \mathbf{B} \mathbf{V}_{2t} + (\mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{B} \mathbf{V}_{2t}) \otimes (\mathbf{X}_t^{-1} \mathbf{e}_i)\} \\ &= -d_{ii} \mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{B} \mathbf{V}_{2t} - (\mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{B} \mathbf{V}_{2t}) \otimes (\mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{e}_i) \\ &= -d_{ii} \mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{B} \mathbf{V}_{2t} - (\mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{e}_i) (\mathbf{V}_{2t} \mathbf{B}' \mathbf{X}_t^{-1} \mathbf{e}_i)', \end{aligned}$$

and the result follows. □

## B.2 Proposition 2

(i)

$$\frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{it}} = d_{ii} \exp(h_{it}), \quad i = 1, \dots, p, \quad (42)$$

(ii)

$$\frac{\partial^2 \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{it}^2} = d_{ii} \exp(h_{it}) - d_{ii}^2 \exp(2h_{it}), \quad i = 1, \dots, p. \quad (43)$$

where  $d_{ii}$  is the  $(i, i)$ -th element of  $(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1}$ .

**Proof:**

(i) Let  $\mathbf{X}_t = \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}$ . Then, using a chain rule as in the proof of Proposition 1,

$$\begin{aligned} \frac{\partial \log |\mathbf{X}_t|}{\partial h_{it}} &= \frac{\partial \log |\mathbf{X}_t|}{\partial \text{vec}(\mathbf{X}_t)'} \times \frac{\partial \text{vec}(\mathbf{X}_t)}{\partial h_{it}} \\ &= \text{vec}(\mathbf{X}_t^{-1})' \times \text{vec}(\mathbf{e}_i \mathbf{e}'_i) \times \exp(h_{it}) \\ &= \text{tr}(\mathbf{X}_t^{-1} \mathbf{e}_i \mathbf{e}'_i) \times \exp(h_{it}) = d_{ii} \exp(h_{it}). \end{aligned}$$

(ii) Since

$$\begin{aligned} \frac{\partial d_{ii}}{\partial h_{it}} &= \text{vec}(\mathbf{e}_i \mathbf{e}'_i)' \frac{\partial \text{vec}(\mathbf{X}_t^{-1})}{\partial h_{it}} \\ &= \text{vec}(\mathbf{e}_i \mathbf{e}'_i) (-\mathbf{X}_t^{-1'} \otimes \mathbf{X}_t^{-1}) \text{vec}(\mathbf{e}_i \mathbf{e}'_i) \exp(h_{it}) \\ &= -(\mathbf{e}'_i \otimes \mathbf{e}'_i) (\mathbf{X}_t^{-1} \otimes \mathbf{X}_t^{-1}) (\mathbf{e}_i \otimes \mathbf{e}_i) \exp(h_{it}) \\ &= -(\mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{e}_i \otimes \mathbf{e}'_i \mathbf{X}_t^{-1} \mathbf{e}_i) \exp(h_{it}) = -d_{ii}^2 \exp(h_{it}), \end{aligned}$$

using  $\partial \text{vec}(\mathbf{X}^{-1})/\partial \text{vec}(\mathbf{X})' = -\mathbf{X}^{-1'} \otimes \mathbf{X}^{-1}$  (e.g. 17.30(d) of Seber (2008)), we obtain

$$\frac{\partial^2 \log |\mathbf{X}_t|}{\partial h_{it}^2} = \frac{\partial d_{ii}}{\partial h_{it}} \exp(h_{it}) + d_{ii} \exp(h_{it}) = d_{ii} \exp(h_{it}) - d_{ii}^2 \exp(2h_{it}),$$

and the result follows. □

### B.3 Proposition 3

(i)

$$\frac{\partial \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{p+i,t}} = \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \exp(h_{p+i,t}), \quad i = 1, \dots, q, \quad (44)$$

(ii)

$$\begin{aligned} & \frac{\partial^2 \log |\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}|}{\partial h_{p+i,t}^2} \\ &= \mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i \exp(h_{2t,i}) - \{\mathbf{b}'_i (\mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t})^{-1} \mathbf{b}_i\}^2 \exp(2h_{2t,i}), \\ & \quad i = 1, \dots, q. \end{aligned} \quad (45)$$

**Proof:**

(i) Let  $\mathbf{X}_t = \mathbf{B} \mathbf{V}_{2t} \mathbf{B}' + \mathbf{V}_{1t}$ . Then, as in the proof of Proposition 2,

$$\begin{aligned} \frac{\partial \log |\mathbf{X}_t|}{\partial h_{p+i,t}} &= \frac{\partial \log |\mathbf{X}_t|}{\partial \text{vec}(\mathbf{X}_t)'} \times \frac{\partial \text{vec}(\mathbf{X}_t)}{\partial h_{p+i,t}} \\ &= \text{vec}(\mathbf{X}_t^{-1'})' \times \frac{\partial \text{vec}(\mathbf{B} \mathbf{V}_{2t} \mathbf{B}')}{\partial h_{p+i,t}} \\ &= \text{vec}(\mathbf{X}_t^{-1})' \times \text{vec}(\mathbf{b}_i \mathbf{b}'_i) \times \exp(h_{p+i,t}) \\ &= \text{tr}(\mathbf{X}_t^{-1} \mathbf{b}_i \mathbf{b}'_i) \times \exp(h_{p+i,t}) = \mathbf{b}'_i \mathbf{X}_t^{-1} \mathbf{b}_i \exp(h_{p+i,t}). \end{aligned}$$

(ii) Similar to the proof in Proposition 2, using

$$\begin{aligned} \frac{\partial \mathbf{b}'_i \mathbf{X}_t^{-1} \mathbf{b}_i}{\partial h_{p+i,t}} &= \text{vec}(\mathbf{b}_i \mathbf{b}'_i)' \frac{\partial \text{vec}(\mathbf{X}_t^{-1})}{\partial h_{p+i,t}} \\ &= \text{vec}(\mathbf{b}_i \mathbf{b}'_i) (-\mathbf{X}_t^{-1'} \otimes \mathbf{X}_t^{-1}) \text{vec}(\mathbf{b}_i \mathbf{b}'_i) \exp(h_{p+i,t}) \\ &= \{\mathbf{b}'_i \mathbf{X}_t^{-1} \mathbf{b}_i\}^2 \exp(h_{p+i,t}), \end{aligned}$$

and the result follows. □