Multivariate Stochastic Volatility Model with Realized Volatilities and Pairwise Realized Correlations

Yuta Yamauchi
Graduate School of Economics, The University of Tokyo

Yasuhiro Omori
The University of Tokyo

Multivariate Stochastic Volatility Model with Realized Volatilities and Pairwise Realized Correlations

Yuta Yamauchi[*] and Yasuhiro Omori[†]

March 13, 2019

Abstract

Although stochastic volatility and GARCH (generalized autoregressive conditional heteroscedasticity) models have successfully described the volatility dynamics of univariate asset returns, extending them to the multivariate models with dynamic correlations has been difficult due to several major problems. First, there are too many parameters to estimate if available data are only daily returns, which results in unstable estimates. One solution to this problem is to incorporate additional observations based on intraday asset returns, such as realized covariances. Second, since multivariate asset returns are not synchronously traded, we have to use the largest time intervals such that all asset returns are observed in order to compute the realized covariance matrices. However, in this study, we fail to make full use of the available intraday informations when there are less frequently traded assets. Third, it is not straightforward to guarantee that the estimated (and the realized) covariance matrices are positive definite.

Our contributions are the following: (1) we obtain the stable parameter estimates for the dynamic correlation models using the realized measures, (2) we make full use of intraday informations by using pairwise realized correlations, (3) the covariance matrices are guaranteed to be positive definite, (4) we avoid the arbitrariness of the ordering of asset returns, (5) we propose the flexible correlation structure model (e.g., such as setting some correlations to be zero if necessary), and (6) the parsimonious specification for the leverage effect is proposed. Our proposed models are applied to the daily returns of nine U.S. stocks with their realized volatilities and pairwise realized correlations and are shown to outperform the existing models with respect to portfolio performances.

---

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

[†]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.

1 Introduction

Modelling the time-varying volatility and the correlations of multivariate time series is one of the most important problems in financial risk management, and there are numerous studies that model the time-varying volatility of univariate time series using the GARCH or stochastic volatility (SV) models. However, the extension of their models to multivariate model with dynamic correlations has not been straightforward due to the following several major problems.

First, there are too many parameters to estimate if the only available data are daily returns, which results in unstable estimates. An intuitive solution to reduce the number of parameters is to introduce the factor structure assuming that a small number of common factors describe the dynamics of time-varying covariance matrices as discussed in the factor stochastic volatility models (e.g. Pitt and Shephard (1999), Chib et al. (2006) and Lopes and Carvalho (2007)). However, factor modelling requires the a priori selection of the number of factors and we need to restrict the structure of the factors in order to identify the parameters (e.g. Lopes and West (2004)). Furthermore, the estimation results and the predictive performance of the model are usually subject to the ordering of the asset returns.

An alternative effective approach is to incorporate additional observations based on the intraday asset returns, such as the realized covariances, which have recently become available in financial markets. In univariate SV models, the realized SV (RSV) models that estimate the time-varying volatilities using the daily returns and realized volatility simultaneously have been proposed to achieve more accurate parameter estimates than those of the SV models using only daily returns and the RSV models outperform SV models in forecasting volatilities (e.g. Takahashi et al. (2009), Dobrev and Szerszen (2010), Koopman and Scharth (2013), Zheng and Song (2014), Takahashi et al. (2016)). Although the realized volatilities are subject to microstructure noises and nontrading hours and hence are biased estimates of the integrated volatilities, such biases are automatically adjusted within the proposed model. Similarly, the univariate GARCH model is extended to the realized GARCH models which incorporates the realized volatilities into the variance equations and it is shown to lead to substantial improvements in the empirical fit and quantile forecasts over the standard GARCH model that only uses daily returns (Hansen et al. (2012)).

The extension to the multivariate RSV model is also considered in the Cholesky RSV model (Shirota et al. (2017)). In this model, the Cholesky decompositions of the realized

covariance matrices are used as additional sources for measurement equations, and it models the dynamics of the logarithm of the diagonal elements and the off-diagonal elements of Cholesky decomposed covariance matrices respectively. It is shown that the portfolio performances of the proposed model outperformed other SV models without realized measures in the empirical studies, but it should also be noted that the performance of the Cholesky RSV models may depend on the ordering of the asset returns in the vector of the response.

Second, high-frequency data are not always observed at the same time points, which causes difficulties in the extension of the univariate RSV model to the multivariate RSV model. For example, in the Cholesky RSV model, it is implicitly assumed that the all multivariate assets are traded every few minutes when computing the realized covariance matrices. If the multivariate assets are not traded synchronously, we have to use the largest time intervals so that all asset returns are observed when computing the realized covariance matrices. This nonsynchronous trading leads us to ignore some of the frequently traded asset return data, and hence we would fail to make full use of the available intraday informations when there are less frequently traded assets.

Third, it is not straightforward to guarantee that the estimated (and the realized) covariance matrices are positive definite. The model parameters may be difficult to estimate in practice under the constraints that satisfy the positive definiteness. Using the Cholesky decomposition of the time-varying covariance matrices is one way to guarantee the positive definiteness (Shirota et al. (2017)), but it also requires that the multivariate assets are traded synchronously in order to compute the realized covariance matrices as mentioned above. Additionally, the interpretation of each latent variable of the decomposition is not straightforward since it does not correspond to each pair of asset returns and it is subject to the ordering of the asset returns. If we use each element of the realized covariances for each pair of asset returns, it may result in the nonpositive definite covariance matrices.

To overcome these difficulties, we propose a multivariate realized SV (MRSV) model with pairwise realized correlations, in which we incorporate the dynamic latent correlation variables in addition to latent volatility variables with realized measures for each pairwise correlation and the volatilities in the framework of multivariate SV models with realized volatilities. The model parameters are estimated using Markov chain Monte Carlo simulations, and we sample the latent correlation variables one at a time given the others so that we keep the covariance matrices positive definite. The realized Beta GARCH model proposed by Hansen et al. (2014) is a promising multivariate GARCH model with realized measures

3

for volatilities and co-volatilities in which they used measurement equations for the pairwise realized correlations with market returns and modelled dynamics of the Fisher transformed conditional correlation coefficients. However, they focused on the pairwise correlations between the market return and an individual asset return, assuming that the individual asset returns are conditionally independent given the market return. Another useful approach for the joint modelling of returns and realized covariances are based on Wishart processes (e.g. Jin and Maheu (2013), Windle et al. (2014), Jin and Maheu (2016), So et al. (2016)). The covariance matrix is assumed to follow a Wishart distribution whose scale matrix depends on the past realized covariance matrices which are computed using larger time intervals than necessary in order for it to be positive definite.

Our approach, on the other hand, is based on simultaneously modeling the individual volatilities and pairwise covariances, rather than the covariance matrix, and we are able to make full use of the available intraday information, even when there are less frequently traded assets. This finding implies that our model still can be constructed even though some of the realized measures are missing. Additionally, our model is far more flexible in the sense that it is possible to restrict any correlation coefficients to be zero for very high dimensional asset returns data, which reduces the number of parameters and may improve the forecasting performances. Among the multivariate SV models in the literature, our model is a natural extension of a univariate RSV model to a multivariate model, and it gives us a straightforward interpretation of the estimated parameters.

Furthermore, we extend our model to incorporate the leverage effect, which is well-known to exist in stock markets. The leverage effect refers to the negative correlation between an asset return and its volatility. In other words, a decrease in the stock return is followed by an increase in its volatility. In forecasting the means and covariances of asset returns for portfolio optimization, it is expected that incorporating the leverage effect in econometric models improves the predictive accuracy. However, it may increase the number of parameters that are to be estimated and the realized measures are not available for such an effect. Thus we also consider the parsimonious parameterization for the leverage effect.

Our contributions are as follows: (1) we obtain the stable parameter estimates for the dynamic correlation models using the realized measures, (2) we make full use of the intraday informations by using pairwise realized correlations, (3) the estimated covariance matrices are guaranteed to be positive definite, (4) we avoid the arbitrariness of the ordering of asset returns, (5) we propose the flexible correlation structure model and set some correlations

4

to be zero if necessary, and (6) we introduce the parsimonious specification for the leverage effect.

The structure of this paper is as follows. Section 2 introduces the multivariate realized SV model with daily returns, realized volatilities, and pairwise realized correlations. Section 3 describes the estimation algorithms using the Markov Chain Monte Carlo simulation. Section 4 extends it to incorporate the leverage effect. Finally, in Section 5, the proposed model is applied to nine U.S. stock return data and the model with the leverage effect is shown to outperform other competing models with regard to the portfolio performances.

2   Multivariate realized stochastic volatility model

This section introduces the multivariate realized stochastic volatility (MRSV) model, which uses realized measures for the volatility and pairwise correlations of asset returns. By using the additional information of the realized measure for asset returns, we can overcome the curse of dimensionality when estimating the dynamic covariance matrices. The Cholesky RSV model proposed by Shirota et al. (2017) also uses the realized measure of variances and covariances (which we call the realized covariance matrix) in order to estimate the latent covariance matrix of asset returns. However, the realized covariance matrix is less informative when there are less frequent asset returns. This finding is observed because we require the synchronous observations of all asset return series in order to compute the realized covariance matrix. In order to utilize the full information of the realized measure for the correlations, we propose using the realized measures for the latent pairwise correlations. It should be noted that the pairwise correlation can be computed if that pair of series is synchronously observed. We call the realized measure for the correlation coefficient the pairwise realized correlation. Using the pairwise realized correlations, in order to guarantee the positive definiteness of the latent covariance matrices, we propose the MCMC algorithm in which we sample latent correlation coefficients from the conditional posterior distribution so that the matrices are positive definite.

2.1   Multivariate stochastic volatility model with dynamic correlations

First, we define the multivariate SV (MSV) model without the realized measures. Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})'$ and $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{pt})'$ denote a $p \times 1$ stock return vector and its corresponding

log volatility latent vector at time t. The basic MSV model is given by

$$\boldsymbol{y}_t = \boldsymbol{m}_t + \mathbf{V}_t^{1/2}\,\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_t), \qquad t = 1, \ldots, T, \tag{1}$$

$$\boldsymbol{h}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Omega}), \quad t = 1, \ldots, T-1, \tag{2}$$

$$\boldsymbol{m}_{t+1} = \boldsymbol{m}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_m), \quad t = 1, \ldots, T-1, \tag{3}$$

$$\boldsymbol{h}_1 \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Omega}_0), \quad \boldsymbol{m}_1 \sim \mathrm{N}(\mathbf{0}, \kappa\,\boldsymbol{\Sigma}_m), \tag{4}$$

where $\mathbf{R}_t = \{\rho_{ij,t}\}$ is a correlation matrix, $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{pt})'$, $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{pt})'$, and

$$\boldsymbol{\nu}_t = (\nu_{1t}, \ldots, \nu_{pt})', \quad \mathbf{V}_t = \mathrm{diag}(\exp(h_{1t}), \ldots, \exp(h_{pt})),$$

$$\boldsymbol{\Sigma}_m = \mathrm{diag}(\boldsymbol{\sigma}_m^2), \quad \boldsymbol{\sigma}_m^2 = (\sigma_{m,1}^2, \ldots, \sigma_{m,p}^2)', \quad \boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{\phi}), \quad \boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)'.$$

We assume that h_{it} follows a stationary autoregressive process (with its coefficient $|\phi_j| < 1$) and that the mean process $\boldsymbol{m}_t = (m_{1t}, \ldots, m_{pt})'$ follows a random walk process. We denote a diagonal matrix $\mathbf{A}$ with diagonal elements $\boldsymbol{a} = (a_{11}, \ldots, a_{mm})'$ as $\mathbf{A} = \mathrm{diag}(\boldsymbol{a})$. For the initial distributions of $\boldsymbol{m}_1$ and $\boldsymbol{h}_1$, we set κ to some large constant for $\boldsymbol{m}_1$ for simplicity and set $\boldsymbol{\Omega}_0$ to satisfy the stationary condition $\boldsymbol{\Omega}_0 = \boldsymbol{\Phi}\,\boldsymbol{\Omega}_0\,\boldsymbol{\Phi} + \boldsymbol{\Omega}$ for $\boldsymbol{h}_1$ such that

$$\mathrm{vec}(\boldsymbol{\Omega}_0) = (\mathbf{I}_{p^2} - \boldsymbol{\Phi} \otimes \boldsymbol{\Phi})^{-1}\,\mathrm{vec}(\boldsymbol{\Omega}), \tag{5}$$

where $\mathbf{I}_{p^2}$ denotes a $p^2 \times p^2$ unit matrix. In order to model the dynamics of the correlation matrix, we consider the following Fisher transformation $g_{ij,t+1}$ of the correlation coefficient $\rho_{ij,t}$, and assume that it follows a random walk process for simplicity:

$$g_{ij,t+1} = g_{ij,t} + \zeta_{ij,t}, \qquad \zeta_{ij,t} \sim \text{i.i.d. } \mathrm{N}(0, \sigma_{\zeta,ij}^2), \quad t = 1, \ldots, T-1, \tag{6}$$

$$g_{ij,1} \sim \mathrm{N}(0, \kappa\sigma_{\zeta,ij}^2), \qquad g_{ij,t} = \log(1 + \rho_{ij,t}) - \log(1 - \rho_{ij,t}), \tag{7}$$

for $i, j = 1, \ldots, p$ $(j < i)$ and we denote $\boldsymbol{\rho}_t = (\rho_{21,t}, \ldots, \rho_{pp-1,t})'$, $\boldsymbol{g}_t = (g_{21,t}, \ldots, g_{pp-1,t})'$, $\boldsymbol{\zeta}_t = (\zeta_{21,t}, \ldots, \zeta_{pp-1,t})'$, and $\boldsymbol{\sigma}_\zeta^2 = (\sigma_{\zeta,21}^2, \ldots, \sigma_{\zeta,pp-1}^2)'$.

Non-arbitrary ordering of asset returns and the flexible correlation structure. We note that above specifications (1) – (7) are independent of the ordering of the asset returns in $\boldsymbol{y}_t$, while the conventional factor SV models or the Cholesky SV models (Shirota et al. (2017)) may be affected by the ordering. Further, it allows us to model the structure of the correlations in a flexible way. For example, we can easily restrict some correlation coefficients to be zero when the dimension of $\boldsymbol{y}_t$ is very high.

Remark 1. It is easy to assume that $\boldsymbol{m}_t$ and $g_{ij,t}$ follow stationary autoregressive processes. However, since it imposes the mean reversion properties on these processes, we would rather consider random walk processes without such properties for simplicity. For the long term prediction, we may need such a stationarity condition.

2.2 Realized stochastic volatilities and pairwise realized correlations

Realized measures as an additional source of information. In the above MSV models, there are too many parameters to estimate using only daily asset returns, and the parameter estimates are often unstable. Recently, high frequency data in the financial markets have become available, and they play a more important role in the finance-related empirical studies, since the realized measures of the variances and covariances, are more informative estimators of the true variances and covariances (see e.g. Andersen et al. (2001), Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen and Shephard (2004)).

Let $x_{it} = \log RV_{it}$ and $w_{ij,t} = \log\{(1 + RCOR_{ij,t})/(1 - RCOR_{ij,t})\}$ where RV_{it} and $RCOR_{ij,t}$ are the realized measures of the volatility of the i-th asset return and the correlation between i-th and j-th asset returns at time t. Thus we introduce the following additional measurement equations based on the realized measures:

$$x_{it} = \xi_i + h_{it} + u_{it} \qquad u_{it} \sim \mathrm{N}(0, \sigma_{u,i}^2), \quad t = 1, \ldots, T, \tag{8}$$

$$w_{ij,t} = \delta_{ij} + g_{ij,t} + v_{ij,t} \qquad v_{ij,t} \sim \mathrm{N}(0, \sigma_{v,ij}^2), \quad t = 1, \ldots, T, \tag{9}$$

for $i, j = 1, \ldots, p$ $(i > j)$. The terms ξ_j and δ_{ij} are included in order to adjust the biases due to the microstructure noise, nontrading hours, nonsynchronous trading and so forth. The multivariate realized stochastic volatility model with pairwise realized correlations is defined by (1) – (9). We denote $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{pt})'$, $\boldsymbol{w}_t = (w_{21,t}, \ldots, w_{pp-1,t})'$, $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)'$, $\boldsymbol{\delta} = (\delta_{21}, \ldots, \delta_{pp-1})'$, $\boldsymbol{u}_t = (u_{1t}, \ldots, u_{pt})'$, $\boldsymbol{v}_t = (v_{21,t}, \ldots, v_{pp-1,t})'$, $\boldsymbol{\sigma}_u^2 = (\sigma_{u,1}^2, \ldots, \sigma_{u,p}^2)'$, and $\boldsymbol{\sigma}_v^2 = (\sigma_{v,21}^2, \ldots, \sigma_{v,pp-1}^2)'$.

Use of pairwise realized correlations. Given the realized correlation $RCOR_{ij,t}$, we will use the pairwise realized correlations. If there is less frequent series of asset returns, the realized covariance matrix may lose a large part of the information since it is calculated only when all the series are synchronously observed. On the other hand, the pairwise realized correlation coefficients can be respectively calculated for each pair of series of returns; therefore, we can use the full information of the realized measures for the correlations. Moreover, we can estimate the parameters even if we cannot obtain the realized measures for some pairs.

Bias corrections of the realized measures. The realized volatilities and pairwise realized correlations have more information about the true volatilities and correlations, but there may be biases due to the market microstructure noise, nontrading hours, nonsynchronous trading and so forth. In order to correct these biases in the realized measures, we model the observation equations of the realized volatilities and pairwise realized correlations with bias adjustment terms, ξ_j and δ_{ij}. Although daily returns have relatively less information about the true volatilities and correlations, they are less subject to the biases that are caused by the high frequency data. Therefore, we can estimate the biases in the realized measures using the information of daily returns and also get additional information with regard to the true volatilities and correlations using the realized measures.

3 Markov chain Monte Carlo estimation

3.1 Prior distributions for parameters

Since there are many latent variables in our proposed model and hence it is difficult to evaluate the likelihood, we take the Bayesian approach and estimate the model parameters using the Markov chain Monte Carlo simulation. First we assume the prior distribution of $\boldsymbol{\theta} \equiv (\boldsymbol{\phi}, \boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{\sigma}_u^2, \boldsymbol{\sigma}_v^2, \boldsymbol{\sigma}_\zeta^2, \boldsymbol{\Sigma}_m, \boldsymbol{\Omega})$ as follows. For the prior distributions of μ_i, ξ_i and δ_{ij}, we assume multivariate independent normal distributions. The prior distributions of $\sigma_{u,i}^2, \sigma_{v,ij}^2, \sigma_{\zeta,ij}^2$ and $\sigma_{m,i}^2$ are assumed to be independent inverse gamma distributions. For ϕ_i and $\boldsymbol{\Omega}$, we assume $(1 + \phi_i)/2 \sim \text{Beta}(a, b)$ and an inverse Wishart distribution respectively. In summary, we assume the following prior distributions:

$$\mu_i \sim \text{N}(m_\mu, s_\mu^2), \quad \xi_i \sim \text{N}(m_\xi, s_\xi^2), \quad \delta_{ij} \sim \text{N}(m_\delta, s_\delta^2), \tag{10}$$

$$\sigma_{u,i}^2 \sim \text{IG}\left(\frac{n_u}{2}, \frac{d_u}{2}\right), \quad \sigma_{v,ij}^2 \sim \text{IG}\left(\frac{n_v}{2}, \frac{d_v}{2}\right), \quad \sigma_{\zeta,ij}^2 \sim \text{IG}\left(\frac{n_\zeta}{2}, \frac{d_\zeta}{2}\right), \tag{11}$$

$$\frac{1 + \phi_i}{2} \sim \text{Beta}(a, b), \quad \sigma_{m,i}^2 \sim \text{IG}\left(\frac{n_m}{2}, \frac{d_m}{2}\right), \quad \boldsymbol{\Omega} \sim \text{IW}(\nu, \boldsymbol{S}), \tag{12}$$

for $i, j = 1, \ldots, p$ $(j < i)$, and $a, b, m_\mu, s_\mu, m_\xi, s_\xi, m_\delta, s_\delta, n_u, d_u, n_v, d_v, n_\zeta, d_\zeta, n_m, d_m, \nu, \boldsymbol{S}$ are hyperparameters.

Remark 2. The particle MCMC may be a possible alternative estimation method to the MCMC below for the univariate models, but it may not be appropriate for the multivariate models since the discrete approximation to the high dimensional state distribution often results in the degeneracy of the particles.

3.2 Markov chain Monte Carlo algorithm

Let $\boldsymbol{g} = (\boldsymbol{g}_1', \ldots, \boldsymbol{g}_T')'$, $\boldsymbol{h} = (\boldsymbol{h}_1', \ldots, \boldsymbol{h}_T')'$ and $\boldsymbol{m} = (\boldsymbol{m}_1', \ldots, \boldsymbol{m}_T')'$. Further, let $\boldsymbol{w} = (\boldsymbol{w}_1', \ldots, \boldsymbol{w}_T')'$, $\boldsymbol{x} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_T')'$ and $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_T')'$. In order to conduct the statistical analysis of the parameters, we implement the Markov chain Monte Carlo simulation in nine blocks. The MCMC sampling algorithm is described in more details in the following subsections. Let $\boldsymbol{\theta}_{\backslash \boldsymbol{\beta}}$ denote the parameter $\boldsymbol{\theta}$ excluding $\boldsymbol{\beta}$. Then,

1. Initialize $\boldsymbol{g}, \boldsymbol{h}, \boldsymbol{m}$ and $\boldsymbol{\theta}$.

2. Generate $\boldsymbol{g} \,|\, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

3. Generate $\boldsymbol{h} \,|\, \boldsymbol{\theta}, \boldsymbol{m}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

4. Generate $\boldsymbol{m} \,|\, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

5. Generate $\boldsymbol{\phi} \,|\, \boldsymbol{\theta}_{\backslash \boldsymbol{\phi}}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

6. Generate $(\boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\delta}) \,|\, \boldsymbol{\theta}_{\backslash (\boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\delta},)}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

7. Generate $(\boldsymbol{\sigma}_u^2, \boldsymbol{\sigma}_v^2, \boldsymbol{\sigma}_\zeta^2, \boldsymbol{\Sigma}_m) \,|\, \boldsymbol{\theta}_{\backslash (\boldsymbol{\sigma}_u^2, \boldsymbol{\sigma}_v^2, \boldsymbol{\sigma}_\zeta^2, \boldsymbol{\Sigma}_m)}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

8. Generate $\boldsymbol{\Omega} \,|\, \boldsymbol{\theta}_{\backslash \boldsymbol{\Omega}}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{g}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$.

9. Go to Step 2.

3.2.1 Generation of $\boldsymbol{g}_t$ for the dynamic correlation matrix $\mathbf{R}_t$

The conditional posterior probability density function of $g_{ij,t}$ given other parameters and latent variables is

$$\pi(g_{ij,t}|\cdot) \propto \exp\left\{ -\frac{1}{2\sigma_{t*}^2}(g_{ij,t} - m_{t*})^2 + r(g_{ij,t}) \right\}, \tag{13}$$

$$r(g_{ij,t}) = -\frac{1}{2}\log|\mathbf{R}_t| - \frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'(\mathbf{V}_t^{1/2}\,\mathbf{R}_t\,\mathbf{V}_t^{1/2})^{-1}(\boldsymbol{y}_t - \boldsymbol{m}_t), \tag{14}$$

where

$$m_{t*} = \begin{cases} \sigma_{t*}^2\left\{ \sigma_{\zeta,ij}^{-2} g_{ij,2} + \sigma_{v,ij}^{-2}(w_{ij,1} - \delta_{ij}) \right\}, & t = 1, \\ \sigma_{t*}^2\left\{ \sigma_{\zeta,ij}^{-2}(g_{ij,t-1} + g_{ij,t+1}) + \sigma_{v,ij}^{-2}(w_{ij,t} - \delta_{ij}) \right\}, & t = 2, \ldots, T-1, \\ \sigma_{t*}^2\left\{ \sigma_{\zeta,ij}^{-2} g_{ij,T-1} + \sigma_{v,ij}^{-2}(w_{ij,T} - \delta_{ij}) \right\}, & t = T, \end{cases} \tag{15}$$

9

and

$$
\sigma_{t*}^2 = \begin{cases} \left\{(\kappa^{-1}+1)\sigma_{\zeta,ij}^{-2} + \sigma_{v,ij}^{-2}\right\}^{-1}, & t=1, \\ \left(2\sigma_{\zeta,ij}^{-2} + \sigma_{v,ij}^{-2}\right)^{-1}, & t=2,\ldots,T-1, \\ \left(\sigma_{\zeta,ij}^{-2} + \sigma_{v,ij}^{-2}\right)^{-1}, & t=T. \end{cases} \tag{16}
$$

Positive definiteness of $\mathbf{R}_t$. We use an identity matrix for the initial value of $\mathbf{R}_t$ when implementing the MCMC. Thus, given the current correlation matrix $\mathbf{R}_t$, we generate each correlation coefficient $\rho_{ij,t}$ (or equivalently $g_{ij,t}$) so that we guarantee that the proposed $\mathbf{R}_t^*$ is the correlation matrix. We first state the condition for $\rho_{ij,t}$ to guarantee that the proposed $\mathbf{R}_t^*$ is positive definite given the other elements of $\mathbf{R}_t$ and other $\rho_{ij,s}$ $(s \neq t)$.

Proposition 1.   Suppose that $\mathbf{R}_t = \{\rho_{ij,t}\}$ is a correlation matrix and let $\boldsymbol{\rho}_{it}$ denote the transpose of the i-th row vector of $\mathbf{R}_t$ excluding 1, $\boldsymbol{\rho}_{it} = (\rho_{i1,t},\ldots,\rho_{ii-1,t},\rho_{ii+1,t},\ldots,\rho_{ip,t})'$, and $\mathbf{R}_{it}$ denotes the submatrix excluding the i-th row and the i-th column from $\mathbf{R}_t$. The condition for $\rho_{ij,t}$ to guarantee that $\mathbf{R}_t$ is positive definite is $\rho_{ij,t} \in (L_{ijt}, U_{ijt})$ where bounds L_{ijt} and U_{ijt} are given by

$$
\frac{-\boldsymbol{b}_j'\,\boldsymbol{\rho}_{i,-j,t} \pm \sqrt{(\boldsymbol{b}_j'\,\boldsymbol{\rho}_{i,-j,t})^2 - a_j(\boldsymbol{\rho}_{i,-j,t}'\,\mathbf{C}_j\,\boldsymbol{\rho}_{i,-j,t}-1)}}{a_j}, \tag{17}
$$

and $\boldsymbol{\rho}_{i,-j,t}$ is the vector excluding the j-th element of $\boldsymbol{\rho}_{it}$, a_j is the (j,j)-th element of $\mathbf{R}_{it}^{-1}$, $\boldsymbol{b}_j$ is the vector excluding a_j from the j-th column of $\mathbf{R}_{it}^{-1}$, and $\mathbf{C}_j$ is the matrix excluding the j-th row and j-th column from $\mathbf{R}_{it}^{-1}$.

Proof: See Appendix B.1

Thus we propose a candidate $g_{ij,t}^{\dagger}$ from normal distribution truncated on the interval (a_{ijt}, b_{ijt}), $TN_{(a_{ijt},b_{ijt})}(m_{t*}, \sigma_{t*}^2)$, and accept it with probability $\min\{1, \exp(r(g_{ij,t}^{\dagger})-r(g_{ij,t}))\}$, where

$$
(a_{ijt}, b_{ijt}) \equiv \left(\log \frac{1+L_{ij,t}}{1-L_{ij,t}}, \log \frac{1+U_{ij,t}}{1-U_{ij,t}}\right). \tag{18}
$$

3.2.2   Generation of $\boldsymbol{h}_t$ for the dynamic volatility $\mathbf{V}_t$

We use a single-move sampler for $\boldsymbol{h}_t$ in which we sample $\boldsymbol{h}_t$ given the other parameters and latent variables. Such a sampler is efficient when the realized measures are available as the

additional information source for $\boldsymbol{h}_t$. The conditional posterior probability density function of $\boldsymbol{h}_t$ is given by

$$\pi(\boldsymbol{h}_t \,|\cdot) \propto \exp\left[-\frac{1}{2}(\boldsymbol{h}_t - \boldsymbol{m}_{t*})' \, \boldsymbol{\Omega}_{t*}^{-1}(\boldsymbol{h}_t - \boldsymbol{m}_{t*}) + l(\boldsymbol{h}_t) \right], \tag{19}$$

$$l(\boldsymbol{h}_t) = -\frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'(\mathbf{V}_t^{1/2}\,\mathbf{R}_t\,\mathbf{V}_t^{1/2})^{-1}(\boldsymbol{y}_t - \boldsymbol{m}_t), \tag{20}$$

where

$$\boldsymbol{m}_{t*} = \begin{cases} \boldsymbol{\Omega}_{1*}\left[\boldsymbol{\Omega}_0^{-1}\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{\Omega}^{-1}\left\{\boldsymbol{h}_2 - (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu}\right\} \right. \\ \qquad\qquad\qquad\qquad\qquad \left. + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_1 - \boldsymbol{\xi}) - \frac{1}{2}\,\mathbf{1}_p\right], & t = 1, \\ \boldsymbol{\Omega}_{t*}\left[\boldsymbol{\Omega}^{-1}\left\{(\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{h}_{t-1}\right\}\right. \\ \quad \left. + \boldsymbol{\Phi}\,\boldsymbol{\Omega}^{-1}\left\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu}\right\} + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi}) - \frac{1}{2}\,\mathbf{1}_p\right], & t = 2, \ldots, T - 1, \\ \boldsymbol{\Omega}_{T*}\left[\boldsymbol{\Omega}^{-1}\left\{(\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{h}_{T-1}\right\} + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_T - \boldsymbol{\xi}) - \frac{1}{2}\,\mathbf{1}_p\right], & t = T, \end{cases} \tag{21}$$

$$\boldsymbol{\Omega}_{t*} = \begin{cases} \left[\boldsymbol{\Omega}_0^{-1} + \boldsymbol{\Phi}\,\boldsymbol{\Omega}^{-1}\,\boldsymbol{\Phi} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = 1, \\ \left[\boldsymbol{\Omega}^{-1} + \boldsymbol{\Phi}\,\boldsymbol{\Omega}^{-1}\,\boldsymbol{\Phi} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = 2, \ldots, T - 1, \\ \left[\boldsymbol{\Omega}^{-1} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = T, \end{cases} \tag{22}$$

where $\mathbf{1}_p$ denotes a $p \times 1$ vector with all elements equal to one. Therefore, we generate a candidate $\boldsymbol{h}_t^{\dagger}$ from $\mathrm{N}(\boldsymbol{m}_{t*}, \boldsymbol{\Omega}_{t*})$, and accept it with probability $\min\{1, \exp(l(\boldsymbol{h}_t^{\dagger}) - l(\boldsymbol{h}_t))\}$. See Appendix A.1 for the generations of $\boldsymbol{\theta}$ and $\boldsymbol{m}_t$.

4 Extension to incorporate the leverage effect

This section extends our model in order to incorporate the leverage effect. The leverage effect, which corresponds to the well-known negative correlation between asset returns and their volatilities in the stock market, is expected to improve the performance of the forecast of the mean processes and volatility processes of asset returns.

4.1 Matrix variate normal distribution

We first define the matrix variate normal distribution and show its probability density function, which will be used in modelling the leverage effect.

Definition 1. The random matrix $\mathbf{X}$ $(p \times n)$ is said to have a matrix variate normal distribution with mean matrix $\mathbf{M}$ $(p \times n)$ and covariance matrix $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ where $\boldsymbol{\Psi}$ $(p \times p)$ and $\boldsymbol{\Sigma}$ $(n \times n)$ are positive definite matrices if $\mathrm{vec}\,(\mathbf{X}') \sim \mathrm{N}(\mathrm{vec}\,(\mathbf{M}'), \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$ and we denote $\mathbf{X} \sim \mathrm{N}_{p,n}(\mathbf{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$.

4.2 Modeling the leverage effect

We extend our proposed model to incorporate the leverage effect as follows. The joint distribution of $(\boldsymbol{y}_t, \boldsymbol{h}_{t+1})$ is given by

$$
\begin{pmatrix} \boldsymbol{y}_t \\ \boldsymbol{h}_{t+1} \end{pmatrix} \sim \mathrm{N} \left(\begin{pmatrix} \boldsymbol{m}_t \\ \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) \end{pmatrix}, \begin{pmatrix} \mathbf{V}_t^{1/2} \mathbf{R}_t \mathbf{V}_t^{1/2} & \mathbf{V}_t^{1/2} \mathbf{R}_t^{1/2} \boldsymbol{\Lambda}' \\ \boldsymbol{\Lambda} \mathbf{R}_t^{1/2'} \mathbf{V}_t^{1/2} & \boldsymbol{\Psi} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}' \end{pmatrix} \right). \tag{23}
$$

The marginal distributions of $\boldsymbol{y}_t$ and $\boldsymbol{h}_{t+1}$ given $\boldsymbol{h}_t$ are the same as before with $\boldsymbol{\Omega} = \boldsymbol{\Psi} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}'$, but we note that

$$
\boldsymbol{h}_{t+1} | \boldsymbol{y}_t, \boldsymbol{h}_t, \boldsymbol{\theta} \sim N \left(\boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\Lambda} \mathbf{R}_t^{-1/2} \mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t), \boldsymbol{\Psi} \right).
$$

If $\boldsymbol{\Lambda} = \mathbf{O}$, it reduces to the model without leverage effect. The matrix $\boldsymbol{\Lambda}$ is the coefficient of the leverage for $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2} \mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$. We assume that the prior distribution of $\boldsymbol{\Lambda}$ given $\boldsymbol{\Psi}$ is $\mathrm{N}_{p,p}(\mathbf{M}_0, \boldsymbol{\Psi} \otimes \boldsymbol{\Gamma}_0)$. That is, $\boldsymbol{\Lambda} | \boldsymbol{\Psi} \sim \mathrm{N}_{p,p}(\mathbf{M}_0, \boldsymbol{\Psi} \otimes \boldsymbol{\Gamma}_0)$.

Remark 3. There are several ways to choose $\mathbf{R}_t^{1/2}$. For example, we can use the spectral decomposition of the correlation matrix $\mathbf{R}_t = \mathbf{P}_t \mathbf{Q}_t \mathbf{P}_t'$ and $\mathbf{R}_t^{1/2} = \mathbf{P}_t \mathbf{Q}_t^{1/2}$ where the i-th diagonal element of the diagonal matrix $\mathbf{Q}_t$ is the i-th largest eigenvalue of $\mathbf{R}_t$ and the i-th column of $\mathbf{P}_t$ is the corresponding i-th eigenvector (and we set the first elements of the eigenvectors to be positive for the identification purpose). Thus the i-th element of $\boldsymbol{z}_t$ can be interpreted as the i-th market factor among p asset returns. Alternatively, Cholesky decomposition, $\mathbf{R}_t = \mathbf{R}_t^{1/2} \mathbf{R}_t^{1/2'}$, can be used so that $\mathbf{R}_t^{1/2}$ is a lower triangular matrix where all the diagonal elements are equal to one, but we note that it is affected by the ordering of the asset returns.

4.2.1 Generation of $\boldsymbol{\Lambda}$

The conditional posterior distribution of $\boldsymbol{\Lambda}$ is derived in the following Proposition and we generate $\mathrm{vec}(\boldsymbol{\Lambda}) | \cdot \sim \mathrm{N}(\mathrm{vec}(\mathbf{M}_1'), \boldsymbol{\Psi} \otimes \boldsymbol{\Gamma}_1)$.

Proposition 2. Suppose that the prior distribution of $\boldsymbol{\Lambda}$ given $\boldsymbol{\Psi}$ is $\mathrm{N}_{p,p}(\mathbf{M}_0, \boldsymbol{\Psi} \otimes \boldsymbol{\Gamma}_0)$. Then the conditional posterior distribution of $\boldsymbol{\Lambda}$ given other parameters and latent variables

is $N_{p,p}(\mathbf{M}_1, \boldsymbol{\Psi} \otimes \boldsymbol{\Gamma}_1)$ where

$$\mathbf{M}_1 = \left(\mathbf{A} + \boldsymbol{\Gamma}_0^{-1}\right)^{-1}\left(\mathbf{B} + \boldsymbol{\Gamma}_0^{-1}\mathbf{M}_0\right), \quad \boldsymbol{\Gamma}_1 = \left(\mathbf{A} + \boldsymbol{\Gamma}_0^{-1}\right)^{-1}, \tag{24}$$

$$\mathbf{A} = \sum_{t=1}^{T-1} \boldsymbol{z}_t \boldsymbol{z}_t', \quad \mathbf{B} = \sum_{t=1}^{T-1} \boldsymbol{z}_t \boldsymbol{\eta}_t', \tag{25}$$

and $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$ and $\boldsymbol{\eta}_t = \boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu})$.

Proof: See Appendix B.2.

For the generations of other parameters and latent variables, see Appendix A.2 .

4.3 Parsimonious specification of the leverage effect

This subsection proposes the parsimonious specification for $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_p]$, in order to reduce the number of leverage parameters from p^2 to pq ($q \ll p$) by setting $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_q, \mathbf{0}, \ldots, \mathbf{0}]$ since we do not have additional measurement equations for the leverage effect. Using the spectral decomposition to compute $\mathbf{R}_t^{1/2}$, we can interpret that the i-th column corresponds to the i-th market factor among asset returns ($i = 1, \ldots, q$). The number of factors, q, is expected to be small, e.g., $q = 1$ or $q = 2$.

4.3.1 Generation of $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_q, \mathbf{0}, \ldots, \mathbf{0}]$

The following proposition and the corollary shows the conditional posterior distribution of the parameters for the leverage effect under parsimonious specifications.

Proposition 3. Let $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_q, \mathbf{0}, \ldots, \mathbf{0}]$ and $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1', \ldots, \boldsymbol{\lambda}_q')'$. If the prior distribution of $\boldsymbol{\lambda}$ is assumed to be normal, $\boldsymbol{\lambda} \sim N(\boldsymbol{m}_0, \boldsymbol{\Gamma}_0)$, then the conditional posterior distribution of $\boldsymbol{\lambda}$ is $\boldsymbol{\lambda} | \cdot \sim N(\boldsymbol{m}_1, \boldsymbol{\Gamma}_1)$ where

$$\boldsymbol{m}_1 = \boldsymbol{\Gamma}_1\left\{\boldsymbol{\Gamma}_0^{-1}\boldsymbol{m}_0 + (\mathbf{I}_q \otimes \boldsymbol{\Psi}^{-1}\mathbf{B}')\mathrm{vec}\left(\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_q\}\right)\right\}, \quad \boldsymbol{\Gamma}_1 = \left(\boldsymbol{\Gamma}_0^{-1} + \mathbf{A}_{1:q,1:q} \otimes \boldsymbol{\Psi}^{-1}\right)^{-1},$$

$\mathbf{A}, \mathbf{B}$ are defined in (25), $\mathbf{A}_{1:q,1:q}$ denotes the first q rows and the q columns of $\mathbf{A}$, $\mathrm{vec}(\mathbf{X}) \equiv (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_m')'$ denotes a vectorization of the matrix $\mathbf{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$, and $\otimes$ denotes Kronecker product.

Proof: See Appendix B.3.

Corollary 1. Let $q = 1$ and $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}, \boldsymbol{0}, \ldots, \boldsymbol{0}]$. If the prior distribution of $\boldsymbol{\lambda}$ is assumed to be normal, $\boldsymbol{\lambda} \sim \mathrm{N}(\boldsymbol{m}_0, \boldsymbol{\Gamma}_0)$, then the conditional posterior distribution of $\boldsymbol{\lambda}$ is $\boldsymbol{\lambda} \,|\, \cdot \sim \mathrm{N}(\boldsymbol{m}_1, \boldsymbol{\Gamma}_1)$ where

$$
\begin{aligned}
\boldsymbol{m}_1 &= \boldsymbol{\Gamma}_1 \left\{ \boldsymbol{\Gamma}_0^{-1} \boldsymbol{m}_0 + \boldsymbol{\Psi}^{-1} \boldsymbol{b} \right\}, \quad \boldsymbol{\Gamma}_1 = \left(\boldsymbol{\Gamma}_0^{-1} + a \times \boldsymbol{\Psi}^{-1} \right)^{-1}, \\
a &= \sum_{t=1}^{T-1} z_{1t}^2, \quad \boldsymbol{b} = \sum_{t=1}^{T-1} z_{1t} \left\{ \boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) \right\},
\end{aligned}
\tag{26}
$$

and z_{1t} is the first element of $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2} \mathbf{V}_t^{-1/2} (\boldsymbol{y}_t - \boldsymbol{m}_t)$.

4.3.2 Generation of $\boldsymbol{\Psi}$

See Appendix A.2.

5 Empirical studies

This section applies our proposed model to the daily returns of nine U.S. stocks ($p = 9$) with realized volatilities and pairwise realized correlations. The nine series of stock returns are JP Morgan (JPM), International Business Machine (IBM), Microsoft (MSFT), Exxon Mobil (XOM), Alcoa (AA), American Express (AXP), Du Pont (DD), General Electric (GE), and Coca Cola (KO). The sample period is from February 1, 2001 to December 31, 2009, and the number of observation is $T = 2242$. 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.
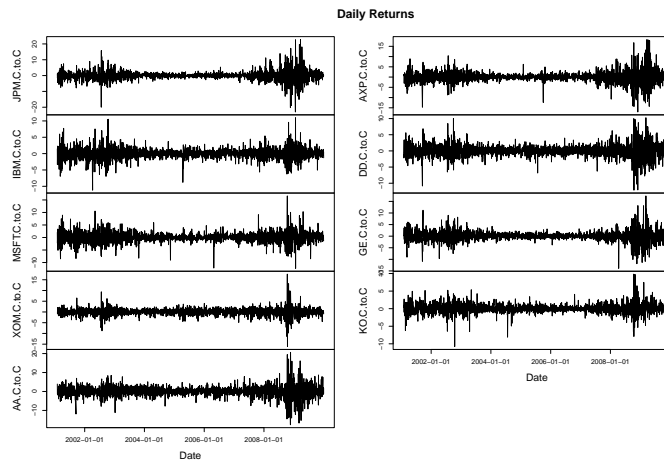

Figure 1: Time series plots of nine US stock (close-to-close) returns.

14

Time series plots of y_{it} are shown in Figure 1, and show that there is a very high volatility period in 2008 (the financial crisis when Lehman Brothers filed for Chapter 11 bankruptcy protection). Additionally, there are other relatively high volatility periods in 2001 (the dot-com bubble and the September 11 attacks) and in 2002 (the market turmoil during which Worldcom filed for Chapter 11 bankruptcy protection). The realized volatilities and pairwise realized correlations are computed from the realized covariance matrices for these assets which can be downloaded from the Oxford Man Institute website (see, Section 5 of Noureldin et al. (2012) for details). The prior distributions are assumed to be vague and flat in order to reflect the fact that we have little information with regard to the parameters:

$$\mu_i \sim \mathrm{N}(0, 10^4), \quad \xi_i \sim \mathrm{N}(0, 10^4), \quad \delta_{ij} \sim \mathrm{N}(0, 10^4), \quad \lambda_i \sim \mathrm{N}(0, 10^4),$$

$$\frac{1 + \phi_i}{2} \sim \mathrm{Beta}(1, 1), \quad \sigma_{u,i}^2 \sim \mathrm{IG}(10^{-5}/2, 10^{-5}/2), \quad \sigma_{v,ij}^2 \sim \mathrm{IG}(10^{-5}/2, 10^{-5}/2),$$

$$\sigma_{\zeta,ij}^2 \sim \mathrm{IG}(10^{-6}/2, 10^{-6}/2), \quad \sigma_{m,i}^2 \sim \mathrm{IG}(10^{-5}/2, 10^{-5}/2), \qquad \boldsymbol{\Psi} \sim \mathrm{IW}(10, \mathbf{I}_9),$$

for $i = 1, \ldots, p$, $j = 1, \ldots, i-1$. The proposed model is estimated and we use the parsimonious specification of the leverage effect with the number of factors $q = 1$.

5.1 Estimation results

We run 12,000 MCMC iterations and the first 2,000 iterations are discarded as the burn-in period. Table 1 shows the posterior means, 95% credible intervals and inefficiency factors[1] for $\boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\sigma}_u, \boldsymbol{\sigma}_m$ and $\boldsymbol{\lambda}$. The inefficiency factors are relatively small (less than 130) in the multivariate stochastic volatility models and our algorithm works well. The posterior means and posterior standard deviations for $\boldsymbol{\delta}$, $\boldsymbol{\sigma}_v$ and $\boldsymbol{\sigma}_\zeta$ are shown in Tables 2, 3 and 4, respectively.

Mean processes and volatilities. The posterior means of $\sigma_{m,i}$ are around $0.067 \sim 0.098$, which reflects that the magnitude of the mean process, $\boldsymbol{m}_t$, is much smaller than that of the stochastic volatility component, $\mathbf{V}_t^{1/2}\boldsymbol{\epsilon}_t$, as we expected. The unconditional means of the log volatilities, μ_i, are estimated to be from 0.203 to 1.582 and the posterior mean of μ_5 (corresponding to Alcoa) is much larger than those of others. The stock returns of Alcoa are found to be the most volatile among others, while those of Coca Cola are the least volatile.

---

[1]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 becomes, the closer the MCMC sampling is to the uncorrelated sampling.

Table 1: Posterior means, 95% credible intervals and inefficiency
factors for $\boldsymbol{\mu}$, $\boldsymbol{\xi}$, $\boldsymbol{\phi}$, $\boldsymbol{\sigma}_u$, $\boldsymbol{\sigma}_m$ and $\boldsymbol{\lambda}$. Spectral decomposition is used to compute $\mathbf{R}_t^{-1/2}$.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
μ_1	1.221	[1.020,1.430]	8	$\sigma_{u,1}$	0.285	[0.270,0.300]	40
μ_2	0.644	[0.495,0.794]	16	$\sigma_{u,2}$	0.286	[0.271,0.303]	57
μ_3	0.914	[0.759,1.070]	9	$\sigma_{u,3}$	0.291	[0.278,0.304]	14
μ_4	0.684	[0.541,0.827]	11	$\sigma_{u,4}$	0.274	[0.261,0.287]	20
μ_5	1.582	[1.430,1.730]	8	$\sigma_{u,5}$	0.318	[0.304,0.332]	17
μ_6	1.136	[0.918,1.360]	9	$\sigma_{u,6}$	0.307	[0.293,0.321]	20
μ_7	0.873	[0.726,1.020]	12	$\sigma_{u,7}$	0.291	[0.278,0.304]	20
μ_8	0.861	[0.670,1.050]	10	$\sigma_{u,8}$	0.303	[0.289,0.317]	18
μ_9	0.203	[0.055,0.352]	15	$\sigma_{u,9}$	0.294	[0.281,0.308]	18
ξ_1	−0.520	[-0.582,-0.470]	102	$\sigma_{m,1}$	0.087	[0.067,0.108]	118
ξ_2	−0.554	[-0.610,-0.495]	101	$\sigma_{m,2}$	0.070	[0.053,0.089]	126
ξ_3	−0.549	[-0.594,-0.501]	90	$\sigma_{m,3}$	0.076	[0.053,0.106]	129
ξ_4	−0.442	[-0.487,-0.394]	88	$\sigma_{m,4}$	0.087	[0.067,0.106]	117
ξ_5	−0.537	[-0.582,-0.494]	78	$\sigma_{m,5}$	0.098	[0.073,0.135]	129
ξ_6	−0.586	[-0.651,-0.533]	109	$\sigma_{m,6}$	0.088	[0.071,0.118]	119
ξ_7	−0.428	[-0.480,-0.376]	102	$\sigma_{m,7}$	0.087	[0.063,0.111]	123
ξ_8	−0.535	[-0.589,-0.474]	100	$\sigma_{m,8}$	0.078	[0.060,0.108]	125
ξ_9	−0.322	[-0.376,-0.263]	93	$\sigma_{m,9}$	0.067	[0.048,0.088]	124
ϕ_1	0.914	[0.904,0.924]	19	λ_1	−0.0626	[-0.0852,-0.0403]	7
ϕ_2	0.888	[0.874,0.901]	22	λ_2	−0.0541	[-0.0757,-0.0330]	7
ϕ_3	0.900	[0.887,0.913]	19	λ_3	−0.0430	[-0.0638,-0.0216]	7
ϕ_4	0.890	[0.876,0.904]	21	λ_4	−0.0518	[-0.0722,-0.0311]	6
ϕ_5	0.907	[0.895,0.920]	20	λ_5	−0.0424	[-0.0625,-0.0219]	7
ϕ_6	0.926	[0.916,0.935]	25	λ_6	−0.0518	[-0.0735,-0.0303]	6
ϕ_7	0.899	[0.886,0.911]	22	λ_7	−0.0536	[-0.0736,-0.0331]	9
ϕ_8	0.908	[0.897,0.920]	21	λ_8	−0.0538	[-0.0767,-0.0308]	8
ϕ_9	0.903	[0.889,0.916]	21	λ_9	−0.0436	[-0.0637,-0.0235]	10

Since all posterior means of the autoregressive coefficients, ϕ_i, are approximately 0.9, the log volatilities are found to have high persistence. The elements of $\boldsymbol{\Psi}$ (the conditional covariance matrix of $\boldsymbol{h}_{t+1}$ given $\boldsymbol{y}_t$) are all approximately 0.1 and the probability that ψ_{ij} is positive is greater than 0.975 for all is and js. The log volatilities, $h_{i,t+1}$, are positively correlated with each other given $\boldsymbol{y}_t$. Figure 2 shows the 95% credible intervals for h_{1t} with $x_{1t} - \xi_1$ where ξ_1 is the estimated posterior mean of the first bias correction term. The figures for

h_{it} $(i = 2, \ldots, 9)$ are similar and hence are omitted. The estimated 95 % credible intervals have smaller fluctuation than those of the bias-adjusted realized measures. These estimates succeeded at automatically extracting the mean trends of the volatilities and adjusting the measurement errors. Overall, the 95% credible intervals captures the traceplot of the (bias-corrected) realized volatilities, suggesting that our proposed model is successful at describing the dynamics of the latent log volatilities.
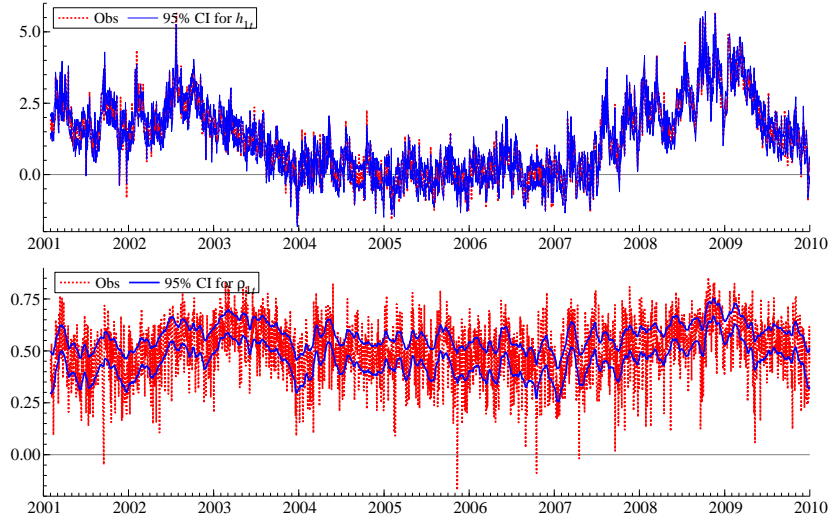
Figure 2: Top: 95% credible intervals (solid) of h_{1t}, and $x_{1t} - \xi_1$ (dotted). Bottom: 95% credible intervals (solid) of $\rho_{21,t}$ and $\{\exp(w_{21,t} - \delta_{21}) - 1\}/\{\exp(w_{21,t} - \delta_{21}) + 1\}$ (dotted).

Table 2: Posterior means (posterior standard deviations) of $\boldsymbol{\delta}$.

δ_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	**-0.313**							
	(0.029)							
$i = 3$	**-0.265**	**-0.333**						
	(0.031)	(0.053)						
$i = 4$	**-0.301**	**-0.180**	**-0.149**					
	(0.048)	(0.035)	(0.065)					
$i = 5$	**-0.410**	**-0.227**	**-0.246**	**-0.287**				
	(0.065)	(0.027)	(0.038)	(0.043)				
$i = 6$	**-0.629**	**-0.265**	**-0.295**	**-0.353**	**-0.391**			
	(0.040)	(0.032)	(0.048)	(0.045)	(0.036)			
$i = 7$	**-0.472**	**-0.276**	**-0.219**	**-0.312**	**-0.532**	**-0.477**		
	(0.047)	(0.029)	(0.030)	(0.025)	(0.035)	(0.032)		
$i = 8$	**-0.526**	**-0.370**	**-0.319**	**-0.340**	**-0.440**	**-0.563**	**-0.478**	
	(0.053)	(0.041)	(0.065)	(0.037)	(0.029)	(0.048)	(0.033)	
$i = 9$	**-0.194**	**-0.081**	**-0.098**	**-0.249**	**-0.150**	**-0.246**	**-0.158**	**-0.172**
	(0.043)	(0.033)	(0.026)	(0.050)	(0.029)	(0.030)	(0.026)	(0.057)

*Bold figures indicate that the 95% credible interval does not include zero.

Table 3: Posterior means (posterior standard deviations) of $\boldsymbol{\sigma}_v$.

$\sigma_{v,ij}$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
$i=2$	0.329							
	(0.006)							
$i=3$	0.325	0.325						
	(0.005)	(0.005)						
$i=4$	0.329	0.329	0.310					
	(0.005)	(0.006)	(0.005)					
$i=5$	0.329	0.322	0.314	0.328				
	(0.005)	(0.005)	(0.005)	(0.005)				
$i=6$	0.350	0.331	0.315	0.319	0.332			
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)			
$i=7$	0.338	0.338	0.319	0.334	0.344	0.338		
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)		
$i=8$	0.334	0.320	0.304	0.322	0.317	0.332	0.326	
	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	
$i=9$	0.314	0.329	0.307	0.317	0.320	0.321	0.337	0.328
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)

Table 4: Posterior means (posterior standard deviations) of $\boldsymbol{\sigma}_\zeta$.

$\sigma_{\zeta,ij}$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
$i=2$	0.047							
	(0.005)							
$i=3$	0.039	0.043						
	(0.004)	(0.004)						
$i=4$	0.061	0.062	0.061					
	(0.004)	(0.006)	(0.005)					
$i=5$	0.047	0.044	0.042	0.047				
	(0.004)	(0.004)	(0.004)	(0.004)				
$i=6$	0.058	0.046	0.044	0.069	0.047			
	(0.005)	(0.005)	(0.005)	(0.006)	(0.004)			
$i=7$	0.041	0.044	0.041	0.061	0.053	0.049		
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)		
$i=8$	0.048	0.052	0.055	0.071	0.049	0.054	0.049	
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	
$i=9$	0.049	0.042	0.040	0.064	0.043	0.045	0.044	0.048
	(0.005)	(0.005)	(0.004)	(0.006)	(0.004)	(0.005)	(0.005)	(0.004)

Biases in realized volatilities and correlations. The bias correction terms, ξ_i, of the realized volatilities are estimated to be negative, thereby indicating that the realized volatilities have downward biases and underestimate the volatilities by ignoring the overnight nontrading hours. Since the realized volatilities tend to overestimate the volatilities due to the microstructure noises, the effect of nontrading hours seems to dominate in the direction of the biases. We also note that the magnitudes of the biases depend on the series of stock returns.

Table 2 shows the estimation result of the bias term $\boldsymbol{\delta}$ of the correlation coefficients. All δ_{ij} are estimated to be negative, and the posterior probability that δ_{ij} is negative is greater than 0.975. This implies that the realized correlations underestimate the latent correlations, thereby suggesting the existence of the Epps effect.

Dynamic correlations. The posterior means of the standard deviations of the disturbance terms in the state equations corresponding to the dynamic correlations, $\sigma_{\zeta,ij}$, are shown in Table 4. They are $0.039 \sim 0.071$ and are much smaller than the posterior means of the standard deviations for the measurement errors of the realized measures, $\sigma_{u,i}$ and $\sigma_{v,ij}$ (as shown in Tables 1 and 3) which are found to be similar for all is and js at approximately 0.30. Figure 2 shows the time series plots of the 95% credible intervals of the selected dynamic correlations, $\rho_{21,t}$ with $\{\exp(w_{21,t} - \delta_{21}) - 1\}/\{\exp(w_{21,t} - \delta_{21}) + 1\}$ where δ_{21} is the estimated posterior mean. The figures for the other $\rho_{ij,t}$ are similar and hence are omitted. Again, the estimated 95% credible intervals of $\rho_{ij,t}$ have much smaller fluctuation those of bias-adjusted realized measures, $\{\exp(x_{ij,t} - \delta_{ij}) - 1\}/\{\exp(x_{ij,t} - \delta_{ij}) + 1\}$. These intervals seem to extract the mean trends of the bias-adjusted realized measures that have relatively large noises in the measurement equation. The correlations between the asset returns are found to be time-varying in the sample period, and they seem to increase after the financial crisis in 2008. This result corresponds to our intuition that each asset return has a larger positive correlation with others when the market faces stress, rather than when it is in a usual period.

Leverage effect and the selection of the number of factors q. The parameters for the leverage effect, λ_i, are estimated to be negative in Table 1 and the posterior probability that λ_i is negative is greater than 0.975 for all is. This implies the existence of the leverage effect. Table 5 also shows the estimation results for the correlation between the first element of $\boldsymbol{z}_t = \mathbf{V}_t^{-1/2}\mathbf{R}_t^{-1/2}\boldsymbol{y}_t$ and $h_{i,t+1}$, *i.e.*, $\rho_i^* = Corr(z_{1t}, h_{i,t+1}) = \lambda_{ii}/\sqrt{\lambda_{ii}^2 + \psi_{ii}}$ for $i = 1, \ldots, 9$. The posterior means of ρ_i^* are estimated to be negative ranging from -0.22 to -0.15. If we regard z_{1t} as the market factor, a decrease in the market return (z_{1t}) is followed by an increase in the log volatility ($h_{i,t+1}$), which implies the existence of the leverage effect. The estimation results for $\boldsymbol{\Psi}$ are omitted in order to save space where the posterior probability that $\psi_{ij} > 0$ is found to be greater than 0.975 for all is and js.

To investigate whether the number of factors is $q = 1$, we also fit the proposed model using $q = 2$ and we set $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{0}, \ldots, \boldsymbol{0}]$ for the leverage effect. Table 6 shows the

posterior means, the 95% credible intervals and inefficiency factors for $\rho_{1i}^* = Corr(z_{1t}, h_{i,t+1})$ and $\rho_{2i}^* = Corr(z_{2t}, h_{i,t+1})$ for $i = 1, \ldots, 9$ where $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$. The estimation results for ρ_{1i}^* are almost the same as those for ρ_i^* in Table 5. Conversely, the posterior means of ρ_{2i}^* are close to zeros, and the 95% credible intervals include zero. This suggests that one factor $(q = 1)$ is enough to describe the leverage effect for our dataset.

Table 5: Posterior means, 95% credible intervals and inefficiency factors for $\rho_i^* = Corr(z_{1t}, h_{i,t+1})$ where $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$ and $q = 1$. Spectral decomposition is used.

Par.	Mean	95% interval	IF
ρ_1^*	-0.199	[-0.306,-0.113]	8
ρ_2^*	-0.187	[-0.302,-0.100]	8
ρ_3^*	-0.153	[-0.261,-0.067]	7
ρ_4^*	-0.199	[-0.327,-0.101]	7
ρ_5^*	-0.179	[-0.322,-0.076]	8
ρ_6^*	-0.177	[-0.288,-0.089]	8
ρ_7^*	-0.224	[-0.374,-0.114]	10
ρ_8^*	-0.171	[-0.278,-0.087]	8
ρ_9^*	-0.187	[-0.334,-0.083]	11

Table 6: Posterior means, 95% credible intervals and inefficiency factors for $\rho_{1i}^* = Corr(z_{1t}, h_{i,t+1})$ and $\rho_{2i}^* = Corr(z_{2t}, h_{i,t+1})$ where $q = 2$. Spectral decomposition is used.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
ρ_{11}^*	-0.195	[-0.305,-0.107]	6	ρ_{21}^*	-0.009	[-0.103,0.071]	35
ρ_{12}^*	-0.183	[-0.305,-0.091]	10	ρ_{22}^*	0.016	[-0.084,0.096]	41
ρ_{13}^*	-0.147	[-0.254,-0.063]	8	ρ_{23}^*	-0.001	[-0.141,0.106]	64
ρ_{14}^*	-0.199	[-0.334,-0.097]	6	ρ_{24}^*	0.037	[-0.054,0.116]	41
ρ_{15}^*	-0.187	[-0.343,-0.080]	11	ρ_{25}^*	0.025	[-0.096,0.114]	51
ρ_{16}^*	-0.166	[-0.276,-0.079]	7	ρ_{26}^*	-0.013	[-0.109,0.070]	31
ρ_{17}^*	-0.225	[-0.400,-0.111]	8	ρ_{27}^*	-0.012	[-0.125,0.081]	45
ρ_{18}^*	-0.173	[-0.284,-0.084]	8	ρ_{28}^*	0.021	[-0.079,0.099]	39
ρ_{19}^*	-0.183	[-0.337,-0.075]	13	ρ_{29}^*	0.054	[-0.066,0.143]	55

Cholesky and spectral decompositions for computing $\mathbf{R}_t^{-1/2}$. We also estimated our proposed models with $q = 1$ and 2 using the Cholesky decomposition instead of the spectral decomposition. The estimation results using the Cholesky decomposition are very similar to those using the spectral decomposition (and hence are omitted) except for the parameters of the

leverage effect. Table 7 shows the estimation results for the correlation, ρ_i^*, with $q = 1$. All posterior means are estimated to be negative and the posterior probability that ρ_i^* is negative is greater than 0.975 for all is. However, we note that the absolute values of ρ_i^* are smaller than those in the model using the spectral decomposition.

Table 7: Posterior means, 95% credible intervals and inefficiency factors for $\rho_i^* = Corr(z_{1t}, h_{i,t+1})$ where $\boldsymbol{z}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$ and $q = 1$. Cholesky decomposition is used.

Par.	Mean	95% interval	IF
ρ_1^*	-0.170	[-0.266,-0.092]	7
ρ_2^*	-0.118	[-0.199,-0.049]	4
ρ_3^*	-0.091	[-0.174,-0.021]	8
ρ_4^*	-0.118	[-0.206,-0.045]	7
ρ_5^*	-0.123	[-0.239,-0.044]	7
ρ_6^*	-0.119	[-0.205,-0.048]	7
ρ_7^*	-0.141	[-0.244,-0.061]	9
ρ_8^*	-0.144	[-0.237,-0.067]	10
ρ_9^*	-0.148	[-0.271,-0.058]	8

Table 8: Posterior means, 95% credible intervals and inefficiency factors for $\rho_{1i}^* = Corr(z_{1t}, h_{i,t+1})$ and $\rho_{2i}^* = Corr(z_{2t}, h_{i,t+1})$ where $q = 2$. Cholesky decomposition is used.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
ρ_{11}^*	-0.168	[-0.271,-0.087]	7	ρ_{21}^*	-0.029	[-0.089, 0.024]	5
ρ_{12}^*	-0.132	[-0.230,-0.056]	6	ρ_{22}^*	-0.113	[-0.200,-0.043]	7
ρ_{13}^*	-0.100	[-0.189,-0.029]	7	ρ_{23}^*	-0.073	[-0.147,-0.009]	6
ρ_{14}^*	-0.120	[-0.212,-0.046]	4	ρ_{24}^*	-0.019	[-0.080, 0.036]	7
ρ_{15}^*	-0.128	[-0.240,-0.045]	9	ρ_{25}^*	-0.044	[-0.117, 0.018]	8
ρ_{16}^*	-0.125	[-0.215,-0.051]	8	ρ_{26}^*	-0.028	[-0.093, 0.029]	5
ρ_{17}^*	-0.149	[-0.262,-0.064]	7	ρ_{27}^*	-0.081	[-0.160,-0.016]	8
ρ_{18}^*	-0.145	[-0.239,-0.066]	6	ρ_{28}^*	-0.011	[-0.071, 0.042]	5
ρ_{19}^*	-0.151	[-0.269,-0.058]	10	ρ_{29}^*	-0.003	[-0.069, 0.052]	7

Table 8 shows the estimation results for the correlations, ρ_{1i}^* and ρ_{2i}^*, with $q = 2$. The estimation results for ρ_{1i}^* are similar to those for ρ_i^* in Table 7, but all posterior means of ρ_{2i}^* are estimated to be negative, and the posterior probability that ρ_{2i}^* is negative is greater than 0.975 for $i = 2, 3$ and 7. This implies that we need to include more factors when we use the Cholesky decomposition. Since the number of factors q depends on the order of the

asset return, we have to find the order that minimizes q for the parsimonious specification. Conversely, the spectral decomposition does not depend on the order of the asset returns and it is much faster at finding a parsimonious specification. We will compare these models using different decompositions with regard to their portfolio performances.

5.2   Comparison of portfolio performances

In order to compare the forecasting performance of our proposed models and other existing models, we consider the minimum-variance portfolio strategy (see Han (2006)). We denote the conditional mean and the conditional covariance matrix of the stock return $\boldsymbol{y}_{t+1}$ given the information set $\mathcal{F}_t$ at time t as

$$\boldsymbol{m}_{t+1|t} \; \equiv \; \mathrm{E}[\boldsymbol{y}_{t+1}\,|\,\mathcal{F}_t] = \boldsymbol{m}_t, \quad \boldsymbol{\Sigma}_{t+1|t} \equiv \mathrm{Var}[\boldsymbol{y}_{t+1}\,|\,\mathcal{F}_t] = \mathbf{V}_{t+1}^{1/2}\,\mathbf{R}_{t+1}\,\mathbf{V}_{t+1}^{1/2} + \boldsymbol{\Sigma}_m\,.$$

Let $r_{p,t+1}$ denote the portfolio return at time $t+1$. Further, we denote the conditional mean and conditional variance of $r_{p,t+1}$ given the information set $\mathcal{F}_t$ at time t by

$$\mu_{p,t+1} \; \equiv \; \mathrm{E}[\boldsymbol{w}_t'\,\boldsymbol{y}_{t+1} + (1 - \boldsymbol{w}_t'\,\mathbf{1}_p)r_f|\,\mathcal{F}_t] = \boldsymbol{w}_t'\,\boldsymbol{m}_{t+1|t} + (1 - \boldsymbol{w}_t'\,\mathbf{1}_p)r_f,$$

$$\sigma_{p,t+1}^2 \; \equiv \; \mathrm{Var}[\boldsymbol{w}_t'\,\boldsymbol{y}_{t+1} + (1 - \boldsymbol{w}_t'\,\mathbf{1}_p)r_f|\,\mathcal{F}_t] = \boldsymbol{w}_t'\,\boldsymbol{\Sigma}_{t+1|t}\,\boldsymbol{w}_t,$$

where r_f is the risk free asset return, and $\boldsymbol{w}_t$ is a portfolio weight vector for the stock return $\boldsymbol{y}_{t+1}$. In the minimum-variance strategy, we minimize the conditional variance $\sigma_{p,t+1}^2$ for the target level μ_p^* of the conditional expected return $\mu_{p,t+1}$. Then the optimal weight $\boldsymbol{w}_t$ is given by

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

The portfolio performances are compared based on the rolling forecast:

Step 1. First, we estimate the parameters using the first 1742 observations from February 1, 2001 to January 8, 2008 and forecast the mean, the volatility and the correlation of the multiple stock returns for January 9, 2008. We use them to obtain the optimal weights of the assets for the above portfolio strategies and the federal funds (FF) rate is used for the risk free asset return r_f.

Step 2. Next, we drop the first observation (February 1, 2001) from the sample period and add the new observation (January 9, 2008). The new sample period is from February 2, 2001 to January 9, 2008. We estimate the parameters using these observations and

forecast the mean, the volatility and the correlation for January 10, 2008. We use them to obtain the optimal weights in a similar manner.

Step 3. We iterate these rolling forecasts until December 31, 2009 to obtain the 500 one-day ahead forecasts and the corresponding weights.

To compute the optimal weight $\hat{\boldsymbol{w}}_t$, we also need the estimates of $\boldsymbol{m}_{t+1|t}$ and $\boldsymbol{\Sigma}_{t+1|t}$. We let N denote the number of MCMC iterations, and $(\theta^{(i)}, \{\boldsymbol{h}_t^{(i)}\}_{t=1}^T, \{\mathbf{R}_t^{(i)}\}_{t=1}^T, \{\boldsymbol{m}_t^{(i)}\}_{t=1}^T)$ denote the i-th MCMC sample $(i = 1, \ldots, N)$. Using $\boldsymbol{m}_{t+1|t}^{(i)}, \mathbf{V}_{t+1|t}^{(i)}, \mathbf{R}_{t+1|t}^{(i)}, \boldsymbol{\Sigma}_m^{(i)}$, we estimate $\boldsymbol{m}_{t+1|t}$ and $\boldsymbol{\Sigma}_{t+1|t}$ by

$$\hat{\boldsymbol{m}}_{t+1|t} = \frac{1}{N} \sum_{i=1}^N \boldsymbol{m}_{t+1|t}^{(i)}, \quad \hat{\boldsymbol{\Sigma}}_{t+1|t} = \frac{1}{N} \sum_{i=1}^N \boldsymbol{\Sigma}_{t+1|t}^{(i)} = \frac{1}{N} \sum_{i=1}^N \left(\mathbf{V}_{t+1|t}^{(i)1/2} \mathbf{R}_{t+1|t}^{(i)} \mathbf{V}_{t+1|t}^{(i)1/2} + \boldsymbol{\Sigma}_m^{(i)} \right).$$

In our empirical study, we set $N = 1500$ and we discard 500 samples as the burn-in period for each MCMC rolling estimation (Steps 2 and 3)[2]. We compare the following multivariate stochastic volatility models as follows.

1. MSV model: Basic multivariate stochastic volatility model without leverage, realized variances and correlations.

2. CRSV model: Cholesky realized stochastic volatility model with leverage proposed in Shirota et al. (2017)

3. MRSV model: Multivariate stochastic volatility model without leverage and with realized variances and pairwise realized correlations.

4. MRSV-L1-C model: Multivariate stochastic volatility model with leverage, realized variances and pairwise realized correlations. The parsimonious specification is assumed to model the leverage effect, $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \mathbf{0}, \ldots, \mathbf{0}]$ with $q = 1$. The Cholesky decomposition is used to compute $\mathbf{R}_t^{-1/2}$.

5. MRSV-L2-C model: Multivariate stochastic volatility model with leverage, realized variances and pairwise realized correlations. The parsimonious specification is assumed to model the leverage effect, $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \mathbf{0}, \ldots, \mathbf{0}]$ with $q = 2$. The Cholesky decomposition is used to compute $\mathbf{R}_t^{-1/2}$.

6. MRSV-L1-S model: Multivariate stochastic volatility model with leverage, realized variances and pairwise realized correlations. The parsimonious specification is assumed to model the

---

[2]The number of samples being discarded as the burn-in period is sufficient after we obtain the MCMC posterior samples from the previous sample period since we use the posterior means of the parameters and latent variables for the initial values of the next MCMC runs.

leverage effect, $\mathbf{\Lambda} = [\boldsymbol{\lambda}_1, \mathbf{0}, \dots, \mathbf{0}]$ with $q = 1$. The spectral decomposition is used to compute $\mathbf{R}_t^{-1/2}$.

7. MRSV-L1-S (constant mean) model: The mean vector $\boldsymbol{m}_t$ of the return is assumed to be constant in the MRSV-L1-S model.

8. DCC-GARCH model: DCC-GARCH model proposed in Engle (2002)[3].

9. HEAVY model: a scalar HEAVY model proposed in Noureldin et al. (2012) for each element of the spectral decomposition of the realized covariance matrix. [4].

10. HAR model: HAR model proposed in Corsi (2009)[5].

11. Equally weighted portfolio model: The weights of the assets are fixed to be equal in the model.

Table 9: The cumulative values of realized objective functions.

	$\mu_p^* = 0.004$	$\mu_p^* = 0.01$	$\mu_p^* = 0.1$
MSV	1.172	6.536	1184
CRSV	0.748	4.448	730
MRSV	0.526	2.943	510
MRSV-L1-C	0.272	1.601	262
MRSV-L2-C	0.264	1.552	255
MRSV-L1-S	**0.249**	**1.430**	**232**
MRSV-L1-S (constant mean)	0.568	3.032	543
DCC-GARCH	2.662	11.962	2537
HEAVY	2.888	14.824	3058
HAR	2.537	13.110	2680
Equal weight	1425	1425	1425

*Bold figures indicates the optimal values. The cumulative variances are computed as $\sum_{t=1742}^{2241} \hat{\boldsymbol{\omega}}_t' \mathbf{\Sigma}_{t+1} \hat{\boldsymbol{\omega}}_t$ where $\mathbf{\Sigma}_{t+1}$ is evaluated using the realized covariance at time $t + 1$. The results for CRSV model are reproduced from Shirota *et al.* (2017).

Cumulative realized objective functions. Table 9 shows the cumulative values of the realized objective functions. The MRSV-L1-S model outperforms the other models. Among the MRSV models, the models with leverage outperform the models without leverage, thereby indicating the existence and the importance of the leverage effect. If we assume a constant mean for $\boldsymbol{y}_t$, the performance becomes poor in this prediction period, which implies that the

---

[3]The parameters are estimated by the maximum likelihood method.

[4]The parameters are estimated by the two step estimation. The mean of the return is estimated by the corresponding sample mean during the sample period.

[5]The mean of the return is estimated by the corresponding sample mean during the sample period.

random walk process is more flexible for describing the dynamics of the mean of the return vector. The MRSV-L2-C model outperforms the MRSV-L1-C model, but its performance is not as good as that of the MRSV-L1-S model. We could improve the performance of the MRSV models using the Cholesky decomposition by changing the order of the assets or increasing the number of nonzero columns of $\mathbf{\Lambda}$, but it is more efficient to use the spectral decomposition to compute $\mathbf{R}_t^{-1/2}$. Finally, in comparison with the DCC-GARCH model, the HEAVY model, the HAR model and the equally weighted model, we found that the classes of the MSV and MRSV models perform much better.

Figure 3: Time series plot of the portfolio weights n MRSV-L1-S: $\mu_p^* = 0.1$.
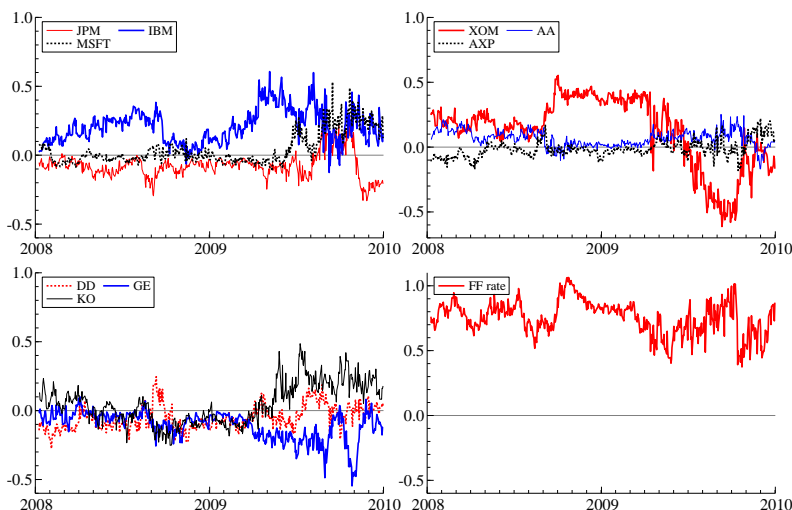
Time series plots of the portfolio weights. Figure 3 shows the time series plots of the portfolio weights in the MRSV-L1-S. The weights for Exxon Mobil are large among all stocks, but towards the end of the period, the weights for IBM and Microsoft tend to become large. However, the weights for the FF rate $(1 - \sum_{i=1}^{9} w_{it})$ are the largest throughout the forecasting period.

Comparison of performances before and after the financial crisis. In order to illustrate the portfolio performances of the models (excluding the CRSV model) in more detail before and after the financial crisis, we divide the forecasting periods into two subperiods: (1) Jan 9, 2008 - July 31, 2008 and (2) Aug 1, 2008 - Dec 31, 2009. As shown in Table 10, in both subperiods (1) and (2), the portfolio performances are similar to those that we found in the whole period.

Table 10: The cumulative values of realized objective functions in two subperiods.

(1) Jan 9, 2008 - July 31, 2008

	$\mu_p^* = 0.004$	$\mu_p^* = 0.01$	$\mu_p^* = 0.1$
MSV	0.436	0.357	253
MRSV	0.175	0.110	87
MRSV-L1-C	0.072	0.042	33
MRSV-L2-C	0.073	0.044	34
MRSV-L1-S	**0.072**	**0.041**	**33**
MRSV-L1-S (constant mean)	0.213	0.139	107
DCC-GARCH	1.265	1.074	769
HEAVY	1.099	1.250	810
HAR	0.958	1.038	683
Equal weight	176	176	176

(2) Aug 1, 2008 - Dec 31, 2009

	$\mu_p^* = 0.004$	$\mu_p^* = 0.01$	$\mu_p^* = 0.1$
MSV	0.736	6.179	932
MRSV	0.351	2.833	423
MRSV-L1-C	0.200	1.559	228
MRSV-L2-C	0.191	1.508	220
MRSV-L1-S	**0.177**	**1.389**	**199**
MRSV-L1-S (constant mean)	0.354	2.893	436
DCC-GARCH	1.397	10.889	1768
HEAVY	1.789	13.574	2248
HAR	1.579	12.072	1997
Equal weight	1248	1248	1248

Remark 4. As suggested by an anonymous referee and the Editor, we conducted a predictive ability test based on Giacomini and White (2006) in order to investigate whether the realized objective function of each model is significantly different from that of the MRSV-L1-S in Tables 9 and 10. We found that the differences are all significant except for the MRSV-L1-C models ($\mu_p^* = 0.004, 0.01, 0.1$) and MRSV-L2-C models ($\mu_p^* = 0.004, 0.1$) in the subperiod (1) (which is before the financial crisis).

6   Conclusions

The multivariate SV model with flexible dynamic correlation structures that uses the Markov chain Monte Carlo estimation method is proposed. By making full use of the realized vari-

ances and realized pairwise correlations, we obtain stable parameter estimates where the covariance matrices are guaranteed to be positive definite. The spectral decomposition is used for the correlation matrices in order to avoid the arbitrariness of the ordering of asset returns. The parsimonious specification for the leverage effect is also proposed. Our models are applied to the daily returns of nine U.S. stocks with their realized volatilities and pairwise realized correlations and are shown to outperform the existing models with regard to portfolio optimizations under a minimum-variance strategy.

Acknowledgements

Appendix

A MCMC algorithm

A.1 MRSV model without the leverage effect

A.1.1 Joint posterior density

The joint posterior probability density function is given by

$$
\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{g}, \boldsymbol{h}, \boldsymbol{m} \,|\, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}) \\
&\propto \prod_{t=1}^{T} |\mathbf{V}_t^{1/2}\mathbf{R}_t\mathbf{V}_t^{1/2}|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'(\mathbf{V}_t^{1/2}\mathbf{R}_t\mathbf{V}_t^{1/2})^{-1}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\} \\
&\quad\times |\,\boldsymbol{\Omega}_0\,|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Omega}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right\} \\
&\quad\times \prod_{t=1}^{T-1} |\boldsymbol{\Omega}|^{-1/2} \exp\left[-\frac{1}{2}\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\mu} - \boldsymbol{\Phi}\,\boldsymbol{h}_t\}'\boldsymbol{\Omega}^{-1}\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\mu} - \boldsymbol{\Phi}\,\boldsymbol{h}_t\}\right] \\
&\quad\times \prod_{t=1}^{T} |\,\boldsymbol{\Sigma}_u\,|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)'\boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)\right\} \\
&\quad\times \prod_{i>j}^{p} \sigma_{\zeta,ij}^{-1} \exp\left(-\frac{g_{ij,1}^2}{2\kappa\sigma_{\zeta,ij}^2}\right) \prod_{t=1}^{T-1} \sigma_{\zeta,ij}^{-1} \exp\left\{-\frac{(g_{ij,t+1} - g_{ij,t})^2}{2\sigma_{\zeta,ij}^2}\right\} \\
&\quad\times \prod_{i>j}^{p}\prod_{t=1}^{T} \sigma_{v,ij}^{-1} \exp\left\{-\frac{(w_{ij,t} - \delta_{ij} - g_{ij,t})^2}{2\sigma_{v,ij}^2}\right\} \\
&\quad\times \prod_{i=1}^{p} \sigma_{m,i}^{-1} \exp\left(-\frac{m_{i,1}^2}{2\kappa\sigma_{m,i}^2}\right) \prod_{t=1}^{T-1} \sigma_{m,i}^{-1} \exp\left\{-\frac{(m_{i,t+1} - m_{i,t})^2}{2\sigma_{m,i}^2}\right\} \times \pi(\boldsymbol{\theta}),
\end{aligned} \tag{27}
$$

where $\Sigma_u = \mathrm{diag}(\sigma_{u,1}^2, \ldots, \sigma_{u,p}^2)$ and $\pi(\boldsymbol{\theta})$ is a prior probability density function of parameters.

A.1.2 Generation of ϕ

It can be shown that the conditional posterior probability density function of ϕ is

$$\pi(\boldsymbol{\phi}|\cdot) \propto k(\boldsymbol{\phi}) \times \exp\left(-\frac{1}{2}(\boldsymbol{\phi} - \boldsymbol{\mu_\phi})'\boldsymbol{\Sigma_\phi}^{-1}(\boldsymbol{\phi} - \boldsymbol{\mu_\phi})\right) \times I\{|\phi_i| < 1, i = 1, \ldots, p\}, \tag{28}$$

where $I(B)$ is an indicator function such that $I(B) = 1$ if B is true and 0 otherwise,

$$k(\boldsymbol{\phi}) = |\boldsymbol{\Omega}_0|^{-1/2} \prod_{i=1}^{p} \left(\frac{1+\phi_i}{2}\right)^{a-1} \left(\frac{1-\phi_i}{2}\right)^{b-1} \exp\left(-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Omega}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right), \tag{29}$$

$$\boldsymbol{\mu_\phi} = \boldsymbol{\Sigma_\phi}\boldsymbol{b}, \quad \boldsymbol{\Sigma_\phi}^{-1} = \boldsymbol{\Omega}^{-1} \odot \mathbf{A}, \tag{30}$$

$$\mathbf{A} = \sum_{t=1}^{T-1}(\boldsymbol{h}_t - \boldsymbol{\mu})(\boldsymbol{h}_t - \boldsymbol{\mu})', \quad \boldsymbol{b} = \mathrm{diagonal}\left\{\sum_{t=1}^{T-1}(\boldsymbol{h}_t - \boldsymbol{\mu})(\boldsymbol{h}_{t+1} - \boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}\right\}, \tag{31}$$

$\odot$ is Hadamard product, and diagonal($\mathbf{B}$) denotes a column vector with diagonal elements of $\mathbf{B}$. We propose a candidate $\boldsymbol{\phi}^\dagger \sim \mathrm{TN}_R(\boldsymbol{\mu_\phi}, \boldsymbol{\Sigma_\phi})$, where $R = \{\boldsymbol{\phi} : |\phi_i| < 1, i = 1, \ldots, p\}$, and accept it with probability $\min\{1, k(\boldsymbol{\phi}^\dagger)/k(\boldsymbol{\phi})\}$.

A.1.3 Generation of $\boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\delta}$

The $\boldsymbol{\mu}, \boldsymbol{\xi}$ and $\boldsymbol{\delta}$ are conditionally independent and we generate them from the following normal distributions:

$$\boldsymbol{\mu}|\cdot \sim N(\tilde{\boldsymbol{m}}_{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\mu}}), \quad \boldsymbol{\xi}|\cdot \sim N(\tilde{\boldsymbol{m}}_{\boldsymbol{\xi}}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}), \quad \boldsymbol{\delta}|\cdot \sim N(\tilde{\boldsymbol{m}}_{\boldsymbol{\delta}}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}}), \tag{32}$$

where

$$\tilde{\boldsymbol{m}}_{\boldsymbol{\mu}} = \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\mu}}\left[s_\mu^{-2}\boldsymbol{m}_\mu + \boldsymbol{\Omega}_0^{-1}\boldsymbol{h}_1 + (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\Omega}^{-1}\sum_{t=1}^{T-1}(\boldsymbol{h}_{t+1} - \boldsymbol{\Phi}\,\boldsymbol{h}_t)\right], \tag{33}$$

$$\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\mu}} = \left[s_\mu^{-2}\mathbf{I}_p + \boldsymbol{\Omega}_0^{-1} + (T-1)(\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\Omega}^{-1}(\mathbf{I}_p - \boldsymbol{\Phi})\right]^{-1}, \tag{34}$$

$$\tilde{\boldsymbol{m}}_{\boldsymbol{\xi}} = \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\left[s_\xi^{-2}\boldsymbol{m}_\xi + \boldsymbol{\Sigma}_u^{-1}\sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{h}_t)\right], \quad \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}} = (s_\xi^{-2}\mathbf{I}_p + T\,\boldsymbol{\Sigma}_u^{-1})^{-1}, \tag{35}$$

$$\tilde{\boldsymbol{m}}_{\boldsymbol{\delta}} = \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}}\left[s_\delta^{-2}\boldsymbol{m}_\delta + \boldsymbol{\Sigma}_v^{-1}\sum_{t=1}^{T}(\boldsymbol{w}_t - \boldsymbol{g}_t)\right], \quad \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}} = (s_\delta^{-2}\mathbf{I}_p + T\,\boldsymbol{\Sigma}_v^{-1})^{-1}. \tag{36}$$

A.1.4 Generation of $(\boldsymbol{\sigma}_u^2, \boldsymbol{\sigma}_v^2, \boldsymbol{\sigma}_\zeta^2, \boldsymbol{\Sigma}_m)$

The $\sigma_{u,i}^2, \sigma_{v,ij}^2, \sigma_{\zeta,ij}^2, \sigma_{m,i}^2$ are conditionally independent and we generate them from inverse gamma distributions:

$$\sigma_{u,i}^2 \sim \mathrm{IG}\left(\frac{\tilde{n}_{ui}}{2}, \frac{\tilde{d}_{ui}}{2}\right), \quad \sigma_{v,ij}^2 \sim \mathrm{IG}\left(\frac{\tilde{n}_{v,ij}}{2}, \frac{\tilde{d}_{v,ij}}{2}\right), \quad \sigma_{\zeta,ij}^2 \sim \mathrm{IG}\left(\frac{\tilde{n}_{\zeta,ij}}{2}, \frac{\tilde{d}_{\zeta,ij}}{2}\right), \quad \sigma_{m,i}^2 \sim \mathrm{IG}\left(\frac{\tilde{n}_{mi}}{2}, \frac{\tilde{d}_{mi}}{2}\right),$$

where

$$\tilde{n}_{ui} = n_u + T, \quad \tilde{d}_{ui} = d_u + \sum_{t=1}^{T}(x_{it} - \xi_i - h_{it})^2, \tag{37}$$

$$\tilde{n}_{v,ij} = n_v + T, \quad \tilde{d}_{v,ij} = d_v + \sum_{t=1}^{T}(w_{ij,t} - \delta_{ij} - g_{ij,t})^2, \tag{38}$$

$$\tilde{n}_{\zeta,ij} = n_\zeta + T, \quad \tilde{d}_{\zeta,ij} = d_\zeta + \kappa^{-1} g_{ij,1}^2 + \sum_{t=1}^{T-1}(g_{ij,t+1} - g_{ij,t})^2, \tag{39}$$

$$\tilde{n}_{mi} = n_m + T, \quad \tilde{d}_{mi} = d_m + \kappa^{-1} m_{i1}^2 + \sum_{t=1}^{T-1}(m_{i,t+1} - m_{it})^2, \tag{40}$$

for $i, j = 1, \ldots, p \ (i > j)$.

A.1.5 Generation of $\boldsymbol{\Omega}$

The conditional posterior probability density function of $\boldsymbol{\Omega}$ is

$$\pi(\boldsymbol{\Omega}|\boldsymbol{h}, \boldsymbol{\mu}, \boldsymbol{\phi}) \propto m(\boldsymbol{\Omega}) \times |\boldsymbol{\Omega}|^{-(\tilde{\nu}+p+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Omega}^{-1}\tilde{\mathbf{S}}\right)\right\}, \tag{41}$$

where

$$m(\boldsymbol{\Omega}) = |\boldsymbol{\Omega}_0|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Omega}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right), \quad \tilde{\nu} = \nu + T - 1, \tag{42}$$

$$\tilde{\boldsymbol{S}} = \boldsymbol{S} + \sum_{t=1}^{T-1}\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\mu} - \boldsymbol{\Phi}\,\boldsymbol{h}_t\}\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\mu} - \boldsymbol{\Phi}\,\boldsymbol{h}_t\}'. \tag{43}$$

We propose a candidate $\boldsymbol{\Omega}^{\dagger} \sim \text{IW}(\tilde{\nu}, \tilde{\boldsymbol{S}})$, and accept it with probability $\min\{1, m(\boldsymbol{\Omega}^{\dagger})/m(\boldsymbol{\Omega})\}$.

A.1.6 Generation of $\boldsymbol{\Psi}$

As a prior distribution of $\boldsymbol{\Psi}$, we assume $\boldsymbol{\Psi} \sim \text{IW}(\nu_\psi, \mathbf{S}_\psi)$. Then the conditional posterior probability density function of $\boldsymbol{\Psi}$ is

$$\pi(\boldsymbol{\Psi}|\cdot) \propto m(\boldsymbol{\Psi}) \times |\boldsymbol{\Psi}|^{-(\tilde{\nu}_\psi+p+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}^{-1}\tilde{\mathbf{S}}_\psi\right)\right\}, \tag{44}$$

where

$$m(\boldsymbol{\Psi}) = |\boldsymbol{\Omega}_0|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Omega}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right), \tag{45}$$

$$\tilde{\nu}_\psi = \nu_\psi + p + T - 1, \tag{46}$$

$$\begin{aligned}
\tilde{\mathbf{S}}_\psi = {} & \mathbf{S}_\psi + (\boldsymbol{\Lambda} - \mathbf{M}_0)\boldsymbol{\Gamma}_0^{-1}(\boldsymbol{\Lambda} - \mathbf{M}_0)' \\
& + \sum_{t=1}^{T-1}\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\} \\
& \qquad\qquad \left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}'.
\end{aligned} \tag{47}$$

Thus we propose a candidate $\boldsymbol{\Psi}^{\dagger}$ from $\text{IW}(\tilde{\nu}_\psi, \tilde{\mathbf{S}}_\psi)$, and accept it with probability $\min\{1, m(\boldsymbol{\Psi}^{\dagger})/m(\boldsymbol{\Psi})\}$.

A.1.7   Generation of $\boldsymbol{\mu}$

Noting that

$$\mathrm{E}[\boldsymbol{y}_t \,|\, \boldsymbol{h}_t] = \boldsymbol{m}_t, \quad \mathrm{Var}[\boldsymbol{y}_t \,|\, \boldsymbol{h}_t] = \mathbf{V}_t^{1/2}\,\mathbf{R}_t\,\mathbf{V}_t^{1/2} \equiv \boldsymbol{\Gamma}_t, \tag{48}$$

it can be shown that the conditional posterior distribution of $\boldsymbol{m}_t$ is the same as that of the following linear Gaussian state space model:

$$\boldsymbol{y}_t \;=\; \boldsymbol{m}_t + \hat{\boldsymbol{\epsilon}}_t, \quad \hat{\boldsymbol{\epsilon}}_t \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_t), \tag{49}$$

$$\boldsymbol{m}_{t+1} \;=\; \boldsymbol{m}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_m), \tag{50}$$

where $\hat{\boldsymbol{\epsilon}}_t$ and $\boldsymbol{\nu}_t$ are independent. Thus we generate $\boldsymbol{m}$ simultaneously at one time using a simulation smoother (e.g. de Jong and Shephard (1995), Durbin and Koopman (2002)).

A.2   MRSV model with the leverage effect

We need to modify the sampling procedures of $\boldsymbol{g}, \boldsymbol{h}, \boldsymbol{m}, \boldsymbol{\phi}$ and $\boldsymbol{\mu}$ for the model with the leverage effect. Generations of other parameters are the same as in the previous section.

A.2.1   Generation of $\boldsymbol{g}_t$

We only need to modify (14) in Section 3.2.1 as follows.

$$r(g_{ij,t}) = -\frac{1}{2}\log|\mathbf{R}_t| - \frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'(\mathbf{V}_t^{1/2}\,\mathbf{R}_t\,\mathbf{V}_t^{1/2})^{-1}(\boldsymbol{y}_t - \boldsymbol{m}_t)$$
$$- \frac{1}{2}\,\boldsymbol{y}_t'\,\mathbf{V}^{-1/2}\,\mathbf{R}_t^{-1/2'}\,\boldsymbol{\Lambda}'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t) + \boldsymbol{y}_t'\,\mathbf{V}_t^{-1/2}\,\mathbf{R}_t^{-1/2'}\,\boldsymbol{\Lambda}'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\eta}_t, \tag{51}$$

for $t = 1, \dots, T-1$ where $\boldsymbol{\eta}_t = \boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu})$.

A.2.2   Generation of $\boldsymbol{h}_t$

The conditional posterior probability density function of $\boldsymbol{h}_t$ is given by

$$\pi(\boldsymbol{h}_t \,|\, \cdot) \propto \exp\left[-\frac{1}{2}(\boldsymbol{h}_t - \boldsymbol{m}_{t*})'\,\boldsymbol{\Omega}_{t*}^{-1}(\boldsymbol{h}_t - \boldsymbol{m}_{t*}) + l(\boldsymbol{h}_t)\right], \tag{52}$$

where

$$
l(\boldsymbol{h}_t)
$$
$$
= \begin{cases}
-\frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'\,\mathbf{V}^{-1/2}\,\mathbf{R}_t^{-1}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t) - \frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{m}_t)'\,\mathbf{V}_t^{-1/2}\,\mathbf{R}_t^{-1/2'}\,\boldsymbol{\Lambda}'\,\boldsymbol{\Psi}^{-1} \\
\quad \times \left\{\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t) - 2(\boldsymbol{h}_{t+1} - (\mathbf{I} - \boldsymbol{\Phi})\,\boldsymbol{\mu} - \boldsymbol{\Phi}\,\boldsymbol{h}_t)\right\}, \\
\hspace{9cm} t = 1, \dots, T-1, \\
-\frac{1}{2}(\boldsymbol{y}_T - \boldsymbol{m}_T)'\,\mathbf{V}_T^{-1/2}\,\mathbf{R}_T^{-1}\,\mathbf{V}_T^{-1/2}(\boldsymbol{y}_T - \boldsymbol{m}_T), \hspace{2.5cm} t = T,
\end{cases}
\tag{53}
$$

and

$$
\boldsymbol{m}_{t*} = \begin{cases}
\boldsymbol{\Psi}_{1*}\left[\boldsymbol{\Omega}_0^{-1}\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{\Psi}^{-1}\left\{\boldsymbol{h}_2 - (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu}\right\} + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_1 - \boldsymbol{\xi}) - \tfrac{1}{2}\,\mathbf{1}_p\right], & t = 1, \\[2ex]
\boldsymbol{\Psi}_{t*}\Bigg[\boldsymbol{\Psi}^{-1}\left\{\boldsymbol{\Lambda}\,\mathbf{R}_{t-1}^{-1/2}\,\mathbf{V}_{t-1}^{-1/2}(\boldsymbol{y}_{t-1} - \boldsymbol{m}_{t-1}) + (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{h}_{t-1}\right\} \\[1ex]
\qquad\quad + \boldsymbol{\Phi}\,\boldsymbol{\Psi}^{-1}\left\{\boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu}\right\} + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi}) - \tfrac{1}{2}\,\mathbf{1}_p\Bigg], \\[1ex]
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t = 2,\ldots,T-1, \\[2ex]
\boldsymbol{\Psi}_{T*}\Bigg[\boldsymbol{\Psi}^{-1}\left\{\boldsymbol{\Lambda}\,\mathbf{R}_{T-1}^{-1/2}\,\mathbf{V}_{T-1}^{-1/2}(\boldsymbol{y}_{T-1} - \boldsymbol{m}_{T-1}) + (\mathbf{I}_p - \boldsymbol{\Phi})\,\boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{h}_{T-1}\right\} \\[1ex]
\qquad\quad + \boldsymbol{\Sigma}_u^{-1}(\boldsymbol{x}_T - \boldsymbol{\xi}) - \tfrac{1}{2}\,\mathbf{1}_p\Bigg], & t = T,
\end{cases} \tag{54}
$$

$$
\boldsymbol{\Psi}_{t*} = \begin{cases}
\left[\boldsymbol{\Omega}_0^{-1} + \boldsymbol{\Phi}\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Phi} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = 1, \\[1ex]
\left[\boldsymbol{\Psi}^{-1} + \boldsymbol{\Phi}\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Phi} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = 2,\ldots,T-1, \\[1ex]
\left[\boldsymbol{\Psi}^{-1} + \boldsymbol{\Sigma}_u^{-1}\right]^{-1}, & t = T.
\end{cases} \tag{55}
$$

We generate a candidate $\boldsymbol{h}_t^\dagger$ from $\mathrm{N}(\boldsymbol{m}_{t*}, \boldsymbol{\Omega}_{t*})$, and accept it with probability $\min\{1, \exp(l(\boldsymbol{h}_t^\dagger) - l(\boldsymbol{h}_t))\}$.

A.2.3   Generation of $\boldsymbol{m}$

Noting that

$$
E[\boldsymbol{y}_t \,|\, \boldsymbol{h}_t, \boldsymbol{h}_{t+1}, \boldsymbol{\theta}] = \boldsymbol{m}_t + \mathbf{V}_t^{1/2}\mathbf{R}_t^{-1/2}\boldsymbol{\Lambda}'(\boldsymbol{\Psi} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}')^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu})\}, \tag{56}
$$

$$
\mathrm{Var}[\boldsymbol{y}_t \,|\, \boldsymbol{h}_t, \boldsymbol{h}_{t+1}, \boldsymbol{\theta}] = \mathbf{V}_t^{1/2}\,\mathbf{R}_t\,\mathbf{V}_t^{-1/2} - \mathbf{V}_t^{1/2}\mathbf{R}_t^{-1/2}\boldsymbol{\Lambda}'(\boldsymbol{\Psi} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}')^{-1}\boldsymbol{\Lambda}\mathbf{R}_t^{-1/2\prime}\mathbf{V}_t^{1/2} \equiv \boldsymbol{\Gamma}_t, \tag{57}
$$

we define

$$
\hat{\boldsymbol{y}}_t = \boldsymbol{y}_t - \mathbf{V}_t^{1/2}\mathbf{R}_t^{-1/2\prime}\boldsymbol{\Lambda}'(\boldsymbol{\Psi} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}')^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu})\}, \tag{58}
$$

and consider the linear Gaussian state space model (49) and (50) with $\boldsymbol{\Gamma}_t$ in (57). We generate $\boldsymbol{m}$ simultaneously using a simulation smoother.

A.2.4   Generation of ϕ

In Section A.1.2, we replace $\boldsymbol{\mu}_\phi$ and $\boldsymbol{\Sigma}_\phi$ as follows.

$$
\boldsymbol{\mu}_\phi = \boldsymbol{\Sigma}_\phi \boldsymbol{b}, \quad \boldsymbol{\Sigma}_\phi^{-1} = \boldsymbol{\Psi}^{-1} \odot \mathbf{A},
$$

where

$$
\mathbf{A} = \sum_{t=1}^{T-1}(\boldsymbol{h}_t - \boldsymbol{\mu})(\boldsymbol{h}_t - \boldsymbol{\mu})', \quad \boldsymbol{b} = \mathrm{diagonal}\left[\sum_{t=1}^{T-1}(\boldsymbol{h}_t - \boldsymbol{\mu})\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}'\boldsymbol{\Psi}^{-1}\right].
$$

A.2.5 Generation of $\boldsymbol{\Psi}$

As the prior distribution of $\boldsymbol{\Lambda}$ is changed, the conditional posterior probability density function of $\boldsymbol{\Psi}$ is now replaced by

$$\pi(\boldsymbol{\Psi}|\cdot) \propto m(\boldsymbol{\Psi}) \times |\boldsymbol{\Psi}|^{-(\tilde{\nu}_\psi + p + 1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}^{-1}\tilde{\mathbf{S}}_\psi\right)\right\}, \tag{59}$$

where

$$m(\boldsymbol{\Psi}) = |\boldsymbol{\Omega}_0|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Omega}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right), \tag{60}$$

$$\tilde{\nu}_\psi = \nu_\psi + T - 1, \tag{61}$$

$$\tilde{\mathbf{S}}_\psi = \mathbf{S}_\psi + \sum_{t=1}^{T-1}\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}$$
$$\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}'. \tag{62}$$

Thus we propose a candidate $\boldsymbol{\Psi}^\dagger$ from $\mathrm{IW}(\tilde{\nu}_\psi, \tilde{\mathbf{S}}_\psi)$, and accept it with probability $\min\{1, m(\boldsymbol{\Psi}^\dagger)/m(\boldsymbol{\Psi})\}$.

A.2.6 Generation of $\boldsymbol{\mu}$

We generate $\boldsymbol{\mu}|\cdot \sim N(\tilde{\boldsymbol{m}}_\mu, \tilde{\boldsymbol{\Psi}}_\mu)$, where

$$\tilde{\boldsymbol{m}}_\mu = \tilde{\boldsymbol{\Psi}}_\mu\left[s_\mu^{-2}\boldsymbol{m}_\mu + \boldsymbol{\Omega}_0^{-1}\boldsymbol{h}_1 + (\mathbf{I} - \boldsymbol{\Phi})\boldsymbol{\Psi}^{-1}\sum_{t=1}^{T-1}\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\Phi}\,\boldsymbol{h}_t - \boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}\right]$$

$$\tilde{\boldsymbol{\Psi}}_\mu = \left[s_\mu^{-2}\mathbf{I} + \boldsymbol{\Omega}_0^{-1} + (T-1)(\mathbf{I} - \boldsymbol{\Phi})\boldsymbol{\Psi}^{-1}(\mathbf{I} - \boldsymbol{\Phi})\right]^{-1}.$$

B Proofs

B.1 Proof of Proposition 1

Proof: Since $\mathbf{R}_{it}$ is positive definite, its principal submatrices are all positive definite. Further, noting that $|\mathbf{R}_t| = |\mathbf{R}_{it}| \times |1 - \boldsymbol{\rho}_{it}'\,\mathbf{R}_{it}^{-1}\,\boldsymbol{\rho}_{it}|$, the condition for $\mathbf{R}_t$ to be positive definite is $-\boldsymbol{\rho}_{it}'\mathbf{R}_{it}^{-1}\boldsymbol{\rho}_{it} > 0$, which reduces to

$$-a_j\rho_{ij,t}^2 - 2\boldsymbol{b}_j'\,\boldsymbol{\rho}_{i,-j,t}\,\rho_{ij,t} - \boldsymbol{\rho}_{i,-j,t}'\,\mathbf{C}_j\,\boldsymbol{\rho}_{i,-j,t} + 1 > 0, \tag{63}$$

Therefore the inequality (63) implies that the lower and upper bounds for $\rho_{ij,t}$ are given by (17).

$\square$

B.2 Proof of Proposition 2

We note that the probability density function for $\mathbf{X} \sim \mathrm{N}_{p,n}(\mathbf{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$ is given by

$$f(\mathbf{X}) = (2\pi)^{-np/2}|\boldsymbol{\Psi}|^{-n/2}|\boldsymbol{\Sigma}|^{-p/2} \times \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}^{-1}(\mathbf{X} - \mathbf{M})\,\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \mathbf{M})'\right)\right\}. \tag{64}$$

Proof: Since

$$
\begin{aligned}
f(\boldsymbol{h}_{t+1}|\boldsymbol{y}_t,\boldsymbol{g}_t,\boldsymbol{h}_t,\boldsymbol{m}_t,\boldsymbol{\theta}) & \\
\propto \quad |\boldsymbol{\Psi}|^{-1/2}\exp\Big\{ & -\frac{1}{2}\Big(\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\boldsymbol{\Phi}(\boldsymbol{h}_t-\boldsymbol{\mu})-\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t-\boldsymbol{m}_t)\Big)' \\
& \boldsymbol{\Psi}^{-1}\Big(\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\boldsymbol{\Phi}(\boldsymbol{h}_t-\boldsymbol{\mu})-\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t-\boldsymbol{m}_t)\Big)\Big\},
\end{aligned}
\tag{65}
$$

and

$$
\pi(\boldsymbol{\Lambda}|\boldsymbol{\Psi}) \quad \propto \quad |\boldsymbol{\Psi}|^{-p/2}|\boldsymbol{\Gamma}_0|^{-p/2}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}^{-1}(\boldsymbol{\Lambda}-\mathbf{M}_0)\boldsymbol{\Gamma}_0^{-1}(\boldsymbol{\Lambda}-\mathbf{M}_0)'\right)\right\},
\tag{66}
$$

the conditional posterior probability density function of $\boldsymbol{\Lambda}$ is

$$
\begin{aligned}
\pi(\boldsymbol{\Lambda}\,|\cdot) \propto \exp\Bigg[-\frac{1}{2}\Bigg\{ & \sum_{t=1}^{T-1}(\boldsymbol{y}_t-\boldsymbol{m}_t)'\,\mathbf{V}_t^{-1/2}\,\mathbf{R}_t^{-1/2'}\,\boldsymbol{\Lambda}'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t-\boldsymbol{m}_t) \\
& -2\sum_{t=1}^{T-1}(\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\boldsymbol{\Phi}(\boldsymbol{h}_t-\boldsymbol{\mu}))'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Lambda}\,\mathbf{R}_t^{-1/2}\,\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t-\boldsymbol{m}_t)\Bigg\}\Bigg]\times\pi(\boldsymbol{\Lambda}\,|\boldsymbol{\Psi}) \\
\propto \exp\Bigg[-\frac{1}{2} & \left\{\operatorname{tr}(\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Lambda}(\mathbf{A}+\boldsymbol{\Gamma}_0^{-1})\,\boldsymbol{\Lambda}')-2\operatorname{tr}(\boldsymbol{\Psi}^{-1}\,\boldsymbol{\Lambda}(\mathbf{B}+\boldsymbol{\Gamma}_0^{-1}\,\mathbf{M}_0))\right\}\Bigg] \\
\propto \exp\Bigg[-\frac{1}{2} & \operatorname{tr}\left\{\boldsymbol{\Psi}^{-1}(\boldsymbol{\Lambda}-\mathbf{M}_1)(\mathbf{A}+\boldsymbol{\Gamma}_0^{-1})(\boldsymbol{\Lambda}-\mathbf{M}_1)'\right\}\Bigg],
\end{aligned}
\tag{67}
$$

and the result follows.

$\square$

B.3 Proof of Proposition 3

Proof: Using $\boldsymbol{\Lambda}=\sum_{j=1}^{q}\boldsymbol{e}_j'\otimes\boldsymbol{\lambda}_j$ where $\boldsymbol{e}_j$ is the $p\times 1$ unit vector with the j-th element being one, the posterior probability density function of $\boldsymbol{\Lambda}$ is

$$
\begin{aligned}
\pi(\boldsymbol{\lambda}|\cdot) & \\
\propto \exp\Bigg[-\frac{1}{2}\Bigg\{ & \sum_{t=1}^{T-1}\boldsymbol{z}_t'\left(\sum_{j=1}^{q}\boldsymbol{e}_j\otimes\boldsymbol{\lambda}_j'\right)\boldsymbol{\Psi}^{-1}\left(\sum_{j=1}^{q}\boldsymbol{e}_j'\otimes\boldsymbol{\lambda}_j\right)\boldsymbol{z}_t-2\sum_{t=1}^{T-1}\boldsymbol{\eta}_t'\,\boldsymbol{\Psi}^{-1}\left(\sum_{j=1}^{q}\boldsymbol{e}_j'\otimes\boldsymbol{\lambda}_j\right)\boldsymbol{z}_t\Bigg\}\Bigg] \\
& \times\pi(\boldsymbol{\lambda}) \\
\propto \exp\Bigg[-\frac{1}{2}\Bigg\{ & \sum_{t=1}^{T-1}\left(\sum_{j=1}^{q}\boldsymbol{z}_t'\boldsymbol{e}_j\otimes\boldsymbol{\lambda}_j'\right)\boldsymbol{\Psi}^{-1}\left(\sum_{j=1}^{q}\boldsymbol{e}_j'\boldsymbol{z}_t\otimes\boldsymbol{\lambda}_j\right)-2\sum_{t=1}^{T-1}\boldsymbol{\eta}_t'\,\boldsymbol{\Psi}^{-1}\left(\sum_{j=1}^{q}\boldsymbol{e}_j'\boldsymbol{z}_t\otimes\boldsymbol{\lambda}_j\right)\Bigg\}\Bigg] \\
& \times\pi(\boldsymbol{\lambda}) \\
\propto \exp\Bigg[-\frac{1}{2}\Bigg\{ & \sum_{t=1}^{T-1}\sum_{i=1}^{q}\sum_{j=1}^{q}(\boldsymbol{e}_i'\boldsymbol{z}_t\,\boldsymbol{z}_t'\,\boldsymbol{e}_j)\,\boldsymbol{\lambda}_i'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\lambda}_j-2\sum_{t=1}^{T-1}\sum_{j=1}^{q}\boldsymbol{e}_j'\boldsymbol{z}_t\,\boldsymbol{\eta}_t'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\lambda}_j\Bigg\}\Bigg]\times\pi(\boldsymbol{\lambda})
\end{aligned}
$$

$$\propto \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{q}\sum_{j=1}^{q}(\boldsymbol{e}_i'\,\mathbf{A}\,\boldsymbol{e}_j)\,\boldsymbol{\lambda}_i'\,\boldsymbol{\Psi}^{-1}\,\boldsymbol{\lambda}_j -2\sum_{j=1}^{q}\boldsymbol{\lambda}_j'\,\boldsymbol{\Psi}^{-1}\,\mathbf{B}'\,\boldsymbol{e}_j\right\}\right]\times\pi(\boldsymbol{\lambda})$$

$$\propto \exp\left[-\frac{1}{2}\left\{\boldsymbol{\lambda}'\left(\boldsymbol{\Gamma}_0^{-1}+\ \mathbf{A}_{1:q,1:q}\otimes\boldsymbol{\Psi}^{-1}\right)\boldsymbol{\lambda} -2\,\boldsymbol{\lambda}'\left(\boldsymbol{\Gamma}_0^{-1}\boldsymbol{m}_0+(\mathbf{I}_q\otimes\boldsymbol{\Psi}^{-1}\,\mathbf{B}')\mathrm{vec}\left(\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_q\}\right)\right)\right\}\right],$$

(68)

and the result follows.

$\square$

References

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

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

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

Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics 134*(2), 341–371.

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

de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika 82*, 339–350.

Dobrev, D. P. and P. J. Szerszen (2010). The information content of high-frequency data for estimating equity return models and forecasting risk. International Finance Discussion Papers, Board of Governors of the Federal Reserve System (U.S.).

Doornik, J. (2006). *Ox: Object Oriented Matrix Programming*. London: Timberlake Consultants Press.

Durbin, J. and S. J. Koopman (2002). Simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20-3*, 339–350.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *74*, 1545–1578.

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 an realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

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

Jin, X. and J. M. Maheu (2013). Modeling realized covariances and returns. *Journal of Financial Econometrics 11*(2), 335–369.

Jin, X. and J. M. Maheu (2016). Bayesian semiparametric modeling of realized covariance matrices. *Journal Econometrics 192*(1), 19–31.

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.

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.

Lopes, H. F. and M. West (2004). Bayesian model assessment in factor analysis. *Statistica Sinica 14*(1), 41–68.

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

Pitt, M. and N. Shephard (1999). Time varying covariances: a factor stochastic volatility approach. *Bayesian statistics 6*, 547–570.

Shirota, S., Y. Omori, H. F. Lopes, and H. Piao (2017). Cholesky realized stochastic volatility model. *Econometrics and Statistics 3*, 34–59.

So, M. K. P., R. W. M. Li, M. Asai, and Y. Jiang (2016). Stochastic multivariate mixture covariance model. *Journal of Forecasting*. in press.

Takahashi, M., Y. Omori, and T. Watanabe (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and 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.

Windle, J., C. M. Carvalho, et al. (2014). A tractable state-space model for symmetric positive-definite matrices. *Bayesian Analysis 9*(4), 759–792.

Zheng, T. and T. Song (2014). A realized stochastic volatility model with Box-Cox transformation. *Journal of Business and Economic Statistics 32*(4), 593–605.