CIRJE-F-1075

Multiple-lock Dynamic Equicorrelations with Realized Measures, Leverage and Endogeneity

Yuta Kurose
Osaka University

Yasuhiro Omori
The University of Tokyo

January 2018

Multiple-block dynamic equicorrelations with realized measures, leverage and endogeneity

Yuta Kurose* and Yasuhiro Omori†

January, 2018

Abstract

The single equicorrelation structure among several daily asset returns is promising and attractive to reduce the number of parameters in multivariate stochastic volatility models. However, such an assumption may not be realistic as the number of assets may increase, for example, in the portfolio optimizations. As a solution to this oversimplification, the multiple-block equicorrelation structure is proposed for high dimensional financial time series, where we assume common correlations within a group of asset returns, but allow different correlations for different groups. The realized volatilities and realized correlations are also jointly modelled to obtain stable and accurate estimates of parameters, latent variables and leverage effects. Using a state space representation, we describe an efficient estimation method of Markov chain Monte Carlo simulation. Illustrative examples are given using simulated data, and empirical studies using U.S. daily stock returns data show that our proposed model outperforms other competing models in portfolio performances.

Key words: Asymmetry, leverage effect, Markov chain Monte Carlo, multiple-block equicorrelation, multivariate stochastic volatility, realized correlation, realized volatility.

*Center for Mathematical Modeling and Data Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka-shi, Osaka 560-8531, JAPAN.

†Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN.

1 Introduction

Modeling multivariate volatilities has been one of the most attractive and challenging research topics in financial econometrics. The stylized facts such as volatility clustering, dynamic correlations and leverage effects in the financial time series have been considered in various multivariate extensions of the univariate time-varying variance models for generalized autoregressive conditional heteroskedasticity (GARCH) models (e.g. Bauwens, Laurent, and Rombouts (2006)) and stochastic volatility (SV) models (e.g. Asai, McAleer, and Yu (2006), Chib, Omori, and Asai (2009)).

This paper focuses on the multivariate stochastic volatility (MSV) model with leverage effect (e.g. Daníelsson (1998), Asai and McAleer (2006), Asai and McAleer (2009), Chan, Kohn, and Kirby (2006), Ishihara and Omori (2012), Ishihara, Omori, and Asai (2016), Nakajima (2015) and Trojan (2014)) in line with the dynamic equicorrelation stochastic volatility (DESV) model proposed by Kurose and Omori (2016), where the correlations between asset returns are assumed to be time-varying and common for all pairs of asset returns. Although such an equicorrelation assumption is simple and useful, it may be too strong and counter-intuitive, especially when the number of dependent variables is very large. As a solution to this oversimplification, a multiple-block DESV model is proposed where we assume common correlations within a group of asset returns, but allow different correlations for different groups. Furthermore, the equicorrelation structure is assumed between variables in the i-th and the j-th groups. That is, we assume the common correlation between one asset return in the i-th group and another asset return in the j-th group. As discussed in Elton and Gruber (1973) in their the empirical studies, one should divide observed variables into traditional (well-known) groups (such as industrial sectors) rather than pseudo (temporary, randomly determined) groups. Similar multiple-block dynamic equicorrelation structures are discussed in Engle and Kelly (2012) for GARCH models, while Lucas, Schwaab, and Zhang (2016) introduce a related multiple-block equicorrelation structure and propose the dynamic generalized hyperbolic (GH) skew-t-error model for generalized autoregressive score (GAS) models. On the other hand, for multivariate stochastic volatility models, Asai, Caporin, and McAleer (2015) propose a simplified block-type parameterization with static correlations.

There are several major difficulties in constructing multivariate volatility models. Firstly, it is necessary to guarantee that the covariance matrix of the high dimensional asset return

vector is positive definite, especially when time-varying correlation structure is incorporated. It would become the computational burden to check the positive definiteness as the number of asset returns increases. However, in a single block DESV model, such a condition reduces to a simple inequality, the equicorrelation is greater than $-(p-1)^{-1}$ where p is the number of asset returns (Kurose and Omori (2016)). In a multiple-block DESV model, we show that it is sufficient to check the positive definiteness of the matrix whose dimension is equal to the number of blocks (K) instead of the number of asset returns. Since the number of blocks is much smaller than that of the asset returns ($K << p$), it is expected to reduce the computational burden significantly to check the positive definiteness.

Secondly, we usually have too many parameters and latent variables in multivariate time-varying covariance models, which may result in their unstable estimates. This is partly because the information from the daily asset returns is not sufficient to identify all of these parameters and variables. To overcome this difficulty, we need either decrease the number of parameters and latent variables by simplifying the volatility model, or increase our source of information by adding more measurement equations. A single-block DESV model, which assumes the single equicorrelation structure among several daily asset returns, is promising and attractive to reduce the number of parameters in multivariate stochastic volatility models. However, such an assumption may not be realistic as the number of assets may increase, which is often the case in the portfolio optimizations. As a solution to this oversimplification, the multiple-block equicorrelation structure is proposed for high dimensional financial time series, where we assume common correlations within a group of asset returns, but allow different correlations for different groups. At the same time, we also incorporate additional source of information by jointly modeling the realized volatilities and realized correlations to obtain stable and accurate estimates of parameters, latent variables and leverage effects.

The realized measures, such as realized volatilities and realized correlations computed from high-frequency data, recently attracted a great deal of attention in the statistical modeling of latent variances and correlations of daily asset returns. The realized volatility, defined as the sum of the squared intraday returns over a specified time interval such as a day, is shown to be a consistent estimator of the latent stochastic variance of the log-price process as the sampling frequency diverges under the ideal market assumption (e.g., Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2001)). Further, Barndorff-Nielsen and

Shephard (2004) defined the realized covariance matrix as the sum of the outer product of intraday returns over a specified period and showed that it is a consistent estimator of the quadratic cross-variance of the multivariate log-price process under the ideal market assumption. The high-frequency data, however, has several major problems in the real market, which causes biases in the realized measures (e.g., Hansen and Lunde (2006) and Ubukata and Oya (2009)). These are the market microstructure noises such as the presence of bid-ask spread, discrete trading and price adjustment (e.g., O'Hara (1995) and Hasbrouck (2007))), non-trading hours (Hansen and Lunde (2005)) and non-synchronous trading for multivariate assets. Ignoring the overnight hours would underestimate the volatilities, while increasing sampling frequencies may result in observations with more microstructure noises and underestimate the correlation between asset returns (called Epps effect) since any two assets are rarely traded simultaneously (Epps (1979)).

Thus various realized measures discussed in the past literature can be used as additional measurement equations in our proposed framework (for the covariance matrices, e.g., Malliavin and Mancino (2002), Hayashi and Yoshida (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Kunitomo and Sato (2013) and Zhang (2011)). Noting that the daily returns are not heavily affected by those noises above, we consider the simultaneous modelling of the daily returns and the realized measures in line with Takahashi, Watanabe, and Omori (2009) who proposed such a joint modeling known as realized stochastic volatility (RSV) model (see also Dobrev and Szerszen (2010), Koopman and Scharth (2013), Zheng and Song (2014), Takahashi, Watanabe, and Omori (2016)). Hansen, Huang, and Shek (2012) similarly proposed the realized GARCH models, and these approaches are shown to lead to the substantial improvement in such as the prediction of volatilities and quantile forecasts (Watanabe (2012)). The daily returns successfully eliminate those biases in the realized measures which provide additional information on latent variables and model parameters.

We estimate model parameters and latent variables by Markov chain Monte Carlo (MCMC) simulation using Bayesian approach since it is difficult to evaluate the likelihood function using the high-dimensional numerical integration. For sampling latent correlation variables, we employ the multiple-trial Metropolized independent algorithm to increase the acceptance rate in Metropolis-Hastings algorithm. For sampling latent log volatility variables, we use a simple single-move sampler, which draws a single latent variable at a time given the other

latent variables and the parameters, when our MCMC samples are not highly autocorrelated. But, when those samples are highly autocorrelated, we also consider the multi-move sampler which draws a block of latent variables at a time given other blocks and the parameters to improve the sampling efficiencies.

The rest of this article is organized as follows. In Section 2, we propose the multivariate realized stochastic volatility model with multiple-block dynamic equicorrelation and cross leverage effect. Section 3 describes an efficient Bayesian estimation method for the proposed model. In Section 4, we illustrate our estimation method using simulated data. Section 5 applies our proposed model to the multivariate asset return data and the corresponding high-frequency data. Section 6 concludes this paper.

2 Multiple-block dynamic equicorrelations with realized measures, leverage and endogeneity

This section describes the multiple-block dynamic equicorrelation model with realized measures, leverage and endogeneity. We first introduce the basic multivariate stochastic volatility model in Section 2.1, and then extend it to consider the multiple-block dynamic correlations in Section 2.2. Finally, we incorporate the additional measurement equations based on realized measures in Section 2.3 to complete our proposed model.

2.1 Basic multivariate stochastic volatility model

Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})'$, $\boldsymbol{m}_t = (m_{1t}, \ldots, m_{pt})'$ and $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{pt})'$ denote a daily asset return vector, its mean vector and its corresponding log variance vector at time t $(t = 1, \ldots, n)$. The log variance vector $\boldsymbol{h}_t$ is assumed to follow a stationary first order autoregressive process with a diagonal coefficient matrix $\boldsymbol{\Phi}$, and, the initial log variance vector $\boldsymbol{h}_1$ is assumed to follow the stationary distribution setting $\boldsymbol{\Psi}_0 = \boldsymbol{\Phi}\boldsymbol{\Psi}_0\boldsymbol{\Phi} + \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\eta}}$ or $\mathrm{vec}(\boldsymbol{\Psi}_0) = (\mathbf{I}_{p^2} - \boldsymbol{\Phi} \otimes \boldsymbol{\Phi})^{-1}\mathrm{vec}(\boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\eta}})$ where $\mathbf{I}_p$ is a $p \times p$ unit matrix and $\mathrm{vec}(\mathbf{A})$ denotes a vectorization of a matrix $\mathbf{A}$. Thus we define a basic multivariate SV model:

$$\begin{aligned}
\boldsymbol{y}_t &= \boldsymbol{m}_t + \mathbf{V}_t^{1/2}\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_p(\mathbf{0}_p, \mathbf{R}_t), \quad t = 1, ..., n, && (1)\\
\boldsymbol{h}_{t+1} &= \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}_p(\mathbf{0}_p, \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\eta}}), \quad t = 1, ..., n-1, && (2)\\
\boldsymbol{m}_{t+1} &= \boldsymbol{m}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \sim \mathrm{N}_p(\mathbf{0}_p, \boldsymbol{\Omega}_{\boldsymbol{m}}), \quad t = 1, ..., n-1, && (3)\\
& \boldsymbol{h}_1 \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Psi}_0), \quad \boldsymbol{m}_1 \sim \mathrm{N}_p(\mathbf{0}_p, \kappa\mathbf{I}_p), && (4)
\end{aligned}$$

where $\mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Psi}_0)$ denotes a p-variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Psi}_0$, $\mathbf{0}_p$ is a p-dimensional zero vector, $\mathbf{R}_t$ is a correlation matrix at time t, and

$$\mathbf{V}_t = \mathrm{diag}\{\exp(h_{1t}), \ldots, \exp(h_{pt})\}, \quad t = 1, \ldots, n, \tag{5}$$

$$\boldsymbol{\Phi} = \mathrm{diag}(\phi_1, \ldots, \phi_p), \quad \boldsymbol{\Omega}_{\boldsymbol{m}} = \mathrm{diag}(\omega^2_{m_1}, \ldots, \omega^2_{m_p}), \tag{6}$$

For simplicity, the elements of the mean process $\boldsymbol{m}_t$, are assumed to follow independent random walk processes with their initial variances are set to some large constant κ.

2.2 Multiple-block dynamic equicorrelations

Suppose that we are able to divide $\boldsymbol{y}_t$ into K blocks (or groups) so that there are p_k elements in k-th block for $k = 1, \ldots, K$ ($\sum_{k=1}^K p_k = p$). We assume the equicorrelation structure within and between blocks. That is, the correlation coefficient between two asset returns within the k-th block is common, $\rho_{kk,t}$, and the correlation coefficient between one asset return in the k-th block and another asset return in the l-th block is also common, $\rho_{kl,t}$:

$$\mathbf{R}_t = \begin{bmatrix} \mathbf{R}_{11,t} & \cdots & \mathbf{R}_{1K,t} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{K1,t} & \cdots & \mathbf{R}_{KK,t} \end{bmatrix}, \quad \mathbf{R}_{kl,t} = \begin{cases} (1-\rho_{kk,t})\mathbf{I}_{p_k} + \rho_{kk,t}\mathbf{1}_{p_k}\mathbf{1}'_{p_k}, & k = l, \\ \rho_{kl,t}\mathbf{1}_{p_k}\mathbf{1}'_{p_l}, & k \neq l, \end{cases} \tag{7}$$

for $t = 1, \ldots, n$, where $\mathbf{1}_p$ denotes a $p \times 1$ vector with all elements equal to one. Then we consider the transformation of $\rho_{kl,t}$

$$g_{kl,t} = \log(1+\rho_{kl,t}) - \log(1-\rho_{kl,t}), \quad k, l = 1, \ldots, K, \ (l < k). \tag{8}$$

Further we assume that $\boldsymbol{g}_t = (g_{11,t}, \ldots, g_{K1,t}, g_{22,t}, \ldots, g_{K2,t}, \ldots, g_{KK,t})'$ follows a stationary autoregressive process with a diagonal coefficient matrix $\boldsymbol{\Theta} = \mathrm{diag}(\theta_1, \ldots, \theta_q)$ where $q = K(K+1)/2$ and θ_j corresponds to the j-th element of $\boldsymbol{g}_t$:

$$\boldsymbol{g}_{t+1} = \boldsymbol{\gamma} + \boldsymbol{\Theta}(\boldsymbol{g}_t - \boldsymbol{\gamma}) + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim \mathrm{N}_q(\mathbf{0}_q, \boldsymbol{\Sigma}), \tag{9}$$

$$\boldsymbol{g}_1 \sim \mathrm{N}_q(\mathbf{0}_q, \boldsymbol{\Sigma}_0), \quad \mathrm{vec}(\boldsymbol{\Sigma}_0) = (\mathbf{I}_{q^2} - \boldsymbol{\Theta} \otimes \boldsymbol{\Theta})^{-1}\mathrm{vec}(\boldsymbol{\Sigma}). \tag{10}$$

We note that $\boldsymbol{g}_t$'s are subject to the condition that the correlation matrix $\mathbf{R}_t$ is positive definite, which is given in Proposition 1[1].

Remark 1. If $p_k = 1$, then ρ_{kk} is not identified. Then we remove $g_{kk,t}$ from $\boldsymbol{g}_t$ and set $q - 1 \to q$.

[1] If $K = 1$, $\mathbf{R}_t$ is an equicorrelation matrix and is positive definite if and only if $-(p_k - 1)^{-1} < \rho_{kk,t} < 1$.

Proposition 1 (Positive definiteness of $\mathbf{R}_t$). Multiple-block equicorrelation matrix $\mathbf{R}_t$ defined in (7) is positive definite if all eigenvalues of $\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P}$ are positive where

$$\begin{aligned}\tilde{\mathbf{A}}_t &= \text{diag}(1-\rho_{11,t},\ldots,1-\rho_{KK,t}), \quad \tilde{\mathbf{B}}_t = \begin{pmatrix}\rho_{11,t} & \cdots & \rho_{1K,t}\\ \vdots & \ddots & \vdots \\ \rho_{K1,t} & \cdots & \rho_{KK,t}\end{pmatrix},\\ \mathbf{P} &= \text{diag}(p_1,\ldots,p_K), \quad |\rho_{kl}|<1, \quad k,l=1,\ldots,K.\end{aligned}$$

Proof: Define the $p \times p$ matrix $\mathbf{C}$ as

$$\mathbf{C} = \begin{pmatrix}\mathbf{1}_{p_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{1}_{p_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}\\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{1}_{p_K}\end{pmatrix}.$$

Then $\mathbf{R}_t = \mathbf{A}_t + \mathbf{B}_t$, where

$$\begin{aligned}\mathbf{A}_t &= \text{diag}\left\{(1-\rho_{11,t})\mathbf{1}'_{p_1},\cdots,(1-\rho_{KK,t})\mathbf{1}'_{p_K}\right\},\\ \mathbf{B}_t &= \begin{pmatrix}\rho_{11,t}\mathbf{1}_{p_1}\mathbf{1}'_{p_1} & \cdots & \rho_{1K,t}\mathbf{1}_{p_1}\mathbf{1}'_{p_K}\\ \vdots & \ddots & \vdots \\ \rho_{K1,t}\mathbf{1}_{p_K}\mathbf{1}'_{p_1} & \cdots & \rho_{KK,t}\mathbf{1}_{p_K}\mathbf{1}'_{p_K}\end{pmatrix} = \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}'.\end{aligned}$$

The determinant of $\mathbf{R}_t$ is given by

$$\begin{aligned}|\mathbf{R}_t| &= |\mathbf{A}_t+\mathbf{B}_t| = |\mathbf{A}_t| \times |\mathbf{I}_p + \mathbf{A}_t^{-1}\mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}'| = |\mathbf{A}_t| \times |\mathbf{I}_K + \tilde{\mathbf{B}}_t\mathbf{C}'\mathbf{A}_t^{-1}\mathbf{C}|\\ &= \prod_{k=1}^{K}(1-\rho_{kk,t})^{p_k} \times |\mathbf{I}_K + \tilde{\mathbf{B}}_t\mathbf{P}\tilde{\mathbf{A}}_t^{-1}|\\ &= \prod_{k=1}^{K}(1-\rho_{kk,t})^{p_k-1} \times |\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P}|,\end{aligned}$$

and the result follows. □

2.3 Realized volatilities and realized correlations

Let RV_{it} and $RCor_{ij,t}$ denote the realized volatility of the i-th asset and the realized correlation between the i-th and j-th asset returns at time t, and let $x_{it} \equiv \log RV_{it}$ and let $x^*_{kl,t}$ denote the sample mean of $\log\{(1+RCor_{ij,t})/(1-RCor_{ij,t})\}$ where the i-th and j-th asset returns belong to the k-th and l-th blocks, respectively. Further, we let $\boldsymbol{x}_t = (x_{1t},\ldots,x_{pt})'$

and $\boldsymbol{x}_t^* = (x_{11,t}^*, \ldots, x_{K1,t}^*, x_{22,t}^* \ldots, x_{K2,t}^*, \ldots, x_{KK,t}^*)'$. We incorporate the additional measurement equations based on realized measures as follows:

$$\begin{aligned} x_{it} &= \xi_i + h_{it} + w_{it}, \quad w_{it} \sim \mathrm{N}(0, \delta_i^2), \\ x_{kl,t}^* &= \xi_{kl} + g_{kl,t} + w_{kl,t}^*, \quad w_{kl,t}^* \sim \mathrm{N}(0, \delta_{kl}^{*2}), \end{aligned}$$

for $i = 1, \ldots, p$ and $k, l = 1, \ldots, K,\ (l < k)$. We set

$$\boldsymbol{x}_t = \boldsymbol{\xi} + \boldsymbol{h}_t + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim \mathrm{N}_p\left(\mathbf{0}_p, \boldsymbol{\Psi}_{\boldsymbol{ww}}\right), \tag{11}$$

$$\boldsymbol{x}_t^* = \boldsymbol{\xi}^* + \boldsymbol{g}_t + \boldsymbol{w}_t^*, \quad \boldsymbol{w}_t^* \sim \mathrm{N}_q(\mathbf{0}_q, \boldsymbol{\Delta}^*), \quad \boldsymbol{\Delta}^* = \mathrm{diag}(\boldsymbol{\delta}^*)^2, \tag{12}$$

where

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)', \quad \boldsymbol{\xi}^* = (\xi_1^*, \ldots, \xi_q^*)', \quad \boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)', \quad \boldsymbol{\delta}^* = (\delta_1^*, \ldots, \delta_q^*)',$$

and ξ_i^*, δ_j^* are relabelled to represent the i-th element of $\boldsymbol{\xi}^*$ and j-th element of $\boldsymbol{\delta}^*$.

We can consider the correlation between w_{it} and w_{jt} by using non-diagonal covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{ww}}$ and further will take account of the correlation among w_{it}, ϵ_{jt} and η_{kt} to describe the endogeneity as discussed in Koopman and Scharth (2013). On the other hand, since $x_{ij,t}^*$'s are subject to relatively large measurement noises in empirical studies, we assume no correlation between $w_{ij,t}^*$'s for simplicity.

Bias correction in realized measures. We note that $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^*$ are adjustment vectors to correct the biases for the realized volatilities and correlations. The positive element of $\boldsymbol{\xi}$ indicates that the realized measure of the variance has an upward bias due to, for example, the market microstructure noise, while the negative elements indicates that it has a downward bias due to such as the existence of non-trading hours. Further, the non-zero element of $\boldsymbol{\xi}^*$ would be caused by the market microstructure noise or the non-synchronous trading effect.

Finally, we consider correlations between $\boldsymbol{z}_t \equiv \left((\mathbf{R}_t^{-1/2}\boldsymbol{\epsilon}_t)', \boldsymbol{w}_t', \boldsymbol{\eta}_t'\right)'$ to incorporate the leverage effect and endogeneity as we shall discuss below:

$$\boldsymbol{z}_t = \begin{pmatrix} \boldsymbol{z}_{1t} \\ \boldsymbol{z}_{2t} \end{pmatrix} \sim \mathrm{N}_{3p}\left(\mathbf{0}_{3p}, \boldsymbol{\Psi}\right), \quad \boldsymbol{\Psi} = \begin{pmatrix} \mathbf{I}_p & \boldsymbol{\Psi}_{12} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} \end{pmatrix}, \tag{13}$$

where

$$\boldsymbol{z}_{1t} = \mathbf{R}_t^{-\frac{1}{2}}\boldsymbol{\epsilon}_t (\equiv \boldsymbol{\varepsilon}_t), \quad \boldsymbol{z}_{2t} = \begin{pmatrix} \boldsymbol{w}_t \\ \boldsymbol{\eta}_t \end{pmatrix}, \tag{14}$$

$$\boldsymbol{\Psi}_{12} = \mathrm{Cov}(\boldsymbol{z}_{1t}, \boldsymbol{z}_{2t}) = (\boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}, \boldsymbol{\Psi}_{\boldsymbol{\varepsilon\eta}}), \quad \boldsymbol{\Psi}_{22} = \mathrm{Var}(\boldsymbol{z}_{2t}) = \begin{pmatrix} \boldsymbol{\Psi}_{\boldsymbol{ww}} & \boldsymbol{\Psi}_{\boldsymbol{w\eta}} \\ \boldsymbol{\Psi}_{\boldsymbol{\eta w}} & \boldsymbol{\Psi}_{\boldsymbol{\eta\eta}} \end{pmatrix}. \tag{15}$$

Leverage effect and endogeneity. We consider correlations among the error terms for the measurement equation of the daily asset return (through $\boldsymbol{\varepsilon}_t = \mathbf{R}_t^{-1/2}\boldsymbol{\epsilon}_t$), the measurement equation of the realized measure of the variance ($\boldsymbol{w}_t$) and the state equation of the log variance ($\boldsymbol{\eta}_t$). First, we extend the univariate SV model with leverage by incorporating the correlation between $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ (see, e.g., Kurose and Omori (2016)). Second, there may exist some dependence between the daily return and the realized measure since they are functions of intraday returns with market microstructure noises, discretizations and jumps (see Peters and de Vilder (2006), Koopman and Scharth (2013) and Chaussé and Xu (2016)). Thus, based on the literature that addresses such an endogeneity between the realized measure of the variance and the asset return, we consider correlations between $\boldsymbol{w}_t$ and $\boldsymbol{\varepsilon}_t$. Further, we also take account of the correlation between $\boldsymbol{w}_t$ and $\boldsymbol{\eta}_t$. On the other hand, we assume that there is no correlation between $\boldsymbol{w}_t^*$ and the other error terms to avoid overparameterization leading to unstable parameter estimates.

The above equations (1)-(15) define the multiple-block dynamic equicorrelation stochastic volatility model with realized measures, leverage and endogeneity (multiple-block RDESV-LE).

3 Markov chain Monte Carlo estimation

3.1 Prior distributions and joint posterior density

For prior distributions of $\boldsymbol{\vartheta} = (\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\xi}, \boldsymbol{\xi}^*, \boldsymbol{\Delta}^*, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}_{\boldsymbol{m}})$, we assume

$$\boldsymbol{\mu} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\mu}0}, \mathbf{S}_{\boldsymbol{\mu}0}), \quad \boldsymbol{\gamma} \sim \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\gamma}0}, \mathbf{S}_{\boldsymbol{\gamma}0}), \tag{16}$$

$$\boldsymbol{\xi} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\xi}0}, \mathbf{S}_{\boldsymbol{\xi}0}), \quad \boldsymbol{\xi}^* \sim \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\xi}^*0}, \mathbf{S}_{\boldsymbol{\xi}^*0}), \quad \delta_j^{*2} \sim \mathrm{IG}(\alpha_{\delta_j 0}^*/2, \beta_{\delta_j 0}^*/2), \tag{17}$$

$$(\phi_i + 1)/2 \sim \mathrm{Be}(a_{\phi_i}, b_{\phi_i}), \quad (\theta_j + 1)/2 \sim \mathrm{Be}(a_{\theta_j}, b_{\theta_j}), \tag{18}$$

$$\boldsymbol{\Sigma} \sim \mathrm{IW}_q(n_{\boldsymbol{\Sigma}0}, \mathbf{S}_{\boldsymbol{\Sigma}0}), \quad \omega_{m_i}^2 \sim \mathrm{IG}(\alpha_{\omega_{m_i}0}/2, \beta_{\omega_{m_i}0}/2), \tag{19}$$

for $i = 1, \ldots, p$, $j = 1, \ldots, q$, where $\mathrm{Be}(a, b)$ denotes a beta distribution with parameters a, b, $\mathrm{IG}(\alpha, \beta)$ denotes an inverse gamma distribution with shape parameter α and rate parameter β and $\mathrm{IW}(n, \mathbf{S})$ denotes an inverse Wishart distribution with parameters $(n, \mathbf{S})$.

For $\boldsymbol{\Psi}$, we consider the prior distribution of the inverse matrix $\boldsymbol{\Psi}^{-1}$ as follows. It is easy

to show that

$$\boldsymbol{\Psi}^{-1} = \begin{pmatrix} \boldsymbol{\Psi}^{11} & \boldsymbol{\Psi}^{12} \\ \boldsymbol{\Psi}^{21} & \boldsymbol{\Psi}^{22} \end{pmatrix}, \quad \boldsymbol{\Psi}^{11} = \mathbf{I}_p + \boldsymbol{\Psi}^{12}(\boldsymbol{\Psi}^{22})^{-1}\boldsymbol{\Psi}^{21},$$

where $\boldsymbol{\Psi}^{12} = \boldsymbol{\Psi}^{21\prime}$ and $\boldsymbol{\Psi}^{22}$ are the $p \times 2p$ and $2p \times 2p$ matrices, respectively. Thus we assume

$$\boldsymbol{\Psi}^{22} \sim \mathrm{W}_{2p}(n_0, \mathbf{S}_0), \tag{20}$$

$$\mathrm{vec}(\boldsymbol{\Psi}^{21})|\boldsymbol{\Psi}^{22} \sim \mathrm{N}_{2p^2}\left(\mathrm{vec}(\boldsymbol{\Psi}^{22}\boldsymbol{\Delta}_0), \boldsymbol{\Lambda}_0 \otimes \boldsymbol{\Psi}^{22}\right), \tag{21}$$

where $\mathrm{W}(n, \mathbf{S})$ denotes a Wishart distribution with parameters $(n, \mathbf{S})$. Let $\pi(\boldsymbol{\vartheta})$ denote the prior probability density function of $\boldsymbol{\vartheta}$. Then, the joint posterior density function is

$$\begin{aligned}
&\pi(\boldsymbol{\vartheta}, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n) \\
&\propto \pi(\boldsymbol{\vartheta}) \times |\boldsymbol{\Psi}|^{-\frac{n-1}{2}} \exp\left(\sum_{t=1}^{n-1} l_t - \frac{1}{2}\sum_{t=1}^{n-1} \boldsymbol{z}_{2t}' \boldsymbol{\Psi}_{22}^{-1} \boldsymbol{z}_{2t} \right) \\
&\quad \times |\mathbf{I}_p - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}} \boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1} \boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}|^{-\frac{1}{2}} |\boldsymbol{\Psi}_{\boldsymbol{ww}}|^{-\frac{1}{2}} \exp\left\{ l_n - \frac{1}{2} \boldsymbol{z}_{2n}' \boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1} \boldsymbol{z}_{2n} \right\} \\
&\quad \times |\boldsymbol{\Delta}^*|^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2}\sum_{t=1}^{n} (\boldsymbol{x}_t^* - \boldsymbol{\xi}^* - \boldsymbol{g}_t)' \boldsymbol{\Delta}^{*-1} (\boldsymbol{x}_t^* - \boldsymbol{\xi}^* - \boldsymbol{g}_t) \right\} \\
&\quad \times |\boldsymbol{\Sigma}|^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{n-1} \{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \boldsymbol{\Theta}(\boldsymbol{g}_t - \boldsymbol{\gamma})\}' \boldsymbol{\Sigma}^{-1} \{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \boldsymbol{\Theta}(\boldsymbol{g}_t - \boldsymbol{\gamma})\} \right] \\
&\quad \times |\boldsymbol{\Psi}_0|^{-\frac{1}{2}} |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left\{ (\boldsymbol{h}_1 - \boldsymbol{\mu})' \boldsymbol{\Psi}_0^{-1} (\boldsymbol{h}_1 - \boldsymbol{\mu}) + (\boldsymbol{g}_1 - \boldsymbol{\gamma})' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{g}_1 - \boldsymbol{\gamma}) \right\} \right] \\
&\quad \times |\boldsymbol{\Omega}_{\boldsymbol{m}}|^{-\frac{n-1}{2}} \exp\left\{ -\frac{1}{2\kappa} \boldsymbol{m}_1' \boldsymbol{m}_1 - \frac{1}{2}\sum_{t=1}^{n-1} (\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)' \boldsymbol{\Omega}_{\boldsymbol{m}}^{-1} (\boldsymbol{m}_{t+1} - \boldsymbol{m}_t) \right\},
\end{aligned} \tag{22}$$

where, for $t = 1, \ldots, n-1$,

$$\begin{aligned}
l_t = &-\frac{1}{2}\Big(\boldsymbol{z}_{1t} - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{z}_{2t}\Big)' (\mathbf{I}_p - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21})^{-1} \Big(\boldsymbol{z}_{1t} - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{z}_{2t}\Big) \\
&- \frac{1}{2}\log|\mathbf{R}_t| - \frac{1}{2}\sum_{i=1}^{p} h_{it}, \\
&\boldsymbol{z}_{1t} = \boldsymbol{\varepsilon}_t = \mathbf{R}_t^{-\frac{1}{2}} \mathbf{V}_t^{-\frac{1}{2}} (\boldsymbol{y}_t - \boldsymbol{m}_t), \quad \boldsymbol{z}_{2t} = \begin{pmatrix} \boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t \\ \boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu}) \end{pmatrix},
\end{aligned}$$

and, for $t = n$,

$$\begin{aligned}
l_n = &-\frac{1}{2}\Big(\boldsymbol{z}_{1n} - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{z}_{2n}\Big)' (\mathbf{I}_p - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}})^{-1} \Big(\boldsymbol{z}_{1n} - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{z}_{2n}\Big) \\
&- \frac{1}{2}\log|\mathbf{R}_n| - \frac{1}{2}\sum_{i=1}^{p} h_{in}, \\
&\boldsymbol{z}_{1n} = \boldsymbol{\varepsilon}_n = \mathbf{R}_n^{-\frac{1}{2}} \mathbf{V}_n^{-\frac{1}{2}} (\boldsymbol{y}_n - \boldsymbol{m}_n), \quad \boldsymbol{z}_{2n} = \boldsymbol{x}_n - \boldsymbol{\xi} - \boldsymbol{h}_n.
\end{aligned}$$

We implement the MCMC algorithm in six blocks:

1. Initialize $\{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.
2. Generate $\boldsymbol{\vartheta}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n$.
3. Generate $\{\boldsymbol{h}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.
4. Generate $\{\boldsymbol{g}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.
5. Generate $\{\boldsymbol{m}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.
6. Go to 2.

3.2 Spectral decomposition of $\mathbf{R}_t$

As shown in the previous subsection, we need to evaluate $\mathbf{R}_t^{-1/2}$ in the posterior probability density. This subsection describes the spectral decomposition of $\mathbf{R}_t$ to compute $\mathbf{R}_t^{-1/2}$, where the eigenvalues and eigenvectors are given in the following proposition.

Proposition 2 (Eigenvalues and eigenvectors of $\mathbf{R}_t$). Let $\lambda_{kt} = 1-\rho_{kk,t}$ $(k = 1, \ldots, K)$, and further let λ_{kt}^* and $\boldsymbol{u}_{kt} = (u_{1k,t}, \ldots, u_{Kk,t})'$ $(k = 1, \ldots, K)$ denote eigenvalues and eigenvectors of the matrix $\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P}$. Then λ_{kt}'s and λ_{kt}^*'s are eigenvalues of $\mathbf{R}_t$ where λ_{kt} is a repeated eigenvalue of multiplicity $p_k - 1$. The orthonormal eigenvectors corresponding to λ_{kt} are

$$\boldsymbol{e}_{km} = \begin{pmatrix} \mathbf{0}_{\sum_{j=1}^{k-1} p_j} \\ \boldsymbol{a}_{km} \\ \mathbf{0}_{\sum_{j=k+1}^{K} p_j} \end{pmatrix}, \quad \text{where} \quad \boldsymbol{a}_{km} = \frac{1}{\sqrt{(p_k - m)(p_k - m + 1)}} \begin{pmatrix} \mathbf{0}_{m-1} \\ p_k - m \\ -\mathbf{1}_{p_k - m} \end{pmatrix},$$

for $m = 1, \ldots, p_k - 1$, and $k = 1, \ldots, K$. The eigenvector corresponding to λ_{kt}^* is

$$\boldsymbol{e}_{kt}^* = \frac{1}{\sqrt{\sum_{j=1}^K u_{jk,t}^2 p_j}} \begin{pmatrix} u_{1k,t}\mathbf{1}_{p_1} \\ \vdots \\ u_{Kk,t}\mathbf{1}_{p_K} \end{pmatrix}, k = 1, \ldots, K,$$

and orthonormal if λ_{kt}^*'s are distinct eigenvalues.

Proof: Since the determinant of the correlation matrix $\mathbf{R}_t$ is

$$|\mathbf{R}_t| \;=\; \prod_{k=1}^K (1 - \rho_{kk,t})^{p_k - 1} \times |\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P}|,$$

$\lambda_{kt} = 1-\rho_{kk,t}$ and λ^*_{kt} $(k = 1, \ldots, K)$ are eigenvalues of $\mathbf{R}_t$ where λ_{kt} is a repeated eigenvalue of multiplicity $p_k - 1$. Noting that

$$\left\{(1-\rho_{kk,t})\mathbf{I}_{p_k} + \rho_{kk,t}\mathbf{1}_{p_k}\mathbf{1}'_{p_k}\right\}\boldsymbol{a}_{km} = (1-\rho_{kk,t})\boldsymbol{a}_{km},$$
$$\rho_{jk,t}\mathbf{1}_{p_j}\mathbf{1}'_{p_k}\boldsymbol{a}_{km} = \mathbf{0}_{p_j}, \quad j \neq k,$$

we obtain

$$\mathbf{R}_t\boldsymbol{e}_{km} = \lambda_{kt}\boldsymbol{e}_{km}, \quad m = 1, \ldots, p_k - 1,$$

for $k = 1, \ldots, K$. Similarly, since

$$\left\{(1-\rho_{jj,t})\mathbf{I}_{p_j} + \rho_{jj,t}\mathbf{1}_{p_j}\mathbf{1}'_{p_j}\right\}(u_{jk,t}\mathbf{1}_{p_j}) = \{1 + (p_j - 1)\rho_{jj,t}\}u_{jk,t}\mathbf{1}_{p_j},$$
$$\rho_{jm,t}\mathbf{1}_{p_j}\mathbf{1}'_{p_m}(u_{mk,t}\mathbf{1}_{p_m}) = (p_m\rho_{jm,t})u_{mk,t}\mathbf{1}_{p_j}, \quad j \neq m,$$

and

$$\sum_{m\neq j}(p_m\rho_{jm,t})u_{mk,t} + \{1 + (p_j - 1)\rho_{jj,t}\}u_{jk,t} = \lambda^*_{kt}u_{jk,t}, \quad j = 1, \ldots, K,$$

we obtain

$$\mathbf{R}_t\boldsymbol{e}^*_{kt} = \lambda^*_{kt}\boldsymbol{e}^*_{kt}, \quad k = 1, \ldots, K,$$

for $k = 1, \ldots, K$. Finally, it is easy to see that

$$\boldsymbol{e}'_{jm}\boldsymbol{e}_{kn} = 0, \quad (m \neq n \text{ or } j \neq k), \quad \boldsymbol{e}'_{jm}\boldsymbol{e}^*_{kt} = 0,$$

and that $\boldsymbol{e}^{*\prime}_{jt}\boldsymbol{e}^*_{kt} = 0$ if $\lambda^*_{jt} \neq \lambda^*_{kt}$ since $\mathbf{R}_t$ is symmetric. □

Using Proposition 2, we set

$$\mathbf{R}_t = \sum_{k=1}^{K}\left\{\lambda^*_{kt}\boldsymbol{e}^*_{kt}\boldsymbol{e}^{*\prime}_{kt} + \lambda_{kt}\left(\sum_{m=1}^{p_k-1}\boldsymbol{e}_{km}\boldsymbol{e}'_{km}\right)\right\}, \tag{23}$$

$$\mathbf{R}_t^{-1/2} = \mathrm{diag}(\lambda^*_{1t}, \ldots, \lambda^*_{Kt}, \lambda_{1t}\mathbf{1}'_{p_1-1}, \lambda_{2t}\mathbf{1}'_{p_2-1}, \ldots, \lambda_{Kt}\mathbf{1}'_{p_K-1})^{-1/2}$$
$$\times [\boldsymbol{e}^*_{1t}, \ldots, \boldsymbol{e}^*_{Kt}, \boldsymbol{e}_{11}, \ldots, \boldsymbol{e}_{1,p_1-1}, \boldsymbol{e}_{21} \ldots, \boldsymbol{e}_{2,p_2-1}, \ldots, \boldsymbol{e}_{K,p_K-1}]', \tag{24}$$

where $\lambda^*_{1t} \geq \cdots \geq \lambda^*_{Kt} > 0$. If λ^*_{kt} is a repeated eigenvalue, then we orthogonalize the corresponding eigenvectors. Further, we set the first element of the eigenvector to be positive for identification purpose only.

Remark 2. Alternatively, the Cholesky decomposition may be used to compute $\mathbf{R}_t^{-1/2}$. In our proposed equicorrelations model, we consider the groups of asset returns which may be highly correlated with each other and, in such a case, some eigenvalues could be very close to zero where Cholesky decompositions may become unstable. Thus, we use the spectral decomposition where most eigenvalues and eigenvectors are obtained analytically as in Proposition 2.

Remark 3. When there is no leverage effect ($\mathbf{\Psi}_{12} = \mathbf{O}$ and $\mathbf{\Psi}_{\boldsymbol{\varepsilon w}} = \mathbf{O}$), Proposition 3 (see Appendix A) derives the inverse matrix of $\mathbf{R}_t$ to reduce the computational burden.

3.3 Generation of $\{\boldsymbol{h}_t\}_{t=1}^n$

A simple sampling method for $\{\boldsymbol{h}_t\}_{t=1}^n$, which we use is a single-move sampler that draws a single latent variable $\boldsymbol{h}_t$ at a time given the other $\boldsymbol{h}_t$'s and the parameters using Metropolis-Hastings (MH) algorithm. To construct a proposal density for $\boldsymbol{h}_t$, we consider the following Taylor expansion around $x_{it} - \xi_i$,

$$\exp\left(-\frac{h_{it}}{2}\right) \approx \exp\left(-\frac{x_{it}-\xi_i}{2}\right) - \frac{1}{2}\exp\left(-\frac{x_{it}-\xi_i}{2}\right)\{h_{it} - (x_{it}-\xi_i)\},$$

and obtain

$$\boldsymbol{z}_{1t} = \mathbf{R}_t^{-\frac{1}{2}}\mathbf{V}_t^{-\frac{1}{2}}(\boldsymbol{y}_t - \boldsymbol{m}_t) \quad \approx \quad \tilde{\boldsymbol{z}}_{1t} = \mathbf{R}_t^{-\frac{1}{2}}(\tilde{\boldsymbol{y}}_t - \mathbf{F}_t\boldsymbol{h}_t),$$

where

$$\mathbf{F}_t = \frac{1}{2}\mathrm{diag}\left((y_{1t}-m_{1t})\exp\left(-\frac{x_{1t}-\xi_1}{2}\right), \ldots, (y_{pt}-m_{pt})\exp\left(-\frac{x_{pt}-\xi_p}{2}\right)\right), \quad (25)$$

$$\tilde{\boldsymbol{y}}_t = 2\mathbf{F}_t\left(\mathbf{1}_p + \frac{\boldsymbol{x}_t - \boldsymbol{\xi}}{2}\right). \quad (26)$$

Let $\tilde{\boldsymbol{z}}_t = (\tilde{\boldsymbol{z}}_{1t}', \boldsymbol{z}_{2t}')'$. Then the conditional posterior density of $\boldsymbol{h}_t$ is

$$\begin{aligned}
\pi(\boldsymbol{h}_t|\cdot) &\propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{z}_t'\mathbf{\Psi}^{-1}\boldsymbol{z}_t + \mathbf{1}_p'\boldsymbol{h}_t + \boldsymbol{z}_{t-1}'\mathbf{\Psi}^{-1}\boldsymbol{z}_{t-1}\right)\right\} \\
&\approx \exp\left\{-\frac{1}{2}\left(\tilde{\boldsymbol{z}}_t'\mathbf{\Psi}^{-1}\tilde{\boldsymbol{z}}_t + \mathbf{1}_p'\boldsymbol{h}_t + \boldsymbol{z}_{t-1}'\mathbf{\Psi}^{-1}\boldsymbol{z}_{t-1}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{h}_t - \boldsymbol{m}_{\boldsymbol{h}t})'\mathbf{S}_{\boldsymbol{h}t}^{-1}(\boldsymbol{h}_t - \boldsymbol{m}_{\boldsymbol{h}t})\right\}, \quad (27)
\end{aligned}$$

where

$$\boldsymbol{m}_{\boldsymbol{h}t} = \boldsymbol{m}_{\boldsymbol{h}t,1} + \boldsymbol{m}_{\boldsymbol{h}t,2}, \quad \mathbf{S}_{\boldsymbol{h}t}^{-1} = \mathbf{S}_{\boldsymbol{h}t,1}^{-1} + \mathbf{S}_{\boldsymbol{h}t,2}^{-1}, \quad (28)$$

and

$$
\begin{aligned}
\mathbf{S}_{\boldsymbol{h}t,1}^{-1} &= \begin{pmatrix} \mathbf{R}_t^{-1/2}\mathbf{F}_t \\ \mathbf{I}_p \\ \boldsymbol{\Phi} \end{pmatrix}' \boldsymbol{\Psi}^{-1} \begin{pmatrix} \mathbf{R}_t^{-1/2}\mathbf{F}_t \\ \mathbf{I}_p \\ \boldsymbol{\Phi} \end{pmatrix}, \\
\mathbf{S}_{\boldsymbol{h}t,2} &= \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\eta}} - \begin{pmatrix} \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{w}} \end{pmatrix} \begin{pmatrix} \mathbf{I}_p & \boldsymbol{\Psi}_{\boldsymbol{\varepsilon}\boldsymbol{w}} \\ \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{w}} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\Psi}_{\boldsymbol{\varepsilon}\boldsymbol{\eta}} \\ \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\eta}} \end{pmatrix}, \\
\boldsymbol{m}_{\boldsymbol{h}t,1} &= \mathbf{S}_{\boldsymbol{h}t} \left\{ -\frac{1}{2}\mathbf{1}_p + \begin{pmatrix} \mathbf{F}_t'\mathbf{R}_t^{-1/2\prime} & \mathbf{I}_p & \boldsymbol{\Phi} \end{pmatrix} \boldsymbol{\Psi}^{-1} \begin{pmatrix} \mathbf{R}_t^{-1/2}\tilde{\boldsymbol{y}}_t \\ \boldsymbol{x}_t - \boldsymbol{\xi} \\ \boldsymbol{h}_{t+1} - (\mathbf{I}_p - \boldsymbol{\Phi})\boldsymbol{\mu} \end{pmatrix} \right\}, \\
\boldsymbol{m}_{\boldsymbol{h}t,2} &= \mathbf{S}_{\boldsymbol{h}t}\mathbf{S}_{\boldsymbol{h}t,2}^{-1} \\
&\times \left\{ \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{h}_{t-1} - \boldsymbol{\mu}) + \begin{pmatrix} \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{\eta}\boldsymbol{w}} \end{pmatrix} \begin{pmatrix} \mathbf{I}_p & \boldsymbol{\Psi}_{\boldsymbol{\varepsilon}\boldsymbol{w}} \\ \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{w}} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{z}_{1,t-1} \\ \boldsymbol{x}_{t-1} - \boldsymbol{\xi} - \boldsymbol{h}_{t-1} \end{pmatrix} \right\}.
\end{aligned}
$$

Define $k(\boldsymbol{h}_t) = -\frac{1}{2}(\boldsymbol{z}_t'\boldsymbol{\Psi}^{-1}\boldsymbol{z}_t - \tilde{\boldsymbol{z}}_t'\boldsymbol{\Psi}^{-1}\tilde{\boldsymbol{z}}_t)$. We propose a candidate $\boldsymbol{h}_t^\dagger \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{h}t}, \mathbf{S}_{\boldsymbol{h}t})$, and accept it with probability $\min[1, \exp\{k(\boldsymbol{h}_t^\dagger) - k(\boldsymbol{h}_t)\}]$.

Remark 4. If the above single-move sampler produces the highly autocorrelated MCMC samples, we use more efficient multi-move sampler as discussed in Appendix C.2.

Remark 5. We note that there might be alternative efficient proposals based on efficient importance sampling (Richard and Zhang (2007) or particle Gibbs (Andrieu, Doucet, and Holenstein (2010)). The investigation of the performance of those samplers is left for the future work.

3.4 Generation of $\{\boldsymbol{g}_t\}_{t=1}^n$

We consider a single-move sampler for sampling $\boldsymbol{g}_t$ where we sample $\boldsymbol{g}_t$ one at a time given other parameters and latent variables. For $t = 1, \ldots, n$, the conditional posterior density of $\boldsymbol{g}_t$ is

$$
\pi(\boldsymbol{g}_t|\cdot) \propto \exp(l_t) \times \exp\left\{ -\frac{1}{2}(\boldsymbol{g}_t - \boldsymbol{m}_{\boldsymbol{g}t})'\mathbf{S}_{\boldsymbol{g}t}^{-1}(\boldsymbol{g}_t - \boldsymbol{m}_{\boldsymbol{g}t}) \right\},
$$

where

$$
\mathbf{S}_{\boldsymbol{g}t} = \{\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Theta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Theta} + \boldsymbol{\Delta}^{*-1}\}^{-1}, \tag{29}
$$

$$
\boldsymbol{m}_{\boldsymbol{g}t} = \mathbf{S}_{\boldsymbol{g}t}[\boldsymbol{\Theta}\boldsymbol{\Sigma}^{-1}\{\boldsymbol{g}_{t+1} - (\mathbf{I}_q - \boldsymbol{\Theta})\boldsymbol{\gamma}\} + \boldsymbol{\Sigma}^{-1}\{\boldsymbol{\gamma} + \boldsymbol{\Theta}(\boldsymbol{g}_{t-1} - \boldsymbol{\gamma})\} + \boldsymbol{\Delta}^{*-1}(\boldsymbol{x}_t^* - \boldsymbol{\xi}^*)]. \tag{30}
$$

Given other parameters and latent variables, we generate a candidate from a truncated normal distribution, $\boldsymbol{g}_t^\dagger \sim \mathrm{TN}_{\{\mathbf{R}_t^\dagger > 0\}}(\boldsymbol{m}_{\boldsymbol{g}t}, \mathbf{S}_{\boldsymbol{g}t})$, where we guarantee the positive definiteness of the proposed correlation matrix $\mathbf{R}_t^\dagger$, *i.e.*, $\mathbf{R}_t^\dagger > 0$. Let l_t and $l_t^\dagger$ denote the conditional log likelihood in Section 3.1 evaluated at the current sample $\boldsymbol{g}_t$ and the candidate $\boldsymbol{g}_t^\dagger$ respectively. Then we accept it with probability $\min\{1, \exp(l_t^\dagger - l_t)\}$.

Remark 6. As we shall see in empirical studies, the realized correlations may be subject to measurement errors with relatively large variances. It may be difficult to propose a good candidate in such a case, and the following multiple-trial Metropolized independent sampler would improve the acceptance rate in the MH algorithm (See Liu, Liang, and Wong (2000) and Liu (2001)).

Multiple-Trial Metropolized Independent Sampler:

1. Generate a trial set of i.i.d. samples $\boldsymbol{g}_t^{(k)} \sim \mathrm{TN}_{\{\mathbf{R}_t^{(k)} > 0\}}(\boldsymbol{m}_{\boldsymbol{g}t}, \mathbf{S}_{\boldsymbol{g}t})$, $k = 1, \ldots, K_{\mathrm{MT}}$.
2. Compute $w(\boldsymbol{g}_t^{(k)}) = \exp(l_t^{(k)})$ where $l_t^{(k)}$ is l_t evaluated at $\boldsymbol{g}_t = \boldsymbol{g}_t^{(k)}$.
3. Draw $\boldsymbol{g}_t^\dagger$ from the trial set with probability proportional to $w(\boldsymbol{g}_t^{(k)})$.
4. Accept $\boldsymbol{g}_t^\dagger$ with probability
$$\min\left\{1, \frac{\sum_{k=1}^{K_{\mathrm{MT}}} w(\boldsymbol{g}_t^{(k)})}{w(\boldsymbol{g}_t) + \sum_{k=1}^{K_{\mathrm{MT}}} w(\boldsymbol{g}_t^{(k)}) - w(\boldsymbol{g}_t^\dagger)}\right\}.$$

3.5 Generation of $\{\boldsymbol{m}_t\}_{t=1}^n$

Define $\hat{\boldsymbol{y}}_t$ and $\hat{\boldsymbol{\Sigma}}_t$ as

$$\hat{\boldsymbol{y}}_t \equiv \begin{cases} \boldsymbol{y}_t - \mathbf{V}_t^{1/2}\mathbf{R}_t^{1/2}\boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{z}_{2t}, & t = 1, \ldots, n-1, \\ \boldsymbol{y}_n - \mathbf{V}_n^{1/2}\mathbf{R}_n^{1/2}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{z}_{2n}, & t = n, \end{cases}$$

$$\hat{\boldsymbol{\Sigma}}_t \equiv \begin{cases} \mathbf{V}_t^{1/2}\mathbf{R}_t^{1/2}(\mathbf{I}_p - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21})\mathbf{R}_t^{1/2\prime}\mathbf{V}_t^{1/2}, & t = 1, \ldots, n-1, \\ \mathbf{V}_n^{1/2}\mathbf{R}_n^{1/2}(\mathbf{I}_p - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}})\mathbf{R}_n^{1/2\prime}\mathbf{V}_n^{1/2}, & t = n. \end{cases}$$

Then it can be shown that the conditional posterior distribution of $\{\boldsymbol{m}_t\}_{t=1}^m$ is the same as that of the following linear Gaussian state space model:

$$\begin{aligned} \hat{\boldsymbol{y}}_t &= \boldsymbol{m}_t + \hat{\boldsymbol{\epsilon}}_t, \quad \hat{\boldsymbol{\epsilon}}_t \sim \mathrm{N}_p(\mathbf{0}_p, \hat{\boldsymbol{\Sigma}}_t), \quad t = 1, \ldots, n, \\ \boldsymbol{m}_{t+1} &= \boldsymbol{m}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \sim \mathrm{N}_p(\mathbf{0}_p, \boldsymbol{\Omega}_{\boldsymbol{m}}), \quad t = 1, \ldots, n-1, \\ & \quad \boldsymbol{m}_1 \sim \mathrm{N}_p(\mathbf{0}_p, \kappa\mathbf{I}_p). \end{aligned}$$

Thus we generate all $\{\boldsymbol{m}_t\}_{t=1}^n$ at a time using a simulation smoother (de Jong and Shephard (1995) or Durbin and Koopman (2002)).

3.6 Generation of $\boldsymbol{\Psi}$

First define

$$
\begin{aligned}
\boldsymbol{\Xi} &= \begin{pmatrix} \boldsymbol{\Xi}_{11} & \boldsymbol{\Xi}_{12} \\ \boldsymbol{\Xi}_{21} & \boldsymbol{\Xi}_{22} \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{n-1} \boldsymbol{z}_{1t}\boldsymbol{z}_{1t}' & \sum_{t=1}^{n-1} \boldsymbol{z}_{1t}\boldsymbol{z}_{2t}' \\ \sum_{t=1}^{n-1} \boldsymbol{z}_{2t}\boldsymbol{z}_{1t}' & \sum_{t=1}^{n-1} \boldsymbol{z}_{2t}\boldsymbol{z}_{2t}' \end{pmatrix}, \\
g(\boldsymbol{\Psi}) &= |\boldsymbol{\Psi}_0|^{-1/2} \exp\left\{ -\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Psi}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu}) \right\} \\
&\quad \times |\mathbf{I}_p - \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}|^{-\frac{1}{2}} |\boldsymbol{\Psi}_{\boldsymbol{ww}}|^{-\frac{1}{2}} \exp\left\{ l_n - \frac{1}{2}\boldsymbol{z}_{2n}'\boldsymbol{\Psi}_{\boldsymbol{ww}}^{-1}\boldsymbol{z}_{2n} \right\}.
\end{aligned}
$$

Then it can be shown that the conditional posterior density of $\boldsymbol{\Psi}^{12}$ and $\boldsymbol{\Psi}^{22}$ is

$$
\begin{aligned}
&\pi(\boldsymbol{\Psi}^{12}, \boldsymbol{\Psi}^{22}|\cdot) \\
&\propto g(\boldsymbol{\Psi}) \times |\boldsymbol{\Psi}|^{-\frac{n-1}{2}} \exp\left\{ -\frac{1}{2}\sum_{t=1}^{n-1} \boldsymbol{z}_t'\boldsymbol{\Psi}^{-1}\boldsymbol{z}_t \right\} \times |\boldsymbol{\Psi}^{22}|^{\frac{n_0-2p-1}{2}} \exp\left\{ -\frac{1}{2}\mathrm{tr}(\mathbf{S}_0^{-1}\boldsymbol{\Psi}^{22}) \right\} \\
&\quad \times |\boldsymbol{\Psi}^{22}|^{-\frac{p}{2}} \exp\left\{ -\frac{1}{2}\mathrm{vec}(\boldsymbol{\Psi}^{21} - \boldsymbol{\Psi}^{22}\boldsymbol{\Delta}_0)' \left(\boldsymbol{\Lambda}_0^{-1} \otimes \boldsymbol{\Psi}^{22-1}\right) \mathrm{vec}(\boldsymbol{\Psi}^{21} - \boldsymbol{\Psi}^{22}\boldsymbol{\Delta}_0) \right\} \\
&\propto g(\boldsymbol{\Psi}) \times |\boldsymbol{\Psi}^{22}|^{\frac{n_1-2p-1}{2}} \exp\left\{ -\frac{1}{2}\mathrm{tr}(\mathbf{S}_1^{-1}\boldsymbol{\Psi}^{22}) \right\} \\
&\quad \times |\boldsymbol{\Psi}^{22}|^{-\frac{p}{2}} \exp\left[-\frac{1}{2}\{\mathrm{vec}(\boldsymbol{\Psi}^{21} - \boldsymbol{\Psi}^{22}\boldsymbol{\Delta}_1)\}'(\boldsymbol{\Lambda}_1 \otimes \boldsymbol{\Psi}^{22})^{-1}\mathrm{vec}(\boldsymbol{\Psi}_{21} - \boldsymbol{\Psi}_{22}\boldsymbol{\Delta}_1) \right],
\end{aligned}
$$

where

$$
\begin{aligned}
n_1 &= n_0 + n - 1, \quad \mathbf{S}_1 = (\mathbf{S}_0^{-1} + \boldsymbol{\Xi}_{22} + \boldsymbol{\Delta}_0\boldsymbol{\Lambda}_0^{-1}\boldsymbol{\Delta}_0' - \boldsymbol{\Delta}_1\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\Delta}_1')^{-1}, \\
\boldsymbol{\Lambda}_1 &= (\boldsymbol{\Lambda}_0^{-1} + \boldsymbol{\Xi}_{11})^{-1}, \quad \boldsymbol{\Delta}_1 = (-\boldsymbol{\Xi}_{21} + \boldsymbol{\Delta}_0\boldsymbol{\Lambda}_0^{-1})\boldsymbol{\Lambda}_1,
\end{aligned}
$$

using $\mathrm{tr}(\mathbf{AB}) = \mathrm{vec}(\mathbf{A}')'\mathrm{vec}(\mathbf{B})$ and $\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}' \otimes \mathbf{A})\mathrm{vec}(\mathbf{X})$, for $\mathbf{X}(n \times n)$, $\mathbf{A}(m \times n)$ and $\mathbf{B}(n \times m)$ (see, e.g., Gupta and Nagar (2000), Ishihara, Omori, and Asai (2016)). We generate a candidate $\boldsymbol{\Psi}^\dagger$ in three steps:

1. Generate $\boldsymbol{\Psi}^{22\dagger} \sim \mathrm{W}(n_1, \mathbf{S}_1)$.
2. Generate $\mathrm{vec}(\boldsymbol{\Psi}^{21\dagger})|\boldsymbol{\Psi}^{22\dagger} \sim \mathrm{N}_{2p^2}(\mathrm{vec}(\boldsymbol{\Psi}^{22\dagger}\boldsymbol{\Delta}_1), \boldsymbol{\Lambda}_1 \otimes \boldsymbol{\Psi}^{22\dagger})$.
3. Compute $\boldsymbol{\Psi}_{21}^\dagger = -\boldsymbol{\Psi}^{22\dagger\,-1}\boldsymbol{\Psi}^{21\dagger}$, $\boldsymbol{\Psi}_{22}^\dagger = \boldsymbol{\Psi}^{22\dagger\,-1} + \boldsymbol{\Psi}_{21}^\dagger\boldsymbol{\Psi}_{21}^{\dagger\prime}$ and accept $\boldsymbol{\Psi}^\dagger$ with probability $\min[1, g(\boldsymbol{\Psi}^\dagger)/g(\boldsymbol{\Psi})]$.

4 Simulation study

This section gives illustrative examples using simulated data for our proposed model. We focus on three-block cases ($K = 3$) and use the following true parameters (with the subscript ‡) that are based on our empirical studies in Section 5 to generate 2,500 observations ($n = 2500$).

$$\begin{aligned}
\boldsymbol{\mu}_\ddagger = \boldsymbol{\xi}_\ddagger = \mathbf{0}_p, \quad \boldsymbol{\xi}^*_\ddagger = \mathbf{0}_q, \quad \boldsymbol{\gamma}_\ddagger &= (1.9, 0.5, 0.5, 1.9, 0.5, 1.9)', \\
\boldsymbol{\Phi}_\ddagger = 0.97\mathbf{I}_p, \quad \boldsymbol{\Theta}_\ddagger = 0.97\mathbf{I}_q, \quad \boldsymbol{\Omega}_{\boldsymbol{m}\ddagger} = 0.001\mathbf{I}_p, \quad \boldsymbol{\Delta}^*_\ddagger &= 0.1\mathbf{I}_q, \quad \boldsymbol{\Sigma}_\ddagger = 0.01\mathbf{I}_q, \\
\Psi_{\boldsymbol{ww}\ddagger} = 0.07\mathbf{I}_p + 0.03\mathbf{1}_p\mathbf{1}_p', \quad \Psi_{\boldsymbol{\eta\eta}\ddagger} &= 0.02\mathbf{I}_p + 0.04\mathbf{1}_p\mathbf{1}_p', \quad \Psi_{\boldsymbol{\eta w}\ddagger} = \mathbf{O}_{p\times p}, \\
\Psi_{\boldsymbol{w\varepsilon}\ddagger} = \Psi_{\boldsymbol{\eta\varepsilon}\ddagger} &= \left(-0.1\mathbf{1}_p,\ 0.05\mathbf{1}_p,\ -0.05\mathbf{1}_p,\ \mathbf{O}_{p\times(p-3)}\right).
\end{aligned}$$

For prior distributions, we assume

$$\begin{aligned}
&\boldsymbol{\mu} \sim \mathrm{N}_p(\boldsymbol{\mu}_\ddagger, 100\mathbf{I}_p), \quad \boldsymbol{\xi} \sim \mathrm{N}_p(\boldsymbol{\xi}_\ddagger, 100\mathbf{I}_p), \quad \boldsymbol{\xi}^* \sim \mathrm{N}_q(\boldsymbol{\xi}^*_\ddagger, 100\mathbf{I}_q), \quad \boldsymbol{\gamma} \sim \mathrm{N}_q(\boldsymbol{\gamma}_\ddagger, 100\mathbf{I}_q), \\
&\frac{\phi_i + 1}{2} \sim \mathrm{Be}(20, 1.5), \quad \frac{\theta_j + 1}{2} \sim \mathrm{Be}(20, 1.5), \quad \omega^2_{m_i 0} \sim \mathrm{IG}(5, 3\omega^2_{m_i\ddagger}), \quad \delta^{*2}_j \sim \mathrm{IG}(10^{-4}, 10^{-4}), \\
&\boldsymbol{\Sigma} \sim \mathrm{IW}_q(1, \boldsymbol{\Sigma}^{-1}_\ddagger), \quad \boldsymbol{\Psi}^{22} \sim \mathrm{W}_{2p}(2p, (2p\boldsymbol{\Psi}_{22\ddagger})^{-1}), \quad \mathrm{vec}(\boldsymbol{\Psi}^{21})|\boldsymbol{\Psi}^{22} \sim \mathrm{N}_{2p^2}(\mathbf{0}_{2p^2}, 10\mathbf{I}_p \otimes \boldsymbol{\Psi}^{22}),
\end{aligned}$$

for $i = 1, \ldots, p$, $j = 1, \ldots, q$ and set $\kappa = 10$ for the initial distribution of $\boldsymbol{m}_1$.

4.1 Example 1: Endogeneity and leverage

As the first example, we consider the case $p = 6$ where $p_1 = 3$, $p_2 = 2$, $p_3 = 1$ as in our empirical study. Since the third block consists of only one asset return, $\rho_{33,t}$ is not identified and we remove $g_{33,t}$ from $\boldsymbol{g}_t$. Thus $q = 5$ and we set $\boldsymbol{\gamma}_\ddagger = (1.9, 0.5, 0.5, 1.9, 0.5)'$.

We generate 100,000 MCMC samples after discarding the first 5,000 samples as the burn-in period. The multi-move sampler is used to generate $\{\boldsymbol{h}_t\}_{t=1}^n$ where we set the number of the blocks to 401 ($N = 400$), and the multiple-trial Metropolized independent sampler is used to generate $\{\boldsymbol{g}_t\}_{t=1}^n$ with the number of trials equal to 3 ($K_{\mathrm{MT}} = 3$). The acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{\boldsymbol{g}_t\}_{t=1}^n$ in the independent MH algorithms are 0.516 and 0.963, respectively.

Tables 1 and 2 show the posterior means, 95% credible intervals and estimated inefficiency factors (IF)[2] for $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, $\boldsymbol{\xi}$, $\boldsymbol{\xi}^*$, $\boldsymbol{\Omega}_{\boldsymbol{m}}$, $\boldsymbol{\Delta}^*$ and for diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{ww}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta w}}$,

[2]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.

$\boldsymbol{\Sigma}$. The posterior means are all close to the true values, and 95% credible intervals include true values except those of ω^2_{m5} and ω^2_{m6} which are very small numbers for the variance of the mean processes, suggesting that our proposed algorithm works well. For other parameters ($\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$, off-diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{ww}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta w}}$ and $\boldsymbol{\Sigma}$), similar estimation results are obtained and hence omitted.

Parameter estimates under misspecified model. We also investigate how parameter estimates are affected when the existence of the endogeneity is ignored. The parameters are re-estimated under the misspecified model where $\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}} = \boldsymbol{\Psi}_{\boldsymbol{\eta w}} = \mathbf{O}$ and $\boldsymbol{\Psi}_{\boldsymbol{ww}}$ is a diagonal matrix. Table 3 shows the posterior means, 95% credible intervals and estimated inefficiency factors for diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{ww}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$ and $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$. Their posterior means are not close to true values, and 95% credible intervals do not include the true values except for diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$ whose true values are zeros. The diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{ww}}$ are underestimated while those of $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$ are overestimated. The posterior means of the diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$ are all shrunk toward zero. Overall, parameter estimates of (diagonal and off-diagonal elements of) the covariance matrices for $\boldsymbol{\Psi}_{\boldsymbol{ww}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$ and $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$ are found to be biased when we ignore the existence of the endogeneity.

Further, to see how the estimates of the cross leverage effect is affected by the model misspecification, we compute the correlation between $\boldsymbol{y}_t$ and $\boldsymbol{h}_{t+1}$ for $t = 1, \ldots, n-1$ using the MCMC samples of $\boldsymbol{\Psi}_{\boldsymbol{\varepsilon\eta}}$, $\mathbf{R}_t$ and $\Psi_{\boldsymbol{\eta\eta}}$. Table 4 shows the sample means of the posterior means of the correlation between y_{it} and $h_{j,t+1}$ under both the true and misspecified models. Under the true model, they are around $-0.09 \sim -0.35$ indicating the existence of the cross leverage effect for all combinations. On the other hand, under the misspecified model, they are all shrunk toward zero suggesting no cross leverage effect. Thus, the misspecified model (ignoring the existence of the endogeneity) may lead us to underestimate the cross leverage effects. This is consistent with the simulation study of Chaussé and Xu (2016) for the univariate case.

Table 1: Example 1. Posterior means, 95% credible intervals and inefficiency factors for $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, $\boldsymbol{\xi}$, $\boldsymbol{\xi}^*$, $\boldsymbol{\Omega_m}$ and $\boldsymbol{\Delta}^*$.

Par.	True	Mean	95% interval	IF	Par.	True	Mean	95% interval	IF
μ_1	0	-0.255	(-0.627, 0.113)	11	ξ_1	0	0.049	(-0.008, 0.102)	372
μ_2	0	-0.066	(-0.399, 0.267)	16	ξ_2	0	0.032	(-0.026, 0.088)	433
μ_3	0	-0.104	(-0.446, 0.240)	11	ξ_3	0	0.045	(-0.009, 0.097)	363
μ_4	0	-0.045	(-0.305, 0.216)	20	ξ_4	0	0.003	(-0.057, 0.059)	426
μ_5	0	0.120	(-0.157, 0.399)	18	ξ_5	0	0.003	(-0.054, 0.060)	385
μ_6	0	-0.157	(-0.531, 0.215)	7	ξ_6	0	-0.011	(-0.068, 0.043)	259
γ_1	1.9	1.857	(1.709, 2.005)	148	ξ_1^*	0	0.014	(-0.049, 0.075)	795
γ_2	0.5	0.404	(0.279, 0.528)	69	ξ_2^*	0	0.023	(-0.027, 0.073)	423
γ_3	0.5	0.460	(0.307, 0.609)	70	ξ_3^*	0	0.040	(-0.018, 0.101)	462
γ_4	1.9	1.889	(1.717, 2.064)	139	ξ_4^*	0	-0.029	(-0.109, 0.047)	691
γ_5	0.5	0.453	(0.325, 0.580)	111	ξ_5^*	0	0.049	(-0.013, 0.113)	467
ϕ_1	0.97	0.974	(0.967, 0.981)	13	$\omega_{m_1}^2 \times 10^3$	1	1.129	(0.745, 1.661)	88
ϕ_2	0.97	0.971	(0.964, 0.978)	16	$\omega_{m_2}^2 \times 10^3$	1	1.275	(0.810, 1.897)	133
ϕ_3	0.97	0.972	(0.965, 0.979)	21	$\omega_{m_3}^2 \times 10^3$	1	0.922	(0.527, 1.500)	181
ϕ_4	0.97	0.963	(0.954, 0.972)	17	$\omega_{m_4}^2 \times 10^3$	1	1.189	(0.747, 1.788)	127
ϕ_5	0.97	0.966	(0.958, 0.973)	20	$\omega_{m_5}^2 \times 10^3$	1	0.442	(0.214, 0.816)	232
ϕ_6	0.97	0.974	(0.967, 0.981)	20	$\omega_{m_6}^2 \times 10^3$	1	0.573	(0.325, 0.943)	221
θ_1	0.97	0.969	(0.956, 0.981)	17	δ_1^{*2}	0.1	0.098	(0.091, 0.106)	10
θ_2	0.97	0.963	(0.948, 0.976)	20	δ_2^{*2}	0.1	0.096	(0.089, 0.103)	14
θ_3	0.97	0.968	(0.954, 0.980)	22	δ_3^{*2}	0.1	0.099	(0.091, 0.106)	17
θ_4	0.97	0.971	(0.958, 0.982)	28	δ_4^{*2}	0.1	0.100	(0.093, 0.108)	21
θ_5	0.97	0.961	(0.946, 0.974)	23	δ_5^{*2}	0.1	0.095	(0.088, 0.102)	14

Table 2: Example 1. Posterior means, 95% credible intervals and inefficiency factors for diagonal elements of $\boldsymbol{\Psi_{ww}}$, $\boldsymbol{\Psi_{\eta\eta}}$, $\boldsymbol{\Psi_{\eta w}}$ and $\boldsymbol{\Sigma}$.

Par.	True	Mean	95% interval	IF	Par.	True	Mean	95% interval	IF
$\boldsymbol{\Psi_{ww}}_{11}$	0.1	0.106	(0.091, 0.122)	200	$\boldsymbol{\Psi_{\eta\eta}}_{11}$	0.06	0.060	(0.053, 0.069)	39
$\boldsymbol{\Psi_{ww}}_{22}$	0.1	0.093	(0.080, 0.107)	186	$\boldsymbol{\Psi_{\eta\eta}}_{22}$	0.06	0.062	(0.054, 0.070)	53
$\boldsymbol{\Psi_{ww}}_{33}$	0.1	0.090	(0.078, 0.103)	146	$\boldsymbol{\Psi_{\eta\eta}}_{33}$	0.06	0.061	(0.053, 0.069)	54
$\boldsymbol{\Psi_{ww}}_{44}$	0.1	0.109	(0.093, 0.126)	221	$\boldsymbol{\Psi_{\eta\eta}}_{44}$	0.06	0.060	(0.052, 0.069)	61
$\boldsymbol{\Psi_{ww}}_{55}$	0.1	0.102	(0.087, 0.118)	312	$\boldsymbol{\Psi_{\eta\eta}}_{55}$	0.06	0.061	(0.053, 0.069)	60
$\boldsymbol{\Psi_{ww}}_{66}$	0.1	0.096	(0.081, 0.112)	314	$\boldsymbol{\Psi_{\eta\eta}}_{66}$	0.06	0.063	(0.055, 0.072)	49
$\boldsymbol{\Psi_{\eta w}}_{11}$	0	-0.001	(-0.014, 0.013)	261	$\boldsymbol{\Sigma}_{11}$	0.01	0.010	(0.008, 0.013)	47
$\boldsymbol{\Psi_{\eta w}}_{22}$	0	-0.009	(-0.019, 0.003)	246	$\boldsymbol{\Sigma}_{22}$	0.01	0.011	(0.008, 0.014)	58
$\boldsymbol{\Psi_{\eta w}}_{33}$	0	-0.009	(-0.018, 0.003)	200	$\boldsymbol{\Sigma}_{33}$	0.01	0.012	(0.009, 0.016)	64
$\boldsymbol{\Psi_{\eta w}}_{44}$	0	0.000	(-0.013, 0.015)	258	$\boldsymbol{\Sigma}_{44}$	0.01	0.012	(0.009, 0.016)	78
$\boldsymbol{\Psi_{\eta w}}_{55}$	0	0.002	(-0.011, 0.017)	358	$\boldsymbol{\Sigma}_{55}$	0.01	0.011	(0.009, 0.015)	61
$\boldsymbol{\Psi_{\eta w}}_{66}$	0	-0.004	(-0.016, 0.011)	372					

Table 3: Example 1. Misspecified model. Posterior means, 95% credible intervals and inefficiency factors for diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{ww}}$, $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$ and $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$.

Par.	True	Mean	95% interval	IF	Par.	True	Mean	95% interval	IF
$\boldsymbol{\Psi}_{\boldsymbol{ww}11}$	0.1	0.076	(0.070, 0.082)	30	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}11}$	0.06	0.105	(0.095, 0.116)	43
$\boldsymbol{\Psi}_{\boldsymbol{ww}22}$	0.1	0.073	(0.067, 0.079)	27	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}22}$	0.06	0.103	(0.093, 0.113)	37
$\boldsymbol{\Psi}_{\boldsymbol{ww}33}$	0.1	0.072	(0.066, 0.078)	29	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}33}$	0.06	0.097	(0.088, 0.107)	33
$\boldsymbol{\Psi}_{\boldsymbol{ww}44}$	0.1	0.075	(0.069, 0.082)	42	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}44}$	0.06	0.103	(0.093, 0.114)	43
$\boldsymbol{\Psi}_{\boldsymbol{ww}55}$	0.1	0.076	(0.070, 0.082)	20	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}55}$	0.06	0.094	(0.085, 0.103)	35
$\boldsymbol{\Psi}_{\boldsymbol{ww}66}$	0.1	0.072	(0.067, 0.078)	22	$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}66}$	0.06	0.103	(0.093, 0.113)	38
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}11}$	-0.1	-0.021	(-0.037, -0.006)	10					
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}22}$	0.05	-0.007	(-0.024, 0.010)	17					
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}33}$	-0.05	0.002	(-0.017, 0.020)	17					
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}44}$	0	0.004	(-0.012, 0.021)	12					
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}55}$	0	-0.017	(-0.033, -0.001)	15					
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}66}$	0	-0.014	(-0.030, 0.002)	8					

Table 4: Means of the correlation between y_{it} and $h_{j,t+1}$.

	True model					
	$h_{1,t+1}$	$h_{2,t+1}$	$h_{3,t+1}$	$h_{4,t+1}$	$h_{5,t+1}$	$h_{6,t+1}$
y_{1t}	-0.317	-0.337	-0.267	-0.266	-0.347	-0.282
y_{2t}	-0.287	-0.305	-0.246	-0.226	-0.299	-0.269
y_{3t}	-0.290	-0.314	-0.236	-0.275	-0.339	-0.291
y_{4t}	-0.257	-0.229	-0.275	-0.233	-0.290	-0.296
y_{5t}	-0.275	-0.255	-0.272	-0.271	-0.295	-0.273
y_{6t}	-0.118	-0.116	-0.093	-0.095	-0.113	-0.124

	Misspecified model					
	$h_{1,t+1}$	$h_{2,t+1}$	$h_{3,t+1}$	$h_{4,t+1}$	$h_{5,t+1}$	$h_{6,t+1}$
y_{1t}	-0.058	-0.088	-0.030	-0.040	-0.103	-0.041
y_{2t}	-0.038	-0.070	-0.010	-0.011	-0.066	-0.027
y_{3t}	-0.054	-0.075	-0.018	-0.050	-0.102	-0.057
y_{4t}	-0.035	-0.028	-0.058	-0.030	-0.068	-0.073
y_{5t}	-0.028	-0.027	-0.038	-0.038	-0.061	-0.041
y_{6t}	-0.033	-0.039	-0.024	-0.023	-0.043	-0.044

4.2 Example 2: Higher dimensional case

Next, we consider the higher dimensional case $p = 30$ with $p_1 = p_2 = p_3 = 10$ and $q = 6$. As in the previous subsection, we generate 100,000 MCMC samples after discarding the first 5,000 samples as the burn-in period. The single-move sampler is used to generate $\{\boldsymbol{h}_t\}_{t=1}^n$, and the multiple-trial Metropolized independent sampler is used to generate $\{\boldsymbol{g}_t\}_{t=1}^n$ with the number of trials equal to 4 ($K_{\text{MT}} = 4$). The acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{\boldsymbol{g}_t\}_{t=1}^n$ in the independent MH algorithms are 0.925 and 0.939, respectively.

Tables 5 shows the posterior means, 95% credible intervals and estimated inefficiency factors for the selected model parameters. The posterior means are all close to the true values, suggesting that our proposed algorithm works well. For other parameters, similar estimation results are obtained and hence omitted.

Table 5: Example 2. Posterior means, 95% credible intervals and inefficiency factors for the selected model parameters.

	True	Mean	95% interval	IF
μ_1	0	-0.168	(-0.459, 0.130)	15
γ_6	1.9	1.909	(1.739, 2.078)	44
ξ_1	0	-0.000	(-0.051, 0.049)	490
ξ_6^*	0	-0.001	(-0.048, 0.044)	630
ϕ_1	0.97	0.967	(0.960, 0.974)	15
θ_6	0.97	0.973	(0.962, 0.984)	9
$\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}11}$	-0.1	-0.094	(-0.112, -0.077)	37
$\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}11}$	-0.1	-0.105	(-0.122, -0.088)	50
$\boldsymbol{\Psi}_{\boldsymbol{ww}11}$	0.1	0.108	(0.097, 0.119)	73
$\boldsymbol{\Psi}_{\boldsymbol{\eta w}11}$	0	0.007	(-0.001, 0.017)	110
$\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}11}$	0.06	0.068	(0.060, 0.076)	34
δ_6^{*2}	0.1	0.102	(0.095, 0.110)	7
$\boldsymbol{\Sigma}_{66}$	0.01	0.011	(0.009, 0.014)	32
$\omega_{m_1}^2 \times 10^3$	1	1.158	(0.753, 1.713)	106

IF of ξ_6^* is the maximum among all parameters (including the omitted parameters).

5 Empirical study

5.1 Data

We use daily close-to-close returns and their realized measures for some of the most liquid stocks in the Dow Jones Industrial Average (DJIA) index, obtained from Oxford-Man Institute[3]. These are: (1) Bank of America (BAC), (2) JP Morgan (JPM), (3) American Express (AXP), (4) International Business Machines (IBM), (5) Microsoft (MSFT), (6) Coca Cola (KO). The sample period is from February 1, 2001 to December 31, 2009[4]. Taking account of the sample correlations among the six series, we divide them into three groups: Group 1 (Finance) consists of BAC, JPM and AXP, Group 2 (Information Technology) consists of IBM and MSFT, and Group 3 (Food) consists of KO. Thus we apply our proposed model to this dataset with three blocks ($K = 3$). Note that each block corresponds to the industrial sector and that the block sizes are three, two and one ($p_1 = 3, p_2 = 2, p_3 = 1,\ p = 6,\ q = 5$) respectively.

5.2 Estimation results

Assuming the same prior distributions for parameters as in Section 4, we implement the MCMC algorithm to conduct the inference on the parameters of interest where the multi-move sampler and the multiple-trial Metropolized independent sampler are used as in Section 4.1. We generate 100,000 MCMC samples from the posterior distributions of the parameters in the model after discarding the first 30,000 samples as the burn-in period. The acceptance rates in the independent MH algorithms for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{\boldsymbol{g}_t\}_{t=1}^n$ are 0.421 and 0.953, which indicates that generated candidates are accepted with relatively high probability and our sampling algorithm works well. Table 6 reports posterior means, 95% credible intervals and inefficiency factors for $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, $\boldsymbol{\xi}$, $\boldsymbol{\xi}^*$, $\boldsymbol{\Omega}_{\boldsymbol{m}}$ and $\boldsymbol{\Delta}^*$.

5.2.1 Bias correction of realized measures

The posterior probability with which the bias adjustment term ξ_i is negative is greater than 0.975 for $i = 1, \ldots, 6$. This suggests that the realized measure of variances tends to have a

[3]See Noureldin, Shephard, and Sheppard (2012) for the detailed explanation. The asset return is calculated as $y_t = (\log p_t - \log p_{t-1}) \times 100$, where p_t is the asset price at time t.

[4]There are 2,241 observations in total. We remove the observation for October 3, 2005 from the dataset since it is considered as an outlier.

downward bias due to the non-trading hours and that the effect of non-trading hours is more influential than that of microstructure noises. In a similar manner, the posterior distributions of ξ_j^* suggest that the realized measure of the correlation tends to have a downward bias as described in Section 1 (Epps effect). These estimation results show that our model can detect and adjust the bias of the realized measures due to various problems in the financial market.

Table 6: U.S. stock returns. Posterior means, 95% credible intervals and inefficiency factors for $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, $\boldsymbol{\xi}$, $\boldsymbol{\xi}^*$, $\boldsymbol{\Omega_m}$ and $\boldsymbol{\Delta}^*$.

Par.	Mean	95% interval	IF	Par.	Mean	95% interval	IF
μ_1	0.766	(0.459, 1.068)	24	ξ_1	-0.474	(-0.536, -0.413)	623
μ_2	1.107	(0.846, 1.365)	29	ξ_2	-0.419	(-0.480, -0.360)	599
μ_3	1.058	(0.779, 1.333)	21	ξ_3	-0.520	(-0.580, -0.462)	471
μ_4	0.547	(0.367, 0.726)	73	ξ_4	-0.461	(-0.522, -0.397)	628
μ_5	0.872	(0.685, 1.059)	40	ξ_5	-0.502	(-0.561, -0.445)	450
μ_6	0.107	(-0.068, 0.280)	53	ξ_6	-0.232	(-0.295, -0.170)	421
γ_1	1.678	(1.592, 1.764)	300	ξ_1^*	-0.627	(-0.691, -0.563)	536
γ_2	1.066	(0.999, 1.132)	879	ξ_2^*	-0.248	(-0.307, -0.187)	1091
γ_3	0.844	(0.769, 0.917)	802	ξ_3^*	-0.182	(-0.250, -0.111)	924
γ_4	1.309	(1.214, 1.399)	608	ξ_4^*	-0.327	(-0.411, -0.239)	694
γ_5	0.798	(0.721, 0.876)	718	ξ_5^*	-0.096	(-0.169, -0.024)	813
ϕ_1	0.960	(0.952, 0.967)	67	$\omega_{m_1}^2 \times 10^3$	0.225	(0.122, 0.409)	269
ϕ_2	0.953	(0.944, 0.961)	83	$\omega_{m_2}^2 \times 10^3$	0.233	(0.118, 0.441)	299
ϕ_3	0.959	(0.952, 0.967)	78	$\omega_{m_3}^2 \times 10^3$	0.231	(0.119, 0.433)	222
ϕ_4	0.938	(0.926, 0.948)	98	$\omega_{m_4}^2 \times 10^3$	0.232	(0.118, 0.431)	236
ϕ_5	0.942	(0.930, 0.952)	89	$\omega_{m_5}^2 \times 10^3$	0.243	(0.123, 0.457)	212
ϕ_6	0.940	(0.929, 0.951)	105	$\omega_{m_6}^2 \times 10^3$	0.234	(0.113, 0.433)	240
θ_1	0.808	(0.777, 0.837)	67	δ_1^{*2}	0.026	(0.022, 0.031)	63
θ_2	0.679	(0.646, 0.712)	64	δ_2^{*2}	0.012	(0.010, 0.014)	72
θ_3	0.613	(0.570, 0.654)	122	δ_3^{*2}	0.019	(0.015, 0.023)	157
θ_4	0.649	(0.608, 0.690)	83	δ_4^{*2}	0.042	(0.036, 0.048)	102
θ_5	0.614	(0.575, 0.653)	50	δ_5^{*2}	0.028	(0.024, 0.032)	72

Table 7 shows the estimation results for the covariance matrix $\boldsymbol{\Psi_{ww}}$. The posterior means of off-diagonal elements of $\boldsymbol{\Psi_{ww}}$ are all positive and the posterior probability of the positive covariance between w_{it} and w_{jt} $(i \neq j)$ is greater than 0.975 for all i and j $(i \neq j)$. That is, all measurement errors of the realized volatilities are positively correlated.

Table 7: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Psi_{ww}}$.

$$\begin{bmatrix}
\mathbf{0.106} \\ (0.092, 0.121) & & & & & \\
\mathbf{0.049} & \mathbf{0.115} & & & & \\ (0.037, 0.061) & (0.100, 0.130) & & & & \\
\mathbf{0.030} & \mathbf{0.037} & \mathbf{0.116} & & & \\ (0.019, 0.041) & (0.025, 0.049) & (0.102, 0.132) & & & \\
\mathbf{0.044} & \mathbf{0.045} & \mathbf{0.038} & \mathbf{0.124} & & \\ (0.033, 0.055) & (0.034, 0.057) & (0.027, 0.050) & (0.109, 0.141) & & \\
\mathbf{0.041} & \mathbf{0.045} & \mathbf{0.035} & \mathbf{0.050} & \mathbf{0.113} & \\ (0.031, 0.052) & (0.034, 0.056) & (0.024, 0.046) & (0.039, 0.062) & (0.099, 0.127) & \\
\mathbf{0.035} & \mathbf{0.029} & \mathbf{0.024} & \mathbf{0.030} & \mathbf{0.026} & \mathbf{0.101} \\ (0.025, 0.046) & (0.019, 0.040) & (0.013, 0.035) & (0.019, 0.042) & (0.016, 0.036) & (0.087, 0.115)
\end{bmatrix}$$

*The (i, j)-th element corresponds to that of $\boldsymbol{\Psi_{ww}}$. Bold figures indicate that the 95% credible interval does not include zero.

5.2.2 Volatility clustering

In Table 6, the posterior means of autoregressive coefficients (ϕ_i's) for the h_{it}'s are all found to be very high ($0.93 \sim 0.96$), which implies that log volatilities are highly persistent. Figure 1 shows the time series plot of the posterior means of $\exp(h_{it}/2)$, the square root of the estimated time-varying variances. These trajectories sharply increased in September 2008, corresponding to the financial crisis during which Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008). We also observe the increase in July 2002 resulted from the market turmoil during which Worldcom filed for Chapter 11 bankruptcy protection (July 21, 2002), while those increases in April 2001 and in September 2001 are due to the collapse of the dot-com bubble and to the September 11 attacks, respectively.

Further, we note that Figure 1 indicates the co-movement of the volatilities. In fact, as shown in Table 8, the posterior means of off-diagonal elements of $\boldsymbol{\Psi_{\eta\eta}}$ are all positive and the posterior probability of the positive covariance between h_{it} and h_{jt} $(i \neq j)$ is greater than 0.975 for all i and j $(i \neq j)$. That is, all log volatilities are positively correlated.

Figure 1: Time series plot of posterior means of $\exp(h_{it}/2)$'s.

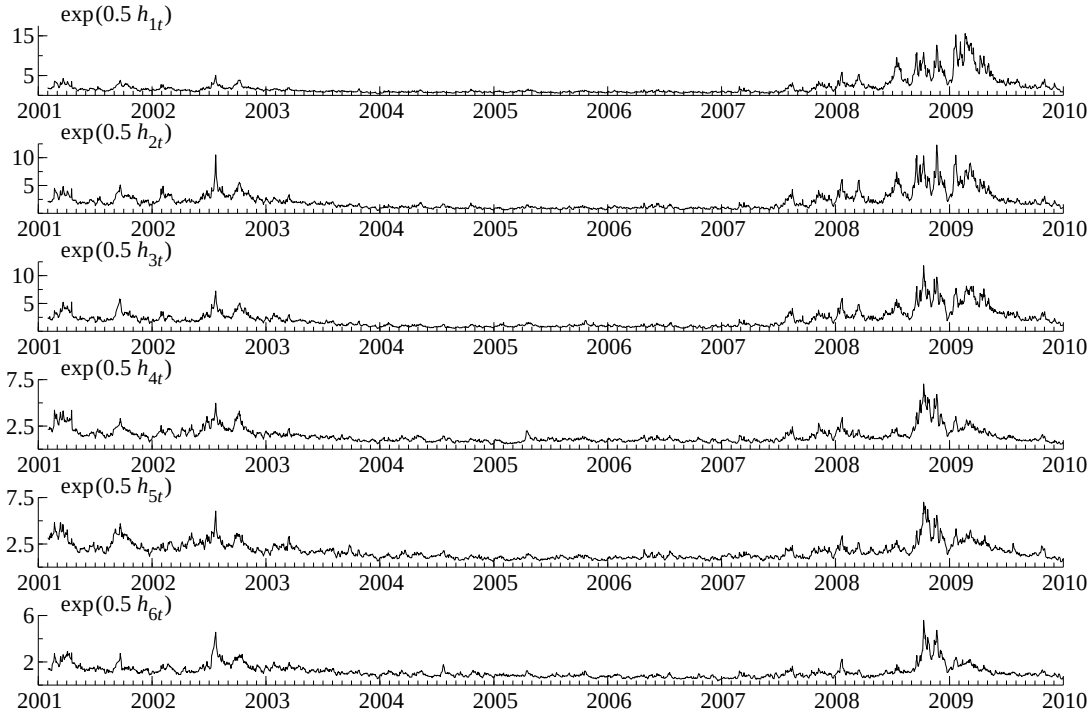


Table 8: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$.

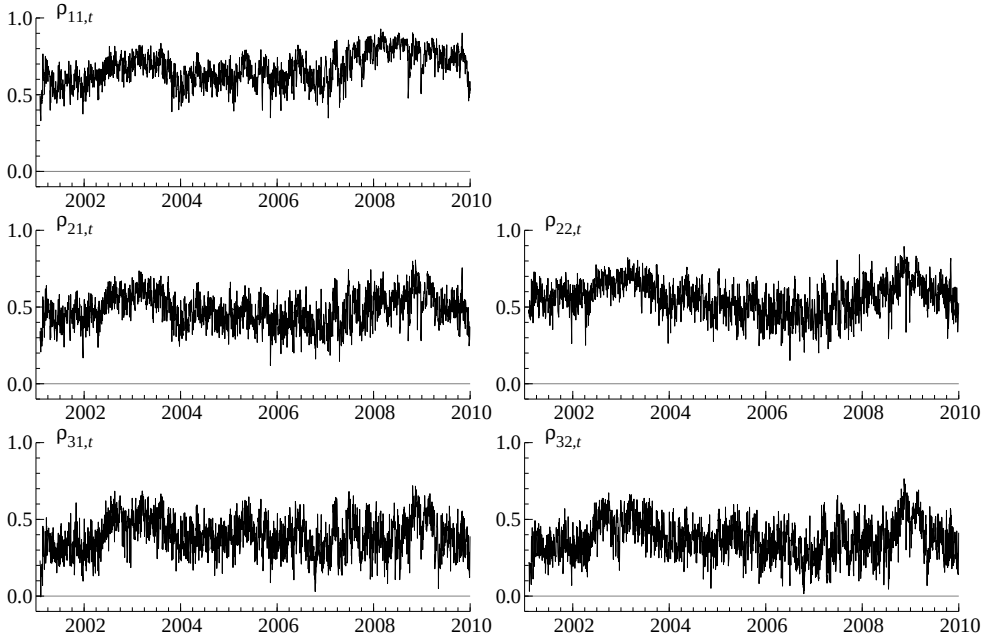
0.087					
(0.075, 0.101)					
0.078	**0.086**				
(0.067, 0.090)	(0.073, 0.100)				
0.069	**0.070**	**0.073**			
(0.058, 0.080)	(0.059, 0.081)	(0.062, 0.086)			
0.054	**0.061**	**0.053**	**0.067**		
(0.045, 0.064)	(0.051, 0.072)	(0.044, 0.063)	(0.055, 0.079)		
0.051	**0.056**	**0.051**	**0.052**	**0.065**	
(0.042, 0.061)	(0.046, 0.067)	(0.042, 0.061)	(0.043, 0.063)	(0.053, 0.078)	
0.047	**0.053**	**0.049**	**0.048**	**0.045**	**0.057**
(0.039, 0.057)	(0.044, 0.062)	(0.040, 0.058)	(0.039, 0.057)	(0.036, 0.054)	(0.047, 0.068)

*The (i, j)-th element corresponds to that of $\boldsymbol{\Psi}_{\boldsymbol{\eta\eta}}$. Bold figures indicate that the 95% credible interval does not include zero.

5.2.3 Multiple-block dynamic equicorrelation

The averages of the posterior means of the block equicorrelations $(\rho_{11,t}, \rho_{21,t}, \rho_{31,t}, \rho_{22,t}, \rho_{32,t})$ are $(0.670, 0.480, 0.391, 0.562, 0.371)'$ [5] . The posterior means of autoregressive coefficients (θ_j's) for $\boldsymbol{g}_t$'s are high ($0.61 \sim 0.81$) and the block equicorrelations are found to be persistent. In Figure 2, these equicorrelations are found to be not constant but time-varying as expected. The diagonal-block equicorrelations, $\rho_{11,t}$ and $\rho_{22,t}$, for Groups 1 and 2, fluctuate at higher levels than those of off-diagonal block equicorrelations, which seems to be consistent with those posterior means of $\boldsymbol{\gamma}$ reported in Table 6. The posterior means of γ_1 and γ_4 (corresponding to $\rho_{11,t}$ and $\rho_{22,t}$) are 1.678 and 1.309, and they are larger than those of γ_2, γ_3 and γ_5 (corresponding to $\rho_{21,t}$, $\rho_{31,t}$ and $\rho_{32,t}$).

Figure 2: Time series plot of posterior means of the $\rho_{ij,t}$'s.



Also, as shown in Table 9, the 95% credible intervals of off-diagonal elements of $\boldsymbol{\Sigma}$ are all positive and do not include zero, which means unobserved multiple-block equicorrelation factors are correlated positively with each other.

[5] We note that $\boldsymbol{\gamma}$ is not the exact mean of $\{\boldsymbol{g}_t\}_{t=1}^n$ though the obtained posterior means of those are very similar.

Table 9: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Sigma}$.

$$\begin{bmatrix} \mathbf{0.070} & & & & \\ (0.061, 0.079) & & & & \\ \mathbf{0.050} & \mathbf{0.053} & & & \\ (0.046, 0.056) & (0.048, 0.058) & & & \\ \mathbf{0.041} & \mathbf{0.044} & \mathbf{0.064} & & \\ (0.036, 0.046) & (0.040, 0.049) & (0.057, 0.071) & & \\ \mathbf{0.039} & \mathbf{0.056} & \mathbf{0.039} & \mathbf{0.081} & \\ (0.033, 0.044) & (0.051, 0.061) & (0.033, 0.044) & (0.071, 0.091) & \\ \mathbf{0.033} & \mathbf{0.044} & \mathbf{0.058} & \mathbf{0.053} & \mathbf{0.064} \\ (0.028, 0.038) & (0.040, 0.049) & (0.053, 0.063) & (0.047, 0.058) & (0.057, 0.071) \end{bmatrix}$$

*The (i, j)-th element corresponds to that of $\boldsymbol{\Sigma}$. Bold figures indicate that the 95% credible interval does not include zero.

5.2.4 Endogeneity and leverage

Figure 3 shows the posterior means of dynamic correlations between the return of i-th asset at time t (y_{it}) and the j-th log variance at time $t+1$ ($h_{j,t+1}$). The sample means of the posterior means also shown in Table 10. The trajectories of these correlations are negative at almost all of the time points as expected and far below zero, which indicates the existence of the cross leverage effects. We also estimated our proposed model without considering the endogeneity, where $\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}} = \boldsymbol{\Psi}_{\boldsymbol{\eta w}} = \mathbf{O}$ and $\boldsymbol{\Psi}_{\boldsymbol{ww}}$ is a diagonal matrix. The sample means of the posterior means of the dynamic correlations under the restricted model are shown in Table 11, and all cross leverage effects are shrunk toward zero. This is consistent with the result obtained in our simulation study in Section 4.1.

Table 12 shows the posterior means and the 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$. The covariances in the first column are all negative and their posterior probability of the negative correlation is greater than 0.975. The cross leverage effect seems to exist through the first element of $\boldsymbol{\varepsilon}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$ since most covariances in other columns are around zeros. In most cases, the first eigenvalue is the largest among all eigenvalues, and hence the first element of $\boldsymbol{\varepsilon}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$ is the first principal component, probably representing the market factor.

Figure 3: Posterior means of the correlation between y_{it} and $h_{j,t+1}$.

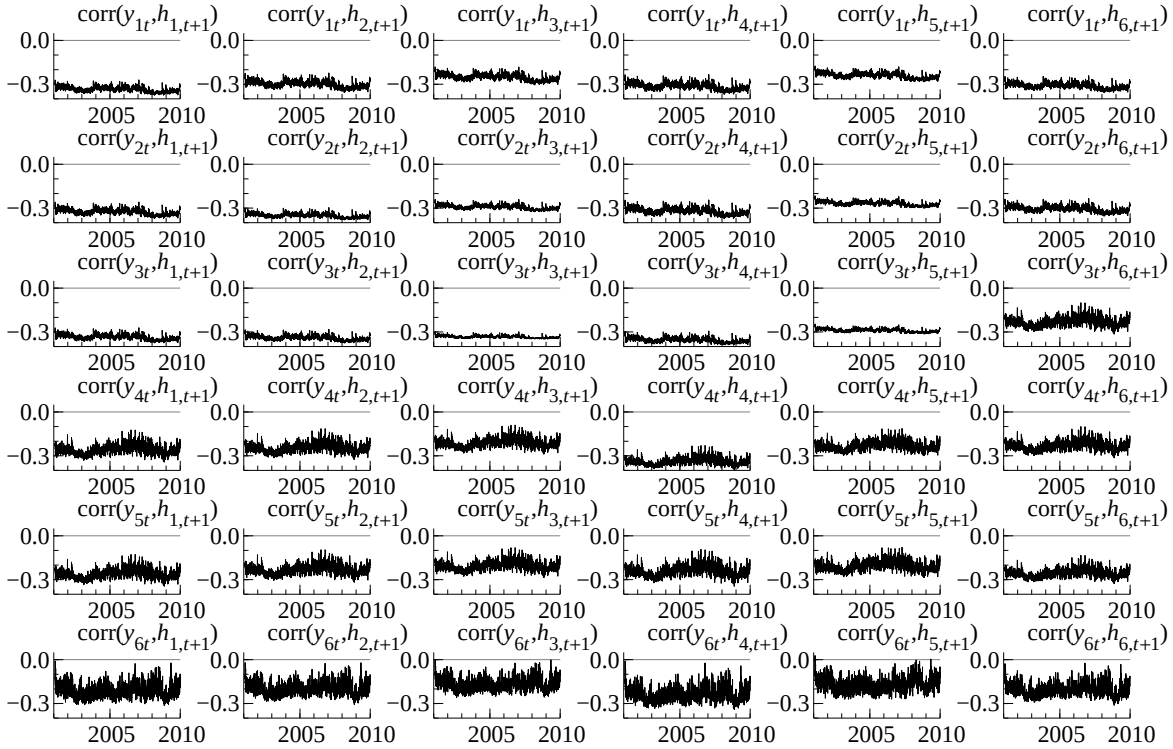


Table 10: Sample means of posterior means of the correlation between y_{it} and $h_{j,t+1}$.

	$h_{1,t+1}$	$h_{2,t+1}$	$h_{3,t+1}$	$h_{4,t+1}$	$h_{5,t+1}$	$h_{6,t+1}$
y_{1t}	-0.335	-0.300	-0.252	-0.314	-0.240	-0.311
y_{2t}	-0.327	-0.351	-0.293	-0.321	-0.269	-0.306
y_{3t}	-0.336	-0.341	-0.334	-0.356	-0.290	-0.343
y_{4t}	-0.260	-0.248	-0.214	-0.340	-0.233	-0.233
y_{5t}	-0.263	-0.234	-0.208	-0.242	-0.208	-0.257
y_{6t}	-0.213	-0.198	-0.167	-0.238	-0.170	-0.206

Table 11: Sample means of posterior means of the correlation between y_{it} and $h_{j,t+1}$. Model without endogeneity.

	$h_{1,t+1}$	$h_{2,t+1}$	$h_{3,t+1}$	$h_{4,t+1}$	$h_{5,t+1}$	$h_{6,t+1}$
y_{1t}	-0.124	-0.091	-0.076	-0.107	-0.073	-0.138
y_{2t}	-0.116	-0.121	-0.100	-0.099	-0.076	-0.132
y_{3t}	-0.137	-0.127	-0.138	-0.136	-0.098	-0.166
y_{4t}	-0.101	-0.089	-0.077	-0.161	-0.100	-0.106
y_{5t}	-0.136	-0.116	-0.100	-0.117	-0.109	-0.141
y_{6t}	-0.095	-0.082	-0.068	-0.104	-0.069	-0.105

Table 12: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$.

$$\begin{bmatrix}
\mathbf{-0.114} & -0.013 & -0.009 & 0.000 & -0.004 & 0.001 \\
(-0.136, -0.091) & (-0.036, 0.009) & (-0.041, 0.024) & (-0.019, 0.020) & (-0.026, 0.018) & (-0.020, 0.023) \\
\mathbf{-0.109} & -0.017 & -0.019 & 0.015 & 0.013 & -0.004 \\
(-0.131, -0.087) & (-0.040, 0.005) & (-0.051, 0.015) & (-0.005, 0.035) & (-0.009, 0.035) & (-0.025, 0.016) \\
\mathbf{-0.089} & -0.020 & -0.010 & \mathbf{0.028} & 0.000 & -0.002 \\
(-0.110, -0.067) & (-0.041, 0.003) & (-0.042, 0.021) & (0.009, 0.047) & (-0.021, 0.021) & (-0.022, 0.018) \\
\mathbf{-0.103} & 0.000 & -0.001 & 0.014 & -0.005 & \mathbf{-0.027} \\
(-0.124, -0.082) & (-0.021, 0.021) & (-0.030, 0.029) & (-0.005, 0.032) & (-0.026, 0.015) & (-0.046, -0.008) \\
\mathbf{-0.080} & -0.014 & 0.007 & 0.016 & 0.001 & -0.007 \\
(-0.101, -0.059) & (-0.035, 0.009) & (-0.026, 0.041) & (-0.003, 0.035) & (-0.020, 0.023) & (-0.026, 0.012) \\
\mathbf{-0.087} & -0.008 & -0.014 & 0.009 & -0.007 & 0.006 \\
(-0.106, -0.068) & (-0.028, 0.011) & (-0.043, 0.016) & (-0.008, 0.027) & (-0.027, 0.012) & (-0.012, 0.025)
\end{bmatrix}$$

*The (i, j)-th element corresponds to that of $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$. Bold figures indicate that the 95% credible interval does not include zero.

Table 13: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}$.

$$\begin{bmatrix}
\mathbf{-0.070} & -0.014 & -0.017 & \mathbf{-0.021} & 0.001 & -0.015 \\
(-0.091, -0.049) & (-0.037, 0.008) & (-0.048, 0.016) & (-0.040, -0.002) & (-0.021, 0.023) & (-0.035, 0.006) \\
\mathbf{-0.057} & \mathbf{-0.028} & \mathbf{-0.043} & -0.004 & 0.013 & -0.016 \\
(-0.078, -0.037) & (-0.052, -0.002) & (-0.074, -0.009) & (-0.024, 0.015) & (-0.008, 0.035) & (-0.036, 0.004) \\
\mathbf{-0.035} & -0.015 & 0.000 & 0.006 & 0.001 & -0.007 \\
(-0.056, -0.014) & (-0.039, 0.009) & (-0.034, 0.034) & (-0.014, 0.025) & (-0.021, 0.023) & (-0.027, 0.013) \\
\mathbf{-0.072} & 0.007 & -0.026 & -0.001 & 0.008 & \mathbf{-0.026} \\
(-0.092, -0.051) & (-0.015, 0.031) & (-0.056, 0.005) & (-0.019, 0.018) & (-0.013, 0.029) & (-0.046, -0.007) \\
\mathbf{-0.039} & -0.016 & -0.032 & 0.011 & \mathbf{0.024} & -0.012 \\
(-0.059, -0.019) & (-0.039, 0.010) & (-0.064, 0.002) & (-0.008, 0.030) & (0.002, 0.045) & (-0.032, 0.007) \\
\mathbf{-0.051} & -0.003 & -0.015 & -0.006 & -0.003 & 0.004 \\
(-0.070, -0.032) & (-0.025, 0.020) & (-0.045, 0.015) & (-0.024, 0.012) & (-0.023, 0.017) & (-0.015, 0.022)
\end{bmatrix}$$

*The (i, j)-th element corresponds to that of $\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}$. Bold figures indicate that the 95% credible interval does not include zero.

Table 13 shows the posterior means and the 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}}$. As in the estimation results for $\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}}$, the covariances in the first column are all negative and their posterior probabilities of the negative correlation are greater than 0.975. Since other covariances in other columns are around zeros, the measurement errors of the log realized volatilities are negatively correlated with the asset returns through the first element of $\boldsymbol{\varepsilon}_t = \mathbf{R}_t^{-1/2}\mathbf{V}_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t)$.

Figure 4 shows the posterior means of dynamic correlations between the return of i-th asset at time t (y_{it}) and the j-th log realized measure of the variance at time t (x_{jt}). The sample means of the posterior means also shown in Table 14. The most trajectories of these correlations are negative at almost all of the time points while some of them fluctuate around zero, which indicates the existence of the negative (sometimes very weak or no) correlation between the two quantities. This is consistent with the empirical study of Peters and de Vilder (2006) or Chaussé and Xu (2016) for the univariate case.

Figure 4: Posterior means of the correlation between y_{it} and x_{jt}.

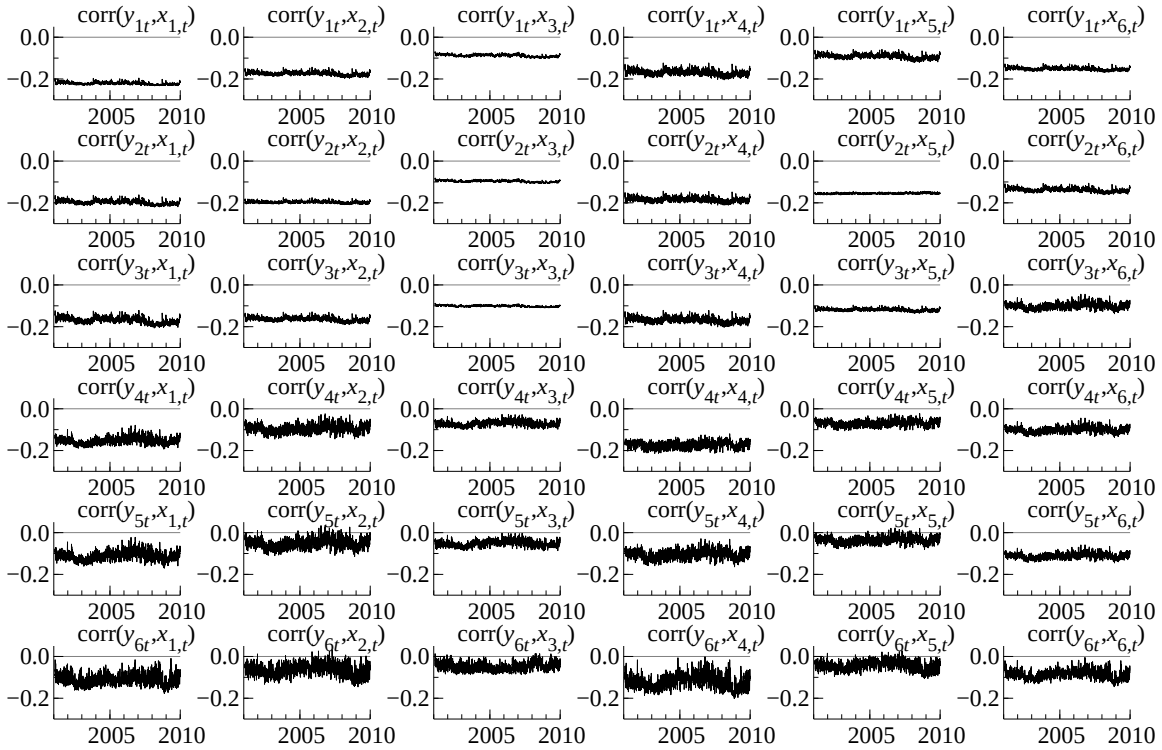


Table 14: Sample means of posterior means of the correlation between y_{it} and x_{jt}.

	x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}
y_{1t}	-0.222	-0.175	-0.088	-0.170	-0.092	-0.152
y_{2t}	-0.198	-0.197	-0.097	-0.185	-0.155	-0.139
y_{3t}	-0.170	-0.165	-0.102	-0.169	-0.119	-0.137
y_{4t}	-0.154	-0.095	-0.071	-0.173	-0.069	-0.101
y_{5t}	-0.112	-0.051	-0.053	-0.104	-0.035	-0.111
y_{6t}	-0.111	-0.065	-0.049	-0.122	-0.044	-0.087

For $\boldsymbol{\Psi}_{\boldsymbol{\eta w}}$, the posterior means and the 95% credible intervals are shown in Table 15. The diagonal elements of covariances are all negative and most of their posterior probabilities of the negative correlation are greater than 0.975. Other covariances of off-diagonal elements are around zeros and the measurement error of the log realized volatility is negatively correlated only with the corresponding latent log volatility.

Table 15: U.S. stock returns. Posterior means with 95% credible intervals for $\boldsymbol{\Psi}_{\boldsymbol{\eta w}}$.

$$\begin{bmatrix}
\underset{(-0.030,-0.009)}{\mathbf{-0.020}} & \underset{(-0.020,0.001)}{-0.010} & \underset{(-0.023,-0.003)}{\mathbf{-0.013}} & \underset{(-0.002,0.018)}{0.008} & \underset{(-0.009,0.010)}{0.000} & \underset{(-0.010,0.009)}{-0.001} \\
\underset{(-0.018,0.002)}{-0.009} & \underset{(-0.026,-0.003)}{\mathbf{-0.015}} & \underset{(-0.020,0.001)}{-0.010} & \underset{(-0.004,0.017)}{0.006} & \underset{(-0.009,0.011)}{0.001} & \underset{(-0.010,0.010)}{0.000} \\
\underset{(-0.016,0.002)}{-0.007} & \underset{(-0.015,0.004)}{-0.006} & \underset{(-0.033,-0.013)}{\mathbf{-0.023}} & \underset{(-0.005,0.015)}{0.005} & \underset{(-0.011,0.008)}{-0.002} & \underset{(-0.010,0.008)}{-0.001} \\
\underset{(-0.002,0.017)}{0.007} & \underset{(-0.007,0.013)}{0.003} & \underset{(-0.010,0.010)}{0.000} & \underset{(-0.015,0.007)}{-0.004} & \underset{(-0.009,0.010)}{0.000} & \underset{(-0.007,0.012)}{0.002} \\
\underset{(-0.005,0.013)}{0.004} & \underset{(-0.007,0.012)}{0.002} & \underset{(-0.010,0.009)}{0.000} & \underset{(-0.003,0.016)}{0.006} & \underset{(-0.027,-0.007)}{\mathbf{-0.017}} & \underset{(-0.011,0.007)}{-0.002} \\
\underset{(-0.002,0.015)}{0.006} & \underset{(-0.006,0.012)}{0.003} & \underset{(-0.010,0.008)}{-0.001} & \underset{(-0.001,0.018)}{0.009} & \underset{(-0.002,0.016)}{0.007} & \underset{(-0.024,-0.005)}{\mathbf{-0.015}}
\end{bmatrix}$$

*The (i,j)-th element corresponds to that of $\boldsymbol{\Psi}_{\boldsymbol{\eta w}}$. Bold figures indicate that the 95% credible interval does not include zero.

5.3 Model comparison based on portfolio performances

This subsection conducts a model comparison based on the portfolio performances. In modeling time-varying variances of asset returns, it is important to forecast the future covariance matrices of the time series for the financial risk management. To evaluate such a forecasting performance, we conduct out-of-sample covariance forecasts and give the minimum variance portfolios. It has often been implemented to investigate such a forecasting performance by the well-known mean-variance optimization (e.g., Luenberger (1997)).

Let $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$ denote the conditional mean and covariance matrix of $\boldsymbol{y}_{t+1}$ given the information at time t ($\mathcal{F}_t$) and the parameter $\boldsymbol{\vartheta}$. We make a minimum variance (MV) portfolio, and the MV portfolio weight ($\boldsymbol{w}$) is obtained as the solution to the problem:

$$\min_{\boldsymbol{w}} \boldsymbol{w}'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})\boldsymbol{w} \text{ s.t. } \boldsymbol{w}'\mathbf{1}_p = 1 \text{ and } \boldsymbol{w}'\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta}) \geq q_0, \tag{31}$$

where q_0 is the target value. It indicates that we make the expected portfolio returns exceed q_0. The optimal weight is given by

$$\boldsymbol{w}_{\mathrm{MV}} = \frac{c - q_0 b}{ac - b^2}\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})^{-1}\mathbf{1}_p + \frac{q_0 a - b}{ac - b^2}\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta}), \tag{32}$$

where

$$a = \mathbf{1}_p'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})^{-1}\mathbf{1}_p, \tag{33}$$

$$b = \mathbf{1}_p'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta}), \tag{34}$$

$$c = \mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta}). \tag{35}$$

We implement the rolling forecast as follows:

1. Estimate the model parameters using the data from 2001 to 2006 (we set the data as $\{\boldsymbol{y}_t\}_{t=1}^n$ and the posterior mean of the parameter vector as $\overline{\boldsymbol{\vartheta}}$.)

2. For the next year including n_1 trading days, i.e., $t = n+1, \ldots, n+n_1-1$,

   (a) Use the efficient sequential Monte Carlo method described in Appendix C.1 to compute $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \overline{\boldsymbol{\vartheta}})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \overline{\boldsymbol{\vartheta}})$.

   (b) Compute the hedge portfolio weight described above and the "realized" returns, $\boldsymbol{w}_{\mathrm{MV}}'\boldsymbol{y}_{t+1}$.

3. Include the new observations of the next year to our estimation period and remove the observations of the oldest year. Re-estimate the parameters of interest using the six-year-data (relabelled as $\{\boldsymbol{y}_t\}_{t=1}^{n}$).

4. Go to Step 2.

In the end, we calculate the standard deviations of the "realized" returns (756 in total). The numerical standard error of the estimate is obtained by repeating the sequential Monte Carlo ten times. We compare the following nine models:

1. Models with realized measures where $\boldsymbol{\Psi}_{w\varepsilon} = \boldsymbol{\Psi}_{\eta w} = \boldsymbol{\Psi}_{\eta\varepsilon} = \mathbf{O}$ and $\boldsymbol{\Psi}_{ww}$ is a diagonal matrix.
   (a) Univariate models (RSV).
   (b) 1-block dynamic equicorrelation model (1B-RDESV).
   (c) 3-block dynamic equicorrelation model (3B-RDESV).
2. Models with realized measures and leverage where $\boldsymbol{\Psi}_{w\varepsilon} = \boldsymbol{\Psi}_{\eta w} = \mathbf{O}$ and $\boldsymbol{\Psi}_{ww}$ is a diagonal matrix.
   (a) Univariate models (RSV-L).
   (b) 1-block dynamic equicorrelation model (1B-RDESV-L).
   (c) 3-block dynamic equicorrelation model (3B-RDESV-L).
3. Models with realized measures, leverage and endogeneity.
   (a) Univariate models (RSV-LE).
   (b) 1-block dynamic equicorrelation model (1B-RDESV-LE).
   (c) 3-block dynamic equicorrelation model (3B-RDESV-LE).

To compute $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \overline{\boldsymbol{\vartheta}})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \overline{\boldsymbol{\vartheta}})$, we implement the sequential Monte Carlo as described in Appendix C.1 where we set the number of particles $M = 5,000$ with the block length $L = 3$ for multivariate models and $L = 1$ for univariate models. The target value is set $q_0 = -10, 0, 10, 20$ and 30 annually.

Table 16: Out-of-sample portfolio standard deviations. Year 2007.

	Target return q_0				
Model	-10	0	10	20	30
RSV	1.661	1.405	1.206	1.068	0.993
	(0.034)	(0.025)	(0.017)	(0.011)	(0.006)
1B-RDESV	1.632	1.310	1.064	0.908	0.856
	(0.095)	(0.069)	(0.046)	(0.025)	(0.009)
3B-RDESV	1.555	1.250	1.023	0.885	0.845
	(0.068)	(0.055)	(0.039)	(0.022)	(0.013)
RSV-L	1.673	1.404	1.204	1.078	1.027
	(0.041)	(0.030)	(0.020)	(0.011)	(0.006)
1B-RDESV-L	1.110	1.204	1.305	1.408	1.510
	(0.061)	(0.077)	(0.091)	(0.103)	(0.113)
3B-RDESV-L	1.164	0.958	**0.839**	**0.812**	0.864
	(0.029)	(0.017)	(0.009)	(0.005)	(0.014)
RSV-LE	1.583	1.332	1.146	1.031	0.986
	(0.053)	(0.039)	(0.025)	(0.015)	(0.011)
1B-RDESV-LE	0.957	0.891	0.852	0.838	0.843
	(0.018)	(0.012)	(0.007)	(0.004)	(0.006)
3B-RDESV-LE	**0.948**	**0.886**	0.843	0.818	**0.807**
	(0.014)	(0.011)	(0.007)	(0.004)	(0.002)

*Standard errors in parentheses. Bold figures indicate the minimum of the column.

Table 17: Out-of-sample portfolio standard deviations. Year 2008.

	Target return q_0				
Model	-10	0	10	20	30
RSV	2.505	2.547	2.766	3.089	3.463
	(0.108)	(0.176)	(0.244)	(0.305)	(0.359)
1B-RDESV	2.545	2.272	2.105	2.036	2.051
	(0.080)	(0.053)	(0.031)	(0.016)	(0.012)
3B-RDESV	2.456	2.144	**1.968**	**1.925**	1.988
	(0.078)	(0.040)	(0.022)	(0.031)	(0.058)
RSV- L	2.620	2.443	2.417	2.504	2.670
	(0.076)	(0.034)	(0.059)	(0.117)	(0.171)
1B-RDESV-L	2.210	2.158	2.223	2.364	2.550
	(0.037)	(0.040)	(0.085)	(0.139)	(0.194)
3B-RDESV-L	2.252	2.422	2.710	3.058	3.429
	(0.094)	(0.165)	(0.224)	(0.274)	(0.315)
RSV-LE	2.609	2.486	2.428	2.421	2.454
	(0.028)	(0.031)	(0.054)	(0.082)	(0.109)
1B-RDESV-LE	2.125	2.054	2.003	1.969	**1.948**
	(0.021)	(0.017)	(0.014)	(0.013)	(0.012)
3B-RDESV-LE	**2.089**	**2.046**	2.021	2.011	2.011
	(0.023)	(0.022)	(0.023)	(0.027)	(0.032)

*Standard errors in parentheses. Bold figures indicate the minimum of the column.

Table 18: Out-of-sample portfolio standard deviations. Year 2009.

	Target return q_0				
Model	-10	0	10	20	30
RSV	1.736	1.598	1.603	1.712	1.882
	(0.038)	(0.042)	(0.076)	(0.120)	(0.165)
1B-RDESV	1.708	1.546	1.507	1.571	1.709
	(0.106)	(0.060)	(0.037)	(0.049)	(0.069)
3B-RDESV	1.612	1.433	1.391	1.462	1.612
	(0.075)	(0.041)	(0.055)	(0.094)	(0.129)
RSV-L	2.274	1.927	1.694	1.584	1.589
	(0.137)	(0.080)	(0.040)	(0.031)	(0.059)
1B-RDESV-L	1.697	1.929	2.206	2.496	2.786
	(0.052)	(0.090)	(0.128)	(0.162)	(0.193)
3B-RDESV-L	1.499	1.394	1.426	1.566	1.776
	(0.087)	(0.036)	(0.057)	(0.087)	(0.105)
RSV-LE	1.958	1.761	1.621	1.534	1.496
	(0.129)	(0.099)	(0.073)	(0.051)	(0.033)
1B-RDESV-LE	1.857	1.744	1.656	1.591	1.544
	(0.064)	(0.053)	(0.042)	(0.033)	(0.025)
3B-RDESV-LE	**1.434**	**1.361**	**1.318**	**1.301**	**1.305**
	(0.031)	(0.022)	(0.013)	(0.009)	(0.014)

*Standard errors in parentheses. Bold figures indicate the minimum of the column.

Tables 16, 17 and 18 report the out-of-sample portfolio standard deviations using the six-year rolling estimation window for the years 2007, 2008 and 2009. For the year 2007, three-block dynamic equicorrelations models with realized measures, leverage and endogeneity (3B-RDESV-L and 3B-RDESV-LE) outperform other competing models in the sense that their standard deviation of the "realized" returns are the smallest among others, but the endogeneity may not be important for the larger target values, $q_0 = 10$ and 20. For the year 2008, three-block or one-block dynamic equicorrelations models with realized measures, leverage and endogeneity (3B-RDESV-LE and 1B-RDESV-LE) and three-block dynamic equicorrelations model (3B-RDESV) seem to outperform other models, and the leverage and endogeneity may not be important for the larger target values, $q_0 = 10$ and 20. All standard deviations of the "realized" returns are much larger than those in 2007, due to the financial crisis in 2008. Finally, for the year 2009, three-block dynamic equicorrelations model with realized measures, leverage and endogeneity (3B-RDESV-LE) also seems to outperform other models for all the target values.

Overall, the standard deviations of the "realized" returns are found to be smaller than

those of the other competing models for our three-block dynamic equicorrelations models with realized measures, leverage and endogeneity. It indicates that our proposed model with the time-varying multiple-block equicorrelation structure with leverage and endogeneity shows good out-of-sample forecasting performance with respect to dynamic MV portfolio, even when the market is turbulent due to the aftermath of 2008 financial crisis.

6 Conclusion

The dynamic equicorrelation SV model of Kurose and Omori (2016), which incorporates the dynamic equicorrelation and cross leverage effect for the SV model, is extended to have the multiple-block dynamic equicorrelation structure. We also extend it to consider the simultaneous modeling of the multivariate daily returns and the related realized measure. Bayesian estimation scheme via Markov chain Monte Carlo method is described to conduct the statistical inference regarding the parameters. Numerical examples are provided and the proposed model is applied to the multivariate stock returns data. We find the persistence in both the volatilities and the correlations, and the existence of cross leverage effects. The estimation result suggests that the realized measures are endogenous and correlated with the asset return and that the biases in the realized measure owing to market microstructure noise, non-trading hours and non-synchronous trading are adjusted within the proposed model. In the model comparison based on the minimum variance portfolio performances, our models are found to outperform competing models regarding the standard deviations of the realized returns.

Acknowledgements

The authors would like to thank Yuzo Maruyama, Teruo Nakatsuma, Makoto Takahashi and Hisashi Tanizaki for their valuable comments. This work is supported by the research grant from Ishii Memorial Securities Research Promotion Foundation and by the Grants-in-Aid for Scientific Research (A) 21243018 and (A) 26245028 from the Japanese Ministry of Education, Science, Sports, Culture and Technology. The computational results are obtained using Ox version 5.10 (see Doornik (2007)).

References

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

Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, Ser. B 72*, 269–342.

Asai, M., M. Caporin, and M. McAleer (2015). Forecasting Value-at-Risk using block structure multivariate stochastic volatility models. *International Review of Economics & Finance 40*, 40–50.

Asai, M. and M. McAleer (2006). Asymmetric multivariate stochastic volatility. *Econometric Reviews 25*, 453–473.

Asai, M. and M. McAleer (2009). Multivariate stochastic volatility, leverage and news impact surfaces. *Econometrics Journal 12*, 292–309.

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

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

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Ser. B 63*, 167–241.

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*, 885–925.

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

Chan, D., R. Kohn, and C. Kirby (2006). Multivariate stochastic volatility models with correlated errors. *Econometric Reviews 25*, 245–274.

Chaussé, P. and D. Xu (2016). GMM estimation of a realized stochastic volatility model: A Monte Carlo study. *Econometric Reviews*, in press.

Chib, S., Y. Omori, and M. Asai (2009). Multivariate stochastic volatility. In T. G. Andersen, R. A. Davis, J. P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, Volume 4, pp. 365–400. Springer-Verlag.

Daníelsson, J. (1998). Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models. *Journal of Empirical Finance 5*, 155–173.

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. Working Paper.

Doornik, J. A. (2007). *An Object-oriented Martrix Programming Language: Ox 5*. Timberlake Consultants Press.

Doucet, A., M. Briersa, and S. Sénécal (2006). Efficient block sampling strategies for sequential Monte Carlo methods. *Journal of Computational and Graphical Statistics 15*, 693–711.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–615.

Elton, E. J. and M. J. Gruber (1973). Estimating the dependence structure of share prices–implications for portfolio selection. *Journal of Finance 28*, 1203–1232.

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

Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association 74*, 291–298.

Gupta, A. K. and D. K. Nagar (2000). *Matrix Variate Distributions*. Chapman and Hall/CRC.

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

Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics 20*, 873–889.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics 24*, 127–161.

Hasbrouck, J. (2007). *Empirical Market Microstructure: The Institutions, Economics And Econometrics of Securities Trading*. Oxford University Press.

Hayashi, T. and N. Yoshida (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli 11*, 359–379.

Ishihara, T. and Y. Omori (2012). Efficient Bayesian estimation of a multivariate stochastic volatility model with cross leverage and heavy-tailed errors. *Computational Statistics & Data Analysis 56*, 3674–3689.

Ishihara, T., Y. Omori, and M. Asai (2016). Matrix exponential stochastic volatility with cross leverage. *Computational Statistics & Data Analysis 100*, 331–350.

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.

Kunitomo, N. and S. Sato (2013). Separating Information Maximum Likelihood estimation of the integrated volatility and covariance with micro-market noise. *North American Journal of Economics and Finance 26*, 282–309.

Kurose, Y. and Y. Omori (2016). Dynamic equicorrelation stochastic volatility. *Computational Statistics & Data Analysis 100*, 795–813.

Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer.

Liu, J. S., F. Liang, and W. H. Wong (2000). The multiple-try method and local optimization in Metropolis sampling. *Journal of the American Statistical Association 95*, 121–134.

Lucas, A., B. Schwaab, and X. Zhang (2016). Modeling financial sector joint tail risk in the Euro area. *Journal of Applied Econometrics*, in press.

Luenberger, D. G. (1997). *Investment Science*. Oxford University Press.

Malliavin, P. and M. E. Mancino (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics 6*, 49–61.

Nakajima, J. (2015). Bayesian analysis of multivariate stochastic volatility with skew distribution. *Econometric Reviews*, in press.

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

O'Hara, M. (1995). *Market Microstructure Theory*. Wiley.

Omori, Y. and T. Watanabe (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis 52*, 2892–2910.

Peters, R. T. and R. G. de Vilder (2006). Testing the continuous semimartingale hypothesis for the S&P 500. *Journal of Business & Economic Statistics 24*, 444–454.

Richard, J.-F. and W. Zhang (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics 141*, 1385–1411.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–667.

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

Takahashi, M., T. Watanabe, and Y. Omori (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis 53*, 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*, 437–457.

Trojan, S. (2014). Multivariate stochastic volatility with dynamic cross leverage. Working Paper.

Ubukata, M. and K. Oya (2009). Estimation and testing for dependence in market microstructure noise. *Journal of Financial Econometrics 7*, 106–151.

Watanabe, T. (2012). Quantile forecasts of financial returns using realized GARCH models. *Japanese Economic Review 63*, 68–80.

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

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

Appendix

A Inverse matrix of $\mathbf{R}_t$

Proposition 3 (Inverse matrix of $\mathbf{R}_t$). The inverse of the multiple-block equicorrelation matrix $\mathbf{R}_t$ is given by

$$\mathbf{R}_t^{-1} = \mathbf{A}_t^{-1} + \mathbf{D}_t,$$

where

$$\begin{aligned}
\mathbf{A}_t^{-1} &= \operatorname{diag}\left\{(1-\rho_{11,t})^{-1}\mathbf{1}_{p_1}', \cdots, (1-\rho_{KK,t})^{-1}\mathbf{1}_{p_K}'\right\}, \\
\mathbf{D}_t &= \begin{pmatrix} d_{11,t}\mathbf{1}_{p_1}\mathbf{1}_{p_1}' & \cdots & d_{1K,t}\mathbf{1}_{p_1}\mathbf{1}_{p_K}' \\ \vdots & \ddots & \vdots \\ d_{K1,t}\mathbf{1}_{p_K}\mathbf{1}_{p_1}' & \cdots & d_{KK,t}\mathbf{1}_{p_K}\mathbf{1}_{p_K}' \end{pmatrix} = \mathbf{C}\tilde{\mathbf{D}}_t\mathbf{C}', \\
\tilde{\mathbf{D}}_t &= \begin{pmatrix} d_{11,t} & \cdots & d_{1K,t} \\ \vdots & \ddots & \vdots \\ d_{K1,t} & \cdots & d_{KK,t} \end{pmatrix} = -(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1},
\end{aligned}$$

and $\tilde{\mathbf{A}}_t, \tilde{\mathbf{B}}_t, \mathbf{C}, \mathbf{P}$ are defined in Proposition 1.

Proof: Noting that $\mathbf{C}'\mathbf{C} = \mathbf{P}$, $\mathbf{C}\tilde{\mathbf{A}}_t = \mathbf{A}_t\mathbf{C}$ and $\tilde{\mathbf{A}}_t^{-1}\mathbf{C}' = \mathbf{C}'\mathbf{A}_t^{-1}$, we obtain

$$\begin{aligned}
\mathbf{R}_t(\mathbf{A}_t^{-1} + \mathbf{D}_t) &= (\mathbf{A}_t + \mathbf{B}_t)(\mathbf{A}_t^{-1} + \mathbf{D}_t) \\
&= (\mathbf{A}_t + \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}')\left\{\mathbf{A}_t^{-1} - \mathbf{C}(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}\mathbf{C}'\right\} \\
&= \mathbf{I}_p + \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}'\mathbf{A}_t^{-1} - (\mathbf{A}_t\mathbf{C} + \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{P})(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}\mathbf{C}' \\
&= \mathbf{I}_p + \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}'\mathbf{A}_t^{-1} - \mathbf{C}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}\mathbf{C}' = \mathbf{I}_p,
\end{aligned}$$

where $\mathbf{A}_t, \mathbf{B}_t$ are defined in the proof of Proposition 1. Similarly, using $\mathbf{C}'\mathbf{A}_t = \tilde{\mathbf{A}}_t\mathbf{C}'$ and

$\mathbf{A}_t^{-1}\mathbf{C} = \mathbf{C}\tilde{\mathbf{A}}_t^{-1}$,

$$\begin{aligned}
(\mathbf{A}_t^{-1} + \mathbf{D}_t)\mathbf{R}_t &= (\mathbf{A}_t^{-1} + \mathbf{D}_t)(\mathbf{A}_t + \mathbf{B}_t) \\
&= \left\{\mathbf{A}_t^{-1} - \mathbf{C}(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}\mathbf{C}'\right\}(\mathbf{A}_t + \mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}') \\
&= \mathbf{I}_p + \mathbf{A}_t^{-1}\mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}' - \mathbf{C}(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}\tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}(\mathbf{C}'\mathbf{A}_t + \mathbf{P}\tilde{\mathbf{B}}_t\mathbf{C}') \\
&= \mathbf{I}_p + \mathbf{A}_t^{-1}\mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}' - \mathbf{C}(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}(\mathbf{I}_K + \tilde{\mathbf{B}}_t\tilde{\mathbf{A}}_t^{-1}\mathbf{P})\tilde{\mathbf{B}}_t\mathbf{C}' \\
&= \mathbf{I}_p + \mathbf{A}_t^{-1}\mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}' - \mathbf{C}(\tilde{\mathbf{A}}_t + \tilde{\mathbf{B}}_t\mathbf{P})^{-1}(\mathbf{I}_K + \tilde{\mathbf{B}}_t\mathbf{P}\tilde{\mathbf{A}}_t^{-1})\tilde{\mathbf{B}}_t\mathbf{C}' \\
&= \mathbf{I}_p + \mathbf{A}_t^{-1}\mathbf{C}\tilde{\mathbf{B}}_t\mathbf{C}' - \mathbf{C}\tilde{\mathbf{A}}_t^{-1}\tilde{\mathbf{B}}_t\mathbf{C}' = \mathbf{I}_p,
\end{aligned}$$

and the result follows. □

B Conditional posterior distributions and MCMC algorithm

B.1 Generation of $\boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\xi}^*$ and $\boldsymbol{\gamma}$

The joint conditional posterior distribution of $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ is

$$\begin{pmatrix}\boldsymbol{\xi} \\ \boldsymbol{\mu}\end{pmatrix} | \cdot \sim \mathrm{N}_{2p}(\boldsymbol{m}_{\boldsymbol{\xi}\boldsymbol{\mu}}, \mathbf{S}_{\boldsymbol{\xi}\boldsymbol{\mu}}), \tag{36}$$

where

$$\begin{aligned}
\mathbf{S}_{\boldsymbol{\xi}\boldsymbol{\mu}} = \Bigg\{ &\mathrm{diag}\left(\mathbf{S}_{\boldsymbol{\xi}0}^{-1} + (\boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{w}} - \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\varepsilon}}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}\boldsymbol{w}})^{-1}, \mathbf{S}_{\boldsymbol{\mu}0}^{-1} + \boldsymbol{\Psi}_0^{-1}\right) \\
&+ (n-1)\mathrm{diag}(\mathbf{I}_p, \mathbf{I}_p - \boldsymbol{\Phi})(\boldsymbol{\Psi}_{22} - \boldsymbol{\Psi}_{21}\boldsymbol{\Psi}_{12})^{-1}\mathrm{diag}(\mathbf{I}_p, \mathbf{I}_p - \boldsymbol{\Phi})\Bigg\}^{-1}, \qquad (37) \\
\boldsymbol{m}_{\boldsymbol{\xi}\boldsymbol{\mu}} = \mathbf{S}_{\boldsymbol{\xi}\boldsymbol{\mu}}\Bigg[&\begin{pmatrix} \mathbf{S}_{\boldsymbol{\xi}0}^{-1}\boldsymbol{m}_{\boldsymbol{\xi}0} + (\boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{w}} - \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\varepsilon}}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}\boldsymbol{w}})^{-1}(\boldsymbol{x}_n - \boldsymbol{h}_n - \boldsymbol{\Psi}_{\boldsymbol{w}\boldsymbol{\varepsilon}}\boldsymbol{z}_{1n}) \\ \mathbf{S}_{\boldsymbol{\mu}0}^{-1}\boldsymbol{m}_{\boldsymbol{\mu}0} + \boldsymbol{\Psi}_0^{-1}\boldsymbol{h}_1 \end{pmatrix} \\
&+ \mathrm{diag}(\mathbf{I}_p, \mathbf{I}_p - \boldsymbol{\Phi})(\boldsymbol{\Psi}_{22} - \boldsymbol{\Psi}_{21}\boldsymbol{\Psi}_{12})^{-1}\sum_{t=1}^{n-1}\left(\begin{pmatrix}\boldsymbol{x}_t - \boldsymbol{h}_t \\ \boldsymbol{h}_{t+1} - \Phi\boldsymbol{h}_t\end{pmatrix} - \boldsymbol{\Psi}_{21}\boldsymbol{z}_{1t}\right)\Bigg]. \qquad (38)
\end{aligned}$$

Similarly, the conditional posterior distributions of $\boldsymbol{\xi}^*$ and $\boldsymbol{\gamma}$ are

$$\begin{aligned}
\boldsymbol{\xi}^*|\cdot \;&\sim\; \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\xi}}^*, \mathbf{S}_{\boldsymbol{\xi}}^*), \\
&\mathbf{S}_{\boldsymbol{\xi}}^* = \left(\mathbf{S}_{\boldsymbol{\xi}0}^{*-1} + n\boldsymbol{\Delta}^{*-1}\right)^{-1}, \;\; \boldsymbol{m}_{\boldsymbol{\xi}}^* = \mathbf{S}_{\boldsymbol{\xi}}^*\left\{\mathbf{S}_{\boldsymbol{\xi}0}^{*-1}\boldsymbol{m}_{\boldsymbol{\xi}0}^* + \boldsymbol{\Delta}^{*-1}\sum_{t=1}^{n}(\boldsymbol{x}_t^* - \boldsymbol{g}_t)\right\},
\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\gamma}|\cdot &\sim \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\gamma}}, \mathbf{S}_{\boldsymbol{\gamma}}),\\ &\mathbf{S}_{\boldsymbol{\gamma}} = \left\{\mathbf{S}_{\boldsymbol{\gamma}0}^{-1} + \boldsymbol{\Sigma}_0^{-1} + (n-1)(\mathbf{I}_q - \boldsymbol{\Theta})\boldsymbol{\Sigma}^{-1}(\mathbf{I}_q - \boldsymbol{\Theta})\right\}^{-1},\\ &\boldsymbol{m}_{\boldsymbol{\gamma}} = \mathbf{S}_{\boldsymbol{\gamma}}\left\{\mathbf{S}_{\boldsymbol{\gamma}0}^{-1}\boldsymbol{m}_{\boldsymbol{\gamma}0} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{g}_1 + (\mathbf{I}_q - \boldsymbol{\Theta})\boldsymbol{\Sigma}^{-1}\sum_{t=1}^{n-1}(\boldsymbol{g}_{t+1} - \boldsymbol{\Theta}\boldsymbol{g}_t)\right\}.\end{aligned}$$

B.2 Generation of $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$

Generation of $\boldsymbol{\Phi}$. Let

$$\begin{aligned}\mathbf{A} &= \sum_{t=1}^{n-1}(\boldsymbol{h}_t - \boldsymbol{\mu})(\boldsymbol{h}_t - \boldsymbol{\mu})',\\ \mathbf{B} &= \sum_{t=1}^{n-1}(\boldsymbol{h}_t - \boldsymbol{\mu})\left\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \begin{pmatrix}\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{\eta w}}\end{pmatrix}\begin{pmatrix}\mathbf{I}_p & \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\\ \boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{ww}}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{z}_{1t}\\ \boldsymbol{w}_t\end{pmatrix}\right\}'\mathbf{S}^{-1},\\ \mathbf{S} &= \boldsymbol{\Psi}_{\boldsymbol{\eta\eta}} - \begin{pmatrix}\boldsymbol{\Psi}_{\boldsymbol{\eta\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{\eta w}}\end{pmatrix}\begin{pmatrix}\mathbf{I}_p & \boldsymbol{\Psi}_{\boldsymbol{\varepsilon w}}\\ \boldsymbol{\Psi}_{\boldsymbol{w\varepsilon}} & \boldsymbol{\Psi}_{\boldsymbol{ww}}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon\eta}}\\ \boldsymbol{\Psi}_{\boldsymbol{w\eta}}\end{pmatrix},\end{aligned}$$

and $\boldsymbol{b}$ denote a vector whose i-th element is equal to the (i,i)-th element of $\mathbf{B}$. The conditional posterior density of $\boldsymbol{\phi} = \boldsymbol{\Phi}\mathbf{1}_p$ is given by

$$\pi(\boldsymbol{\phi}|\cdot) \propto g(\boldsymbol{\phi}) \times \exp\left\{-\frac{1}{2}(\boldsymbol{\phi} - \boldsymbol{m}_{\boldsymbol{\phi}})'\mathbf{S}_{\boldsymbol{\phi}}^{-1}(\boldsymbol{\phi} - \boldsymbol{m}_{\boldsymbol{\phi}})\right\}, \tag{39}$$

where

$$\begin{aligned}g(\boldsymbol{\phi}) &= \prod_{j=1}^{p}\pi(\phi_j) \times |\boldsymbol{\Psi}_0|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\boldsymbol{\Psi}_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right\}, &(40)\\ \mathbf{S}_{\boldsymbol{\phi}}^{-1} &= \mathbf{A} \odot \mathbf{S}^{-1}, \quad \boldsymbol{m}_{\boldsymbol{\phi}} = \mathbf{S}_{\boldsymbol{\phi}}\boldsymbol{b}, &(41)\end{aligned}$$

and $\odot$ denotes Hadamard product. Generate a candidate $\boldsymbol{\phi}^\dagger \sim \mathrm{TN}_{\{-1<\phi_i<1,\forall i\}}(\boldsymbol{m}_{\boldsymbol{\phi}}, \mathbf{S}_{\boldsymbol{\phi}})$ and accept it with probability $\min[1, g(\boldsymbol{\phi}^\dagger)/g(\boldsymbol{\phi})]$.

Generation of $\boldsymbol{\Theta}$. Similarly, let

$$\mathbf{A} = \sum_{t=1}^{n-1}(\boldsymbol{g}_t - \boldsymbol{\gamma})(\boldsymbol{g}_t - \boldsymbol{\gamma})', \quad \mathbf{B} = \sum_{t=1}^{n-1}(\boldsymbol{g}_t - \boldsymbol{\gamma})(\boldsymbol{g}_{t+1} - \boldsymbol{\gamma})'\boldsymbol{\Sigma}^{-1},$$

and $\boldsymbol{b}$ denote a vector whose i-th element is equal to the (i,i)-th element of $\mathbf{B}$. Then the conditional posterior density of $\boldsymbol{\theta} = \boldsymbol{\Theta}\mathbf{1}_q$ can be shown as

$$\pi(\boldsymbol{\theta}|\cdot) \propto g(\boldsymbol{\theta}) \times \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{m}_{\boldsymbol{\theta}})'\mathbf{S}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta} - \boldsymbol{m}_{\boldsymbol{\theta}})\right\}, \tag{42}$$

where

$$g(\boldsymbol{\theta}) = \prod_{j=1}^{q} \pi(\theta_j) \times |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}(\boldsymbol{g}_1 - \boldsymbol{\gamma})' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{g}_1 - \boldsymbol{\gamma}) \Big\}, \tag{43}$$

$$\mathbf{S}_{\boldsymbol{\theta}}^{-1} = \mathbf{A} \odot \boldsymbol{\Sigma}^{-1}, \quad \boldsymbol{m}_{\boldsymbol{\theta}} = \mathbf{S}_{\boldsymbol{\theta}} \boldsymbol{b}. \tag{44}$$

Generate a candidate $\boldsymbol{\theta}^\dagger \sim \mathrm{TN}_{\{-1<\theta_j<1, \forall j\}}(\boldsymbol{m}_{\boldsymbol{\theta}}, \mathbf{S}_{\boldsymbol{\theta}})$ and accept it with probability $\min[1, g(\boldsymbol{\theta}^\dagger)/g(\boldsymbol{\theta})]$.

B.3 Generation of $\boldsymbol{\Delta}^*$ and $\boldsymbol{\Omega}_{\boldsymbol{m}}$

The conditional posterior distributions of δ_j^{*2} and $\omega_{m_i}^2$ are

$$\delta_j^{*2}|\cdot \sim \mathrm{IG}(\alpha_{\delta_j 1}^*/2, \beta_{\delta_j 1}^*/2), \quad j = 1, \ldots, q, \tag{45}$$

$$\alpha_{\delta_j 1}^* = \alpha_{\delta_j 0}^* + n, \quad \beta_{\delta_j 1}^* = \beta_{\delta_j 0}^* + \sum_{t=1}^{n} (x_{kl,t}^* - \xi_j^* - g_{kl,t})^2. \tag{46}$$

where $\xi_j^* = \xi_{kl}^*$ and

$$\omega_{m_i}^2| \sim \mathrm{IG}(\alpha_{m_i 1}/2, \beta_{m_i 1}/2), \quad i = 1, \ldots, p, \tag{47}$$

$$\alpha_{m_i 1} = \alpha_{m_i 0} + n - 1, \quad \beta_{m_i 1} = \beta_{m_i 0} + \sum_{t=1}^{n-1} (m_{i,t+1} - m_{it})^2. \tag{48}$$

B.4 Generation of $\boldsymbol{\Sigma}$

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

$$\pi(\boldsymbol{\Sigma}|\cdot) \propto g(\boldsymbol{\Sigma}) \times |\boldsymbol{\Sigma}|^{-\frac{n_{\boldsymbol{\Sigma}}+q+1}{2}} \exp\Big\{ -\frac{1}{2} \mathrm{tr}(\mathbf{S}_{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\Sigma}^{-1}) \Big\}, \tag{49}$$

where $n_{\boldsymbol{\Sigma}} = n_{\boldsymbol{\Sigma}0} + n - 1$ and

$$\mathbf{S}_{\boldsymbol{\Sigma}}^{-1} = \mathbf{S}_{\boldsymbol{\Sigma}0}^{-1} + \sum_{t=1}^{n-1} \{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \boldsymbol{\Theta}(\boldsymbol{g}_t - \boldsymbol{\gamma})\} \{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \boldsymbol{\Theta}(\boldsymbol{g}_t - \boldsymbol{\gamma})\}', \tag{50}$$

$$g(\boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2} (\boldsymbol{g}_1 - \boldsymbol{\gamma}) \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{g}_1 - \boldsymbol{\gamma}) \Big\}. \tag{51}$$

Generate a candidate $\boldsymbol{\Sigma}^\dagger \sim \mathrm{IW}(n_{\boldsymbol{\Sigma}}, \mathbf{S}_{\boldsymbol{\Sigma}})$ and accept it with probability $\min[1, g(\boldsymbol{\Sigma}^\dagger)/g(\boldsymbol{\Sigma})]$.

C Efficient algorithms

C.1 Block particle filtering

We compute $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$ by integrating out latent variables numerically, using the block sampling strategy for the particle filtering, which is an efficient sequential Monte Carlo method discussed in Doucet, Briersa, and Sénécal (2006).

In a standard sequential Monte Carlo method, (i) we generate particles of $(\boldsymbol{h}_{t+1}, \boldsymbol{g}_{t+1}, \boldsymbol{m}_{t+1})$ given $\mathcal{F}_t$, (ii) compute the weights of the particles given $\mathcal{F}_{t+1}$ to construct the discrete approximation of the filtering density, $\hat{f}(\boldsymbol{h}_{t+1}, \boldsymbol{g}_{t+1}, \boldsymbol{m}_{t+1}|\mathcal{F}_{t+1}, \boldsymbol{\vartheta})$, and (iii) resample the particles. In our proposed model, the filtering density of the latent variables $\boldsymbol{h}_{t+1}$ and $\boldsymbol{g}_{t+1}$ heavily relies on the new observations at time $t+1$ (especially on the realized measures), and the resampling particles may concentrate in the narrow region of the high filtering density, which results in the path degeneracy and hence inaccurate estimates of $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t, \boldsymbol{\vartheta})$. To avoid this path degeneracy problem, we consider the conditional joint density functions,

$$
\begin{aligned}
&f(\boldsymbol{h}_{t-L+1:t}, \boldsymbol{g}_{t-L+1:t}, \boldsymbol{m}_{t-L+1:t}|Y_t, X_t, X_t^*, \boldsymbol{\vartheta}) \\
&\quad \propto \quad f(\boldsymbol{y}_{t-L+1:t}, \boldsymbol{x}_{t-L+1:t}, \boldsymbol{h}_{t-L+1:t}|\boldsymbol{y}_{t-L}, \boldsymbol{x}_{t-L}, \boldsymbol{h}_{t-L}, \boldsymbol{g}_{t-L:t}, \boldsymbol{m}_{t-L:t}, \boldsymbol{\vartheta}) f(\boldsymbol{m}_{t-L+1:t}|\boldsymbol{m}_{t-L}, \boldsymbol{\vartheta}) \\
&\qquad \times f(\boldsymbol{g}_{t-L+1:t}|\boldsymbol{x}^*_{t-L+1:t}, \boldsymbol{g}_{t-L}, \boldsymbol{\vartheta}) f(\boldsymbol{h}_{t-L}, \boldsymbol{g}_{t-L}, \boldsymbol{m}_{t-L}|Y_{t-1}, X_{t-1}, X^*_{t-1}, \boldsymbol{\vartheta}), \\
&f(\boldsymbol{h}_{t+1}, \boldsymbol{g}_{t+1}, \boldsymbol{m}_{t+1}|Y_t, X_t, X_t^*, \boldsymbol{\vartheta}) \\
&\quad \propto \quad f(\boldsymbol{h}_{t+1}|\boldsymbol{h}_t, \boldsymbol{g}_t, \boldsymbol{m}_t, \boldsymbol{y}_t, \boldsymbol{x}_t, \boldsymbol{\vartheta}) f(\boldsymbol{m}_{t+1}|\boldsymbol{m}_t, \boldsymbol{\vartheta}) f(\boldsymbol{g}_{t+1}|\boldsymbol{g}_t, \boldsymbol{\vartheta}),
\end{aligned}
$$

where $Y_t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$, $X_t = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t\}$, $X_t^* = \{\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_t^*\}$, and implement the sequential Monte Carlo as follows:

1. For $i = 1, \ldots, M$, conditional on $(\boldsymbol{h}_{t-L}^{(i)}, \boldsymbol{g}_{t-L}^{(i)}, \boldsymbol{m}_{t-L}^{(i)}, \boldsymbol{\vartheta})$, generate samples from an importance density approximating $f(\boldsymbol{h}_{t-L+1:t}, \boldsymbol{g}_{t-L+1:t}, \boldsymbol{m}_{t-L+1:t}|Y_t, X_t, X_t^*, \boldsymbol{\vartheta})$:

   (i) Generate $\boldsymbol{m}_{t-L+1:t}^{(i)} \sim f(\boldsymbol{m}_{t-L+1:t}|\boldsymbol{m}_{t-L}^{(i)}, \boldsymbol{\vartheta})$ using the state equation (3).

   (ii) Using the state equation (9) and the observation equation (12), implement the simulation smoother (e.g., de Jong and Shephard (1995)) and generate $\boldsymbol{g}_{t-L+1:t}^{(i)} \sim f(\boldsymbol{g}_{t-L+1:t}|\boldsymbol{x}^*_{t-L+1:t}, \boldsymbol{g}_{t-L}^{(i)}, \boldsymbol{\vartheta})$.

   (iii) Using the the state equation (2) and the observation equations (11) and (53), generate $\boldsymbol{h}_{t-L+1:t}^{(i)} \sim g(\boldsymbol{h}_{t-L+1:t}|\boldsymbol{y}_{t-L:t}, \boldsymbol{x}_{t-L:t}, \boldsymbol{m}_{t-L:t}^{(i)}, \boldsymbol{g}_{t-L:t}^{(i)}, \boldsymbol{h}_{t-L}^{(i)}, \boldsymbol{\vartheta})$ using the simulation smoother.

2. For $i = 1, \ldots, M$, compute

$$
\begin{aligned}
\pi_i &= \frac{\tilde{\pi}_i}{\sum_{j=1}^M \tilde{\pi}_j}, \\
\tilde{\pi}_i &= \frac{f(\boldsymbol{y}_{t-L+1:t}, \boldsymbol{x}_{t-L+1:t}, \boldsymbol{h}_{t-L+1:t}^{(i)}|\boldsymbol{y}_{t-L}, \boldsymbol{x}_{t-L}, \boldsymbol{h}_{t-L}^{(i)}, \boldsymbol{g}_{t-L:t}^{(i)}, \boldsymbol{m}_{t-L:t}^{(i)}, \boldsymbol{\vartheta})}{g(\boldsymbol{h}_{t-L+1:t}^{(i)}|\boldsymbol{y}_{t-L:t}, \boldsymbol{x}_{t-L:t}, \boldsymbol{m}_{t-L:t}^{(i)}, \boldsymbol{g}_{t-L:t}^{(i)}, \boldsymbol{h}_{t-L}^{(i)}, \boldsymbol{\vartheta})}.
\end{aligned}
$$

3. For $i = 1, \ldots, M$, set $\hat{f}(\boldsymbol{h}_{t-L+1:t}^{(i)}, \boldsymbol{g}_{t-L+1:t}^{(i)}, \boldsymbol{m}_{t-L+1:L}^{(i)} | Y_t, X_t, X_t^*, \boldsymbol{\vartheta}) = \pi_i$.

4. For $i = 1, \ldots, M$, resample
$\boldsymbol{h}_{t-L+1:t}^{(i)}, \boldsymbol{g}_{t-L+1:t}^{(i)}, \boldsymbol{m}_{t-L+1:t}^{(i)} \sim \hat{f}(\boldsymbol{h}_{t-L+1:t}, \boldsymbol{g}_{t-L+1:t}, \boldsymbol{m}_{t-L+1:t} | Y_t, X_t, X_t^*, \boldsymbol{\vartheta})$.

5. For $i = 1, \ldots, M$, generate $\boldsymbol{h}_{t+1}^{(i)} \sim f(\boldsymbol{h}_{t+1} | \boldsymbol{h}_t^{(i)}, \boldsymbol{g}_t^{(i)}, \boldsymbol{m}_t^{(i)}, Y_t, X_t, \boldsymbol{\vartheta})$, $\boldsymbol{g}_{t+1}^{(i)} \sim f(\boldsymbol{g}_{t+1} | \boldsymbol{g}_t^{(i)}, \boldsymbol{\vartheta})$, $\boldsymbol{m}_{t+1}^{(i)} \sim f(\boldsymbol{m}_{t+1} | \boldsymbol{m}_t^{(i)}, \boldsymbol{\vartheta})$. Estimate $\mathrm{E}(\boldsymbol{y}_{t+1} | \mathcal{F}_t, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1} | \mathcal{F}_t, \boldsymbol{\vartheta})$ by

$$\hat{\boldsymbol{\mu}}_{t+1|t} = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{m}_{t+1}^{(i)}, \quad \hat{\boldsymbol{\Sigma}}_{t+1|t} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{V}_{t+1}^{1/2(i)} \mathbf{R}_{t+1}^{(i)} \mathbf{V}_{t+1}^{1/2(i)} + \boldsymbol{\Omega}_{\boldsymbol{m}}. \tag{52}$$

6. Discard $\boldsymbol{h}_{t-L+2:t+1}^{(i)}, \boldsymbol{g}_{t-L+2:t+1}^{(i)}, \boldsymbol{m}_{t-L+2:t+1}^{(i)}$, $i = 1, \ldots, M$.

7. Replace t with $t+1$ and go to 1.

Remark 7. The above sequential Monte Carlo method generates a block of state variables of the length L, and discards the last $L-1$ state variables to avoid the path degeneracy problem. This improves the estimation accuracy of $\mathrm{E}(\boldsymbol{y}_{t+1} | \mathcal{F}_t, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{t+1} | \mathcal{F}_t, \boldsymbol{\vartheta})$. We note that it reduces to the standard sequential Monte Carlo method for $L = 1$.

C.2 Multi-move sampler for $\{\boldsymbol{h}_t\}_{t=1}^n$

We first divide $\{\boldsymbol{h}_t\}_{t=1}^n$ into $N+1$ blocks, $(\boldsymbol{h}_{k_{m-1}+1}, \ldots, \boldsymbol{h}_{k_m})$, $m = 1, \ldots, N$ with $k_0 = 0$, $k_{N+1} = n$, $k_i - k_{i-1} \geq 2$, using stochastic knots $k_m = \mathrm{int}[n(m+U_m)/(N+2)]$, where U_m's are independent uniform random variables on $(0,1)$ (see e.g. Omori and Watanabe (2008), Shephard and Pitt (1997)). Then consider the approximating linear Gaussian state space model where we replace Equations (1) and (14) with

$$\tilde{\boldsymbol{y}}_t = \mathbf{F}_t \boldsymbol{h}_t + \tilde{\boldsymbol{\epsilon}}_t, \quad \tilde{\boldsymbol{z}}_{1t} = \mathbf{R}_t^{-\frac{1}{2}} \tilde{\boldsymbol{\epsilon}}_t, \quad t = k_{m-1}+1, \ldots, k_m. \tag{53}$$

where $\tilde{\boldsymbol{y}}_t$, $\mathbf{F}_t$, and $\tilde{\boldsymbol{z}}_{1t}$ are defined in Section 3.3. Generate a candidate for $(\boldsymbol{h}_{k_{m-1}+1}, \ldots, \boldsymbol{h}_{k_m})$ given other blocks from using the simulation smoother (e.g. de Jong and Shephard (1995), Durbin and Koopman (2002)) and conduct MH algorithm.

D Supplementary results of model comparison

D.1 Cases with $K = 2$

We check the forecasting ability for two-block models ($K = 2$) and justify our proposed model with $K = 3$ in our empirical studies.

We use the same data series as in Section 5.3: (1) Bank of America (BAC), (2) JP Morgan (JPM), (3) American Express (AXP), (4) International Business Machines (IBM), (5) Microsoft (MSFT), (6) Coca Cola (KO).

Case 1. We divide them into two groups: Group 1_{A} (Finance and Information Technology) consists of BAC, JPM, AXP, IBM and MSFT, and Group 2_{A} (Food) consists of KO. We call the models as two-block (A) dynamic equicorrelations models. Thus we apply our proposed model to the dataset with two blocks ($K = 2$) and conduct out-of-sample covariance forecasts based on the portfolio performances.

Tables 19, 20 and 21 report the out-of-sample portfolio standard deviations using the six-year rolling estimation window for the years 2007, 2008 and 2009. For the year 2007, two-block (A) dynamic equicorrelations models do not outperform competing models listed in Table 16 in Section 5.3. For the year 2008, two-block (A) dynamic equicorrelations models with realized measures, leverage and endogeneity (2B-RDESV-LE(A) and 2B-RDESV-L(A)) seem to perform a little better than other models listed in Table 17 for $q_0 = -10, 0$ and 30. For the year 2009, two-block (A) dynamic equicorrelations model with realized measures, leverage (2B-RDESV-LE(A)) do not seem to outperform other models listed in Table 18 for all the target values.

Case 2. We also divide the data series into other two groups: Group 1_{B} (Finance) consists of BAC, JPM and AXP, and Group 2_{B} (Information Technology and Food) consists of MSFT, IBM and KO. We call the models as two-block (B) dynamic equicorrelations models.

Tables 22, 23 and 24 report the out-of-sample portfolio standard deviations using the six-year rolling estimation window for the years 2007, 2008 and 2009. For the year 2007, two-block (B) dynamic equicorrelations models with realized measures, leverage and endogeneity (2B-RDESV-LE(B)) seem to outperform other models listed in Table 16 for target value $q_0 = -10$. For the year 2008, two-block (B) dynamic equicorrelations models with realized measures, leverage and endogeneity (2B-RDESV-LE(B) and 2B-RDESV-L(B)) seem to perform a little better than other models listed in Table 17 for $q_0 = -10$ and 0. For the year 2009, two-block(B) dynamic equicorrelations model with realized measures, leverage (2B-RDESV-L(B)) seems to outperform other models listed

in Table 18 for $q_0 = 0$.

Case 3. Finally, we divide the data series into two groups: Group 1_C (Finance and Food) consists of BAC, JPM, AXP and KO, and Group 2_C (Information Technology) consists of MSFT and IBM. We call the models as two-block (C) dynamic equicorrelations models.

Tables 25, 26 and 27 report the out-of-sample portfolio standard deviations using the six-year rolling estimation window for the years 2007, 2008 and 2009. For the years 2007, 2008 and 2009, two-block (C) dynamic equicorrelations models with realized measures, leverage and endogeneity do not seem to outperform other models listed in Tables 16, 17 and 18 for all the target values.

Overall, our three-block dynamic equicorrelations models show good out-of-sample forecasting performance with respect to dynamic MV portfolio, even when the market is turbulent due to the aftermath of 2008 financial crisis.

Table 19: Out-of-sample portfolio standard deviations. Year 2007.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(A)	1.542	1.242	1.022	0.897	0.873
	(0.066)	(0.050)	(0.035)	(0.018)	(0.008)
2B-RDESV-L(A)	1.197	1.002	0.877	0.827	0.844
	(0.023)	(0.018)	(0.013)	(0.006)	(0.005)
2B-RDESV-LE(A)	0.990	0.917	0.867	0.837	0.824
	(0.016)	(0.012)	(0.008)	(0.005)	(0.003)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 16.

Table 20: Out-of-sample portfolio standard deviations. Year 2008.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(A)	2.291	2.068	2.022	2.112	2.289
	(0.025)	(0.018)	(0.026)	(0.065)	(0.113)
2B-RDESV-L(A)	2.134	**2.032**	2.024	2.088	2.200
	(0.025)	(0.012)	(0.028)	(0.048)	(0.072)
2B-RDESV-LE(A)	**2.078**	**2.019**	1.976	1.948	**1.931**
	(0.015)	(0.012)	(0.010)	(0.008)	(0.008)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 17.

Table 21: Out-of-sample portfolio standard deviations. Year 2009.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(A)	1.626	1.441	1.523	1.790	2.139
	(0.110)	(0.039)	(0.041)	(0.082)	(0.124)
2B-RDESV-L(A)	1.467	1.361	1.358	1.434	1.559
	(0.022)	(0.007)	(0.015)	(0.028)	(0.044)
2B-RDESV-LE(A)	1.800	1.678	1.580	1.504	1.447
	(0.065)	(0.053)	(0.043)	(0.034)	(0.025)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 18.

Table 22: Out-of-sample portfolio standard deviations. Year 2007.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(B)	1.646	1.328	1.082	0.919	0.853
	(0.087)	(0.064)	(0.043)	(0.024)	(0.010)
2B-RDESV-L(B)	1.149	0.971	0.865	0.830	0.857
	(0.031)	(0.022)	(0.013)	(0.009)	(0.015)
2B-RDESV-LE(B)	**0.945**	0.887	0.847	0.822	0.811
	(0.012)	(0.008)	(0.005)	(0.003)	(0.002)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 16.

Table 23: Out-of-sample portfolio standard deviations. Year 2008.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(B)	2.443	2.164	2.022	2.001	2.076
	(0.085)	(0.053)	(0.021)	(0.029)	(0.059)
2B-RDESV-L(B)	2.175	**2.027**	2.007	2.086	2.234
	(0.025)	(0.015)	(0.027)	(0.039)	(0.052)
2B-RDESV-LE(B)	**2.080**	**2.034**	2.007	1.994	1.993
	(0.015)	(0.010)	(0.009)	(0.013)	(0.019)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 17.

Table 24: Out-of-sample portfolio standard deviations. Year 2009.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(B)	1.839	1.578	1.505	1.572	1.734
	(0.172)	(0.113)	(0.053)	(0.072)	(0.131)
2B-RDESV-L(B)	1.484	**1.353**	1.358	1.462	1.628
	(0.056)	(0.030)	(0.034)	(0.064)	(0.093)
2B-RDESV-LE(B)	1.655	1.546	1.461	1.396	1.350
	(0.048)	(0.041)	(0.035)	(0.030)	(0.025)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 18.

Table 25: Out-of-sample portfolio standard deviations. Year 2007.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(C)	1.807	1.463	1.188	0.989	0.883
	(0.068)	(0.057)	(0.045)	(0.032)	(0.019)
2B-RDESV-L(C)	1.221	1.017	0.886	0.833	0.851
	(0.034)	(0.023)	(0.012)	(0.005)	(0.010)
2B-RDESV-LE(C)	1.056	0.977	0.918	0.876	0.850
	(0.027)	(0.021)	(0.015)	(0.011)	(0.007)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 16.

Table 26: Out-of-sample portfolio standard deviations. Year 2008.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(C)	2.535	2.216	2.046	2.015	2.095
	(0.069)	(0.042)	(0.024)	(0.024)	(0.040)
2B-RDESV-L(C)	2.210	2.076	2.063	2.142	2.280
	(0.024)	(0.013)	(0.021)	(0.046)	(0.077)
2B-RDESV-LE(C)	2.264	2.180	2.114	2.063	2.025
	(0.021)	(0.018)	(0.015)	(0.013)	(0.010)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 17.

Table 27: Out-of-sample portfolio standard deviations. Year 2009.

	Target return q_0				
Model	-10	0	10	20	30
2B-RDESV(C)	2.055	1.665	1.540	1.661	1.932
	(0.079)	(0.037)	(0.018)	(0.050)	(0.094)
2B-RDESV-L(C)	1.565	1.443	1.424	1.489	1.608
	(0.039)	(0.017)	(0.011)	(0.027)	(0.046)
2B-RDESV-LE(C)	2.006	1.868	1.755	1.663	1.588
	(0.069)	(0.059)	(0.051)	(0.043)	(0.036)

*Standard errors in parentheses. Bold figures indicate that they are smaller than the minimum of Table 18.

D.2 Comparison with CRSV model

In this subsection, we conduct the one-step-ahead forecasting by rolling the sample period described as in Shirota, Omori, Lopes, and Piao (2017).

We use daily close-to-close returns and their realized measures for some of the most liquid stocks in the Dow Jones Industrial Average (DJIA) index, obtained from Oxford-Man Institute (Noureldin, Shephard, and Sheppard (2012)). These 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[6]. Taking account of the sample correlations among the nine series, we divide them into four groups: Group 1 (Finance) consists of JPM and AXP, Group 2 (Information Technology and Others) consists of IBM, MSFT and GE, Group 3 (Manufacturing) consists of XOM, AA and DD, and Group 4(Food) consists of KO. Thus we apply our proposed model to this dataset with four blocks ($K = 4$). We follow Shirota, Omori, Lopes, and Piao (2017) and modify the equation (4) in Section 2.1 as

$$\boldsymbol{m}_1 \sim \mathrm{N}(\mathbf{0}_p, \lambda_m \boldsymbol{\Omega}_{\boldsymbol{m}}),$$

and set $\lambda_m = 100$. We also set $\omega_{m_i 0}^2 \sim \mathrm{IG}(10^{-5}/2, 10^{-5}/2),\ i = 1, \ldots, p$.

First we generate the model parameters and latent variables using 1742 ($n = 1742$) observations from February 1, 2001 to January 8, 2008 and also generate the latent variables for the next day given them. Next, we include the new observations of the next day to our estimation period and remove the observations of the oldest day. The parameters of interest are re-estimated using the new data and the latent variables are generated for the next day. This is repeated until all one-step-ahead forecasts are conducted through December 31, 2009 (500 in total).

Let N_{iter} denote the number of MCMC iterations for parameter estimation. We generate MCMC samples of $(\{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta})$ as described in Section 3[7] and add a step to the i-th MCMC iteration as follows:

- Generate
  $\boldsymbol{h}_{n+1}^{(i)}, \boldsymbol{g}_{n+1}^{(i)}, \boldsymbol{m}_{n+1}^{(i)} | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t^{(i)}\}_{t=1}^n, \{\boldsymbol{g}_t^{(i)}\}_{t=1}^n, \{\boldsymbol{m}_t^{(i)}\}_{t=1}^n, \boldsymbol{\vartheta}^{(i)}$.

We estimate $\mathrm{E}(\boldsymbol{y}_{n+1}|\mathcal{F}_n, \boldsymbol{\vartheta})$ and $\mathrm{Var}(\boldsymbol{y}_{n+1}|\mathcal{F}_n, \boldsymbol{\vartheta})$ by

$$\hat{\boldsymbol{\mu}}_{n+1|n} = \frac{1}{N_{\mathrm{iter}}} \sum_{i=1}^{N_{\mathrm{iter}}} \boldsymbol{m}_{n+1}^{(i)} \text{ and } \hat{\boldsymbol{\Sigma}}_{n+1|n} = \frac{1}{N_{\mathrm{iter}}} \sum_{i=1}^{N_{\mathrm{iter}}} (\mathbf{V}_{n+1}^{1/2(i)} \mathbf{R}_{n+1}^{(i)} \mathbf{V}_{n+1}^{1/2(i)} + \boldsymbol{\Omega}_{\boldsymbol{m}}^{(i)}).$$

[6]Noting that the asset return of American Express (AXP) is extremely low on October 3, 2005 when the company conducted a stock split, we regard the data on that day as an outlier and remove it from the original data by Noureldin, Shephard, and Sheppard (2012) in Section 5.3.

[7]The estimation algorithm are partly modified for this forecasting study. See the remark at the end of this subsection.

We set $N_{\text{iter}} = 3000$ and generate MCMC samples after discarding first 1,000 samples as the burn-in period. We also set the initial values for the model parameters and latent variables as the posterior means of previous MCMC iteration.

In this subsection, we make three hedge portfolios: a minimum variance portfolio, a mean-variance portfolio and a maximum expected return portfolio[8].

Table 28 reports the cumulative objective functions (realized portfolio variances, realized utility functions and realized returns) for (1) minimum variance strategy with the target value $\mu_p^* = 0.004, 0.01$ and 0.1, (2) mean-variance strategy with $\gamma = 6, 10$ and 15, and (3) maximum return strategy with $\sigma_p^{2*} = 0.001, 0.01$ and 0.1 for 4B-RDESV-LE and 1B-RDESV-LE models. For the minimum variance and maximum return strategies, 4B-RDESV-LE and 1B-RDESV-LE models outperform the models of Shirota, Omori, Lopes, and Piao (2017)[9]. On the other hand, for mean-variance strategy, CRSV models of Shirota, Omori, Lopes, and Piao (2017) outperform our model. As discussed in Shirota, Omori, Lopes, and Piao (2017), the realized measures are not useful to improve the performance in the mean-variance strategy in general. Under all settings in each strategy, 4B-RDESV-LE model outperforms 1B-RDESV-LE model.

Figures 5, 6 and 7 illustrate the time series plots of the portfolio weight for three strategies in 4B-RDESV-LE model. Similarly to Shirota, Omori, Lopes, and Piao (2017), we find that the portfolio weights fluctuate more drastically under the riskest setting in each strategy ($\mu_p^* = 0.1$, $\gamma = 6$ and $\sigma_p^{2*} = 0.1$) than under the other settings in each strategy. We also notice that all weights of the stocks for the mean-variance strategy are very small and the weights of the risk free asset are very large. Time series plots of cumulative realized returns and realized portfolio variances are illustrated in Figures 8 and 9.

Therefore, in comparison with the models proposed by Shirota, Omori, Lopes, and Piao (2017), our method is shown to perform well in the analysis of multivariate stock returns.

Remark 8. If we include the data on October 3, 2005 to the data set, we face the difficulty that the sampling method of $\boldsymbol{g}_t$ does not work for the day. Therefore, we need to generate an element of $\boldsymbol{g}_t$ of the day given the other elements one by one and conduct the multiple-trial

[8]See Shirota, Omori, Lopes, and Piao (2017) for details. We follow Shirota, Omori, Lopes, and Piao (2017) and use federal funds rate for risk-free rate.

[9]See Table 18 of Shirota, Omori, Lopes, and Piao (2017).

Metropolized independent sampler with many trials.

Table 28: The value of cumulative objective functions for three strategies.

Minimum Variance	$\mu_p^* = 0.004$	$\mu_p^* = 0.01$	$\mu_p^* = 0.1$
CRSV	0.748	4.448	730.4
4B-RDESV-LE	**0.437**	**2.571**	**461.0**
1B-RDESV-LE	0.576	3.422	632.6
Mean-Variance	$\gamma = 6$	$\gamma = 10$	$\gamma = 15$
CRSV	0.981	1.153	1.239
4B-RDESV-LE	0.905	1.092	1.186
1B-RDESV-LE	0.877	1.075	1.175
Maximum Return	$\sigma_p^{2*} = 0.001$	$\sigma_p^{2*} = 0.01$	$\sigma_p^{2*} = 0.1$
CRSV	1.172	0.662	-0.947
4B-RDESV-LE	**1.537**	**1.892**	**3.013**
1B-RDESV-LE	1.384	1.409	1.485

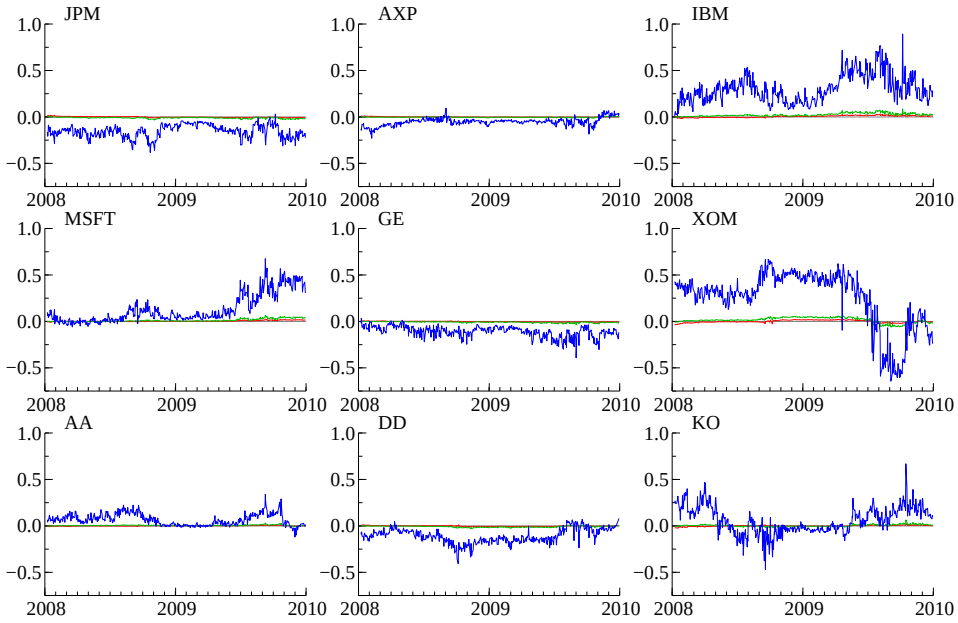


Figure 5: Time series plot of the portfolio weight for minimum variance strategy: $\mu_p^* = 0.004$ (red), $\mu_p^* = 0.01$ (green) and $\mu_p^* = 0.1$ (blue).

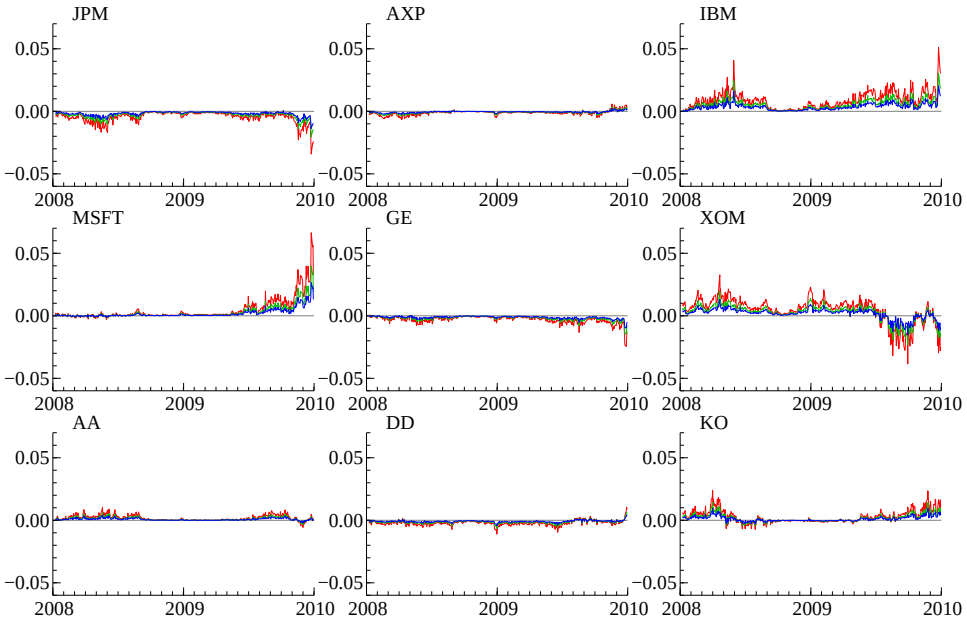


Figure 6: Time series plot of the portfolio weight for mean-variance strategy: $\gamma = 6$ (red), $\gamma = 10$ (green) and $\gamma = 15$ (blue).

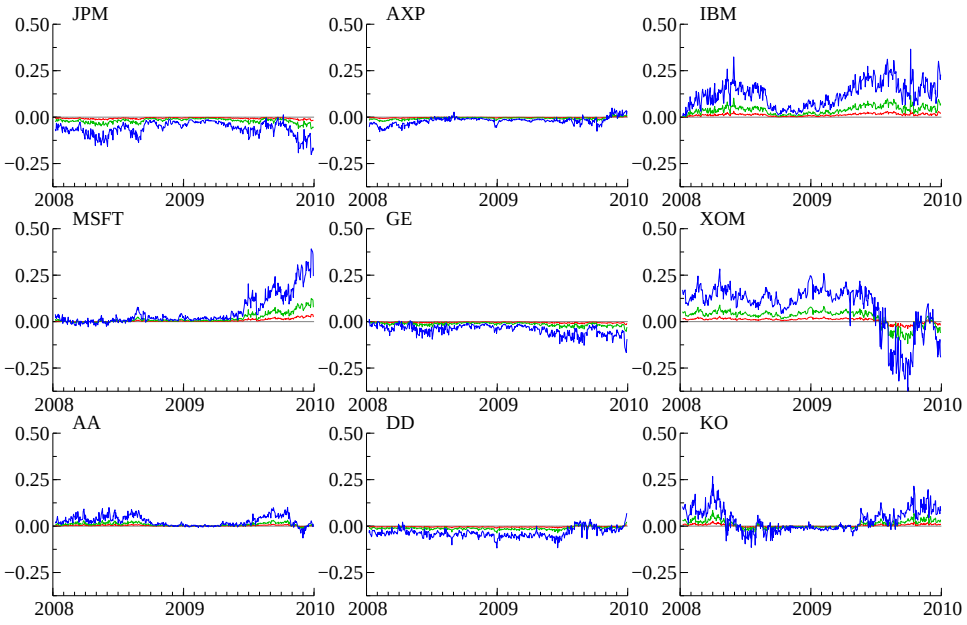


Figure 7: Time series plot of the portfolio weight for maximum return strategy: $\sigma_p^{2*} = 0.001$ (red), $\sigma_p^{2*} = 0.01$ (green) and $\sigma_p^{2*} = 0.1$ (blue).

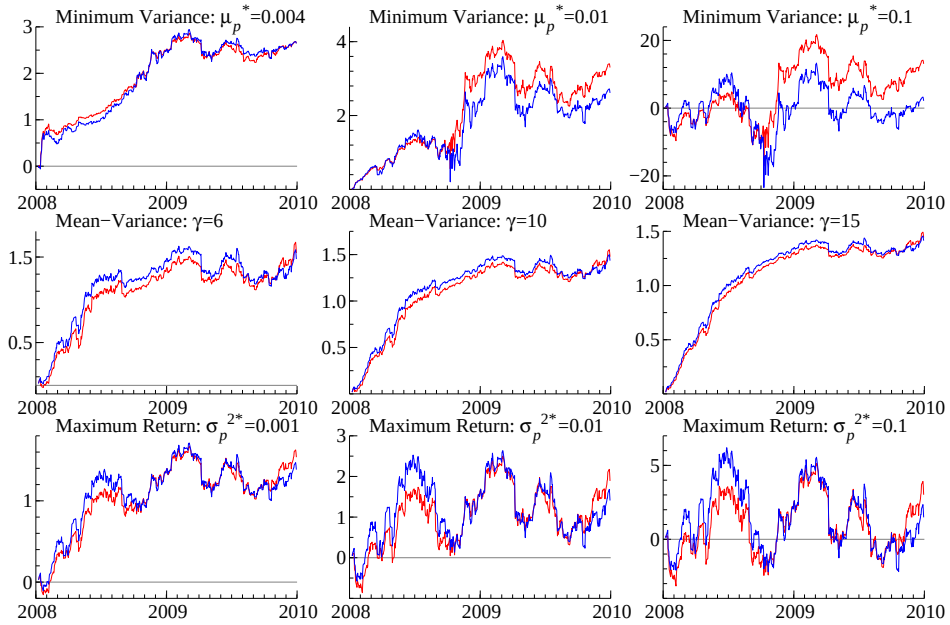


Figure 8: Cumulative realized return: 4B-RDESV-LE (red) and 1B-RDESV-LE (blue).

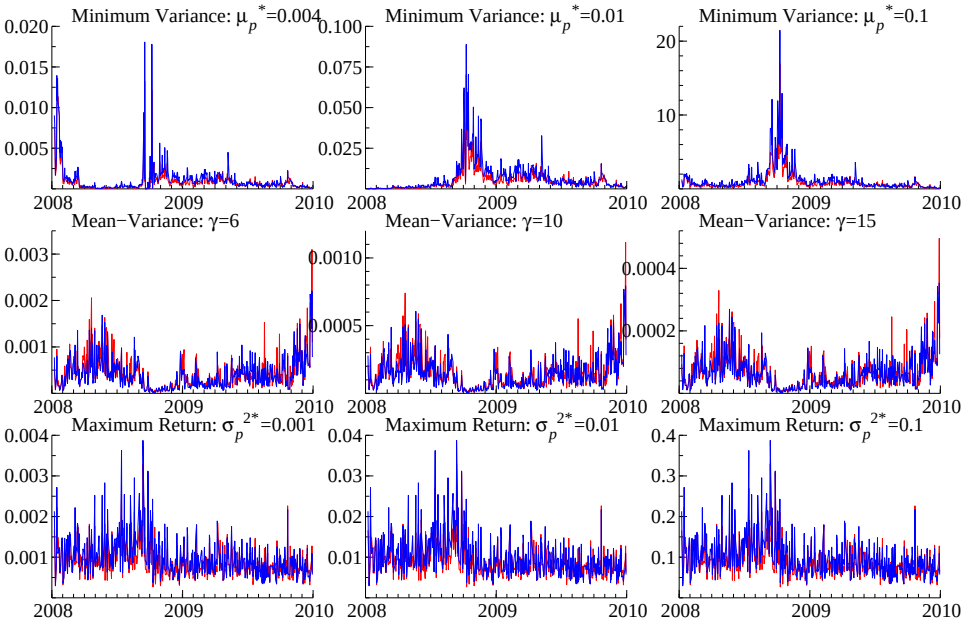


Figure 9: Realized portfolio variance 4B-RDESV-LE (red) and 1B-RDESV-LE (blue).