

CIRJE-F-979

**Cholesky Realized Stochastic Volatility Model**

Shinichiro Shirota

Duke University

Yasuhiro Omori

The University of Tokyo

Hedibert. F. Lopes

Inspere Institute of Education and Research

Haixiang Piao

Nippon Life Insurance Company

July 2015

CIRJE Discussion Papers can be downloaded without charge from:

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

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

# Cholesky Realized Stochastic Volatility Model\*

Shinichiro Shirota,<sup>†</sup> Yasuhiro Omori,<sup>‡</sup> Hedibert. F. Lopes,<sup>§</sup>  
and Haixiang Piao.<sup>¶</sup>

July 8, 2015

Abstract

Multivariate stochastic volatility models are expected to play important roles in financial applications such as asset allocation and risk management. However, these models suffer from two major difficulties: (1) there are too many parameters to estimate using only daily asset returns and (2) estimated covariance matrices are not guaranteed to be positive definite. Our approach takes advantage of realized covariances to attain the efficient estimation of parameters by incorporating additional information for the co-volatilities, and considers Cholesky decomposition to guarantee the positive definiteness of the covariance matrices. In this framework, we propose a flexible modeling for stylized facts of financial markets such as dynamic correlations and leverage effects among volatilities. Taking a Bayesian approach, we describe Markov Chain Monte Carlo implementation with a simple but efficient sampling scheme. Our model is applied to nine U.S. stock returns data, and the model comparison is conducted based on portfolio performances.

*Keywords:* Cholesky stochastic volatility model; Dynamic leverage; Single move sampler; Realized covariance

---

\* *Corresponding author:* Shinichiro Shirota, Department of Statistical Science, Duke University, Box 90251, Durham NC 27708-0251, [ss571@stat.duke.edu](mailto:ss571@stat.duke.edu)

<sup>†</sup>Department of Statistical Science, Duke University. E-mail: [ss571@stat.duke.edu](mailto:ss571@stat.duke.edu).

<sup>‡</sup>Faculty of Economics, The University of Tokyo, Japan. E-mail: [omori@e.u-tokyo.ac.jp](mailto:omori@e.u-tokyo.ac.jp).

<sup>§</sup>Inspere Institute of Education and Research, Brazil. E-mail: [hedibertFL@insper.edu.br](mailto:hedibertFL@insper.edu.br).

<sup>¶</sup>Nippon Life Insurance Company.

# 1 Introduction

Modeling time-varying co-volatilities of multiple asset returns has become increasingly important in recent years for financial risk management. Although there is by now large literature on univariate volatility models such as GARCH and stochastic volatility (SV) models, their extension to multivariate models has not been straightforward. The major concern in this field is the flexible and intuitive modeling of time-varying variances and correlations, but these multivariate volatility models suffer from two major difficulties: (1) there are too many parameters to estimate using only daily asset returns and (2) estimated covariance matrices are not guaranteed to be positive definite. This paper proposes a promising solution to overcome these problems for a multivariate SV model using a Cholesky decomposition of the covariance matrices and additional information of realized covariances.

Among multivariate SV models, factor models are intuitive to describe the high dimensional asset returns with common volatility dynamics and have been successful to reduce the number of parameters to estimate (Harvey et al. (1994), Pitt and Shephard (2003), Aguilar and West (2000), Chib et al. (2006) and Lopes and Carvalho (2007)). However, we have to decide the number of factor a priori and choose the factor structure for parameter identification (Geweke and Zhou (1996), Aguilar and West (2000), Lopes and West (2004)). For more flexible modeling, this paper considers Cholesky decomposition of the covariance matrices, which guarantees the positive definiteness.

The Cholesky stochastic volatility (CSV) models introduce the dynamic structure to the diagonal and off-diagonal components of Cholesky decomposed covariance matrices. Pourahmadi (1999) proposes to model components of (time-invariant) Cholesky decomposed inverse covariance matrices as a linear function of predictors, and Fox and Dunson (2011) propose a Bayesian nonparametric approach for the covariance regression. For the time-varying covariance structure, Lopes et al. (2012) consider the Cholesky stochastic volatility model which incorporate dynamic structures to each diagonal and off-diagonal components of Cholesky decomposed covariance matrices. This approach enables us to utilize parallel computing methods due to the conditional independence property of each rows of decomposed components, thus it is efficient and fast even in high dimensional cases. In the class of GARCH models (see e.g., Bauwens et al. (2006) for a recent survey), Dellaportas and Pourahmadi

(2012) propose Cholesky-GARCH model where conditional variances are assumed to follow GARCH(1,1) process, but non-diagonal elements of the lower triangular matrix of Cholesky decomposition are constant over time.

On the other hand, high-frequency (intraday) data of asset prices have become available recently in the financial market, and various realized measures have been proposed to estimate daily volatilities, which attracts attention in financial econometrics. They are, for example, realized volatility, realized kernel and realized covariance (Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2001), Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen and Shephard (2004)). These measures have more information regarding true volatilities or covariance matrices than those estimators based solely on daily returns, but have some biases primarily due to market microstructure noises and non-trading hours. To adjust these biases in SV models, realized stochastic volatility (RSV) models are proposed by Takahashi et al. (2009) where they consider simultaneous modeling of daily returns and realized volatilities, since daily returns are less subject to these biases. The RSV models are expected to provide more accurate estimates of volatilities while removing biases of realized measures than conventional SV models with solely daily returns.

Several econometric models for realized covariances have been proposed in the literature. Jin and Maheu (2013) incorporates the realized covariance information into Wishart Autoregressive processes by extending models of Philipov and Glickman (2006) and Asai and McAleer (2009). Windle and Carvalho (2014) propose a state space model whose observations and latent states take values on the manifold of symmetric positive-definite matrices. This paper considers multivariate RSV models in the context of CSV models. We incorporate realized covariances into the CSV models as additional information resources for true covariance matrices. Furthermore, we extend these original models so as to incorporate dynamic leverage effects and correlations among volatilities. In empirical studies of nine U.S. stock returns, we compare our proposed model with standard CSV models based on portfolio performances with several strategies.

The organization of the paper is as follows. Section 2 introduces CSV and RSV models. In Section 3, we describe Bayesian estimation procedure using Markov chain Monte Carlo (MCMC) simulation. We also discuss prior specifications for each parameters. Section 4 illustrate our estimation algorithm using simulated data. In Section 5, we apply our model to nine U.S. stock returns data and show the empirical

estimation results. Finally, in Section 6, we compare the performances of proposed models with those of standard CSV models based on the different types of portfolio strategies.

## 2 Cholesky realized stochastic volatility model

In this section we introduce our Cholesky realized SV (CRSV) model. However, we start by briefly reviewing Lopes et al. (2012) Cholesky stochastic volatility (CSV) model and Barndorff-Nielsen et al. (2011) multivariate realized SV (RSV) model. These models tackle the two major difficulties in modeling multivariate volatility: positive-definiteness of estimated covariance matrices and the curse of dimensionality when estimating highly parameterized models for daily asset returns.

### 2.1 Cholesky stochastic volatility model

The Cholesky decomposition is unique and guarantees positive definiteness of the covariance matrix when the diagonal components of the decomposed covariance are positive. More specifically, let  $\mathbf{y}_t = (y_{1t}, \dots, y_{pt})'$  be a  $p$ -dimensional vector of assets returns, such that

$$\mathbf{y}_t \sim \mathcal{N}(\mathbf{m}_t, \Sigma_t) \quad (1)$$

where  $\mathbf{m}_t = (m_{1t}, \dots, m_{pt})'$  and  $\Sigma_t$  are, respectively, the mean vector and the covariance matrix at time  $t$ . We consider the Cholesky decomposition of  $\Sigma_t$  as follows, for  $t = 1, \dots, n$ :

$$\Sigma_t = \mathbf{H}_t^{*-1} \mathbf{V}_t \mathbf{H}_t^{*-1'}, \quad (2)$$

where  $\mathbf{V}_t = \text{diag}\{\exp(h_{11,t}), \dots, \exp(h_{pp,t})\}$  and

$$\mathbf{H}_t^* = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -h_{21,t} & 1 & 0 & \dots & 0 \\ -h_{31,t} & -h_{32,t} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -h_{p1,t} & -h_{p2,t} & \dots & -h_{p,p-1,t} & 1 \end{pmatrix}, \quad (3)$$

**Recursive conditional regressions.** It follows that

$$\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t) \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_t),$$

where the quantities  $h_{ij,t}$ 's ( $i > j$ ) are the regression coefficients in  $p$  recursive conditional regressions:

$$y_{it} | \{y_{jt}\}_{j=1}^{i-1} \sim \mathcal{N} \left( m_{i,t} + \sum_{j=1}^{i-1} h_{ij,t}(y_{jt} - m_{jt}), \exp(h_{ii,t}) \right), \quad i = 1, \dots, p.$$

Lopes et al. (2012) propose the Cholesky stochastic volatility (CSV) models by assuming  $\mathbf{m}_t = \mathbf{0}$  and autoregressive processes for  $\mathbf{h}_t = (h_{11,t}, h_{22,t}, \dots, h_{pp,t})'$  and  $\mathbf{h}_t^* = (h_{21,t}, h_{31,t}, \dots, h_{pp-1,t})'$ ,

$$\mathbf{h}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{D}), \quad (4)$$

$$\mathbf{h}_{t+1}^* = \boldsymbol{\mu}^* + \boldsymbol{\Phi}^*(\mathbf{h}_t^* - \boldsymbol{\mu}^*) + \boldsymbol{\eta}_t^*, \quad \boldsymbol{\eta}_t^* \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{D}^*), \quad (5)$$

where  $\boldsymbol{\mu} = (\mu_{11}, \mu_{22}, \dots, \mu_{pp})'$ ,  $\boldsymbol{\mu}^* = (\mu_{21}, \mu_{31}, \dots, \mu_{pp-1})'$ ,  $\boldsymbol{\Phi} = \text{diag}(\boldsymbol{\phi})$ ,  $\boldsymbol{\phi} = (\phi_{11}, \phi_{22}, \dots, \phi_{pp})'$ ,  $\boldsymbol{\Phi}^* = \text{diag}(\boldsymbol{\phi}^*)$ ,  $\boldsymbol{\phi}^* = (\phi_{21}, \phi_{31}, \dots, \phi_{pp-1})'$ ,  $\mathbf{D} = \text{diag}(\tau_{11}^2, \tau_{22}^2, \dots, \tau_{pp}^2)$ , and  $\mathbf{D}^* = \text{diag}(\tau_{21}^2, \tau_{31}^2, \dots, \tau_{pp-1}^2)$ . They implement the highly efficient estimation based on the mixture sampler (see, Kim et al. (1998) and Omori et al. (2007)) by using the normal mixture approximation. The CSV model enables us to utilize parallel computing procedures for estimating each components due to the conditional independence of each rows.

## 2.2 Realized stochastic volatility model

Another major difficulty in the multivariate volatility model is that there are too many parameters to estimate using only daily asset returns. In addition to daily returns, recently, the high frequency datasets have become available and attracted attentions in financial econometrics. Using high frequency data, Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002) propose more accurate volatility estimator which is called the realized volatility. However, in the presence of the market microstructure noises, it becomes a biased estimator and Barndorff-Nielsen et al. (2008) further propose the realized kernel which is a robust estimator to such noises. Several other extensions of these estimators have been proposed under different assumptions for the stochastic processes of assets. For multivariate asset returns, the realized covariance,  $\mathbf{RC}_t$ , is defined by

$$\mathbf{RC}_t = \sum_{j=1}^m \mathbf{r}_{j,t} \mathbf{r}_{j,t}'$$

where  $\mathbf{r}_{j,t} = \mathbf{y}_{\frac{j}{m},t} - \mathbf{y}_{\frac{j-1}{m},t}$ ,  $j = 1, \dots, m$ ,  $t = 1, \dots, n$ , and  $\mathbf{y}_{\frac{j}{m},t}$  is a  $p \times 1$  log-price vector at  $j$ -th time of the day  $t$ . Barndorff-Nielsen and Shephard (2004) show that, in the absence of the market microstructure noise, it converges to the quadratic covariation of  $\mathbf{y}$  as  $m \rightarrow \infty$ . Barndorff-Nielsen et al. (2011) propose the multivariate realized kernel which is robust to such noises. Although various extensions of these estimators have been proposed, their properties depend on assumptions imposed on the price processes.

The realized measures have more information regarding true volatilities and covariance matrices, while there may be a bias due to the market microstructure noise. On the other hand, daily returns have less information about true volatilities, while they are less affected by these noises. Thus Takahashi et al. (2009) propose realized stochastic volatility models (RSV), which is the simultaneous modeling of daily returns and realized volatilities, and there is a growing literature on similar simultaneous modeling (Koopman and Scharth (2013), Venter and de Jongh (2014), Shirota et al. (2014) and Zheng and Song (2014)). In addition to the standard stochastic volatility model with leverage

$$y_t = \exp\{h_t/2\}\epsilon_t, \quad (6)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad (7)$$

for  $t = 1, \dots, n$  and

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \text{i.i.d. } \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{pmatrix}\right), \quad (8)$$

where  $y_t$  is an asset returns at time  $t$ , Takahashi et al. (2009) consider another measurement equation for the logarithm of the realized volatility at time  $t$

$$x_t = \xi + h_t + u_t, \quad (9)$$

where  $u_t \sim \text{i.i.d. } \mathcal{N}(0, \sigma_u^2)$  and independent of  $(\epsilon_t, \eta_t)'$ . This model automatically adjusts the bias of the realized measure without any additional adjustment such as selecting the optimal sampling frequency to compute the realized volatility. The parameter  $\xi$  is the bias adjustment term to account for the effects of the market microstructure noise and non-trading hours simultaneously. When it is negative (positive), the realized volatility is considered to underestimate (overestimate) the latent volatility. Although we could extend it by replacing  $h_t$  with  $\psi h_t$  in (9), where  $\psi$  is

another adjustment coefficient, it has been pointed out that this extension does not necessarily improve the forecasting performances in the empirical studies.

### 2.3 Cholesky realized stochastic volatility model

We now extend the CSV model in two directions. Firstly, we consider additional measurement equations to incorporate the information of the realized measures. Secondly, we incorporate leverage effects that are often observed to exist in the empirical studies of stock markets.

Consider the Cholesky decomposition of the realized covariance  $\mathbf{RC}_t$  given by

$$\mathbf{RC}_t = \mathbf{X}_t^{*-1} \text{diag}\{\exp(x_{11,t}), \exp(x_{22,t}), \dots, \exp(x_{pp,t})\} \mathbf{X}_t^{*-1'}$$

where

$$\mathbf{X}_t^* = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -x_{21,t} & 1 & 0 & \cdots & 0 \\ -x_{31,t} & -x_{32,t} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -x_{p1,t} & -x_{p2,t} & \cdots & -x_{pp-1,t} & 1 \end{pmatrix},$$

and defining  $\mathbf{x}_t = (x_{11,t}, x_{22,t}, \dots, x_{pp,t})'$  and  $\mathbf{x}_t^* = (x_{21,t}^*, x_{31,t}^*, \dots, x_{pp-1,t}^*)'$ . The information of the realized covariances are added to the following measurement equations:

$$\mathbf{x}_t = \boldsymbol{\xi} + \mathbf{h}_t + \mathbf{u}_t, \quad \mathbf{u}_t \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{C}), \quad (10)$$

$$\mathbf{x}_t^* = \boldsymbol{\xi}^* + \mathbf{h}_t^* + \mathbf{u}_t^*, \quad \mathbf{u}_t^* \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{C}^*), \quad (11)$$

where  $\boldsymbol{\xi} = (\xi_{11}, \xi_{22}, \dots, \xi_{pp})'$ ,  $\boldsymbol{\xi}^* = (\xi_{21}, \xi_{31}, \dots, \xi_{pp-1})'$ ,  $\mathbf{C} = \text{diag}(\sigma_{u,11}^2, \sigma_{u,22}^2, \dots, \sigma_{u,pp}^2)'$ , and  $\mathbf{C}^* = \text{diag}(\sigma_{u,21}^2, \sigma_{u,31}^2, \dots, \sigma_{u,pp-1}^2)'$ .

The  $\mathbf{u}_t$  and  $\mathbf{u}_t^*$  are measurement error terms for realized covariances, which are assumed to be independent of each other. Further, we assume the common bias  $\xi$  for realized covariances, *i.e.*,  $\boldsymbol{\xi} = \xi \mathbf{1}_p$  where  $\mathbf{1}_p$  is a  $p \times 1$  vector with all elements equal to one so that

$$\mathbf{RC}_t = \exp(\xi) \mathbf{X}_t^{*-1} \text{diag}(\exp(h_{11,t} + u_{1t}), \dots, \exp(h_{pp,t} + u_{pt})) \mathbf{X}_t^{*-1'}$$

We note that the common scale bias  $\exp(\xi)$  does not affect the correlation structure of realized covariances, while the bias adjustment terms for off-diagonal components  $\boldsymbol{\xi}^*$  may affect the correlation structure of covariance matrices. Also we could consider



additional bias adjustment coefficient matrices for  $\mathbf{h}$  and  $\mathbf{h}^*$ , but assume that they are unit matrices for simplicity.

Second, we consider leverage and cross leverage effects which are observed to exist in stock markets to improve the model predictive performance (e.g., Ishihara and Omori (2012), Ishihara et al. (2014), Trojan (2014)). The cross leverage effects are defined as the negative correlation between the  $i$ -th asset return at time  $t$  and the  $j$ -th log volatility at time  $t + 1$  for  $i \neq j$ . Given  $\mathbf{y}_t, \mathbf{h}_t$  and  $\mathbf{m}_t$ , we incorporate dynamic leverage effects through the following equations:

$$\mathbf{h}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu}) + \mathbf{R}\mathbf{V}_t^{-1/2}\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t) + \boldsymbol{\zeta}_t,$$

where  $\boldsymbol{\zeta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ , with  $\boldsymbol{\Omega} = \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'}$  and

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -s_{21} & 1 & 0 & \cdots & 0 \\ -s_{31} & -s_{32} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -s_{p1} & -s_{p2} & \cdots & -s_{pp-1} & 1 \end{pmatrix}.$$

The  $p \times p$  matrix  $\mathbf{R}$  captures the influence of daily returns at time  $t$  on the diagonal elements of covariance matrices at time  $t + 1$ . We could consider similar effects for  $h_{t+1}^*$ , but their interpretations are not clear in empirical studies. Moreover, they will introduce a large number of many parameters (order  $p^4$ ) to estimate such effects. Thus we focus on leverage effects only for  $\mathbf{h}_{t+1}$ . Further, the matrix  $\mathbf{S}$  describes the dependence among  $\mathbf{h}_{t+1}$  in a sequential regression form. Since time-varying variances of multiple asset returns move in the similar direction in financial market, we expect them to have high correlation and hence incorporate the correlation through the matrix  $\mathbf{S}$ . Setting  $\mathbf{R} = \mathbf{O}$  and  $\mathbf{S} = \mathbf{I}_p$  where  $\mathbf{I}_p$  denotes a  $p \times p$  identity matrix, it reduces to the model without leverage as in Lopes et al. (2012).

Finally, we assume a random walk process for the mean process of  $\mathbf{y}_t$  to allow possible dynamic movement in mean levels rather than setting  $\mathbf{m}_t \equiv \mathbf{0}$ . We note that the introduction of such a mean process is important to improve the portfolio performance in empirical studies.

In summary, we propose the Cholesky realized stochastic volatility (CRSV) model

given by three measurement equations

$$\mathbf{y}_t = \mathbf{m}_t + \mathbf{H}_t^{*-1} \mathbf{V}_t^{1/2} \boldsymbol{\epsilon}_t, \quad (12)$$

$$\mathbf{x}_t = \xi \mathbf{1}_p + \mathbf{h}_t + \mathbf{u}_t, \quad \mathbf{u}_t \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{C}), \quad (13)$$

$$\mathbf{x}_t^* = \boldsymbol{\xi}^* + \mathbf{h}_t^* + \mathbf{u}_t^*, \quad \mathbf{u}_t^* \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{C}^*), \quad (14)$$

for  $t = 1, \dots, n$  and three state equations

$$\mathbf{h}_{t+1} = \boldsymbol{\mu} + \Phi(\mathbf{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad (15)$$

$$\mathbf{h}_{t+1}^* = \boldsymbol{\mu}^* + \Phi^*(\mathbf{h}_t^* - \boldsymbol{\mu}^*) + \boldsymbol{\eta}_t^*, \quad \boldsymbol{\eta}_t^* \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{D}^*), \quad (16)$$

$$\mathbf{m}_{t+1} = \mathbf{m}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_m), \quad \boldsymbol{\Omega}_m = \text{diag}(\sigma_{m1}^2, \dots, \sigma_{mp}^2), \quad (17)$$

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

$$\begin{pmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \text{i.i.d. } \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \mathbf{I}_p & \mathbf{R}' \\ \mathbf{R} & \mathbf{R}\mathbf{R}' + \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'} \end{pmatrix}\right).$$

For initial values of state variables, we assume, for simplicity,

$$\mathbf{h}_1 \sim \mathcal{N}(\boldsymbol{\mu}, \lambda \mathbf{S}^{-1} \mathbf{D} \mathbf{S}^{-1'}), \quad \mathbf{h}_1^* \sim \mathcal{N}(\boldsymbol{\mu}^*, \lambda^* \mathbf{D}^*), \quad \mathbf{m}_1 \sim \mathcal{N}(\mathbf{0}^*, \lambda_m \boldsymbol{\Omega}_m),$$

where  $\lambda, \lambda^*, \lambda_m$  are set to some known large constant.

## 3 Posterior inference

### 3.1 Prior distributions

We consider the vague prior for each component of parameters if we do not have sufficient prior information on parameters. For example, we assume univariate normal distribution for  $\xi$  and independent multivariate normal distributions for  $\boldsymbol{\xi}^*$ ,  $\boldsymbol{\mu}$  and  $\boldsymbol{\mu}^*$  with large variance. For the vectorized components of  $\mathbf{S}$  and  $\mathbf{R}$ , we assume independent multivariate normal distribution with large variance. Furthermore, for the components of  $\mathbf{C}$ ,  $\mathbf{C}^*$ ,  $\mathbf{D}$ ,  $\mathbf{D}^*$  and  $\boldsymbol{\Omega}_m$ , we consider inverse gamma distribution with large variance for the conjugacy property. For the prior distribution for  $\phi$ , we assume Beta distribution for  $(\phi + 1)/2$  that is often used in the previous empirical studies for the univariate models.

As for the prior distribution of  $\phi^*$ , we need more careful discussion. Lopes et al. (2012) introduce variable selection priors for  $\phi^*$ . This approach is flexible in the sense

that it includes several dynamic patterns for  $\mathbf{h}_t^*$  such as the constant plus noise, and the random walk process. Since the constant plus noise process has almost similar path to that of the random walk with a small error variance, we consider the prior distribution for  $\phi^*$  to include random walk processes for  $\mathbf{h}_t^*$ . Noting that the random walk process is nonstationary with the unit root, we considered three types of unit root priors for  $\phi^*$  proposed in the literature: (1) the beta prior defined not on  $(-1, 1)$ , but on a slightly extended range so as to allow for slightly explosive values (Lubrano (1995)):

$$\pi(\phi_i^*) \propto (1 + v - \phi_i^{*2})^{-1/2}, \quad -\sqrt{1+v} < \phi_i^* < \sqrt{1+v},$$

where  $v$  is a small positive number, (2) the reference prior (Berger and Yang (1994)) given by

$$\pi(\phi_i^*) = \begin{cases} \left(2\pi\sqrt{1-\phi_i^{*2}}\right)^{-1}, & |\phi_i^*| < 1, \\ \left(2\pi|\phi_i^*|\sqrt{\phi_i^{*2}-1}\right)^{-1}, & |\phi_i^*| \geq 1, \end{cases}$$

and (3) the uniform prior on  $(-\sqrt{1+v}, \sqrt{1+v})$ ,  $\phi_i^* \sim \mathcal{U}(-\sqrt{1+v}, \sqrt{1+v})$ , by setting  $\varphi = \frac{\sqrt{1+v}+\phi}{2\sqrt{1+v}} \sim \mathcal{B}(1, 1)$ . In empirical studies, we take  $v = 0.3$  since it is sufficient to cover the support of posterior probability density function of  $\phi_i^*$ , and any differences are not found in the sensitivity analysis for these three different prior settings. Thus we focus on the simple uniform prior on  $(-\sqrt{1+v}, \sqrt{1+v})$  with  $v = 0.3$ .

Since we include the realized covariance as an additional source of information of covariance matrices in CRSV models, the estimation of  $\phi^*$  is more stable and efficient than in CSV models. However, the off-diagonal components of realized covariance are still noisy as shown in the empirical studies, and allowing for  $\mathbf{h}_t^*$  to follow the smooth random walk process is more appropriate to estimate  $\phi_*$  rather than imposing the conventional stationarity conditions.

### 3.2 Customized MCMC scheme

Taking Bayesian approach, we implement Markov chain Monte Carlo simulation and estimate posterior distributions of parameters to conduct statistical inferences. Let  $\mathbf{z} = \{\mathbf{y}, \mathbf{x}\}$  denote the set of observations where  $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ ,  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ , and  $\mathbf{x}^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_n^*\}$ . Further let  $\theta$  denote the set of parameters  $\{\xi, \xi^*, \mu, \mu^*, \phi, \phi^*, \mathbf{R}, \mathbf{S}, \mathbf{C}, \mathbf{C}^*, \mathbf{D}, \mathbf{D}^*, \Omega_m\}$ , and define the set of latent variables

as  $\mathbf{m} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n\}$ ,  $\mathbf{h} = \{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n\}$ , and  $\mathbf{h}^* = \{\mathbf{h}_1^*, \mathbf{h}_2^*, \dots, \mathbf{h}_n^*\}$ . We generate random samples from the posterior distribution with the probability density function  $\pi(\boldsymbol{\theta}, \mathbf{m}, \mathbf{h}, \mathbf{h}^* | \mathbf{z})$  given by

$$\begin{aligned}
& \log \pi(\boldsymbol{\theta}, \mathbf{m}, \mathbf{h}, \mathbf{h}^* | \mathbf{z}) \tag{18} \\
&= \text{const.} + \log \pi(\boldsymbol{\theta}) - \frac{1}{2} \sum_{t=1}^n \sum_{i=1}^p h_{ii,t} - \frac{n}{2} \sum_{i=1}^p \sum_{j=1}^i \log \sigma_{u,ij}^2 - \frac{n}{2} \sum_{i=1}^p \log \sigma_{mi}^2 - \frac{n}{2} \sum_{i=1}^p \sum_{j=1}^i \log \tau_{ij}^2 \\
&\quad - \frac{1}{2} \sum_{t=1}^n (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^* \mathbf{V}_t^{-1} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) - \frac{1}{2} \sum_{t=1}^n (\mathbf{x}_t - \xi \mathbf{1}_p - \mathbf{h}_t)' \mathbf{C}^{-1} (\mathbf{x}_t - \xi \mathbf{1}_p - \mathbf{h}_t) \\
&\quad - \frac{1}{2} \sum_{t=1}^n (\mathbf{x}_t^* - \boldsymbol{\xi}^* - \mathbf{h}_t^*)' \mathbf{C}^{*-1} (\mathbf{x}_t^* - \boldsymbol{\xi}^* - \mathbf{h}_t^*) - \frac{1}{2} \sum_{t=1}^n (\mathbf{m}_{t+1} - \mathbf{m}_t)' \boldsymbol{\Omega}_m^{-1} (\mathbf{m}_{t+1} - \mathbf{m}_t) \\
&\quad - \frac{1}{2} \sum_{t=1}^{n-1} \{ \mathbf{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu}) - \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \}' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \\
&\quad \quad \quad \{ \mathbf{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu}) - \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \} \\
&\quad - \frac{1}{2} \sum_{t=1}^{n-1} \{ \mathbf{h}_{t+1}^* - \boldsymbol{\mu}^* - \boldsymbol{\Phi}^*(\mathbf{h}_t^* - \boldsymbol{\mu}^*) \}' \mathbf{D}^{*-1} \{ \mathbf{h}_{t+1}^* - \boldsymbol{\mu}^* - \boldsymbol{\Phi}^*(\mathbf{h}_t^* - \boldsymbol{\mu}^*) \} \\
&\quad - \frac{1}{2\lambda_m} \mathbf{m}_1' \boldsymbol{\Omega}^{-1} \mathbf{m}_1 - \frac{1}{2\lambda} (\mathbf{h}_1 - \boldsymbol{\mu})' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} (\mathbf{h}_1 - \boldsymbol{\mu}) - \frac{1}{2\lambda^*} (\mathbf{h}_1^* - \boldsymbol{\mu}^*)' \mathbf{D}^{*-1} (\mathbf{h}_1^* - \boldsymbol{\mu}^*),
\end{aligned}$$

where  $\pi(\boldsymbol{\theta})$  denote the prior density function of  $\boldsymbol{\theta}$ .

The MCMC algorithm is implemented in ten blocks:

1. Initialize  $\boldsymbol{\theta}, \mathbf{m}, \mathbf{h}$  and  $\mathbf{h}^*$ .
2. Generate  $\mathbf{h} | \boldsymbol{\theta}, \mathbf{m}, \mathbf{h}^*, \mathbf{z}$
3. Generate  $\mathbf{h}^* | \boldsymbol{\theta}, \mathbf{m}, \mathbf{h}, \mathbf{z}$
4. Generate  $\mathbf{m} | \boldsymbol{\theta}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$
5. Generate  $(\xi, \boldsymbol{\xi}^*, \boldsymbol{\mu}, \boldsymbol{\mu}^*) | \boldsymbol{\theta}_{\setminus(\xi, \boldsymbol{\xi}^*, \boldsymbol{\mu}, \boldsymbol{\mu}^*)}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$
6. Generate  $(\boldsymbol{\phi}, \boldsymbol{\phi}^*) | \boldsymbol{\theta}_{\setminus(\boldsymbol{\phi}, \boldsymbol{\phi}^*)}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$
7. Generate  $\mathbf{R} | \boldsymbol{\theta}_{\setminus \mathbf{R}}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$
8. Generate  $\mathbf{S} | \boldsymbol{\theta}_{\setminus \mathbf{S}}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$

9. Generate  $(\mathbf{C}, \mathbf{C}^*, \mathbf{D}, \mathbf{D}^*, \boldsymbol{\Omega}_m) | \boldsymbol{\theta}_{\setminus (\mathbf{C}, \mathbf{C}^*, \mathbf{D}, \mathbf{D}^*, \boldsymbol{\Omega}_m)}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}$

10. Go to Step 2

where  $\boldsymbol{\theta}_{\setminus \boldsymbol{\delta}}$  denote the parameter  $\boldsymbol{\theta}$  excluding  $\boldsymbol{\delta}$ . Details of the MCMC algorithm are described in Appendix A.

## 4 Simulation exercise

**Simulation set up.** To illustrate the MCMC estimation for our proposed model, we generate  $n = 2,000$  observations where the dimension of asset returns is  $p = 9$ , which is the same dimension as in empirical studies. We set the true values of parameters as  $\mu_{ii} = \mu_{ij} = 0$ ,  $\xi = \xi_{ij} = -0.5$ ,  $\phi_{ii} = 0.97$ ,  $\phi_{ij} = 0.99$ ,  $s_{ij} = 0.2$ ,  $\tau_{ii} = 0.1$ ,  $\tau_{ij} = 0.01$ ,  $\sigma_{u,ii} = 0.1$ ,  $\sigma_{u,ij} = 0.1$ ,  $\sigma_{mi} = 0.002$ , for  $i = 1, \dots, p$ , and  $j = 1, \dots, i - 1$ , and  $\mathbf{R} = \{-0.05 \times \mathbf{1}_p, \mathbf{O}\}$ .

**Prior set up.** The prior distributions of parameters are assumed to be

$$\begin{aligned} \mu_{ii} &\sim \mathcal{N}(0, 10^2), & \mu_{ij} &\sim \mathcal{N}(0, 10^2), & \xi &\sim \mathcal{N}(0, 10^2), & \xi_{ij} &\sim \mathcal{N}(0, 10^2), \\ \phi_{ii} &\sim \mathcal{U}(-1, 1), & \phi_{ij} &\sim \mathcal{U}(-\sqrt{1.3}, \sqrt{1.3}), & s_{ij} &\sim \mathcal{N}(0, 10^2), \\ \tau_{ii}^2 &\sim \mathcal{IG}(5, 0.05), & \sigma_{u,ii}^2 &\sim \mathcal{IG}(5, 0.05), & \sigma_{u,ij}^2 &\sim \mathcal{IG}(5, 0.05), \\ \tau_{ij}^2 &\sim \mathcal{IG}(10^{-4}, 10^{-4}), & \sigma_{mi}^2 &\sim \mathcal{IG}(10^{-6}, 10^{-6}), \end{aligned}$$

for  $i = 1, \dots, p$ , and  $j = 1, \dots, i - 1$ , and  $r_{ij} \sim \mathcal{N}(0, 10^2)$  for  $i, j = 1, \dots, p$ . We assume fairly flat prior distributions for  $\boldsymbol{\mu}$ ,  $\boldsymbol{\mu}^*$ ,  $\xi$ ,  $\boldsymbol{\xi}^*$ ,  $\mathbf{r}$ ,  $\mathbf{s}$  and flat prior distributions for  $\boldsymbol{\phi}$ ,  $\boldsymbol{\phi}^*$ . For error variances of the diagonal components ( $\tau_{ii}^2$ ) and realized measures ( $\sigma_{u,ii}^2, \sigma_{u,ij}^2$ ), we assume inverse gamma prior distributions  $\mathcal{IG}(5, 0.05)$  which are often used in the literature. For error variances of no-diagonal components ( $\tau_{ij}^2$ ) and the mean return process  $\mathbf{m}_t$ , we put  $\mathcal{IG}(\epsilon, \epsilon)$  prior distributions with small  $\epsilon$  as discussed in the previous section. For the hyperparameters,  $(\lambda, \lambda^*, \lambda_m)$ , we set  $\lambda = \lambda^* = \lambda_m = 100$  to remove the influence of initial values.

**MCMC set up.** The number of iteration for MCMC is 90,000. The first 60,000 samples are discarded as the burn-in period. The detailed estimation results for each parameters are shown in Appendix A1. Tables 1 to 3 show the estimation results for  $(\boldsymbol{\mu}, \boldsymbol{\phi}, \xi, \mathbf{C}, \mathbf{D}, \boldsymbol{\Omega}_m)$ ,  $\mathbf{R}$  and  $\mathbf{S}$  for diagonal components  $\mathbf{h}$ , while Tables 4 to 8 show

the estimation results for  $(\boldsymbol{\mu}^*, \boldsymbol{\phi}^*, \boldsymbol{\xi}^*, \mathbf{C}^*, \mathbf{D}^*)$  for non-diagonal components  $\mathbf{h}^*$ . The posterior means of parameters are close to their true values and 95% credible intervals include these true values in most cases.

Figure 1 shows estimated 95% credible intervals for  $\mathbf{h}_t$  with true values. The estimated intervals cover the true values and capture the dynamic behavior of these diagonal components. Further, estimated dynamic correlations between asset returns for  $(\rho_{21,t}, \dots, \rho_{62,t})$  are shown in Figure 2. These true values are also included in 95% credible intervals and well estimated.

## 5 Application to U.S. stock returns

The proposed model is applied to daily returns and realized covariances of nine U.S. stocks ( $p = 9$ ): 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 realized covariance for these assets can be downloaded from Oxford Man Institute website<sup>1</sup> (see, Heber et al. (2009), Noureldin et al. (2012)). The daily returns for  $i$ -th stock are defined as  $y_{it} = 100 \times (\log p_{it} - \log p_{i,t-1})$ , where  $p_{it}$  is the close value of  $i$ -th asset at time  $t$ . The realized covariance is calculated via 5 minutes intraday returns with subsampling. The number of observations is  $n = 2242$  (February 1, 2001 to December 31, 2009). We note that this dataset includes the high volatility period after the financial crisis in 2008 as shown in Figure 3.

### 5.1 Estimation results

**Estimation efficiencies for CSV and CRSV models.** First, we compare estimation results of CSV and CRSV models, using the single-move sampler. The number of iteration for MCMC is 90,000. The first 60,000 samples are discarded as the burn-in period. Table 9 shows a part of estimation results for diagonal components. It is shown that the estimation of the CRSV model is more efficient than that of the CSV model in terms of inefficiency factors (IF)<sup>2</sup>. Since we incorporate diagonal and

<sup>1</sup>The dataset also includes Bank of America (BA). However, since it has the extremely high volatility period after the financial crisis, it is excluded from our empirical studies.

<sup>2</sup>The inefficiency factor is the ratio of the numerical variance of the estimate from the MCMC samples relative to that from hypothetical uncorrelated samples, and is defined as  $1 + 2 \sum_{s=1}^{\infty} \rho_s$

off-diagonal components of decomposed realized covariance as additional source of information regarding latent covariance variables. It improves sampling efficiencies as well as standard deviations of parameters.

**Estimation results for diagonal elements  $\mathbf{h}_t$ .** Table 9 shows estimation results (posterior means, posterior standard deviations, 95% credible intervals and inefficiency factors) for  $\boldsymbol{\mu}$ ,  $\boldsymbol{\phi}$ ,  $\xi$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  and  $\boldsymbol{\Omega}_m$ .

Since posterior means of  $\phi_{ii}$ 's are over 0.9, the persistence of diagonal elements  $\mathbf{h}_t$  are high for all  $i$ , but they are relatively smaller than those in univariate SV models. This is consistent with empirical studies in the past literature (Takahashi et al. (2009)). Estimates of some means of diagonal elements, such as  $\mu_{11}$  and  $\mu_{55}$ , are much larger than others (especially  $\mu_{88}$  and  $\mu_{99}$ ), suggesting that there are differences among the magnitude of conditional volatilities. The bias adjustment term  $\xi$  is estimated to be  $-0.366$  with 95% credible interval  $[-0.387, -0.345]$ . It implies the scale bias for the realized covariance is close to  $\exp(-0.366) \approx 0.693$ . This is partly because non-trading hours are not considered to calculate the realized variances. Thus, the construction of covariance estimators solely from the realized covariance would overestimate true volatilities, while one could adjust such a scale bias within CRSV models. Finally, the variances of mean process  $\mathbf{m}_t$  are estimated to be small as we expected, suggesting that there are only small fluctuations in means of  $\mathbf{y}_t$ .

Figure 4 shows estimated posterior 95% credible intervals for  $\xi \mathbf{1}_p + \mathbf{h}_t$  with blue lines, where the observed  $\mathbf{x}_t (= \xi \mathbf{1}_p + \mathbf{h}_t + \mathbf{u}_t)$  are shown with red lines. The 95% credible intervals are much narrower in comparison with the sharp fluctuation of  $\mathbf{x}_t$ , and succeeded to extract the smooth mean trends of  $\xi + h_{ii,t}$  for all  $i$  after eliminating measurement errors.

Table 11 shows posterior means of elements of  $\mathbf{R}$ , and the red figure indicates that 95% credible interval does not include 0 (in other words,  $\Pr(r_{ij} < 0 | \text{data}) > 0.975$ ). The elements of the first column of  $\mathbf{R}$  are found to be much smaller than those of other columns, which implies credible negative effects of  $y_{1t} - m_{1t}$  on  $\mathbf{h}_{t+1}$ . Figure 6 shows the time series plots of dynamic leverage effects of the first element  $\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t)$  on  $\mathbf{h}_{t+1}$ , *i.e.*, the first column  $\mathbf{l}_{1t} = (L_{11,t}, \dots, L_{91,t})'$  of  $\mathbf{L}_t = \mathbf{R}\mathbf{V}_t^{-1/2}\mathbf{H}_t^*$ . Posterior

---

where  $\rho_s$  is the sample autocorrelation at lag  $s$ . It suggests the relative number of correlated draws necessary to attain the same variance of the posterior sample mean from the uncorrelated draws (Chib (2001)).

means of all  $l_{1j,t}$ 's are negative (shown with red lines) and indicate leverage effects on conditional log volatilities. The effect  $l_{11,t}$  on the first components  $h_{11,t+1}$  is the largest in its absolute value and credible in the sense that 95% credible intervals (shown with blue lines) are below zero. Table 12 shows the posterior means of  $\mathbf{S}$ . There are some positive posterior means with  $\Pr(s_{ij} > 0 | \text{data}) > 0.975$ , and hence there is a credible positive dependence among  $\mathbf{h}_{t+1}$  given  $\mathbf{y}_t$ .

**Estimation results for non-diagonal elements  $\mathbf{h}_t^*$ .** Tables 13–17 show estimation results (posterior means, posterior standard deviations) for  $\boldsymbol{\mu}^*$ ,  $\boldsymbol{\phi}^*$ ,  $\boldsymbol{\xi}^*$ ,  $\mathbf{C}^*$  and  $\mathbf{D}^*$ . The posterior means of  $\boldsymbol{\mu}^*$  are all positive, which suggests the positive dependence among  $\mathbf{y}_t - \mathbf{m}_t$ . Taking account of their standard deviations, all but  $\mu_{9j}$ 's are credibly positive. The  $\phi_{ij}$ 's are estimated to be close to one as expected, and the error standard deviations  $\tau_{ij}$ 's are relatively small. The small  $\tau_{ij}$  indicates that the corresponding  $h_{ij,t}$  are close to constant. Figure 5 shows estimated posterior 95% credible intervals for off-diagonal components,  $\xi_{ij} + h_{ij,t}$  for  $i = 2, \dots, 6$  and  $j = 1, 2$  with blue lines, and the red lines show the corresponding off-diagonal components of realized covariance  $x_{ij,t} = \xi_{ij} + h_{ij,t} + u_{ij,t}$ . While observed  $x_{ij,t}$ 's seem to have large noises, the 95% credible intervals are much narrower and stable. This is consistent with the posterior estimates of  $\sigma_{u,ij}$  whose magnitude is much larger than that of the  $\tau_{ij}$ . We note that, for some off-diagonal components such as  $h_{61}$ ,  $h_{52}$  and  $h_{54}$ , there are sharp increases in trends after the financial crisis. Also, there seems to be no clear bias direction in  $\mathbf{x}_t^*$  since some bias terms  $\xi_{ij}$  are estimated to be positive while others are negative.

Finally, Figure 7 shows the posterior means (with red lines) and 95% credible intervals (with blue lines) for time-varying correlation  $\rho_{ij,t}$ 's ( $i = 2, \dots, 6$  and  $j = 1, 2$ ). We observe the co-movement of the correlation processes among stock returns with the upward trend after financial crisis in 2008.

## 5.2 Model Comparison based on portfolio performances

For univariate volatility models, it is straightforward to evaluate the forecasting performances: we compare the predictive mean square error (PMSE) using several kinds of loss functions (such as mean square loss, mean absolute loss and quasi likelihood loss. see Patton (2011)). However, for multivariate volatility models, it is not straight-



forward to choose appropriate distance measures between two covariance matrices. Furthermore, it certainly is not obvious that all elements of the difference should be treated as equally important. Engle and Colacito (2006), on the other hand, propose the model comparison based on the variances of portfolios given expected returns. Univariate loss functions may be used for such a comparison of variances of portfolios. In this section, we compare the forecasting performance based on the three portfolio strategies: (1) minimum-variance strategy, (2) mean-variance strategy, and (3) maximum-expected return strategy (see, *e.g.* Han (2006)) for three multivariate SV models: CSV model, CRSV model without leverage (by setting  $\mathbf{R} = \mathbf{O}$  and  $\mathbf{S} = \mathbf{I}$ ) and CRSV model.

**Predictive mean and covariance.** The procedures for volatility forecasting are as follows. We first estimate parameters using 1742 ( $n = 1742$ ) observations from February 1, 2001 to January 8, 2008, and forecast the volatility for January 9, 2008. Then we shift the sample period one day (from February 2, 2001 to January 9, 2008) to estimate parameters and forecast the volatility for January 10, 2008. We continue this one-day ahead forecast by rolling the sample period until we forecast the volatility for December 31, 2009, which results in 500 one-day ahead forecasts.

Let  $N$  denote the number of MCMC iterations used in the parameter estimation, and  $(\boldsymbol{\theta}^{(i)}, \{\mathbf{h}_t^{(i)}\}_{t=1}^n, \{\mathbf{h}_t^{*(i)}\}_{t=1}^n, \{\mathbf{m}_t^{(i)}\}_{t=1}^n)$  denote the MCMC sample of  $(\boldsymbol{\theta}, \{\mathbf{h}_t\}_{t=1}^n, \{\mathbf{h}_t^*\}_{t=1}^n, \{\mathbf{m}_t\}_{t=1}^n)$  at the  $i$ -th iteration ( $i = 1, \dots, N$ ). Further, let  $\mathbf{m}_{t+1|t} \equiv E[\mathbf{y}_{t+1}|\mathcal{F}_t]$  and  $\boldsymbol{\Sigma}_{t+1|t} \equiv \text{Var}[\mathbf{y}_{t+1}|\mathcal{F}_t]$  denote the conditional mean and covariance of stock returns  $\mathbf{y}_{t+1}$ , given the current information set  $\mathcal{F}_t$  at time  $t$ . Then, the one-step-ahead volatility forecast is obtained by adding the following several steps to each MCMC iteration:

1. Generate  $\mathbf{h}_{n+1}^{(i)}, \mathbf{h}_{n+1}^{*(i)}, \mathbf{m}_{n+1}^{(i)} | \{\mathbf{z}_t\}_{t=1}^n, \boldsymbol{\theta}^{(i)}, \{\mathbf{h}_t^{(i)}\}_{t=1}^n, \{\mathbf{h}_t^{*(i)}\}_{t=1}^n, \{\mathbf{m}_t^{(i)}\}_{t=1}^n$ .

## 2. Store

$$\begin{aligned}
\mathbf{V}_{n+1|n}^{(i)} &= \text{diag} \left( \exp(h_{11,n+1}^{(i)}), \exp(h_{22,n+1}^{(i)}), \dots, \exp(h_{pp,n+1}^{(i)}) \right), \\
\mathbf{H}_{n+1|n}^{*(i)} &= \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -h_{21,n+1}^{(i)} & 1 & 0 & \dots & 0 \\ -h_{31,n+1}^{(i)} & -h_{32,n+1}^{(i)} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -h_{p1,n+1}^{(i)} & -h_{p2,n+1}^{(i)} & \dots & -h_{pp-1,n+1}^{(i)} & 1 \end{pmatrix}, \\
\mathbf{\Omega}_m^{(i)} &= \text{diag}(\sigma_{m1}^{2(i)}, \dots, \sigma_{mp}^{2(i)}),
\end{aligned}$$

to compute the posterior predictive means of  $\mathbf{m}_{n+1|n}$  and  $\mathbf{\Sigma}_{n+1|n}$  given by

$$\begin{aligned}
\hat{\mathbf{m}}_{n+1|n} &= \frac{1}{N} \sum_{i=1}^N \mathbf{m}_{n+1|n}^{(i)}, \\
\hat{\mathbf{\Sigma}}_{n+1|n} &= \frac{1}{N} \sum_{i=1}^N \mathbf{\Sigma}_{n+1|n}^{(i)} = \frac{1}{N} \sum_{i=1}^N \left( \mathbf{H}_{n+1|n}^{*-1(i)} \mathbf{V}_{n+1|n}^{(i)} \mathbf{H}_{n+1|n}^{*-1'(i)} + \mathbf{\Omega}_m^{(i)} \right).
\end{aligned}$$

We set  $N = 3,000$  with the burn-in period 1,000 for each one-step-ahead prediction. The initial values for parameters and latent variables are set to their posterior means of previous MCMC iteration.

**Portfolio performance.** We describe three strategies: (1) minimum-variance strategy, (2) mean-variance strategy, and (3) maximum-expected return strategy. Let  $\mu_{p,t+1}$  and  $\sigma_{p,t+1}^2$  denote the conditional mean and variance of the portfolio returns,  $r_{p,t+1}$ , and let  $r_f$  denote the risk free asset return where we use the federal funds rate for  $r_f$ . If we let  $\boldsymbol{\omega}_t$  denote the vector of portfolio weights for stock returns, then

$$\begin{aligned}
\mu_{p,t+1} &= \boldsymbol{\omega}'_t \mathbf{m}_{t+1|t} + (1 - \boldsymbol{\omega}'_t \mathbf{1}) r_f, \\
\sigma_{p,t+1}^2 &= \boldsymbol{\omega}'_t \mathbf{\Sigma}_{t+1|t} \boldsymbol{\omega}_t.
\end{aligned}$$

- *Minimum-variance strategy*

This strategy is to minimize the conditional variance  $\sigma_{p,t+1}^2$  for given levels of conditional expected return  $\mu_{p,t+1} = \mu_p^*$ . The investors solve the following quadratic problem at time  $t$ :

$$\min_{\boldsymbol{\omega}_t} \sigma_{p,t+1}^2 \quad \text{s.t.} \quad \mu_{p,t+1} = \mu_p^*,$$

where  $\mu_p^*$  is the target expected return. The solution is

$$\begin{aligned}\hat{\omega}_t &= \Sigma_{t+1|t}^{-1}(\mathbf{m}_{t+1|t} - r_f \mathbf{1}) \frac{\mu_p^* - r_f}{\kappa_t}, \\ \kappa_t &= (\mathbf{m}_{t+1|t} - r_f \mathbf{1})' \Sigma_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}).\end{aligned}$$

- *Mean-variance strategy*

This strategy is to maximize an expected mean-variance utility function. Then the investors solve the following utility maximization problem,

$$\max_{\mathbf{w}_t} \left\{ \mu_{p,t+1} - \frac{\gamma}{2} \sigma_{p,t+1}^2 \right\},$$

where  $\gamma$  is the coefficient of the absolute risk aversion. The high value of the risk aversion  $\gamma$  implies that people tend to be risk averse. The solution is

$$\hat{\omega}_t = \frac{1}{\gamma} \Sigma_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}).$$

Thus the risk aversion coefficient  $\gamma$  can influence not to the relative weights of individual risky assets but to the sum of their weights.

- *Maximum expected return strategy*

This strategy is to maximize the conditional expected return  $\mu_{p,t+1}$  for a given level of conditional volatility  $\sigma_{p,t+1}^2 = \sigma_p^{*2}$ . The investors solve the following problem at time  $t$ :

$$\max_{\mathbf{w}_t} \mu_{p,t+1} \quad \text{s.t.} \quad \sigma_{p,t+1}^2 = \sigma_p^{*2},$$

where  $\sigma_p^{*2}$  the target level of variance. The solution is given by

$$\begin{aligned}\hat{\omega}_t &= \Sigma_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}) \sqrt{\frac{\sigma_p^{*2}}{\kappa_t}}, \\ \kappa_t &= (\mathbf{m}_{t+1|t} - r_f \mathbf{1})' \Sigma_{t+1|t}^{-1} (\mathbf{m}_{t+1|t} - r_f \mathbf{1}).\end{aligned}$$

### *Estimation results.*

The cumulative realized returns are shown in Table 18 for (1) minimum-variance strategy with  $\mu_p^* = 0.004, 0.01$  and  $0.1$ , (2) mean-variance strategy with  $\gamma = 6, 10$  and  $15$ , and (3) maximum-return strategy with  $\sigma_p^* = 0.001, 0.01$  and  $0.1$ . For the minimum-variance strategy, CRSV models outperform CSV models, while CRSV model without leverage ( $\mathbf{R} = \mathbf{O}$  and  $\mathbf{S} = \mathbf{I}$ ) shows the better performance than CRSV

model with leverage (no restriction on  $\mathbf{R}$  and  $\mathbf{S}$ ). For the mean-variance strategy, CSV and CRSV models outperform CRSV without leverage model, and there is little difference between two models. For maximum-return strategies, CRSV model without leverage outperforms other two models. Overall, CRSV models outperform CSV model.

Figures 8 to 10 show the time series plots of the weight for the  $i$ -th stock,  $\omega_{it}$ , for three strategies in CRSV models. Under the riskiest settings in each strategy (*i.e.*,  $\mu_p^* = 0.1, \gamma = 6, \sigma_p^* = 0.1$ ), it is found that the portfolio weights fluctuate more drastically than those under two other less risky settings in each strategy during the prediction period. Also we note that the weights for the mean-variance strategy do not change so much during the prediction period, and the weights for the risk free asset are more than 95% weight under all levels of the risk aversion. Furthermore, the weights for the maximum-return strategy take more stable values relatively than those for the minimum-variance strategy.

Furthermore, Figure 14 illustrates the cumulative realized return for three models. Since the weights for risk assets under the mean-variance strategy are close to zero, the differences of cumulative realized returns among three models are very small for this strategy. For the minimum-variance and the maximum-return strategies, CSV model demonstrates the higher performance from the end of 2008 to the middle of 2009 than other two models, but its performance fluctuates drastically during the prediction period. On the other hand, CRSV models show more stable performance during the period.

## 6 Conclusion

In this paper, we propose CRSV model, simultaneous modeling of multiple daily asset returns and realized covariance, where the well-known bias problems in realized measures due to market microstructure noise and non-trading hours biases are automatically solved within our proposed framework. Furthermore, we extend them to incorporate the leverage effects that have been observed in the literature in univariate SV models. Taking a Bayesian approach, the efficient MCMC algorithm is described for the parameter estimation. In our empirical studies, our proposed CRSV models capture the dynamic behaviors of diagonal and off-diagonal components of Cholesky decomposed covariance matrices and are shown to outperform CSV models. The dy-

dynamic leverage effects  $\mathbf{R}$  and the correlation matrix  $\mathbf{S}$  for  $\mathbf{h}_{t+1}$  given  $\mathbf{y}_t$  are shown to improve the portfolio performances. Furthermore, the cumulative returns for CRSV models are less volatile than that of the CSV model.

As our future work, we could consider several directions. First, more parsimonious modeling by detecting the sparsity of off-diagonal components may improve the forecasting performances of the multivariate SV model, using such as Bayesian threshold dynamic modeling (Nakajima and West (2013a), Nakajima and West (2013b)). Second, the long memory property of diagonal components may need be taken in to account. Since the realized volatilities are well-known to have high persistence, a superposition model can be used to describe such a long range dependence and to improve the goodness of fit to the data. Finally, although our specification of stock returns is a simple random walk process, the more sophisticated mean structures, in such as factor SV models, may be useful to improve the portfolio performances.

## References

- Aguilar, O. and M. West (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics* 18, 338–357.
- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.
- Asai, M. and M. McAleer (2009). The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics* 150, 182–192.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realized 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, Series B* 63, 167–241.

- Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253–280.
- Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometrics analysis of realized co-variation: high frequency 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.
- Berger, J. O. and R. Y. Yang (1994). Non informative priors and bayesian testing for the ar(1) model. *ET* 10, 461–482.
- Chib, S. (2001). Markov chain monte carlo methods: computation and inference. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Econometrics*, Volume 5, pp. 3569–3649. Amsterdam: North Holland Press.
- Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134, 341–371.
- de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika* 82, 339–350.
- Dellaportas, P. and M. Pourahmadi (2012). Cholesky-garch models with applications to finance. *Statistics and Computing* 22, 849–855.
- Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika* 89, 603–616.
- Engle, R. and R. Colacito (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business and Economic Statistics* 24, 238–253.
- Fox, E. and D. Dunson (2011). Bayesian nonparametric covariance regression. Discussion Paper.
- Geweke, J. and G. Zhou (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies* 9, 557–587.
- Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies* 19, 237–271.

- Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61, 247–264.
- Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009). *Oxford-Man Institute’s realized library, version 0.1*. Oxford-Man Institute, University of Oxford.
- Ishihara, T. and Y. Omori (2012). Efficient bayesian estimation of a multivariate stochastic volatility model with cross leverage and heavy-tailed errors. *Computational Statistics and Data Analysis* 56, 3674–3689.
- Ishihara, T., Y. Omori, and M. Asai (2014). Matrix exponential stochastic volatility with cross leverage. *Computational Statistics and Data Analysis*.
- Jin, X. and J. M. Maheu (2013). Modeling realized covariances and returns. *Journal of Financial Economics* 11, 335–369.
- Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with arch models. *Review of Economic Studies* 65, 361–393.
- Koopman, S. J. and M. Scharth (2013). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics*, 76–115.
- Lopes, H. F. and C. M. Carvalho (2007). Factor stochastic volatility with time varying loadings and markov switching regimes. *Journal of Statistical Planning and Inference* 137, 3082–3091.
- Lopes, H. F., R. E. McCulloch, and R. S. Tsay (2012). Cholesky stochastic volatility models for high-dimensional time series. Discussion Paper.
- Lopes, H. F. and M. West (2004). Bayesian model assessment in factor analysis. *Statistica Sinica* 14, 41–67.
- Lubrano, M. (1995). Testing for unit roots in a bayesian framework. *Journal of Econometrics* 69, 81–109.
- Nakajima, J. and M. West (2013a). Bayesian analysis of latent threshold dynamic models. *Journal of Business and Economic Statistics* 31, 151–164.
- Nakajima, J. and M. West (2013b). Dynamic factor volatility modeling: A bayesian latent threshold approach. *Journal of Financial Economics* 11, 116–153.

- Noureldin, D., N. Shephard, and K. Sheppard (2012). Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics* 27, 907–933.
- Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: fast likelihood inference. *Journal of Econometrics* 140-2, 425–449.
- Omori, Y. and T. Watanabe (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics and Data Analysis* 52-6, 2892–2910.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160, 246–256.
- Philipov, A. and M. E. Glickman (2006). Factor multivariate stochastic volatility via wishart processes. *Econometric Reviews* 25, 311–334.
- Pitt, M. K. and N. Shephard (2003). *Time varying covariance: A factor stochastic volatility approach (with discussion)*. Oxford: Oxford University Press.
- Pourahmadi, M. (1999). Joint mean-variance models with applications to longitudinal data: Unconstrained parametrisation. *Biometrika* 86, 677–690.
- Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika* 84, 653–667.
- Shirota, S., T. Hizu, and Y. Omori (2014). Realized stochastic volatility with leverage and long memory. *Computational Statistics and Data Analysis* 76, 618–641.
- Takahashi, M., Y. Omori, and T. Watanabe (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and Data Analysis* 53-6, 2404–2426.
- Trojan, S. (2014). Multivariate stochastic volatility with dynamic cross leverage.
- Venter, J. H. and P. J. de Jongh (2014). Extended stochastic volatility models incorporating realized measures. *Computational Statistics and Data Analysis* 76, 687–707.
- Watanabe, T. and Y. Omori (2004). A multi-move sampler for estimating non-gaussian time series models: Comments on shephard and pitt (1997) . *Biometrika* 91, 246–248.
- Windle, J. and C. M. Carvalho (2014). A tractable state-space model for symmetric positive-definite matrices. *Bayesian Analysis* 9, 759–792.



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

## Appendix

### A The details of MCMC algorithm

#### A.1 Generation of $\mathbf{h}$

Lopes et al. (2012) implement the mixture sampler (Kim et al. (1998), Omori et al. (2007)) for diagonal components of the decomposed covariance matrix, which transforms the nonlinear measurement equation of individual SV models into a linear equation by taking the logarithm of squared daily returns and approximates the distribution of the non-normal disturbances ( $\log\chi_1^2$  distribution) by mixture of normals. Given the component of mixture normals, the model reduces to the linear and Gaussian state space model, and hence we are able to sample  $\{h_{ii,t}\}_{t=1}^n$  for  $i = 1, \dots, p$  simultaneously using a simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002)). The mixture sampler approach is the highly efficient estimation strategy, but it needs to correct the approximation error by the additional reweighting step. Alternatively, the multi-move sampler which samples a block of  $\{h_{ijt}\}_{t=s}^{t=s+m}$  given other  $h_{ijs}$ 's is also known to be efficient to sample these latent volatility variables in the univariate SV models (Shephard and Pitt (1997), Watanabe and Omori (2004), Omori and Watanabe (2008)). However, such estimation approaches based on the univariate SV model cannot be applied to our model unless  $\mathbf{R}$  is diagonal and  $\mathbf{S} = \mathbf{I}_p$ . Thus, in this paper, we use the single-move sampler which sample  $\mathbf{h}_t$  given other  $\mathbf{h}_s$ 's instead and improve the estimation efficiency by incorporating additional measurement equations of the logarithm of the realized variances,  $\mathbf{x}_t$ , without approximating the error distribution by normal mixtures. We shall show that this is a simple but efficient estimation strategy.

We conduct MH algorithm using a single-move sampler to generate  $\mathbf{h}_t$  given other parameters and latent variables. Let  $I(A)$  denote an indicator function such that  $I(A) = 1$  when  $A$  is true and 0 otherwise. The log conditional posterior density is

$$\log \pi(\mathbf{h}_t | \boldsymbol{\theta}, \mathbf{h}_{-t}, \mathbf{h}^*, \mathbf{m}, \mathbf{z}) = \text{const.} + g(\mathbf{h}_t) - \frac{1}{2}(\mathbf{h}_t - \mathbf{f}_t)' \mathbf{F}_t^{-1} (\mathbf{h}_t - \mathbf{f}_t),$$

where

$$\begin{aligned}
g(\mathbf{h}_t) &= -\frac{1}{2}(\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^* \mathbf{V}_t^{-1/2} (\mathbf{I} + \mathbf{R}' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \mathbf{R}) \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \\
&\quad + (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^* \mathbf{V}_t^{-1/2} \mathbf{R}' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \{ \mathbf{h}_{t+1} - \boldsymbol{\mu} - \Phi(\mathbf{h}_t - \boldsymbol{\mu}) \} I(t < n), \\
\mathbf{F}_t^{-1} &= \Phi \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \Phi I(t < n) + \lambda^{-I(t=1)} \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} + \mathbf{C}^{-1}, \\
\mathbf{f}_t &= \mathbf{F}_t \left[ -\frac{1}{2} \mathbf{1}_p + \Phi \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \{ \mathbf{h}_{t+1} - (\mathbf{I}_p - \Phi) \boldsymbol{\mu} \} I(t < n) + \mathbf{C}^{-1} (\mathbf{x}_t - \xi \mathbf{1}_p) \right. \\
&\quad \left. + \lambda^{-I(t=1)} \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \left\{ (\mathbf{I}_p - \Phi) \boldsymbol{\mu} + \Phi \mathbf{h}_{t-1} + \mathbf{R} \mathbf{V}_{t-1}^{-1/2} \mathbf{H}_{t-1}^* (\mathbf{y}_{t-1} - \mathbf{m}_{t-1}) \right\} \right],
\end{aligned}$$

with  $\mathbf{y}_0 \equiv \mathbf{m}_0$  and  $\mathbf{h}_0 \equiv \boldsymbol{\mu}$ . We generate a candidate  $\mathbf{h}_t^\dagger \sim \mathcal{N}(\mathbf{f}_t, \mathbf{F}_t)$  and accept it with probability  $\min\{1, \exp(g(\mathbf{h}_t^\dagger) - g(\mathbf{h}_t))\}$ .

## A.2 Generation of $\mathbf{h}^*$

Similar to the generation of  $\mathbf{h}$ , we consider the simple single-move sampling as follows. Consider the representation  $\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t) \equiv (\mathbf{y}_t - \mathbf{m}_t) + \mathbf{A}_t \mathbf{h}_t^*$  where  $\mathbf{A}_t$  is the  $p \times k$  matrix composed of the linear functions of  $\mathbf{y}_t - \mathbf{m}_t$ . Then generate

$$\mathbf{h}_t^* \sim \mathcal{N}(\mathbf{f}_t^*, \mathbf{F}_t^*),$$

where

$$\begin{aligned}
\mathbf{F}_t^{*-1} &= \mathbf{A}_t' \mathbf{V}_t^{-1} \mathbf{A}_t + \lambda^{*-I(t=1)} \mathbf{D}^{*-1} + \mathbf{C}^{*-1} \\
&\quad + \{ \mathbf{A}_t' \mathbf{V}_t^{-1/2} \mathbf{R}' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{A}_t + \Phi^* \mathbf{D}^{*-1} \Phi^* \} I(t < n),
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{f}_t^* &= \mathbf{F}_t^* \left[ -\mathbf{A}_t' \mathbf{V}_t^{-1} \mathbf{y}_t + \mathbf{C}^{*-1} (\mathbf{x}_t^* - \xi^*) + \lambda^{*-I(t=1)} \mathbf{D}^{*-1} \{ (\mathbf{I}_K - \Phi^*) \boldsymbol{\mu}^* + \Phi^* \mathbf{h}_{t-1}^* \} \right. \\
&\quad \left. + \mathbf{A}_t' \mathbf{V}_t^{-1/2} \mathbf{R}' \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \left\{ \mathbf{h}_{t+1} - \boldsymbol{\mu} - \Phi(\mathbf{h}_t - \boldsymbol{\mu}) - \mathbf{R} \mathbf{V}_t^{-1/2} (\mathbf{y}_t - \mathbf{m}_t) \right\} I(t < n) \right. \\
&\quad \left. + \Phi^* \mathbf{D}^{*-1} \{ \mathbf{h}_{t+1}^* - (\mathbf{I}_K - \Phi^*) \boldsymbol{\mu}^* \} I(t < n) \right],
\end{aligned}$$

with  $\mathbf{h}_0^* \equiv \boldsymbol{\mu}^*$  and  $K = p(p-1)/2$ . Although above algorithm is simple and efficient, we note that we could instead use a simulation smoother to generate  $\{h_{ijt}\}_{t=1}^n$  for  $i > j$  given  $\{h_{ikt}\}_{t=1}^n$  ( $k \neq j$ ) as in Lopes et al. (2012) when  $\mathbf{R}$  is diagonal and  $\mathbf{S} = \mathbf{I}_p$ .

### A.3 Generation of $m$

Noting that  $\epsilon_t | \eta_t \sim \mathcal{N}(\mathbf{R}'(\mathbf{R}\mathbf{R}' + \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'})^{-1}\eta_t, \mathbf{I}_p - \mathbf{R}'(\mathbf{R}\mathbf{R}' + \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'})^{-1}\mathbf{R})$ , we define

$$\begin{aligned}\tilde{\mathbf{y}}_t &= \mathbf{y}_t - \mathbf{H}_t^{*-1}\mathbf{V}_t^{1/2}\mathbf{R}'(\mathbf{R}\mathbf{R}' + \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'})^{-1}\{\mathbf{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu})\}I(t < n), \\ \boldsymbol{\Gamma}_t &= \mathbf{H}_t^{*-1}\mathbf{V}_t^{1/2}\{\mathbf{I}_p - \mathbf{R}'(\mathbf{R}\mathbf{R}' + \mathbf{S}^{-1}\mathbf{D}\mathbf{S}^{-1'})^{-1}\mathbf{R}\}I(t < n)\mathbf{V}_t^{1/2}\mathbf{H}_t^{*-1'},\end{aligned}$$

and consider the following linear Gaussian state space model

$$\begin{aligned}\tilde{\mathbf{y}}_t &= \mathbf{m}_t + \tilde{\boldsymbol{\epsilon}}_t, \quad \tilde{\boldsymbol{\epsilon}}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_t), \\ \mathbf{m}_{t+1} &= \mathbf{m}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_m),\end{aligned}$$

where  $\tilde{\boldsymbol{\epsilon}}_t$  and  $\boldsymbol{\nu}_t$  are independent. Generate  $\mathbf{m}$  using a simulation smoother.

### A.4 Generation of $(\xi, \boldsymbol{\xi}^*, \boldsymbol{\mu}, \boldsymbol{\mu}^*)$

We assume that prior distributions are

$$\xi \sim \mathcal{N}(\xi_0, \sigma_{\xi_0}^2), \quad \boldsymbol{\xi}^* \sim \mathcal{N}(\boldsymbol{\xi}_0^*, \boldsymbol{\Xi}_0^*), \quad \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}_0, \tilde{\mathbf{M}}_0), \quad \boldsymbol{\mu}^* \sim \mathcal{N}(\boldsymbol{\mu}_0^*, \tilde{\mathbf{M}}_0^*).$$

Then the conditional posterior distributions of  $(\xi, \boldsymbol{\xi}^*, \boldsymbol{\mu}, \boldsymbol{\mu}^*)$  are independent given other parameters,  $\mathbf{m}, \mathbf{h}$  and  $\mathbf{h}^*$ . The posterior distributions are

$$\xi \sim \mathcal{N}(\xi_1, \sigma_{\xi_1}^2), \quad \boldsymbol{\xi}^* \sim \mathcal{N}(\boldsymbol{\xi}_1^*, \boldsymbol{\Xi}_1^*), \quad \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}_1, \tilde{\mathbf{M}}_1), \quad \boldsymbol{\mu}^* \sim \mathcal{N}(\boldsymbol{\mu}_1^*, \tilde{\mathbf{M}}_1^*),$$

where

$$\begin{aligned}\sigma_{\xi_1}^{-2} &= \sigma_{\xi_0}^{-2} + n\mathbf{1}'_p\mathbf{C}^{-1}\mathbf{1}_p, \quad \xi_1 = \sigma_{\xi_1}^2 \left\{ \sigma_{\xi_0}^{-2}\xi_0 + \mathbf{1}'_p\mathbf{C}^{-1} \sum_{t=1}^n (\mathbf{x}_t - \mathbf{h}_t) \right\}, \\ \boldsymbol{\Xi}_1^{*-1} &= \boldsymbol{\Xi}_0^{*-1} + n\mathbf{C}^{*-1}, \quad \boldsymbol{\xi}_1^* = \boldsymbol{\Xi}_1^* \left\{ \boldsymbol{\Xi}_0^{*-1}\boldsymbol{\xi}_0^* + \mathbf{C}^{*-1} \sum_{t=1}^n (\mathbf{x}_t^* - \mathbf{h}_t^*) \right\},\end{aligned}$$

and

$$\begin{aligned}\tilde{\mathbf{M}}_1^{-1} &= \tilde{\mathbf{M}}_0^{-1} + (n-1)(\mathbf{I}_p - \boldsymbol{\Phi})\mathbf{S}'\mathbf{D}^{-1}\mathbf{S}(\mathbf{I}_p - \boldsymbol{\Phi}) + \lambda^{-1}\mathbf{S}'\mathbf{D}^{-1}\mathbf{S}, \\ \boldsymbol{\mu}_1 &= \tilde{\mathbf{M}}_1 \left[ \tilde{\mathbf{M}}_0^{-1}\boldsymbol{\mu}_0 + (\mathbf{I}_p - \boldsymbol{\Phi})\mathbf{S}'\mathbf{D}^{-1}\mathbf{S} \sum_{t=1}^{n-1} \left\{ \mathbf{h}_{t+1} - \boldsymbol{\Phi}\mathbf{h}_t - \mathbf{R}\mathbf{V}_t^{-1/2}\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t) \right\} \right. \\ &\quad \left. + \lambda^{-1}\mathbf{S}'\mathbf{D}^{-1}\mathbf{S}\mathbf{h}_1 \right], \\ \tilde{\mathbf{M}}_1^{*-1} &= \tilde{\mathbf{M}}_0^{*-1} + (n-1)(\mathbf{I}_K - \boldsymbol{\Phi}^*)\mathbf{D}^{*-1}(\mathbf{I}_K - \boldsymbol{\Phi}^*) + \lambda^{*-1}\mathbf{D}^{*-1}, \\ \boldsymbol{\mu}_1^* &= \tilde{\mathbf{M}}_1^* \left[ \tilde{\mathbf{M}}_0^{*-1}\boldsymbol{\mu}_0^* + (\mathbf{I}_K - \boldsymbol{\Phi}^*)\mathbf{D}^{*-1} \sum_{t=1}^{n-1} (\mathbf{h}_{t+1}^* - \boldsymbol{\Phi}^*\mathbf{h}_t^*) + \lambda^{*-1}\mathbf{D}^{*-1}\mathbf{h}_1^* \right].\end{aligned}$$

## A.5 Generation of $\phi$ and $\phi^*$

Given other parameters and variables,  $\phi$  and  $\phi^*$  are conditionally independent. For the prior distribution of  $\phi$ , we assume  $(\phi_i + 1)/2 \sim \mathcal{B}(a_{0i}, b_{0i})$  for  $i = 1, \dots, p$ , and denote the prior probability density function by  $\pi(\phi)$ . Then

$$\log \pi(\phi | \boldsymbol{\theta}_{\setminus \phi, \phi^*}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}) = \text{const.} + \log \pi(\phi) - (\phi - \boldsymbol{\mu}_\phi)' \boldsymbol{\Sigma}_\phi^{-1} (\phi - \boldsymbol{\mu}_\phi),$$

where

$$\boldsymbol{\Sigma}_\phi^{-1} = (\mathbf{S}' \mathbf{D}^{-1} \mathbf{S}) \odot \left\{ \sum_{t=1}^{n-1} (\mathbf{h}_t - \boldsymbol{\mu})(\mathbf{h}_t - \boldsymbol{\mu})' \right\}, \quad \boldsymbol{\mu}_\phi = \boldsymbol{\Sigma}_\phi \mathbf{b}_\phi, \quad (19)$$

and  $\mathbf{b}_\phi$  is the column vector with diagonal elements of

$$\mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \sum_{t=1}^{n-1} \left\{ \mathbf{h}_{t+1} - \boldsymbol{\mu} - \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \right\} (\mathbf{h}_t - \boldsymbol{\mu})'.$$

We generate a candidate  $\phi^\dagger$  from a truncated multivariate normal distribution  $\mathcal{TN}_{R_\phi}(\boldsymbol{\mu}_\phi, \boldsymbol{\Sigma}_\phi)$  where  $R_\phi = \{\phi \mid |\phi_i| < 1, i = 1, \dots, p\}$  and accept it with probability  $\min\{1, \pi(\phi^\dagger)/\pi(\phi)\}$ .

Let  $\pi(\phi^*)$  denote the prior probability density function of  $\phi^*$ . Then

$$\log \pi(\phi^* | \boldsymbol{\theta}_{\setminus \phi, \phi^*}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z}) = \text{const.} + \log \pi(\phi^*) - (\phi^* - \boldsymbol{\mu}_{\phi^*})' \boldsymbol{\Sigma}_{\phi^*}^{-1} (\phi^* - \boldsymbol{\mu}_{\phi^*}),$$

where

$$\boldsymbol{\Sigma}_{\phi^*}^{-1} = \mathbf{D}^{*-1} \odot \left\{ \sum_{t=1}^{n-1} (\mathbf{h}_t^* - \boldsymbol{\mu}^*)(\mathbf{h}_t^* - \boldsymbol{\mu}^*)' \right\}, \quad \boldsymbol{\mu}_{\phi^*} = \boldsymbol{\Sigma}_{\phi^*} \mathbf{b}_{\phi^*}, \quad (20)$$

and  $\mathbf{b}_{\phi^*}$  is the column vector with diagonal elements of

$$\mathbf{D}^{*-1} \sum_{t=1}^{n-1} (\mathbf{h}_{t+1}^* - \boldsymbol{\mu}^*) (\mathbf{h}_t^* - \boldsymbol{\mu}^*)'.$$

When we take a uniform prior on  $(-\sqrt{1+v}, \sqrt{1+v})$ , we generate a candidate  $\phi^{*\dagger}$  from a truncated multivariate normal distribution  $\mathcal{TN}_{R_{\phi^*}}(\boldsymbol{\mu}_{\phi^*}, \boldsymbol{\Sigma}_{\phi^*})$  where  $R_{\phi^*} = \{\phi^* \mid |\phi_i^*| < \sqrt{1+v}, i = 1, \dots, K\}$ .

## A.6 Generation of $\mathbf{R}$

Let  $\mathbf{r} = \text{vec}(\mathbf{R}) = (\mathbf{r}'_1, \dots, \mathbf{r}'_p)'$  where  $\mathbf{r}_j$  denotes the  $j$ -th column of  $\mathbf{R}$ . For the prior distribution of  $\mathbf{r}$ , we assume

$$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\gamma}_0, \boldsymbol{\Gamma}_0).$$

Noting that  $\text{vec}(\mathbf{AB}) = (\mathbf{B}' \otimes \mathbf{I}_n)\text{vec}(\mathbf{A})$  for an  $n \times m$  matrix  $\mathbf{A}$  and an  $m \times q$  matrix  $\mathbf{B}$ , we have

$$\text{vec} \left\{ \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \right\} = \left\{ (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^{*'} \mathbf{V}_t^{-1/2} \otimes \mathbf{I}_p \right\} \mathbf{r},$$

and obtain the conditional posterior distribution of  $\mathbf{r}$ :

$$\mathbf{r} | \boldsymbol{\theta}_{\setminus \mathbf{R}}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z} \sim \mathcal{N}(\boldsymbol{\gamma}_1, \boldsymbol{\Gamma}_1),$$

where

$$\begin{aligned} \boldsymbol{\Gamma}_1^{-1} &= \boldsymbol{\Gamma}_0^{-1} + \left\{ \sum_{t=1}^{n-1} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^{*'} \mathbf{V}_t^{-1/2} \right\} \otimes (\mathbf{S}' \mathbf{D}^{-1} \mathbf{S}), \\ \boldsymbol{\gamma}_1 &= \boldsymbol{\Gamma}_1 \left[ \boldsymbol{\Gamma}_0^{-1} \boldsymbol{\gamma}_0 + \left\{ \mathbf{Q} \otimes (\mathbf{S}' \mathbf{D}^{-1} \mathbf{S}) \right\} \text{vec}(\mathbf{I}_p) \right], \\ \mathbf{Q} &= \sum_{t=1}^{n-1} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) (\mathbf{h}_{t+1} - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{h}_t - \boldsymbol{\mu}))'. \end{aligned}$$

**Alternative parsimonious specification.** In the empirical studies of the factor-based models, it is often that only the first factor (which corresponds to the market factor) shows the leverage effects. Taking account of such empirical results which we shall also see in our empirical studies, it is useful to consider the parsimonious parameterization  $\mathbf{R} = \{\mathbf{r}_1, \mathbf{0}, \dots, \mathbf{0}\}$ . It implies only the first component of the standardized vector  $\mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t)$ , which is expected to include the market factor, produces the leverage effect. Thus, we assume

$$\mathbf{R} = \{\mathbf{r}_1, \mathbf{0}, \dots, \mathbf{0}\}, \quad \mathbf{r}_1 \sim \mathcal{N}(\boldsymbol{\gamma}_{10}, \boldsymbol{\Gamma}_{10}),$$

and then  $\text{vec}(\mathbf{R}) = (\mathbf{r}'_1, \mathbf{0}', \dots, \mathbf{0}')' = \mathbf{e}_1 \otimes \mathbf{r}_1$  where  $\mathbf{e}_1 = (1, 0, \dots, 0)'$  is a  $p \times 1$  vector. Noting that

$$\begin{aligned} \text{vec} \left\{ \mathbf{R} \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \right\} &= \left\{ (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^{*'} \mathbf{V}_t^{-1/2} \otimes \mathbf{I}_p \right\} \{\mathbf{e}_1 \otimes \mathbf{r}_1\} \\ &= \left\{ (\mathbf{y}_t - \mathbf{m}_t)' \mathbf{H}_t^{*'} \mathbf{V}_t^{-1/2} \mathbf{e}_1 \right\} \mathbf{r}_1, \end{aligned}$$

we obtain the conditional posterior distribution of  $\mathbf{r}_1$ :

$$\mathbf{r}_1 | \boldsymbol{\theta}_{\setminus \mathbf{R}}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z} \sim \mathcal{N}(\boldsymbol{\gamma}_{11}, \boldsymbol{\Gamma}_{11}),$$

where

$$\begin{aligned} \boldsymbol{\Gamma}_{11}^{-1} &= \boldsymbol{\Gamma}_{10}^{-1} + \left[ \sum_{t=1}^{n-1} \{ \mathbf{e}_1' \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \}^2 \right] \mathbf{S}' \mathbf{D}^{-1} \mathbf{S}, \\ \boldsymbol{\gamma}_{11} &= \boldsymbol{\Gamma}_{11} \left[ \boldsymbol{\Gamma}_{10}^{-1} \boldsymbol{\gamma}_{10} + \mathbf{S}' \mathbf{D}^{-1} \mathbf{S} \sum_{t=1}^{n-1} \{ \mathbf{e}_1' \mathbf{V}_t^{-1/2} \mathbf{H}_t^* (\mathbf{y}_t - \mathbf{m}_t) \} (\mathbf{h}_{t+1} - \boldsymbol{\mu} - \Phi(\mathbf{h}_t - \boldsymbol{\mu})) \right]. \end{aligned}$$

## A.7 Generation of $\mathbf{S}$

Let  $\mathbf{s} = (s_{21}, \dots, s_{p1}, s_{32}, \dots, s_{pp-1})'$  denote a stacked vector of lower off-diagonal elements of  $\mathbf{S}$ . For the prior distribution of  $\mathbf{s}$ , we assume

$$\mathbf{s} \sim \mathcal{N}(\boldsymbol{\delta}_0, \boldsymbol{\Delta}_0).$$

Using the representation of  $\mathbf{S}\tilde{\mathbf{h}}_{t+1} = \tilde{\mathbf{h}}_{t+1} + \mathbf{B}_{t+1}\mathbf{s}$  where  $\tilde{\mathbf{h}}_{t+1} = \mathbf{h}_{t+1} - \boldsymbol{\mu} - \Phi(\mathbf{h}_t - \boldsymbol{\mu}) - \mathbf{R}\mathbf{V}_t^{-1/2}\mathbf{H}_t^*(\mathbf{y}_t - \mathbf{m}_t)$ , we obtain the conditional posterior distribution of  $\mathbf{s}$ :

$$\mathbf{s} | \boldsymbol{\theta}_{\setminus \mathbf{S}}, \mathbf{m}, \mathbf{h}, \mathbf{h}^*, \mathbf{z} \sim \mathcal{N}(\boldsymbol{\delta}_1, \boldsymbol{\Delta}_1),$$

where

$$\boldsymbol{\Delta}_1^{-1} = \boldsymbol{\Delta}_0^{-1} + \sum_{t=1}^n \lambda^{-I(t=1)} \mathbf{B}_t' \mathbf{D}^{-1} \mathbf{B}_t, \quad \boldsymbol{\delta}_1 = \boldsymbol{\Delta}_1 \left[ \boldsymbol{\Delta}_0^{-1} \boldsymbol{\delta}_0 - \sum_{t=1}^n \lambda^{-I(t=1)} \mathbf{B}_t' \mathbf{D}^{-1} \tilde{\mathbf{h}}_t \right].$$

## A.8 Generation of $\mathbf{C}, \mathbf{C}^*, \mathbf{D}, \mathbf{D}^*$ and $\boldsymbol{\Omega}_m$

Note that elements of  $\mathbf{C}, \mathbf{C}^*, \mathbf{D}, \mathbf{D}^*$  and  $\boldsymbol{\Omega}_m$  are conditionally independent. We assume their prior distributions are

$$\tau_{ij}^2 \sim \mathcal{IG}(a_{\tau,ij}/2, b_{\tau}/2), \quad \sigma_{u,ij}^2 \sim \mathcal{IG}(a_{u,ij}/2, b_{u,ij}/2), \quad \sigma_{mi}^2 \sim \mathcal{IG}(a_{mi}/2, b_{mi}/2),$$

for  $i = 1, \dots, p$  and  $j = 1, \dots, i$ .

In empirical studies, the diagonal components,  $\mathbf{h}_t$ , fluctuate more drastically than off-diagonal components,  $\mathbf{h}_t^*$ . Actually, the time series plots of  $\mathbf{h}_t^*$  are found to be almost constant or those of the random walk process with small variance. Hence, we shall assume  $\mathcal{IG}(\epsilon, \epsilon)$  prior for  $\mathbf{D}^*$ ,  $(\tau_{ij}^2, i < j)$  with e.g.,  $\epsilon = 10^{-4}$ . Also we expect the

variance of the random walk process for the mean vector  $\mathbf{m}_t$  to be small and hence assume the  $\mathcal{IG}(\epsilon, \epsilon)$  prior, with e.g.,  $\epsilon = 10^{-6}$ .

The conditional posterior distributions of  $\tau_{ij}^2$ 's are

$$\tau_{ii}^2 | \cdot \sim \mathcal{IG} \left( \frac{a_{\tau,ii} + n}{2}, \frac{b_{\tau,ii} + \sum_{t=1}^n w_{it}^2}{2} \right), \quad \tau_{ij}^2 | \cdot \sim \mathcal{IG} \left( \frac{a_{\tau,ij} + n}{2}, \frac{b_{\tau,ij} + \sum_{t=1}^n w_{ij,t}^{*2}}{2} \right),$$

for  $i = 1, \dots, p$  and  $j = 1, \dots, i-1$ , where

$$\mathbf{w}_t = \begin{cases} \lambda^{-1/2} \mathbf{S}(\mathbf{h}_1 - \boldsymbol{\mu}), & t = 1, \\ \mathbf{S} \left\{ \mathbf{h}_t - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{h}_{t-1} - \boldsymbol{\mu}) - \mathbf{R}\mathbf{V}_{t-1}^{-1/2} \mathbf{H}_{t-1}^* (\mathbf{y}_{t-1} - \mathbf{m}_{t-1}) \right\}, & t = 2, \dots, n, \end{cases}$$

$$w_{ij,t}^* = \begin{cases} \lambda^{*-1/2} (h_{ij,1} - \mu_{ij}), & t = 1, \\ h_{ij,t+1} - \mu_{ij} - \phi_{ij}(h_{ij,t} - \mu_{ij}), & t = 2, \dots, n. \end{cases}$$

Similarly, the conditional posterior distributions of  $\sigma_{u,ij}^2$ 's are

$$\sigma_{u,ii}^2 | \cdot \sim \mathcal{IG} \left( \frac{a_{u,ii} + n}{2}, \frac{b_{u,ii} + \sum_{t=1}^n (x_{ii,t} - \xi - h_{ii,t})^2}{2} \right),$$

$$\sigma_{u,ij}^2 | \cdot \sim \mathcal{IG} \left( \frac{a_{u,ij} + n}{2}, \frac{b_{u,ij} + \sum_{t=1}^n (x_{ij,t} - \xi_{ij} - h_{ij,t})^2}{2} \right),$$

for  $i = 1, \dots, p$  and  $j = 1, \dots, i-1$ . Finally, the conditional posterior distributions of  $\sigma_{mi}^2$ 's are

$$\sigma_{mi}^2 | \cdot \sim \mathcal{IG} \left( \frac{a_{mi} + n}{2}, \frac{b_{mi} + \lambda_m^{-1} m_{i1}^2 + \sum_{t=1}^{n-1} (m_{i,t+1} - m_{it})^2}{2} \right),$$

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

## B Tables and Figures

### B.1 Simulated data

Table 1: Estimation results for  $(\boldsymbol{\mu}, \boldsymbol{\phi}, \xi, \mathbf{C}, \mathbf{D}, \boldsymbol{\Omega}_m)$  for simulated data

Parameter	True	Mean	95% interval	IF	Parameter	True	Mean	95% interval	IF
$\mu_{11}$	0	0.047	[-0.099, 0.207]	11.0	$\phi_{11}$	0.97	0.971	[0.963, 0.979]	5.9
$\mu_{22}$	0	0.148	[-0.002, 0.313]	8.4	$\phi_{22}$	0.97	0.970	[0.962, 0.978]	4.7
$\mu_{33}$	0	0.104	[-0.028, 0.247]	13.1	$\phi_{33}$	0.97	0.967	[0.959, 0.974]	6.0
$\mu_{44}$	0	0.043	[-0.102, 0.201]	11.3	$\phi_{44}$	0.97	0.968	[0.960, 0.975]	4.2
$\mu_{55}$	0	0.169	[0.048, 0.296]	13.0	$\phi_{55}$	0.97	0.959	[0.951, 0.967]	3.7
$\mu_{66}$	0	0.128	[-0.041, 0.306]	7.8	$\phi_{66}$	0.97	0.969	[0.962, 0.977]	5.4
$\mu_{77}$	0	0.065	[-0.104, 0.243]	7.5	$\phi_{77}$	0.97	0.969	[0.962, 0.976]	5.1
$\mu_{88}$	0	0.155	[-0.044, 0.364]	6.5	$\phi_{88}$	0.97	0.970	[0.963, 0.977]	5.3
$\mu_{99}$	0	0.142	[-0.059, 0.346]	5.7	$\phi_{99}$	0.97	0.968	[0.961, 0.974]	2.9
$\tau_{11}$	0.1	0.098	[0.093, 0.104]	26.8	$\xi$	-0.5	-0.496	[-0.518, -0.478]	569.0
$\tau_{22}$	0.1	0.100	[0.095, 0.106]	21.2					
$\tau_{33}$	0.1	0.094	[0.089, 0.100]	25.5					
$\tau_{44}$	0.1	0.098	[0.093, 0.105]	21.4					
$\tau_{55}$	0.1	0.101	[0.095, 0.107]	17.8					
$\tau_{66}$	0.1	0.104	[0.098, 0.110]	26.9					
$\tau_{77}$	0.1	0.096	[0.090, 0.103]	33.0					
$\tau_{88}$	0.1	0.104	[0.098, 0.110]	22.3					
$\tau_{99}$	0.1	0.107	[0.101, 0.113]	22.6					
$\sigma_{u,11}$	0.1	0.103	[0.098, 0.108]	18.7	$\sigma_{m1}$	0.002	0.0021	[0.0009, 0.0043]	561.8
$\sigma_{u,22}$	0.1	0.099	[0.095, 0.104]	17.1	$\sigma_{m2}$	0.002	0.0020	[0.0008, 0.0047]	592.5
$\sigma_{u,33}$	0.1	0.101	[0.096, 0.106]	22.9	$\sigma_{m3}$	0.002	0.0016	[0.0006, 0.0050]	654.3
$\sigma_{u,44}$	0.1	0.104	[0.098, 0.108]	22.7	$\sigma_{m4}$	0.002	0.0022	[0.0009, 0.0045]	569.5
$\sigma_{u,55}$	0.1	0.0977	[0.093, 0.102]	17.4	$\sigma_{m5}$	0.002	0.0010	[0.0003, 0.0025]	642.1
$\sigma_{u,66}$	0.1	0.0978	[0.092, 0.103]	18.7	$\sigma_{m6}$	0.002	0.0029	[0.0014, 0.0053]	528.9
$\sigma_{u,77}$	0.1	0.100	[0.095, 0.106]	32.1	$\sigma_{m7}$	0.002	0.0022	[0.0011, 0.0041]	533.3
$\sigma_{u,88}$	0.1	0.094	[0.088, 0.100]	25.0	$\sigma_{m8}$	0.002	0.0015	[0.0006, 0.0037]	620.4
$\sigma_{u,99}$	0.1	0.094	[0.088, 0.104]	20.1	$\sigma_{m9}$	0.002	0.0010	[0.0004, 0.0028]	636.2



Table 2: Posterior means of  $\mathbf{R}$  for simulated data: red font means that 95% credible interval includes true value.

$r_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$
$i = 1$	-0.0480	-0.0008	-0.0025	0.0011	0.0010	-0.0042	-0.0042	0.0015	-0.0015
$i = 2$	-0.0522	0.0016	0.0009	0.0009	-0.0039	0.0023	-0.0044	-0.0004	0.0010
$i = 3$	-0.0524	-0.0020	-0.0017	0.0029	-0.0003	0.0014	0.0010	-0.0024	-0.0012
$i = 4$	-0.0510	-0.0054	-0.0002	0.0054	-0.0008	0.0007	-0.0030	-0.0053	-0.0008
$i = 5$	-0.0532	-0.0041	-0.0030	0.0005	0.0021	-0.0014	-0.0003	0.0035	0.0065
$i = 6$	-0.0485	-0.0019	-0.0059	0.0013	0.0010	-0.0002	-0.0098	0.0000	0.0020
$i = 7$	-0.0497	-0.0040	-0.0049	-0.0020	0.0034	-0.0053	-0.0045	0.0039	-0.0028
$i = 8$	-0.0520	0.0021	-0.0036	-0.0012	0.0039	-0.0017	-0.0039	-0.0026	-0.0049
$i = 9$	-0.0464	-0.0036	0.0002	0.0003	0.0046	0.0040	-0.0020	0.0006	-0.0005

Table 3: Posterior means of  $\mathbf{S}$  for simulated data: red font means that 95% credible interval includes true value.

$s_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$
$i = 1$	1								
$i = 2$	0.212	1							
$i = 3$	0.266	0.126	1						
$i = 4$	0.245	0.189	0.178	1					
$i = 5$	0.179	0.193	0.194	0.244	1				
$i = 6$	0.180	0.218	0.214	0.193	0.204	1			
$i = 7$	0.163	0.253	0.274	0.223	0.213	0.100	1		
$i = 8$	0.180	0.236	0.255	0.218	0.179	0.204	0.173	1	
$i = 9$	0.277	0.183	0.165	0.189	0.183	0.266	0.224	0.133	1

Table 4: Posterior means of  $\boldsymbol{\mu}^*$  for simulated data: red font means that 95% credible interval includes true value.

$\mu_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.024							
$i = 3$	-0.009	0.010						
$i = 4$	-0.032	0.003	-0.007					
$i = 5$	-0.001	-0.033	0.008	-0.006				
$i = 6$	0.012	-0.045	0.028	0.034	-0.013			
$i = 7$	-0.050	0.017	-0.016	0.023	-0.014	-0.017		
$i = 8$	-0.016	-0.059	0.031	0.028	-0.008	-0.013	-0.001	
$i = 9$	-0.003	0.005	-0.057	-0.012	0.011	0.035	-0.025	-0.012

Table 5: Posterior means of  $\phi^*$  for simulated data: red font means that 95% credible interval includes true value.

$\phi_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.993							
$i = 3$	0.987	0.991						
$i = 4$	0.987	0.986	0.990					
$i = 5$	0.990	0.989	0.986	0.994				
$i = 6$	0.986	0.988	0.989	0.991	0.989			
$i = 7$	0.975	0.989	0.986	0.989	0.980	0.988		
$i = 8$	0.990	0.989	0.987	0.994	0.989	0.988	0.982	
$i = 9$	0.987	0.985	0.989	0.989	0.992	0.980	0.993	0.989

Table 6: Posterior means of  $\xi^*$  for simulated data: red font means that 95% credible interval includes true value.

$\mu_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	-0.4916							
$i = 3$	-0.5022	-0.5005						
$i = 4$	-0.4920	-0.4966	-0.4979					
$i = 5$	-0.5137	-0.4788	-0.4961	-0.4917				
$i = 6$	-0.5032	-0.4815	-0.5288	-0.5299	-0.5356			
$i = 7$	-0.4754	-0.5009	-0.4889	-0.5212	-0.5070	-0.5098		
$i = 8$	-0.5059	-0.4681	-0.5165	-0.5216	-0.5122	-0.5069	-0.4843	
$i = 9$	-0.5021	-0.4793	-0.4557	-0.4775	-0.4989	-0.5061	-0.4843	-0.4721

Table 7: Posterior means of  $C^*$  for simulated data: red font means that 95% credible interval includes true value.

$\mu_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.099							
$i = 3$	0.099	0.103						
$i = 4$	0.101	0.099	0.099					
$i = 5$	0.101	0.100	0.101	0.099				
$i = 6$	0.101	0.099	0.100	0.097	0.103			
$i = 7$	0.098	0.098	0.100	0.098	0.101	0.099		
$i = 8$	0.098	0.098	0.099	0.099	0.100	0.101	0.101	
$i = 9$	0.100	0.099	0.099	0.098	0.100	0.100	0.100	0.0972

Table 8: Posterior means of  $\mathbf{D}^*$  for simulated data: red font means that 95% credible interval includes true value.

$\tau_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.0098							
$i = 3$	0.0096	0.0098						
$i = 4$	0.0101	0.0120	0.0097					
$i = 5$	0.0106	0.0112	0.0099	0.0078				
$i = 6$	0.0088	0.0105	0.0115	0.0105	0.0093			
$i = 7$	0.0119	0.0103	0.0105	0.0102	0.0117	0.0105		
$i = 8$	0.0098	0.0087	0.0106	0.0089	0.0097	0.0107	0.0110	
$i = 9$	0.0104	0.0102	0.0103	0.0108	0.0099	0.0119	0.0098	0.0106

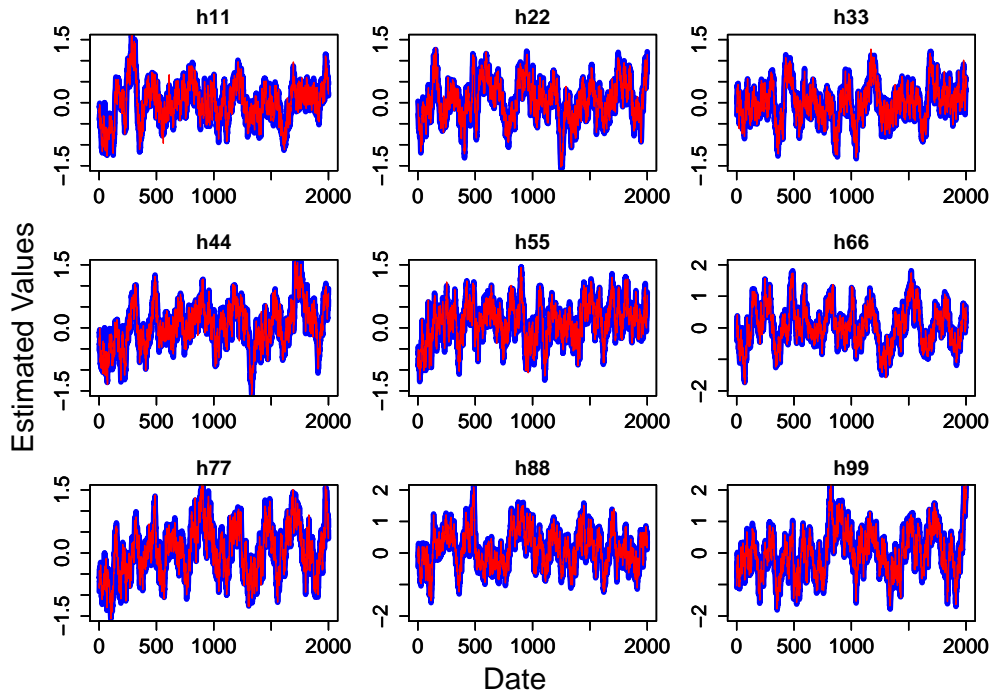


Figure 1: 95% credible intervals (blue) and true values (red) for  $h_{11,t}, \dots, h_{99,t}$ .

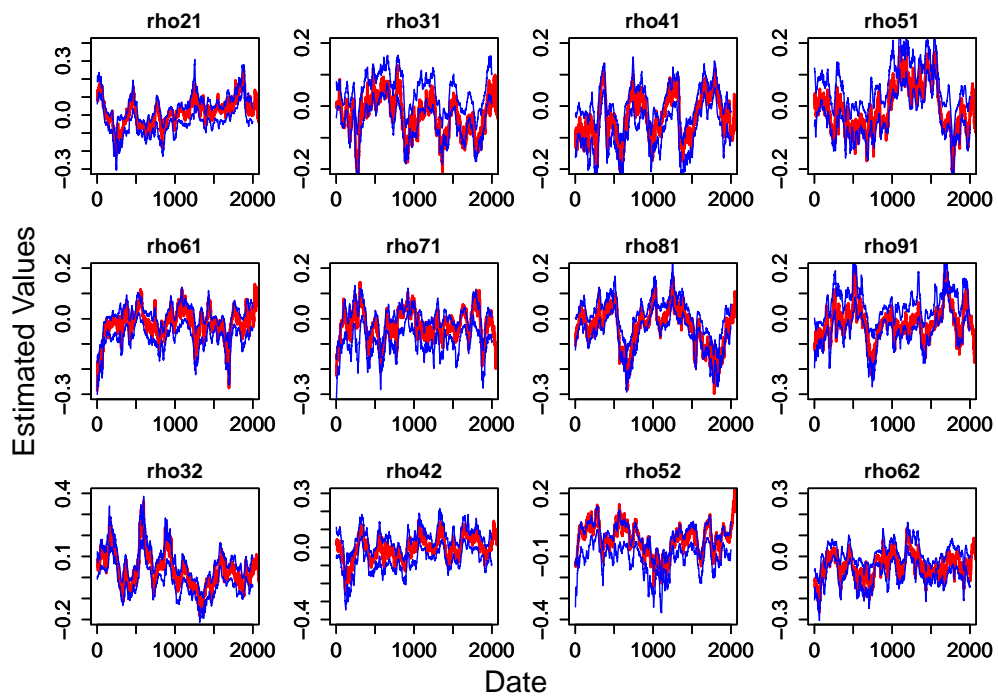


Figure 2: 95% credible intervals (blue) and true values (red) for dynamic correlations  $\rho_{21,t}, \dots, \rho_{62,t}$

## B.2 U.S. stock returns data

Table 9: Estimation results for  $\mu, \phi, \mathbf{D}$  for U.S. stock returns  
CRSV and CSV models

CRSV	Mean	Stdev	95% interval	IF	CSV	Mean	Stdev	95% interval	IF
$\mu_{11}$	1.097	0.133	[0.833, 1.360]	3.5	$\mu_{11}$	1.086	0.241	[0.603, 1.560]	21.1
$\mu_{22}$	0.278	0.088	[0.104, 0.452]	4.3	$\mu_{22}$	0.192	0.099	[-0.004, 0.389]	31.8
$\mu_{33}$	0.405	0.091	[0.225, 0.585]	4.1	$\mu_{33}$	0.127	0.106	[-0.085, 0.329]	146.6
$\mu_{44}$	0.290	0.084	[0.126, 0.456]	5.5	$\mu_{44}$	0.259	0.072	[0.117, 0.400]	80.4
$\mu_{55}$	1.095	0.092	[0.913, 1.279]	5.0	$\mu_{55}$	1.023	0.078	[0.870, 1.178]	86.5
$\mu_{66}$	0.435	0.132	[0.178, 0.695]	3.0	$\mu_{66}$	0.225	0.114	[0.002, 0.451]	61.5
$\mu_{77}$	0.296	0.086	[0.126, 0.469]	5.4	$\mu_{77}$	-0.070	0.080	[-0.232, 0.085]	226.7
$\mu_{88}$	0.125	0.106	[-0.083, 0.333]	3.9	$\mu_{88}$	-0.120	0.102	[-0.320, 0.081]	91.5
$\mu_{99}$	-0.169	0.083	[-0.332, -0.004]	5.3	$\mu_{99}$	-0.418	0.095	[-0.608, -0.232]	87.4
$\phi_{11}$	0.947	0.004	[0.938, 0.955]	103.1	$\phi_{11}$	0.982	0.005	[0.970, 0.991]	431.1
$\phi_{22}$	0.930	0.006	[0.917, 0.941]	82.6	$\phi_{22}$	0.921	0.014	[0.888, 0.946]	363.3
$\phi_{33}$	0.936	0.006	[0.924, 0.947]	78.9	$\phi_{33}$	0.867	0.018	[0.830, 0.902]	341.5
$\phi_{44}$	0.935	0.006	[0.923, 0.948]	76.5	$\phi_{44}$	0.899	0.015	[0.868, 0.926]	367.7
$\phi_{55}$	0.944	0.006	[0.933, 0.956]	114.7	$\phi_{55}$	0.875	0.020	[0.826, 0.908]	413.5
$\phi_{66}$	0.955	0.004	[0.947, 0.963]	122.0	$\phi_{66}$	0.931	0.007	[0.914, 0.944]	310.8
$\phi_{77}$	0.942	0.005	[0.931, 0.953]	118.2	$\phi_{77}$	0.849	0.021	[0.796, 0.882]	394.3
$\phi_{88}$	0.941	0.005	[0.931, 0.951]	94.3	$\phi_{88}$	0.892	0.011	[0.869, 0.912]	283.7
$\phi_{99}$	0.938	0.006	[0.925, 0.950]	72.3	$\phi_{99}$	0.848	0.017	[0.811, 0.878]	287.2
$\tau_{11}$	0.328	0.010	[0.308, 0.347]	89.1	$\tau_{11}$	0.179	0.027	[0.134, 0.243]	599.2
$\tau_{22}$	0.140	0.009	[0.122, 0.159]	151.2	$\tau_{22}$	0.333	0.035	[0.269, 0.405]	467.7
$\tau_{33}$	0.129	0.008	[0.113, 0.146]	129.9	$\tau_{33}$	0.535	0.047	[0.440, 0.627]	425.6
$\tau_{44}$	0.141	0.007	[0.127, 0.157]	86.6	$\tau_{44}$	0.251	0.026	[0.204, 0.305]	474.3
$\tau_{55}$	0.113	0.009	[0.096, 0.131]	229.0	$\tau_{55}$	0.345	0.045	[0.269, 0.453]	544.3
$\tau_{66}$	0.109	0.008	[0.093, 0.125]	196.2	$\tau_{66}$	0.251	0.028	[0.194, 0.302]	545.6
$\tau_{77}$	0.087	0.008	[0.072, 0.104]	234.7	$\tau_{77}$	0.316	0.048	[0.247, 0.426]	573.0
$\tau_{88}$	0.127	0.009	[0.108, 0.146]	198.1	$\tau_{88}$	0.308	0.034	[0.253, 0.383]	510.4
$\tau_{99}$	0.128	0.010	[0.109, 0.148]	187.8	$\tau_{99}$	0.419	0.045	[0.343, 0.515]	449.7

Table 10: Estimation result for  $\boldsymbol{\mu}, \boldsymbol{\phi}, \xi, \mathbf{C}, \mathbf{D}, \boldsymbol{\Omega}_m$   
for U.S. stock returns. CRSV model.

Parameter	Mean	95% interval	IF	Parameter	Mean	95% interval	IF
$\mu_{11}$	1.097	[0.833, 1.360]	3.5	$\phi_{11}$	0.947	[0.938, 0.955]	103.1
$\mu_{22}$	0.278	[0.104, 0.452]	4.3	$\phi_{22}$	0.930	[0.917, 0.941]	82.6
$\mu_{33}$	0.405	[0.225, 0.585]	4.1	$\phi_{33}$	0.936	[0.924, 0.947]	78.9
$\mu_{44}$	0.290	[0.126, 0.456]	5.5	$\phi_{44}$	0.935	[0.923, 0.948]	76.5
$\mu_{55}$	1.095	[0.913, 1.279]	5.0	$\phi_{55}$	0.944	[0.933, 0.956]	114.7
$\mu_{66}$	0.435	[0.178, 0.695]	3.0	$\phi_{66}$	0.955	[0.947, 0.963]	122.0
$\mu_{77}$	0.296	[0.126, 0.469]	5.4	$\phi_{77}$	0.942	[0.931, 0.953]	118.2
$\mu_{88}$	0.125	[-0.083, 0.333]	3.9	$\phi_{88}$	0.941	[0.931, 0.951]	94.3
$\mu_{99}$	-0.169	[-0.332, -0.004]	5.3	$\phi_{99}$	0.938	[0.925, 0.950]	72.3
$\tau_{11}$	0.328	[0.308, 0.347]	89.1	$\xi$	-0.366	[-0.387, -0.345]	156.9
$\tau_{22}$	0.140	[0.122, 0.159]	151.2				
$\tau_{33}$	0.129	[0.113, 0.146]	129.9				
$\tau_{44}$	0.141	[0.127, 0.157]	86.6				
$\tau_{55}$	0.113	[0.096, 0.131]	229.0				
$\tau_{66}$	0.109	[0.093, 0.125]	196.2				
$\tau_{77}$	0.087	[0.072, 0.104]	234.7				
$\tau_{88}$	0.127	[0.108, 0.146]	198.1				
$\tau_{99}$	0.128	[0.109, 0.148]	187.8				
$\sigma_{u,11}$	0.326	[0.312, 0.341]	60.9	$\sigma_{m1}$	0.0010	[0.0004, 0.0027]	583.0
$\sigma_{u,22}$	0.314	[0.299, 0.328]	63.9	$\sigma_{m2}$	0.0017	[0.0005, 0.0036]	561.2
$\sigma_{u,33}$	0.316	[0.302, 0.330]	56.0	$\sigma_{m3}$	0.0012	[0.0003, 0.0034]	623.8
$\sigma_{u,44}$	0.310	[0.298, 0.322]	27.6	$\sigma_{m4}$	0.0034	[0.0013, 0.0071]	554.2
$\sigma_{u,55}$	0.356	[0.342, 0.370]	57.1	$\sigma_{m5}$	0.0015	[0.0005, 0.0039]	654.1
$\sigma_{u,66}$	0.361	[0.347, 0.375]	54.6	$\sigma_{m6}$	0.0012	[0.0004, 0.0053]	626.8
$\sigma_{u,77}$	0.341	[0.327, 0.354]	56.8	$\sigma_{m7}$	0.0011	[0.0003, 0.0026]	565.4
$\sigma_{u,88}$	0.333	[0.318, 0.348]	61.9	$\sigma_{m8}$	0.0009	[0.0004, 0.0020]	528.2
$\sigma_{u,99}$	0.313	[0.299, 0.327]	57.4	$\sigma_{m9}$	0.0014	[0.0005, 0.0039]	608.7

Table 11: Posterior means of  $\mathbf{R}$  for U.S. stock returns. CRSV model: red font indicates that 95% credible interval does not include 0.

$r_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$
$i = 1$	-0.0585	-0.0119	-0.0113	-0.0177	-0.0042	-0.0152	0.0047	0.0063	-0.0152
$i = 2$	-0.0291	-0.0319	0.0031	-0.0062	-0.0029	-0.0169	-0.0010	-0.0052	-0.0086
$i = 3$	-0.0167	-0.0144	-0.0095	-0.0007	-0.0002	-0.0064	-0.0022	-0.0040	0.0002
$i = 4$	-0.0319	-0.0049	-0.0180	-0.0206	-0.0063	-0.0141	-0.0107	-0.0038	-0.0029
$i = 5$	-0.0276	-0.0049	-0.0086	-0.0110	-0.0063	-0.0019	-0.0026	-0.0015	-0.0025
$i = 6$	-0.0315	0.0016	-0.0000	-0.0211	-0.0005	-0.0247	0.0006	-0.0115	0.0028
$i = 7$	-0.0271	-0.0125	-0.0081	-0.0064	-0.0075	-0.0072	-0.0002	0.0026	-0.0047
$i = 8$	-0.0295	-0.0004	-0.0035	-0.0110	-0.0020	-0.0038	0.0042	-0.0306	-0.0106
$i = 9$	-0.0289	-0.0004	-0.0044	-0.0139	0.0007	-0.0127	-0.0062	-0.0082	-0.0219

Table 12: Posterior means of  $\mathbf{S}$  for U.S. stock returns. CRSV model: red font indicates that 95% credible interval does not include 0.

$s_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$
$i = 1$	1								
$i = 2$	0.761	1							
$i = 3$	0.227	0.591	1						
$i = 4$	0.273	0.206	0.227	1					
$i = 5$	0.407	-0.023	-0.045	0.397	1				
$i = 6$	0.657	-0.188	0.227	0.239	-0.095	1			
$i = 7$	-0.043	-0.091	0.195	0.236	0.193	0.402	1		
$i = 8$	0.134	0.159	0.055	-0.146	0.309	0.579	-0.125	1	
$i = 9$	-0.072	0.337	-0.112	0.015	-0.071	0.148	0.507	0.0639	1

Table 13: Posterior means (standard deviation) of  $\mu^*$  for U.S. stock returns. CRSV model: red font indicates that 95% credible interval does not include 0.

$\mu_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.333							
$i = 3$	0.210	0.484						
$i = 4$	0.170	0.177	0.120					
$i = 5$	0.263	0.168	0.151	0.424				
$i = 6$	0.419	0.160	0.105	0.151	0.070			
$i = 7$	0.139	0.114	0.085	0.135	0.184	0.113		
$i = 8$	0.135	0.145	0.111	0.106	0.057	0.167	0.119	
$i = 9$	0.058	0.050	0.060	0.128	0.014	0.055	0.037	0.058

Table 14: Posterior means (standard deviation) of  $\phi^*$ :  
U.S. stock returns. CRSV model.

$\phi_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.941 (0.016)							
$i = 3$	0.971 (0.011)	0.987 (0.007)						
$i = 4$	0.968 (0.010)	0.962 (0.016)	0.961 (0.012)					
$i = 5$	0.984 (0.005)	0.997 (0.002)	0.975 (0.013)	0.997 (0.002)				
$i = 6$	0.994 (0.003)	0.976 (0.010)	0.958 (0.015)	0.970 (0.012)	0.984 (0.007)			
$i = 7$	0.993 (0.005)	0.996 (0.005)	0.917 (0.045)	0.993 (0.004)	0.985 (0.006)	0.986 (0.011)		
$i = 8$	0.958 (0.031)	0.964 (0.030)	0.978 (0.010)	0.975 (0.011)	0.991 (0.004)	0.979 (0.007)	0.983 (0.037)	
$i = 9$	0.993 (0.005)	0.484 (0.146)	0.943 (0.070)	0.989 (0.007)	0.979 (0.013)	0.995 (0.005)	0.965 (0.016)	0.990 (0.005)



Table 15: Posterior means (standard deviation) of  $\xi^*$ :  
U.S. stock returns. CRSV model.

$\xi_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	-0.0431 (0.0161)							
$i = 3$	-0.0013 (0.0120)	-0.0682 (0.0173)						
$i = 4$	-0.0250 (0.0115)	0.0321 (0.0180)	0.0668 (0.0147)					
$i = 5$	-0.0724 (0.0132)	0.0398 (0.0167)	0.0168 (0.0249)	-0.1304 (0.0183)				
$i = 6$	-0.1736 (0.0195)	0.0253 (0.0089)	0.0374 (0.0147)	-0.0109 (0.0224)	0.0123 (0.0139)			
$i = 7$	-0.0361 (0.0107)	0.0315 (0.0066)	0.0299 (0.0136)	0.0291 (0.0194)	-0.0504 (0.0092)	0.0088 (0.0135)		
$i = 8$	-0.0264 (0.0080)	0.0044 (0.0231)	0.0173 (0.0174)	0.0298 (0.0160)	-0.0013 (0.0073)	-0.0524 (0.0107)	0.0014 (0.0196)	
$i = 9$	-0.0023 (0.0096)	0.0444 (0.0179)	0.0085 (0.0164)	-0.0387 (0.0145)	0.0099 (0.0098)	0.0032 (0.0052)	0.0329 (0.0120)	0.0380 (0.0083)

Table 16: Posterior means (standard deviation) of  $\mathbf{C}^*$ :  
U.S. stock returns. CRSV model.

$\sigma_{u,ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.126 (0.002)							
$i = 3$	0.134 (0.002)	0.182 (0.003)						
$i = 4$	0.138 (0.002)	0.179 (0.003)	0.151 (0.002)					
$i = 5$	0.202 (0.003)	0.272 (0.004)	0.231 (0.003)	0.245 (0.004)				
$i = 6$	0.150 (0.002)	0.208 (0.003)	0.178 (0.002)	0.189 (0.003)	0.123 (0.002)			
$i = 7$	0.140 (0.002)	0.195 (0.003)	0.153 (0.002)	0.173 (0.002)	0.116 (0.002)	0.162 (0.002)		
$i = 8$	0.124 (0.002)	0.183 (0.003)	0.153 (0.002)	0.175 (0.003)	0.103 (0.002)	0.139 (0.002)	0.150 (0.001)	
$i = 9$	0.114 (0.001)	0.142 (0.006)	0.126 (0.002)	0.140 (0.002)	0.092 (0.001)	0.130 (0.002)	0.130 (0.002)	0.138 (0.002)

Table 17: Posterior means (standard deviation) of  $\mathbf{D}^*$ :  
U.S. stock returns. CRSV model.

$\tau_{ij}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$i = 2$	0.0283 (0.0041)							
$i = 3$	0.0143 (0.0028)	0.0198 (0.0042)						
$i = 4$	0.0177 (0.0027)	0.0205 (0.0046)	0.0220 (0.0035)					
$i = 5$	0.0119 (0.0018)	0.0097 (0.0019)	0.0123 (0.0038)	0.0132 (0.0026)				
$i = 6$	0.0125 (0.0017)	0.0117 (0.0027)	0.0127 (0.0029)	0.0232 (0.0051)	0.0062 (0.0014)			
$i = 7$	0.0048 (0.0011)	0.0062 (0.0022)	0.0182 (0.0065)	0.0098 (0.0018)	0.0093 (0.0018)	0.0072 (0.0027)		
$i = 8$	0.0071 (0.0025)	0.0097 (0.0045)	0.0088 (0.0022)	0.0181 (0.0040)	0.0044 (0.0007)	0.0070 (0.0014)	0.0101 (0.0018)	
$i = 9$	0.0042 (0.0010)	0.0512 (0.0151)	0.0069 (0.0028)	0.0082 (0.0020)	0.0040 (0.0010)	0.0034 (0.0013)	0.0077 (0.0017)	0.0057 (0.0012)

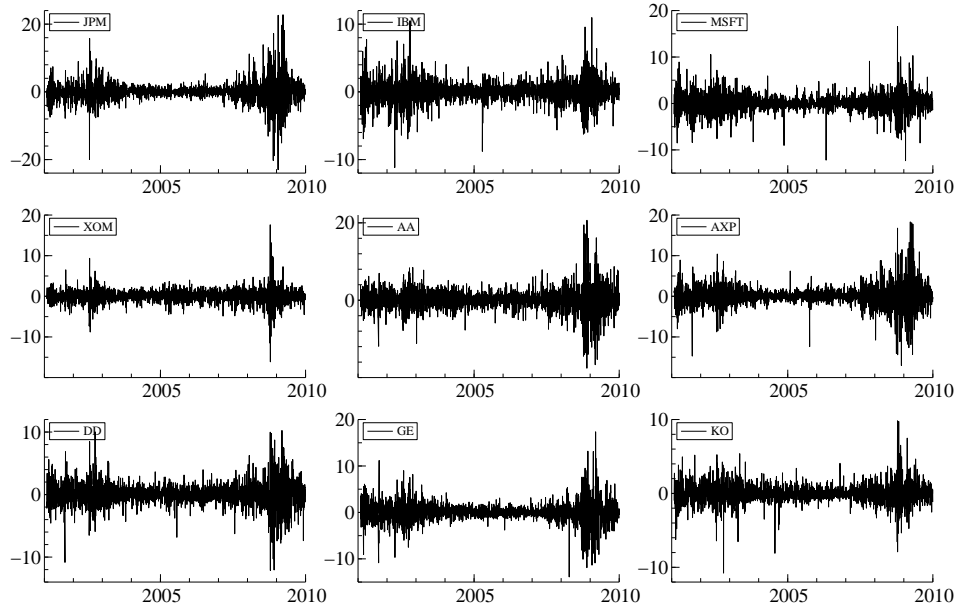


Figure 3: Time series plot of daily returns,  $y_{t1}, \dots, y_{9t}$ .

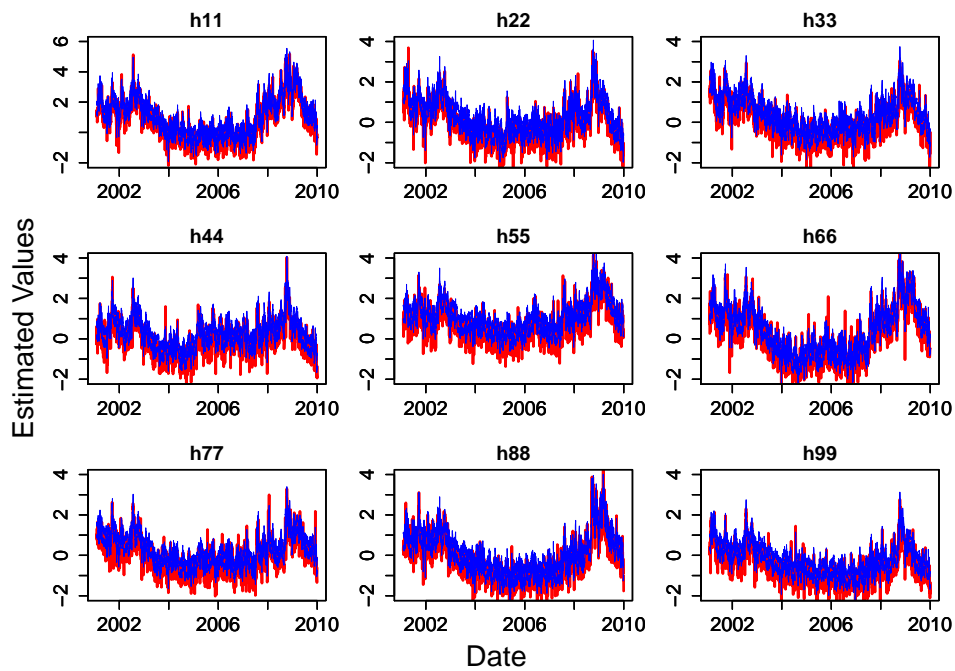


Figure 4: 95% credible intervals (blue) for  $\xi + h_{11,t}, \dots, \xi + h_{99,t}$ , and  $x_{1t}, \dots, x_{9t}$  (red)

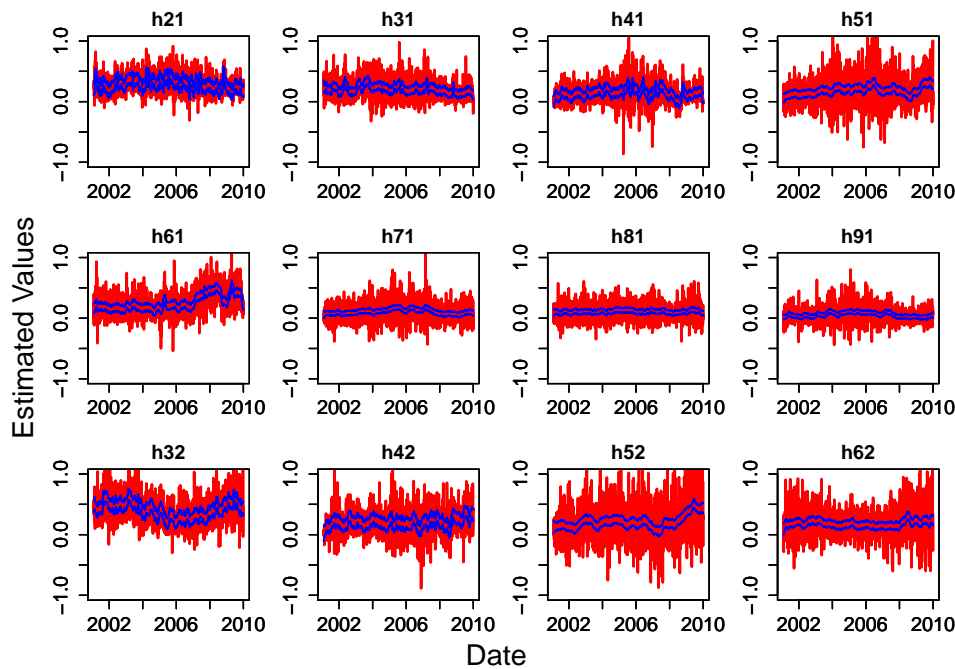


Figure 5: 95% credible intervals (blue) for  $\xi_{21}^* + h_{21,t}, \dots, \xi_{62}^* + h_{62,t}$ , and  $x_{21,t}^*, \dots, x_{62,t}^*$  (red)

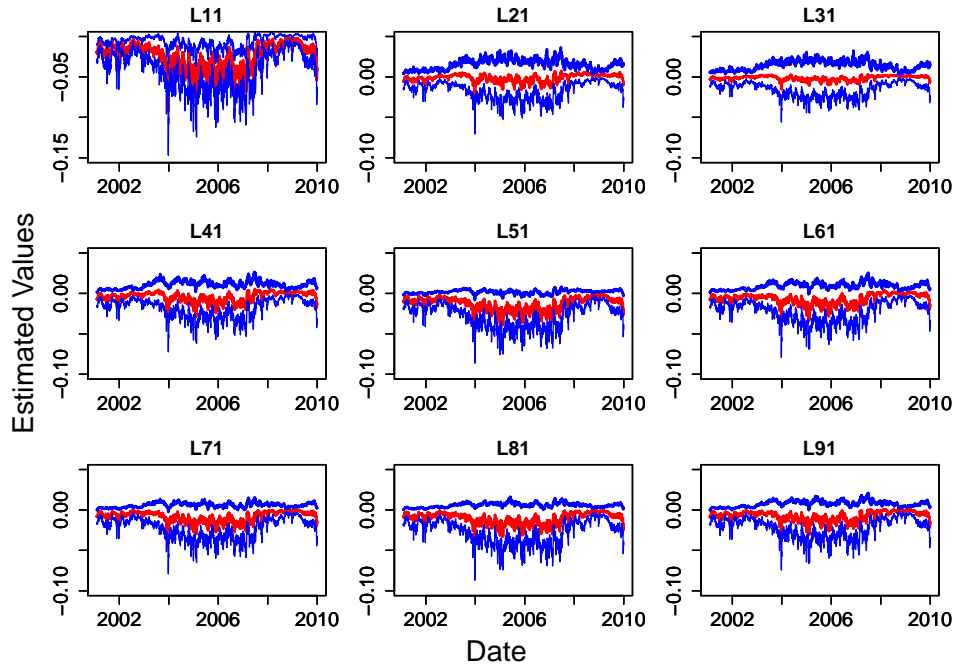


Figure 6: Posterior means (red) and 95% credible intervals (blue) for  $L_{11,t}, \dots, L_{91,t}$  where  $\mathbf{L}_t = \mathbf{R}\mathbf{V}_t^{-1/2}\mathbf{H}_t^*$ :

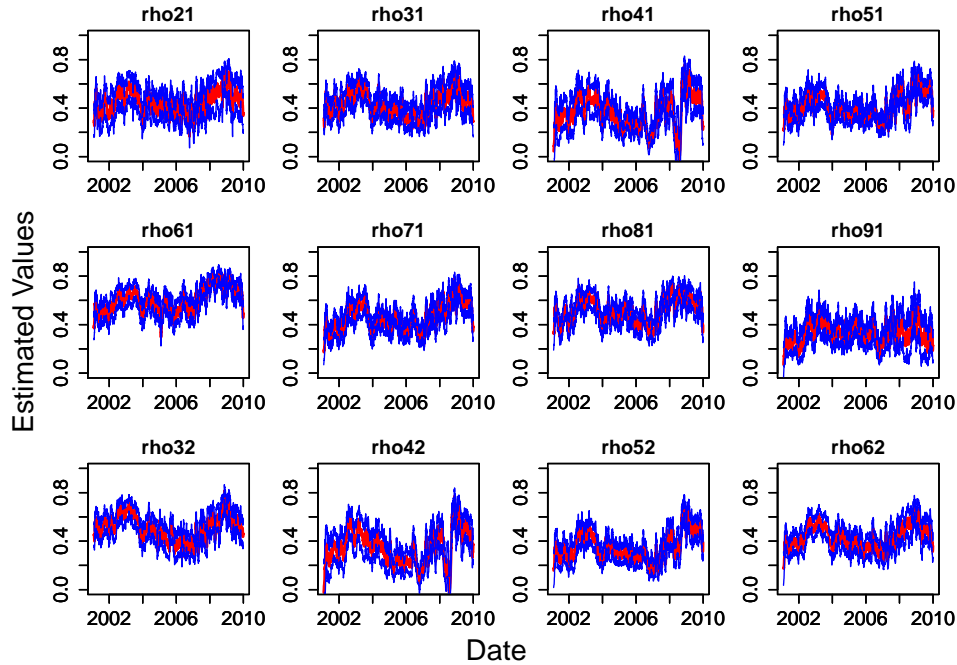


Figure 7: Posterior means (red) and 95% credible intervals (blue) for  $\rho_{21,t}, \dots, \rho_{62,t}$ .

### B.3 Portfolio performance: U.S. stock returns data

Table 18: Cumulative realized returns for three strategies.

Minimum-Variance	$\mu_p^*=0.004$	$\mu_p^*=0.01$	$\mu_p^*=0.1$
CSV	-0.538	-2.282	-28.45
CRSV ( $\mathbf{R} = \mathbf{O}, \mathbf{S} = \mathbf{I}$ )	2.367	3.052	13.31
CRSV	1.609	1.683	2.786
Mean-Variance	$\gamma = 6$	$\gamma = 10$	$\gamma = 15$
CSV	1.256	1.316	1.347
CRSV ( $\mathbf{R} = \mathbf{O}, \mathbf{S} = \mathbf{I}$ )	1.108	1.227	1.287
CRSV	1.279	1.330	1.356
Maximum-Return	$\sigma_p^{*2} = 0.001$	$\sigma_p^{*2} = 0.01$	$\sigma_p^{*2} = 0.1$
CSV	1.124	0.511	-1.427
CRSV ( $\mathbf{R} = \mathbf{O}, \mathbf{S} = \mathbf{I}$ )	1.142	0.569	-1.243
CRSV	1.172	0.662	-0.947

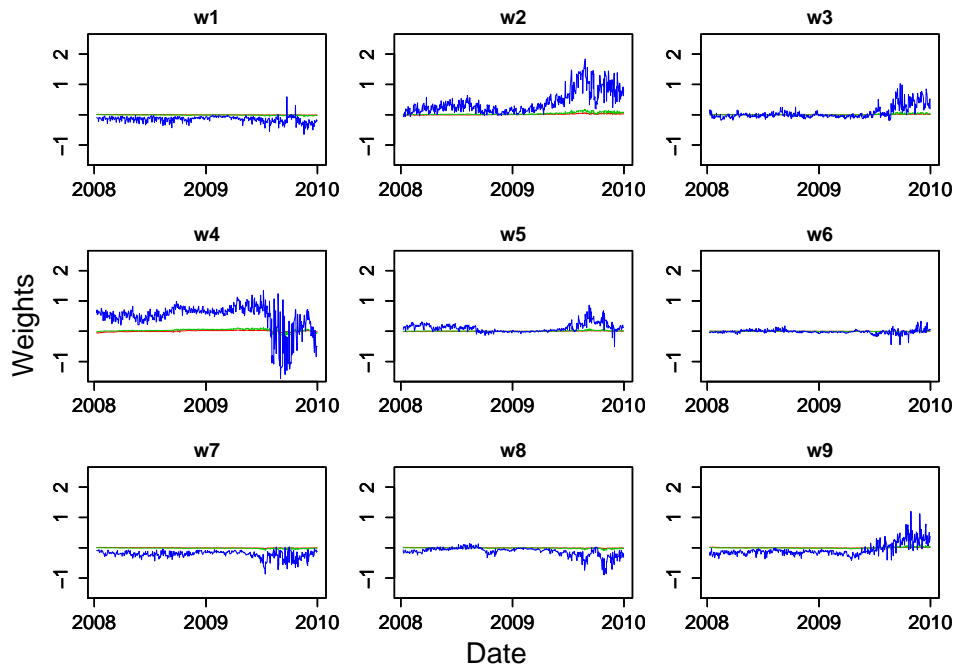


Figure 8: Time series plot of the portfolio weight  $\omega_{it}$  for minimum-variance strategy:  $\mu_p^* = 0.004$  (red),  $\mu_p^* = 0.01$  (green) and  $\mu_p^* = 0.1$  (blue). CRSV model.

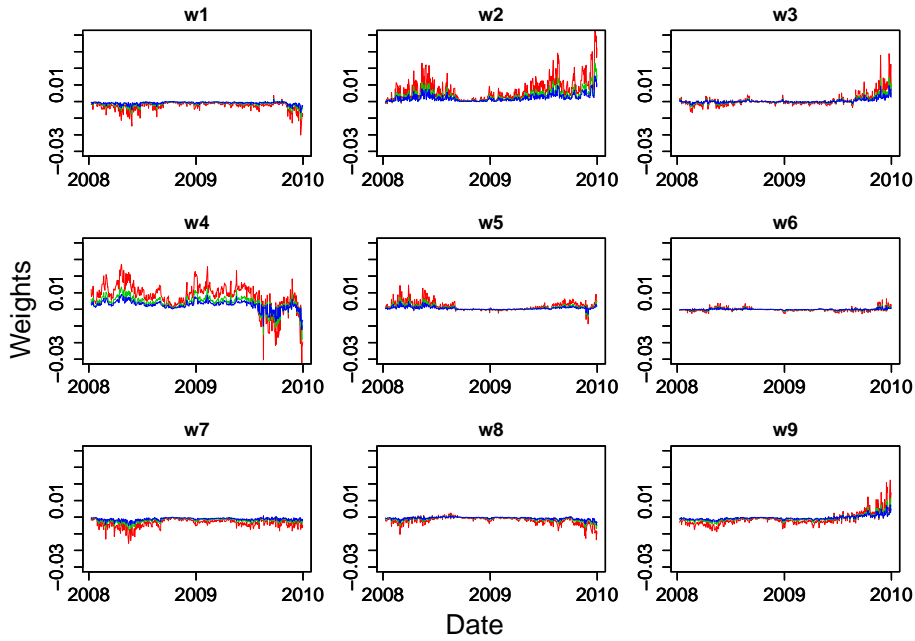


Figure 9: Time series plot of the portfolio weight  $\omega_{it}$  for mean-variance strategy:  $\gamma = 6$  (red),  $\gamma = 10$  (green) and  $\gamma = 15$  (blue). CRSV model.

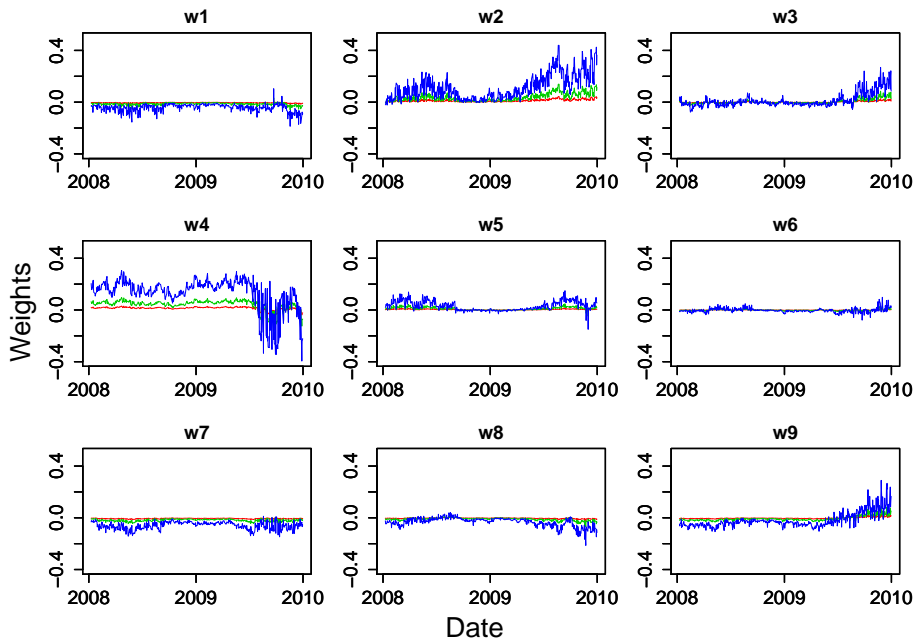


Figure 10: Time series plot of the portfolio weight  $\omega_{it}$  for maximum return strategy:  $\sigma_p^{*2} = 0.001$  (red),  $\sigma_p^{*2} = 0.01$  (green) and  $\sigma_p^{*2} = 0.1$  (blue). CRSV model.

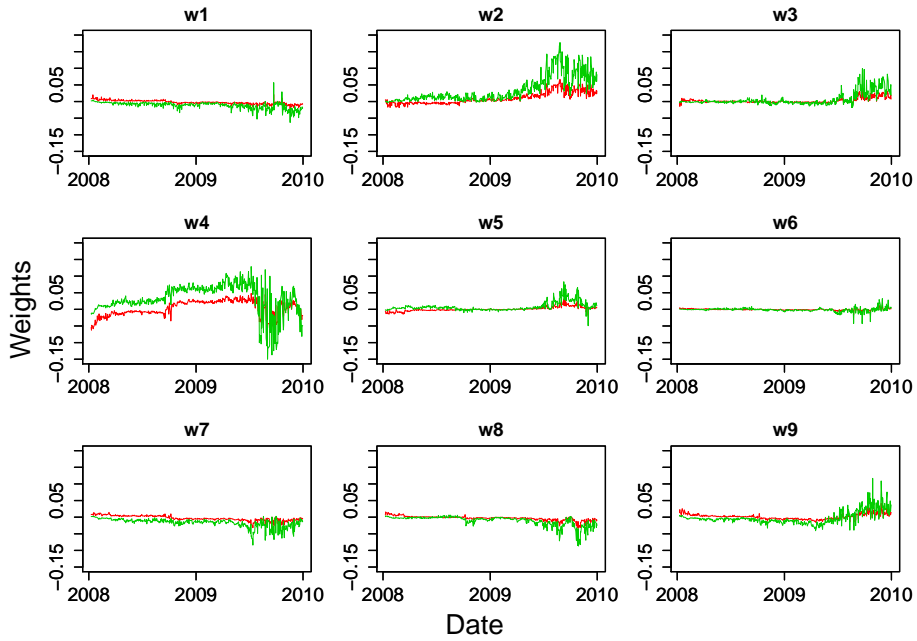


Figure 11: Time series plot of the portfolio weight  $\omega_{it}$  for minimum-variance strategy:  $\mu_p^* = 0.004$  (red) and  $\mu_p^* = 0.01$  (green). CRSV model.

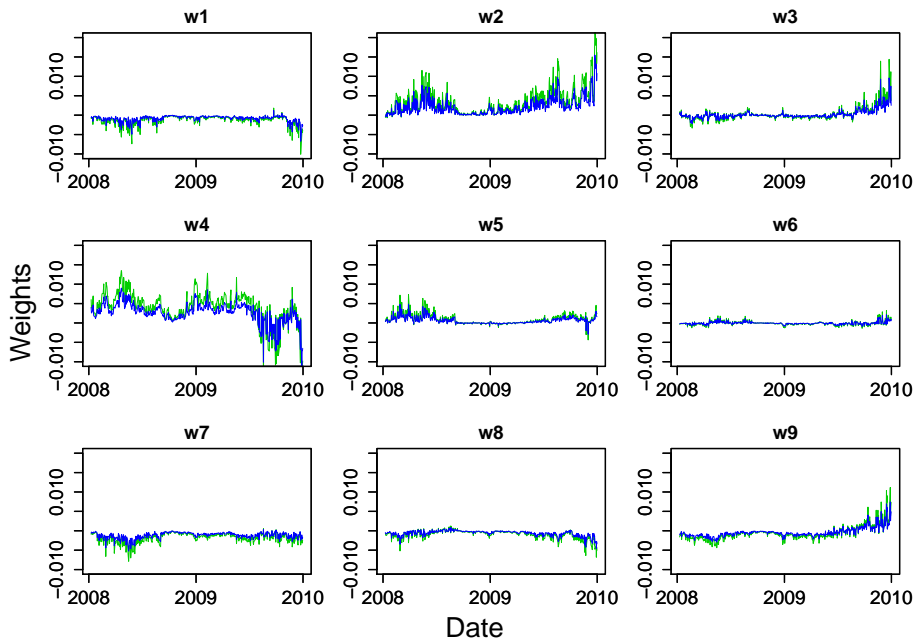


Figure 12: Time series plot of the portfolio weight  $\omega_{it}$  for mean-variance strategy:  $\gamma = 10$  (green) and  $\gamma = 15$  (blue). CRSV model.

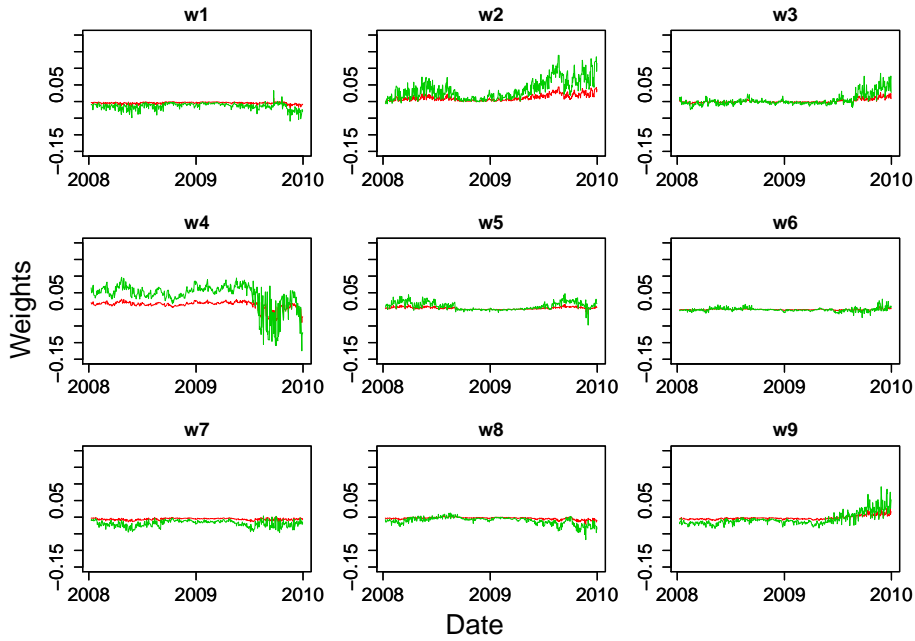


Figure 13: Time series plot of the portfolio weight  $\omega_{it}$  for maximum return strategy:  $\sigma_p^{*2} = 0.001$  (red) and  $\sigma_p^{*2} = 0.01$  (green). CRSV model.

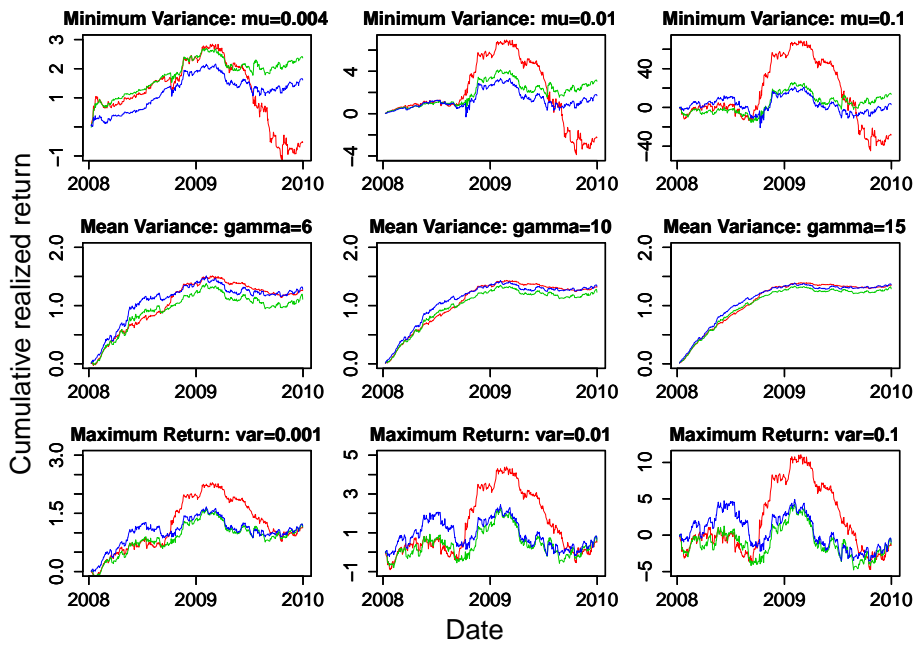


Figure 14: Cumulative realized return: red (CSV), green (CRSV ( $\mathbf{R} = \mathbf{O}, \mathbf{S} = \mathbf{I}$ )) and blue (CRSV)



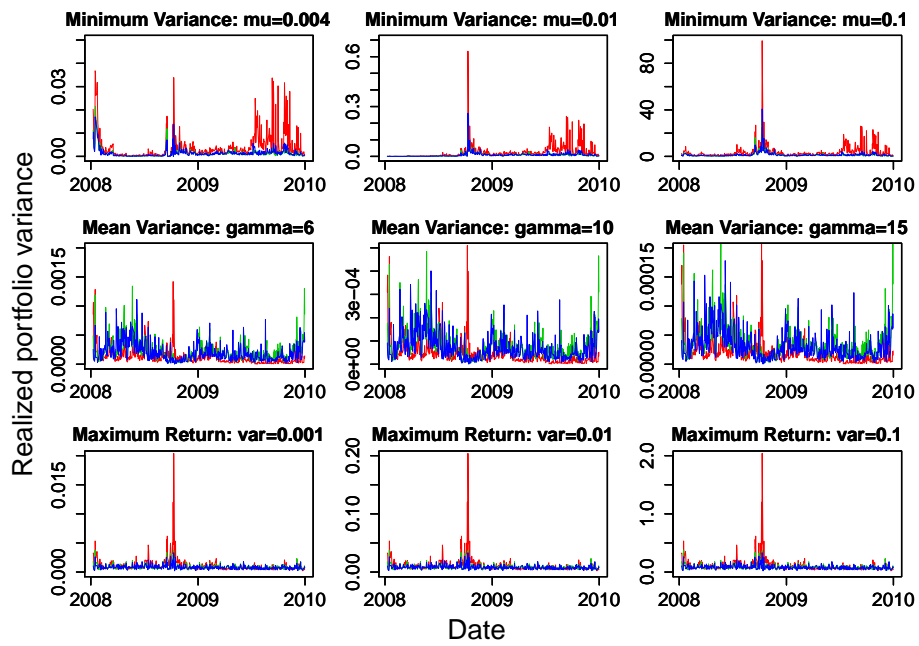


Figure 15: Realized portfolio variance  $\hat{\omega}'_t \hat{\Sigma}_t \hat{\omega}_t$ , where  $\hat{\Sigma}_t$  is realized covariance at time  $t$ : red (CSV), green (CRSV ( $\mathbf{R} = \mathbf{O}, \mathbf{S} = \mathbf{I}$ )) and blue (CRSV)