博士論文

論文題目
Efficient Bayesian Estimation of Time Series Models and Its Applications
(時系列モデルの効率的なベイズ推定とその応用)

氏名　黒瀬 雄大

Contents

Acknowledgement

I would like to express my sincere gratitude to all of the people who helped me complete this thesis.

It is difficult to overstate my gratitude to my adviser, Professor Yasuhiro Omori, Faculty of Economics, University of Tokyo for his patience and enthusiasm. I could not have finished the work without his warm encouragement and insightful suggestions.

Besides my adviser, I would like to appreciate Professors and ex-Professors of Statistics at Faculty of Economics, University of Tokyo: Naoto Kunitomo, Akimichi Takemura, Yoshihiro Yajima, Hidehiko Ichimura, Tatsuya Kubokawa, Katsumi Shimotsu, Yuzo Maruyama and Yoichi Arai for their generous supports and thoughtful suggestions. My sincere thank goes to Professor Toshiaki Watanebe (Hitotsubashi University), Professor Teruo Nakatsuma (Keio University) and Associate Professor Haruhisa Nishino (Chiba University) for giving me presentation opportunity at their seminars. I am also indebted to Professor Kazuya Kamiya (University of Tokyo) for his beneficial support.

I thank my colleagues and ex-colleagues at Graduate School of Economics, University of Tokyo for their emotional supports and stimulating advice. I am especially grateful to Koji Miyawaki and Tsunehiro Ishihara.

Finally, I would like to thank my family for their patience and warm-hearted encouragement throughout my life.

Chapter 1

Overview

The last decades have brought dynamic changes in the way that researchers analyze various types of data using a Bayesian approach. The goal to Bayesian inference is to obtain a set of draws from the posterior distributions with enough draws so that the quantities of interest can be estimated accurately. As recently as thirty years ago, the computations associated with Bayesian modeling was very challenging except for the case with natural conjugate priors have been assumed so that it was often easy to draw from the posterior distribution directly. The realistic models, which often required integrations in terms of many unknown parameters, was too complex to obtain random draws from the posterior distribution of the model parameters. However, around 1990, the Markov chain Monte Carlo (MCMC) "revolution" occurred in Bayesian computing and such a situation of Bayesian modeling have been changed completely. The most popular of these methods are known as the Gibbs sampler and the Metropolis-Hastings (MH) algorithm, which can greatly facilitate calculation of the integrals. With the development of high-power computing, researchers in diverse areas are trying to conduct a statistical modeling using Bayesian approach. Details are, e.g., in Gelman, Carlin, Stern, and Rubin (2003), Koop, Poirier, and Tobias (2007) and Meyn and Tweedie (2009).

When we turn to the time series modeling of economic and financial data, we must face at the existence of so many time-dependent latent variables, e.g., the mean, the variance and the covariance or the quantile. To obtain the estimates of the model parameters of interest, we are often required to calculate the integral of the likelihood function in terms of a large number of such latent variables that makes maximizing the likelihood very difficult. On the other hand, since MCMC methods enables evaluation of the integral, the author decided to take the Bayesian approach in this thesis.

In sampling the unobserved parameters in time series models, early studies suggest that they are often time-dependent and highly-correlated. Many researchers take a Bayesian approach and propose MCMC algorithms to draw samples from the posterior of the parameters, but part of them are known to be inefficient. (Since MCMC samples are not independent of each other, if the correlation becomes higher, we need to obtain the larger number of samples to estimate the posterior density of the model parameters with accuracy and we say the algorithm is more inefficient.)

This thesis provides Bayesian analyses of four time series models and also develops the fast and efficient estimation algorithm using MCMC method. It is noted that our models with time-dependent latent variables are given in the state-space form (Durbin and Koopman (2001) and Brockwell and Davis (1991)). Furthermore, empirical studies are presented using economic and financial time series datasets. Organizations of this thesis is as follows.

In Chapter 2, a smoothing spline is used to propose a novel model for the time-varying quantile of the univariate time series using a state space approach. A correlation is further incorporated between the dependent variable and its one-step-ahead quantile. Using a Bayesian approach, an efficient Markov chain Monte Carlo algorithm is described where we use the multi-move sampler, which generates simultaneously latent time-varying quantiles. Numerical examples are provided to show its high sampling efficiency, in comparison with the simple algorithm that generates one latent quantile at a time given other latent quantiles. Furthermore, using Japanese inflation rate data, an empirical analysis is provided with model comparisons.

Chapter 3 proposes a multivariate stochastic volatility (MSV) model with dynamic equicorrelation and cross leverage effect and conducts the estimation. Using a Bayesian approach, an efficient Markov chain Monte Carlo algorithm is described where we use the multi-move sampler, which generates multiple latent variables simultaneously. Numerical examples are provided to show its sampling efficiency in comparison with the simple algorithm that generates one latent variable at a time given other latent variables. Furthermore, the proposed model is applied to the multivariate daily stock price index data. The empirical study shows that our novel model provides a substantial improvement in forecasting with respect to out-of-sample hedging performances.

In Chapter 4, we propose simultaneous modeling of the multivariate daily returns and the related realized measure with dynamic equicorrelation and cross leverage effect. We describe the simple (but efficient) MCMC sampling algorithm for our model. It is explained that the biases in the realized measure owing to market microstructure noise, non-trading hours and non-synchronous trading can be estimated within our modeling framework. Numerical examples are provided and the proposed model is applied to the multivariate stock data. We find the persistence in the volatilities, the dynamics of correlations among the returns and the existence of weak cross leverage effects.

In Chapter 5, the MSV model of Chapter 3, which incorporates the SV model with dynamic equicorrelation and cross leverage effect, is extended with the dynamic block-equicorrelation structure. We also extend the model and propose simultaneous modeling of the multivariate daily returns and the related realized measure in line with chapter 4. With the Bayesian estimation scheme via Markov chain Monte Carlo method, it makes it possible to obtain the estimation results for the parameters in our proposed model. Numerical examples are provided and the proposed model is applied to the multivariate stock data. We find the persistence in the volatilities and the dynamic diagonal-block-equicorrelations and the existence of weak cross leverage effects. The biases in the realized measure owing to market microstructure noise, non-trading hours and non-synchronous trading can also be estimated.

Chapter 2

Bayesian analysis of time-varying quantiles using a smoothing spline

2.1 Introduction

Time-varying quantiles have been receiving attention recently, and various econometric models have been proposed. Tail quantiles are especially important for financial risk management or policy evaluation because they are useful to describe the extreme behavior of the dependent variable in serious events, such as the financial crisis. The Value at Risk (VaR) is an example of tail quantiles in a financial time series, which is one of the well-known risk measures that are associated with an asset or a portfolio of many assets.

As discussed in the vast literature of financial econometrics, time-varying variances are found to exist in empirical studies of financial time series (see, e.g., Engle (1995), Shephard (2005)). As the variances of the dependent variables change over time, the corresponding quantiles vary over time. When we focus on the tail behavior of financial time series, it is necessary to describe the time-dependent structure that is appropriate for the tail quantile or τ-quantile (which is roughly defined as the value that will be exceeded by the dependent variable with probability τ) with τ close to zero or one. For the i.i.d. observations, the estimate is given by the 100τ-th percentile of the samples, and in the generalized linear model, it is given by the minimizer of some loss function (Koenker and Bassett (1978)). Based on Koenker and Bassett (1978), several models have been proposed for time-varying quantiles: Conditional Autoregressive Value at Risk (CAViaR) model (Engle and Manganelli (2004)), Quantile Autoregressive (QAR) model (Koenker and Xiao (2006)) and Dynamic Additive Quantile (DAQ) model (Gourieroux and Jasiak (2008)).

On the other hand, Koenker and Machado (1999) noted that solving the loss function of Koenker and Bassett (1978) is equivalent to obtaining the maximal likelihood estimate of the τ-quantile assuming some distribution for the dependent variable. For a Bayesian inference in the quantile regression, Yu and Moyeed (2001) took this approach to obtain the posterior distribution of the quantile and the parameters using the Markov chain Monte Carlo (MCMC) method. Gerlach, Chen, and Chan (2011) proposed a threshold-CAViaR model that extends the CAViaR model for the Bayesian analysis of time-varying quantiles.

In modeling time-varying quantiles, it is important to forecast the future tail behavior of the time series for risk management as well as to describe its past movement. Overfitting the model to the past dataset could result in a poor future forecasting for practical applications. Thus, recently, the backtesting procedure has often been implemented to investigate such a forecasting performance, e.g., by checking that the proportion of observations that exceed the estimated one-step-ahead quantiles is equal to its expected value.

One approach to avoid overfitting to the past dataset is to make the time-varying quantile function so smooth that it produces stable predictions. In this paper, we use the smoothing spline for that purpose as discussed in De Rossi and Harvey (2009) and propose an efficient Bayesian estimation using the MCMC method in which we exploit a state space representation to apply a simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002)) to generate the latent time-varying quantiles from the posterior distributions. The model is further extended to incorporate a correlation between the dependent variable and its one-step-ahead quantile.

The rest of this article is organized as follows. In Section 2.2, we propose the time-varying quantile model using the smoothing spline. Section 2.3 describes an efficient Bayesian estimation method for the proposed model using a multi-move sampling method. A single-move sampling method that is simple but inefficient is also described as a benchmark. We show that a normal variance-mean mixture representation of the measurement error leads us to exploit the efficient sampling method for the linear Gaussian state space model. Section 2.4 illustrates our estimation method using simulated data and shows that our MCMC algorithm is efficient. Section 2.5 applies the proposed time-varying quantile model to the inflation rate based on the domestic Corporate Goods Price Index (CGPI) of Japan. The backtesting procedure and the model comparison using DIC (Deviance Information Criterion) are conducted using our models and the CAViaR model. Section 2.6 concludes this paper.

2.2 Time-varying quantile model

2.2.1 Quantile regression model

Let y_t denote the dependent variable at time t ($t = 1, \ldots, n$) whose distribution function is given by $F(y) = \Pr(y_t \le y)$. For any fixed $0 < \tau < 1$, we define a τ-quantile as

$$\xi(\tau) := F^{-1}(\tau) = \inf\{y | F(y) \ge \tau\}, \tag{2.1}$$

and we define the loss function, called a "check function" as

$$\rho_\tau(u) = (\tau - \mathrm{I}(u < 0))(u), \tag{2.2}$$

where $\mathrm{I}(\cdot)$ is an indicator function. Then, the expected loss,

$$\mathrm{E}(\rho_\tau(y_t - \xi(\tau))), \tag{2.3}$$

is minimized when $\xi(\tau)$ satisfies $F(\xi(\tau)) = \tau$. Using this loss function, Koenker and Bassett (1978) considered a quantile regression for i.i.d. observations assuming that $\xi(\tau) = \boldsymbol{x}'\boldsymbol{b}$, where $\boldsymbol{x}$ is a vector of explanatory variables and $\boldsymbol{b}$ is a corresponding regression coefficient vector (see, e.g., Koenker (2005) for the asymptotic property of the minimum loss estimator and the numerical method using linear programing).

Yu and Moyeed (2001) assumed an asymmetric double exponential (or asymmetric Laplace) density that corresponds to the loss function and described an MCMC algorithm for the quantile regression using a Bayesian approach, where y_t is i.i.d. with the probability density function given $\xi(\tau)$,

$$f(y_t|\xi(\tau)) = \frac{\tau(1-\tau)}{\lambda} \exp\Big(-\frac{1}{\lambda}\rho_\tau(y_t - \xi(\tau))\Big), \tag{2.4}$$

and the first and second moments of $\epsilon_t = y_t - \xi(\tau)$ are (see, e.g., Kotz, Kozubowski, and Podgórski (2001))

$$\lambda\frac{-\tau^2 + (1-\tau)^2}{(1-\tau)\tau}, \quad 2\lambda^2\frac{\tau^3 + (1-\tau)^3}{(1-\tau)^2\tau^2}. \tag{2.5}$$

Note that $\xi(\tau)$, which maximizes the logarithm of this density function also minimizes the expected loss function, (2.3).

This paper extends the static model (2.4) to describe the time-varying quantiles, and we let ξ_t denote the time-varying τ-th quantile, where we suppress τ in a parenthesis and add a subscript t to emphasize that it depends on time t.

2.2.2 Time-varying quantile model using a smoothing spline

We assume that $\xi_t = h(t)$ changes slowly over time t and, hence, that $h(t)$ is a smooth function of t. That is, $h(t)$ is of a C^{m-1}-class, and its m-th derivative is square integrable and is the smoothing spline function, which minimizes

$$\sum_{t=1}^{n} \rho_\tau(y_t - h(t)) + \lambda_m \int [h^{(m)}(t')]^2 dt', \tag{2.6}$$

for given m and λ_m. Furthermore, we assume that (i) h and its first $(m-1)$ derivatives at time $t = 1$ follow the m-variate normal distribution with mean $\mathbf{0}_m$ and covariance matrix κE_m, where $\mathbf{0}_m$ is an m-dimensional zero vector, E_m is an identity matrix of size m, and κ is some known constant,

$$(h(1), h'(1), \ldots, h^{(m-1)}(1))' \sim \mathrm{N}(\mathbf{0}_m, \kappa E_m), \tag{2.7}$$

and that (ii)

$$h(t) = \sum_{j=1}^{m} \frac{(t-1)^{j-1}}{(j-1)!} h^{(j-1)}(1) + \sigma_\eta \int_1^t \frac{(t-s)^{m-1}}{(m-1)!} dW_s, \tag{2.8}$$

where W_s is a Wiener process.

Under these additional assumptions (i)(ii), if $\lambda_m = \lambda/(2\sigma_\eta^2)$, then the mode of the distribution of $(h(1), h(2), \ldots, h(n)|y_1, \ldots, y_n)$ converges to the solution of the smoothing spline problem as $\kappa \to \infty$ (De Rossi and Harvey (2009)). Noting that equation (2.8) can be represented in the following state space form (see, e.g., Wecker and Ansley (1983), Kohn and Ansley (1987)),

$$\boldsymbol{h}(t+1) = T\boldsymbol{h}(t) + \boldsymbol{\eta}(t), \boldsymbol{\eta}(t) \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{2.9}$$

$$\boldsymbol{h}(t) = (h(t), dh(t)/dt, \ldots, d^{m-1}h(t)/dt^{m-1})', \tag{2.10}$$

$$(T)_{ij} = \begin{cases} 1/(j-i)! & j \geq i, \\ 0 & j < i, \end{cases} \tag{2.11}$$

$$(Q)_{ij} = \frac{1}{(m-i)!(m-j)!(2m-i-j+1)}, \tag{2.12}$$

we propose a time-varying quantile model using the smoothing spline (TQSS model) in the state space representation with an asymmetric double exponential measurement error:

$$y_t = Z\boldsymbol{\xi}_t + \epsilon_t, \epsilon_t \sim \mathrm{aDE}_\tau(\lambda), \tag{2.13}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{2.14}$$

where

$$\boldsymbol{\xi}_t = (\xi_t, \tilde{\boldsymbol{\xi}}_t')', \ \ \tilde{\boldsymbol{\xi}}_t = (d\xi_t/dt, \ldots, d^{m-1}\xi_t/dt^{m-1})', \ \ (\xi_1, \xi_1^{(1)}, \ldots, \xi_1^{(m-1)})' \sim \mathrm{N}(\mathbf{0}_m, \kappa E_m),$$

$$Z = (1, \mathbf{0}_{m-1}').$$

Yue and Rue (2011) consider an additive mixed quantile regression model for longitudinal data with quantile functions including such a smooth function.

2.3 Bayesian estimation

2.3.1 Prior and posterior densities

For prior distributions of σ_η^2 and λ, we assume

$$\sigma_\eta^2 \sim \mathrm{IG}(\alpha_0/2, \beta_0/2), \quad \lambda \sim \mathrm{IG}(\alpha_0^*, \beta_0^*), \tag{2.15}$$

where $\mathrm{IG}(a, b)$ denotes an inverted gamma distribution with shape parameter a and scale parameter b. Let $I_t = \mathrm{I}(y_t - \xi_t < 0), \ t = 1, \ldots, n$. Then, the joint posterior density function is

$$\begin{aligned} f(\sigma_\eta^2, \lambda, \{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n) &\propto \prod_{t=1}^n f(y_t|\boldsymbol{\xi}_t, \lambda) \times f(\boldsymbol{\xi}_1) \prod_{t=1}^{n-1} f(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \sigma_\eta^2) \times f(\sigma_\eta^2) f(\lambda) \\ &\propto \lambda^{-n} \exp\left(- \frac{\sum_{t=1}^n (\tau - I_t)(y_t - Z\boldsymbol{\xi}_t)}{\lambda} \right) \times \exp\left(- \frac{1}{2\kappa} \boldsymbol{\xi}_1' \boldsymbol{\xi}_1 \right) \\ &\times (\sigma_\eta^2)^{-m\frac{n-1}{2}} \exp\left(- \frac{1}{2} \sum_{t=1}^{n-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' (\sigma_\eta^2 Q)^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t) \right) \\ &\times (\sigma_\eta^2)^{-(\frac{\alpha_0}{2}+1)} \exp\left(- \frac{\beta_0}{2\sigma_\eta^2} \right) \times \lambda^{-(\alpha_0^*+1)} \exp\left(- \frac{\beta_0^*}{\lambda} \right). \end{aligned} \tag{2.16}$$

We implement the MCMC algorithm in five blocks:

1. Initialize $\sigma_\eta^2, \lambda, \{\boldsymbol{\xi}_t\}_{t=1}^n$.

2. Generate $\sigma_\eta^2 | \{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n \sim \mathrm{IG}(\alpha_1/2, \beta_1/2)$, where

$$\alpha_1 = \alpha_0 + m(n-1), \ \beta_1 = \beta_0 + \sum_{t=1}^{n-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' Q^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t). \tag{2.17}$$

3. Generate $\lambda | \{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n \sim \mathrm{IG}(\alpha_1^*, \beta_1^*)$, where

$$\alpha_1^* = \alpha_0^* + n, \ \beta_1^* = \beta_0^* + \sum_{t=1}^n (\tau - I_t)(y_t - Z\boldsymbol{\xi}_t). \tag{2.18}$$

4. For $t = 1, \ldots, n$, generate $\{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$ as in Section 2.3.2.

5. Go to 2.

2.3.2 Generation of latent time-varying quantiles

2.3.2.1 Single-move sampling method

A simple sampling method for $\{\boldsymbol{\xi}_t\}_{t=1}^n$ is a single-move sampler that draws a single latent variable $\boldsymbol{\xi}_t$ at a time given the other $\boldsymbol{\xi}_t$'s and the parameters. The other method is a multi-move sampler that draws all of $\boldsymbol{\xi}_t$'s simultaneously. A single-move sampler is simpler than a multi-move sampler, but a multi-move sampler is known to be more efficient (de Jong and Shephard (1995)). As a benchmark, we describe the single-move sampling method as follows (see Appendix 2.A.1 for details):

Step4. For $t = 1, \ldots, n$,

4.a Generate $I_t | \tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

4.b Generate $\xi_t | I_t, \tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

4.c Generate $\tilde{\boldsymbol{\xi}}_t | \xi_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

2.3.2.2 Efficient multi-move sampling

A simulation smoother, an efficient sampler for the state variables was proposed by de Jong and Shephard (1995) and by Durbin and Koopman (2002) for the linear Gaussian state space model. However, in the time-varying quantile model, the measurement error is non-Gaussian, and such a simulation smoother cannot be applied directly. For non-Gaussian measurement models, it is usually necessary to approximate the non-Gaussian likelihood by the Gaussian likelihood in the previous literature for the MCMC implementation (e.g., Shephard and Pitt (1997), Watanabe and Omori (2004), Kim, Shephard, and Chib (1998), Omori, Chib, Shephard, and Nakajima (2007)), and in our proposed model, the error distribution is asymmetric double exponential. Noting that it is a normal variance-mean mixture with a generalized inverted Gaussian distribution (e.g., Kotz, Kozubowski, and Podgórski (2001), Tsionas (2003), Kozumi and Kobayashi (2011) and Yue and Rue (2011)), we rewrite

$$\epsilon_t = a v_t + b\sqrt{\lambda v_t} u_t, \tag{2.19}$$

$$a = \frac{1-2\tau}{\tau(1-\tau)},\ b^2 = \frac{2}{\tau(1-\tau)}, \tag{2.20}$$

where u_t is a standard normal variable and where v_t is an exponential variable with scale parameter (or mean) λ. Since the error distribution is normal conditionally on v_t, we can apply a simulation smoother for the linear Gaussian state space model.

Thus, to sample $\{\boldsymbol{\xi}_t\}_{t=1}^n$ efficiently, we rewrite the state space model (2.13)-(2.14) as follows:

$$y_t = Z\boldsymbol{\xi}_t + av_t + b\sqrt{\lambda v_t}u_t, u_t \sim \mathrm{N}(0,1), \quad v_t \sim \mathrm{Exp}(\lambda), \tag{2.21}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q). \tag{2.22}$$

The conditional joint posterior density of $\{\boldsymbol{\xi}_t\}_{t=1}^n$ and $\{v_t\}_{t=1}^n$ given $\{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$ is

$$\begin{aligned}
& f(\{\boldsymbol{\xi}_t\}_{t=1}^n, \{v_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \\
& \propto \prod_{t=1}^n f(y_t|\boldsymbol{\xi}_t, v_t, \lambda) \times \prod_{t=1}^n f(v_t|\lambda) \times f(\boldsymbol{\xi}_1) \prod_{t=1}^{n-1} f(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \sigma_\eta^2) \\
& \propto \prod_{t=1}^n v_t^{-\frac{1}{2}} \exp\left\{ -\sum_{t=1}^n \frac{(y_t - Z\boldsymbol{\xi}_t - av_t)^2}{2b^2\lambda v_t} \right\} \times \exp\left\{ -\frac{\sum_{t=1}^n v_t}{\lambda} \right\} \\
& \times \exp\left(-\frac{1}{2\kappa}\boldsymbol{\xi}_1'\boldsymbol{\xi}_1 - \frac{1}{2}\sum_{t=1}^{n-1}(\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)'(\sigma_\eta^2 Q)^{-1}(\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t) \right).
\end{aligned} \tag{2.23}$$

Thus, we generate v_t and $\boldsymbol{\xi}_t$ in two blocks:

Step4.

4.a' Generate $v_t|y_t, \boldsymbol{\xi}_t, \lambda \sim \mathrm{GIG}(1/2, \delta_t, \gamma)$[1] for $t = 1, \ldots, n$, where

$$\delta_t^2 = \frac{(y_t - Z\boldsymbol{\xi}_t)^2}{b^2\lambda},\ \gamma^2 = \frac{2}{\lambda} + \frac{a^2}{b^2\lambda}. \tag{2.24}$$

4.b' Generate $\{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n$ using a simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002))[2] .

Note that in Step 3 we generate $\lambda|\{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$ and in Step 4.a' we generate $v_t|y_t, \boldsymbol{\xi}_t, \lambda$ for $t = 1, \ldots, n$ using the collapsed Gibbs sampler (see, e.g., Chen, Shao, and Ibrahim (2000)).

[1]An efficient algorithm for random sampling from a generalized inverted Gaussian distribution ($\mathrm{GIG}(\nu, \delta, \gamma)$) is available from Dagpunar (1989), but the method does not work when δ or γ is too small. In that case, we adopt a rejection method using $\mathrm{Gamma}(\nu, \gamma^2/2)$ as a sampling method.

[2]For large m, the determinant of $\sigma_\eta^2 Q$ becomes so small that the round-off error for calculation is not negligible and that the covariance matrix of state ($\boldsymbol{\xi}_t$) estimation error can be non-positive definite. In such a case, we should use a square root filter by Morf and Kailath (1975) instead of an ordinary Kalman filter; see also Durbin and Koopman (2001).

2.3.3 Extension to the correlated errors

This subsection extends our model to describe a correlation between the dependent variable at time t and the latent time-varying quantile at time $(t+1)$. Based on the literature that addresses time-varying variances, the asymmetry or the leverage effect, which implies a decrease in the dependent variable at time t followed by an increase in the latent time-varying variance, is known to occur frequently in empirical studies. Similarly, as discussed in Engle and Manganelli (2004), the time-varying quantile at time $(t+1)$ is also influenced by the observation at time t in their analysis of stock returns data. Thus, we incorporate a correlation between y_t and ξ_{t+1} using the state space representation as follows[3]:

$$y_t = Z\boldsymbol{\xi}_t + av_t + b\sqrt{\lambda v_t}u_t, u_t \sim \mathrm{N}(0,1), \quad v_t \sim \mathrm{Exp}(\lambda), \tag{2.25}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{2.26}$$

where

$$\begin{pmatrix} b\sqrt{\lambda v_t}u_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_t), \ \Sigma_t = \left(\begin{array}{c|cc} b^2\lambda v_t & \rho b\sigma_\eta\sqrt{\lambda v_t q_{11}} & \mathbf{0}'_{m-1} \\ \hline \rho b\sigma_\eta\sqrt{\lambda v_t q_{11}} & & \\ \mathbf{0}_{m-1} & \multicolumn{2}{c}{\sigma_\eta^2 Q} \end{array}\right), \tag{2.27}$$

$$|\Sigma_t| = b^2\lambda v_t \sigma_\eta^{2m}(|Q| - \rho^2 q_{11}|Q_{22}|), \tag{2.28}$$

q_{11} denotes the $(1,1)$ element of Q and Q_{22} denotes the matrix obtained by excluding the first row and column of Q. Assuming a uniform prior distribution for the correlation parameter, ρ[4],

$$\rho \sim \mathrm{U}(-c_m, c_m), \quad c_m = \left\{|Q|/(q_{11}|Q_{22}|)\right\}^{\frac{1}{2}}, \tag{2.29}$$

(e.g., $c_2 = \frac{1}{2}, c_3 = \frac{1}{6}$) and assuming the same prior distributions for σ_η^2 and λ as in the previous subsection, the joint posterior density is

[3]Unlike the case without correlation, the mode of the distribution of $\{\xi_t\}_{t=1}^n|\{y_t\}_{t=1}^n$ may not necessarily converge to the solution to the smoothing spline problem (2.6).

[4]Given $\boldsymbol{\xi}_t$, the correlation coefficient between y_t and ξ_{t+1} is $\sqrt{\frac{(1-\tau)^2+\tau^2}{2\tau(1-\tau)}}\frac{\sqrt{\pi}}{2}\rho$. When $m=2$, the correlation coefficient of (y_t, ξ_{t+1}) is included in $(-0.208, 0.208)$ for $\tau = 0.1, 0.9$ and in $(-0.444, 0.444)$ for $\tau = 0.5$.

$$
\begin{aligned}
&f(\rho,\sigma_\eta^2,\lambda,\{v_t\}_{t=1}^n,\{\boldsymbol{\xi}_t\}_{t=1}^n|\{y_t\}_{t=1}^n)\\
&\propto f(\boldsymbol{\xi}_1)\prod_{t=1}^n f(y_t,\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t,\rho,\sigma_\eta^2,\lambda,v_t)\times\prod_{t=1}^n f(v_t|\lambda)\pi(\rho)\pi(\sigma_\eta^2)\pi(\lambda)\\
&\propto \exp\left\{-\frac{\boldsymbol{\xi}_1'\boldsymbol{\xi}_1}{2\kappa}\right\}\times\prod_{t=1}^n|\Sigma_t|^{-\frac{1}{2}}\exp\left\{-\sum_{t=1}^n\frac{1}{2}\begin{pmatrix}y_t-Z\boldsymbol{\xi}_t-av_t\\ \boldsymbol{\xi}_{t+1}-T\boldsymbol{\xi}_t\end{pmatrix}'\Sigma_t^{-1}\begin{pmatrix}y_t-Z\boldsymbol{\xi}_t-av_t\\ \boldsymbol{\xi}_{t+1}-T\boldsymbol{\xi}_t\end{pmatrix}\right\}\\
&\times\lambda^{-n}\exp\left\{-\frac{\sum_{t=1}^n v_t}{\lambda}\right\}\times(\sigma_\eta^2)^{-(\frac{\alpha_0}{2}+1)}\exp\left\{-\frac{\beta_0}{2\sigma_\eta^2}\right\}\times(\lambda)^{-(\alpha_0^*+1)}\exp\left\{-\frac{\beta_0^*}{\lambda}\right\}. \quad (2.30)
\end{aligned}
$$

We implement the MCMC algorithm in seven blocks:

1. Initialize $\rho,\sigma_\eta^2,\lambda,\{v_t\}_{t=1}^n,\{\boldsymbol{\xi}_t\}_{t=1}^n$.
2. Generate $\rho|\{y_t\}_{t=1}^n,\sigma_\eta^2,\lambda,\{v_t\}_{t=1}^n,\{\boldsymbol{\xi}_t\}_{t=1}^n$.
3. Generate $\sigma_\eta^2|\{y_t\}_{t=1}^n,\rho,\lambda,\{v_t\}_{t=1}^n,\{\boldsymbol{\xi}_t\}_{t=1}^n$.
4. Generate $\lambda|\{y_t\}_{t=1}^n,\rho,\sigma_\eta^2,\{v_t\}_{t=1}^n,\{\boldsymbol{\xi}_t\}_{t=1}^n$.
5. Generate $v_t|y_t,\boldsymbol{\xi}_t,\rho,\sigma_\eta^2,\lambda$ for $t=1,\ldots,n$.
6. Generate $\{\boldsymbol{\xi}_t\}_{t=1}^n|\{y_t\}_{t=1}^n,\rho,\sigma_\eta^2,\lambda,\{v_t\}_{t=1}^n$ using a simulation smoother as in the previous subsection.
7. Go to 2.

Generation of ρ. Let

$$
l(\rho):=-\frac{n}{2}\log(|Q|-\rho^2q_{11}|Q_{22}|)-\sum_{t=1}^n\frac{1}{2}\begin{pmatrix}y_t-Z\boldsymbol{\xi}_t-av_t\\ \boldsymbol{\xi}_{t+1}-T\boldsymbol{\xi}_t\end{pmatrix}'\Sigma_t^{-1}\begin{pmatrix}y_t-Z\boldsymbol{\xi}_t-av_t\\ \boldsymbol{\xi}_{t+1}-T\boldsymbol{\xi}_t\end{pmatrix}, \quad (2.31)
$$

which is the logarithm of the posterior density of ρ excluding the constant. To approximate the conditional posterior density by the truncated normal distribution, we use a Taylor expansion of the log posterior density of ρ to the second order around its mode $\hat\rho$ and let

$$
m_\rho=\hat\rho+s_\rho^2 l'(\hat\rho),\ s_\rho^2=-1/l''(\hat\rho), \quad (2.32)
$$

where $l'(\rho)=dl(\rho)/d\rho$ and $l''(\rho)=d^2l(\rho)/d\rho^2$. Then, we propose a candidate $\rho^\dagger$ from the truncated normal distribution over the interval $(-c_m,c_m)$, $\mathrm{TN}_{(-c_m,c_m)}(m_\rho,s_\rho^2)$ and accept it with probability

$$
\min\left[1,\frac{\exp\{l(\rho^\dagger)-(\rho-m_\rho)^2/(2s_\rho^2)\}}{\exp\{l(\rho)-(\rho^\dagger-m_\rho)^2/(2s_\rho^2)\}}\right].
$$

Generation of σ_η^2. Propose a candidate $\sigma_\eta^{2\dagger} \sim \mathrm{IG}(\alpha_1/2, \beta_1/2)$, where

$$\alpha_1 = \alpha_0 + mn,\ \beta_1 = \beta_0 + \sum_{t=1}^{n} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' Q^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t), \tag{2.33}$$

and accept it with probability $\min\left[1, \exp\left(g(\sigma_\eta^{2\dagger}) - g(\sigma_\eta^2)\right)\right]$, where

$$g(\sigma_\eta^2) = -\frac{1}{2}\sum_{t=1}^{n} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{\beta_1}{2\sigma_\eta^2}. \tag{2.34}$$

Generation of λ. Propose a candidate $\lambda^\dagger \sim \mathrm{IG}(\alpha_1^*, \beta_1^*)$, where

$$\alpha_1^* = \alpha_0^* + \frac{3}{2}n,\ \beta_1^* = \beta_0^* + \sum_{t=1}^{n} \frac{(y_t - Z\boldsymbol{\xi}_t - av_t)^2}{2b^2 v_t} + \sum_{t=1}^{n} v_t, \tag{2.35}$$

and accept it with probability $\min\left[1, \exp\left(g(\lambda^\dagger) - g(\lambda)\right)\right]$, where

$$g(\lambda) = -\frac{1}{2}\sum_{t=1}^{n} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{\beta_1^*}{\lambda}. \tag{2.36}$$

Generation of v_t. Propose a candidate $v_t^\dagger \sim \mathrm{GIG}(1/2, \delta_t, \gamma)$, where

$$\delta_t^2 = \frac{(y_t - Z\boldsymbol{\xi}_t)^2}{b^2\lambda},\ \gamma^2 = \frac{2}{\lambda} + \frac{a^2}{b^2\lambda}, \tag{2.37}$$

and accept it with probability $\min\left[1, \exp\left(g(v_t^\dagger) - g(v_t)\right)\right]$, where

$$g(v_t) = -\frac{1}{2} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{1}{2}(v_t^{-1}\delta_t^2 + v_t\gamma^2). \tag{2.38}$$

2.4 Illustrative examples using simulated data

This section illustrates our proposed time-varying quantile models using simulated data. We consider the case $m = 2$ and assume that the variance parameter for the initial state κ is 100. The sensitivity analysis for the selection of κ is also investigated.

First, we show the high efficiency of our multi-move sampling method in comparison with the single-move sampling method using the TQSS model given by (2.13)-(2.14). Using the following parameters based on our empirical studies in Section 2.5,

$$\sigma_\eta^2 = 4.0 \times 10^{-3}, \quad \lambda = 3.5 \times 10^{-2}, \quad \text{for } \tau = 0.1,$$
$$\sigma_\eta^2 = 1.0 \times 10^{-4}, \quad \lambda = 4.0 \times 10^{-2}, \quad \text{for } \tau = 0.9,$$

we generate three hundred observations ($n = 300$) for each τ. For prior distributions, we assume

$$\sigma_\eta^2 \sim \text{IG}(0.1, 0.00005), \quad \lambda \sim \text{IG}(0.1, 0.1).$$

Using the single-move sampler (Section 2.3.2.1), we generate 600,000 (450,000) MCMC samples after discarding the first 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$). Also, the multi-move sampler (Section 2.3.2.2) is used to generate 30,000 (15,000) MCMC samples after discarding the first 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$).

Table 2.1: TQSS model ($\tau = 0.1$).
Posterior means, standard deviations, 95% credible intervals and inefficiency factors (IF).

		True	Mean	Stdev	95% interval	IF
single-move	$\sigma_\eta^2 \times 10^3$	4	3.398	0.743	[2.206, 5.091]	133
	$\lambda \times 10^2$	3.5	3.296	0.215	[2.900, 3.741]	46
multi-move	$\sigma_\eta^2 \times 10^3$	4	3.374	0.766	[2.175, 5.169]	31
	$\lambda \times 10^2$	3.5	3.302	0.214	[2.908, 3.748]	2

Table 2.2: TQSS model ($\tau = 0.9$).
Posterior means, standard deviations, 95% credible intervals and inefficiency factors (IF).

		True	Mean	Stdev	95% interval	IF
single-move	$\sigma_\eta^2 \times 10^4$	1	0.976	0.343	[0.499, 1.817]	530
	$\lambda \times 10^2$	4	4.018	0.241	[3.572, 4.517]	51
multi-move	$\sigma_\eta^2 \times 10^4$	1	0.939	0.317	[0.495, 1.700]	44
	$\lambda \times 10^2$	4	4.025	0.242	[3.578, 4.537]	2

Tables 2.1 and 2.2 report the true values, posterior means, posterior standard deviations, 95% credible intervals and estimates of inefficiency factors (IF). The inefficiency factor is defined as $1+2\sum_{g=1}^{\infty}\rho(g)$, where $\rho(g)$ is the sample autocorrelation at lag g. This is interpreted as the ratio of the numerical variance of the posterior mean from the chain to the variance of the posterior mean from hypothetical uncorrelated draws. The smaller the inefficiency factor becomes, the closer the MCMC sampling is to the uncorrelated sampling.

The posterior means are all close to the true values, which suggests that our proposed algorithms work well.

The inefficiency factors for the single-move sampler are much larger than those for the

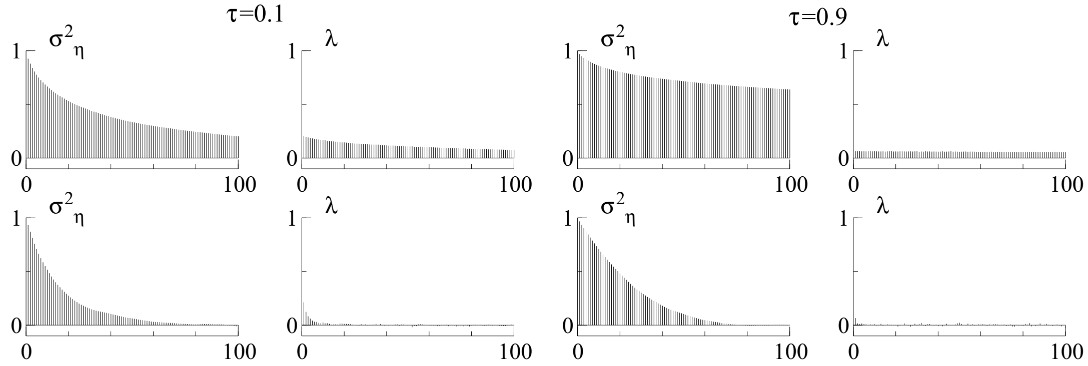

Figure 2.1: TQSS model. Sample autocorrelation functions of MCMC samples for the single-move sampler (top) and the multi-move sampler (bottom).

multi-move sampler, which suggests that our multi-move sampling method is highly efficient compared with the single-move sampling method. Figure 2.1 shows the sample autocorrelation functions where autocorrelations decay slowly for the single-move sampler and vanish quickly for the multi-move sampler, which also implies the efficiency of our multi-move sampler.

Next, we illustrate our estimation method for the TQSS model with correlations and investigate the sensitivity analysis with respect to the selection of the initial variance parameter κ.

Setting the parameters,

$$\rho = -0.35, \quad \sigma_\eta^2 = 4.0 \times 10^{-3}, \quad \lambda = 3.0 \times 10^{-2}, \quad \text{for } \tau = 0.1,$$
$$\rho = -0.33, \quad \sigma_\eta^2 = 1.5 \times 10^{-4}, \quad \lambda = 4.0 \times 10^{-2}, \quad \text{for } \tau = 0.9,$$

we generate three hundred observations ($n = 300$) for each $\tau = 0.1, 0.9$. For prior distributions, we assume a uniform distribution for ρ, $\rho \sim \mathrm{U}(-1/2, 1/2)$, and the same distributions for σ_η^2 and λ. We generate 1,200,000 (600,000) MCMC samples after discarding 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$) using the multi-move sampler (Section 2.3.3) with $\kappa = 10, 100$ and 1000.

The estimation results are summarized in Table 2.3. For all κ, the posterior means of the parameters are close to the true values, which suggests that our sampler works well. The inefficiency factors are larger overall than those for the model without correlations due to the inefficiency of sampling the new parameter ρ. The sampling method for ρ needs to be

Table 2.3: TQSS model with correlations ($\tau = 0.1, 0.9$). Posterior means, standard deviations, 95% credible intervals and inefficiency factors for $\kappa = 10, 100$ and 1000.

		True	Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.35	-0.128	0.266	[-0.488, 0.431]	475
($\kappa = 100$)	$\sigma_\eta^2 \times 10^3$	4	3.362	0.761	[2.167, 5.125]	73
	$\lambda \times 10^2$	3	3.006	0.204	[2.629, 3.429]	42
	ρ	-0.35	-0.127	0.264	[-0.484, 0.434]	487
($\kappa = 10$)	$\sigma_\eta^2 \times 10^3$	4	3.364	0.756	[2.166, 5.106]	66
	$\lambda \times 10^2$	3	3.006	0.202	[2.633, 3.424]	40
	ρ	-0.35	-0.111	0.271	[-0.484, 0.443]	496
($\kappa = 1000$)	$\sigma_\eta^2 \times 10^3$	4	3.354	0.758	[2.160, 5.103]	63
	$\lambda \times 10^2$	3	3.011	0.203	[2.636, 3.430]	41
$\tau = 0.9$	ρ	-0.33	-0.047	0.283	[-0.484, 0.459]	461
($\kappa = 100$)	$\sigma_\eta^2 \times 10^4$	1.5	1.875	0.617	[0.995, 3.363]	109
	$\lambda \times 10^2$	4	3.768	0.228	[3.346, 4.241]	17
	ρ	-0.33	-0.038	0.279	[-0.473, 0.456]	464
($\kappa = 10$)	$\sigma_\eta^2 \times 10^4$	1.5	1.892	0.626	[1.009, 3.433]	105
	$\lambda \times 10^2$	4	3.768	0.228	[3.347, 4.241]	13
	ρ	-0.33	-0.052	0.290	[-0.485, 0.470]	509
($\kappa = 1000$)	$\sigma_\eta^2 \times 10^4$	1.5	1.877	0.621	[0.991, 3.378]	122
	$\lambda \times 10^2$	4	3.766	0.228	[3.345, 4.237]	14

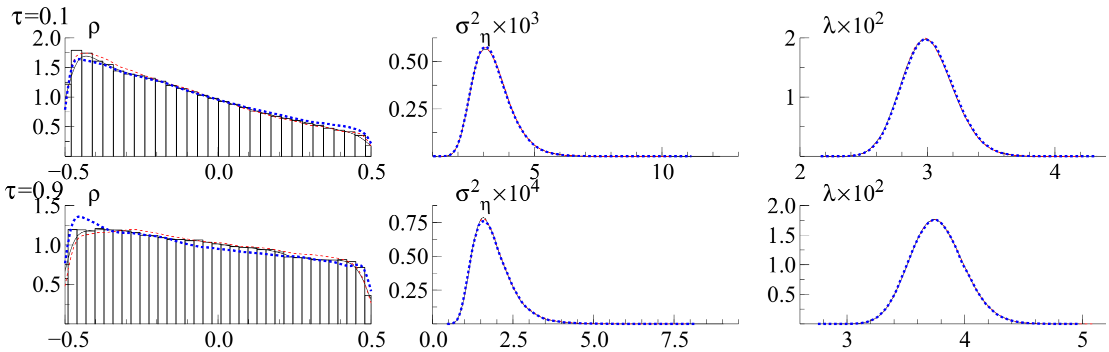

Figure 2.2: Time-varying quantile model with correlations ($\tau = 0.1, 0.9$). Estimated posterior densities for $\kappa = 10$ (thin dotted line), $\kappa = 100$ (solid line with histogram), $\kappa = 1000$ (thick dotted line).

improved but will be left for future work.

Taking into account the posterior standard deviations, the estimation results are robust with respect to the selection of κ. Figure 2.2 shows the estimated posterior densities of the parameters for $\kappa = 10, 100$ and 1000. They also show the robustness of the estimation results with respect to the selection of κ.

2.5 Empirical study

2.5.1 Data

Economic agents are known to consider not only the mean of the response distribution but also many aspects of the distribution in their decision making processes. In macroeconomics, for example, the upper tail of the distribution of the inflation rate is usually one of the greatest concerns to the central bank because monetary policy measures are deployed that focus on (violent) inflation in an attempt to keep the price movement stable. On the other hand, the lower tail of the distribution of the inflation rate would attract increasing attention from those advanced countries who face the risk of deflation, which may cause a serious recession.

Thus, this section applies our proposed model to the upper and the lower tails of the distribution of the inflation rate of Japan, using the rate of change for the domestic Corporate Goods Price Index (CGPI) excluding the consumption tax for all commodities of Japan (reported by Bank of Japan). The rate of change is calculated as $y_t = 100 \times (\log p_t - \log p_{t-1})$, where p_t is the CGPI at time t. We consider the following two sample periods:

$$\begin{aligned}\text{Period (I)}: &\ \text{February, 1985 – June, 2008 (281 months) and}\\ \text{Period (II)}: &\ \text{February, 1985 – January, 2010 (300 months).}\end{aligned}$$

Period (I) is set before Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008) to eliminate the event's significant negative impact on the economy. We consider two time-varying quantiles with $\tau = 0.1, 0.9$ as mentioned above. Table 2.4 shows the summary statistics for the inflation rate based on the CGPI of Japan (y_t). The sample mean of the inflation rate for Period (II) is slightly lower than that for Period (I), while the maximum and the absolute value of the minimum for Period (II) are much larger than those for Period (I). This suggests the existence of the greater uncertainty in the inflation rate during Period (II).

2.5.2 Estimation results

Using the same prior distributions for the parameters as in Section 2.4, $m = 2$ and $\kappa = 100$, we implement the MCMC algorithm to conduct a Bayesian inference on parameters of interest.

Table 2.4: Summary statistics for the inflation rate based on CGPI of Japan (y_t).

	Nob	Mean	Stdev	Max	Min
Period (I) [1985.2-2008.6]	281	-0.038	0.279	1.201	-0.975
Period (II) [1985.2-2010.1]	300	-0.059	0.361	2.160	-1.989

We generate 60,000 (30,000, 30,000) MCMC samples from the posterior distributions of the parameters in the model without correlations and generate 60,000 (240,000, 240,000) MCMC samples from the posterior distributions of the parameters in the model with correlations, after discarding the first 1,000 (1,000, 1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.5$, $\tau = 0.9$).

Table 2.5: Period (I). TQSS model.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	$\sigma_\eta^2 \times 10^4$	3.469	3.355	[0.523, 12.66]	210
	$\lambda \times 10^2$	3.402	0.261	[2.902, 3.931]	77
$\tau = 0.5$	$\sigma_\eta^2 \times 10^5$	5.338	3.143	[1.703, 13.77]	119
	$\lambda \times 10^2$	7.626	0.487	[6.723, 8.638]	6
$\tau = 0.9$	$\sigma_\eta^2 \times 10^4$	1.424	0.695	[0.570, 3.214]	87
	$\lambda \times 10^2$	3.159	0.204	[2.783, 3.579]	5

Table 2.6: Period (I). TQSS model with correlations.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.253	0.211	[-0.492, 0.293]	228
	$\sigma_\eta^2 \times 10^4$	4.299	3.268	[0.629, 13.19]	259
	$\lambda \times 10^2$	3.316	0.263	[2.825, 3.855]	115
$\tau = 0.5$	ρ	0.0192	0.294	[-0.473, 0.476]	391
	$\sigma_\eta^2 \times 10^5$	4.970	2.876	[1.444, 12.26]	223
	$\lambda \times 10^2$	7.640	0.485	[6.745, 8.642]	18
$\tau = 0.9$	ρ	-0.114	0.278	[-0.485, 0.455]	389
	$\sigma_\eta^2 \times 10^4$	1.473	0.759	[0.562, 3.435]	196
	$\lambda \times 10^2$	3.150	0.205	[2.770, 3.576]	26

The estimation results for Period (I) are given in Tables 2.5 and 2.6 for the two models ($\tau = 0.1, 0.5, 0.9$). The parameter estimates are quite similar for both models taking into account the posterior standard deviations. The estimate of the variance parameter of the

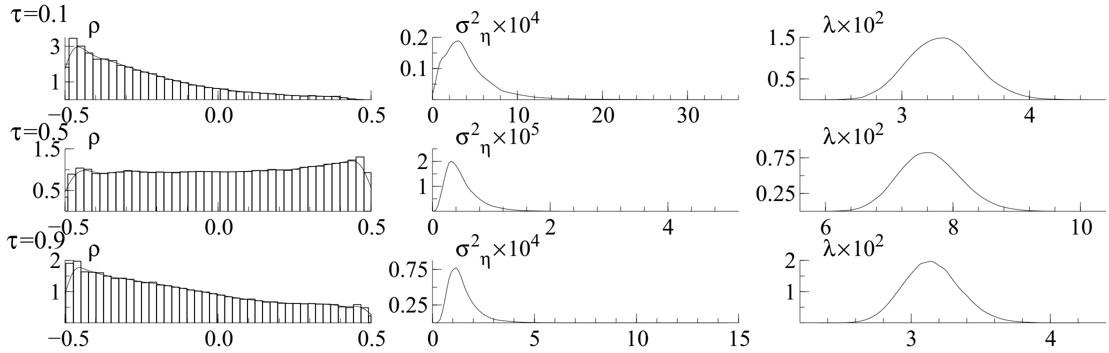

Figure 2.3: Period (I). TQSS model with correlations.

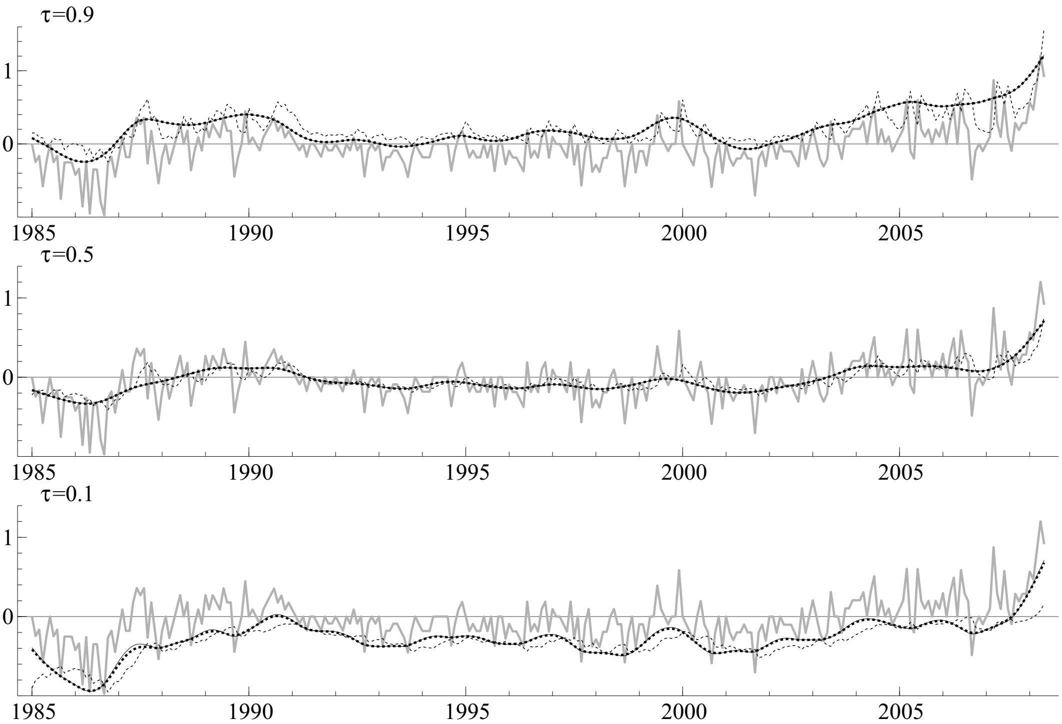

Figure 2.4: Period (I). Time series plot of the inflation rate (gray line), posterior means of time-varying quantiles using TQSS models without correlations (thick dotted line), with correlations (solid line) and CAViaR model (thin dotted line).

state equation σ_η^2 for $\tau = 0.1$ is much larger than that for $\tau = 0.9$, indicating that the lower tail quantile is more uncertain than the upper tail quantile. The posterior means of the correlations are negative for $\tau = 0.1, 0.9$, but they are not credible since their 95% credible intervals include zeros. The estimated posterior densities are shown in Figure 2.3 in which ρ seems to have a large posterior probability on negative values especially for $\tau = 0.1$[5].

Furthermore, Figure 2.4 shows the time series plot of the posterior means of ξ_t and y_t for $\tau = 0.1, 0.5, 0.9$. The estimated quantile posterior means vary smoothly and capture the changes in the level and the magnitude of the inflation rate over the sample period. The estimates for the two TQSS models with and without correlations are found to be quite similar (solid and thick dotted lines, respectively).

Table 2.7: Period (II). TQSS model.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	$\sigma_\eta^2 \times 10^3$	5.418	2.255	[1.570, 10.48]	114
	$\lambda \times 10^2$	3.274	0.294	[2.758, 3.913]	67
$\tau = 0.5$	$\sigma_\eta^2 \times 10^5$	4.226	3.480	[0.996, 14.03]	153
	$\lambda \times 10^2$	9.717	0.600	[8.607, 10.94]	12
$\tau = 0.9$	$\sigma_\eta^2 \times 10^4$	1.425	0.813	[0.525, 3.288]	103
	$\lambda \times 10^2$	4.084	0.254	[3.619, 4.612]	7

Table 2.8: Period (II). TQSS model with correlations.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.351	0.143	[-0.496, 0.031]	219
	$\sigma_\eta^2 \times 10^3$	5.679	2.114	[2.157, 10.21]	192
	$\lambda \times 10^2$	3.159	0.289	[2.631, 3.771]	133
$\tau = 0.5$	ρ	0.0410	0.293	[-0.474, 0.485]	389
	$\sigma_\eta^2 \times 10^5$	4.397	4.312	[1.005, 16.21]	374
	$\lambda \times 10^2$	9.722	0.602	[8.605, 10.96]	39
$\tau = 0.9$	ρ	-0.092	0.276	[-0.486, 0.448]	391
	$\sigma_\eta^2 \times 10^4$	1.436	0.683	[0.528, 3.166]	215
	$\lambda \times 10^2$	4.077	0.255	[3.606, 4.604]	25

Tables 2.7 and 2.8 show the estimation results for Period (II), and the estimated posterior densities are shown in Figure 2.5. The results between the two models are quite similar, and the posterior means of σ_η^2 for $\tau = 0.1$ are found to be much larger than those for $\tau = 0.9$ as for Period (I). We note that the estimate of σ_η^2 for $\tau = 0.1$ for Period (II) is ten times as large as the corresponding estimate for Period (I). This is because the inflation rate fluctuated

[5]Period (I). The posterior mean of $\mathrm{corr}(y_t, \xi_{t+1})$ given $\boldsymbol{\xi}_t$ is -0.105 (0.0170, -0.0473) for $\tau = 0.1$ (0.5, 0.9).

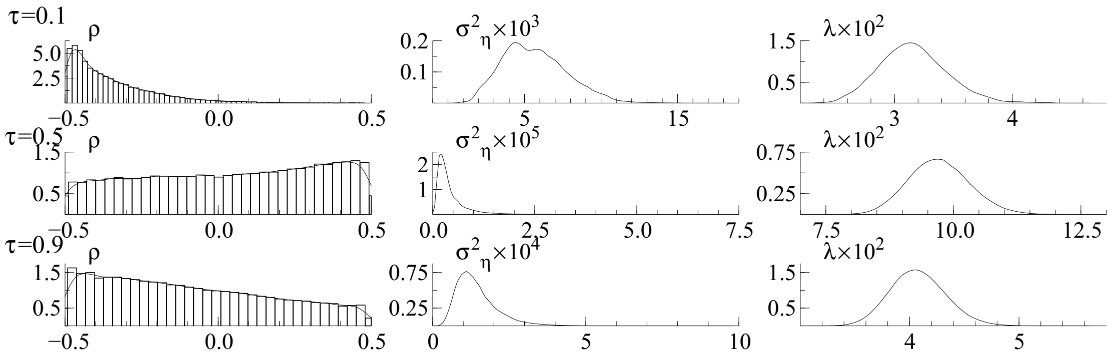

Figure 2.5: Period (II). TQSS model with correlations.

greatly around September 2008 for Period (II). It suggests that the deflationary impact in Japan was more serious than the inflation during this period[6].

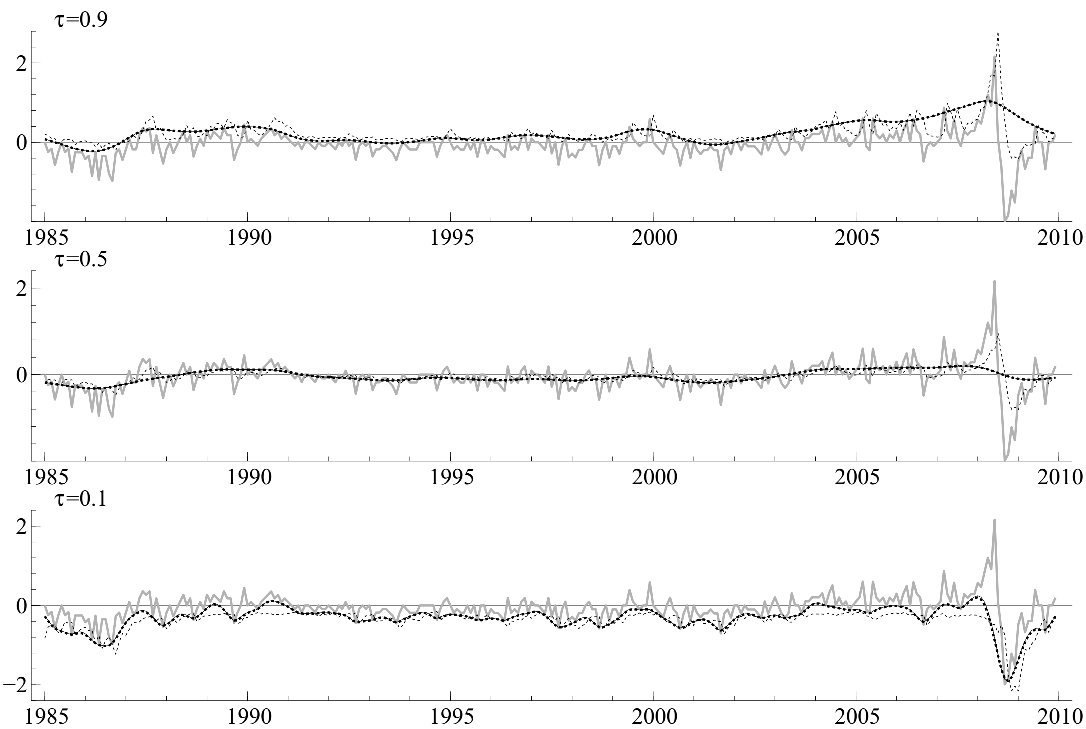

Figure 2.6: Period (II). Time series plot of the inflation rate (gray line), posterior means of time-varying quantiles using TQSS models without correlations (thick dotted line), with correlations (solid line) and CAViaR model (thin dotted line).

[6]Period (II). The posterior mean of corr(y_t, ξ_{t+1}) given $\boldsymbol{\xi}_t$ is -0.146 (0.0363, -0.0382) for $\tau = 0.1$ (0.5, 0.9).

Figure 2.6 shows the time series plot of the posterior means of ξ_t and y_t for $\tau = 0.1, 0.5, 0.9$. The estimated quantile posterior means are similar for the two models, analogous to Figure 2.4 for Period (I). However, in contrast to the plot for Period (I), there is a sharp downward spike around the end of Period (II) for $\tau = 0.1$, corresponding to the large estimate of the state variance parameter σ_η^2.

As a benchmark model for the time-varying quantiles, we also estimate the CAViaR (Conditional Autoregressive Value at Risk) model by Engle and Manganelli (2004). The asymmetric slope type model is chosen among models in the CAViaR class because it is able to capture the effect of asymmetry in time-varying quantiles (see Appendix 2.A.2 for more details). The posterior means of ξ_t based on the CAViaR (asymmetric slope) model are also shown in Figures 2.4 and 2.6 (thin dotted lines). The plots seem to be rough and sensitive to the change in the inflation rate compared with those estimates from our proposed models based on the smoothing spline.

2.5.3 Model Comparison

This subsection conducts a model comparison of our proposed models and the CAViaR model based on backtesting (Kupiec (1995)) and the DIC (Spiegelhalter, Best, Carlin, and van der Linde (2002)).

Backtesting. Whether the future observation exceeds the τ-quantile ($(1-\tau)$-VaR) is an important issue in terms of decision making for economic policy and for financial risk management. To evaluate such forecasting performance, we conduct the likelihood test by Kupiec (1995), which is often used for backtesting for VaR in finance. First, we fix $n_0(< n)$ and set $s = 0$. Then, we (i) estimate the parameters and the latent quantiles using observations $\{y_t\}_{t=s+1}^{s+n_0}$ and (ii) compute the posterior mean of the predictive distribution of the one-step-ahead quantile, ξ_{s+n_0+1}. We repeat (i) and (ii) for $s = 0, 1, \ldots, n - n_0$ and compute the number of times N when the posterior mean $\hat{\xi}_{s+n_0+1}$ exceeds the observed data y_{s+n_0+1} for $s = 0, 1, \ldots, n - n_0$. If the null hypothesis that $\Pr(y_t < \xi_t) = \tau$ is true (and the probability is independent for each t),

$$2\left\{ \log\left(\left(\frac{N}{n-n_0}\right)^N \left(1 - \frac{N}{n-n_0}\right)^{n-n_0-N} \right) - \log(\tau^N (1-\tau)^{n-n_0-N}) \right\} \tag{2.39}$$

is asymptotically distributed as $\chi^2(1)$[7] .

[7] Since the time-varying quantiles are not independent of one another, the asymptotic result may not hold.

Table 2.9: p-values of the likelihood test.

	Model	Period (I)	Period (II)
$\tau = 0.1$	TQSS	0.021	0.000
	TQSSC	0.021	0.000
	CAViaR	0.417	0.743
τ=0.9	TQSS	0.045	0.032
	TQSSC	0.045	0.032
	CAViaR	0.009	0.001

TQSSC: TQSS model with correlations.

Table 2.9 shows the p-values of the one-sided likelihood ratio tests using $n_0 = 200$ observations for both Periods (I) and (II) ($n = 281$ for Period (I) and $n = 300$ for Period (II)). The number of the MCMC iterations and the prior distributions for each estimation are the same as those of the previous subsection. For Period (I), the null hypothesis is rejected for the CAViaR model ($\tau = 0.9$) since its p-value is smaller than 0.01. For Period (II), we reject the null hypotheses for the TQSS models ($\tau = 0.1$) and for the CAViaR model ($\tau = 0.9$). Thus, for the upper tail quantile ($\tau = 0.9$), our proposed models show good forecasting performances with respect to VaR, while the CAViaR model fails for both periods. On the other hand, for the lower tail quantile ($\tau = 0.1$), the CAViaR model performs well for both periods, while the performance of our proposed models depends on the sample period.

Model selection based on DIC. The Deviance Information Criterion (DIC) is used as a Bayesian measure of fit or adequacy and is defined as

$$\mathrm{DIC} = \mathrm{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})] + p_D, \tag{2.40}$$

where $D(\boldsymbol{\theta}) = -2 \log f(Y_n|\boldsymbol{\theta})$, $p_D = \mathrm{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})] - D(\mathrm{E}_{\boldsymbol{\theta}|Y_n}[\boldsymbol{\theta}])$ represents model complexity as a penalty, $Y_n = \{y_t\}_{t=1}^n$ and $\boldsymbol{\theta}$ denotes the parameters. We estimate $\mathrm{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})]$ using the sample analogue $\overline{D(\boldsymbol{\theta}_{(d)})} = \frac{1}{d^*}\sum_{d=1}^{d^*} D(\boldsymbol{\theta}_{(d)})$, where $\boldsymbol{\theta}_{(d)}$s are resampled from the posterior distribution. We set d^* equal to 600 ($\tau = 0.1$) and 300 ($\tau = 0.9$) for the TQSS model, 600 ($\tau = 0.1$) and 2,400 ($\tau = 0.9$) for the TQSS model with correlations and 240 for the CAViaR model. Because we need to compute $D(\boldsymbol{\theta})$ numerically, we use the particle filter (e.g., Doucet, de Freitas, and Gordon (2001)), where we set the number of particles $M = 10,000$

The implications of the likelihood test should be used with caution.

(see Appendix 2.A.3 for details). The numerical standard error of the estimate is obtained by repeating the particle filter forty times.

Table 2.10: DIC (standard errors in parentheses).

	Model	Period (I)		Period (II)	
$\tau = 0.1$	TQSS	141.16	(0.19)	329.86	(2.34)
	TQSSC	138.31	(0.25)	249.42	(0.68)
	CAViaR	141.69	(0.14)	283.44	(0.10)
$\tau = 0.9$	TQSS	92.93	(0.24)	259.29	(1.52)
	TQSSC	91.51	(0.19)	240.72	(2.26)
	CAViaR	102.84	(0.09)	190.11	(0.10)

TQSSC: TQSS model with correlations

Table 2.10 shows the sample means of forty DICs for each model with the standard errors in parentheses. The DICs of the TQSS model with correlations are the smallest and outperform the other competing models for $\tau = 0.1, 0.9$ for Period (I) and for $\tau = 0.1$ for Period (II). However, for $\tau = 0.9$ for Period (II), the CAViaR model outperforms the two TQSS models. This is partly because the quantile estimates of the CAViaR model follow the fluctuations of the inflation rate more quickly than those of the TQSS models as seen in Figure 2.6.

In summary, for Period (I), as illustrated in Figure 2.4, the trajectories of the estimated quantiles of the two TQSS models are smooth to obtain good forecasting performances for the VaR. In addition, the TQSS model with correlation attains the smallest DIC. However, with respect to Period (II), there is no model for which the null hypothesis of the backtesting is accepted using both $\tau = 0.1$ and 0.9. The DIC suggests that a different model should be used depending on τ[8].

2.6 Conclusion

This article proposed a novel smoothing spline model for time-varying quantiles. Taking a Bayesian approach, the efficient MCMC algorithm is described using a normal variance-mean mixture representation of the measurement error term where we exploit a simulation smoother for the linear Gaussian state space model. Its high efficiency is illustrated using simulated

[8]The TQSS model with correlation and with $m = 1$ is not considered here since it does not pass the likelihood ratio test at all. In terms of DIC, the model may perform well depending on the period and τ.

data in comparison with the single-move sampler. The model is extended to incorporate a correlation between the dependent variable and its one-step-ahead quantile. Furthermore, in comparison with the CAViaR model, our method is shown to perform well regarding both one-ahead predictions and goodness-of-fit in the analysis of Japanese inflation rate.

Appendix

2.A.1 Generation of $\{\boldsymbol{\xi}_t\}$ using a single-move sampler

In Section 2.3.2.1, the joint posterior density of $\boldsymbol{\xi}_t, I_t|\{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_{-t}\}, \sigma_\eta^2, \lambda$ is given by

$$f(\boldsymbol{\xi}_t, I_t|\{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \propto \exp\left(-\frac{1}{2}\{\boldsymbol{\xi}_t' S^{-1}\boldsymbol{\xi}_t - 2\boldsymbol{\xi}_t'\tilde{\boldsymbol{m}}_{tI_t}\}\right) \times \exp\left(-\frac{(\tau - I_t)y_t}{\lambda}\right), \quad t = 2, \ldots, n-1, \tag{2.41}$$

where

$$S = (T'(\sigma_\eta^2 Q)^{-1}T + (\sigma_\eta^2 Q)^{-1})^{-1}, \tag{2.42}$$

$$\tilde{\boldsymbol{m}}_{tI_t} = \left(T'(\sigma_\eta^2 Q)^{-1}\boldsymbol{\xi}_{t+1} + (\sigma_\eta^2 Q)^{-1}T\boldsymbol{\xi}_{t-1} + \frac{(\tau - I_t)}{\lambda}Z'\right). \tag{2.43}$$

Let

$$\tilde{\boldsymbol{m}}_{tI_t} = \begin{pmatrix} m_{tI_t} \\ \tilde{\boldsymbol{m}}_t \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{pmatrix},$$

where m_{tI_t}, S^{11} are scalars, $\tilde{\boldsymbol{m}}_t, (S^{12})', S^{21}$ are $(m-1)\times 1$ vectors and S^{22} is an $(m-1)\times(m-1)$ matrix. We generate $(\boldsymbol{\xi}_t, I_t)$ as follows:

a. $I_t|\tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$.

b. $\xi_t|I_t, \tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$.

c. $\tilde{\boldsymbol{\xi}}_t|\xi_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$.

Noting that

$$\begin{aligned} f(\xi_t, I_t|\tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \propto &\exp\left(-\frac{S^{11}}{2}\{\xi_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{tI_t})\}^2\right) \\ &\times \exp\left(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{tI_t})^2\right) \times \exp\left(-\frac{(\tau - I_t)y_t}{\lambda}\right), \end{aligned} \tag{2.44}$$

define

$$g(0) := \Phi\left(\frac{y_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})}{(S^{11})^{-\frac{1}{2}}}\right) \times \exp\left(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})^2 - \frac{\tau y_t}{\lambda}\right), \tag{2.45}$$

$$g(1) := \left(1-\Phi\Big(\frac{y_t-(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t1})}{(S^{11})^{-\frac{1}{2}}}\Big)\right)\times\exp\left(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t1})^2-\frac{(\tau-1)y_t}{\lambda}\right), \tag{2.46}$$

where Φ is a cumulative normal distribution function.

Generation of I_t. Generate $I_t|\cdot \sim \text{Bernoulli}(p)$ with $p = g(1)/(g(0)+g(1))$[9] .

Generation of ξ_t. Given I_t, generate

$$\xi_t|\cdot \sim \begin{cases} \text{TN}_{(-\infty,y_t]}((S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t0}),(S^{11})^{-1}) & \text{if } I_t=0, \\ \text{TN}_{(y_t,\infty)}((S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t1}),(S^{11})^{-1}) & \text{if } I_t=1. \end{cases} \tag{2.49}$$

Generation of $\tilde{\boldsymbol{\xi}}_t$. Generate

$$\tilde{\boldsymbol{\xi}}_t|\xi_t,\{\boldsymbol{\xi}_{-t}\},\{y_t\}_{t=1}^n,\sigma_\eta^2,\lambda \sim \text{N}((S^{22})^{-1}(-S^{21}\xi_t+\tilde{\boldsymbol{m}}_t),(S^{22})^{-1}). \tag{2.50}$$

2.A.2 CAViaR model

As a benchmark for the model comparison, we consider the following asymmetric slope CAViaR model discussed in Engle and Manganelli (2004):

$$\xi_{t+1} = \beta_1+\beta_2\xi_t+\beta_3 y_t^+ + \beta_4 y_t^-, \tag{2.51}$$

where $y_t^+ = \max(y_t,0)$ and $y_t^- = -\min(y_t,0)$ to model asymmetry of the dynamics of the quantile. We assume the following prior distributions:

$$\lambda \sim \text{IG}(0.1,0.1),\ \ \boldsymbol{\beta} \sim \text{TN}_{(0\le\beta_2<1)}(\mathbf{0}_4,100E_4),\ \ \xi_1 \sim \text{N}(0,100). \tag{2.52}$$

For $\tau = 0.1$ and for $\tau = 0.9$, respectively, we generate 240,000 MCMC draws after discarding 10,000 draws as the burn-in period.

[9]When $\zeta_0 = [y_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})]/(S^{11})^{-\frac{1}{2}} \ll 0$, $\Phi(\zeta_0)$ may be almost 0 and $\exp\left(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t0})^2-\frac{\tau y_t}{\lambda}\right)$ becomes huge. In this case, the computed value of g is inaccurate and we need to approximate the values of g using a partial fractional expansion (Stuart and Ord (1994)),

$$\Phi(\zeta_0) = \frac{1}{\sqrt{2\pi}}(-\zeta_0^{-1}+\zeta_0^{-3}-3\zeta_0^{-5}+15\zeta_0^{-7}-\ldots)\exp\Big(-\frac{\zeta_0^2}{2}\Big), \tag{2.47}$$

and obtain

$$g_t(0) \approx \frac{1}{\sqrt{2\pi}}\exp\Big(\log(-\zeta_0^{-1}+\zeta_0^{-3}-3\zeta_0^{-5}+15\zeta_0^{-7})-\frac{\zeta_0^2}{2}+\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t0})^2-\frac{\tau}{\lambda}y_t\Big). \tag{2.48}$$

The approximation error is

$$\text{error} < \frac{105}{\sqrt{2\pi}}\zeta_0^{-9}\exp\Big(-\frac{\zeta_0^2}{2}\Big) < 6.34\times10^{-14}$$

when $\zeta_0 \le -6$.

2.A.3 Particle filter

Let $f(\boldsymbol{\xi}_t|Y_t,\boldsymbol{\theta})$ denote the density function of $\boldsymbol{\xi}_t$ given $(Y_t,\boldsymbol{\theta})$ where $Y_t=\{y_1,\ldots,y_t\}$, and let $\hat{f}(\boldsymbol{\xi}_t|Y_t,\boldsymbol{\theta})$ denote the discrete approximation to $f(\boldsymbol{\xi}_t|Y_t,\boldsymbol{\theta})$.

We draw M samples from the conditional joint distribution of $(\boldsymbol{\xi}_{t+1},\boldsymbol{\xi}_t,v_t)$ given $(Y_{t+1},\boldsymbol{\theta})$ with the density

$$f(\boldsymbol{\xi}_{t+1},\boldsymbol{\xi}_t,v_t|Y_{t+1},\boldsymbol{\theta}) \propto f(y_{t+1}|\boldsymbol{\xi}_{t+1},\boldsymbol{\theta})f(\boldsymbol{\xi}_{t+1}|y_t,\boldsymbol{\xi}_t,v_t,\boldsymbol{\theta})f(v_t|\boldsymbol{\theta})f(\boldsymbol{\xi}_t|Y_t,\boldsymbol{\theta}), \tag{2.53}$$

where

$$f(y_t|\boldsymbol{\xi}_t,\boldsymbol{\theta}) = \int f(y_t|\boldsymbol{\xi}_t,v_t,\boldsymbol{\theta})f(v_t|\boldsymbol{\theta})dv_t,\ \ f(\boldsymbol{\xi}_{t+1}|y_t,\boldsymbol{\xi}_t,v_t,\boldsymbol{\theta}) = f(y_t,\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t,v_t,\boldsymbol{\theta})/f(y_t|\boldsymbol{\xi}_t,v_t,\boldsymbol{\theta}). \tag{2.54}$$

We implement the particle filter:

1. (a) Generate
$$\boldsymbol{\xi}_1^{(i)} \sim \mathrm{N}(\boldsymbol{m}_1,V_1),\ i=1,\ldots,M,$$
where $\boldsymbol{m}_1,V_1$ are some constant vector and some constant positive-definite matrix (we adopt the posterior mean and the covariance matrix of $\boldsymbol{\xi}_1$), respectively.

   (b) We calculate
$$w_1^{(i)} := \frac{f(y_1|\boldsymbol{\xi}_1^{(i)},\boldsymbol{\theta})f(\boldsymbol{\xi}_1^{(i)})}{g(\boldsymbol{\xi}_1^{(i)})},$$
where $g(\cdot)$ is a normal density with mean $\boldsymbol{m}_1$ and covariance matrix V_1 and set $\bar{w}_1=\frac{1}{M}\sum_{i=1}^M w^{(i)},\ \ \pi_1^{(i)} := \hat{f}(\boldsymbol{\xi}_1^{(i)}|y_1,\boldsymbol{\theta}^*) = w^{(i)}/\sum_{j=1}^M w^{(j)}$.

2. For $t=1,\ldots,n-1$, we generate $(\boldsymbol{\xi}_{t+1}^{(i)},\boldsymbol{\xi}_t^{(i)}), i=1,\ldots,M$:

   (a) Generate $\boldsymbol{\xi}_t^{(i)} \sim \hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t,\boldsymbol{\theta})$ and $v_t^{(i)} \sim f(v_t^{(i)}|\boldsymbol{\theta})$.

   (b) Generate $\boldsymbol{\xi}_{t+1}^{(i)} \sim f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t,\boldsymbol{\xi}_t^{(i)},v_t^{(i)},\boldsymbol{\theta})$.

   (c) Compute
$$\begin{aligned} w_{t+1}^{(i)} &:= \frac{f(y_{t+1}|\boldsymbol{\xi}_{t+1}^{(i)},\boldsymbol{\theta})f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t,\boldsymbol{\xi}_t^{(i)},v_t^{(i)},\boldsymbol{\theta})f(v_t^{(i)}|\boldsymbol{\theta})\hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t,\boldsymbol{\theta})}{f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t,\boldsymbol{\xi}_t^{(i)},v_t^{(i)},\boldsymbol{\theta})f(v_t^{(i)}|\boldsymbol{\theta})\hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t,\boldsymbol{\theta})} \\ &= f(y_{t+1}|\boldsymbol{\xi}_{t+1}^{(i)},\boldsymbol{\theta}), \end{aligned}$$
and set $\bar{w}_{t+1}=\frac{1}{M}\sum_{i=1}^M w^{(i)} \to f(y_{t+1}|Y_t,\boldsymbol{\theta}^*),\ \hat{f}(\boldsymbol{\xi}_{t+1}^{(i)}|Y_{t+1},\boldsymbol{\theta}^*) = w^{(i)}/\sum_{j=1}^M w^{(j)} := \pi_{t+1}^{(i)}$.

3. We obtain

$$\sum_{t=0}^{n-1} \log \bar{w}_{t+1} \to \sum_{t=0}^{n-1} \log f(y_{t+1}|Y_t, \boldsymbol{\theta}^*) = \log f(Y_n|\boldsymbol{\theta}^*) \text{ as } M \to \infty.$$

Chapter 3

Dynamic equicorrelation stochastic volatility

3.1 Introduction

Over the last several decades, various multivariate volatility models have been proposed to model asset returns with time-varying variances. Two popular examples are generalized autoregressive conditional heteroskedasticity (GARCH) models (Bauwens, Laurent, and Rombouts (2006)) and multivariate stochastic volatility (SV) models (Asai, McAleer, and Yu (2006), Chib, Omori, and Asai (2009)).

They are proposed to model the volatility clustering and the dynamic correlations, which are found to exist in empirical studies of financial time series (Bauwens, Hafner, and Laurent (2012)). Dynamic conditional correlation (DCC) models (Engle (2002)) and BEKK models (Engle and Kroner (1995)) are such widely used multivariate GARCH models. They simplify the multivariate covariance structure since there is an increasing difficulty in estimating too many parameters for dynamic correlations for high dimensional data. To overcome the difficulty, Engle and Kelly (2012), Vargas (2009), Jin and Tang (2009) and Clements, Coleman-Fenn, and Smith (2011) proposed the dynamic equicorrelation (DECO) model, which is based on a DCC model with all correlations equal but time-varying. Making reference to Elton and Gruber (1973), they argue that the dynamic equicorrelation assumption gives a superior portfolio allocation. In a Bayesian context, Ledoit and Wolf (2004) proposed the covariance matrix estimator obtained by shrinking the sample correlation matrix to an equicorrelated matrix for the purpose of the portfolio optimization. Hafner and Reznikova (2012) applied the shrinkage methods to the DCC models and improved the estimation results of the DCC model. Lucas, Schwaab, and Zhang (2012) proposed the dynamic generalized hyperbolic (GH) skew-t-error model with generalized autoregressive score (GAS) equicorrelation structure. For an asset allocation, an equicorrelated factor model is sometimes considered as a mean of dimension reduction (e.g., McNeil, Frey, and Embrechts (2005)).

In volatilities of stock returns, we often observe the asymmetry or the cross leverage effect, which implies a decrease in the i-th dependent variable at date t followed by an increase in the j-th latent stochastic variance at date $(t+1)$. The simple univariate SV model with leverage effect is given in a state space form as:

$$y_t = m + \exp(h_t/2)\epsilon_t, \quad t = 1, \dots, n, \tag{3.1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \dots, n-1, \tag{3.2}$$

$$h_1 \sim \mathrm{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right), \tag{3.3}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & q \\ q & \sigma_\eta^2 \end{pmatrix} \right), \tag{3.4}$$

$$|\phi| < 1, \tag{3.5}$$

where y_t denotes a (univariate) asset return, h_t is a log-variance of y_t. The negative value of q implies the existence of the leverage effect. It can be extended to the multivariate SV model with cross leverage effect (Daníelsson (1998), Asai and McAleer (2006), Asai and McAleer (2009), Chan, Kohn, and Kirby (2006), Ishihara, Omori, and Asai (2011), Ishihara and Omori (2012) and Nakajima (2012)). The major difficulty in constructing such multivariate models is to make the covariance matrices positive definite, especially when some dynamic correlation structure between the asset returns is incorporated. It is desirable to model the dynamic covariance structure as simply as possible since the parameter estimation becomes difficult in the sense that there are too many latent variables to be integrated out analytically to obtain the likelihood function.

In this article, we propose the multivariate SV model with dynamic equicorrelation and cross leverage effect (DESV model) and describe an efficient Bayesian estimation using the Markov chain Monte Carlo (MCMC) method to generate the latent stochastic volatilities and dynamic equicorrelation from the posterior distributions. As discussed in Shephard and Pitt (1997), Watanabe and Omori (2004) and Omori and Watanabe (2008), we divide all latent variables into several blocks and generate one block given other blocks (multi-move sampler or block-sampler). This is known to be more efficient than the simple single-move sampler, which draws the single latent stochastic volatility (the single dynamic equicorrelation factor) at a time given the other latent variables and the parameters. It means that we only need to generate a fewer number of MCMC samples to estimate the posterior distribution of the interested parameters.

The rest of this article is organized as follows. In Section 3.2, we propose the multivariate stochastic volatility model with dynamic equicorrelation and cross leverage effect. Section 3.3 describes an efficient Bayesian estimation method for the proposed model using a multi-move sampling method. A single-move sampling method that is simple but inefficient is also described as a benchmark. In Section 3.4, we illustrate our estimation method using simulated data and show that our MCMC algorithm is efficient. Section 3.5 applies our proposed DESV model to the trivariate asset return data based on industrial sector indices of TOPIX (Tokyo stock price index). Section 3.6 concludes this article.

3.2 Equicorrelation model

3.2.1 Equicorrelation matrix

Suppose that a p-dimensional random variable has an equicorrelation structure where the $p \times p$ equicorrelation matrix takes the form

$$R = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \dots & \rho & 1 \end{pmatrix} \tag{3.6}$$

$$= (1-\rho)I_p + \rho J_p, \tag{3.7}$$

I_p is a unit matrix of size p, and $J_p = \mathbf{1}_p\mathbf{1}_p'$ ($\mathbf{1}_p$ denotes a p-dimensional vector with all elements equal to one). The matrix R is positive definite if and only if ρ satisfies the condition $-(p-1)^{-1} < \rho < 1$. It is noted that the lower bound of ρ is depending on p, that is, as p becomes larger the lower bound approaches to zero. The determinant and the inverse are given by

$$|R| = (1-\rho)^{p-1}\{1+(p-1)\rho\}, \tag{3.8}$$

$$R^{-1} = \frac{1}{1-\rho}\left(I_p - \frac{\rho}{1-\rho+p\rho}J_p\right). \tag{3.9}$$

We note that the eigenvalues of R are $1-\rho$ (multiplicity $p-1$) and $1+(p-1)\rho$. The eigenvector $\boldsymbol{x} = (x_1, \dots, x_p)'$ associated with $1-\rho$ satisfies the condition $\sum_{i=1}^{p} x_i = 0$, while the eigenvector associated with $1 + (p-1)\rho$ satisfies the condition $x_1 = \cdots = x_p$. Thus the spectral decomposition of R is given by

$$R = \{1 + (p-1)\rho\}\boldsymbol{r}_1\boldsymbol{r}_1' + (1-\rho)\boldsymbol{r}_2\boldsymbol{r}_2' + \cdots + (1-\rho)\boldsymbol{r}_p\boldsymbol{r}_p',$$

where $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_p$ are the associated orthonormalized eigenvectors.

Let $R_D = \mathrm{diag}\{1+(p-1)\rho, 1-\rho, \ldots, 1-\rho\}$ and $R_O = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_p)$. Then, $R = R_O R_D R_O'$ and we denote $R^{1/2} = R_O R_D^{1/2}$ where $R_D^{1/2} = \mathrm{diag}\{(1+(p-1)\rho)^{1/2}, (1-\rho)^{1/2}, \ldots, (1-\rho)^{1/2}\}$. Alternatively we could decompose as $R = R_c R_c'$ (e.g., Choleski decomposition), but the spectral decomposition is advantageous in that it does not depend on the order of the random variables in the vector.

3.2.2 Multivariate SV model with dynamic equicorrelation

Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})'$ denote a p-dimensional asset return vector at time t $(t = 1, \ldots, n)$. Let $\boldsymbol{m}_t = (m_{1t}, \ldots, m_{pt})'$ and $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{pt})'$ denote p-dimensional vectors of unobserved variables and g_t an unobserved variable. We consider the multivariate SV model given by

$$\boldsymbol{y}_t = \boldsymbol{m}_t + V_t^{1/2}\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_p(\mathbf{0}_p, R_t), \quad t = 1, \ldots, n, \tag{3.10}$$

$$\boldsymbol{h}_{t+1} = \boldsymbol{\mu} + \Phi(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}_p(\mathbf{0}_p, \Omega), \quad t = 1, \ldots, n-1, \tag{3.11}$$

$$\boldsymbol{h}_1 = \boldsymbol{\mu} + \boldsymbol{\eta}_0, \quad \boldsymbol{\eta}_0 \sim \mathrm{N}_p(\mathbf{0}_p, \Omega_0), \tag{3.12}$$

$$g_{t+1} = \gamma + \theta(g_t - \gamma) + \zeta_t, \quad \zeta_t \sim \mathrm{N}(0, \sigma^2), \quad t = 1, \ldots, n-1, \tag{3.13}$$

$$g_1 = \gamma + \zeta_0, \quad \zeta_0 \sim \mathrm{N}(0, \sigma^2/(1-\theta^2)), \tag{3.14}$$

$$\boldsymbol{m}_{t+1} = \boldsymbol{m}_t + \boldsymbol{\eta}_{\boldsymbol{m}t}, \quad \boldsymbol{\eta}_{\boldsymbol{m}t} \sim \mathrm{N}_p(\mathbf{0}_p, \Omega_{\boldsymbol{m}}), \quad t = 1, \ldots, n-1, \tag{3.15}$$

$$\boldsymbol{m}_1 = \boldsymbol{\eta}_{\boldsymbol{m}0}, \quad \boldsymbol{\eta}_{\boldsymbol{m}0} \sim \mathrm{N}_p(\mathbf{0}_p, \kappa I_p), \tag{3.16}$$

where $\mathbf{0}_p$ is a p-dimensional zero vector,

$$V_t = \text{diag}\{\exp(h_{1t}), \dots, \exp(h_{pt})\},\ t = 1, \dots, n, \tag{3.17}$$

$$R_t = (1 - \rho_t) I_p + \rho_t J_p,\ t = 1, \dots, n, \tag{3.18}$$

$$\rho_t = \frac{\exp(g_t)}{\exp(g_t) + 1},\ t = 1, \dots, n, \tag{3.19}$$

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)', \tag{3.20}$$

$$\Phi = \text{diag}(\phi_1, \dots, \phi_p), \tag{3.21}$$

$$\Omega_{\boldsymbol{m}} = \text{diag}(\omega^2_{m_1}, \dots, \omega^2_{m_p}), \tag{3.22}$$

$$\begin{pmatrix} R_t^{-\frac{1}{2}} \boldsymbol{\epsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \text{N}_{2p}(\mathbf{0}_{2p}, \Psi),\ t = 1, \dots, n-1, \tag{3.23}$$

$$\Psi = \begin{pmatrix} I_p & Q' \\ Q & \Omega \end{pmatrix}, \tag{3.24}$$

and Ω_0, the covariance matrix of the initial latent variable $\boldsymbol{h}_1$, satisfies the stationary condition $\Omega_0 = \Phi \Omega_0 \Phi + \Omega$ such that

$$\text{vec}(\Omega_0) = (I_{p^2} - \Phi \otimes \Phi)^{-1} \text{vec}(\Omega). \tag{3.25}$$

We also set $\text{Var}(g_1) = \sigma^2/(1 - \theta^2)$, the variance of the initial latent variable g_1, for assuming the stationary condition and set κ, the variance of the initial latent variable $m_{j,1}$ ($j = 1, \dots, p$), equal to some large known constant. The latent vector, $\boldsymbol{h}_t$, is a vector of log-variances of the returns, and the latent variable, g_t, is the transformed equicorrelation of $\boldsymbol{y}_t$. For the identifiability, we set the diagonal elements of the covariance matrix of $\boldsymbol{\epsilon}_t$ equal to 1.

Notice that we define ρ_t, $t = 1, \dots n$, so as to take values on the unit interval, $(0, 1)$. As shown in the previous subsection, the equicorrelation matrix R_t is positive definite if and only if ρ_t is in $(-(p-1)^{-1}, 1)$. It means that as p becomes large, the negative region of the parameter space becomes smaller. Therefore it is reasonable to restrict the parameter space of ρ_t to the positive region so that the parameter space is independent of p.

We assume, for simplicity, that $\{\boldsymbol{m}_t\}_{t=1}^n$, the latent sequence of the expectations of $\boldsymbol{y}_t$, $t = 1, \dots, n$, follows a simple random walk. Empirical studies suggest that the expectation of the asset return is nearly 0 in most cases (e.g., McNeil, Frey, and Embrechts (2005)), but it is practically important to include non-zero asset means especially for portfolio optimizations as we shall see in Section 3.5.3.

3.3 Bayesian estimation

3.3.1 Priors and posterior densities

For prior distributions of $\{\boldsymbol{\mu}, \gamma, \Phi, \theta, \sigma^2, \Omega_{\boldsymbol{m}}\}$, we assume that

$$
\begin{aligned}
\boldsymbol{\mu} &\sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\mu}0}, S_{\boldsymbol{\mu}0}), && (3.26)\\
\gamma &\sim \mathrm{N}(m_{\gamma 0}, s^2_{\gamma 0}), && (3.27)\\
(\phi_j + 1)/2 &\sim \mathrm{Be}(a_{\phi_j}, b_{\phi_j}), \quad j = 1, \ldots, p, && (3.28)\\
(\theta + 1)/2 &\sim \mathrm{Be}(a_\theta, b_\theta), && (3.29)\\
\sigma^2 &\sim \mathrm{IG}(\alpha_{\sigma^2 0}/2, \beta_{\sigma^2 0}/2), && (3.30)\\
\omega^2_{m_j} &\sim \mathrm{IG}(\alpha_{m_j 0}/2, \beta_{m_j 0}/2), \quad j = 1, \ldots, p, && (3.31)
\end{aligned}
$$

where $\mathrm{Be}(a, b)$ denotes a beta distribution with parameters a, b and $\mathrm{IG}(\alpha, \beta)$ denotes an inverted gamma distribution with shape parameter α and scale parameter β. For a prior distribution of Ψ, we let

$$
\Psi^{-1} = \begin{pmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{pmatrix},
$$

where Ψ^{11}, Ψ^{12} and Ψ^{22} are $p \times p$ matrices, respectively. Noting that $\Psi^{11} = I_p + \Psi^{12}(\Psi^{22})^{-1}\Psi^{21}$, we assume

$$
\begin{aligned}
\Psi^{22} &\sim \mathrm{W}_p(n_0, S_0), && (3.32)\\
\Psi^{21} | \Psi^{22} &\sim \mathrm{N}_{p \times p}(\Psi^{22}\Delta_0, \Lambda_0 \otimes \Psi^{22}), && (3.33)
\end{aligned}
$$

where $\mathrm{W}(n, S)$ denotes a Wishart distribution with parameters (n, S).

Then, the joint posterior density function is

$$
\begin{aligned}
& f(\boldsymbol{\vartheta}, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n | \{\boldsymbol{y}_t\}_{t=1}^n) \\
& \propto \pi(\boldsymbol{\vartheta}) \times \exp\left(\sum_{t=1}^n l_t\right) \times |\Omega_0|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\Omega_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu}) \right\} \\
& \quad \times |\Omega|^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{n-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\} \right] \\
& \quad \times \left(\frac{\sigma^2}{1 - \theta^2}\right)^{-1/2} \exp\left\{ -\frac{(g_1 - \gamma)^2}{2\sigma^2/(1 - \theta^2)} \right\} \times (\sigma^2)^{-\frac{n-1}{2}} \exp\left[-\frac{\sum_{t=1}^{n-1}\{g_{t+1} - (1 - \theta)\gamma - \theta g_t\}^2}{2\sigma^2} \right] \\
& \quad \times \exp\left(-\frac{1}{2\kappa}\boldsymbol{m}_1'\boldsymbol{m}_1 \right) \times |\Omega_{\boldsymbol{m}}|^{-\frac{n-1}{2}} \exp\left\{ -\frac{1}{2}\sum_{t=1}^{n-1}(\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)'\Omega_{\boldsymbol{m}}^{-1}(\boldsymbol{m}_{t+1} - \boldsymbol{m}_t) \right\},
\end{aligned}
\tag{3.34}
$$

where

$$
\begin{aligned}
l_t = &-\frac{1}{2}\left[R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}(\boldsymbol{y}_t-\boldsymbol{m}_t)-Q'\Omega^{-1}\{\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\Phi(\boldsymbol{h}_t-\boldsymbol{\mu})\}\right]' \\
&\times(I_p-Q'\Omega^{-1}Q)^{-1}\left[R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}(\boldsymbol{y}_t-\boldsymbol{m}_t)-Q'\Omega^{-1}\{\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\Phi(\boldsymbol{h}_t-\boldsymbol{\mu})\}\right] \\
&-\frac{1}{2}\log\{(1-\rho_t)^{p-1}(1+(p-1)\rho_t)\}-\frac{1}{2}\textstyle\sum_{j=1}^{p}h_{jt},
\end{aligned} \tag{3.35}
$$

and $\boldsymbol{\vartheta}=\{\boldsymbol{\mu},\gamma,\Phi,\theta,\Omega,Q,\sigma^2,\Omega_{\boldsymbol{m}}\}$.

We implement the MCMC algorithm in twelve blocks:

1. Initialize $\{\boldsymbol{h}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\mu},\gamma,\Phi,\theta,\Omega,Q,\sigma^2,\Omega_{\boldsymbol{m}}$.
2. Generate $\boldsymbol{\mu}|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{h}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\Phi,\Omega,Q$.
3. Generate $\gamma|\{\boldsymbol{y}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\theta,\sigma^2$.
4. Generate $\Phi|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{h}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\mu},\Omega,Q$.
5. Generate $\theta|\{\boldsymbol{y}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\gamma,\sigma^2$.
6. Generate $\Omega,Q|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{h}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\mu},\Phi$.
7. Generate $\sigma^2|\{\boldsymbol{y}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\gamma,\theta$.
8. Generate $\Omega_{\boldsymbol{m}}|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n$.
9. Generate $\{\boldsymbol{h}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\mu},\Phi,\Omega,Q$.
10. Generate $\{g_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{h}_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\mu},\gamma,\Phi,\theta,\Omega,Q,\sigma^2$.
11. Generate $\{\boldsymbol{m}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n,\{\boldsymbol{h}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\boldsymbol{\mu},\Phi,\Omega,Q,\Omega_{\boldsymbol{m}}$.
12. Go to 2.

3.3.2 Generation of latent variable $\{\boldsymbol{h}_t\}_{t=1}^n$

3.3.2.1 Single-move sampling method

A simple sampling method for $\{\boldsymbol{h}_t\}_{t=1}^n$, is a single-move sampler that draws a single latent variable $\boldsymbol{h}_t$ at a time given the other $\boldsymbol{h}_t$'s and the parameters. The other method is a multi-move sampler that draws multiple $\boldsymbol{h}_t$'s simultaneously. A single-move sampler is simpler than a multi-move sampler, but a multi-move sampler is known to be more efficient (Shephard and Pitt (1997), Watanabe and Omori (2004) and Omori and Watanabe (2008)). As a benchmark, we first describe the single-move sampling method. We generate a candidate based on the following approximation of l_t:

$$
\begin{aligned}
l_t &\approx -\frac{1}{2}(\boldsymbol{y}_t-\boldsymbol{m}_t)'V_t^{-1/2}R_t^{-1}V_t^{-1/2}(\boldsymbol{y}_t-\boldsymbol{m}_t)-\frac{1}{2}\log|R_t|-\frac{1}{2}\sum_{j=1}^{p}h_{jt}\\
&= -\frac{1}{2}\begin{pmatrix}(y_{1t}-m_{1t})\exp(-\frac{1}{2}h_{1t})\\ \vdots \\ (y_{pt}-m_{pt})\exp(-\frac{1}{2}h_{pt})\end{pmatrix}' R_t^{-1}\begin{pmatrix}(y_{1t}-m_{1t})\exp(-\frac{1}{2}h_{1t})\\ \vdots \\ (y_{pt}-m_{pt})\exp(-\frac{1}{2}h_{pt})\end{pmatrix}-\frac{1}{2}\log|R_t|-\frac{1}{2}\sum_{j=1}^{p}h_{jt}\\
&\approx -\frac{1}{2}\begin{pmatrix}(y_{1t}-m_{1t})(1-\frac{h_{1t}}{2})\\ \vdots \\ (y_{pt}-m_{pt})(1-\frac{h_{pt}}{2})\end{pmatrix}' R_t^{-1}\begin{pmatrix}(y_{1t}-m_{1t})(1-\frac{h_{1t}}{2})\\ \vdots \\ (y_{pt}-m_{pt})(1-\frac{h_{pt}}{2})\end{pmatrix}-\frac{1}{2}\log|R_t|-\frac{1}{2}\sum_{j=1}^{p}h_{jt}.
\end{aligned}
$$

For $t=1,\dots,n$, we propose a candidate $\boldsymbol{h}_t|\{\boldsymbol{h}_s\}_{-t},\{\boldsymbol{y}_t\}_{t=1}^n,\{g_t\}_{t=1}^n,\{\boldsymbol{m}_t\}_{t=1}^n,\boldsymbol{\vartheta}\sim \mathrm{N}(\boldsymbol{m}_{\boldsymbol{h}t},S_{\boldsymbol{h}t})$, where

$$
S_{\boldsymbol{h}t}=\{\mathrm{diag}\{(\boldsymbol{y}_t-\boldsymbol{m}_t)/2\}R_t^{-1}\mathrm{diag}\{(\boldsymbol{y}_t-\boldsymbol{m}_t)/2\}+\Phi\Omega^{-1}\Phi+\Omega^{-1}\}^{-1},
$$

$$
\begin{aligned}
\boldsymbol{m}_{\boldsymbol{h}t}=S_{\boldsymbol{h}t}[&\mathrm{diag}\{(\boldsymbol{y}_t-\boldsymbol{m}_t)/2\}R_t^{-1}\mathrm{diag}\{(\boldsymbol{y}_t-\boldsymbol{m}_t)/2\}2\mathbf{1}_p-0.5\mathbf{1}_p+\Phi\Omega^{-1}\{\boldsymbol{h}_{t+1}-(I_p-\Phi)\boldsymbol{\mu}\}\\
&+\Omega^{-1}\{\boldsymbol{\mu}+\Phi(\boldsymbol{h}_{t-1}-\boldsymbol{\mu})\}],
\end{aligned}
$$

and accept it using MH algorithm.

3.3.2.2 Efficient multi-move sampling method

A simulation smoother, an efficient sampler for the state variables was proposed by de Jong and Shephard (1995) and by Durbin and Koopman (2002) for the linear Gaussian state space model. However, such a simulation smoother cannot be applied directly to our nonlinear model. As discussed in Shephard and Pitt (1997), Watanabe and Omori (2004) and Omori and Watanabe (2008), we approximate the nonlinear Gaussian likelihood function by the linear Gaussian likelihood function and implement the MH algorithm.

In this algorithm, we first divide $\{\boldsymbol{h}_t\}_{t=1}^n$ into $K+1$ blocks, $(\boldsymbol{h}_{k_{m-1}+1},\dots,\boldsymbol{h}_{k_m})$, $m=1,\dots,K$ with $k_0=0$, $k_{K+1}=n$, $k_i-k_{i-1}\geq 2$, using stochastic knots $k_m=\mathrm{int}[n(m+U_m)/(K+2)]$, where U_m's are independent uniform random variables on $(0,1)$. Next, we generate $(\boldsymbol{h}_{k_{m-1}+1},\dots,\boldsymbol{h}_{k_m})$ given other blocks by generating $(\underline{\boldsymbol{\eta}}_{k_{m-1}},\dots,\underline{\boldsymbol{\eta}}_{k_m-1})$, where $\underline{\boldsymbol{\eta}}_t=\Omega^{-1/2}\boldsymbol{\eta}_t$.

The conditional posterior density of $\underline{\boldsymbol{\eta}} = (\underline{\boldsymbol{\eta}}_s', \ldots, \underline{\boldsymbol{\eta}}_{s+r-1}')'$ is given by

$$
\begin{aligned}
& f(\underline{\boldsymbol{\eta}}_s, \ldots, \underline{\boldsymbol{\eta}}_{s+r-1} | \{\boldsymbol{y}_t\}_{t=1}^n, \boldsymbol{h}_s, \boldsymbol{h}_{s+r+1}, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}) \\
& \propto \prod_{t=s}^{s+r} f(\boldsymbol{y}_t | \boldsymbol{h}_t, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}) \prod_{t=s}^{s+r-1} f(\boldsymbol{\eta}_t | \boldsymbol{\vartheta}) \times (f(\boldsymbol{h}_{s+r+1} | \boldsymbol{h}_{s+r}, \boldsymbol{\vartheta}))^{\mathrm{I}_{\{s+r<n\}}} \\
& \propto \exp\left(L - \frac{1}{2} \sum_{t=s}^{s+r-1} \underline{\boldsymbol{\eta}}_t' \underline{\boldsymbol{\eta}}_t \right),
\end{aligned}
$$

where

$$
L = \begin{cases} \displaystyle\sum_{t=s}^{s+r} l_t - \frac{1}{2} \{\boldsymbol{h}_{s+r+1} - (I_p - \Phi)\boldsymbol{\mu} - \Phi \boldsymbol{h}_{s+r}\}' \Omega^{-1} \{\boldsymbol{h}_{s+r+1} - (I_p - \Phi)\boldsymbol{\mu} - \Phi \boldsymbol{h}_{s+r}\} \\ \quad \text{if } s + r < n, \\ \displaystyle\sum_{t=s}^{n} l_t \quad \text{if } s + r = n. \end{cases}
$$

Using Taylor expansion around the conditional posterior mode of $\underline{\boldsymbol{\eta}} = (\underline{\boldsymbol{\eta}}_s', \ldots, \underline{\boldsymbol{\eta}}_{s+r-1}')'$, we approximate L and construct a proposal density (linear and Gaussian state-space model) as

$$
\begin{aligned}
& \log f(\underline{\boldsymbol{\eta}}_s, \ldots, \underline{\boldsymbol{\eta}}_{s+r-1} | \boldsymbol{h}_s, \boldsymbol{h}_{s+r+1}, \boldsymbol{y}_s, \ldots, \boldsymbol{y}_{s+r}) \\
& \approx \text{const.} - \frac{1}{2} \sum_{t=s}^{s+r-1} \underline{\boldsymbol{\eta}}_t' \underline{\boldsymbol{\eta}}_t + \hat{L} + \left.\frac{\partial L}{\partial \underline{\boldsymbol{\eta}}'}\right|_{\underline{\boldsymbol{\eta}} = \hat{\underline{\boldsymbol{\eta}}}} (\underline{\boldsymbol{\eta}} - \hat{\underline{\boldsymbol{\eta}}}) + \frac{1}{2} (\underline{\boldsymbol{\eta}} - \hat{\underline{\boldsymbol{\eta}}})' \mathrm{E}\left.\left(\frac{\partial^2 L}{\partial \underline{\boldsymbol{\eta}} \partial \underline{\boldsymbol{\eta}}'} \right)\right|_{\underline{\boldsymbol{\eta}} = \hat{\underline{\boldsymbol{\eta}}}} (\underline{\boldsymbol{\eta}} - \hat{\underline{\boldsymbol{\eta}}}) \\
& = \text{const.} - \frac{1}{2} \sum_{t=s}^{s+r-1} \underline{\boldsymbol{\eta}}_t' \underline{\boldsymbol{\eta}}_t + \hat{L} + \hat{\boldsymbol{d}}' (\boldsymbol{h} - \hat{\boldsymbol{h}}) + \frac{1}{2} (\boldsymbol{h} - \hat{\boldsymbol{h}})' \mathrm{E}\left.\left(\frac{\partial^2 L}{\partial \boldsymbol{h} \partial \boldsymbol{h}'} \right)\right|_{\boldsymbol{h} = \hat{\boldsymbol{h}}} (\boldsymbol{h} - \hat{\boldsymbol{h}}) \\
& = \log f^*(\underline{\boldsymbol{\eta}}_s, \ldots, \underline{\boldsymbol{\eta}}_{s+r-1} | \boldsymbol{h}_s, \boldsymbol{h}_{s+r+1}, \boldsymbol{y}_s, \ldots, \boldsymbol{y}_{s+r}),
\end{aligned}
$$

where $\boldsymbol{h} = (\boldsymbol{h}_{s+1}', \ldots, \boldsymbol{h}_{s+r}')'$ and

$$
\boldsymbol{d} = (\boldsymbol{d}_{s+1}', \ldots, \boldsymbol{d}_{s+r}')', \quad \boldsymbol{d}_t = \partial L / \partial \boldsymbol{h}_t, \tag{3.36}
$$

$$
-\mathrm{E}\left(\frac{\partial^2 L}{\partial \boldsymbol{h} \partial \boldsymbol{h}'} \right) = \begin{pmatrix} A_{s+1} & B_{s+2}' & O & \cdots & O \\ B_{s+2} & A_{s+2} & B_{s+3}' & \cdots & O \\ O & B_{s+3} & A_{s+3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & B_{s+r}' \\ O & \cdots & O & B_{s+r} & A_{s+r} \end{pmatrix}, \tag{3.37}
$$

$$
A_t = -\mathrm{E}\left(\frac{\partial^2 L}{\partial \boldsymbol{h}_t \partial \boldsymbol{h}_t'} \right), \; t = s+1, \ldots, s+r, \tag{3.38}
$$

$$B_t = -\mathrm{E}\left(\frac{\partial^2 L}{\partial \boldsymbol{h}_t \partial \boldsymbol{h}'_{t-1}}\right),\ t = s+2, \ldots, s+r,\ B_{s+1} = O \tag{3.39}$$

(see Appendix 3.A.1 for the derivation of $\boldsymbol{d}_t, A_t, B_t$). We generate $\{\underline{\boldsymbol{\eta}}_t\}_{t=s}^{s+r-1}$ in two steps:

Step 1 (Disturbance smoother).

(a) Initialize $\hat{\underline{\boldsymbol{\eta}}}_t,\ t = s, \ldots, s+r-1$.

(b) Compute $\hat{\boldsymbol{d}}_t, \hat{A}_t, \hat{B}_t,\ t = s+1, \ldots, s+r$. ($\boldsymbol{d}_t, A_t, B_t$ evaluated at $\hat{\underline{\boldsymbol{\eta}}}_t$)

(c) For $t = s+2, \ldots, s+r$, compute

$$\begin{aligned}
C_t &= \hat{A}_t - \hat{B}_t C_{t-1}^{-1} \hat{B}'{}_t,\ C_{s+1} = \hat{A}_{s+1},\ C_t = F_t F_t', \\
M_t &= \hat{B}_t F'{}_{t-1}^{-1},\ M_{s+1} = O,\ M_{s+r+1} = O, \\
\boldsymbol{b}_t &= \hat{\boldsymbol{d}}_t - M_t F_{t-1}^{-1} \boldsymbol{b}_{t-1},\ \boldsymbol{b}_{s+1} = \hat{\boldsymbol{d}}_{s+1}.
\end{aligned}$$

(d) For $t = s+1, \ldots, s+r$, define
$\hat{\boldsymbol{y}}_t = \hat{\boldsymbol{\gamma}}_t + C_t^{-1} \boldsymbol{b}_t,\ \ \hat{\boldsymbol{\gamma}}_t = \hat{\boldsymbol{h}}_t + F_t'^{-1} M'_{t+1} \hat{\boldsymbol{h}}_{t+1}$.

(e) Consider the linear and Gaussian state-space model:

$$\hat{\boldsymbol{y}}_t = Z_t \boldsymbol{h}_t + G_t \boldsymbol{u}_t, \tag{3.40}$$

$$\boldsymbol{h}_{t+1} = (I_p - \Phi)\boldsymbol{\mu} + \Phi \boldsymbol{h}_t + H_t \boldsymbol{u}_t, \tag{3.41}$$

$$Z_t = I_p + F_t'^{-1} M'_{t+1} \Phi,\ G_t = F_t'^{-1}[I_p, M'_{t+1}\mathrm{chol}(\Omega)], \tag{3.42}$$

$$H_t = [O, \mathrm{chol}(\Omega)]. \tag{3.43}$$

(chol(Ω) denotes Choleski decomposition of Ω.)

(f) Apply Kalman filter and disturbance smoother (Koopman (1993)) and update $\hat{\underline{\boldsymbol{\eta}}}$.

(g) Go to (b) until $\hat{\underline{\boldsymbol{\eta}}}$ converges to the mode.

Step 2. (a) Update $\hat{\underline{\boldsymbol{\eta}}}$ using the disturbance smoother and find a linear and Gaussian state-space model (3.40)-(3.43).

(b) Generate $\underline{\boldsymbol{\eta}}^{\dagger} \sim f^*$ (the linear and Gaussian state-space model) using Kalman filter and simulation smoother (de Jong and Shephard (1995) or Durbin and Koopman (2002)). Conduct AR-MH algorithm.

3.3.3 Generation of latent variables $\{g_t\}_{t=1}^n$ and $\{\boldsymbol{m}_t\}_{t=1}^n$

We consider the two following sampling methods for $\{g_t\}_{t=1}^n$ corresponding to the last subsection.

Single-move sampling method. Generate g_t given $\{g_s\}_{-t}, \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$ using a random walk MH algorithm for $t = 1, \ldots, n$.

Multi-move sampling method. We divide $\{g_t\}_{t=1}^n$ into $K+1$ blocks using stochastic knots and generate one block given other blocks using the block sampler as mentioned in the last subsection. See Appendix A.1 for the derivation of $\boldsymbol{d}_t, A_t$ and B_t.

We generate all of $\boldsymbol{m}_t$'s simultaneously using a simulation smoother (de Jong and Shephard (1995) or Durbin and Koopman (2002)).

3.3.4 Generation of $\boldsymbol{\mu}, \gamma, \Phi, \theta, \sigma^2, \Omega_{\boldsymbol{m}}$

Generation of $\boldsymbol{\mu}$. The conditional distribution of $\boldsymbol{\mu}$ is

$$\boldsymbol{\mu}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \Phi, \Omega, Q \sim \mathrm{N}(\boldsymbol{m}_{\boldsymbol{\mu}}, S_{\boldsymbol{\mu}}), \tag{3.44}$$

where

$$S_{\boldsymbol{\mu}} = \{S_{\boldsymbol{\mu}0}^{-1} + \Omega_0^{-1} + (n-1)(I_p - \Phi)(\Omega - QQ')^{-1}(I_p - \Phi)\}^{-1}, \tag{3.45}$$

$$\boldsymbol{m}_{\boldsymbol{\mu}} = S_{\boldsymbol{\mu}}\left\{S_{\boldsymbol{\mu}0}^{-1}\boldsymbol{m}_{\boldsymbol{\mu}0} + \Omega_0^{-1}\boldsymbol{h}_1 + (I_p - \Phi)(\Omega - QQ')^{-1}\sum_{t=1}^{n-1}(\boldsymbol{h}_{t+1} - \Phi\boldsymbol{h}_t - Q\boldsymbol{z}_t)\right\}, \tag{3.46}$$

$$\boldsymbol{z}_t = R_t^{-1/2}V_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t).$$

Generation of γ. The conditional distribution of γ is

$$\gamma|\{\boldsymbol{y}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \theta, \sigma^2 \sim \mathrm{N}(m_\gamma, s_\gamma^2), \tag{3.47}$$

where

$$s_\gamma^2 = \{s_{\gamma0}^{-2} + (1-\theta^2)\sigma^{-2} + (n-1)(1-\theta)^2\sigma^{-2}\}^{-1}, \tag{3.48}$$

$$m_\gamma = s_\gamma^2\left[s_{\gamma0}^{-2}m_{\gamma0} + (1-\theta^2)\sigma^{-2}g_1 + (1-\theta)\sigma^{-2}\left\{\sum_{t=2}^{n} g_t - \theta\sum_{t=1}^{n-1} g_t\right\}\right]. \tag{3.49}$$

Generation of Φ*.* The conditional posterior density of $\boldsymbol{\phi} = \Phi \mathbf{1}_p$ is given by

$$f(\boldsymbol{\phi}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\mu}, \Omega, Q) \propto g(\boldsymbol{\phi}) \times \exp\left\{ -\frac{1}{2}(\boldsymbol{\phi}-\boldsymbol{m}_{\boldsymbol{\phi}})' S_{\boldsymbol{\phi}}^{-1}(\boldsymbol{\phi}-\boldsymbol{m}_{\boldsymbol{\phi}})\right\}, \tag{3.50}$$

where

$$S_{\boldsymbol{\phi}} = \left\{ \sum_{t=1}^{n-1} ((\boldsymbol{h}_t - \boldsymbol{\mu})(\boldsymbol{h}_t - \boldsymbol{\mu})') \odot (\Omega - QQ')^{-1} \right\}^{-1}, \tag{3.51}$$

$$(\boldsymbol{m}_{t,\boldsymbol{\phi}})_i = \left\{ (\Omega - QQ')^{-1}(\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - Q\boldsymbol{z}_t)(\boldsymbol{h}_t - \boldsymbol{\mu})' \right\}_{i,i}, \quad \boldsymbol{m}_{\boldsymbol{\phi}} = S_{\boldsymbol{\phi}} \sum_{t=1}^{n-1} \boldsymbol{m}_{t,\boldsymbol{\phi}}, \tag{3.52}$$

$$g(\boldsymbol{\phi}) = \prod_{j=1}^{p} \pi(\phi_j) \times |\Omega_0|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})' \Omega_0^{-1} (\boldsymbol{h}_1 - \boldsymbol{\mu}) \right\}, \tag{3.53}$$

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

Generation of θ*.* The conditional posterior density of θ is given by

$$f(\theta|\{\boldsymbol{y}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \gamma, \sigma^2) \propto g(\theta) \times \exp\left\{ -\frac{1}{2s_\theta^2}(\theta - m_\theta)^2 \right\}, \tag{3.54}$$

where

$$s_\theta^2 = \sigma^2 \left\{ \sum_{t=1}^{n-1} (g_t - \gamma)^2 \right\}^{-1}, \quad m_\theta = s_\theta^2 \sum_{t=1}^{n-1} \sigma^{-2}(g_{t+1} - \gamma)(g_t - \gamma), \tag{3.55}$$

$$g(\theta) = \pi(\theta) \times (1 - \theta^2)^{1/2} \exp\left\{ -\frac{(g_1 - \gamma)^2}{2\sigma^2/(1 - \theta^2)} \right\}. \tag{3.56}$$

Generate a candidate $\theta^\dagger \sim \mathrm{TN}_{(-1,1)}(m_\theta, s_\theta^2)$ and accept it with probability $\min[1, g(\theta^\dagger)/g(\theta)]$.

Generation of σ^2*.* The conditional distribution of σ^2 is

$$\sigma^2|\{\boldsymbol{y}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \gamma, \theta \sim \mathrm{IG}(\alpha_{\sigma^2 1}/2, \beta_{\sigma^2 1}/2), \tag{3.57}$$

where $\alpha_{\sigma^2 1} = \alpha_{\sigma^2 0} + n$ and

$$\beta_{\sigma^2 1} = \beta_{\sigma^2 0} + (g_1 - \gamma)^2(1 - \theta^2) + \sum_{t=1}^{n-1} \{g_{t+1} - \gamma - \theta(g_t - \gamma)\}^2. \tag{3.58}$$

Generation of $\Omega_{\boldsymbol{m}}$*.* The conditional distribution of $\omega_{m_j}^2$, $j = 1, \ldots, p$, is

$$\omega_{m_j}^2|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n \sim \mathrm{IG}(\alpha_{m_j 1}/2, \beta_{m_j 1}/2), \tag{3.59}$$

where $\alpha_{m_j1} = \alpha_{m_j0} + n - 1$ and

$$\beta_{m_j1} = \beta_{m_j0} + \sum_{t=1}^{n-1}(m_{j,t+1} - m_{jt})^2. \tag{3.60}$$

3.3.5 Generation of Ω, Q

The conditional posterior density of Ψ^{12} and Ψ^{22} is given by

$$\begin{aligned}
& f(\Psi^{12}, \Psi^{22}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\mu}, \Phi) \\
& \propto \prod_{t=1}^{n-1} f(\boldsymbol{z}_t, \boldsymbol{\eta}_t|\boldsymbol{\vartheta}) \times \pi(\Psi^{21}, \Psi^{22}) \\
& \propto |\Omega_0|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\eta}_0'\Omega_0^{-1}\boldsymbol{\eta}_0 \right) \times |\Psi^{22}|^{(n_1-p-1)/2} \exp\left\{ -\frac{1}{2}\mathrm{tr}(S_1^{-1}\Psi^{22}) \right\} \\
& \quad \times |\Psi^{22}|^{-p/2} \exp\left[-\frac{1}{2}\{\mathrm{vec}(\Psi^{21} - \Psi^{22}\Delta_1)\}'(\Lambda_1 \otimes \Psi^{22})^{-1}\mathrm{vec}(\Psi_{21} - \Psi_{22}\Delta_1) \right]
\end{aligned} \tag{3.61}$$

(see, e.g., Gupta and Nagar (2000) and Ishihara, Omori, and Asai (2011)), where

$$n_1 = n_0 + n - 1,\ S_1 = (S_0^{-1} + \Xi_{22} + \Delta_0\Lambda_0^{-1}\Delta_0' - \Delta_1\Lambda_1^{-1}\Delta_1')^{-1}, \tag{3.62}$$

$$\Lambda_1 = (\Lambda_0^{-1} + \Xi_{11})^{-1},\ \Delta_1 = (-\Xi_{21} + \Delta_0\Lambda_0^{-1})\Lambda_1, \tag{3.63}$$

$$\Xi = \begin{pmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{21} & \Xi_{22} \end{pmatrix} = \sum_{t=1}^{n-1} \begin{pmatrix} \boldsymbol{z}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \begin{pmatrix} \boldsymbol{z}_t \\ \boldsymbol{\eta}_t \end{pmatrix}'. \tag{3.64}$$

We generate a candidate $\Psi^\dagger$ in three steps:

1. Generate $(\Psi^{22})^\dagger \sim \mathrm{W}(n_1, S_1)$.

2. Generate $(\Psi^{21})^\dagger|(\Psi^{22})^\dagger \sim \mathrm{N}_{p\times p}((\Psi^{22})^\dagger\Delta_1, \Lambda_1 \otimes (\Psi^{22})^\dagger)$.

3. Compute $Q^\dagger = -((\Psi^{22})^\dagger)^{-1}(\Psi^{21})^\dagger$, $\Omega^\dagger = ((\Psi^{22})^\dagger)^{-1} + Q^\dagger(Q^\dagger)'$ and accept $\Psi^\dagger$ with probability

$$\min\left[1, \exp\left\{ -\frac{1}{2}\log|\Omega_0^\dagger| - \frac{1}{2}\boldsymbol{\eta}_0'(\Omega_0^\dagger)^{-1}\boldsymbol{\eta}_0 + \frac{1}{2}\log|\Omega_0| + \frac{1}{2}\boldsymbol{\eta}_0'\Omega_0^{-1}\boldsymbol{\eta}_0 \right\}\right].$$

3.4 Illustrative example using simulated data

This section illustrates our proposed DESV model using simulated data. We consider a trivariate case ($p = 3$) and investigate the efficiency of our multi-move sampling method in comparison with the single-move sampling method. Using the following parameters based on our empirical studies in Section 3.5,

$$\boldsymbol{\mu}_* = \mathbf{0}_3,\ \gamma_* = 1.7,\ \Phi_* = 0.97I_3,\ \theta_* = 0.97, \Omega_* = 0.015I_3 + 0.015J_3,$$
$$Q_* = (-0.1 \times \mathbf{1}_3, \mathbf{0}_3, \mathbf{0}_3),\ \sigma_*^2 = 0.05,\ \Omega_{\boldsymbol{m}*} = 0.001I_3,$$

we generate 2,000 observations ($n = 2000$). For prior distributions, we assume

$$\boldsymbol{\mu} \sim \mathrm{N}(\boldsymbol{\mu}_*, 100I_3),\ \gamma \sim \mathrm{N}(\gamma_*, 100),\ \frac{\phi_i + 1}{2} \sim \mathrm{Be}(20, 1.5),\ i = 1, 2, 3,$$
$$\frac{\theta + 1}{2} \sim \mathrm{Be}(20, 1.5),\ \sigma^2 \sim \mathrm{IG}(5, 3\sigma_*^2),\ \omega_{m_j 0}^2 \sim \mathrm{IG}(5, 3\omega_{m_j *}^2),\ j = 1, 2, 3,$$
$$\Psi^{22} \sim \mathrm{W}(6, 6^{-1}\Omega_*),\ \Psi^{21}|\Psi^{22} \sim \mathrm{N}_{3\times 3}(O, 10I_3 \otimes \Psi^{22}).$$

We set $\kappa = 10$. Using the single-move sampler, we generate 250,000 MCMC samples after discarding the first 10,000 samples as the burn-in period. Also, the multi-move sampler is used to generate 50,000 MCMC samples after discarding the first 10,000 samples as the burn-in period. We set the number of the blocks to 301 ($K = 300$) based on several trials.

Table 3.1 reports the true values, posterior means, 95% credible intervals and estimated inefficiency factors (IF). The inefficiency factor is defined as $1 + 2\sum_{g=1}^{\infty} \rho(g)$, where $\rho(g)$ is the sample autocorrelation at lag g. This is interpreted as the ratio of the numerical variance of the posterior mean from the chain to the variance of the posterior mean from hypothetical uncorrelated draws. The smaller the inefficiency factor becomes, the closer the MCMC sampling is to the uncorrelated sampling.

The posterior means are all close to the true values, which suggests that our proposed algorithms work well. The inefficiency factors for the single-move sampler (the maximum is 1274) are larger than those for the multi-move sampler (the maximum is 201), which suggests that our proposed multi-move sampling method is efficient compared with the single-move sampling method[1].

[1]The acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{g_t\}_{t=1}^n$ are 0.641 and 0.802 using the multi-move sampler and 0.639 and 0.614 using the single-move sampler.

Table 3.1: Posterior means, 95% credible intervals and inefficiency factors.

	True	Mean	95% interval	IF (m-move)	(s-move)
μ_1	0	-0.059	(-0.431, 0.319)	6	49
μ_2	0	-0.053	(-0.366, 0.272)	6	68
μ_3	0	0.098	(-0.129, 0.340)	20	166
γ	1.7	1.597	(1.322, 1.868)	9	149
ϕ_1	0.97	0.979	(0.967, 0.989)	91	277
ϕ_2	0.97	0.976	(0.962, 0.988)	73	273
ϕ_3	0.97	0.966	(0.949, 0.981)	122	606
θ	0.97	0.959	(0.934, 0.978)	54	644
Ω_{11}	0.03	0.028	(0.018, 0.040)	**201**	1082
Ω_{21}	0.015	0.010	(0.003, 0.018)	104	**1274**
Ω_{31}	0.015	0.014	(0.007, 0.023)	107	796
Ω_{22}	0.03	0.025	(0.016, 0.037)	182	1017
Ω_{32}	0.015	0.011	(0.004, 0.020)	147	683
Ω_{33}	0.03	0.027	(0.017, 0.040)	185	1006
Q_{11}	-0.1	-0.065	(-0.096, -0.033)	110	400
Q_{21}	-0.1	-0.063	(-0.093, -0.033)	121	302
Q_{31}	-0.1	-0.081	(-0.112, -0.051)	109	462
Q_{12}	0	-0.009	(-0.042, 0.026)	78	432
Q_{22}	0	-0.004	(-0.040, 0.032)	112	604
Q_{32}	0	-0.015	(-0.049, 0.020)	118	677
Q_{13}	0	-0.009	(-0.047, 0.028)	107	417
Q_{23}	0	-0.030	(-0.068, 0.007)	131	477
Q_{33}	0	-0.010	(-0.047, 0.027)	170	527
σ^2	0.05	0.052	(0.031, 0.081)	106	1130
$\omega_{m_1}^2 \times 10^3$	1	0.889	(0.536, 1.419)	139	184
$\omega_{m_2}^2 \times 10^3$	1	0.727	(0.418, 1.146)	55	191
$\omega_{m_3}^2 \times 10^3$	1	0.810	(0.452, 1.305)	163	178

The maximum IF's are indicated in bold type.

3.5 Empirical study

3.5.1 Data

This section applies our proposed model to returns of subindices of Tokyo stock price index (TOPIX) — three industrial sector indices: (1) Machinery, (2) Electric Appliances and (3) Precision Instruments. The sample period is from January 4, 2005 to December 28, 2012 (1964 observations in total). The asset return is calculated as $y_t = (\log p_t - \log p_{t-1}) \times 100$, where p_t is the asset price at time t. Figure 3.1 shows the time series plot of the three returns. The trajectories are relatively similar to each other, so it is expected that the equicorrelation parameters are estimated to be positive.

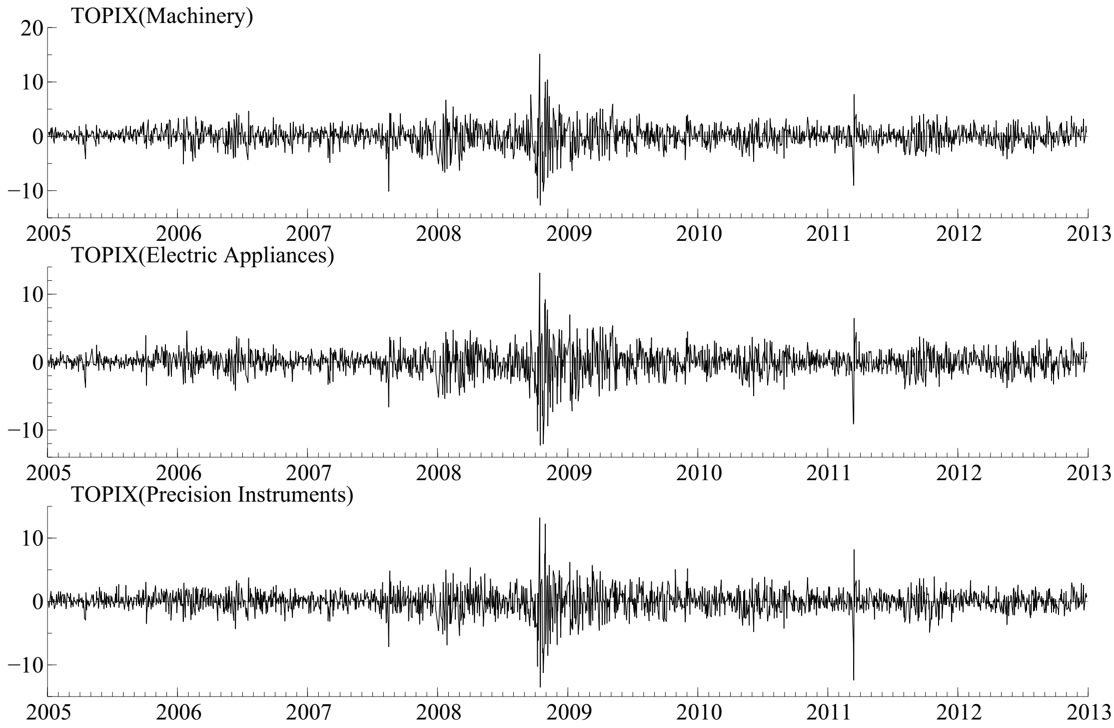

Figure 3.1: Time series plot of TOPIX.

3.5.2 Estimation results

Using the same prior distributions for the parameters as in Section 3.4, we implement the MCMC algorithm to conduct a Bayesian inference on the parameters of interest. We generate

100,000 MCMC samples from the posterior distributions of the parameters in the model after discarding the first 10,000 samples as the burn-in period.

Table 3.2: Posterior means, 95% credible intervals and inefficiency factors.

	Mean	95% interval	IF
μ_1	0.767	(0.478, 1.040)	29
μ_2	0.614	(0.268, 0.939)	17
μ_3	0.683	(0.375, 0.973)	33
γ	1.724	(1.491, 1.946)	19
ϕ_1	0.971	(0.959, 0.982)	450
ϕ_2	0.978	(0.968, 0.987)	476
ϕ_3	0.976	(0.965, 0.985)	360
θ	0.943	(0.897, 0.975)	316
Ω_{11}	0.030	(0.020, 0.043)	260
Ω_{21}	0.027	(0.018, 0.038)	270
Ω_{31}	0.026	(0.017, 0.037)	240
Ω_{22}	0.025	(0.016, 0.037)	393
Ω_{32}	0.024	(0.015, 0.035)	352
Ω_{33}	0.024	(0.015, 0.036)	423
Q_{11}	-0.108	(-0.138, -0.080)	262
Q_{21}	-0.086	(-0.113, -0.060)	241
Q_{31}	-0.084	(-0.110, -0.059)	193
Q_{12}	-0.015	(-0.039, 0.012)	187
Q_{22}	-0.005	(-0.032, 0.021)	368
Q_{32}	-0.002	(-0.029, 0.024)	373
Q_{13}	0.008	(-0.026, 0.043)	178
Q_{23}	0.004	(-0.031, 0.039)	582
Q_{33}	-0.014	(-0.047, 0.019)	256
σ^2	0.061	(0.028, 0.111)	361
$\omega^2_{m_1} \times 10^3$	0.261	(0.131, 0.486)	166
$\omega^2_{m_2} \times 10^3$	0.205	(0.110, 0.364)	130
$\omega^2_{m_3} \times 10^3$	0.224	(0.118, 0.408)	251

Table 3.2 reports the summary of the estimation results. Figure 3.2 shows the posterior means of $\exp(h_{j,t}/2)$, square root of the estimated time-varying variances and ρ_t, the dynamic equicorrelation of $\boldsymbol{y}_t$.

The posterior means of μ_j's are similar and the levels of the volatilities are not different from one another. We also find that the posterior means of the diagonal elements of Ω are similar and it is consistent with the observed fluctuations shown in Figure 3.2. The credible intervals of off-diagonal elements of Ω are positive and do not include zero, which means that

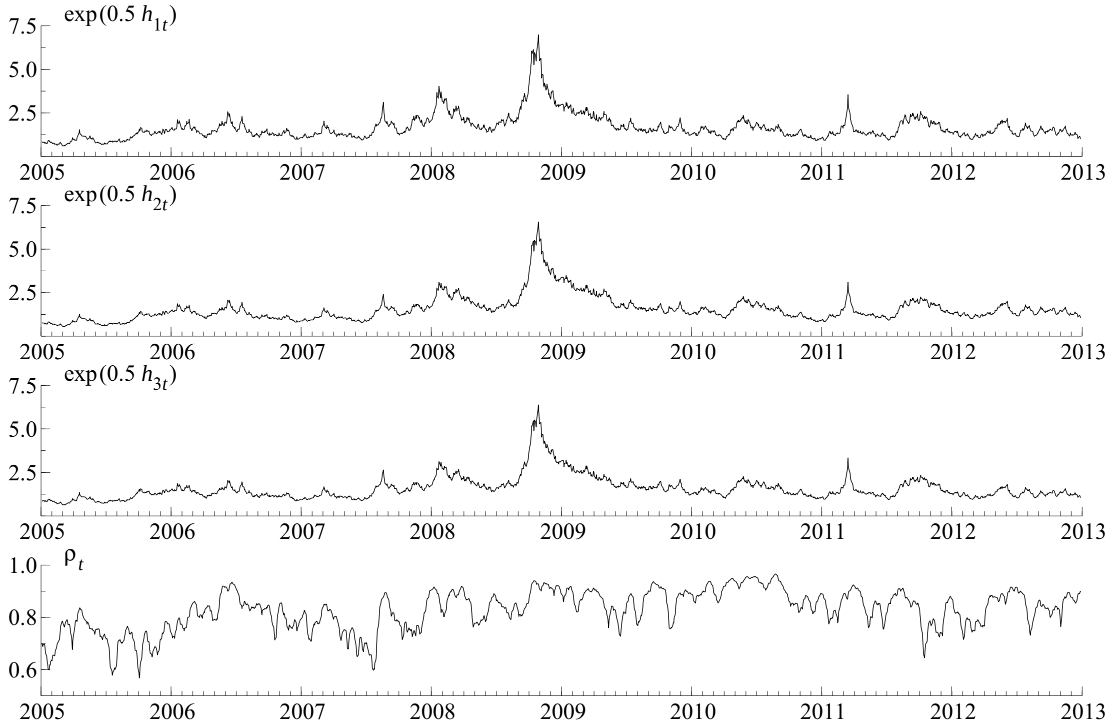

Figure 3.2: Posterior means of square root of the variances and the dynamic equicorrelation.

unobserved volatilities are correlated positively each other.

The posterior means of autoregressive coefficients (ϕ_j's) are very high (over 0.97), which shows that the log volatilities follow highly persistent processes. In addition, the top three panels of Figure 3.2 indicate the comovement of the volatilities. The trajectories sharply increased in September 2008, corresponding to the financial crisis during which Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008). We also observe the increase in March 2011, resulted from Tohoku Region Pacific Coast Earthquake.

The posterior means of autoregressive coefficient (θ) is very high and the equicorrelation parameter is highly persistent, too. The bottom panel of Figure 3.2 shows that the equicorrelation parameter varies at a high level (far from zero) and greatly, which means that it is time-varying and far from constant.

The posterior probability with which Q_{11} is negative is over 0.975 and this is similar for Q_{21} and Q_{31}. The 95% credible intervals of the other elements of Q include zeros. As stated in Section 3.2, Q is the covariance between $\boldsymbol{z}_t$, the transformed error term in the observed equation of our model and $\boldsymbol{\eta}_{t+1}$, the error term in the state equation for $t = 1, \ldots, n-1$. To verify the existence of the cross leverage effects, we should transform Q using R_t and Ω to obtain the posterior means of the (time-varying) correlations between $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\eta}_{t+1}$.

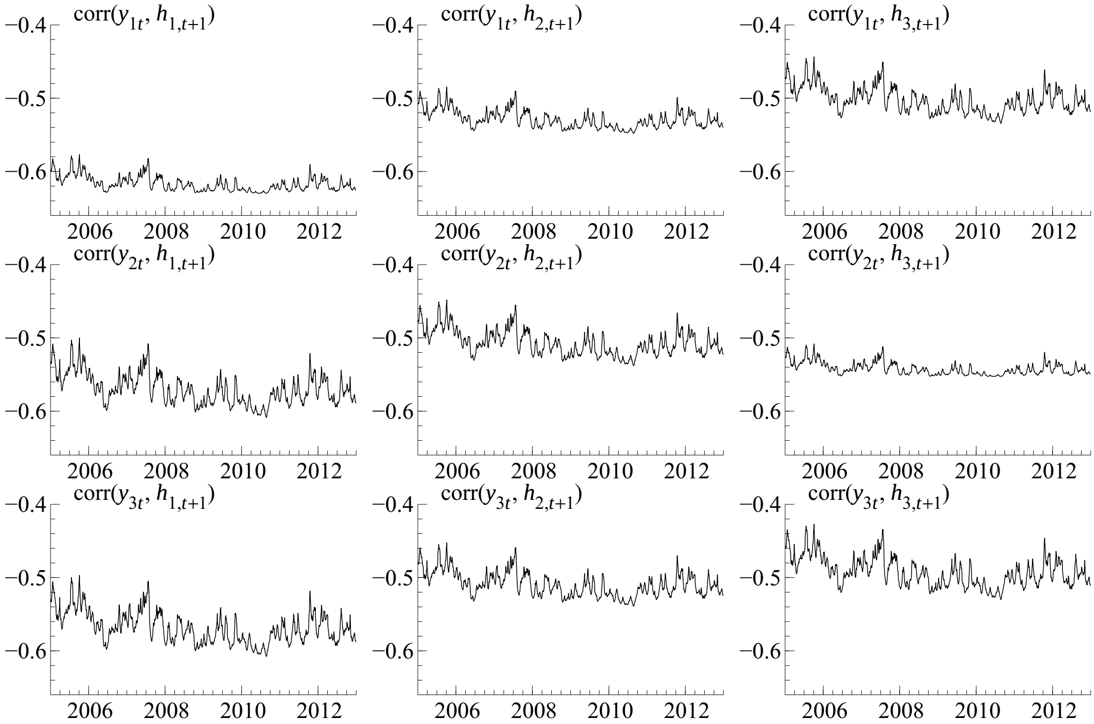

Figure 3.3: Posterior means of correlation of $(y_{it}, h_{j,t+1})$.

Figure 3.3 shows the posterior means of dynamic correlations between the return of i-th asset at time t ($y_{i,t}$) and the j-th log variance at time $t+1$ ($h_{j,t+1}$). Their trajectories are negative and far from zero, which indicates the existence of the cross leverage effect.

We note that the leverage effect of asset 1 is the strongest (the mean is -0.617) and the leverage effect of asset 3 is the weakest (the mean is -0.492) among the cross leverage effects. In addition, the correlation between y_{1t} and $h_{2,t+1}$ is apparently weaker than the correlation between y_{2t} and $h_{1,t+1}$. It indicates that cross leverage effects from asset i to j, $i \neq j$, are not symmetric.

In conclusion, using our multivariate DESV model, we can therefore detect the volatility clustering, the dynamic equicorrelation and the cross leverage effects of the three subindices.

Figure 3.4 shows the posterior means of the expectations of the asset returns. We find

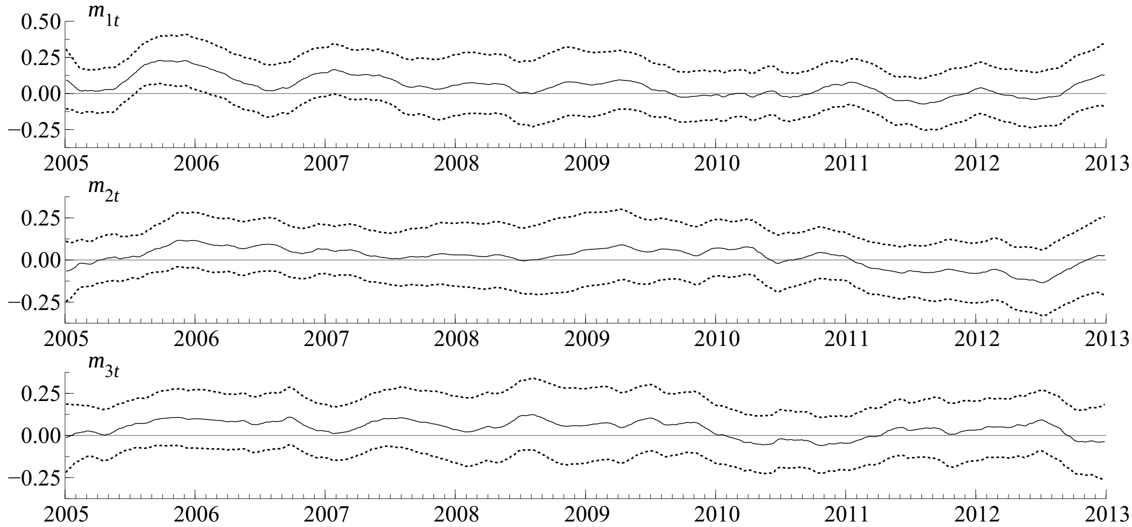

Figure 3.4: Posterior means (solid lines) and 95% credible intervals (between the two outmost dotted lines) of the expectations.

that the 95% credible intervals include zero at almost all of the time points as expected. It seems to fluctuate slowly, which is consistent with the small values of the variances ($\omega_{m_j}^2$'s) reported in Table 3.2.

Note that the acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{g_t\}_{t=1}^n$ in the independent MH algorithms are 0.645 and 0.791, respectively. It indicates that the generated candidates are accepted with relatively high probability and our sampling algorithm works well.

3.5.3 Model comparison

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

Suppose that $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)$ denote, respectively, the conditional mean and covariance of a p-dimensional vector $\boldsymbol{y}_{t+1}$, the asset return at time $t+1$, given $\mathcal{F}_t$, the information at time t. In this study, we make two hedge portfolios: a global minimum variance (GMV) portfolio and a minimum variance (MV) portfolio. The GMV portfolio weights ($\boldsymbol{w}$) are obtained as the solution to the problem:

$$\min_{\boldsymbol{w}} \boldsymbol{w}'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)\boldsymbol{w} \text{ s.t. } \boldsymbol{w}'\mathbf{1}_p = 1. \tag{3.65}$$

We set the MV portfolio weights ($\boldsymbol{w}$) as the solution to the problem:

$$\min_{\boldsymbol{w}} \boldsymbol{w}'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)\boldsymbol{w} \text{ s.t. } \boldsymbol{w}'\mathbf{1}_p = 1 \text{ and } \boldsymbol{w}'\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t) \geq q_0, \tag{3.66}$$

where q_0 is the target value. It indicates that we make the expected returns exceed q_0 for this case. The optimal weights are given by

$$\boldsymbol{w}_{\mathrm{GMV}} = \frac{1}{a}\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathbf{1}_p, \tag{3.67}$$

$$\boldsymbol{w}_{\mathrm{MV}} = \frac{c - q_0 b}{ac - b^2}\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathbf{1}_p + \frac{q_0 a - b}{ac - b^2}\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t), \tag{3.68}$$

where

$$a = \mathbf{1}_p'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathbf{1}_p, \tag{3.69}$$

$$b = \mathbf{1}_p'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t), \tag{3.70}$$

$$c = \mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)'\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)^{-1}\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t). \tag{3.71}$$

We implement the rolling forecast as follows:

1. Estimate the parameters of interest using the data from January 2005 to December 2010. (We set the data as $\{\boldsymbol{y}_t\}_{t=1}^n$.)
2. For the next 3 months including n_1 trading days, i.e., $t = n+1, \ldots, n+n_1-1$,
   (a) use the particle filter (e.g., Doucet, de Freitas, and Gordon (2001)) to compute $\mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)$ and $\mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t)$ numerically (Note that they cannot be obtained analytically). See Appendix 3.A.2 for details.
   (b) compute the two hedge portfolio weights described above and the "realized" returns, $\boldsymbol{w}_{\mathrm{GMV}}'\boldsymbol{y}_{t+1}$, $\boldsymbol{w}_{\mathrm{MV}}'\boldsymbol{y}_{t+1}$.
3. Include the new observations of the next three months to our estimation period and remove the old observations of the first three months. Re-estimate the parameters of interest using the six-year-data (relabeled as $\{\boldsymbol{y}_t\}_{t=1}^n$).
4. Go to 2.

This is repeated until all one-step-ahead forecasts and portfolio choices are conducted through December 2012. In the end, we calculate the standard deviations of the "realized" returns (493 in total). The numerical standard error of the estimate is obtained by repeating the particle filter forty times.

As a benchmark for the model comparison, we also estimate the univariate SV model with leverage effect introduced in Section 3.1. We set the number of particles $M = 100,000$ and the target value $q_0 = -10, 0, 10, 20$ annually.

Table 3.3: Out-of-sample portfolio standard deviations (standard errors in parentheses).

	GMV	MV(-10)	MV(0)	MV(10)	MV(20)
DESV	1.477	1.745	1.835	2.094	2.446
	(0.001)	(0.053)	(0.072)	(0.115)	(0.154)
univariate	1.484	2.581	3.158	3.935	4.756
	(0.000)	(0.062)	(0.109)	(0.153)	(0.194)

Table 3.3 shows the out-of-sample portfolio standard deviations using the six-year rolling estimation window. The prior distributions for each estimation are the same as those of the previous subsection. For each of the hedging strategies, the standard deviation based on our multivariate model is smaller than that of the univariate model. We note that MV portfolio strategy with $q_0 = 20$ (annually) makes the biggest difference between the two and GMV portfolio strategy makes the smallest difference between the two. Thus, our proposed model with the time-varying covariance structure show good out-of-sample forecasting performances with respect to dynamic GMV and MV portfolios. It suggests that for asset returns with time-varying variances we should model the covariance structure between the asset returns and the one-step-ahead variances including the dynamic correlations between the asset returns.

3.6 Conclusion

This article proposed the novel multivariate stochastic volatility model with dynamic equicorrelation and cross leverage effect. We took a Bayesian approach and described the efficient MCMC algorithm by dividing the latent variables of our nonlinear model into several blocks and approximating them to those of linear Gaussian state-space models. Its sampling efficiency is illustrated using simulated data in comparison with the single-move sampler.

An empirical study is provided using industrial sector index data of TOPIX. We find the persistence in the volatilities and equicorrelations and the existence of strong cross leverage effects. In comparison with the univariate model, our DESV model achieves efficient out-of-sample portfolios with significantly lower variances.

Appendix

3.A.1 Computations for the block-sampler

Noting that

$$\frac{\partial R_t^{-1/2}}{\partial g_t} = -\frac{1}{2}\mathrm{diag}\begin{bmatrix}(p-1)\{1+(p-1)\rho_t\}^{-3/2}(1-\rho_t)\rho_t \\ -(1-\rho_t)^{-1/2}\rho_t\cdot\mathbf{1}_{p-1}\end{bmatrix}\cdot R_O', \tag{3.72}$$

$$\begin{aligned}&\frac{\partial^2 R_t^{-1/2}}{\partial g_t^2}\\ &= -\frac{1}{2}\mathrm{diag}\begin{bmatrix}-\frac{3}{2}(p-1)^2\{1+(p-1)\rho_t\}^{-5/2}(1-\rho_t)\rho_t+(p-1)\{1+(p-1)\rho_t\}^{-3/2}(1-2\rho_t) \\ \{-\frac{1}{2}(1-\rho_t)^{-3/2}\rho_t-(1-\rho_t)^{-1/2}\}\cdot\mathbf{1}_{p-1}\end{bmatrix}\cdot R_O',\end{aligned} \tag{3.73}$$

$$\frac{\partial V_t^{-1/2}}{\partial h_{j_1 t}} = \mathrm{diag}\left\{\mathbf{0}', -\frac{1}{2}\exp\left(-\frac{1}{2}h_{j_1 t}\right), \mathbf{0}'\right\},\ \ j_1 = 1,\ldots,p, \tag{3.74}$$

$$\frac{\partial^2 V_t^{-1/2}}{\partial h_{j_1 t}^2} = \mathrm{diag}\left\{\mathbf{0}', \frac{1}{4}\exp\left(-\frac{1}{2}h_{j_1 t}\right), \mathbf{0}'\right\},\ \ j_1 = 1,\ldots,p, \tag{3.75}$$

and

$$\frac{\partial^2 V_t^{-1/2}}{\partial h_{j_1 t}\partial h_{j_2 t}} = O,\ \ j_1 = 1,\ldots,p,\ j_2 = 1,\ldots,p,\ j_1\neq j_2, \tag{3.76}$$

we compute $\boldsymbol{d}_t, A_t, B_t$ as below.

3.A.1.1 Computations for $\{\boldsymbol{h}_t\}_{t=1}^n$

Derivation of $\boldsymbol{d}_t,\ t = s+1,\ldots,s+r.$ We can obtain

$$(\boldsymbol{d}_t)_{j_1} = \frac{\partial l_t}{\partial h_{j_1 t}} + \frac{\partial l_{t-1}}{\partial h_{j_1 t}},\quad j_1 = 1,\ldots,p, \tag{3.77}$$

where

$$\begin{aligned}\frac{\partial l_t}{\partial h_{j_1 t}} = &-\left\{R_t^{-\frac{1}{2}}\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1 t}}(\boldsymbol{y}_t-\boldsymbol{m}_t)+Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1 t}}\right\}'(I_p-Q'\Omega^{-1}Q)^{-1}\\ &\cdot\left[R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}(\boldsymbol{y}_t-\boldsymbol{m}_t)-Q'\Omega^{-1}\{\boldsymbol{h}_{t+1}-\boldsymbol{\mu}-\Phi(\boldsymbol{h}_t-\boldsymbol{\mu})\}\right]-\frac{1}{2},\end{aligned} \tag{3.78}$$

$$\frac{\partial l_{t-1}}{\partial h_{j_1t}} = \left(Q'\Omega^{-1}\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\right)'(I_p - Q'\Omega^{-1}Q)^{-1}$$
$$\cdot\left[R_{t-1}^{-\frac{1}{2}}V_{t-1}^{-\frac{1}{2}}(\boldsymbol{y}_{t-1} - \boldsymbol{m}_{t-1}) - Q'\Omega^{-1}\{\boldsymbol{h}_t - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_{t-1} - \boldsymbol{\mu})\}\right]. \tag{3.79}$$

Derivation of $A_t,\ t = s+1,\dots,s+r.$ We can obtain

$$(A_t)_{[j_1,j_1]} = -\mathrm{E}\left(\frac{\partial^2 l_t}{\partial h_{j_1t}^2}\right) - \frac{\partial^2 l_{t-1}}{\partial h_{j_1t}^2},\quad j_1 = 1,\dots,p, \tag{3.80}$$

$$(A_t)_{[j_1,j_2]} = -\mathrm{E}\left(\frac{\partial^2 l_t}{\partial h_{j_1t}\partial h_{j_2t}}\right) - \frac{\partial^2 l_{t-1}}{\partial h_{j_1t}\partial h_{j_2t}},\quad j_1 = 1,\dots,p,\ ,j_2 = 1,\dots,p,\ j_1 \neq j_2, \tag{3.81}$$

where

$$\begin{aligned}
&\mathrm{E}(\partial^2 l_t/\partial h_{j_1t}^2)\\
&= -\mathrm{tr}\Bigg[\Bigg\{\frac{\partial^2 V_t^{-\frac{1}{2}}}{\partial h_{j_1t}^2}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}\\
&\qquad + \frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1t}}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}R_t^{-\frac{1}{2}}\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1t}}\Bigg\}\mathrm{E}\{(\boldsymbol{y}_t - \boldsymbol{m}_t)(\boldsymbol{y}_t - \boldsymbol{m}_t)'\}\Bigg]\\
&\quad - \mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t)'\Bigg[-\frac{\partial^2 V_t^{-\frac{1}{2}}}{\partial h_{j_1t}^2}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\\
&\qquad + 2\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1t}}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\Bigg]\\
&\quad - \left(\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\right)'\Phi\Omega^{-1}Q(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}},
\end{aligned} \tag{3.82}$$

$$\begin{aligned}
&\mathrm{E}(\partial^2 l_t/\partial h_{j_1t}\partial h_{j_2t})\\
&= -\mathrm{tr}\Bigg\{\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1t}}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}R_t^{-\frac{1}{2}}\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_2t}}\cdot\mathrm{E}\{(\boldsymbol{y}_t - \boldsymbol{m}_t)(\boldsymbol{y}_t - \boldsymbol{m}_t)'\}\Bigg\}\\
&\quad - \mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t)'\Bigg\{\frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_1t}}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_2jt}}\\
&\qquad + \frac{\partial V_t^{-\frac{1}{2}}}{\partial h_{j_2t}}R_t^{-\frac{1}{2}'}(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\Bigg\}\\
&\quad - \left(\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\right)'\Phi\Omega^{-1}Q(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Phi\frac{\partial \boldsymbol{h}_t}{\partial h_{j_2t}},
\end{aligned} \tag{3.83}$$

$$\frac{\partial^2 l_{t-1}}{\partial h_{j_1t}\partial h_{j_2t}} = -\left(Q'\Omega^{-1}\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1t}}\right)'(I_p - Q'\Omega^{-1}Q)^{-1}\left(Q'\Omega^{-1}\frac{\partial \boldsymbol{h}_t}{\partial h_{j_2t}}\right). \tag{3.84}$$

We note that

$$\boldsymbol{z}_t = R_t^{-1/2} V_t^{-1/2}(\boldsymbol{y}_t - \boldsymbol{m}_t) \sim \mathrm{N}(Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}, I_p - Q'\Omega^{-1}Q), \tag{3.85}$$

and hence that

$$\mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t) = V_t^{1/2} R_t^{1/2} Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}, \tag{3.86}$$

$$\mathrm{Var}(\boldsymbol{y}_t - \boldsymbol{m}_t) = V_t^{1/2} R_t^{1/2} (I_p - Q'\Omega^{-1}Q)(R_t^{1/2})' V_t^{1/2}, \tag{3.87}$$

$$\mathrm{E}\{(\boldsymbol{y}_t - \boldsymbol{m}_t)(\boldsymbol{y}_t - \boldsymbol{m}_t)'\} = \mathrm{Var}(\boldsymbol{y}_t - \boldsymbol{m}_t) + \mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t)\mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t)'. \tag{3.88}$$

Derivation of $B_t,\ t = s+2, \ldots, s+r.$ We can obtain

$$(B_t)_{[j_1, j_2]} = -\mathrm{E}\left(\frac{\partial^2 l_{t-1}}{\partial h_{j_1 t} \partial h_{j_2, t-1}}\right), \quad j_1 = 1, \ldots, p,\ j_2 = 1, \ldots, p, \tag{3.89}$$

where

$$\begin{aligned} \mathrm{E}\left(\frac{\partial^2 l_{t-1}}{\partial h_{j_1 t} \partial h_{j_2, t-1}}\right) &= \left(Q'\Omega^{-1}\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1 t}}\right)' (I_p - Q'\Omega^{-1}Q)^{-1} R_{t-1}^{-\frac{1}{2}} \frac{\partial V_{t-1}^{-\frac{1}{2}}}{\partial h_{j_2, t-1}} \mathrm{E}(\boldsymbol{y}_{t-1} - \boldsymbol{m}_{t-1}) \\ &\quad + \left(Q'\Omega^{-1}\frac{\partial \boldsymbol{h}_t}{\partial h_{j_1 t}}\right)' (I_p - Q'\Omega^{-1}Q)^{-1} \left(Q'\Omega^{-1}\Phi \frac{\partial \boldsymbol{h}_{t-1}}{\partial h_{j_2, t-1}}\right). \end{aligned} \tag{3.90}$$

3.A.1.2 Computations for $\{g_t\}_{t=1}^n$

Derivation of $\boldsymbol{d}_t,\ t = s+1, \ldots, s+r.$ We can obtain

$$\boldsymbol{d}_t = \frac{\partial l_t}{\partial g_t}, \tag{3.91}$$

where

$$\begin{aligned} \frac{\partial l_t}{\partial g_t} &= -\left\{\frac{\partial R_t^{-\frac{1}{2}}}{\partial g_t} \cdot V_t^{-\frac{1}{2}} (\boldsymbol{y}_t - \boldsymbol{m}_t)\right\}' (I_p - Q'\Omega^{-1}Q)^{-1} \\ &\quad \cdot \left[R_t^{-\frac{1}{2}} V_t^{-\frac{1}{2}} (\boldsymbol{y}_t - \boldsymbol{m}_t) - Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\right] + \frac{1}{2} p(p-1)\rho_t^2 \{1 + (p-1)\rho_t\}^{-1}. \end{aligned} \tag{3.92}$$

Derivation of $A_t,\ t = s+1, \ldots, s+r.$ We can obtain

$$A_t = -\mathrm{E}\left(\frac{\partial^2 l_t}{\partial g_t^2}\right), \tag{3.93}$$

where

$$
\begin{aligned}
&\mathrm{E}(\partial^2 l_t/\partial g_t^2)\\
&= -\mathrm{tr}\left[\left\{V_t^{-\frac{1}{2}}\left(\frac{\partial^2 R_t^{-\frac{1}{2}}}{\partial g_t^2}\right)'(I_p - Q'\Omega^{-1}Q)^{-1}R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}\right.\right.\\
&\qquad \left.\left. + V_t^{-\frac{1}{2}}\left(\frac{\partial R_t^{-\frac{1}{2}}}{\partial g_t}\right)'(I_p - Q'\Omega^{-1}Q)^{-1}\frac{\partial R_t^{-\frac{1}{2}}}{\partial g_t}V_t^{-\frac{1}{2}}\right\}\mathrm{E}\{(\boldsymbol{y}_t - \boldsymbol{m}_t)(\boldsymbol{y}_t - \boldsymbol{m}_t)'\}\right]\\
&\quad + \mathrm{E}(\boldsymbol{y}_t - \boldsymbol{m}_t)'V_t^{-\frac{1}{2}}\left(\frac{\partial^2 R_t^{-\frac{1}{2}}}{\partial g_t^2}\right)'(I_p - Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\\
&\quad + \frac{1}{2}p(p-1)\rho_t^2(1-\rho_t)\{1+(p-1)\rho_t\}^{-2}\{2+(p-1)\rho_t\}. \qquad (3.94)
\end{aligned}
$$

Derivation of $B_t,\ t = s+2, \ldots, s+r$. We can obtain $B_t = 0$.

3.A.2 Particle filter

Let $f(\boldsymbol{h}_t|Y_t, \boldsymbol{\vartheta})$ denote the density function of $\boldsymbol{h}_t$ given $(Y_t, \boldsymbol{\vartheta})$ where $Y_t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$, and let $\hat{f}(\boldsymbol{h}_t|Y_t, \boldsymbol{\vartheta})$ denote the discrete approximation to $f(\boldsymbol{h}_t|Y_t, \boldsymbol{\vartheta})$.

We draw M samples from the conditional joint distribution of $(\boldsymbol{h}_{t+1}, \boldsymbol{h}_t, g_{t+1}, g_t, \boldsymbol{m}_{t+1}, \boldsymbol{m}_t)$ given $(Y_{t+1}, \boldsymbol{\vartheta})$ with the density

$$
\begin{aligned}
&f(\boldsymbol{h}_{t+1}, \boldsymbol{h}_t, g_{t+1}, g_t, \boldsymbol{m}_{t+1}, \boldsymbol{m}_t|Y_{t+1}, \boldsymbol{\vartheta})\\
&\propto f(\boldsymbol{y}_{t+1}|\boldsymbol{h}_{t+1}, g_{t+1}, \boldsymbol{m}_{t+1}, Y_t, \boldsymbol{\vartheta})f(\boldsymbol{h}_{t+1}|\boldsymbol{h}_t, g_t, \boldsymbol{m}_t, Y_t, \boldsymbol{\vartheta})f(g_{t+1}|g_t, \boldsymbol{\vartheta})f(\boldsymbol{m}_{t+1}|\boldsymbol{m}_t, \boldsymbol{\vartheta})\\
&\quad \times f(\boldsymbol{h}_t, g_t, \boldsymbol{m}_t|Y_t, \boldsymbol{\vartheta}). \qquad (3.95)
\end{aligned}
$$

We implement the particle filter:

1. (a) Generate

$$
\boldsymbol{h}_1^{(i)} \sim \mathrm{N}_p(\boldsymbol{\mu}_{\boldsymbol{h}_1}, \Sigma_{\boldsymbol{h}_1}), \quad g_1^{(i)} \sim \mathrm{N}(\mu_{g_1}, \sigma_{g_1}^2), \quad \boldsymbol{m}_1^{(i)} \sim \mathrm{N}_p(\boldsymbol{\mu}_{\boldsymbol{m}_1}, \Sigma_{\boldsymbol{m}_1}), \quad i = 1, \ldots, M,
$$

where $\boldsymbol{\mu}_{\boldsymbol{h}_1}, \boldsymbol{\mu}_{\boldsymbol{m}_1}$ are some constant vectors, $\Sigma_{\boldsymbol{h}_1}, \Sigma_{\boldsymbol{m}_1}$ are some constant positive-definite matrices, μ_{g_1} is some constant and $\sigma_{g_1}^2$ is some positive constant (we adopt the posterior mean vectors of $\boldsymbol{h}_1, \boldsymbol{m}_1$, the posterior covariance matrices of $\boldsymbol{h}_1, \boldsymbol{m}_1$, the posterior mean of g_1 and the posterior variance of g_1), respectively.

(b) Compute

$$\pi_i = \frac{\tilde{\pi}_i}{\sum_{j=1}^{M} \tilde{\pi}_i}, \quad \tilde{\pi}_i = \frac{f(\boldsymbol{y}_1|\boldsymbol{h}_1, g_1, \boldsymbol{m}_1, \boldsymbol{\vartheta}) f(\boldsymbol{h}_1, g_1, \boldsymbol{m}_1|\boldsymbol{\vartheta})}{g(\boldsymbol{h}_1, g_1, \boldsymbol{m}_1|\boldsymbol{\vartheta})}, \tag{3.96}$$

where $g(\cdot)$ is a density generating $\boldsymbol{h}_1^{(i)}, g_t^{(i)}$ and $\boldsymbol{m}_1^{(i)}$.

(c) Set $\hat{f}(\boldsymbol{h}_1^{(i)}, g_1^{(i)}, \boldsymbol{m}_1^{(i)}|Y_1, \boldsymbol{\theta}) = \pi_i$.

2. For $t = 1, \ldots, n + n_1 - 1$,

(a) generate $\boldsymbol{h}_t^{(i)}, g_t^{(i)}, \boldsymbol{m}_t^{(i)} \sim \hat{f}(\boldsymbol{h}_t, g_t, \boldsymbol{m}_t|Y_t, \boldsymbol{\vartheta})$.

(b) generate $\boldsymbol{h}_{t+1}^{(i)} \sim f(\boldsymbol{h}_{t+1}|\boldsymbol{h}_t^{(i)}, g_t^{(i)}, \boldsymbol{m}_t^{(i)}, Y_t, \boldsymbol{\vartheta})$, $g_{t+1}^{(i)} \sim f(g_{t+1}|g_t^{(i)}, \boldsymbol{\vartheta})$, $\boldsymbol{m}_{t+1}^{(i)} \sim f(\boldsymbol{m}_{t+1}|\boldsymbol{m}_t^{(i)}, \boldsymbol{\vartheta})$.

(c) compute

$$\pi_i = \frac{\tilde{\pi}_i}{\sum_{j=1}^{M} \tilde{\pi}_j}, \quad \tilde{\pi}_i = f(\boldsymbol{y}_{t+1}|\boldsymbol{h}_{t+1}^{(i)}, g_{t+1}^{(i)}, \boldsymbol{m}_{t+1}^{(i)}, Y_t, \boldsymbol{\vartheta}). \tag{3.97}$$

(d) set $\hat{f}(\boldsymbol{h}_{t+1}^{(i)}, g_{t+1}^{(i)}, \boldsymbol{m}_{t+1}^{(i)}|Y_{t+1}, \boldsymbol{\vartheta}) = \pi_i$.

Note that $\boldsymbol{y}_{t+1}|\boldsymbol{h}_{t+1}, g_{t+1}, \boldsymbol{m}_{t+1}, \sim \mathrm{N}_p(\boldsymbol{m}_{t+1}, V_{t+1}^{1/2} R_{t+1} V_{t+1}^{1/2})$, $t = 1, \ldots, n + n_1 - 1$. For $t = n, \ldots, n + n_1 - 1$, we obtain

$$\hat{\boldsymbol{\mu}}_{t+1|t} = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{m}_{t+1}^{(i)} \to \mathrm{E}(\boldsymbol{y}_{t+1}|\mathcal{F}_t), \tag{3.98}$$

$$\hat{\Sigma}_{t+1|t} = \frac{1}{M} \sum_{i=1}^{M} V_{t+1}^{1/2(i)} R_{t+1}^{(i)} V_{t+1}^{1/2(i)} + \Omega_m \to \mathrm{Var}(\boldsymbol{y}_{t+1}|\mathcal{F}_t). \tag{3.99}$$

Chapter 4

Realized stochastic volatility with dynamic equicorrelation and cross leverage

4.1 Introduction

Recently a great deal of attention has been attracted to the statistical modeling with the realized measure as a measure of latent stochastic variables, e.g., the variance of the dependent variable, by using high-frequency data. This is partly because a large number of such high-frequency data has become available in financial markets. The most commonly used measure of this type is the realized variance, defined as the sum of the squared intraday returns over a specified time interval such as a day. Early studies, e.g., Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2001) showed that this measure provide a consistent estimator of the latent stochastic variance of the log-price process as the sampling frequency diverges under the ideal market assumption. Barndorff-Nielsen and Shephard (2004) introduced the realized covariance matrix as the sum of the outer product of intraday returns over a specified period and showed that it is a consistent estimator of the quadratic cross-variance of the multivariate log-price process under the ideal market assumption. The high-frequency data, however, has three major problems in the real market, which results in a bias in the realized measure. First, high-frequency data is influenced by the market microstructure noise and daily realized measures have the larger bias for the higher sampling frequency. (e.g., Hansen and Lunde (2006) and Ubukata and Oya (2009)). It is known that the market microstructure noise is caused mainly by (i) the presence of bid-ask spread, (ii) discrete trading and (iii) price adjustment (e.g., O'Hara (1995) and Hasbrouck (2007)). Second, there are non-trading hours when we cannot obtain high-frequency data, that is, the real markets are open in the limited hours in a day and in the limited days in a week. If we calculate, e.g., the realized variance as the sum of the squared intraday returns when the market is open, we may underestimate the latent one-day variance (Hansen and Lunde (2005)). The third is inherent in multivariate financial data: non-synchronous trading problem. Any two assets rarely traded simultaneously and the two transaction prices are not recorded at the same time, which affects the estimation of the covariates or correlations. Epps (1979) reported that those estimation have a tendency to diminish as the sampling frequency increases and this phenomenon is known as Epps effect.

To overcome these difficulties, a vast literature is devoted to improving the estimator (for the multivariate case, e.g., Malliavin and Mancino (2002), Hayashi and Yoshida (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Kunitomo and Sato (2008) and Zhang (2011)). Whereas we cannot ignore the contamination of the financial data obtained at short intervals as stated above, the daily returns are not heavily affected by such noise. Therefore the daily returns may provide us additional information to eliminate the bias contained in the realized measure. Takahashi, Watanabe, and Omori (2009) extended a univariate stochastic volatility (SV) model and proposed simultaneous modeling of the daily returns and realized measure known as realized stochastic volatility model or RSV model (Dobrev and Szerszen (2010) and Koopman and Scharth (2012)). Hansen, Huang, and Shek (2012) introduced a framework that combines a GARCH structure for returns with an integrated model for realized measures and named their models as realized GARCH models. Since they demonstrated their contemporaneous models offer a substantial improvement in understanding the financial time series structure, we adopt their approach and incorporate the SV model with the information of realized measures.

The simple univariate SV model with leverage effect is given in a state space form as:

$$y_t = m + \exp(h_t/2)\epsilon_t, \quad t = 1, \ldots, n, \tag{4.1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, n-1, \tag{4.2}$$

$$h_1 \sim \mathrm{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right), \tag{4.3}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & q \\ q & \sigma_\eta^2 \end{pmatrix} \right), \tag{4.4}$$

$$|\phi| < 1, \tag{4.5}$$

where y_t denotes a (univariate) asset return, h_t is a log-variance of y_t. In volatilities of stock returns, we often observe the asymmetry or the cross leverage effect, which implies a decrease in the i-th dependent variable at date t followed by an increase in the j-th latent stochastic variance at date $(t+1)$. The negative value of q implies the existence of the leverage effect. If ϕ takes a value close to 1, it indicates the existence of the volatility clustering frequently found in the financial time series. It can be extended to the multivariate SV model with cross leverage effect; e.g., Chaper 3 proposes the multivariate SV model with dynamic equicorrelation and cross leverage effect (DESV model), which is assumed that all correlations equal but time-varying. Studies of the dynamic equicorrelation (DECO) model, which is based on a Dynamic conditional correlation (DCC) model, include Engle and Kelly (2012) and Clements, Coleman-Fenn, and Smith (2011). Making reference to Elton and Gruber (1973), they argued that the dynamic equicorrelation assumption gives a superior portfolio allocation.

In this study, we extend the DESV model and RSV model to the multivariate model by introducing the multivariate realized SV model with (i) dynamic equicorrelation and (ii) cross leverage effect. Because it is difficult to evaluate the likelihood function using the high-dimensional numerical integration, Bayesian efficient estimation methods have been proposed (e.g., Shephard and Pitt (1997), Watanabe and Omori (2004) and Omori and Watanabe (2008)). Chapter 3 also describes an efficient Bayesian estimation using the Markov chain Monte Carlo (MCMC) method to generate the latent stochastic variances and dynamic equicorrelation factor from the posterior distributions. It is highly efficient compared with the benchmarking simple method, single-move sampling method, which draws a single latent variable at a time given the other latent variables and the parameters. It means that we only need to generate a fewer number of MCMC samples to estimate the posterior distribution of the interested parameters. With the additional information to estimate unobserved variables by using the realized measures, we do not need to require the efficient (but complicated) estimation method and conduct the simple estimation in this study.

The rest of this article is organized as follows. In Section 4.2, we propose the multivariate realized stochastic volatility model with dynamic equicorrelation and cross leverage effect. Section 4.3 describes an efficient Bayesian estimation method for the proposed model. In Section 4.4, we illustrate our estimation method using simulated data and shows that our

MCMC algorithm works well. Section 4.5 applies our proposed multivariate RSV model to the multivariate asset return data and the corresponding realized measure. Section 4.6 concludes this article.

4.2 Multivariate realized SV model with dynamic equicorrelation and cross leverage

Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})'$ denote a p-dimensional asset return vector at time t ($t = 1, \ldots, n$). We set $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{pt})'$ be a p-dimensional vector of the logarithm of the realized measure of the variances and $RCor_t$ be the realized measure of the correlation matrix corresponding to $\boldsymbol{y}_t$, respectively. Let $\boldsymbol{x}^*_t = (x^*_{1,t}, \ldots, x^*_{p(p-1)/2,t})'$, denote a $p(p-1)/2$-dimensional vector, where

$$x^*_{jt} = \log\left(\frac{\rho^*_{jt}}{1-\rho^*_{jt}}\right), \quad \boldsymbol{\rho}^*_t = (\rho^*_{1,t}, \ldots, \rho^*_{p(p-1)/2,t})',$$

and where $\boldsymbol{\rho}^*_t$ is made by extracting and vectorizing the lower off-diagonal elements of $RCor_t$. Let $\boldsymbol{m}_t = (m_{1t}, \ldots, m_{pt})'$ and $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{pt})'$ denote p-dimensional vectors of unobserved variables and g_t an unobserved variable. We consider the multivariate RSV model given by

$$\boldsymbol{y}_t = \boldsymbol{m}_t + V_t^{1/2}\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_p(\boldsymbol{0}_p, R_t), \quad t = 1, ..., n, \tag{4.6}$$

$$\boldsymbol{x}_t = \boldsymbol{\xi} + \boldsymbol{h}_t + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim \mathrm{N}(\boldsymbol{0}_p, \Delta), \quad t = 1, ..., n, \tag{4.7}$$

$$\boldsymbol{x}^*_t = \boldsymbol{\xi}^* + \boldsymbol{\lambda}^* g_t + \boldsymbol{w}^*_t, \quad \boldsymbol{w}^*_t \sim \mathrm{N}_{p(p-1)/2}(\boldsymbol{0}_{p(p-1)/2}, \Delta^*), \quad t = 1, ..., n, \tag{4.8}$$

$$\boldsymbol{h}_{t+1} = \boldsymbol{\mu} + \Phi(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathrm{N}_p(\boldsymbol{0}_p, \Omega), \quad t = 1, ..., n-1, \tag{4.9}$$

$$\boldsymbol{h}_1 = \boldsymbol{\mu} + \boldsymbol{\eta}_0, \quad \boldsymbol{\eta}_0 \sim \mathrm{N}_p(\boldsymbol{0}_p, \Omega_0), \tag{4.10}$$

$$g_{t+1} = \gamma + \theta(g_t - \gamma) + \zeta_t, \quad \zeta_t \sim \mathrm{N}(0, \sigma^2), \quad t = 1, ..., n-1, \tag{4.11}$$

$$g_1 = \gamma + \zeta_0, \quad \zeta_0 \sim \mathrm{N}(0, \sigma_0^2), \tag{4.12}$$

$$\boldsymbol{m}_{t+1} = \boldsymbol{m}_t + \boldsymbol{\eta}_{\boldsymbol{m}t}, \quad \boldsymbol{\eta}_{\boldsymbol{m}t} \sim \mathrm{N}_p(\boldsymbol{0}_p, \Omega_{\boldsymbol{m}}), \quad t = 1, ..., n-1, \tag{4.13}$$

$$\boldsymbol{m}_1 = \boldsymbol{\eta}_{\boldsymbol{m}0}, \quad \boldsymbol{\eta}_{\boldsymbol{m}0} \sim \mathrm{N}_p(\boldsymbol{0}_p, \kappa I_p), \tag{4.14}$$

where $\boldsymbol{0}_p$ is a p-dimensional zero vector,

$$V_t = \mathrm{diag}\{\exp(h_{1t}), \ldots, \exp(h_{pt})\}, \; t = 1, \ldots, n, \tag{4.15}$$

$$R_t = (1-\rho_t)I_p + \rho_t J_p, \; t = 1, \ldots, n, \tag{4.16}$$

$$\rho_t = \frac{\exp(g_t)}{\exp(g_t)+1}, \; t = 1, \ldots, n, \tag{4.17}$$

I_p is a unit matrix of size p, $J_{p_1p_2} = \mathbf{1}_{p_1}\mathbf{1}_{p_2}'$ ($\mathbf{1}_p$ denotes a p-dimensional vector with all elements equal to one),

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)', \quad \boldsymbol{\xi}^* = (\xi_1^*, \ldots, \xi_{p(p-1)/2}^*)', \tag{4.18}$$

$$\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_{p(p-1)/2}^*)', \tag{4.19}$$

$$\Delta = \mathrm{diag}(\delta_1^2, \ldots, \delta_p^2), \quad \Delta^* = \mathrm{diag}(\delta_1^{*2}, \ldots, \delta_{p(p-1)/2}^{*2}), \tag{4.20}$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)', \quad \Phi = \mathrm{diag}(\phi_1, \ldots, \phi_p), \tag{4.21}$$

$$\Omega_{\boldsymbol{m}} = \mathrm{diag}(\omega_{m_1}^2, \ldots, \omega_{m_p}^2), \tag{4.22}$$

$$\begin{pmatrix} R_t^{-\frac{1}{2}}\boldsymbol{\epsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \mathrm{N}_{2p}(\mathbf{0}_{2p}, \Psi), \ t = 1, \ldots, n-1, \tag{4.23}$$

$$\Psi = \begin{pmatrix} I_p & Q' \\ Q & \Omega \end{pmatrix}, \tag{4.24}$$

and Ω_0, the covariance matrix of the initial latent variable $\boldsymbol{h}_1$, satisfies the stationary condition $\Omega_0 = \Phi\Omega_0\Phi + \Omega$ such that

$$\mathrm{vec}(\Omega_0) = (I_{p^2} - \Phi \otimes \Phi)^{-1}\mathrm{vec}(\Omega). \tag{4.25}$$

We also set $\mathrm{Var}(g_1) = \sigma^2/(1-\theta^2)$, the variance of the initial latent variable g_1, for assuming the stationary condition and set κ, the variance of the initial latent variable $m_{i,1}$ ($i = 1, \ldots, p$), equal to some large known constant. The latent vector, $\boldsymbol{h}_t$, is a vector of log-variances of the returns, and the latent variable, g_t, is the transformed equicorrelation of $\boldsymbol{y}_t$. For the identifiability, we set the diagonal elements of the covariance matrix of $\boldsymbol{\epsilon}_t$ equal to 1.

Notice that we define ρ_t, $t = 1, \ldots n$, so as to take values on the unit interval, $(0, 1)$. As shown in the previous chapter[1], the equicorrelation matrix R_t is positive definite if and only if ρ_t is in $(-(p-1)^{-1}, 1)$. It means that as p becomes large, the negative region of the parameter space becomes smaller. Therefore it is reasonable to restrict the parameter space of ρ_t to the positive region so that the parameter space is independent of p.

We assume, for simplicity, that $\{\boldsymbol{m}_t\}_{t=1}^n$, the latent sequence of the expectations of $\boldsymbol{y}_t$, $t = 1, \ldots, n$, follows a simple random walk.

The parameter $\boldsymbol{\xi}$ corrects the bias due to the market microstructure noise and the effect of non-trading hours. If the elements of $\boldsymbol{\xi}$ are positive, the realized measure of the variance has an upward bias that may be due to the market microstructure noise and if the elements

[1] See Section 3.2.1 for the details of R_t.

of $\boldsymbol{\xi}$ are negative, it has a downward bias due to the existence of non-trading hours. When the mean of $\boldsymbol{\xi}^*$ is not equal to $\mathbf{0}_{p(p-1)/2}$ or the mean of $\boldsymbol{\lambda}^*$ is not equal to $\mathbf{1}_{p(p-1)/2}$, the realized measure of the correlation has a bias due to the market microstructure noise or the non-synchronous trading effect.

4.3 Bayesian estimation

4.3.1 Priors and posterior densities

For prior distributions of $(\boldsymbol{\xi}, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^*, \Delta, \Delta^*, \boldsymbol{\mu}, \boldsymbol{\gamma}, \Phi, \theta, \sigma^2, \Omega_{\boldsymbol{m}})$, we assume that

$$\boldsymbol{\xi} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\xi}0}, S_{\boldsymbol{\xi}0}), \tag{4.26}$$

$$\boldsymbol{\xi}^* \sim \mathrm{N}_{p(p-1)/2}(\boldsymbol{m}^*_{\boldsymbol{\xi}0}, S^*_{\boldsymbol{\xi}0}), \tag{4.27}$$

$$\boldsymbol{\lambda}^* \sim \mathrm{N}_{p(p-1)/2}(\boldsymbol{m}^*_{\boldsymbol{\lambda}0}, S^*_{\boldsymbol{\lambda}0}), \tag{4.28}$$

$$\delta_i^2 \sim \mathrm{IG}(\alpha_{\delta_i 0}/2, \beta_{\delta_i 0}/2), \quad i = 1, \ldots, p, \tag{4.29}$$

$$\delta_j^{*2} \sim \mathrm{IG}(\alpha^*_{\delta_j 0}/2, \beta^*_{\delta_j 0}/2), \quad j = 1, \ldots, p(p-1)/2 \tag{4.30}$$

$$\boldsymbol{\mu} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\mu}0}, S_{\boldsymbol{\mu}0}), \tag{4.31}$$

$$\gamma \sim \mathrm{N}(\boldsymbol{m}_{\gamma 0}, s^2_{\gamma 0}), \tag{4.32}$$

$$(\phi_i + 1)/2 \sim \mathrm{Be}(a_{\phi_i}, b_{\phi_i}), \quad i = 1, \ldots, p, \tag{4.33}$$

$$(\theta + 1)/2 \sim \mathrm{Be}(a_\theta, b_\theta), \tag{4.34}$$

$$\sigma^2 \sim \mathrm{IG}(\alpha_{\sigma 0}/2, \beta_{\sigma 0}/2), \tag{4.35}$$

where $\mathrm{Be}(a, b)$ denotes a beta distribution with parameters a, b and $\mathrm{IG}(\alpha, \beta)$ denotes an inverted gamma distribution with shape parameter α and scale parameter β. For a prior distribution of Ψ, we let

$$\Psi^{-1} = \begin{pmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{pmatrix},$$

where Ψ^{11}, Ψ^{12} and Ψ^{22} are $p \times p$ matrices. Noting that $\Psi^{11} = I_p + \Psi^{12}(\Psi^{22})^{-1}\Psi^{21}$, we assume

$$\Psi^{22} \sim \mathrm{W}_p(n_0, S_0), \tag{4.36}$$

$$\Psi^{21} | \Psi^{22} \sim \mathrm{N}_{p \times p}((\Psi^{22}\Delta_0), \Lambda_0 \otimes \Psi^{22}), \tag{4.37}$$

where $\mathrm{W}(n, S)$ denotes a Wishart distribution with parameters (n, S).

Then, the joint posterior density function is

$$
\begin{aligned}
&f(\boldsymbol{\vartheta}, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n) \\
&\propto \pi(\boldsymbol{\vartheta}) \times \exp\left(\sum_{t=1}^n l_t\right) \times \exp\left\{-\frac{1}{2}\sum_{t=1}^n (\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)'\Delta^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}\sum_{t=1}^n (\boldsymbol{x}_t^* - \boldsymbol{\xi}^* - \boldsymbol{\lambda}^* g_t)'\Delta^{*-1}(\boldsymbol{x}_t^* - \boldsymbol{\xi}^* - \boldsymbol{\lambda}^* g_t)\right\} \\
&\quad \times |\Omega_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\Omega_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right\} \\
&\quad \times |\Omega|^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{n-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\right] \\
&\quad \times \left(\frac{\sigma^2}{1-\theta^2}\right)^{-\frac{1}{2}} \exp\left\{-\frac{(g_1 - \gamma)^2}{2\sigma^2/(1-\theta^2)}\right\} \times (\sigma^2)^{-\frac{n-1}{2}} \exp\left[-\frac{\sum_{t=1}^{n-1}\{g_{t+1} - (1-\theta)\gamma - \theta g_t\}^2}{2\sigma^2}\right] \\
&\quad \times \exp\left(-\frac{1}{2\kappa}\boldsymbol{m}_1'\boldsymbol{m}_1\right) \times |\Omega_{\boldsymbol{m}}|^{-\frac{n-1}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{n-1}(\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)'\Omega_{\boldsymbol{m}}^{-1}(\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)\right\},
\end{aligned}
\tag{4.38}
$$

where

$$
\begin{aligned}
l_t = &-\frac{1}{2}\Big[R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}(\boldsymbol{y}_t - \boldsymbol{m}_t) - Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\Big]' \\
&\times (I_p - Q'\Omega^{-1}Q)^{-1}\Big[R_t^{-\frac{1}{2}}V_t^{-\frac{1}{2}}(\boldsymbol{y}_t - \boldsymbol{m}_t) - Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\Big] \\
&- \frac{1}{2}\log|R_t| - \frac{1}{2}\textstyle\sum_{i=1}^p h_{it}.
\end{aligned}
$$

and $\boldsymbol{\vartheta} = (\boldsymbol{\xi}, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^*, \Delta, \Delta^*, \boldsymbol{\mu}, \gamma, \Phi, \theta, \Omega, Q, \sigma^2, \Omega_{\boldsymbol{m}})$.

We implement the MCMC algorithm in five blocks:

1. Initialize $\{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

2. Generate $\boldsymbol{\vartheta}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n$.

3. Generate $\{\boldsymbol{h}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

4. Generate $\{g_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

5. Generate $\{\boldsymbol{m}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

6. Go to 2.

4.3.2 Generation of latent variable $\{\boldsymbol{h}_t\}_{t=1}^n$

A simple sampling method for $\{\boldsymbol{h}_t\}_{t=1}^n$ which we use is a single-move sampler that draws a single latent variable $\boldsymbol{h}_t$ at a time given the other $\boldsymbol{h}_t$'s and the parameters. For $t = 1, \ldots, n$, we propose a candidate $\boldsymbol{h}_t | \{\boldsymbol{h}_s\}_{-t}, \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta} \sim \mathrm{N}(\boldsymbol{m}_{\boldsymbol{h}t}, S_{\boldsymbol{h}t})$, where

$$
\begin{aligned}
S_{\boldsymbol{h}t} &= \{\Phi\Omega^{-1}\Phi + (\Omega - QQ')^{-1} + \Delta^{-1}\}^{-1}, \\
\boldsymbol{m}_{\boldsymbol{h}t} &= S_{\boldsymbol{h}t}[-0.5\mathbf{1}_p + \Phi\Omega^{-1}\{\boldsymbol{h}_{t+1} - (I_p - \Phi)\boldsymbol{\mu}\} \\
&\qquad + (\Omega - QQ')^{-1}\{\boldsymbol{\mu} + \Phi(\boldsymbol{h}_{t-1} - \boldsymbol{\mu}) + QR_{t-1}^{-\frac{1}{2}}V_{t-1}^{-\frac{1}{2}}(\boldsymbol{y}_{t-1} - \boldsymbol{m}_{t-1})\} + \Delta^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi})],
\end{aligned}
$$

and accept it using MH algorithm.

4.3.3 Generation of latent variable $\{g_t\}_{t=1}^n$

For $t = 1, \ldots, n$, the conditional distribution of g_t is

$$
\begin{aligned}
&f(g_t | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{g_t\}_{-t}, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}) \\
&\propto \exp(l_t) \times \exp\left\{ -\frac{1}{2s_{gt}^2}(g_t - m_{gt})^2 \right\},
\end{aligned}
$$

where

$$
s_{gt}^2 = (\sigma^{-2} + \theta^2\sigma^{-2} + \boldsymbol{\lambda}^{*'}\Delta^{*-1}\boldsymbol{\lambda}^*)^{-1}, \tag{4.39}
$$

$$
m_{gt} = s_{gt}^2[\theta\sigma^{-2}\{g_{t+1} - (1-\theta)\gamma\} + \sigma^{-2}\{\gamma\theta(g_{t-1} - \gamma)\} + \boldsymbol{\lambda}^{*'}\Delta^{*-1}(\boldsymbol{x}_t^* - \boldsymbol{\xi}^*)]. \tag{4.40}
$$

Generate a candidate $g_t^\dagger \sim \mathrm{N}(m_{g_t}, s_g^2)$ and accept it with probability
$\min[1, \exp\{l_t(g_t^\dagger) - l_t(g_t)\}]$.

4.3.4 Generation of $\{\boldsymbol{m}_t\}_{t=1}^n$ and $\boldsymbol{\vartheta}$

Generation of $\{\boldsymbol{m}_t\}_{t=1}^n$. We generate all of $\boldsymbol{m}_t$'s simultaneously using a simulation smoother (de Jong and Shephard (1995) or Durbin and Koopman (2002)).

Generation of $\boldsymbol{\xi}$ *and* $\boldsymbol{\xi}^*$. The conditional distributions of $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^*$ are

$$
\boldsymbol{\xi} | \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \Delta \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\xi}}, S_{\boldsymbol{\xi}}), \tag{4.41}
$$

$$
\boldsymbol{\xi}^* | \{\boldsymbol{x}_t^*\}_{t=1}^n, \{g_t\}_{t=1}^n, \boldsymbol{\lambda}^*, \Delta^* \sim \mathrm{N}_{p(p-1)/2}(\boldsymbol{m}_{\boldsymbol{\xi}}^*, S_{\boldsymbol{\xi}}^*), \tag{4.42}
$$

where

$$S_{\boldsymbol{\xi}} = \{(S_{\boldsymbol{\xi}0})^{-1} + n\Delta^{-1}\}^{-1},\ \ \boldsymbol{m}_{\boldsymbol{\xi}} = S_{\boldsymbol{\xi}}\left[(S_{\boldsymbol{\xi}0})^{-1}\boldsymbol{m}_{\boldsymbol{\xi}0} + \Delta^{-1}\sum_{t=1}^{n}(\boldsymbol{x}_t - \boldsymbol{h}_t)\right], \tag{4.43}$$

$$S^*_{\boldsymbol{\xi}} = \{S^{*-1}_{\boldsymbol{\xi}0} + n\Delta^{*-1}\}^{-1},\ \ \boldsymbol{m}^*_{\boldsymbol{\xi}} = S^*_{\boldsymbol{\xi}}\left\{S^{*-1}_{\boldsymbol{\xi}0}\boldsymbol{m}^*_{\boldsymbol{\xi}0} + \Delta^{*-1}\sum_{t=1}^{n}(\boldsymbol{x}^*_t - \boldsymbol{\lambda}^* g_t)\right\}. \tag{4.44}$$

Generation of $\boldsymbol{\lambda}^*$. The conditional distribution of $\boldsymbol{\lambda}^*$ is

$$\boldsymbol{\lambda}^*|\{\boldsymbol{x}^*_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \boldsymbol{\xi}^*, \Delta^* \sim \mathrm{N}_{p(p-1)/2}(\boldsymbol{m}^*_{\boldsymbol{\lambda}}, S^*_{\boldsymbol{\lambda}}), \tag{4.45}$$

where

$$S^*_{\boldsymbol{\lambda}} = \left(S^{*-1}_{\boldsymbol{\lambda}0} + \sum_{t=1}^{n} g_t^2\Delta^{*-1}\right)^{-1},\ \ \boldsymbol{m}^*_{\boldsymbol{\lambda}} = S^*_{\boldsymbol{\lambda}}\left\{S^{*-1}_{\boldsymbol{\lambda}0}\boldsymbol{m}^*_{\boldsymbol{\lambda}0} + \Delta^{*-1}\sum_{t=1}^{n}(\boldsymbol{x}^*_t - \boldsymbol{\xi}^*)g_t\right\}. \tag{4.46}$$

Generation of Δ *and* Δ^*. The conditional distributions of $\delta_i^2,\ i = 1,\dots,p$ and $\delta_j^{*2},\ j = 1,\dots,p(p-1)/2$ are given by

$$\delta_i^2|\{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \boldsymbol{\xi}, \sim \mathrm{IG}(\alpha_{\delta_i}/2, \beta_{\delta_i}/2),$$

$$\delta_j^{*2}|\{\boldsymbol{x}^*_t\}_{t=1}^n, \{g_t\}_{t=1}^n, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^* \sim \mathrm{IG}(\alpha^*_{\delta_j}/2, \beta^*_{\delta_j}/2), \tag{4.47}$$

where $\alpha_{\delta_i} = \alpha_{\delta_i 0} + n$, $\alpha^*_{\delta_j} = \alpha^*_{\delta_j 0} + n$,

$$\beta_{\delta_i} = \beta_{\delta_i 0} + \sum_{t=1}^{n}\{x_{it} - \xi_i - h_{it}\}^2,\ \ \beta^*_{\delta_j} = \beta^*_{\delta_j 0} + \sum_{t=1}^{n}(x^*_{jt} - \xi^*_j - \lambda^*_j g_t)^2. \tag{4.48}$$

We implement the sampling procedure for $\boldsymbol{\mu}, \gamma, \Phi, \theta, \Omega, Q, \sigma^2, \Omega_{\boldsymbol{m}}$ as mentioned in the last chapter.

4.4 Illustrative example using simulated data

This section illustrates the extended DESV model using simulated data. We consider a trivariate case ($p = 3$). Using the following parameters based on our empirical studies in Section 4.5,

$$\boldsymbol{\xi}_{\ddagger} = \boldsymbol{0}_3,\ \boldsymbol{\xi}^*_{\ddagger} = \boldsymbol{0}_3,\ \boldsymbol{\lambda}^*_{\ddagger} = \boldsymbol{1}_3,\ \Delta_{\ddagger} = 0.1I_3,\ \Delta^*_{\ddagger} = 0.1I_3,$$

$$\boldsymbol{\mu}_{\ddagger} = \boldsymbol{0}_3,\ \gamma_{\ddagger} = 1,\ \Phi_{\ddagger} = 0.97I_3,\ \theta_{\ddagger} = 0.97,$$

$$\Omega_{\ddagger} = 0.1I_3 + 0.1J_{33},\ Q_{\ddagger} = O_{3\times 3},\ \Sigma_{\ddagger} = 0.01,\ \Omega_{\boldsymbol{m}*} = 0.001I_3,$$

we generate 2,500 observations ($n = 2500$). For prior distributions, we assume

$$
\begin{aligned}
&\boldsymbol{\xi} \sim \mathrm{N}_3(\boldsymbol{\xi}_{\ddagger}, 100I_3),\ \boldsymbol{\xi}^* \sim \mathrm{N}_3(\boldsymbol{\xi}^*_{\ddagger}, 100I_3),\ \boldsymbol{\lambda}^* \sim \mathrm{N}_3(\boldsymbol{\lambda}^*_{\ddagger}, 100I_3),\\
&\delta_i^2 \sim \mathrm{IG}(5, 3\delta_{i\ddagger}^2),\ i = 1, \ldots, 3,\ \ \delta_j^{*2} \sim \mathrm{IG}(5, 3\delta_{j\ddagger}^{*2}),\ j = 1, \ldots, 3,\\
&\boldsymbol{\mu} \sim \mathrm{N}(\boldsymbol{\mu}_{\ddagger}, 100I_3),\ \gamma \sim \mathrm{N}(\gamma_{\ddagger}, 100),\ \frac{\phi_i + 1}{2} \sim \mathrm{Be}(20, 1.5),\ \ i = 1, \ldots, 3,\\
&\frac{\theta + 1}{2} \sim \mathrm{Be}(20, 1.5),\ \ \sigma^2 \sim \mathrm{IG}(5, 3\sigma_{\ddagger}^2),\\
&\Psi^{22} \sim \mathrm{W}(12, 12^{-1}\Omega_{\ddagger}),\ \Psi^{21}|\Psi^{22} \sim \mathrm{N}_{3\times 3}(O, 10I_3 \otimes \Psi^{22}),\ \ \omega_{m_i 0}^2 \sim \mathrm{IG}(5, 3\omega_{m_i\ddagger}^2),\ i = 1, \ldots, 3.
\end{aligned}
$$

Using the sampling method described in the last section, we generate 100,000 MCMC samples after discarding the first 10,000 samples as the burn-in period.

Table 4.1 reports the true values, posterior means, 95% credible intervals and estimated inefficiency factors (IF)[2]. The posterior means are all close to the true values, which suggests that our proposed algorithm works well. Note that the acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{g_t\}_{t=1}^n$ in the independent MH algorithms are 0.799 and 0.897, respectively. It indicates that the generated candidates are accepted with relatively high probability and our sampling algorithm also works well.

[2]The inefficiency factor is defined as $1 + 2\sum_{g=1}^{\infty} \rho(g)$, where $\rho(g)$ is the sample autocorrelation at lag g. This is interpreted as the ratio of the numerical variance of the posterior mean from the chain to the variance of the posterior mean from hypothetical uncorrelated draws. The smaller the inefficiency factor becomes, the closer the MCMC sampling is to the uncorrelated sampling.

Table 4.1: Posterior means, 95% credible intervals and inefficiency factors.

	True	Mean	95% interval	IF
ξ_1	0	0.047	(-0.009, 0.107)	242
ξ_2	0	-0.016	(-0.075, 0.045)	304
ξ_3	0	-0.008	(-0.067, 0.051)	259
ξ_1^*	0	-0.029	(-0.214, 0.113)	1178
ξ_2^*	0	-0.033	(-0.218, 0.108)	1184
ξ_3^*	0	-0.035	(-0.219, 0.107)	1188
λ_1^*	1	1.055	(0.933, 1.210)	991
λ_2^*	1	1.057	(0.935, 1.210)	994
λ_3^*	1	1.063	(0.941, 1.216)	997
δ_1^2	0.1	0.097	(0.084, 0.111)	20
δ_2^2	0.1	0.091	(0.078, 0.104)	15
δ_3^2	0.1	0.091	(0.079, 0.104)	15
δ_1^{*2}	0.1	0.103	(0.096, 0.110)	3
δ_2^{*2}	0.1	0.095	(0.089, 0.102)	3
δ_3^{*2}	0.1	0.096	(0.090, 0.103)	4
μ_1	0	-0.300	(-1.005, 0.406)	2
μ_2	0	-0.373	(-1.035, 0.289)	3
μ_3	0	-0.178	(-0.762, 0.403)	5
γ	1	0.974	(0.781, 1.169)	142
ϕ_1	0.97	0.975	(0.966, 0.983)	3
ϕ_2	0.97	0.972	(0.962, 0.981)	2
ϕ_3	0.97	0.970	(0.961, 0.979)	3
θ	0.97	0.967	(0.954, 0.979)	3
Ω_{11}	0.2	0.201	(0.181, 0.222)	15
Ω_{21}	0.1	0.102	(0.089, 0.116)	2
Ω_{31}	0.1	0.089	(0.076, 0.102)	7
Ω_{22}	0.2	0.213	(0.191, 0.235)	8
Ω_{32}	0.1	0.092	(0.079, 0.106)	5
Ω_{33}	0.2	0.190	(0.171, 0.210)	8
Q_{11}	0.06	0.069	(0.046, 0.092)	9
Q_{21}	0.06	0.049	(0.025, 0.072)	4
Q_{31}	0.06	0.047	(0.024, 0.069)	7
Q_{12}	0	-0.004	(-0.027, 0.020)	4
Q_{22}	0	0.021	(-0.003, 0.045)	3
Q_{32}	0	-0.013	(-0.035, 0.011)	3
Q_{13}	0	-0.019	(-0.042, 0.006)	3
Q_{23}	0	-0.014	(-0.038, 0.010)	3
Q_{33}	0	-0.005	(-0.028, 0.019)	5
σ^2	0.02	0.017	(0.012, 0.023)	771
$\omega_{m_1}^2 \times 10^3$	1	0.989	(0.662, 1.422)	91
$\omega_{m_2}^2 \times 10^3$	1	1.320	(0.927, 1.833)	55
$\omega_{m_3}^2 \times 10^3$	1	1.200	(0.843, 1.653)	93

4.5 Empirical Study

4.5.1 Data

We use daily close-to-close returns and the associated realized measure of the covariance for some of the most liquid stocks in the Dow Jones Industrial Average (DJIA) index, obtained from Oxford-Man Institute[3]. These are: (1) Bank of America (BAC), (2) JP Morgan (JPM), (3) International Business Machines (IBM), (4) Microsoft (MSFT). The sample period is from February 1, 2001 to December 31, 2009 (2242 observations in total). Note that the asset return is calculated as $y_t = (\log p_t - \log p_{t-1}) \times 100$, where p_t is the asset price at time t.

4.5.2 Estimation results

Using the same prior distributions for the parameters as in Section 4.4, we implement the MCMC algorithm to conduct a Bayesian inference on the parameters of interest. We generate 100,000 MCMC samples from the posterior distributions of the parameters in the model after discarding the first 100,000 samples as the burn-in period.

Tables 4.2 and 4.3 report the summary of the estimation results. Figure 4.1 shows the posterior means of $\exp(h_{i,t}/2)$, square root of the estimated time-varying variances and ρ_t, dynamic equicorrelation of $\boldsymbol{y}_t$.

The posterior probability with which the bias adjustment term ξ_1 is negative is over 0.975 and this is similar for ξ_2, ξ_3 and ξ_4. It suggests that the realized measure of variances which we use tends to have a downward bias due to the non-trading hours and that the effect of non-trading hours is more important than that of microstructure noise. Note that the posterior probability with which γ, the mean of $\{g_t\}_{t=1}^n$ is negative is over 0.975. The posterior distributions of $\boldsymbol{\xi}^*$ and $\boldsymbol{\lambda}^*$ suggest that the realized measure of the correlation which we use tends to have a downward bias as described in Section 4.1 (Epps effect). These estimation results show that our model can detect and eliminate the bias contained in the realized measure owing to the problems in the financial market.

Among the posterior means of μ's, μ_2 is the largest and μ_3 is the smallest, but the difference between them is not huge and the levels of the volatilities are not so different from one another. We find that the posterior means of Ω_{11} and Ω_{22} are apparently larger than

[3]See Noureldin, Shephard, and Sheppard (2012) for the detailed explanation.

Table 4.2: Posterior means, 95% credible intervals and inefficiency factors.

	Mean	95% interval	IF
ξ_1	-0.611	(-0.673, -0.549)	273
ξ_2	-0.493	(-0.555, -0.431)	419
ξ_3	-0.596	(-0.655, -0.536)	251
ξ_4	-0.682	(-0.741, -0.624)	192
ξ_1^*	-0.244	(-0.330, -0.160)	114
ξ_2^*	-0.885	(-0.992, -0.780)	124
ξ_3^*	-0.934	(-1.044, -0.828)	124
ξ_4^*	-0.837	(-0.938, -0.741)	117
ξ_5^*	-0.894	(-1.003, -0.788)	124
ξ_6^*	-0.520	(-0.611, -0.432)	125
λ_1^*	0.694	(0.602, 0.800)	69
λ_2^*	0.885	(0.774, 1.014)	77
λ_3^*	0.916	(0.803, 1.046)	82
λ_4^*	0.812	(0.709, 0.932)	74
λ_5^*	0.911	(0.798, 1.040)	82
λ_6^*	0.751	(0.657, 0.860)	80
δ_1^2	0.088	(0.079, 0.097)	9
δ_2^2	0.083	(0.074, 0.091)	27
δ_3^2	0.103	(0.093, 0.112)	42
δ_4^2	0.105	(0.095, 0.115)	22
δ_1^{*2}	0.663	(0.621, 0.709)	2
δ_2^{*2}	0.705	(0.658, 0.755)	2
δ_3^{*2}	0.571	(0.531, 0.614)	3
δ_4^{*2}	0.715	(0.668, 0.764)	2
δ_5^{*2}	0.623	(0.580, 0.669)	4
δ_6^{*2}	0.464	(0.432, 0.498)	2
μ_1	0.893	(0.481, 1.300)	5
μ_2	1.174	(0.819, 1.526)	14
μ_3	0.681	(0.450, 0.909)	13
μ_4	1.052	(0.805, 1.298)	10
γ	0.269	(0.164, 0.368)	127
ϕ_1	0.959	(0.950, 0.968)	34
ϕ_2	0.954	(0.944, 0.963)	49
ϕ_3	0.937	(0.923, 0.949)	73
ϕ_4	0.939	(0.926, 0.952)	40
θ	0.626	(0.596, 0.656)	7

Table 4.3: Posterior means, 95% credible intervals and inefficiency factors.

	Mean	95% interval	IF
Ω_{11}	0.161	(0.145, 0.178)	11
Ω_{21}	0.146	(0.133, 0.161)	10
Ω_{31}	0.114	(0.103, 0.126)	17
Ω_{41}	0.118	(0.106, 0.130)	15
Ω_{22}	0.149	(0.133, 0.165)	35
Ω_{32}	0.118	(0.106, 0.130)	17
Ω_{42}	0.116	(0.105, 0.129)	32
Ω_{33}	0.114	(0.101, 0.129)	64
Ω_{43}	0.110	(0.098, 0.121)	9
Ω_{44}	0.123	(0.109, 0.137)	59
Q_{11}	0.059	(0.035, 0.083)	5
Q_{21}	0.047	(0.025, 0.070)	8
Q_{31}	0.049	(0.028, 0.070)	8
Q_{41}	0.035	(0.013, 0.058)	5
Q_{12}	0.014	(-0.007, 0.037)	5
Q_{22}	0.009	(-0.012, 0.030)	6
Q_{32}	-0.014	(-0.033, 0.006)	8
Q_{42}	0.013	(-0.007, 0.034)	10
Q_{13}	0.005	(-0.018, 0.029)	12
Q_{23}	0.011	(-0.011, 0.034)	8
Q_{33}	0.011	(-0.009, 0.033)	10
Q_{43}	0.021	(-0.001, 0.043)	7
Q_{14}	-0.002	(-0.023, 0.020)	11
Q_{24}	-0.013	(-0.033, 0.008)	6
Q_{34}	0.012	(-0.007, 0.031)	9
Q_{44}	0.005	(-0.014, 0.026)	10
σ^2	0.406	(0.307, 0.519)	87
$\omega^2_{m_1} \times 10^3$	0.229	(0.119, 0.405)	177
$\omega^2_{m_2} \times 10^3$	0.269	(0.136, 0.519)	309
$\omega^2_{m_3} \times 10^3$	0.248	(0.122, 0.465)	151
$\omega^2_{m_4} \times 10^3$	0.228	(0.120, 0.421)	155

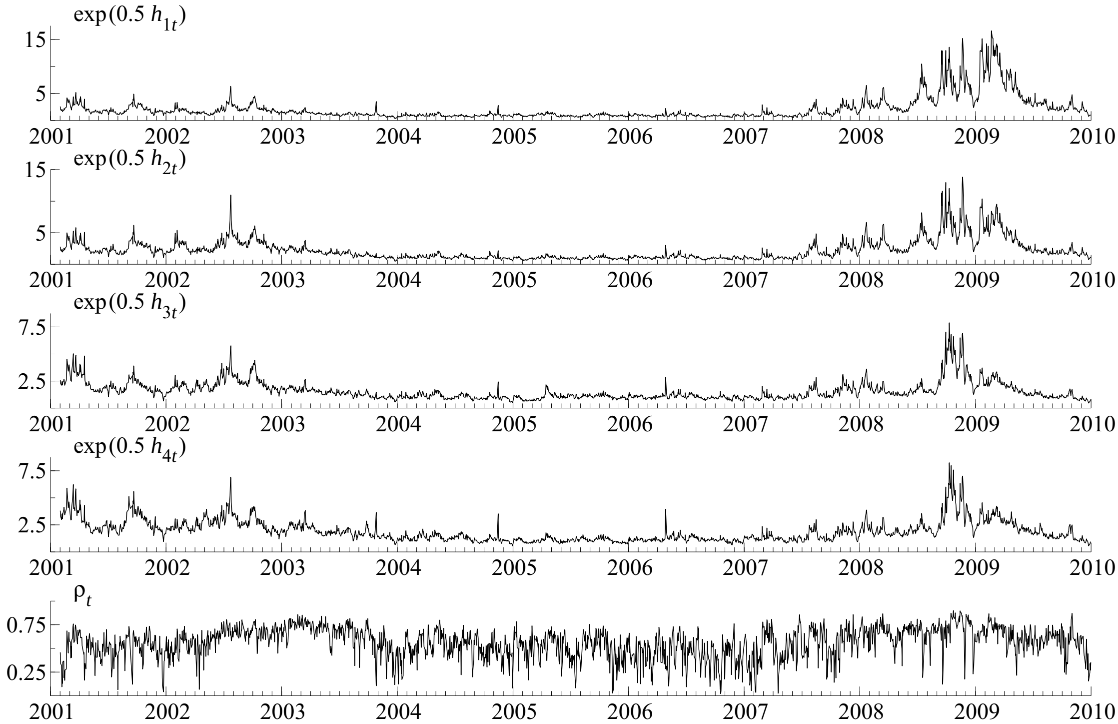

Figure 4.1: Posterior means of square root of the variances and the dynamic equicorrelation.

those of Ω_{33} and Ω_{44} and it is consistent with the observed fluctuations shown in Figure 4.1. The credible intervals of off-diagonal elements of Ω are positive and do not include zero, which means that unobserved volatilities are correlated positively each other.

The posterior means of autoregressive coefficients (ϕ_i's) are very high (over 0.9), which shows that the log volatilities follow highly persistent processes. In addition, Figure 4.1 indicates the comovement of the volatilities. The trajectories sharply increased in September 2008, corresponding to the financial crisis during which Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008). We also observe the increase in July 2002 resulted from the market turmoil during which Worldcom filed for Chapter 11 bankruptcy protection (July 21, 2002). The increases in April 2001 and in September 2001 are due to the collapse of dot com bubble and to September 11 attacks.

The posterior means of autoregressive coefficient (θ) is not high and the equicorrelation parameter is not so persistent. The bottom panel of Figure 4.1 shows that the equicorrelation parameter varies greatly, which means that it is time-varying and far from constant.

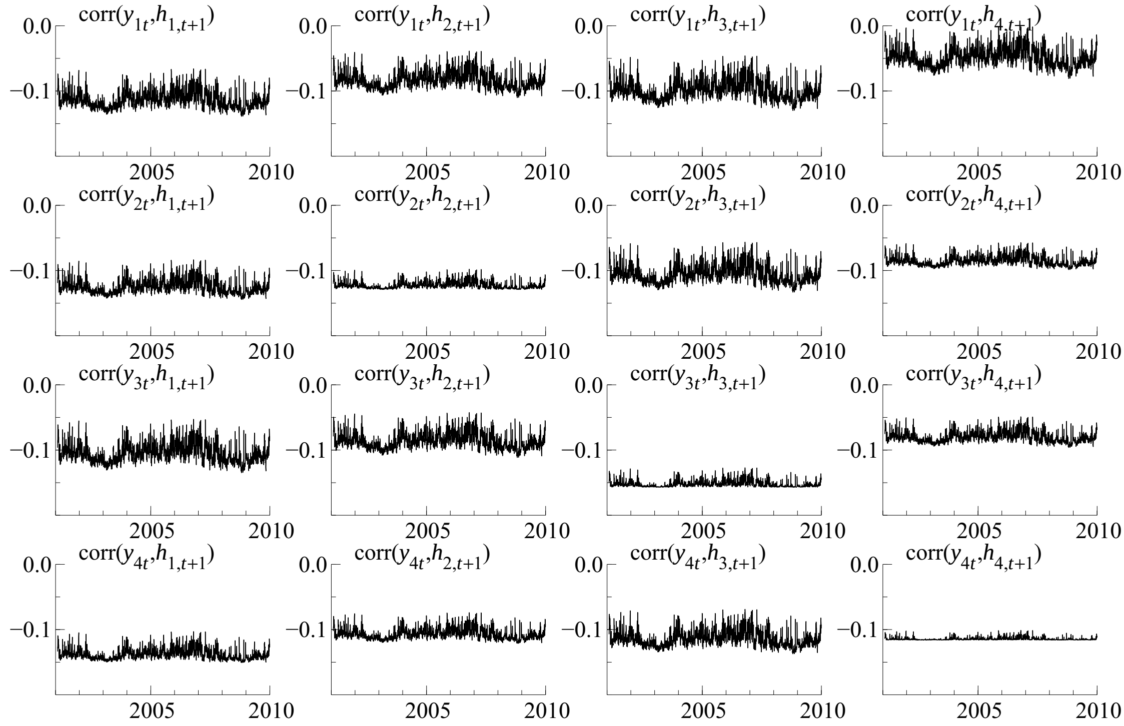

Figure 4.2: Posterior means of correlation of $(y_{it}, h_{j,t+1})$.

Figure 4.2 shows the posterior means of dynamic correlations between the return of i-th asset at time t ($y_{i,t}$) and the j-th log variance at time $t+1$ ($h_{j,t+1}$). As stated in Section 4.2, Q is the covariance between $\boldsymbol{z}_t$, the transformed error term in the observed equation of our model and $\boldsymbol{\eta}_{t+1}$, the error term in the state equation for $t = 1, \ldots, n-1$. To verify the existence of the cross leverage effects, we should transform Q using R_t and Ω to obtain the posterior means of the (time-varying) correlations between $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\eta}_{t+1}$. Their trajectories are negative as expected and but not far from zero, which indicates the existence of the weak cross leverage effect. While the strong existence of leverage effect is frequently reported in the analysis of the stock index, Takeuchi-Nogimori (2012) shows that the variance of the common stock has a weak asymmetry. It is consistent with our estimation results about the asymmetry stated above.

We note that the leverage effect of asset 3 is the strongest (the mean is -0.152) and the cross leverage effect of asset 1 to asset 4 is the weakest (the mean is -0.049) among the cross leverage effects. In addition, the correlation between y_{1t} and $h_{4,t+1}$ is apparently weaker than the correlation between y_{4t} and $h_{1,t+1}$. It indicates that cross leverage effects from asset i to j, $i \neq j$, are not symmetric.

In conclusion, using our multivariate RSV model, we can therefore detect the volatility clustering, the dynamic equicorrelation and the (weak) cross leverage effects of the four stocks.

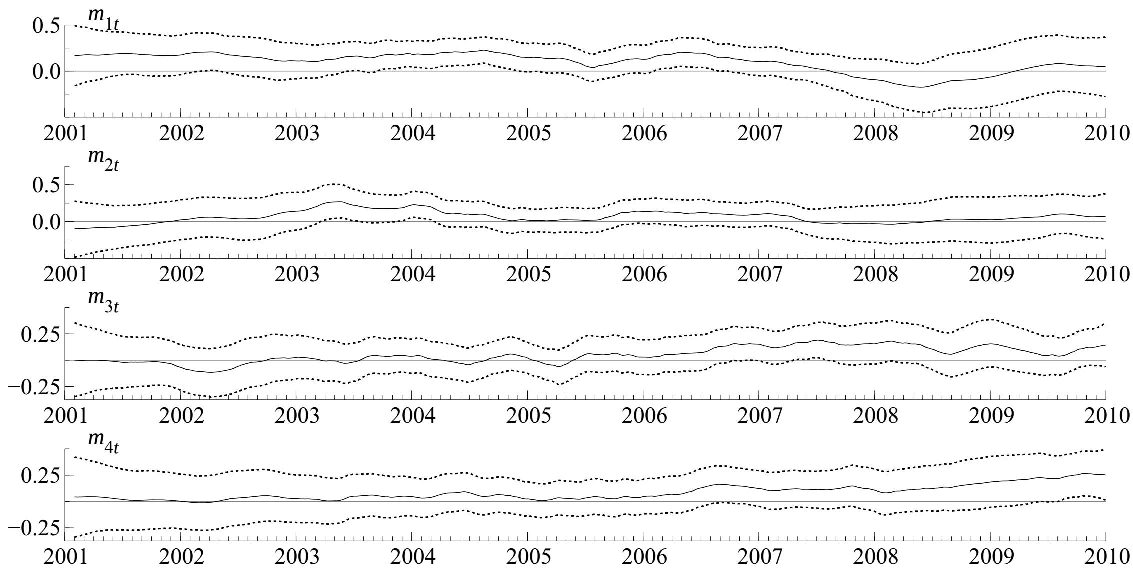

Figure 4.3: Posterior means (solid lines) and 95% credible intervals (between the two outmost dotted lines) of the expectations.

Figure 4.3 shows the posterior means of the expectations of the asset returns. We find that the 95% credible intervals include zero at almost all of the time points. It seems to fluctuate slowly, which is consistent with the small values of the variances ($\omega_{m_i}^2$'s) reported in Table 4.3.

Note that the acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{g_t\}_{t=1}^n$ in the independent MH algorithms are 0.825 and 0.975, respectively. It indicates that the generated candidates are accepted with relatively high probability and our sampling algorithm works well.

4.6 Conclusion

In this paper, we proposed simultaneous modeling of the multivariate daily returns and the related realized measure with dynamic equicorrelation and cross leverage effect. We took a

Bayesian approach and described the simple (but efficient) MCMC sampling algorithm for our model. We showed that the biases in the realized measure owing to market microstructure noise, non-trading hours and non-synchronous trading can be estimated within our modeling framework. Numerical examples are provided and the proposed model is applied to the multivariate stock data. We find the persistence in the volatilities, the dynamics of correlations among the returns and the existence of weak cross leverage effects.

Chapter 5

Dynamic block-equicorrelation, realized stochastic volatility and cross leverage

5.1 Introduction

Over the last several decades, a number of multivariate volatility models have been proposed to model asset returns with time-varying variances. Two popular examples are generalized autoregressive conditional heteroskedasticity (GARCH) models (Bauwens, Laurent, and Rombouts (2006)) and multivariate stochastic volatility (SV) models (Asai, McAleer, and Yu (2006), Chib, Omori, and Asai (2009)).

They are proposed to model the volatility clustering and the dynamic correlations, which are found to exist in empirical studies of financial time series (Bauwens, Hafner, and Laurent (2012)). In volatilities of stock returns, we often observe the asymmetry or the cross leverage effect, which implies a decrease in the i-th dependent variable at date t followed by an increase in the j-th latent stochastic variance at date $(t+1)$. The simple univariate SV model with leverage effect is given in a state space form as:

$$y_t = m + \exp(h_t/2)\epsilon_t, \quad t = 1, \ldots, n, \tag{5.1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, n-1, \tag{5.2}$$

$$h_1 \sim \mathrm{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right), \tag{5.3}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & q \\ q & \sigma_\eta^2 \end{pmatrix} \right), \tag{5.4}$$

$$|\phi| < 1, \tag{5.5}$$

where y_t denotes a (univariate) asset return, h_t is a log-variance of y_t. The negative value of q implies the existence of the leverage effect. It can be extended to the multivariate SV model with cross leverage effect (Daníelsson (1998), Asai and McAleer (2006), Asai and McAleer (2009), Chan, Kohn, and Kirby (2006), Ishihara, Omori, and Asai (2011), Ishihara and Omori (2012) and Nakajima (2012)). The major difficulty in constructing such multivariate models is to make the covariance matrices positive definite, especially when some dynamic correlation structure between the asset returns is incorporated. It is desirable to model the dynamic covariance structure as simply as possible since the parameter estimation becomes difficult in the sense that there are too many latent variables to be integrated out analytically to obtain the likelihood function.

Chapter 3 proposes the multivariate SV model with dynamic equicorrelation and cross leverage effect (DESV model), which is assumed that all correlations equal but time-varying. Studies of the dynamic equicorrelation (DECO) model, which is based on a Dynamic conditional correlation (DCC) model, include Engle and Kelly (2012) and Clements, Coleman-Fenn, and Smith (2011). Making reference to Elton and Gruber (1973), they argue that the dynamic equicorrelation assumption gives a superior portfolio allocation. As shown in Chapter 3, the equicorrelation assumption is very simple but useful. Meanwhile, the assumption seems to be too strong and runs counter to the intuition, especially when the number of dependent variables is very large. Although we can extend it to the general model that all correlations between the dependent variables are different from each other and time-varying, it is not easy to estimate all the dynamic correlations and may result in overfitting to the past data. Thus the model in Chapter 3 should be extended to the block-equicorrelation model: we divide the dependent variables into several blocks and assume the dependent variables within the block have the dynamic equicorrelation structure. Between the variables in block i and in block j, we also assume the dynamic block-equicorrelation structure. Elton and Gruber (1973) discussed this extension and the groupings with the empirical work.

In modeling block-equicorrelation structure, we face a difficulty in keeping the block-equicorrelation matrices positive definite due to the existence of inter-block-equicorrelation structure since we cannot define the support of the dynamic block-equicorrelation factors analytically. With the equicorrelation (i.e., 1-block) case, we assume that the state space of the transformed dynamic equicorrelation factor is $\mathbb{R}$, and the models can be expressed analytically in terms of the dynamic equicorrelation factor, ρ_t. For the model of this article, we cannot express the likelihood function simply (and analytically) in terms of ρ_t's, so we cannot conduct the efficient estimation for the (transformed) dynamic block-equicorrelation factor without additional information.

Takahashi, Watanabe, and Omori (2009) proposed simultaneous modeling of the daily returns and the realized measure by adding the observation equation

$$x_t = \xi + h_t + \zeta_t, \quad t = 1, \ldots, n, \tag{5.6}$$

to the simple univariate SV model (5.1)-(5.5), where x_t denotes a realized measure of h_t (log-variance of y_t). We refer to this model as the realized stochastic volatility model or RSV model (Dobrev and Szerszen (2010) and Koopman and Scharth (2012)). Hansen, Huang, and Shek (2012) introduced a framework that combines a GARCH structure for returns with an integrated model for realized measures and named their models as realized GARCH models. Since they demonstrated their contemporaneous models offer a substantial improvement in understanding the financial time series structure, we adopt their approach and incorporate the SV model with the information of the realized measure.

As addressed in Section 4.1, the realized measure is biased as a measure of the latent stochastic variable. To overcome the difficulty, a vast literature is devoted to improving the estimator (for the multivariate case, e.g., Malliavin and Mancino (2002), Hayashi and Yoshida (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Kunitomo and Sato (2008) and Zhang (2011)). Whereas we cannot ignore the contamination of the financial data obtained at short intervals as stated in Section 4.1, the daily returns are not heavily affected by such noise. Therefore the daily returns may provide us additional information to eliminate the bias contained in the realized measure. Takahashi, Watanabe, and Omori (2009) and Hansen, Huang, and Shek (2012) estimated the parameters in the distribution of financial returns and in the model jointly, which also makes it possible to adjust the bias of the realized measure simultaneously.

In this study, we extend the DESV model and RSV model to the multivariate model by introducing the multivariate realized SV model with dynamic "block-equicorrelation" and cross leverage effect. Because it is difficult to evaluate the likelihood function using the high-dimensional numerical integration, Bayesian efficient estimation methods have been proposed (e.g., Shephard and Pitt (1997), Watanabe and Omori (2004) and Omori and Watanabe (2008)). Chapter 3 also describes an efficient Bayesian estimation using the Markov chain Monte Carlo (MCMC) method to generate the latent stochastic variances and dynamic equicorrelation factor from the posterior distributions. It is highly efficient compared with the benchmarking simple method, single-move sampling method, which draws a single latent variable at a time given the other latent variables and the parameters. It means that we only need to generate a fewer number of MCMC samples to estimate the posterior distribution of the interested parameters. With the additional information to estimate unobserved variables by using the realized measures, we do not need to require the efficient (but complicated) estimation method and conduct the simple estimation in this study.

The rest of this article is organized as follows. In Section 5.2, we propose the multivariate realized stochastic volatility model with dynamic block-equicorrelation and cross leverage effect. Section 5.3 describes an efficient Bayesian estimation method for the proposed model. In Section 5.4, we illustrate our estimation method using simulated data and shows that our MCMC algorithm works well. Section 5.5 applies our proposed DESV model to the multivariate asset return data and the corresponding realized measure. Section 5.6 concludes this paper.

5.2 Multivariate realized SV model with dynamic block-equicorrelation and cross leverage

Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})'$ denote a p-dimensional asset return vector at time t $(t = 1, \ldots, n)$. We divide $\boldsymbol{y}_t$ into d blocks and let the l-th block consist of p_l elements of $\boldsymbol{y}_t$ (Note that $\sum_{l=1}^{d} p_l = p$). We also set $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{pt})'$ be a p-dimensional vector of the logarithm of the realized measure of the variances and $RCor_t$ be the realized measure of the correlation matrix corresponding to $\boldsymbol{y}_t$, respectively. We divide $RCor_t$ into d^2 submatrices corresponding to the subvectors of $\boldsymbol{y}_t$ and denote them as $RCor_t^{l_1 l_2}$, $l_1 = 1, \ldots, d$, $l_2 = 1, \ldots, d$. For $l = 1, \ldots, d$, define $\boldsymbol{x}_t^{ll*} = (x_{1,t}^{ll*}, \ldots, x_{p_l(p_l-1)/2,t}^{ll*})'$, where

$$x_{jt}^{ll*} = \log\left(\frac{\rho_{jt}^{ll*}}{1-\rho_{jt}^{ll*}}\right), \quad \boldsymbol{\rho}_t^{ll*} = (\rho_{1,t}^{ll*}, \ldots, \rho_{p_l(p_l-1)/2,t}^{ll*})',$$

and where $\boldsymbol{\rho}_t^{ll*}$ is made by extracting and vectorizing the lower off-diagonal elements of $RCor_t^{ll}$. We also define $\boldsymbol{x}_t^{l_1l_2*} = (x_{1,t}^{l_1l_2*},\ldots,x_{p_{l_1}p_{l_2},t}^{l_1l_2*})'$, where

$$x_{jt}^{l_1l_2*} = \log\left(\frac{1+\rho_{j,t}^{l_1l_2*}}{1-\rho_{j,t}^{l_1l_2*}}\right),\ \ \boldsymbol{\rho}_t^{l_1l_2*} = \mathrm{vec}(RCor_t^{l_1l_2}) = (\rho_{1,t}^{l_1l_2*},\ldots,\rho_{p_{l_1}p_{l_2},t}^{l_1l_2*})',$$

for $l_1 = l_2,\ldots,d,\ l_2 = 1,\ldots,d,\ l_1 > l_2$. Let $\boldsymbol{m}_t = (m_{1t},\ldots,m_{pt})'$ and $\boldsymbol{h}_t = (h_{1t},\ldots,h_{pt})'$ denote p-dimensional vectors of unobserved variables and $\boldsymbol{g}_t = (g_t^{11},\ldots,g_t^{d1},g_t^{22},\ldots,g_t^{d2},\ldots,g_t^{dd})'$ a q-dimensional vector of unobserved variables ($q = d(d+1)/2$). We consider the multivariate RSV model given by

$$\boldsymbol{y}_t = \boldsymbol{m}_t + V_t^{1/2}\boldsymbol{\epsilon}_t,\ \ \boldsymbol{\epsilon}_t \sim \mathrm{N}_p(\mathbf{0}_p, R_t),\ \ t=1,...,n, \tag{5.7}$$

$$\boldsymbol{x}_t = \boldsymbol{\xi} + \boldsymbol{h}_t + \boldsymbol{w}_t,\ \ \boldsymbol{w}_t \sim \mathrm{N}_p(\mathbf{0}_p, \Delta),\ \ t=1,...,n, \tag{5.8}$$

$$\boldsymbol{x}_t^{l_1l_2*} = \boldsymbol{\xi}^{l_1l_2*} + \boldsymbol{\lambda}^{l_1l_2*}g_t^{l_1l_2} + \boldsymbol{w}_t^{l_1l_2*},\ \ \boldsymbol{w}_t^{l_1l_2*} \sim \mathrm{N}.(\mathbf{0}., \Delta^{l_1l_2*}),$$
$$l_1 = l_2,\ldots,d,\ l_2 = 1,...,d,\ t=1,...,n, \tag{5.9}$$

$$\boldsymbol{h}_{t+1} = \boldsymbol{\mu} + \Phi(\boldsymbol{h}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t,\ \ \boldsymbol{\eta}_t \sim \mathrm{N}_p(\mathbf{0}_p, \Omega),\ \ t=1,...,n-1, \tag{5.10}$$

$$\boldsymbol{h}_1 = \boldsymbol{\mu} + \boldsymbol{\eta}_0,\ \ \boldsymbol{\eta}_0 \sim \mathrm{N}_p(\mathbf{0}_p, \Omega_0), \tag{5.11}$$

$$\boldsymbol{g}_{t+1} = \boldsymbol{\gamma} + \Theta(\boldsymbol{g}_t - \boldsymbol{\gamma}) + \boldsymbol{\zeta}_t,\ \ \boldsymbol{\zeta}_t \sim \mathrm{TN}_q(\mathbf{0}_q, \Sigma),\ \ t=1,...,n-1, \tag{5.12}$$

$$\boldsymbol{g}_1 = \boldsymbol{\gamma} + \boldsymbol{\zeta}_0,\ \ \boldsymbol{\zeta}_0 \sim \mathrm{TN}_q(\mathbf{0}_q, \Sigma_0), \tag{5.13}$$

$$\boldsymbol{m}_{t+1} = \boldsymbol{m}_t + \boldsymbol{\eta}_{\boldsymbol{m}t},\ \ \boldsymbol{\eta}_{\boldsymbol{m}t} \sim \mathrm{N}_p(\mathbf{0}_p, \Omega_{\boldsymbol{m}}),\ \ t=1,...,n-1, \tag{5.14}$$

$$\boldsymbol{m}_1 = \boldsymbol{\eta}_{\boldsymbol{m}0},\ \ \boldsymbol{\eta}_{\boldsymbol{m}0} \sim \mathrm{N}_p(\mathbf{0}_p, \kappa I_p), \tag{5.15}$$

where $\mathbf{0}_p$ is a p-dimensional zero vector,

$$V_t = \mathrm{diag}\{\exp(h_{1t}),\ldots,\exp(h_{pt})\},\ t=1,\ldots,n, \tag{5.16}$$

$$R_t = \begin{bmatrix} (1-\rho_t^{11})I_{p_1} + \rho_t^{11}J_{p_1p_1} & \rho_t^{12}J_{p_1p_2} & \cdots & \rho_t^{1d}J_{p_1p_d} \\ \rho_t^{12}J_{p_2p_1} & (1-\rho_t^{22})I_{p_2} + \rho_t^{22}J_{p_2p_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_t^{d-1,d}J_{p_{d-1}p_d} \\ \rho_t^{1d}J_{p_dp_1} & \cdots & \rho_t^{d-1,d}J_{p_dp_{d-1}} & (1-\rho_t^{dd})I_{p_d} + \rho_t^{dd}J_{p_dp_d} \end{bmatrix},$$

I_p is a unit matrix of size p, $J_{p_1p_2} = \mathbf{1}_{p_1}\mathbf{1}'_{p_2}$ ($\mathbf{1}_p$ denotes a p-dimensional vector with all elements equal to one)[1],

$$\text{vech}\begin{pmatrix} \rho_t^{11} & \cdots & \rho_t^{1d} \\ \vdots & \ddots & \vdots \\ \rho_t^{1d} & \cdots & \rho_t^{dd} \end{pmatrix} = \begin{pmatrix} \frac{\exp(g_t^{11})-c_{11}}{\exp(g_t^{11})+1} \\ \vdots \\ \frac{\exp(g_t^{dd})-c_{dd}}{\exp(g_t^{dd})+1} \end{pmatrix}, \tag{5.17}$$

$$c_{ij} = \begin{cases} 0 & \text{if } i = j, \\ 1 & \text{otherwise,} \end{cases} \tag{5.18}$$

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)', \quad \boldsymbol{\xi}^{l_1l_2*} = \begin{cases} (\xi_1^{l_1l_2*}, \ldots, \xi_{p_{l_1}(p_{l_1}-1)/2}^{l_1l_2*})', & l_1 = l_2, \\ (\xi_1^{l_1l_2*}, \ldots, \xi_{p_{l_1}p_{l_2}}^{l_1l_2*})', & l_1 > l_2, \end{cases} \tag{5.19}$$

$$\boldsymbol{\lambda}^{l_1l_2*} = \begin{cases} (\lambda_1^{l_1l_2*}, \ldots, \lambda_{p_{l_1}(p_{l_1}-1)/2}^{l_1l_2*})', & l_1 = l_2, \\ (\lambda_1^{l_1l_2*}, \ldots, \lambda_{p_{l_1}p_{l_2}}^{l_1l_2*})', & l_1 > l_2, \end{cases} \tag{5.20}$$

$$\Delta = \text{diag}(\delta_1^2, \ldots, \delta_p^2), \quad \Delta^* = \begin{cases} \text{diag}(\delta_1^{l_1l_2*2}, \ldots, \delta_{p_{l_1}(p_{l_1}-1)/2}^{l_1l_2*2}), & l_1 = l_2, \\ \text{diag}(\delta_1^{l_1l_2*2}, \ldots, \delta_{p_{l_1}p_{l_2}}^{l_1l_2*2}), & l_1 > l_2, \end{cases} \tag{5.21}$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)', \quad \boldsymbol{\gamma} = (\gamma^{11}, \ldots, \gamma^{dd})', \tag{5.22}$$

$$\Phi = \text{diag}(\phi_1, \ldots, \phi_p), \quad \Theta = \text{diag}(\theta^{11}, \ldots, \theta^{dd}), \tag{5.23}$$

$$\Sigma = \text{diag}((\sigma^{11})^2, \ldots, (\sigma^{dd})^2), \tag{5.24}$$

$$\Omega_{\boldsymbol{m}} = \text{diag}(\omega_{m_1}^2, \ldots, \omega_{m_p}^2), \tag{5.25}$$

$$\begin{pmatrix} R_t^{-\frac{1}{2}}\boldsymbol{\epsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \text{N}_{2p}(\mathbf{0}_{2p}, \Psi), \ t = 1, \ldots, n-1, \tag{5.26}$$

$$\Psi = \begin{pmatrix} I_p & Q' \\ Q & \Omega \end{pmatrix}, \tag{5.27}$$

and Ω_0, the covariance matrix of the initial latent variable $\boldsymbol{h}_1$, satisfies the stationary condition $\Omega_0 = \Phi\Omega_0\Phi + \Omega$ such that

$$\text{vec}(\Omega_0) = (I_{p^2} - \Phi \otimes \Phi)^{-1}\text{vec}(\Omega). \tag{5.28}$$

$\Sigma_0 = \text{diag}((\sigma_0^{11})^2, \ldots, (\sigma_0^{dd})^2)$, the covariance matrix of the initial latent variable $\boldsymbol{g}_1$, satisfies the condition $\Sigma_0 = \Theta\Sigma_0\Theta + \Sigma$ such that

$$(\sigma_0^{l_1l_2})^2 = (1 - (\theta^{l_1l_2})^2)^{-1}(\sigma^{l_1l_2})^2, \ \ l_1 = l_2, \ldots, d, \ l_2 = 1, \ldots, d. \tag{5.29}$$

We also set κ, the variance of the initial latent variable $m_{i,1}$ ($i = 1, \ldots, p$), equal to some large known constant. The latent vector, $\boldsymbol{h}_t$, is a vector of log-variances of the returns and

[1] See Appendix 5.A.1 for the details of R_t.

the latent vector $\boldsymbol{g}_t$, is a vector of the transformed dynamic block-equicorrelation factors of $\boldsymbol{y}_t$. $\boldsymbol{\zeta}_t$, $t = 0, \ldots, n-1$, are truncated normally distributed so as to make R_{t+1} positive definite. For the identifiability, we set the diagonal elements of the covariance matrix of $\boldsymbol{\epsilon}_t$ equal to 1.

Notice that we define ρ_t^{ll}, $t = 1, \ldots n$, so as to take values on the unit interval, $(0, 1)$. It means that the (equi)correlation between the returns in the same group is restricted to be positive[2]. If an element of $RCor_t$ is not positive, we replace it to some small positive value.

We assume, for simplicity, that $\{\boldsymbol{m}_t\}_{t=1}^n$, the latent sequence of the expectations of $\boldsymbol{y}_t$, $t = 1, \ldots, n$, follows a simple random walk.

The parameter $\boldsymbol{\xi}$ corrects the bias due to the market microstructure noise and the effect of non-trading hours. If the elements of $\boldsymbol{\xi}$ are positive, the realized measure of the variance has an upward bias that may be due to the market microstructure noise and if the elements of $\boldsymbol{\xi}$ are negative, it has a downward bias due to the existence of non-trading hours. When the mean of $\boldsymbol{\xi}^{l_1 l_2 *}$ is not equal to $\mathbf{0}$. or the mean of $\boldsymbol{\lambda}^{l_1 l_2 *}$ is not equal to $\mathbf{1}$., the realized measure of the correlation has a bias due to the market microstructure noise or the non-synchronous trading effect.

[2]Suppose $d = 1$, i.e., R_t is a equicorrelation matrix. In this case, R_t is positive definite if and only if ρ_t is in $(-(p-1)^{-1}, 1)$. It means that as p becomes large, the negative region of the parameter space becomes smaller. Therefore it is reasonable to restrict the parameter space of ρ_t^{ll} to the positive region so that the parameter space is independent of p^{ll}.

5.3 Bayesian estimation

5.3.1 Priors and posterior densities

Denote $\boldsymbol{\xi}^* = \{\boldsymbol{\xi}^{l_1 l_2 *}\}_{l_1,l_2}$, $\boldsymbol{\lambda}^* = \{\boldsymbol{\lambda}^{l_1 l_2 *}\}_{l_1,l_2}$, $\Delta^* = \{\Delta^{l_1 l_2 *}\}_{l_1,l_2}$. For prior distributions of $(\boldsymbol{\xi}, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^*, \Delta, \Delta^*, \boldsymbol{\mu}, \gamma, \Phi, \Theta, \Sigma, \Omega_{\boldsymbol{m}})$, we assume that

$$\boldsymbol{\xi} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\xi}0}, S_{\boldsymbol{\xi}0}), \tag{5.30}$$

$$\boldsymbol{\xi}^{l_1 l_2 *} \sim \mathrm{N}.(\boldsymbol{m}^*_{\boldsymbol{\xi}^{l_1 l_2}0}, S^*_{\boldsymbol{\xi}^{l_1 l_2}0}), \tag{5.31}$$

$$\boldsymbol{\lambda}^{l_1 l_2 *} \sim \mathrm{N}.(\boldsymbol{m}^*_{\boldsymbol{\lambda}^{l_1 l_2}0}, S^*_{\boldsymbol{\lambda}^{l_1 l_2}0}), \tag{5.32}$$

$$\delta_i^2 \sim \mathrm{IG}(\alpha_{\delta_i 0}/2, \beta_{\delta_i 0}/2), \quad i = 1, \ldots, p, \tag{5.33}$$

$$(\delta_j^{l_1 l_2 *})^2 \sim \mathrm{IG}(\alpha^*_{\delta_j^{l_1 l_2}0}/2, \beta^*_{\delta_j^{l_1 l_2}0}/2), \tag{5.34}$$

$$\boldsymbol{\mu} \sim \mathrm{N}_p(\boldsymbol{m}_{\boldsymbol{\mu}0}, S_{\boldsymbol{\mu}0}), \tag{5.35}$$

$$\boldsymbol{\gamma} \sim \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\gamma}0}, S_{\boldsymbol{\gamma}0}), \tag{5.36}$$

$$(\phi_i + 1)/2 \sim \mathrm{Be}(a_{\phi_i}, b_{\phi_i}), \quad i = 1, \ldots, p, \tag{5.37}$$

$$(\theta^{l_1 l_2} + 1)/2 \sim \mathrm{Be}(a_{\theta^{l_1 l_2}}, b_{\theta^{l_1 l_2}}), \tag{5.38}$$

$$(\sigma^{l_1 l_2})^2 \sim \mathrm{IG}(\alpha_{\sigma^{l_1 l_2}0}/2, \beta_{\sigma^{l_1 l_2}0}/2), \tag{5.39}$$

where $\mathrm{Be}(a,b)$ denotes a beta distribution with parameters a, b and $\mathrm{IG}(\alpha, \beta)$ denotes an inverted gamma distribution with shape parameter α and scale parameter β. For a prior distribution of Ψ, we let

$$\Psi^{-1} = \begin{pmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{pmatrix},$$

where Ψ^{11}, Ψ^{12} and Ψ^{22} are $p \times p$ matrices. Noting that $\Psi^{11} = I_p + \Psi^{12}(\Psi^{22})^{-1}\Psi^{21}$, we assume

$$\Psi^{22} \sim \mathrm{W}_p(n_0, S_0), \tag{5.40}$$

$$\Psi^{21} | \Psi^{22} \sim \mathrm{N}_{p \times p}((\Psi^{22}\Delta_0), \Lambda_0 \otimes \Psi^{22}), \tag{5.41}$$

where $\mathrm{W}(n, S)$ denotes a Wishart distribution with parameters (n, S).

Then, the joint posterior density function is

$$
\begin{aligned}
&f(\boldsymbol{\vartheta}, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n) \\
&\propto \pi(\boldsymbol{\vartheta}) \times \exp\left(\sum_{t=1}^n l_t\right) \times \exp\left\{-\frac{1}{2}\sum_{t=1}^n (\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)'\Delta^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi} - \boldsymbol{h}_t)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}\sum_{t=1}^n \sum_{l_1 \geq l_2} (\boldsymbol{x}_t^{l_1 l_2 *} - \boldsymbol{\xi}^{l_1 l_2 *} - \boldsymbol{\lambda}^{l_1 l_2 *} g_t^{l_1 l_2})'\Delta^{l_1 l_2 * -1}(\boldsymbol{x}_t^{l_1 l_2 *} - \boldsymbol{\xi}^{l_1 l_2 *} - \boldsymbol{\lambda}^{l_1 l_2 *} g_t^{l_1 l_2})\right\} \\
&\quad \times |\Omega_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{h}_1 - \boldsymbol{\mu})'\Omega_0^{-1}(\boldsymbol{h}_1 - \boldsymbol{\mu})\right\} \\
&\quad \times |\Omega|^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{n-1} \{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\right] \\
&\quad \times |\Sigma_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{g}_1 - \boldsymbol{\gamma})'\Sigma_0^{-1}(\boldsymbol{g}_1 - \boldsymbol{\gamma})\right\} \\
&\quad \times |\Sigma|^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{n-1} \{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \Theta(\boldsymbol{g}_t - \boldsymbol{\gamma})\}'\Sigma^{-1}\{\boldsymbol{g}_{t+1} - \boldsymbol{\gamma} - \Theta(\boldsymbol{g}_t - \boldsymbol{\gamma})\}\right] \\
&\quad \times \exp\left(-\frac{1}{2\kappa}\boldsymbol{m}_1'\boldsymbol{m}_1\right) \times |\Omega_{\boldsymbol{m}}|^{-\frac{n-1}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{n-1} (\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)'\Omega_{\boldsymbol{m}}^{-1}(\boldsymbol{m}_{t+1} - \boldsymbol{m}_t)\right\},
\end{aligned} \tag{5.42}
$$

where

$$
\begin{aligned}
l_t = &-\frac{1}{2}\left[R_t^{-\frac{1}{2}} V_t^{-\frac{1}{2}}(\boldsymbol{y}_t - \boldsymbol{m}_t) - Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\right]' \\
&\times (I_p - Q'\Omega^{-1}Q)^{-1}\left[R_t^{-\frac{1}{2}} V_t^{-\frac{1}{2}}(\boldsymbol{y}_t - \boldsymbol{m}_t) - Q'\Omega^{-1}\{\boldsymbol{h}_{t+1} - \boldsymbol{\mu} - \Phi(\boldsymbol{h}_t - \boldsymbol{\mu})\}\right] \\
&-\frac{1}{2}\log|R_t| - \frac{1}{2}\textstyle\sum_{i=1}^p h_{it}.
\end{aligned}
$$

and $\boldsymbol{\vartheta} = (\boldsymbol{\xi}, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^*, \Delta, \Delta^*, \boldsymbol{\mu}, \boldsymbol{\gamma}, \Phi, \Theta, \Omega, Q, \Sigma, \Omega_{\boldsymbol{m}})$.

We implement the MCMC algorithm in six blocks:

1. Initialize $\{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

2. Generate $\boldsymbol{\vartheta}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n$.

3. Generate $\{\boldsymbol{h}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

4. Generate $\{\boldsymbol{g}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

5. Generate $\{\boldsymbol{m}_t\}_{t=1}^n|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\vartheta}$.

6. Go to 2.

5.3.2 Generation of latent variable $\{\boldsymbol{h}_t\}_{t=1}^n$

A simple sampling method for $\{\boldsymbol{h}_t\}_{t=1}^n$, which we use is a single-move sampler that draws a single latent variable $\boldsymbol{h}_t$ at a time given the other $\boldsymbol{h}_t$'s and the parameters. For $t = 1, \ldots, n$, we propose a candidate $\boldsymbol{h}_t | \{\boldsymbol{h}_s\}_{-t}, \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta} \sim \mathrm{N}(\boldsymbol{m}_{\boldsymbol{h}t}, S_{\boldsymbol{h}t})$, where

$$
\begin{aligned}
S_{\boldsymbol{h}t} &= \{\Phi\Omega^{-1}\Phi + (\Omega - QQ')^{-1} + \Delta^{-1}\}^{-1}, \\
\boldsymbol{m}_{\boldsymbol{h}t} &= S_{\boldsymbol{h}t}[-0.5\mathbf{1}_p + \Phi\Omega^{-1}\{\boldsymbol{h}_{t+1} - (I_p - \Phi)\boldsymbol{\mu}\} \\
&\qquad + (\Omega - QQ')^{-1}\{\boldsymbol{\mu} + \Phi(\boldsymbol{h}_{t-1} - \boldsymbol{\mu}) + QR_{t-1}^{-\frac{1}{2}}V_{t-1}^{-\frac{1}{2}}(\boldsymbol{y}_{t-1} - \boldsymbol{m}_{t-1})\} + \Delta^{-1}(\boldsymbol{x}_t - \boldsymbol{\xi})],
\end{aligned}
$$

and accept it using MH algorithm.

5.3.3 Generation of latent variable $\{\boldsymbol{g}_t\}_{t=1}^n$

For $t = 1, \ldots, n$, the conditional distribution of $g_t^{l_1 l_2}$ is

$$
\begin{aligned}
& f(g_t^{l_1 l_2} | \{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{h}_t\}_{t=1}^n, (\boldsymbol{g}_t)_{-l_1 l_2}, \{\boldsymbol{g}_s\}_{-t}, \{\boldsymbol{m}_t\}_{t=1}^n, \boldsymbol{\vartheta}) \\
& \propto \exp(l_t) \times \exp\left\{ -\frac{1}{2s^2_{g^{l_1 l_2}t}} (g_t^{l_1 l_2} - m_{g^{l_1 l_2}t})^2 \right\},
\end{aligned}
$$

where

$$
\begin{aligned}
s^2_{g^{l_1 l_2}t} &= \{(\sigma^{l_1 l_2})^{-2} + \theta^{l_1 l_2 2}(\sigma^{l_1 l_2})^{-2} + \boldsymbol{\lambda}^{l_1 l_2 *\prime}\Delta^{l_1 l_2 * -1}\boldsymbol{\lambda}^{l_1 l_2 *}\}^{-1}, && (5.43) \\
m_{g^{l_1 l_2}t} &= s^2_{g^{l_1 l_2}t}[\theta^{l_1 l_2}(\sigma^{l_1 l_2})^{-2}\{g_{t+1}^{l_1 l_2} - (1 - \theta^{l_1 l_2})\gamma^{l_1 l_2}\} + (\sigma^{l_1 l_2})^{-2}\{\gamma^{l_1 l_2}\theta^{l_1 l_2}(g_{t-1}^{l_1 l_2} - \gamma^{l_1 l_2})\} \\
&\qquad + \boldsymbol{\lambda}^{l_1 l_2 *\prime}\Delta^{l_1 l_2 * -1}(\boldsymbol{x}_t^{l_1 l_2 *} - \boldsymbol{\xi}^{l_1 l_2 *})]. && (5.44)
\end{aligned}
$$

Generate a candidate $g_t^{l_1 l_2 \dagger} \sim \mathrm{TN}_{\{R_t(\boldsymbol{g}_t) > 0\}}(m_{g^{l_1 l_2}t}, s^2_{g^{l_1 l_2}t})$ and accept it with probability $\min[1, \exp\{l_t(g_t^{l_1 l_2 \dagger}) - l_t(g_t^{l_1 l_2})\}]$.

Since the supports of $\boldsymbol{g}_t$, $t = 1, \ldots, n$, are not expressed analytically, we must construct the transition density q with caution. Note that, for example, it is not easy to conduct the random-walk MH algorithm with normal density because the normalizing constants of $q(\boldsymbol{g}_t, \boldsymbol{g}_t^c)$ and $q(\boldsymbol{g}_t^c, \boldsymbol{g}_t)$ are not equal; To calculate the acceptance rate, we must conduct a complicated numerical computation to obtain the two normalizing constants. For the transition density described above, the normalizing constants of $q(g_t^{l_1 l_2}, (g_t^{l_1 l_2})^c)$ and $q((g_t^{l_1 l_2})^c, g_t^{l_1 l_2})$ are coincided, which suggests that we do not need to know them to conduct our algorithm.

5.3.4 Generation of $\{\boldsymbol{m}_t\}_{t=1}^n$ and $\boldsymbol{\vartheta}$

Generation of $\{\boldsymbol{m}_t\}_{t=1}^n$. We generate all of $\boldsymbol{m}_t$'s simultaneously using a simulation smoother (de Jong and Shephard (1995) or Durbin and Koopman (2002)).

Generation of $\boldsymbol{\xi}^*$. The conditional distribution of $\boldsymbol{\xi}^{l_1l_2*}$ is

$$\boldsymbol{\xi}^{l_1l_2*}|\{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\lambda}^*, \Delta^* \sim \mathrm{N}_\cdot(\boldsymbol{m}^*_{\boldsymbol{\xi}^{l_1l_2}}, S^*_{\boldsymbol{\xi}^{l_1l_2}}), \tag{5.45}$$

where

$$S^*_{\boldsymbol{\xi}^{l_1l_2}} = \{S^{*-1}_{\boldsymbol{\xi}^{l_1l_2}0} + n\Delta^{l_1l_2*-1}\}^{-1}, \tag{5.46}$$

$$\boldsymbol{m}^*_{\boldsymbol{\xi}^{l_1l_2}} = S^*_{\boldsymbol{\xi}^{l_1l_2}}\left\{S^{*-1}_{\boldsymbol{\xi}^{l_1l_2}0}\boldsymbol{m}^*_{\boldsymbol{\xi}^{l_1l_2}0} + \Delta^{l_1l_2*-1}\sum_{t=1}^n(\boldsymbol{x}_t^{l_1l_2*} - \boldsymbol{\lambda}^{l_1l_2*}g_t^{l_1l_2})\right\}, \tag{5.47}$$

Generation of $\boldsymbol{\lambda}^*$. The conditional distribution of $\boldsymbol{\lambda}^{l_1l_2*}$ is

$$\boldsymbol{\lambda}^{l_1l_2*}|\{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\xi}^*, \Delta^* \sim \mathrm{N}_\cdot(\boldsymbol{m}^*_{\boldsymbol{\lambda}^{l_1l_2}}, S^*_{\boldsymbol{\lambda}^{l_1l_2}}), \tag{5.48}$$

where

$$S^*_{\boldsymbol{\lambda}^{l_1l_2}} = \left\{S^{*-1}_{\boldsymbol{\lambda}^{l_1l_2}0} + \sum_{t=1}^n(g_t^{l_1l_2})^2\Delta^{l_1l_2*-1}\right\}^{-1}, \tag{5.49}$$

$$\boldsymbol{m}^*_{\boldsymbol{\lambda}^{l_1l_2}} = S^*_{\boldsymbol{\lambda}^{l_1l_2}}\left\{S^{*-1}_{\boldsymbol{\lambda}^{l_1l_2}0}\boldsymbol{m}^*_{\boldsymbol{\lambda}^{l_1l_2}0} + \Delta^{l_1l_2*-1}\sum_{t=1}^n(\boldsymbol{x}_t^{l_1l_2*} - \boldsymbol{\xi}^{l_1l_2*})g_t^{l_1l_2}\right\}. \tag{5.50}$$

Generation of Δ^*. The conditional distribution of $\delta_j^{l_1l_2*2}$ is given by

$$\delta_j^{l_1l_2*2}|\{\boldsymbol{x}_t^*\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\xi}^*, \boldsymbol{\lambda}^* \sim \mathrm{IG}(\alpha^*_{\delta_j^{l_1l_2}}/2, \beta^*_{\delta_j^{l_1l_2}}/2), \tag{5.51}$$

where $\alpha^*_{\delta_j^{l_1l_2}} = \alpha^*_{\delta_j^{l_1l_2}0} + n$,

$$\beta^*_{\delta_j^{l_1l_2}} = \beta^*_{\delta_j^{l_1l_2}0} + \sum_{t=1}^n(x_{jt}^{l_1l_2*} - \xi_j^{l_1l_2*} - \lambda_j^{l_1l_2*}g_t^{l_1l_2})^2. \tag{5.52}$$

Generation of $\boldsymbol{\gamma}$. The conditional distribution of $\boldsymbol{\gamma}$ is

$$\boldsymbol{\gamma}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \Theta, \Sigma \sim \mathrm{N}_q(\boldsymbol{m}_{\boldsymbol{\gamma}}, S_{\boldsymbol{\gamma}}), \tag{5.53}$$

where

$$S_{\boldsymbol{\gamma}} = \{S_{\boldsymbol{\gamma}0}^{-1} + \Sigma_0^{-1} + (n-1)(I_q - \Theta)\Sigma^{-1}(I_q - \Theta)\}^{-1}, \tag{5.54}$$

$$\boldsymbol{m}_{\boldsymbol{\gamma}} = S_{\boldsymbol{\gamma}}\left[S_{\boldsymbol{\gamma}0}^{-1}\boldsymbol{m}_{\boldsymbol{\gamma}0} + \Sigma_0^{-1}\boldsymbol{g}_1 + (I_q - \Theta)\Sigma^{-1}\left\{\sum_{t=2}^{n}\boldsymbol{g}_t - \Theta\sum_{t=1}^{n-1}\boldsymbol{g}_t\right\}\right]. \tag{5.55}$$

Generation of Θ*.* The conditional posterior density of $\boldsymbol{\theta} = \Theta\mathbf{1}_p$ is given by

$$f(\boldsymbol{\theta}|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\gamma}, \Sigma) \propto g(\boldsymbol{\theta}) \times \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{m}_{\boldsymbol{\theta}})' S_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta} - \boldsymbol{m}_{\boldsymbol{\theta}})\right\}, \tag{5.56}$$

where

$$S_{\boldsymbol{\theta}} = \left[\sum_{t=1}^{n-1}\{(\boldsymbol{g}_t - \boldsymbol{\gamma})(\boldsymbol{g}_t - \boldsymbol{\gamma})'\} \odot \Sigma^{-1}\right]^{-1}, \tag{5.57}$$

$$(\boldsymbol{m}_{t,\boldsymbol{\theta}})_i = \left\{\Sigma^{-1}(\boldsymbol{g}_{t+1} - \boldsymbol{\gamma})(\boldsymbol{g}_t - \boldsymbol{\gamma})'\right\}_{i,i}, \quad \boldsymbol{m}_{\boldsymbol{\theta}} = S_{\boldsymbol{\theta}}\sum_{t=1}^{n-1}\boldsymbol{m}_{t,\boldsymbol{\theta}}, \tag{5.58}$$

$$g(\boldsymbol{\theta}) = \prod_{l_1 \geq l_2}\pi(\theta^{l_1 l_2}) \times |\Sigma_0|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{g}_1 - \boldsymbol{\gamma})'\Sigma_0^{-1}(\boldsymbol{g}_1 - \boldsymbol{\gamma})\right\}, \tag{5.59}$$

and $\odot$ denotes Hadamard product. Generate a candidate $\boldsymbol{\theta}^\dagger \sim \mathrm{TN}_{(-1,1)}(\boldsymbol{m}_{\boldsymbol{\theta}}, S_{\boldsymbol{\theta}})$ and accept it with probability $\min[1, g(\boldsymbol{\theta}^\dagger)/g(\boldsymbol{\theta})]$.

Generation of Σ*.* The conditional distribution of $(\sigma^{l_1 l_2})^2$ is

$$(\sigma^{l_1 l_2})^2|\{\boldsymbol{y}_t\}_{t=1}^n, \{\boldsymbol{g}_t\}_{t=1}^n, \boldsymbol{\gamma}, \Theta \sim \mathrm{IG}(\alpha_{\sigma^{l_1 l_2}}/2, \beta_{\sigma^{l_1 l_2}}/2), \tag{5.60}$$

where $\alpha_{\sigma^{l_1 l_2}} = \alpha_{\sigma^{l_1 l_2}0} + n$ and

$$\beta_{\sigma_j^{l_1 l_2}} = \beta_{\sigma^{l_1 l_2}0} + (g_t^{l_1 l_2} - \gamma^{l_1 l_2})^2\{1 - (\theta^{l_1 l_2})^2\} + \sum_{t=1}^{n-1}\{g_{t+1}^{l_1 l_2} - \gamma^{l_1 l_2} - \theta^{l_1 l_2}(g_t^{l_1 l_2} - \gamma^{l_1 l_2})\}^2. \tag{5.61}$$

We implement the sampling procedure for $\boldsymbol{\xi}, \Delta, \boldsymbol{\mu}, \gamma, \Phi, \theta, \Omega, Q, \sigma^2, \Omega_{\boldsymbol{m}}$ as mentioned in the last chapter.

5.4 Illustrative example using simulated data

This section illustrates the extended DESV model using simulated data. We consider a two-block case ($d = 2,\ q = 3$) and each block consists of two components ($p_1 = p_2 = 2$, $p = 4$).

Using the following parameters based on our empirical studies in Section 5.5,

$$
\begin{aligned}
&\boldsymbol{\xi}_{\ddagger} = \mathbf{0}_4,\ \boldsymbol{\xi}_{\ddagger}^{11*} = \boldsymbol{\xi}_{\ddagger}^{22*} = 0,\ \boldsymbol{\xi}_{\ddagger}^{21*} = \mathbf{0}_4,\ \boldsymbol{\lambda}_{\ddagger}^{11*} = \boldsymbol{\lambda}_{\ddagger}^{22*} = 1,\ \boldsymbol{\lambda}_{\ddagger}^{21*} = \mathbf{1}_4,\\
&\Delta_{\ddagger} = 0.1I_4,\ \Delta_{\ddagger}^{11*} = \Delta_{\ddagger}^{22*} = 0.1, \Delta_{\ddagger}^{21*} = 0.1I_4,\\
&\boldsymbol{\mu}_{\ddagger} = \mathbf{0}_4,\ \boldsymbol{\gamma}_{\ddagger} = (1, 0.5, 1)',\ \Phi_{\ddagger} = 0.97I_4,\ \Theta_{\ddagger} = 0.97I_3,\\
&\Omega_{\ddagger} = 0.1I_4 + 0.1J_{44},\ Q_{\ddagger} = O_{4\times 4},\ \Sigma_{\ddagger} = 0.01I_3,\ \Omega_{\boldsymbol{m}*} = 0.001I_4,
\end{aligned}
$$

we generate 2,500 observations ($n = 2500$). For prior distributions, we assume

$$
\begin{aligned}
&\boldsymbol{\xi} \sim \mathrm{N}_4(\boldsymbol{\xi}_{\ddagger}, 100I_4),\ \boldsymbol{\xi}^{11*} \sim \mathrm{N}(\boldsymbol{\xi}_{\ddagger}^{11*}, 100),\ \boldsymbol{\xi}^{22*} \sim \mathrm{N}(\boldsymbol{\xi}_{\ddagger}^{22*}, 100),\ \boldsymbol{\xi}^{21*} \sim \mathrm{N}_4(\boldsymbol{\xi}_{\ddagger}^{21*}, 100I_4),\\
&\boldsymbol{\lambda}^{11*} \sim \mathrm{N}(\boldsymbol{\lambda}_{\ddagger}^{11*}, 100),\ \boldsymbol{\lambda}^{22*} \sim \mathrm{N}(\boldsymbol{\lambda}_{\ddagger}^{22*}, 100),\ \boldsymbol{\lambda}^{21*} \sim \mathrm{N}_4(\boldsymbol{\lambda}_{\ddagger}^{21*}, 100I_4),\\
&\delta_i^2 \sim \mathrm{IG}(5, 3\delta_{i\ddagger}^2),\ \ i = 1, \ldots, 4,\\
&\delta_1^{11*2} \sim \mathrm{IG}(5, 3\delta_{\ddagger}^{11*2}),\ \ \delta_1^{22*2} \sim \mathrm{IG}(5, 3\delta_{\ddagger}^{22*2}),\ \ \delta_j^{21*2} \sim \mathrm{IG}(5, 3\delta_{\ddagger}^{21*2}),\ j = 1, \ldots, 4\\
&\boldsymbol{\mu} \sim \mathrm{N}(\boldsymbol{\mu}_{\ddagger}, 100I_4),\ \boldsymbol{\gamma} \sim \mathrm{N}(\boldsymbol{\gamma}_{\ddagger}, 100I_3),\ \frac{\phi_i + 1}{2} \sim \mathrm{Be}(20, 1.5),\ \ i = 1, \ldots, 4,\\
&\frac{\theta^{l_1 l_2} + 1}{2} \sim \mathrm{Be}(20, 1.5),\ \ (\sigma^{l_1 l_2})^2 \sim \mathrm{IG}(5, 3(\sigma^{l_1 l_2 *})_{\ddagger}^2),\ \ l_1 = 1, 2,\ l_2 = l_1, 2,\\
&\Psi^{22} \sim \mathrm{W}(12, 12^{-1}\Omega_{\ddagger}),\ \Psi^{21}|\Psi^{22} \sim \mathrm{N}_{4\times 4}(O, 10I_4 \otimes \Psi^{22}),\ \ \omega_{m_i 0}^2 \sim \mathrm{IG}(5, 3\omega_{m_i \ddagger}^2),\ i = 1, \ldots, 4.
\end{aligned}
$$

Using the sampling method described in the last section, we generate 300,000 MCMC samples after discarding the first 10,000 samples as the burn-in period.

Tables 1 and 2 report the true values, posterior means, 95% credible intervals and estimated inefficiency factors (IF)[3]. The posterior means are all close to the true values, which suggests that our proposed algorithm works well. Note that the acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{\boldsymbol{g}_t\}_{t=1}^n$ in the independent MH algorithms are 0.789 and 0.930, respectively. It indicates that the generated candidates are accepted with relatively high probability and our sampling algorithm also works well.

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

Table 5.1: Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

	True	Mean	95% interval	IF
ξ_1	0	0.050	(-0.008, 0.108)	235
ξ_2	0	-0.017	(-0.075, 0.041)	188
ξ_3	0	-0.013	(-0.070, 0.045)	418
ξ_4	0	0.014	(-0.042, 0.071)	250
ξ_1^{11*}	0	0.108	(-0.179, 0.322)	2201
ξ_1^{21*}	0	0.044	(-0.085, 0.138)	952
ξ_2^{21*}	0	-0.008	(-0.148, 0.093)	940
ξ_3^{21*}	0	0.046	(-0.087, 0.142)	951
ξ_4^{21*}	0	0.023	(-0.111, 0.120)	951
ξ_1^{22*}	0	0.030	(-0.187, 0.203)	1780
λ_1^{11*}	1	0.976	(0.769, 1.291)	2118
λ_1^{21*}	1	0.970	(0.830, 1.166)	799
λ_2^{21*}	1	1.047	(0.896, 1.259)	791
λ_3^{21*}	1	0.992	(0.849, 1.193)	805
λ_4^{21*}	1	1.002	(0.858, 1.205)	806
λ_1^{22*}	1	0.930	(0.784, 1.119)	1490
δ_1^2	0.1	0.106	(0.093, 0.119)	25
δ_2^2	0.1	0.100	(0.088, 0.113)	25
δ_3^2	0.1	0.101	(0.089, 0.114)	12
δ_4^2	0.1	0.099	(0.087, 0.112)	13
δ_1^{11*2}	0.1	0.103	(0.096, 0.111)	18
δ_1^{21*2}	0.1	0.099	(0.093, 0.106)	2
δ_2^{21*2}	0.1	0.103	(0.097, 0.110)	2
δ_3^{21*2}	0.1	0.099	(0.092, 0.105)	2
δ_4^{21*2}	0.1	0.097	(0.091, 0.104)	1
δ_1^{22*2}	0.1	0.097	(0.090, 0.104)	10
μ_1	0	-0.386	(-1.097, 0.336)	2
μ_2	0	-0.190	(-0.751, 0.372)	2
μ_3	0	-0.148	(-0.856, 0.567)	5
μ_4	0	-0.327	(-1.040, 0.398)	3
γ^{11}	1	0.867	(0.718, 1.008)	469
γ^{21}	0.5	0.510	(0.384, 0.636)	184
γ^{22}	1	0.982	(0.746, 1.222)	247
ϕ_1	0.97	0.975	(0.966, 0.982)	4
ϕ_2	0.97	0.969	(0.960, 0.978)	6
ϕ_3	0.97	0.975	(0.966, 0.983)	3
ϕ_4	0.97	0.975	(0.966, 0.983)	3
θ^{11}	0.97	0.959	(0.943, 0.973)	31
θ^{21}	0.97	0.962	(0.949, 0.974)	5
θ^{22}	0.97	0.979	(0.969, 0.988)	10

Table 5.2: Posterior means, 95% credible intervals and inefficiency factors.

	True	Mean	95% interval	IF
Ω_{11}	0.2	0.203	(0.183, 0.224)	16
Ω_{21}	0.1	0.100	(0.087, 0.114)	4
Ω_{31}	0.1	0.102	(0.088, 0.115)	9
Ω_{41}	0.1	0.099	(0.086, 0.113)	5
Ω_{22}	0.2	0.190	(0.171, 0.210)	16
Ω_{32}	0.1	0.103	(0.090, 0.117)	7
Ω_{42}	0.1	0.104	(0.091, 0.118)	6
Ω_{33}	0.2	0.194	(0.174, 0.215)	10
Ω_{43}	0.1	0.099	(0.086, 0.113)	4
Ω_{44}	0.2	0.197	(0.178, 0.217)	13
Q_{11}	0	0.018	(-0.012, 0.050)	17
Q_{21}	0	-0.019	(-0.048, 0.011)	26
Q_{31}	0	0.000	(-0.029, 0.031)	28
Q_{41}	0	-0.004	(-0.034, 0.027)	26
Q_{12}	0	0.004	(-0.022, 0.031)	7
Q_{22}	0	0.001	(-0.024, 0.026)	12
Q_{32}	0	0.023	(-0.003, 0.049)	28
Q_{42}	0	0.015	(-0.011, 0.041)	75
Q_{13}	0	-0.003	(-0.040, 0.037)	41
Q_{23}	0	0.006	(-0.030, 0.043)	20
Q_{33}	0	0.003	(-0.032, 0.039)	157
Q_{43}	0	-0.027	(-0.066, 0.012)	85
Q_{14}	0	-0.021	(-0.062, 0.021)	115
Q_{24}	0	-0.009	(-0.045, 0.029)	69
Q_{34}	0	-0.018	(-0.054, 0.019)	124
Q_{44}	0	0.007	(-0.029, 0.044)	19
$(\sigma^{11})^2$	0.01	0.011	(0.005, 0.019)	1407
$(\sigma^{21})^2$	0.01	0.011	(0.007, 0.015)	593
$(\sigma^{22})^2$	0.01	0.012	(0.007, 0.018)	970
$\omega_{m_1}^2 \times 10^3$	1	0.850	(0.584, 1.212)	117
$\omega_{m_2}^2 \times 10^3$	1	0.964	(0.599, 1.480)	124
$\omega_{m_3}^2 \times 10^3$	1	0.616	(0.395, 0.930)	88
$\omega_{m_4}^2 \times 10^3$	1	1.076	(0.726, 1.542)	108

5.5 Empirical Study

5.5.1 Data

We use daily close-to-close returns and the associated realized measure of the covariance for some of the most liquid stocks in the Dow Jones Industrial Average (DJIA) index, obtained from Oxford-Man Institute[4]. These are: (1) Bank of America (BAC), (2) JP Morgan (JPM), (3) International Business Machines (IBM), (4) Microsoft (MSFT). The sample period is from February 1, 2001 to December 31, 2009 (2242 observations in total). We divide them into two groups: Group 1 (Finance) consists of (1) and (2), and Group 2 (Information Technology) consists of (3) and (4). In the following, we apply our proposed model with two blocks ($d = 2,\ q = 3$) to the data. Note that each block corresponds to the group based on the industrial sector described above and consists of two components ($p_1 = p_2 = 2$, $p = 4$). Also note that the asset return is calculated as $y_t = (\log p_t - \log p_{t-1}) \times 100$, where p_t is the asset price at time t.

5.5.2 Estimation results

Using the same prior distributions for the parameters as in Section 5.4, we implement the MCMC algorithm to conduct a Bayesian inference on the parameters of interest. We generate 300,000 MCMC samples from the posterior distributions of the parameters in the model after discarding the first 100,000 samples as the burn-in period.

Tables 5.3 and 5.4 report the summary of the estimation results. Figure 5.1 shows the posterior means of $\exp(h_{i,t}/2)$, square root of the estimated time-varying variances and Figure 5.2 shows the posterior means of $\rho_t^{l_1 l_2}$, the dynamic block-equicorrelation factors of $\boldsymbol{y}_t$.

The posterior probability with which the bias adjustment term ξ_1 is negative is over 0.975 and this is similar for ξ_2, ξ_3 and ξ_4. It suggests that the realized measure of variances which we use tends to have a downward bias due to the non-trading hours and that the effect of non-trading hours is more important than that of microstructure noise. Note that the averages of the posterior means of $\{\boldsymbol{g}_t\}_{t=1}^n$ are $(0.880, 0.969, 0.177)'$. The posterior distributions of $\boldsymbol{\xi}^*$ and $\boldsymbol{\lambda}^*$ suggest that the realized measure of the correlation which we use tends to have a downward bias as described in Section 4.1 (Epps effect)[5]. These estimation results show that

[4]See Noureldin, Shephard, and Sheppard (2012) for the detailed explanation.

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

Table 5.3: Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

	Mean	95% interval	IF
ξ_1	-0.465	(-0.522, -0.407)	1913
ξ_2	-0.433	(-0.490, -0.377)	1617
ξ_3	-0.462	(-0.519, -0.404)	1621
ξ_4	-0.518	(-0.576, -0.461)	743
ξ_1^{11*}	-1.180	(-1.450, -0.960)	1434
ξ_1^{21*}	0.131	(0.023, 0.220)	1128
ξ_2^{21*}	0.083	(-0.026, 0.174)	1131
ξ_3^{21*}	0.169	(0.065, 0.256)	1116
ξ_4^{21*}	0.134	(0.026, 0.223)	1150
ξ_1^{22*}	-0.428	(-0.520, -0.344)	868
λ_1^{11*}	1.185	(1.002, 1.412)	1245
λ_1^{21*}	0.686	(0.609, 0.785)	1202
λ_2^{21*}	0.707	(0.628, 0.807)	1217
λ_3^{21*}	0.658	(0.582, 0.754)	1182
λ_4^{21*}	0.679	(0.602, 0.777)	1219
λ_1^{22*}	0.565	(0.472, 0.681)	978
δ_1^2	0.081	(0.073, 0.090)	103
δ_2^2	0.082	(0.074, 0.090)	203
δ_3^2	0.099	(0.089, 0.109)	149
δ_4^2	0.098	(0.088, 0.108)	67
δ_1^{11*2}	0.512	(0.477, 0.549)	53
δ_1^{21*2}	0.062	(0.056, 0.067)	15
δ_2^{21*2}	0.056	(0.051, 0.061)	14
δ_3^{21*2}	0.070	(0.064, 0.076)	13
δ_4^{21*2}	0.066	(0.060, 0.071)	13
δ_1^{22*2}	0.502	(0.454, 0.550)	460
μ_1	0.714	(0.190, 1.233)	24
μ_2	1.091	(0.644, 1.532)	29
μ_3	0.544	(0.284, 0.804)	84
μ_4	0.876	(0.596, 1.153)	34
γ^{11}	0.916	(0.643, 1.172)	304
γ^{21}	1.010	(0.938, 1.083)	551
γ^{22}	0.191	(-0.052, 0.422)	290
ϕ_1	0.969	(0.961, 0.977)	158
ϕ_2	0.963	(0.954, 0.972)	173
ϕ_3	0.947	(0.934, 0.959)	210
ϕ_4	0.951	(0.939, 0.963)	118
θ^{11}	0.976	(0.961, 0.988)	122
θ^{21}	0.634	(0.593, 0.674)	9
θ^{22}	0.901	(0.847, 0.945)	887

Table 5.4: Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

	Mean	95% interval	IF
Ω_{11}	0.147	(0.131, 0.163)	134
Ω_{21}	0.142	(0.129, 0.156)	27
Ω_{31}	0.104	(0.093, 0.115)	93
Ω_{41}	0.099	(0.089, 0.110)	57
Ω_{22}	0.146	(0.131, 0.162)	150
Ω_{32}	0.112	(0.100, 0.124)	49
Ω_{42}	0.106	(0.095, 0.118)	112
Ω_{33}	0.105	(0.091, 0.119)	233
Ω_{43}	0.095	(0.084, 0.106)	123
Ω_{44}	0.100	(0.086, 0.114)	125
Q_{11}	0.066	(0.044, 0.088)	73
Q_{21}	0.054	(0.032, 0.076)	124
Q_{31}	0.053	(0.033, 0.074)	80
Q_{41}	0.043	(0.022, 0.063)	92
Q_{12}	0.014	(-0.010, 0.040)	383
Q_{22}	0.007	(-0.016, 0.031)	101
Q_{32}	0.039	(0.016, 0.061)	180
Q_{42}	0.020	(-0.001, 0.043)	336
Q_{13}	0.003	(-0.026, 0.034)	461
Q_{23}	0.006	(-0.021, 0.034)	241
Q_{33}	-0.036	(-0.062, -0.010)	210
Q_{43}	0.002	(-0.022, 0.028)	108
Q_{14}	0.018	(-0.011, 0.049)	455
Q_{24}	-0.031	(-0.060, -0.003)	193
Q_{34}	-0.011	(-0.039, 0.019)	332
Q_{44}	-0.011	(-0.036, 0.015)	115
$(\sigma^{11})^2$	0.017	(0.009, 0.030)	820
$(\sigma^{21})^2$	0.114	(0.086, 0.142)	1130
$(\sigma^{22})^2$	0.205	(0.091, 0.375)	1530
$\omega_{m_1}^2 \times 10^3$	0.232	(0.121, 0.419)	174
$\omega_{m_2}^2 \times 10^3$	0.257	(0.240, 0.481)	210
$\omega_{m_3}^2 \times 10^3$	0.245	(0.123, 0.459)	151
$\omega_{m_4}^2 \times 10^3$	0.243	(0.125, 0.447)	182

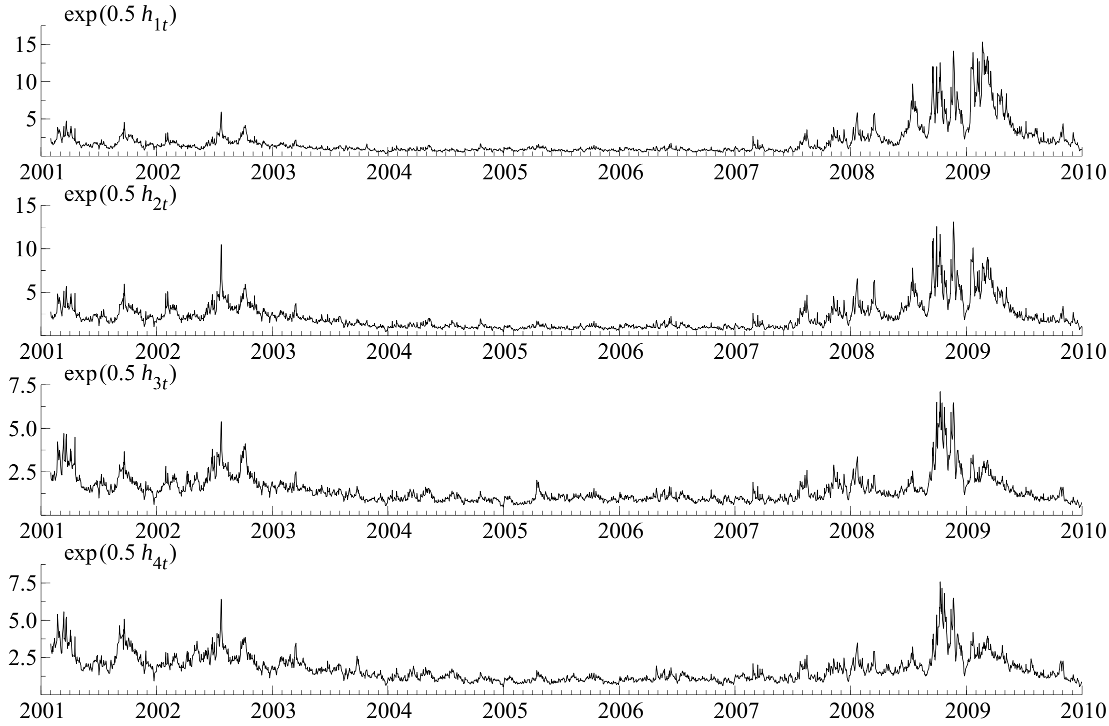

Figure 5.1: Posterior means of square root of the variances.

our model can detect and eliminate the bias of the realized measure owing to the problems in the financial market.

Among the posterior means of μ's, μ_2 is the largest and μ_3 is the smallest, but the difference between them is not huge and the levels of the volatilities are not so different from one another. We find that the posterior means of Ω_{11} and Ω_{22} are apparently larger than those of Ω_{33} and Ω_{44} and it is consistent with the observed fluctuations shown in Figure 5.1. The credible intervals of off-diagonal elements of Ω are positive and do not include zero, which means that unobserved volatilities are correlated positively each other.

The posterior means of autoregressive coefficients (ϕ_i's) are very high (over 0.94), which shows that the log volatilities follow highly persistent processes. In addition, Figure 5.1 indicates the comovement of the volatilities. The trajectories sharply increased in September 2008, corresponding to the financial crisis during which Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008). We also observe the increase in July 2002 resulted from the market turmoil during which Worldcom filed for Chapter 11 bankruptcy

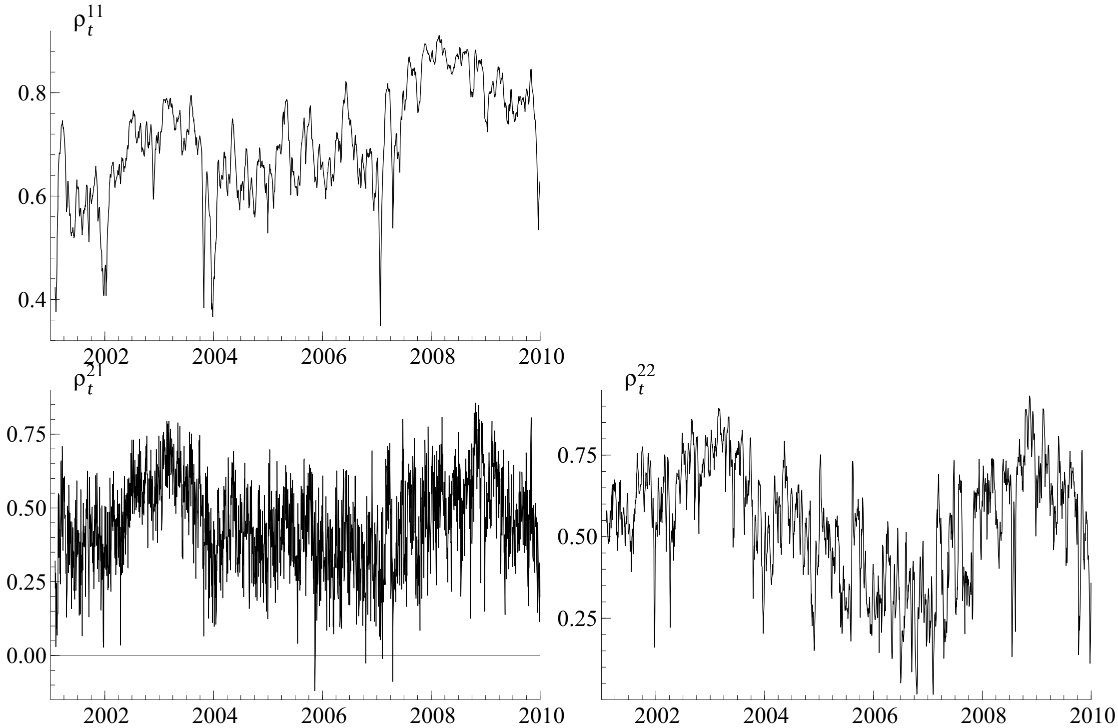

Figure 5.2: Posterior means of the dynamic block-equicorrelations.

protection (July 21, 2002). The increases in April 2001 and in September 2001 are due to the collapse of dot com bubble and to September 11 attacks.

The posterior means of autoregressive coefficients of θ^{11} and θ^{22} are very high (over 0.9) and the diagonal-block-equicorrelation parameter is highly persistent, too. However, that of θ^{21} is not so high and the off-diagonal-block-equicorrelation parameter is not highly persistent. Figure 5.2 shows that ρ_t^{11}, the block-equicorrelation factor of Group 1 is at a higher level than the other factors over the period and the trajectory is smooth but far from constant. ρ_t^{21}, the off-diagonal-block-equicorrelation factor (Group 1 - Group 2) varies at a relatively low level (and is negative at some time points) and greatly, which means that it is time-varying and far from constant. ρ_t^{22}, the block-equicorrelation factor of Group 2 varies at a slightly higher level than ρ_t^{21} and greatly, which means that it is time-varying and far from constant. It seems to be consistent with the values of $\gamma^{l_1 l_2}$'s and $(\sigma^{l_1 l_2})^2$'s reported in Tables 5.3 and 5.4.

Figure 5.3 shows the posterior means of dynamic correlations between the return of i-th

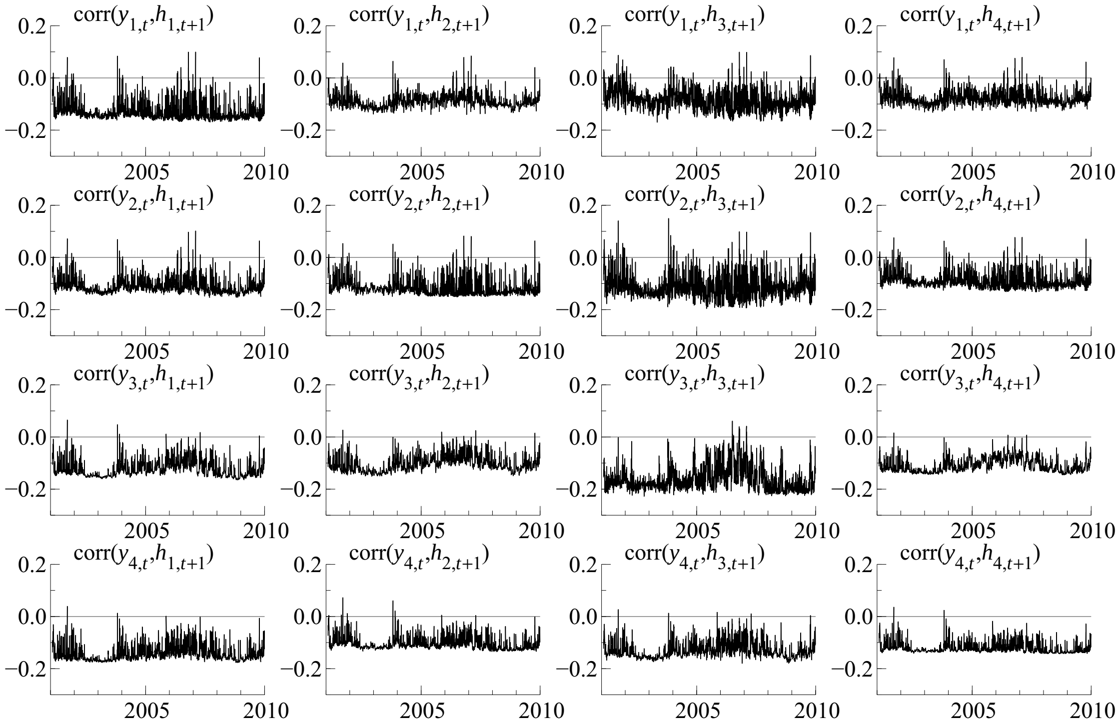

Figure 5.3: Posterior means of correlation of $(y_{it}, h_{j,t+1})$.

asset at time t $(y_{i,t})$ and the j-th log variance at time $t+1$ $(h_{j,t+1})$. As stated in Section 5.2, Q is the covariance between $\boldsymbol{z}_t$, the transformed error term in the observed equation of our model and $\boldsymbol{\eta}_{t+1}$, the error term in the state equation for $t = 1, \ldots, n-1$. To verify the existence of the cross leverage effects, we should transform Q using R_t and Ω to obtain the posterior means of the (time-varying) correlations between $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\eta}_{t+1}$. Their trajectories are negative at almost all of the time points as expected and but not far from zero, which indicates the existence of the weak cross leverage effect. While the existence of strong leverage effect is frequently reported in the analysis of the stock index, Takeuchi-Nogimori (2012) shows that the variance of the common stock has a weak asymmetry. It is consistent with our estimation results about the asymmetry stated above.

We note that the leverage effect of asset 3 is the strongest (the mean is -0.164) and the cross leverage effect of asset 1 to asset 4 is the weakest (the mean is -0.078) among the cross leverage effects. In addition, the correlation between y_{1t} and $h_{4,t+1}$ is apparently weaker than the correlation between y_{4t} and $h_{1,t+1}$. It indicates that cross leverage effects from asset i to

j, $i \neq j$, are not symmetric.

In conclusion, using our multivariate RSV model, we can therefore detect the volatility clustering, the dynamic block-equicorrelation and the (weak) cross leverage effects of the four stocks.

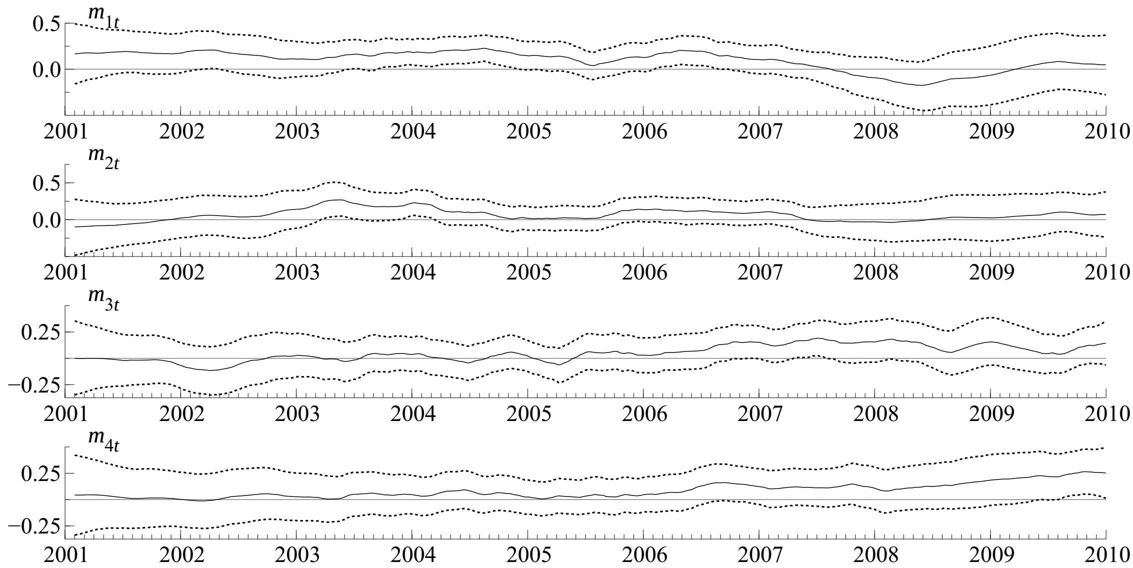

Figure 5.4: Posterior means (solid lines) and 95% credible intervals (between the two outmost dotted lines) of the expectations.

Figure 5.4 shows the posterior means of the expectations of the asset returns. We find that the 95% credible intervals include zero at almost all of the time points. It seems to fluctuate slowly, which is consistent with the small values of the variances ($\omega_{m_i}^2$'s) reported in Table 5.2.

Note that the acceptance rates for $\{\boldsymbol{h}_t\}_{t=1}^n$ and $\{\boldsymbol{g}_t\}_{t=1}^n$ in the independent MH algorithms are 0.657 and 0.724, respectively. It indicates that the generated candidates are accepted with relatively high probability and our sampling algorithm works well.

5.6 Conclusion

The MSV model of Chapter 3, which incorporates the SV model with dynamic equicorrelation and cross leverage effect, is extended with the dynamic block-equicorrelation structure. We also extend the model and propose simultaneous modeling of the multivariate daily returns and the related realized measure in line with chapter 4. With the Bayesian estimation scheme via Markov chain Monte Carlo method, it makes it possible to obtain the estimation results for the parameters in our proposed model. Numerical examples are provided and the proposed model is applied to the multivariate stock data. We find the persistence in the volatilities and the dynamic diagonal-block-equicorrelations and the existence of weak cross leverage effects. The biases in the realized measure owing to market microstructure noise, non-trading hours and non-synchronous trading can also be estimated.

Appendix

5.A.1 Block-equicorrelation matrix

Suppose that a random variable vector has a block equicorrelation structure where the correlation matrix takes the form

$$R = \begin{bmatrix} (1-\rho^{11})I_{p_1} + \rho^{11}J_{p_1p_1} & \rho^{12}J_{p_1p_2} & \cdots & \rho^{1d}J_{p_1p_d} \\ \rho^{12}J_{p_2p_1} & (1-\rho^{22})I_{p_2} + \rho^{22}J_{p_2p_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho^{d-1,d}J_{p_{d-1}p_d} \\ \rho^{1d}J_{p_dp_1} & \cdots & \rho^{d-1,d}J_{p_dp_{d-1}} & (1-\rho^{dd})I_{p_d} + \rho^{dd}J_{p_dp_d} \end{bmatrix}.$$

The determinant is given by

$$|R| = \prod_{l=1}^{d}(1-\rho^{ll})^{p_l-1}|R_A + R_B P|,$$

where

$$R_A = \mathrm{diag}(1-\rho^{11}, 1-\rho^{22}, \ldots, 1-\rho^{dd}),$$
$$R_B = \begin{pmatrix} \rho^{11} & \cdots & \rho^{1d} \\ \vdots & \ddots & \vdots \\ \rho^{1d} & \cdots & \rho^{dd} \end{pmatrix},$$
$$P = \mathrm{diag}(p_1, \ldots, p_d),$$

and the inverse is given by

$$R^{-1} = \begin{bmatrix} (1-\rho^{11})^{-1}I_{p_1} + r^{11}J_{p_1p_1} & r^{12}J_{p_1p_2} & \cdots & r^{1d}J_{p_1p_d} \\ r^{12}J_{p_2p_1} & (1-\rho^{22})^{-1}I_{p_2} + r^{22}J_{p_2p_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & r^{d-1,d}J_{p_{d-1}p_d} \\ r^{1d}J_{p_dp_1} & \cdots & r^{d-1,d}J_{p_dp_{d-1}} & (1-\rho^{dd})^{-1}I_{p_d} + r^{dd}J_{p_dp_d} \end{bmatrix}.$$

where

$$R_C = -(R_A + R_B P)^{-1} R_B R_A^{-1}, \quad R_C = \begin{pmatrix} r^{11} & \cdots & r^{1d} \\ \vdots & \ddots & \vdots \\ r^{1d} & \cdots & r^{dd} \end{pmatrix}.$$

We can easily confirm that the inverse is a symmetric matrix.

We note that the eigenvalues of R are $1-\rho^{ll}$ (multiplicity p_d-1) for $l = 1, \ldots, d$, and those of $R_A + R_B P$. We define $R^{1/2}$ using the spectral decomposition of R. Alternatively we could decompose as $R = R_c R_c'$ (e.g., Choleski decomposition), but the spectral decomposition is advantageous in that it does not depend on the order of the random variables in the vector.

For more details about 2-block-equicorrelation matrix, see Engle and Kelly (2012).

Lucas, Schwaab, and Zhang (2012) introduce a different type of block-equicorrelation structure and propose the dynamic generalized hyperbolic (GH) skew-t-error model with generalized autoregressive score (GAS) equicorrelation/block-equicorrelation structure.

Bibliography

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

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

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

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

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

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

Bauwens, L., C. M. Hafner, and S. Laurent (Eds.) (2012). *Handbook of Volatility Models and Their Applications*. Wiley.

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

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2 ed.). Springer.

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

Chen, M., Q. Shao, and J. G. Ibrahim (2000). *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag.

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

Clements, A. E., C. A. Coleman-Fenn, and D. R. Smith (2011). Forecasting equicorrelation. NCER Working Paper Series 72.

Dagpunar, J. S. (1989). An easily implemented generalized inverse Gaussian generator. *Communications in Statistics - Simulation and Computation 18*, 703–710.

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

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

De Rossi, G. and A. Harvey (2009). Quantiles, expectiles and splines. *Journal of Econometrics 152*, 179–185.

Dobrev, D. P. and P. J. Szerszen (2010). The information content of high-frequency data for estimating equity return models and forecasting risk. Working Paper.

Doucet, A., N. de Freitas, and N. Gordon (Eds.) (2001). *Sequential Monte Carlo Methods in Practice*. Springer.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

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

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

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business &*

Economic Statistics 20, 339–350.

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

Engle, R. F. (Ed.) (1995). *ARCH: Selected Readings*. Oxford University Press.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*, 122–150.

Engle, R. F. and S. Manganelli (2004). CAViaR: Conditional Autoregressive Value at Risk by regression quantiles. *Journal of Business & Economic Statistics 22*, 367–381.

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

Gelman, A., J. B. Carlin, H. Stern, and D. B. Rubin (2003). *Bayesian Data Analysis* (2 ed.). Chapman and Hall/CRC.

Gerlach, R. H., C. W. S. Chen, and N. Y. C. Chan (2011). Bayesian time-varying quantile forecasting for Value-at-Risk in financial markets. *Journal of Business & Economic Statistics 29*, 481–492.

Gourieroux, C. and J. Jasiak (2008). Dynamic quantile models. *Journal of Econometrics 147*, 198–205.

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

Hafner, C. M. and O. Reznikova (2012). On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis 56*, 3533–3545.

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

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

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

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

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

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

Ishihara, T., Y. Omori, and M. Asai (2011). Matrix exponential stochastic volatility with cross leverage. Discussion Paper CIRJE-F-812.

Jin, X. and Y. Tang (2009). The conditional dynamic dependence between herding and return: Evidence from US equitty market. Working Paper.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies 65*, 361–393.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

Koenker, R. and J. A. F. Machado (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association 94*, 1296–1310.

Koenker, R. and Z. Xiao (2006). Quantile autoregression. *Journal of the American Statistical Association 101*, 980–990.

Kohn, R. and C. F. Ansley (1987). A new algorithm for spline smoothing based on smoothing a stochastic process. *SIAM Journal on Scientific and Statistical Computing 8*, 33–48.

Koop, G., D. J. Poirier, and J. L. Tobias (2007). *Bayesian Econometric Methods*. Cambridge University Press.

Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika 80*, 117–126.

Koopman, S. J. and M. Scharth (2012). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics*, in press.

Kotz, S., T. J. Kozubowski, and K. Podgórski (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser.

Kozumi, H. and G. Kobayashi (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation 81*, 1565–1578.

Kunitomo, N. and S. Sato (2008). Separating information maximum likelihood estimation of the integrated volatility and covariance with micro-market noise. Discussion Paper CIRJE-F-581.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives 3*, 73–84.

Ledoit, O. and M. Wolf (2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management 30*(4), 110–119.

Lucas, A., B. Schwaab, and X. Zhang (2012). Measuring credit risk in a large banking system: econometric modeling and empirics. Discussion Paper.

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

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

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques And Tools*. Princeton University Press.

Meyn, S. and R. L. Tweedie (2009). *Markov Chains and Stochastic Stability* (2 ed.). Cambridge University Press.

Morf, M. and T. Kailath (1975). Square-root algorithms for least-squares estimation. *IEEE Transactions on Automatic Control 4*, 487–497.

Nakajima, J. (2012). Bayesian analysis of multivariate stochastic volatility with skew distribution. Working Paper.

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

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

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics 140*, 425–449.

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

Shephard, N. (Ed.) (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press.

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

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Ser. B 64*, 583–639.

Stuart, A. and K. Ord (1994). *Kendall's Advanced Theory of Statistics* (6 ed.), Volume 1. Arnold.

Takahashi, M., T. Watanabe, and Y. Omori (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis 53*, 2404–2426.

Takeuchi-Nogimori, A. (2012). An empirical analysis of volatility clustering and asymmetry for the common stocks using FIEGARCH and EGARCH models (in Japanese). *Journal of the Japan Statistical Society (Series-J) 42*, 1–23.

Tsionas, E. G. (2003). Bayesian quantile inference. *Journal of Statistical Computation and Simulation 73*, 659–674.

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

Vargas, G. A. (2009). Dynamic equicorrelation with non-Gaussian innovations. Working Paper.

Watanabe, T. and Y. Omori (2004). A multi-move sampler for estimating non-Gaussian time series models: Comments on Shephard & Pitt (1997). *Biometrika 91*, 246–248.

Wecker, W. E. and C. F. Ansley (1983). The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association 78*, 81–89.

Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. *Statistics & Probability Letters 54*, 437–447.

Yue, Y. R. and H. Rue (2011). Bayesian inference for additive mixed quantile regression models. *Computational Statistics & Data Analysis 55*, 84–96.

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