A Bayesian Method for Foreign Currency Portfolio Optimization of Conditional Value-at-Risk

Dongwhan Kim *

Kyu Ho Kang †

March 2017

Abstract

This paper presents a Bayesian method for implementing a conditional value-at-risk (C-VaR) portfolio optimization for foreign currency investments. Our portfolio strategy seeks to minimize the C-VaR with an expected return constraint. This consists of two stages. The first stage is to simulate the posterior predictive densities of the currency returns. To this end, we propose a new multivariate stochastic volatility (MSV) model with time-varying conditional correlations, and provide an efficient MCMC algorithm for the posterior inference. In the second stage, given the currency return density forecasts, we conduct optimal portfolio choice minimizing the C-VaR through a numerical optimization. We subsequently evaluate our portfolio strategy in terms of out-of-sample C-VaR prediction. For an empirical application, we use data on weekly returns for USD/EUR, USD/JPY, and USD/KRW. According to our out-of-sample portfolio choice experiments, the MSV with fat tail produces the most accurate C-VaR forecasts. In addition, we find that the optimal portfolio weights change over time drastically even when the transaction cost is considered. (JEL classification codes: C11, C53, G11)

Keywords: Fat tail, stochastic volatility, time-varying conditional correlation, Bayesian MCMC method

*Department of Economics, Korea University, Seoul, South Korea 136-701. E-mail: dh.kim1230@gmail.com

†*Corresponding author* Department of Economics, Korea University, Seoul, South Korea 136-701. E-mail: kyuho@korea.ac.kr

1 Introduction

Portfolio optimization is one of the most important tasks in the exchange risk management of foreign asset investment. For instance, it is well-known that international bond portfolio return tends to be much more sensitive to the currency composition rather than the maturity structure. There are three practical issues that should be considered carefully for the successful risk management of foreign currency portfolio: estimation risk, adequate risk measure, and predictive accuracy of underlying asset returns. To handle these issues, we suggest a new Bayesian method of conditional value-at-risk (C-VaR) portfolio optimization, and examine this with an application. Our method can be characterized by the corresponding three features distinguished from existing literature: Bayesian approach, the use of C-VaR, and development of a new multivariate stochastic volatility (MSV) model. Below we briefly discuss the importance of each of these features.

First, our econometric approach is Bayesian, in which the portfolio selection is done by maximizing the investor's expected utility function given the joint posterior predictive asset return density. There are two advantages from using the Bayesian approach. The first advantage is to easily reflect parameter and model uncertainty in the return prediction. The predictive distribution is obtained by integrating out the parameters and models as well as the future innovations to the returns. This can be done by the posterior predictive return density simulation. In fact, the estimation risk matters in portfolio optimization because the plug-in method borrowed in classical works involves estimation bias as Best and Grauer (1991) and Black and Litterman (1992) show. Ando (2009) empirically demonstrates that the portfolio performance can be improved by incorporating the estimation risk. The second advantage is that, using the posterior predictive return samples, we are able to numerically calculate the investor's expected utility, so that the expected utility maximization problem can be solved with respect to the portfolio weights even if the closed form solution is not feasible. Moreover, one can incorporate the prior information, which is not contained in the observed data such as macroeconomic insights and statistical intuitions, into the future return prediction.

The portfolio risk management requires an adequate measure of risk. Since 1990s, the value-at-risk (VaR) and C-VaR have been widely used as risk measures replacing the traditional variance measure. VaR is defined as the threshold point of a specific lower percentile on the return distribution, and C-VaR is the expected loss beyond VaR level in the lower tail of the distribution. These risk measures are particularly useful when investors concern on the downside risk of return on the distribution or the return distributions are often non-normal with negative kurtosis or even asymmetric. On the other hand, despite of its popularity, VaR is challenged in some aspects. Theoretically, VaR has a unfavorable property, the lack of sub-additivity (Artzner, Delbaen, Eber, and Heath (1999)). Additionally, it has the practical problem of neglecting the remaining risk beyond the threshold point. Indeed, Rockafellar and Uryasev (2002) indicate that VaR is biased for the optimum portfolio, which minimizes loss in unfavorable situations. Meanwhile, C-VaR is coherent satisfying the sub-additivity condition as Alexander and Baptista (2004) prove. Under the C-VaR, the risk of a diversified portfolio is smaller than the weighted average risk of its individual assets. Meanwhile, diversification does not necessarily lead to a reduction in VaR that is not always coherent. For this reason we use C-VaR as an alternative in this paper.[1]

The precise estimates of the portfolio C-VaR require accurate predictive return joint density forecasts because the portfolio return is a linear combination of individual asset returns. For this, we develop a new multivariate factor stochastic volatility (MSV) model of currency returns for predictive density simulation. Our MSV model is a modified version of Yu and Meyer (2006), in which the conditional correlation among the returns, as well as the volatilities, are stochastic and time-varying. Allowing for time-varying volatilities and correlations is essential for the following two reasons. The first reason is the stylized fact that a number of financial asset returns including foreign currencies and stocks have a time-varying variance-covariance structure.[2] The second reason is that unlike the VaR, the C-VaR is sub-additive and this risk measure always generates a lower risk level for a diversified portfolio than its individual assets. Because the diversification benefit depends on the correlation structure, modeling potentially time-varying conditional correlation is necessary.

Specifically, in the model the currency returns are fully determined by a linear combination of latent factors without errors. Each of the factors is assumed to follow a first-order autoregressive process with stochastic volatility. By assuming a one-to-one mapping between the observed asset returns and latent factors, we are able to identify the factor stochastic volatilities in a stable way. In addition, our estimation does not suffer from the curse of dimensionality which involves excessive computational burden

[1] Alexander and Baptista (2004), Hoogerheide and van Dijk (2010), Krokhmal, Palmquist, and Uryasev (2002), Pajor and Osiewalski (2012) compare C-VaR to VaR and introduce its methodology and application in place of VaR.

[2] For instance, see Harvey, Ruiz, and Shephard (1994), Diebold and Nerlove (1989), Schwert (1989), Kim, Shephard, and Chib (1998a), and Bollerslev (1990).

or incapability with large number of assets.

As mentioned above, we deal with the three requirements for foreign currency portfolio risk management. To the best of our knowledge, this paper is the first work that attempts C-VaR currency portfolio optimization using a MSV model in a Bayesian framework. Our Bayesian C-VaR portfolio analysis consists of two steps. The first step is to simulate the posterior predictive densities of the currency returns. To this end, we propose a new MSV model with time-varying conditional correlations, and provide an efficient MCMC algorithm for the posterior inference. In the second step, given the currency return density forecasts, we conduct the optimal portfolio choice minimizing the C-VaR through a numerical optimization. We evaluate our portfolio strategy in terms of out-of-sample C-VaR prediction.

Our work differs from existing studies in several aspects. First, there are many studies on Bayesian portfolio analysis. Since Zellner and Chetty (1965) pioneered the use of predictive distribution in portfolio selection, most studies have employed Bayesian portfolio analysis to consider the estimation risk and utilize investors' prior information into posterior updating. For example, Jorion (1986) introduces hyper-parametric prior with Bayes-Stein shrinkage method, and Black and Litterman (1992) and Pástor (2000) use a Bayesian approach with informative prior obtained from asset pricing theories. Tu and Zhou (2010) examine the prior from economic objectives under parameter uncertainty in stock portfolio choice. Although our work also relies on the Bayesian approach, our interest is to bring the recent issues such as C-VaR and MSV modeling into the Bayesian portfolio choice context.

Second, regarding the C-VaR portfolio optimization, Rockafellar and Uryasev (2000) develop the methodology and its application to portfolio selection with equity and bond. Krokhmal et al. (2002) also introduce the C-VaR as an objective or constraint in portfolio choice, and construct a strategic frontier between C-VaR and expected return. Recently, Wang, Chen, Jin, and Zhou (2010), which is one of the studies most similar to our work, analyze the C-VaR portfolio of foreign currencies using a GARCH-Copula model. Unlike our work, however, their econometric approach is not Bayesian and parameter uncertainty is not considered in the C-VaR portfolio optimization.

Finally, MSV models developed by Harvey et al. (1994) have been commonly used for multivariate asset return prediction. One technical problem with estimating MSV models is the so-called curse of dimensionality. To reduce this computational burden, Jacquier, Polson, Rossi, et al. (1995), Pitt and Shephard (1999), Chib, Nardari, and Shephard (2006), and Yu and Meyer (2006) propose a parsimonious specification of

MSV models, which is called a factor stochastic volatility model. In this model multiple asset returns are generated by the sum of the measurement errors and dynamic common latent factors. The common factors follow a vector-autoregressive process with stochastic volatility. Those studies also provide a Bayesian Markov chain Monte Carlo (MCMC) algorithm for estimation. Our MSV model is a simplified version of Yu and Meyer (2006). By assuming no measurement errors and the same number of factors as the returns, we are able to achieve a lower computational cost. In addition, we compare the Student-t errors with the normal errors since Ishihara and Omori (2016) find a strong evidence in favor of the fat-tail property of financial asset returns.

In the paper, we present a detailed process through an application to weekly USD/Euro (EUR), USD/Japanese Yen (JPY), and USD/Korean Won (KRW) data. We begin by estimating various MSV models for predictive return density simulation. Subsequently, we also conduct C-VaR portfolio optimization using the predictive densities from each of the prediction models. The models are compared statistically and economically based on the out-of-sample experiment. The statistical comparison is done by the posterior predictive likelihood which measures the predictive accuracy of return density forecasts. The economical comparison, which is our primary interest, is the mean absolute value of the C-VaR forecast errors. According to the estimation results, among the competing models the MSV with t-errors and time-varying conditional correlation produces the most accurate C-VaR forecasts indicating maximum currency-hedging benefits.

Although this work is originally motivated by the foreign currency risk management, our Bayesian portfolio analysis is generally applicable for other financial assets. There are two reasons why we concentrate on foreign currency investment. The first reason is that the variance-covariance structure of foreign currencies is strongly time-varying as reported in many studies, including Kim et al. (1998a), Kulikova and Taylor (2013), and Kastner, Frühwirth-Schnatter, and Lopes (2014). The second reason is that the return distribution of foreign currencies tends to follow a non-normal asymmetric distribution, which is suitable for C-VaR risk measure.

The remainder of the paper is organized as follows. Section 2 discusses the details of C-VaR portfolio optimization. In Sections 3 and 4, we propose our MSV model and provide a Bayesian MCMC algorithm for estimation and prediction. Section 5 illustrates the empirical application and reports the results of the out-of-sample experiments. Finally, Section 6 concludes the paper.

2 C-VaR and Portfolio Selection

Throughout this paper we consider k different foreign assets and currencies. The investment horizon is denoted by h, and the vectors of h-period-ahead foreign asset returns expressed in the domestic currency and the corresponding portfolio weights are denoted by y and x, respectively. We also consider a downside risk-averse investor whose loss function equals 1 when the realized portfolio return is less than the (1-β) level VaR, and equals 0 otherwise. β is a investor-specific credibility level. For loss function minimization, the investor optimizes the international portfolio by minimizing the (1-β) level C-VaR under the expected return constraint with respect to the portfolio weight x given the predictive distribution of returns:

$$\min_x \zeta + (1-\beta)^{-1} \int [\max\{f(x,y) - \zeta, 0\}] \times p(y)dy \tag{2.1}$$

subject to

$$\mathbb{E}[x'y] \geq \bar{\mu} + c(x), \tag{2.2}$$

$$\sum\nolimits_{i=1}^{k} x_i = 1, \text{ and } x_i \geq 0 \; for \; all \; i = 1, 2, .., k \tag{2.3}$$

where $x'y$ is the portfolio return and ζ is the (1-β) level VaR that is a function of x. $c(x)$ is the transaction cost and $\bar{\mu}$ is a minimum expected portfolio return that the investor can tolerate. $\bar{\mu}$ and β are both chosen by the investor a priori depending on his preference. $f(x,y) = -x'y$ is the loss function, and $p(y)$ is the joint probability density function of y. The equation (2.1) is the C-VaR objective, which is the sum of the VaR and the expected loss beyond the VaR. Equation (2.2) is the expected return constraint. Note that one can include the domestic currency as a risk-free asset into the portfolio by substituting the restriction $\sum_{i=1}^{k} x_i = 1$ with $\sum_{i=1}^{k} x_i \leq 1$.

Unfortunately, this optimization problem is not analytically solved in general. Instead, we rely on a simulation method. Suppose that $\{y^{(i)}\}_{i=1}^{n}$ is the samples drawn from the joint return distribution where n is a large simulation size. Given $\{y^{(i)}\}_{i=1}^{n}$, we are able to easily obtain the optimal portfolio weights x^* through Algorithm 1. The first four steps numerically calculate the C-VaR along with the VaR as a function of the portfolio weights. Then, we minimize the C-VaR with respect to the portfolio weights using a grid search.

Algorithm 1: C-VaR Portfolio Optimization

Given the h-period-ahead return samples $\{y^{(i)}\}_{i=1}^{n}$,

Step 1: Propose a portfolio vector $x = (x_1, ..., x_k)'$ satisfying the constraint in equation (2.3)

Step 2: Obtain the samples of the portfolio return,

$$\{x'y^{(i)}\}_{i=1}^{n}$$

Step 3: Calculate the expected return of portfolio,

$$\hat{\mu}(x) = n^{-1} \sum_{i=1}^{n} x'y^{(i)}$$

and check whether it satisfies the constraint in equation (2.2).

Step 4: If $\hat{\mu}(x) \geq \bar{\mu}$, then sequentially calculate the VaR and C-VaR of the portfolio. Otherwise, the proposed value x is discarded.

Step 5: Repeat steps 1 to 4 a number of times.

Step 6: Select x^* minimizing the C-VaR.

3 Multivariate Stochastic Volatility Models

3.1 Joint Return Process

This section describes our predictive models for multivariate density forecasting. Suppose that $y_t = [y_{1t}, ..., y_{kt}]'$ is a $k \times 1$ vector of foreign asset returns in terms of the domestic currency a time t. Note that the foreign asset return in the domestic currency is the sum of the asset return in the foreign currency and the currency return. The foreign asset returns are assumed to be linear to a $k \times 1$ vector of time-varying latent factors, $\mathbf{f}_t$ such that

$$\underbrace{\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{kt} \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_k \end{bmatrix}}_{\delta} + \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ \gamma_{21} & 1 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ \gamma_{k1} & \gamma_{k2} & \dots & 1 \end{bmatrix}}_{\Gamma} \underbrace{\begin{bmatrix} f_{1t} \\ f_{2t} \\ \vdots \\ f_{kt} \end{bmatrix}}_{\mathbf{f}_t}. \tag{3.1}$$

Let $\mathcal{ST}(a, b, \nu)$ denote the Student-t distribution where a is its mean, b is the scale parameter, and ν is the degrees of freedom. Then, each of the latent factors follows a first-order autoregressive process with stochastic volatility such that

$$f_{i,t}|\phi_i, f_{it-1}, V_{it} \sim \mathcal{ST}\left(\phi_i f_{i,t-1}, V_{i,t}^2, \nu\right) \tag{3.2}$$

where

$$\alpha_{i,t}|\mu_i, \varphi_i, \alpha_{i,t-1}, \sigma_i^2 \sim \mathcal{N}\left(\mu_i + \varphi_i \alpha_{i,t-1}, \sigma_i^2\right), \tag{3.3}$$
$$\text{and } V_{it} = \exp(\alpha_{i,t}/2) \text{ for } i = 1, ..., k.$$

Suppose that λ_t follows a gamma distribution, $\mathcal{G}(\nu/2, \nu/2)$. Then, the equation (3.2) can be rewritten as

$$f_{i,t}|\phi_i, f_{i,t-1}, V_{i,t}, \lambda_t \sim \mathcal{N}\left(\phi_i f_{i,t-1}, \lambda_t^{-1} V_{i,t}^2\right),$$

by the method of composition.

The lower triangular elements of the Γ matrix are to be estimated. Non-zero estimates indicate the presence of common factors among the returns. Because of the presence of common factors with time-varying volatility, the conditional correlations of the returns also change over time. As such, we note that, given the observed return y_t, the latent factors are easily obtained as

$$\mathbf{f}_t = \Gamma^{-1}(y_t - \delta) \tag{3.4}$$

because the foreign asset returns are observed without errors and the dimensions of the observed asset returns and latent factors are the same. Suppose that $\mathcal{F}_t = \{y_i\}_{i=0}^t$ denotes the information up to time t. By plugging equation (3.4) into the factor process in equation (3.2) we can obtain the joint conditional distribution of the returns as

$$y_t|\delta, \Gamma, \phi, V_t, \mathcal{F}_{t-1}, \lambda_t \sim \mathcal{N}\left(\delta - \Gamma\phi\Gamma^{-1}\delta + \Gamma\phi\Gamma^{-1}y_{t-1}, \lambda_t^{-1}\Gamma V_t V_t' \Gamma'\right) \tag{3.5}$$

where $\phi = diag(\phi_1, \phi_2, .., \phi_k)$ and $V_t = diag\left(V_{1,t}, V_{2,t}, .., V_{k,t}\right)$. For the case of $k = 3$, conditioned on the lagged returns, the stochastic volatilities, and the parameters, the conditional variance-covariance matrix of the returns at time t is

$$\lambda_t^{-1}\Gamma V_t V_t' \Gamma' = \lambda_t^{-1} \begin{bmatrix} V_{1t} & \gamma_{21}V_{1,t} & \gamma_{31}V_{1,t} \\ \gamma_{21}V_{1t} & \gamma_{21}^2 V_{1,t} + V_{2,t} & \gamma_{21}\gamma_{31}V_{1,t} + \gamma_{32}V_{2,t} \\ \gamma_{31}V_{1t} & \gamma_{21}\gamma_{31}V_{1,t} + \gamma_{32}V_{2,t} & \gamma_{31}^2 V_{1,t} + \gamma_{32}^2 V_{2,t} + V_{3,t} \end{bmatrix},$$

and the conditional correlation is given by

$$\begin{bmatrix} 1 & \frac{\gamma_{21}V_{1,t}}{\sqrt{V_{1,t}}\sqrt{\gamma_{21}^2V_{1,t}+V_{2,t}}} & \frac{\gamma_{31}V_{1,t}}{\sqrt{V_{1,t}}\sqrt{\gamma_{31}^2V_{1,t}+\gamma_{32}^2V_{2,t}+V_{3,t}}} \\ \frac{\gamma_{21}V_{1,t}}{\sqrt{V_{1,t}}\sqrt{\gamma_{21}^2V_{1,t}+V_{2,t}}} & 1 & \frac{\gamma_{21}\gamma_{31}V_{1,t}+\gamma_{32}V_{2,t}}{\sqrt{\gamma_{21}^2V_{1,t}+V_{2,t}}\sqrt{\gamma_{31}^2V_{1,t}+\gamma_{32}^2V_{2,t}+V_{3,t}}} \\ \frac{\gamma_{31}V_{1,t}}{\sqrt{V_{1,t}}\sqrt{\gamma_{31}^2V_{1,t}+\gamma_{32}^2V_{2,t}+V_{3,t}}} & \frac{\gamma_{21}\gamma_{31}V_{1,t}+\gamma_{32}V_{2,t}}{\sqrt{\gamma_{21}^2V_{1,t}+V_{2,t}}\sqrt{\gamma_{31}^2V_{1,t}+\gamma_{32}^2V_{2,t}+V_{3,t}}} & 1 \end{bmatrix}.$$

Hence, the conditional correlations along with the volatilities change over time, which generates time-varying diversification effects.

3.2 Prior

Bayesian modeling is completed by specifying a prior for the parameters. The unconditional mean of the returns, δ_i is assumed to be normally distributed. The mean is given by zero because the log of nominal exchange rates tend to follow a random-walk process without drift.

$$\delta_i \sim \mathcal{N}(b_{0,\delta}, B_{0,\delta}) \equiv \mathcal{N}(0,1),\ i = 1,2,..,k. \tag{3.6}$$

A priori we assume the presence of k common factors among the returns. The prior variance of γ_{ij}'s is set to be large, so that our prior belief is not too strong.

$$\gamma_{ij} \sim \mathcal{N}(b_{ij,0,\gamma}, B_{ij,0,\gamma}) \equiv \mathcal{N}(0,1),\ i,j = 2,...,k,\ i > j. \tag{3.7}$$

Next, the autoregressive coefficient ϕ_i is constrained to lie on the interval (-1, 1), and this is assumed to follow a beta distribution such that

$$\frac{\phi_i + 1}{2} \sim beta(a_{0,\phi}, b_{0,\phi}) \equiv beta(5,5),\ i = 1,2,..,k. \tag{3.8}$$

The financial asset return persistence is typically small, so our prior mean of ϕ_i's is zero. Meanwhile, the exchange rate volatilities are known to be persistent. For the normal error models, the prior mean of φ_i is set to be close to one.

$$\varphi_i \sim \mathcal{N}(b_{0,\varphi}, B_{0,\varphi}) \equiv \mathcal{N}(0.9,1),\ i = 1,2,..,k. \tag{3.9}$$

Meanwhile, we specify a stronger prior for the fat-tailed models such that

$$\varphi_i \sim \mathcal{N}(b_{0,\varphi}, B_{0,\varphi}) \equiv \mathcal{N}(0.9,0.01),\ i = 1,2,..,k. \tag{3.10}$$

Otherwise, the stochastic volatilities are not well-identified because of the fat-tail property. The intercept term μ_i is identified by the log unconditional variance of the returns given the other parameters. Considering the scale of the returns, our prior mean of μ_i's is -0.5.

$$\mu_i \sim \mathcal{N}(b_{0,\mu}, B_{0,\mu}) \equiv \mathcal{N}(-0.5, 1), \ i = 1, 2, .., k. \tag{3.11}$$

The conditional variance of the log conditional factor variance $\alpha_{it} (= \log V_{it}^2)$ follows an inverse-gamma distribution,

$$\sigma_i^2 \sim \mathcal{IG}(a_{0,\sigma}, b_{0,\sigma}) \equiv \mathcal{IG}(2, 0.1), \ i = 1, 2, .., k. \tag{3.12}$$

Finally, the degree of freedom of the assets following Student-t distribution, λ_t follows a gamma distribution,

$$\lambda_t \sim \mathcal{G}(\nu/2, \nu/2) \equiv \mathcal{G}(5, 5), \ t = 1, 2, .., T. \tag{3.13}$$

The degree of freedom is fixed at $\nu = 10$.

3.3 Candidate Prediction Models

Suppose that the MSV model explained in the previous section is denoted by SVCt. In order to account for the model uncertainty, we consider another four model specifications as special cases of the SVCt model. The following describes each of the competing models:

- FV: First-order vector-autoregressive model with constant variance-covariance
- SVn: MSV model with normal errors and zero conditional correlations
- SVCn: MSV model with normal errors and time-varying conditional correlations
- SVt: MSV model with Student-t errors and zero conditional correlations
- SVCt: MSV model with Student-t errors and time-varying conditional correlations

Note that the γ_{ij}'s are constrained to be zero in the SVn and SVt models. In this paper we deal with the model uncertainty by choosing the best model in term of the out-of-sample C-VaR prediction performance. Although our models are flexible in fitting financial asset returns, they have several deficiencies. For example, the MSV model with non-zero constant correlations is not included as this is not a restricted model of the SVCt model. Moreover, they do not take account of a jump or drastic regime shifts in the return process, either.

4 Posterior Sampling

In this section we illustrate the MCMC sampling procedure of the SVCt model that is the most general specification among the candidate models. The set of model parameters is denoted by

$$\theta = \{\gamma,\ \delta,\ \mu,\ \varphi,\ \phi,\ \sigma^2,\ \Lambda\}$$

where

$$\begin{aligned} \gamma &= \{\gamma_{ij} | i, j = 1, 2, 3, .., k., i > j\}, \\ \mu &= \{\mu_i\}_{i=1}^{k},\ \varphi = \{\varphi_i\}_{i=1}^{k},\ \sigma^2 = \{\sigma_i^2\}_{i=1}^{k}, \Lambda = \{\lambda_t\}_{t=1}^{T}. \end{aligned}$$

The time series of the factors and the log stochastic volatilities are denoted by $\mathbf{F} = \{\mathbf{f}_t\}_{t=1}^{T}$ and $\mathbf{A} = \{\{\alpha_{i,t}\}_{i=1}^{k}\}_{t=1}^{T}$, respectively.

This section presents our posterior simulation of the parameters and the time series of the stochastic volatilities, $(\theta, \mathbf{A})$. Given the observations $\mathbf{Y} = \{y_t\}_{t=1}^{T}$, these are simulated from their joint posterior distribution

$$\theta, \mathbf{A} | \mathbf{Y}.$$

Its density is given by

$$\pi(\theta, \mathbf{A}|\mathbf{Y}) \propto p(\mathbf{Y}|\theta, \mathbf{A}) \times p(\mathbf{A}|\theta) \times \pi(\theta),$$

which is proportional to the product of the conditional return density, the conditional volatility density, and prior density of the parameters.

From equation (3.5) the conditional densities of the returns, $p(\mathbf{Y}|\theta, \mathbf{A})$ are obtained as

$$\begin{aligned} & p(\mathbf{Y}|\theta, \mathbf{A}) \\ &= \prod_{t=1}^{T} \mathcal{N}\left(y_t | \delta - \Gamma\phi\Gamma^{-1}\delta + \Gamma\phi\Gamma^{-1}y_{t-1}, \lambda_t^{-1}\Gamma V_t V_t' \Gamma'\right) \end{aligned} \tag{4.1}$$

where the initial returns y_0 are assumed to be observed from the data. The conditional density of $\mathbf{A}$, $f(\mathbf{A}|\theta)$ is obtained from equation (3.3) as follows:

$$f(\mathbf{A}|\theta) = \prod_{i=1}^{k} \left[\prod_{t=1}^{T} \mathcal{N}\left(\alpha_{i,t} | \mu_i + \varphi_i \alpha_{i,t-1}, \sigma_i^2\right)\right].$$

Finally, the prior density of the parameters, $\pi(\theta)$, is simply the multiplication of the densities of each parameter in equations (3.6) to (3.13), because all parameters are assumed to be independent a priori. It is important to note that once the posterior draws for $(\theta, \mathbf{A})$ are obtained from the posterior simulation, the posterior draws for the factors and stochastic volatilities are immediately retained as

$$\mathbf{F} = \{\Gamma^{-1}(y_t - \delta)\}_{t=1}^{T} \text{ and}$$
$$\mathbf{V} = \{\exp(\alpha_{1t}/2), \exp(\alpha_{2t}/2), .., \exp(\alpha_{kt}/2)\}_{t=1}^{T},$$

respectively. Particularly, it is a great advantage that simulation stage with a high computational cost such as forward and backward recursions is not necessary for factor sampling. This enables us to identify the stochastic volatilities in a more stable way.

We simulate $(\theta, \mathbf{A})$ in multiple blocks as direct simulation from the joint posterior distribution is not feasible. Particularly, the stochastic volatility process is sampled based on the method of Kim, Shephard, and Chib (1998b). The key idea of their approach is to approximate the log squared normal errors by a mixture of normal distributions. The mixture is governed by state variables. Therefore, in each MCMC cycle, the time series of stochastic volatilities and state variables are simulated sequentially. Suppose that the time series of the state variables are denoted by

$$\mathbf{S} = \{\{s_{i,t}\}_{t=1}^{T}\}_{i=1}^{k}.$$

Our MCMC algorithm can be summarized as follows.

Algorithm 2: Posterior MCMC simulation

Step 0: Initialize θ and $\mathbf{S}$.

Step 1: Sample $\mathbf{A}|\theta, \mathbf{S}, \mathbf{Y}$.

Step 2: Sample $\mu, \varphi, \sigma^2|\mathbf{A}, \mathbf{Y}$.

Step 3: Sample $\mathbf{S}|\theta, \mathbf{Y}$.

Step 4: Sample $\gamma|\delta, \phi, \Lambda, \mathbf{A}, \mathbf{Y}$.

Step 5: Sample $\delta, \phi|\gamma, \Lambda, \mathbf{A}, \mathbf{Y}$.

Step 6: Sample $\Lambda|\delta, \gamma, \phi, \mathbf{A}, \mathbf{Y}$

Step 7: Sample $y_{T+1}|\theta, \mathbf{A}, \mathbf{Y}$.

We begin by sampling the stochastic volatilities given the data and initialized parameters and state variables. Subsequently, the parameters in the stochastic volatility process are simulated. Given the parameters and data, the state variables are sampled. Next, the parameters in the factor loadings are simulated. After simulating the unconditional mean of the returns, persistence coefficients, and the time series of the scale parameters, the posterior predictive distribution of the returns are sampled. In the following, we explain the details of each MCMC stage.

4.1 Sampling $\mathbf{A}|\theta, \mathbf{S}, \mathbf{Y}$

To sample the stochastic volatility processes given the parameters and data, we obtain the time series of the factors

$$\mathbf{F}= \{\mathbf{f_t}\}_{t=1}^{T}= \{\Gamma^{-1}(y_t - \delta)\}_{t=1}^{T}, \tag{4.2}$$

and transform the factor process as follows. For $i = 1, 2, .., k$ and $t = 1, 2, .., T$,

$$\tilde{f}_{i,t} = f_{i,t} - \phi_i f_{i,t-1} = \exp(\alpha_{i,t}/2)\varepsilon_{i,t} \tag{4.3}$$

It follows that the log of squares is expressed as a sum of the log squared volatility and log squared factor shocks,

$$f^*_{i,t} = \alpha_{i,t} + \varepsilon^*_{i,t}$$

with $f^*_{i,t} = \log(\tilde{f}^2_{i,t})$ and $\varepsilon^*_{i,t} = \log(\varepsilon^2_{i,t})$. According to the work of Kim et al. (1998b), the distribution of $\varepsilon^*_{i,t}$ can be approximated by a mixture of seven normal distributions.

q	$\Pr(s_{i,t} = q)$	$m_{s_{i,t}}$	$R_{s_{i,t}}$
1	0.00730	−10.12999	5.79596
2	0.10556	−3.97281	2.61369
3	0.00002	−8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	−1.08819	1.26261

As a result, the model can be expressed in a state-space representation,

$$f^*_{i,t}|\alpha_{i,t}, \theta, s_{i,t} \sim \mathcal{N}(\alpha_{i,t} + m_{s_{i,t}}, R_{s_{i,t}}),$$
$$\alpha_{i,t}|\theta, \alpha_{i,t-1} \sim \mathcal{N}(\mu_i + \varphi_i\alpha_{i,t-1}, \sigma^2_i)$$

where the initial $\alpha_{i,0}$ follows the unconditional distribution of $\alpha_{i,t}$,

$$\alpha_{i,0} \sim \mathcal{N}\left(\frac{\mu_i}{1-\varphi_i}, \frac{\sigma_i^2}{1-\varphi_i^2}\right).$$

Stochastic volatility sampling consists of two steps. The first step is to run the Kalman filter and obtain the filtered distribution

$$\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t$$

for $i = 1, 2, .., k$. As $\alpha_{i,t}$'s conditioned on $(\theta, \mathcal{F}_t)$ are normally distributed, the objective of this step is to calculate the conditional mean and variance at each time t

$$\begin{aligned}
&\mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t) = \mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) \\
&+ \left[Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1})/Var(f_{i,t}^*|\theta, \mathbf{S}, \mathcal{F}_{t-1})\right] \times (f_{i,t}^* - \mathbb{E}(f_{i,t}^*|\theta, \mathbf{S}, \mathcal{F}_{t-1})), \\
&Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t) = Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) \\
&- \left[Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1})/Var(f_{i,t}^*|\theta, \mathbf{S}, \mathcal{F}_{t-1})\right] \times Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1})
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}(\alpha_{i,0}|\theta, \mathbf{S}, \mathcal{F}_0) &= \frac{\mu_i}{1-\varphi_i},\ Var(\alpha_{i,0}|\theta, \mathbf{S}, \mathcal{F}_0) = \frac{\sigma_i^2}{1-\varphi_i^2}, \\
\mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) &= \mu_i + \varphi_i \mathbb{E}(\alpha_{i,t-1}|\theta, \mathbf{S}, \mathcal{F}_{t-1}), \\
Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) &= \varphi_i^2 Var(\alpha_{i,t-1}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) + \sigma_i^2, \\
\mathbb{E}(f_{i,t}^*|\theta, \mathbf{S}, \mathcal{F}_{t-1}) &= \mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) + m_{s_{i,t}}, \\
\text{and } Var(f_{i,t}^*|\theta, \mathbf{S}, \mathcal{F}_{t-1}) &= Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t-1}) + R_{s_{i,t}}.
\end{aligned}$$

The second step is the backward recursion. At time T, $\alpha_{i,T}$ is sampled from its filtered distribution,

$$\alpha_{i,T}|\theta, \mathbf{S}, \mathbf{Y} \equiv \alpha_{i,T}|\theta, \mathbf{S}, \mathcal{F}_T \sim \mathcal{N}(\mathbb{E}(\alpha_{i,T}|\theta, \mathbf{S}, \mathbf{Y}),\ Var(\alpha_{i,T}|\theta, \mathbf{S}, \mathbf{Y}))$$

as $\mathcal{F}_T = \mathbf{Y}$ by definition. Given $\alpha_{i,t+1}$, $\alpha_{i,t}$ $(t = T-1, T-2, .., 1)$ is sampled from its conditional distribution,

$$\begin{aligned}
\alpha_{i,t}|\theta, \mathbf{S}, \mathbf{Y} &\equiv \alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_{t+1} \equiv \alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t, \alpha_{i,t+1} \\
&\sim \mathcal{N}\left(\mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t, \alpha_{i,t+1}),\ Var(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t, \alpha_{i,t+1})\right)
\end{aligned}$$

The conditional mean and variance are given by

$\mathbb{E}(\alpha_{i,t}|\theta, \mathbf{S}, \mathcal{F}_t, \alpha_{i,t+1})$

$$= \mathbb{E}(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t) + [Var(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t) \times \varphi / Var(\alpha_{i,t+1}|\theta,\mathbf{S},\mathcal{F}_t)] \times (\alpha_{i,t+1} - \mathbb{E}(\alpha_{i,t+1}|\theta,\mathbf{S},\mathcal{F}_t)),$$

$$Var(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t,\alpha_{i,t+1})$$

$$= Var(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t) - [Var(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t)/Var(\alpha_{i,t+1}|\theta,\mathbf{S},\mathcal{F}_t)] \times Var(\alpha_{i,t}|\theta,\mathbf{S},\mathcal{F}_t),$$

respectively. The samples drawn from this backward recursion are taken as posterior draws for $\{\alpha_{i,t}\}_{t=1}^T$.

The Kalman filtering and backward recursion are repeated for each $i = 1, 2, .., k$, which completes the joint sampling of $\mathbf{A} = \{\alpha_{1,t}, \alpha_{2,t}, .., \alpha_{k,t}\}_{t=1}^T$ given $(\theta, \mathbf{S}, \mathbf{Y})$.

4.2 Sampling $\mu, \varphi, \sigma^2|\mathbf{A}, \mathbf{Y}$

Given $(\mathbf{A}, \mathbf{Y})$, the full conditional distributions of (μ, φ) and σ^2 are tractable because the stochastic process for $\alpha_{i,t}$ in equation (3.3) is a standard linear regression and their priors are conjugate. Suppose that $\beta_0 = (\ b'_{0,\mu}\quad b'_{0,\varphi}\ )'$ is the prior mean of (μ_i, φ_i) and

$$B_0 = \begin{pmatrix} B_{0,\mu} & 0 \\ 0 & B_{0,\varphi} \end{pmatrix}$$

is the prior variance-covariance. Then, (μ_i, φ_i) is first sampled from its full conditional distribution,

$$\begin{pmatrix} \mu_i \\ \varphi_i \end{pmatrix} |\mathbf{A},\sigma_i^2 \sim \mathcal{N}(B_1^{-1} A_1,\ B_1^{-1}),$$

where

$$\begin{aligned} B_1 &= B_0^{-1} + \sum_{t=2}^{T} (1\ \alpha_{i,t-1})' \times (1\ \alpha_{i,t-1})/\sigma_i^2, \\ A_1 &= B_0^{-1}\beta_0 + \sum_{t=2}^{T} (1\ \alpha_{i,t-1})' \times \alpha_{i,t}/\sigma_i^2. \end{aligned}$$

Now, given (μ_i, φ_i) and $\mathbf{A}$, σ_i^2 is drawn from the inverse gamma distribution,

$$\sigma_i^2|\alpha_{i,t}, \alpha_{i,t-1} \sim \mathcal{IG}\left(\frac{a_{0,\sigma} + T}{2}, \frac{b_{0,\sigma} + \sum_{t=1}^{T}(\alpha_{i,t} - \mu_i - \varphi_i\alpha_{i,t-1})^2}{2}\right).$$

Sampling $(\mu, \varphi, \sigma^2)\,|\mathbf{A}, \mathbf{Y}$ is completed by repeating $(\mu_i, \varphi_i, \sigma_i^2)$ sampling for $i = 1, 2, .., k$.

4.3 Sampling $\mathbf{S}|\theta, \mathbf{A}, \mathbf{Y}$

Next, the state variables that govern the distribution of the log squared errors over time are sampled. Given $(\theta, \mathbf{Y})$, $\mathbf{F}$ is obtained as in equation (4.2), and

$$\{\{f_{i,t}^* = \log((f_{i,t} - \phi_i f_{i,t-1})^2)\}_{t=1}^T\}_{i=1}^k$$

is constructed. Then, the full conditional mass of $s_{i,t} = q$ is proportional to the product of the conditional density of $f^*_{i,t}$ and the prior mass of $s_{i,t} = q$.

$$\begin{aligned}\mathsf{Pr}(s_{i,t} = q|\theta, \mathbf{A}, \mathbf{Y}) &= \mathsf{Pr}(s_{i,t} = q|f^*_{i,t}, \alpha_{i,t}) \\ &\propto p(f^*_{i,t}|\alpha_{i,t}, s_t = q) \times \mathsf{Pr}(s_{i,t} = q) \\ &= \mathcal{N}\left(f^*_{i,t}|\alpha_{i,t} + m_q,\ R_q\right) \times \mathsf{Pr}(s_{i,t} = q),\ q = 1, 2, .., 7\end{aligned} \tag{4.4}$$

Therefore, $s_{i,t} = q$ is drawn with the probability

$$\mathsf{Pr}(s_{i,t} = q|\theta, \mathbf{A}, \mathbf{Y}) = \frac{\mathcal{N}\left(f^*_{i,t}|\alpha_{i,t} + m_q,\ R_q\right) \times \mathsf{Pr}(s_{i,t} = q)}{\sum_{j=1}^{7} \mathcal{N}\left(f^*_{i,t}|\alpha_{i,t} + m_j,\ R_j\right) \times \mathsf{Pr}(s_{i,t} = j)}$$

independently of the other state variables. The simulation of $\mathbf{S}$ is done by sampling $s_{i,t}$ for all $i = 1, 2, .., k$ and $t = 1, 2, .., T$.

4.4 Sampling $\gamma|\delta, \phi, \Lambda, \mathbf{A}, \mathbf{Y}$

The full conditional density of γ is given by

$$\begin{aligned}\pi(\gamma|\delta, \phi, \Lambda, \mathbf{A}, \mathbf{Y}) &\propto p(\mathbf{Y}|\theta, \mathbf{A}) \times \pi(\gamma) \\ &= p(\mathbf{Y}|\theta, \mathbf{A}) \times \left[\prod_{i>j} \mathcal{N}(\gamma_{ij}|b_{ij,0,\gamma}, B_{ij,0,\gamma})\right]\end{aligned} \tag{4.5}$$

The complete likelihood $p(\mathbf{Y}|\theta, \mathbf{A})$ and the prior density of $\pi(\gamma)$ are given in equations (4.1) and (3.7), respectively. Because the full conditional distribution is not tractable, we rely on the random-walk Metropolis-Hastings (RW-MH) algorithm. The efficiency of the RW-MH method depends on the variance-covariance of the proposal distribution, $\mathbf{R}_{RW}$, which should be chosen carefully. At the gth MCMC cycle, the inverse of negative hessian computed at the (g-1)th MCMC draw $\gamma^{(g-1)}$,

$$\mathbf{R}_{RW} = \left[-\frac{\partial\partial\left(\log p(\mathbf{Y}|\theta^{(g-1)}, \mathbf{A}) + \log \pi(\gamma^{(g-1)})\right)}{\partial\gamma^{(g-1)}\partial\gamma^{(g-1)\prime}}\right]^{-1}$$

is used as $\mathbf{R}_{RW}$, and a candidate γ^* is drawn from the normal distribution,

$$\gamma^* \sim \mathcal{N}(\gamma^{(g-1)}, \mathbf{R}_{RW}).$$

Then, the candidate is accepted and retained as the gth MCMC draw $\gamma^{(g)}$ with the usual MH probability

$$\min\left\{\frac{p(\mathbf{Y}|\gamma^*, \delta^{(g-1)}, \phi^{(g-1)}, \mathbf{A}) \times \pi(\gamma^*)}{p(\mathbf{Y}|\gamma^{(g-1)}, \delta^{(g-1)}, \phi^{(g-1)}, \mathbf{A}) \times \pi(\gamma^{(g-1)})}, 1\right\}.$$

If it is rejected, $\gamma^{(g-1)}$ is saved, instead.

4.5 Sampling $\delta, \phi | \Gamma, \Lambda, \mathbf{A}, \mathbf{Y}$

As shown in equation (3.5), the conditional expectation of the returns is nonlinear to (δ, ϕ), and their full conditional distribution is not feasible. Like γ, (δ, ϕ) are sampled through the RW-MH method. Moreover, the target density of this block in each MCMC iteration is given by

$$\begin{aligned} p(\delta, \phi | \Gamma, \Lambda, \mathbf{A}, \mathbf{Y}) &\propto p(\mathbf{Y}|\theta, \mathbf{A}) \times \pi(\delta) \times \pi(\phi) \\ &\propto p(\mathbf{Y}|\theta, \mathbf{A}) \times \left[\prod_{i=1}^{k} \mathcal{N}(\delta_i | b_{0,\delta}, B_{0,\delta})\right] \times \left[\prod_{i=1}^{k} beta(\frac{\phi_i + 1}{2} | a_{0,\phi}, b_{0,\phi})\right] \end{aligned}$$

where $\pi(\delta)$ and $\pi(\phi)$ are the prior densities of δ and ϕ, respectively.

4.6 Sampling $\Lambda | \delta, \phi, \Gamma, \mathbf{A}, \mathbf{Y}$

Given $(\theta, \mathbf{Y})$, $\Lambda = \{\lambda_t\}_{t=1}^{T}$ is sampled via a single move. The full conditional distribution of λ_t is tractable as y_t is normally distributed given θ and its gamma prior is conjugate. The full conditional distribution is obtained as a gamma distribution such as

$$\begin{aligned} p(\lambda_t | \delta, \phi, \Gamma, \mathbf{A}, \mathbf{Y}) &\propto p(\mathbf{Y}|\theta, \mathbf{A}) \times \pi(\lambda_t) \\ &\propto p(y_t|\theta, \mathbf{A}) \times \mathcal{G}(\lambda_t | \nu/2, \nu/2) \\ &\propto \mathcal{G}\left(\lambda_t | \frac{\nu + 1}{2}, \frac{\nu + \tilde{y}_t' \Sigma_t^{-1} \tilde{y}_t}{2}\right) \end{aligned}$$

where $\tilde{y}_t = y_t - \delta + \Gamma\phi\Gamma^{-1}\delta - \Gamma\phi\Gamma^{-1}y_{t-1}$ and $\Sigma_t = \Gamma \mathbf{V}_t \mathbf{V}_t' \Gamma'$.

4.7 Sampling the Posterior Predictive Distribution

Once the posterior draws for $(\theta, \mathbf{A})$ are obtained from the Steps 1 to 5 in each MCMC cycle, the h-period-ahead posterior predictive draws of the returns can be simulated by the following algorithm.

Algorithm 3: Posterior predictive distribution simulation

For $j = 1, 2, .., h$,

Step 1: Sample $\alpha_{i,T+j} | \theta, \mathbf{A} \sim N(\mu_i + \varphi_i \alpha_{i,T+j-1}, \sigma_i^2)$ for $i = 1, 2, .., k$

Step 2: Sample $y_{T+j}|\delta, \phi, \Gamma, V_{T+j}, \mathbf{Y}$ from

$$\mathcal{ST}\left(\delta - \Gamma\phi\Gamma^{-1}\delta + \Gamma\phi\Gamma^{-1}y_{T+j-1}, \Gamma V_{T+j}V'_{T+j}\Gamma', \nu\right)$$

where $V_{T+j} = diag\left(\exp(\alpha_{1,T+j}/2), \exp(\alpha_{2,T+j}/2), .., \exp(\alpha_{k,T+j}/2)\right)$

Step 3: Retain y_{T+j} as a h-period-ahead posterior predictive draw

4.8 Predictive Density Accuracy Measure

To minimize the estimation risk in portfolio selection, the joint predictive distribution of the asset returns should be accurate. Otherwise, the decision-making based on a poor return prediction may result in a much heavier loss than expected especially when the risk is underestimated or the expected return is overestimated. Thus, the predictive accuracy is a prerequisite of a successful risk management. We evaluate the out-of-sample predictive density accuracy of the prediction models. As suggested by Eklund and Karlsson (2007), the density prediction performance is measured by the posterior predictive likelihood(PPL). This is the product of the posterior predictive densities over the out-of-sample periods. The posterior predictive density(PPD) is the conditional density of the realized one-period-ahead returns y^*_{T+1} conditioned on the observation. The PPD, $p(y^*_{T+1}|\mathbf{Y})$ is computed by integrating out the conditional density $p(y^*_{T+1}|\theta, \mathbf{A}, \mathbf{Y})$ over the parameters and stochastic volatilities. That is,

$$p(y^*_{T+1}|\mathbf{Y}) = \int p(y^*_{T+1}|\theta, \mathbf{A}, \mathbf{Y}) \times \pi(\theta, \mathbf{A}|\mathbf{Y})d(\theta, \mathbf{A}). \tag{4.6}$$

However, the integration cannot be done analytically, and we rely on the numerical approximation as

$$p(y^*_{T+1}|\mathbf{Y}) \fallingdotseq \frac{1}{n_1}\sum_{g=1}^{n_1} p(y^*_{T+1}|\theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y}) \tag{4.7}$$

where $(\theta^{(g)}, \mathbf{A}^{(g)})$ are the posterior draws. Further, $p(y^*_{T+1}|\theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y})$ is not feasible analytically, either. For this reason, it is also computed by numerically integrating out the one-period-ahead stochastic volatilities from the conditional density $p(y^*_{T+1}|V_{T+1}, \theta, \mathbf{A}, \mathbf{Y})$ as follows:

$$\begin{aligned} p(y^*_{T+1}|\theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y}) &= \int p(y^*_{T+1}|V_{T+1}, \theta, \mathbf{A}, \mathbf{Y}) \times p(V_{T+1}|\theta, \mathbf{A}, \mathbf{Y})dV_{T+1} \\ &\fallingdotseq \frac{1}{1,000}\sum_{j=1}^{1,000} p(y^*_{T+1}|V^{(j)}_{T+1}, \theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y}) \end{aligned} \tag{4.8}$$

where

$$\alpha_{i,T+1}^{(j)}|\theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y} \sim \mathcal{N}(\mu^{(g)} + \varphi^{(g)}\alpha_{i,T}^{(g)},\ \sigma_i^{2(g)}),\ i = 1, 2, .., k,$$

$$\begin{aligned} V_{T+1}^{(j)} &= diag(\exp(\alpha_{1,T+1}^{(j)}/2), \exp(\alpha_{2,T+1}^{(j)}/2), .., \exp(\alpha_{k,T+1}^{(j)}/2)), \\ \lambda_{T+1}^{(g)} &\sim \mathcal{G}(5, 5), \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} & p(y_{T+1}^{*}|V_{T+1}^{(j)}, \theta^{(g)}, \mathbf{A}^{(g)}, \mathbf{Y}) \\ = \ & \mathcal{N}(y_{T+1}^{*}|\delta^{(g)} - \Gamma^{(g)}\phi^{(g)}\Gamma^{(g)-1}\delta^{(g)} + \Gamma^{(g)}\phi^{(g)}\Gamma^{(g)-1}y_T,\ \lambda_{T+1}^{(g)-1}\Gamma^{(g)}V_{T+1}^{(j)}V_{T+1}^{(j)\prime}\Gamma^{(g)\prime}). \end{aligned}$$

Note that the PPL is model-dependent, and we denote the PPL of the model $\mathcal{M}$ by PPL($\mathcal{M}$), which is computed as

$$\text{PPL}(\mathcal{M}) = \prod_{h=1}^{H} p(y_{T+h}^{*}|\mathcal{F}_{T+h-1}, \mathcal{M})$$

where the out-of-sample size is H. Using the log PPL we evaluate the relative out-of-sample density prediction performance among the models.

The PPL is one of the standard Bayesian model choice criteria. Nevertheless, using the PPL is not desirable for the model choice in this work. Our primary model selection criterion is the C-VaR forecasting performance. The aim of this work is to precisely quantify and minimize the extreme loss or downside risk whereas the PPL measures the predictive accuracy over the entire support of the return distribution.

5 Application

5.1 Data

The underlying assets considered in our application are the foreign government bonds whose the maturity is one week. The one-week bond yield is almost zero, and its variation is much smaller than that of the exchange rate. For this reason, we use currency return data only in this application. Of course, in case of different underling assets such as stocks, long-term government bonds, and the corporate bonds, the underlying asset return in terms of the foreign currency should be added to the currency return.

The data used in this paper are three foreign exchange rates: USD per EUR, USD per JPY, and USD per KRW. Our choice of these foreign currencies is based on the fact that each of them has its own unique characteristics as an investment asset. EUR is the second most traded currency and JPY is regarded as a safety asset during financial crises. KRW is a representative currency of the developing countries and it tends to depreciate sensitively to regional shocks as well as changes in global economic factors. It should be noted that, although we consider only three foreign currencies in this application, our methodology is generic to the number and components of the currencies and underlying financial assets.

Our exchange rate data is weekly, ranging from the first week of 1999 to the 15th week of 2016, as plotted in Figure 1. The total observations are 900 weekly currency returns for each currency, and these data can be obtained from the FRB of St. Louis FRED Economic Data. For the in-sample analysis, we use the entire 900 return observations in order to sample the model parameters, stochastic volatilities, and conditional correlation. For the out-of-sample portfolio choice experiment, we simulate the one-week-ahead predictive distributions of the currency returns for the past 150 weeks. The investment horizon is set to be one week, and this can be easily generalized to a longer horizon. The rolling window size is 750 weeks. The first out-of-sample forecast is the 22nd week of 2013 and the last one is the 15th week of 2016.

Given the joint return density forecasts, the C-VaR portfolio optimization is conducted. The credibility levels considered in this paper are 50%, 75%, and 90%. Although the higher levels such as 95% and 99% are more desirable for extreme risk management, they are excluded because the out-of-sample size seems to be too small for the C-VaR forecasts to be compared among the alternative prediction models.

5.2 Estimation Results

Our MCMC size is 10,000 and the first 5,000 draws are discarded to ensure the convergence of the Markov chain. Figure 2 displays the prior and posterior distributions of each parameter in the SVCt model, which is the most general specification among the competing models. As shown in the figure, the posterior densities are much sharper than the corresponding prior densities. This means that the data is informative and the posterior densities are determined by the information in the data rather than the prior information.

Table 1 reports the estimation results for the parameters in the four MSV models. In this table, δ_i;s are estimated to be small for all assets and models as its 90 % credibility interval includes 0, except for the KRW. ϕ_i's, the autoregressive coefficients of the factors are also estimated to be close to zero, which means a very strong mean-revering property of the currency returns, which is well-known. Meanwhile, the estimates of the autoregressive coefficients of the log stochastic volatilities are highly positive, and the volatilities are found to be persistent except for the second factor. In addition, all γ's in the SVCn and SVCt models are very precisely estimated to be non-zero, which indicates the presence of co-movement among the currencies.

On the other hand, we check the convergence of the Markov chain and the efficiency of the sampling scheme in terms of the serial correlation of the MCMC draws following Chib and Ramamurthy (2010) and Chib and Kang (2013). Figure 3 displays the results for the autocorrelation functions for the SVCt model. As seen from this figure, the autocorrelations decay to zero fast indicating high efficiency of the sampling algorithm. Of course, the parameters sampled from the MH method reveal the higher persistence than those sampled by the Gibbs-sampler. Formally, the inefficiency for each parameter sampling is measured by the inefficiency factor, which is computed as

$$1 + 2\sum_{l=1}^{200} \hat{\rho}(l)$$

where $\hat{\rho}(l)$ is the estimate of the lth order autocorrelation. By definition, a small inefficiency factor means a well-mixing sampler. The results for the inefficiency factors of the SVCt model can be found in the last column of Table 1.

Figure 4 shows the time series of the estimated stochastic volatilities and conditional correlations over the entire sample period. As expected, all currency returns reveal a strong evidence for time-varying volatilities. Further, the conditional correlations as well as the volatilities dramatically change over time between 0.1 and 0.8. Overall, the correlation between EUR and JPY is higher than the other correlations. The drastic changes in the volatilities and correlations imply that the diversification effect is not fixed and the optimal portfolio shares must be dynamic, not static.

5.3 Economic Evaluation: Out-of-Sample Portfolio Performance

We follow the approach of Gerlach and Chen (2016) and evaluate our method in terms of risk management. The economic evaluation measure used in this paper is the mean absolute error(MAE) of the one-week-ahead C-VaR forecasts, which is computed as

$$\left(\sum_{t\in OSP} I(y_t^* < \widehat{\mathrm{VaR}}_t)\right)^{-1} \times \sum_{t\in OSP}\left[|y_t^* - \widehat{\text{C-VaR}}_t| \times I(y_t^* < \widehat{\mathrm{VaR}}_t)\right]$$

where OSP is the set of the out-of-sample periods, $I(\cdot)$ is an indicator function, y_t^* is the realized portfolio return, and $\widehat{\mathrm{VaR}}_t$ and $\widehat{\text{C-VaR}}_t$ are one-week-ahead VaR and C-VaR forecasts at time t, respectively. This MAE of the C-VaR forecasts is a proxy of the average loss caused by the C-VaR prediction error during the out-of-sample periods. As the MAE of the C-VaR forecasts is larger, the currency-hedging benefits compared to the costs of implementing the hedges decrease.

Optimal Portfolio Weights Before we compare the portfolio performance of each of the models, we discuss the results for the optimal portfolio weights obtained from the C-VaR minimization. Figure 5 and 6 show the time series of the portfolio weights when the transaction cost is zero and 0.1%, respectively. The transaction cost is assumed to occur in both buying and selling assets. For instance, suppose that we sell 0.2 shares of EUR and buy same amount of JPY. Then, the cost is $0.2\times0.1\%+0.2\times0.1\%=0.04\%$, and the expected portfolio return decreases by this amount. Those portfolio weights minimize the 10% C-VaR under the constraint that the expected portfolio return is greater than -0.2%.

There are two interesting findings resulting from these figures. First, not surprisingly, the FV model yields almost constant weights over time as the variance-covariance matrix is constrained to be fixed during the rolling windows. However, the optimal weights based on the stochastic volatility models dramatically change over time. Until the early 2015, the share of the JPY was smaller than the others because the JPY was more volatile than the other currencies as shown in Figure 7. This figure plots the predictive volatilities and correlations of the currency returns over the out-of-sample periods. Since EUR and KRW became more volatile than the JPY because of the prolonged EU economy downturn and the announcement of the U.S. Fed tapering policy, however, the JPY has been more attractive as a safe asset for risk management.

Second, the changes in the portfolio weights when the transaction cost is 0.1% are much smoother than those when the cost is zero. For instance, the weight on JPY in March 2015 changed from 0.42 to 1.0 based on the SVCt density forecasts when the transaction cost is zero. Meanwhile, when the transaction cost is 0.1%, it changed to 0.7, not 1.0. This is because the responses of the portfolio weights on the changes in the predictive joint distribution of the currency returns are less sensitive as the transaction cost increases.

C-VaR Forecasts We now evaluate the out-of-sample foreign currency portfolio performance based on the accuracy of the C-VaR forecasting, reported in Table 2. We concentrate on the benchmark case, in which the credibility level is 90% and the transaction cost is 0.1% because our focus is on tail risk management and non-zero transaction cost is more realistic.

The accuracy of the C-VaR forecasting is measured by the MAE of the C-VaR forecasts. The C-VaR prediction error is the difference between the C-VaR forecasts and the realized portfolio return when the realized portfolio return is less than the VaR forecasts. For this reason, the evaluation of the VaR forecasting should be conducted a priori. The models with a poor VaR prediction performance are not qualified to be used for C-VaR forecasting. The VaR prediction performance is usually measured by the coverage ratio, which is the frequency for which the realized portfolio returns are below the 10%, 25%, and 50% quantiles of the predictive portfolio return distributions during the out-of-sample periods, respectively. The corresponding results are reported in Table 2(a). If the coverage ratio of a model is closer to (1 - the credibility level), the model is regarded as better for VaR forecasting. A coverage ratio higher than (1 - credibility level) implies underestimation of the downside risk. As one can see from the table, the FV and SVCt models seem to outperform the others, whereas the SVn and SVt models tend to underestimate the downside risk because of the restriction on the conditional correlation. This also indicates that the conditional correlations among the foreign currency returns are non-zero and play a critical role in the foreign currency risk management.

Most importantly, Table 2(b) reports the predictive accuracy of the C-VaR forecasts. The values in this table are the MAE of the C-VaR forecasts. Obviously, the SVCt model outperforms the other models for the benchmark case while the SVCn model seems to perform best for the lower credibility level. This is attributed to the fact that the Student-t error helps to capture the extreme event due to the fat tail property. In addition, the stochastic volatility models with time-varying correlations are preferred to the benchmark FV model. For example, the MAE of the SVCt is 0.279 whereas that of the FV model is 0.443.

Figure 8 plots the time series of the realized portfolio returns, VaR forecasts, and C-VaR forecasts from the FV and SVCt models. As expected, the FV model produces almost constant VaR and C-VaR forecasts over time since the predictive variance-covariance generated from the model is barely time-varying. Meanwhile, the risk measures estimated from the SVCt model change dramatically according to the density forecasts. As a result, their portfolio performance is remarkably different. For instance, the realized return from the FV model in October 2014 and March 2015 is much lower than the C-VaR forecasts compared to the realized return from the SVCt model. This happened because the FV model underestimated the size of risk during the periods. In contrast, unlike the SVCt model, the FV model seems to overestimate the risk in May and June 2014 when the portfolio return is stable.

Finally, Table 3 reports a brief summary of the 10%, 25%, and 50% C-VaR minimizations. The first column of the table shows the average realized portfolio returns over the out-of-sample periods. The second and third columns are average realized C-VaR and predicted C-VaR forecasts, respectively, which exceed the VaR forecasts. For instance, in the case of a 10% C-VaR portfolio, these average values in the second and third columns are computed using the realized returns and C-VaR forecasts only when the realized portfolio return is less than the 10% VaR.

For the 10% C-VaR minimization, the SVn model produces the highest average C-VaR returns, and the SVn model yields the highest average C-VaR forecasts. In case of 25% C-VaR minimization, SVn and SVt seem to perform better than the others in terms of the C-VaR returns and forecasts, while the FV model is the best for 50% C-VaR minimization. However, it is important to notice that the smaller average realized C-VaR and forecasts do not indicate a better out-of-sample portfolio performance, because C-VaR realization is small when the risk is underestimated. This is why we evaluate the portfolio performance in terms of the MAE of the C-VaR forecasts, not average realized portfolio return lower than VaR.

5.4 Statistical Evaluation: Out-of-Sample Prediction Performance

We compare the models in terms of the predictive density forecasting to supplement the economic model selection. Figure 9 plots the log PPLs of the competing models over time, which are obtained from the most recent 100 weeks' density forecasts. This figure demonstrates the relative out-of-sample density prediction accuracy among the models. The result is consistent with that of the C-VaR prediction comparison. After 2013, the SVCt model seems to produce the best density forecasts consistently over the out-of-sample periods. Until 2013, all models except the SVCn reveal similar predictive performance. The performance of the SVCn model improves substantially, so that it becomes better than that of the FV model. After all, incorporating the time-varying volatilities and conditional correlations is found to be critical in improving multiple foreign currency return density forecasting.

6 Concluding Remarks

The contribution of our study is the proposal of a Bayesian method of C-VaR portfolio optimization for foreign currency investment. This consists of two stages. In the first stage, we estimate various multivariate stochastic volatility models for the joint predictive currency return density simulation. Subsequently, given the density forecasts, the best model is chosen based on both predictive density accuracy and C-VaR portfolio performance. Using the best model, one should conduct the C-VaR portfolio optimization to manage the left-tail risk.

Our out-of-sample experiment based on weekly USD/EUR, USD/JPY, and USD/KRW return data indicates that the fat-tailed stochastic model with time-varying conditional correlations produces the most accurate C-VaR forecasts, as well as density forecasts. Particularly, the SVCt model leads to the smallest MAE of the C-VaR forecasts maximizing currency-hedging benefits. Meanwhile, the stochastic volatility models with normal error or correlation tend to underestimate the left tail risk. In addition, the optimal portfolio weights change over time dramatically, and their movement can be smoother by incorporating the transaction cost. Lastly, we would like to emphasize that our Bayesian approach for the C-VaR minimization portfolio choice can be generally used for other investment assets such as stocks, bonds, commodities, and so forth.

References

Alexander, G. J. and Baptista, A. M. (2004), "A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model," *Management Science*, 50, 1261–1273.

Ando, T. (2009), "Bayesian portfolio selection using a multifactor model," *International Journal of Forecasting*, 25, 550–566.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999), "Coherent measures of risk," *Mathematical Finance*, 9, 203–228.

Best, M. J. and Grauer, R. R. (1991), "Sensitivity analysis for mean-variance portfolio problems," *Management Science*, 37, 980–989.

Black, F. and Litterman, R. (1992), "Global portfolio optimization," *Financial Analysts Journal*, 48, 28–43.

Bollerslev, T. (1990), "Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model," *The Review of Economics and Statistics*, 498–505.

Chib, S. and Kang, K. H. (2013), "Change Points in Affine Arbitrage-free Term Structure Models," *Journal of Financial Econometrics*, 11(2), 302–334.

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

Chib, S. and Ramamurthy, S. (2010), "Tailored randomized-block MCMC methods for analysis of DSGE models," *Journal of Econometrics*, 155(1), 19–38.

Diebold, F. X. and Nerlove, M. (1989), "The dynamics of exchange rate volatility: a multivariate latent factor ARCH model," *Journal of Applied Econometrics*, 4, 1–21.

Eklund, J. and Karlsson, S. (2007), "Forecast Combination and Model Averaging Using Predictive Measures," *Econometric Reviews*, 26, 329–363.

Gerlach, R. and Chen, C. W. S. (2016), "Bayesian Expected Shortfall Forecasting Incorporating the Intraday Range," *Journal of Financial Econometrics*, 14, 128–58.

Harvey, A., Ruiz, E., and Shephard, N. (1994), "Multivariate stochastic variance models," *The Review of Economic Studies*, 61, 247–264.

Hoogerheide, L. and van Dijk, H. K. (2010), "Bayesian forecasting of value at risk and expected shortfall using adaptive importance sampling," *International Journal of Forecasting*, 26, 231–247.

Ishihara, T. and Omori, Y. (2016), "Portfolio optimization using dynamic factor and stochastic volatility: evidence on Fat-tailed errors and leverage," *The Japanese Economic Review*.

Jacquier, E., Polson, N. G., Rossi, P. E., et al. (1995), "Models and priors for multivariate stochastic volatility," Tech. rep., CIRANO.

Jorion, P. (1986), "Bayes-Stein estimation for portfolio analysis," *Journal of Financial and Quantitative Analysis*, 21, 279–292.

Kastner, G., Frühwirth-Schnatter, S., and Lopes, H. F. (2014), "Analysis of Exchange Rates via Multivariate Bayesian Factor Stochastic Volatility Models," in *The Contribution of Young Researchers to Bayesian Statistics*, Springer Proceedings in Mathematics and Statistics, pp. 181–185.

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

— (1998b), "Stochastic volatility: Likelihood inference and comparison with ARCH models," *Review of Economic Studies*, 65, 361–393.

Krokhmal, P., Palmquist, J., and Uryasev, S. (2002), "Portfolio optimization with conditional value-at-risk objective and constraints," *Journal of Risk*, 4, 43–68.

Kulikova, M. and Taylor, D. (2013), "Stochastic volatility models for exchange rates and their estimation using quasi-maximum-likelihood methods: an application to the South African Rand," *Journal of Applied Statistics*, 40, 495–507.

Pajor, A. and Osiewalski, J. (2012), "Bayesian value-at-risk and expected shortfall for a large portfolio (multi-and univariate approaches)," *Acta Physica Polonica A*, 121, 101–109.

Pástor, L. (2000), "Portfolio selection and asset pricing models," *The Journal of Finance*, 55, 179–223.

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

Rockafellar, R. T. and Uryasev, S. (2000), "Optimization of conditional value-at-risk," *Journal of Risk*, 2, 493–517.

— (2002), "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, 26, 1443–1471.

Schwert, G. W. (1989), "Why does stock market volatility change over time?" *The Journal of Finance*, 44, 1115–1153.

Tu, J. and Zhou, G. (2010), "Incorporating economic objectives into Bayesian priors: Portfolio choice under parameter uncertainty," *Journal of Financial and Quantative Analysis*, 45, 959–986.

Wang, Z.-R., Chen, X.-H., Jin, Y.-B., and Zhou, Y.-J. (2010), "Estimating risk of foreign exchange portfolio: Using VaR and CVaR based on GARCH–EVT-Copula model," *Physica A: Statistical Mechanics and its Applications*, 389, 4918–4928.

Yu, J. and Meyer, R. (2006), "Multivariate stochastic volatility models: Bayesian estimation and model comparison," *Econometric Reviews*, 25, 361–384.

Zellner, A. and Chetty, V. K. (1965), "Prediction and decision problems in regression models from the Bayesian point of view," *Journal of the American Statistical Association*, 60, 608–616.

Figure 1: Currency Returns This figure plots the weekly currency returns of EUR, JPY, and KRW from January 1999 to May 2016.

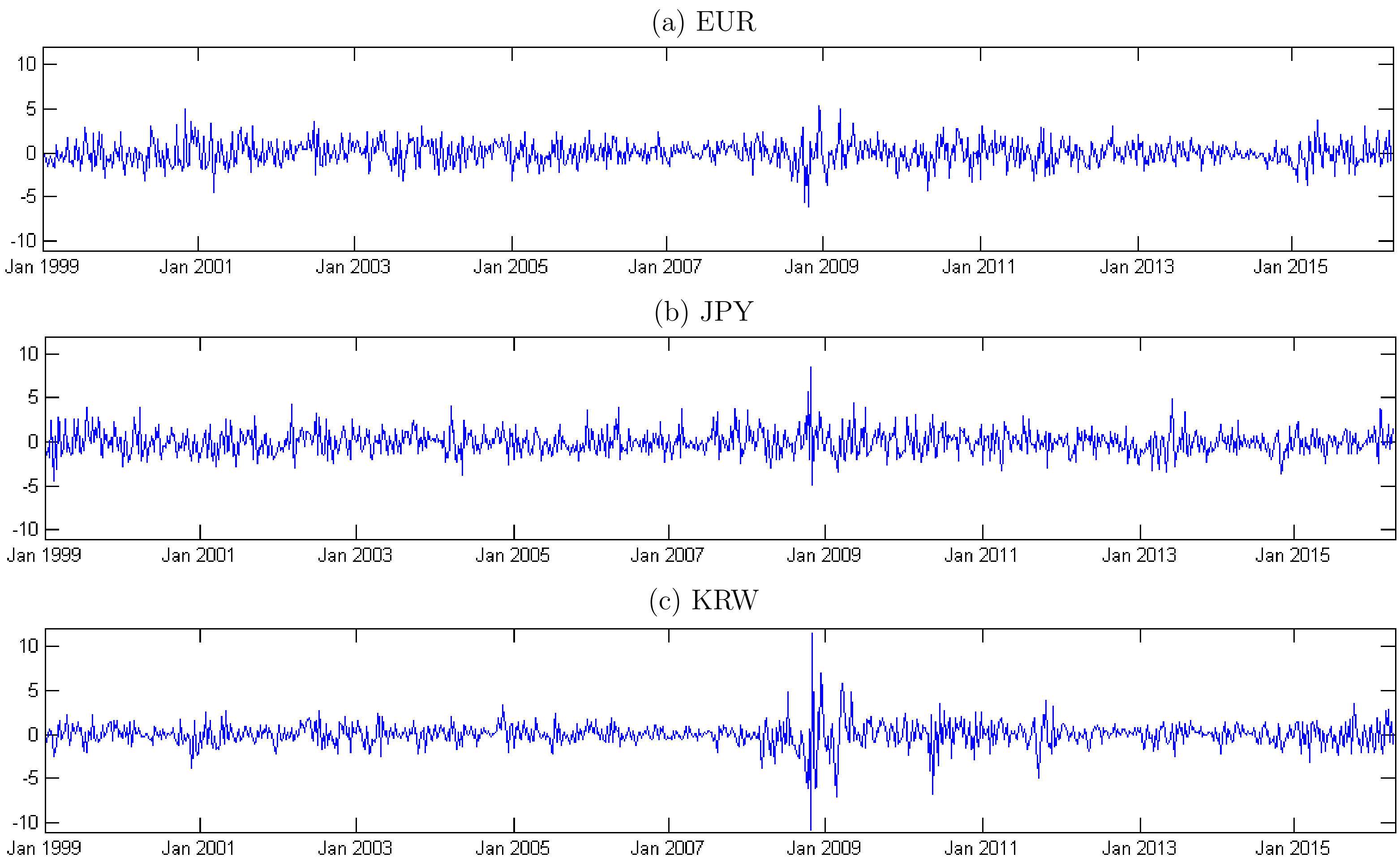


Figure 2: Prior-Posterior densities: SVCt model This figure plots the prior and posterior distributions of the parameters in the SVCt model. The solid and dashed lines are the posterior and prior densities, respectively.

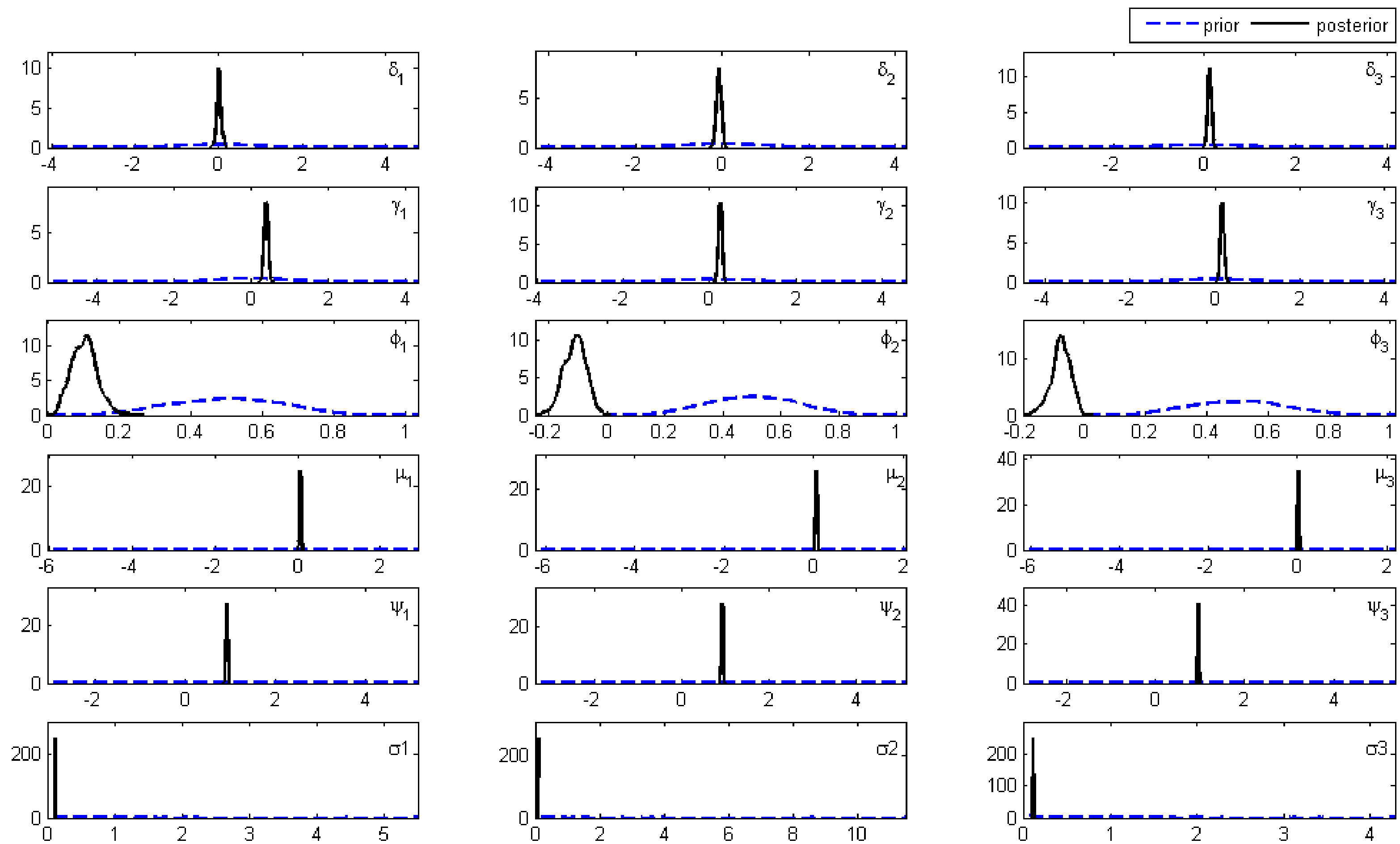


Figure 3: Autocorrelation functions of the MCMC draws: SVCt model This figure plots the autocorrelation functions of the posterior draws from the SVCt model. The MCMC size is 10,000.

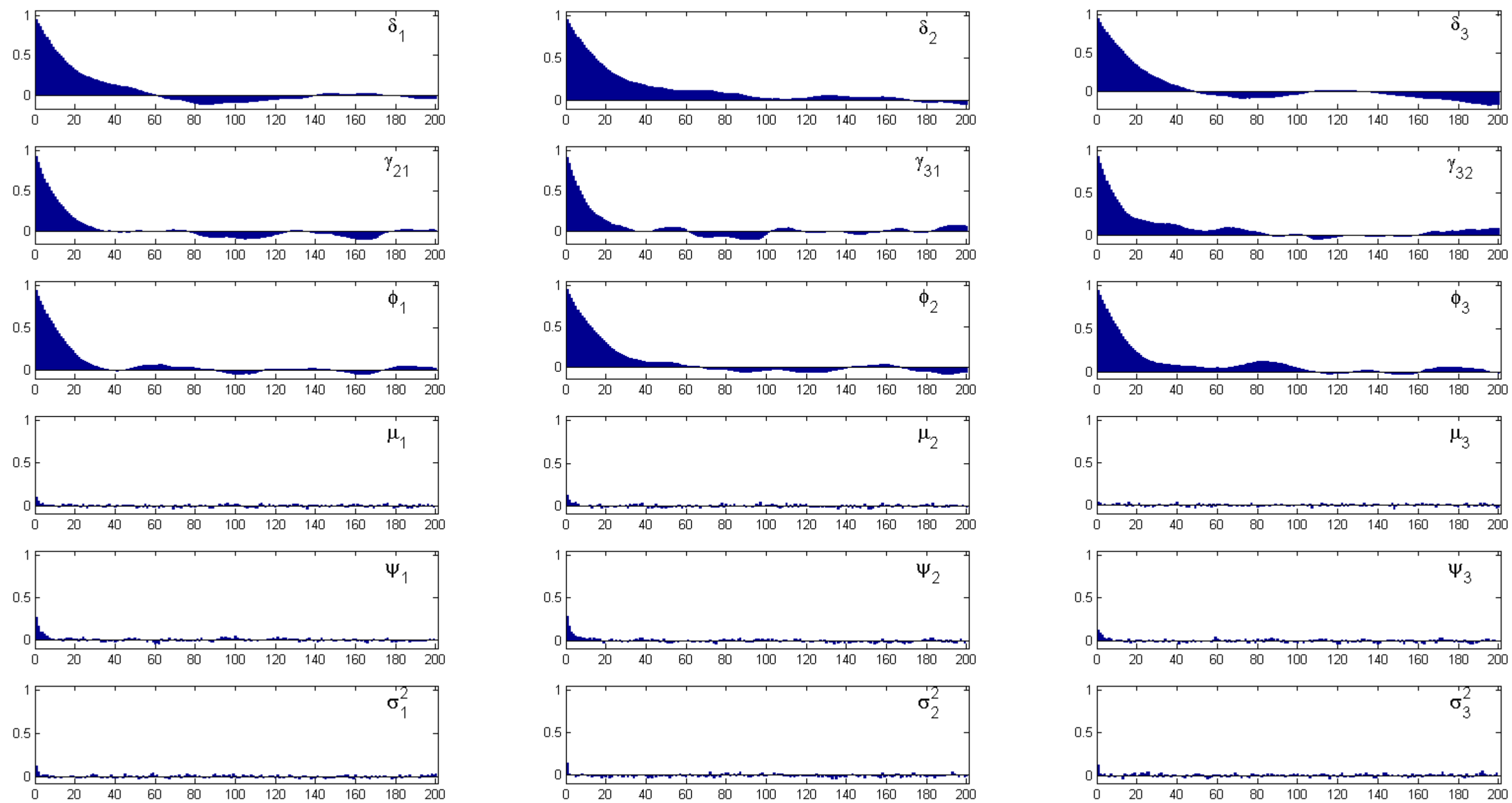
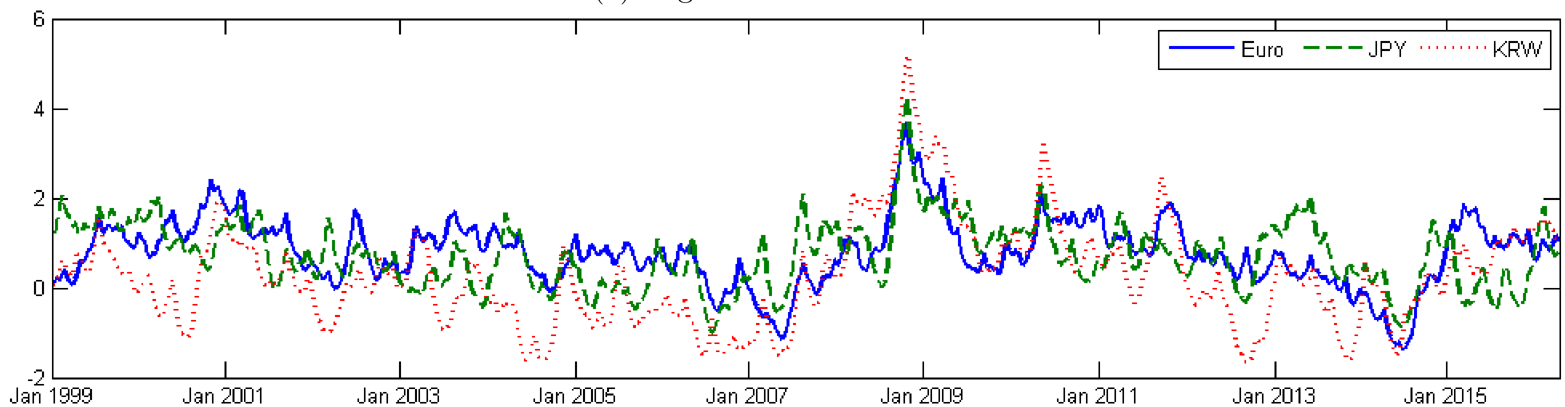


Figure 4: Stochastic Volatilities and Conditional Correlation: SVCt model This figure plots the posterior mean of the volatilities and conditional correlations of the currency returns.

(a) Log stochastic volatilities



(b) Conditional correlations

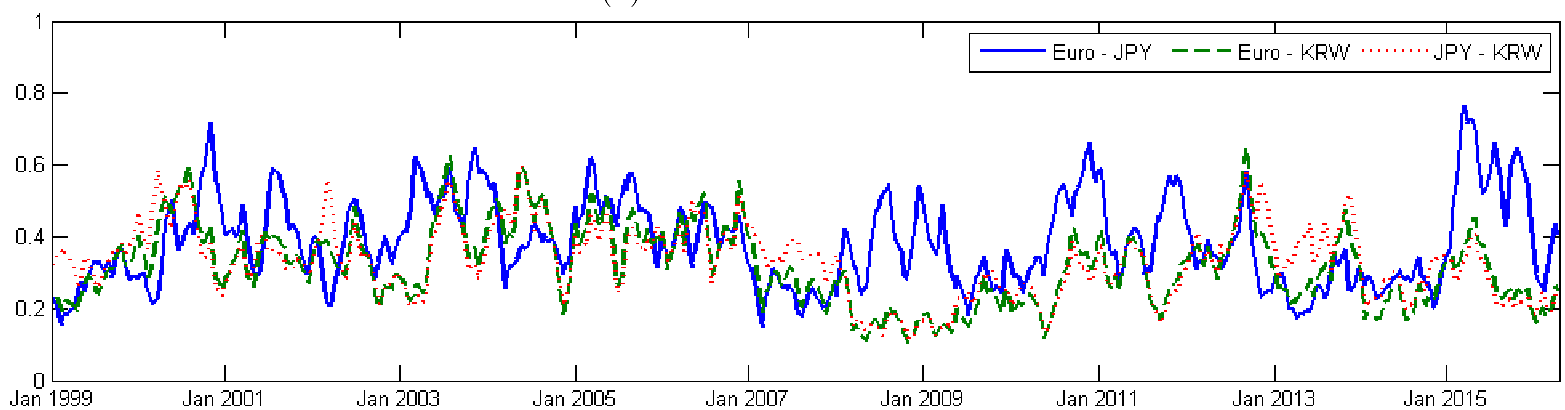


Figure 5: Optimal C-VaR Portfolio Weights: credibility level = 0.9 and transaction cost = 0% This figure plots the time series of the optimal portfolio weights over the out-of-sample periods.

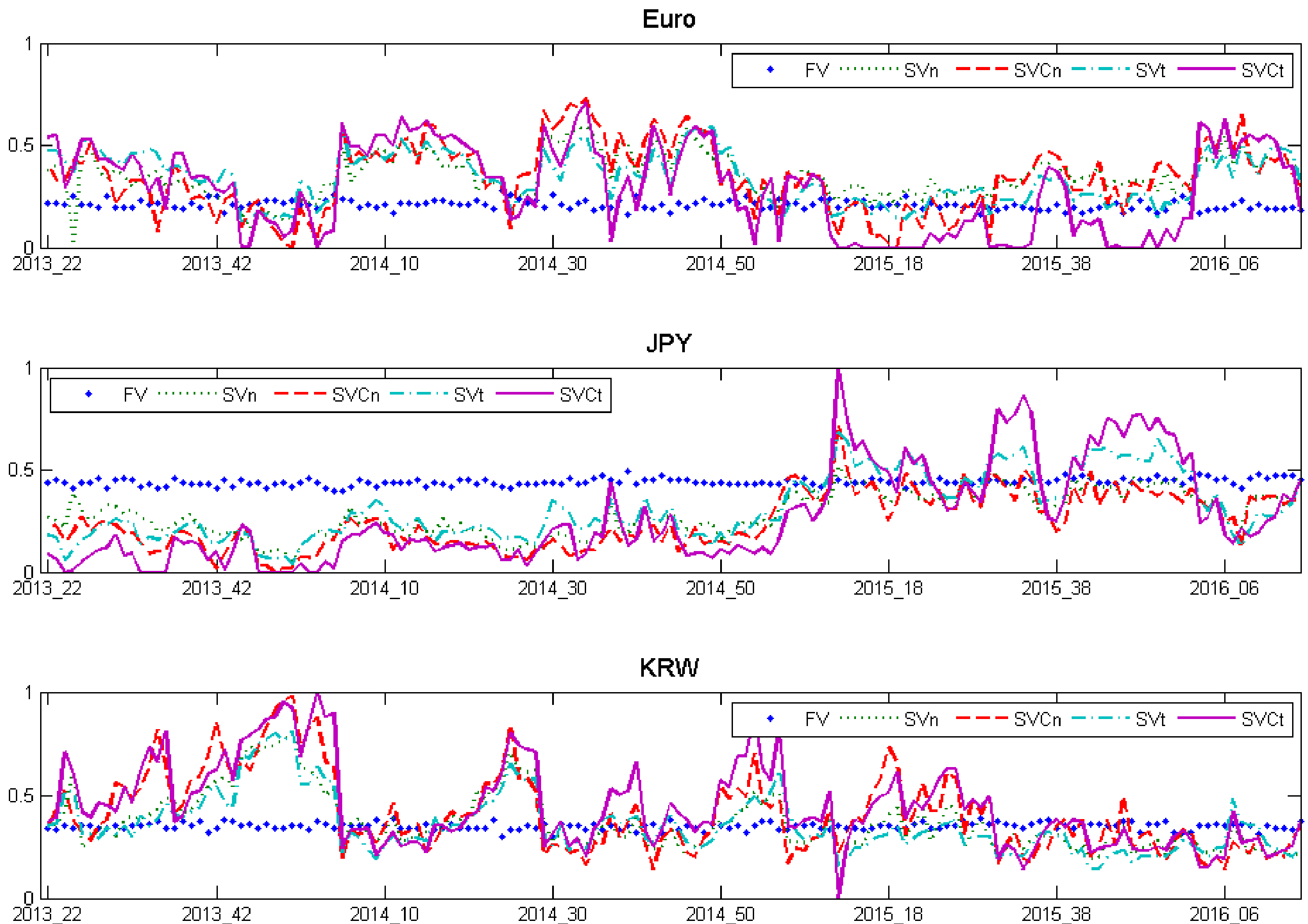


Figure 6: Optimal C-VaR Portfolio Weights: credibility level = 0.9 and transaction cost = 0.1% This figure plots the time series of the optimal portfolio weights over the out-of-sample periods.

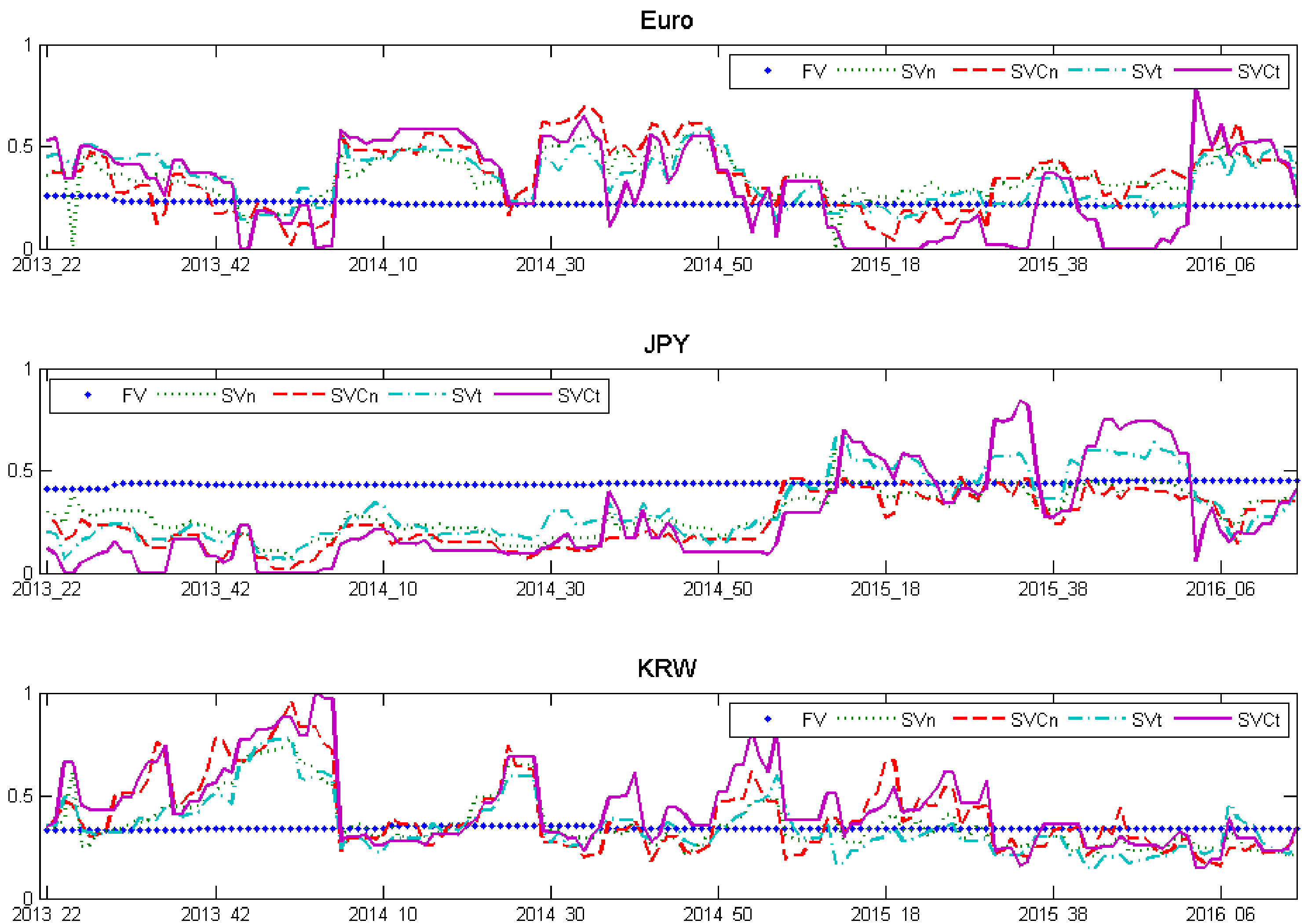
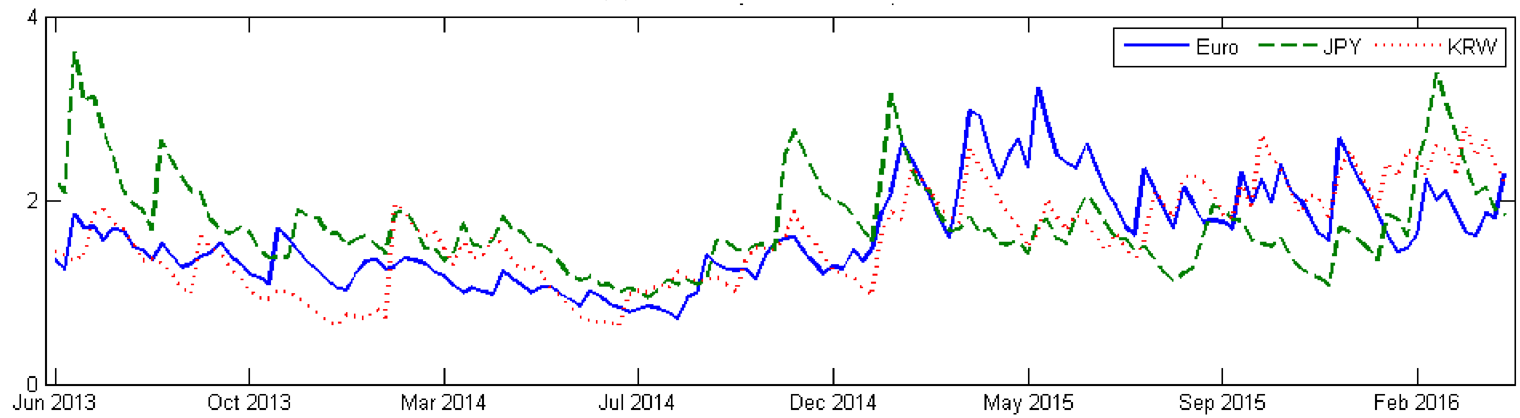


Figure 7: Posterior Predictive Volatilities and Conditional Correlations: SVCt model This figure plots the one-week-ahead posterior mean of the predictive volatilities and conditional correlations of the currency returns over the out-of-sample periods.

(a) Stochastic volatilities



(b) Conditional correlations

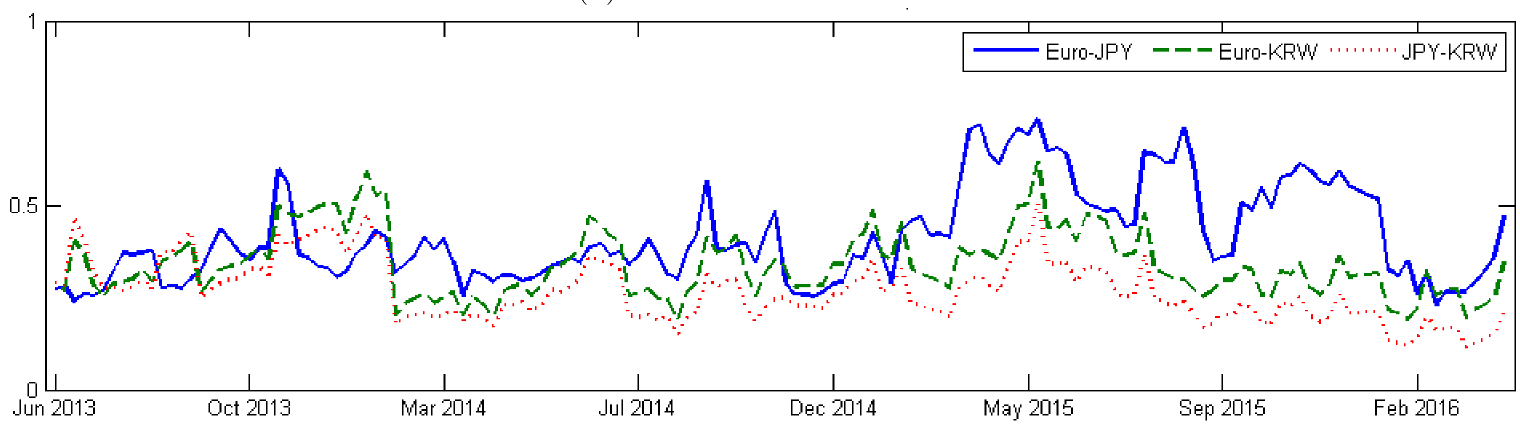
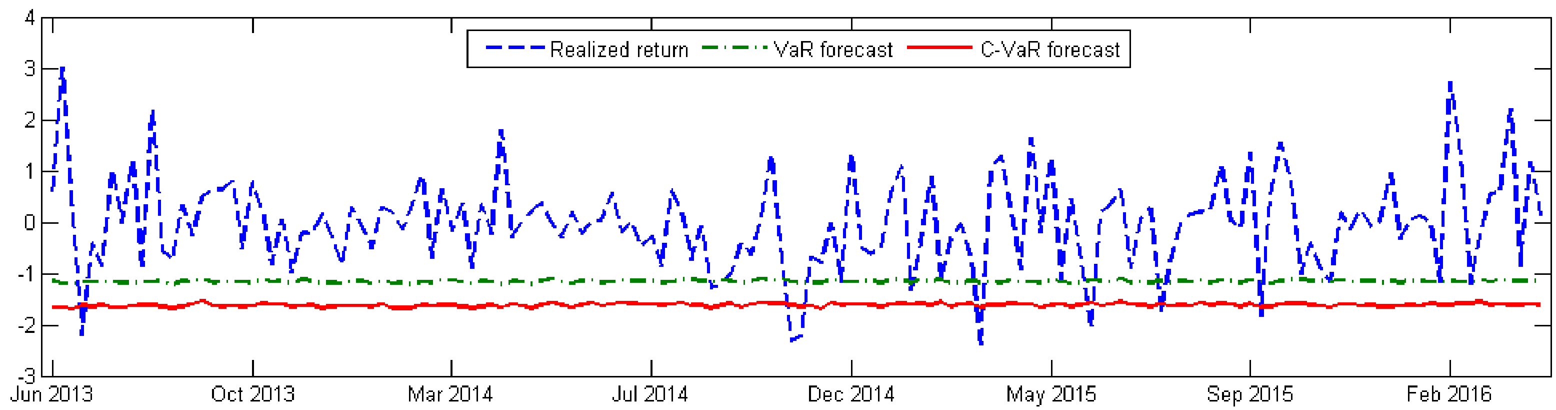


Figure 8: VaR and C-VaR forecasts This figure plots the one-week-ahead VaR and C-VaR forecasts of the portfolio return over the out-of-sample periods.

(a) FV model



(b) SVCt model

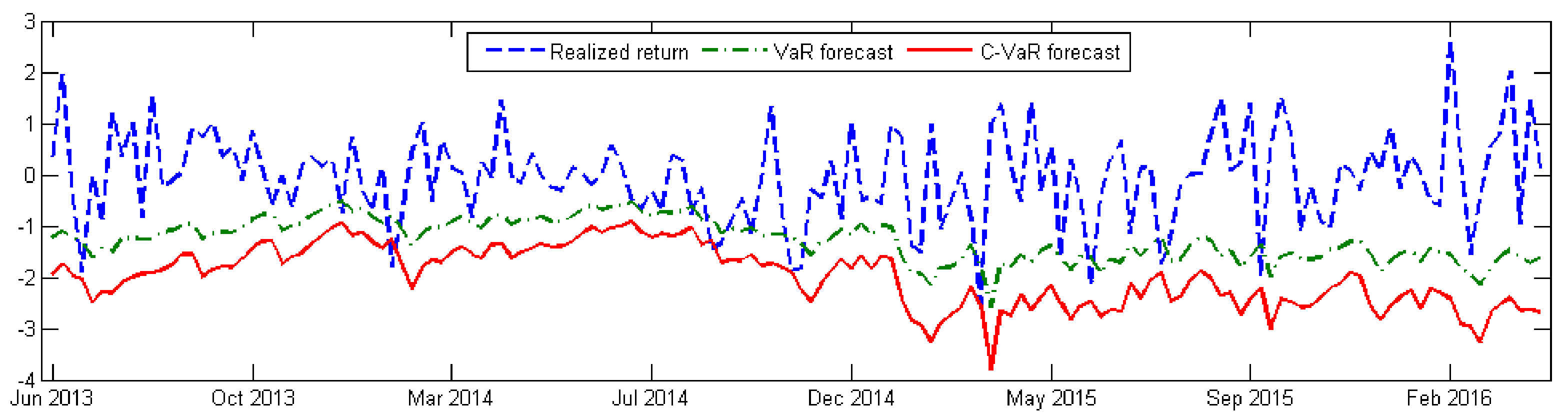


Figure 9: Posterior Predictive Likelihood This figure plots the PPLs of the alternative prediction models over time. The out-of-sample size is 150 weeks and the rolling window size is 750 weeks.

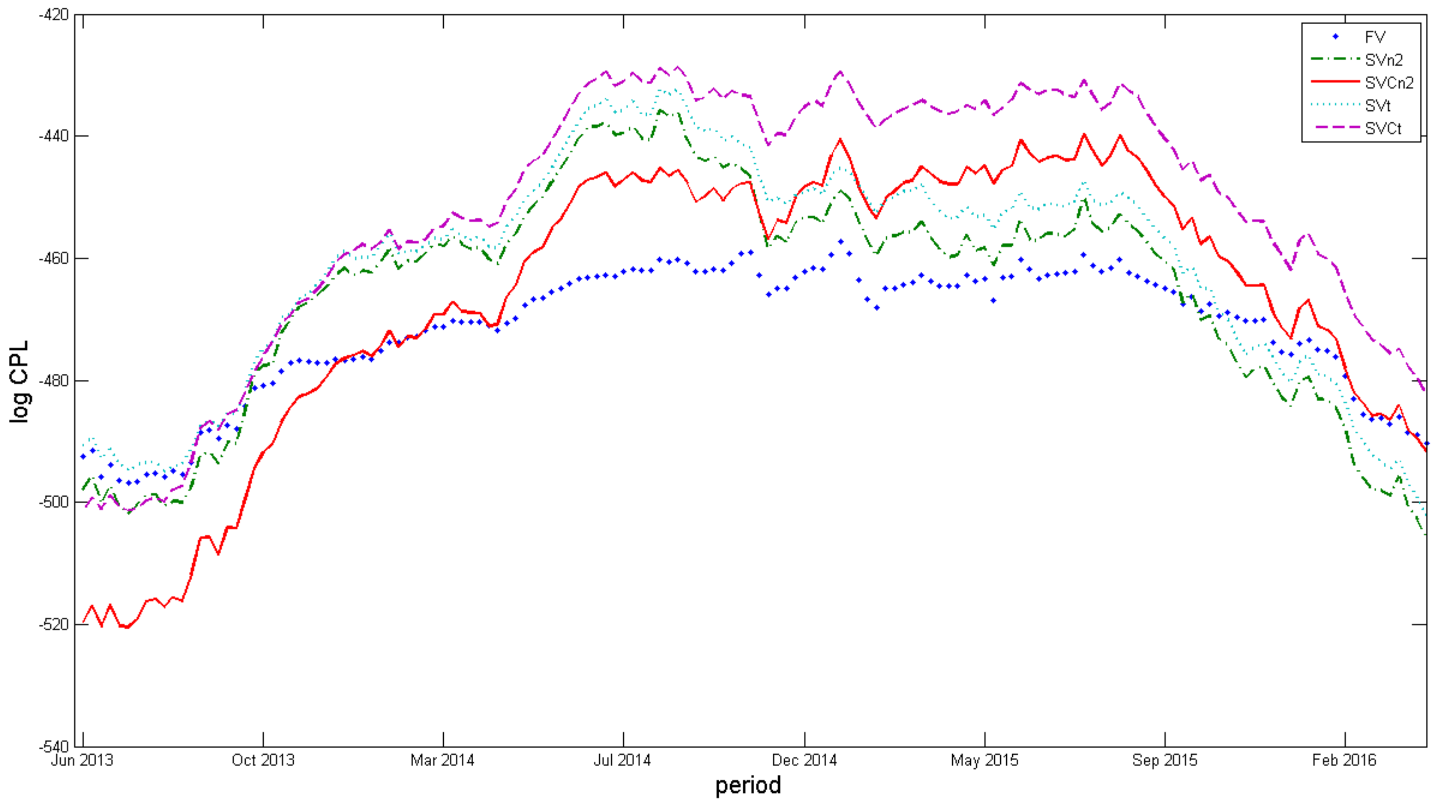


Table 1: Stochastic Volatility Parameters This table reports the posterior mean of the stochastic volatility parameters. The standard errors are in the parentheses. Ineff. is the inefficiency factor of the parameters in the SVCt model.

	SVn	SVCn	SVt	SVCt	Ineff.
δ_1	0.0123	0.0192	0.0034	0.0058	30.309
	(0.0363)	(0.0434)	(0.0591)	(0.0547)	
δ_2	-0.0322	-0.0553	-0.0492	-0.0691	42.298
	(0.0449)	(0.0454)	(0.0568)	(0.0532)	
δ_3	0.0807	0.1004	0.0976	0.1113	28.905
	(0.0322)	(0.0353)	(0.0423)	(0.0383)	
γ_{21}	0.0000	0.3634	0.0000	0.3867	19.533
	(0.0000)	(0.0322)	(0.0000)	(0.0473)	
γ_{31}	0.0000	0.2657	0.0000	0.2521	17.525
	(0.0000)	(0.0274)	(0.0000)	(0.0369)	
γ_{32}	0.0000	0.1532	0.0000	0.1773	26.722
	(0.0000)	(0.0268)	(0.0000)	(0.0381)	
ϕ_1	0.0846	0.0825	0.1127	0.1028	25.416
	(0.0239)	(0.0270)	(0.0388)	(0.0355)	
ϕ_2	0.0593	0.0647	-0.0115	-0.1111	29.671
	(0.0217)	(0.0242)	(0.0973)	(0.0375)	
ϕ_3	0.0971	0.0956	0.1075	-0.0771	30.435
	(0.0311)	(0.0325)	(0.0364)	(0.0314)	
μ_1	0.0257	0.0253	0.0671	0.0634	1.249
	(0.0161)	(0.0172)	(0.0164)	(0.0158)	
μ_2	0.1638	0.0995	0.0872	0.0592	1.985
	(0.0548)	(0.0434)	(0.0180)	(0.0154)	
μ_3	-0.0036	-0.0339	0.0198	0.0119	0.998
	(0.0177)	(0.0270)	(0.0118)	(0.0114)	
φ_1	0.9354	0.9352	0.9184	0.9196	2.378
	(0.0420)	(0.0499)	(0.0147)	(0.0144)	
φ_2	0.4185	0.4477	0.8954	0.9153	3.354
	(0.1446)	(0.1488)	(0.0166)	(0.0144)	
φ_3	0.8957	0.8371	0.9520	0.9550	1.720
	(0.0321)	(0.0493)	(0.0099)	(0.0094)	
σ_1^2	0.0515	0.0516	0.1032	0.1032	1.182
	(0.0367)	(0.0430)	(0.0015)	(0.0016)	
σ_2^2	0.5179	0.6683	0.1039	0.1038	0.884
	(0.1478)	(0.2049)	(0.0016)	(0.0016)	
σ_3^2	0.2416	0.4510	0.1037	0.1038	1.439
	(0.0798)	(0.1590)	(0.0016)	(0.0016)	

Table 2: Predictive Accuracy Comparison of VaR and C-VaR This table summarizes the results of the out-of-sample VaR and C-VaR prediction. The credibility levels(β) considered here are 90%, 75%, and 50%. *Cost* indicates the transaction cost.

(a) Coverage ratio

	Cost=0%			Cost=0.1%		
1-β	10%	25%	50%	10%	25%	50%
FV	10.0%	27.3%	58.0%	10.7%	27.3%	58.0%
SVn	23.3%	34.7%	56.7%	23.3%	34.0%	57.3%
SVCn	12.7%	29.3%	54.0%	13.3%	28.0%	57.3%
SVt	12.7%	31.3%	56.7%	14.7%	32.0%	56.0%
SVCt	10.7%	28.7%	54.0%	9.3%	28.7%	54.0%

(b) Mean absolute error of the C-VaR prediction

	Cost=0%			Cost=0.1%		
1-β	10%	25%	50%	10%	25%	50%
FV	0.438	0.390	0.475	0.443	0.386	0.470
SVn	0.337	0.350	0.440	0.338	0.348	0.454
SVCn	0.312	0.347	0.428	0.315	0.346	0.452
SVt	0.353	0.343	0.459	0.367	0.335	0.453
SVCt	0.317	0.441	0.488	0.279	0.429	0.477

Table 3: Summary of the Realized Portfolio Returns and C-VaR forecasts *mean* is the average of the realized portfolio returns in percentage. *average loss* and *average forecasts* are the averages of the realized portfolio returns and C-VaR forecasts, respectively, when the realized return is less than the VaR forecasts. The credibility levels(β) considered here are 90%, 75%, and 50%.

(a) Transaction cost = 0%

	10% C-VaR Portfolio			25% C-VaR Portfolio			50% C-VaR Portfolio		
	mean	average loss	average forecasts	mean	average loss	average forecasts	mean	average loss	average forecasts
FV	-0.044	1.633	1.619	-0.041	1.123	1.148	-0.043	0.630	0.687
SVn	-0.060	1.180	1.189	-0.058	0.969	0.814	-0.056	0.648	0.479
SVCn	-0.053	1.476	1.559	-0.048	1.067	1.069	-0.040	0.678	0.634
SVt	-0.050	1.333	1.588	-0.052	1.017	1.056	-0.052	0.643	0.622
SVCt	-0.048	1.503	1.981	-0.048	1.105	1.318	-0.044	0.678	0.779

(b) Transaction cost = 0.1%

	10% C-VaR Portfolio			25% C-VaR Portfolio			50% C-VaR Portfolio		
	mean	average loss	average forecasts	mean	average loss	average forecasts	mean	average loss	average forecasts
FV	-0.043	1.600	1.621	-0.044	1.115	1.151	-0.044	0.624	0.691
SVn	-0.070	1.186	1.193	-0.069	0.993	0.819	-0.066	0.644	0.486
SVCn	-0.062	1.459	1.565	-0.057	1.104	1.077	-0.056	0.647	0.645
SVt	-0.061	1.256	1.590	-0.063	1.000	1.060	-0.065	0.651	0.627
SVCt	-0.068	1.575	1.989	-0.066	1.109	1.327	-0.058	0.694	0.790