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

Dongwhan Kim
Kyu Ho Kang
(Korea University)

May 2017

- International bond(currency) portfolio
- Euro bond return in USD = bond return in Euro + currency return in USD
  - |currency return in USD| >> |bond return in Euro|
- High exchange rate risk
  - Left tail risk

Review: Minimum Variance Portfolio

- $y : 2 \times 1 =$ vector of returns
- Predictive return distribution

$$y \sim \mathcal{N}(\mu, \Sigma)$$

- $x : 2 \times 1 =$ vector of portfolio weights
- Portfolio optimization:

$$\underset{x}{\text{Min}} \;\; Var(x'y) = w'\Sigma w$$

s.t.

$$x'1 = 1 \text{ and } \mathbb{E}(x'y) \geq \bar{\mu}.$$

Motivation

- Challenges in practice
  1. Fast and accurate predictive return distribution computation
  2. Choice of adequate risk measure
  3. Feasible solution for the portfolio optimization

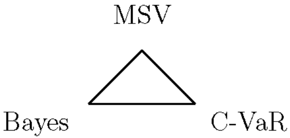
A Bayesian method for foreign portfolio selection

1. Fast and accurate predictive return distribution computation
   - **Multivariate stochastic volatility (MSV) model + Bayesian approach**
2. Choice of adequate risk measure
   - **C-VaR**
3. Feasible solution for the portfolio optimization
   - **Bayesian approach**

What's new?

To the best of our knowledge,

- this paper is the first work that attempts **C-VaR** portfolio optimization using **a MSV model** in a **Bayesian framework**



- A new MSV model for foreign currency return density prediction

Why MSV + Bayesian?

MSV

- Time-varying volatility
- Time-varying conditional correlation
- Fat-tail
- Fast inference
- No factor identification problem

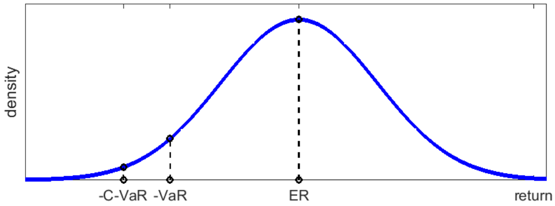
Bayesian approach

- Estimation risk: parameter and model uncertainties
- Density forecasting
- Numerical solution

Why C-VaR?

- Left tail risk
- VaR = minimum loss
- C-VaR (expected shortfall) = expected loss
  - extreme risk beyond the VaR
  - sub-additivity property

Figure: **VaR vs C-VaR**



- β = confidence level
- ζ = (1-β) level VaR
- $y : k \times 1$ = vectors of h-period-ahead returns
- $x : k \times 1$ = portfolio weight
- $x'y$ = portfolio return
- C-VaR Minimization:

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

subject to

$$\mathbb{E}[x'y] \geq \bar{\mu},$$

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

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 expected return constraint

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 expected return constraint

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.

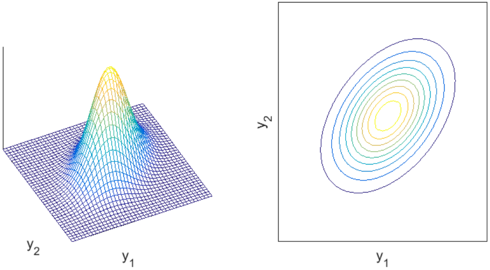
Step 5: Repeat steps 1 to 4 a number of times.

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

- Two asset returns: y_1 and y_2
- **Step 1**: Model($\mathcal{M}$)-dependent joint predictive distribution,

$$(y_1, y_2)|\mathcal{M}$$

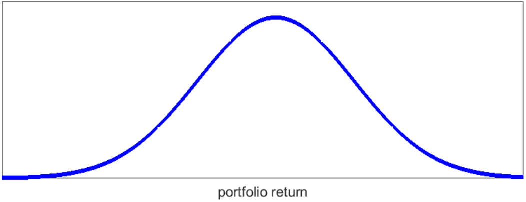
Figure: **Joint predictive distribution**



- Predictive portfolio return distribution conditioned on a portfolio weight (x),

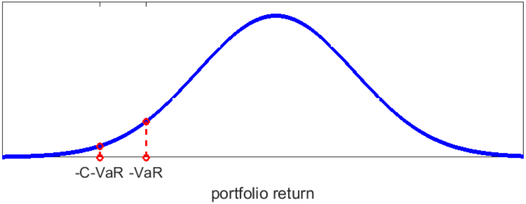
$$x \times y_1 + (1 - w) \times y_2 | \mathcal{M}, x$$

Figure: **Portfolio return distribution**



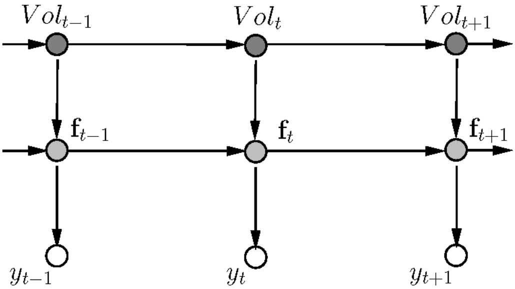
- **Step 2**: VaR($\mathcal{M}, x$) and C-VaR($\mathcal{M}, x$)
- **Step 3**: Min C-VaR($\mathcal{M}, x$) w.r.t the portfolio weight (x)
- **Step 4**: Model choice based on out-of-sample portfolio performance evaluation

Figure: **VaR and C-VaR**



DGP

Figure: **Data generating process**



- $y_t = [y_{1t}, ..., y_{kt}]'$ is a $k \times 1$ vector of foreign currency returns a time t.
- $\mathbf{f}_t = k \times 1$ vector of time-varying latent factors
- Measurement equation:

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

- Factor process with stochastic volatility:

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

where

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

- Suppose that $\lambda_t \sim \mathcal{G}(\nu/2, \nu/2)$. By the method of composition,

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

- one-to-one mapping between the returns and factors

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

- $\mathcal{F}_t = \{y_i\}_{i=0}^{t}$ = the information up to time t.
- $\phi = diag(\phi_1, \phi_2, .., \phi_k)$ and $V_t = diag\left(V_{1t}, V_{2t}, .., V_{kt}\right)$
- the joint conditional distribution of the returns as

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

- $k = 3$
- The conditional variance-covariance matrix of the returns at time t,

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

- The unconditional mean of the returns

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

- Factor loadings

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

- The autoregressive coefficient

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

- Log SV persistence

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

- The intercept term in the log SV process

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

- The conditional variance of the log SV

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

Candidate Prediction Models

- VAR(1): benchmark
- SVn: normal errors + zero conditional correlations
- SVCn: normal errors + time-varying conditional correlations
- SVt: fat tail + zero conditional correlations
- SVCt: fat tail + time-varying conditional correlations

- The set of model parameters

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

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 factors, $\mathbf{F} = \{\mathbf{f}_t\}_{t=1}^T$
- The log SV, $\mathbf{A} = \{\{\alpha_{i,t}\}_{i=1}^k\}_{t=1}^T$
- The observations, $\mathbf{Y} = \{y_t\}_{t=1}^T$

- Joint posterior distribution

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

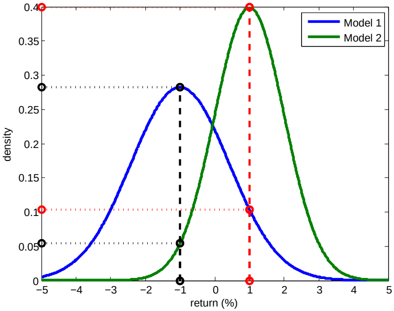
- Its density is given by

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

Evaluation

- Statistical evaluation
  - Predictive density accuracy: posterior predictive likelihood (PPL)
- Economic evaluation
  - VaR predictive accuracy
  - C-VaR predictive accuracy
  - High predictive accuracy = Efficient currency-hedging

Posterior predictive likelihood



Predictive density accuracy

- The posterior predictive likelihood (PPL)

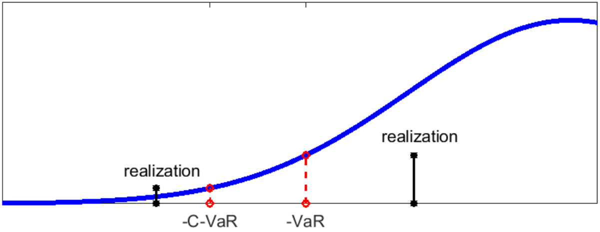
$$\mathrm{PPL}(\mathcal{M}) = \prod_{h=1}^{H} p(y^*_{T+h} | \mathcal{F}_{T+h-1}, \mathcal{M})$$

- This is the product of the posterior predictive densities over the out-of-sample periods

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

Economic evaluation

- VaR predictive accuracy
- C-VaR predictive accuracy



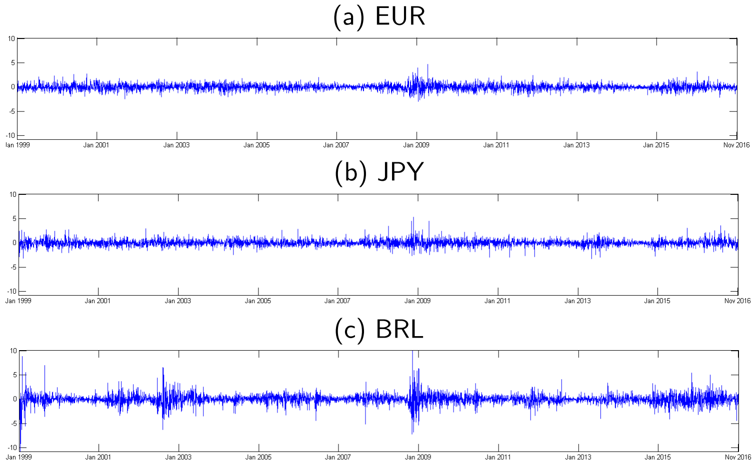
- OSP is the set of the out-of-sample periods
- y_t^* is the realized portfolio return, and $\widehat{\mathrm{VaR}_t}$ and $\widehat{\mathrm{C\text{-}VaR}_t}$ are one-week-ahead VaR and C-VaR forecasts at time t, respectively.
- MAE of the one-week-ahead C-VaR forecasts

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

- Large MAE of the C-VaR forecasts = small currency-hedging benefits

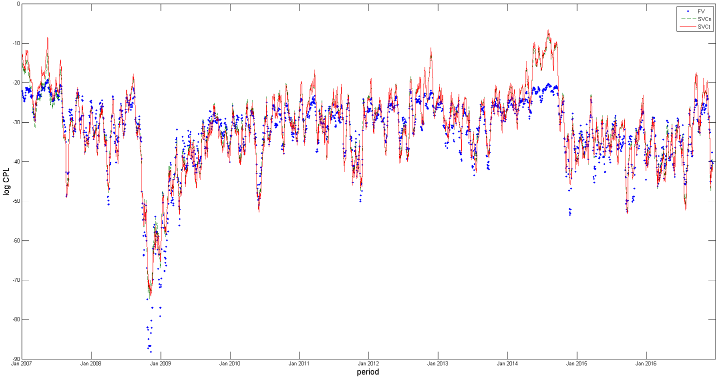
- Daily USD/EUR, USD/JPY, and USD/BRL returns
- January 1999 to December 2016
- The rolling window size = eight years
- Out of sample size = 10 years
- Transaction cost = 0% or 0.1%

Figure: **Currency Returns**



Statistical evaluation

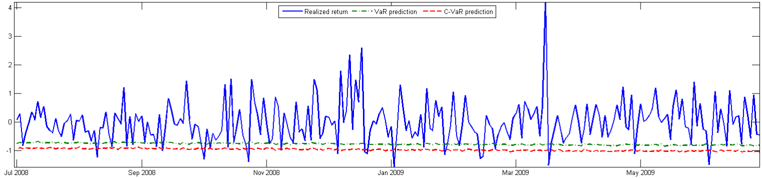
Figure: **Posterior Predictive Likelihood**



Economic evaluation

Figure: **VaR and C-VaR forecasts**

(a) VAR



(b) SVCn

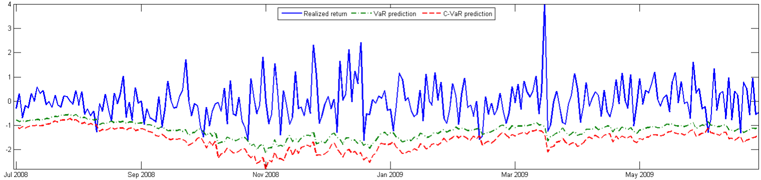


Table: **Predictive Accuracy Comparison of VaR and C-VaR**

(a) Coverage ratio

	Cost=0%		Cost=0.1%	
1-β	1%	5%	1%	5%
FV	1.53%	4.50%	1.61%	4.50%
SVn	0.36%	2.17%	0.36%	2.13%
SVCn	0.88%	3.69%	0.80%	3.73%
SVt(ν=20)	0.12%	2.07%	0.12%	2.01%
SVCt(ν=20)	0.56%	3.13%	0.60%	3.21%

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

	Cost=0%		Cost=0.1%	
1-β	1%	5%	1%	5%
FV	0.236	0.226	0.238	0.225
SVn	0.254	0.195	0.281	0.188
SVCn	0.206	0.223	0.229	0.212
SVt(ν=20)	0.220	0.218	0.193	0.215
SVCt(ν=20)	0.226	0.216	0.203	0.212

Figure: **Optimal C-VaR Portfolio Weights: credibility level = 0.9 and transaction cost = 0%**

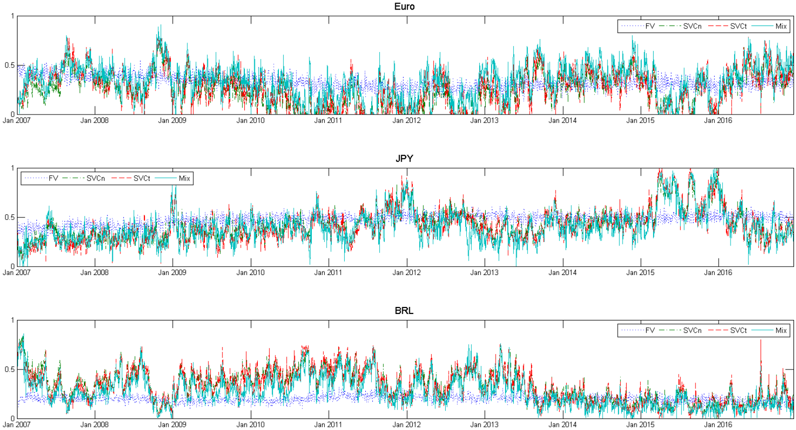
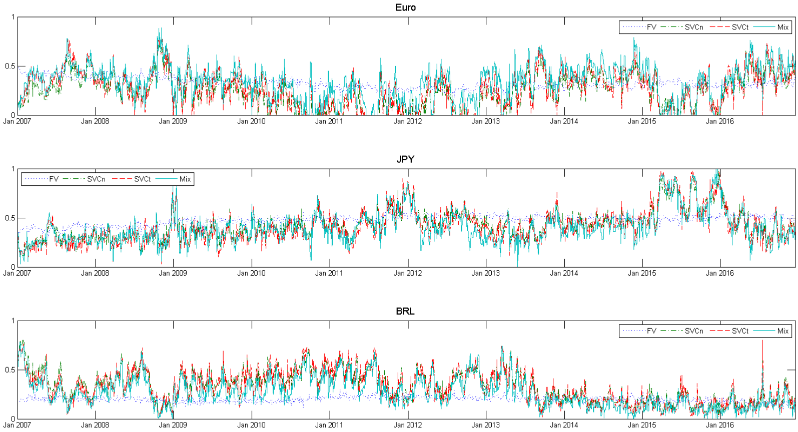


Figure: **Optimal C-VaR Portfolio Weights: credibility level = 0.9 and transaction cost = 0.1%**



What to be done in this work

- Larger number of currencies
- Risk-free asset
- Short-sale
- Constant conditional correlation SV model

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

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

- The conditional density of $\mathbf{A}$, $f(\mathbf{A}|\theta)$

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

- the prior density of the parameters, $\pi(\theta)$

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 $y_{T+1}|\theta, \mathbf{A}, \mathbf{Y}$.

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

- Given the factors

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

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

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

- 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 $\varepsilon^*_{i,t} = \log(\varepsilon^2_{i,t})$.

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

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

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

- Gibbs-sampling

$$\alpha_{i,t} | \theta, \alpha_{i,t-1} \sim \mathcal{N}(\mu_i + \varphi_i \alpha_{i,t-1}, \sigma_i^2)$$

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

- Tailored Metropolis-Hastings algorithm
- The full conditional density of γ is given by

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

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

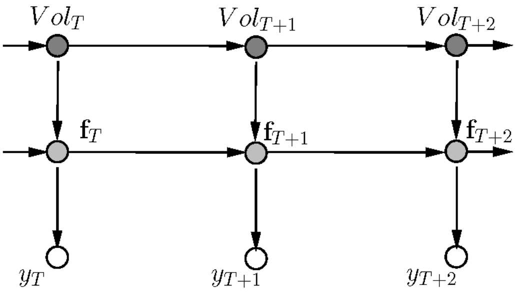
- Tailored Metropolis-Hastings algorithm
- The full conditional density

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

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 is obtained as a gamma distribution such as
$$\lambda_t|\delta, \phi, \Gamma, \mathbf{A}, \mathbf{Y} \sim \mathcal{G}\left(\frac{\nu+1}{2}, \frac{\nu + \tilde{y}_t' \Sigma_t^{-1} \tilde{y}_t}{2}\right)$$
where $\tilde{y}_t = y_t - \delta + \Gamma\phi\Gamma^{-1}\delta - \Gamma\phi\Gamma^{-1}y_{t-1}$ and $\Sigma_t = \Gamma\mathbf{V}_t\mathbf{V}_t'\Gamma'$.

Posterior predictive density simulation



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

Table: **Summary of the Realized Portfolio Returns and C-VaR forecasts**

(a) Transaction cost = 0%

	1% C-VaR Portfolio			5% C-VaR Portfolio		
	mean	average loss	average forecasts	mean	average loss	average forecasts
FV	-0.0045	-1.456	-1.281	-0.0050	-1.112	-0.992
SVn	-0.0004	-1.694	-1.935	-0.0002	-1.262	-1.401
SVCn	-0.0001	-1.529	-1.557	0.0014	-1.103	-1.138
SVt(ν=20)	0.0025	-1.858	-2.244	0.0028	-1.232	-1.561
SVCt(ν=20)	0.0037	-1.634	-1.829	0.0025	-1.103	-1.281

(b) Transaction cost = 0.1%

	1% C-VaR Portfolio			5% C-VaR Portfolio		
	mean	average loss	average forecasts	mean	average loss	average forecasts
FV	-0.0071	-1.443	-1.284	-0.0041	-1.118	-0.994
SVn	-0.0114	-1.592	-1.940	-0.0062	-1.242	-1.407
SVCn	-0.0089	-1.589	-1.560	-0.0048	-1.087	-1.142
SVt(ν=20)	-0.0130	-1.836	-2.249	-0.0059	-1.229	-1.567
SVCt(ν=20)	-0.0086	-1.580	-1.832	-0.0055	-1.100	-1.286

Figure: **Stochastic Volatilities and Conditional Correlation: SVCn model**

(a) Log stochastic volatilities

(b) Conditional correlations

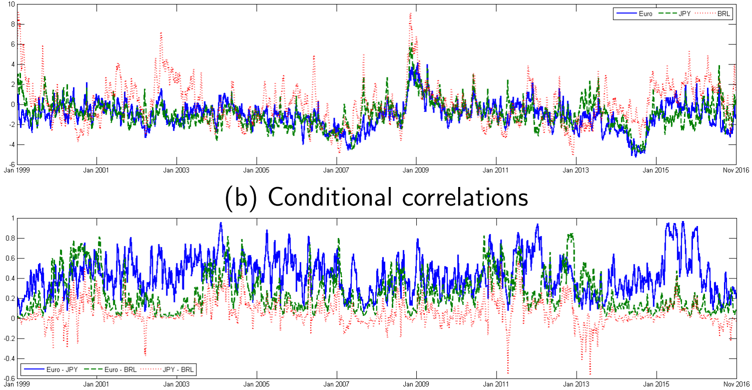


Figure: **Posterior Predictive Volatilities and Conditional Correlations: SVCn**

(a) Stochastic volatilities