Vector autoregression models with skewness and heavy tails

Sune Karlsson[a], Stepan Mazur[a,b] and Hoang Nguyen[a]

[a]Unit of Statistics, School of Business, Örebro University, Sweden
[b]Department of Economics and Statistics, School of Business and Economics, Linnaeus University, Sweden

March 21, 2022

Online Appendix

A Bayesian estimation

This section summarizes the model setup and MCMC samplers for the models where the specification or the MCMC sampler is not discussed in detail in the paper.

A.1 Gaussian VAR with constant variance

The model is $\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\mathbf{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$ with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$ and $\mathbf{\Sigma} = \mathrm{diag}(\tau_1^2, \ldots, \tau_k^2)$.

A.1.1 Prior distribution

The priors for $\mathbf{B}$ and $\mathbf{A}$ are as described in section 3.1. The prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$.

A.1.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of Section 3.2, setting $\widetilde{\mathbf{u}}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \mathbf{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2 + 1)/2)$ for $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{A}\widetilde{\mathbf{u}}_t$

A.2 Gaussian VAR with SV

The model is $\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$ with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$ with $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt})$ and the log volatilities evolving as

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, \ldots, k,$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.2.1 Prior distribution

The priors for $\mathbf{B}$, $\mathbf{A}$, $\boldsymbol{\sigma}^2$ and $\mathbf{h}_0$ are as described in section 3.1.

A.2.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of Section 3.2, setting $\widetilde{\mathbf{u}}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$
3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2
4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2

A.3 VAR with multivariate $t$distribution and constant variance

The model is $\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \sqrt{\xi_t}\mathbf{A}^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$ with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \mathrm{diag}(\tau_1^2, \ldots, \tau_k^2)$ and $\xi_t \sim \mathcal{IG}(\nu/2, \nu/2)$.

A.3.1 Prior distribution

The priors for $\mathbf{B}$ and $\mathbf{A}$ are as described in section 3.1. The prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$ and the prior for the single ν is $\mathcal{G}(2, 0.1)$ as in section 3.1.

A.3.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = (\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)/\sqrt{\xi_t}$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2 + 1)/2)$ for $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{A}\widetilde{\mathbf{u}}_t$
4. The full conditional posterior for ν is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^T \mathcal{IG}(\xi_t, \nu/2, \nu/2)$ and is sampled as in step 5 of section 3.2
5. Sample ξ_t from the full conditional inverse Gamma posterior, $\mathcal{IG}((\nu+k)/2, (\nu+\sum_{i=1}^k u_{it}^2/\tau_i^2)/2)$ for $\mathbf{u}_t = \mathbf{A}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)$

A.4 VAR with multivariate $t$distribution and SV

The model is $\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \sqrt{\xi_t}\mathbf{A}^{-1}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$ with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\xi_t \sim \mathcal{IG}(\nu/2, \nu/2)$, $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt})$ and the log volatilities evolving as

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, \ldots, k,$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.4.1 Prior distribution

The priors for $\mathbf{B}$, $\mathbf{A}$, $\mathbf{h}_0$ and $\boldsymbol{\sigma}^2$ are as described in section 3.1 and the prior for the single ν is $\mathcal{G}(2, 0.1)$ as in section 3.1.

A.4.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = (\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)/\sqrt{\xi_t}$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$
3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2
4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2
5. The full conditional posterior for ν is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^{T}\mathcal{IG}(\xi_t, \nu/2, \nu/2)$ and is sampled as in step 5 of section 3.2
6. Sample ξ_t from the full conditional inverse Gamma posterior, $\mathcal{IG}((\nu+k)/2, (\nu+\sum_{i=1}^{k} u_{it}^2/h_{it})/2)$ for $\mathbf{u}_t = \mathbf{A}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)$

A.5 VAR with multivariate skew-t distribution and constant variance

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + (\xi_t - \mu_\xi)\boldsymbol{\gamma} + \xi_t^{1/2}\mathbf{A}^{-1}\boldsymbol{\Sigma}^{1/2}\epsilon_t$$

with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \mathrm{diag}(\tau_1^2, \ldots, \tau_k^2)$ and $\xi_t \sim \mathcal{IG}(\nu/2, \nu/2)$

A.5.1 Prior distribution

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$ and $\mathbf{A}$ are as described in section 3.1. The prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$ and the prior for the single ν is $\mathcal{G}(2, 0.1)$ as in section 3.1.

A.5.2 MCMC

1. Sample $\mathbf{B}$ and $\boldsymbol{\gamma}$ jointly as $\boldsymbol{b} = (vec(\mathbf{B})', \boldsymbol{\gamma}')'$. Rewrite the model as
$$\begin{aligned}
\xi_t^{-1/2}\mathbf{A}\mathbf{y}_t &= \xi_t^{-1/2}\mathbf{A}\mathbf{B}\mathbf{x}_t + \xi_t^{-1/2}(\xi_t - \mu_\xi)\mathbf{A}\boldsymbol{\gamma} + \boldsymbol{\Sigma}^{1/2}\epsilon_t \qquad (1)\\
&= \left(\xi_t^{-1/2}\mathbf{x}_t' \otimes \mathbf{A} \quad \xi_t^{-1/2}(\xi_t - \mu_\xi)\mathbf{A}\right)\begin{pmatrix} vec(\mathbf{B}) \\ \boldsymbol{\gamma}\end{pmatrix} + \boldsymbol{\Sigma}^{1/2}\epsilon_t\\
\widetilde{\mathbf{y}}_t &= \widetilde{\mathbf{X}}_t\boldsymbol{b} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t.
\end{aligned}$$
a multivariate regression model with error variance $\boldsymbol{\Sigma}$ where $\widetilde{\mathbf{y}}_t = \xi_t^{-1/2}\mathbf{A}\mathbf{y}_t$ and $\widetilde{\mathbf{X}}_t = \left(\xi_t^{-1/2}\mathbf{x}_t' \otimes \mathbf{A} \quad \xi_t^{-1/2}(\xi_t - \mu_\xi)\mathbf{A}\right)$. $\boldsymbol{b}$ is then sampled from the full conjugate normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = \xi_t^{-1/2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t - \xi_t^{-1/2}(\xi_t - \mu_\xi)\boldsymbol{\gamma})$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2 + 1)/2)$ for $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{A}\widetilde{\mathbf{u}}_t$
4. The full conditional posterior for ν is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^T \mathcal{IG}(\xi_t, \nu/2, \nu/2)$ and is sampled as in step 5 of section 3.2
5. To sample from the full conditional posterior for ξ_t, write
$$\mathbf{u}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t + \mu_\xi\boldsymbol{\gamma} = \xi_t\boldsymbol{\gamma} + \xi_t^{1/2}\mathbf{A}^{-1}\boldsymbol{\Sigma}^{1/2}\epsilon_t.$$
$\mathbf{u}_t$ is thus conditionally normal, $\mathbf{u}_t \sim \mathcal{N}(\xi_t\boldsymbol{\gamma}, \xi_t\mathbf{A}^{-1}\mathbf{H}_t\mathbf{A}^{-1\prime})$ and the likelihood contribution is
$$\begin{aligned}
&\xi_t^{-k/2}\exp\left\{-\frac{1}{2\xi_t}(\mathbf{u}_t - \xi_t\boldsymbol{\gamma})'\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}(\mathbf{u}_t - \xi_t\boldsymbol{\gamma})\right\} \qquad (2)\\
&\propto \xi_t^{-k/2}\exp\left\{\frac{1}{2}\left(\frac{\mathbf{u}_t'\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}\mathbf{u}_t}{\xi_t} + \xi_t\boldsymbol{\gamma}'\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}\boldsymbol{\gamma}\right)\right\}.
\end{aligned}$$
With the independent inverse Gamma distribution for ξ_t we have conditional independence and the full conditional posterior is generalized inverse Gaussian (GIG)
$$\begin{aligned}
\pi(\xi_t|\boldsymbol{\Psi}) &\propto \xi_t^{-k/2}\exp\left\{-\frac{1}{2}\left(\frac{q_t^2}{\xi_t} + \xi_t p_t^2\right)\right\}\xi_t^{-(\nu/2+1)}\exp\left\{-\frac{\nu}{2\xi_t}\right\}\\
&= \xi_t^{\lambda-1}\exp\left\{-\frac{1}{2}\left(\frac{\chi}{\xi_t} + \psi\xi_t\right)\right\}
\end{aligned}$$
where $q_t^2 = \mathbf{u}_t'\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}\mathbf{u}_t$ and $p_t^2 = \boldsymbol{\gamma}'\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A}\boldsymbol{\gamma}$. The conditional distribution of $\pi(\xi_t|\boldsymbol{\Psi})$ is $GIG(\lambda, \psi, \chi)$ with $\lambda = -(\nu + k)/2$, $\chi = q_t^2 + \nu$ and $\psi = p_t^2$.

A.6 VAR with multivariate skew-t distribution and SV

The model is
$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + (\xi_t - \mu_\xi)\boldsymbol{\gamma} + \xi_t^{1/2}\mathbf{A}^{-1}\mathbf{H}_t^{1/2}\epsilon_t$$
with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\xi_t \sim \mathcal{IG}(\nu/2, \nu/2)$, $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt})$ and the log volatilities evolving as
$$\log h_{it} = \log h_{it-1} + \sigma_i\eta_{it}, \quad i = 1, \ldots, k,$$
where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.6.1 Prior distribution

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$, $\mathbf{A}$, $\mathbf{h}_0$ and $\boldsymbol{\sigma}^2$ are as described in section 3.1 and the prior for the single ν is $\mathcal{G}(2, 0.1)$ as in section 3.1.

A.6.2 MCMC

1. Sample Sample $\mathbf{B}$ and $\boldsymbol{\gamma}$ jointly from the full conditional normal posterior. To this end, rewrite the model as in (1) to yield $\widetilde{\mathbf{y}}_t = \widetilde{\mathbf{X}}_t \boldsymbol{b} + \mathbf{H}_t^{1/2} \boldsymbol{\epsilon}_t$, a heteroskedastic multivariate regression.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = \xi_t^{-1/2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t - \xi_t^{-1/2}(\xi_t - \mu_\xi)\boldsymbol{\gamma})$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$
3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2
4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2
5. The full conditional posterior for ν is proportional to $\mathcal{G}(2, 0.1) \prod_{t=1}^T \mathcal{IG}(\xi_t, \nu/2, \nu/2)$ and is sampled as in step 5 of section 3.2

   (d) Sample from $\pi(\xi_t|\boldsymbol{\Psi})$ for $t = 1, \ldots, T$
6. To sample ξ_t from the full conditional posterior, write

$$\mathbf{u}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t + \mu_\xi \gamma = \xi_t \gamma + \xi_t^{1/2} \mathbf{A}^{-1} \mathbf{H}_t^{1/2} \epsilon_t.$$

   and replace $\boldsymbol{\Sigma}$ with $\mathbf{H}_t$ in (2). The full conditional posterior is thus generalized inverse Gaussian, $GIG(\lambda, \psi, \chi)$ with $\lambda = -(\nu + k)/2$, $\chi = q_t^2 + \nu$ and $\psi = p_t^2$ for $q_t^2 = \mathbf{u}_t' \mathbf{A}' \mathbf{H}_t^{-1} \mathbf{A} \mathbf{u}_t$ and $p_t^2 = \boldsymbol{\gamma}' \mathbf{A}' \mathbf{H}_t^{-1} \mathbf{A} \boldsymbol{\gamma}$.

A.7 VAR with MT distribution and constant variance

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{W}_t^{1/2} \mathbf{A}^{-1} \boldsymbol{\Sigma}^{1/2} \boldsymbol{\epsilon}_t$$

with $\mathbf{W}_t = \mathrm{diag}(\xi_{1t}, \ldots, \xi_{kt})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \mathrm{diag}(\tau_1^2, \ldots, \tau_k^2)$ and $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$

A.7.1 Prior

The priors for $\mathbf{B}$, ν_i and $\mathbf{A}$ are as described in section 3.1 and the prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$.

A.7.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = \mathbf{W}_t^{-1/2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2 + 1)/2)$ for $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{A}\widetilde{\mathbf{u}}_t$

4. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1) \prod_{t=1}^{T} \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_1/2)$ and is sampled as in step 5 of section 3.2

5. Sample ξ_{it} as in step 6 of Section 3.2, with

$$\pi(\mathbf{W}_t|\mathbf{\Psi}) \propto \prod_{i=1}^{k} \xi_{it}^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}_t^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)\right) \mathcal{IG}\left(\xi_{it}; \frac{\nu_i}{2}, \frac{\nu_i}{2}\right).$$

where $\mathbf{\Omega}_t = \mathbf{W}_t^{1/2}\mathbf{A}^{-1}\mathbf{\Sigma}\mathbf{A}^{-1\prime}\mathbf{W}_t^{1/2}$.

A.8 VAR with MT distribution and SV

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{W}_t^{1/2}\mathbf{A}^{-1}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$$

with $\mathbf{W}_t = \mathrm{diag}(\xi_{1t}, \ldots, \xi_{kt})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$ $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{kt})$ and the log volatilities evolving as

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, \ldots, k,$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.8.1 Prior

The priors for $\mathbf{B}$, ν_i, $\mathbf{A}$ and $\mathbf{h}_0$ are as described in section 3.1.

A.8.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.

2. Sample $\mathbf{A}$ from the full conditional normal posterior as in step 2 of section 3.2, setting $\widetilde{\mathbf{u}}_t = \mathbf{W}_t^{-1/2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)$ and $\mathbf{A}\widetilde{\mathbf{u}}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$

3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2

4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2

5. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1) \prod_{t=1}^{T} \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_1/2)$ and is sampled as in step 5 of section 3.2

6. Sample ξ_{it} as in step 6 of Section 3.2, with

$$\pi(\mathbf{W}_t|\mathbf{\Psi}) \propto \prod_{i=1}^{k} \xi_{it}^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}_t^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)\right) \mathcal{IG}\left(\xi_{it}; \frac{\nu_i}{2}, \frac{\nu_i}{2}\right).$$

where $\mathbf{\Omega}_t = \mathbf{W}_t^{1/2}\mathbf{A}^{-1}\mathbf{H}_t\mathbf{A}^{-1\prime}\mathbf{W}_t^{1/2}$.

A.9 VAR with MST distribution and constant variance

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + (\mathbf{W}_t - \overline{\mathbf{W}})\boldsymbol{\gamma} + \mathbf{W}_t^{1/2}\mathbf{A}^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$$

with $\mathbf{W}_t = \mathrm{diag}(\xi_{1t}, \dots, \xi_{kt})$, $\overline{\mathbf{W}} = \mathrm{diag}(\mu_{\xi,1}\dots, \mu_{\xi,k})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \mathrm{diag}(\tau_1^2, \dots, \tau_k^2)$ and $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$

A.9.1 Prior

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$, ν_i and $\mathbf{A}$ are as described in section 3.1 and the prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$.

A.9.2 MCMC

1. Sample $\mathbf{B}$ and $\boldsymbol{\gamma}$ jointly from the full conditional normal posterior as in step 1 of Section 3.2, replacing $\mathbf{H}_t$ with $\boldsymbol{\Sigma}$.
2. Sample $\mathbf{A}$ as in step 2 of Section 3.2, replacing $\mathbf{H}_t$ with $\boldsymbol{\Sigma}$.
3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2 + 1)/2)$ for $\tilde{\tilde{\mathbf{u}}}_t = \mathbf{A}\tilde{\mathbf{u}}_t$ and $\tilde{\mathbf{u}}_t = \mathbf{W}_t^{-1/2}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t - (\mathbf{W}_t - \overline{\mathbf{W}})\gamma)$
4. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^T \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_1/2)$ and is sampled as in step 5 of section 3.2
5. Sample ξ_{it} as in step 6 of Section 3.2, replacing $\mathbf{H}_t$ with $\boldsymbol{\Sigma}$ in the expression for $\boldsymbol{\Omega}_t$

A.10 VAR with OT distribution and constant variance

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$$

with $\mathbf{W}_t = \mathrm{diag}(\xi_{1t}, \dots, \xi_{kt})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \mathrm{diag}(\tau_1^2, \dots, \tau_k^2)$ and $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$

A.10.1 Prior

The priors for $\mathbf{B}$, ν_i and $\mathbf{A}$ are as described in section 3.1 and the prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$.

A.10.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.
2. Let $\mathbf{u}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t$ we then have a triangular equation system

$$\mathbf{A}\mathbf{u}_t = \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$$

   with $\xi_{it}\tau_i^2$ for $i = 2, \dots, k$. Rows $2, \dots, k$ in $\mathbf{A}$ are then sampled from the full conditional posteriors as in step 2 of Section 3.2.

3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2/\xi_{it} + 1)/2)$ for $\tilde{\tilde{\mathbf{u}}}_t = \mathbf{A}\mathbf{u}_t$

4. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1) \prod_{t=1}^T \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_i/2)$ and is sampled as in step 4 of section 3.2

5. Sample ξ_{it} from the full conditional inverse Gamma distribution, $\mathcal{IG}((\nu_i + \tilde{\tilde{u}}_{it}^2/\tau_i)/2, (\nu_i+1)/2)$ for $\tilde{\tilde{\mathbf{u}}}_t = \mathbf{A}\mathbf{u}_t$

A.11 VAR with OT distribution and SV

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\mathbf{W}_t^{1/2}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t.$$

with $\mathbf{W}_t = \text{diag}(\xi_{1t}, \ldots, \xi_{kt})$, $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\mathbf{H}_t = \text{diag}(h_{1t}, \ldots, h_{kt})$ and the log volatilities evolving as

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, \ldots, k,$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.11.1 Prior distribution

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$, ν_i, $\mathbf{A}$, $\mathbf{h}_0$ and $\boldsymbol{\sigma}^2$ are as described in section 3.1.

A.11.2 MCMC

1. Sample $\mathbf{B}$ from the full conditional normal posterior.

2. Sample $\mathbf{A}$ as in step 2 of Section A.10.2 with $\xi_{it} h_{it}$ as the known error variance of equations $i = 2, \ldots, k$ of the triangular equation system.

3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2 with $\tilde{\tilde{\mathbf{u}}}_t = \mathbf{W}_t^{-1/2}\mathbf{A}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)$

4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2

5. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1) \prod_{t=1}^T \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_1/2)$ and is sampled as in step 5 of section 3.2

6. Sample ξ_{it} from the full conditional inverse Gamma distribution, $\mathcal{IG}((\nu_i + \tilde{\tilde{u}}_{it}^2/h_{it})/2, (\nu_i+1)/2)$ for $\tilde{\tilde{\mathbf{u}}}_t = \mathbf{A}\mathbf{u}_t$

A.12 VAR with OST distribution and constant variance

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\left((\mathbf{W}_t - \overline{\mathbf{W}})\boldsymbol{\gamma} + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t\right).$$

with $\mathbf{W}_t = \text{diag}(\xi_{1t}, \ldots, \xi_{kt})$, $\overline{\mathbf{W}} = \text{diag}(\mu_{\xi,1} \ldots, \mu_{\xi,k})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\boldsymbol{\Sigma} = \text{diag}(\tau_1^2, \ldots, \tau_k^2)$ and $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$

A.12.1 Prior

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$, ν_i and $\mathbf{A}$ are as described in section 3.1 and the prior for τ_i^2 is inverse Gamma, $\mathcal{IG}(1/2, 1/2)$.

A.12.2 MCMC

1. Sample $\mathbf{B}$ and $\boldsymbol{\gamma}$ jointly as $\boldsymbol{b} = (\mathrm{vec}(\mathbf{B})', \boldsymbol{\gamma}')'$. Rewrite the model as
$$\begin{aligned}\mathbf{A}\mathbf{y}_t =& \mathbf{A}\mathbf{B}\mathbf{x}_t + (\mathbf{W}_t - \overline{\mathbf{W}})\gamma + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t \\ =& \left(\mathbf{x}_t' \otimes \mathbf{A} \quad (\mathbf{W}_t - \overline{\mathbf{W}})\right)\begin{pmatrix}\mathrm{vec}(\mathbf{B}) \\ \gamma\end{pmatrix} + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t \\ \widetilde{\mathbf{y}}_t =& \widetilde{\mathbf{X}}_t\boldsymbol{b} + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t.\end{aligned} \tag{3}$$
a heteroskedastic multivariate regression model with error variance-covariance $\mathrm{diag}(\xi_{1t}\tau_1^2, \ldots, \xi_{kt}\tau_k^2)$ where $\widetilde{\mathbf{y}}_t = \mathbf{A}\mathbf{y}_t$ and $\widetilde{\mathbf{X}}_t = \left(\mathbf{x}_t' \otimes \mathbf{A} \quad (\mathbf{W}_t - \overline{\mathbf{W}})\right)$. $\boldsymbol{b}$ is then sampled from the full conditional normal posterior.

2. Let $\mathbf{u}_t = \mathbf{y}_t - \mathbf{B}\mathbf{x}_t$ we then have a triangular equation system
$$\mathbf{A}\mathbf{u}_t = (\mathbf{W}_t - \overline{\mathbf{W}})\gamma + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$$
with $\tilde{u}_{it} = u_{it} - (\xi_{it} - \mu_{\xi,i})\gamma_i$ as dependent variable, $(-u_{1,t}, \ldots, -u_{i-1,t})$ as explanatory variables and know variance $\xi_{it}\tau_i^2$ for $i = 2, \ldots, k$. Rows $2, \ldots, k$ in $\mathbf{A}$ are then sampled from the full conditional posteriors.

3. Sample τ_i^2 from the full conditional inverse Gamma posterior, $\mathcal{IG}((T+1)/2, (\sum_{t=1}^T \tilde{\tilde{u}}_{it}^2/\xi_{it} + 1)/2)$ for $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{A}\mathbf{u}_t - (\mathbf{W}_t - \overline{\mathbf{W}})\boldsymbol{\gamma}$

4. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^T \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_i/2)$ and is sampled as in step 4 of section 3.2

5. To sample from the full conditional distribution for ξ_{it} write
$$\mathbf{u}_t = \mathbf{A}\left(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t\right) + \overline{\mathbf{W}}\boldsymbol{\gamma} = \mathbf{W}_t\boldsymbol{\gamma} + \mathbf{W}_t^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t.$$
That is, the conditional distribution of $\mathbf{u}_t$ is normal, $\mathbf{u}_t \sim \mathcal{N}(\mathbf{W}_t\boldsymbol{\gamma}, \mathbf{W}_t\boldsymbol{\Sigma})$ and the likelihood contribution is
$$|\mathbf{W}_t| \exp\left(-\frac{1}{2}(\mathbf{u}_t - \mathbf{W}_t\gamma)'\boldsymbol{\Sigma}^{-1}\mathbf{W}_t^{-1}(\mathbf{u}_t - \mathbf{W}_t\gamma)\right) \tag{4}$$
with
$$(\mathbf{u}_t - \mathbf{W}_t\gamma)'\boldsymbol{\Sigma}^{-1}\mathbf{W}_t^{-1}(\mathbf{u}_t - \mathbf{W}_t\gamma) = \sum_{i=1}^k \frac{u_{it}^2/\tau_i^2}{\xi_{it}} - 2\sum_{i=1}^k \frac{u_{it}\gamma_i}{\tau_i^2} + \sum_{i=1}^k \frac{\gamma_i^2}{\tau_i^2}\xi_{it}.$$
With the independent inverse Gamma distribution for ξ_{it} we have conditional independence and the full conditional posterior is generalized inverse Gaussian (GIG),
$$\begin{aligned}\pi(\xi_{it}|\boldsymbol{\Psi}) &\propto \xi_{it}^{-1/2}\exp\left\{-\frac{1}{2}\left(\frac{q_{it}^2}{\xi_{it}} + \xi_{it}p_{it}^2\right)\right\}\xi_{it}^{-(\nu_i/2+1)}\exp\left\{-\frac{\nu_i}{2\xi_{it}}\right\} \\ &= \xi_{it}^{\lambda-1}\exp\left\{-\frac{1}{2}\left(\frac{\chi}{\xi_{it}} + \psi\xi_{it}\right)\right\}\end{aligned}$$

for $q_{it}^2 = u_{it}^2/\tau_i^2$, $p_{it}^2 = \gamma_i^2/\tau_i^2$, $\lambda = -(\nu_i+1)/2$, $\chi = q_{it}^2 + \nu_i$ and $\psi = p_{it}^2$. That is a $GIG(\lambda, \psi, \chi)$ distribution.

A.13 VAR with OST distribution and SV

The model is

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \mathbf{A}^{-1}\left((\mathbf{W}_t - \overline{\mathbf{W}})\boldsymbol{\gamma} + \mathbf{W}_t^{1/2}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t\right).$$

with $\mathbf{W}_t = \mathrm{diag}(\xi_{1t}, \dots, \xi_{kt})$, $\xi_{it} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$, $\overline{\mathbf{W}} = \mathrm{diag}(\mu_{\xi,1} \dots, \mu_{\xi,k})$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I})$, $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \dots, h_{kt})$ and the log volatilities evolving as

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, \dots, k,$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$.

A.13.1 Prior distribution

The priors for $\mathbf{B}$, $\boldsymbol{\gamma}$, ν_i, $\mathbf{A}$, $\mathbf{h}_0$ and $\boldsymbol{\sigma}^2$ are as described in section 3.1.

A.13.2 MCMC

1. Sample Sample $\mathbf{B}$ and $\boldsymbol{\gamma}$ jointly from the full conditional normal posterior. To this end, rewrite the model as in (3) to yield $\widetilde{\mathbf{y}}_t = \widetilde{\mathbf{X}}_t \boldsymbol{b} + \mathbf{W}_t^{1/2}\mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$, a heteroskedastic multivariate regression.

2. Sample $\mathbf{A}$ as in step 2 of Section A.12.2 with $\xi_{it} h_{it}^2$ as the known error variance of equations $i = 2, \dots, k$ of the triangular equation system.

3. Sample $\mathbf{h}_{0:T}$ as in step 3 of Section 3.2 with $\widetilde{\widetilde{\mathbf{u}}}_t = \mathbf{W}_t^{-1/2}\left[\mathbf{A}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t) - (\mathbf{W}_t - \overline{\mathbf{W}})\boldsymbol{\gamma}\right]$

4. Sample $\boldsymbol{\sigma}^2$ as in step 4 of Section 3.2

5. The full conditional posterior for ν_i is proportional to $\mathcal{G}(2, 0.1)\prod_{t=1}^T \mathcal{IG}(\xi_{it}, \nu_i/2, \nu_1/2)$ and is sampled as in step 5 of section 3.2

6. Sample ξ_{it} as in step 5 of Section A.12.2 with $q_{it}^2 = u_{it}^2/h_{it}$ and $p_{it}^2 = \gamma_i^2/h_{it}$

B Posterior comparison

Figures 1 and 2 compares the posterior distribution of the skewness, γ_i, and heavy tail, ν_i parameters for the MST and OST specifications with and without stochastic volatilities.

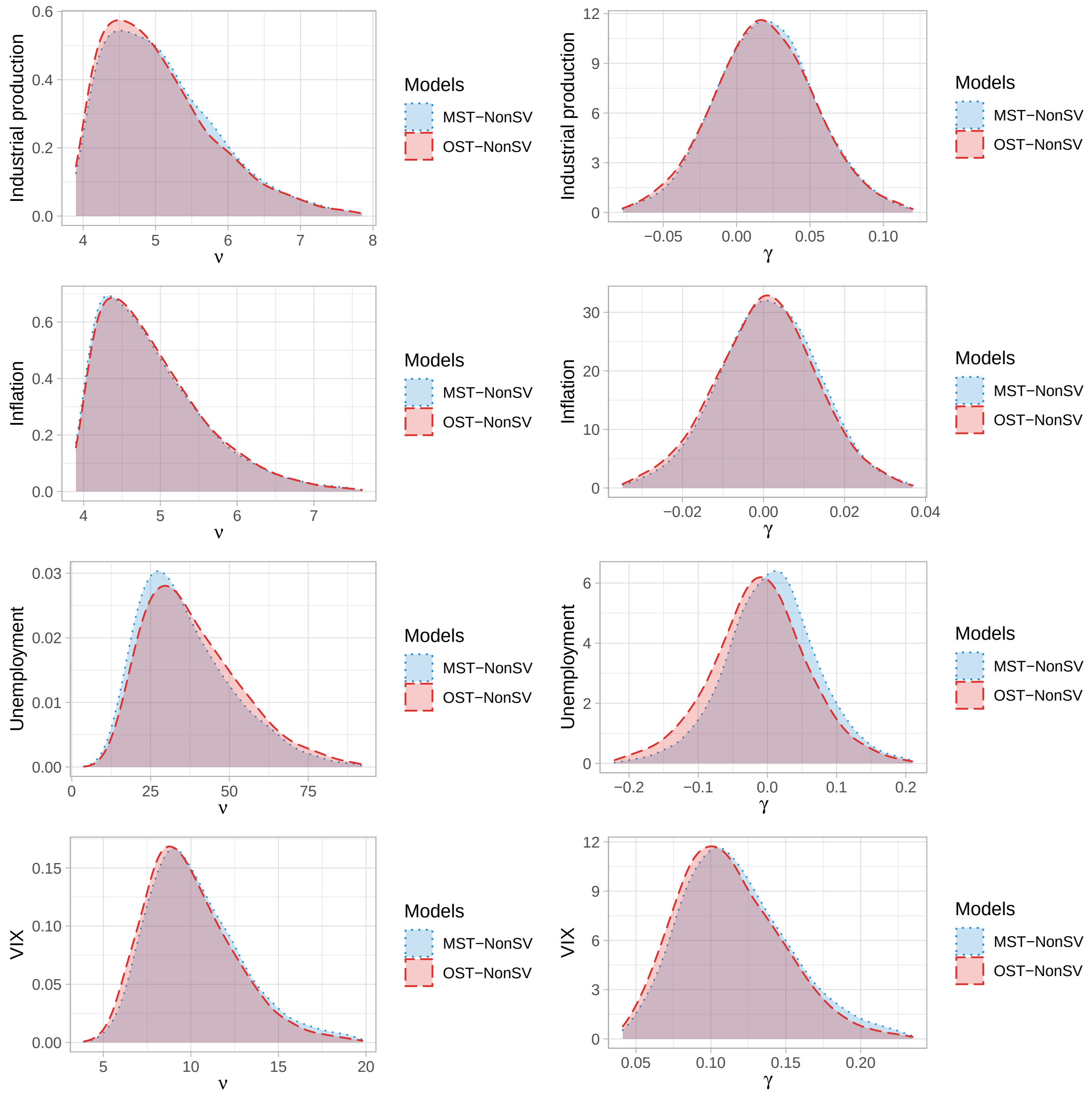

Figure 1: The plots show the posterior samples of the heavy tail and skewness parameters of the VAR models with MST and OST innovations.

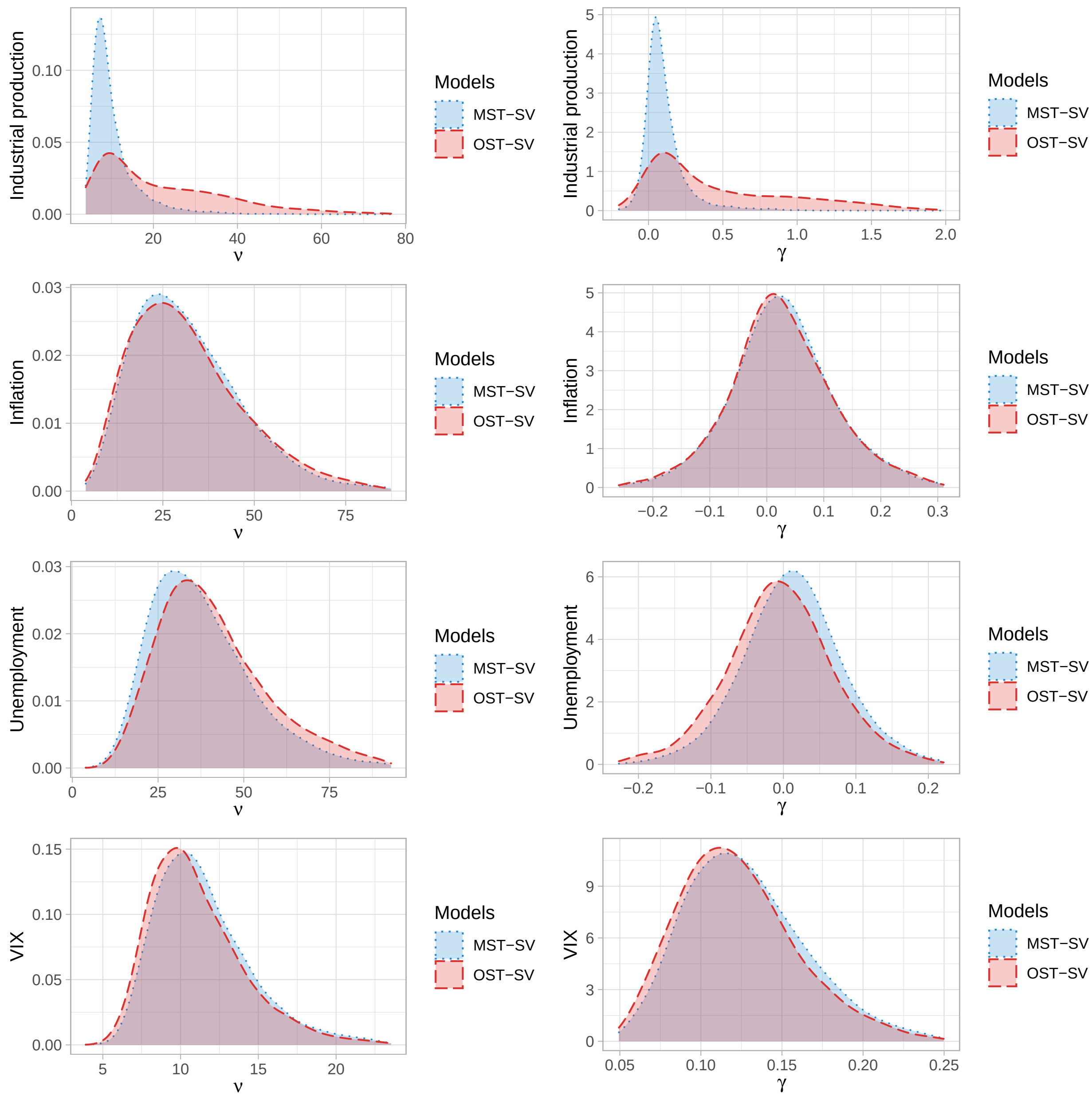

Figure 2: The plots show the posterior samples of the heavy tail and skewness parameters of the the VAR models with MST-SV and OST-SV innovations.

C Sensitivity to the order of variables

Figures 3 and 4 compare the posterior distribution of the skewness (γ_i) and heavy tail (ν_i) for two orderings of the variables for the MST and OST specifications where the results may be sensitive to the ordering. The orderings are the original order used in the paper; Industrial production, Inflation, Unemployment, VIX, and the alternative ordering; Inflation, VIX, Industrial production, Unemployment. The posterior distribution is largely unaffected by the ordering.

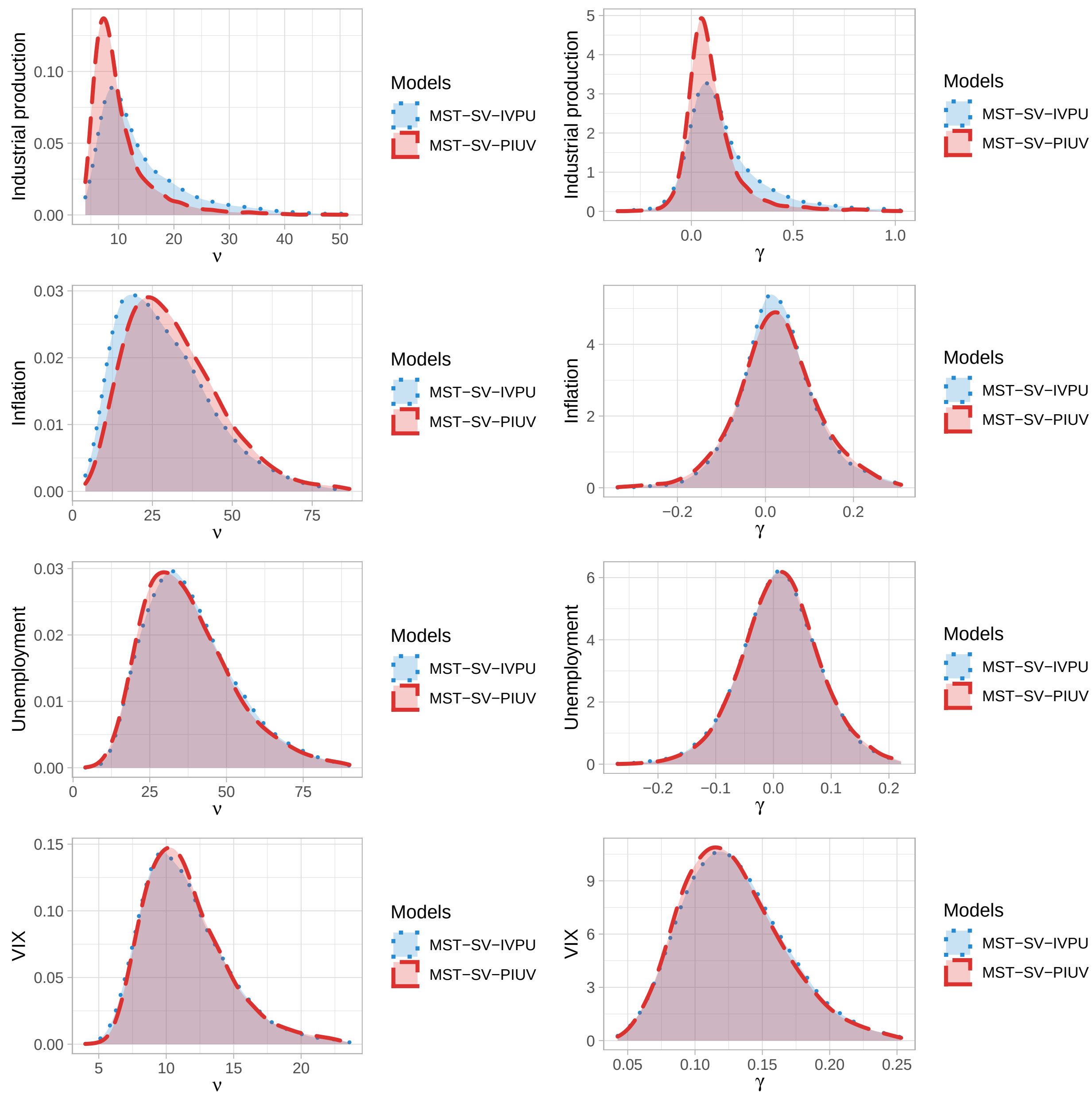

Figure 3: Posterior distribution of the heavy tail (left column) and skewness (right column) parameters of the VAR models with MST-SV. The order PIUV stands for Industrial production, Inflation, Unemployment, VIX and The order IVPU stands for Inflation, VIX, Industrial production, Unemployment.

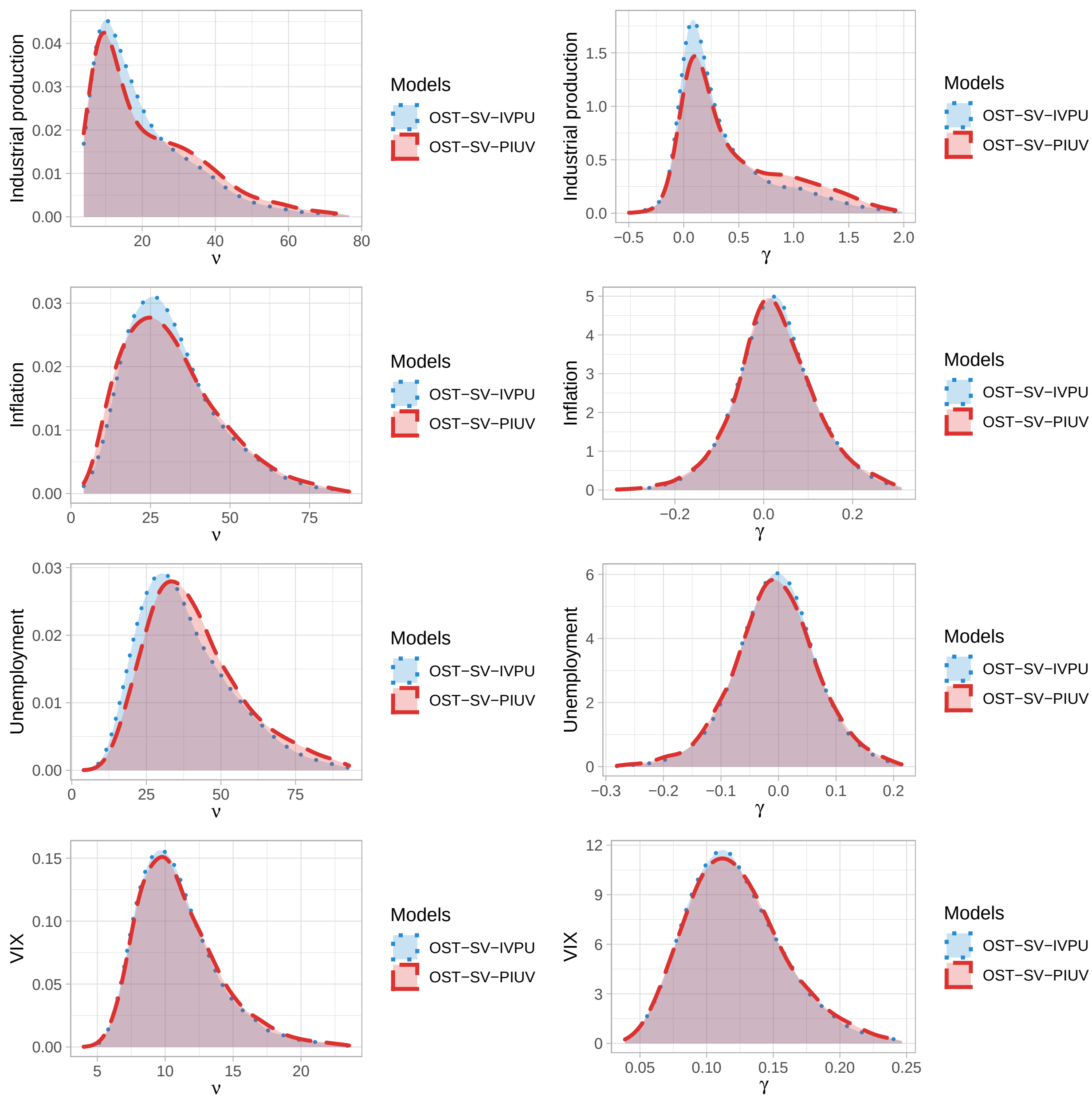

Figure 4: Posterior distribution of the heavy tail (left column) and skewness (right column) parameters of the VAR models with OST-SV. The order PIUV stands for Industrial production, Inflation, Unemployment, VIX and The order IVPU stands for Inflation, VIX, Industrial production, Unemployment.