Supplementary material for "Bayesian SAR model with stochastic volatility and multiple time-varying weights"

Michele Costola* Matteo Iacopini† Casper Wichers‡

December 29, 2024

Abstract

This supplementary material includes the detailed derivation of the posterior distributions in Section A. The details of the real data investigated in the empirical applications are presented in Section B. Section C reports additional results from the application. Section D presents the model selection criteria. Section E describes the settings in the simulation study. Finally, Section F analyses the residual correlation of the SAR-SV model.

A Posterior full conditional distributions

The conditional distribution for $\boldsymbol{\beta}$ is standard and fast algorithms to sample from it are used. In high-dimensional set-ups with numerous covariates, the introduction of global-local shrinkage priors for $\boldsymbol{\beta}$ based on scale mixtures of Gaussian distributions (e.g., see Bhadra et al., 2016) does not affect significantly the computational cost, as only draws from standard distributions are required.

The main computational challenge concerns sampling the layer-specific weights, $\boldsymbol{\delta}$, and the network weights, ρ_j, since there is no distribution conjugate with the likelihood in eq. (13). Adopting a Bayesian approach, we introduce the restrictions on the parameters $\boldsymbol{\delta}$ and $\boldsymbol{\rho}$ discussed in Section 3 through an appropriate choice of the prior distributions in eq. (14), which leads to non-standard posterior full conditional distributions.

Therefore, we sample each ρ_j from its full conditional distribution using a slice sampler algorithm (Neal, 2003). This approach allows us to sample from an arbitrary univariate distribution (known up to a proportionality constant) via introducing an auxiliary (slice) variable. Specifically, denoting by f the full conditional distribution of ρ_j and $\rho_j^{(m)}$ its

*Ca' Foscari University of Venice, Italy. michele.costola@unive.it
†LUISS University, Italy. miacopini@luiss.it
‡ADC, Data & AI Consultancy, The Netherlands. casperwichers@gmail.com

value at the mth iteration, we obtain a draw of $\rho_j^{(m+1)}$ as

$$u_j^{(m)}|\rho_j^{(m)} \sim \mathcal{U}\big(0, f(\rho_j^{(m)})\big), \qquad \rho_j^{(m+1)}|u_j^{(m)} \sim \mathcal{U}\big(\mathbb{A}^{(m+1)}\big). \tag{1}$$

where $\mathbb{A}^{(m+1)} = \big\{x : f(x) \geq u_j^{(m)}\big\}$. By sampling from univariate distributions, this approach does not suffer from the tuning issue and possible low acceptance rate of traditional Metropolis-Hastings techniques. Also, any draw from this step belongs to the interval $(-1, 1)$ by the (transformed) beta prior in eq. (14) incorporating this restriction.

For the layer-specific weights in the vector $\boldsymbol{\delta}$, we ensure that A3 is satisfied by assuming a Dirichlet prior in eq. (14), then jointly sampling the vector. Specifically, we design an independent Metropolis-Hastings (iMH) algorithm with a Dirichlet proposal distribution and use a preliminary short run of the MCMC algorithm to tune the hyperparameter.[1]

Below we report the detailed derivation of the posterior full conditional distribution for each parameter.

Posterior full conditional distribution of $\boldsymbol{\beta}$ Recall that $\boldsymbol{\beta} = (\boldsymbol{\alpha}_0', \text{vec}(B)')'$, $X_t = (\mathbf{1}_n, \mathbf{f}_t' \otimes I_n)$, $\Sigma_t = \text{diag}(\exp(h_{1,t}), \dots, \exp(h_{n,t}))$, and $\mathbf{y}_t^* = A_t \mathbf{y}_t$. Then the posterior full conditional distribution of $\boldsymbol{\beta}$ is obtained as:

$$\begin{aligned} P(\boldsymbol{\beta}|\mathbf{y}, \mathbf{h}, \boldsymbol{\rho}, \boldsymbol{\delta}) &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \underline{\boldsymbol{\mu}}_\beta)' \underline{\Sigma}_\beta^{-1} (\boldsymbol{\beta} - \underline{\boldsymbol{\mu}}_\beta)\right\} \\ &\quad \times \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} (\mathbf{y}_t^* - X_t\boldsymbol{\beta})' \Sigma_t^{-1} (\mathbf{y}_t^* - X_t\boldsymbol{\beta})\right\} \\ &\propto \exp\left\{-\frac{1}{2}\Big[\boldsymbol{\beta}' \underline{\Sigma}_\beta^{-1} \boldsymbol{\beta} - 2\boldsymbol{\beta}' \underline{\Sigma}_\beta^{-1} \underline{\boldsymbol{\mu}}_\beta + \sum_{t=1}^{T} \boldsymbol{\beta}' X_t' \Sigma_t^{-1} X_t \boldsymbol{\beta} - 2\boldsymbol{\beta}' X_t' \Sigma_t^{-1} \mathbf{y}_t^*\Big]\right\}, \end{aligned}$$

which is proportional to the Gaussian distribution $\mathcal{N}_{k_\beta}(\boldsymbol{\beta}|\overline{\boldsymbol{\mu}}_\beta, \overline{\Sigma}_\beta)$ with

$$\overline{\Sigma}_\beta = \Big(\underline{\Sigma}_\beta^{-1} + \sum_{t=1}^{T} X_t' \Sigma_t^{-1} X_t\Big)^{-1} \qquad \overline{\boldsymbol{\mu}}_\beta = \overline{\Sigma}_\beta \Big(\underline{\Sigma}_\beta^{-1} \underline{\boldsymbol{\mu}}_\beta + \sum_{t=1}^{T} X_t' \Sigma_t^{-1} \mathbf{y}_t^*\Big).$$

Posterior full conditional distribution of SV hyperparameters The hyperparameters driving the stochastic volatility processes, $(\mu_{h,j}, \phi_{h,j}, \sigma_{h,j}^2)$, for each $j = 1, \dots, n$, are sampled from their full conditional distributions using the `stochvol` package in `R` (Hosszejni and Kastner, 2019).

Posterior full conditional distribution of $\mathbf{h}_j$ The path of the stochastic volatility $\mathbf{h}_j = (h_{j,1}, \dots, h_{j,T})$, for each $j = 1, \dots, n$, is sampled using a forward-filtering backward-

[1] The proposed method works well in both simulation studies and the real-data application. As the number of network layers d increases, the acceptance rate of the iMH algorithm tends to decay, suggesting the use of entry-wise updates (see Debarsy and LeSage, 2022, for a possible approach).

sampling approach using the `stochvol` package in `R` (Hosszejni and Kastner, 2019).

Posterior full conditional distribution of $\boldsymbol{\delta}$ For the layer-specific weights $\boldsymbol{\delta}$, combining the Dirichlet prior distribution with the multivariate normal likelihood in eq. (14) yields:

$$P(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\beta}, \mathbf{h}, \boldsymbol{\rho}) \propto \left(\prod_{i=1}^{d} \delta_i^{c_i - 1} \right) \times \left(\prod_{t=1}^{T} |A_t| \exp\left\{ -\frac{1}{2} \mathbf{e}_t' \Sigma_t^{-1} \mathbf{e}_t \right\} \right),$$

where $A_t = I_n - R(\sum_{i=1}^{d} \delta_i W_{i,t})$ and $\mathbf{e}_t = A_t \mathbf{y}_t - X_t \boldsymbol{\beta}$ depend on $\boldsymbol{\delta}$. We obtain draws from this distribution by using an independent Metropolis-Hastings algorithm with a Dirichlet proposal.

Posterior full conditional of ρ_j The posterior full conditional distribution of the country-specific network weight, ρ_j, for each $j = 1, \ldots, n$, is:

$$P(\rho_j|\mathbf{y}, \boldsymbol{\rho}_{-j}, \boldsymbol{\beta}, \mathbf{h}, \boldsymbol{\delta}) \propto \left(\frac{\rho_j + 1}{2}\right)^{\underline{a}_\rho - 1} \left(1 - \frac{\rho_j + 1}{2}\right)^{\underline{b}_\rho - 1} \times \prod_{t=1}^{T} |A_t| \exp\left\{ -\frac{1}{2} \mathbf{e}_t' \Sigma_t^{-1} \mathbf{e}_t \right\}.$$

where $A_t = I_n - R(\sum_{i=1}^{d} \delta_i W_{i,t})$ and $\mathbf{e}_t = A_t \mathbf{y}_t - X_t \boldsymbol{\beta}$ depend on $\boldsymbol{\rho}$. We obtain draws from this distribution by using a slice sampler algorithm (Neal, 2003).

A.1 SAR-SV model with time-varying $\rho_{j,t}$

To introduce time-variation and temporal dependence in the country-specific network weights, $\rho_{j,t}$, we consider the following prior structure

$$\frac{\rho_{j,t} + 1}{2} \Big| \zeta_{j,t-1} \sim \mathcal{B}e\left(\frac{\rho_{j,t} + 1}{2} \,\Big|\, \underline{a}_\rho + \zeta_{j,t-1},\ \nu + \eta - \zeta_{j,t-1} \right) \tag{2}$$

$$\zeta_{j,t-1} | \rho_{j,t-1} \sim \mathcal{B}in\left(\zeta_{j,t-1} \,\big|\, \eta,\ \frac{\rho_{jt-1} + 1}{2} \right) \tag{3}$$

$$\frac{\rho_{j,1} + 1}{2} \sim \mathcal{B}e\left(\frac{\rho_{j,1,} + 1}{2} \,\Big|\, \underline{a}_\rho, \nu \right). \tag{4}$$

This choice is an adaptation of the proposal by Nieto-Barajas et al. (2012) and has the advantage of constraining $\rho_{j,t} \in (-1, 1)$ at any period $t = 1, \ldots, T$, since marginally $\frac{\rho_{j,t}+1}{2} \sim \mathcal{B}e(\underline{a}_\rho, \nu)$. Conversely, Hauzenberger and Pfarrhofer (2021) used an unrestricted model with $\rho_t \in \mathbb{R}$ and assumed a Gaussian random walk process to govern the temporal evolution of this parameter. This choice allows for a minor computational advantage but still calls for the use of a Metropolis-Hastings step to sample ρ_t at each period t. Conversely, we adopted the more conventional approach and imposed the constraint $\rho_{j,t} \in (-1, 1)$. The main difference is that our specification grants the invertibility of the

matrix A_t, which is instead not guaranteed by allowing $|\rho_{j,t}| \geq 1$ as in Hauzenberger and Pfarrhofer (2021).

Posterior full conditional distribution of $\rho_{j,t}$

$$
\begin{aligned}
P\Big(\frac{\rho_{j,t}+1}{2}|\zeta_{j,t},\zeta_{j,t-1},\mathbf{y}_t,\boldsymbol{\beta},\mathbf{h}_t,\boldsymbol{\rho}_{-j,t},\eta,\nu\Big) &\propto \left(\frac{\rho_{j,t}+1}{2}\right)^{\zeta_{j,t-1}}\left(1-\frac{\rho_{j,t}+1}{2}\right)^{\nu+\eta-\zeta_{j,t-1}-1} \\
&\times \left(\frac{\rho_{j,t}+1}{2}\right)^{\zeta_{j,t}}\left(1-\frac{\rho_{j,t}+1}{2}\right)^{\eta-\zeta_{j,t}} \times |A_t|\exp\Big\{-\frac{1}{2}\mathbf{e}_t'\Sigma_t^{-1}\mathbf{e}_t\Big\} \\
&\propto \left(\frac{\rho_{j,t}+1}{2}\right)^{\zeta_{j,t}+\zeta_{j,t-1}+\underline{a}_\rho-1}\left(1-\frac{\rho_{j,t}+1}{2}\right)^{\nu+2\eta-\zeta_{j,t}-\zeta_{j,t-1}-1} |A_t|\exp\Big\{-\frac{1}{2}\mathbf{e}_t'\Sigma_t^{-1}\mathbf{e}_t\Big\} \\
P\Big(\frac{\rho_{j,1}+1}{2}|\zeta_{j,1},\mathbf{y}_1,\boldsymbol{\beta},\mathbf{h}_1,\boldsymbol{\rho}_{-j,1},\eta,\nu\Big) &\propto \\
&\propto \left(\frac{\rho_{j,1}+1}{2}\right)^{\zeta_{j,1}+\underline{a}_\rho-1}\left(1-\frac{\rho_{j,1}+1}{2}\right)^{\nu+\eta-\zeta_{j,1}-1} |A_1|\exp\Big\{-\frac{1}{2}\mathbf{e}_1'\Sigma_1^{-1}\mathbf{e}_1\Big\} \\
P\Big(\frac{\rho_{j,T}+1}{2}|\zeta_{j,T-1},\mathbf{y}_T,\boldsymbol{\beta},\mathbf{h}_T,\boldsymbol{\rho}_{-j,T},\eta,\nu\Big) &\propto \\
&\propto \left(\frac{\rho_{j,T}+1}{2}\right)^{\zeta_{j,T-1}+\underline{a}_\rho-1}\left(1-\frac{\rho_{j,T}+1}{2}\right)^{\nu+\eta-\zeta_{j,T-1}-1} |A_T|\exp\Big\{-\frac{1}{2}\mathbf{e}_T'\Sigma_T^{-1}\mathbf{e}_T\Big\},
\end{aligned}
$$

where $A_t = I_n - R_t(\sum_{i=1}^d \delta_i W_{i,t})$, with $R_t = \text{diag}(\rho_{1,t},\dots,\rho_{N,t})$, and $\mathbf{e}_t = A_t\mathbf{y}_t - X_t\boldsymbol{\beta}$. Notice that, for every $t = 1,2,\dots,T$, the posterior full conditional distribution is of the same type as the one in the constant ρ_j case. Therefore, we can exploit the same slice sampler approach to obtain draws of $\rho_{j,t}$ for each t.

Posterior full conditional distribution of $\zeta_{j,t}$

$$
\begin{aligned}
P(\zeta_{j,t}|\rho_{j,t},\rho_{j,t-1},\eta,\nu) &\propto \frac{\eta!}{\zeta_{j,t}!(\eta-\zeta_{j,t})!}\left(\frac{\rho_{j,t}+1}{2}\right)^{\zeta_{j,t}}\left(1-\frac{\rho_{j,t}+1}{2}\right)^{\eta-\zeta_{j,t}} \\
&\times \frac{\Gamma(\nu+\eta+1)}{\Gamma(\zeta_{j,t}+1)\Gamma(\nu+\eta-\zeta_t)}\left(\frac{\rho_{j,t+1}+1}{2}\right)^{\zeta_{j,t}-1}\left(1-\frac{\rho_{j,t+1}+1}{2}\right)^{\nu+\eta-\zeta_{j,t}-1} \\
&\propto \frac{1}{(\Gamma(\zeta_{j,t}+1))^2\Gamma(\eta-\zeta_{j,t}+1)\Gamma(\nu+\eta-\zeta_{j,t})}\left(\frac{(\rho_{j,t}+1)(\rho_{j,t+1}+1)}{(1-\rho_{j,t})(1-\rho_{j,t+1})}\right)^{\zeta_{j,t}},
\end{aligned}
$$

which can be evaluated by computing the (unnormalised) mass function for each $\zeta_{j,t} = 0,1,\dots,\eta$ and rescaling. Then, $\zeta_{j,t}$ is directly sampled from the resulting discrete distribution.

Posterior full conditional distribution of η

$$P(\eta|\boldsymbol{\rho},\boldsymbol{\zeta}) \propto \prod_{t=1}^{T-1}\prod_{j=1}^{n} \frac{\Gamma(\eta+1)\Gamma(\underline{a}_\rho+\nu+\eta)}{\Gamma(\eta-\zeta_{j,t}+1)\Gamma(\nu+\eta-\zeta_{j,t})} \times \left[\prod_{t=1}^{T-1}\prod_{j=1}^{n}\left(1-\frac{\rho_{j,t}+1}{2}\right)\left(1-\frac{\rho_{j,t+1}+1}{2}\right)\right]^{\eta} \times \text{TPoi}(\eta|\underline{\lambda},\eta^*),$$

for $\eta = \zeta^*,\ldots,\eta^*$ and $\zeta^* = \max_{(t,j)\in\{1,\ldots,T-1\}\times\{1,\ldots,n\}}\zeta_{j,t}$. We evaluate the function for each admissible η, then renormalise and sample from the resulting discrete distribution. The truncated (from above) Poisson distribution $\text{TPoi}(\eta|\underline{\lambda},\eta^*)$ has the probability mass function

$$P(\eta|\underline{\lambda},\eta^*) = \frac{\underline{\lambda}^{\eta}\exp(-\underline{\lambda})}{\eta!} \times \left(\frac{\Gamma(\lfloor\eta^*+1\rfloor,\underline{\lambda})}{\lfloor\eta^*\rfloor!}\right)^{-1}, \qquad \eta = 0,1,\ldots,\eta^*,$$

where $\Gamma(x,y)$ is the upper incomplete gamma function and $\lfloor x\rfloor$ is the floor function.

Posterior full conditional distribution of ν Assuming a Gamma prior for $\nu \sim \mathcal{G}a(\nu|1,1)$, where we use the shape-scale parameterization, we obtain the posterior full conditional distribution

$$P(\nu|\boldsymbol{\rho},\boldsymbol{\zeta}) \propto \nu^{n+\underline{a}_\nu-1}\exp(-\nu/\underline{b}_\nu) \times \left[\prod_{j=1}^{n}\left\{\left(1-\frac{\rho_{j,1}+1}{2}\right)\prod_{t=1}^{T-1}\left(1-\frac{\rho_{j,t+1}+1}{2}\right)\right\}\right]^{\nu}\prod_{j=1}^{n}\prod_{t=1}^{T}\frac{\Gamma(\underline{a}_\rho+\nu+\eta)}{\Gamma(\nu+\eta-\zeta_{j,t})}.$$

We use a random walk Metropolis-Hastings (RWMH) step to sample from this distribution using a gamma proposal $\nu^* \sim \mathcal{G}a(\nu^*|1,\nu^{(s)})$, where $\nu^{(s)}$ is the value of ν at the current iteration s.

Posterior full conditional distribution of $\boldsymbol{\delta}$, $\boldsymbol{\beta}$, $\mathbf{h}_j$, the SV hyperparameters For all these parameters, the posterior full conditional distributions are the same as in the constant ρ_j case, except that the matrix R in the definition of A_t is now replaced by $R_t = \text{diag}(\rho_{1,t},\ldots,\rho_{N,t})$, that is $A_t = I_n - R_t(\sum_{i=1}^{d}\delta_i W_{i,t})$.

A.2 Model with constant volatility

In the special case of a constant volatility model, we assume an improper prior on the country-specific log-variance, which results in a flat prior on σ_j^2 (see also Dittrich et al., 2017), $P(\sigma_j^2) \propto \sigma_j^{-2}$. For adequately large samples and given the functional form of the likelihood, the full conditional posterior is proper, and this prior is not influential on

posterior estimates. The set of all prior distributions for the parameters of the model is:

$$\begin{aligned} \boldsymbol{\beta} &\sim \mathcal{N}_{k_\beta}(\boldsymbol{\beta}|\underline{\boldsymbol{\mu}}_\beta, \underline{\Sigma}_\beta), & \sigma_j^2 &\sim \mathcal{IG}(\sigma_j^2|\underline{a}_\sigma, \underline{b}_\sigma) \\ \boldsymbol{\delta} &\sim \mathcal{D}ir(\boldsymbol{\delta}|\underline{\mathbf{c}}), & \frac{\rho_j+1}{2} &\sim \mathcal{B}e\Big(\frac{\rho_j+1}{2}\Big|\underline{a}_\rho, \underline{b}_\rho\Big). \end{aligned} \tag{5}$$

Let $\boldsymbol{\sigma}^2_{-j}$ denote all the elements of $\boldsymbol{\sigma}^2$, respectively, except the jth. The MCMC algorithm to sample from the posterior distributions is structured as:

1) sample $\boldsymbol{\beta}|\mathbf{y}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}, \boldsymbol{\delta}$ from a Gaussian distribution;
2) sample $\sigma_j^2|\mathbf{y}, \boldsymbol{\sigma}^2_{-j}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}$ from an inverse gamma distribution;
3) sample $\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}$ from $P(\boldsymbol{\delta}|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho})$ using an independent Metropolis-Hastings (iMH) algorithm;
4) sample $\rho_j|\mathbf{y}, \boldsymbol{\rho}_{-j}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\delta}$ from $P(\rho_j|\mathbf{y}, \boldsymbol{\rho}_{-j}, \boldsymbol{\beta}, \sigma_j^2, \boldsymbol{\delta})$ using a slice sampler algorithm (Neal, 2003).

The conditional distributions for $\boldsymbol{\beta}, \boldsymbol{\delta}, \rho_j$ are the same discussed for the SV model (with minor changes due to the use of σ_j^2 in place of $\mathbf{h}_j$).

The posterior full conditional distribution of the country-specific innovation variance, σ_j^2, for each $j = 1, \ldots, n$, is:

$$\begin{aligned} P(\sigma_j|\mathbf{y}, \boldsymbol{\sigma}^2_{-j}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}) &\propto (\sigma_j^2)^{-(1+\underline{a}_\sigma)} \exp(-\underline{b}_\sigma \sigma_j^2) \times \prod_{t=1}^{T} |\Sigma|^{-1/2} \exp\Big\{-\frac{1}{2}\mathbf{e}_t'\Sigma^{-1}\mathbf{e}_t\Big\} \\ &\propto \big(\sigma_j^2\big)^{-1-\underline{a}_\sigma - T/2} \exp\Big\{-\frac{1}{\sigma_j^2}\Big(\underline{b}_\sigma + \frac{1}{2}\sum_{t=1}^{T} e_{jt}^2\Big)\Big\}, \end{aligned}$$

which is proportional to the inverse Gamma distribution $\mathcal{IG}(\overline{a}, \overline{b})$ with parameters $\overline{a} = \underline{a}_\sigma + T/2$ and $\overline{b} = \underline{b}_\sigma + \sum_{t=1}^{T} e_{jt}^2/2$.

B Additional Data description

In the following, we present complement information for the application's control variables and stock indices.

B.1 Control variables

For returns, we consider the OECD Business Confidence Index (BCI) and the OECD Consumer Confidence Index (CCI) for the G7 area as control variables. The BCI offers insights into future trends by drawing from surveys on production, orders, and inventories of finished goods within the industrial sector. It can predict potential shifts in economic activity. Readings above 100 signify heightened confidence in upcoming business performance, while values below 100 indicate a more pessimistic outlook toward

future performance. The CCI provides information on how households will approach spending and saving in the future. It is based on their outlook on personal finances, the overall economy, employment, and their savings ability. A reading above 100 indicates increased confidence in the future economy, suggesting people are less inclined to save and more likely to make significant purchases next year. Conversely, values below 100 suggest a negative outlook on the economy's future, potentially leading to an inclination to increase savings and reduce consumption.[2] We incorporate changes in the BCI and CCI into the model as they measure shifts in expectations within both the business and consumer sectors, which might contribute to explaining fluctuations in market prices. Figure 1 presents the BCI and CCI on the level and on the first difference over the considered period.

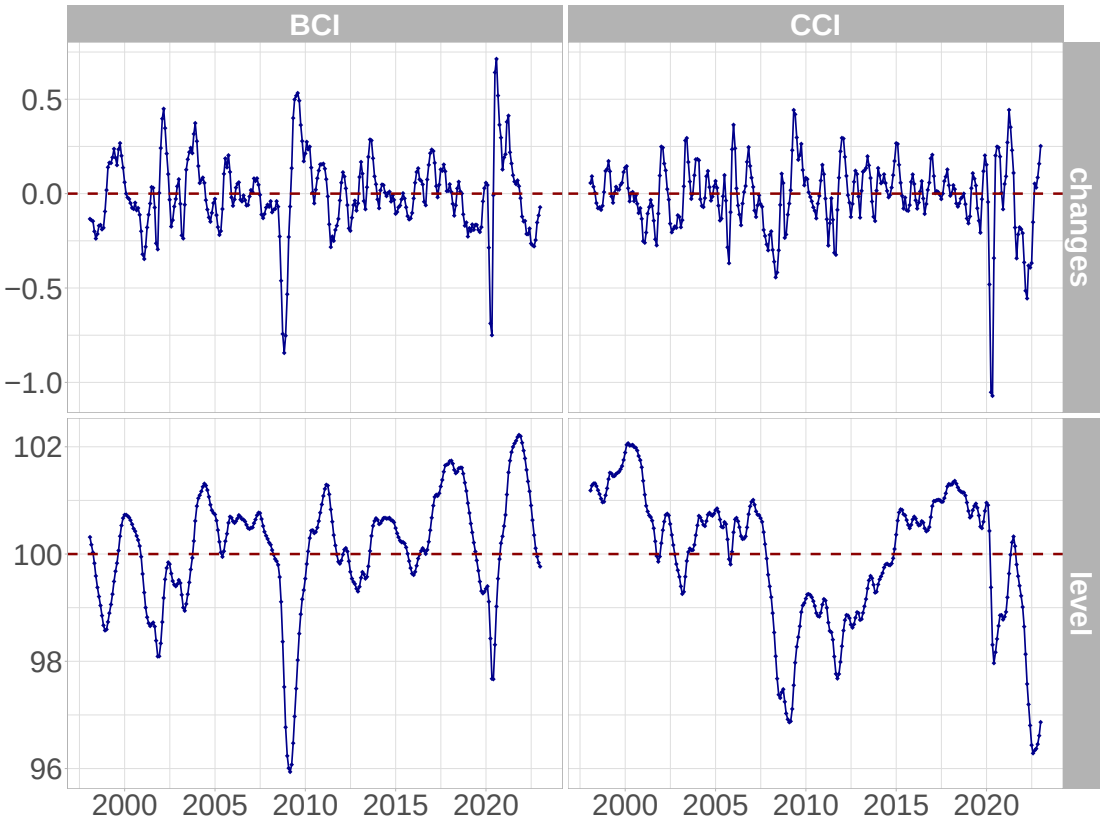

Figure 1: The OECD Business Confidence Index (BCI) and the OECD Consumer Confidence Index (CCI) for the G7 area over the considered period. The top panels show the indicators on the levels where readings above (below) 100 in the red dashed line imply optimistic (pessimistic) business/consumer confidence in future developments. The bottom panels show their first difference.

[2]OECD (2023), Consumer confidence index (CCI) (indicator). doi:10.1787/46434d78-en. OECD (2023), Business Confidence Index (BCI) (indicator). doi:10.1787/3092dc4f-en.

B.2 Stock indices

Monthly logarithmic returns are shown in Figure 2. To highlight the presence of common risk factors shared among the G7 nations, Table 1 presents the results of a Principal Component Analysis (PCA) conducted on the returns of these stock indices. The first principal component (PC1) accounts for a substantial portion of the variability, explaining approximately 77.2%. This implies that a significant part of the overall variation in the data can be attributed to a single systematic factor. Subsequent components like PC2 and PC3 elucidate progressively smaller portions of the variance. Despite the diminishing explanatory power of each additional component, they still contribute to a better understanding of the interconnections among these economies.

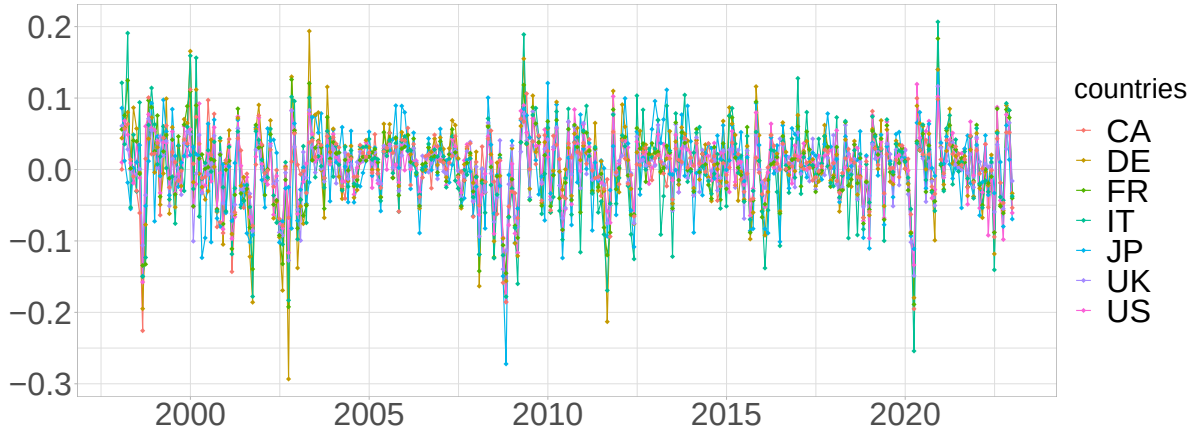

Figure 2: Log returns of G7 stock indices over the considered period: Canada (cyan), France (green), Germany (purple), Italy (magenta), Japan (blue), the United Kingdom (orange), and the United States (red).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Standard deviation	2.3250	0.7262	0.6601	0.4850	0.4415	0.3730	0.2518
Proportion of Variance	0.7720	0.0754	0.0622	0.0336	0.0278	0.0199	0.0091
Cumulative Proportion	0.7720	0.8474	0.9096	0.9432	0.9711	0.9909	1.0000

Table 1: PCA results for G7 stock market indices: standard deviations (first row), variance proportions (second row), and cumulative proportions (third row).

Figure 3 presents the obtained realised volatilities.

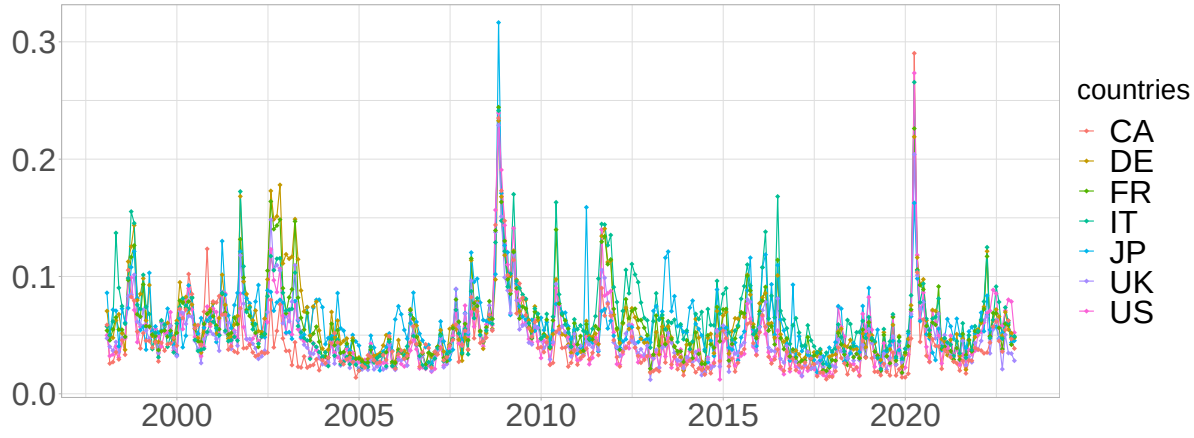

Figure 3: Realised volatilities of G7 stock indices over the considered period: Canada (cyan), France (green), Germany (purple), Italy (magenta), Japan (blue), the United Kingdom (orange), and the United States (red).

C Complementary results of the application

In this section, we present complement and additional findings based on the application for the case of returns and volatilities.

C.1 The case of returns

We present the model outcomes after excluding the SAR component, a step taken to validate the effectiveness of the control variables. The results reveal that, apart from ΔBCI in Italy, all coefficients exhibit statistical significance, underscoring the effectiveness of BCI and CCI as control variables in elucidating the dynamics of stock market indices.

	const	ΔBCI						
		US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.0058**	**0.0460**	**0.0586**	**0.0409**	**0.0310**	**0.0440**	0.0357	**0.0536**
CI 2.5%	0.0041	0.0219	0.0213	0.0064	0.0049	0.0096	-0.0007	0.0278
CI 97.5%	0.0074	0.0718	0.0964	0.0767	0.0575	0.0765	0.0743	0.0786
		ΔCCI						
		US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$		**0.0509**	**0.0437**	**0.0785**	**0.0345**	**0.0654**	**0.0845**	**0.0280**
CI 2.5%		0.0250	0.0039	0.0422	0.0074	0.0326	0.0492	0.0022
CI 97.5%		0.0769	0.0838	0.1134	0.0616	0.0985	0.1185	0.0539

Table 2: Posterior mean $\hat{\beta}$ and 95% credible interval of the coefficients $\boldsymbol{\beta}$, for the static model excluding the SAR component. Coefficients statistically significant (i.e., such that the 95% credible interval does not include zero) are in bold.

Finally, we present the country-specific direct effects for each country over time for returns. Figure 4 displays the distribution of country-specific direct effects on the left panel, illustrating the average response across countries to shocks in stock market returns.

The right panel depicts the distribution of indirect spillover effects, which represent the average response across countries resulting from interactions with other countries over time. Additionally, Figure 5 displays the posterior mean of the country-specific direct effects for the case of returns.

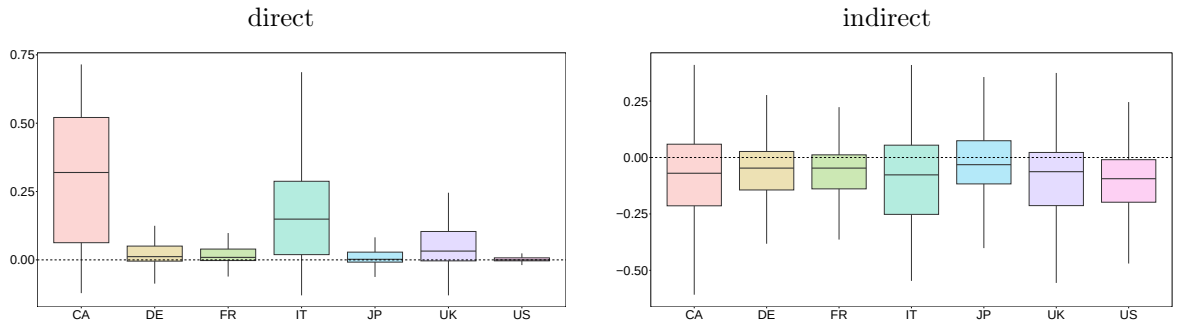

Figure 4: Distribution (over time) of the country-specific direct (left) and indirect (right) spillover effects.

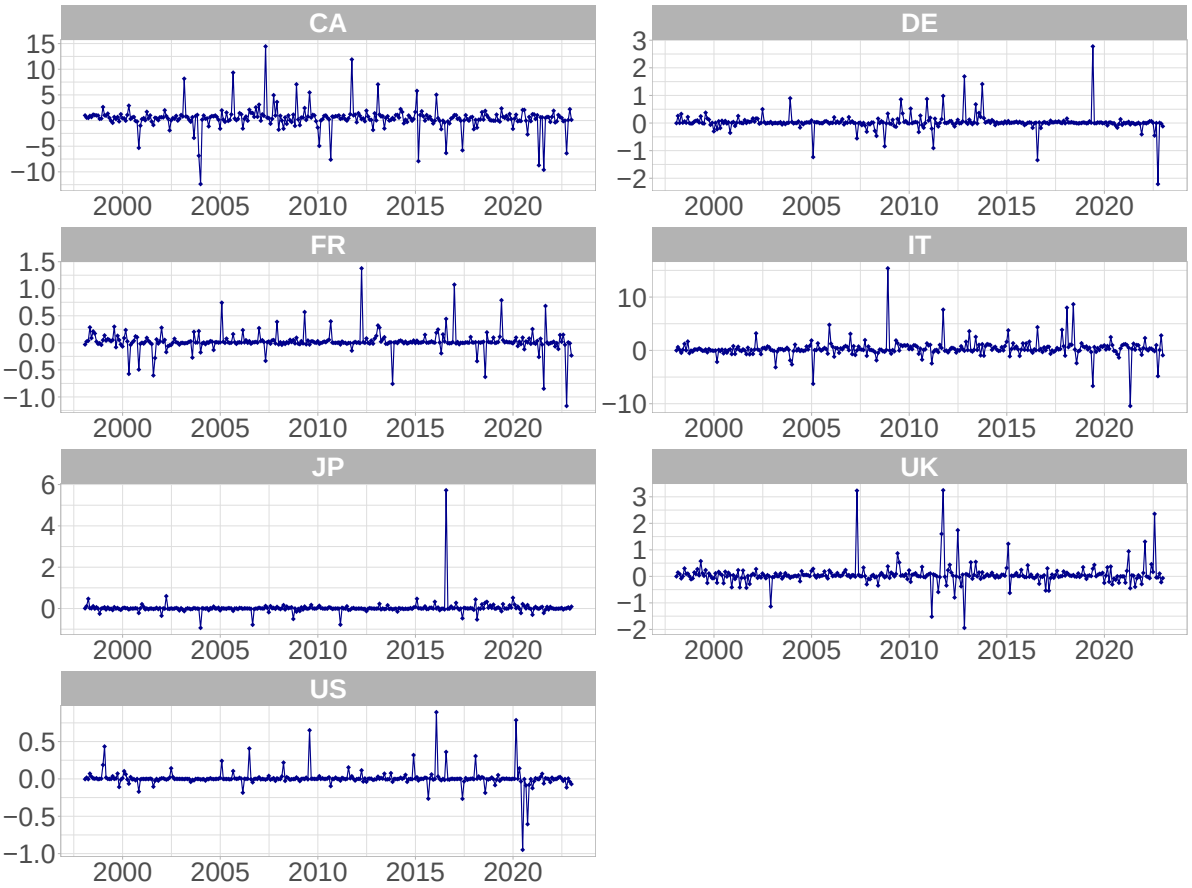

Figure 5: Posterior mean of the country-specific direct effects for each country j over time for the case of the returns.

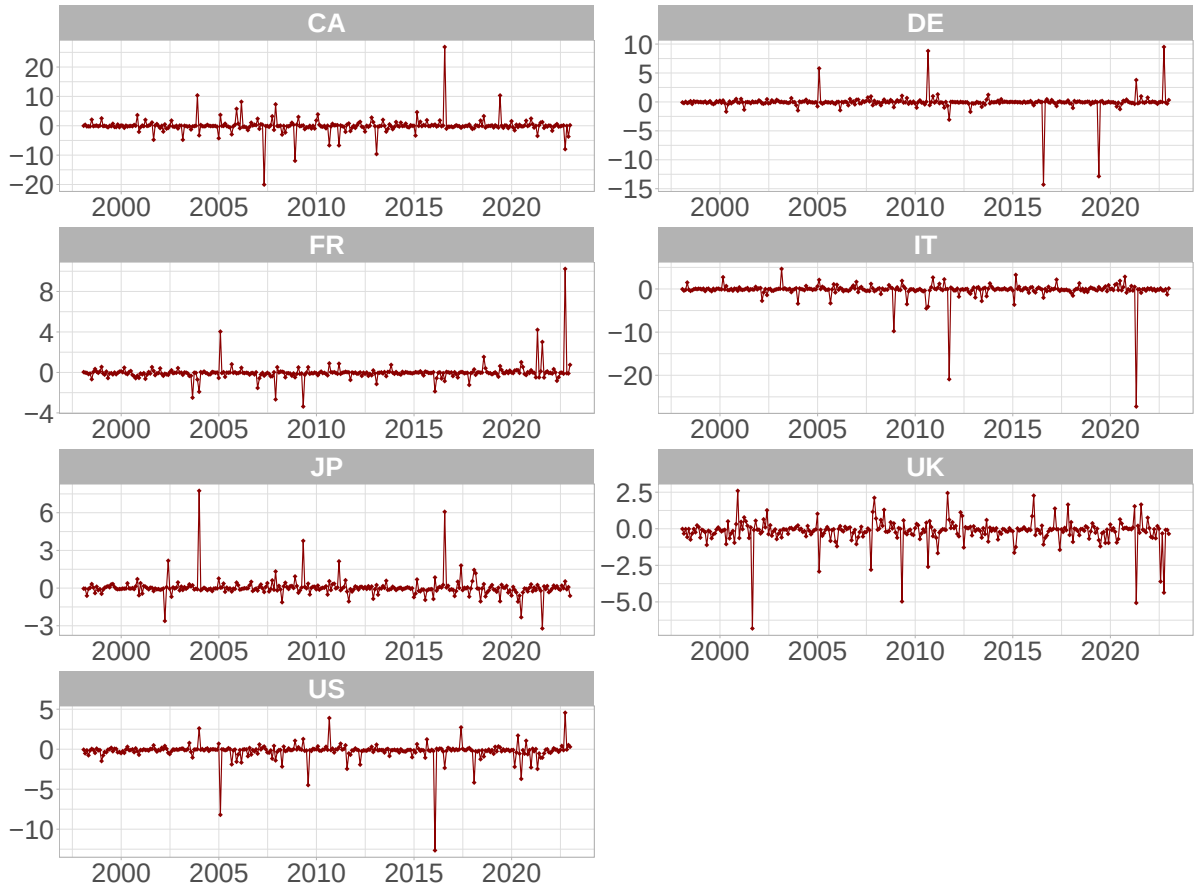

Figure 6: Posterior mean of the country-specific indirect effects for each country j over time for the case of the returns.

C.2 The case of returns with fixed effects

The following subsection presents the results for fixed effects in the model, as reported in Table 3, Table 4, and Figure 7.

				const			
	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.0065**	-0.0004	0.0026	-0.0000	-0.0004	-0.0023	0.0005
CI 2.5%	0.0029	-0.0050	-0.0001	-0.0022	-0.0026	-0.0058	-0.0020
CI 97.5%	0.0101	0.0042	0.0053	0.0023	0.0018	0.0011	0.0031
				ΔBCI			
	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.0408**	**0.0284**	-0.0092	0.0002	-0.0013	-0.0053	**0.0245**
CI 2.5%	0.0149	0.0008	-0.0250	-0.0136	-0.0135	-0.0248	0.0069
CI 97.5%	0.0656	0.0552	0.0059	0.0140	0.0115	0.0131	0.0419
				ΔCCI			
	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.0468**	-0.0039	**0.0248**	0.0011	**0.0142**	**0.0293**	-0.0070
CI 2.5%	0.0194	-0.0336	0.0093	-0.0136	0.0004	0.0107	-0.0242
CI 97.5%	0.0746	0.0255	0.0403	0.0169	0.0288	0.0475	0.0104

Table 3: Data: G7 returns. Posterior mean ($\hat{\beta}$) and 95% credible interval of the coefficients $\boldsymbol{\beta}$. Statistically significant coefficients (those with a 95% credible interval excluding zero) are in bold.

	US	JP	DE	UK	FR	IT	CA
$\hat{\rho}$	0.0770	**0.7302**	**0.9942**	**0.7053**	**0.9886**	**0.9910**	**0.7017**
CI 2.5%	-0.1050	0.5798	0.9728	0.6429	0.9538	0.9582	0.6300
CI 97.5%	0.2418	0.8740	1.0000	0.7656	1.0000	1.0000	0.7705

Table 4: Data: G7 returns. Posterior mean ($\hat{\rho}$) of country-specific spatial weight (ρ_j) for each country j, along with the 95% credible interval. Statistically significant coefficients (those with 95% credible intervals not including zero) are in bold.

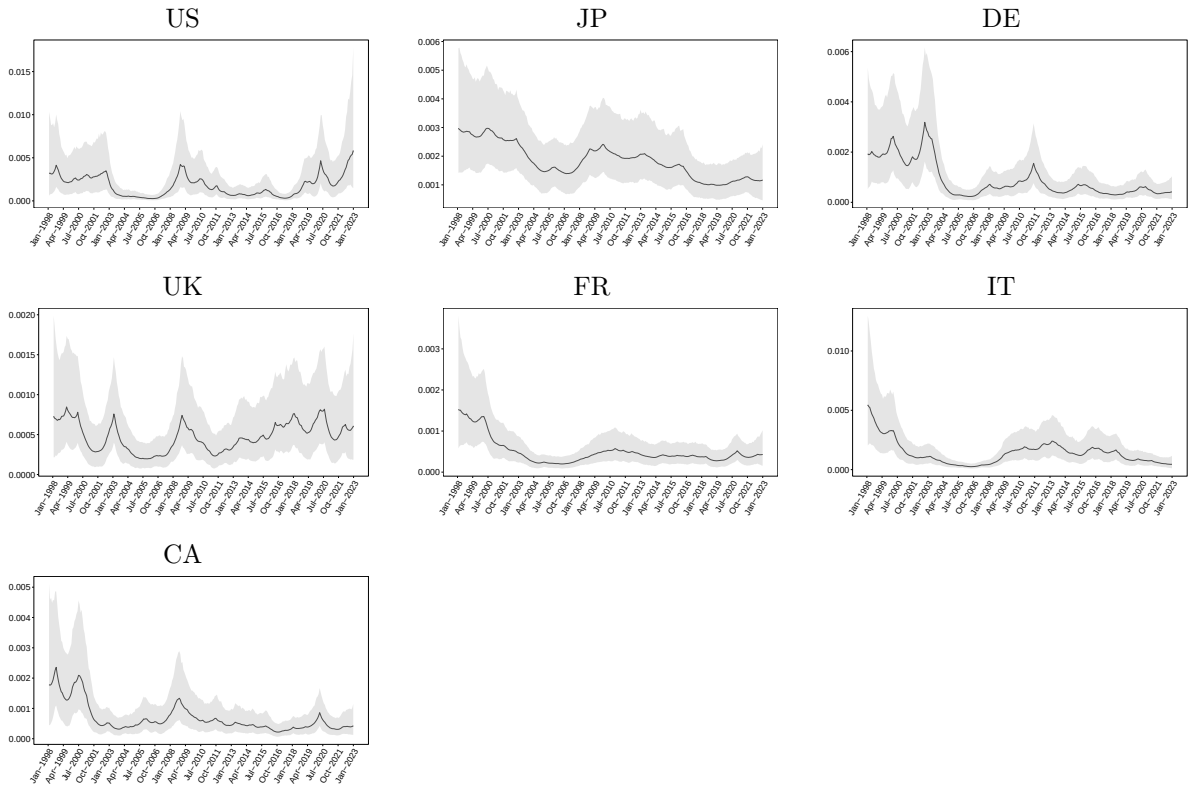

Figure 7: Data: G7 returns. Posterior mean of the country-specific variance of G7 returns, $\exp(h_{j,t})$, for each country j (line), with 95% credible intervals (grey shade).

C.3 The case of returns with time-varying $\rho_{j,t}$

Figures 8-9-10 report the results for the empirical application using a time-varying network weight, $\rho_{j,t}$. Specifically, Figure 8 show the estimated path (posterior mean) of each $\rho_{j,t}$. Then, Figure 9 plots the distribution of direct and indirect spillover effects, whereas Figure 10 illustrates the pairwise cross-correlations of direct, indirect and overall effects.

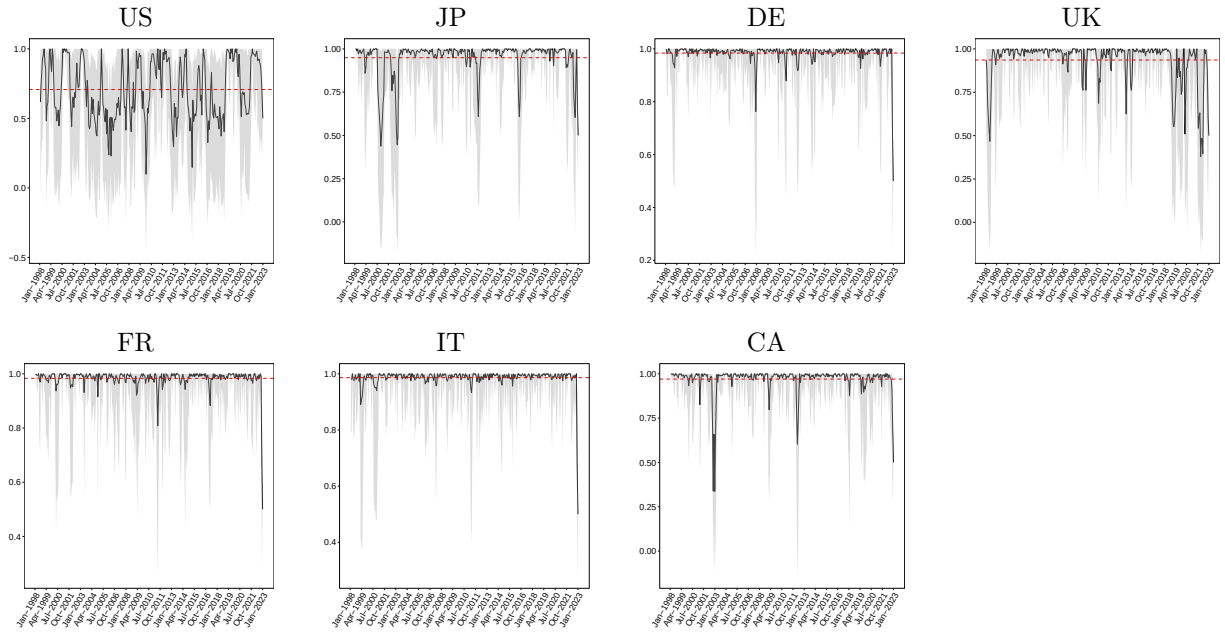

Figure 8: Data: G7 returns. Posterior mean of the country-specific network weight, $\rho_{j,t}$, for each country j (solid black line), with temporal average (dashed line) and 95% credible intervals (solid gray lines).

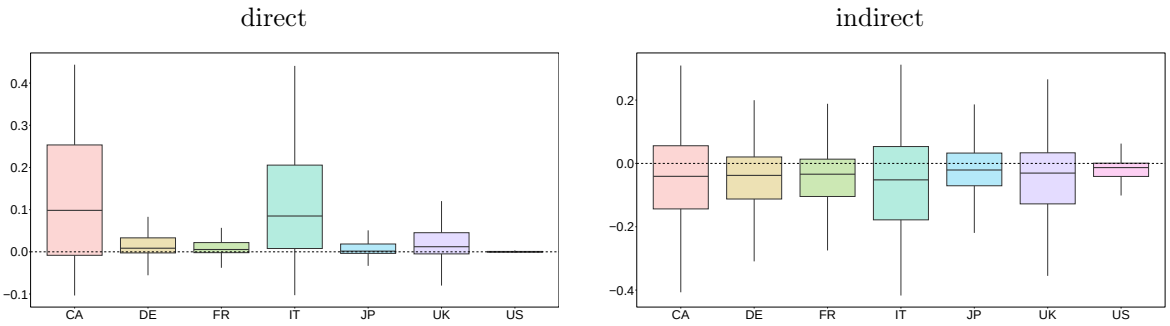

Figure 9: Data: G7 returns. Distribution (over time) of the country-specific direct (left) and indirect (right) spillover effects.

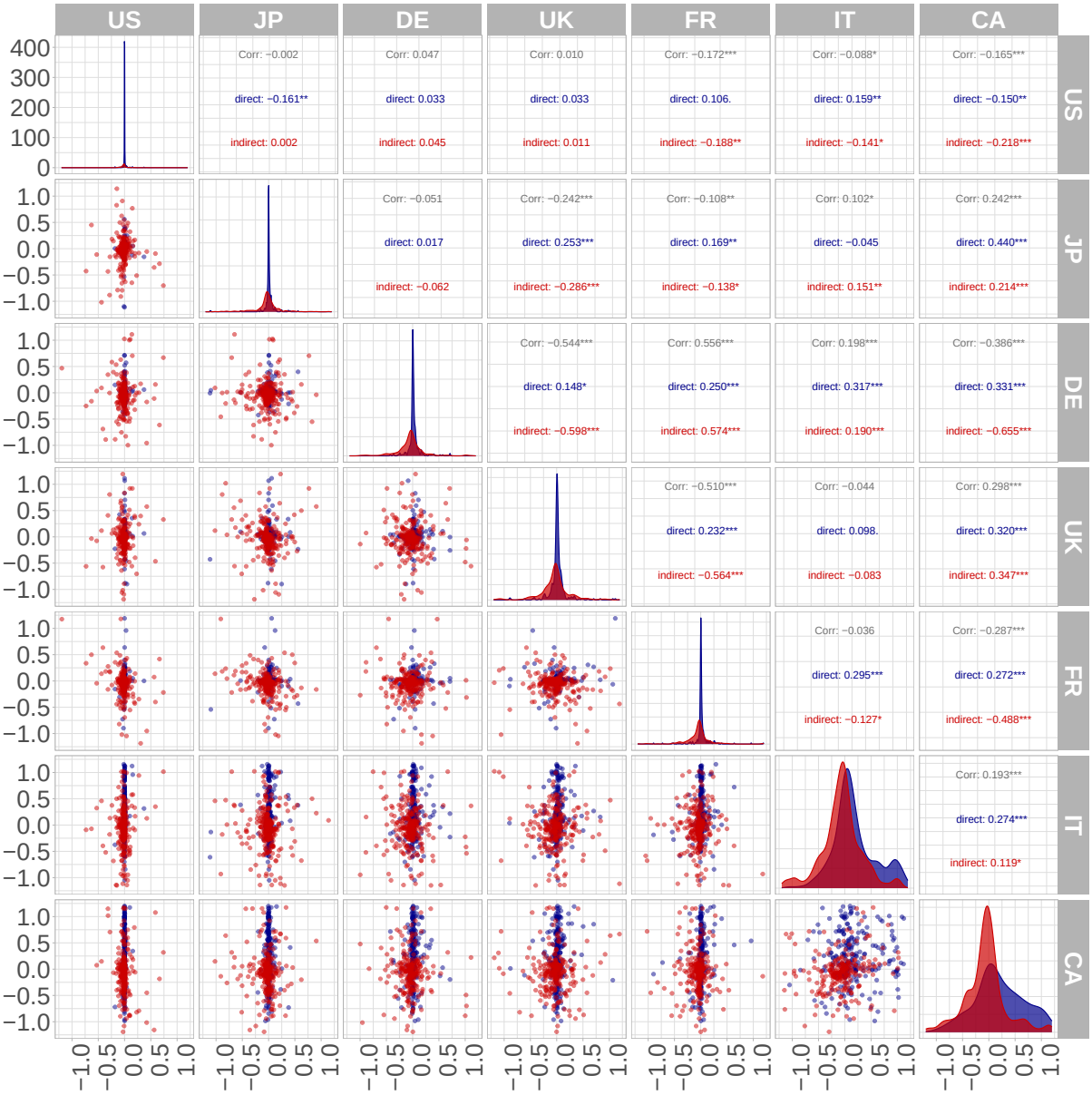

Figure 10: Data: G7 returns. Upper: pairwise cross-correlation of overall (grey), direct (blue), and indirect (red) effects over time between any pair of countries. Stars denote significance at 10% (*), 5% (**), 1% (***). Diagonal: kernel density estimate of the distribution of direct (blue) and indirect (red) effects over time for each country, visualised on the interval $(-1.2, 1.2)$. Bottom: scatter of the bivariate distribution of direct (blue) and indirect (red) effects over time for each pair of countries, visualised on the region $(-1.2, 1.2) \times (-1.2, 1.2)$.

C.4 The case of volatilities

In the alternative application, we consider the realised volatility as follows:

$$\text{RV}_{i,t} = \sum_{l=1}^{M} r_{l,t}^2, \tag{6}$$

where i ={CA, FR, DE, IT, JP, UK, US}, t denotes the monthly time index, and M represents the total number of market days in a given month t.

Table 5 presents the posterior mean and associated 95% credible intervals of the realised volatility MSCI_{t-1}. Findings reveal statistically significant coefficients for the United States, Japan, Germany, and the United Kingdom. Figure 11 presents the posterior distribution of the country-specific spatial weight (ρ_j) for each country (j). Additionally, Figure 12 displays the posterior mean of each country's time-varying volatility throughout the period, showing pronounced spikes during specific country-related events. This suggests structural changes in the parameters guiding the stochastic volatility model, as shown in Grassi and de Magistris (2015).

Finally, we present the country-specific direct effects for each country over time, focusing on volatilities. Figure 13 displays the distribution of country-specific direct effects on the left panel, illustrating the average response across countries to shocks in market volatilities. The right panel shows the distribution of indirect spillover effects, depicting the average response across countries resulting from interactions with other countries over time. Furthermore, Figure 14 displays the posterior mean of the country-specific direct effects for the case of volatilities. Finally, Figure 16 illustrates the cross-correlation of overall (grey), direct (blue), and indirect (red) effects between pairs of countries over time.

		realised volatility MSCI_{t-1}						
	const	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.0063**	**0.6605**	**0.4268**	0.0972	0.0351	0.0012	0.0093	0.0248
CI 2.5%	0.0052	0.5238	0.2752	0.0461	-0.0215	-0.0480	-0.0750	-0.0444
CI 97.5%	0.0073	0.8159	0.5671	0.1551	0.0896	0.0525	0.1064	0.0869

Table 5: Data: G7 volatilities. Posterior mean ($\hat{\beta}$) and 95% credible interval of the coefficients $\boldsymbol{\beta}$. Statistically significant coefficients (those with a 95% credible interval excluding zero) are in bold.

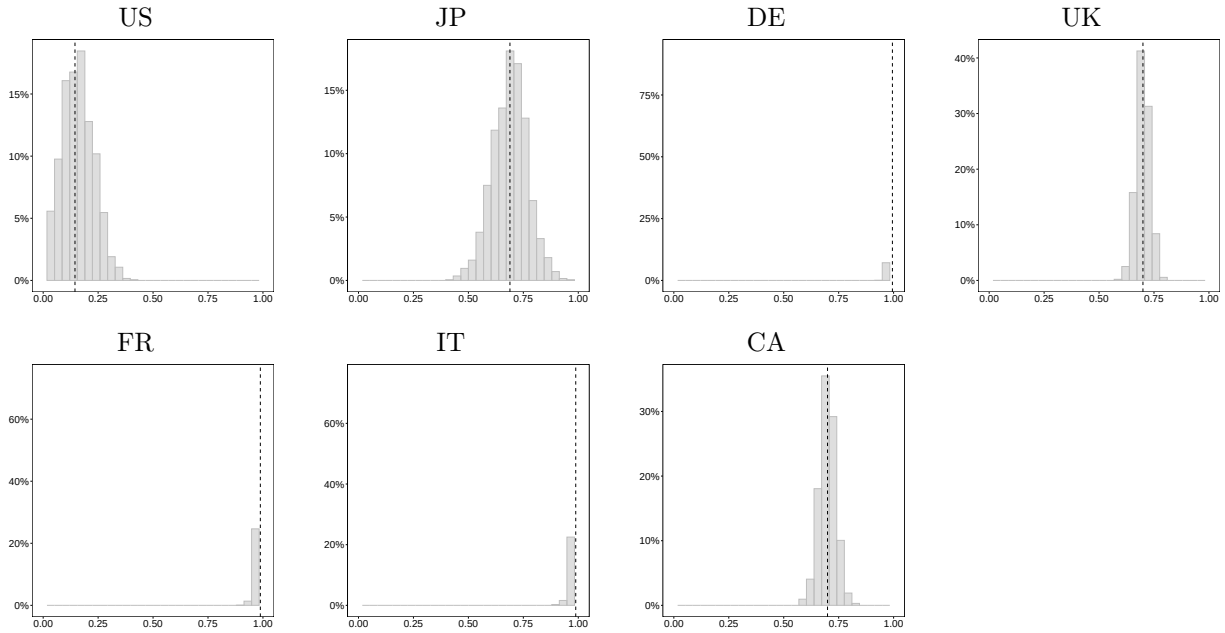

Figure 11: Posterior distribution of the country-specific spatial weight ρ_j, for each country j, with the posterior mean (dashed line) for G7 volatilities.

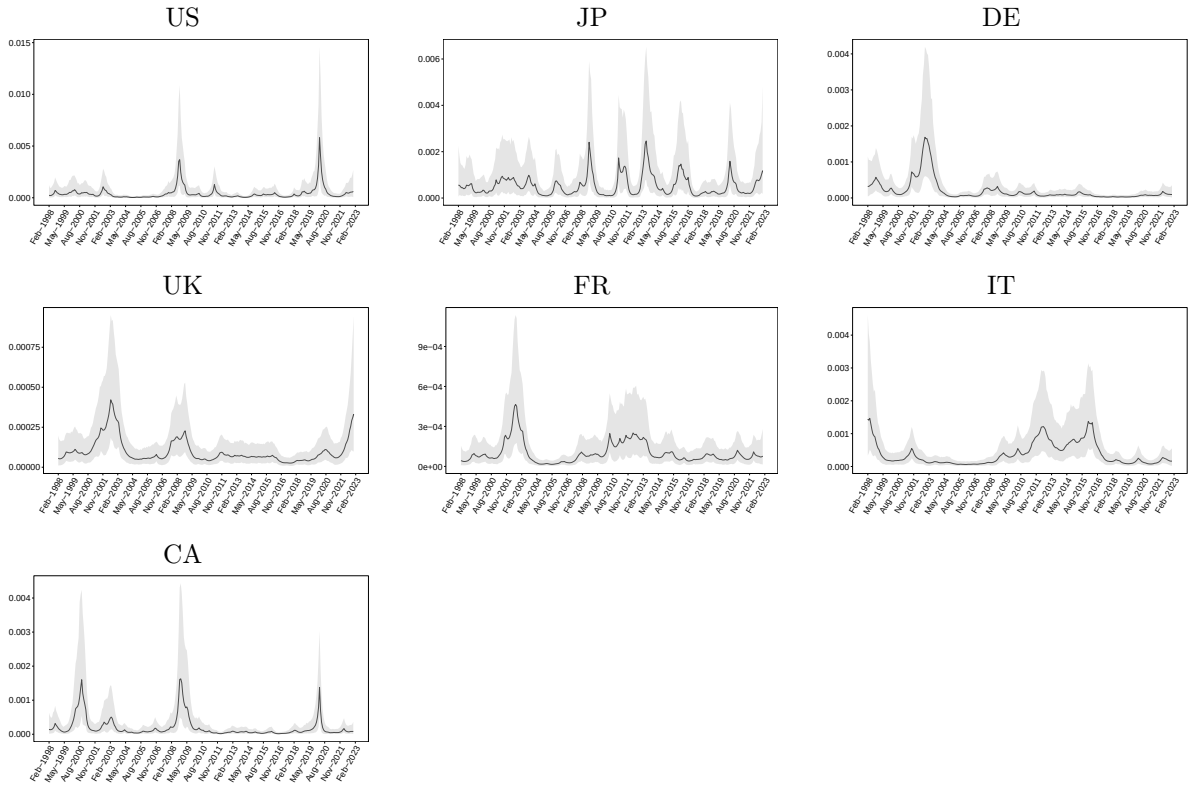

Figure 12: Data: G7 volatilities. Posterior mean of the country-specific variance, $\exp(h_{j,t})$, for each country j (line), with 95% credible intervals (grey shade).

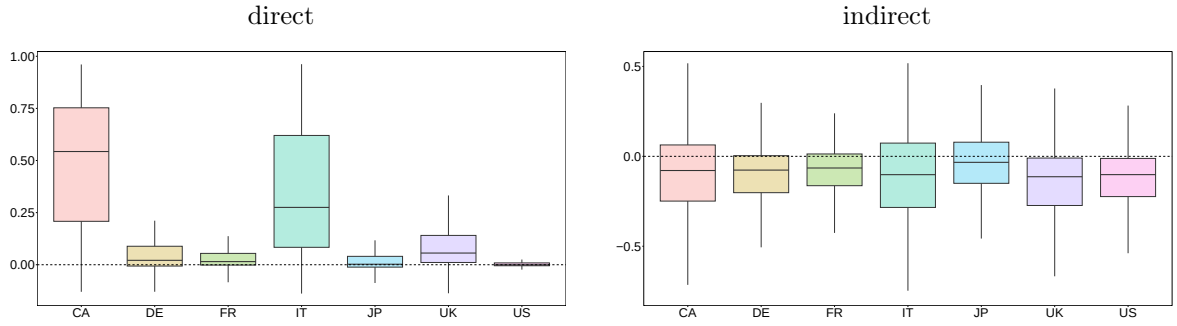

Figure 13: Distribution (over time) of the country-specific direct (left) and indirect (right) spillover effects.

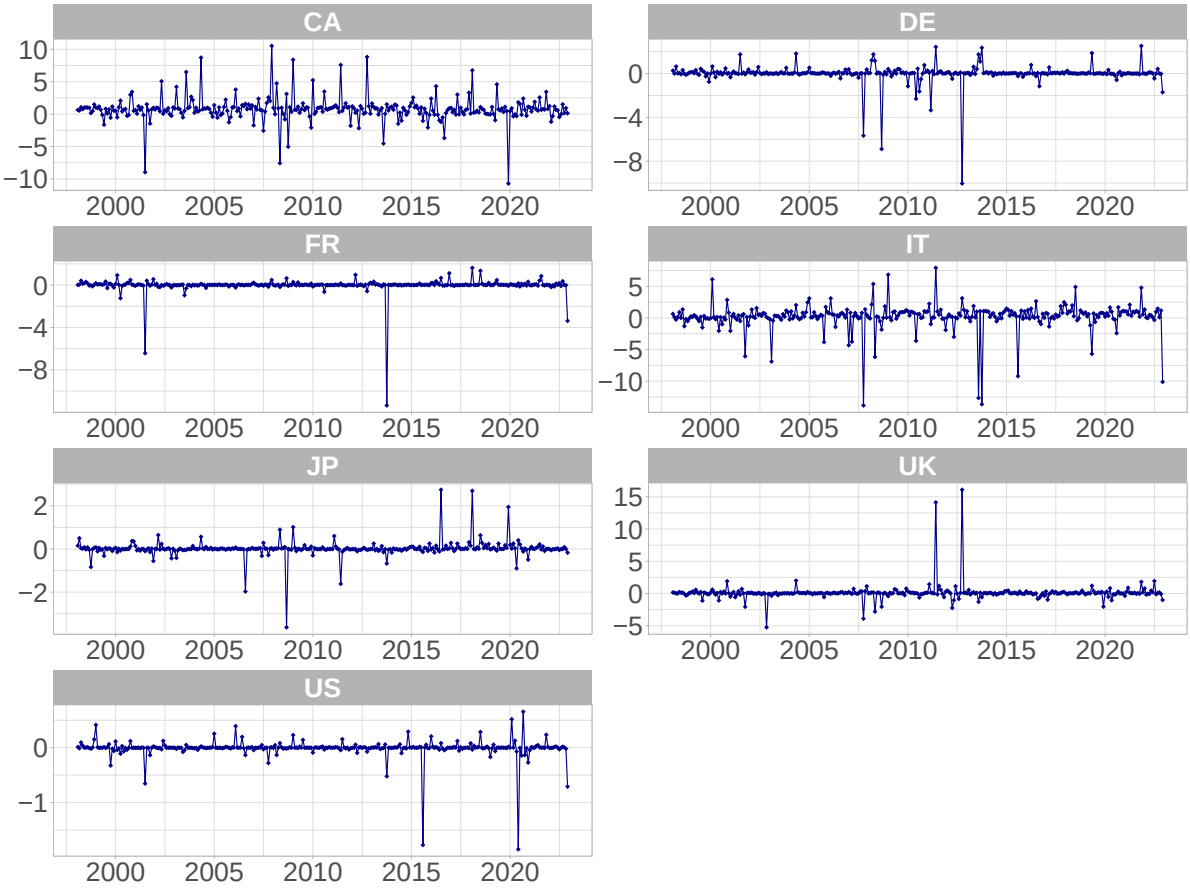

Figure 14: Posterior mean of the country-specific direct effects for each country j over time for the case of the volatilities.

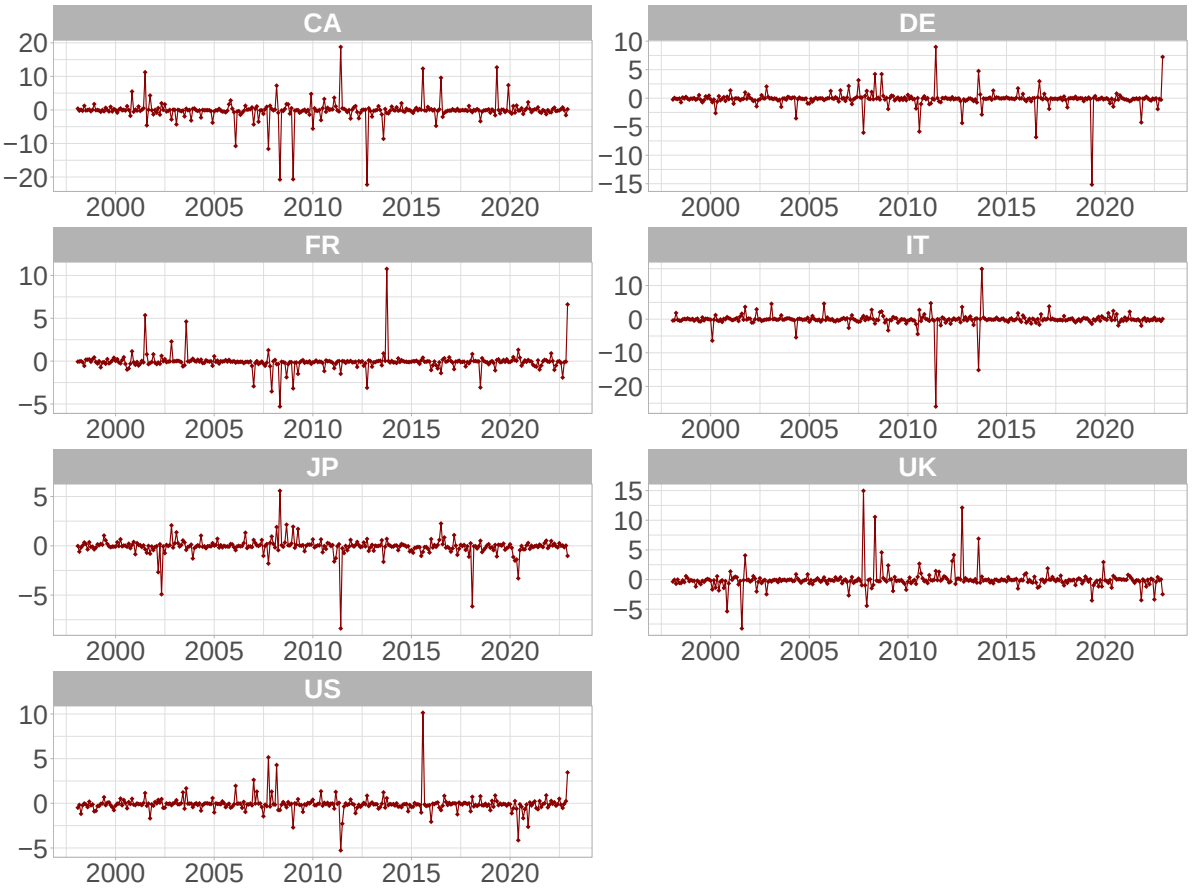

Figure 15: Posterior mean of the country-specific indirect effects for each country j over time for the case of the volatilities.

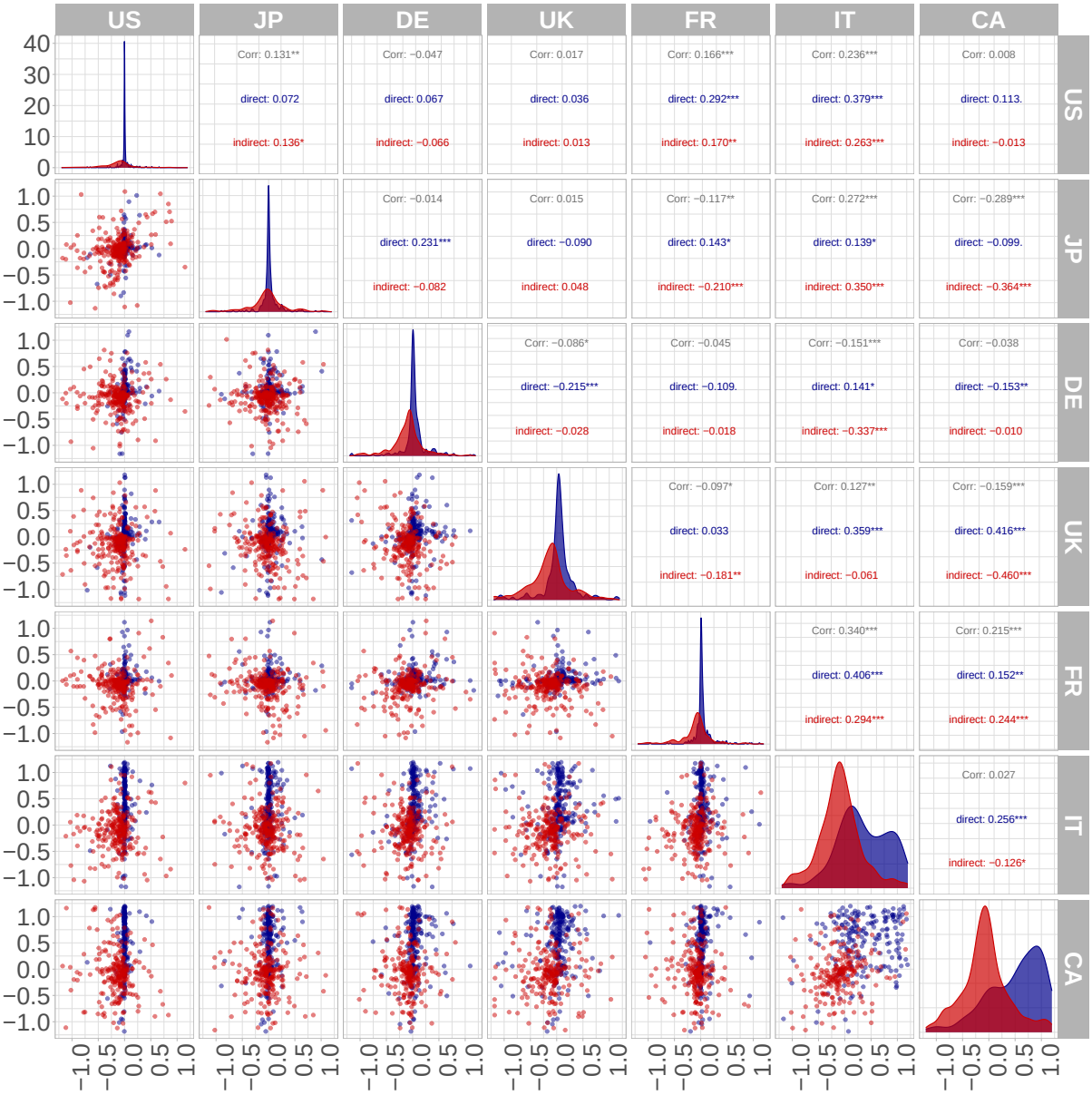

Figure 16: Upper: pairwise cross-correlation of overall (grey), direct (blue), and indirect (red) effects over time between any pair of countries. Stars denote significance at 10% (*), 5% (**), 1% (***). Diagonal: kernel density estimate of the distribution of direct (blue) and indirect (red) effects over time for each country, visualised on the interval $(-1.2, 1.2)$. Bottom: scatter of the bivariate distribution of direct (blue) and indirect (red) effects over time for each pair of countries, visualised on the region $(-1.2, 1.2) \times (-1.2, 1.2)$.

C.5 The case of volatilities with fixed effects

The following subsection presents the results for fixed effects in the model, as reported in Table 6, Table 7, and Figure 17.

				const			
	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**-1.6692**	**-2.1695**	**-0.4267**	**-0.6649**	-0.2573	**-0.8365**	**-0.8312**
CI 2.5%	-2.5849	-2.6781	-0.7201	-0.9893	-0.5405	-1.3077	-1.3269
CI 97.5%	-0.8422	-1.6780	-0.1105	-0.3327	0.0167	-0.3852	-0.3731
				realised volatility MSCI_{t-1}			
	US	JP	DE	UK	FR	IT	CA
$\hat{\beta}$	**0.7012**	**0.0854**	**0.0632**	**0.0560**	0.0412	0.0657	**0.0741**
CI 2.5%	0.5455	0.0034	0.0126	0.0010	-0.0093	-0.0141	0.0005
CI 97.5%	0.8714	0.1678	0.1171	0.1143	0.0968	0.1515	0.1490

Table 6: Data: G7 volatilities. Posterior mean ($\hat{\beta}$) and 95% credible interval of the coefficients $\boldsymbol{\beta}$. Statistically significant coefficients (those with a 95% credible interval excluding zero) are in bold.

	US	JP	DE	UK	FR	IT	CA
$\hat{\rho}$	-0.0011	**0.4889**	**0.8197**	**0.8577**	**0.8787**	**0.7516**	**0.8539**
CI 2.5%	-0.2500	0.4067	0.7482	0.8031	0.8204	0.6509	0.7710
CI 97.5%	0.2144	0.5743	0.8810	0.8996	0.9283	0.8470	0.9244

Table 7: Data: G7 volatilities. Posterior mean ($\hat{\rho}$) of country-specific spatial weight (ρ_j) for each country j, along with the 95% credible interval. Statistically significant coefficients (those with 95% credible intervals not including zero) are in bold.

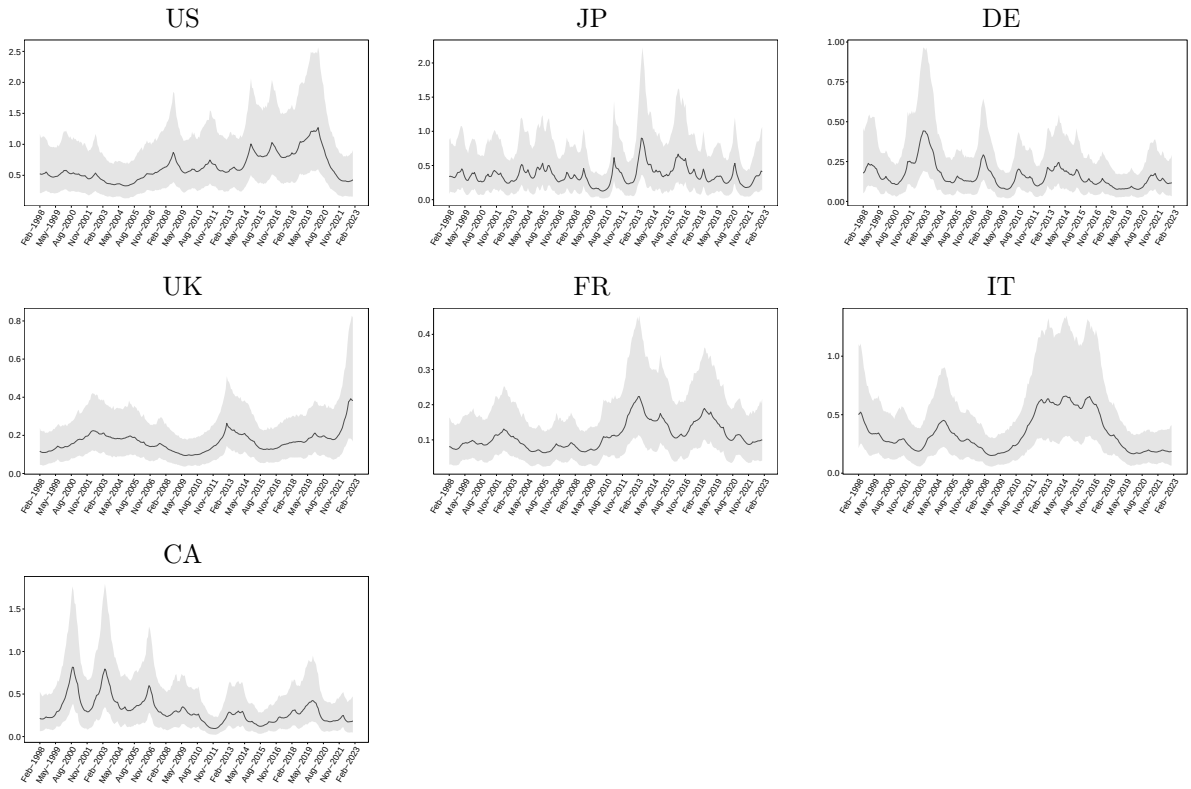

Figure 17: Data: G7 volatilities. Posterior mean of the country-specific variance of G7 returns, $\exp(h_{j,t})$, for each country j (line), with 95% credible intervals (grey shade).

C.6 The case of volatilities with time-varying $\rho_{j,t}$

Figures 18-19-20 report the results for the empirical application using a time-varying network weight, $\rho_{j,t}$. Specifically, Figure 18 show the estimated path (posterior mean) of each $\rho_{j,t}$. Then, Figure 19 plots the distribution of direct and indirect spillover effects, whereas Figure 20 illustrates the pairwise cross-correlations of direct, indirect and overall effects.

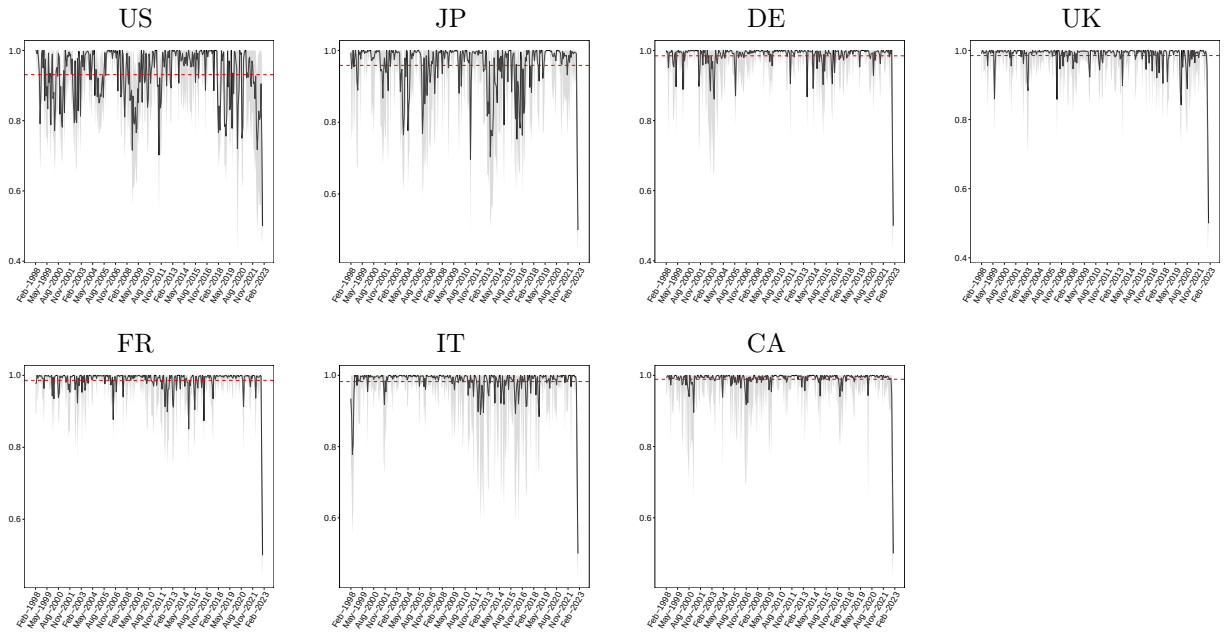

Figure 18: Data: G7 volatilities. Posterior mean of the country-specific network weight, $\rho_{j,t}$, for each country j (solid black line), with temporal average (dashed line) and 95% credible intervals (solid gray lines).

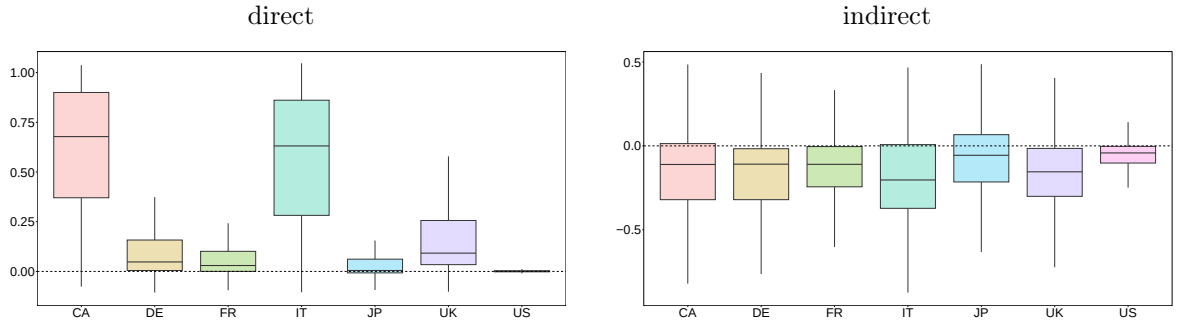

Figure 19: Data: G7 volatilities. Distribution (over time) of the country-specific direct (left) and indirect (right) spillover effects.

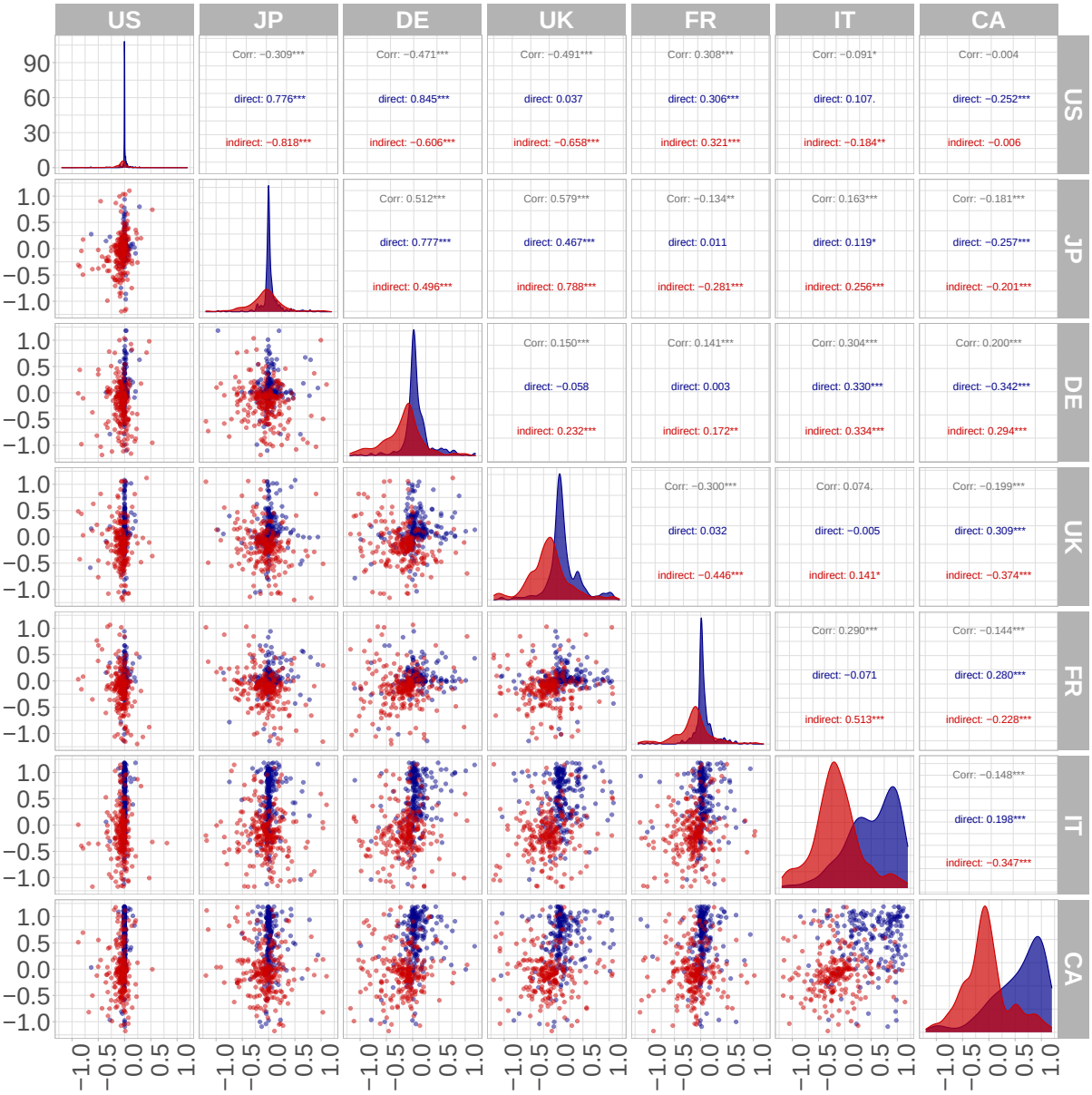

Figure 20: Data: G7 volatilities. Upper: pairwise cross-correlation of overall (grey), direct (blue), and indirect (red) effects over time between any pair of countries. Stars denote significance at 10% (*), 5% (**), 1% (***). Diagonal: kernel density estimate of the distribution of direct (blue) and indirect (red) effects over time for each country, visualised on the interval $(-1.2, 1.2)$. Bottom: scatter of the bivariate distribution of direct (blue) and indirect (red) effects over time for each pair of countries, visualised on the region $(-1.2, 1.2) \times (-1.2, 1.2)$.

D Model selection metrics

A likelihood-free measure of fitness for the SAR-SV model (and competitors) is the MAE($\mathbf{y}$) for the observed data, defined as

$$\text{MAE}_\ell(\mathbf{y}) = \frac{1}{nT}\sum_{t=1}^{T}\sum_{j=1}^{n}\left|y_{j,t} - \hat{y}_{j,t}^{(\ell)}\right|, \tag{7}$$

where $\hat{y}_{j,t}^{(\ell)}$ is the fitted value from model ℓ.

Alternative model selection criteria include the marginal likelihood and the information criteria, and their computation for the proposed SAR-SV model is explained below.

D.1 Marginal likelihood

An alternative approach to model selection in a full Bayesian way relies on posterior model probabilities obtained from the marginal likelihood (ML).[3] In contrast to non-nested model tests that rely on model estimates and associated predictions. Since in the presence of model misspecification the resulting estimates will be biased, it would be desirable to perform model comparison and selection unconditional on the parameter estimate (Debarsy and LeSage, 2018). The Bayesian approach via marginal likelihood addresses this issue by integrating over all model parameters, which allow us to make inference about the best model unconditional on any specific parameter estimate.

The ML is defined as:

$$p(\mathbf{y}|\mathcal{M}_i) = \iint p(\mathbf{y}|\mathbf{h},\boldsymbol{\theta},\mathcal{M}_i)p(\mathbf{h}|\boldsymbol{\theta},\mathcal{M}_i)p(\boldsymbol{\theta}|\mathcal{M}_i)\mathrm{d}\mathbf{h}\mathrm{d}\boldsymbol{\theta} = \int p(\mathbf{y}|\boldsymbol{\theta},\mathcal{M}_i)p(\boldsymbol{\theta}|\mathcal{M}_i)\mathrm{d}\boldsymbol{\theta}, \tag{8}$$

where $\mathcal{M}_i$ denotes model i, $\boldsymbol{\theta}$ is the collection of all parameters and $p(\mathbf{y}|\boldsymbol{\theta},\mathcal{M}_i)$ is the integrated likelihood of model i. Computing the marginal likelihood for models with stochastic volatility is a nontrivial task since they are inherently mixture models. Following Chan and Eisenstat (2018), we design a two-step procedure to approximate the marginal likelihood based on an improved cross-entropy method where the integrated likelihood is computed using importance sampling. In the following description, we remove the conditioning on $\mathcal{M}_i$ for notational simplicity.

First, we aim at minimising the Kullback–Leibler divergence between the ideal importance sampling density (i.e., the posterior density) and a candidate density, $p^*(\boldsymbol{\theta};\mathbf{v}) \in \mathcal{F}$, indexed by a parameter vector $\mathbf{v}$. We do this by solving the minimisation problem

$$\mathbf{v}^* = \arg\min_{\mathbf{v}} \frac{1}{M}\sum_{m=1}^{M}\log\left(p^*(\boldsymbol{\theta}^{(m)};\mathbf{v})\right),$$

[3] The higher the value of the marginal likelihood, the better the model's performance.

where $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(M)}$ are posterior draws of the parameters and the candidate density $p^*(\boldsymbol{\theta}^{(m)}; \mathbf{v})$ is defined as the product of densities:

$$p^*(\boldsymbol{\theta}; \mathbf{v}) = p_1(\boldsymbol{\theta}_1; \mathbf{v}_1) \times \ldots \times p_J(\boldsymbol{\theta}_J; \mathbf{v}_J),$$

with J being the number of blocks of parameters. The main advantage of this approach is the reduction in the computation time due to disentangling a high-dimensional problem into a collection of B low-dimensional ones. Recalling that $\boldsymbol{\theta} = (\boldsymbol{\mu}_h', \boldsymbol{\phi}_h', \boldsymbol{\sigma}_h^{2\prime}, \boldsymbol{\beta}', \boldsymbol{\delta}', \boldsymbol{\rho}')'$, we exploit the natural blocking suggested by the SAR-SV model and consider the family

$$\mathcal{F} = \Bigg\{ \prod_{j=1}^{n} f_{\mathcal{N}}(\mu_{h,j}; v_{\mu,1}, v_{\mu,2}) \times f_{\mathcal{N}}(\phi_{h,j}; v_{\phi,1}, v_{\phi,2}) \times f_{IG}(\sigma_{h,j}^2; v_{\sigma,1}, v_{\sigma,2}) \\ \times f_{\mathcal{N}}(\boldsymbol{\beta}; \mathbf{v}_{\beta,1}, \mathbf{v}_{\beta,2}) \times f_{\mathcal{D}ir}(\boldsymbol{\delta}; \mathbf{v}_{\delta}) \times \prod_{j=1}^{n} f_{\mathcal{B}e}\Big(\frac{\rho+1}{2}; v_{\rho,1}, v_{\rho,2}\Big) \Bigg\}.$$

Having obtained the optimal density $p^*(\boldsymbol{\theta}; \mathbf{v}^*)$, in the second step we draw $\ell = 1, \ldots, N$ samples $\boldsymbol{\theta}^{(\ell)} \sim p^*(\boldsymbol{\theta}; \mathbf{v}^*)$ and approximate the marginal likelihood via the average

$$\widehat{p}(\mathbf{y}) = \frac{1}{N} \sum_{\ell=1}^{N} \frac{\widehat{p}(\mathbf{y}|\boldsymbol{\theta}^{(\ell)}) p(\boldsymbol{\theta}^{(\ell)})}{p^*(\boldsymbol{\theta}^{(\ell)}; \mathbf{v}^*)}, \tag{9}$$

where $\widehat{p}(\mathbf{y}|\boldsymbol{\theta})$ is an estimate of the integrated likelihood computed using the approach previously described.

For a model without stochastic volatility, the marginal likelihood can be computed similarly. In this case, first recall that $p(\mathbf{y}|\boldsymbol{\theta})$ is available in closed form as the observed likelihood. Second, the integral over the parameters can be approximated numerically using an importance sampling procedure analogous to the one described above.

In particular, we minimise the Kullback–Leibler divergence between the ideal importance sampling density (i.e., the posterior density) and a candidate density, $p^*(\boldsymbol{\theta}; \mathbf{v}) \in \mathcal{F}$ by solving the problem

$$\mathbf{v}^* = \arg\min_{\mathbf{v}} \frac{1}{M} \sum_{m=1}^{M} \log\big(p^*(\boldsymbol{\theta}^{(m)}; \mathbf{v})\big),$$

where $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(M)}$ are posterior draws of the parameters and the candidate density $p^*(\boldsymbol{\theta}^{(m)}; \mathbf{v})$ is defined as the product of densities. We exploit the natural blocking suggested by the SAR-SV model, thus obtaining the family

$$\mathcal{F} = \Bigg\{ f_{\mathcal{N}}(\boldsymbol{\beta}; \mathbf{v}_{\beta,1}, \mathbf{v}_{\beta,2}) \times f_{\mathcal{D}ir}(\boldsymbol{\delta}; \mathbf{v}_{\delta}) \times \prod_{j=1}^{n} f_{\mathcal{B}e}\Big(\frac{\rho+1}{2}; v_{\rho,1}, v_{\rho,2}\Big) \Bigg\}.$$

Finally, we draw samples $\boldsymbol{\theta}^{(\ell)} \sim p^*(\boldsymbol{\theta}; \mathbf{v}^*)$, $\ell = 1, \ldots, N$, from the optimal density and approximate the marginal likelihood via the average

$$\widehat{p}(\mathbf{y}) = \frac{1}{N} \sum_{\ell=1}^{N} \frac{p(\mathbf{y}|\boldsymbol{\theta}^{(\ell)}) p(\boldsymbol{\theta}^{(\ell)})}{p^*(\boldsymbol{\theta}^{(\ell)}; \mathbf{v}^*)}. \tag{10}$$

E Simulation study settings

This simulation study investigates the ability of the marginal likelihood (ML) to recover the "true model" generating the data. We consider the following data generating processes (DGPs):

1. a SAR model with a one-layer time constant network and homoskedastic innovations (SAR-W);
2. a SAR model with a two-layer time-varying networks and homoskedastic innovations (SAR-H);
3. the proposed model, with a two-layer time-varying networks and stochastic volatility (SAR-SV).

For each DGP, we consider $n = 7$ response variables and $p = 2$ covariates, over $T = 1000$ periods. At each time t, we generate the covariates by drawing from a uniform distribution $f_{kt} \sim \mathcal{U}(-2.0, 2.0)$, for each $k = 1, \ldots, p$, and the associated coefficients are drawn from a standard Gaussian distribution, that is $\beta_l \sim \mathcal{N}(0, 1)$, for each $l = 1, \ldots, k_\beta$ The network weights are sampled as $\rho_j \sim \mathcal{U}(0.50, 0.999)$, for each $j = 1, \ldots, n$, whereas the layer-specific weights are set to $\boldsymbol{\delta} = (0.75, 0.25)$. Concerning the structural variance, the homoskedastic models draw $\sigma_j^2 \sim \mathcal{IG}(4.0, 3.5)$, for each $j = 1, \ldots, n$, whereas the SAR-SV model simulates the path of the log-volatility independently for each $j = 1, \ldots, n$ using as hyperparameters $(\mu_h, \phi_h, \sigma_h^2) = (-10.0,\ 0.98,\ 0.28^2)$. Finally, each layer of the network at each time t is generated by randomly drawing the entries from two uniform distributions, that is $W_{1t,lk} \sim \mathcal{U}(0, 1)$ in layer one and $W_{2t,lk} \sim \mathcal{U}(0, 0.5)$ in layer two, for each $l, k = 1, \ldots, n$. For the SAR-W model, a one-layer time constant network W is obtained by first generating the two-layer sequence as above, then computing $W = \frac{1}{T} \sum_{t=1}^{T} |W_{1t} - W_{2t}|$.

F Residual correlation

To validate the use of the proposed multi-layer time-varying spatial matrices as a tool to capture the co-movements among the observables, we first compare the empirical correlation among the observables $\mathbf{y}_t$ and the correlation of the residuals from the proposed SAR-SV model. To this end, Figure 21 and Figure 22 show the empirical correlation

among $\mathbf{y}_t$ and the estimated correlation for the residuals series, respectively. In either case, the estimated value of the Spearman correlation is reported together with the associated p-value.

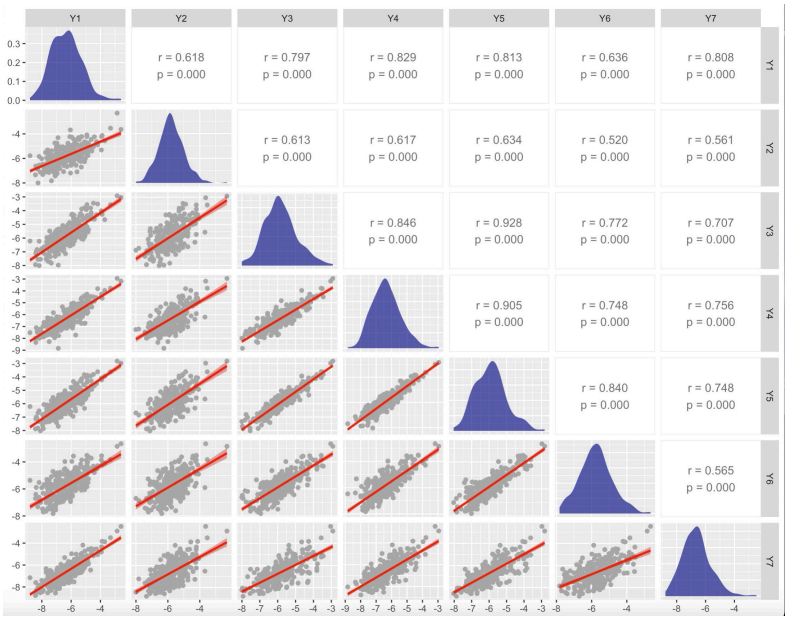

Figure 21: Upper triangular: empirical Spearman correlation (r) and associated p-value (p) among the observables $\mathbf{y}_t$. Lower triangular: pairwise scatterplots of the observables (dots) and robust linear regression fit (solid line) and 95% confidence interval (shade). Diagonal: kernel density estimate of the marginal distribution of each obsevable.

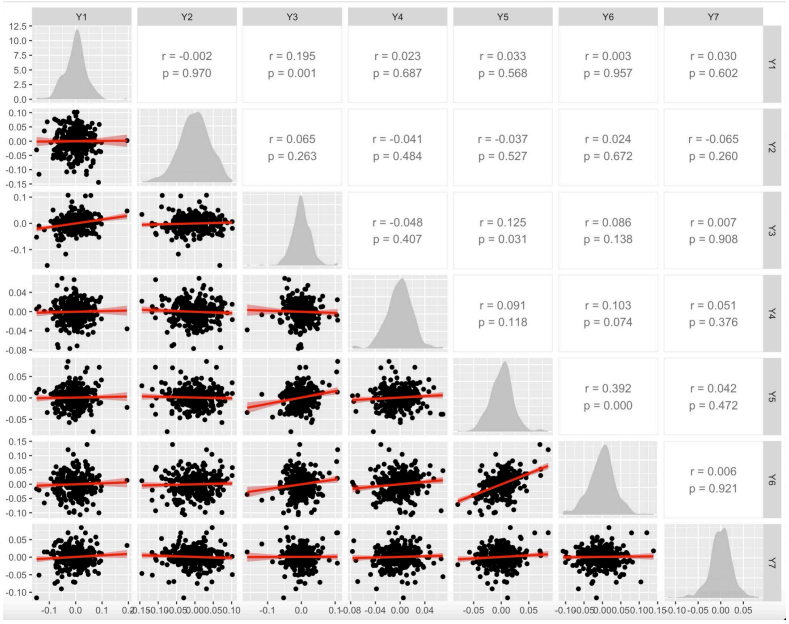

Figure 22: Upper triangular: empirical Spearman correlation (r) and associated p-value (p) among the residuals from our SAR-SV model. Lower triangular: pairwise scatterplots of the residuals (dots) and robust linear regression fit (solid line) and 95% confidence interval (shade). Diagonal: kernel density estimate of the marginal distribution of each residual series.

We find that all correlations among the observed series are statistically significant at any significance level. Conversely, almost every pairwise correlation among the residuals from our model is not statistically different from zero. It is worth mentioning that in two cases where the residual correlation is found nonzero, that is $\varrho(e_1, e_3)$ and $\varrho(e_3, e_5)$, the results are highly affected by few outliers. To better highlight this point, we report on the left of Figure 23 the same scatterplots as in Figure 22, whereas on the right we show the scatterplots, Spearman correlation and p-value for the same series without the outliers. The results show that, once the outliers are removed, the correlation among these residual series is not statistically different from zero.

Therefore, we conclude that the proposed SAR-SV model with multi-layer time-varying spatial weights is able to capture the co-movements among the observed series.

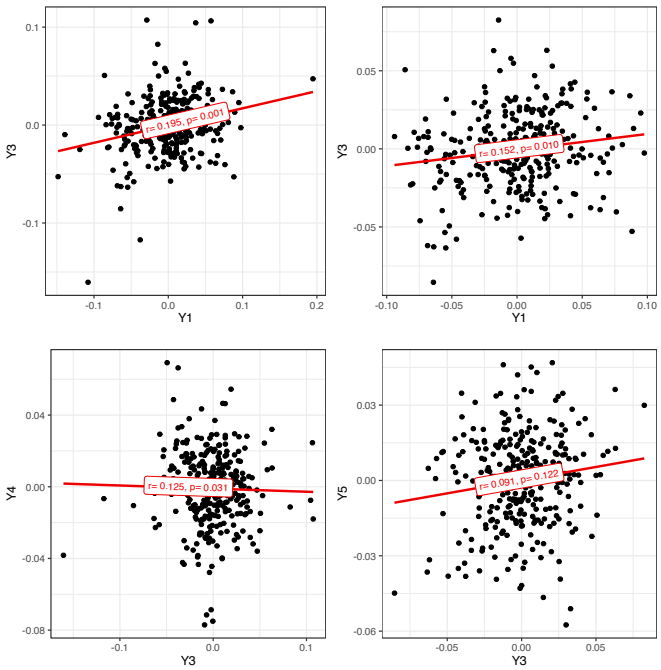

Figure 23: In each plot, the black dots are the scatter of the corresponding series, the solid line represents the robust linear regression fit, r is the estiamted Spearman correlation, and p the associated p-value. First row: pairwise scatterplot of the residual e_1 against e_3, including outliers (left) and without the outliers (right). Second row: pairwise scatterplot of the residual e_3 against e_5, including outliers (left) and without the outliers (right).

References

Bhadra, A., Datta, J., Polson, N. G., and Willard, B. (2016). Default Bayesian analysis with global-local shrinkage priors. *Biometrika*, 103(4):955–969.

Chan, J. C. and Eisenstat, E. (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics*, 33(4):509–532.

Debarsy, N. and LeSage, J. (2018). Flexible dependence modeling using convex combinations of different types of connectivity structures. *Regional Science and Urban Economics*, 69:48–68.

Debarsy, N. and LeSage, J. P. (2022). Bayesian model averaging for spatial autoregressive models based on convex combinations of different types of connectivity matrices. *Journal of Business & Economic Statistics*, 40(2):547–558.

Dittrich, D., Leenders, R. T. A., and Mulder, J. (2017). Bayesian estimation of the network autocorrelation model. *Social Networks*, 48:213–236.

Grassi, S. and de Magistris, P. S. (2015). It's all about volatility of volatility: Evidence from a two-factor stochastic volatility model. *Journal of Empirical Finance*, 30:62–78.

Hauzenberger, N. and Pfarrhofer, M. (2021). Bayesian state-space modeling for analyzing heterogeneous network effects of US monetary policy. *The Scandinavian Journal of Economics*, 123(4):1261–1291.

Hosszejni, D. and Kastner, G. (2019). Modeling univariate and multivariate stochastic volatility in r with stochvol and factorstochvol. *arXiv preprint arXiv:1906.12123*.

Neal, R. M. (2003). Slice sampling. *The Annals of Statistics*, 31(3):705–767.

Nieto-Barajas, L. E., Müller, P., Ji, Y., Lu, Y., and Mills, G. B. (2012). A time-series DDP for functional proteomics profiles. *Biometrics*, 68(3):859–868.