Markov Switching Panel with Endogenous Synchronization Effects *

Komla M. Agudze[†] Monica Billio[‡]
Roberto Casarin[‡] Francesco Ravazzolo[§¶]

[†]World Bank Group, Washington, US

[‡]Ca' Foscari University of Venice, Italy

[§]Free University of Bozen-Bolzano, Italy

[¶]BI Norwegian Business School, Norway

March 7, 2021

Abstract This paper introduces a new dynamic panel model with multi-layer network effects. Series-specific latent Markov chain processes drive the dynamics of the observable processes, and several types of interaction effects among the hidden chains allow for various degrees of endogenous synchronization of both latent and observable processes. The interaction is driven by a multi-layer network with exogenous and endogenous connectivity layers. We provide some theoretical properties of the model, develop a Bayesian inference framework and an efficient Markov Chain Monte Carlo algorithm for estimating parameters, latent states, and endogenous network layers. An application to the US-state coincident indicators shows that the synchronization in the US economy is generated by network effects among the states. The inclusion of a multi-layer network provides a new tool for measuring the effects of the public policies that impact the connectivity between the US states, such as mobility restrictions or job support schemes. The proposed new model and the related inference are general and may find application in a wide spectrum of datasets where the extraction of endogenous interaction effects is relevant and of interest.

Keywords: Bayesian inference, interacting Markov chains, multi-layer networks, panel Markov-switching.

*Corresponding author: `Francesco.Ravazzolo@unibz.it` (F. Ravazzolo). Address: Free University of Bozen-Bolzano, Piazza dell'Universita' 1, 39100 Bozen-Bolzano Italy. Tel.: +39 0471 013133. Fax: +39 0471 013009. Other contacts: `billio@unive.it` (M. Billio) and `r.casarin@unive.it` (R. Casarin). The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank Group, its International Finance Corporation (IFC) and its Executive Board.

JEL code: C11, C13, C15, C23, C55.

1 Introduction

This paper introduces a new Markov-switching model for panel data. The dynamics of each series is driven by its own hidden Markov chain process, which interacts with the other chains of the panel within a multi-layer network. One of the most interesting theoretical properties of the proposed model is that the network interaction effects naturally allow for the endogenous synchronization of the time series, irrespective of the exogenous, or endogenous nature, of the interaction network. Moreover the multi-layer structure of the networks allows us to specify or identify different channels of interaction. The model is well suited to fitting complex dynamics and extracting the nonlinear dependence structures present in panel data. Below, we provide a more detailed discussion of the contributions of this paper in terms of modelling, inference and application.

Markov switching (MS) models have been extensively used in macroeconomics and finance to extract the different phases, or regimes, of the market. Originally, MS models were applied to univariate series Hamilton (1989), or to small sets of series, and the hidden Markov chains were assumed with constant transition probabilities, thus not varying over time. These assumptions have been challenged by recent literature. The use of large databases has been shown to be very important in achieving better forecasting Bańbura et al. (2010) and fitting Stock and Watson (2014) performances. Secondly, time-varying MS processes provide more accurate fitting. This is particularly true when different units of the panel (e.g. countries or states) are grouped together Kaufmann (2010) and common factors drive the transition, (e.g., see Kaufmann, 2015; Billio et al., 2016). So far, modelling solutions have focused on medium-sized panels with series-specific chains driven by a common factor (e.g., see Billio et al., 2016; Casarin et al., 2018)), or on reducing the number of latent chains by assuming a clustering model for the observed series (e.g., see Hamilton and Owyang, 2012; Kaufmann, 2015; Francis et al., 2017; Owayng et al., 2019). We expand upon this literature in various directions.

First, we assume dependent hidden Markov chains while preserving tractability in large-sized panels. Modelling dependence through interaction effects has been successfully used in many contexts, such as in random fields, graphical models and MS literature (e.g., see Brémaud, 2013). In this paper, we provide a new general model for Markov chains interaction based on network theory. See Allen and Babus (2009) for a review of network models in economics and finance. The interaction of the chains on a network produces endogenous synchronization effects with time-varying clustering of the series. This model feature expands upon the literature that assumes static clustering effects. The tractability is preserved thanks to a linear

parametrization of the transition probabilities and to the use of networks to represent the dependence structure. Linear parametrization has several advantages. The first advantage relates to inference aspects. The multivariate logistic transformation, widely used in the literature, implies a non-linear transformation of the parameters which makes the inference task more difficult. In a Bayesian setting, non-linear parametrization can lead to the poor performance of the Markov Chain Monte Carlo (MCMC) sampler used for posterior approximation (e.g., see Scott (2011)). Our model, on the other hand, is not exposed to these inferential issues. The drawback of linear parametrization concerns the constraints one needs to introduce to the parameters, but these constraints can be easily handled through the use of standard prior distributions. The second advantage is that linear parametrization allows us to provide some of the theoretical properties of the multiple-chain model under a broader class of interaction mechanisms, including idiosyncratic, local and global interactions (Föllmer and Horst (2001)).

Secondly, the model allows for multiple interaction effects thanks to the use of multi-layer networks. The network layers are defined as a set of nodes representing the unit-specific chains, and a set of edges representing pairs of chains. Each edge encodes a dependence relationship between panel units (e.g., the US states) and each layer corresponds to different dependence structures among the chains (e.g., the business cycles of the US states). We consider both exogenous layers observed by the researcher, such as geographical, sectoral and economic proximity of the units, or endogenously estimated from a set of node- and edge-specific covariates. The specification of the layers depends on the researcher's understanding of the economic and social relationships (e.g., information, trade, labor and financial interactions) among the economic actors and of their effects on the economic activity. A model selection procedure could be applied to choose the most relevant layers. Another feature of our model is the chain interaction mechanism. We assume that the chains can interact either globally or locally. In particular, a global interaction effect assesses the importance of common movements in the panel, while a local one captures the co-movements of a chain with other chains, which are directly, or indirectly, connected to the layer. The increased data availability allows researchers to employ multiple-layer networks and calls for the development of new econometric tools for networks. To the best of our knowledge, this is the first model with Markov switching processes interacting within a multi-layer network. In this sense, this study contributes to the recent, and expanding, stream of literature on network econometrics (e.g. Billio et al. (2012); Ahelegbey et al. (2016); Bianchi et al. (2019); Diebold and Yilmaz (2014, 2015)).

Another contribution of this paper consists in the proposal of a new efficient MCMC algorithm for the posterior approximation based on the Metropolis-adjusted Langevin (MALA) sampling method (Girolami and Calderhead (2011)), which exploits the information on the gradient of the target distribution. This method

has been successfully applied to many fields, such as statistics, physics and recently even in econometrics by Burda and Maheu (2013), Burda (2015) and Virbickaite et al. (2015).

Finally, we contribute to the study of US regional business cycles. We collect monthly coincident indices from 50 US states. With respect to previous studies (see Hamilton and Owyang (2012), Camacho and Leiva-Leon (2014), Leiva-Leon (2014)), we are able to identify the role of the global and local interaction factors in the cycle co-movements. Also, we shed light on the role of the exogenous and endogenous channels of transmission of shocks across the US states and show that the network interaction effects play a key role in making slowdowns and recessions deeper and longer, contrary to what is predicted by an aggregate index. We find evidence that not only direct, but also indirect, network effects contribute to the spread of recession and expansion across states.

The remainder of the paper is structured into seven sections. Section 2 describes the panel Markov switching model with interacting chains and the regime switching transition probabilities. In section 3, we discuss some properties of the proposed model and present examples of the model outputs. In section 4 we provide the Bayesian estimation procedure. Section F studies the model in simulation exercises and section 5 applies it to regional US business cycles. Section 6 concludes.

2 Panel Markov switching with interacting chains

Denote with $\{y_{it}\}$, $t = 1, \ldots, T$ the i-th time series, $i = 1, \ldots, n$ of our panel MS model with interacting hidden Markov chains (PMS-IC). We assume each series follows a conditionally linear and Gaussian process with mean and variance driven by a unit-specific Markov chain s_{it}, $t = 1, \ldots, T$ with values in the finite set $\{0, \ldots, K-1\}$.

The measurement equation is written as:

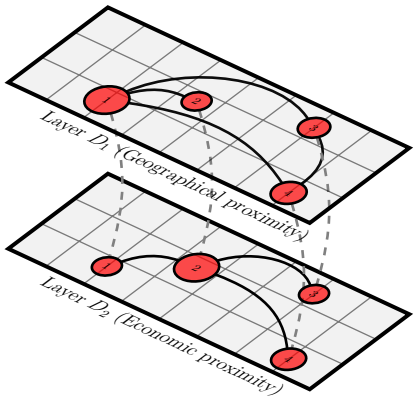
$$y_{it} = \begin{cases} \mu_{i1} + \sigma_{i1}\varepsilon_{it}, \quad \varepsilon_{it} \overset{iid}{\sim} \mathcal{N}(0,1) \quad \text{if } s_{it} = 0 \\ \vdots \\ \mu_{iK} + \sigma_{iK}\varepsilon_{it}, \quad \varepsilon_{it} \overset{iid}{\sim} \mathcal{N}(0,1) \quad \text{if } s_{it} = K-1 \end{cases} \tag{1}$$

where K represents the number of unobserved latent regimes and the symbol $\mathbb{I}_A(X)$ is the indicator function, which takes value 1 if $X \in A$ and 0 otherwise.

The $(K \times K)$ transition matrix P_{it} of the i-th chain is time-varying and has l-th row and k-th column element $p_{it,lk}$ defined as:

$$p_{it+1,lk} = \mathbb{P}\left(s_{it+1} = l | s_{it} = k, \mathbf{s}_{-i,t}\right) \tag{2}$$

which represents the conditional probability that unit i moves to the regime $l \in \{1, \ldots, K\}$ at time $t+1$. $\mathbf{s}_t = (s_{1t}, s_{2t}, \ldots, s_{nt}) \in \mathcal{S}$ denotes a state vector



$$A_1 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad W_{1,2} = \begin{pmatrix} 3 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad W_{2,2} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 3 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

Figure 1: Illustration of 2-layer networks $\mathcal{G} = (V, D_1, D_2)$, with four nodes in the vertex set $V = \{1, 2, 3, 4\}$. Left: the graphical representation with nodes (circles), edges (solid lines) and layers (different hyperplanes). The size of a circle is proportional to the node degree, that is the number of incident edges. Centre: network representation with binary matrices A_r. Right: matrices $W_{r,q}$ those elements are the number of walks and cycles of length $q = 2$ between pairs of nodes in a layer $r = 1, 2$.

with all chain states at time t, with $\mathcal{S} = \{0, \ldots, K-1\}^n$ and $\mathbf{s}_{-i,t} = \{s_{jt}, j = 1, \ldots, n; \quad j \neq i\}$ a state vector with the states of all chains but the i-th chain.

2.1 Chains interaction on multi-layer networks

We represent different connectivity channels between units with a multi-layer network defined by an ordered sequence of sets $\mathcal{G} = (V, D_1, \ldots, D_N)$. The vertex set $V = \{1, \ldots, n\}$ represents the units of the panel, each edge set $D_r \subset V \times V$ represents the unit pairs $\{i, j\}$, $i, j \in V$, which are connected in the layer r. Each connectivity layer corresponds to a different shock transmission channel between panel units and may either be exogenous or endogenous. The connectivity structure of a layer D_r is encoded by a binary matrix, called adjacency matrix, A_r, with (i, j)-th entry given by

$$a_{r,ij} = \begin{cases} 1 & \text{if } \{i, j\} \in D_r \\ 0 & \text{if } \{i, j\} \notin D_r \end{cases}$$

Figure 1 provides an illustration of a 2-layer network with layers of geographical and economic proximity of panel units (panel a) and the corresponding adjacency matrices (panel b).

The network representation of the interaction channels allows for either direct or indirect connections. We denote with $V_{rq}(i) \subset V$ the sets of nodes directly ($q = 1$) or indirectly ($q > 1$) connected to the node i in the layer D_r, i.e. $V_q(i) = \{j \in$

$V | \exists\ D(i,j) \subset D_r\}$ where $D(i,j) = (\{j_0, j_1\}, \{j_1, j_2\}, \ldots, \{j_{q-1}, j_q\}, \{j_q, j_{q+1}\})$ is the sequence of connections, i.e. a walk on the graph, joining the unit $j_0 = i$ to the unit $j_{q+1} = j$ in the layer D_r. If shocks propagate at a higher frequency than the observation frequency, then the indirect connections do matter. For example, if unit 1 is connected to unit 2 and 2 is connected to 3, and we observe the variables of interest at a monthly frequency, we can expect that a shock originating from unit 3 and spreading faster than monthly, e.g. at a weekly frequency, will hit unit 1 through unit 2 in the same month. In our model we define q-th as the order interaction effect between two nodes if there exists at least one walk of length q joining the two nodes. In formulas, let $W_{r,q} = (A_r)^q$, then $i, j \in V_q(i)$ if the (i,j)-th entry of $W_{r,q} > 0$. See panel (c) of Figure 1 for an illustration.

We assume the unit-specific Markov chains interact on a multi-layer network and propose the following linear parametrization of the transition probability:

$$p_{it+1,lk} = \alpha p_{lk} + \sum_{r=1}^{N} \sum_{q=1}^{M} \beta_{rq} m_{irq,k}\left(\mathbf{s}_t\right) + \gamma m_k\left(\mathbf{s}_t\right) \tag{3}$$

where $p_{lk} \in [0,1]$ represents the fixed transition probability, $m_{irq,k}\left(\mathbf{s}_t\right) \in [0,1]$ the local interaction factor and $m_k\left(\mathbf{s}_t\right) \in [0,1]$ the global interaction factor, which is constant across units. The following constrains exist on the interaction parameters: $0 < \alpha \leq 1$, $0 \leq \beta_{rq} < 1$, $r = 1, \ldots, N$, $q = 1, \ldots, M$, $0 \leq \gamma < 1$ and $\alpha + \sum_{r=1}^{N} \sum_{q=1}^{M} \beta_{rq} + \gamma = 1$ and on the time-invariant transition probabilities: $p_{l1} + \ldots + p_{lK} = 1\ \forall i, l$ are sufficient conditions for the positivity of the transition probability.

The fixed transition probability p_{lk} determines the long-run behavior of the processes in our PMS-IC (see Section 3) and if the parameter α is equal to 1 then we obtain a Markov-switching panel model with time-homogeneous transition common to the panel units.

The local interaction factor $m_{irq,k}\left(\mathbf{s}_t\right)$ reflects the connectivity of the unit i in the layer r and is given by the proportion of neighboring units, i.e. the elements of $V_{rq}(i)$, which are in the state k at time t:

$$m_{irq,k}\left(\mathbf{s}_t\right) = \frac{1}{|V_{rq}(i)|} \sum_{j \in V_{rq}(i)} \mathbb{I}_{\{k\}}(s_{jt}) \tag{4}$$

If a unit i has in its neighbor $V_{rq}(i)$ a high proportion of units in regime k at time t, then it will have a larger probability of transiting to regime k at time $t+1$. The local interaction factor with $q > 1$ is a non-linear term which accounts for shocks transmission due to indirect connections.

The global interaction factor $m_k\left(\mathbf{s}_t\right)$ is given by the proportion of chains in

regime k at time t that is:

$$m_k(\mathbf{s}_t) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}_{\{k\}}(s_{jt}) \tag{5}$$

If a large proportion of units are affected by a global shock at time t, it is highly likely that the other units, including unit i, will be affected by the shock.

In a macro perspective, our PMS-IC has the potential to analyze the co-movement of regional business cycles. It does not only help to characterize the unit-specific cycles but also shows the importance of global and local factors (global and local interaction), and of the long-term business cycle (fixed transition).

2.2 Endogenous network interaction effects

For the endogenous layers of the multi-layer network, we specify a random network model with node-specific covariates. Without loss of generality, in the following we assume the first layer D_1 is endogenous and the remaining are exogenous. In the random layer, the edge existence is given by the Bernoulli variable $a_{1,ij}$ which is equal to 1 if $\{i,j\} \in D_1$ and 0 otherwise. The edge probability $\mathbb{P}(a_{1,ij}=1) = q_{ij}$ given by

$$q_{ij} = \frac{\exp(g(\mathbf{x}_i,\mathbf{x}_j)'\boldsymbol{\varphi})}{1+\exp(g(\mathbf{x}_i,\mathbf{x}_j)'\boldsymbol{\varphi})} \tag{6}$$

where $g(\mathbf{x}_i,\mathbf{x}_j) \in \mathbb{R}^n$ is the distance function between the n-dimensional feature vectors $\mathbf{x}_i = (x_{1i},\ldots,x_{ni})$, which is a map from $\mathbb{R}^n$ to $\mathbb{R}^r$, and $\boldsymbol{\varphi} \in \mathbb{R}^r$ is a parameter vector. In this model, the ability of a vertex i to attract edges in the network (node fitness) varies across nodes and is driven by the node-specific feature vector $\mathbf{x}_i$ and the fitness parameter $\boldsymbol{\varphi}$. The larger the value of the l-the element φ_l, the higher the fitness of the nodes is with respect to the l feature.

Following Holmes and Held (2006) and Hamilton and Owyang (2012), we use a convenient data augmentation representation of the edge indicator and introduce the latent variables

$$\zeta_{ij} = g(\mathbf{x}_i,\mathbf{x}_j)'\boldsymbol{\varphi} + 2\psi_{ij}\eta_{ij}, \quad \eta_{ij} \overset{iid}{\sim} \mathcal{N}(0,1), \tag{7}$$

$$\psi_{ij} \overset{iid}{\sim} f(\psi_{ij}), \quad f(\psi_{ij}) = 8\sum_{j=1}^{\infty}(-1)^{j+1}j^2\psi\exp(-2j^2\psi) \tag{8}$$

where $f(\psi_{ij})$ is the Kolmogorov-Smirnov distribution. The probability in Eq. (6) corresponds to the distribution of the random variable $a_{ij} = \mathbb{I}(\zeta_{ij} > 0)$.

3 Properties of the model

3.1 Theoretical properties

The use of linear parametrization for the time-varying transition probabilities allows us to study the convergence of the ergodic probabilities. Let $\mathcal{S} = \{0, 1, \ldots, K-1\}^n$ be the finite set of all configurations of $\mathbf{s}_t = (s_{it})_{1 \leq i \leq n}$ with $s_{it} \in \{0, 1, \ldots, K-1\}$, $i = 1, \ldots, n$. The following result provides a characterization of the macroscopic dynamic of the set of Markov chains for diverging n and shows that the interacting transition kernel defined in equation (3) is generating a deterministic sequence of empirical averages. These quantities can be used to find the limiting behavior of the set of chains as t tends to infinity and to give an interpretation of the parameters of the transition probabilities.

The relationship between the local interactions factor and the global one is summarized by the following. Let us define the empirical averages:

$$\overline{m}\left(\mathbf{s}_t\right) = \left(\overline{m}_0\left(\mathbf{s}_t\right), \ldots, \overline{m}_{K-1}\left(\mathbf{s}_t\right)\right)'$$

and the proportion of regime in some finite neighborhood $V_{rq}(i)$ of i:

$$\overline{m}_{irq}\left(\mathbf{s}_t\right) = \left(\overline{m}_{irq,0}\left(\mathbf{s}_t\right), \ldots, \overline{m}_{irq,K-1}\left(\mathbf{s}_t\right)\right)'$$

where

$$\overline{m}_k\left(\mathbf{s}_t\right) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{k\}}(s_{it}), \quad k = 0, 1, \ldots, K-1$$

$$\overline{m}_{irq,k}\left(\mathbf{s}_t\right) = \frac{1}{|V_{rq}(i)|} \sum_{j \in V_{rq}(i)} \mathbb{I}_{\{k\}}(s_{jt}), \quad k = 0, 1, \ldots, K-1$$

with s_{it} being a Markov chain with transition probabilities

$$\pi_i(s_{it+1} = k | \mathbf{s}_t) = \alpha p_{s_{it}k} + \sum_{r=1}^{N} \sum_{q=1}^{M} \beta_{rq} m_{irq,k}\left(\mathbf{s}_t\right) + \gamma m_k\left(\mathbf{s}_t\right)$$

where $\sum_{k=0}^{K-1} \pi_i(s_{it+1} = k | s_t) = 1$. Then the sequence of empirical averages of $\overline{m}_i\left(\mathbf{s}_t\right)$ converge to $\overline{m}\left(\mathbf{s}_t\right)$.

The theoretical relationship between the global interaction factor and the fixed transition probability matrix is given in Proposition 3.1.

Proposition 3.1. *Let* $\mathcal{S}_1 = \{\mathbf{s}_t \in \mathcal{S} | \exists m_{t+1}\}$ *and* $\Pi(\cdot|\mathbf{s}_t) = \prod_{i=1}^{\infty} \pi_i(\cdot|\mathbf{s}_t)$ *the product kernel of the population of chains, then*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{k\}} s_{it+1} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \pi_i\left(s_{it+1} = k | \mathbf{s}_t\right) \quad \Pi(\cdot|\mathbf{s}_t) - a.s., \tag{9}$$

the sequence of empirical averages almost surely satisfies the recurrence relation

$$\overline{m}_k\left(\mathbf{s}_{t+1}\right) = \alpha \sum_{j=0}^{K-1} p_{jk}\overline{m}_j\left(\mathbf{s}_t\right) + (1-\alpha)\overline{m}_k\left(\mathbf{s}_t\right) \tag{10}$$

and the global interaction process $\{\overline{m}\left(\mathbf{s}_t\right)\}_{t\in N}$ *converges almost surely to the unique invariant probability of the fixed transition probability matrix*

$$P = \begin{pmatrix} p_{01} & \cdots & p_{K-1,0} \\ \vdots & \ddots & \vdots \\ p_{0,K-1} & \cdots & p_{K-1,K-1} \end{pmatrix}$$

Proof. See Appendix A.1.

The convergence of the Markov chain process of the population of time series with time-varying transition probabilities is present in the following.

Proposition 3.2. *The process* $\{\mathbf{s}_t\}_{t\in N}$ *converges in law to the unique product kernel*

$$\Pi_{\overline{m}}(\cdot|\mathbf{s}_t) = \prod_{i=1}^{\infty} \pi_{\overline{m}i}(\cdot|\mathbf{s}_t)$$

Proof. See Appendix A.2.

3.2 Numerical illustration

We provide a qualitative description of the synchronization effects of our PMS-IC model through some simulated examples. Consider a PMS-IC model with $n = 50$ time series of length $T = 5000$, one layer ($N = 1$), only linear interaction effects ($M = 1$), and three regimes ($K = 3$). Figure 2 shows the heat-maps of a realization of the dynamic panel (50 trajectories of length 5000 each) with weak ($\alpha = 0.7$, $\beta_{11} = 0$ and $\gamma = 0.3$, top chart) and strong ($\alpha = 0.3$, $\beta_{11} = 0$ and $\gamma = 0.7$, bottom chart) global interaction effects. For further details on the parameter settings, see $\mathcal{M}_2$ and $\mathcal{M}_3$ in Table E.1 of the Supplementary Material. Colors represent the value of the series, where white, green and red indicate expansion, moderate expansion and recession regimes, respectively. If two chains are in the same regime, then their trajectories appear with the same color on the heat-map. In the strong interaction case, vertical-colored bars suggest that a large proportion of chains in the population are in the same regime. The results indicate that the level of synchronization increases with the value of both the global and the local interaction parameters. Our PMS-IC model allows for various degrees of synchronization and for time-varying transition probabilities. Also, the convergence of the ergodic averages to

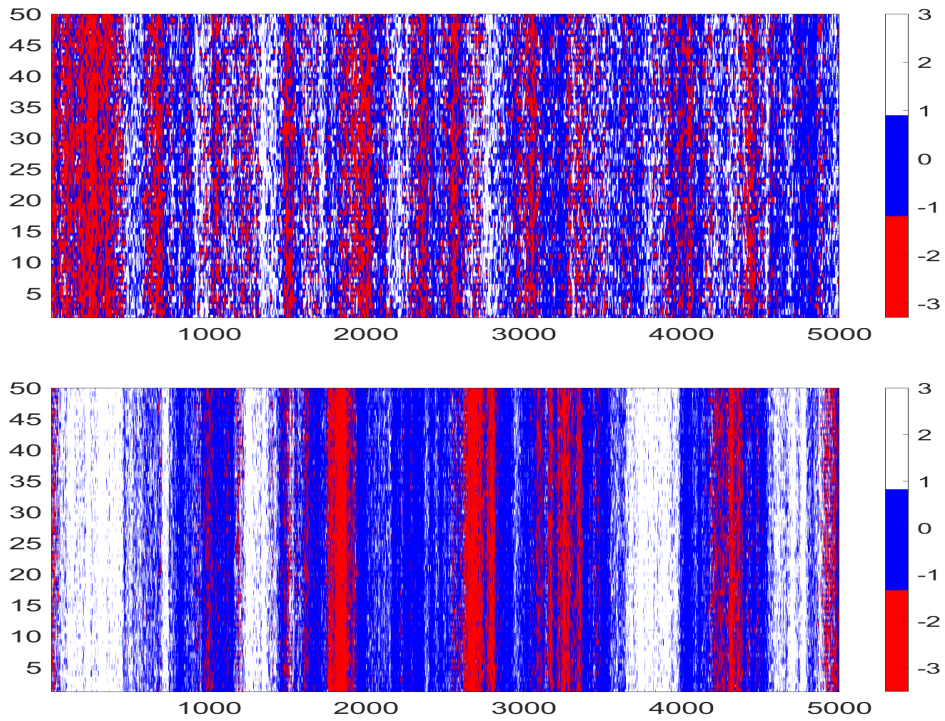


Figure 2: Endogenous synchronization effects. Heat-map of 50 trajectories (vertical axis) of 5000 observations each (horizontal axis) for the weak global interaction model $\mathcal{M}_2$ ($\alpha = 0.3$, $\beta_{11} = 0$ and $\gamma = 0.7$, top plot) and the strong global interaction model $\mathcal{M}_3$ ($\alpha = 0.7$, $\beta_{11} = 0$ and $\gamma = 0.3$, bottom plot). Colors represent the value of the series, with white, green and red indicating expansion, moderate expansion and recession regimes, respectively.

the ergodic probabilities is reached at different speeds and is usually faster for strong global interaction, than for strong local interaction (see Figure E.1-E.5 in the Supplementary Material).

We study the impact of the local connectivity structure of the networks on the endogenous synchronization effects and provide a Monte Carlo estimate of the synchronization level for different values of the local and global interactions parameters. We use the bivariate concordance index of Harding and Pagan (2002)

$$c_{i,j} = \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}_{\{0\}}(s_{it} - s_{jt}) \tag{11}$$

to assess the impact of the local and global parameters on the synchronization of chains. This index describes the fraction of times that two Markov chains, i and j, are in the same state. The relationship between the local and global interaction

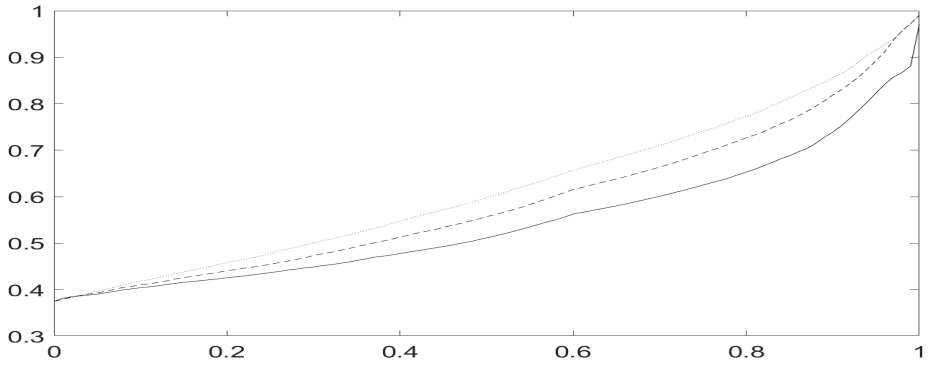


Figure 3: Chains synchronization level, measured by $\overline{c}$ (vertical axis), as a function of the local interaction parameter β (horizontal axis) for different neighbors size: 4 (blue line), 16 (orange line) and 24 (yellow line) units

parameters and the level of synchronization among the chains can be measured by:

$$\overline{c} = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} c_{i,j} \tag{12}$$

which is in the unit interval. The closer the value of $\overline{c}$ is to one, the greater the extent of synchronization within the panel of series.

Three neighborhood systems are designed for a model with only local interactions. In each model, the neighbors are made up of 4, 16 and 24 units. The rate of synchronization increases with the size of the neighborhood and the value of β, as well as γ (different lines in Figure 3).

4 Bayesian inference

4.1 Likelihood function and prior distributions

Let $\boldsymbol{\theta} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K, \boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_K, vec(P), vec(Q), \alpha, \boldsymbol{\beta}, \gamma, \boldsymbol{\varphi})$ be the vector of parameters with $\boldsymbol{\mu}_l = (\mu_{1l}, \ldots, \mu_{nl})$, $\boldsymbol{\sigma}_l = (\sigma_{1l}, \ldots, \sigma_{nl})$, $Q = (q_{11}, \ldots, q_{1n}, \ldots, q_{n1}, \ldots, q_{nn})$ and $\boldsymbol{\beta} = (\beta_{11}, \ldots, \beta_{1M}, \ldots, \beta_{N1}, \ldots, \beta_{NM})$. Let us define the latent state variable $\xi_{k,it} = \mathbb{I}_{\{k\}}(s_{it})$, which takes value 1 if the state of the chain s_{it} is k and 0 otherwise, then the model in Eq. 1 can be written as

$$y_{it} = \sum_{k=1}^{K} \xi_{k,it}\mu_{ik} + \varepsilon_{it} \sum_{k=1}^{K} \xi_{k,it}\sigma_{ik}, \quad \varepsilon_{it} \overset{iid}{\sim} \mathcal{N}(0,1) \tag{13}$$

Let $\mathbf{y} = (y_1, \ldots, y_T)$ be the collection of all observations, and $\Psi = (\psi_{11}, \ldots, \psi_{1n}, \ldots, \psi_{n1}, \ldots, \psi_{nn})$ and $Z = (\zeta_{11}, \ldots, \zeta_{1n}, \ldots, \zeta_{n1}, \ldots, \zeta_{nn})$ the two matrices of latent variables for the latent network $\mathcal{G}$ with connectivity structure given by the adjacency matrix $A = (a_{11}, \ldots, a_{1n}, \ldots, a_{n1}, \ldots, a_{nn})$. Let $\Xi = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_T)$ the $K \times nT$ matrix of latent variables for the panel of hidden Markov chains, where $\Xi_t = (\boldsymbol{\xi}_{1t}, \ldots, \boldsymbol{\xi}_{nt})$ is the $K \times n$ matrix of latent variables at time t with $\boldsymbol{\xi}_{it} = (\xi_{1,it}, \ldots, \xi_{K,it})$ where $\xi_{k,it} = \mathbb{I}_k(s_{it})$.

By using the sequential factorization of the likelihood, the complete likelihood of the PMS-IC model is:

$$L(\mathbf{y}, \Xi, A, Z, \Psi \mid \boldsymbol{\theta}) = f(\mathbf{y}|\Xi, \boldsymbol{\theta}) f(\Xi|A, \boldsymbol{\theta}) f(A|Z) f(Z|\Psi, \theta) f(\Psi) \tag{14}$$

where the factors in the complete-data likelihood are:

$$\begin{aligned}
f(\mathbf{y}|\Xi, \boldsymbol{\theta}) &= \prod_{t=1}^{T} \prod_{i=1}^{n} f(y_{it}|\xi_{it}, \boldsymbol{\theta}) \\
f(\Xi|A, \boldsymbol{\theta}) &= \prod_{t=1}^{n} \prod_{i=1}^{n} f(\xi_{it}|\Xi_{t-1}, A, \boldsymbol{\theta}) \\
f(A|Z) &= \prod_{i=1}^{n} \prod_{j=1}^{n} \mathbb{I}(\zeta_{ij} > 0)^{a_{ij}} (1 - \mathbb{I}(\zeta_{ij} > 0))^{1-a_{ij}} \\
f(Z|\Psi, \boldsymbol{\theta}) &= \prod_{i=1}^{n} \prod_{j=1}^{n} (2\pi\lambda_{ij})^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\lambda_{ij}} (\zeta_{ij} - g(\mathbf{x}_i, \mathbf{x}_j)\boldsymbol{\varphi})^2 \right\} \\
f(\Psi) &= \prod_{i=1}^{n} \prod_{j=1}^{n} f(\psi_{ij})
\end{aligned}$$

where $\lambda_{ij} = 4\psi_{ij}^2$ and $f(\psi_{ij})$ is give in Eq. (8) and

$$\begin{aligned}
f(y_{it}|\Xi, \boldsymbol{\theta}) &= \prod_{l=1}^{K} (2\pi\sigma_{il}^2)^{-\frac{\xi_{l,it}}{2}} \exp\left\{ -\frac{\xi_{l,it}}{2\sigma_{il}^2} (y_{it} - \mu_{il})^2 \right\} \\
f(\boldsymbol{\xi}_{it}|\Xi_{t-1}, A) &= \prod_{l=1}^{K} \prod_{k=1}^{K} p_{it,lk}^{\xi_{l,it}\xi_{k,it-1}} \\
p_{it,lk} &= (\alpha p_{lk} + \sum_{r=1}^{N} \sum_{q=1}^{M} \beta_{rq} m_{irq,k}(\mathbf{s}_t) + \gamma m_{t,k})^{\xi_{l,it}\xi_{k,it-1}}.
\end{aligned}$$

In order to complete the specification of the Bayesian model, we discuss the prior choice. We consider conjugate prior distributions based on proper distributions. We assume independent priors for the unit-specific and the common parameters:

$$\mu_{il} \sim \mathcal{N}(m_{il}, s_{il}^2) \tag{15}$$

$$\sigma_{il}^2 \sim \mathcal{IG}(a_{il}, b_{il}) \quad (16)$$

$$(p_{l1}, \ldots, p_{lK}) \sim \mathcal{D}ir(v_1, \ldots, v_K) \quad (17)$$

$$(\alpha, \boldsymbol{\beta}', \gamma) \sim \mathcal{D}ir(r_1, \mathbf{r}'_{1+NM}, r_{2+NM}) \quad (18)$$

$$\boldsymbol{\varphi} \sim \mathcal{N}(\mathbf{m}, S) \quad (19)$$

with $l = 1, \ldots, K$ and $i = 1, \ldots, n$, where $\mathcal{IG}(a, b)$ denote the inverse gamma distribution with shape and rate parameters a and b, respectively and $\mathcal{D}ir(d_1, \ldots, d_K)$ the K dimensional Dirichlet distribution with parameters: $d_1, \ldots, d_K$.

One of the main problems of Bayesian analysis using Markov switching processes is the non-identifiability of the parameters. That is, the posterior distributions of the parameters of resulting Markov switching model is invariant to permutations in the labelling of the parameters, if this latter follows exchangeable priors. Consequently, the marginal posterior distributions for the parameters are identical for each switching component and the symmetry of the posterior distributions affect the MCMC simulation and the interpretation of the labels switch. For more details about the effects that label switching and non-identification have on the results of a MCMC-based Bayesian inference, see, among others, Celeux (1998), Frühwirth-Schnatter (2001), and Frühwirth-Schnatter (2006). One way to address the label switching problem is to consider the permutation sampler proposed by Frühwirth-Schnatter (2001) under some specific conditions. Another alternative is to impose identification constraints on the parameters. This practice is widely used in macroeconomics because it is naturally related to the interpretation of the different states (e.g. recession and expansion) of the business cycle. We follow the latter approach and impose the identification restrictions $\mu_{i1} \leq \mu_{i2} \leq, \ldots, \leq \mu_{iK}$.

4.2 Posterior simulation

The joint posterior distribution is

$$\pi(\boldsymbol{\theta}, \Xi, A, Z, \Psi \mid \mathbf{y}) \propto L(\mathbf{y}, \Xi, A, Z, \Psi \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta}) \quad (20)$$

where $\pi(\boldsymbol{\theta})$ is the joint prior distribution defined in Eq. (15)-(19). Since the posterior is not tractable, we approximate the posterior quantities of interest by applying a Monte Carlo method. We develop a Gibbs sampling algorithm based on the full conditional posterior distributions. The derivation and the details of the conditional distributions are provided in Appendix B. The model in equation (1) is estimated by adapting the multi-move Gibbs-sampling procedure for the Bayesian estimation of Markov switching models presented in Frühwirth-Schnatter (2006). At the d-th iteration, the Gibbs sampler iterates over the following steps for the parameters.

1. Draw $(\alpha, \boldsymbol{\beta}', \gamma)^{(d)}$, from $f(\alpha, \boldsymbol{\beta}', \gamma \mid \mathbf{y}, \Xi^{(d-1)}, A^{(d-1)}, P^{(d-1)})$.

2. Draw $(p_{l1}, \ldots, p_{lK})^{(d)}$ from $f(p_{l1}, \ldots, p_{lK} \mid \mathbf{y}, \Xi^{(d-1)}, A^{(d-1)}, (\alpha, \beta, \gamma)^{(d)})$, $l = 1, \ldots, K$.

3. Draw $\mu_{il}^{(d)}$,from $f(\mu_{il} \mid \mathbf{y}, \Xi^{(d-1)}, \sigma_{il}^{(d-1)})$, $i = 1, \ldots, n$, $l = 1, \ldots, K$.

4. Draw $\sigma_{il}^{(d)}$, from $f(\sigma_{il} \mid \mathbf{y}, \Xi^{(d-1)}, \mu_{il}^{(d)})$ $i = 1, \ldots, n$, $l = 1, \ldots, K$.

5. Draw $\boldsymbol{\varphi}^{(d)}$, from $f(\boldsymbol{\varphi} \mid Z^{(d-1)}, \Psi^{(d-1)})$

We exploit the following factorization for the joint distribution of the latent variables and the network

$$f(A, Z, \Psi \mid \boldsymbol{\theta}, \Xi) \propto f(\Psi \mid Z, A, \boldsymbol{\theta}, \Xi) \int f(Z, \Psi \mid A, \boldsymbol{\theta}, \Xi) d\Psi \int \int f(A, Z, \Psi \mid \boldsymbol{\theta}, \Xi) dZ d\Psi$$
$$\propto f(\Psi \mid Z, A, \boldsymbol{\theta}) f(Z \mid A, \boldsymbol{\theta}) f(A \mid \Xi, \boldsymbol{\theta}) \propto f(\Psi \mid Z, \boldsymbol{\theta}) f(Z \mid A, \boldsymbol{\theta}) f(A \mid \Xi, \boldsymbol{\theta})$$

and use a collapsed Gibbs step

6. Draw $A^{(d)}$, from $f(A \mid \Xi^{(d-1)}, \boldsymbol{\varphi}^{(d)})$

7. Draw $Z^{(d)}$, from $f(Z \mid A^{(d)}, \boldsymbol{\theta}^{(d)})$

8. Draw $\Psi^{(d)}$, from $f(\Psi \mid Z^{(d)}, \boldsymbol{\varphi}^{(d)})$.

In the last step of the Gibbs, the switching indicator variables are sampled

9. Draw $\Xi_{i,1}^{(d)}, \ldots, \Xi_{i,T}^{(d)}$ from $f(\Xi_{i,1}, \ldots, \Xi_{i,T} \mid \mathbf{y}, A^{(d)}, \boldsymbol{\theta}^{(d)})$, $i = 1, \ldots, n$.

Regarding the sampling methods, the standard sampler based on independent proposal distributions poorly estimates the parameter (α, β, γ). A straightforward implementation of the Metropolis-Hastings (MH) algorithm with the prior distribution applied as proposal distribution becomes inefficient, resulting in a high rate of acceptance, followed by poor mixing of the chain. Thus, we apply a Metropolis-adjusted Langevin (MALA) sampling algorithm as an efficient method for solving the issues described above. We simulate $(\alpha, \boldsymbol{\beta}', \gamma)$ from $f(\alpha, \boldsymbol{\beta}', \gamma | Y, \Xi, A, P)$, where the prior is chosen to be Dirichlet $\mathcal{D}ir(\varphi_1, \boldsymbol{\varphi}'_{1+NM}, \varphi_{2+MN})$. Since by definition $(\alpha, \beta, \gamma) \in [0; 1]^3$; when dealing with random walk proposals, we need to use the transformation of $\alpha, \beta_{rq}, \forall\, r, q$ and γ to the real line, which introduces a Jacobian factor to the acceptance probability of the MALA. We assume $\alpha = g(\tilde{\alpha})$, $\beta = g(\tilde{\beta}_{rq})$, $\gamma = g(\tilde{\gamma})$, where $g(x) = 1/(1 + \exp(-x))$ is the logistic transformation. For the MALA, we need the partial derivatives of the complete log-likelihood with respect to the transformed parameters. A similar strategy is applied to get samples from the full conditional distribution of $(p_{l1}, \ldots, p_{lK})$, $l = 1, \ldots, K$. See Appendix D.

The proposal distribution at the n-th iteration of the MALA is given by the following equation

$$\omega^* = \omega^{(n)} + \frac{\epsilon^2}{2} M \nabla_\omega l\left(\omega^{(n)}\right) + \epsilon M^{\frac{1}{2}} z^{(n)} \tag{21}$$

where $\omega = (\tilde{\alpha}, \tilde{\boldsymbol{\beta}}', \tilde{\gamma})'$, $l(\omega) = \log f(\omega|\mathbf{y}, \Xi, A, P)$ is the log-full conditional, ϵ is the integration step and $z \sim \mathcal{N}(0,1)$. M is a preconditioning matrix which helps to circumvent issues that appear when the elements of ω have very different scales, or if they are strongly correlated. $\sqrt{M}$ can be obtained via Cholesky decomposition, such that $M = UU'$ and $\sqrt{M} = U$. The convergence to stationary distribution $p(\omega)$ is ensured by employing a Metropolis acceptance step. The proposal distribution $q(\omega^*|\omega^n)$ is $\mathcal{N}(\mu(\omega^n, \epsilon), \epsilon^2 I)$ with $\mu(\omega^n, \epsilon) = \omega^n + \frac{\epsilon^2}{2}\nabla_\omega p\left(\omega^n\right)$ and acceptance probability given by $\min\{1, p(\omega^*)q(\omega^n|\omega^*)/p(\omega^n)q(\omega^*|\omega^n)\}$. The choice of the preconditioning matrix does not follow any systematic and principled manner. For instance, Christensen et al. (2005) showed that a global level of preconditioning can be inappropriate for the transient phase of the Markov process.

The full conditional distributions of μ_{il}, $\boldsymbol{\varphi}$, and σ_{il} are normal and inverse gamma and can be simulated exactly.

We sample from the full conditional distribution of the adjacency matrix A of the latent network $\mathcal{G}$ by MH, which explores all configurations of the binary entries of A. At the second step of the collapsed Gibbs, we sample from $f(Z, \Psi|A, \boldsymbol{\theta})$ by adapting the sampler of Holmes and Held (2006) and Hamilton and Owyang (2012).

Regarding the allocation variables $\boldsymbol{\xi}_{i,t}$ $i = 1, \ldots, n$ and $t = 1, \ldots, T$, of the hidden Markov chain process, multi-move sampling cannot be directly implemented since the full conditional posterior distribution of the unit-specific allocation variable depends on the other-country allocations. We apply an efficient MH procedure developed in Billio et al. (2016) for a panel of interacting chains, where the candidates of the allocation variables are generated by forward filtering backward sampling (FFBS). We check the effectiveness of the sampling procedure on simulated data (See Appendix F).

5 US states coincident indices

We apply our model to business cycle of US states and network of states . Not all US states are identical and, for example, the US financial crisis has shown that some states were heavily affected by the crisis, e.g. Michigan, while other states were essentially unaffected, such as Texas, which benefited from high oil prices in 2009 and 2010. Furthermore, the US labor force is often considered to be mobile and keen to move from one region to another when economic situations differ across states. This may strengthen the regional effect by creating a network of states that attract

population and other states that lose population. We believe that our model is very well suited to investigating such mechanisms and deriving state and US global country cycles.

We work with the US-state monthly coincident indices datasets produced by the Federal Reserve Bank of Philadelphia (FED), see Crone and Clayton-Matthews (2005). The database covers 50 states in the US and our sample dates from October 1979 to June 2010. An index of business cycle diffusion is available for each state. A global US index is also constructed. The state-level diffusion indices are constructed on the scale -100 to 100, where a negative number is related to the spread of a national recession and a positive number to national expansion. Owyang et al. (2005) apply a Bayesian univariate independent Markov switching model to the same dataset.[1]

We study the importance of the local interactions across the US states by specifying three connectivity settings. The first setting considers the network generated by contiguous states (geographical network). The second setting considers four statistical US regions defined by the United States Census Bureau: the West, the Midwest, the South and the North-east (economic network). We follow Bernile et al. (2017), Bernile et al. (2015) and Garcia and Norli (2012) and use firm-level information based on the 10-K filings on the Securities and Exchange Commission's EDGAR system to identify economic connections among US states.[2] The final setting estimates a network (endogenous network) driven by the same variables of Hamilton and Owyang (2012), namely the share of total state employment accounted for by small firms (x_{1i} in Eq. (6)), manufacturing employment share (x_{2i}), financial activities employment share (x_{3i}), and barrels of oil produced per 100 dollars of state GDP (x_{4i}).

The three networks are used in the local-interaction of three types of model specifications: single-layer with first-order interaction effects ($\mathcal{M}_{11}^{GEO}$,$\mathcal{M}_{11}^{ECO}$, $\mathcal{M}_{11}^{END}$ models), single-layer with up to the second-order interaction effects ($\mathcal{M}_{12}^{GEO}$,$\mathcal{M}_{12}^{ECO}$, $\mathcal{M}_{12}^{END}$ models), and multi-layer with first-order interaction effects ($\mathcal{M}_{11}^{GEO,ECO}$, $\mathcal{M}_{11}^{END,GEO}$, $\mathcal{M}_{11}^{END,ECO}$ models). In addition, we include a specification with only global interactions, which ignores contagion effects between

[1]More information and data are available at `https://philadelphiafed.org/research-and-data/regional-economy/indexes/coincident`.

[2]The federal securities laws require companies issuing publicly traded securities to disclose information on an ongoing basis. Notably, Section 13 and 15(d) of the Securities Exchange Act of 1934 requires companies with more than 10 million dollars in assets and whose securities are held by more than 500 owners to file an annual report (Form 10-K) providing a comprehensive overview of the company's business and financial condition. As in Bernile et al. (2017), we do not make explicit assumptions about the nature of the economic connections but rather count the number of times a US state is cited in items 1, 2, 6 and 7 of the 10-K filings. Such items are design to focus on the firm's economic activities. By comparing information for firms in two different states, we can derive an economic network between the two states.

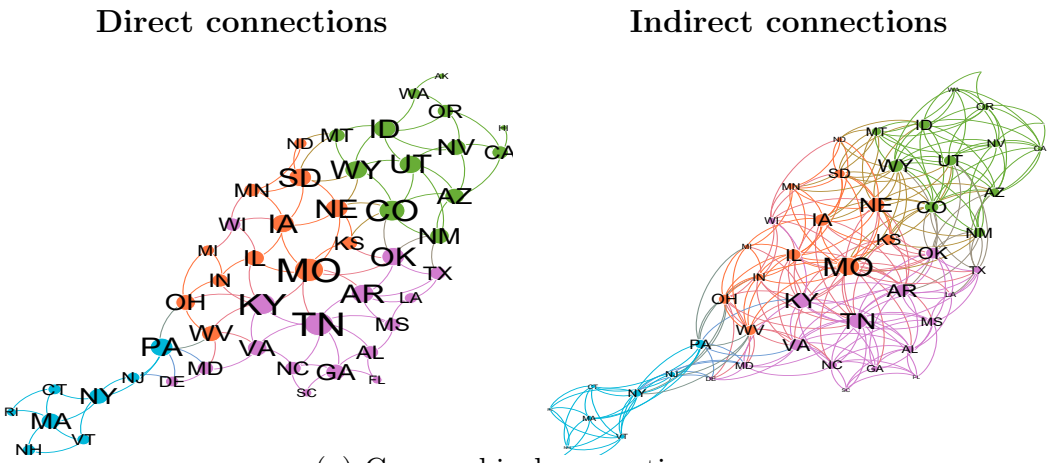
cycles of connected US states ($\mathcal{M}_0$ model).

Figure 4, panels a), b) and c), provides the direct (left) and indirect (right) connectivity structure of the three networks. The geographic and endogenous networks are connected graphs, whereas the economic one is not and presents three connected components. In a connected graph, it is possible to get from every vertex to every other vertex through a series of edges (path). This implies that recessions in a state can propagate to the entire economy only in the geographic and endogenous networks.

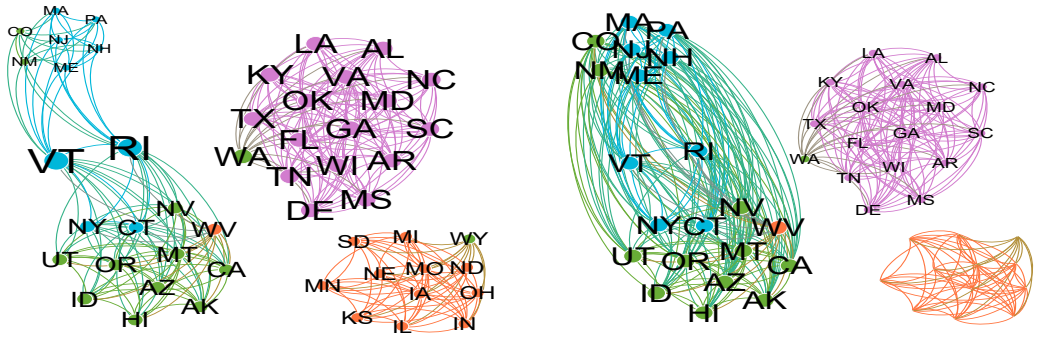
The average number of connections (average degree) in the direct geographical and economic networks are 4.36 and 13.12, respectively; on the contrary, the endogenous network is more dense with an average degree of 25.12. Thus, the propagation of recession and expansion is faster in the endogenous network since the number of direct connections is larger. Nevertheless, the propagation also depends on indirect connections between states. Thus for low-degree networks, such as the geographical and economic ones, indirect connections can become relevant in models, which account for second-order interaction effects ($\mathcal{M}_{12}^{GEO}$,$\mathcal{M}_{12}^{ECO}$). The right column in Fig 4 shows that indirect connectivity differs substantially from the direct ones for the geographical and economic networks, whereas the difference is smaller for the endogenous. For example, following degree centrality (see Table G.1 in Appendix G and thicker circles in Fig. 4), the most central nodes are Missouri (MO) and Tennessee (TN) in the geographical network, and Rhode Island (RI) and Vermont (VT) in the economic network, where they play a key role in connecting the West and North-east regions. In the indirect-connection network (right panel), these central nodes become less central, a new connectivity structure and new most central nodes appear. Similarly, all states in the West and North-east regions appear more connected and central and thus more exposed to contagion effects.

The endogenous network has a substantially different structure than the geographic and economic networks, with several central nodes, such as California (CA), New Hampshire (NH), Montana (MO), Washington (WA), Rhode Island (RI) and Utah (UT). This indicates that the exogenous variables proposed by Hamilton and Owyang (2012) can be used in our framework to capture some relevant contagion channels between US states. As regards the effects of the covariates on the network formation, negative coefficients (small firms employment share (-1.371), manufacturing employment share (-1.828), financial activities employment share (-1.221)) imply that states with similar features and those that are contiguous are more connected. The positive coefficient of the oil production (0.532) implies that oil-producing states, which have different economic features and are not contiguous in the geographical network, are closer in the endogenous network.

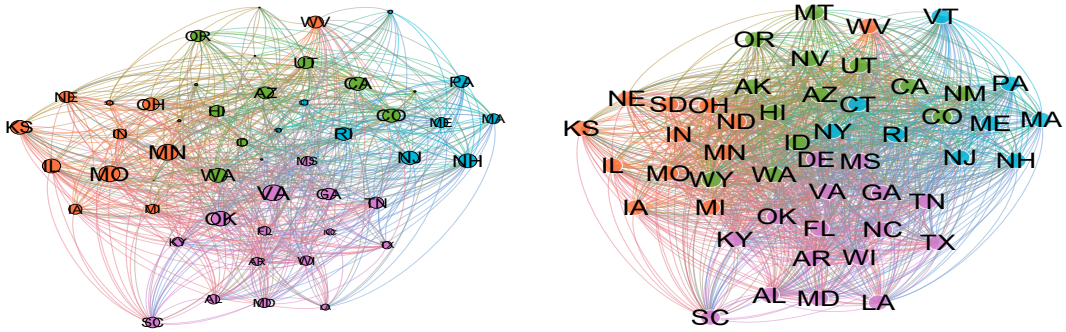
The results in Table 1 indicate that a single-layer specification model with economic and endogenous proximity presents the highest marginal log-likelihoods; models with geographical proximity the second highest, and the model with only



(a) Geographical connections



(b) Economic connections



(c) Endogenous connections

Figure 4: Geographical (a), economic (b) and endogenous (c) networks of direct (left) and 2-step indirect (right) connections. In the network plots: lines indicate connections between pairs of nodes (colored circles); thicker circles and large labels mean that a node has a larger number of connections; colors indicate USCB geographical regions with light blue for North-east, orange for Midwest, pink for South, and green for West. Node labels represent states labels.

(a) First and second order interaction effects

	$m(\mathbf{y})$	α	β_{11}	β_{12}	β_{21}	γ
$\mathcal{M}_0$	$8.797{\cdot}10^2$	0.724				0.276
		(0.712, 0.733)				(0.267, 0.288)
$\mathcal{M}_{11}^{GEO}$	$1.198{\cdot}10^3$	0.159	0.829			0.012
		(0.149, 0.170)	(0.816, 0.841)			(0.004, 0.028)
$\mathcal{M}_{12}^{GEO}$	$3.533{\cdot}10^3$	0.174	0.012	0.801		0.013
		(0.074 ,0.311)	(0.010 , 0.013)	(0.653, 0.915)		(0.004,0.029)
$\mathcal{M}_{11}^{ECO}$	$8.695{\cdot}10^3$	0.028	0.970			0.002
		(0.026, 0.031)	(0.967, 0.973)			(0.001, 0.003)
$\mathcal{M}_{12}^{ECO}$	$1.838\ {\cdot}10^4$	0.026	0.739	0.229		0.006
		(0.026, 0.028)	(0.735, 0.743)	(0.227 ,0.231)		(0.002,0.006)
$\mathcal{M}_{11}^{END}$	$2.220{\cdot}10^4$	0.057	0.936			0.007
		(0.004, 0.233)	(0.751, 0.994)			(0.002, 0.016)
$\mathcal{M}_{12}^{END}$	$1.156{\cdot}10^4$	0.021	0.966	0.001		0.012
		(0.018 ,0.023)	(0.910,0.979)	(0.001,0.002)		(0.01,0.021)

(b) Multi-layer interaction effects

	$m(\mathbf{y})$	α	β_{11}	β_{12}	β_{21}	γ
$\mathcal{M}_{21}^{GEO,ECO}$	$1.767{\cdot}10^4$	0.102	0.098		0.691	0.109
		(0.101 ,0.115)	(0.078,0.112)		(0.611,0.701)	(0.101,0.129)
$\mathcal{M}_{21}^{END,GEO}$	$2.669{\cdot}10^4$	0.005	0.739		0.217	0.039
		(0.001 ,0.020)	(0.7201,0.7518)		(0.203 ,0.236)	(0.023 ,0.057)
$\mathcal{M}_{21}^{END,ECO}$	$2.811{\cdot}10^4$	0.043	0.3211		0.601	0.0349
		(0.028 ,0.061)	(0.305,0.337)		(0.585,0.617)	(0.017,0.052)

Table 1: Comparison of the US states regional data for the PMS-IC models. First column: model label. Second column: marginal log-likelihood ($m(\mathbf{y})$). Remaining columns: posterior mean of β_{rq} $r = 1, \ldots, N$, $q = 1, \ldots, M$ and 95% high posterior density interval (in parenthesis). Panel (a): only global interaction ($\mathcal{M}_0$) and global and local interactions with geographical proximity ($\mathcal{M}_{NM}^{GEO}$), exogenous economic proximity ($\mathcal{M}_{NM}^{ECO}$), endogenous economic proximity ($\mathcal{M}_{NM}^{EN}$), with first ($M = 1$) and second order ($M = 2$) interaction effects and with one layer ($N = 1$). Panel (b): local multilayer interaction ($\mathcal{M}_{NM}^{a,b}$) with first order ($M = 1$) interaction and two layers ($N = 2$) with first layer $a \in \{GEO, ECO\}$ and second layer $b \in \{ECO, END\}$, where GEO is the label of the exogenous geographical proximity graph, ECO the exogenous USCB economic proximity and END the endogenous economic proximity graph.

global interaction gives the lowest marginal likelihood. Among the multi-layer specifications, the best model in terms of marginal log-likelihood is the PMS-IC multiple-layer model that includes an endogenous layer and an exogenous economic layer ($\mathcal{M}_{21}^{END,ECO}$). There is evidence of substantial contagion effect in both the endogenous ($\hat{\beta}_{11} = 0.321$) and the exogenous ($\hat{\beta}_{21} = 0.601$) layers. This indicates that the information on business relations encoded in the economic network are relevant for explaining contagion, confirming the results obtained with single-layer first- and second-order interaction ($\mathcal{M}_{11}^{ECO}$ and $\mathcal{M}_{12}^{ECO}$). Nevertheless, they are not enough to explain contagion effects and a latent endogenous layer is needed to provide additional information about the spread of contagion cross the US states. The results for the model $\mathcal{M}_{11}^{END,GEO}$ indicate that the geographical network has a smaller yet still significant impact on contagion when compared with the endogenous economic network. Another interesting result regards the second-order effect specification. Comparing the marginal log-likelihood of models $\mathcal{M}_{12}^{GEO}$ and $\mathcal{M}_{12}^{ECO}$ with those of models $\mathcal{M}_{11}^{GEO}$ and $\mathcal{M}_{11}^{ECO}$, we can conclude that indirect effects should be accounted for in the specification of network interaction models. As in the geographical network model with second-order interaction ($\hat{\beta}_{11} = 0.011$ and $\hat{\beta}_{12} = 0.801$), higher order interaction effects may be more relevant than direct effects in spreading contagion. In conclusion, data support the estimation of locally-interacting chains which cannot simply proxy with exogenous geographical and economic proximity perspectives or with direct interaction effects. Our findings pose challenges to the view that crises can be restricted and limited locally.

We also study the ability of the model to detect the turning points of the US business cycle. Figure 5 plots the global interaction of recessions of US-state business cycles obtained from our model together with the diffusion index of the US national cycle phases published by the Federal Reserve of Philadelphia. We present two different measures from our model output: i) a recession probability computed as equal average of the smoothed recession probabilities of all states; ii) a recession probability computed as equal average of the filtered recession probabilities of all states. The former indicates how ex-post our model can call recessions; the latter one can be seen as an ex-ante measure of recessions where probabilities for time $T + 1$ are computed at time T (assuming parameters are known). The two probabilities behave very similarly, with the one based on filtered probabilities just marginally more volatile. Usually, filtered probabilities vary more than smoothed, but we remember our indicator is an average of 50 state probabilities. The individual state probabilities are quite different in several periods, see their distribution in the bottom panel of Figure 5. This is particularly evident in the middle of the 1980s, when NBER does not register a recession; however they are more similar when entering and exiting recessions. On average, our PMS-IC with global and local interactions of cycles matches better the national recessions given by the NBER, as the receiver operating characteristic curve (Berge and Jordà, 2011; Aastveit et al.,

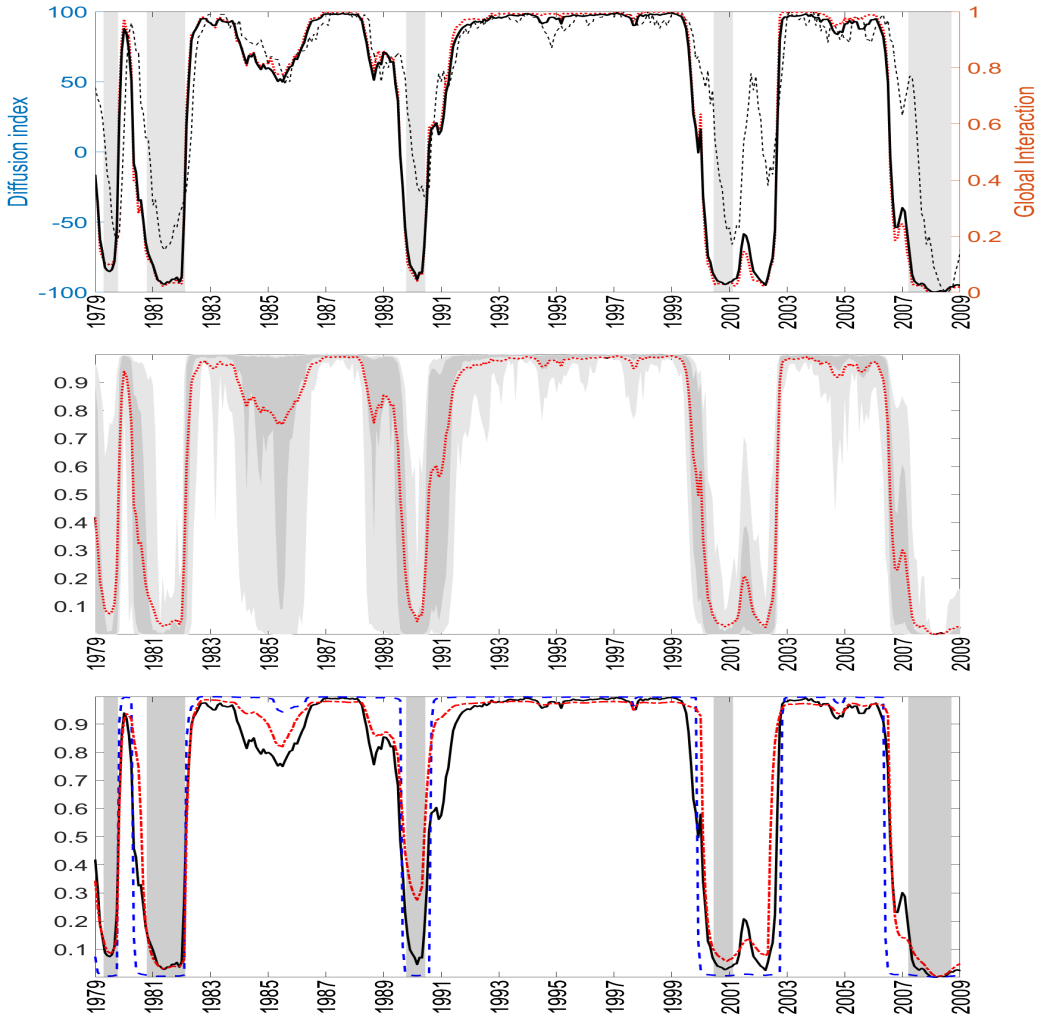


Figure 5: Top: The diffusion index of the Federal Reserve of Philadelphia (black dashed, left axis) and the global interaction factor of the model $\mathcal{M}_{21}^{END,ECO}$ estimated with the smoothed (red dotted, right axis) and the filtered (black solid, right axis) expansion probabilities. For all variables, lower values indicate higher recession probabilities. Gray bars represent the US national recession periods following the official dating of the National Bureau of Economic Research (NBER). Middle: Global interaction factor (red dotted) and cross-sectional inter-quantile ranges, at 90% (dark grey area) and 60% (light grey area), of the US states filtered probabilities. Bottom: Filtered expansion probabilities under different network configuration scenarios: baseline (black solid), employment contraction with mobility restrictions (blue dashed), and job support schemes with mobility restrictions (red dotted).

2016) in Figure G.1 in the Supplementary Material G.

Our model allows for us to assess the effects of policy interventions using connectivity among US states. We illustrate the approach through two counterfactual exercises by recomputing the expansion probabilities under alternative network configurations, while keeping all the other parameters unchanged. The baseline scenario corresponds to the historical recession probability of the model $\mathcal{M}_{21}^{END,ECO}$, where the average degree of USCB and endogenous networks are 13.12 and 25.12 respectively.

In the first exercise, we investigate the effects of negative shocks on employment in small firms, manufacturing and financial activities, and mobility restrictions across US states, which have become relevant during the COVID-19 pandemic event. The employment contraction increases the average degree of the endogenous network to 33.4. Moreover, we design a mobility scenario where US states with high contagion risk (Connecticut, New Jersey, Massachusetts, Rhode Island, Delaware, Minnesota, Maryland, Illinois, Nebraska, Iowa) are restricted in lockdown. The risk level is measured as the per-capita infection rate on June 1, 2020, estimated by combining documented case reports at the county level. See Figure 1 in Chande et al. (2020). The mobility restrictions reduce the average degree of the USCB network to 8.08.

In the second scenario, mobility restrictions are still in place, but job support schemes are implemented and employment does not reduce and therefore the endogenous network is kept unchanged. The results in the bottom panel in Figure 5 indicate that job support schemes (second scenario) reduce the length and intensity of the recessions despite the mobility lockdown. On the contrary, recessions can be longer and deeper when employment is not preserved (first scenario).

6 Conclusion

We propose a new dynamic panel Markov switching model with Markov chains interacting on a multi-layer network and provide a suitable parametrization and an efficient Markov Chain Monte Carlo (MCMC) algorithm for the posterior approximation. Our original model naturally accounts for the endogenous synchronization of the series, time-varying clustering of the hidden states, and time-varying transition probabilities. The synchronization effects are induced by the interaction mechanism. The interaction effects can either be local (i.e. in some neighborhoods) or global (i.e. among all units in the panel), direct or indirect. The underlying network can be exogenously given (e.g., geographical or economic proximity) or endogenously estimated.

We illustrate the usefulness of our PMS-IC model by applying it to the US regional business cycle. We find evidence of local interaction and of endogenous network effects which cannot simply proxy with exogenous proximity perspectives

or with only direct interaction effects. The approach can be used to assess the effects of policy interventions in a counterfactual analysis. Our methodology is motivated by and applied to a macroeconomic dataset, however it is general and could be applied to a wide spectrum of research where endogenous interaction effects are of interest.

References

Aastveit, K. A., Jore, A. S., and Ravazzolo, F. (2016). Identification and real-time forecasting of norwegian business cycles. *International Journal of Forecasting*, 32(2):283–292.

Ahelegbey, D. F., Billio, M., and Casarin, R. (2016). Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics*, 31(2):357 – 386.

Allen, F. and Babus, A. (2009). Networks in finance. In Kleindorfer, P. and Wind, J., editors, *The network challenge: strategy, profit, and risk in an interlinked world*. Wharton School Publishing.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Berge, T. J. and Jordà, O. (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics*, 3(2):246–277.

Bernile, G., Delikouras, S., Korniotis, G. M., and Kumar, A. (2017). The propagation of industrial business cycles. Technical Report 11-2014, Research Collection Lee Kong Chian School Of Business.

Bernile, G., Kumar, A., and Sulaeman, J. (2015). Home away from home: Geography of information and local investors. *The Review of Financial Studies*, 28(7):2009–2049.

Bianchi, D., Billio, M., Casarin, R., and Guidolin, M. (2019). Modeling systemic risk with markov switching graphical sur models. *Journal of Econometrics*, 210(1):58 – 74.

Billio, M., Casarin, R., Ravazzolo, F., and Van Dijk, H. K. (2016). Interconnections between Eurozone and US booms and busts: A Bayesian panel Markov-switching VAR model. *Journal of Applied Econometrics*, 31(7):1352–1370.

Billio, M., Getmansky, M., Lo, A., and Pellizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104:535 – 559.

Brémaud, P. (2013). *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media.

Burda, M. (2015). Constrained Hamiltonian Monte Carlo in BEKK GARCH with targeting. *Journal of Time Series Econometrics*, 7(1):95–113.

Burda, M. and Maheu, J. M. (2013). Bayesian adaptively updated Hamiltonian Monte Carlo with an application to high-dimensional bekk garch models. *Studies in Nonlinear Dynamics and Econometrics*, 17(4):345–372.

Camacho, M. and Leiva-Leon, D. (2014). The propagation of industrial business cycles. Technical Report 2014-48, Bank of Canada.

Casarin, R., Foroni, C., Marcellino, M., and Ravazzolo, F. (2018). Economic uncertainty through the lenses of a mixed-frequency bayesian panel markov switching model. *Annals of Applied Statistics*, 12(4):2559 – 2568.

Celeux, G. (1998). Bayesian inference for mixture: The label switching problem. In Payne, R. and Green, P., editors, *Compstat*. Physica, Heidelberg.

Christensen, O. F., Roberts, G. O., and Rosenthal, J. S. (2005). Scaling limits for the transient phase of local metropolis–hastings algorithms. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):253–268.

Crone, T. M. and Clayton-Matthews, A. (2005). Consistent economic indexes for the 50 states. *Review of Economics and Statistics*, (87):593–603.

Diebold, F. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Diebold, F. and Yilmaz, K. (2015). *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press.

Föllmer, H. and Horst, U. (2001). Convergence of locally and globally interacting Markov chains. *Stochastic Processes and Their Applications*, 96(1):99–121.

Francis, N., Owyang, M. T., and Savascin, O. (2017). An endogenously clustered factor approach to international business cycles. *Journal of Applied Econometrics*, 32(7):1261–1276.

Frühwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association*, 96(453):194–209.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer, New York.

Garcia, D. and Norli, O. (2012). Geographic dispersion and stock returns. *Journal of Financial Economics*, 106(3):547 – 565.

Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(2):123–214.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Hamilton, J. D. and Owyang, M. T. (2012). The propagation of regional recessions. *Review of Economics and Statistics*, 94(4):935–947.

Harding, D. and Pagan, A. (2002). Dissecting the cycle: A methodological investigation. *Journal of Mononetary Economics*, 49(2):365–381.

Holmes, C. C. and Held, L. (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis*, 1(1):145–168.

Kaufmann, S. (2010). Dating and forecasting turning points by Bayesian clustering with dynamic structure: A suggestion with an application to Austrian data. *Journal of Applied Econometrics*, 25(2):309–344.

Kaufmann, S. (2015). K-state switching models with time-varying transition distributions - does loan growth signal stronger effects of variables on inflation? *Journal of Econometrics*, 187(1):82–94.

Leiva-Leon, D. (2014). A new approach to infer changes in the synchronization of business cycle phases. Technical Report 2014-38, Bank of Canada.

Owayng, M., Piger, J., and Soques, D. (2019). Contagious switching. Technical report, Federal Reserve Bank of St. Louis Working Paper 2019-014B.

Owyang, M. T., Piger, J., and Wall, H. J. (2005). Business cycle phases in US states. *Review of Economics and Statistics*, 87(4):604–616.

Scott, S. L. (2011). Data augmentation, frequentist estimation, and the Bayesian analysis of multinomial logit models. *Statistical Papers*, 52(1):87–109.

Stock, J. H. and Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, 178:368–381.

Virbickaite, A., Ausín, M. C., and Galeano, P. (2015). Bayesian inference methods for univariate and multivariate garch models: A survey. *Journal of Economic Surveys*, 29(1):76–96.

Wasserman, S. (1994). *Advances in social network analysis: Research in the social and behavioral sciences*. Sage.

Williams, D. (1991). *Probability with martingales*. Cambridge University Press.

A Properties of the PMS-IC model

A.1 Proof of Proposition 1

Without loss of generality let assume $K = 2$. Define $w_{it,1} = s_{it} - \mu_{it}$ where $\mu_{it} = \pi_{it}(\mathbf{s}_t, 1)$ and $\mathbf{s}_t \in \mathcal{S}_1$. Then $\{w_{it,1}\}_{i\geq 1}$ is a sequence of independent random variables, conditioning on $\mathcal{F}_{t-1} = \sigma(\{\mathbf{s}_u\}_{u\leq t-1})$, such that $\mathbb{E}(w_{it,1}|\mathcal{F}_{t-1}) = 0$ and $\mathbb{V}(w_{it,1}|\mathcal{F}_{t-1}) = \mu_{it}(1-\mu_{it})$ which satisfies $\sum_{i=1}^{\infty} \frac{\mu_{it}(1-\mu_{it})}{i^2} < \infty$. Then by the Strong Law of Large Numbers it follows that $\sum_{i=1}^{n} w_{it,1}$ converge a.s. to zero for $n \to \infty$ (see Williams (1991), p. 118, Theorem 12.8).

From the previous result we have

$$\begin{aligned} \overline{m}_{t+1,1} &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \pi_{it}(\mathbf{s}_t, 1) \qquad (22) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(s_{it}\alpha p_{11} + (1-s_{it})\alpha p_{01} + \sum_{r=1}^{N}\sum_{q=1}^{M} \beta_{rq}\overline{m}_{irqt,1} + \gamma \overline{m}_{t,1} \right) \\ &= (\overline{m}_{t,1}\alpha p_{11} + (1-\overline{m}_{t,1})\alpha p_{01} + (1-\alpha-\gamma)\overline{m}_{t,1} + \gamma\overline{m}_{t,1}). \end{aligned}$$

since $\sum_{r=1}^{N}\sum_{q=1}^{M}\beta_{rq}\overline{m}_{irqt,1} = \overline{m}_{t,1}(1-\alpha-\gamma)/\beta$. The limits of the recursion can be easily find by setting $\overline{m}_t = \overline{m}^*$ and solving the equation

$$\overline{m}^* = \alpha(\overline{m}^* p_{11} + (1-\overline{m}^*)p_{01}) + (1-\alpha)\overline{m}^*$$

in $\overline{m}^*$.

A.2 Proof of Proposition 2

See Föllmer and Horst (2001).

B Parameter full conditional distributions

1. The full conditional distribution of $(\alpha, \boldsymbol{\beta}, \gamma)$ according to the likelihood in equation (15) and the prior in equation (18) density function:

$$\begin{aligned} f(\alpha, \boldsymbol{\beta}, \gamma | \Xi, \boldsymbol{\theta}_{-(\alpha,\beta,\gamma)}) &\propto \left(\alpha^{r_1-1}\beta_1^{r_2-1}\dots\beta_{NM}^{r_{1+NM}-1}\gamma^{r_{2+NM}-1} \right) \\ &\cdot \prod_{k=1}^{K}\prod_{l=1}^{K}\prod_{i=1}^{n}\prod_{t\in\mathcal{T}_{ilk}} \left(\alpha p_{lk} + \sum_{r=1}^{N}\sum_{q=1}^{M}\beta_{rq} m_{irq,k}(\mathbf{s}_t) + \gamma m_{t,k} \right) \end{aligned}$$

where $\mathcal{T}_{ilk} = \{t = 1, \dots, T | \xi_{l,it-1}\xi_{k,it} = 1\}$.

2. The full conditional distribution of the regime dependent intercept μ_{il} (where $l = 1, \ldots, K$ and $i = 1, \ldots, n$) according to the likelihood in equation (15) and the prior in equation (15) has a normal distribution with density function:

$$
\begin{aligned}
f(\mu_{il}|\mathbf{y}, \Xi, \boldsymbol{\theta}_{-\mu_{il}}) &\propto \exp\left(-\frac{1}{2s_{il}^2}\left(\mu_{il} - m_{il}\right)^2\right) \prod_{t=1}^{T} \exp\left(-\frac{\xi_{l,it}}{2\sigma_{il}^2}\left(y_{it} - \mu_{il}\right)^2\right) \\
&\propto \exp\left(-\frac{1}{2}\mu_{il}^2\left(\frac{1}{s_{il}^2} + \frac{T_{il}}{\sigma_{il}^2}\right) - 2\mu_{il}\left(\frac{m_{il}}{s_{il}^2} + \frac{1}{\sigma_{il}^2}\sum_{t\in\mathcal{T}_{il}} y_{it}\right)\right) \\
&\propto \mathcal{N}(\overline{m}_{il}, \overline{s}_{il}^2)
\end{aligned}
$$

with $\overline{m}_{il} = \overline{s}_{il}^2\left(\frac{m_{il}}{s_{il}^2} + \frac{1}{\sigma_{il}^2}\sum_{t\in\mathcal{T}_{il}} y_{it}\right)$ and $\overline{s}_{il}^2 = \left(\frac{1}{s_{il}^2} + \frac{T_{il}}{\sigma_{il}^2}\right)^{-1}$.
We defined $\mathcal{T}_{il} = \{t = 1, \ldots, T | s_{it} = l\}$, $T_{il} = \text{card}(\mathcal{T}_{il})$. The notation $\boldsymbol{\theta}_{-r}$ indicates that element r is excluded from the vector $\boldsymbol{\theta}$.

3. The full conditional distribution of the regime-specific volatilities σ_{il}, $l = 1, \ldots, K$ and $i = 1, \ldots, n$, according to the likelihood in equation (15) and the prior in equation (16) has inverse gamma distribution with density function:

$$
\begin{aligned}
f\left(\sigma_{il}^2|\mathbf{y}, \Xi, \boldsymbol{\theta}_{-\sigma_{il}}\right) &\propto \left(\frac{1}{\sigma_{il}^2}\right)^{a_{il}+T_{il}+1} \exp\left(-\frac{1}{\sigma_{il}^2}\left(b_{il} + \sum_{t\in\mathcal{T}_{il}}(y_{it} - \mu_{il})^2\right)\right) \\
&\propto \mathcal{IG}\left(a_{il} + T_{il}, b_{il} + \sum_{t\in\mathcal{T}_{il}}\left(y_{it} - \mu_{il}\right)^2\right)
\end{aligned}
$$

4. The full conditional distribution of each l-th row of the transition matrix $P_{l,1:K} = (P_{l1}, \ldots, P_{lK})$ according to the likelihood in equation (15) and the prior in equation (17) density function:

$$
\begin{aligned}
&f\left(p_{l,1:K}|\mathbf{y}, \Xi, \theta_{-\left(p_{l,1:K}\right)}\right) \\
&\propto \left(\prod_{k=1}^{K} p_{lk}^{(v_k-1)}\right) \prod_{t=1}^{T}\prod_{i=1}^{n}\prod_{k=1}^{K}\left(\alpha p_{lk} + \sum_{r=1}^{N}\sum_{q=1}^{M}\beta_{rq} m_{irq,k}\left(\mathbf{s}_t\right) + \gamma m_k\left(\mathbf{s}_t\right)\right)^{\xi_{k,it}\xi_{l,it-1}} \\
&\propto \left(\prod_{k=1}^{K} p_{lk}^{(v_k-1)}\right) \prod_{k=1}^{K}\prod_{i=1}^{n}\prod_{t\in\mathcal{T}_{ilk}}\left(\alpha p_{lk} + \sum_{r=1}^{N}\sum_{q=1}^{M}\beta_{rq} m_{irq,k}\left(\mathbf{s}_t\right) + \gamma m_k\left(\mathbf{s}_t\right)\right) \\
&\propto \prod_{k=1}^{K}\left(p_{lk}^{(v_k-1)} \prod_{i=1}^{n}\prod_{t\in\mathcal{T}_{ilk}}\left(\alpha p_{lk} + \sum_{r=1}^{N}\sum_{q=1}^{M}\beta_{rq} m_{irq,k}\left(\mathbf{s}_t\right) + \gamma m_k\left(\mathbf{s}_t\right)\right)\right)
\end{aligned}
$$

5. If we stack vertically the elements of Z and η_{ij} forming the two n^2-dimensional column vectors $\mathbf{z}$ and $\boldsymbol{\eta}$, then the latent regression can be written as $\mathbf{z} = X\boldsymbol{\varphi} + 2W\boldsymbol{\eta}$, where X is a $n^2 \times 4$ matrix and $W = \text{diag}\{\text{vec}(\Lambda)\}$ is a n^2-dimensional diagonal matrix where $\Lambda = ((\lambda_{11}, \ldots, \lambda_{n1})', \ldots, (\lambda_{1n}, \ldots, \lambda_{nn})')$ and $\lambda_{ij} = 4\psi_{ij}^2$. The full conditional distribution of the latent regression parameters $\boldsymbol{\varphi}$ is

$$f(\boldsymbol{\varphi}|X, \Lambda, Z) \propto \mathcal{N}(\bar{\mathbf{m}}, \bar{S})$$

where

$$\bar{\mathbf{m}} = \bar{S}\left(\underline{S}^{-1}\underline{m} + X'W^{-1}\mathbf{z}\right), \quad \bar{S} = (\underline{S}^{-1} + X'W^{-1}X)$$

6. The full conditional distribution of A is

$$f(A|\Xi, \boldsymbol{\theta}) \propto f(\Xi|A, \boldsymbol{\theta}) f(A|\boldsymbol{\theta})$$

with support set $\{0, 1\}^{n^2}$. We sample from $f(A|\Xi, Z, \boldsymbol{\theta})$ by Metropolis-Hasting. Given two nodes i and j selected at random in the vertex set of $\mathcal{G}$, if there exists an edge between them, then we propose to remove the link. Also, we propose to add an edge between i and j if there is no path linking i to j (we verify this using the reachability matrix, see Wasserman (1994)). The validity of the proposed MCMC is granted by the fact that (i) the probability of selecting a node is strictly positive for all nodes, and (ii) the chain may remain in the current state with positive probability. Together, these conditions guarantee irreducibility and aperiodicity of the Markov chain.

C Latent variables full conditional distributions

C.1 State allocation variables

In our inference procedure, a multi-move sampling is needed to sample from the joint posterior distribution $f(\Xi_{i,1:T}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta})$. We apply a forward-filtering and backward-sampling (FFBS) algorithm. By means of dynamic factorization, the full conditional distribution of the unit specific hidden state is

$$\begin{aligned}
\mathbb{P}(\Xi_{i,1:T}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) &= \mathbb{P}(\boldsymbol{\xi}_{iT}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \mathbb{P}(\Xi_{i,1:T-1}|\xi_{iT}, \Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \\
&= \mathbb{P}(\boldsymbol{\xi}_{iT}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \prod_{t=1}^{T-1} \mathbb{P}(\boldsymbol{\xi}_{i,t}|\Xi_{i,t+1:T}, \Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \\
\propto \mathbb{P}(\boldsymbol{\xi}_{iT}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) &\prod_{t=1}^{T-1} \mathbb{P}(\boldsymbol{\xi}_{i,t+1}|\boldsymbol{\xi}_{i,t}, \Xi_{-i,t}, A) \mathbb{P}(\boldsymbol{\xi}_{i,t}|\Xi_{-i,1:T}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta})
\end{aligned}$$

$$\propto \left(\mathbb{P}(\boldsymbol{\xi}_{iT}|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \prod_{t=1}^{T-1} \mathbb{P}(\boldsymbol{\xi}_{i,t}|\Xi_{-i,1:t}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta}) \right) \left(\prod_{t=1}^{T-1} \mathbb{P}(\boldsymbol{\xi}_{i,t+1}|\boldsymbol{\xi}_{i,t}, \Xi_{-i,t}, A, \boldsymbol{\theta}) \right)$$

With this factorization, a Metropolis-Hasting (MH) procedure is needed to take into account the proportionality factor with the FFBS algorithm as proposal for the unit specific hidden state. The filtering probability for unit i at time t, $t = 1, \ldots, T$, algorithm gives the prediction probability, the one step-ahead forecast density and the updated probability. Recalling the one-to-one mapping between s_{it} and $\boldsymbol{\xi}_{l,it} = \mathbb{I}_l(s_{it})$, the prediction probability of s_{it} is:

$$\begin{aligned}
&\mathbb{P}(s_{it} = 1|\Xi_{-i,1:t}, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) = \\
&= \sum_{k=1}^{K} \mathbb{P}(s_{it} = l|s_{it-1} = k, \Xi_{-it-1}, A) \mathbb{P}(s_{i,t-1} = k|\Xi_{-i,1:t-1}, A, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) \\
&= \sum_{k=1}^{K} P_{it-1,kl} \mathbb{P}(s_{i,t-1} = k|\Xi_{-i,1:t-1}, A, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) \qquad (23)
\end{aligned}$$

for $l = 1, \ldots, K$, where $P_{it-1,lk}$ is the conditional probability that unit i moves from regime k at time $t-1$ to regime l at time t. $\Xi_{-i,t} = \{\boldsymbol{\xi}_{jt}, j = 1, \ldots, nj \neq i\}$. We initialize for $t = 1$, $\mathbb{P}(s_{i,0} = k|\mathbf{y}_0, \boldsymbol{\theta})$ to be equal to the ergodic probabilities.
The filtered probability for all $l = 1, \ldots, K$ is computed as:

$$\begin{aligned}
\mathbb{P}(s_{it} = l|\Xi_{-i,1:t}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta}) &\propto \mathbb{P}(s_{it} = l|\Xi_{-i,1:t-1}, A, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) f(y_{it}|s_{it} = l, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) \\
&\propto \mathbb{P}(s_{it} = l|\Xi_{-i,1:t-1}, A, \mathbf{y}_{1:t-1}, \boldsymbol{\theta}) \mathcal{N}(\mu_{il}, \sigma_{il}^2) \qquad (24)
\end{aligned}$$

The smoothing probabilities are obtained recursively and backward in time, once all the filtered probabilities $\mathbb{P}(s_{it} = l|\Xi_{-i,1:t}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta})$ for $t = 1, \ldots, T$ are calculated. If $t = T$, smoothing probability and filtered probability are equal. For $t = T-1, T-2, \ldots, 1$ and for all $l = 1, \ldots, K$ the smoothing algorithm proceeds as follows:

$$\begin{aligned}
&\mathbb{P}(s_{it} = l|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) = \sum_{k=1}^{K} \mathbb{P}(s_{it} = l, s_{it+1} = k|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \\
&= \sum_{k=1}^{K} \mathbb{P}(s_{it} = l|s_{it+1} = k, \Xi_{-i,1:T}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta}) \mathbb{P}(s_{it+1} = k|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta}) \\
&= \sum_{k=1}^{K} \frac{p_{it,lk} \mathbb{P}(s_{it} = l|\Xi_{-i,1:T}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta}) \mathbb{P}(s_{it+1} = k|\Xi_{-i,1:T}, A, \mathbf{y}, \boldsymbol{\theta})}{\sum_{j=1}^{K} p_{it,jk} \mathbb{P}(s_{it} = j|\Xi_{-i,1:T}, A, \mathbf{y}_{1:t}, \boldsymbol{\theta})} \qquad (25)
\end{aligned}$$

C.2 Network latent regression variables

1. The full conditional distribution of the latent regression variable Z is

$$f(\zeta_{11t}, \ldots, \zeta_{nnt}|A, \boldsymbol{\theta}) \propto \prod_{i=1}^{n}\prod_{j=1}^{n} f(\zeta_{ij}|a_{ij}, \boldsymbol{\varphi}) \tag{26}$$

where $f(\zeta_{ij}|a_{ij}, \boldsymbol{\varphi})$ is a logistic distribution with mean $g(\mathbf{x}_i, \mathbf{x}_j)'\boldsymbol{\varphi}$ and truncated by $\zeta_{ij} \geq 0$ if $a_{ij} = 1$ and by $\zeta_{ij} < 0$ if $a_{ij} = 0$. Random samples from this distribution can generated exactly by applying the following transform $\zeta_{ij} = g(\mathbf{x}_i, \mathbf{x}_j)'\boldsymbol{\varphi} - \log(u_{ij}^{-1} - 1)$ where u_{ij} is generated from the following mixture of two uniforms distributions

$$u_{ij} \sim \mathcal{U}_{[0,d]}\mathbb{I}(a_{ij} = 0) + \mathcal{U}_{d,1}\mathbb{I}(a_{ij} = 1) \tag{27}$$

where $\mathcal{U}_{[a,b]}$ denotes the uniform distribution over the interval $[a, b]$ and $d = (1 + \exp(g(\mathbf{x}_i, \mathbf{x}_j)'\boldsymbol{\varphi}))^{-1}$.

2. The full conditional distribution of the latent regression variable Λ is

$$f(\Lambda|Z, \boldsymbol{\theta}) \propto \prod_{i=1}^{n}\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\lambda_{ij}}} \exp\{-\frac{1}{2\lambda_{ij}}(\zeta_{ij} - g(\mathbf{x}_i, \mathbf{x}_j)'\boldsymbol{\varphi})^2\} f(\lambda_{ij}) \tag{28}$$

where $f(\lambda_{ij})$ is equal to $f(\psi_{ij})/(8\sqrt(\lambda_{ij}))$ with $\psi_{ij} = \sqrt(\lambda_{ij})/4$. As in Holmes and Held (2006) we apply an accept/reject sampling method with a Generalized Inverse Gaussian proposal distribution

$$\lambda_{ij}^* \sim \mathcal{GIG}(1/2, 1, r^2)$$

where $r_{ij}^2 = (\zeta_{ij} - g(\mathbf{x}_i, \mathbf{x}_j)'\boldsymbol{\varphi})^2$.

D Metropolis-adjusted Langevin algorithm

The implementation of the Metropolis-adjusted Langevin algorithm (MALA) requires some necessary expressions which we discuss in this section. We consider the logistic transformation of the parameters $\tilde{p}_{lk}$, $k = 1, \ldots, K$, $\tilde{\alpha}$, $\tilde{\beta_{rg}}$ and $\tilde{\gamma}$ to impose constrains on the parameters space, that is $\alpha = g(\tilde{\alpha})/c$, $\beta_r q = g(\tilde{\beta}_{rq})/c$, $\gamma = g(\tilde{\gamma})/c$, $p_{lk} = g(\tilde{p_{lk}})/d$ where $g(x) = \exp(x)$, $c = g(\tilde{\alpha}) + g(\tilde{\beta}_{11}) + \ldots + g(\tilde{\beta}_{NM}) + g(\tilde{\gamma})$, $d = g(\tilde{p_{l1}}) + \ldots + g(\tilde{p_{lK}})$, and $\tilde{p_{lk}}$, $k = 1, \ldots, K$, $\tilde{\alpha}$, $\tilde{\beta}_{rq}$, $\tilde{\gamma}$ take value in the set of the real numbers.

The partial derivatives of the complete data log-likelihood with respect to the new parametrization are:

$$\frac{\partial \log L}{\partial \tilde{\alpha}} = g(\tilde{\alpha})(1-g(\tilde{\alpha}))\frac{1}{c^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\xi_{l,it}\xi_{k,it-1}\frac{p_{lk}}{p_{it,lk}} \tag{29}$$

$$\frac{\partial \log L}{\partial \tilde{\beta}} = g(\tilde{\beta}_{rq})(1-g(\tilde{\beta}_{rq}))\frac{1}{c^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\xi_{l,it}\xi_{k,it-1}\frac{m_{irq,k}(\mathbf{s}_t)}{p_{it,lk}} \tag{30}$$

$$\frac{\partial \log L}{\partial \tilde{\gamma}} = g(\tilde{\gamma})(1-g(\tilde{\gamma}))\frac{1}{c^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\xi_{l,it}\xi_{k,it-1}\frac{m_{k}(\mathbf{s}_t)}{p_{it,lk}} \tag{31}$$

$$\frac{\partial \log L}{\partial \tilde{p_{lk}}} = g(\tilde{p}_{lk})(1-g(\tilde{p}_{lk}))\frac{1}{d^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\xi_{l,it}\xi_{k,it-1}\frac{\alpha}{p_{it,lk}} \tag{32}$$

where $\xi_{l,it} = \mathbb{I}_{\{l\}}(s_{it})$.

Supplementary Material

E Numerical illustrations

We isolate the contribution of the global and local interaction mechanism and specify six parameter settings which are summarized in Table E.1.

Model label	α	β	γ
$\mathcal{M}_1$ (No interaction)	1.00	0.00	0.00
$\mathcal{M}_2$ (Weak global interaction)	0.70	0.00	0.30
$\mathcal{M}_3$ (Strong global interaction)	0.30	0.00	0.70
$\mathcal{M}_4$ (Weak local interaction)	0.70	0.30	0.00
$\mathcal{M}_5$ (Strong local interaction)	0.30	0.70	0.00
$\mathcal{M}_6$ (Both local and global interaction)	0.50	0.25	0.25

Table E.1: Parameter values of idiosyncratic, global and local interaction mechanisms for each model $\mathcal{M}_i$, $i = 1, \ldots, 6$.

The difference between the six settings is on the choice of the underlying parameters α, β and γ. The values of α, β and γ in our settings were carefully chosen in order to represent a wide variety of possible interaction effects.

In model $\mathcal{M}_1$, all the interactions are null and the overall effect is given by the fixed component of the transition matrix. We assume a weak global interaction among the Markov chains in $\mathcal{M}_2$ and a stronger one in $\mathcal{M}_3$. On the other side, we consider a weak global interaction among the Markov chains in $\mathcal{M}_4$ and a stronger one in $\mathcal{M}_5$. The global and local effect are simulated in $\mathcal{M}_6$ assuming for them an equal weight. In all the experiments, we consider a population of 50 time series following PMS-IC generated with 5000 time horizon. Furthermore, we assume the following model specification with three regimes (ie, $K = 3$):

$$y_{it} = \sum_{k=1}^{K} \mathbb{I}_{\{k\}}(s_{it}) \left[\mu_{ik} + \sigma_{ik}\varepsilon_{it}\right], \quad \varepsilon_{it} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$$

for $\quad i = 1, \ldots, 50 \quad$ and $\quad t = 1, \ldots, 5000$, and the fixed transition matrix:

$$P = \begin{pmatrix} 0.98 & 0.02 & 0.00 \\ 0.01 & 0.98 & 0.01 \\ 0.00 & 0.02 & 0.98 \end{pmatrix} \tag{E.33}$$

$\mu_{i1} = -2, \quad \mu_{i2} = 0, \quad \mu_{i3} = 2, \quad \sigma_{i1} = .3, \quad \sigma_{i2} = .05, \quad \sigma_{i3} = .3$. Note that the ergodic probability associate with P is $\pi_i = (.25 \quad .5 \quad .25)$.

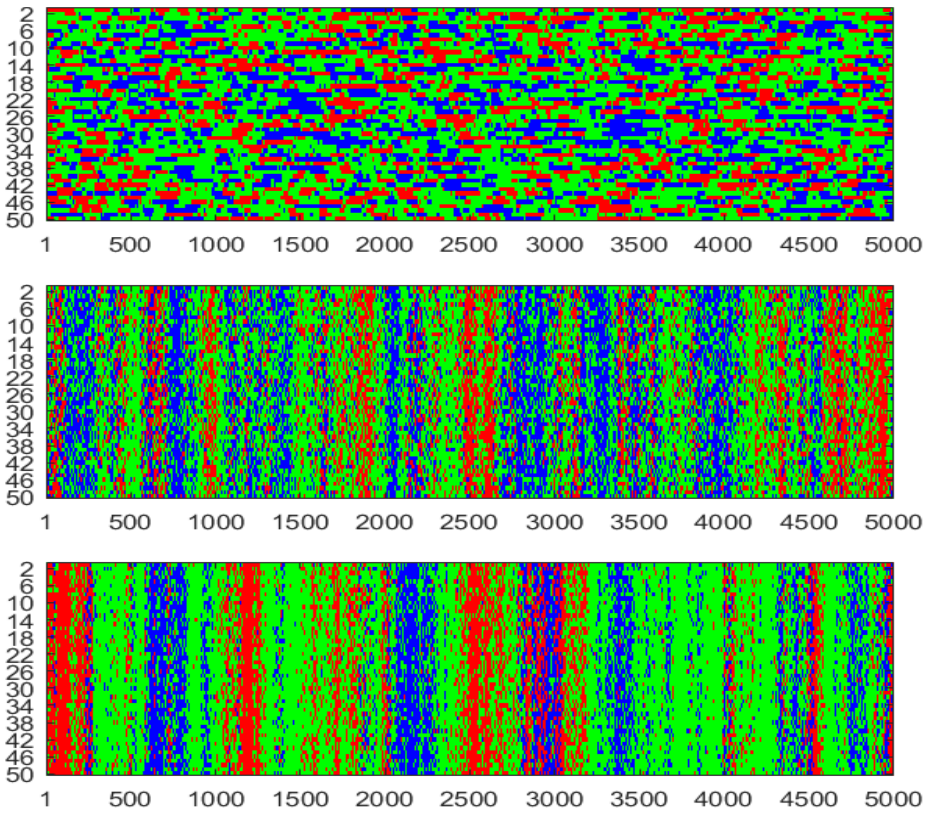


Figure E.1: Endogenous synchronization effects. Heat-map of 50 trajectories (vertical axis) of 5000 observations each (horizontal axis), in absence of interaction ($\gamma = 0$, top plot), and in presence of weak (mid plot, $\gamma = 0.3$) and strong (bottom plot, $\gamma = 0.7$) global interaction effects. Colors represent the value of the series, where blue, green and red indicate expansion, moderate expansion and recession regimes, respectively.

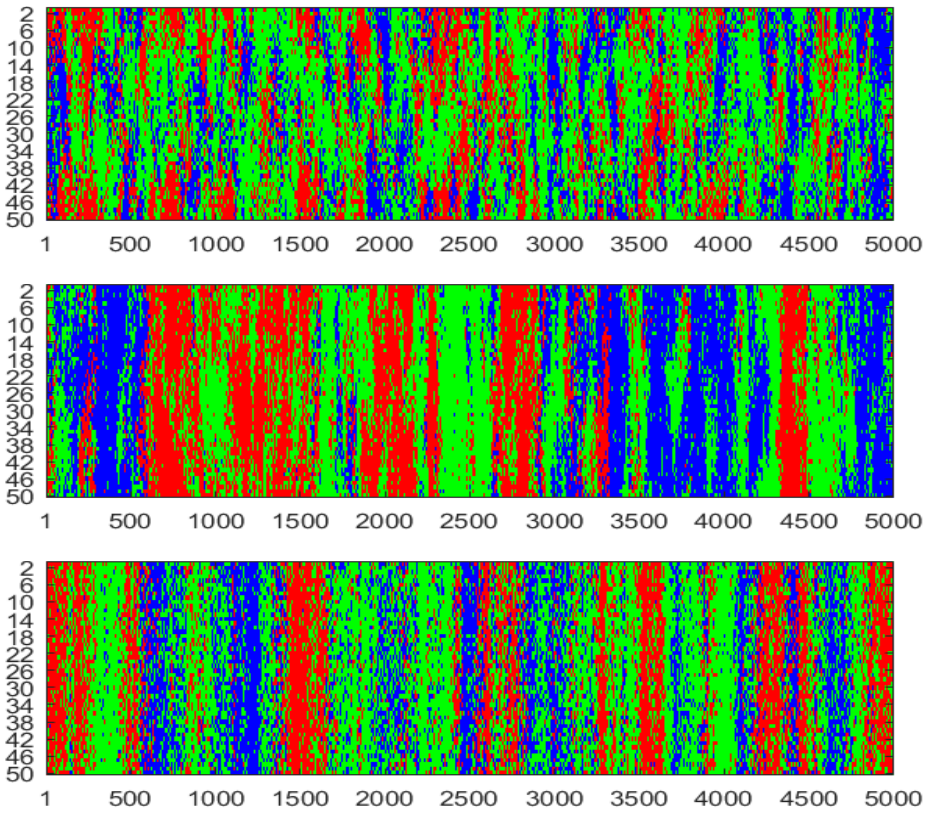


Figure E.2: Endogenous synchronization effects. Heat-map of 50 trajectories (vertical axis) of 5000 observations each (horizontal axis), in presence of weak ($\beta = 0.3$, top plot), and strong (mid plot, $\gamma = 0.7$) local interaction effects and of both local and global interaction effects (bottom plot, $\beta = 0.25$ and $\gamma = 0.25$). Colors represent the value of the series, where blue, green and red indicate expansion, moderate expansion and recession regimes, respectively.

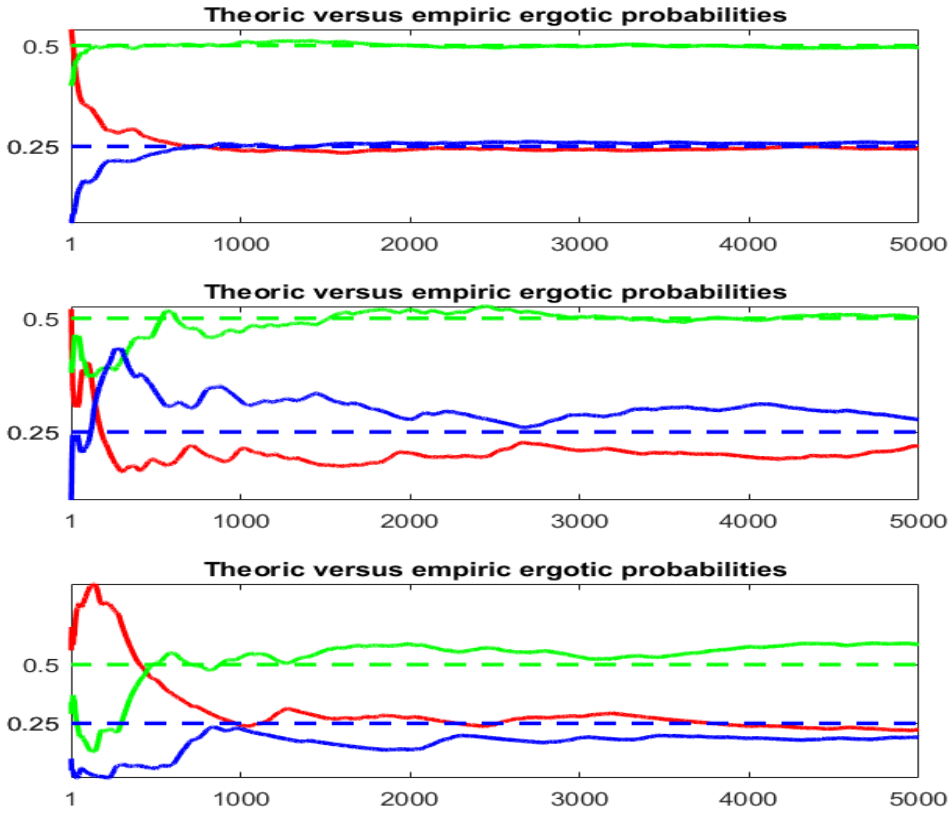


Figure E.3: Endogenous synchronization effects. Empirical (solid lines) and theoretical (dashed lines) ergodic probabilities, in absence of interaction ($\gamma = 0$, top plot), and in presence of weak (mid plot, $\gamma = 0.3$) and strong (bottom plot, $\gamma = 0.7$) global interaction effects. Blue, green and red indicate expansion, moderate expansion and recession regime probabilities, respectively.

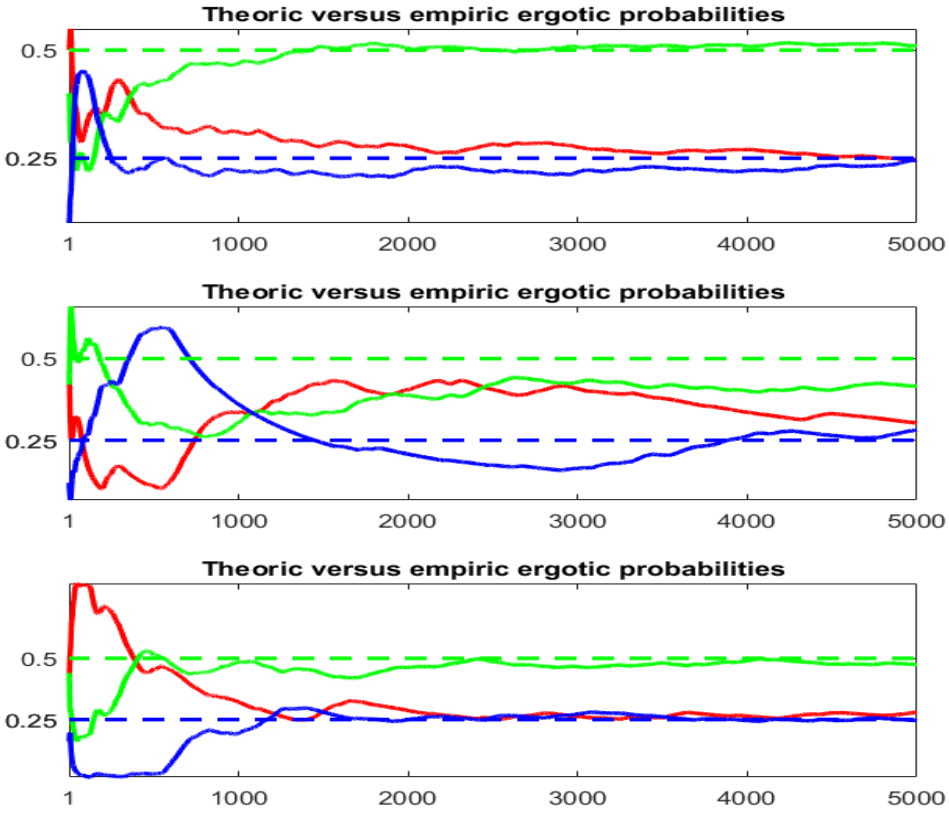
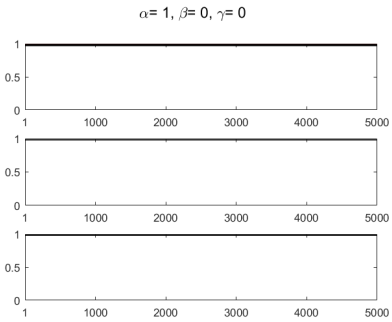
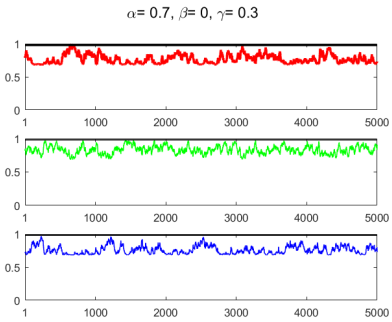


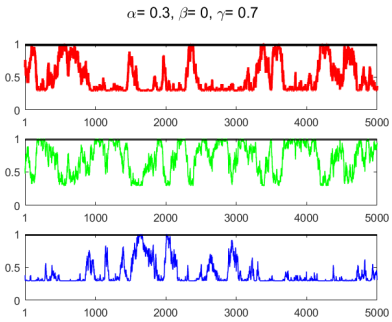
Figure E.4: Endogenous synchronization effects. Empirical (solid lines) and theoretical (dashed lines) ergodic probabilities, in presence of weak ($\beta = 0.3$, top plot), and strong (mid plot, $\gamma = 0.7$) local interaction effects and of both local and global interaction effects (bottom plot, $\beta = 0.25$ and $\gamma = 0.25$). Blue, green and red indicate expansion, moderate expansion and recession regime probabilities, respectively.



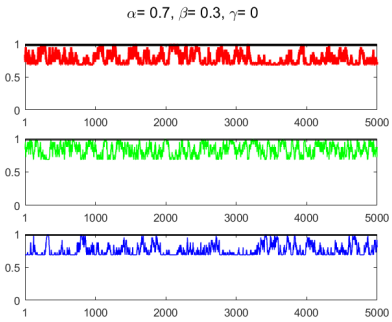
(a) No interaction



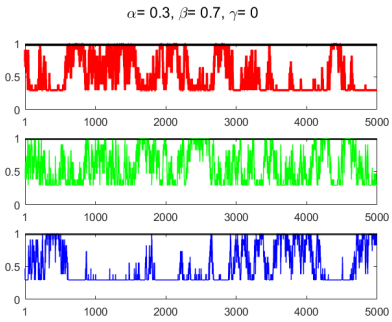
(b) Weak global interactions



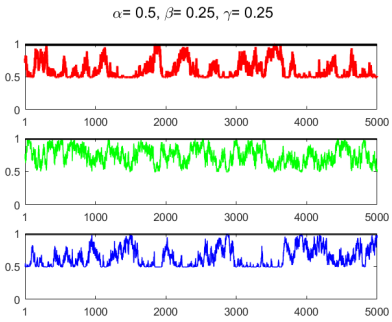
(c) Strong global interactions



(d) Weak local interactions



(e) Strong local interactions



(f) Both local and global interactions

Figure E.5: For each parameter setting, the trajectory of the elements of both fixed and time varying transition matrices for the first chain in the panel. Colors blue is for $p_{1t,11}$; colors green refer $p_{1t,22}$ and colors red refer to $p_{1t,33}$. For all plots, horizontal black lines represent the fixed transition probabilities defined in equation (E.33).

F Simulation experiments

We simulate data in the six parameter settings described in section 3.2 (see Table E.1 for parameter settings) in order to assess the efficiency of the proposed MCMC algorithm for the posterior approximation. We assess the efficiency using the mean square error (MSE) for the parameters and the hidden states. MSE is evaluated on

Setting label	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
The unit-specific regression parameters (in total 50 parameters for each regime)						
$\mu_{.,1}$	1.323e-02	6.144e-04	0.762e-03	0.866e-03	8.078e-04	6.954e-04
$\mu_{.,2}$	1.977e-02	0.750e-04	0.009e-03	0.009e-03	0.161e-04	0.119e-04
$\mu_{.,3}$	0.357e-02	4.031e-04	0.731e-03	0.632e-03	4.402e-04	9.807e-04
$\sigma_{.,1}$	0.740e-03	4.000e-04	0.001e-03	0.313e-03	5.044e-04	0.000e-04
$\sigma_{.,2}$	0.120e-03	1.600e-04	0.000e-04	0.005e-03	0.090e-04	0.000e-04
$\sigma_{.,3}$	0.109e-03	2.000e-04	0.000e-04	0.757e-03	2.013e-04	0.001e-04
Idiosyncratic, local and global parameters						
(α, β, γ)	1.976e-07	3.5535e-04	5.5682e-04	1.540e-02	3.09e-02	4.110e-02
The unit-specific Markov chains (in total 50 parameters for each regime)						
Regime 1	5.000e-04	1.200e-03	2.000e-03	1.400e-03	1.600e-03	1.900e-03
Regime 2	1.300e-04	2.200e-03	3.800e-03	2.100e-03	2.100e-03	2.800e-03
Regime 3	6.000e-04	1.600e-03	2.400e-03	1.200e-03	2.200e-03	1.600e-03

Table F.1: Mean square error (MSE) for the parameters estimated using the proposed MCMC algorithm for the PMS-IC model.

5000 iterations after convergence (1000 draws). Table F.1 reports for each model the average MSE for the 50 units in our panel.

The first evidence is that the precision of the inference of the unit-specific regression parameters decreases with the parameters β and γ. The second evidence is that precision of the inferences of the unit-specific Markov chains increases with the parameters β and γ.

G Additional empirical results

State	Geographical Region	Degree Geo Prox	Degree USCB Net	Eigencentrality Geo Prox	Eigencentrality USCB Net
Alabama	South	4	16	0.31	1.00
Alaska	West	1	13	0.03	0.43
Arizona	West	5	13	0.49	0.43
Arkansas	South	6	16	0.65	1.00
California	West	4	13	0.23	0.43
Colorado	West	7	8	0.80	0.16
Connecticut	NorthEast	3	13	0.10	0.43
Delaware	South	3	16	0.17	1.00
Florida	South	2	16	0.13	1.00
Georgia	South	5	16	0.33	1.00
Hawaii	West	1	13	0.05	0.43
Idaho	West	6	13	0.47	0.43
Illinois	MidWest	5	11	0.57	0.18
Indiana	MidWest	4	11	0.37	0.18
Iowa	MidWest	6	11	0.65	0.18
Kansas	MidWest	4	11	0.59	0.18
Kentucky	South	7	16	0.76	1.00
Louisiana	South	3	16	0.26	1.00
Maine	NorthEast	1	8	0.02	0.16
Maryland	South	4	16	0.29	1.00
Massachusetts	NorthEast	5	8	0.14	0.16
Michigan	MidWest	3	11	0.21	0.18
Minnesota	MidWest	4	11	0.33	0.18
Mississippi	South	4	16	0.38	1.00
Missouri	MidWest	8	11	1.00	0.18
Montana	West	4	13	0.36	0.43
Nebraska	MidWest	6	11	0.77	0.18
Nevada	West	5	13	0.40	0.43
New Hampshire	NorthEast	3	8	0.07	0.16
New Jersey	NorthEast	3	8	0.16	0.16
New Mexico	West	5	8	0.55	0.16
New York	NorthEast	5	13	0.20	0.43
North Carolina	South	4	16	0.33	1.00
North Dakota	MidWest	3	11	0.24	0.18
Ohio	MidWest	5	11	0.41	0.18
Oklahoma	South	6	16	0.72	1.00
Oregon	West	4	13	0.25	0.43
Pennsylvania	NorthEast	6	8	0.35	0.16
Rhode Island	NorthEast	2	20	0.06	0.52
South Carolina	South	2	16	0.13	1.00
South Dakota	MidWest	6	11	0.56	0.18
Tennessee	South	8	16	0.79	1.00
Texas	South	4	16	0.40	1.00
Utah	West	6	13	0.63	0.43
Vermont	NorthEast	3	20	0.11	0.52
Virginia	South	5	16	0.50	1.00
Washington	West	3	16	0.15	1.00
West Virginia	South	4	16	0.33	1.00
Wisconsin	MidWest	5	13	0.45	0.43
Wyoming	West	6	11	0.66	0.18

Table G.1: Statistics, including degree and eigen centrality, for each node of the geographical proximity network and USCB network. Geographical regions in which each State is classified are also provided.

State	μ_{i1}	Bayesian CI	μ_{i2}	Bayesian CI	σ_{i1}	Bayesian CI	σ_{i2}	Bayesian CI
Alabama	-0.153	(-0.191, -0.115)	0.345	(0.330, 0.360)	0.247	(0.228, 0.266)	0.147	(0.140, 0.154)
Alaska	-1.032	(-1.125, -0.942)	0.234	(0.201, 0.266)	0.260	(0.210, 0.317)	0.367	(0.350, 0.383)
Arizona	-0.067	(-0.115, -0.016)	0.629	(0.600, 0.659)	0.356	(0.333, 0.380)	0.235	(0.221, 0.250)
Arkansas	-0.068	(-0.095, -0.040)	0.356	(0.342, 0.371)	0.187	(0.173, 0.201)	0.134	(0.127, 0.140)
California	0.002	(-0.024, 0.027)	0.380	(0.369, 0.391)	0.184	(0.172, 0.197)	0.095	(0.090, 0.101)
Colorado	-0.095	(-0.134, -0.055)	0.425	(0.410, 0.442)	0.230	(0.213, 0.248)	0.143	(0.135, 0.151)
Connecticut	-0.111	(-0.139, -0.083)	0.349	(0.333, 0.367)	0.182	(0.170, 0.196)	0.150	(0.142, 0.159)
Delaware	-0.084	(-0.113, -0.051)	0.409	(0.390, 0.429)	0.221	(0.206, 0.237)	0.176	(0.167, 0.186)
Florida	-0.001	(-0.057, 0.059)	0.425	(0.413, 0.438)	0.396	(0.367, 0.426)	0.116	(0.110, 0.123)
Georgia	-0.108	(-0.150, -0.067)	0.453	(0.432, 0.474)	0.263	(0.244, 0.284)	0.196	(0.186, 0.207)
Hawaii	-0.072	(-0.101, -0.044)	0.416	(0.386, 0.443)	0.211	(0.199, 0.223)	0.163	(0.150, 0.177)
Idaho	-0.216	(-0.290, -0.152)	0.526	(0.507, 0.543)	0.410	(0.381, 0.442)	0.144	(0.135, 0.155)
Illinois	-0.230	(-0.267, -0.193)	0.334	(0.316, 0.350)	0.251	(0.232, 0.271)	0.149	(0.141, 0.157)
Indiana	-0.410	(-0.477, -0.346)	0.359	(0.338, 0.379)	0.414	(0.380, 0.449)	0.193	(0.184, 0.203)
Iowa	-0.199	(-0.244, -0.153)	0.331	(0.314, 0.346)	0.295	(0.274, 0.316)	0.150	(0.142, 0.158)
Kansas	-0.280	(-0.347, -0.208)	0.306	(0.283, 0.329)	0.303	(0.274, 0.332)	0.198	(0.187, 0.209)
Kentucky	-0.255	(-0.303, -0.209)	0.344	(0.325, 0.361)	0.303	(0.280, 0.329)	0.174	(0.165, 0.183)
Louisiana	-0.483	(-0.585, -0.387)	0.280	(0.249, 0.308)	0.373	(0.336, 0.411)	0.235	(0.220, 0.251)
Maine	-0.209	(-0.255, -0.162)	0.509	(0.475, 0.544)	0.373	(0.351, 0.396)	0.293	(0.275, 0.313)
Maryland	-0.160	(-0.198, -0.123)	0.378	(0.356, 0.399)	0.223	(0.204, 0.241)	0.191	(0.180, 0.201)
Massachusetts	-0.158	(-0.196, -0.119)	0.379	(0.360, 0.398)	0.224	(0.206, 0.241)	0.171	(0.162, 0.182)
Michigan	-0.651	(-0.762, -0.539)	0.427	(0.392, 0.462)	0.714	(0.661, 0.772)	0.346	(0.327, 0.366)
Minnesota	-0.048	(-0.076, -0.018)	0.359	(0.346, 0.373)	0.204	(0.189, 0.220)	0.127	(0.120, 0.134)
Mississippi	-0.223	(-0.261, -0.190)	0.352	(0.330, 0.374)	0.209	(0.194, 0.225)	0.191	(0.181, 0.202)
Missouri	-0.171	(-0.205, -0.138)	0.325	(0.307, 0.343)	0.216	(0.201, 0.232)	0.160	(0.151, 0.168)
Montana	-0.373	(-0.421, -0.319)	0.378	(0.356, 0.401)	0.338	(0.314, 0.365)	0.212	(0.201, 0.224)
Nebraska	-0.085	(-0.118, -0.054)	0.335	(0.320, 0.350)	0.198	(0.183, 0.214)	0.140	(0.133, 0.147)
Nevada	-0.294	(-0.372, -0.211)	0.634	(0.611, 0.656)	0.515	(0.474, 0.553)	0.204	(0.193, 0.216)
New Hampshire	-0.079	(-0.132 , -0.028)	0.484	(0.460 , 0.508)	0.277	(0.256 , 0.299)	0.205	(0.194 , 0.218)
New Jersey	-0.093	(-0.123 , -0.064)	0.347	(0.331 , 0.365)	0.216	(0.200 , 0.232)	0.156	(0.147 , 0.164)
New Mexico	-0.083	(-0.140 , -0.018)	0.351	(0.333 , 0.371)	0.243	(0.221 , 0.265)	0.155	(0.147 , 0.163)
New York	-0.100	(-0.131 , -0.069)	0.270	(0.259 , 0.281)	0.186	(0.172 , 0.201)	0.100	(0.095 , 0.106)
North Carolina	-0.122	(-0.159 , -0.084)	0.421	(0.406 , 0.437)	0.250	(0.232 , 0.271)	0.145	(0.137 , 0.153)
North Dakota	-0.040	(-0.066 , -0.015)	0.291	(0.278 , 0.303)	0.148	(0.137 , 0.161)	0.119	(0.114 , 0.126)
Ohio	-0.376	(-0.491 , -0.265)	0.297	(0.275 , 0.319)	0.688	(0.634 , 0.749)	0.206	(0.196 , 0.217)
Oklahoma	-0.352	(-0.404 , -0.300)	0.308	(0.286 , 0.329)	0.291	(0.265 , 0.319)	0.213	(0.202 , 0.225)
Oregon	-0.188	(-0.272 , -0.107)	0.536	(0.510 , 0.561)	0.465	(0.432 , 0.503)	0.146	(0.130 , 0.162)
Pennsylvania	-0.166	(-0.204 , -0.130)	0.304	(0.288 , 0.320	0.248	(0.230 , 0.266)	0.149	(0.140 , 0.158
Rhode Island	-0.368	(-0.422 , -0.312)	0.356	(0.333 , 0.382)	0.308	0.282 , 0.336)	0.225	(0.212 , 0.236)
South Carolina	-0.219	(-0.263 , -0.171)	0.417	(0.396 , 0.438)	0.293	(0.270 , 0.315)	0.202	(0.192 , 0.213)
South Dakota	-0.060	(-0.090 , -0.027)	0.382	(0.367 , 0.396)	0.197	(0.181 , 0.213)	0.135	(0.128 , 0.142)
Tennessee	-0.123	(-0.177 , -0.074)	0.362	(0.345 0.379)	0.243	(0.222 , 0.264)	0.137	(0.128 0.145)
Texas	-0.111	(-0.147 , -0.073)	0.406	(0.392 , 0.420)	0.216	(0.198 , 0.234)	0.134	(0.127 , 0.141)
Utah	0.001	(-0.030 , 0.032)	0.466	(0.453 , 0.480)	0.230	(0.214 , 0.245)	0.126	(0.119 , 0.133)
Vermont	-0.146	(-0.195 , -0.095)	0.481	(0.452 , 0.510)	0.335	(0.312 , 0.358)	0.243	(0.229 , 0.258)
Virginia	-0.081	(-0.106 , -0.054)	0.410	(0.394 , 0.426)	0.176	(0.163 , 0.191)	0.146	(0.137 , 0.154)
Washington	-0.132	(-0.171 , -0.092)	0.355	(0.341 , 0.370)	0.224	(0.205 , 0.243)	0.137	(0.130 , 0.144)
West Virginia	-0.186	(-0.231 , -0.139)	0.336	(0.320 , 0.353)	0.290	(0.267 , 0.314)	0.150	(0.143 , 0.158)
Wisconsin	-0.680	(-0.885 , -0.474)	0.396	(0.361 , 0.432)	1.100	(0.998 , 1.201)	0.325	(0.303 , 0.347)
Wyoming	-0.770	(-1.004 , -0.532)	0.366	(0.344 , 0.386)	0.989	(0.878 , 1.108)	0.213	(0.199 , 0.226)

Table G.2: PMS-IC estimations coefficient for the US-States coincident indices.

Following Berge and Jordà (2011) and Aastveit et al. (2016), we compute the receiver operating characteristic (ROC) to evaluate the accuracy of the filtered probabilities and the diffusion index to predict the NBER phases. We compare the PMS-IC with endogenous network filtered probabilities to a transformation of the diffusion index of the Federal Reserve of Philadelphia based on a normal CDF. Let d_t be the diffusion index at time t, the probability is computed as $\Phi((d_t - \hat{\mu})/\hat{\sigma})$, where μ and σ are the empirical mean and standard deviation of the index d_t and $\Phi()$ is the of a normal distribution. The results can be summarized by calculating the area under the ROC curve. Right panel in Fig. G.1 indicates that our PMS-IC with global and local interactions of cycles better matches the national recessions given by the NBER (higher areas). The national diffusion index also captures the same downturn points, however, our model shows that the degree of synchronization of the US states cycles seems to call deeper and longer slowdowns and recessions than the FED diffusion index. Left panel in Fig. G.1 confirms substantial discrepancy across states.

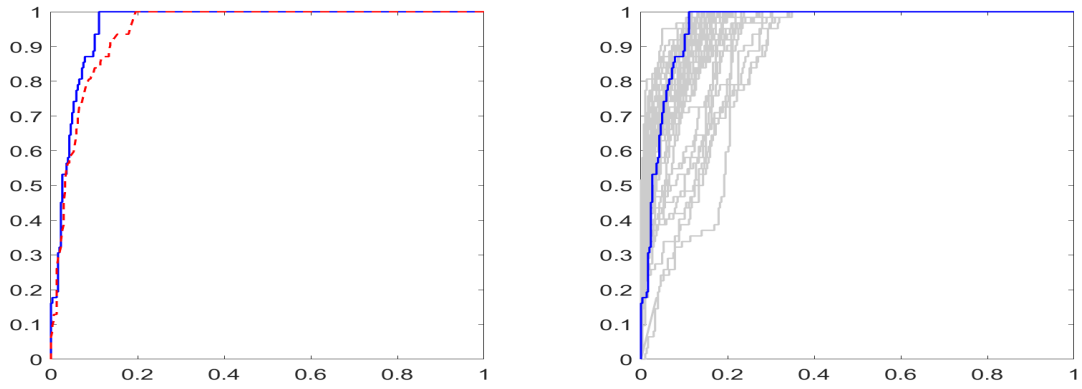


Figure G.1: Receiver operating characteristic: false positive rate (horizontal axis) versus true positive rate (vertical axis). Results given by the PMS-IC with endogenous network filtered probabilities (right panel, blue solid line), the diffusion index (right panel, red dashed line), and the country-specific filtered probabilities (left plot, gray dashed lines).