CENTRE FOR APPLIED MACRO – AND PETROLEUM ECONOMICS (CAMP)

CAMP Working Paper Series
No 1/2018

Markov Switching Panel with Network Interaction Effects

Komla Mawulom AGUDZE, Monica BILLIO, Roberto CASARIN and Francesco RAVAZZOLO

BI NORWEGIAN BUSINESS SCHOOL

Markov Switching Panel with Network Interaction Effects

Komla Mawulom AGUDZE[a], Monica BILLIO[a], Roberto CASARIN[a,*], Francesco RAVAZZOLO[b]

[a] *Department of Economics, University of Venice*
[b] *Free University of Bozen-Bolzano and CAMP, BI Norwegian Business School*

Abstract

The paper introduces a new dynamic panel model for large data sets of time series, each of them characterized by a series-specific Markov switching process. By introducing a neighbourhood system based on a network structure, the model accounts for local and global interactions among the switching processes. We develop an efficient Markov Chain Monte Carlo (MCMC) algorithm for the posterior approximation based on the Metropolis adjusted Langevin sampling method. We study efficiency and convergence of the proposed MCMC algorithm through several simulation experiments. In the empirical application, we deal with US states coincident indices, produced by the Federal Reserve Bank of Philadelphia, and find evidence that local interactions of state-level cycles with geographically and economically networks play a substantial role in the common movements of US regional business cycles.

Keywords: Bayesian inference, interacting Markov chains, Metropolis adjusted Langevin, panel Markov-switching.

*We thank for their useful comments conference and seminar participants at the 13th Dynare Conference in Tokyo.
Corresponding author: Francesco Ravazzolo. Address: Free University of Bozen-Bolzano, Italy. E-mail: `Francesco.Ravazzolo@unibz.it`

January 3, 2018

1. Introduction

The paper introduces a new dynamic panel model for large data sets of time series with series-specific Markov switching processes, which interact through a network. The interaction can be local, meaning in some neighbourhoods of a series-specific Markov chain; global, thus regarding all the series; or a combination of both. The local interactions among the Markov chains are based on a system of neighbourhoods: we assume that the state transition of each chain depends on the previous state of the series in its neighbourhood. Moreover, the neighbourhood system or the network structure can be known a priori by the researcher or inferred by the system.

Markov switching (MS) models have been extensively used in macroeconomics and finance to extract the different phases or regimes of the market. Originally, the MS model was applied to a univariate series Hamilton (1989), or a small set of series, and the Markov chain was assumed with constant transition probabilities, thus not varying over time. These assumptions have been challenged by the recent literature. First of all, the use of large database has been proved very important for forecasting, see e.g. Bańbura et al. (2010) in VAR framework, Stock and Watson (2014) for turning point application, and Casarin et al. (2015) for forecast combinations. Secondly, time-varying MS processes provide more accurate fitting of the cycle, see e.g. Kaufmann (2010). This is particularly true when different countries or states are grouped in panels and time-varying interacting mechanisms are necessary to reinforce the estimation of the regimes, see Kaufmann (2015) and Billio et al. (2016). The solution has been to focus on medium size panel (Billio et al. (2016)) or to use a number of Markov chains that are smaller than the number of series (e.g., Kaufmann (2015), Hamilton and Owyang (2012)).

We extend this literature and introduce a multivariate set-up with multiple chains where the dependence between the chains is modelled via a network. Therefore, we extend the literature to allow for state-specific business cycles and interaction effects that can also provide endogenous synchronization of the business cycle.

More in detail, we propose a convenient parametrization of the transition matrix. We assume that the transition probabilities of each chain do not only depend on their own past values but involve also the past regimes of other chains in the panel and we use a linear regression to specify such relationship. The proposed parametrization has several advantages. The first one relates to inference aspects. The multivariate logistic transforma-

tion, widely used in the literature to model transition probabilities, implies a non-linear transformation of the parameters which makes the inference task more difficult. In a Bayesian setting, the non-linear parametrization can lead to a poor performance of the Markov Chain Monte Carlo (MCMC) sampler used for posterior approximation (e.g., see Scott (2011)). Our endogenous linear time-varying transition model instead is not exposed to these inferential difficulties. The drawback of the linear parametrization relies on the constraints one needs to introduce on the parameters, but in our models these constraints can be easily handled through the use of standard prior distributions defined on the parameter space.

The second advantage of the linear assumption for the time-varying transition is that it allows us to provide some theoretical properties of our multiple-chain model under the assumption of a broader class of interaction mechanisms, which allow for idiosyncratic, local and global interactions (Föllmer and Horst (2001)). In particular, the global interaction parameter assesses the importance of common movement in all the Markov chains, while the local one captures the commovement of a chain with the neighbouring chains. In the application, we assume the neighbouring system is generated by an endogenously given network, that is a set of nodes (the chains) and a set of edges (pair of chains) defining the pairwise interaction between the chains. To the best of our knowledge this is the first paper which provides a Markov switching model with network interaction effects. In this sense, this study contributes to the recent and expanding stream of literature on network econometrics (e.g. Billio et al. (2012), Diebold and Yilmaz (2015), and Diebold and Yilmaz (2014)).

The paper also contributes to the literature on Markov switching dynamic panels by developing an efficient MCMC algorithm for the posterior approximation. The standard approach based on Metropolis-Hastings algorithm becomes quickly inefficient due to difficulties in setting the scale of the multivariate proposal distribution. The scale of the posterior distribution can change across different directions of the parameter space and the handtuning of the proposal scale can be a very difficult task. To solve this problem we consider the Metropolis adjusted Langevin (MALA) sampling method (Girolami and Calderhead (2011)) which is exploiting the information on the gradient of the target distribution. This method has been successufly applied in many fields such as statistics, physics and recently also econometrics by Burda and Maheu (2013), Burda (2015) and Virbickaite et al. (2015).

To assess the effectiveness of our model, we apply it to US regional business cycles. We collect monthly coincident indices from 50 states to assess the importance of the global and regional components in US business cycle synchronization. We identify the role of the global and regional interactions in the cycle co-movement and are able to shed light on how the co-movements propagate to the rest of the economy (see Hamilton and Owyang (2012), Camacho and Leiva-Leon (2014), Leiva-Leon (2014)). More specifically, we consider a panel with a network interaction based on the division into statistical regions defined by the geographical classification of the United States Census Bureau, precisely the West, the Midwest, the South and the North-east, and economic connections among States based on 10-K fillings on the Securities and Exchange Commission's EDGAR system. Therefore, such model accounts for both geographical and economic interconnections of the states. We compare it to a panel with only a global interaction among the states and a model that considers geographical proximity interaction. Our findings show that the model with the network of states, and therefore global and regional interactions, receives higher support from the data than the other two models. This confirms evidence that US states react differently to business cycle shocks. Moreover, we show that the interaction of the network of US states cycles plays a key role in making the slowdowns and the recessions deeper and longer, differently of what is predicted by an aggregate index. Indeed, our model can both measure the effects of state-specific-recessions and strengthen the consequences of a national recession via endogenous synchronization of regional cycles. For example, our model can proxy a mobile labour force that moves from one state to another one depending on economic conditions, helping to identify local features, but also amplifying national cycles. Finally, we confirm at state level what has been documented at country level, i.e. that the uncertainty is higher during deep recessions and this can slow the recovery itself.

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 5 studies the model in simulation exercises and section 6 applies it to regional US business cycle. Section 7 concludes.

2. Panel Markov switching with interacting chains

In our panel MS model with interacting chains (PMS-IC), we assume each series $\{X_{it}\}$, $t = 1, \ldots, T$ in a panel of N units $i = 1, \ldots, N$, is a conditionally linear and Gaussian process with mean and variance driven by a unit-specific Markov chain S_{it} which takes value in the finite set $\{0, \ldots, K-1\}$.

The measurement equation is written as:

$$X_{it} = \sum_{k=1}^{K} \mathbb{I}_{\{k\}}(S_{it}) \left[\Psi_{ik}' Z_{it} + \sigma_{ik}\varepsilon_{it}\right], \quad \varepsilon_{it} \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \tag{1}$$

where $Z_{it}' = (1, Z_{2it}, \ldots, Z_{mit})$ is a vector of covariates and Ψ_{ik} a m-dimensional parameters vector. K represents the number of unobserved latent regimes and the symbol $\mathbb{I}_E(X)$ is the indicator function which takes value 1 if $X \in E$ 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:

$$P_{it,lk} = \mathbb{P}\left(S_{it+1} = l | S_{it} = k, S_{-i,t}\right) \tag{2}$$

which represents the conditional probability that unit i moves to the regime $l \in \{1, \ldots, K\}$ at time $t+1$. $S_t = (S_{1t}, S_{2t}, \ldots, S_{Nt}) \in \mathcal{S}$ includes all configurations at time t, with $\mathcal{S} = \{0, \ldots, K-1\}^N$ and $S_{-i,t} = \{S_{jt}, j = 1, \ldots, N; \quad j \neq i\}$.

Following Kaufmann (2010) and Billio et al. (2016), we assume that the transition probabilities of each chain do not only depend on their own past values, but involve also the past regimes of other chains in the panel. In this paper we propose a new interaction mechanism:

$$P_{it+1,lk} = \alpha p_{lk} + \beta m_{i,k}\left(S_t\right) + \gamma m_k\left(S_t\right) \tag{3}$$

with, $0 < \alpha \leqslant 1$, $\quad 0 \leqslant \beta < 1$, $\quad 0 \leqslant \gamma < 1$, interactions parameters such that: $\quad \alpha + \beta + \gamma = 1$ and p_{lk} fixed transition probability parameter such that $\sum_{k=1}^{K} p_{lk} = 1$. The second term on the right-hand side represents the local interactions mechanism and the third term defines the global interactions. In our model, $m_{i,k}\left(S_t\right)$ is the local interaction factor and measures the proportion of chains belonging to the neighbourhood of chain i which are

in the state k at time t, that is:

$$m_{i,k}\left(S_t\right) = \frac{1}{|N(i)|} \sum_{j \in N(i)} \mathbb{I}_{\{k\}}(S_{jt}) \tag{4}$$

where $N(i) \in \mathcal{N}$ is the neighbourhood of the chain i, and $\mathcal{N} = \{N(i)\}_{i=1,\ldots,N}$ is a neighbourhood system with $N(i)$ subset of $D = \{1, \ldots, N\}$ (see Brémaud (2013), Chap. 7). $\mathcal{G} = (\mathcal{N}, D)$ is called graph or topology.

The global interaction factor $m_k\left(S_t\right)$ is given by the proportion of chains in regime k at time t that is:

$$m_k\left(S_t\right) = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{\{k\}}(S_{jt}) \tag{5}$$

These specifications of the global and local interactions allow us to assess the dependence among the time series in the panel. To model dependence through interaction effects is appealing in many contexts. In financial econometrics, the interactions represent linkages between financial institutions and the PMS-IC synchronization has the interpretation of contagion effects. One way of capturing these effects is via the network of connections of the individual series. Allen and Babus (2009) provide a review of network model application in economics and finance. Vesper (2013) complements existing measures of systemic risk, by introducing the combination of MS models with a latent network structure of financial institutions. In macro perspective, our PMS-IC has the potential to analyse the co-movement of regional business cycles. It does not only help to characterize the unit-specific cycles but also shows the importance of a global component (global interactions), a regional component (local interactions) and fixed time independent transitions in business cycles synchronization. In this vein, Kose et al. (2003, 2008) document the common dynamic properties of the world business cycles fluctuations employing a Bayesian dynamic latent factor model. Their results suggest that the regional components play only minor role in explaining cycles fluctuations. However, recent studies suggest that the world component is not enough to explain the business cycle synchronization. Francis et al. (2017) find that when the regional component is defined differently from simple geography, its effect becomes more important. Aastveit et al. (2015, 2016) explicitly introduce regional factors into a global dynamic factor model. They find that both the global and

the regional factors are relevant in explaining the business cycle variation. Leiva-Leon (2014) proposed a new model that combines several bivariate MS models and network of synchronisation in order to create a link of interdependencies business cycles. However, the model cannot assess the importance of global and regional components in the cycles fluctuations.

The behaviour of unit i of the panel at time $t+1$ is influenced by an idiosyncratic interaction, an empirical average of all the system at t and the situation in some neighbourhood $N(i)$ at t. For instance, an individual that has in its network at t a high proportion of series in regime k will tend to transit in regime k or will remain in regime k at time $t+1$. α reflects the intensity of idiosyncratic interaction from time t to time $t+1$ and γ reflects the intensity of global interactions between the collection of time series from time t to time $t+1$. Parameter β captures the level of local interactions.

The transition probabilities satisfy the condition:

$$\sum_{k=1}^{K} P_{it,lk} = 1, \quad \forall i \in \{1,\dots,N\}, \quad \forall l \in \{0,\dots,K-1\}, \quad \forall t \in \{1,\dots,T\}$$

Therefore, the linear parametrization of the regime switching transition matrix presents twofold advantages. First, the parametrizations allow for idiosyncratic, global interactions and interaction in the neighbourhood of the Markov chains. Hence, if the population of time series presents high global interactions at time t then the panel MS model will more likely exhibit episodes of contagion at time $t+1$. Secondly, formulating linearly over unit specific interaction and global interaction, the new endogenous transition matrix can be seen as solution to the usual critique to exogenous transition matrix and as well as generalisation of the fixed transition matrix. In fact, if the parameter α which measure the intensity of unit specific interaction is equal to 1 then we are back to the case of exogenous fixed transition matrix.

3. Properties of the model

The use of linear parametrization to model the evolution of the endogenous time-varying transition matrix insures the convergence of the Markov processes transition probabilities to unique ergodic probabilities.

Let $\mathcal{S} = \{0,1,\dots,K-1\}^N$ be the finite set of all configurations of $S_t = (S_{it})_{1 \leqslant i \leqslant N}$ with $S_{it} \in \{0,1,\dots,K-1\}$, $i = 1\dots,N$. 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 behaviour of the set of chains as t tends to infinity and to give an interpretation to the parameters of the transition probabilities.

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

$$\overline{m}\left(S_t\right) = \left(\overline{m}_0\left(S_t\right), \ldots, \overline{m}_{K-1}\left(S_t\right)\right)\prime$$

and the proportion of regime in some finite neighbourhood $N(i)$ of i:

$$\overline{m}_i\left(S_t\right) = \left(\overline{m}_{i0}\left(S_t\right), \ldots, \overline{m}_{iK-1}\left(S_t\right)\right)\prime$$

where

$$\overline{m}_k\left(S_t\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{k\}}(S_{it}), \quad k = 0, 1, \ldots, K-1$$

$$\overline{m}_{ik}\left(S_t\right) = \frac{1}{|N(i)|} \sum_{j\in N(i)} \mathbb{I}_{\{k\}}(S_{jt}), \quad k = 0, 1, \ldots, K-1$$

with S_{it} is a Markov chain with transition probabilities

$$\pi_i(S_{it+1} = k|S_t) = \alpha p_{S_{it}k} + \beta m_{i,k}\left(S_t\right) + \gamma m_k\left(S_t\right)$$

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

The theoretical relationship between the global interaction factor and the fixed transition probability matrix is given in Proposition 1.

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

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{k\}} S_{it+1} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \pi_i\left(S_{it+1} = k|S_t\right) \quad \Pi(\cdot|S_t) - a.s., \tag{6}$$

the sequence of empirical averages satisfies almost surely the recurrence relation

$$\overline{m}_k\left(S_{t+1}\right) = \alpha \sum_{j=0}^{K-1} p_{jk} \overline{m}_j\left(S_t\right) + (1-\alpha)\overline{m}_k\left(S_t\right) \tag{7}$$

and the global interaction process $\{\overline{m}\,(S_t)\}_{t\in\mathcal{N}}$ 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.

Finally, a second proposition on the convergence of the Markov chain process of the population of time series with time-varying transition probabilities is present in the following with Proof in Appendix A.2.

Proposition 2. The process $\{S_t\}_{t\in\mathcal{N}}$ converges in law to the unique product kernel

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

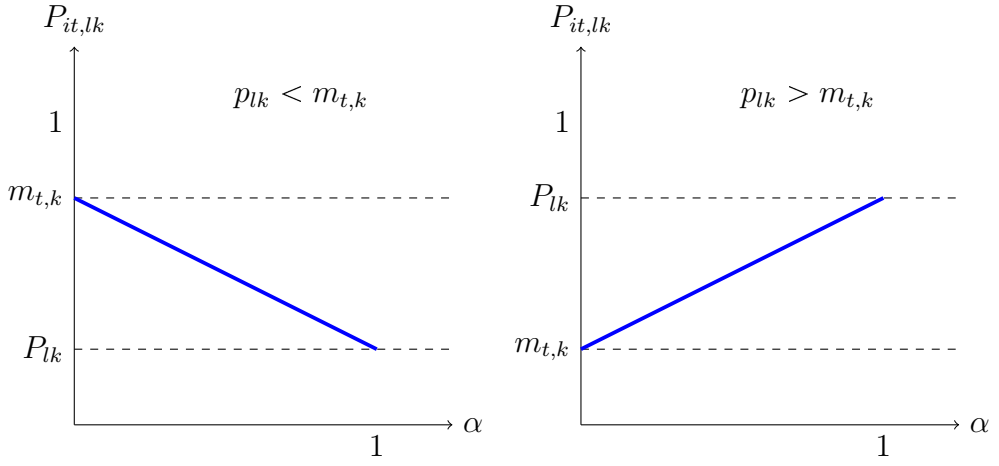
3.1. Examples of model outputs

Figure 1 shows how the endogenous transition probabilities vary with the parameter α. If the fixed component of the transition probabilities is larger than the global interaction term at time t (Figure 1 right) then the probability to stay in one state or to switch to another state is increasing with α.

In the opposite case, if the fixed component of the transition probabilities is less than the global interaction at time t (Figure 1 left) then the time-varying transition probability is decreasing with respect to α. It comes out that in presence of important global interactions and persistence over time, the PMS-IC model exhibits a large scale globalisation of episodes of one regime.

3.1 *Examples of model outputs*

Figure 1: Different shapes of the time varying probabilities $P_{it,lk} = \alpha P_{lk} + \gamma m_{t,k}$ at a given time t as function of α. Note that in this case the presence of local interaction is not allowed, $\beta = 0$.



In order to give a qualitatively description of the dynamic behaviour of our PMS-IC model, we provide some simulated examples. We isolate the contribution of the global and local interaction mechanism and specify six parameter settings which are summarized in Table 1.

Settings label	α	β	γ
Setting 1 (No interaction)	1.00	0.00	0.00
Setting 2 (Weak global interaction)	0.70	0.00	0.30
Setting 3 (Strong global interaction)	0.30	0.00	0.70
Setting 4 (Weak local interaction)	0.70	0.30	0.00
Setting 5 (Strong local interaction)	0.30	0.70	0.00
Setting 6 (Both local and global interaction)	0.50	0.25	0.25

Table 1: Parameter values of idiosyncratic, global and local interaction mechanisms.

The difference between the six settings is on the choice of the underlying parameters α, β and γ. We distinguish six cases. In Model 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 for Model 2 and a stronger one for Model 3. On the other side, we consider a weak global interaction among the Markov chains for Model 4 and a stronger one for Model 5. The global and local effect are simulated in Model 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, $K = 3$):

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

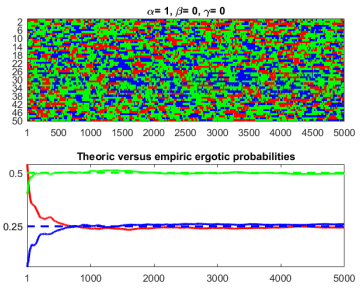
for $i = 1, \ldots, 50$ and $t = 1, \ldots, 5000$, and the fixed transition component:

$$P = \begin{pmatrix} 0.98 & 0.02 & 0.00 \\ 0.01 & 0.98 & 0.01 \\ 0.00 & 0.02 & 0.98 \end{pmatrix} \tag{8}$$

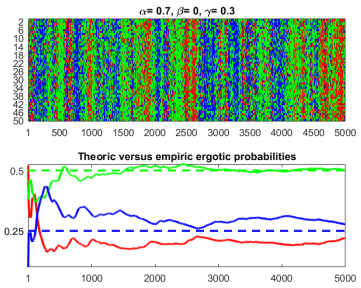
$\mu_{i1} = -2, \quad \mu_{i2} = 0, \quad \mu_{i3} = 2, \quad \sigma_{i1} = .3, \quad \sigma_{i2} = .05, \quad \sigma_{i3} = .3$

Note that the ergodic probability associate with P is $\pi_i = (.25 \quad .5 \quad .25)$.

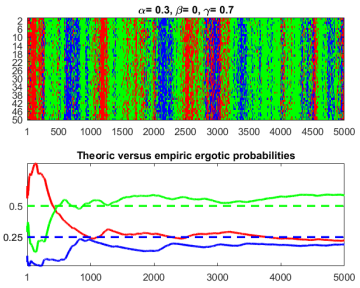
3.1 *Examples of model outputs*



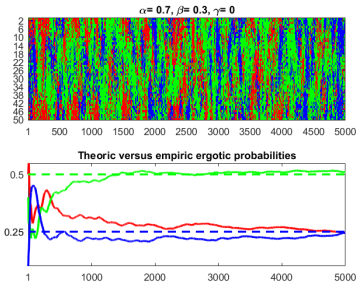
(a) Setting 1: no interaction



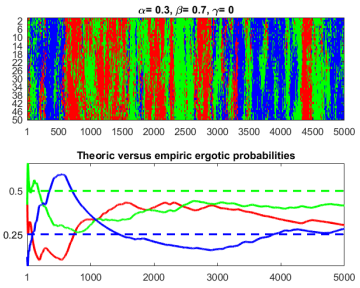
(b) Setting 2: weak global interactions



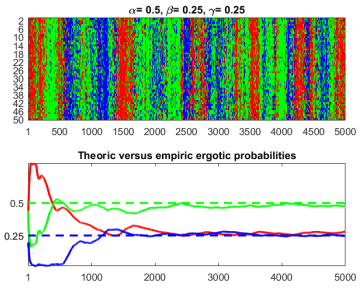
(c) Setting 3: strong global interactions



(d) Setting 4: weak local interactions



(e) Setting 5: strong local interactions



(f) Setting 6: both local and global interactions

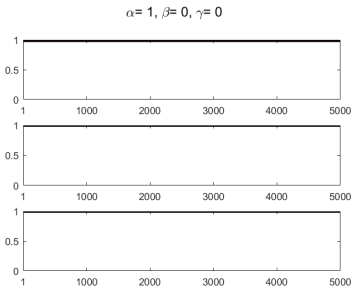
Figure 2: Population of 50 time series for different settings (different panels). For each setting, the top plot is a heat-map of the time series. In each plot, colors blue mean that the series is in expansion regime; colors green refer to moderate expansion regime and colors red refer to recession regime. For each setting: the bottom plot describes the evolution of the global interaction chain of the linear time varying transition matrix over simulations together with the horizontal lines given by the elements of theoretic ergotic probabilities of the fixed transition matrix.

Figure 2 shows the different impact of the parameter values on the level of synchronization among the series. Indeed, the second aspect we study is

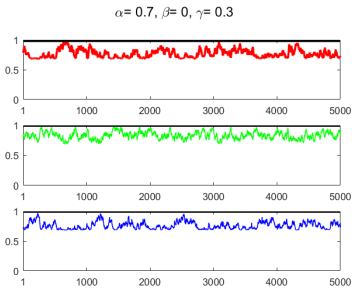
the convergence of the global interaction chain to the invariant transition matrix of the fixed transition matrix. Figure 2 highlights the ability of our PMS-IC model to account for various degree of synchronization of chains. The value of α, β and γ corresponding to our six parameters settings were carefully chosen in order to represent a wide variety of possible interactions. A look at the heat-maps shows that the level of synchronization has augmented with the level of the local and global parameters. In all the different cases, the convergence to the ergodic of the fixed transition probabilities is reached with different speed.

The shape of the time varying transition is given in Figure 3. Without loss of generality, we present only the simulation of the probabilities to stay in regime 1, regime 2 and regime 3 for the first unit in the panel. The evidence emerging from these plots is that the global and local parameters play an important role: the higher the levels of these parameters the higher the volatility of the time varying transition probabilities is.

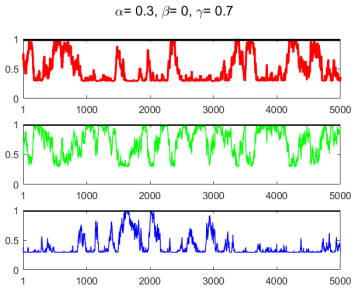
3.1 *Examples of model outputs*



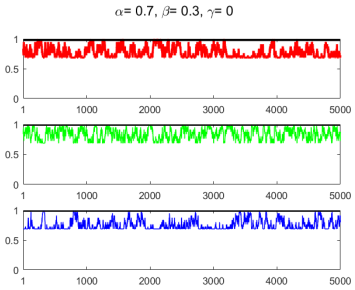
(a) Setting 1: no interaction



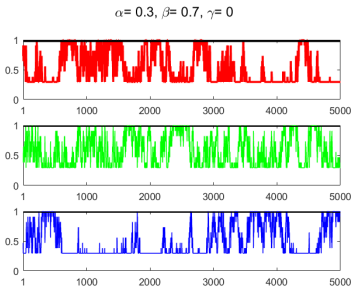
(b) Setting 2: weak global interactions



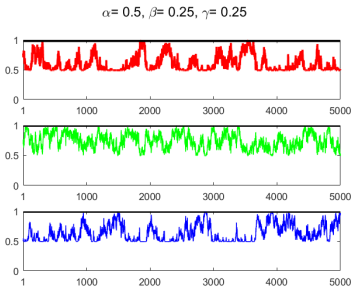
(c) Setting 3: strong global interactions



(d) Setting 4: weak local interactions



(e) Setting 5: strong local interactions



(f) Setting 6: both local and global interactions

Figure 3: Evolution of elements of both fixed and time varying transition matrices for the first chain of the panel for each setting. Colours 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 (8).

We 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) to assess the impact of the local and global parameters on the synchronization of chains.

3.1 *Examples of model outputs*

This index describes the fraction of times that two chains i and j spend in the same phase. Let us represent it by the following equation:

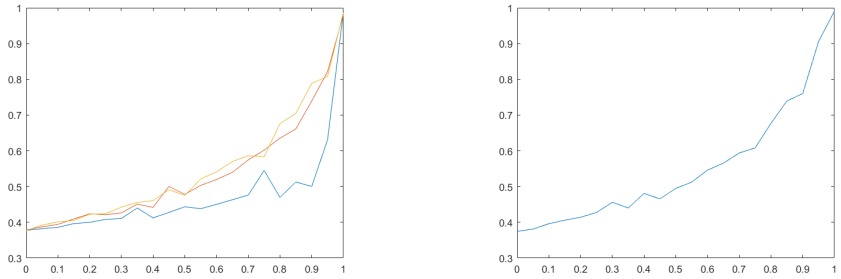
$$c_{i,j} = \frac{1}{T} \sum_{t=1}^{T} \sum_{k=0}^{K-1} \mathbb{I}_{\{k\}}(S_{it}) \mathbb{I}_{\{k\}}(S_{jt}) \tag{9}$$

The relationship between the local and global interactions parameters and the level of synchronization among the chains can be measured by:

$$\overline{c} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} c_{i,j} \tag{10}$$

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.

A panel of 200 series is simulated from our PMS-IC using the settings presented for the underlying specification detailed in equation (8). A system of three neighbourhoods is designed for the model with only local interactions. The neighbours selected are made up with 4, 16 and 24 units. Figures 4a and 4b reveal that the rate of synchronization increases with the size of the neighbourhood and the value of β as well as γ. Hence, the level of synchronization is positively related to the parameter β (resp. γ) that reflect the importance of common movement with the unit-specific chains in the neighbourhood (resp. of the importance of common movement in all the unit-specific chains).



(a) $0 \leqslant \beta \leqslant 1$ (b) $0 \leqslant \gamma \leqslant 1$

Figure 4: Relationship between the local parameter β, the global parameters γ and the synchronization of chains. The vertical axis represents the value taken by $\overline{c}$ as a function of β the local interaction parameter for panel (a) and as a function of γ the global interaction parameter for panel (b). The horizontal axis represents the value of α.

4. Bayesian inference

4.1. Likelihood function and prior distributions

Let $\theta = (\mu_1, \ldots, \mu_K, \sigma_1, \ldots, \sigma_K, vec(P_1)', \ldots, vec(P_N)', \alpha, \beta, \gamma)$ be the vector of parameters with $\mu_l = (\mu_{1l}, \ldots, \mu_{Nl})$, $\sigma_l = (\sigma_{1l}, \ldots, \sigma_{Nl})$. Let us define $\xi_{k,it} = \mathbb{I}_{\{k\}}(S_{it})$, and $\xi_{k,it}$ indicates the regime k that the observation X_{it} belongs to. By using the sequential factorization of the likelihood, the complete likelihood of the PMS-IC model is:

$$\begin{aligned}\mathcal{L}(X_{1:T}, S_{1:T} \mid \theta) &= \prod_{i=1}^{N}\prod_{t=1}^{T}\prod_{l=1}^{K}\prod_{k=1}^{K} f(X_{it}|S_{it},\theta)^{\xi_{l,it}} P_{it,lk}^{\xi_{l,it}\xi_{k,it-1}} \\ &= \prod_{i=1}^{N}\prod_{t=1}^{T}\prod_{l=1}^{K}\prod_{k=1}^{K} (2\pi\sigma_{il}^2)^{-\frac{\xi_{l,it}}{2}} \exp\{-\frac{\xi_{l,it}}{2\sigma_{il}^2}(X_{it}-\mu_{il})^2\} \\ &\qquad (\alpha P_{lk} + \beta m_{it,k} + \gamma m_{t,k})^{\xi_{l,it}\xi_{k,it-1}}\end{aligned} \tag{11}$$

In order to complete the specification of the Bayesian model, we discuss the prior choice. A variety of priors can be used to estimate the panel Markov switching model. We consider conjugate priors which are based on proper distributions. We assume conjugate independent priors for the unit specific parameters:

$$\mu_{il} \sim \mathcal{N}(m_{il}, \tau_{il}^2) \tag{12}$$

$$\sigma_{il}^2 \sim \mathcal{IG}(\alpha_{il}, \beta_{il}) \tag{13}$$

$$(P_{l1}, \ldots, P_{lK}) \sim \mathcal{D}ir(\delta_1, \ldots, \delta_K) \tag{14}$$

$$(\alpha, \beta, \gamma) \sim \mathcal{D}ir(\varphi_1, \varphi_2, \varphi_3) \tag{15}$$

with $l = 1, \ldots, K$ and $i = 1, \ldots, N$, where $\mathcal{IG}(\alpha, \beta)$ denote the inverse gamma distribution with parameters: α and β and $\mathcal{D}ir(\delta_1, \ldots, \delta_K)$ the K dimensional Dirichlet distribution with parameters: $\delta_1, \ldots, \delta_K$.

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 parameters of Markov switching model resulting is invariant to permutations in the labelling of the parameters, if this latter follow 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

4.2 Posterior simulation

Bayesian inference, see among other Celeux (1998), Frühwirth-Schnatter (2001), and Frühwirth-Schnatter (2006). One way to address the label switching problem is to consider under some specific condition the permutation sampler proposed by Frühwirth-Schnatter (2001). 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 identification restrictions that $\mu_{i1} \leqslant \mu_{i2} \leqslant, \ldots, \leqslant \mu_{iK}$.

4.2. Posterior simulation

The posterior distribution in a general form follows the following rule:

$$\pi(\theta \mid X, S) \propto \mathcal{L}(X, S \mid \theta) p(S \mid X, \theta) \pi(\sigma) \pi(\mu) \tag{16}$$

where $\pi(\theta \mid X)$ is the posterior distribution, $\mathcal{L}()$ the completed likelihood function and $\pi(\theta)$ the prior distribution. We develop a sampling algorithm based on conditional posterior distributions. Full details of the Algorithm are provided in Appendix C. The model in equation (1) is estimated by adapting the multi-move Gibbs-sampling procedure for Bayesian estimation of Markov switching models presented in Frühwirth-Schnatter (2006). The Gibbs sampler iterates according to the following steps:

1. Draw $S_i^{(d)}$ from $f(S_i^{(d)} \mid X, \theta^{(d-1)})$, $i = 1, \ldots, N$.
2. Draw $(\alpha, \beta, \gamma)^{(d)}$, from $f((\alpha, \beta, \gamma)^{(d)} \mid X, S^{(d)}, (p_{it,l1}, \ldots, p_{it,lK})^{(d-1)})$. where $S^{(d)} = (S_1^{(d)}, \ldots, S_N^{(d)})$.
3. Draw $(p_{l1}, \ldots, p_{lK})^{(d)}$ from $f((p_{l1}, \ldots, p_{lK})^{(d)} \mid X, S^{(d)}, (\alpha, \beta, \gamma)^{(d-1)})$, $i = 1, \ldots, N$, $l = 1, \ldots, K$.
4. Draw $\mu_{il}^{(d)}$,from $f(\mu_{il}^{(d)} \mid X, S_i^{(d)}, \sigma_{il}^{(d-1)}$, $i = 1, \ldots, N$, $l = 1, \ldots, K$.
5. Draw $\sigma_{il}^{(d)}$, from $f(\sigma_{il}^{(d)} \mid X, S_i^{(d)}, \mu_{il}^{(d-1)})$ $i = 1, \ldots, N$, $l = 1, \ldots, K$.

4.3. Gibbs iterations mains issues

Firstly, the multi-move sampling of the hidden state cannot be directly implemented:

1. the full conditional posterior distribution of the unit specific hidden state depends on proportionality factor that need to be taken into account by introducing Metropolis-Hastings adjustment, where
2. the proposal distribution is identical to forward filtering backward sampling (FFBS) of the states.

Secondly, the standard sampler based on independent proposal poorly estimates the parameter (α, β, γ):

1. the posterior distribution of (α, β, γ) is prior dependent, and
2. a straightforward implementation of Metropolis-Hastings algorithm with the proposal distribution equal to the prior distribution becomes inefficient, resulting in high rate of acceptance followed by poor mixing of the chain.

In the following section we present a Metropolis adjusted Langevin sampling algorithm as an efficient option to solve the issues described above when using independent proposal. In this case, the Gibbs sampler changes slightly according to the following steps:

1. Draw $S_i^{(d)}$ from Metropolis-Hastings adjusted FFBS (see Appendix D).
2. Draw $(\alpha, \beta, \gamma)^{(d)}$, from Metropolis adjusted Langevin algorithm (MALA) .
3. Draw $(p_{l1}, \ldots, p_{lK})^{(d)}$ from $f((p_{l1}, \ldots, p_{lK})^{(d)} \mid X, S^{(d)}, (\alpha, \beta, \gamma)^{(d-1)})$, $i = 1, \ldots, N$, $l = 1, \ldots, K$ by using a Metropolis-Hastings algorithm.
4. Draw $\mu_{il}^{(d)}$,from $f(\mu_{il}^{(d)} \mid X, S_i^{(d)}, \sigma_{il}^{(d-1)}$, $i = 1, \ldots, N$, $l = 1, \ldots, K$.
5. Draw $\sigma_{il}^{(d)}$, from $f(\sigma_{il}^{(d)} \mid X, S_i^{(d)}, \mu_{il}^{(d-1)})$ $i = 1, \ldots, N$, $l = 1, \ldots, K$.

We simulate (α, β, γ) from $f\left(\alpha, \beta, \gamma | X_{1-T}, S_{1-T}, \theta_{-(\alpha,\beta,\gamma)}\right)$ where the prior is chosen to be Dirichlet $\mathcal{D}ir(\varphi_1, \varphi_2, \varphi_3)$. Since by definition $(\alpha, \beta, \gamma) \in [0; 1]^3$; when dealing with random walk proposals we need to use transformation of α, β and γ to the real line which introduces a Jacobian factor into the acceptance probability of the MALA. We assume

$$\alpha = \frac{1}{1 + exp(-\alpha')}, \beta = \frac{1}{1 + exp(-\beta')}, \gamma = \frac{1}{1 + exp(-\gamma')}$$

For the MALA we need the partial derivatives of the complete log-likelihood with respect to the transformed parameters (see Appendix E).

The proposal mechanism of the MALA is given by the following equation

$$\omega^* = \omega^n + \frac{\epsilon^2}{2} M \nabla_\omega l\left(\omega^n\right) + \epsilon \sqrt{M} z^n \tag{17}$$

where $\omega = (\alpha, \beta, \gamma)'$, $l = log\{\mathcal{L}(X, S, \theta)\}$ is the complete data joint log-likelihood, 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 invariant distribution $p(\omega)$ is ensure by employing Metropolis acceptance probability after each integration step. The proposal density is

$$q(\omega^* | \omega^n) = \mathcal{N}\{\omega^* | \mu(\omega^n, \epsilon), \epsilon^2 \mathcal{I}\}$$

with $\mu(\omega^n, \epsilon) = \omega^n + \frac{\epsilon^2}{2} \nabla_\omega L(\omega^n)$ and acceptance probability of standard form if 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 Markov process.

5. Simulation experiments

We simulate data in the six parameter settings described in section 3.1 (see Table 1 for parameter settings) in order to assess the efficiency of the MCMC algorithm for the posterior approximation.

We assess the efficiency using the mean square error (MSE) for the parameters and the hidden states.

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

MSE is evaluated on 5000 iterations after convergence (1000 draws). Table 2 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.

6. US States coincident indices

We apply our model to US States business cycles and network of States. Not all the US States are identical and, for example, the recent US financial crisis has shown that some States were heavily affected by the crisis, e.g. Michigan, and other States were basically not affected, such as Texas that benefited from high oil prices in 2009 and 2010. Furthermore, the US labour force is often considered to be mobile and keen to change from one region to another one when economic situations differ across places. This can strengthen the regional effect by creating network of States that attract population and other States that loose population. We believe our model is very suitable to investigate such mechanisms.

We work with the US states monthly coincident indices datasets produced by the Federal Reserve Bank of Philadelphia. The database covers 50 States of US and our sample dates from October 1979 to June 2010. For each State an index of business cycle diffusion is available. The State-level diffusion indices are constructed on the scale -100 to 100 where a negative number is related to the spread of national recession and the positive one to national expansion. Owyang et al. (2005) apply a Bayesian univariate independent Markov switching model to the same dataset.

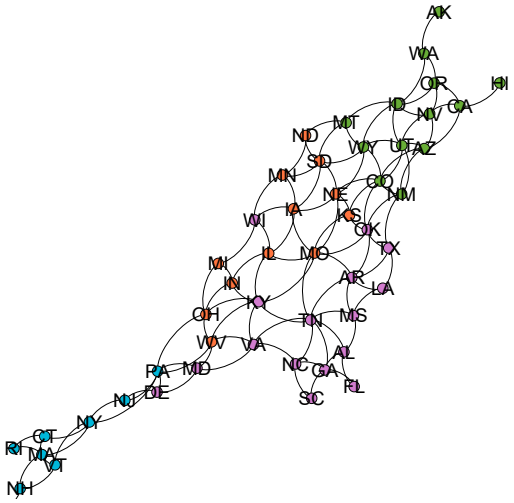
In order to check the importance of the local and global interactions across the US states, we consider three settings. The first setting assumes only global interactions among the US states coincident indices. This model implies that accounting for specific regional business cycle effects is not important for the aggregate country cycle. The second setting considers a local chain for each State composed by neighbouring States. This group is then different for each State and implies that networks among boarder States is a relevant information for the country cycle. The third setting considers the US in four large statistical regions defined by the United States Census Bureau, the West, the Midwest, the South and the North-east and creates local chains based on the economic relationship among them. 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 fillings on the

Securities and Exchange Commission's EDGAR system to identify economic connections among US states.[1] In this model, the network is among four different geographical regions and possible exchanges among nodes of some of the four regions.

To sum up, the first setting discards regional business cycles and the last two settings allow different level of grouping, the first one only geographical proximity, the second one geographical and economic networks, and can provide several interesting findings on how each State relates with neighbours and similar entities. Figure 5 provides the structure of the networks in the two models. The average number of connections (average degree) in the geographical proximity network is 4.36; on the contrary, the economic network is more dense with average degree of 13.12. Moreover, the former network exhibits one connected component; whereas the latter one shows three connected components. Following both degree and eigenvector centrality the most central nodes in the economic network are Rhode Island (RI) and Vermont (VT), whereas in the geographical proximity network are Missouri (MO) and Tennessee (TN) following the degree and Colorado (CO) and Missouri following eigenvector centrality (see Table 4 in Appendix F). In the economic network Rhode Island and Vermont plays a central role in connecting West and North-east regions. We also notice that the components of the second network do not necessarily correspond to the standard United States Census Bureau geographical regions because we also account for economic connections among States. For example, Wisconsin (WV) in the Midwest is connected to West region; Washington (WA) in the West is connected only to states in the South region; Wyoming (WY) in the West is connected only to states in the Midwest region.

[1]The federal securities laws require companies issuing publicly publicly traded securities to disclose information on an ongoing basis. Notably, Section 13 or 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 assumption about the nature of the economic connections but count the number of times a U.S. state is cited in items 1, 2, 6 and 7 of the 10-K fillings. Such items are design to focus on firm's economic activities. By comparing information for firms in two different states, we derive an economic network between the two states.

Panel A: Geographical Proximity



Panel B: Economic Network

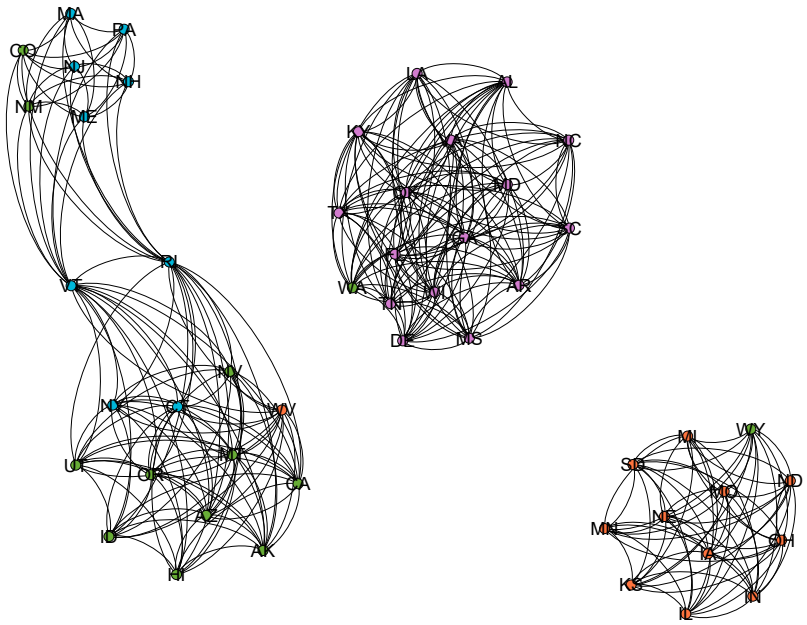


Figure 5: Panel A: Geographical proximity network; panel B: economic network. Lines indicate connections between pairs of nodes (colored circles). Thicker circles mean that a node has a larger number of connections. Colors indicate USCB geographical regions: light blue for North-east; orange for Midwest; pink for South; green for West. Label for each state is provided in the nodes.

To discriminate among the models $K = 1, 2, 3$ we use a Bayes factor approach based on the marginal likelihood:

$$BF_{ij} = \frac{p(y|K = i)}{p(y|K = j)}$$

where $p(y|K = i)$ and $p(y|K = j)$ is the marginal likelihood for models i and j, and ij. If $BF_{ij} > 1$, model i will be preferred; if $BF_{ij} < 1$, model j will be preferred. Therefore, the model with maximum marginal (log-)likelihood among $K = 1, 2, 3$ is preferred.

The results presented in Table 3 favour the hypothesis that both global and local interactions of the cycles prevail. The model with network of geographical and economic regions presents the highest marginal log-likelihood; the model with geographical proximity the second highest, even if the difference with model 2 is large. Therefore, data support the modeling of regional chains and the regional component becomes substantially more important when defined differently from a geographical proximity perspective and based of network of States depending on economic regions with stronger connections.

	Marginal log-likelihood	α	β	γ
Setting 1: Only global interactions	879.726	0.724	–	0.276
		(0.712, 0.733)	–	(0.267, 0.288)
Setting 2: Regions defined by the neighbouring States	1.1990e+03	0.1590	0.8293	0.0116
		(0.1492, 0.1700)	(0.8159, 0.8405)	(0.0035, 0.0279)
Setting 3: Regions defined by economic relationship	8.6951e+03	0.0281	0.9702	0.0016
		(0.0260, 0.0307)	(0.9668, 0.9726)	(0.0009, 0.0031)

Table 3: Comparison in terms of marginal log likelihood between different PMS-IC on the US states regional data. In parenthesis the 95% high posterior density interval

We investigate what are the consequences of the network structure and their interactions on global and States cycles. Figure 6 plots the global interaction of recessions of the US-States business cycles obtained from model 3 together with the diffusion index of the US national cycle phases published

by the Federal Reserve Bank of Philadelphia. Our proposed PMS-IC with global and local interactions of cycles is able to capture all the national recessions given by the NBER. The national diffusion index also captures the same downturn points, however, our model shows that the degree of synchronization of the US states cycles plays an important role in making the slowdowns and the recessions deeper and longer than the FED diffusion index reveals. Our model is also faster in calling recession than the aggregate diffusion index and in the five recessions defined by the NBER our model assigns 70%-80% probability of recession in the first quarter of the defined recession period. The diffusion index in all cases gives lower probability at the same period. In exiting the recessions, the two indices provide more similar evidence.

Few more words require the recent US financial crisis. The US financial crisis has been different from the previous ones for its duration, deepness, real and financial consequences. It has been a global and wide spread crisis across the States. The global and regional components explain a substantial portion of the cyclical movements for most States. The linkages between the financial institutions increase the strength of the connections of the States. Our global and local interaction factors catch such effects and show that global and local connections can strengthen the consequences of a national recession.

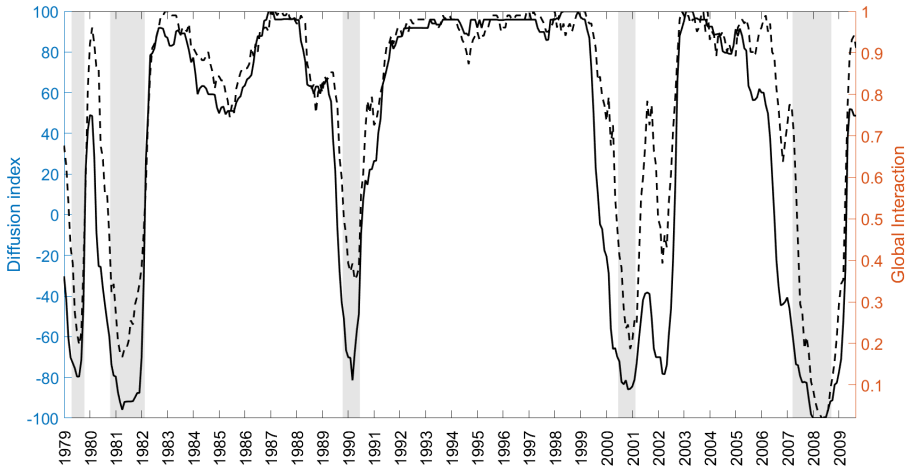


Figure 6: Evolution of the global interactions factor (solid line) from the PMS-IC model and the diffusion index of the Federal Reserve of Philadelphia (dashed line). Gray bars represent the US national recession periods following the official dating of the National Bureau of Economic Research (NBER)

About the properties of the States cycles, the scatter plot in Figure 7 of the estimated mean growth rates and volatilities shows a clear separation between the two phases of the regional cycles. The regression line shows a negative relationship between the State diffusion index mean growth rates and their volatilities during recessions. On the contrary in the expansion regime the volatilities are quite low and they are positively correlated with the mean growth rates. Moreover, Table 5 in the Appendix F shows that for every State, the expansion growth rate is positive and the recession growth rate is negative. And the volatilities of the recession growth rates are always higher than the ones of expansion regimes. Therefore, we document that typical evidence for aggregate county cycle, see, for example, Anas et al. (2007) and Billio et al. (2016), is also valid for regional cycles.

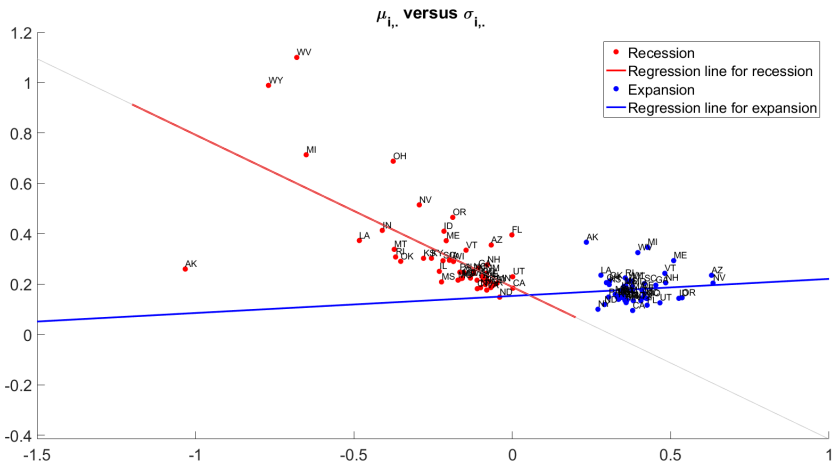


Figure 7: Estimations of monthly State-level coincident index mean growth rates and volatilities with the Bayesian PMS-IC model. The horizontal axis represents the mean growth rates and the vertical axis the volatilities. Sample period: October 1979 to December 2015 (month on month). The label of the US-States is added for clarity purpose

7. Conclusion

This paper models the idea that there might be interaction mechanisms among a population of time series. These interactions can either be local (i.e. in some neighbourhoods) or global (i.e. for all the units in the panel). To investigate these interactions, we propose a new dynamic panel Markov

switching model with interacting chains (PMS-IC) and provide an efficient Markov Chain Monte Carlo algorithm for the posterior approximation.

We introduce linear time-varying transition probabilities for the unit-specific Markov chains. These transition probabilities linearly depend on three factors. The first factor, the fixed transition matrix, assumes that all the series share a time independent common movement. The second factor, the local interaction factor, assumes that each unit-specific Markov chain shares a time dependent common movement with its neighbours. The third factor, the global interaction factor, assumes that all the unit-specific Markov chains share a time dependent common movement.

The Markov Chain Monte Carlo algorithm for the posterior approximation follows a four steps algorithm: (1) run a Metropolis-adjusted Forward-Filtering Backward-Sampling for the hidden states; (2) apply a Metropolis adjusted Langevin (MALA) sampling method for (α, β, γ); (3) use a standard Metropolis-Hastings step to draw the fixed transition probabilities; and (4) draw the unit-specific regression parameters from their posterior distributions.

We illustrate the usefulness of our PMS-IC model by conducting simulation exercises and a regional business cycle application. Our simulation experiments show that the proposed model is able to capture several levels of synchronization. The empirical macro application concerns US regional business cycles. The estimation reveals that a geographical and economic interaction factor plays an important role in the US regional business cycles. Both local and global interactions prevail and models that include both factors result in faster and deeper recessions than those models that only relying in a global factor, such as the FED diffusion index. Moreover, our model correctly captures all the national recessions defined by the NBER.

References

Aastveit, K. A., Bjørnland, H. C., and Thorsrud, L. A. (2015), "What drives oil prices? Emerging versus developed economies," *Journal of Applied Econometrics*, 30, 1013–1028.

Aastveit, K. A., Bjørnland, H. C., and Thorsrud, L. A. (2016), "The world is not enough! Small open economies and regional dependence," *The Scandinavian Journal of Economics*, 118, 168–195.

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

Anas, J., Billio, M., Ferrara, L., and Duca, M. L. (2007), "Business cycle analysis with multivariate Markov switching models," *Growth and Cycle in the Eurozone, BASINGSTOKE, HANTS,*, pp. 249–260.

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

Bernile, G., Kumar, A., and Sulaeman, J. (2015), "Home away from Home: Geography of Information and Local Investors," *The Review of Financial Studies*, 28, 2009–2049.

Bernile, G., Delikouras, S., Korniotis, G. M., and Kumar, A. (2017), "The Propagation of Industrial Business Cycles," Tech. Rep. 11-2014, Research Collection Lee Kong Chian School Of Business.

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.

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, 1352–1370.

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

Burda, M. (2015), "Constrained Hamiltonian Monte Carlo in BEKK GARCH with Targeting," *Journal of Time Series Econometrics*, 7, 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, 345–372.

Camacho, M. and Leiva-Leon, D. (2014), "The Propagation of Industrial Business Cycles," Tech. Rep. 2014-48, Bank of Canada.

Casarin, R., Grassi, S., Ravazzolo, F., and van Dijk, H. K. (2015), "Dynamic Predictive Density Combinations for Large Data Sets in Economics and Finance," Tech. rep., Tinbergen Institute Discussion Paper 15-084/III.

Celeux, G. (1998), "Bayesian Inference for Mixture: The Label Switching Problem," in *Compstat*, eds. R. Payne and P. Green, 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, 253–268.

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, 99–121.

Francis, N., Owyang, M. T., and Savascin, O. (2017), "An endogenously clustered factor approach to international business cycles," *Journal of Applied Econometrics*, forthcoming.

Frühwirth-Schnatter, S. (2001), "Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models," *Journal of the American Statistical Association*, 96, 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, 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, 123–214.

Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

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

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

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, 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, 82–94.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2003), "International business cycles: World, region, and country-specific factors," *American Economic Review*, 75, 1216–1239.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2008), "Understanding the evolution of world business cycles," *Journal of International Economics*, 93, 110–130.

Leiva-Leon, D. (2014), "A New Approach to Infer Changes in the Synchronization of Business Cycle Phases," Tech. Rep. 2014-38, Bank of Canada.

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

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

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

Vesper, A. J. (2013), "Three Essays of Applied Bayesian Modeling: Financial Return Contagion, Benchmarking Small Area Estimates, and Time-Varying Dependence," Ph.D. thesis.

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, 76–96.

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

A. Properties of the PMS-IC

A.1. Proof of Proposition 1

Without loss of generality let assume K=2. Define $W_{it,1} = S_{it} - \mu_{it}$ where $\mu_{it} = \pi_{it}(S_t, 1)$ and $S_t \in \mathcal{S}_\infty$. Then $\{W_{it,1}\}_{i \geqslant 1}$ is a sequence of independent random variables, conditioning on $\mathcal{F}_{t-1} = \sigma(\{S_u\}_{u \leqslant t-1})$, 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

$$\begin{aligned} \overline{m}_{t+1,1} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \pi_{it}(S_t, 1) \qquad (18) \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (S_{it}\alpha p_{11} + (1 - S_{it})\alpha p_{01} + \beta \overline{m}_{it,1} + \gamma \overline{m}_{t,1}) \\ &= (\overline{m}_{t,1}\alpha p_{11} + (1 - \overline{m}_{t,1})\alpha p_{01} + (1 - \alpha - \gamma)\overline{m}_{t,1} + \gamma \overline{m}_{t,1}). \end{aligned}$$

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. Linear time varying transition mechanism

We assume a linear parametrization of the transition probabilities. The transition matrix P_{it} of unit i at time t is linear with respect to the fixed transition matrix P, the local interactions indices $m_{it,k}$, $k = 1, \ldots, K$, and the global interactions index $m_{t,k}$, $k = 1, \ldots, K$.

We define a global interaction mechanism as a map

$$m_{t,k}(S_{1t}, \ldots, S_{Nt}) : \Delta^N_{[0,1]^K} \longrightarrow \Delta_{[0,1]^K}$$

where $\Delta_{[0,1]^K}$ is the standard K-dimensional simplex. The local interaction is defined as a map

$$m_{it,k}\left(S_{1t},\ldots,S_{Nt}\right):\Delta_{[0,1]^K}^{|N(i)|}\longrightarrow\Delta_{[0,1]^K}$$

where $|N(i)|$ is the cardinal of the set $N(i)$ of neighbourhood of unit i, $i=1,\ldots,N$.

The matrix representation of the linear time varying transition probabilities is:

$$\mathbb{P}\left(S_{it}|S_{it-1}\right)=P_{it}=\alpha P+\beta\mathbf{1}_K m_{it}\left(S_{1t},\ldots,S_{Nt}\right)+\gamma\mathbf{1}_K m_t\left(S_{1t},\ldots,S_{Nt}\right),$$

$$0<\alpha\leqslant 1,\quad 0<\beta\leqslant 1,\quad 0<\gamma\leqslant 1,\quad \alpha+\beta+\gamma=1$$

We denote $\mathbf{1}_K=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$, $m_t\left(S_{1t},\ldots,S_{Nt}\right)=(m_{t,1},\ldots,m_{t,K})$, $m_{it}\left(S_{1t},\ldots,S_{Nt}\right)=(m_{it,1},\ldots,m_{it,K})$, with $m_{t,l}=\frac{1}{N}\sum_{j=1}^{N}\mathbb{I}_{\{l\}}(S_{jt-1})$, $\forall k=1,\ldots,K$

P the fixed transition probabilities matrix is defined by

$$P=\begin{pmatrix}P_{11}&\cdots&P_{1K}\\ \vdots&\ddots&\vdots\\ P_{K1}&\cdots&P_{KK}\end{pmatrix}$$

where P_{lk} represents the fixed conditional probability that unit i moves from the latent regime k at time $t-1$ to the latent regime l at time t. So then, $P_{lk}\geqslant 0$, $P_{lk}=\mathbb{P}(S_{it}=l|S_{i,t-1}=k)$.

C. Parameter conditional distributions

1. The posterior distribution of the regime dependent intercept μ_{il} (where $l=1,\ldots,K$ and $i=1,\ldots,N$) according to the likelihood in equation (11) and the prior in equation (12) has a normal distribution with density function:

$$
\begin{aligned}
f(\mu_{il}|X_{1:T}, S_{1:T}, \theta_{-\mu_{il}}) &\propto \exp\{-\frac{1}{2\tau_{il}^2}\left(\mu_{il} - m_{il}\right)^2\} \prod_{t=1}^{T} \exp\{-\frac{\xi_{l,it}}{2\sigma_{il}^2}\left(X_{it} - \mu_{il}\right)^2\} \\
&\propto \exp\{-\frac{1}{2\tau_{il}^2}\left(\mu_{il} - m_{il}\right)^2\} \prod_{t\in\mathcal{T}_i} \exp\{-\frac{1}{2\sigma_{il}^2}\left(X_{it} - \mu_{il}\right)^2\} \\
&\propto \exp\{-\frac{1}{2\tau_{il}^2}(\mu_{il}^2 - 2\mu_{il}m_{il})\} \exp\{-\sum_{t\in\mathcal{T}_{il}} \frac{1}{2\sigma_{il}^2}(\mu_{il}^2 - 2\mu_{il}X_{it})\} \\
&\propto \exp\{-\frac{1}{2}\mu_{il}^2\left(\frac{1}{\tau_{il}^2} + \frac{T_{il}}{\sigma_{il}^2}\right) - 2\mu_{il}\left(\frac{m_{il}}{\tau_{il}^2} + \frac{\sum_{t\in\mathcal{T}_{il}} X_{it}}{\sigma_{il}^2}\right)\} \\
&\propto \mathcal{N}(\overline{m}_{il}, \overline{\tau}_{il}^2)
\end{aligned}
$$

with $\overline{m}_{il} = \overline{\tau}_{il}^2\left(\frac{m_{il}}{\tau_{il}^2} + \frac{\sum_{t\in\mathcal{T}_{il}} X_{it}}{\sigma_{il}^2}\right)$ and $\overline{\tau}_{il}^2 = \left(\frac{1}{\tau_{il}^2} + \frac{T_{il}}{\sigma_{il}^2}\right)^{-1}$.

We defined $\mathcal{T}_{il} = \{t = 1, \ldots, T | S_{it} = l\}$, $T_{il} = card(\mathcal{T}_{il})$, $X_{1:T} = (X_1, \ldots, X_T)$ and $S_{1:T} = (S_1, \ldots, S_T)$. The notation θ_{-r} indicates that element r is excluded from the vector θ.

2. The posterior distribution of the regime dependent volatility σ_{il} (where $l = 1, \ldots, K$ and $i = 1, \ldots, N$) according to the likelihood in equation (11) and the prior in equation (13) has inverse gamma distribution with density function:

$$
\begin{aligned}
lf\left(\sigma_{il}|X_{1:T}, S_{1:T}, \theta_{-\sigma_{il}}\right) &\propto \left(\frac{1}{\sigma_{il}^2}\right)^{(\alpha_{il}+1)} \exp\{-\frac{\beta_{il}}{\sigma_{il}^2}\} \prod_{t\in\mathcal{T}_{il}} \frac{1}{\sigma_{il}^2} \\
&\quad \times \exp\{-\sum_{t\in\mathcal{T}_{il}} \frac{1}{2\sigma_{il}^2}(X_{it} - \mu_{il})^2\} \\
&\propto \left(\frac{1}{\sigma_{il}^2}\right)^{(\alpha_{il}+T_{il}+1)} \exp\{-\frac{1}{\sigma_{il}^2}\{\beta_{il} + \sum_{t\in\mathcal{T}_{il}} (X_{it} - \mu_{il})^2\}\} \\
&\propto \mathcal{IG}\left(\alpha_{il} + T_{il}, \beta_{il} + \sum_{t\in\mathcal{T}_{il}} \left(X_{it} - \mu_{il}\right)^2\right)
\end{aligned}
$$

3. The posterior 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 (11)

and the prior in equation (14) density function:

$$
\begin{aligned}
& f\left(p_{l,1:K}|X_{1:T}, S_{1:T}, \theta_{-\left(p_{l,1:K}\right)}\right) \\
\propto & \left(\prod_{k=1}^{K} p_{lk}^{(\delta_k-1)}\right) \prod_{t=1}^{T}\prod_{i=1}^{N}\prod_{k=1}^{K} \left(\alpha p_{lk} + \beta m_{it,k} + \gamma m_{t,k}\right)^{\xi_{k,it}\xi_{l,it-1}} \\
\propto & \left(\prod_{k=1}^{K} p_{lk}^{(\delta_k-1)}\right) \prod_{k=1}^{K}\prod_{i=1}^{N}\prod_{t\in\mathcal{T}_{ilk}} \left(\alpha p_{lk} + \beta m_{it,k} + \gamma m_{t,k}\right) \\
\propto & \prod_{k=1}^{K} \left(p_{lk}^{(\delta_k-1)} \prod_{i=1}^{N}\prod_{t\in\mathcal{T}_{ilk}} \left(\alpha p_{lk} + \beta m_{it,k} + \gamma m_{t,k}\right)\right)
\end{aligned}
$$

4. The posterior distribution of (α, β, γ) according to the likelihood in equation (11) and the prior in equation (15) density function:

$$
\begin{aligned}
& f\left(\alpha, \beta, \gamma|X_{1-T}, S_{1-T}, \theta_{-(\alpha,\beta,\gamma)}\right) \\
\propto & \left(\alpha^{\varphi_1-1}\beta^{\varphi_2-1}\gamma^{\varphi_3-1}\right) \prod_{t=1}^{T}\prod_{i=1}^{N}\prod_{l=1}^{K}\prod_{k=1}^{K} \left(\alpha p_{lk} + \beta m_{it,k} + \gamma m_{t,k}\right)^{\xi_{k,it}\xi_{l,it-1}} \\
\propto & \left(\alpha^{\varphi_1-1}\beta^{\varphi_2-1}\gamma^{\varphi_3-1}\right) \prod_{k=1}^{K}\prod_{l=1}^{K}\prod_{i=1}^{N}\prod_{t\in\mathcal{T}_{ilk}} \left(\alpha p_{lk} + \beta m_{it,k} + \gamma m_{t,k}\right)
\end{aligned}
$$

D. Forward-filtering backward-sampling (FFBS) algorithm

The FFBS algorithm also known as multi-move sampling, is needed to sample from the joint posterior distribution $f(S_{i,1:T}|S_{-i,1:T}, X_{1:T}, \theta)$. By means of dynamic factorization, the full conditional distribution of the unit specific hidden state is

$$
\begin{aligned}
\mathbb{P}(S_{i,1:T}|S_{-i,1:T}, X_{1:T}, \theta) &= \mathbb{P}(S_{iT}|S_{-i,1:T}, X_{1:T}, \theta)\mathbb{P}(S_{i,1:T-1}|S_{iT}, S_{-i,1:T}, X_{1:T}, \theta) \\
&= \mathbb{P}(S_{iT}|S_{-i,1:T}, X_{1:T}, \theta) \prod_{t=1}^{T-1} \mathbb{P}(S_{i,t}|S_{i,t+1:T}, S_{-i,1:T}, X_{1:T}, \theta) \\
\propto & \mathbb{P}(S_{iT}|S_{-i,1:T}, X_{1:T}, \theta) \prod_{t=1}^{T-1} \mathbb{P}(S_{i,t+1}|S_{i,t}, S_{-i,t})\mathbb{P}(S_{i,t}|S_{-i,1:T}, X_{1:t}, \theta)
\end{aligned}
$$

$$\propto \left(\mathbb{P}(S_{iT}|S_{-i,1:T}, X_{1:T}, \theta) \prod_{t=1}^{T-1} \mathbb{P}(S_{i,t}|S_{-i,1:t}, X_{1:t}, \theta) \right) \left(\prod_{t=1}^{T-1} \mathbb{P}(S_{i,t+1}|S_{i,t}, S_{-i,t}, \theta) \right)$$

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.
The prediction probability for $l = 1, \ldots, K$ is:

$$\begin{aligned} \mathbb{P}(S_{it} = l|S_{-i,1:t}, X_{1:t-1}, \theta) &= \sum_{k=1}^{K} \mathbb{P}(S_{it} = l|S_{it-1} = k, S_{-it-1}) \\ &\quad \times \mathbb{P}(S_{i,t-1} = k|S_{-i,1:t-1}, X_{1:t-1}, \theta) \qquad (19) \\ &= \sum_{k=1}^{K} P_{it-1,kl} \mathbb{P}(S_{i,t-1} = k|S_{-i,1:t-1}, X_{1:t-1}, \theta) \end{aligned}$$

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. $S_{-i,t} = \{S_{jt}, j = 1, \ldots, Nj \neq i\}$. We initialize for $t = 1$, $\mathbb{P}(S_{i,0} = k|X_0, \theta)$ to be equal to the ergodic probabilities. The filtered probability for all $l = 1, \ldots, K$ is computed as:

$$\begin{aligned} \mathbb{P}(S_{it} = l|S_{-i,1:t} X_{1:t}, \theta) \propto \mathbb{P}(S_{it} = l|S_{-i,1:t-1}, X_{1:t-1}, \theta).f(X_{it}|S_{it} = l, X_{1:t-1}, \theta) \\ = \mathbb{P}(S_{it} = l|S_{-i,1:t-1}, X_{1:t-1}, \theta) \mathcal{N}(\mu_{il}, \sigma_{il}^2) \qquad (20) \end{aligned}$$

The smoothing probabilities are obtained recursively and backward in time, once all the filtered probabilities $\mathbb{P}(S_{it} = l|X_{1:t-1}, \theta)$ 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 $k = 1, \ldots, K$ the smoothing algorithm

proceeds as follows:

$$\mathbb{P}(S_{it} = l|S_{-i,1:T}, X_{1:T}, \theta) = \sum_{k=1}^{K} \mathbb{P}(S_{it} = l, S_{it+1} = k|S_{-i,1:T}, X_{1:T}, \theta)$$

$$= \sum_{k=1}^{K} \mathbb{P}(S_{it} = l|S_{it+1} = k, S_{-i,1:T}, X_{1:t}, \theta)\mathbb{P}(S_{it+1} = k|S_{-i,1:T}, X_{1:T}, \theta)$$

$$= \sum_{k=1}^{K} \frac{p_{it,lk}\mathbb{P}(S_{it} = l|S_{-i,1:T}, X_{1:t}, \theta)\mathbb{P}(S_{it+1} = k|S_{-i,1:T}, X_{1:T}, \theta)}{\sum_{j=1}^{K} p_{it,jk}\mathbb{P}(S_{it} = j|S_{-i,1:T}, X_{1:t}, \theta)}$$

E. Metropolis-adjusted Langevin algorithm (MALA)

The implementation of the MALA requires some necessary expressions which we discuss in this section. We consider the logistic transformation on the parameter p_{lk}, $k = 1, \ldots, K$, α, β and γ to deal with restrictions on parameters space.

$$\alpha = \frac{1}{1 + exp(-\tilde{\alpha})}$$
$$\beta = \frac{1}{1 + exp(-\tilde{\beta})}$$
$$\gamma = \frac{1}{1 + exp(-\tilde{\gamma})}$$
$$p_{lk} = \frac{1}{1 + exp(-\tilde{p_{lk}})}$$

where $\tilde{p_{lk}}$, $k = 1, \ldots, K$, $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\gamma}$ take value in the set of the real numbers. The derivatives of the parameter α, β and γ with respect to the logistic

transformation are:

$$\frac{d\alpha}{d\tilde{\alpha}} = \frac{exp(-\tilde{\alpha})}{(1+exp(-\tilde{\alpha}))^2} = \alpha(1-\alpha)$$
$$\frac{d\beta}{d\tilde{\beta}} = \frac{exp(-\tilde{\beta})}{\left(1+exp(-\tilde{\beta})\right)^2} = \beta(1-\beta)$$
$$\frac{d\gamma}{d\tilde{\gamma}} = \frac{exp(-\tilde{\gamma})}{(1+exp(-\tilde{\gamma}))^2} = \gamma(1-\gamma)$$
$$\frac{dp_{lk}}{d\tilde{p_{lk}}} = \frac{exp(-\tilde{p_{lk}})}{(1+exp(-\tilde{p_{lk}}))^2} = p_{lk}(1-p_{lk})$$

The partial derivatives of the complete data log-likelihood

$$L = log\mathcal{L}(X_{1:T}, S_{1:T}, \theta) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{l=1}^{K}\left(-\frac{\xi_{l,it}}{2}\right) log(2\pi\sigma_{il}^2)\left(-\frac{\xi_{l,it}}{2\sigma_{il}^2}\right)$$
$$\times(X_{it}-\mu_{il})^2(\xi_{l,it}\xi_{k,it-1})log(\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}) \quad (21)$$

with respect to the transformed parameters are

$$\begin{aligned}\frac{\partial L}{\partial\tilde{\alpha}} =& \frac{\partial L}{\partial\alpha}\frac{d\alpha}{d\tilde{\alpha}} \\ =& \alpha(1-\alpha)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\frac{p_{lk}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ =& \alpha(1-\alpha)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\mathcal{A} \end{aligned} \quad (22)$$

$$\begin{aligned}\frac{\partial L}{\partial\tilde{\beta}} =& \frac{\partial L}{\partial\beta}\frac{d\beta}{d\tilde{\beta}} \\ =& \beta(1-\beta)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\frac{m_{it,k}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ =& \beta(1-\beta)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\mathcal{B} \end{aligned} \quad (23)$$

$$\begin{aligned}\frac{\partial L}{\partial \tilde{\gamma}} =& \frac{\partial L}{\partial \gamma}\frac{d\gamma}{d\tilde{\gamma}} \\ =& \gamma(1-\gamma)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\frac{m_{t,k}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ =& \gamma(1-\gamma)\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\mathcal{C} \end{aligned} \tag{24}$$

$$\begin{aligned}\frac{\partial L}{\partial \tilde{p_{lk}}} =& \frac{\partial L}{\partial p_{lk}}\frac{dp_{lk}}{d\tilde{p_{lk}}} \\ =& p_{lk}(1-p_{lk})\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\frac{\alpha}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ =& p_{lk}(1-p_{lk})\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{\{l\}}(S_{it})\mathbb{I}_{\{k\}}(S_{it-1})\mathcal{D} \end{aligned} \tag{25}$$

where for ease of presentation we defined

$$\begin{aligned}\mathcal{A} =& \frac{p_{lk}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ \mathcal{B} =& \frac{m_{it,k}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ \mathcal{C} =& \frac{m_{t,k}}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}} \\ \mathcal{D} =& \frac{\alpha}{\alpha p_{lk}+\beta m_{it,k}+\gamma m_{t,k}}\end{aligned}$$

F. 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 4: Statistics, including degree and eigencentrality, 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 5: PMS-IC estimations coefficient for the US-States coincident indices.

Centre for Applied Macro - and Petroleum economics (CAMP)
will bring together economists working on applied macroeconomic issues, with special emphasis on petroleum economics.

BI Norwegian Business School
Centre for Applied Macro - Petroleum economics (CAMP)
N-0442 Oslo

http://www.bi.no/camp

CAMP Working Paper Series
ISSN: 1892-2198