

TI 2015-111/III
Tinbergen Institute Discussion Paper



Interconnections between Eurozone and US Booms and Busts using a Bayesian Panel Markov–Switching VAR mode

*Monica Billio*¹

*Roberto Casarin*¹

*Francesco Ravazzolo*²

*Herman K. van Dijk*³

¹ *University Ca' Foscari of Venice; Italy;*

² *BI Norwegian Business School, and Norges Bank, Norway;*

³ *Erasmus School of Economics, Erasmus University Rotterdam, VU University Amsterdam, and Tinbergen Institute, the Netherlands.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Interconnections between Eurozone and US booms and busts using a Bayesian Panel Markov-Switching VAR model*

Monica Billio[‡] Roberto Casarin[‡]
Francesco Ravazzolo[¶] Herman K. van Dijk[§]

[‡]University Ca' Foscari of Venice and GRETA Assoc.

[¶]Norges Bank and BI Norwegian Business School

[§]Econometric Institute Erasmus University Rotterdam, Econometrics Department
VU University Amsterdam and Tinbergen Institute

September 8, 2015

Abstract

Interconnections between Eurozone and United States booms and busts and among major Eurozone economies are analyzed using a Panel Markov-Switching VAR model. The model accommodates changes in low and high data frequencies and incorporates endogenous time-varying transition matrices of country-specific Markov chains. These country-specific Markov chains depend on their own past history and the history of other chains, thus allowing for interconnections between cycles, and an endogenous common Eurozone cycle is derived by aggregating the country-specific cycles. The model is estimated using a simulation based Bayesian approach in which an efficient multi-move sampling algorithm is defined to draw time-varying Markov-switching chains. Using industrial production growth and credit spread data for all countries, several empirical results have emerged. Recession, slow growth and expansion are empirically identified as three regimes with slow growth becoming persistent in the Eurozone in recent years different from the US. The Eurozone and the US regimes appear not fully synchronized, with evidence of more recessions in the Eurozone. Second, turning point analysis indicates larger synchronization at the beginning of the Great Financial Crisis: this shock affects the US first, leading the Eurozone cycle, and spreads then rapidly among these economies. Third, amplification effects influence recession probabilities for Eurozone countries when shocks occur. The evidence is different for the US where this reinforcement does not exist. In recent years there are more imbalances among regimes in Eurozone countries. Fourth, a credit shock results in substantial negative industrial production growth for several months in Germany, Spain and the US. Our empirical results point towards the existence of a group of core countries: Germany, France and

*We thank the co-editor Fabio Canova, three anonymous referees and conference and seminar participants at the: 68th European Meeting of the Econometric Society 2014, 5th CSDA International Conference on Computational and Financial Econometrics 2012, 5th Italian Congress of Econometrics and Empirical Economics 2013, European Seminar on Bayesian Econometrics 2012, 2nd CAMP Workshop 2012 on “Business cycles, monetary policy and nonlinear relationships”, University Ca' Foscari of Venice Workshop on “Modern Tools in Macro-Econometrics”, Aix Marseille University and Norges Bank. The views expressed in this paper are our own and do not necessarily reflect those of Norges Bank. Authors' research is supported by funding from the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270, by the Institut Europlace of Finance, “Systemic Risk grant”, the Global Risk Institute in Financial Services, the Louis Bachelier Institute, “Systemic Risk Research Initiative”, and by the Italian Ministry of Education, University and Research (MIUR) PRIN 2010-11 grant MISURA.

Netherlands with Italy sometimes belonging in the core and sometimes not and a group of peripheral countries Spain and sometimes Italy. Belgium taking a special position probably due to the de-industrialization in the nineties. US recessionary shocks affect Germany and Spain first in the Eurozone. The empirical results may serve as important information for the specification of a coordinated fiscal and monetary policy between the Eurozone and the US economies and also within the Eurozone economies, aimed at more adequate dynamic intervention with respect to shocks and reducing european imbalances.

JEL codes: C11, C15, C53, E37.

Keywords: Bayesian Modelling, Panel VAR, Markov-switching, International Business Cycles, Interaction mechanisms.

1 Introduction

According to the *Economist*, October 25-31, 2014, "the Eurozone region is marching towards stagnation and deflation." Interconnections between the economies of the aggregate Eurozone, individual countries and the US are important pieces of information for an empirical analysis of these issues. In this paper we investigate interconnections between booms and busts in the Eurozone and the US economies, where the Eurozone is represented by its six largest countries, with a particular focus on the number, persistence and synchronization of regimes, leads and lags in turning points and transmission mechanisms. Our modeling approach allows for shock transmission among different sectors: in particular we investigate the transmission from the financial sector, modeled with the credit spread, to the real sector, modeled with the industrial production index. The shock transmission among countries depends further on endogenous aggregate Eurozone and US business cycle factors.

The analysis is based on a Panel Markov-Switching VAR (PMS-VAR) model that describes cyclical behavior of the Eurozone economy at a country specific level and at an aggregate level and compares these with the US economy. We build on Canova and Ciccarelli (2004) and Canova and Ciccarelli (2009) and extend their panel VAR model in order to model asymmetry and turning points in the business cycles of different countries. See also Bassetti et al. (2014) for a Bayesian nonparametric specification of a multi-country panel VAR model. Our paper also extends Kaufmann (2010), where a panel of univariate Markov-switching (MS) regression models is considered, by constructing a multivariate panel MSVAR structure for the country-specific time series. We build on basic

model structures of Hamilton (1989) and Krolzig (2000) and consider Markov-switching dynamics for low and high frequency components that are specified as conditional means and covariance matrices of country-specific equations (see also Billio et al. (2012), Basturk et al. (2014) and Billio et al. (2013b)). We further build on Kaufmann (2015) and use an endogenous time-varying transition mechanism to model the transition matrix of the country-specific Markov-chains. In our model the transition of such chains depends not only on their past history but endogenously also on the past history of the other chains of the panel. Since only the transition probability matrix connects the different chains, the specification forces spill-over effects to enter nonlinearly in the model.

Our inference is Bayesian and we develop an efficient multi-move Gibbs sampling algorithm, based on forwarding-filtering backward sampling (e.g., see Frühwirth-Schnatter (2006)), to approximate the posterior distribution of the time-varying Markov-switching chains. To avoid potential overfitting problems due to the large number of parameters in the model, we follow the hierarchical prior specification strategy proposed by Canova and Ciccarelli (2009). Our paper relates also to Amisano and Tristani (2013), who propose a panel Markov-switching model to investigate transmission mechanisms in European sovereign bond markets, but our modeling and inference differ since we follow a hierarchical specification of the VAR and Markov-switching parameters. We make use of an endogenous transition that is based on alternative weighting rules with time-varying weights that account for differences in size and importance of the countries and our regime transition also accounts for Harding and Pagan (2002) constraints on minimum phases in order to obtain well defined business cycles.

Our main empirical results can be summarized as follows. We provide substantial empirical evidence on the existence of three regimes in all countries: recession, slow growth and expansion, with slow growth becoming persistent in the Eurozone in recent years, different from the US. The Eurozone and the US regimes appear not fully synchronized, with evidence of more recessions in the Eurozone, in particular during the 90's when the monetary union was planned. Second, turning point analysis indicates larger synchronization at the beginning of the Great Financial Crisis: this shock affects the US first, leading the Eurozone cycle, and spreads then very rapidly among Eurozone economies. Third, amplification effects influence recession probabilities for Eurozone countries when shocks occur, which is different for the US where this reinforcement does not exist. In recent years there are more

imbalances among regimes in Eurozone countries. Fourth, a credit shock, increasing the credit spreads and therefore deteriorating credit conditions, results in substantial negative industrial production growth for several months in Germany, Spain and the US.

The majority of our empirical results point towards the existence of two groups of countries in the Eurozone: a group of core countries: Germany, France and Netherlands with Italy sometimes belonging in the core and sometimes not; and a group of peripheral countries: Spain and sometimes Italy. Belgium taking a special position probably due to the de-industrialization in the nineties. Credit and other recessionary shocks in the US affect Germany and Spain first.

Our empirical results may serve as important information for the specification of a coordinated fiscal and monetary policy between the Eurozone and the US economies and also within the Eurozone economies that should be aimed at more adequate dynamic intervention with respect to shocks and reducing european imbalances.

The contents of this paper is as follows. Section 2 introduces the Panel MS-VAR model. Section 3 discusses the prior choice and the Bayesian inference framework. Section 4 presents empirical evidence within the Eurozone and between the Eurozone and the US economies. Section 5 concludes. The online Appendix contains details on related literature, data, simulation method, extra empirical results and several references.

2 A Panel Markov-Switching VAR model

We introduce a general Panel Markov-Switching VAR (PMS-VAR) model with endogenous transitions and interconnections. The transitions of unit specific Markov-chains are assumed to be dependent on their own past history and on the history of other chains in order to capture typical features of interconnections. Mechanisms such as different weighting schemes for countries and duration of regimes are also suggested. Stochastic parameter restrictions and hierarchical priors, see further Section 3, are used to avoid overfitting.

2.1 Panel VAR specification

Let $\mathbf{y}_{it} \in \mathbb{R}^M$ be a sequence of $t = 1, \dots, T$ time observations on an M -dimensional vector of economic variables for $i = 1, \dots, N$ countries. A general specification of a PMS-VAR

model reads

$$\mathbf{y}_{it} = \mathbf{a}_i(s_{it}) + \sum_{j=1}^N \sum_{p=1}^P A_{ijp}(s_{it}) \mathbf{y}_{jt-p} + \boldsymbol{\varepsilon}_{it}, \quad \boldsymbol{\varepsilon}_{it} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_i(s_{it})) \quad (1)$$

where $\mathcal{N}_M(\cdot, \cdot)$ denotes an M -variate normal distribution function, and $\mathbf{a}_i(s_{it})$, $A_{ijp}(s_{it})$ and $\Sigma_i(s_{it})$ are parameters depending on Markov chains. The $\{s_{it}\}$ are unit-specific and independent K -states Markov-chain processes with values in $\{1, \dots, K\}$ and with time-varying transition probabilities $\mathbb{P}(s_{it} = k | s_{it-1} = l, V_t, \boldsymbol{\alpha}_i^{kl}) = p_{it,kl}$ with $k, l \in \{1, \dots, K\}$, where V_t is a set of G_v common endogenous covariates to all chains and $\boldsymbol{\alpha}_i^{kl}$ is a unit-specific vector of parameters.

The generality of this statistical model stems from the possibility that the coefficients may vary both across units and across time. Moreover interdependencies between units are allowed whenever $A_{ijp}(s_{it}) \neq 0$ for $i \neq j$.

To clearly define parameter shifts and to simplify the exposition of the inference procedure, we introduce the indicator variable $\xi_{ikt} = \mathbb{I}(s_{it} = k)$, where

$$\mathbb{I}(s_{it} = k) = \begin{cases} 1 & \text{if } s_{it} = k \\ 0 & \text{otherwise} \end{cases}$$

for $k = 1, \dots, K$, $i = 1, \dots, N$, and $t = 1, \dots, T$. The vector of indicators $\boldsymbol{\xi}_{it} = (\xi_{i1t}, \dots, \xi_{iKt})'$ collects information about the realizations of the i -th unit-specific Markov chain over the sample period. Using these indicator variables, parameter shifts can be written as

$$\mathbf{a}_i(s_{it}) = \sum_{k=1}^K \mathbf{a}_{i,k} \xi_{ikt}, \quad A_{ijp}(s_{it}) = \sum_{k=1}^K A_{ijp,k} \xi_{ikt}, \quad \Sigma_i(s_{it}) = \sum_{k=1}^K \Sigma_{ik} \xi_{ikt}.$$

where $\mathbf{a}_{i,k} = (a_{i1,k}, \dots, a_{iM,k})' \in \mathbb{R}^M$ are M dimensional column vectors representing the country- and regime-specific VAR intercepts, $A_{ijp,k} \in \mathbb{R}^M \times \mathbb{R}^M$ are M -dimensional matrices of unit- and regime-specific autoregressive coefficients and $\Sigma_{ik} \in \mathbb{R}^M \times \mathbb{R}^M$ are M -dimensional unit- and regime-specific covariance matrices.

The large number of parameters makes our PMS-VAR very flexible. However, overparameterization may lead to overfitting, especially in macroeconomic applications, where time series are characterized by a low number of observations, slowly changing

means and time-varying variances (see, *e.g.*, Basturk et al. (2014)). These issues call for the use of a Bayesian approach to modeling and estimation, since it allows for inclusion of parameter restrictions with different degrees of prior belief. These restrictions should clearly be motivated by the specific application. In our application, using monthly data on the growth of the US Industrial Production Index (IPI) and on credit spread, we allow for different regimes in the intercepts and variances, whereas the autoregressive parameters are assumed to be constant across regimes. Also, prior restrictions on blocks of parameters are introduced. More specifically, we assume: $\mathbb{E}(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}'_{jt}) = O_{M \times M}$ with $O_{n \times m}$ the $(n \times m)$ -dimensional null matrix, and there are no interdependencies among the same variable across units, that is $A_{ijp,k} = A_{ip,k}\mathbb{I}(i = j) + O_{M \times M}(1 - \mathbb{I}(i = j))$, when conditioning on the parameters. The dependence across units will be modeled through a hierarchical prior specification discussed in section 3.1.

There exists empirical evidence for these choices. Clements and Krolzig (1998) find that most forecast errors are due to the constant terms in predictive models. They also suggest to consider MS models with regime-dependent volatility. We follow Krolzig (2000), Billio et al. (2012) and Basturk et al. (2014) and specifically assume that both unit-specific intercepts, $\mathbf{a}_i(s_{it})$, and covariance matrices, $\Sigma_i(s_{it})$, are driven by the regime-switching variables $\{s_{it}\}_t$ but that the autoregressive coefficients $A_{ip,k} = A_{ip}, \forall k$ remain constant. The restricted model is given as:

$$\mathbf{y}_{it} = \mathbf{a}_i(s_{it}) + \sum_{p=1}^P A_{ip}\mathbf{y}_{it-p} + \boldsymbol{\varepsilon}_{it}, \quad \boldsymbol{\varepsilon}_{it} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_i(s_{it})). \quad (2)$$

For the switching behaviour, we make use of an intercept-switching parameterization of the autoregressive model, see McCulloch and Tsay (1994). After a regime change, the mean level approaches the new value smoothly over several time periods. An alternative parameterization of the model is one where, after a regime change, an immediate mean level shift occurs. This parameterization has been used by Hamilton (1989) and Schotman and van Dijk (1991) and has the advantage that parameters can be easily interpreted, but a major drawback is that inference is more involved than for the McCulloch and Tsay (1994) parameterizations. See also Frühwirth-Schnatter (2006), ch. 11.4 and 12.2, and Paap and van Dijk (1998).

Following Frühwirth-Schnatter (2006), in order to simplify the exposition, we consider

the following re-parameterization based on a partitioning of the set of regressors $(1, \mathbf{y}'_{it-1}, \dots, \mathbf{y}'_{it-P})$ into $K+1$ subsets $\bar{\mathbf{x}}_{i0t} = (\mathbf{y}'_{it-1}, \dots, \mathbf{y}'_{it-P})'$ and $\bar{\mathbf{x}}_{ikt} = 1$, $k = 1, \dots, K$, that are an M_0 -dimensional vector of regressors with regime-invariant coefficients and K vectors of M_K regime-specific regressors with regime-dependent coefficients. Under our assumptions, $M_0 = MP$, $M_K = 1$, $\forall k$ and the PMS-VAR model is given as:

$$\mathbf{y}_{it} = X_{i0t}\boldsymbol{\gamma}_{i0} + \xi_{i1t}X_{i1t}\boldsymbol{\gamma}_{i1} + \dots + \xi_{iKt}X_{iKt}\boldsymbol{\gamma}_{iK} + \boldsymbol{\varepsilon}_{it}, \quad \boldsymbol{\varepsilon}_{it} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_i(\boldsymbol{\xi}_{it})) \quad (3)$$

where $X_{i0t} = (I_M \otimes \bar{\mathbf{x}}'_{i0t})$ and $X_{ikt} = I_M$ are the regime-invariant and the regime-specific regressor matrices, respectively, $\boldsymbol{\gamma}_{i0} \in \mathbb{R}^{MM_0}$, $\boldsymbol{\gamma}_{ik} \in \mathbb{R}^{MM_K}$, $k = 1, \dots, K$, $i = 1, \dots, N$, and $\Sigma_i(\boldsymbol{\xi}_{it}) = \Sigma_i(\boldsymbol{\xi}_{it} \otimes I_M)$ and $\Sigma_i = (\Sigma_{i1}, \dots, \Sigma_{iK})$. The relationship between the new parameterization and the previous one is: $\boldsymbol{\gamma}_{i0} = \text{vec}((A_{i1}, \dots, A_{iP})')$, and $\boldsymbol{\gamma}_{ik} = \mathbf{a}_{i,k}$.

2.2 Transition mechanisms

Following Kaufmann (2015) we assume a centered parameterization of the transition probabilities

$$\mathbb{P}(s_{it} = k | s_{it-1} = l, V_t, \boldsymbol{\alpha}_i) = H(V_t, \boldsymbol{\alpha}_i^{kl}), \quad k, l = 1, \dots, K \quad (4)$$

with

$$H(V_t, \boldsymbol{\alpha}_i^{kl}) = \frac{\exp((V_t - c_i)' \boldsymbol{\alpha}_{1i}^{kl} + \alpha_{0i}^{kl})}{\sum_{k=1}^K \exp((V_t - c_i)' \boldsymbol{\alpha}_{1i}^{kl} + \alpha_{0i}^{kl})}, \quad (5)$$

where $\boldsymbol{\alpha}_i^{kl} = (\alpha_{0i}^{kl}, \boldsymbol{\alpha}_{1i}^{kl})'$ and c_i is a vector of threshold parameters that can be chosen to be the average of V_t . For identification purposes, we let K be the reference state and assume $\boldsymbol{\alpha}_{1i}^{Kl} = \mathbf{0}$ and $\alpha_{0i}^{Kl} = 0$, for all $l = 1, \dots, K$. To simplify the exposition we also denote with $\boldsymbol{\alpha}_i = \text{vec}((\boldsymbol{\alpha}_i^{11}, \dots, \boldsymbol{\alpha}_i^{KK}))$ the collection of parameters of the sequence of transition matrices for the i -th unit.

Using recent macroeconomic data there is substantial evidence of a business cycle model with more than two regimes (see also Clements and Krolzig (1998)) and a time-varying error variance. For example, Kim and Murray (2002) and Kim and Piger (2002) propose a three-regime (recession, high-growth, and normal-growth) MS model while Krolzig (2000) suggests a model with regime-dependent volatility for US GDP. We consider data on Eurozone industrial production growth for a period of time including the 2009 recession and find that

three regimes (recession, $k = 1$, slow growth, $k = 2$, and expansion, $k = 3$) are necessary to capture some important features of the US and Eurozone cycles, see Section 4 for more details.

As evidenced in Harding and Pagan (2011) and Harding (2010) the use of simple logit or probit models for modeling the transition probabilities of phases of a business cycle may be inappropriate when the goal is to describe particular features of the business cycle such as duration of regimes. The feature of a minimum phase duration leads to impose restrictions on the parameters of the transition model. Extending the idea of Harding and Pagan (2011) to our PMS-VAR model and focusing on the minimum recession duration, we specify the following transition probabilities

$$\mathbb{P}(s_{it} = k | s_{it-1} = l, s_{it-2}, V_t, \boldsymbol{\alpha}_i) = \begin{cases} H(V_t, \boldsymbol{\alpha}_i^{kl}) & \text{if } s_{it-2} = 1 \\ 1 & \text{if } s_{it-2} \neq 1, k = 1, l = 1 \\ 0 & \text{if } s_{it-2} \neq 1, k \neq 1, l = 1 \\ H(V_t, \boldsymbol{\alpha}_i^{kl}) & \text{if } s_{it-2} \neq 1, \forall k \text{ and } l \neq 1 \end{cases} \quad (6)$$

and we impose the constraint of a minimum duration of two months for the recession phase. Note that this differs from the traditional definition of recession duration as two quarters of negative growth which is usually applied for turning point detection based on GDP. However, we emphasize that we only impose a lower bound on the minimum duration which is motivated by our use of monthly data.

2.3 Interaction mechanisms

We introduce dependence among Markov chains through a set of common covariates V_t . This set contains observable variables and also the state value of the N unit-specific Markov-chains. In order to achieve a parsimonious model, the information content of the N chains is summarized by an auxiliary variable $\boldsymbol{\eta}_t$ resulting from the aggregation of past values of the unit-specific chains.

The elements of $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{Kt})'$ are defined by the weighted average

$$\eta_{kt} = \sum_{i=1}^N \omega_{it} \mathbb{I}(s_{it-1} = k) \quad (7)$$

The weights do not necessarily have to sum up to one. However, we assume $\omega_{it} \geq 0$ and

$\sum_{i=1}^N \omega_{it} = 1$, for all t in order to have a probabilistic interpretation for the η_{kt} . The unit-specific weight ω_{it} , can be driven, for example, by the relative IPI growth rate or size of the i -th unit at time $t - 1$. Distance measures based on other features of the units can also be considered to aggregate the hidden states. When $k = 1$ we get a measure of the relative economic size of the proportion of countries which are in a “recession” regime.

Aggregation weights may be included in the inference procedure. This leads to a more complex latent variable model, both in terms of modeling and computation. Alternatively, one can use completely unobserved combination weights (e.g., in the modeling strategies of Billio et al. (2013a)) or weights which are partially observed and driven by one or some of the variables mentioned above. Given the already large number of latent variables in our model, the latter weight specification strategy is preferred in order to avoid overfitting problems. Other aggregation rules can be easily included in our framework to account for prevailing regimes over time and/or in the cross-section, but this is left for future research.

3 Bayesian Inference

The PMS-VAR model is estimated with a simulation based Bayesian procedure. In order to avoid overfitting due to the large number of parameters, we use hierarchical prior distributions. Moreover, we construct an efficient algorithm to draw latent MS chains which uses Forwarding-Filtering Backward Sampling (FFBS) (see, e.g., Frühwirth-Schnatter (2006)) on unit-specific auxiliary bivariate Markov chains. The auxiliary chains allow for both interaction effects and minimum duration restrictions when sampling from the posterior distribution of the latent MS processes.

3.1 Hierarchical prior

We follow a hierarchical prior specification strategy (see, e.g., Canova and Ciccarelli (2009)), which allows to model dependence between panel units through common latent variables

and to further avoid overfitting. For the parameters of the VAR regression we assume

$$\gamma_{i0} \sim \mathcal{N}_{MM_0}(\boldsymbol{\lambda}_0, \underline{\Sigma}_{i0}), \quad i = 1, \dots, N \quad (8)$$

$$\boldsymbol{\lambda}_0 \sim \mathcal{N}_{MM_0}(\underline{\boldsymbol{\lambda}}_0, \underline{\Sigma}_0) \quad (9)$$

$$\gamma_{ik} \sim \mathcal{N}_{MM_K}(\boldsymbol{\lambda}_k, \underline{\Sigma}_{ik}), \quad (10)$$

$$\boldsymbol{\lambda}_k \sim \mathcal{N}_{MM_K}(\underline{\boldsymbol{\lambda}}_k, \underline{\Sigma}_k), \quad k = 1, \dots, K \quad (11)$$

with $i = 1, \dots, N$. We also assume conditional independence across units, that is: $\mathbb{C}ov(\gamma_{i0}, \gamma_{j0} | \underline{\boldsymbol{\lambda}}_0) = O_{MM_0 \times MM_0}$ and $\mathbb{C}ov(\gamma_{ik}, \gamma_{jk} | \underline{\boldsymbol{\lambda}}_k) = O_{MM_K \times MM_K}$, for $i \neq j$; and for the inverse covariance matrix Σ_{ik}^{-1} we assume independent Wishart priors

$$\Sigma_{ik}^{-1} \sim \mathcal{W}_M(\underline{\nu}_{ik}, \Upsilon_k), \quad i = 1, \dots, N, k = 1, \dots, K \quad (12)$$

$$\Upsilon_k^{-1} \sim \mathcal{W}_M(\underline{\nu}_k, \underline{\Upsilon}_k), \quad k = 1, \dots, K \quad (13)$$

which allow to maintain the assumption of regime-specific degrees of freedom $\underline{\nu}_{ik}$ and precision Υ_k parameters. We finally assume $\mathbb{C}ov(\Sigma_{ik}^{-1}, \Sigma_{ik}^{-1} | \underline{\Upsilon}_k^{-1}) = O_{M^2 \times M^2}$.

It is important to note that the hierarchical prior specification also allows for dependence among units. Moreover, through the specification of the coefficients γ_{ik} it is possible to have regime-specific dependence structures.

When using Markov-switching processes, one has to deal with the identification issue associated with label switching. See, for example, Frühwirth-Schnatter (2001) for a discussion on the effects that the label switching and the lack of identification have on the results of MCMC based Bayesian inference. In the literature, different routes have been proposed for dealing with this problem (see Frühwirth-Schnatter (2006) for a review). One efficient approach is the permutation sampler (see Frühwirth-Schnatter (2001)), which can be applied under the assumption of exchangeability of the posterior density. This assumption is satisfied when one assumes symmetric priors on the transition probabilities of the switching process. As an alternative one may impose identification constraints on the parameters. This practice is followed to a large extent in macroeconomics and it is related to the natural interpretation of the different regimes as different phases (e.g. recession and expansion) of the business cycle. We follow this latter approach and include inequality constraints on the constant terms of the IPI growth rate equations of the system. For

details we refer to Section 4.2.

Modeling dependence among the chains is another issue to deal with. To avoid overfitting here, we use a hierarchical prior specification for the transition matrices. In particular, for the parameters of the l -th row, $\mathbf{p}_{it,l} = (p_{it,1l}, \dots, p_{it,Kl})$, $l = 1, \dots, K$, of the i -th unit transition matrix, at time t , we assume

$$\alpha_i^{kl} \sim \mathcal{N}_{G_v+1}(\boldsymbol{\psi}, \Upsilon_i) \quad i = 1, \dots, N, k = 1, \dots, K - 1 \quad (14)$$

$$\boldsymbol{\psi} \sim \mathcal{N}_{G_v+1}(\underline{\boldsymbol{\psi}}, \underline{\Upsilon}) \quad (15)$$

In the empirical application, we consider the following hyper-parameter specification: $\underline{\boldsymbol{\lambda}}_0 = \mathbf{0}$, $\underline{\Sigma}_{i0} = I_{MM_0}$, $\underline{\Sigma}_0 = 10I_{MM_0}$, $\underline{\boldsymbol{\lambda}}_k = \mathbf{0}$, $\underline{\Sigma}_{ik} = I_{MM_K}$, $\underline{\Sigma}_k = 10I_{MM_K}$, $k = 1, \dots, K$, $\underline{\nu}_{ik} = 5$, $\underline{\nu}_k = 5$, $\underline{\Upsilon}_k = 10I_M$, $\underline{\boldsymbol{\psi}} = \mathbf{0}$, $\Upsilon_i = I_{G_v+1}$, $\underline{\Upsilon} = 10I_{G_v+1}$ where $\mathbf{0}$ is the null vector.

3.2 Posterior simulation

We combine and extend the Gibbs sampler of Krolzig (1997) and Frühwirth-Schnatter (2006) to our PMS-VAR model with prior densities detailed in the previous sections. Under the hierarchical prior setting the full conditional posterior distributions of the equation-specific blocks of parameters are conditionally independent. Thus the Gibbs sampler iterates over different blocks of unit-specific parameters avoiding computational difficulties associated with the inversions of large covariance matrices (see Canova and Ciccarelli (2009)). We derive full conditional densities of the parameters in equation 3 and propose a further blocking step. We separate the unit-specific parameters into two different blocks: the regime-independent and the regime-specific ones.

Let $\mathbf{y}_i = \text{vec}((\mathbf{y}_{i1}, \dots, \mathbf{y}_{iT}))$ be the MT -dimensional vector of observations collected over time for the i -th unit of the panel, $\mathbf{y} = \text{vec}((\mathbf{y}_1, \dots, \mathbf{y}_N)')$ the MTN -dimensional vector of observations collected over time and panel units and $\boldsymbol{\xi} = \text{vec}((\Xi_1, \dots, \Xi_N))$ the KTN -dimensional vector of allocation variables, with $\Xi_i = (\boldsymbol{\xi}_{i1}, \dots, \boldsymbol{\xi}_{iT})$. We define the vector of regression coefficients, $\boldsymbol{\gamma} = \text{vec}((\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N))$ where $\boldsymbol{\gamma}_i = \text{vec}((\boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_{i1}, \dots, \boldsymbol{\gamma}_{iK}))$, the set of covariance matrices, $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_N)$, and the transition probability parameter vector, $\boldsymbol{\alpha} = \text{vec}((\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N))$.

Under the conditional independence assumption, the complete data likelihood function, associated to the PMS-VAR model, writes as

$$p(\mathbf{y}, \boldsymbol{\xi} | \boldsymbol{\gamma}, \Sigma, \boldsymbol{\alpha}) = \prod_{i=1}^N p(\mathbf{y}_i, \boldsymbol{\xi} | \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\alpha}_i) \quad (16)$$

where

$$p(\mathbf{y}_i, \boldsymbol{\xi} | \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\alpha}_i) = (2\pi)^{-\frac{TM}{2}} \prod_{t=1}^T |\Sigma_i(s_{it})|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \mathbf{u}'_{it} \Sigma_i(s_{it})^{-1} \mathbf{u}_{it} \right\} \prod_{k,l=1}^K p_{it,kl}^{\xi_{ikt} \xi_{ilt-1}} \quad (17)$$

with $p_{it,kl} = \mathbb{P}(s_{it} = k | s_{it-1} = l, s_{it-2}, V_t, \boldsymbol{\alpha}_i)$, $\mathbf{u}_{it} = \mathbf{y}_{it} - ((1, \boldsymbol{\xi}'_{it}) \otimes I_M) X_{it} \boldsymbol{\gamma}_i$ and

$$X_{it} = \begin{pmatrix} X_{i0t} & X_{i1t} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ X_{i0t} & \mathbf{0} & \dots & X_{iKt} \end{pmatrix}$$

Let us define $\boldsymbol{\gamma}_{i(-k)} = (\boldsymbol{\gamma}_{i1}, \dots, \boldsymbol{\gamma}_{ik-1}, \boldsymbol{\gamma}_{ik+1}, \dots, \boldsymbol{\gamma}_{iK})$ and $\Sigma_{i(-k)} = (\Sigma_{i1}, \dots, \Sigma_{ik-1}, \Sigma_{ik+1}, \dots, \Sigma_{iK})$. The Gibbs sampler is thus in six blocks. In blocks from one to three, the Gibbs iterates over the unit index, $i = 1, \dots, N$, and simulates the unit-specific parameters

- (i) $\boldsymbol{\gamma}_{i0}$ from $f(\boldsymbol{\gamma}_{i0} | \mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\lambda}_0)$;
- (ii) for $k = 1, \dots, K$
 - (ii.a) $\boldsymbol{\gamma}_{ik}$ from $f(\boldsymbol{\gamma}_{ik} | \mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_{i(-k)}, \Sigma, \boldsymbol{\lambda}_k)$, for $k = 1, \dots, K$;
 - (ii.b) Σ_{ik}^{-1} from $f(\Sigma_{ik}^{-1} | \mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_i, \Sigma_{i(-k)}, \Upsilon_k)$;
- (iii) $\boldsymbol{\alpha}_i^{1k}, \dots, \boldsymbol{\alpha}_i^{K-1k}$ from $f(\boldsymbol{\alpha}_i^{1k}, \dots, \boldsymbol{\alpha}_i^{K-1k} | \mathbf{y}_i, \Xi, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_i)$.

The mixing of the MCMC chain can be further improved by jointly simulating the parameters in the first and second block, while conditioning on the last iteration draws. In blocks four to six, the Gibbs sampler simulates from the full conditionals of the common part of the hierarchical structure and jointly from the full conditional of all the Markov-switching processes, i.e.

- (iv) For $k = 0, 1, \dots, K$:
 - (iv.a) $\boldsymbol{\lambda}_k$ from $f(\boldsymbol{\lambda}_k | \boldsymbol{\gamma}_k, \Sigma_k)$;
 - (iv.b) Υ_k^{-1} from $f(\Upsilon_k^{-1} | \boldsymbol{\gamma}_k, \Sigma_k)$;
- (v) Ξ from $p(\Xi | \mathbf{y}_{1:T}, \boldsymbol{\gamma}, \Sigma, \boldsymbol{\alpha})$

All full conditionals can be deduced from the joint density, that is proportional to the product of the prior densities, given in Section 3.1, and the completed likelihood given in equation 16. Further details on full the conditional distributions and their sampling methods are given in Appendix A.

4 Evidence on interconnections between Eurozone and US booms and busts

The empirical focus is on features of the interconnection between the aggregate Eurozone, euro country specific and US business cycles. The Eurozone is considered at the country level since the academic and economic debate is still open on whether European countries have synchronized and whether regional shocks still play a dominant role. Our analysis aims to contribute to this debate and to provide new evidence.

4.1 Data description

In our PMS-VAR we consider the US and the six largest economies in the Eurozone (Belgium, France, Germany, Italy, Netherlands, and Spain). For each country, we consider two dependent variables: growth of the Industrial Production Index (IPI), labeled as $y_{i1,t}$ and the credit spread (CS), i.e. the corporate bond yield spread over the 10 years government interest rate, $y_{i2,t}$.

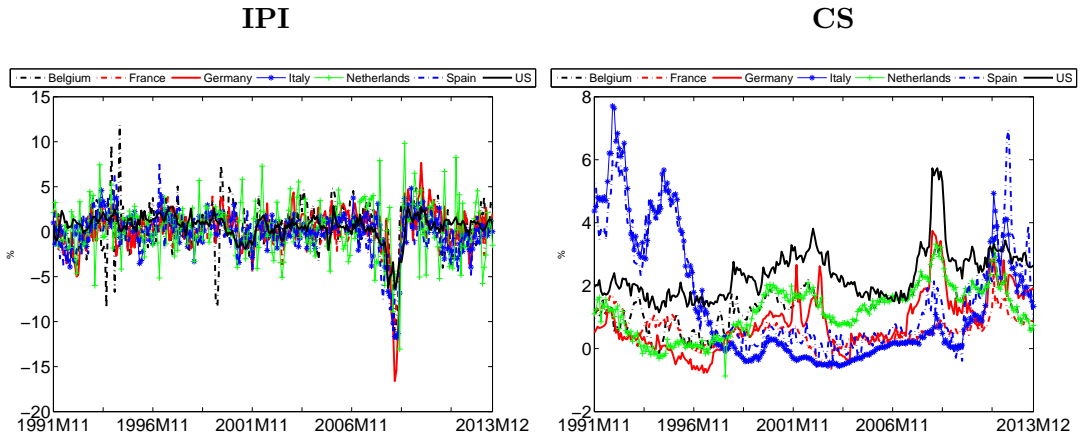


Figure 1: Country-specific endogenous variables: industrial production growth rate (IPI) and credit spread (CS).

All data are sampled at a monthly frequency, from July 1991 to December 2013, and

are seasonally and working day adjusted. They are plotted in Figure 1.

One crucial assumption for our model relates to the composition of the variable V_t . To investigate the interconnectedness between the Eurozone and the US, we specify the set of common endogenous covariates V_t equal to a vector η_{1t} and $\mathbb{I}(s_{US,t-1} = 1)$. The indicator η_{1t} is a weighted average of the number of euro countries in the recession regime (regime 1) at time $t - 1$; $\mathbb{I}(s_{US,t-1} = 1)$ takes the value 1 when the US economy is in recession (i.e. $s_{US,t-1} = 1$) and 0 otherwise. These assumptions allow us to have an endogenous interconnection mechanisms between the two economies and force spillovers to enter nonlinearly. More precisely, we focus on the weighted interconnection indicator given in equation (7) and use economic size unit-specific weights. We follow the Eurostat framework for aggregation of Eurozone variables and derive weights on relative value added, see Eurostat Regulation EC No 1165/98. Value added data are available from the UNData database and Figure 2 displays the weights. The value added data are annual and we transform them to a monthly frequency by using the same values for the 12 months in each calendar year. More details on the data are provided in the online Appendix.

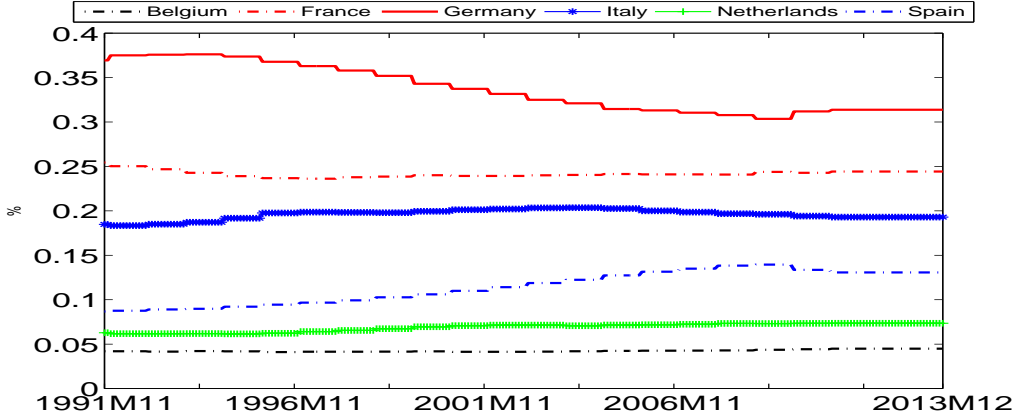


Figure 2: Value added Eurozone weights.

4.2 Evidence on number, persistence and synchronization of regimes

Number of regimes. To avoid issues with possibly non-stationary series, we take the IPI in log-changes, that is, in growth rates. We consider two possible cases of regimes, $K = 2$ and $K = 3$ for all countries in the panel, and discriminate between them using the Bayes factor

based on the predictive likelihood:

$$BF = \frac{p(\mathbf{y}|K = 3)}{p(\mathbf{y}|K = 2)}$$

where $p(\mathbf{y}|K = 3) = K! \prod_{t=1}^{T-1} \prod_{i=1}^N p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 3)$ with $p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 3)$ is the 1-step ahead predictive density for \mathbf{y}_{it+1} conditional on information up to time t and $K = 3$ regimes (see Frühwirth-Schnatter (2006), Chapter 5); $p(\mathbf{y}|K = 2)$ is a similar 1-step ahead predictive density for $K = 2$ regimes. We impose the following restrictions on the intercept of the IPI growth rate $a_{i1,1} < 0$ and $a_{i1,1} < a_{i1,2} < a_{i1,3}$, $i = 1, \dots, N$, in order to identify the regimes, see Section 3.1. We labeled regime 1 as recession; regime 2 as slow growth; and regime 3 as expansion.

We find that the BF is larger than one, therefore supporting the evidence of three regimes. Ferrara (2003), e.g., finds similar evidence for the US cycle. We also consider the number of autoregressive lags p to vary from 1 to 4 and choose $p = 4$ again by comparing BF s.

Empirical identification of regimes. Next, we present evidence of regime identification and a description of the three regimes through a graphical analysis using posterior densities. Figure 3 shows the approximated posterior densities of the parameters $\gamma_{ik} = (a_{i1,k}, a_{i2,k})'$, $(\sigma_{i11,k}, \sigma_{i22,k})$, $k = 1, \dots, K$ and $i = 1, \dots, N$, that represent the value of the unit- and variable-specific time-varying intercepts and volatilities (square root of the diagonal elements of the variance covariance $\Sigma_i(s_{it})$) of the PMS-VAR model. A comparison of such posteriors may provide useful information on whether and how individual countries differ over booms and busts.

The support of the posterior densities for the IPI growth intercept in regime 1, $a_{i1,1}$ is for most countries not concentrated near the zero bound. This provides evidence that there is no overlap with the posterior densities for the IPI growth intercept in regimes 2 and 3 (see the left column in Figure 3) for most countries, which suggests that the recession regime is empirically well identified by the IPI growth data. The slow growth regime 2 intercept is centered around zero and $a_{i1,3}$, the expansion intercept, is positive. However, there are differences between the European countries and the US: the posteriors are in most cases wider for the European countries, in particular the posteriors of $a_{i1,1}$ are more concentrated on the negative part of the support set. Posteriors for the US are more concentrated

and closer to zero. The posteriors for $a_{i1,2}$ and $a_{i1,3}$ overlap substantially, in particular for Belgium, indicating that expansion periods cannot easily be identified from the slow growth regime periods by just looking at the IPI intercepts.

The posterior densities of the credit spread intercept (see the second column in Figure 3) are centered just above zero for all countries as Figure 1 suggests, with larger dispersion for the recession and expansion periods. The overlapping supports of the posterior densities indicate substantial equivalence of credit spread means across regimes.

The posterior densities of the IPI volatility show that there exist large differences in volatility behavior across regimes between the US and European countries. For the US, volatility is higher during recession and expansion periods, and lower and more concentrated in slow growth periods. Volatility increases in regime 1 compared to regime 3. The results are different for the European countries. France, Germany and Italy, the most industrialized countries, have higher volatility in regime 1; Belgium in regime 3; The Netherlands has similar high volatilities in regimes 1 and 3. Spain has the lowest volatility in regime 1. IPI growth in the US exhibits larger switches during recession and expansion periods which increase volatility estimates. Posterior mean estimates suggest such movements are transitory and do not imply large changes in the intercept. On the contrary, Eurozone estimates seem to be dominated by more switches across all regimes, both in intercepts and the volatilities.

The residual volatilities of the credit spread for all seven economies have a posterior $\sigma_{i,22,2}$ that is more concentrated and closer to zero than the posteriors of the volatility in the other two regimes. In all countries, its support does not overlap with the ones of the recession and expansion regimes, indicating that regime 2 is well identified and supported by the data. This confirms the evidence in Gilchrist and Zakrajsek (2012) and Gilchrist and Mojon (2014) that credit spreads provide substantial empirical content for a variety of real activity and lending measures across different countries.

Regime probabilities: persistence and synchronization. The PMS-VAR model produces both country-specific smoothed probabilities for each regime and, further, Eurozone and US aggregate smoothed probabilities. Specifically, the number of Eurozone countries in recession and a similar measure for the US, used in the vector V_t , are reported in the Figure 4. Figure C.1 in the Online Appendix reports the associated recession probabilities of the Eurozone and the US. These figures provide several interesting results and generally shows

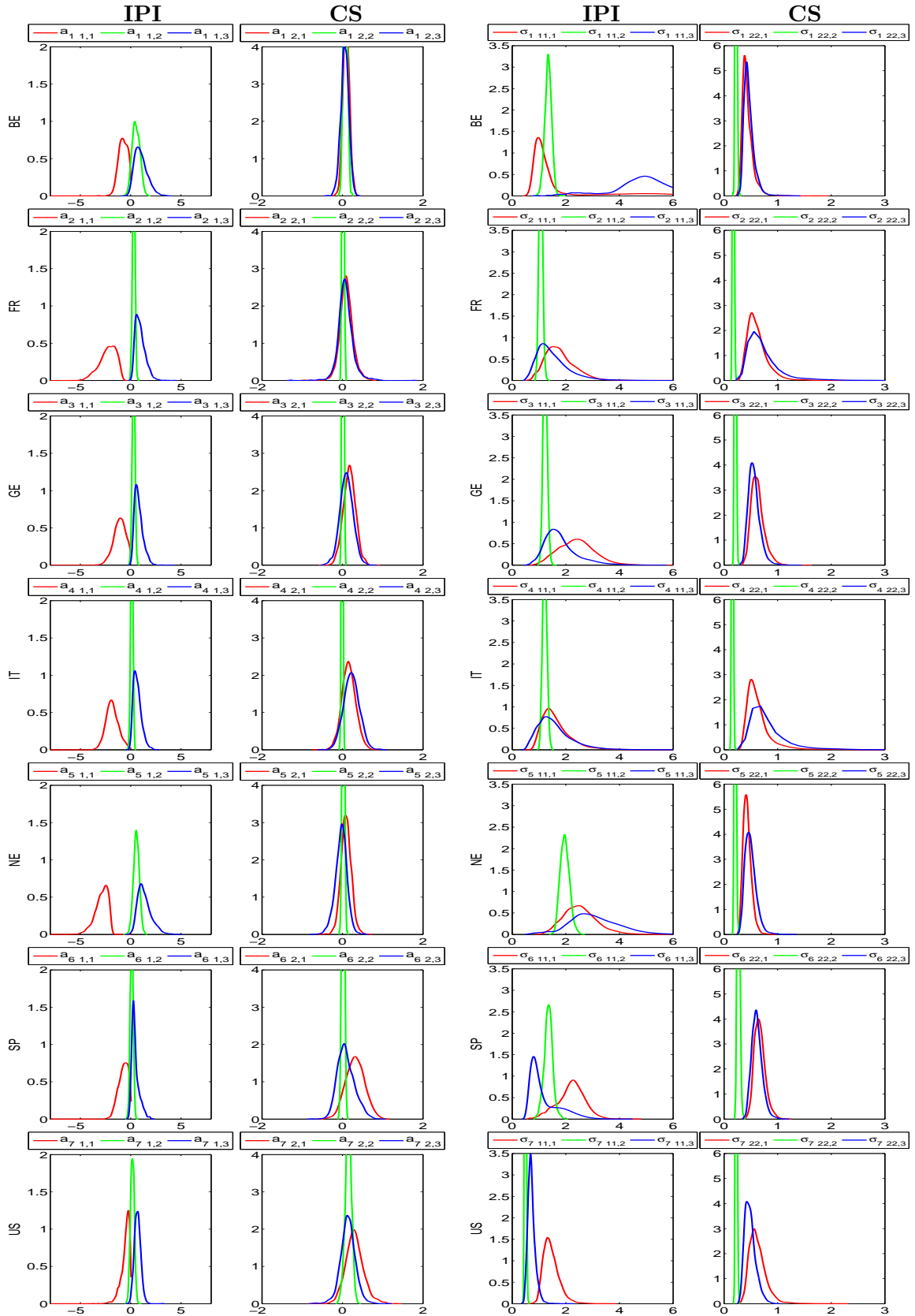


Figure 3: The figures show the kernel density estimate of the posterior densities of the Markov-switching intercepts (left panel), labeled $(a_{i1,k}, a_{i2,k})$ in the Figure, and volatilities (right panel), $(\sigma_{i11,k}, \sigma_{i22,k})$, for the different $i = 1, \dots, N$ countries and $k = 1, \dots, 3$ regimes (in red the first regime, in green the second regime and in blue the third regime) for industrial production growth rate (IPI) and credit spread (CS). The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

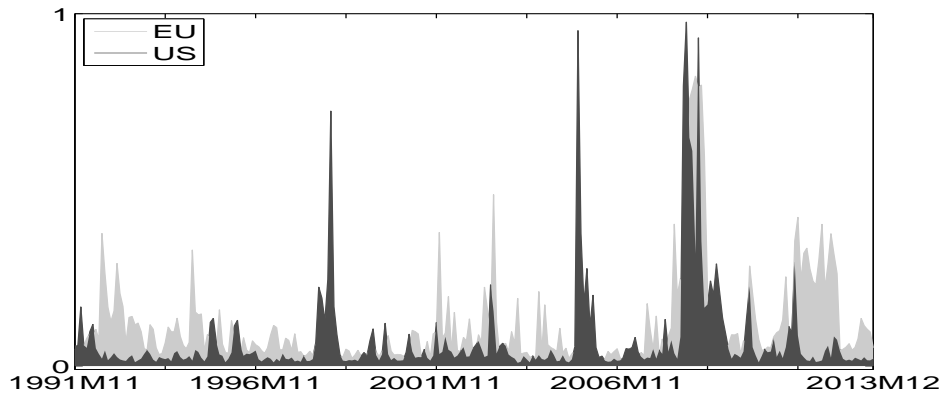


Figure 4: The light grey line shows the fraction of Eurozone countries in the recession regime standardized between 0 and 1 and the black line shows the US state probability for regime one.

that the Eurozone and US economies are not fully aligned.

In the first decade of our sample, the recession probability in the Eurozone is more volatile than in the US, see also Figure C.2 in the Online Appendix, and this may be related to the European Exchange Rate Mechanism (ERM) crisis and the construction of the European Monetary Union. A noticeable exception occurs at beginning of 1999 with the internet bubble in US. In the second decade, the US apparently leads the Eurozone cycle, especially during the Financial Crisis in 2007-2008. Also, the internet bubble started in the USA and spread to the Eurozone only in the final months of 2001 and generates further increases in recession probabilities up to 2003 but all lasting a few months. A new call for recession in the US is found at the end of 2005 and in 2006, possibly related to the rise of interest rates after a long period of low rates and at the beginning of the burst of the real estate bubble. The largest recession probabilities are during the Financial Crisis, with both economies having probabilities close to 1. The US enters the recession phase in December 2007, while the Eurozone recession starts in September 2008. Both economies enter in the second quarter 2009 in the slow growth regime (see also the low probability levels in Figure C.2 and C.4 and the high probability level in Figure C.3 in the Online Appendix). Furthermore, the Eurozone has evidence of a new recession regime at the third quarter in 2011. This recession can be associated to sovereign debt problems for some European countries, in particular Italy and Spain. The role of the credit spread volatility is quite important for detecting the recession because credit conditions deteriorate and become more volatile from 2010 onward in Europe, but also improved after the European Central Bank (ECB) interventions in December 2011 and during 2012 resulting in a recovering phase after 2012.

Looking at the seven country specific smoothed probabilities (reported in Figure C.2-C.4 in the Online Appendix) we observe that the regimes are often highly persistent. Regime 2 is most probable as one may anticipate since its definition fits slow growth as well as moderate expansion periods which are appropriate definitions for most of our sample period. The global Financial Crisis in 2008-2009 and its impact are evident, with most of the countries in recession. There is some evidence of a recession in 1999 in US and in 2001 in Germany and The Netherlands, but all short-lived. Larger differences exist during the European sovereign debt crisis, with US being the only country where the probability of regime 1 does not increase. The third regime has the lowest probabilities, but it shows an interesting increase in some European countries, e.g. Spain, at the end of the sample when the large liquidity provided by the ECB and bailout programs for Spanish banks result in better economic conditions. Finally, probabilities for Belgium seem the least related to US probabilities in the first decade of our sample, but converging in the second part of the sample. The large decline of mining in the 80's is a possible explanation.

The heterogeneity in the cycle patterns of the Eurozone countries is not only evident from regime probabilities, but it is also shown in posterior densities. The dynamic features of the cycle, in terms of posterior densities for the VAR time-varying intercept and for the VAR time-varying variance, are given for each country in Figure D.1-D.4 in the Online Appendix. We provide a short summary of this evidence in this subsection. The Financial Crisis is evident with regime 1 dominant in all the four parameters. The level of the conditional IPI intercept is more negative in Europe than in the US during the crisis. France, Germany, Italy and The Netherlands have a large part of the posterior below -1.5, compared to the 90% interval $[0, -1]$ for the US. The difference is even larger during the European sovereign debt crisis.¹ The intercept of the credit spread is the highest in US, but some Eurozone countries, e.g. Spain, have similar values. High volatilities for the IPI growth in recession are evident, with the US being one of the smallest. Volatilities of the credit spread across regions are, however, more comparable.

The graphical analysis of Figure 3 suggests also a classification of the countries in three major groups: the US alone since its posterior estimates are rather precisely estimated;

¹We note that the US IPI is more persistent with a larger value for the sum of AR coefficients. Therefore, the difference between the US IPI conditional growth and the European one is smaller than comparing only the intercepts. Across European countries, the persistence for the Netherlands is the lowest and this suggests a lower conditional mean for the Dutch IPI. Persistence for credit spreads is high and very similar across countries.

France, Germany, Italy and the Netherlands, the most industrialized European countries, have higher IPI volatility in regime 1; Spain and Belgium have higher IPI volatility in regime 2. Posteriors for the credit spread are more similar across countries and identify the second regime as a low credit risk regime compared to the first and third more volatile regimes. Therefore, our finding is that credit risk across countries is lower in a period of slow growth in our model compared to a recession and expansion regime. A country grouping based on features and synchronization of regime probabilities produces similar results: the US; France, Germany, Italy and the Netherlands; Spain and Belgium as three separate groups.

4.3 Leads and lags in turning points and clustering

To further investigate how countries relate to the aggregate cycle and possible synchronize with it, we study how each member country cycle shows turning points compared to the aggregate European business cycle. The contribution of each country is not necessarily equal to the value added scheme used to aggregate country-specific cycles in V_t because the link from individual countries to the aggregate depends on how the turning points are defined and on which statistics are used to measure the relationships across countries. We note again that the dating of the turning points is based on monthly IPI growth and not on GDP. The estimated turning point dates may, therefore, differ from the official ones published by the NBER.

As a first step in the analysis, we follow Billio et al. (2012) and date the Eurozone business cycle turning points by applying the Bry and Boschan (1971) (BB) rule, that identifies a downward turn (or peak) at time t for the variable of interest y_t , i.e. the log industrial production index, if $\Delta_\kappa y_t > 0, \dots, \Delta_1 y_t > 0$ and $\Delta_1 y_{t+1} < 0, \dots, \Delta_\kappa y_{t+\kappa} < 0$ and a upward turn (or trough) at time t if $\Delta_\kappa y_t < 0, \dots, \Delta_1 y_t < 0$ and $\Delta_1 y_{t+1} > 0, \dots, \Delta_\kappa y_{t+\kappa} > 0$, where Δ_κ denotes the κ -difference operator (see Harding and Pagan (2011)). Larger values of the parameter κ reduce the number of false signals. These definitions are standard in business cycle analysis (see for example Chauvet and Piger (2008)) and are also used (with some adjustments) by the NBER institute for building the reference cycle for the US. We apply an approximation of the BB rule and use only downward, $D_t(\kappa)$, and upward,

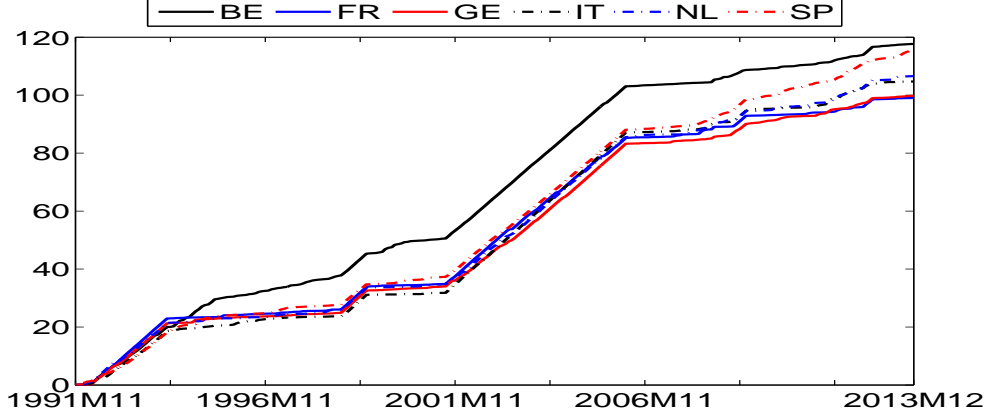


Figure 5: Cumulative concordance statistics of individual countries to predict the Eurozone cycle. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands and Spain.

$U_t(\kappa)$, turn signals, that are given as

$$D_t(\kappa) = \prod_{k=1}^{\kappa} \mathbb{I}(\Delta_k y_t > 0) \mathbb{I}(\Delta_k y_{t+k} < 0) \quad (18)$$

$$U_t(\kappa) = \prod_{k=1}^{\kappa} \mathbb{I}(\Delta_k y_t < 0) \mathbb{I}(\Delta_k y_{t+k} > 0) \quad (19)$$

respectively (see Harding and Pagan (2011)). Our analysis can be extended to include modifications of the BB rule, see for example Mönch and Uhlig (2005), in order to account for asymmetries and time-varying duration across business cycle phases. Censoring rules preventing the algorithm from the detection of false signals may also be used.

Set y_t equal to the aggregate Eurozone IPI growth. The following indicator variable can be computed:

$$z_t = z_{t-1}(1 - D_t(\kappa)) + (1 - z_{t-1})U_t(\kappa)$$

that is equal to 1 in the expansion phases and 0 in the recession phases. We assume z_0 is given. We evaluate synchronization of the different country-specific cycles (or regimes) by the concordance statistics (CS):

$$CS_i = \frac{1}{t+1-\kappa} \sum_{r=1}^{t+1-\kappa} \left(\mathbb{I}(\hat{s}_{i,r} = 1)(1 - z_r) + (1 - \mathbb{I}(\hat{s}_{i,r} = 1))z_r \right) \quad (20)$$

where we define a downward turn when switching to regime 1, i.e. $\mathbb{I}(\hat{s}_{i,r} = 1)$, and upward turn otherwise, i.e. $(1 - \mathbb{I}(\hat{s}_{i,r} = 1))$. This means that an upward turn can be a switch to regime 2 or 3 in our three-regime models. The hidden state estimate $\hat{s}_{i,r}$ is obtained by applying the maximum a posteriori probability (MAP) of the state posterior probabilities. The CS statistic is a nonparametric measure of the proportion of time during which two series, in our case the country-specific cycle and the Eurozone cycle, are in the same regime. This measure ranges between 0 and 1, with 0 representing perfectly counter-cyclical switches, and 1 perfectly synchronous shifts. Figure 5 shows the CS_i cumulated over time and it ranges between 0 and the sample size, that is 261 in our application. The countries with the highest CS have a business cycle which conserves over time a strong similarity to the Eurozone cycle. We identify graphically three clear patterns. Belgium deviates from other countries, but it is the most synchronized with the aggregate cycle from the mid nineties to the beginning of 2000. There is a period of large synchronization for all countries from 2001 to the beginning of 2006. The unfolding from this crisis and the beginning of the European sovereign debt crisis finishes such synchronization with large differences across CS_i . In particular, Italy, Spain and, a bit less, The Netherlands statistics deviate from those of the other countries.

As a second analysis, we apply a k -means clustering algorithm to the posterior mean regime probabilities of the six euro countries, $\hat{p}_{it,kl}$. The benefit of this exercise is that it does not require a definition of an aggregate index and it compares countries over the three regimes and not just the recession one as in the previous paragraphs. The drawback is that results are not standardized to a reference cycle. The k -means algorithm maximizes the difference between clusters and minimizes the difference within a cluster.

The results of the k -means cluster analysis divide the countries in three groups: France, Germany and the Netherlands are core euro country members; Italy and Spain are the periphery countries unfolding differently in the recent recession. Italy was classified together with France, Germany and the Netherlands in previous subsections and is now grouped with Spain. Finally, Belgium differs probably due to the de-industrialization process in the nineties.²

²When restricting the number of clusters to two, Belgium is moved to group 1.

4.4 Amplification of recession probabilities through interconnections

Another relevant result regards the interconnection between recessions in different countries. Posterior estimates of the loadings of V_t (see Table 1) provide information on the interconnection between the Eurozone and the US cycles. Estimates of the coefficients $\alpha_{1i}^{EU,11}$, $i = 1, \dots, 6$, associated with the Eurozone recession indicator, η_{1t} , appearing in the country-specific probability to stay in recession (see Equation 4-5), are all positive, large and significant. This means that there is an amplification effect, that is an increase in the probability to stay in recession at time $t + 1$ due to the fact that the Eurozone countries were in a recession phase in t . The evidence is different for the US where this amplification does not exist, probably due to the leading behavior of the US cycle. The US indicator seems not to have a clear amplification mechanism for the recession probabilities of the euro countries.

For the probability to exit a recession, $p_{it,21}$, the reinforcement exists for the US. Conditional on a recession of the Eurozone, the probability of recovery for the US increases. The faster recovery of the US after the Financial Crisis and the euro sovereign debt crisis in 2011-2012 are examples supportive of this finding. Across European countries, there exist large differences. The US recession indicator reduces the probability of regime 2 for Belgium and The Netherlands whereas the effect is not statistically significant for Germany and Italy. The coefficient is small, but positive and statistically significant for France and Spain. See Figure D.5 in the Online Appendix to see the sensitivity of the recession probability $p_{it,11}$ for values of η_{1t} when the US is not in recession, i.e. $s_{7t} \neq 1$, and when US is in recession, i.e. $s_{7t} = 1$.

The results indicate four groups: the US leads the euro area cycle indicated by positive α_{1i}^{11} and α_{1i}^{12} coefficients. The first countries that are affected by US recessionary shocks are (marginally) Germany and (significantly) Spain. This may suggest that US shocks are transmitted to Europe through Germany and Spain. Then we have France and Italy, which are not (directly) affected by US recessions. Finally, Belgium and the Netherlands respond generally negatively to US shocks.

4.5 Credit shock effects

Our PMS-VAR allows us to investigate how exogenous shocks propagate within and across countries. Unfortunately, the parameters in the reduced form model presented in Section 2

| Country | | $p_{it,11}$ | | $p_{it,21}$ | |
|---------|-------|-----------------------|-----------------------|-----------------------|-----------------------|
| i | Label | $\alpha_{1i}^{EU,11}$ | $\alpha_{1i}^{US,11}$ | $\alpha_{1i}^{EU,21}$ | $\alpha_{1i}^{US,21}$ |
| 1 | BE | 1.42 | -0.20 | -0.03 | -0.13 |
| | | (1.36, 1.48) | (-0.27,-0.14) | (-0.21, 0.08) | (-0.19,-0.08) |
| 2 | FR | 1.73 | 0.03 | 0.21 | 0.11 |
| | | (1.55,1.86) | (-0.11,0.14) | (0.05,0.31) | (0.02,0.20) |
| 3 | GE | 1.67 | 0.10 | 0.02 | 0.10 |
| | | (1.54,1.78) | (-0.05,0.21) | (-0.06,0.09) | (-0.07,0.24) |
| 4 | IT | 1.78 | 0.07 | 0.25 | 0.04 |
| | | (1.58,1.98) | (-0.28,0.32) | (0.03,0.41) | (-0.07,0.13) |
| 5 | NL | 1.80 | -0.10 | 0.04 | -0.31 |
| | | (1.69,1.95) | (-0.21,0.03) | (-0.12,0.28) | (-0.48,-0.09) |
| 6 | SP | 1.51 | 0.45 | 0.47 | 0.21 |
| | | (1.40,1.75) | (0.23,0.73) | (0.21,0.60) | (0.06,0.35) |
| 7 | US | -0.02 | 1.69 | 1.17 | 0.04 |
| | | (-0.10,0.09) | (1.58,1.88) | (0.88,1.32) | (-0.08,0.13) |

Table 1: Posterior mean and 90% credible interval (in parenthesis) for the parameters, $\alpha_{1i} = (\alpha_{1i}^{11}, \alpha_{1i}^{21})'$, with $\alpha_{1i}^{kl} = (\alpha_{1i}^{EU,kl}, \alpha_{1i}^{US,kl})'$, which are the coefficients of the interaction variables η_{1t} and $\mathbb{I}(s_{US,t} = 1)$ driving the Markov-switching transition probabilities.

do not identify uniquely structural parameters and shocks across equations, implying that it is not possible to distinguish regime shifts from one structural equation to another, see e.g. Sims and Zha (2006) for further discussions. Therefore we transform the PMS-VAR to the following structural model:

$$\mathbf{y}'_{it} B_{0,i}(s_{it}) = \mathbf{x}'_{it} B_{+,i}(s_{it}) + \mathbf{u}'_{it}, \quad (21)$$

where $\mathbf{u}_{it} \sim \mathcal{N}_M(\mathbf{0}, I_M)$; $\mathbf{x}_{it} = (1, \mathbf{y}'_{i,t-1}, \dots, \mathbf{y}'_{i,t-P})$. The parameters $A_i(s_{it}) = (\mathbf{a}_i(s_{it}), A_{i1}, \dots, A_{iP})$ are estimated in the reduced form model and $A_i(s_{it})' = B_{+,i}(s_{it})B_{0,i}(s_{it})^{-1}$; $\boldsymbol{\varepsilon}_{it} = B_{0,i}(s_{it})^{-1}\mathbf{u}_{it}$; $E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}'_{it}) = (B_{0,i}(s_{it})B_{0,i}(s_{it})')^{-1}$ with $\boldsymbol{\varepsilon}_{it}$ the residual of the reduced form model. When sufficient restrictions are imposed on $(B_{0,i}(s_{it})B_{0,i}(s_{it})')$, the structural model is identified. Recalling the notation in Section 2, we follow the framework in Sims and Zha (2006) for Markov-Switching models and use a Cholesky decomposition of $\Sigma_i(s_{it}) = (B_{0,i}(s_{it})B_{0,i}(s_{it})')^{-1}$ to identify the structural system.

We investigate the effect of a credit shock to IPI growth in the recession regime. That is, we extend the evidence in Del Negro et al. (2014) to a Markov-switching model for the US and the European countries. Our identification restriction scheme implies that the credit spread responds contemporaneously to a credit shock in each country; while IPI growth responds with one month-lag. The motivation is that financial variables move

faster than real variables for several reasons, including publication delays for real variables. Figure 6 plots the impulse responses (IR) for the seven countries in our panel. The IRs are standardized by imposing that the median response of US credit spread is 1. The response of IPI growth is plotted as a cumulative sum over horizons. The credit shocks play a relevant role for most of the countries, but large differences exist across countries. Responses are large and a statistical significant reduction of IPI growth over several quarters is evident for Germany and the US and to a lesser extent for Spain. The response is, however, not significant for Italy and the Netherlands. Moreover, IPI growth increases in the first months after the shock for France. The large size of the government sector in the French economy may be an explanation for this.³

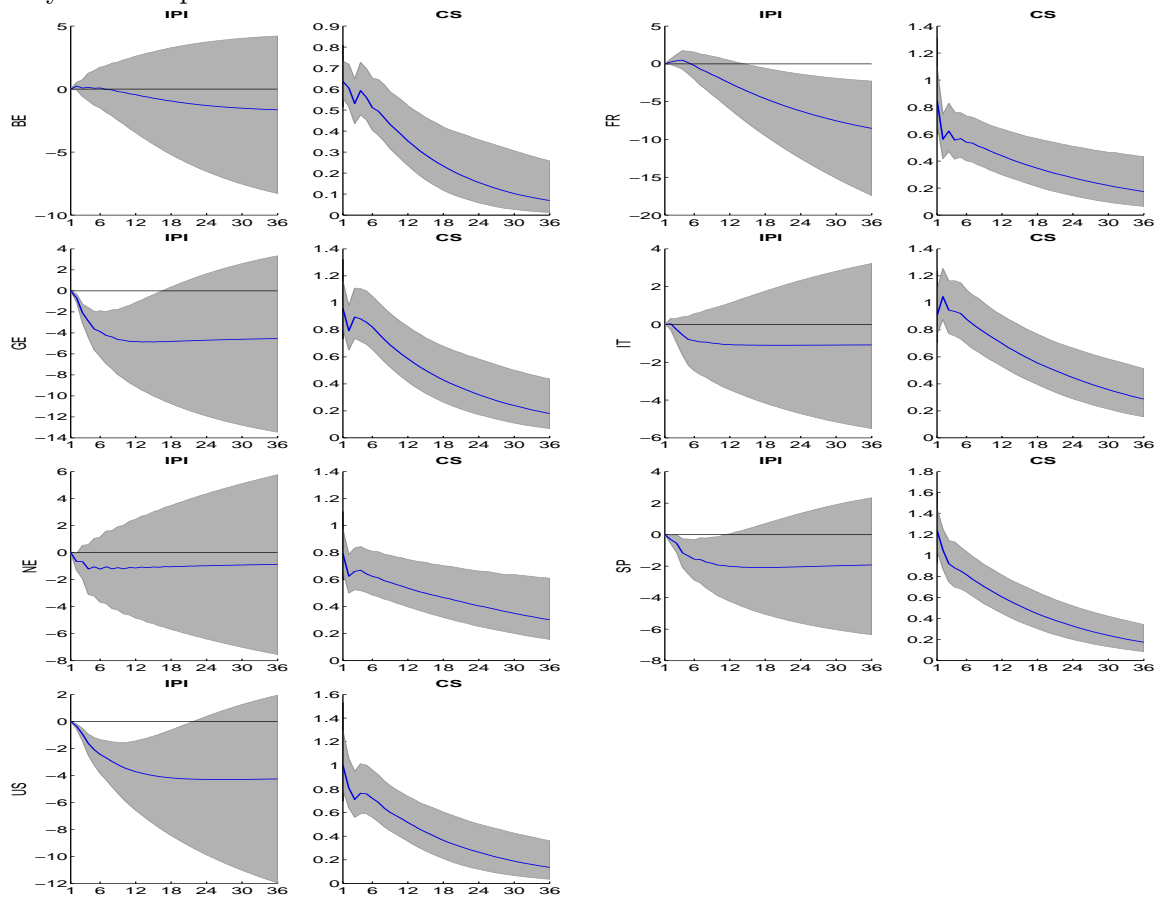


Figure 6: Median response (blue line) and 68% confidence interval to a credit shock in the first regime (recession) for industrial production growth (IPI) and credit spread (CS) for the different countries. IRs are standardized such as the median response of US credit spread in the first regime is 1. The response for IPI is plotted as cumulative sum over horizons. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

³Given that the autoregressive coefficients are state-independent, the impulse responses for regimes 2 and 3 differ only in scale (due to state-specific volatilities) and not in shape.

5 Conclusions

We have proposed a new Bayesian panel VAR model with unit-specific time-varying Markov-switching latent factors and developed a suitable Gibbs sampling procedure for posterior inference. We applied our panel MS-VAR model to the analysis of possible interconnections between Eurozone and US business cycles.

The evidence found is summarized in four major characteristics.

Number, empirical identification and persistence of regimes

Predictive likelihood analysis finds three regimes: recession, slow growth, and expansion. Graphical analysis using posterior densities of conditional means of IPI growth indicate that these densities empirically identify recession regimes and that posterior densities of credit spread volatility identify slow growth regimes. Credit spreads provide accurate predictive content for business cycle regimes. Slow growth becomes persistent in the Eurozone in recent years different from the US.

Lead-lag patterns in turning points

The US and Eurozone regimes are not fully synchronized over the 1991-2013 period, with evidence of more recessions in the Eurozone, in particular during the 90's. The larger synchronization is at beginning of the great Financial Crisis. This shock first affects the US economy and then it spreads among economies rapidly.

Amplification of recession probabilities through interconnections

Amplification effects in the recession probabilities and in the probabilities of exiting recessions occurs for both Eurozone and US with substantial differences in phase transitions within the Eurozone. Recently more imbalance within the Eurozone regimes takes place.

Shock effects

A credit shock results in a statistically significant negative industrial production growth for several months in Germany, Spain and the US.

The majority of our empirical results point towards the existence of two groups of countries in the Eurozone: A group of core countries: Germany, France and Netherlands with Italy sometimes belonging in the core and sometimes not and a group of peripheral countries Spain and sometimes Italy. Belgium taking a special position probably due to the de-industrialization in the nineties. This result is fairly robust when one applies cluster analysis. With respect to transmissions of shocks we found evidence that US credit and other recessionary shocks affect Germany and Spain first and spread then more or less

among other countries.

Our empirical results need to be investigated further using, for instance, forecasting but they may serve already as important information for the specification of a coordinated fiscal and monetary policy between the Eurozone and the US economies and also within the Eurozone economies that should be aimed at more adequate dynamic intervention with respect to shocks and reducing European imbalances.

References

- Amisano, G. and Tristani, O. (2013). Fundamentals and contagion mechanisms in the Euro area sovereign bonds markets. In *Proceedings from the SIS 2013 Conference*.
- Bassetti, F., Casarin, R., and Leisen, F. (2014). Pitman-Yor process prior for Bayesian inference. *Journal of Econometrics*, 180:49–72.
- Basturk, N., Cakmakli, C., Ceyhan, P., and van Dijk, H. K. (2014). Posterior-predictive evidence on US inflation using extended Phillips curve models with non-filtered data. *Journal of Applied Econometrics*, 29(7):1164–1182.
- Billio, M., Casarin, R., Ravazzolo, F., and van Dijk, H. K. (2012). Combination schemes for turning point predictions. *Quarterly Review of Economics and Finance*, 52(4):402–412.
- Billio, M., Casarin, R., Ravazzolo, F., and van Dijk, H. K. (2013a). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics*, 177:213–232.
- Billio, M., Ferrara, L., Guegan, D., and Mazzi, G. (2013b). Evaluation of regime-switching models for real-time business cycle analysis of the Euro area. *Journal of Forecasting*, 32(7):577–586.
- Bry, G. and Boschan, C. (1971). Cyclical analysis of time series: Selected procedures and computer programs. NBER Technical Paper 20.
- Canova, F. and Ciccarelli, M. (2004). Forecasting and turning point predictions in a Bayesian panel VAR model. *Journal of Econometrics*, 120(2):327–359.
- Canova, F. and Ciccarelli, M. (2009). Estimating multicountry VAR models. *International Economic Review*, 50(3):929–959.
- Chauvet, M. and Piger, J. (2008). Comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics*, 26:42–49.
- Clements, M. P. and Krolzig, H. M. (1998). A comparison of the forecast performances of Markov-switching and threshold autoregressive models of US GNP. *Econometrics Journal*, 1:C47–C75.
- Del Negro, M., Giannoni, M., and Schorfheide, F. (2014). Inflation in the great recession and new keynesian models. *American Economic Journal: Macroeconomics*, (forthcoming).
- Ferrara, L. (2003). A three-regime real-time indicator for the US economy. *Economics Letters*, 81:373–378.
- 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). *Mixture and Markov-switching Models*. Springer, New York.
- Gilchrist, S. and Mojon, B. (2014). Credit risk in the euro area. Technical report, NBER Working Paper No. 20041.
- Gilchrist, S. and Zakrajsek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–384.
- Harding, D. (2010). Applying shape and phase restrictions in generalized dynamic categorical models of the business cycle. NCER Working Paper Series 58, National Centre for Econometric Research.
- Harding, D. and Pagan, A. (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics*, 49:365–381.
- Harding, D. and Pagan, A. (2011). An econometric analysis of some models for constructed binary time series. *Journal of Business & Economic Statistics*, 29(1):86–95.
- 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(1):82–94.
- Kim, C. J. and Murray, C. J. (2002). Permanent and transitory components of recessions. *Empirical Economics*, 27:163–183.
- Kim, C. J. and Piger, J. (2002). Common stochastic trends, common cycles, and asymmetry in economic fluctuations. *Journal of Monetary Economics*, 49(6):1189 – 1211.
- Krolzig, H.-M. (1997). *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*. Springer, Berlin.
- Krolzig, H.-M. (2000). Predicting Markov-switching vector autoregressive processes. Nuffield College Economics Working Papers, 2000-WP31.
- McCulloch, R. E. and Tsay, R. S. (1994). Statistical analysis of economic time series via markov switching models. *Journal of Time Series Analysis*, 15:523–539.
- Mönch, E. and Uhlig, H. (2005). Towards a monthly business cycle chronology for the euro area. *Journal of Business Cycle Measurement and Analysis*, 2:43–69.
- Paap, R. and van Dijk, H. K. (1998). Distribution and mobility of wealth of nations. *European Economic Review*, 42:1269–1293.
- Schotman, P. and van Dijk, H. K. (1991). A Bayesian analysis of the unit root in real exchange rates. *Journal of Econometrics*, 49(1-2):195–2381.
- Sims, C. and Zha, T. (2006). Were there regime switches in us monetary policy? *American Economic Review*, 1:54–81.

Online Appendix

A Literature, data and posteriors

A.1 Summary of literature

One of our aims is to provide useful information on business cycle synchronization and heterogeneity across countries and to investigate how shocks transmit across countries and regions. In the literature there is no consensus on the international transmission of shocks. For example, Canova and Marrinan (1998) address the question whether international business cycles originate from common shocks or from a common propagation mechanism. Monfort et al. (2003) try to disentangle common shocks from spill-over effects. To this end, they estimate a Bayesian dynamic factor model for the G7 real output growth, featuring a global common factor and two area specific (North-American and Continental European) common factors, which, being modelled as a VAR process, are interdependent. They find empirical support for the presence of spill-over effects running from North-America to Continental Europe, but not vice versa. Our approach and empirical application aim to contribute to this debate by describing country specific cycles and their Interconnections.

We also contribute to the literature on the analysis of the business cycle of large panel of countries. A complete description of this literature is beyond the scope of our paper but we summarize the issue. A first attempt to model an international business cycle is by Gregory et al. (1997), who consider output, consumption and investment for G7 countries and estimate a dynamic factor model featuring a common cycle, a country-specific component and a series-specific one. The specification extends the Stock and Watson (1991) single index model and allow the authors to conclude that both the common and the country-specific factors capture a significant amount of fluctuations. Kose et al. (2003) reach similar conclusions, using a larger data set on 60 countries and using a Bayesian dynamic factor model. Kose et al. (2012) find, however, that the relative importance of the common factor has been declining over time and that the cycle of emerging economies has become decoupled from that of industrialized countries. Lumsdaine and Prasad (2003) assess the relative importance of country specific versus common shocks, using industrial production growth for a set of 17 countries. They estimate the common component of international fluctuations by aggregation with time-varying weights. In the present paper

we contribute and generalize the literature in this direction by focusing on the business cycle of the Eurozone, represented by the cycles of its six largest economies, and US economies. We measure the cycle by using multivariate series and extract several features of the country-specific business cycles in order to investigate the similarities and differences in booms and busts between the Eurozone cycle at an aggregated level and the US one, and further among the cycles of the Eurozone countries. As regards the relationship with the US cycle, note that the dating of the turning point is based on industrial production and not on the GDP, thus the estimated turning point dates may differ from the one officially published by the NBER, see for the example the dates around the internet bubble.

From a methodological point of view, this paper aims to contribute to the econometric literature on heterogeneity in cross-country panel data models. The more recent approaches have focused on two issues: the estimation of international cycles focusing on the nature of the co-movements using relatively large dimensional data sets; and the introduction of country and time heterogeneity in multi-country vector autoregressive models. The first issue has been considered by Hallin and Liska (2008), Pesaran et al. (2004), and Dees et al. (2007) and the second by Canova and Ciccarelli (2004) and Canova and Ciccarelli (2009). Hallin and Liska (2008) extend the generalized dynamic factor model proposed by Forni et al. (2000, 2001) to a panel of time series with a block structure, where the blocks are represented by countries. They show that the extension provides the means for the analysis of the interblock relationships, allowing the identification of strongly common factors, which are common to all the blocks (e.g. international common factors), strongly idiosyncratic factors, which are idiosyncratic for all blocks, and weakly common/weakly idiosyncratic factors, that are common to at least one block, but idiosyncratic to at least another one.

Multi-country VAR models provide a tool for examining shock propagation across countries. Canova and Ciccarelli (2009) consider Bayesian multi-country VAR models with time varying parameters, lagged interdependencies and country specific effects. They avoid the curse of dimensionality on the number of parameters by a factorial parameterization of the time varying VAR coefficients in terms of a number of continuous random effects that are linear in the number of countries and series. The authors propose a Monte Carlo Markov Chain sampling scheme for posterior approximation. In general, overparameterization may lead to overfitting, especially in macroeconomic applications, where time series are characterized by a low number of observations, slowly changing means and time-varying

variances. These issues call for the use of a Bayesian approach to modeling and estimation, since it allows for inclusion of parameter restrictions with different degrees of prior belief (see, e.g., Litterman (1986), Sims and Zha (1998) for Bayesian VAR, Chib and Greenberg (1995) for Bayesian Seemingly Unrelated Regression and Canova and Ciccarelli (2009) for panel Bayesian VAR).

A.2 Data Summary

IPI growth is one of the main economic indicators that measures changes in output for the manufacturing, mining, and utilities sectors. Although these sectors contribute to a moderate fraction of GDP and several countries have partially shifted from being production oriented to being service and consumer oriented, which reduces the contribution of these sectors even more, they are rather sensitive to variations in interest rates and consumer demand. This makes IPI growth an important variable for forecasting the future economic performance of an economic system. We downloaded IPI data from the OECD database.

Financial shocks have been found to play an important role in economic fluctuations, both as a transmission mechanism of other shocks to the real sector, see Claessens et al. (2009) who link shock transmissions from the financial sector to the real one using a large set of variables; and as a source of shocks itself, see e.g. Furlanetto et al. (2014). Del Negro et al. (2014) discuss how a standard DSGE model extended to include financial frictions measured by credit spread predicts the US 2008 recession reasonably well. These authors define credit spread as the rate entrepreneurs pay in excess of a deposit yield to finance their projects and measure it as the difference between the Baa corporate bond yield minus the 10-Year Treasury Note Yield. We use the same variable for the US. The construction of this variable is more difficult for the European countries. First, at the European country-level corporate bond indices are not easily available. We collect corporate bond yields from the Global Financial Data site (<https://www.globalfinancialdata.com/index.html>). Yield indices are combination of many underlying securities and based upon long-term (10-30 year) bonds of investment grade (AAA to BBB), with average rate A or Baa. Second, the definition of a deposit rate in the various European countries is problematic. We follow Gilchrist and Mojon (2014) and compute the spread over the 10 year German Bund yield for all six Eurozone countries. Government Treasury yields are downloaded from Datastream. Unfortunately, we are obliged to ignore the exchange rate risk before the euro was introduced

as an accounting currency on January 1, 1999.

A.3 Parameter full conditional densities

Updating γ_{i0} . Then the full conditional distribution of the regime-independent parameter γ_{i0} is a normal with density function

$$\begin{aligned} f(\gamma_{i0}|\mathbf{y}_i, \Xi_i, \gamma_i, \Sigma_i, \boldsymbol{\lambda}_0) &\propto & (A.1) \\ &\propto \exp \left\{ -\frac{1}{2} \gamma'_{i0} \left(\sum_{t=1}^T X'_{i0t} \Sigma_{it}^{-1} X_{i0t} + \underline{\Sigma}_{i0}^{-1} \right) \gamma_{i0} + \gamma_{i0} \left(\sum_{t=1}^T X'_{i0t} \Sigma_{it}^{-1} \mathbf{y}_{i0t} + \underline{\Sigma}_{i0}^{-1} \boldsymbol{\lambda}_0 \right) \right\} \\ &\propto \mathcal{N}_{MM_0}(\bar{\gamma}_{i0}, \bar{\Sigma}_{i0}) \end{aligned}$$

where $\mathbf{y}_{i0t} = \mathbf{y}_{it} - (\xi_{i1t} X_{i1t} \gamma_{i1} + \dots + \xi_{iKt} X_{iKt} \gamma_{iK})$, $\bar{\gamma}_{i0} = \bar{\Sigma}_{i0} (\underline{\Sigma}_{i0}^{-1} \boldsymbol{\lambda}_0 + \sum_{t=1}^T X'_{i0t} \Sigma_{it}^{-1} X_{i0t})$ and $\bar{\Sigma}_{i0}^{-1} = (\underline{\Sigma}_{i0}^{-1} + \sum_{t=1}^T X'_{i0t} \Sigma_{it}^{-1} X_{i0t})$.⁴

Updating γ_{ik} . The full conditional distributions of the regime-dependent parameters γ_{ik} , with $k = 1, \dots, K$ are normal with density function

$$\begin{aligned} f(\gamma_{ik}|\mathbf{y}_i, \Xi_i, \gamma_{i0}, \gamma_{i(-k)}, \Sigma, \boldsymbol{\lambda}_k) &\propto & (A.2) \\ &\propto \exp \left\{ -\frac{1}{2} \gamma'_i \left(\sum_{t \in \mathcal{T}_{ik}} X'_{ikt} \Sigma_{it}^{-1} X_{ikt} + \underline{\Sigma}_{ik}^{-1} \right) \gamma_i + \gamma'_i \left(\sum_{t \in \mathcal{T}_{ik}} X'_{ikt} \Sigma_{it}^{-1} \mathbf{y}_{ikt} + \underline{\Sigma}_{ik}^{-1} \boldsymbol{\lambda}_k \right) \right\} \\ &\propto \mathcal{N}_{MM_K}(\bar{\gamma}_{ik}, \bar{\Sigma}_{ik}) \end{aligned}$$

with $\bar{\gamma}_{ik} = \bar{\Sigma}_{ik}^{-1} (\underline{\Sigma}_{ik}^{-1} \boldsymbol{\lambda}_k + \sum_{t \in \mathcal{T}_{ik}} X'_{ikt} \Sigma_{it}^{-1} X_{ikt})$ and $\bar{\Sigma}_{ik}^{-1} = (\underline{\Sigma}_{ik}^{-1} + \sum_{t \in \mathcal{T}_{ik}} X'_{ikt} \Sigma_{it}^{-1} X_{ikt})$, where we defined $\mathcal{T}_{ik} = \{t | \xi_{ikt} = 1, t = 1, \dots, T\}$ and $\mathbf{y}_{ikt} = \mathbf{y}_{it} - X_{i0t} \gamma_{i0}$. An accept/reject method is applied to account for the set of identification constraints on the parameters γ_{ik} , $k = 1, \dots, K$.⁵

Updating Σ_{ik}^{-1} . The full conditional distributions of the regime-dependent inverse variance-covariance matrix Σ_{ik}^{-1} , $k = 1, \dots, K$, are Wishart distributions with density

⁴Anas et al. (2008) also applies constant autoregressive coefficients and only unit-specific intercepts and covariance matrices vary over regimes.

⁵See, for example, Celeux (1998) and Frühwirth-Schnatter (2001) for a discussion on the effects that label switching and the lack of identification have on the results of a MCMC based Bayesian inference.

$$\begin{aligned}
& f(\Sigma_{ik}^{-1} | \mathbf{y}_i, \Xi_i, \gamma_{i0}, \gamma_i, \Sigma_{i(-k)}, \Upsilon_k) \propto \\
& \propto |\Sigma_{ik}^{-1}|^{\frac{\nu_{ik} + T_{ik} - M - 1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left(\left(\Upsilon_k^{-1} + \sum_{t \in \mathcal{T}_{ik}} \mathbf{u}_{ikt} \mathbf{u}'_{ikt} \right) \Sigma_{ik}^{-1} \right) \right\} \\
& \propto \mathcal{W}_M(\bar{\nu}_{ik}, \bar{\Upsilon}_{ik})
\end{aligned} \tag{A.3}$$

where $T_{ik} = \sum_{t=1}^T \mathbb{I}(\xi_{ikt} = 1)$, $\mathbf{u}_{ikt} = \mathbf{y}_{it} - X_{i0t}\gamma_{i0} - X_{ikt}\gamma_{ik}$, $\bar{\nu}_{ik} = \nu_{ik} + T_{ik}$ and $\bar{\Upsilon}_{ik}^{-1} = \Upsilon_k^{-1} + \sum_{t \in \mathcal{T}_{ik}} \mathbf{u}_{ikt} \mathbf{u}'_{ikt}$.

Updating α_i . The full conditional distribution of the parameters in the l -th row of the transition matrix, $\alpha_i^l = \text{vec}((\alpha_i^{1l}, \dots, \alpha_i^{K-1l}))$, is

$$\begin{aligned}
& f(\alpha_i^l | \mathbf{y}_i, \Xi, \gamma_{i0}, \gamma_i) \propto \prod_{k=1}^{K-1} \exp \left\{ -\frac{1}{2} (\alpha_i^{kl} - \boldsymbol{\psi})' \Upsilon^{-1} (\alpha_i^{kl} - \boldsymbol{\psi}) \right\} \\
& \prod_{t=1}^T \prod_{k=1}^K (H(V_t, \alpha_i^{kl}))^{\xi_{ikt} \xi_{il t-1}}
\end{aligned} \tag{A.4}$$

Note that we update the coefficients across the past state and not across the current state as in Kaufmann (2015). As regards the sampling method, we follow Lenk and DeSarbo (2000) and apply a Metropolis-Hastings with proposal centred at the mode of the posterior distribution, with scale proportional to the posterior local curvature. In general, the mode is not known, and its estimate is updated at each iteration by applying an iteration of the Newton-Raphson algorithm.

We consider the quadratic approximation of the full conditional log-density, $\tilde{f}(\alpha_i^l) = \log f(\alpha_i^l | \mathbf{y}_i, \Xi, \gamma_{i0}, \gamma_i)$. The gradient or vector of partial derivatives of the log-posterior is denoted with $D_1(\alpha_i^l) = \nabla_{\alpha_i^l}^{(1)} \tilde{f}(\alpha_i^l) = \text{vec}(\nabla_{\alpha_i^{1l}}^{(1)} \tilde{f}(\alpha_i^l), \dots, \nabla_{\alpha_i^{1K-1l}}^{(1)} \tilde{f}(\alpha_i^l))$. Since

$$\begin{aligned}
\nabla_{\boldsymbol{\alpha}_i^{kl}}^{(1)} \tilde{f}(\boldsymbol{\alpha}_i^l) &= \sum_{t=1}^T \xi_{ilt-1} \left(\xi_{ikt} \frac{1}{H(V_t, \boldsymbol{\alpha}_i^{kl})} \nabla_{\boldsymbol{\alpha}_i^{kl}} (H(V_t, \boldsymbol{\alpha}_i^{kl})) \right. \\
&\quad + \sum_{k' \neq k, K} \xi_{ik't} \frac{1}{H(V_t, \boldsymbol{\alpha}_i^{k'l})} \nabla_{\boldsymbol{\alpha}_i^{k'l}} (H(V_t, \boldsymbol{\alpha}_i^{k'l})) \\
&\quad \left. + \xi_{iKt} \frac{1}{H(V_t, \boldsymbol{\alpha}_i^{Kl})} \nabla_{\boldsymbol{\alpha}_i^{Kl}} (H(V_t, \boldsymbol{\alpha}_i^{Kl})) \right) - \Upsilon^{-1}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi}) \\
&= \sum_{t=1}^T \xi_{ilt-1} \left(\xi_{ikt} (1 - H(V_t, \boldsymbol{\alpha}_i^{kl})) \mathbf{z}_t - \sum_{k' \neq k} \xi_{ik't} H(V_t, \boldsymbol{\alpha}_i^{k'l}) \mathbf{z}_t \right) - \Upsilon^{-1}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi}) \quad (\text{A.5})
\end{aligned}$$

we conclude that

$$D_1(\boldsymbol{\alpha}_i^l) = \sum_{t=1}^T \xi_{ilt-1} \left(E(\boldsymbol{\xi}_{it} - H(V_t, \boldsymbol{\alpha}_i^l)) \right) \otimes \mathbf{z}_t - (I_{K-1} \otimes \Upsilon^{-1})(\boldsymbol{\alpha}_i^l - \boldsymbol{\nu}_{K-1} \otimes \boldsymbol{\psi}) \quad (\text{A.6})$$

where $\mathbf{z}_t = (V_t, 1)'$ is a $(G_v + 1)$ -dimensional vector, $E = (I_{K-1}, \mathbf{0}_{K-1})$ is a selection matrix, with $\mathbf{0}_n$ a n -dimensional null vector, and $H(V_t, \boldsymbol{\alpha}_i^l) = (H(V_t, \boldsymbol{\alpha}_i^{1l}), \dots, H(V_t, \boldsymbol{\alpha}_i^{Kl}))'$.

The Hessian or matrix of second derivatives is denoted with $D_2(\boldsymbol{\alpha}_i^l) = \nabla_{\boldsymbol{\alpha}_i^l}^{(2)} \tilde{f}(\boldsymbol{\alpha}_i^l)$. The k' -th row of the Hessian is $(\nabla_{\boldsymbol{\alpha}_i^{k'l} \boldsymbol{\alpha}_i^{1l}}^{(2)}, \dots, \nabla_{\boldsymbol{\alpha}_i^{k'l} \boldsymbol{\alpha}_i^{K-1l}}^{(2)})$ where

$$\begin{aligned}
\nabla_{\boldsymbol{\alpha}_i^{k'l} \boldsymbol{\alpha}_i^{kl}}^{(2)} &= \sum_{t=1}^T \xi_{ilt-1} \left(H(V_t, \boldsymbol{\alpha}_i^{k'l}) H(V_t, \boldsymbol{\alpha}_i^{kl}) \right. \\
&\quad \left. - H(V_t, \boldsymbol{\alpha}_i^{kl}) \mathbb{I}(k = k') \right) \mathbf{z}_t \mathbf{z}_t' - \mathbb{I}(k = k') \Upsilon^{-1} \quad (\text{A.7})
\end{aligned}$$

thus the Hessian is

$$D_2(\boldsymbol{\alpha}_i^l) = \sum_{t=1}^T \xi_{ilt-1} \left((EH(V_t, \boldsymbol{\alpha}_i^l)(EH(V_t, \boldsymbol{\alpha}_i^l))' - \text{diag}(EH(V_t, \boldsymbol{\alpha}_i^l))) \otimes (\mathbf{z}_t \mathbf{z}_t') - I_{K-1} \otimes \Upsilon^{-1} \right) \quad (\text{A.8})$$

Then the mode of the log full conditional is updated at each iteration of the M.-H. with the Newton-Raphson's rule

$$\hat{\boldsymbol{\alpha}}^{(n+1)} = \hat{\boldsymbol{\alpha}}^{(n)} - D_2(\hat{\boldsymbol{\alpha}}^{(n)})^{-1} D_1(\hat{\boldsymbol{\alpha}}^{(n)}) \quad (\text{A.9})$$

The proposal distribution is adapting over the iterations and at the iteration n the

proposal $\boldsymbol{\alpha}^*$ for the parameter $\boldsymbol{\alpha}_i^l$ is drawn from the normal

$$\boldsymbol{\alpha}^* \sim \mathcal{N}_{(K-1)(G_v+1)}(\hat{\boldsymbol{\alpha}}^{(n+1)}, V^{(n)}) \quad (\text{A.10})$$

where $V^{(n)} = -D_2(\hat{\boldsymbol{\alpha}}^{(n)})$. After an initial, transitory period the adaptation of the proposal stops and $\hat{\boldsymbol{\alpha}}^{(n)}$ and $V^{(n)}$ are not updated. Let $(\boldsymbol{\alpha}_i^l)^{(n-1)}$ be the current value of the chain, then at the iteration n , the proposal is accepted with log-probability

$$\varrho \left((\boldsymbol{\alpha}_i^l)^{(n-1)}, \boldsymbol{\alpha}^* \right) = \min \left\{ 0, \left(\tilde{f}(\boldsymbol{\alpha}_i^*) - \tilde{f}((\boldsymbol{\alpha}_i^l)^{(n-1)}) - \frac{1}{2}((\boldsymbol{\alpha}_i^l)^{(n-1)} - \hat{\boldsymbol{\alpha}}_i^{(n+1)})'(V^{(n)})^{-1} \right. \right. \\ \left. \left. ((\boldsymbol{\alpha}_i^l)^{(n-1)} - \hat{\boldsymbol{\alpha}}_i^{(n+1)}) - \frac{1}{2}(\boldsymbol{\alpha}_i^* - \hat{\boldsymbol{\alpha}}_i^{(n+1)})'(V^{(n)})^{-1}(\boldsymbol{\alpha}_i^* - \hat{\boldsymbol{\alpha}}_i^{(n+1)}) \right) \right\} \quad (\text{A.11})$$

Updating $\boldsymbol{\lambda}_0$. The full conditional distribution of the parameter $\boldsymbol{\lambda}_0$, of the third stage of the hierarchical structure, is a normal distributions with density function

$$f(\boldsymbol{\lambda}_0 | \boldsymbol{\gamma}_0, \boldsymbol{\Sigma}_0) \propto \quad (\text{A.12}) \\ \propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\lambda}'_0 \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{i0}^{-1} + \boldsymbol{\Sigma}_0^{-1} \right) \boldsymbol{\lambda}_0 - 2\boldsymbol{\lambda}'_0 \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{i0}^{-1} \boldsymbol{\gamma}_{i0} + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\lambda}_0 \right) \right] \right\} \\ \propto \mathcal{N}_{MM_0}(\bar{\boldsymbol{\lambda}}_0, \bar{\boldsymbol{\Sigma}}_0)$$

where $\bar{\boldsymbol{\Sigma}}_0^{-1} = \sum_{i=1}^N \boldsymbol{\Sigma}_{i0}^{-1} + \boldsymbol{\Sigma}_0^{-1}$ and $\bar{\boldsymbol{\lambda}}_0 = \bar{\boldsymbol{\Sigma}}_0 \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{i0}^{-1} \boldsymbol{\gamma}_{i0} + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\lambda}_0 \right)$.

Updating $\boldsymbol{\lambda}_k$. The full conditional distributions of the parameters $\boldsymbol{\lambda}_k$, $k = 1, \dots, K$, of the third stage of the hierarchical structure, are normal distributions with density functions

$$f(\boldsymbol{\lambda}_k | \boldsymbol{\gamma}_k, \boldsymbol{\Sigma}_k) \propto \quad (\text{A.13}) \\ \propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\lambda}'_k \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{ik}^{-1} + \boldsymbol{\Sigma}_k^{-1} \right) \boldsymbol{\lambda}_k - 2\boldsymbol{\lambda}'_k \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{ik}^{-1} \boldsymbol{\gamma}_{ik} + \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\lambda}_k \right) \right] \right\} \\ \propto \mathcal{N}_{MM_K}(\bar{\boldsymbol{\lambda}}_k, \bar{\boldsymbol{\Sigma}}_k)$$

where $\bar{\boldsymbol{\Sigma}}_k^{-1} = \sum_{i=1}^N \boldsymbol{\Sigma}_{ik}^{-1} + \boldsymbol{\Sigma}_k^{-1}$ and $\bar{\boldsymbol{\lambda}}_k = \bar{\boldsymbol{\Sigma}}_k \left(\sum_{i=1}^N \boldsymbol{\Sigma}_{ik}^{-1} \boldsymbol{\gamma}_{ik} + \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\lambda}_k \right)$.

Updating Υ_k^{-1} . The full conditional distributions of the Υ_k , $k = 1, \dots, K$, are Wishart

distributions with density

$$\begin{aligned}
f(\Upsilon_k^{-1}|\gamma_k, \Sigma_k) &\propto \\
&\propto |\Upsilon_k^{-1}|^{\frac{\nu_k - M - 1}{2}} \exp\left\{-\frac{1}{2}\text{tr}(\underline{\Upsilon}_k^{-1}\Upsilon_k^{-1})\right\} \prod_{i=1}^N |\Upsilon_k^{-1}|^{\nu_{ik}/2} \exp\left\{-\frac{1}{2}\text{tr}\left(\sum_{i=1}^N \Upsilon_k^{-1}\Sigma_{ik}^{-1}\right)\right\} \\
&\propto \mathcal{W}_M(\bar{\nu}_k, \bar{\Upsilon}_k)
\end{aligned} \tag{A.14}$$

where $\bar{\nu}_k = \sum_{i=1}^N \nu_{ik} + \nu_k$ and $\bar{\Upsilon}_k^{-1} = \underline{\Upsilon}_k^{-1} + \sum_{i=1}^N \Sigma_{ik}^{-1}$.

A.4 Allocation variable full conditional distributions

To sample the hidden states we propose a multi-move strategy. In Krolzig (1997) a multi-move Gibbs sampler (see Carter and Kohn (1994) and Shephard (1994)) is presented for Markov-switching vector autoregressive models as an alternative to the single-move Gibbs sampler introduced, for example, in Albert and Chib (1993). The multi-move procedure, also known as forward-filtering backward sampling (FFBS) algorithm, is particularly useful in highly parameterized model, because it can improve the mixing of the MCMC chain over a large parameter space, thus leading to a more efficient posterior approximation. Unfortunately, the FFBS does not apply easily to our model due to the presence of the chain interaction mechanism. In fact, the FFBS should be iterated jointly for all the Markov-switching processes of the panel implying large matrix operations and, therefore, a high computational cost. We follow a different route and apply here the FFBS to the unit-specific chains, conditioning on the sampled value of other chains in the panel.

Let us define $\boldsymbol{\xi}_{-i,1:T} = (\boldsymbol{\xi}_{1:i-1,1:T}, \boldsymbol{\xi}_{i+1:N,1:T})$, with $\boldsymbol{\xi}_{i,j,1:T} = (\boldsymbol{\xi}_{i,1:T}, \dots, \boldsymbol{\xi}_{j,1:T})$, $i \leq j$, and $p(\boldsymbol{\xi}_{it} = \nu_k | \boldsymbol{\xi}_{it-1} = \nu_l, \boldsymbol{\xi}_{it-2}, V_t, \boldsymbol{\alpha}_i) = p_{it,kl}$, with ν_k the k -th column of the identity matrix. The full conditional distribution of the allocation variables $\boldsymbol{\xi}_{i,1:T}$ is

$$\begin{aligned}
p(\boldsymbol{\xi}_{i,1:T} | \mathbf{y}_{1:T}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}) &\propto \\
&\propto \prod_{i=1}^N \prod_{t=1}^T \left(p(\mathbf{y}_{it} | \mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\Sigma}_i) \prod_{k,l=1}^K p_{it,kl}^{\xi_{it}\xi_{ikt-1}} \right) \\
&\propto \prod_{t=1}^T \left(p(\mathbf{y}_{it} | \mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\Sigma}_i) \prod_{k,l=1}^K p_{it,kl}^{\xi_{ikt}\xi_{ilt-1}} g_{it}(\boldsymbol{\xi}_{i,t-1} | \boldsymbol{\xi}_{-i,t-1}) \right)
\end{aligned} \tag{A.15}$$

where

$$g_{it}(\boldsymbol{\xi}_{it-1}|\boldsymbol{\xi}_{-i,t-1}) = \prod_{\substack{j=1 \\ j \neq i}}^N \prod_{k,l=1}^K \mathbb{P}(\boldsymbol{\xi}_{jt} = \boldsymbol{\nu}_k | \boldsymbol{\xi}_{jt-1} = \boldsymbol{\nu}_l, \boldsymbol{\xi}_{jt-2}, V_t, \boldsymbol{\alpha}_j)^{\xi_{jkt}\xi_{jlt-1}}$$

is a multiplicative factor that depends on the values $\boldsymbol{\xi}_{it-1}$, through some or all of the covariates $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{Kt})'$ appearing in V_t . For example, in our application we considered, $V_t = (1, \eta_{1t}, \mathbb{I}(s_{7t} = 1))'$.

Conditionally on the other unit state variables, the full conditional distribution of the i -th unit allocation variable at time t results from the product of the time-varying transition probability $\mathbb{P}(\boldsymbol{\xi}_{jt}|\boldsymbol{\xi}_{jt-1}, \boldsymbol{\xi}_{jt-2}, V_t, \boldsymbol{\alpha}_j)$ and the augmented likelihood $p(\mathbf{y}_{it}|\mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\xi}_{it-1}, \boldsymbol{\gamma}_i, \Sigma_i) = p(\mathbf{y}_{it}|\mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\gamma}_i, \Sigma_i)g_{it}(\boldsymbol{\xi}_{it-1}|\boldsymbol{\xi}_{-i,t-1})$, where $p(\mathbf{y}_{it}|\mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\gamma}_i, \Sigma_i)$ is the conditional distribution of the variable \mathbf{y}_{it} from our panel VAR model. These terms can be written as functions of the Markov chain $\tilde{s}_{it} = (s_{it-1} - 1)K + s_{it}$ with values in the space $\{1, 2, \dots, K^2\}$. This representation of the chain s_{it} allows us to easily impose the duration constraint when simulating the hidden states. Thus, we introduce a new allocation variable $\tilde{\boldsymbol{\xi}}_{it} = (\tilde{\xi}_{i1t}, \dots, \tilde{\xi}_{iK^2t})'$, where $\tilde{\xi}_{ikt} = \mathbb{I}(\tilde{s}_{it} = k)$. The transition probability of the process $\tilde{\boldsymbol{\xi}}_{it}$, for the case $K = 3$ and a minimum duration constrain of two periods is introduced (see equation 6) is:

$$\mathbb{P}(\tilde{\boldsymbol{\xi}}_{it}|\tilde{\boldsymbol{\xi}}_{it-1}, V_t, \boldsymbol{\alpha}_i) = \begin{pmatrix} h_{it,11} & h_{it,21} & h_{it,31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \end{pmatrix} \quad (\text{A.16})$$

with $h_{it,kl} = H(V_t, \boldsymbol{\alpha}_i^{kl})$, where in the l -th row the matrix the probability $\mathbb{P}(\tilde{s}_{it} = k | \tilde{s}_{it-1} = l, V_t, \boldsymbol{\alpha}_i)$ is given for $k = 1, \dots, K^2$.

In the simulation from the full conditional distribution of the hidden states, we exploit

the following factorization

$$p(\tilde{\boldsymbol{\xi}}_{i,1:T} | \mathbf{y}_{i,1:T}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}, \Sigma, \boldsymbol{\alpha}) \propto \left(p(\tilde{\boldsymbol{\xi}}_{iT} | \mathbf{y}_{i,1:T}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i) \right. \quad (\text{A.17}) \\ \left. \prod_{t=1}^{T-1} p(\tilde{\boldsymbol{\xi}}_{it+1} | \tilde{\boldsymbol{\xi}}_{it}, V_t, \boldsymbol{\alpha}_i) p(\tilde{\boldsymbol{\xi}}_{it} | \mathbf{y}_{i,1:t}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i) \right)$$

This factorization suggests that a Forward-Filtering Backward-Sampling (FFBS) algorithm can be used for the hidden states of the i -th bivariate chain \tilde{s}_{it} of the panel. At the iteration n -th of the Gibbs sampler, since the country specific state variables are updated sequentially over the unit index, a new trajectory $\boldsymbol{\xi}_{i,1:T}^{(n)}$ for the country i is generated from the FFBS algorithm, conditioning on the updated values of the hidden states $\boldsymbol{\xi}_{j,1:T}^{(n)}$, $j < i$, and the previous-iteration values of the hidden states $\boldsymbol{\xi}_{j,1:T}^{(n-1)}$, $j > i$. The steps of the FFBS algorithm is described in Frühwirth-Schnatter (2006), ch. 11.

Note that in the updating step of the FFBS, the filtered probability, $p(\tilde{\boldsymbol{\xi}}_{it} | \mathbf{y}_{i,1:t}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\alpha})$ is proportional to the product of two terms: $p(\mathbf{y}_{it} | \mathbf{y}_{it-P-1:t-1}, \boldsymbol{\xi}_{it}, \boldsymbol{\gamma}_i, \Sigma_i)$ and $git(\boldsymbol{\xi}_{it-1} | \boldsymbol{\xi}_{-i,t-1})$ which are evaluated at the different values of allocation variable. This requires the evaluation of V_t and $\boldsymbol{\eta}_t$ as a function of $\boldsymbol{\xi}_{j,1:T}^{(n)}$, $j < i$, $\boldsymbol{\xi}_{j,1:T}^{(n-1)}$, $j > i$ and $\boldsymbol{\xi}_{it-1} = \boldsymbol{\nu}_k$, $k = 1, \dots, K$. The elements of $\boldsymbol{\eta}_t$ are

$$\eta_{kt} = \omega_{it} \xi_{ikt-1} + \sum_{j < i} \omega_{jt} \xi_{jkt-1}^{(n)} + \sum_{j > i} \omega_{jt} \xi_{jkt-1}^{(n-1)}$$

The relationship $\xi_{ikt} = \sum_{l=K(k-1)+1}^{Kk} \tilde{\xi}_{ilt}$, $k = 1, \dots, K$, is used to find the value of the allocation variable in the single-chain representation. In order to obtain the draws of the hidden state variables from the draw of the allocation variables, the following transform $s_{it} = \sum_{k=1}^K k \xi_{ikt}$ is used.

As discussed in previous sections, when using data-dependent priors the generation of the allocation variables should omit draws that yield to impropriety of the posterior. In our prior settings, the set of non-troublesome grouping, for the i -th unit, is $\mathcal{S}_i = \mathcal{S}_{i,\nu} \cap \mathcal{S}_{i,\sigma} = \mathcal{S}_{i,\sigma}$. Thus, each time the set of allocation variables $\boldsymbol{\xi}_{i,1:T}$, does not assign at least two observations to each component of the dynamic mixture, the entire set $\boldsymbol{\xi}_{i,1:T}$, is rejected and a new set is drawn until a proper set is obtained.

B MCMC convergence issues

We apply the Gibbs sampler described in Section 3 and obtain the posterior densities of the PMS-VAR model parameters. These posterior densities are then approximated through a kernel density estimator applied to a sample of 4,000 random draws from the posterior. In order to generate 4,000 i.i.d. sample from the posterior, we run the Gibbs sampler, for 50,000 iterations, discard the first 10,000 draws to avoid dependence from the initial condition, and finally apply a thinning procedure with a factor of 10, to reduce the dependence between consecutive Markov-chain draws. See Section B in the Online Appendix for further details on the choice of the number of iterations and of the burn in samples.

As regards to the number of iterations, the choice of the initial sample size and the convergence detection of the Gibbs sampler remain open issues (see Robert and Casella (1999)). In our application we choose the sample size on the basis of both a graphical inspection of the MCMC progressive averages and the application of the convergence diagnostic statistics (CD) proposed in Geweke (1992). We let n be the MCMC sample size and $n_1 = 0.1n$, and $n_2 = 0.5n$ the sizes of two non-overlapping sub-samples. For a parameter θ of interest, we let

$$\hat{\theta}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} \theta^{(j)}, \quad \hat{\theta}_2 = \frac{1}{n_2} \sum_{j=n+1-n_2}^n \theta^{(j)}$$

be the MCMC sample means and $\hat{\sigma}_i^2$ their variances estimated as:

$$\frac{\hat{\sigma}_i^2}{n_i} = \hat{\Gamma}(0) + \frac{2n_i}{n_i - 1} \sum_{j=1}^{h_i} K(j/h_i) \hat{\Gamma}(j), \quad (\text{B.1})$$

$$\hat{\Gamma}(j) = \frac{1}{n_i} \sum_{k=j+1}^{n_i} (\theta^{(k)} - \hat{\theta}_i)(\theta^{(k-j)} - \hat{\theta}_i)' \quad (\text{B.2})$$

where we choose $K(x)$ to be the Parzen kernel and $h_1 = n_1^{1/4}$ and $h_2 = n_2^{1/4}$ the bandwidths (see Horvath and Rice (2015)). Then the following statistics

$$CD = \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{\hat{\sigma}_1^2/n_1 + \hat{\sigma}_2^2/n_2}} \quad (\text{B.3})$$

converges in distribution to a standard normal (see Geweke (1992)), under the null hypothesis that the MCMC chain has converged.

C Smoothed probabilities

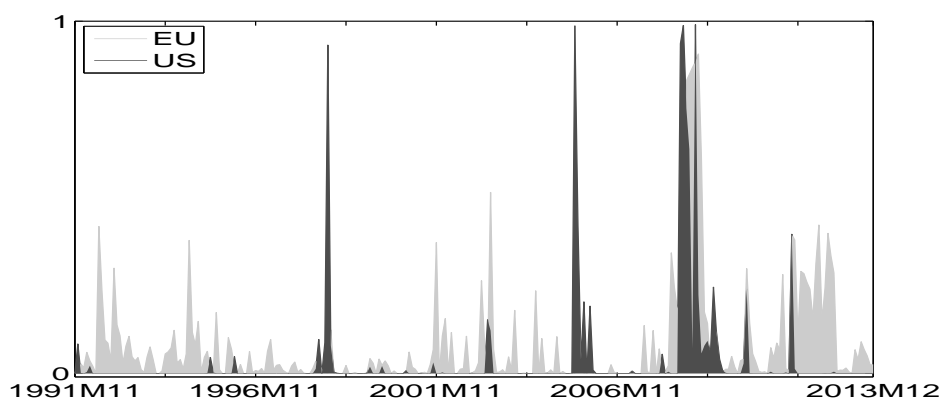


Figure C.1: Smoothed probabilities of the Eurozone and US economies to be in recessions

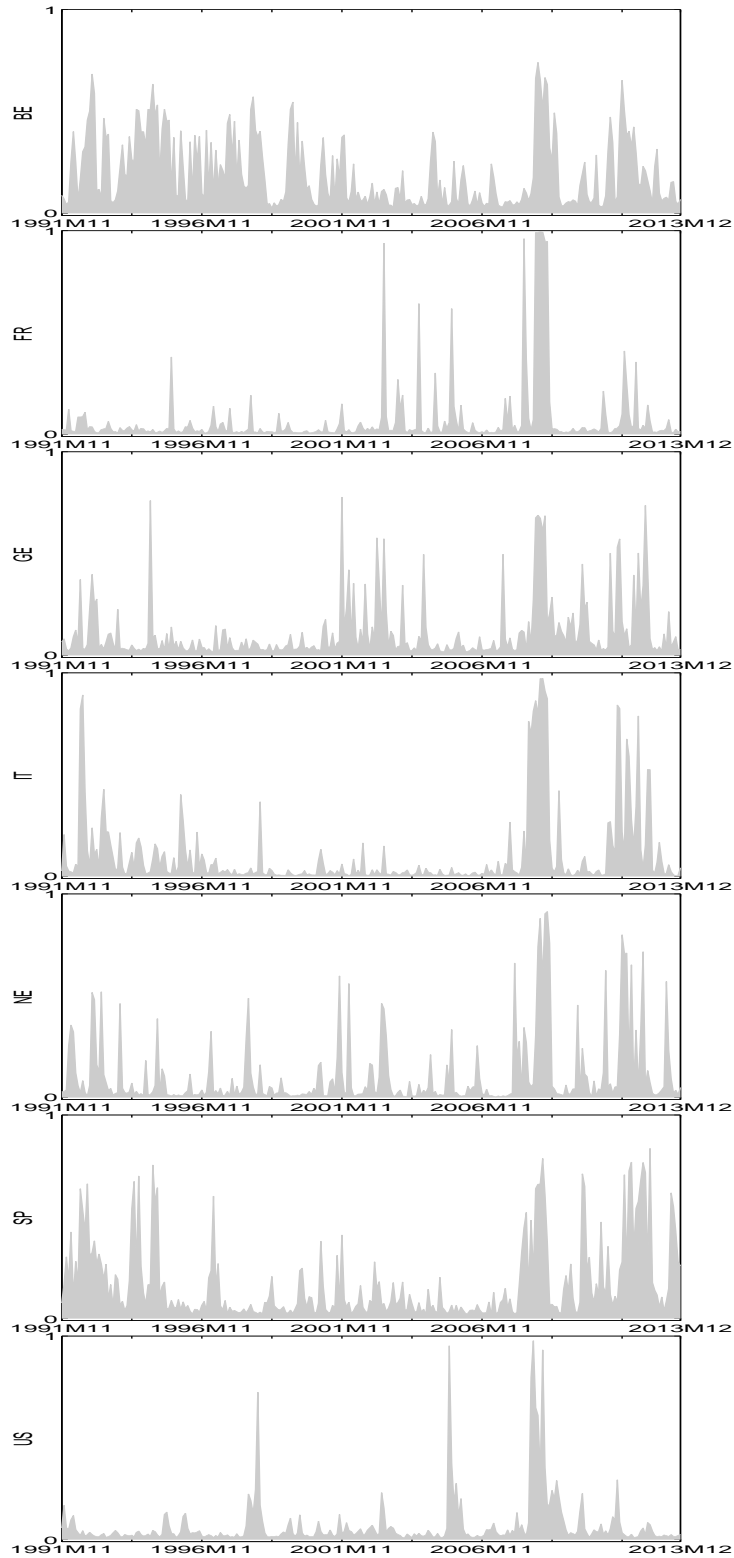


Figure C.2: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

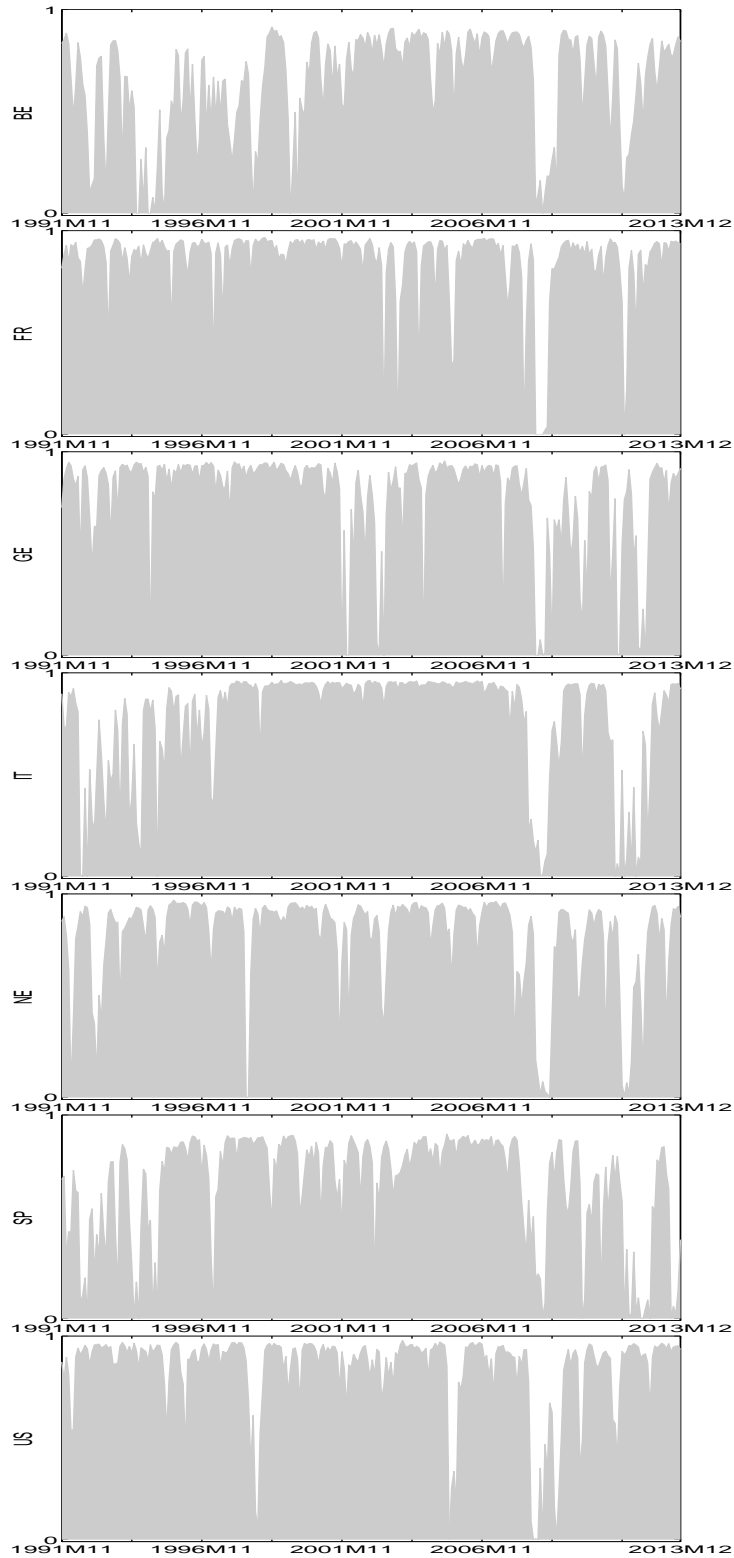


Figure C.3: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

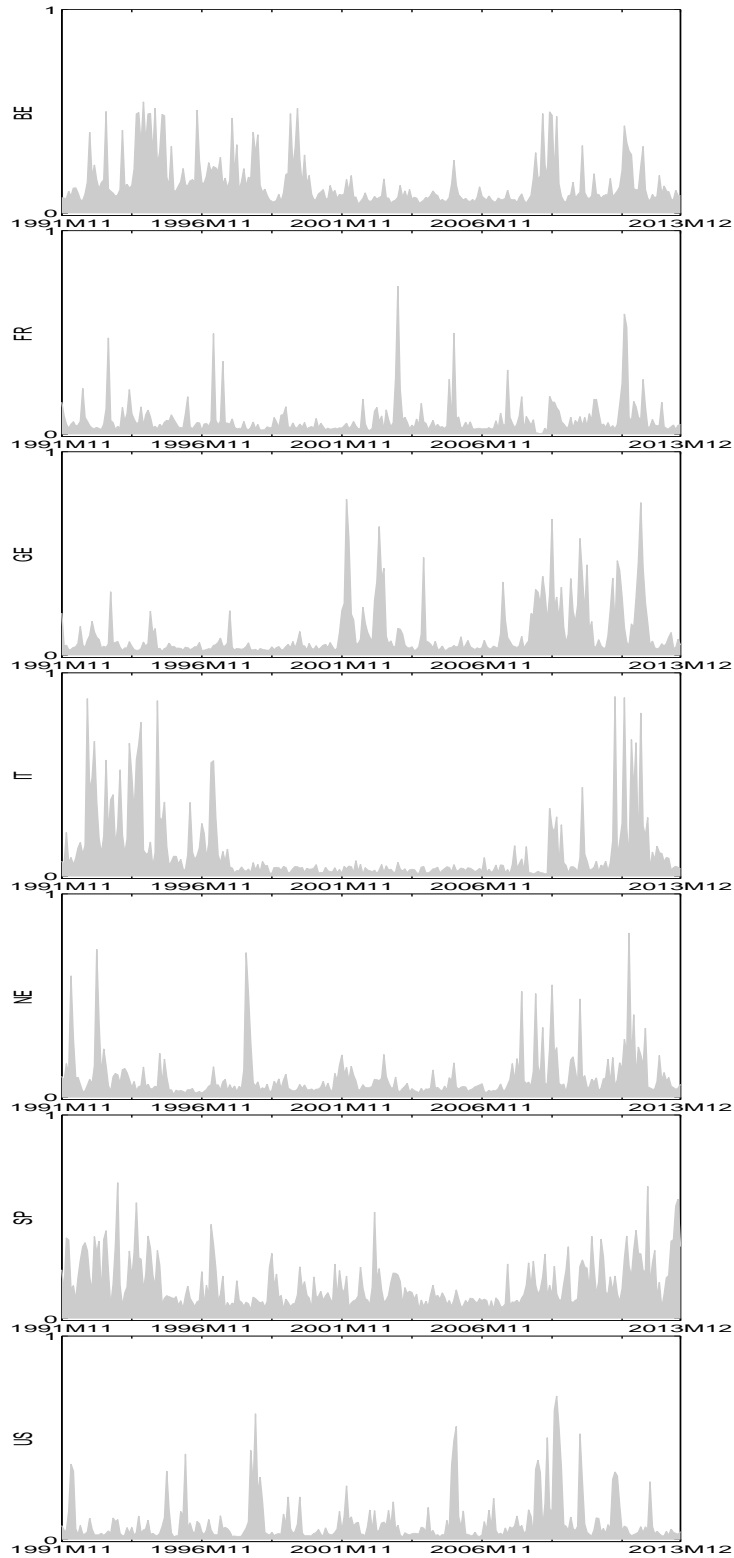


Figure C.4: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

D Cycle dynamic features

This section reports plots of the posterior mean distributions (and 5% and 95% quantiles) for the VAR time-varying intercept and for the VAR time-varying variance in Fig. D.1-D.4, computed as:

$$\widehat{a_{im}(sit)} = \frac{1}{\bar{n}} \sum_{n=1}^{\bar{n}} \sum_{k=1}^K a_{im,k}^{(n)} \xi_{ikt}^{(n)} \quad (\text{D.1})$$

$$\widehat{\Sigma_{im}(sit)} = \frac{1}{\bar{n}} \sum_{n=1}^{\bar{n}} \sum_{k=1}^K \Sigma_{im,k}^{(n)} \xi_{ikt}^{(n)} \quad (\text{D.2})$$

where \bar{n} is the number of MCMC iterations after the burn-in period. The whole set of figures highlights the heterogeneity of the fluctuations in the Eurozone.

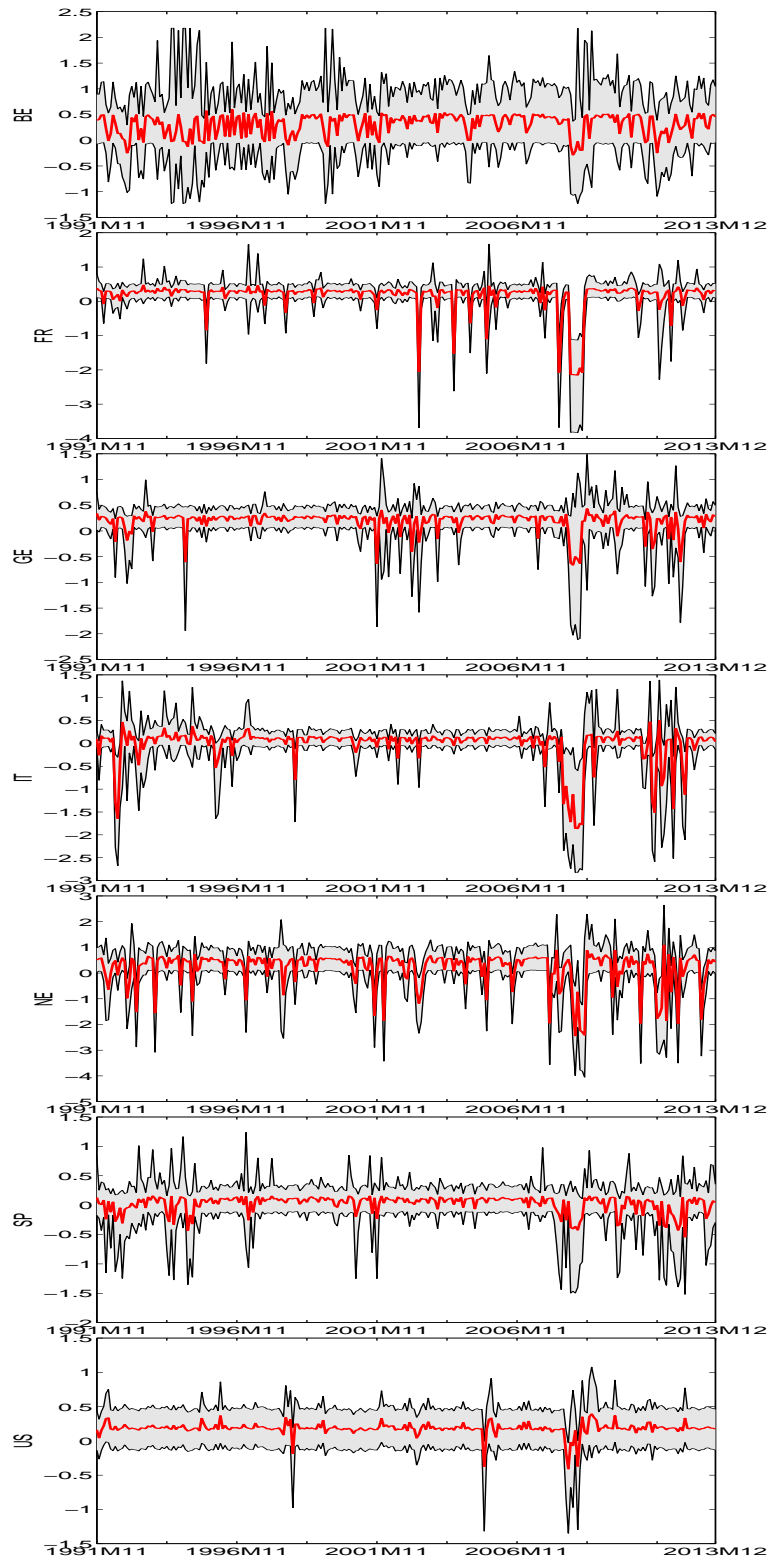


Figure D.1: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying intercept for IPI growth. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

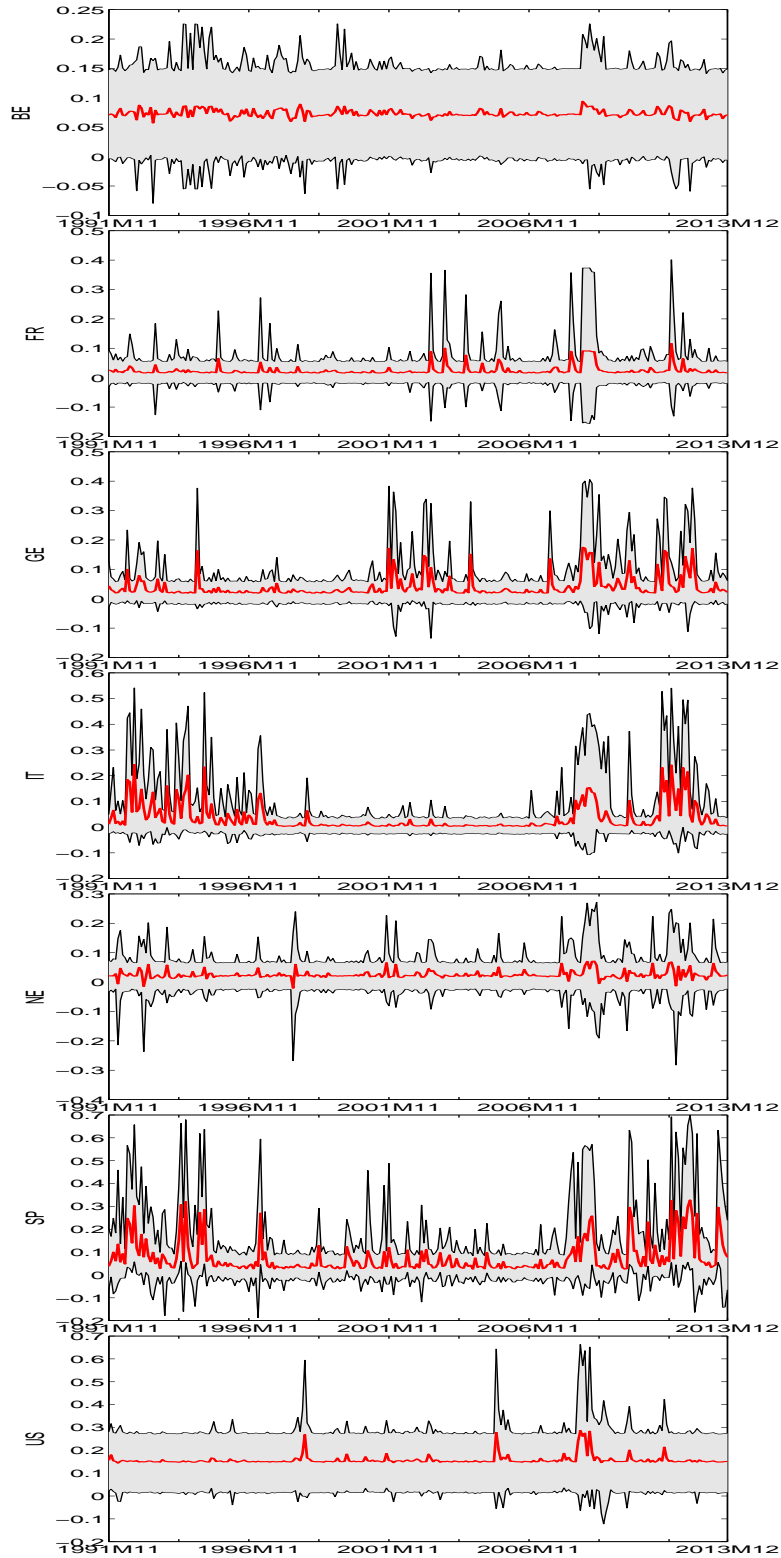


Figure D.2: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying intercept for the credit spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

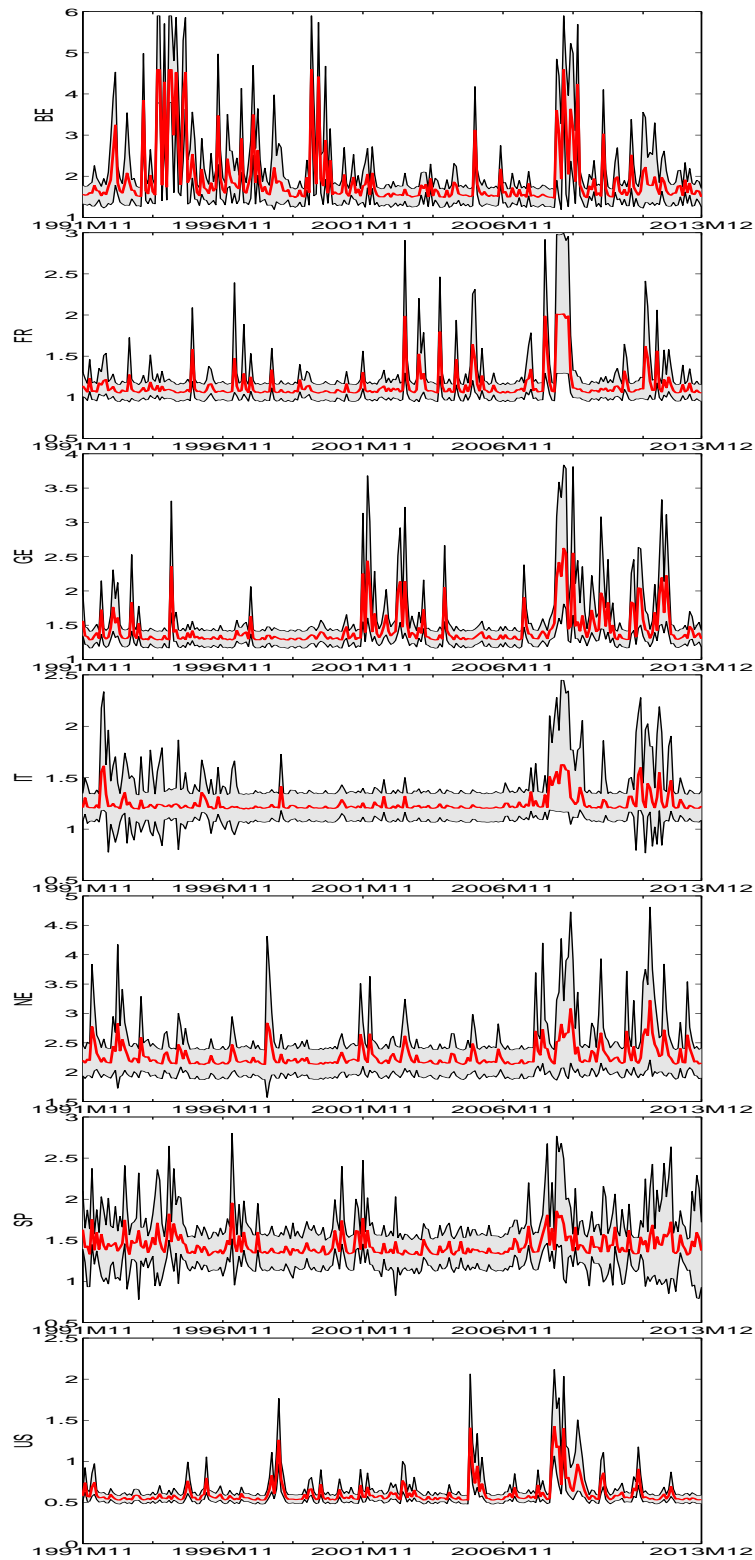


Figure D.3: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying standard deviation IPI growth. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

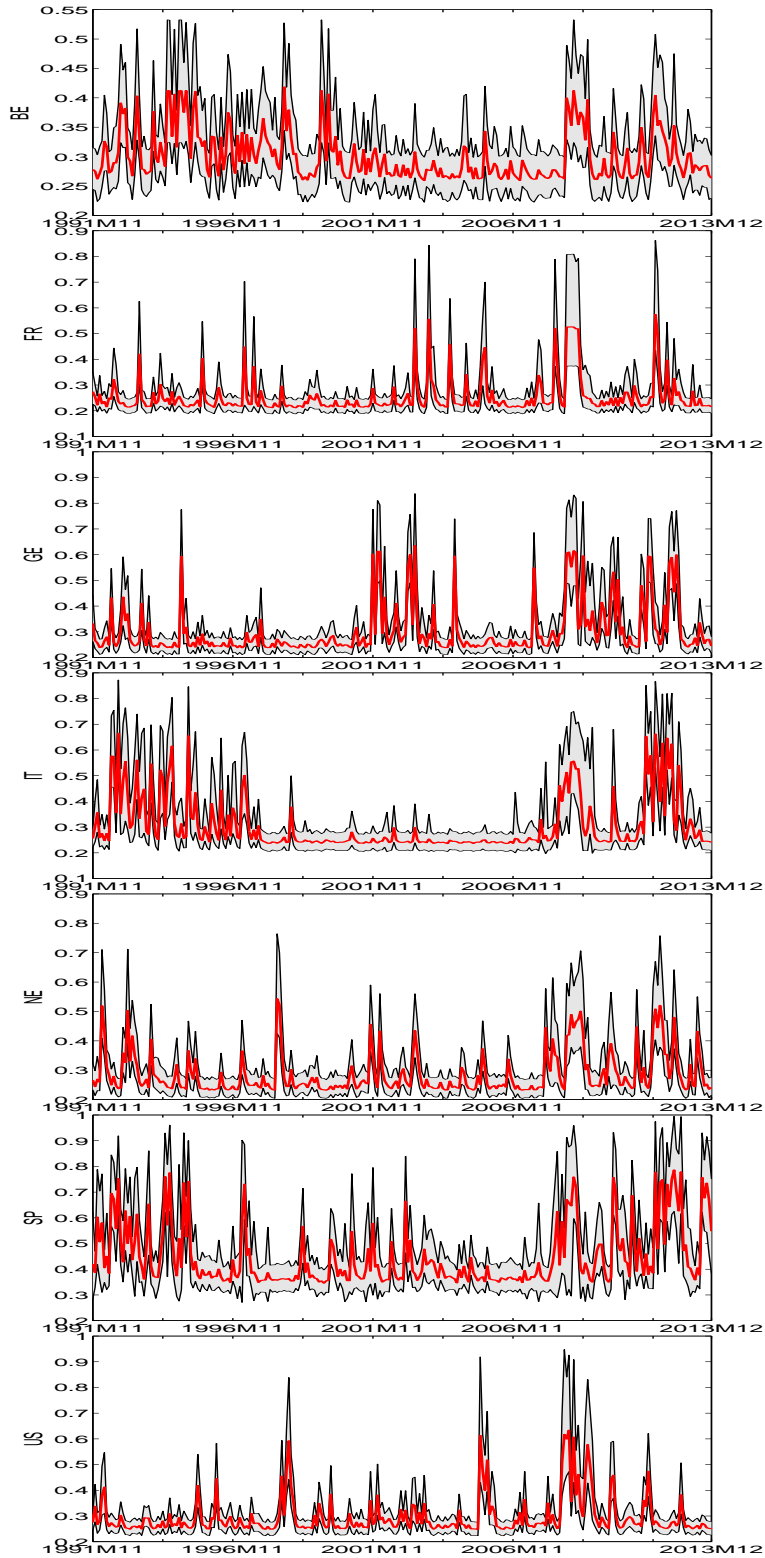


Figure D.4: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying standard deviation for the credit spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

The US in regimes 2 or 3 (moderate or expansion) The US in regime 1 (recession)

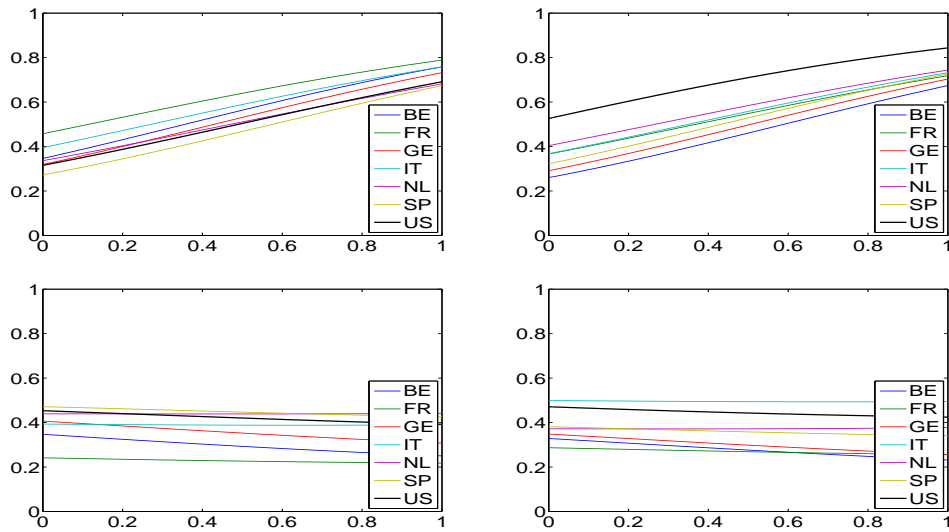


Figure D.5: Reaction of the transition probabilities to stay in recession $p_{it,11}$ (first row) and to exit the recession $p_{it,21}$ (second row) to changes in the weighted aggregate numbers of countries in recession η_{1t} , when conditioning on not recession for the US, i.e. $s_{7,t} \neq 1$ (left column) and recession for the US, i.e. $s_{7,t} = 1$, (right column).

E Smoothed probabilities with different variables

In order to assess the performance of our PMS-VAR model, we also consider two different endogenous variables. We keep industrial production growth and substitute the credit spread with an alternative definition of it as first exercise and the term spread as a second one.

In defining the credit spread, we change the deposit rate from the 10 years German Bund yield to the domestic 10 year bond yield for Belgium, France, Italy, The Netherlands and Spain. We compute then the credit spread as the corporate bond yield spread over the 10 years domestic government interest rate.

The term spread has often been advocated as predictor of recession periods, see e.g. Harvey (1991). Estrella and Hardouvelis (1991) use real GNP growth in US to examine the predictive ability of the term spread. The results show that term spread has significant predictive power on output growth, consumption, and investment. Plosser and Rouwenhorst (1994) find the term structure of interest rates has significant predictive for economic growth in three industrial countries. However, there is no conclusive finding that the term spread consistently contains information in explaining future economic activity. For example, Plosser and Rouwenhorst (1994) find the evidence that term spread contains useful information to forecast real economic activities in US, Canada and Germany, but not in France and UK. Harvey (1991) and Kim and Limpaphayom (1997) examine G7 economies and conclude that the term spread does not consistently contain information about future economic activity. Hamilton and Kim (2002) address the theoretical model toward the nature of the term spread. They nicely present that the term spread forecasting contribution is attributed to two effects: an expectation effect that shows a sign of the public's expectation on the future economic activities and a premium effect that represents the risk of investments in alternative assets. They find that both factors are relevant for predicting real GDP growth but respective contributions differ. The contributions are similar at short horizons but the effect of expected future short rates is larger than the term spread for predicting GDP more than two years ahead.

Figures E.1-E.6 in the Online Appendix show the probabilities for the three regimes when using the the two different endogenous variables. The findings of the previous analysis are qualitatively confirmed. The large differences refer to the final part of the sample, after the Financial Crisis and during the ECB intervention period started in December 2011. The

European Central Bank succeeded in reducing government and corporate yields in most countries, in particular in Italy and Spain, resulting in lower probabilities for regime 1 and higher probabilities for regime 2, in particular when using the alternative definition of credit spread.

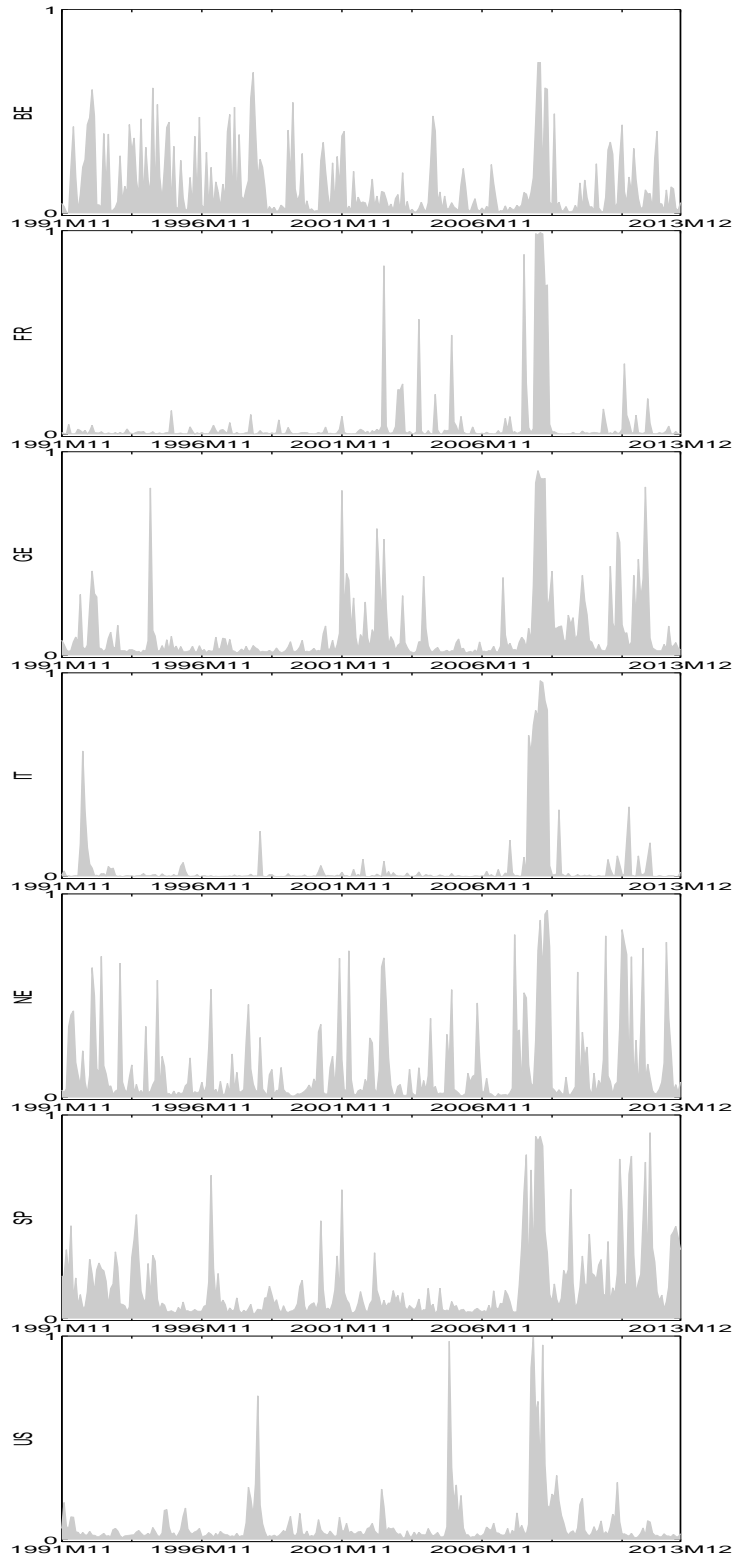


Figure E.1: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using an alternative definition of credit spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

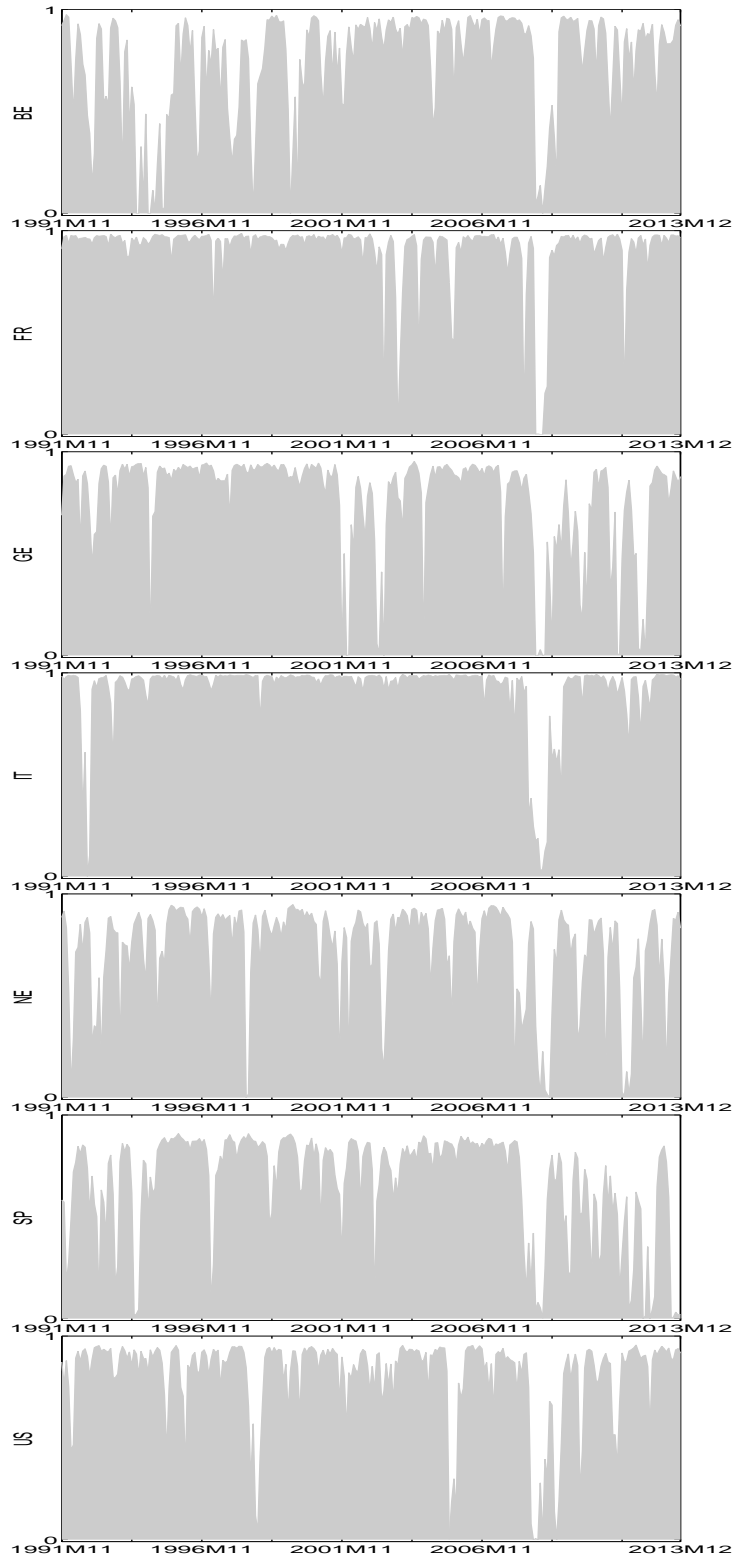


Figure E.2: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using an alternative definition of credit spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

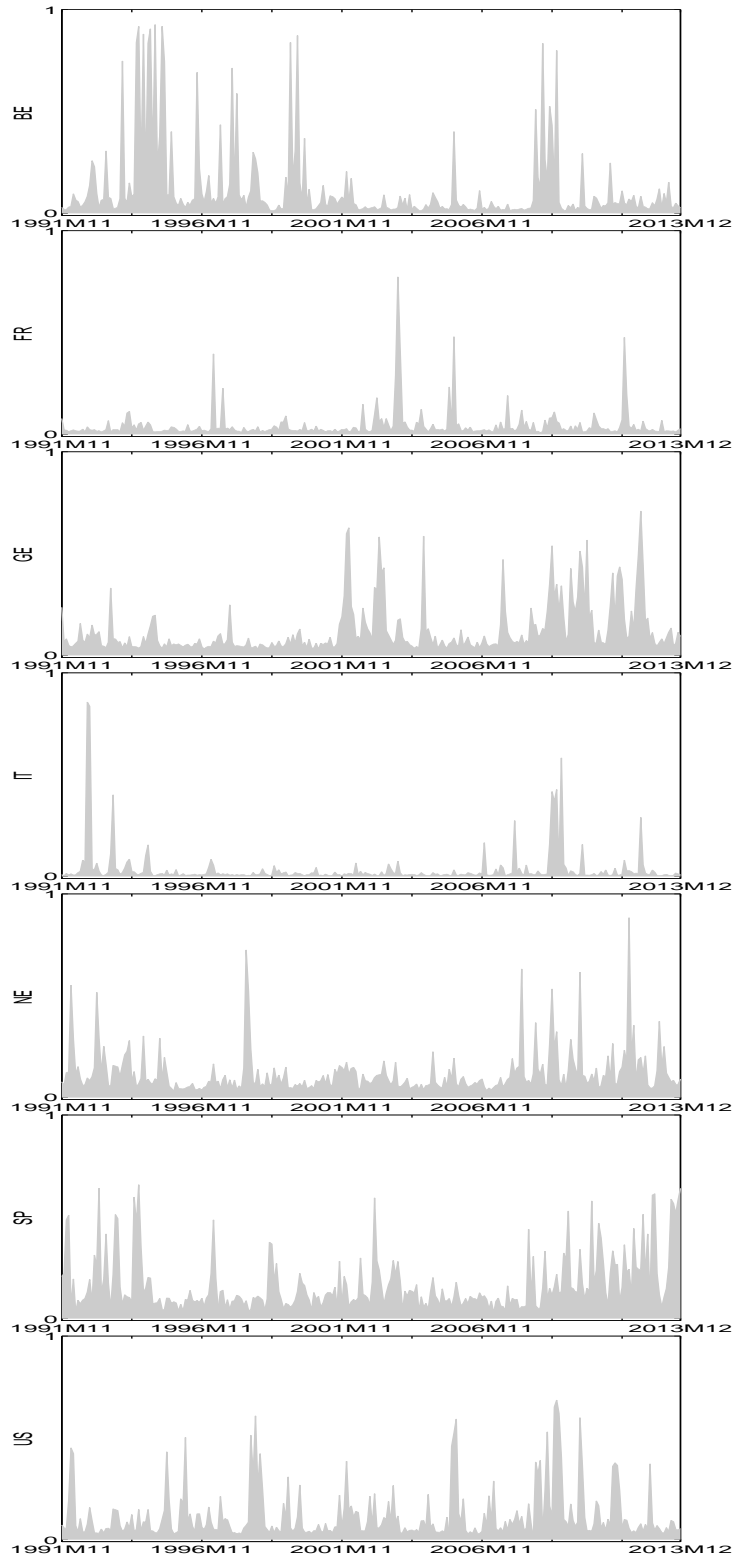


Figure E.3: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using an alternative definition of credit spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

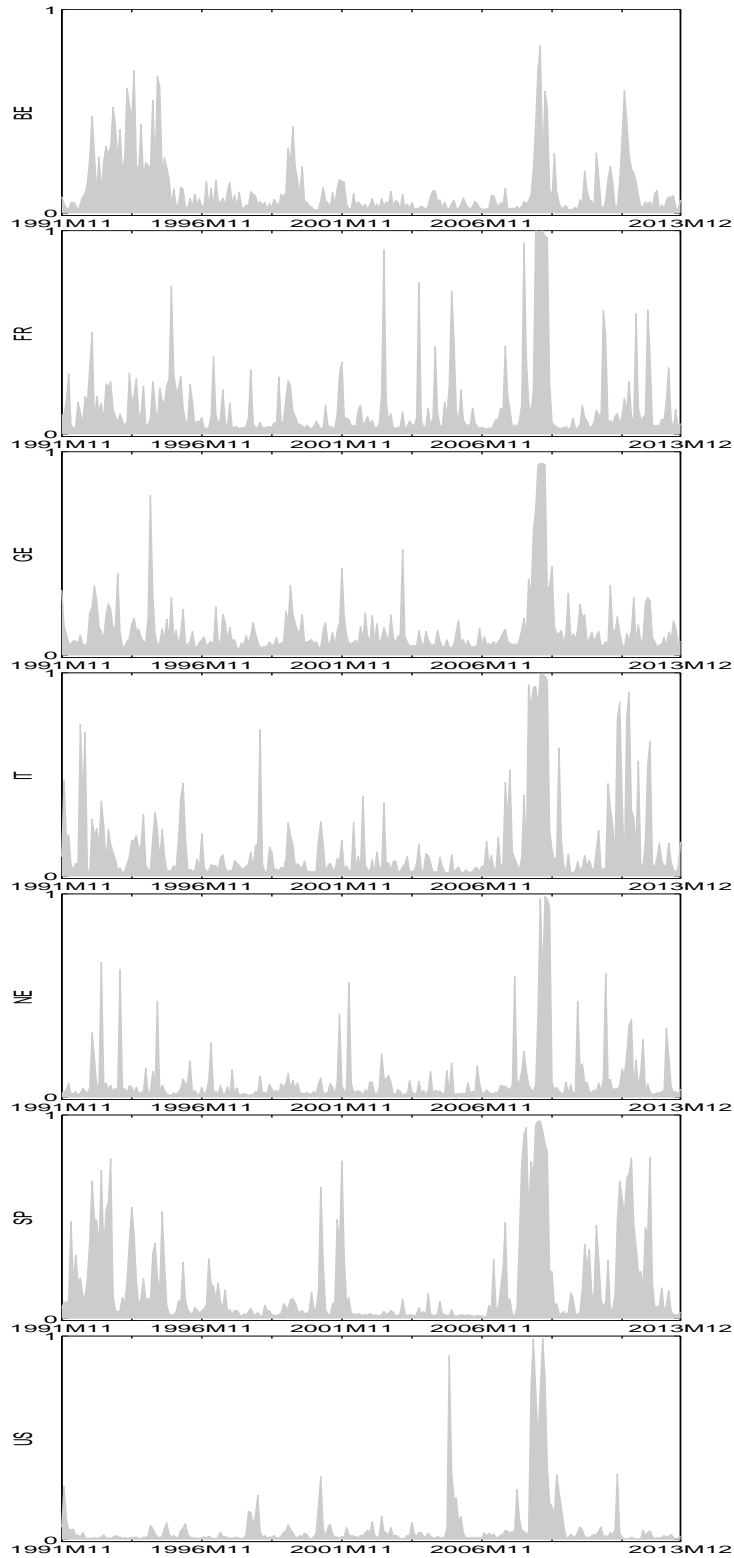


Figure E.4: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using the term spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

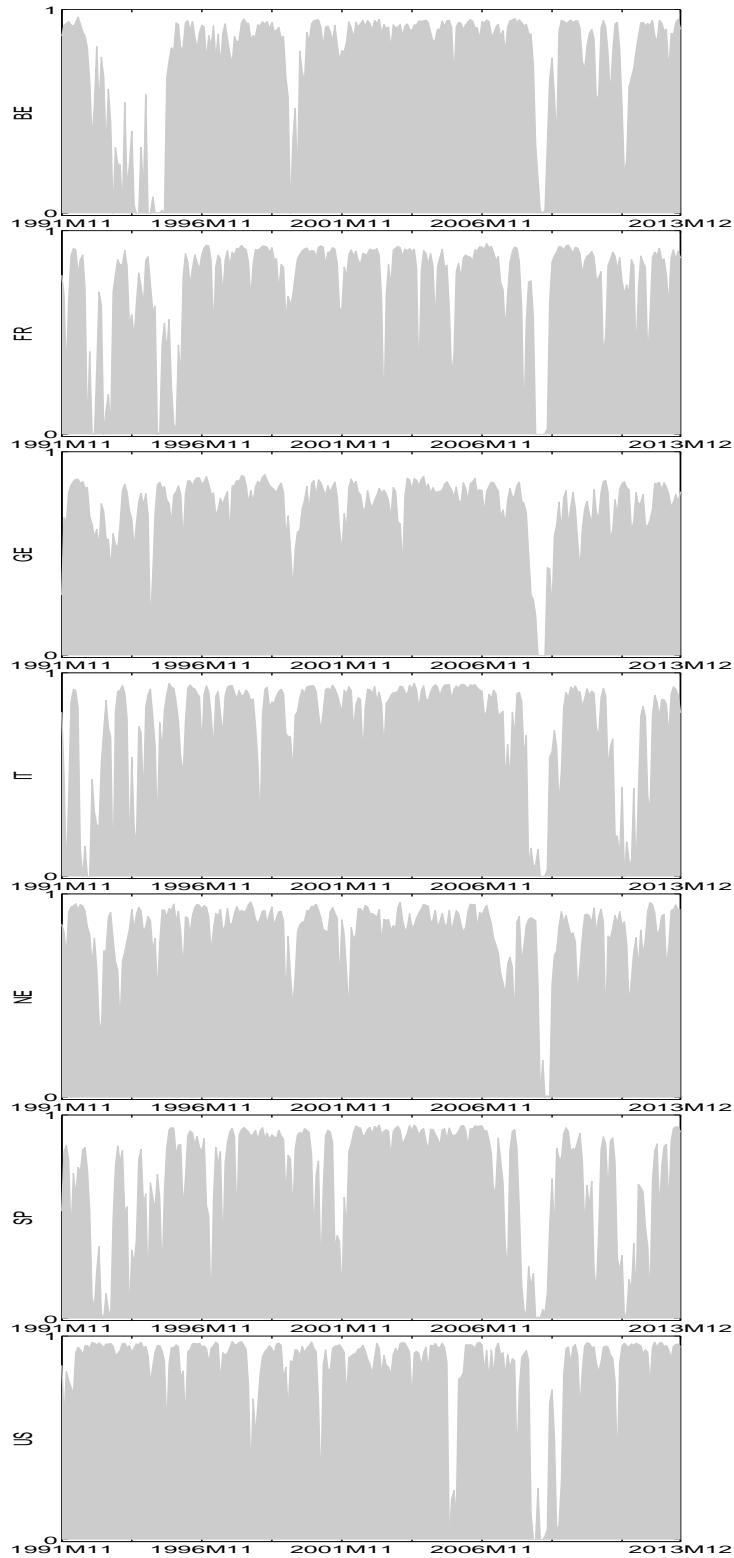


Figure E.5: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using the term spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

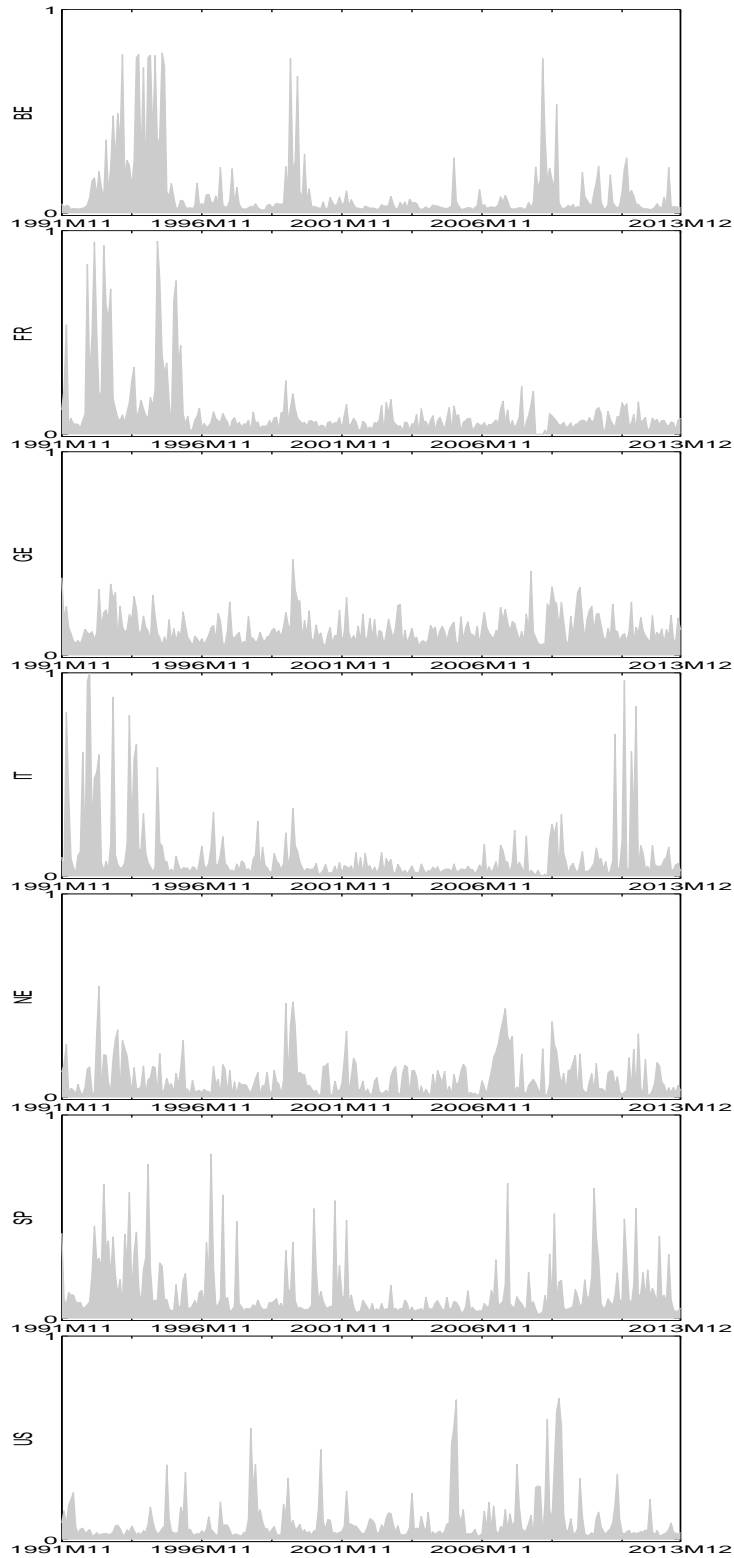


Figure E.6: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \dots, N$ and $t = 1, \dots, T$ using the term spread. The labels “BE”, “FR”, “GE”, “IT”, “NE”, “SP”, “US” indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

References

- Albert, J. H. and Chib, S. (1993). Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *Journal of Business and Economic Statistics*, 11:1–15.
- Anas, J., Billio, M., Ferrara, L., and Mazzi, G. L. (2008). A system for dating and detecting turning points in the Euro area. *The Manchester School*, 76:549–577.
- Canova, F. and Ciccarelli, M. (2004). Forecasting and turning point predictions in a Bayesian panel VAR model. *Journal of Econometrics*, 120(2):327–359.
- Canova, F. and Ciccarelli, M. (2009). Estimating multicountry VAR models. *International Economic Review*, 50(3):929–959.
- Canova, F. and Marrinan, J. (1998). Sources and propagation of international cycles: common shocks or transmission? *Journal of International Economics*, 42:133–167.
- Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.
- Celeux, G. (1998). Bayesian inference for mixture: The label switching problem. Preprint INRIA.
- Chib, S. and Greenberg, E. (1995). Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics*, 68:339–360.
- Claessens, S., Kose, M. A., , and Terrones, M. E. (2009). What happens during recessions, crunches and busts? *Economic Policy*, 24(60):653–700.
- Dees, S., Di Mauro, F., Pesaran, M., and Smith, L. (2007). Exploring the international linkages of the Euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22:1–38.
- Del Negro, M., Giannoni, M., and Schorfheide, F. (2014). Inflation in the great recession and new keynesian models. *American Economic Journal: Macroeconomics*, (forthcoming).
- Estrella, A. and Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, 46(2):555–576.

- Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic factor model: identification and estimation. *The Review of Economics and Statistics*, 82:540–554.
- Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2001). Coincident and leading indicators for the Euro area. *The Economic Journal*, 111:C62–C85.
- 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). *Mixture and Markov-switching Models*. Springer, New York.
- Furlanetto, F., Ravazzolo, F., and Sarferaz, S. (2014). Identification of financial factors in economic fluctuations. Technical Report 9/2014, Norges Bank.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics 4*, pages 169–193. Oxford University Press, Oxford.
- Gilchrist, S. and Mojon, B. (2014). Credit risk in the euro area. Technical report, NBER Working Paper No. 20041.
- Gregory, A., Head, A., and Raynauld, J. (1997). Measuring world business cycles. *International Economic Review*, 38:677–701.
- Hallin, M. and Liska, R. (2008). Dynamic factors in the presence of block structure. Economics Working Papers ECO2008/22, European University Institute.
- Hamilton, J. and Kim, D. (2002). A reexamination of the predictability of economic activity using the yield spread. *Journal of Money Credit and Banking*, 34(2):340–360.
- Harvey, C. (1991). The term structure and world economic growth. *Journal of Fixed Income*, 1:4–17.
- Horvath, L. and Rice, G. (2015). Testing the equality of means when the observations are from functional time series. *Journal of Time Series Analysis*, 36(1):84–108.

- 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.
- Kim, K. A. and Limpaphayom, P. (1997). The effect of economic regimes on the relation between term structure and real activity in japan. *Journal of Economics and Business*, 49(4):379–392.
- Kose, M., Otrok, C., and Whiteman, C. (2003). International business cycles: world, region, and country-specific factors. *American Economic Review*, 93:1216–1239.
- Kose, M. A., Otrok, C., and Prasad, E. (2012). Global business cycles: Convergence or decoupling? *International Economic Review*, 53(2):511–538.
- Krolzig, H.-M. (1997). *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*. Springer, Berlin.
- Lenk, P. J. and DeSarbo, W. S. (2000). Bayesian Inference for Finite Mixtures of Generalized Linear Models with Random Effects. *Psychometrika*, 65:93–119.
- Litterman, R. (1986). Forecasting with Bayesian vector autoregressions-five years of experience. *Journal of Business and Economic Statistics*, 4:25–38.
- Lumsdaine, R. and Prasad, E. (2003). Identifying the common component of international economic fluctuations: a new approach. *Economic Journal*, 113:101–127.
- Monfort, A., Renne, J., R., R., and Vitale, G. (2003). Is economic activity in the G7 synchronized? Common shocks versus spillover effects. CEPR Discussion Paper No. 4119, Centre for Economic Policy Research, London.
- Pesaran, M., Schuermann, T., and Weiner, S. (2004). Modelling regional interdependencies using a global error correcting macroeconometric model. *Journal of Business and Economic Statistics*, 22:129–162.
- Plosser, C. and Rouwenhorst, K. (1994). International term structures and real economic growth. *Journal of Monetary Economics*, 33:133–155.
- Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer Verlag, New York.

- Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika*, 81:115–131.
- Sims, C. and Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39(4):949–968.
- Stock, J. H. and Watson, M. W. (1991). A probability model of the coincident economic indicators. In Lahiri K., M. G., editor, *Leading Economic Indicators*. Cambridge University Press, New York.