tinbergen institute

Interactions between Eurozone and US Booms and Busts:
A Bayesian Panel Markov-switching VAR Model

Monica Billio[1]
Roberto Casarin[1]
Francesco Ravazzolo[2]
Herman K. van Dijk[3]

[1] *University of Venice, GRETA Assoc. and School for Advanced Studies in Venice, Italy;*
[2] *Norges Bank and BI Norwegian Business School, Norway;*
[3] *Erasmus School of Economics, Erasmus University Rotterdam, Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, The Netherlands.*

Interactions between eurozone and US booms and busts:
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

November 1, 2014

Abstract

Interactions between eurozone and United States booms and busts and among major eurozone economies are analyzed by introducing a panel Markov-switching VAR model. The model is well suitable for a multi-country cyclical analysis and accommodates changes in low and high data frequencies and endogenous time-varying transition matrices of the country-specific Markov chains. The transition matrix of each Markov chain depends on its own past history and on the history of other chains, thus allowing for modelling the interactions between cycles. An endogenous common eurozone cycle is derived by aggregating country-specific cycles. The model is estimated using a simulation based Bayesian approach in which an efficient multi-move algorithm is defined to draw time-varying Markov-switching chains. Using real and financial data on industrial production growth and credit spread for all countries, our main empirical results are as follows. Recession, slow recovery and expansion are empirically identified as three regimes with slow recovery becoming persistent in the eurozone in recent years differing from the US. US and eurozone cycles are not fully synchronized over the 1991-2013 period, with evidence of more recessions in the eurozone, in particular during the 90's. Larger synchronization across regions occurs at beginning of the financial crisis but recently more heterogeneity takes place. Cluster analysis yields a group of core countries: Germany, France and Netherlands and a group of peripheral countries Spain and Italy. Reinforcement effects in the recession probabilities and in the probabilities of exiting recessions occur for both eurozone and US with substantial differences in phase

transitions within the eurozone. Finally, credit spreads provide accurate predictive content for business cycle fluctuations. A credit shock results in statistically significant negative industrial production growth for several months in Germany, Spain and US. Our empirical result may serve as important information for the specification of a coordinated policy between the eurozone and the US and within the eurozone.

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." The interconnection between the eurozone and the US economies is not clear. Thus, at the very least, a careful empirical analysis of these issues is necessary. In this paper we investigate interactions 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 similarities and differences in cyclical co-movements, turning points, transmission mechanisms and analysis of shock effects. Our analysis is based on a Bayesian panel Markov-switching model that describes cyclical behavior of the eurozone economy at a country specific level and at an aggregate level and by comparing it with the US economy. Our modeling approach allows also 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. In our empirical application, the shock transmission among countries depends on endogenous aggregate eurozone and US business cycle factors. By comparing such factors and by allowing each country to load on these factors, we can investigate differences among countries business cycles.

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 interactions.

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.

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 this paper, we build on 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. 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 models of Hamilton (1989) and Krolzig (2000) and consider Markov-switching dynamics for low and high frequency components, that is means and covariance matrices of the country-specific equations (see also Billio et al. (2012), Basturk et al. (2013) and Billio et al. (2013b)). We further build on Kaufmann (2011) 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 a country-specific chain may depend not only on its past history but (endogenously) also on the past history of 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.

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. Moreover, to solve potential overfitting problems due to large number of parameters in the model, we follow the hierarchical prior specification strategy proposed by Canova and Ciccarelli (2009). Our paper also relates 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 recovery and expansion, with slow recovery becoming persistent in the eurozone in recent years differing from the US. The first regime is characterized by a negative posterior distribution for the intercept of the industrial production growth. The support of this parameter posterior differs substantially across countries. However, posteriors for the credit spread volatilities are more similar across countries and identify *a posteriori* the low credit risk second regime.

Second, the eurozone and the US cycles appear not fully synchronized, with evidence of more recessions in the eurozone, in particular during the 90's when the monetary union was planned. The larger synchronization is at beginning of the Great Financial Crisis: this shock affects the US first and then spreads very rapidly among economies. As regards the synchronization across eurozone, we identify large heterogeneity, with the global financial crisis ending a period of synchronization and dividing the eurozone in core country and periphery country members.

Third, we find evidence of reinforcement effects in the recession probabilities for the

eurozone when the number of eurozone countries in recession increases. The evidence is different for US where this reinforcement does not exist. The US indicator seems not to have a clear reinforcement mechanism for the recession probabilities of the eurozone countries.

Finally, we document that a credit shock, increasing the credit spreads and therefore deteriorating credit conditions, results in statistically significant negative industrial production growth for several months in Germany, Spain and US.

Our empirical result may serve as important information for the specification of a coordinated economic policy between the eurozone and the US economies and also within the eurozone economies.

The remainder of this paper is organized as follows. Section 2 introduces the Bayesian panel MS-VAR model. Section 3 discusses the prior choice and the Bayesian inference framework. Section 4 presents empirical evidence on such cross-country features before within the eurozone and also between the eurozone and the US economies. Finally, Section 6 concludes.

2 A panel Markov-switching VAR model

In this section, we introduce a general Panel Markov-switching VAR (PMS-VAR) model with endogenous transition and interaction. Moreover, we discuss VAR parameter restrictions needed to avoid overfitting and define the endogenous time-varying transition of the unit specific Markov-chains. We assume that the transitions are dependent on their own past history and on the history of other chains in order to capture the cycle interactions. Alternative interaction mechanisms such as weighting schemes and duration of regimes are also suggested.

2.1 Panel VAR specification

Let $\mathbf{y}_{it} \in \mathbb{R}^M$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$, be a sequence of observations on a K-dimensional vectors of economic variables. N is the number of units (countries) and T the number of time observations. A general specification of the PMS-VAR model reads

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

$i = 1, \ldots, N$, where $\mathcal{N}_M(\boldsymbol{\mu}, \Sigma)$ denotes a M-variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ, and $\mathbf{a}_i(s_{it})$, $A_{ijp}(s_{it})$ and $\Sigma_i(s_{it}))$ are parameters depending on the Markov chain. The $\{s_{it}\}_t$ are unit-specific and independent K-states Markov-chain processes with values in $\{1, \ldots, K\}$ and time-varying transition probability $\mathbb{P}(s_{it} = k | s_{it-1} = l, V_t, \boldsymbol{\alpha}_i^{kl}) = p_{it,kl}$, $k, l \in \{1, \ldots, K\}$, where V_t is a set of G_v common endogenous covariates and $\boldsymbol{\alpha}_i^{kl}$ is a unit-specific vector of parameters.

The generality of this statistical model comes from the possibility that coefficients may vary both across units and across time. Moreover the interdependencies between units are also 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 $k = 1, \ldots, K$, $i = 1, \ldots, N$, and $t = 1, \ldots, T$ and the vector of indicators $\boldsymbol{\xi}_{it} = (\xi_{i1t}, \ldots, \xi_{iKt})'$, which collects the information about the realizations of the i-th unit-specific Markov chain over the sample period. Indicator variables allow us to write parameter shifts 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 $\mathbf{a}_{i,k} = (a_{i1,k}, \ldots, a_{iM,k})' \in \mathbb{R}^M$ are M dimensional column vectors representing the country- and regime-specific VAR intercept, $A_{ijp,k} \in \mathbb{R}^M \times \mathbb{R}^M$ M-dimensional matrices of unit- and regime-specific autoregressive coefficients and $\Sigma_{ik} \in \mathbb{R}^M \times \mathbb{R}^M$ M-dimensional unit- and regime-specific covariance matrices.

The large number of parameters makes our PMS-VAR very flexible. Nevertheless, the overparameterization may lead to an overfitting problem, especially in macroeconomics applications, where time series are characterized by a low number of observations, slowly changing means and time-varying variances (see Basturk et al. (2013)). These issues call for the use of a Bayesian approach to modeling and estimation, since it allows inclusion of parameter restrictions, with different degrees of prior beliefs, through the specification of the prior (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), and thus overfitting problems can be strongly reduced. These restrictions should clearly be motivated by the specific application. In our application on monthly macroeconomic data on the industrial production index growth and on the credit spread we assume Markov-switching in means and variances to model the low and high frequency dynamics and constant autoregressive parameters, constant common variables and block structure for panel in order to avoid overfitting. More specifically, we assume the following restrictions to hold: $\mathbb{E}(\boldsymbol{\varepsilon}_{it} \boldsymbol{\varepsilon}'_{jt}) = O_{M \times M}$ 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. Anyhow, the dependence across units can be modelled through the hierarchical prior specification discussed later on in this paper (see section 3.1).

There are empirical evidences for this type of choice. Clements and Krolzig (1998) find that most forecast errors are due to the constant terms in the prediction models. They also suggest to consider MS models with regime-dependent volatility. In this paper, we

follow Krolzig (2000), Billio et al. (2012) and Basturk et al. (2013) and assume that both unit-specific intercepts, $\mathbf{a}_i(s_{it})$, and volatilities, $\Sigma_i(s_{it})$, are driven by the regime-switching variables $\{s_{it}\}_t$ and assume constant autoregressive coefficients $A_{ip,k} = A_{ip}$, $\forall k$ (see also Anas et al. (2008)). The restricted model considered in the present paper is thus:

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

$i = 1, \ldots, N$. Regarding the switching behaviour, we use an intercept-switching parameterization of the autoregressive model introduced by McCulloch and Tsay (1994). After a regime change, the mean level approaches the new value smoothly over several time periods. We shall notice that an alternative parameterization of the model can be used, in which after a regime change an immediate mean level shift occurs. This parameterization has been used by Hamilton (1989) and has the advantage that parameters can be easily interpreted, but the main drawback is that inference is far more involved than for the McCulloch and Tsay (1994) parameterization. See also Frühwirth-Schnatter (2006), ch. 11.4 and 12.2.

Following Frühwirth-Schnatter (2006), to simplify the exposition of the approximate Bayesian inference, we consider the following re-parameterization based on a partition of the set of regressors $(1, \mathbf{y}'_{it-1}, \cdots, \mathbf{y}'_{it-P})$ into $K+1$ subsets $\bar{\mathbf{x}}_{i0t} = (\mathbf{y}'_{it-1}, \ldots, \mathbf{y}'_{it-P})'$ and $\bar{\mathbf{x}}_{ikt} = 1$, $k = 1, \ldots, K$, that are a 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 writes as

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

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, $\boldsymbol{\gamma}_{i0} \in \mathbb{R}^{MM_0}$, $\boldsymbol{\gamma}_{ik} \in \mathbb{R}^M$, $k = 1, \ldots, K$, $i = 1, \ldots, N$, and $\Sigma_i(\boldsymbol{\xi}_{it}) = \Sigma_i(\boldsymbol{\xi}_{it} \otimes I_M)$ and $\Sigma_i = (\Sigma_{i1}, \cdots, \Sigma_{iK})$. The relationship between the new parameterization and the previous one is: $\boldsymbol{\gamma}_{i0} = (\text{vec}(A_{i1})', \cdots, \text{vec}(A_{iP})')'$, and $\boldsymbol{\gamma}_{ik} = \mathbf{a}_{i,k}$.

2.2 Transition mechanisms

Following Kaufmann (2011) we assume a centered parameterization of the transition probabilities

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

with

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

where $\boldsymbol{\alpha}_i^{kl} = (\alpha_{0i}^{kl}, \boldsymbol{\alpha}_{1i}^{kl\prime})'$ 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 $l = 1, \ldots, K$. To simplify the exposition we also denote with $\boldsymbol{\alpha}_i = \text{vec}\left((\boldsymbol{\alpha}_i^{11}, \ldots, \boldsymbol{\alpha}_i^{KK})\right)$ the collection of parameters of the sequence of transition matrices for the i-th unit.

As regards to the choice of the number M of regimes, we notice that for more recent data one needs an adequate 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 the use of a model with regime-dependent volatility for the US GDP. In our paper we consider data on eurozone industrial production, for a period of time including the 2009 recession and find that three regimes (recession, $k = 1$, slow recovery or moderate expansion, $k = 2$, and expansion, $k = 3$) are necessary to capture some important features of the US and eurozone cycles.

As evidenced in Harding and Pagan (2011) and Harding (2010) the use of simple logit or probit models for modelling the transition probability of the phases of a business cycle may be inappropriate when the goal is to describe the feature of the business cycle. More specifically, minimum phase duration leads to impose restrictions on the parameters of the transition model. Extending the idea of Harding and Pagan (2011) to our panel MS-VAR model and focusing on the minimum recession duration, we specify the following transition probabilities

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

to impose the constraint of a minimum duration of two months for the recession phase.

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 the past values of the unit-specific chains.

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

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

where, in order to have a properly defined vector of probability, we assume $\omega_{it} \geq 0$ and

$\sum_{i=1}^{N} \omega_{it} = 1$, for all t. 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.

We shall notice that the aggregation weights could be included in the inference procedure but leading to a more complex latent variable model, both in terms of modelling and computation. Alternatively, one can use completely unobserved combination weights (e.g., see the modelling strategies in Billio et al. (2013a)) or weights which are partially observed and driven by one or some of the variables mentioned above. Given the high number of latent variables in our model, the latter weight specification strategy should be preferred in order to avoid overfitting problems and to take advantage of all the information available. Also other aggregation rules can be easily included in our framework, to account for prevailing regimes over time and/or in the cross-section, but not being relevant for our application we left them as a topic for future research.

3 Bayesian Inference

The PMS-VAR model is estimated with a simulation based Bayesian procedure. In order to solve potential overfitting problems due to the large number of parameters, we use hierarchical prior distributions. Moreover, we develop an efficient algorithm to draw the latent MS chains, which uses forwarding- filtering backward sampling (e.g., see Frühwirth-Schnatter (2006)) on unit-specific auxiliary bivariate Markov chains. The auxiliary chains allow us to account for both the interaction effects and the 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 us to model dependence between the cross-sectional units through common latent variables and to avoid the potential overfitting problem. For the parameters of the VAR regression we assume

$$\boldsymbol{\gamma}_{i0} \quad \sim \quad \mathcal{N}_{MM_0}(\boldsymbol{\lambda}_0, \underline{\Sigma}_{i0}) \tag{8}$$

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

$$\boldsymbol{\gamma}_{ik} \quad \sim \quad \mathcal{N}_{MM_K}(\boldsymbol{\lambda}_k, \underline{\Sigma}_{ik}), \quad k = 1, \ldots, K \tag{10}$$

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

$i = 1, \ldots, N$. We also assume conditional independence across units, that is: $\mathbb{C}ov(\boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_{j0} | \underline{\boldsymbol{\lambda}}_0) = O_{MM_0 \times MM_0}$ and $\mathbb{C}ov(\boldsymbol{\gamma}_{ik}, \boldsymbol{\gamma}_{jk} | \underline{\boldsymbol{\lambda}}_k) = O_{MM_K \times MM_K}$, for $i \neq j$; and for

the inverse covariance matrix Σ_{ik}^{-1} we assume independent Wishart priors

$$
\begin{aligned}
\Sigma_{ik}^{-1} &\sim \mathcal{W}_M(\underline{\nu}_{ik}/2, \Upsilon_k/2), \quad i = 1, \dots, N && (12) \\
\Upsilon_k^{-1} &\sim \mathcal{W}_M(\underline{\nu}_k/2, \underline{\Upsilon}_k/2), && (13)
\end{aligned}
$$

$k = 1, \dots, K$, that allow us 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 allow us to introduce dependence among units. Moreover, through the specification of the coefficients $\boldsymbol{\gamma}_{ik}$ it is possible to have a regime-specific dependence structure.

When using Markov-switching processes, one should deal with the identification issue associated to the label switching problem. See 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. In the literature, different routes have been proposed for dealing with this problem (see Frühwirth-Schnatter (2006) for a review). One of the most 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 the different phases (e.g. recession and expansion) of the business cycle. We follow this latter approach and include the constraints

$$
\gamma_{ij1} < \gamma_{ij2} < \dots < \gamma_{ijK}
$$

$j = 1, \dots, M$ and $i = 1, \dots, N$, that corresponds to a total ordering, across the different regimes, of the constant terms in the different equations of the system.

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

$$
\begin{aligned}
\boldsymbol{\alpha}_i^{kl} &\sim \mathcal{N}_{G_v+1}(\boldsymbol{\psi}, \Upsilon_i) \quad i = 1, \dots, N, l = 1, \dots, K-1 && (14) \\
\boldsymbol{\psi} &\sim \mathcal{N}_{G_v+1}(\underline{\boldsymbol{\psi}}, \underline{\Upsilon}) && (15)
\end{aligned}
$$

In particular in the empirical application, we consider the following hyper-parameter specification: $\underline{\boldsymbol{\lambda}}_0 = \mathbf{0}$, $\underline{\Sigma}_{i0} = I_{MM_0}$, $\underline{\Sigma}_0 = 10 I_{MM_0}$, $\underline{\boldsymbol{\lambda}}_k = \mathbf{0}$, $\underline{\Sigma}_{ik} = I_{MM_K}$, $\underline{\Sigma}_k = 10 I_{MM_K}$, $k = 1, \dots, K$, $\underline{\nu}_{ik} = 5$, $\underline{\nu}_k = 5$, $\underline{\Upsilon}_k = 10 I_M$, $\underline{\boldsymbol{\psi}} = \mathbf{0}$, $\Upsilon_i = I_{G_v+1}$, $\underline{\Upsilon} = 10 I_{G_v+1}$ 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 can be iterated over different blocks of unit-specific parameters avoiding the computational difficulties associated with the inversions of large covariance matrices (see Canova and Ciccarelli (2009)). We derive the 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 $\mathbf{y}_i = \text{vec}((\mathbf{y}_{i1}, \ldots, \mathbf{y}_{iT}))$ be the set of observations collected over time, $\mathbf{y} = \text{vec}((\mathbf{y}_1, \ldots, \mathbf{y}_N)')$ the set of observations collected over time and panel units and $\boldsymbol{\xi} = \text{vec}((\Xi_1, \ldots, \Xi_N))$ the set of allocation variables, with $\Xi_i = (\boldsymbol{\xi}_{i1}, \ldots, \boldsymbol{\xi}_{iT})$. We define the vector of regression coefficients, $\boldsymbol{\gamma} = \text{vec}((\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_N))$ where $\boldsymbol{\gamma}_i = \text{vec}((\boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_{i1}, \ldots, \boldsymbol{\gamma}_{iK}))$, the set of covariance matrices, $\Sigma = (\Sigma_1, \ldots, \Sigma_N)$, and the transition probability parameter vector, $\boldsymbol{\alpha} = \text{vec}((\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N))$.

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

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

where

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

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

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

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

    (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 $k = 1, \ldots, K$;

(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) $\boldsymbol{\alpha}_i^{k1}, \ldots, \boldsymbol{\alpha}_i^{kK-1}$ from $f(\boldsymbol{\alpha}_i^{k1}, \ldots, \boldsymbol{\alpha}_i^{kK-1}|\mathbf{y}_i, \Xi, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_i)$.

Note that 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 from 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 $k = 1, \ldots, K$:

(iv.a) $\boldsymbol{\lambda}_k$ from $f(\boldsymbol{\lambda}_k|\boldsymbol{\gamma}, \Sigma)$;

(iv.b) Υ_k^{-1} from $f(\Upsilon_k^{-1}|\boldsymbol{\gamma}, \Sigma)$;

(vi) 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 the proposed MCMC algorithm are given in the Appendix A.

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 parametrized 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. We show that the full conditional distribution of the unit-specific chains has a representation in terms of the augmented likelihood. At time t, the augmented likelihood is the product of the likelihood of the observations at time t and a term containing the value at time $t-1$ of all the chains of the panel. The model is thus Markovian of the second order in the hidden state variables and the multivariate chain representation of the hidden state process can be exploited. This representations allows us to apply a FFBS for exact sampling of the unit-specific chains and, also to impose more efficiently the minimum phase duration restriction discussed in the previous sections. Further details on the FFBS procedure are given in the Appendix A.

4 On eurozone and US booms and busts

4.1 Data description

The empirical focus of the paper deals with whether eurozone and US economies differ in periods of booms and busts. We consider the eurozone 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 wants to contribute to this debate and to provides new evidence.

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: the Industrial Production Index (IPI), labelled as $y_{i1,t}$ and the credit spread (CS), i.e. the corporate bond yield spread over the 10 years government interest rate, $y_{i2,t}$. The IPI is one of the main economic indicators that measures changes in output for the manufacturing, mining, and utilities business sectors. Although these sectors contribute only to a fraction of the GDP, and several countries have partially shifted from being production oriented to being service and consumer oriented, which reduces the contribution of these sectors, they are rather sensitive to variations in interest rates and consumer demand. This makes the IPI an important variable for forecasting the future economic performance of an economic system. We download IPI data from the OECD database. Moreover, there is a large stream of literature which is using MS-VAR to extract the cycle from a set of variables, see, e.g., Krolzig (2004). 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, Claessens et al. (2009) link shock transmissions from the financial sector to the real one using a larger 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 the credit spread could predict the US 2008 recession. Del Negro et al. (2014) define the credit spread as the rate entrepreneurs pay in excess of deposit yield to finance their projects and measure it as the differences between the Baa corporate bond yield minus the 10-Year Treasury Note Yield. We use the same variable for US. For European countries, the construction of the variable is more difficult. First, at European country-level corporate bond indices are not easily available. We collect corporate bond yields from Global Financial Data (https://www.globalfinancialdata.com/index.html) and they are combination of many underlying securities. Indices are based upon long-term (10-30 year) bonds of investment grade (AAA to BBB), with average rate A or Baa. Second, the definition of deposit rate in the various European countries is problematic. We follow Gilchrist and Mojon (2014) and compute the spread over the 10 years German Bund yield for all the six eurozone countries. Government Treasure yields are downloaded from Datastream. Unfortunately, we are obliged to ignore exchange rate risk before the euro was introduced as an accounting currency on 1 January 1999.
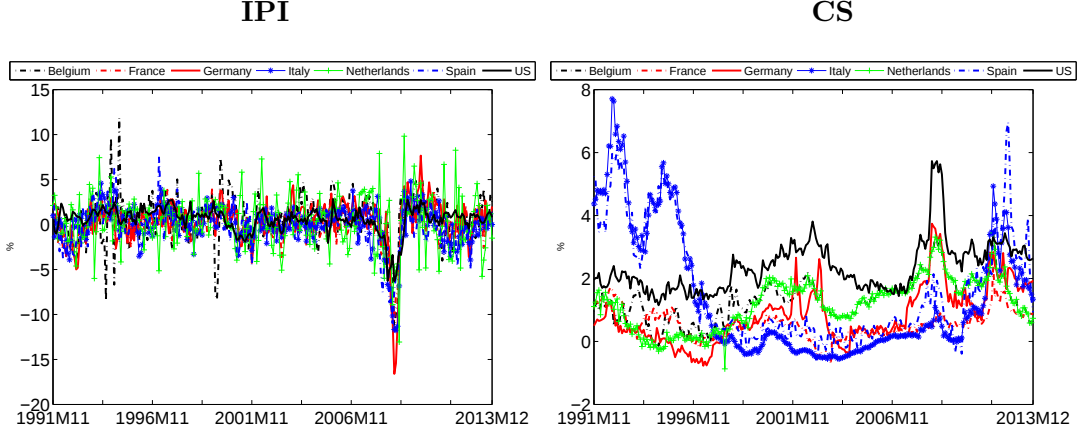
Figure 1: Country-specific endogenous variables: industrial production growth rate (IPI) and credit spread (CS).

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.

Finally, one crucial assumption for our model relates to the composition of the variable V_t. To investigate the interconnectedness between eurozone and US, we specify the set of common endogenous covariates V_t equal to the 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 value 1 when the US economy is in recession and 0 otherwise. Such assumptions allow us to have an endogenous interaction mechanism between the two economies and force spillovers enter nonlinearly. Note that the information synthesis of the euro countries is discussed in Section 2.3. More precisely, we focus on the weighted interaction indicator given in equation (7) and use economic size unit-specific weights. We follow the Eurostat framework for eurozone variables aggregation and derive weights on relative value added, see Eurostat Regulation EC No 1165/98. Value added data are downloaded from the UNData database and Figure 2 displays the weights. The value added data are annual and we transform them to monthly frequency by using the same values for the 12 months in each calendar year.

4.2 Evidence on business cycle regime classification

To avoid issues with possibly non-stationary series, we take the IPI in log-changes. We consider two possible number 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 $p(\mathbf{y}|K = 3) = \prod_{t=1}^{T-1} \prod_{i=1}^{N} p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 3)$ with $p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 3)$ the 1-step ahead predictive density for $\mathbf{y}_{it+1}$ conditional on information up to time t and $K = 3$

14
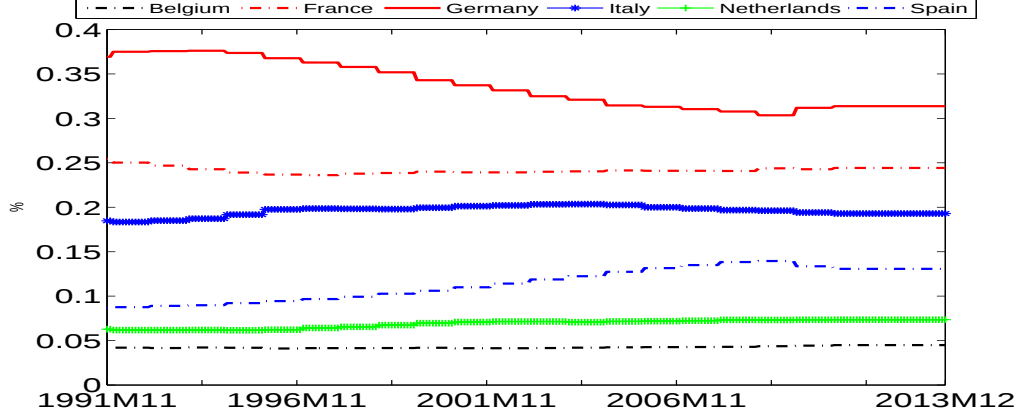
Figure 2: Value added eurozone weights.

regimes; $p(\mathbf{y}|K = 2) = \prod_{t=1}^{T-1} \prod_{i=1}^{N} p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 2)$ with $p(\mathbf{y}_{it+1}|\mathbf{y}_{it}, K = 2)$ the 1-step ahead predictive density for $\mathbf{y}_{it+1}$ conditional on information up to time t and $K = 2$ regimes. We find that the BF is larger than one, therefore supporting 3-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 Bayes factors. Finally, we impose the following restrictions on the intercept of the IPI growth rate $a_{i1,1} < 0$ and $a_{i1,1} < a_{i1,2} < a_{i1,3}$, $i = 1, \ldots, N$, in order to identify the regimes (see Section 3.1). We label regime 1 as recession; regime 2 as slow recovery or moderate expansion; and regime 3 as expansion.

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.

Figures 3 and 4 show the approximated posterior densities of the parameters $\boldsymbol{\gamma}_{ik} = (a_{i1,k}, a_{i2,k})'$, $(\sigma_{i1,k})$ and $(\sigma_{i2,k})$, $k = 1, \ldots, K$ and $i = 1, \ldots, N$, that represent the value of the unit- and variable-specific time-varying intercepts and volatilities of the PMS-VAR model. A comparison of such posteriors provides useful information on whether and how individual countries differ over booms and busts. We recall that the regime identification follows from the parameter constraints $a_{i1,1} < 0$ and $a_{i1,1} < a_{i1,2} < a_{i1,3}$, on the intercept of the IPI growth rate.

15
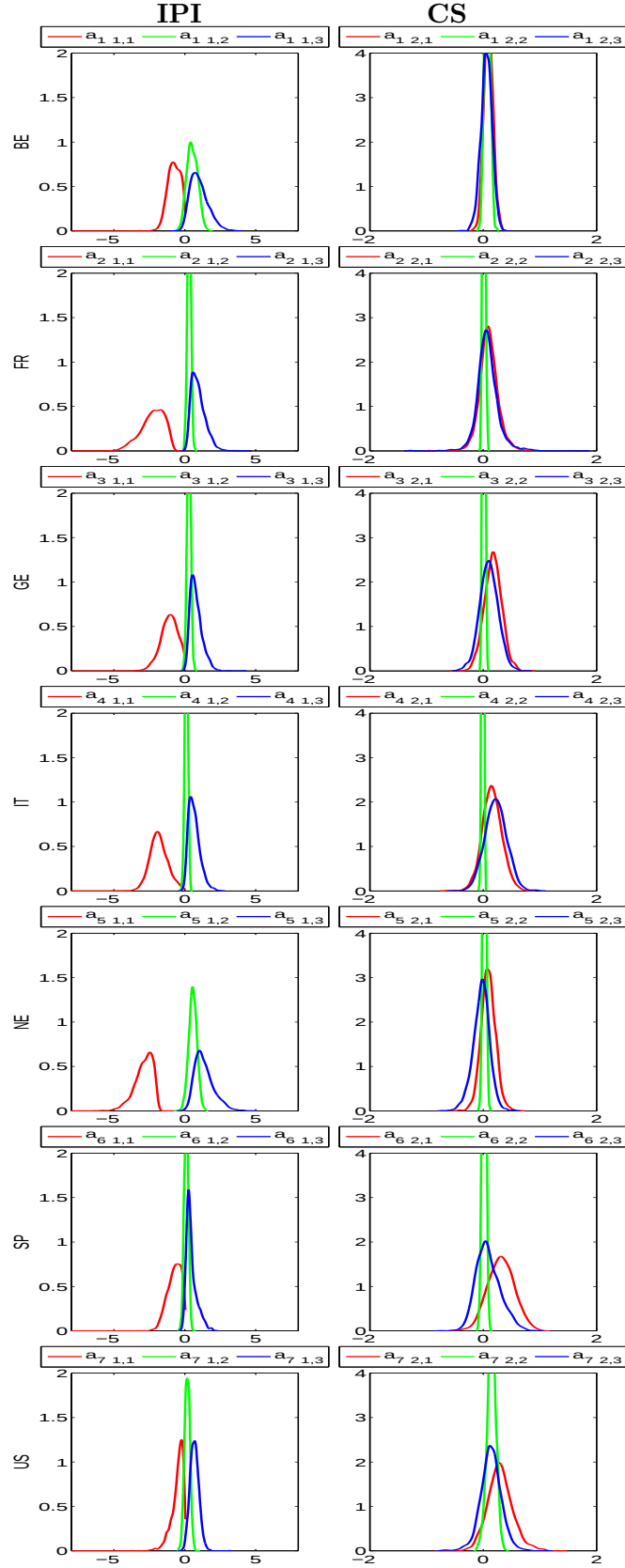
Figure 3: The figures show the kernel of the posterior densities of the Markov-switching intercepts, $\gamma_{ik} = (a_{i1,k}, a_{i2,k})'$, for the different $i = 1, \ldots, N$ countries and $k = 1, \ldots, 3$ regimes (in red regime the first one, in green regime the second one and in blue regime the third one) 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.
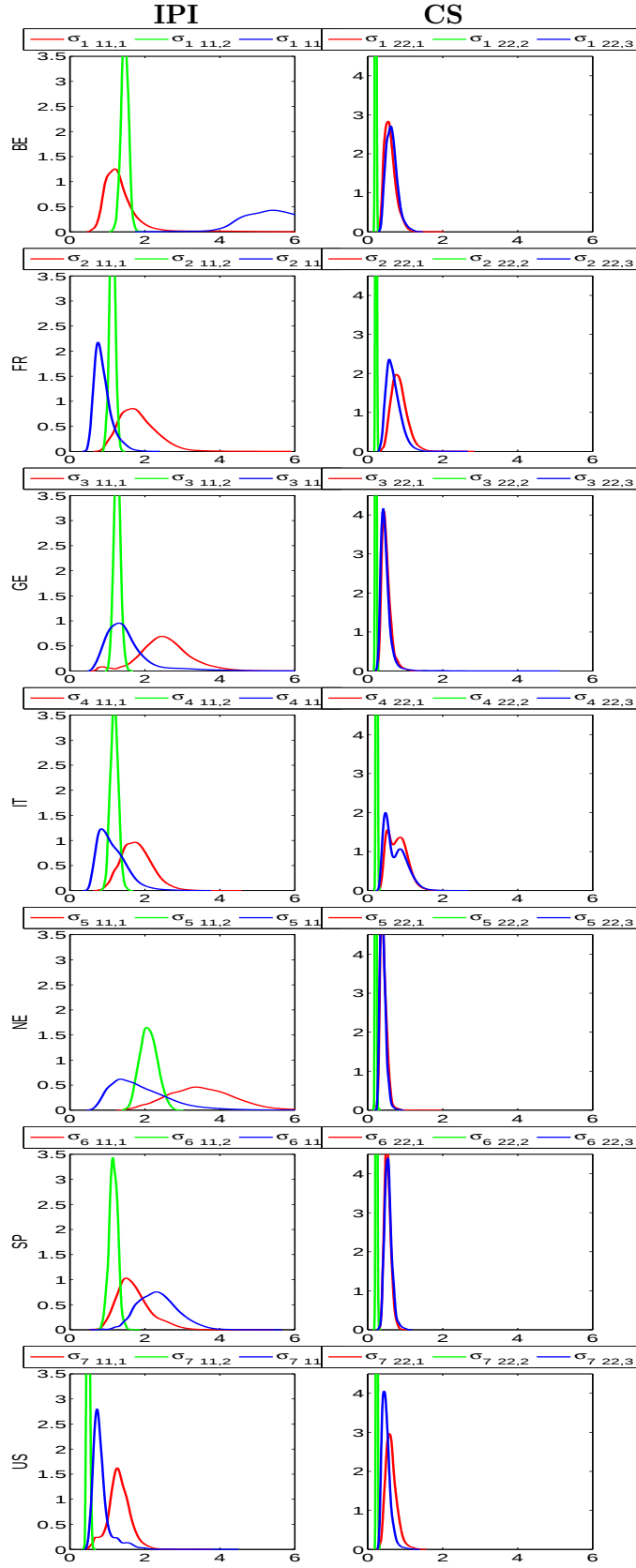
Figure 4: The figures show the kernel of the posterior densities of the Markov-switching volatilities, $\sqrt{\sigma_{ijj,k}}$, for the different $i = 1, \ldots, N$ countries and $k = 1, \ldots, 3$ regimes (in red regime the first one, in green regime the second one and in blue regime the third one) 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.

The posterior densities for the IPI growth intercept in regime 1, $a_{i1,1}$ are not overlapping with posterior densities for the IPI growth intercept in regimes 2 and 3, $a_{i1,m}$, $m = 1, 2, 3$ (see left column in Figure 3), in most of the countries. This suggests that the recession regime is well identified on the IPI growth data. Moreover, for all panel units the support of the posterior density for $a_{i1,1}$, the intercept of the recession regime, is negative as we impose; while $a_{i1,2}$, the moderate regime intercept, is centered around zero; and $a_{i1,3}$, the expansion intercept, is positive. Nevertheless, there are substantial differences between European countries and US: the posteriors are in most cases wider for European countries; and the posteriors of $a_{i1,1}$ are large and negative. Posteriors for US are more concentrated and closer to zero. The posteriors for $a_{i1,2}$ and $a_{i1,3}$ overlap substantially, in particular for the Belgium case, indicating that strong expansion periods cannot be easily identified, at least by just looking to IPI intercepts, from slow recovery and moderate growth in our sample.

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

The differences across regimes and across countries are larger for the posterior densities of the residual volatilities, in particular for the credit spread (see Figure 4). As regards the IPI volatility, there is a large difference of the volatility behavior across regimes between US and European countries. The general pattern is that volatility is higher during recessions and, for many countries, during expansion periods, and lower and more concentrated in recovery and moderate expansion periods, but with important differences across countries. For US, the order is clear with volatility increasing in regime 3 versus regime 2 and in regime 1 versus regime 3. This evidence is less clear, with, e.g., mean volatility in regime 3 in Italy lower than the mean of the volatilities in the other two regimes. The US industrial production has larger switches during recession or expansion periods, which increase volatility estimates. Posterior mean estimates suggest such movements are transitory and do not imply large changes in the intercept. The eurozone estimates seem to be dominated by more switches across regimes, both in the intercept and the volatility. In the next section, we document that the differences are larger at the beginning of the sample, when European countries experienced turbulent period in the early 1990's related to exchange rate crisis, and at the end of the sample, when Europe has experienced the sovereign debt crisis.

The main differences across regimes are for residual volatilities of the credit spread. For all the seven economies, the posterior $\sigma_{i,22,2}$ is more concentrated and closer to zero than the volatility in the other two regimes. Its support set does not overlap with the ones on the recession and expansion regimes, indicating that regime 2 is well identified and supported by data as the Bayes factor analysis also evidences. Posterior volatilities are similar across countries, indicating a clear behavior of such variable over the business cycle.

To sum up, we find some important differences in the parameter posterior densities of

the eurozone and US, both in the intercept and in the regime volatility of the industrial production. The heterogeneity is also important among eurozone economies. Posteriors for the credit spread are more similar across countries and identify the low credit risk regime from the more volatile first and third regimes. This confirms evidence in Gilchrist and Zakrajsek (2012) and Gilchrist and Mojon (2014) that credit spreads provide substantial predictive content for a variety of real activity and lending measures across different countries.

4.3 Synchronization of eurozone and US cycles

The PMS-VAR model allows us to study the business cycles fluctuations of each country in the panel and to analyse the transmission of shocks across cycles. We recall that the regime labeling is: recession, $s_{i,t} = 1$, recovery or moderate expansion, $s_{i,t} = 2$, and expansion, $s_{i,t} = 3$. The PMS-VAR model produces both country-specific smoothed probabilities for each regime and eurozone and US aggregate smoothed probabilities. Specifically, the number of euro countries in recession and the similar measure for the US, used in the vector V_t, are reported in the Figure 5 (Figure C.1 in the Online Appendix reports the associated probabilities of eurozone and US economies to be in recessions). The Figure provides several interesting results and generally shows that eurozone and US economies are not fully aligned.
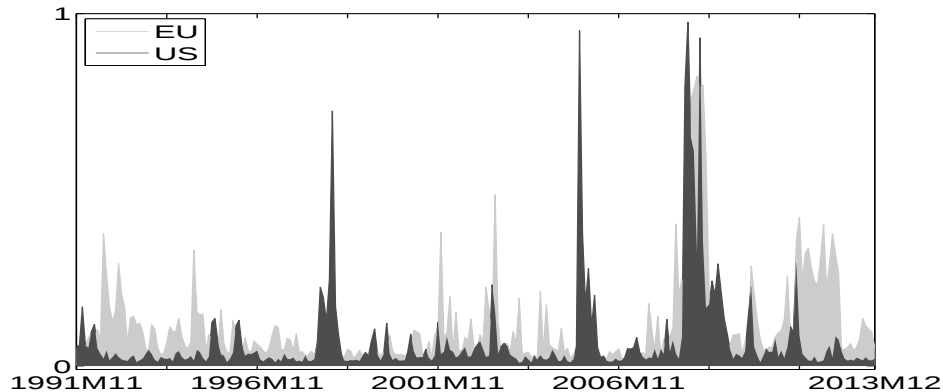

Figure 5: The light grey line shows the fraction of eurozone countries in the recession regime standardized between 0 and 1 and and the black line shows the US transition probability for regime one, $s_{7,t}$, $t = 1, \ldots, T$
.

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 is at beginning of 1999 with the internet bubble in US. In the second decade, US apparently lead the eurozone cycle, especially during the Financial Crisis in 2007-2008. The internet bubble has generated small and short-lasting recessions in both economies, with instabilities up to 2003, and some calls for new recession in the US at the end of 2005 and in 2006. 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, where the eurozone recession starts in September 2008 . Both economies enter in the second quarter 2009 in a new regime generally defined slow recovery in our paper (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), but which is probably more accurate to interpret as stagnation. Furthermore, the eurozone has evidence of a new recession regime from the third quarter in 2011. The recession can be associated to sovereign debt problems for some European countries, in particular Italy and Spain. The role of the credit spread is quite important in detecting the recession because credit conditions deteriorate 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. In general, recessions in the US are shorter than in the eurozone.

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 the most probable as we could anticipate since its definition can fit both stagnation, recovery and (moderate) expansion periods, which are appropriate definitions for most of our sample. 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 of the eurozone is evident not only in regime dynamics but also in the features of the regimes. The dynamic features of the cycle, in terms of posterior distributions 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 section. The Financial Crisis is evident with regime 1 dominant in all the four parameters. The level of IPI growth is much more negative in Europe than US during the crisis. France, Germany, Italy and The Netherlands have large part of the posterior below -1.5, compared to the 90% interval $[0, -1]$ of the US. The difference is even larger during the European sovereign debt crisis. 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 one the smallest. Volatilities of the credit spreads across regions are, on the contrary, more comparable.

4.4 Heterogeneity of country cycles

To further investigate how countries relate to the aggregate and possible synchronize with it, we study how each member country cycle detects turning points of 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 is used to measure the relationships across countries.

As first analysis, we follow Billio et al. (2012) and date th 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 $\Delta_\kappa y_t > 0, \ldots, \Delta_1 y_t > 0$ and $\Delta_1 y_{t+1} < 0, \ldots, \Delta_\kappa y_{t+\kappa} < 0$ and a upward turn (or trough) at time t if $\Delta_\kappa y_t < 0, \ldots, \Delta_1 y_t < 0$ and $\Delta_1 y_{t+1} > 0, \ldots, \Delta_\kappa y_{t+\kappa} > 0$, where Δ_κ denotes the κ-difference operator (see Harding and Pagan (2011)). The parameter κ reduces 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 US.

In the following we apply an approximation of the BB rule and use only downward, $D_t(\kappa)$, and upward, $U_t(\kappa)$, turn signals, that are (see Harding and Pagan (2011))

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

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

respectively. Our analysis can be extended to include modifications of the BB rule (see for example Mönch and Uhlig (2005)), who account for asymmetries and time-varying duration across business cycle phases. Censoring rules preventing the algorithm from the detection of false signals could 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 turning point detection for the different country Markov chains by the concordance statistics (CS):

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

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 given by
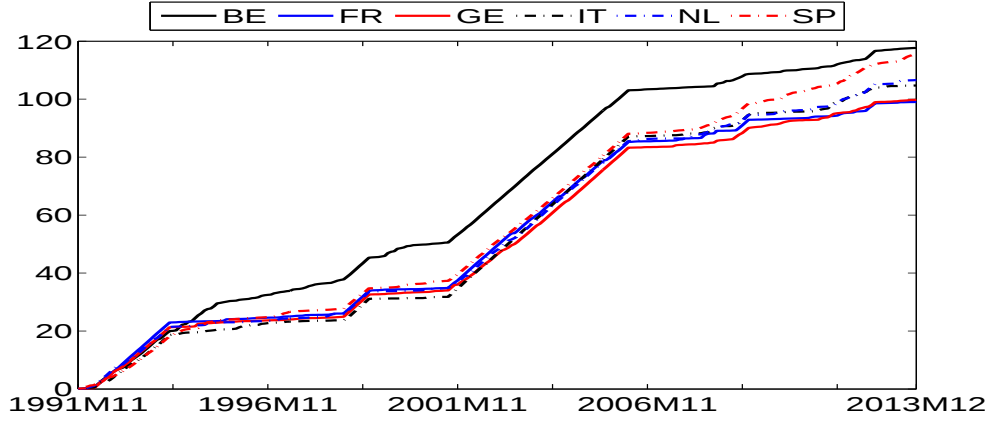
Figure 6: Cumulative concordance statistics of individual countries to predict the eurozone cycle. The labels "BE", "FR", "GE", "IT", "NE", "SP" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands and Spain.

applying the maximum a posteriori probability (MAP) estimator to the state posterior probabilities. The CS statistics 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 6 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 and the euro are aggregate cycle from mid nineties to the beginning of 2000. There is a period of large synchronization 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 second analysis, we apply a k-mean clustering algorithm to the regime probabilities of the six euro countries, $p_{it,kl}$. The benefits of this exercise are 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-mean algorithm maximizes the difference between clusters and minimizes the difference within cluster.

The results of the k-mean cluster analysis divides the countries in three groups:

1. France, Germany and the Netherlands.

2. Italy and Spain.

22

3. Belgium.

The first group can be labeled as core euro country members. The second group can be associated to the periphery countries unfolding differently the recent recession. Finally, Belgium differ for de-industrialization process in the nineties.[1]

Both exercises in this section find large heterogeneity for euro country business cycles, with the Financial Crisis ending a period of synchronization and dividing the eurozone in core euro and periphery members.

4.5 Reinforcement effects on regime probabilities through interaction

The evidence of strong heterogeneity of the cycles is one of the main results of our PMS-VAR model. Another relevant result regards the interaction between the cycles. The posterior estimates of the loadings of V_t (see Table 1) provide further information on the interaction between eurozone and US cycles. Estimates of the coefficients $\alpha_{1i}^{EU,11}$, $i = 1, \ldots, 6$, 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 a reinforcement effect, that is an increase in the probability to stay into the recession regime at time $t + 1$ due to the fact that the eurozone countries were in a recession phase in t. The evidence is different for US where this reinforcement does not exist, probably due to the leading behaviour of the US cycle and therefore its entering and reaching the peak of the recession in advance with respect to the eurozone. The US indicator seems not to have a clear reinforcement mechanism for the recession probabilities of the euro countries. The coefficient is positive, relative large and zero is outside the credible interval for Spain. For Belgium it is, on the contrary, negative; whereas the evidence is not clear for the other countries.

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

4.6 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 do not identify uniquely structural parameters and shocks across equations, implying that it

---

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

Country		$p_{it,11}$		$p_{it,12}$	
i	Label	$\alpha_{1i}^{EU,11}$	$\alpha_{1i}^{US,11}$	$\alpha_{1i}^{EU,12}$	$\alpha_{1i}^{US,12}$
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, $\boldsymbol{\alpha}_{1i} = (\alpha_{1i}^{11}, \alpha_{1i}^{12})'$, with $\alpha_{1i}^{kl} = (\alpha_{1i}^{EU,kl}, \alpha_{1i}^{US,kl})'$, which are the coefficients of the interaction variables η_{1t} and $\mathbb{I}(s_{7,t} = 1)$ driving the Markov-switching transition probabilities.

is not possible to distinguish regime shifts from one structural equation to another, see e.g. Sims and Zha (2006) and Primiceri (2005) for further discussions. Anyhow, we transform the PMS-VAR in the following structural model:

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

where $\boldsymbol{u}_{it} \sim \mathcal{N}_M(\mathbf{0}, I_M)$; $\mathbf{x}_{it} = (1, \mathbf{y}'_{i,t-1}, \ldots, \mathbf{y}'_{i,t-P})$, $A_i(s_{it}) = (\mathbf{a}_i(s_{it}), A_{i1}, \ldots, A_{iP})$ is estimated in the reduced form model and $A_i(s_{it})' = B_{+,i}(s_{it})B_{0,i}(s_{it})^{-1}$; $\boldsymbol{\varepsilon}_{it} = B_{0,i}(s_{it})^{-1}\boldsymbol{u}_{it}$; $E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}'_{it}) = (B_{0,i}(s_{it})B_{0,i}(s_{it})')^{-1}$ 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 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 in regime 1 (recession). Therefore, we extend evidence in Del Negro et al. (2014) to Markov Switching model for US and other six European countries. Our identification restriction scheme implies that the credit spread responds contemporaneously to a credit shock in each country; while IPI responds with one month-lag. The motivation is that financial variables move faster than real variables for several reasons, including publication delays of real variables. Figure 7 plots the impulse responses (IR) for the seven countries in our panel. Plots of IRs are standardized by imposing that the median response of US credit spread is 1. The response for IPI is plotted as 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 over several quarters is evident for Germany and US and to a less

extent for Spain. The response is, however, not significant for Italy and the Netherlands. Moreover, IPI increases in the first months after the shock for France. The large size of the government sector in the French economy can be an explanation for it. Figures F.1 and F.2 in the Online Appendix provide similar evidence for other phases of the cycle (regimes 2 and 3), documenting that credit spreads provide substantial predictive content across regimes, extending evidence in Gilchrist and Zakrajsek (2012) and Gilchrist and Mojon (2014).

5   Robustness

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 has significant predictive for economic growth in three industrial countries. However, there is no conclusive finding that the yield spread consistently contains information in explaining future economic activity. For example, Plosser and Rouwenhorst (1994) find the evidence that yield spreads contain 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 yield 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 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 the term 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 much more important than the term premium for predicting GDP more than two years ahead.
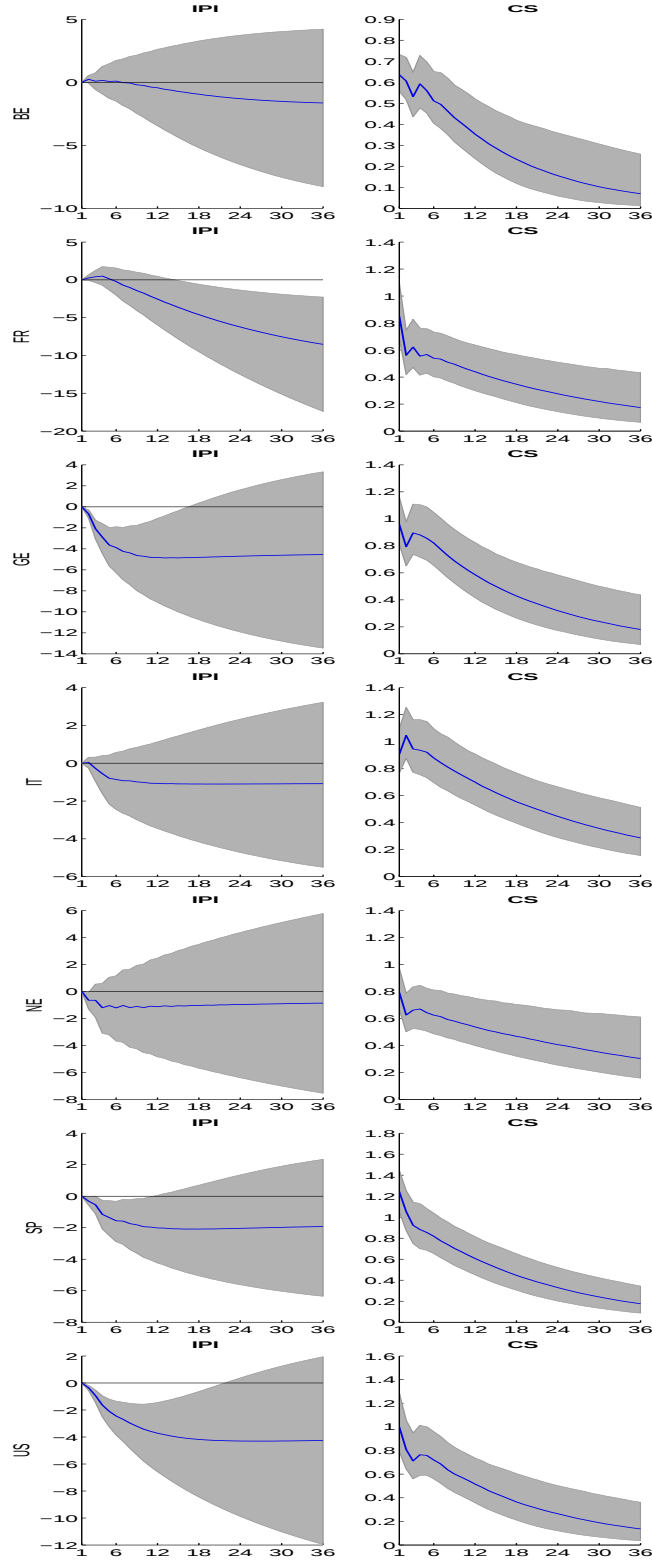
Figure 7: Median response (blue line) and 68% confidence interval to a credit shock in the first regime (recession) for industrial production growth (IPI) and credit spread (CS) for the different countries. Plots of IRs are standardized such as the median response of US credit spread in the first regime 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.

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 ECB succeeded in reducing government and corporate yields in most of the countries, in particular for Italy and Spain, and this results in a lower probability for regime 1 and higher probability for regime 2, in particular when using the alternative definition of credit spread. The exception is Germany since the definition of spread does not change for this country and the term spread has sharply reduced shortly after the US crisis following a pattern similar to the US term spread.

6   Conclusion

We propose a new Bayesian panel VAR model with unit-specific time-varying Markov-switching latent factors and develop a suitable Gibbs sampling procedure for posterior inference. We apply our panel MS-VAR model to the analysis of the interconnections and differences between eurozone and US business cycles.

Our results show that recession, slow recovery and expansion are empirically identified as three regimes with slow recovery becoming persistent in the eurozone in recent years differing from the US. US and eurozone cycles 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 very rapidly and recently more heterogeneity within the eurozone takes place. Cluster analysis yields a group of core countries: Germany, France and Netherlands and a group of peripheral countries Spain and Italy. Reinforcement 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. Finally, credit spreads provide accurate predictive content for business cycle fluctuations. A credit shock results in statistically significant negative industrial production growth for several months in Germany, Spain and US.

Our empirical result need to be investigated further but they may serve already as important information for the specification of a coordinated economic policy between the eurozone and the US economies and also within the eurozone economies.

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.

Amisano, G. and Tristani, O. (2013). Fundamentals and contagion mechanisms in the Euro area sovereign bonds markets. European Central Bank.

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.

Basturk, N., Cakmakli, C., Ceyhan, P., and van Dijk, H. K. (2013). Posterior-predictive evidence on US inflation using extended Phillips curve models with non-filtered data. *Journal of Applied Econometrics*, forthcoming.

Billio, M., Casarin, R., Ravazzolo, F., and Van Dijk, H. (2013a). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics*, 177:213–232.

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., 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.

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.

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.

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.

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.

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). In ation 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.

Ferrara, L. (2003). A three-regime real-time indicator for the US economy. *Economics Letters*, 81:373–378.

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-swithing 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.

Gilchrist, S. and Zakrajsek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720.

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.

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.

Harvey, C. (1991). The term structure and world economic growth. *Journal of Fixed Income*, 1:4–17.

Horvath, L. and Rice, G. (2014). Testing the equality of means when the observations are from functional time series. *Journal of Time Series Analysis*, forthcoming.

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. (2011). K-state switching models with endogenous transition distribution. Swiss National Bank Working Paper 2011-13.

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.

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.

Krolzig, H.-M. (2000). Predicting Markov-switching vector autoregressive processes. Nuffield College Economics Working Papers, 2000-WP31.

Krolzig, H.-M. (2004). Constructing turning point chronologies with Markov-switching vector autoregressive models: the Euro-zone business cycle. In Mazzi, G. L. and Savio, G., editors, *Modern Tools for Business Cycle Analysis*, pages 147–190. Eurostat, Luxembourg.

Lenk, P. J. and DeSarbo, W. S. (2000). Bayesian Inference for Finite Mixtures of Generalized Linear Models with Random Effects. *Psycometrika*, 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.

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.

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.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72:821–852.

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. (2006). Were there regime switches in us monetary policy? *American Economic Review*, 1:54–81.

Sims, C. A. 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.

A Computational details

A.1 Parameter full conditional densities

Updating $\boldsymbol{\gamma}_{i0}$. Then the full conditional distribution of the regime-independent parameter $\boldsymbol{\gamma}_{i0}$ is a normal with density function

$$
\begin{aligned}
f(\boldsymbol{\gamma}_{i0}|\mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\lambda}_0) &\propto \\
&\propto \exp\left\{ -\frac{1}{2}\boldsymbol{\gamma}_{i0}'\left(\sum_{t=1}^{T} X_{i0t}'\Sigma_{it}^{-1}X_{i0t} + \underline{\Sigma}_{i0}^{-1}\right)\boldsymbol{\gamma}_{i0} + \boldsymbol{\gamma}_{i0}\left(\sum_{t=1}^{T} X_{i0t}'\Sigma_{it}^{-1}\mathbf{y}_{i0t} + \underline{\Sigma}_{i0}^{-1}\boldsymbol{\lambda}_0\right)\right\} \\
&\propto \mathcal{N}_{MM_0}(\bar{\boldsymbol{\gamma}}_{i0}, \bar{\Sigma}_{i0})
\end{aligned}
\tag{22}
$$

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

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

$$
\begin{aligned}
f(\boldsymbol{\gamma}_{ik}|\mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_{i(-k)}, \Sigma, \boldsymbol{\lambda}_k) &\propto \\
&\propto \exp\left\{ -\frac{1}{2}\boldsymbol{\gamma}_i'\left(\sum_{t\in\mathcal{T}_{ik}} X_{ikt}'\Sigma_{it}^{-1}X_{ikt} + \underline{\Sigma}_{ik}^{-1}\right)\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_i'\left(\sum_{t\in\mathcal{T}_{ik}} X_{ikt}'\Sigma_{it}^{-1}\mathbf{y}_{ikt} + \underline{\Sigma}_{ik}^{-1}\boldsymbol{\lambda}_k\right)\right\} \\
&\propto \mathcal{N}_{MM_K}(\bar{\boldsymbol{\gamma}}_{ik}, \bar{\Sigma}_{ik})
\end{aligned}
\tag{23}
$$

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

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

$$
\begin{aligned}
f(\Sigma_{ik}^{-1}|\mathbf{y}_i, \Xi_i, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_i, \Sigma_{i(-k)}, \Upsilon_k) &\propto \\
&\propto |\Sigma_{ik}^{-1}|^{\frac{\nu_{ik}+T_{ik}-M-1}{2}} \exp\left\{ -\frac{1}{2}\mathrm{tr}\left(\left(\Upsilon_k^{-1} + \sum_{t\in\mathcal{T}_{ik}} \mathbf{u}_{ikt}\mathbf{u}_{ikt}'\right)\Sigma_{ik}^{-1}\right)\right\} \\
&\propto \mathcal{W}_M(\bar{\nu}_{ik}, \bar{\Upsilon}_{ik})
\end{aligned}
\tag{24}
$$

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

Updating $\boldsymbol{\alpha}_i$. The full conditional distribution of the parameters in the l-th row of the

transition matrix, $\boldsymbol{\alpha}_i^{\cdot l} = \text{vec}((\boldsymbol{\alpha}_i^{1l}, \ldots, \boldsymbol{\alpha}_i^{K-1l}))$, is

$$f(\boldsymbol{\alpha}_i^{\cdot l}|\mathbf{y}_i, \Xi, \boldsymbol{\gamma}_{i0}, \boldsymbol{\gamma}_i) \quad \propto \quad \prod_{t=1}^{T}\prod_{k=1}^{K}(H(V_t, \boldsymbol{\alpha}_i^{kl}))^{\xi_{ikt}\xi_{il\,t-1}} \exp\left\{-\frac{1}{2}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi})\Upsilon^{-1}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi})\right\}$$

Following Lenk and DeSarbo (2000), we apply a Metropolis-Hastings with proposal centred at the mode of the posterior distribution, with scale proportional to the posterior local curvature. Since the mode is not known, it is replaced by the value of the n-th iteration of the Newton-Raphson algorithm.

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

$$
\begin{aligned}
\nabla_{\boldsymbol{\alpha}_i^{kl}}^{(1)}\tilde{f}(\boldsymbol{\alpha}_i^{\cdot l}) &= \sum_{t=1}^{T}\xi_{il\,t-1}\Big(\xi_{ikt}\frac{1}{H(V_t, \boldsymbol{\alpha}_i^{kl})}\nabla_{\boldsymbol{\alpha}_i^{kl}}(H(V_t, \boldsymbol{\alpha}_i^{kl})) \\
&\quad + \sum_{k'\neq k, K}\xi_{ik't}\frac{1}{H(V_t, \boldsymbol{\alpha}_i^{k'l})}\nabla_{\boldsymbol{\alpha}_i^{kl}}(H(V_t, \boldsymbol{\alpha}_i^{k'l})) \\
&\quad + \xi_{iKt}\frac{1}{H(V_t, \boldsymbol{\alpha}_i^{Kl})}\nabla_{\boldsymbol{\alpha}_i^{kl}}(H(V_t, \boldsymbol{\alpha}_i^{Kl}))\Big) - \Upsilon^{-1}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi}) \\
&= \sum_{t=1}^{T}\xi_{il\,t-1}\Big(\xi_{ikt}H(V_t, \boldsymbol{\alpha}_i^{kl})\mathbf{z}_t + \sum_{k'\neq k, K}\xi_{ik't}H(V_t, \boldsymbol{\alpha}_i^{kl})\mathbf{z}_t \\
&\quad + \xi_{iKt}H(V_t, \boldsymbol{\alpha}_i^{kl})\mathbf{z}_t\Big) - \Upsilon^{-1}(\boldsymbol{\alpha}_i^{kl} - \boldsymbol{\psi})
\end{aligned}
\tag{25}
$$

we conclude that

$$D_1(\boldsymbol{\alpha}_i^{\cdot l}) = \sum_{t=1}^{T}\xi_{ilt-1}\Big(E(\boldsymbol{\xi}_{it} - H(V_t, \boldsymbol{\alpha}_i^{\cdot l})) \otimes \mathbf{z}_t\Big) - (I_{K-1} \otimes \Upsilon^{-1})(\boldsymbol{\alpha}_i^{\cdot l} - \boldsymbol{\iota}_{K-1} \otimes \boldsymbol{\psi}) \tag{26}$$

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, and $H(V_t, \boldsymbol{\alpha}_i^{\cdot l}) = (H(V_t, \boldsymbol{\alpha}_i^{1l}), \ldots, H(V_t, \boldsymbol{\alpha}_i^{Kl}))'$.

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

$$
\begin{aligned}
\nabla_{\boldsymbol{\alpha}_i^{k'l}\boldsymbol{\alpha}_i^{kl}}^{(2)} &= \sum_{t=1}^{T}\xi_{il\,t-1}\Big(H(V_t, \boldsymbol{\alpha}_i^{k'l})H(V_t, \boldsymbol{\alpha}_i^{kl}) \\
&\quad - H(V_t, \boldsymbol{\alpha}_i^{kl})\mathbb{I}(k = k')\Big)\mathbf{z}_t\mathbf{z}_t' - \mathbb{I}(k = k')\Upsilon^{-1}
\end{aligned}
\tag{27}
$$

thus the Hessian is

$$D_2(\boldsymbol{\alpha}_i^{\cdot l}) = \sum_{t=1}^{T} \xi_{ilt-1}\Big((EH(V_t, \boldsymbol{\alpha}_i^{\cdot l})(EH(V_t, \boldsymbol{\alpha}_i^{\cdot l}))' - \mathrm{diag}(EH(V_t, \boldsymbol{\alpha}_i^{\cdot l}))\Big) \otimes (\mathbf{z}_t\mathbf{z}_t') - I_{K-1} \otimes \Upsilon \tag{28}$$

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

$$\hat{\boldsymbol{\alpha}}^{(n+1)} = \hat{\boldsymbol{\alpha}}^{(n)} - D_2(\hat{\boldsymbol{\alpha}}^{(n)})^{-1} D_1(\hat{\boldsymbol{\alpha}}^{(n)}) \tag{29}$$

The proposal distribution is adapting over the iterations and at the iteration n the proposal $\boldsymbol{\alpha}^*$ for the parameter $\boldsymbol{\alpha}_i^{\cdot l}$ is drawn from the normal

$$\boldsymbol{\alpha}^* \sim \mathcal{N}_{(K-1)(G_v+1)}(\hat{\boldsymbol{\alpha}}^{(n+1)}, V^{(n)}) \tag{30}$$

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

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

$$\left.\left. ((\boldsymbol{\alpha}_i^{\cdot l})^{(n-1)} - \hat{\boldsymbol{\alpha}}_i^{(n+1)}) - \frac{1}{2}(\boldsymbol{\alpha}_i^* - \hat{\boldsymbol{\alpha}}_i^{(n+1)})'(V^{(n)})^{-1}(\boldsymbol{\alpha}_i^* - \hat{\boldsymbol{\alpha}}_i^{(n+1)})\right)\right\} \tag{31}$$

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

$$f(\boldsymbol{\lambda}_0|\boldsymbol{\gamma}, \Sigma) \propto \tag{32}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\lambda}_0'\left(\sum_{i=1}^{N} \underline{\Sigma}_{i0}^{-1} + \underline{\Sigma}_0^{-1}\right)\boldsymbol{\lambda}_0 - 2\boldsymbol{\lambda}_0'\left(\sum_{i=1}^{N} \underline{\Sigma}_{i0}^{-1}\boldsymbol{\gamma}_{i0} + \underline{\Sigma}_0^{-1}\underline{\boldsymbol{\lambda}}_0\right)\right]\right\}$$

$$\propto \mathcal{N}_{MM_0}(\bar{\boldsymbol{\lambda}}_0, \bar{\Sigma}_0)$$

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

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

$$f(\boldsymbol{\lambda}_k|\boldsymbol{\gamma}, \Sigma) \propto \tag{33}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\lambda}_k'\left(\sum_{i=1}^{N} \underline{\Sigma}_{ik}^{-1} + \underline{\Sigma}_k^{-1}\right)\boldsymbol{\lambda}_k - 2\boldsymbol{\lambda}_k'\left(\sum_{i=1}^{N} \underline{\Sigma}_{ik}^{-1}\boldsymbol{\gamma}_{ik} + \underline{\Sigma}_k^{-1}\underline{\boldsymbol{\lambda}}_k\right)\right]\right\}$$

$$\propto \mathcal{N}_{MM_K}(\bar{\boldsymbol{\lambda}}_k, \bar{\Sigma}_k)$$

where $\bar\Sigma_k^{-1} = \sum_{i=1}^N \underline\Sigma_{ik}^{-1} + \underline\Sigma_k^{-1}$ and $\bar{\boldsymbol\lambda}_k = \bar\Sigma_k \left(\sum_{i=1}^N \underline\Sigma_{ik}^{-1} \boldsymbol\gamma_{ik} + \underline\Sigma_k^{-1} \underline{\boldsymbol\lambda}_k \right)$.

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

$$f(\Upsilon_k^{-1} | \boldsymbol\gamma, \Sigma) \propto \tag{34}$$
$$\propto |\Upsilon_k^{-1}|^{\frac{\underline\nu_k - M - 1}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr}\left(\underline\Upsilon_k^{-1} \Upsilon_k^{-1} \right) \right\} \prod_{i=1}^N |\Upsilon_k^{-1}|^{\underline\nu_{ik}/2} \exp\left\{ -\frac{1}{2}\mathrm{tr}\left(\sum_{i=1}^N \Upsilon_k^{-1} \Sigma_{ik}^{-1} \right) \right\}$$
$$\propto \mathcal{W}_M(\bar\nu_k, \bar\Upsilon_k)$$

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

A.2   Allocation variable full conditional distributions

Let us define $\boldsymbol\xi_{-i,1:T} = (\boldsymbol\xi_{1:i-1,1:T}, \boldsymbol\xi_{i+1:N,1:T})$, with $\boldsymbol\xi_{i:j,1:T} = (\boldsymbol\xi_{i\,1:T}, \ldots, \boldsymbol\xi_{j\,1:T})$, $i \le j$, and $p(\boldsymbol\xi_{it} = \boldsymbol\iota_k | \boldsymbol\xi_{it-1} = \boldsymbol\iota_l, \boldsymbol\xi_{it-2}, V_t, \boldsymbol\alpha_i) = p_{it,kl}$, with $\boldsymbol\iota_k$ the k-th column of the identity matrix. The full conditional distribution of the allocation variables $\boldsymbol\xi_{i\,1:T}$ is

$$p(\boldsymbol\xi_{i\,1:T} | \mathbf{y}_{1:T}, \boldsymbol\xi_{-i,1:T}, \boldsymbol\gamma, \Sigma, \boldsymbol\alpha) \propto \tag{35}$$
$$\propto \prod_{i=1}^N \prod_{t=1}^T \left(p(\mathbf{y}_{it} | \mathbf{y}_{it-P-1:t-1}, \boldsymbol\xi_{it}, \boldsymbol\gamma_i, \Sigma_i) \prod_{k,l=1}^K p_{it,kl}^{\xi_{ilt}\xi_{ikt-1}} \right)$$
$$\propto \prod_{t=1}^T \left(p(\mathbf{y}_{it} | \mathbf{y}_{it-P-1:t-1}, \boldsymbol\xi_{it}, \boldsymbol\gamma_i, \Sigma_i) \prod_{k,l=1}^K p_{it,kl}^{\xi_{ikt}\xi_{ilt-1}} g_{it}(\boldsymbol\xi_{i,t-1} | \boldsymbol\xi_{-i,t-1}) \right)$$

where

$$g_{it}(\boldsymbol\xi_{it-1} | \boldsymbol\xi_{-i,t-1}) = \prod_{\substack{j=1 \\ j \ne i}}^N \prod_{k,l=1}^K \mathbb{P}(\boldsymbol\xi_{jt} = \boldsymbol\iota_k | \boldsymbol\xi_{jt-1} = \boldsymbol\iota_l, \boldsymbol\xi_{jt-2}, V_t, \boldsymbol\alpha_j)^{\xi_{jkt}\xi_{jlt-1}}$$

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

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 distribution $\mathbb{P}(\boldsymbol\xi_{jt} | \boldsymbol\xi_{jt-1}, \boldsymbol\xi_{jt-2}, V_t, \boldsymbol\alpha_j)$ 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. We note that terms are functions of the bivariate Markov chain $\tilde s_{it} = (s_{it}, s_{it-1})$ with values in the product space $\{1, \ldots, K\}^2$. Thus, we introduce the allocation variable $\tilde{\boldsymbol\xi}_{it} = (\tilde\xi_{i1t}, \ldots, \tilde\xi_{iK^2t})'$, where $\tilde\xi_{ikt} = \mathbb{I}(((s_{it} - 1)K + s_{it-1}) = k)$. The transition probability

of the process $\tilde{\boldsymbol{\xi}}_{it}$, for the case $K = 3$, is $\mathbb{P}(\tilde{\boldsymbol{\xi}}_{it}|\tilde{\boldsymbol{\xi}}_{it-1}, V_t, \boldsymbol{\alpha}_j) = I_3 \otimes P_{it}$ where

$$P_{it} = \begin{pmatrix} h_{it,11} & h_{it,21} & h_{it,31} \\ h_{it,12} & h_{it,22} & h_{it,32} \\ h_{it,13} & h_{it,23} & h_{it,33} \end{pmatrix} \tag{36}$$

where $h_{it,kl} = H(V_t, \boldsymbol{\alpha}_i^{kl})$. In the case a minimum duration constrain of two periods is introduced (see equation 6), then the transition probability is:

$$\mathbb{P}(\tilde{\boldsymbol{\xi}}_{it}|\tilde{\boldsymbol{\xi}}_{it-1}, V_t, \boldsymbol{\alpha}_i) = \begin{pmatrix} h_{it,11} & h_{it,21} & h_{it,31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{it,12} & h_{it,22} & h_{it,32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{it,13} & h_{it,23} & h_{it,33} \end{pmatrix} \tag{37}$$

which allows us to easily impose the duration constraint when simulating the hidden states.

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

$$p(\tilde{\boldsymbol{\xi}}_{i\,1:T}|\mathbf{y}_{i\,1:T}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}, \Sigma, \boldsymbol{\alpha}) \propto \Big(p(\tilde{\boldsymbol{\xi}}_{iT}|\mathbf{y}_{i\,1:T}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i) \tag{38}$$
$$\prod_{t=1}^{T-1} p(\tilde{\boldsymbol{\xi}}_{it+1}|\tilde{\boldsymbol{\xi}}_{it}, V_t, \boldsymbol{\alpha}_i) p(\tilde{\boldsymbol{\xi}}_{it}|\mathbf{y}_{i\,1:t}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i) \Big)$$

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 are described in the following. First, the filtered probability distribution for the i-th Markov chain at time t, $t = 1, \ldots, T$, is determined by iterating the prediction step

$$p(\tilde{\boldsymbol{\xi}}_{it} = \boldsymbol{\iota}_k|\mathbf{y}_{i\,1:t-1}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\alpha}) = \sum_{l=1}^{K^2} \tilde{p}_{it,kl} p(\tilde{\boldsymbol{\xi}}_{it-1} = \boldsymbol{\iota}_l|\mathbf{y}_{i\,1:t-1}, \boldsymbol{\xi}_{-i,1:T}, \boldsymbol{\gamma}_i, \Sigma_i, \boldsymbol{\alpha}) \tag{39}$$

where $\tilde{p}_{it,kl}$ is the (l, k)-th element of the transition matrix $\mathbb{P}(\tilde{\boldsymbol{\xi}}_{it}|\tilde{\boldsymbol{\xi}}_{it-1}, V_t, \boldsymbol{\alpha}_j)$, and the

updating step

$$p(\tilde{\boldsymbol{\xi}}_{it}|\mathbf{y}_{i\,1:t},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha}) \propto \tag{40}$$
$$\propto p(\tilde{\boldsymbol{\xi}}_{it}|\mathbf{y}_{i\,1:t-1},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha})p(\mathbf{y}_{it}|\mathbf{y}_{i\,t-1-P:t-1},\tilde{\boldsymbol{\xi}}_{it},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha})$$

Secondly, the smoothed probabilities given by

$$p(\tilde{\boldsymbol{\xi}}_{it}=\boldsymbol{\iota}_k|\mathbf{y}_{i\,1:T},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha}) \propto \tag{41}$$
$$\propto \sum_{l=1}^{K^2} p(\tilde{\boldsymbol{\xi}}_{it}=\boldsymbol{\iota}_k|\tilde{\boldsymbol{\xi}}_{it+1}=\boldsymbol{\iota}_l,\mathbf{y}_{i\,1:t},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha})p(\tilde{\boldsymbol{\xi}}_{it+1}=\boldsymbol{\iota}_l|\mathbf{y}_{i\,1:T},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha})$$

with $k=1,\ldots,K^2$, are evaluated recursively and backward in time for $t=T,T-1,\ldots,1$, with initial condition $p(\tilde{\boldsymbol{\xi}}_{iT}=\boldsymbol{\iota}_k|\mathbf{y}_{i\,1:T},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha})$ given by the last forward filtering iteration. The conditional distribution

$$p(\tilde{\boldsymbol{\xi}}_{it}=\boldsymbol{\iota}_k|\tilde{\boldsymbol{\xi}}_{it+1}=\boldsymbol{\iota}_l,\mathbf{y}_{i\,1:T},\boldsymbol{\xi}_{-i,1:T},\boldsymbol{\gamma}_i,\Sigma_i,\boldsymbol{\alpha}) \propto \tilde{p}_{it+1,lk}p(\tilde{\boldsymbol{\xi}}_{it}=\boldsymbol{\iota}_k|\mathbf{y}_{i\,1:t},\boldsymbol{\xi}_{-i,1:T})$$

represents the building block of the smoothed probability formula and is used in the FFBS algorithm to sample the allocation variables from their joint posterior distribution sequentially and backward in time for $t=T,T-1,\ldots,1$. See Frühwirth-Schnatter (2006), ch. 11-13, where the same set of recursions is given in terms of hidden state variables instead of allocation variables.

Note that in the updating step the two terms of the augmented likelihood, $p(\mathbf{y}_{it}|\mathbf{y}_{it-P-1:t-1},\boldsymbol{\xi}_{it},\boldsymbol{\gamma}_i,\Sigma_i)$ and $g_{it}(\boldsymbol{\xi}_{it-1}|\boldsymbol{\xi}_{-i,t-1})$, are evaluated at the different values of allocation variable, which 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 $\boldsymbol{\xi}_{it-1}=\boldsymbol{\iota}_k$, $k=1,\ldots,K$. 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 $\xi_{ikt}=\sum_{l=K(k-1)+1}^{Kk}\tilde{\xi}_{ilt}$, $k=1,\ldots,K$, 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.

Online Appendix

B MCMC convergence issues

As regards to the number of iterations, we should say that 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 with the non-parametric estimator

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

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

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 (2014)). Then the following statistics

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

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

C Smoothed probabilities
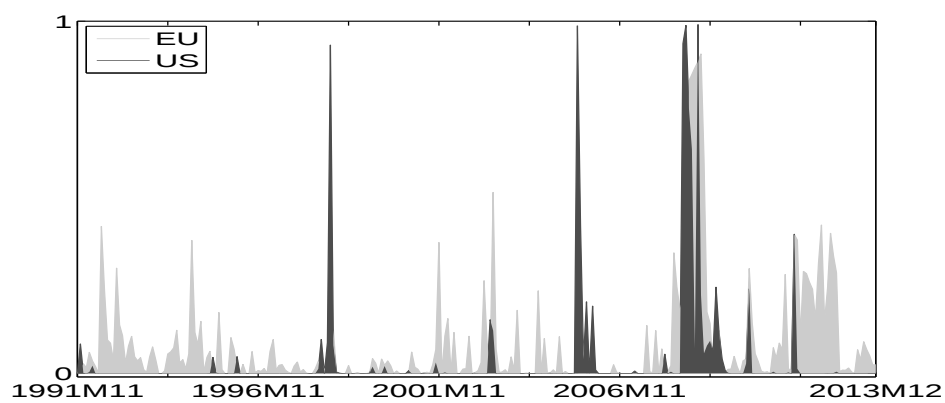
Figure C.1: Smoothed probabilities of the eurozone and US economies to be in recessions
.
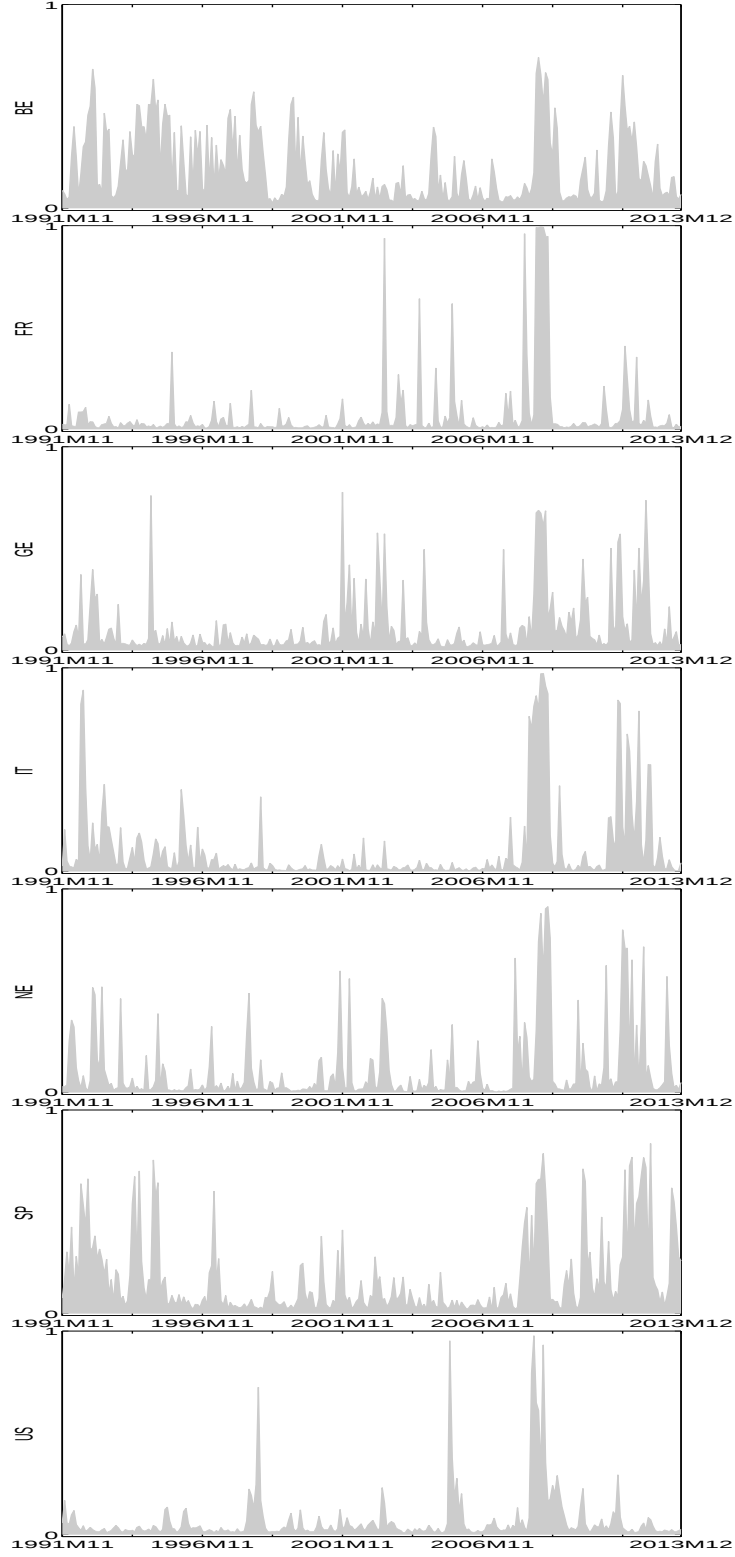
Figure C.2: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.
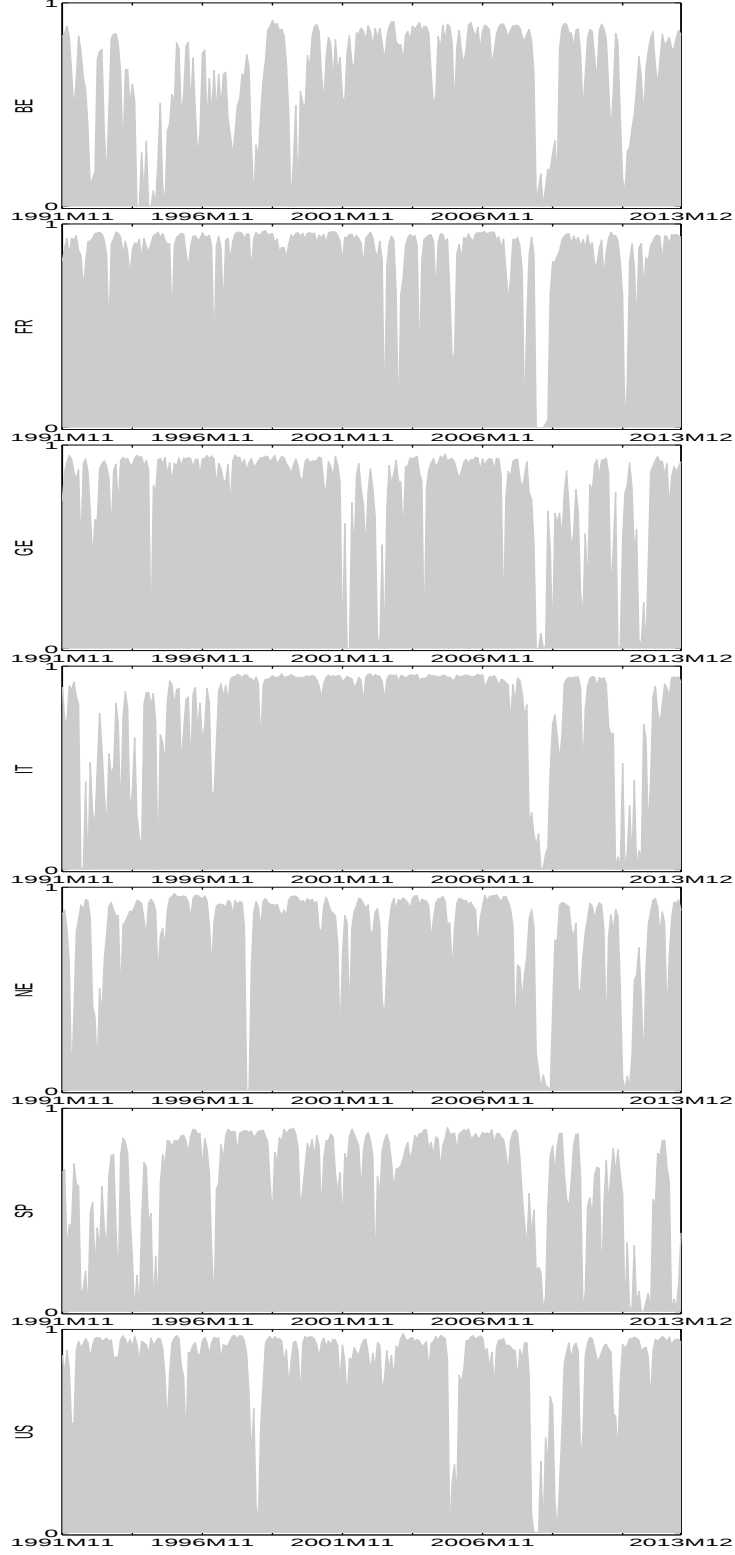
Figure C.3: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.
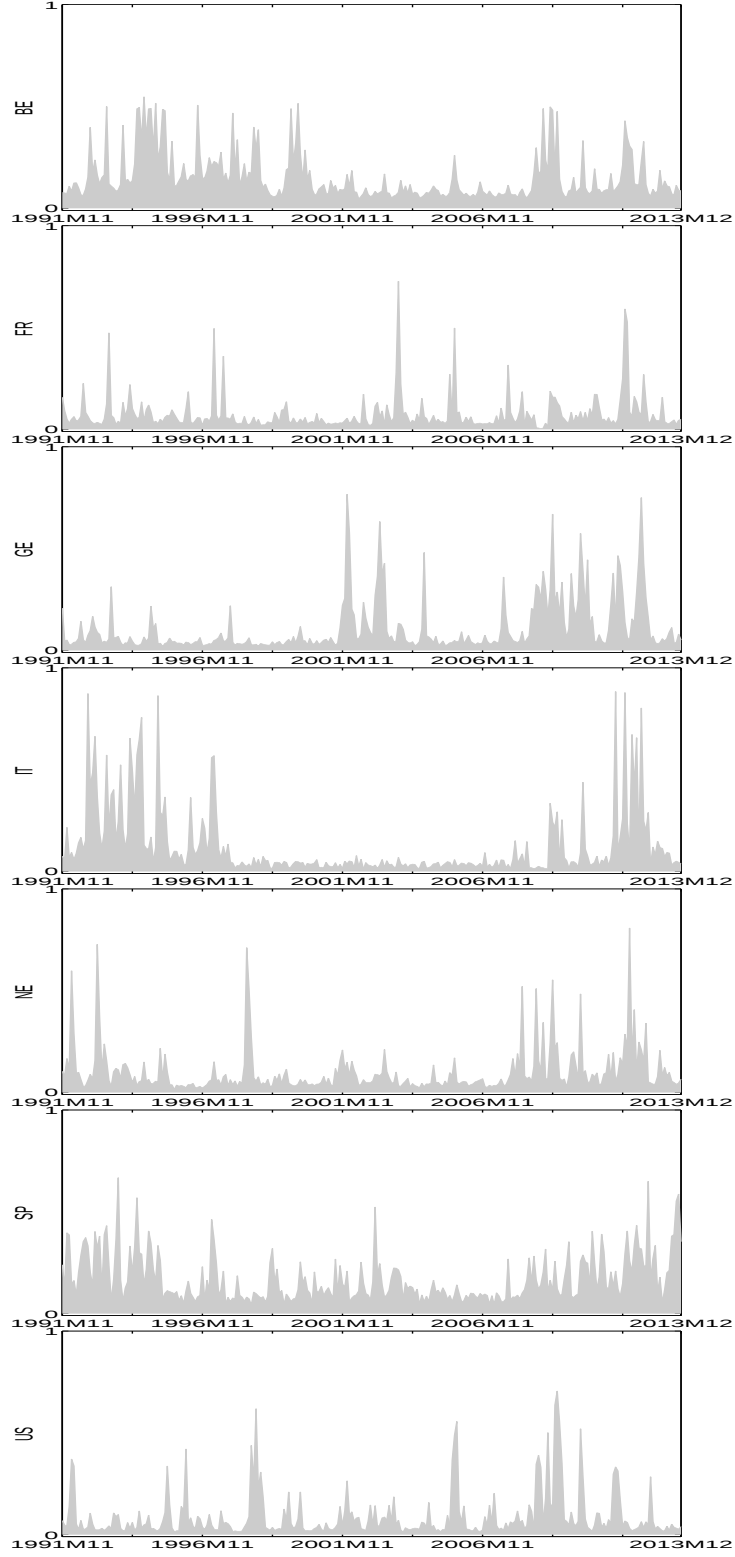
Figure C.4: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

5

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:

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

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

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.
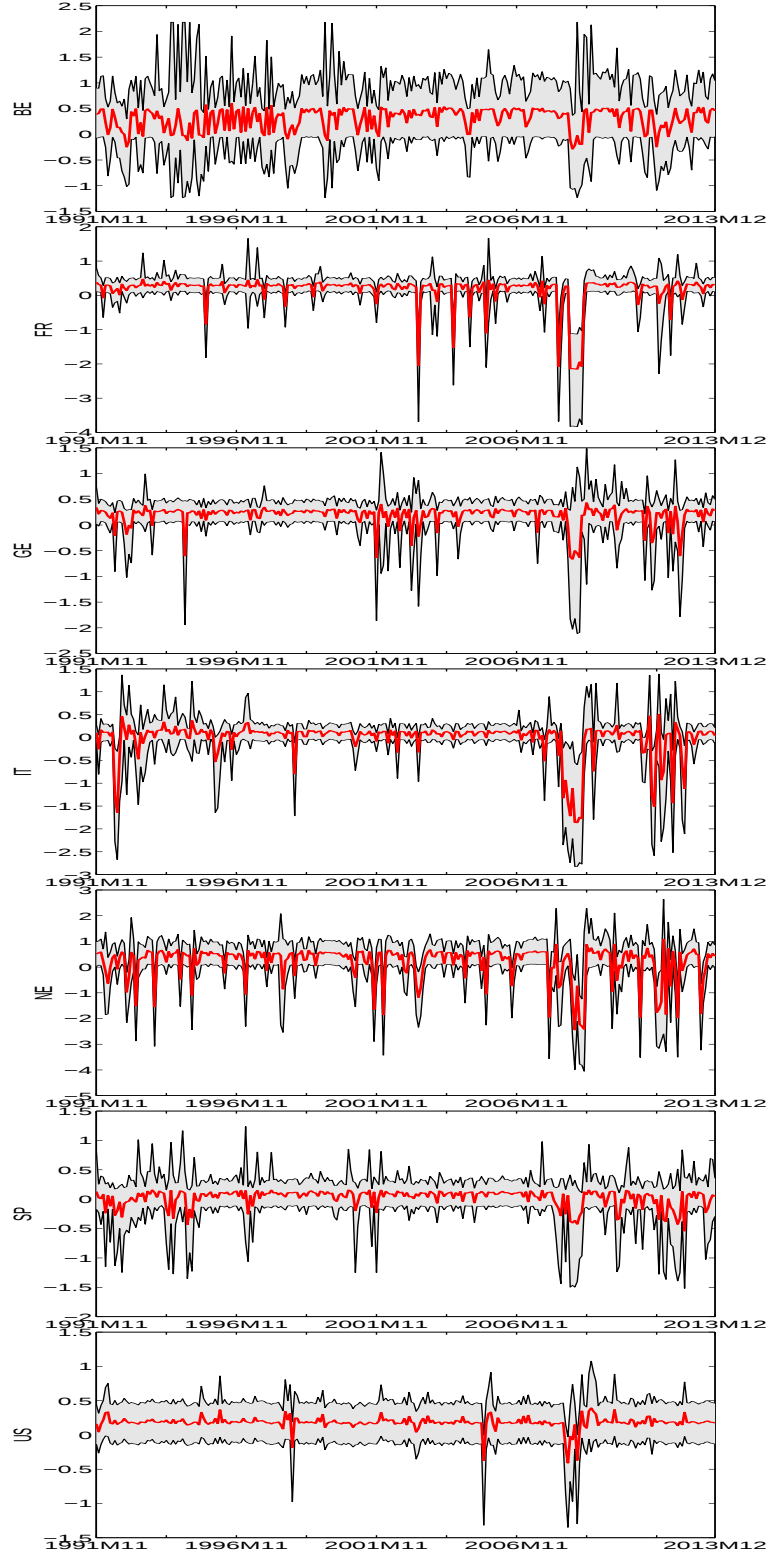
Figure D.1: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying intercept for IPI growth. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.
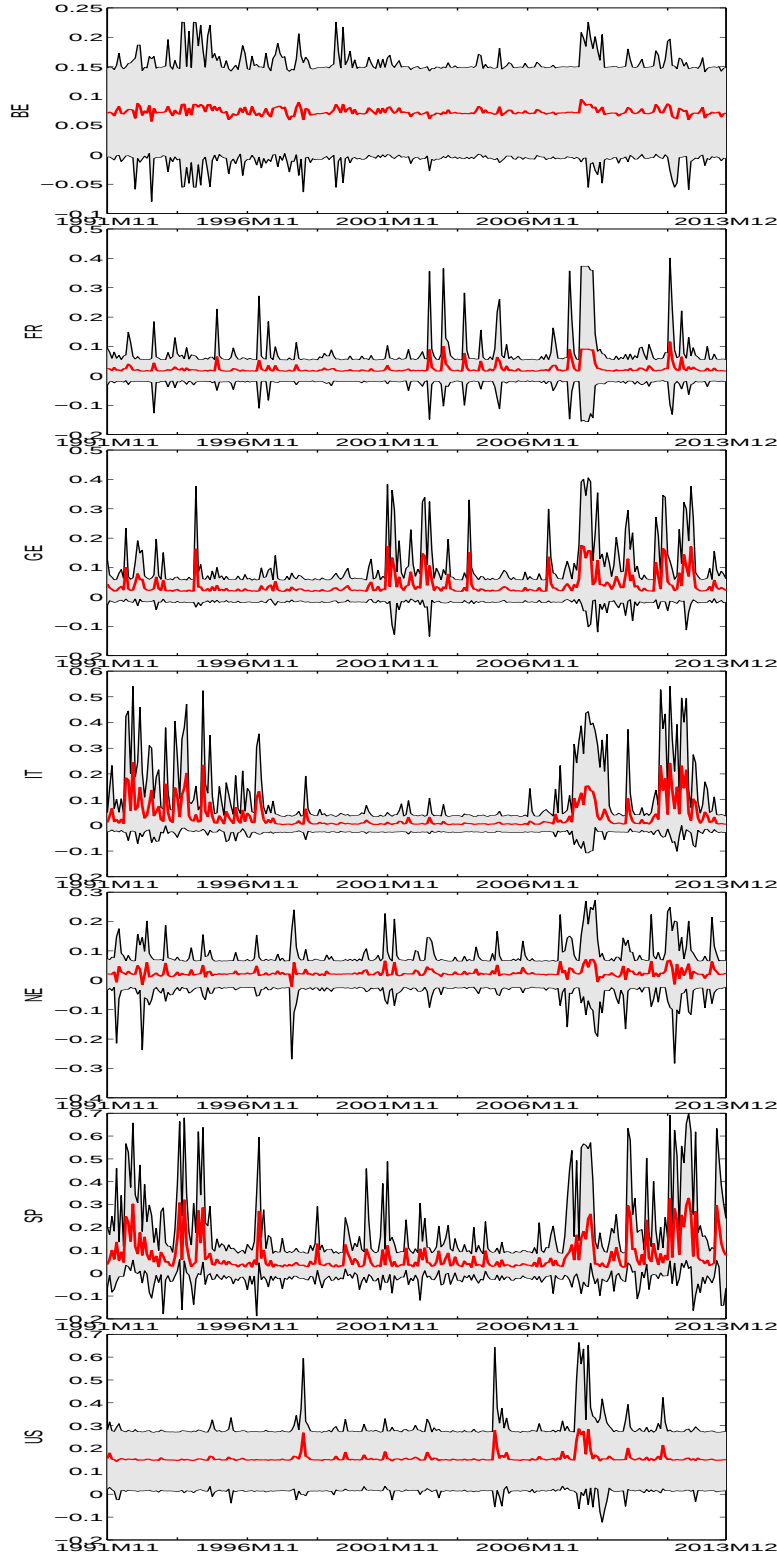
Figure D.2: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying intercept for the credit spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.
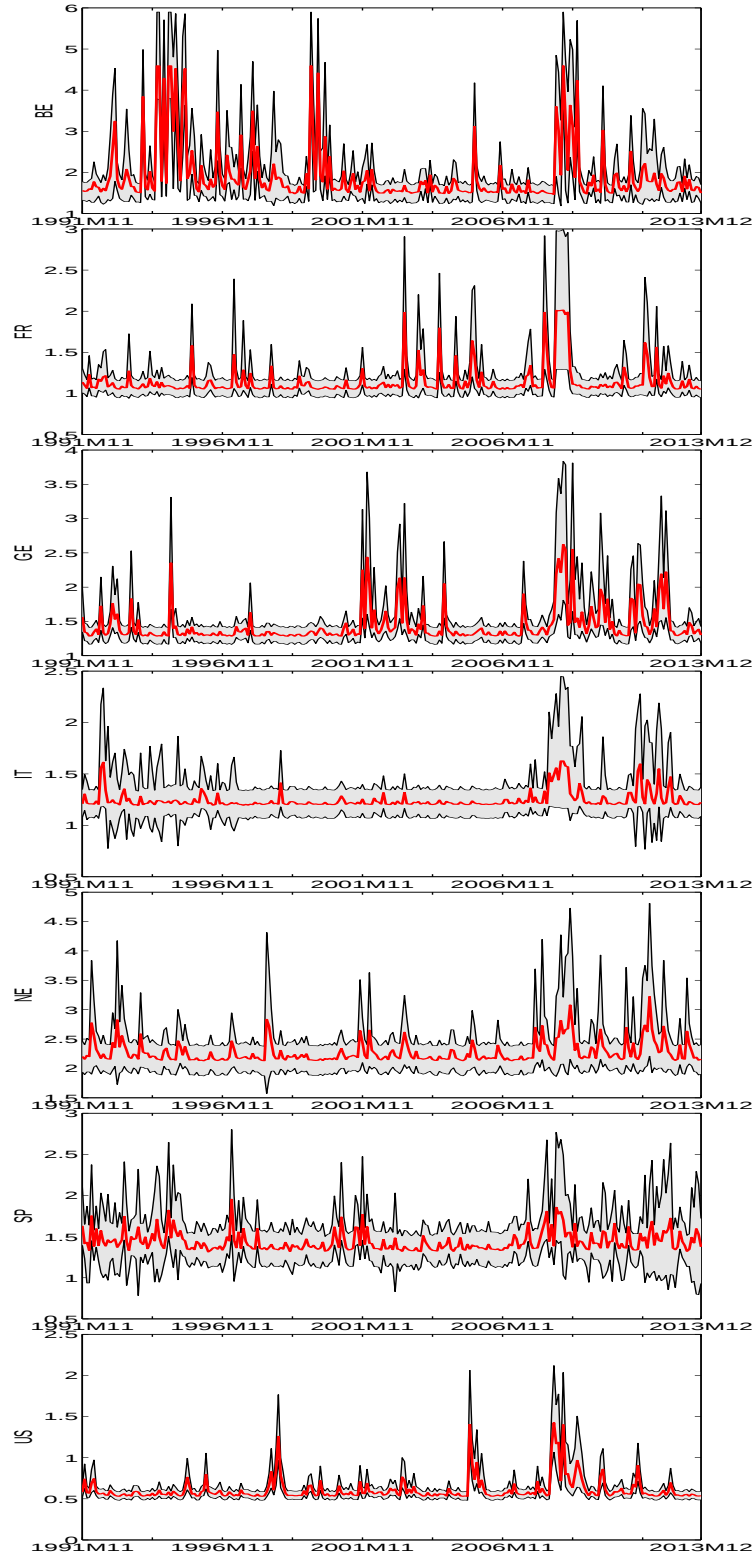
Figure D.3: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying standard deviation IPI growth. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

9
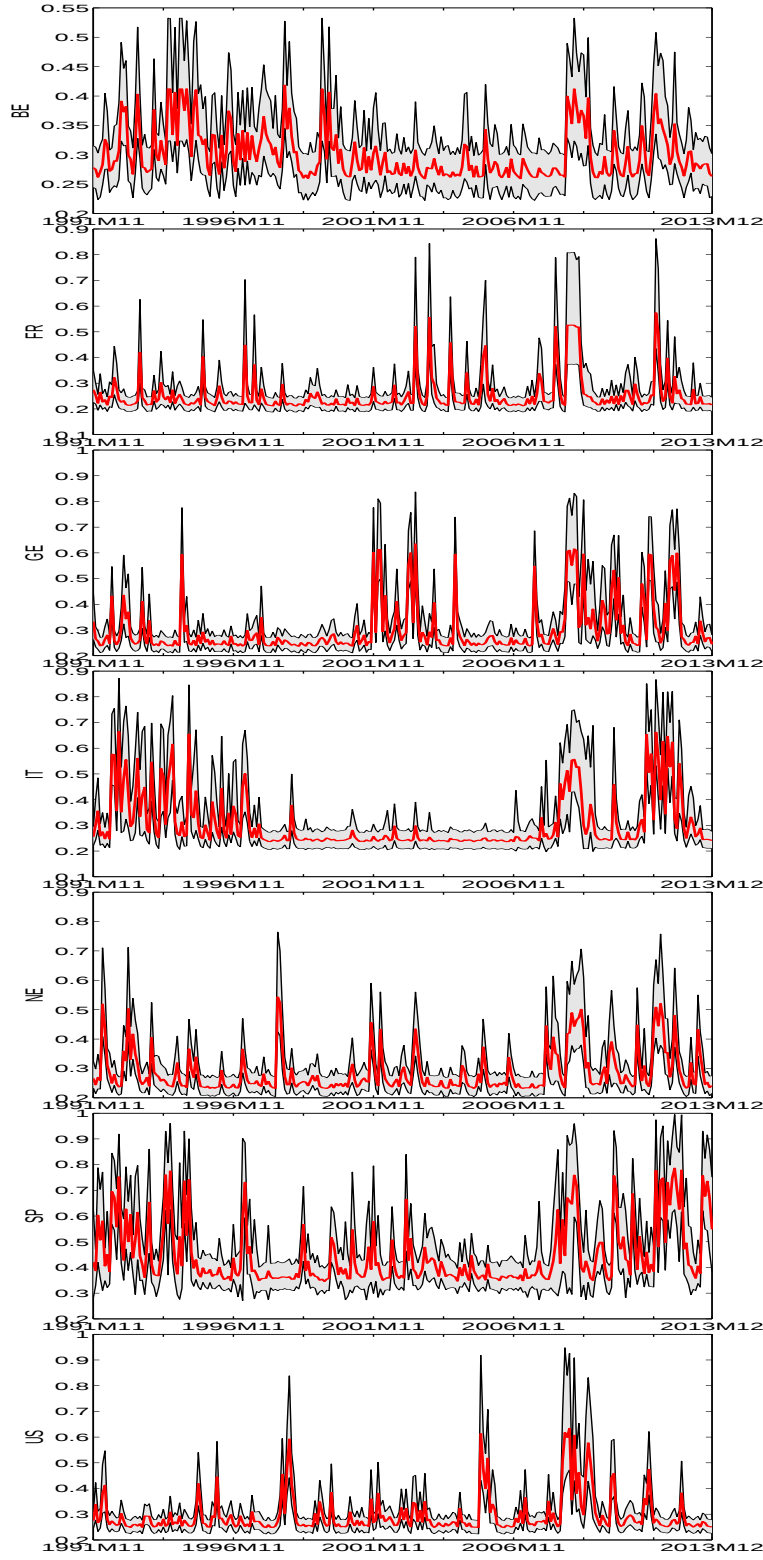
Figure D.4: Posterior mean distributions (in red) and 90% posterior intervals (in grey) for the VAR time-varying standard deviation for the credit spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

The US in regimes 2 or 3 (moderate or expansion)   The US in regime 1 (recession)
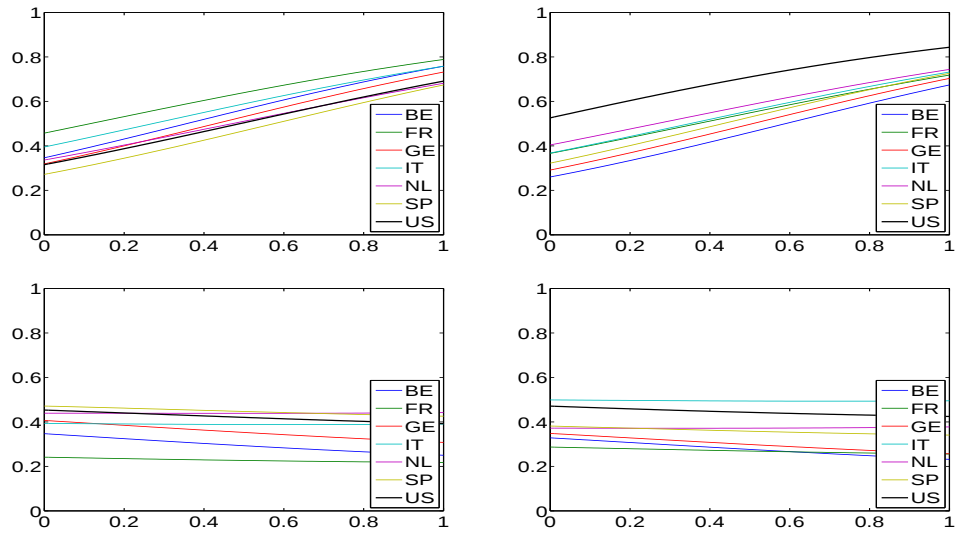
Figure D.5: Reaction of the transition probabilities to stay in recession $p_{it,11}$ (first row) and to exit the recession $p_{it,12}$ (second row) to changes in the weighted aggregate numbers of countries in recession η_{1t}, when conditioning on not recession for the US, i.e. $s_{7,t} \neq 1$ (left column) and recession for the US, i.e. $s_{7,t} = 1$, (right column).

.

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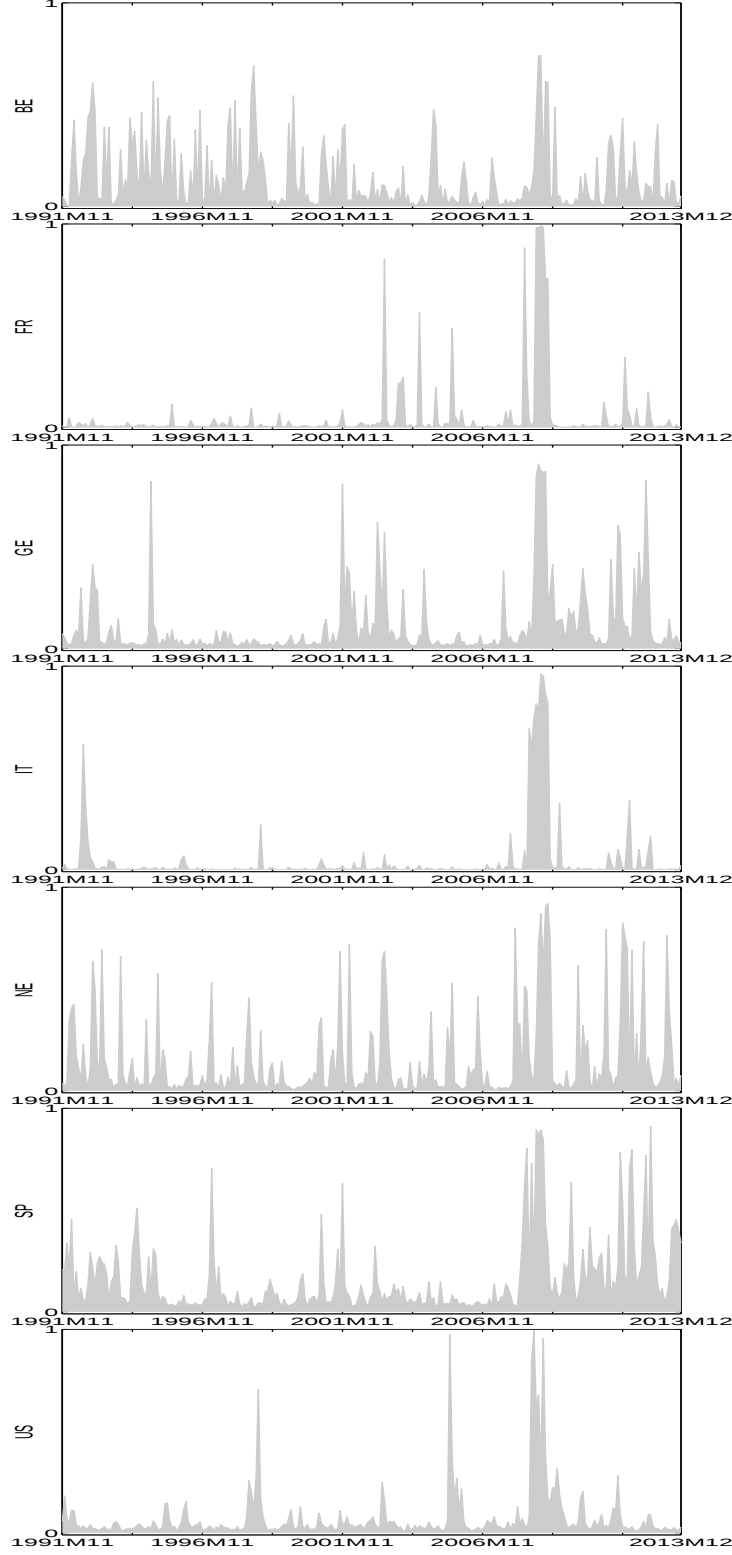
Figure E.1: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using an alternative definition of credit spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy,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

13
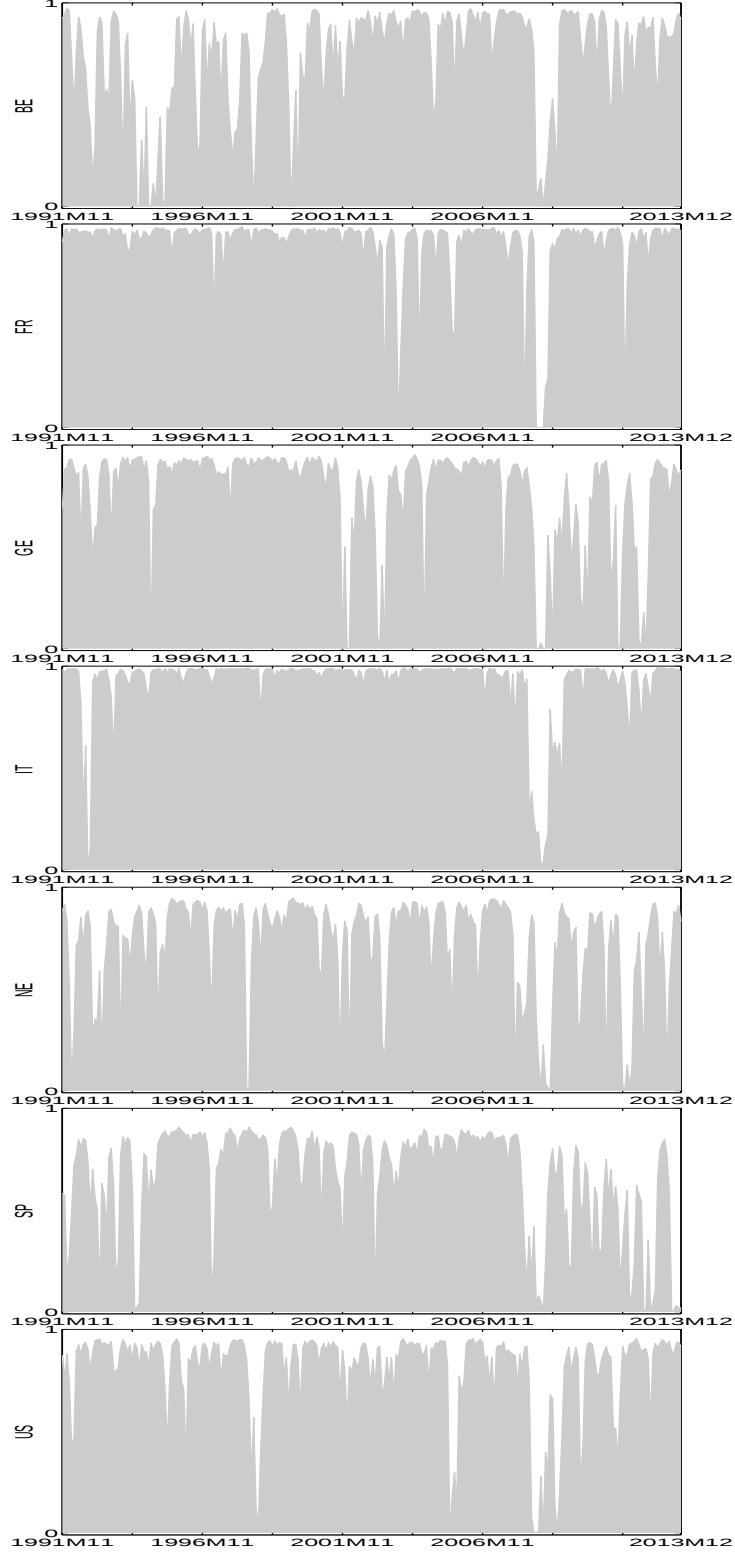
Figure E.2: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using an alternative definition of credit spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

14
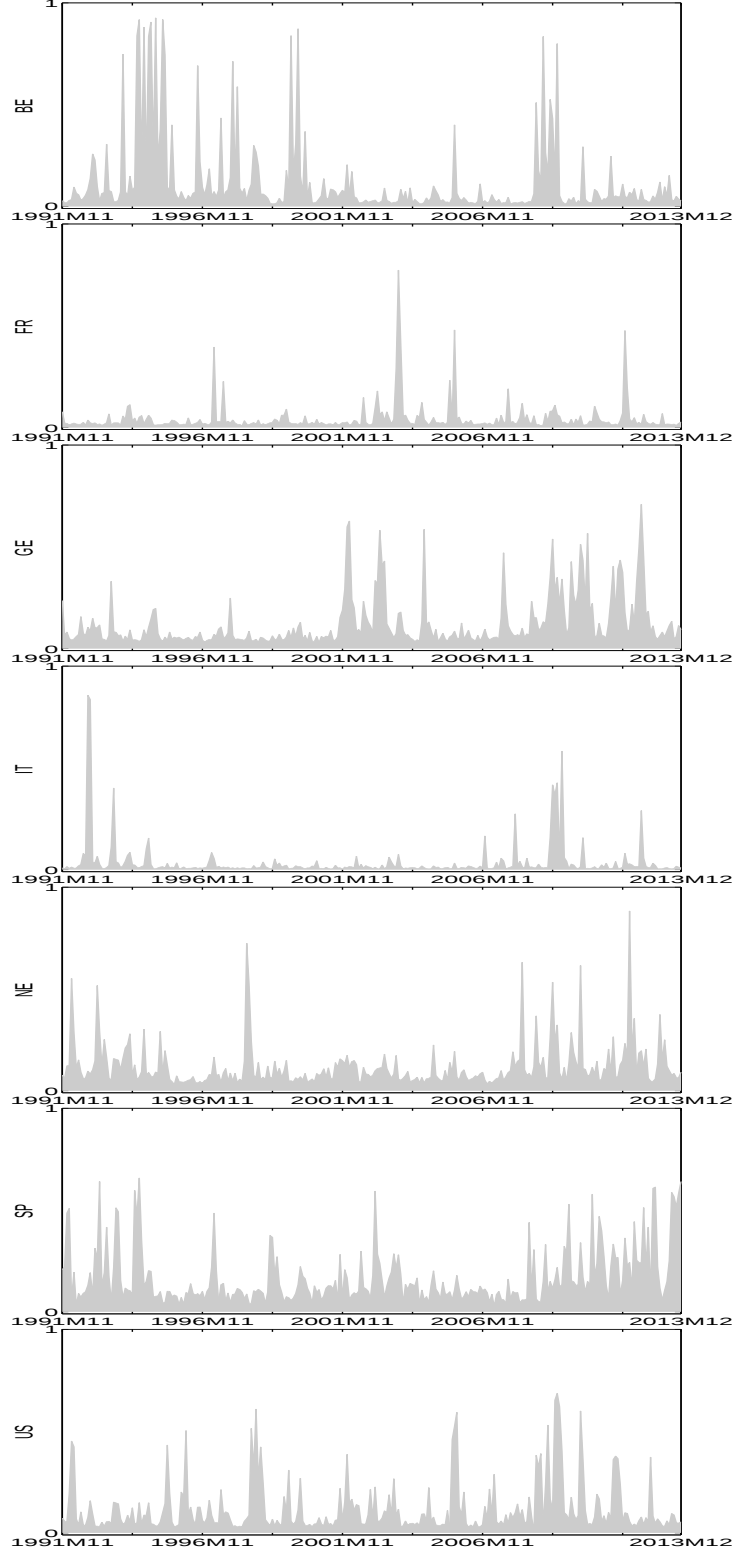
Figure E.3: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using an alternative definition of credit spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.
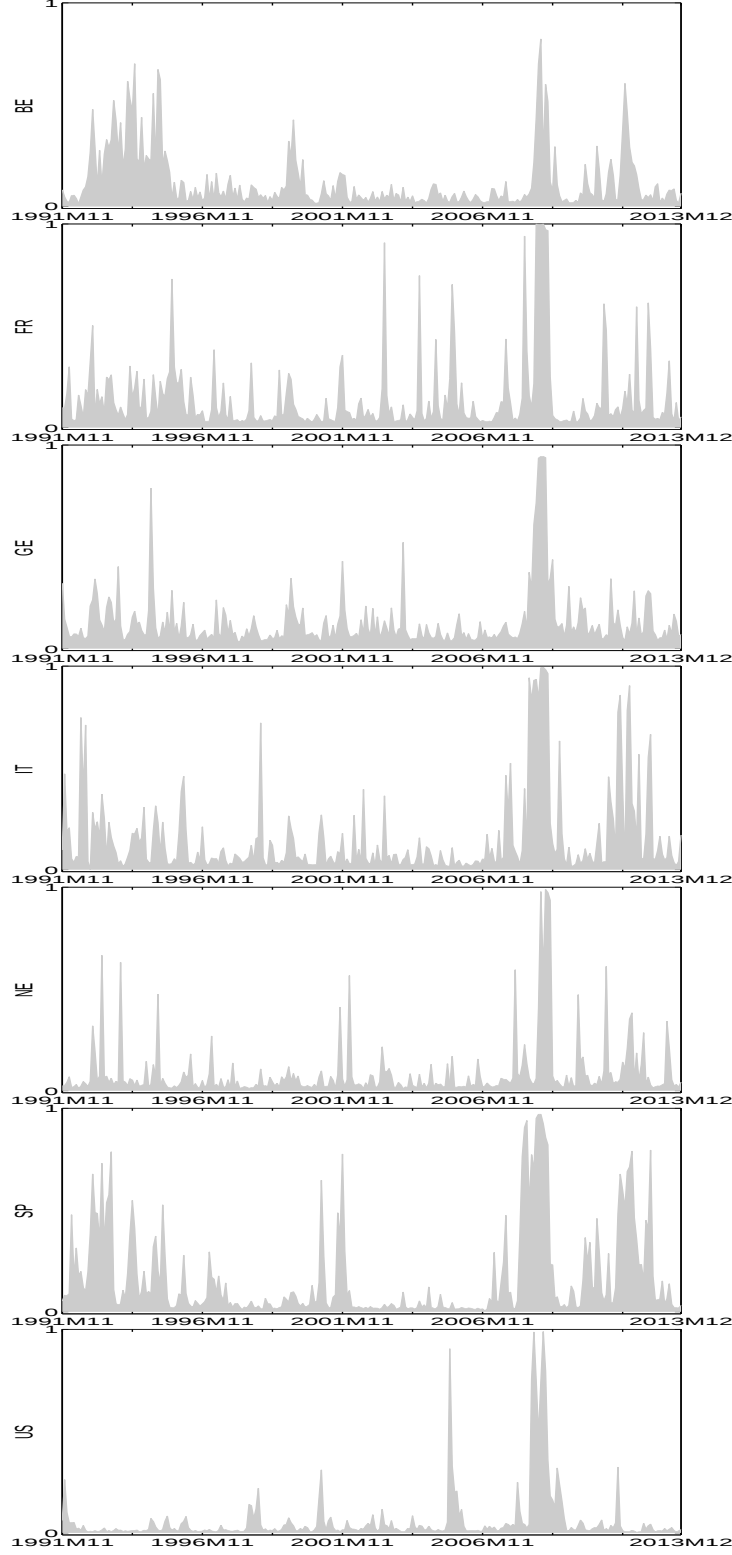
Figure E.4: First regime (recession) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using the term spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

16
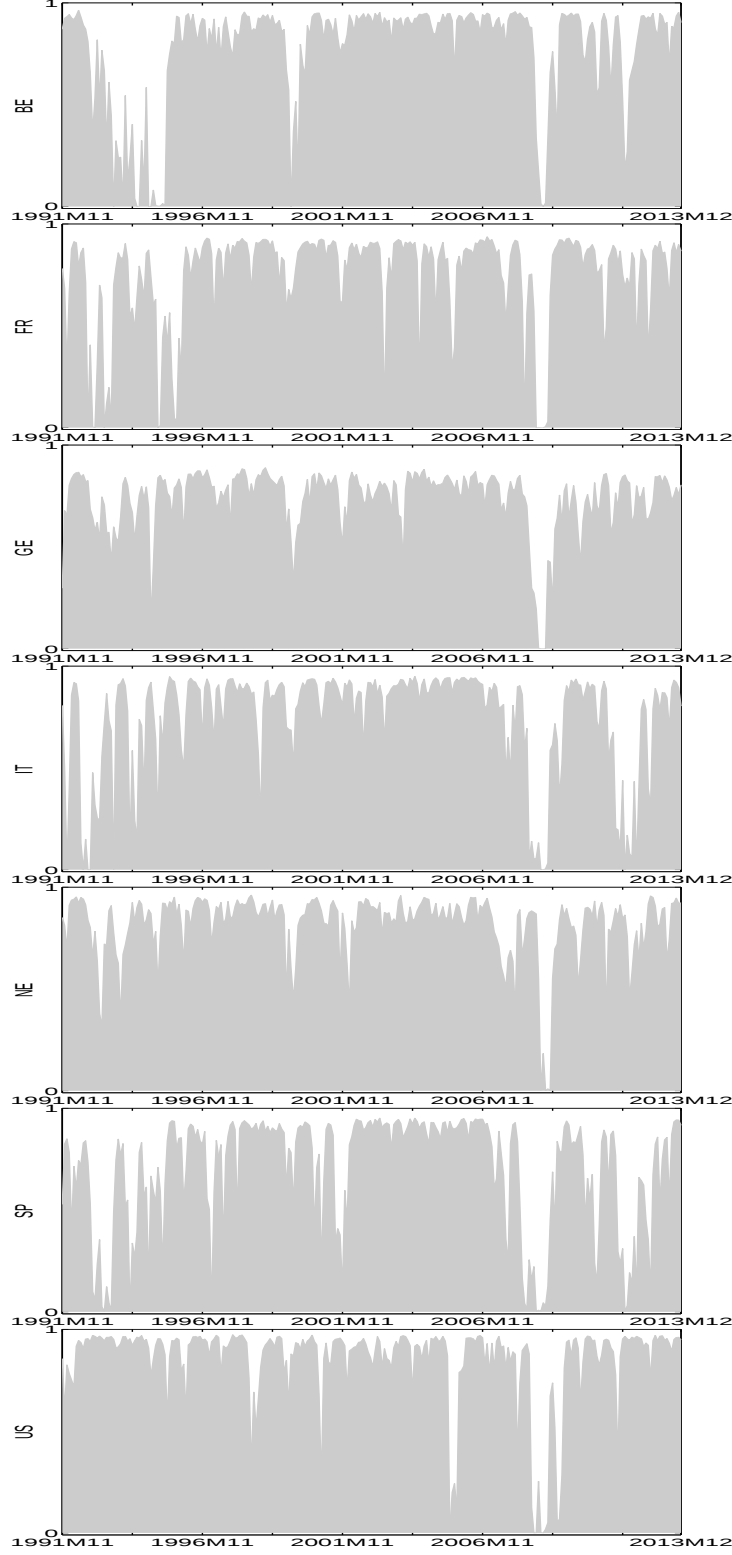
Figure E.5: Second regime (slow recovery and moderate expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using the term spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

17
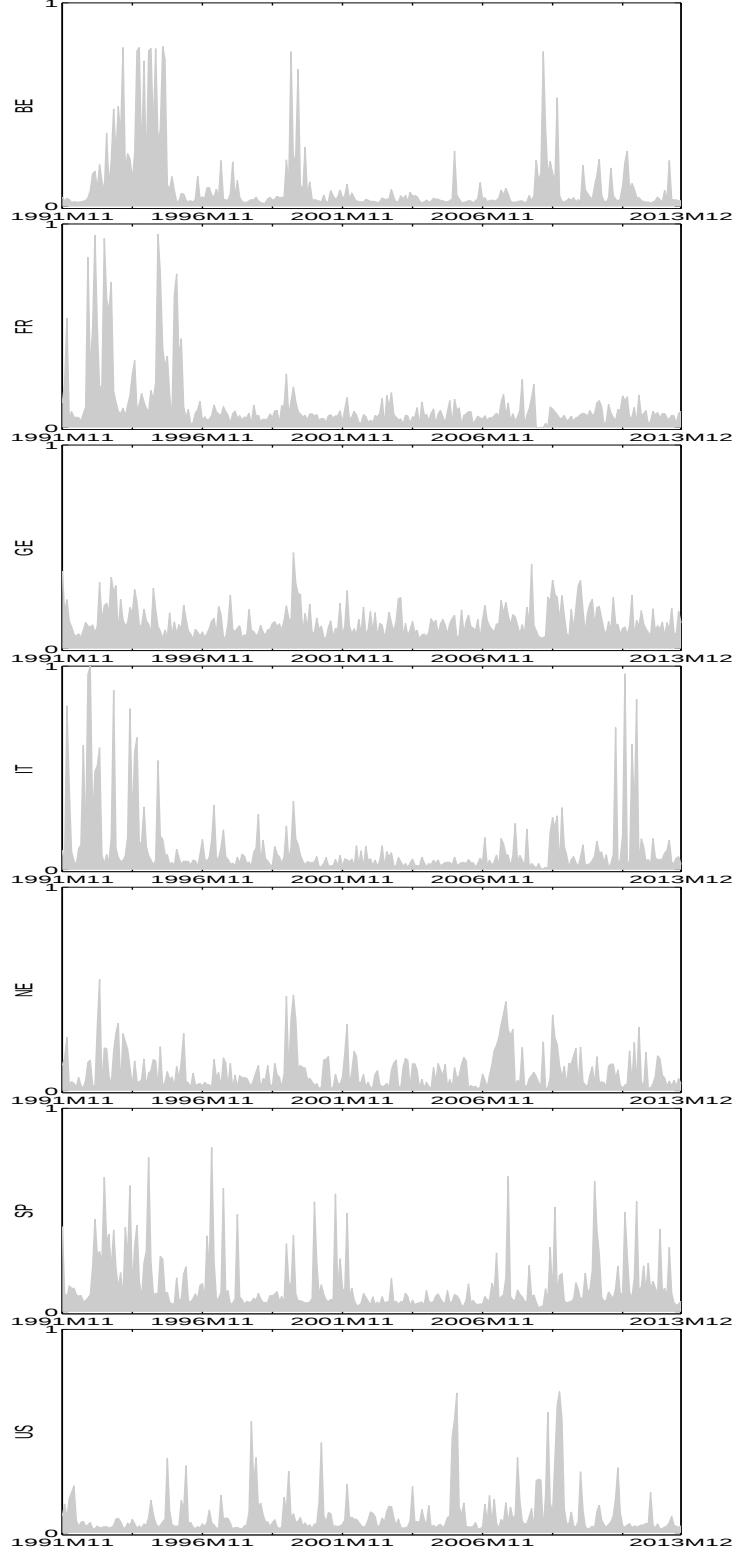
Figure E.6: Third regime (expansion) smoothed probabilities for the Markov-switching processes $s_{i,t}$, $i = 1, \ldots, N$ and $t = 1, \ldots, T$ using the term spread. The labels "BE", "FR", "GE", "IT", "NE", "SP", "US" indicate, respectively, Belgium, France, Germany, Italy, the Netherlands, Spain and the US.

F   Credit spread shocks
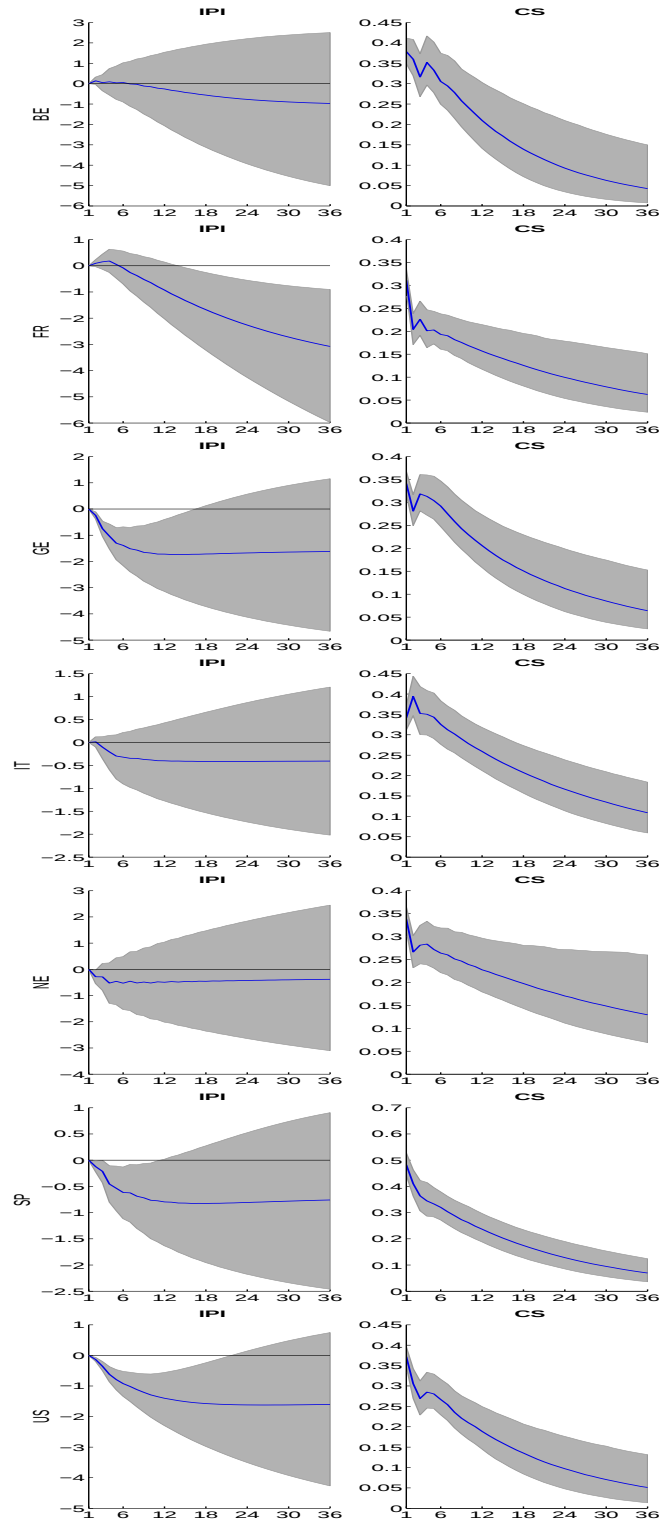
Figure F.1: Median response (blue line) and 68% confidence interval to a credit shock in the second regime (slow recovery and moderate expansion) for the different countries. Plots of IRs are standardized such as the median response of US credit spread in the first regime is 1. The response for IPI is plotted as cumulative sum over horizons.
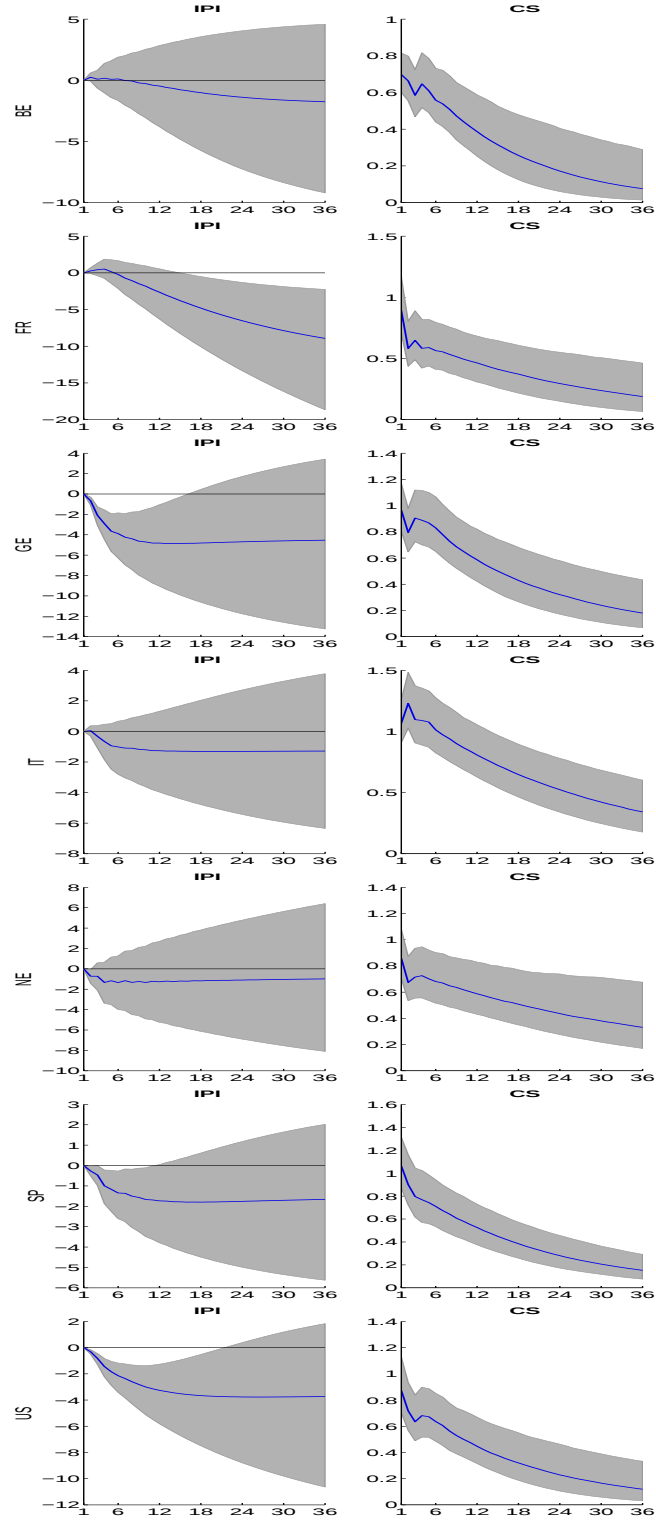
Figure F.2: Median response (blue line) and 68% confidence interval to a credit shock in the third regime (expansion) for the different countries. Plots of IRs are standardized such as the median response of US credit spread in the first regime is 1. The response for IPI is plotted as cumulative sum over horizons.