International Journal of Academic Research in Accounting, Finance and Management Sciences
Vol. 3, No.3, July 2013, pp. 372–383
ISSN: 2225-8329

www.hrmars.com

Assessing Structural Convergence between Romanian Economy and Euro Area: A Bayesian Approach

Alexie ALUPOAIEI[1]
Ana-Maria SANDICA[2]

[1,2]Bucharest Academy of Economic Studies, Romania,
[1]E-mail: alex.alupoaiei@gmail.com, [2]E-mail: anamaria.sandica@gmail.com

Abstract *In this paper we involved a study of structural convergence between Romanian and Euro Zone economies from the view point of synchronization in responses to shocks. For this purpose we called a Bayesian framework in which we estimated a time-varying parameters VAR model. For the identification of structural shocks we started from semi-structural VAR in which we incorporated the standard predictions of DSGE literature for a New-Keynesian model. For several purposes mentioned in the paper, we used two versions of data, replacing GDP and GDP deflator from a standard approach with consumption and its deflator. In this paper we were mainly interested for the response of interest variables to a monetary shock for policy purposes and in a second timeframe for the responses to other types of shocks.*

Key words Business Cycles, consumption, sign restriction convergence, Minnesota Prior, VAR, Gibbs Sampling, Bayesian econometrics

DOI: 10.6007/IJARAFMS/v3-i3/247 URL: http://dx.doi.org/10.6007/IJARAFMS/v3-i3/247

1. Introduction

Given the recent developments from the last two decades in regard with geo-political and economic options, Romania as other countries from Central and Eastern Europe which joined the European Union intend to adopt the euro currency in the short or mid run. This fact involve important change on role of national economic institution, namely the central bank lose its monetary and exchange rate policies. Therefore the attending to currency area is very challenging for national economies as those one must show flexibility to adapt to different economic episodes and also to prove a high degree of homogeneity as compared with other economies from euro zone. In accordance with the two requirements mentioned before, the adoption of euro may pose, after case, costs and benefits to the national economies. Considering the framework of *Optimal Currency Area*, on which is based this analysis, underline that the costs of joining to Euro Area depends on the degree of cycles synchronization, i.e. the higher the synchronization between the countries that intend to adopt the unique currency and the Euro Area the lower the costs of giving up the monetary policy. This paper, as the recent tendencies from academic and policy making fields, the assessment of costs, respectively the benefits resulted from attending the monetary union throughout the similarity of responses to macro shocks, as well as the similarity of demand and supply shocks between Romania and Euro Area. For these purposes we called a Bayesian approach of Structural Vector Auto-Regression methods (SVAR) in order to trace out the impulse-response function and structural matrix. As compared with standard approaches of SVAR models, from technical viewpoints the use of Bayesian econometrics has two main advantages: offers the possibility to incorporate priorly the own knowledge on the underlined debate and secondly the mode of computing error bands using percentiles provide a more accuracy measure of the uncertainty in regard with obtained impulse-response functions. As methodological backgrounds, we based our analysis on *Minnesota Prior Approach* and the approach proposed by Bayoumi and Eichengreen (1993). As compared with the Bayoumi and Eichengreen's approach, in this paper we used a short-run identification scheme of the structural shocks from

VAR model. Another important departure from the analysis of Bayoumi and Eichengreen and other studies conducted for Romanian case is that here we used the evolution of consumption and consumption prices. This aspect could be argued by the *Backus-Kehoe-Kydland puzzle* that underlines a lower degree of correlation across countries between consumption evolutions as compared with GDP's evolutions. Taking into account the high share of consumption in the use of GDP, in this paper we used the evolution of consumption and its deflator as proxy for the business cycle analysis within an IS-LM-AD-AS theoretical framework.

2. Background of the research

2.1. Bayesian econometrics

Modern Bayesian econometric tool of analysis gives to researcher the opportunity to incorporate its own knowledge about observed phenomena into the estimation process of different models of regression. More exactly, in the classical econometrics, the research conducts its analysis on the relation between two variables using the likelihood function in order to estimate the vector of coefficients and variance of error term. Instead the Bayesian econometrician call the *Bayes Law* from the probability theory and writes the posterior distribution, which in the case of a simple regression model $Y_t = \beta X_t + \varepsilon_t$, with $\varepsilon_t \sim N(0, \sigma^2)$, takes the following form:

$$\mathrm{H}\left(\frac{1}{\sigma^2}, \beta|\mathrm{Y_t}\right) = \mathrm{P}\left(\beta, \frac{1}{\sigma^2}\right) \times \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^{\mathrm{T}}} \exp\left(-\frac{\left(\mathrm{Y_t}-\beta\mathrm{X_t}\right)'\left(\mathrm{Y_t}-\beta\mathrm{X_t}\right)}{2\sigma^2}\right) \quad (1)$$

The above relation states that Bayesian econometrician updates its prior beliefs postulated under the form of probabilities distribution for coefficients β and variance of error term $\frac{1}{\sigma^2}$ (the first term from the right hand of equation) with the information on observed data provided by the likelihood function (the second term). In order to infer the two sets of parameters, the researcher has to isolate the marginal distribution of β and $\frac{1}{\sigma^2}$ from the posterior distribution which in fact is a joint distribution:

$$\Omega(\beta|\mathrm{Y_t}) = \int_0^{\infty} \Omega\left(\frac{1}{\sigma^2}, \beta|\mathrm{Y_t}\right) d\frac{1}{\sigma^2} \; ; \Omega\left(\frac{1}{\sigma^2}|\mathrm{Y_t}\right) = \int_0^{\infty} \Omega\left(\frac{1}{\sigma^2}, \beta|\mathrm{Y_t}\right) d\beta \quad (2)$$

The approximation of both the joint distribution and the two marginal distributions it is possible with the help of *Gibbs Sampling* algorithm which is a numerical procedure of Markov Chain Monte Carlo class that draws from conditional distribution. But we will describe in detail this method particularizing on the case of multivariate models as VAR. Therefore in the following lines we will define a bivariate unrestricted VAR with two lags in matrix form:

$$\begin{bmatrix} \mathrm{Y_t} \\ \pi_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{1,11} & a_{1,12} \\ a_{1,21} & a_{2,22} \end{bmatrix} \begin{bmatrix} \mathrm{Y_{t-1}} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} a_{2,11} & a_{2,12} \\ a_{2,21} & a_{2,22} \end{bmatrix} \begin{bmatrix} \mathrm{Y_{t-2}} \\ \pi_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (3)$$

where $\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$ coming from the variance-covariance matrix $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ and $\begin{bmatrix} Y_t \\ \pi_t \end{bmatrix}$ is the vector of explained variables. Specialists from Federal Reserve of Minnesota pioneered by Litterman (1979) used the Bayesian framework to incorporate beliefs that explained variables are coming from different types of forms of a Markov first order autoregressive model, namely a *Random Walk* or an AR model (from here the *Minnesota prior* approach). In their seminal work, they used such an approach in forecasting macro variables considering priory that $\begin{bmatrix} a_{2,11} & a_{2,12} \\ a_{2,21} & a_{2,22} \end{bmatrix}$ equals 0 and $\begin{bmatrix} a_{1,11} & a_{1,12} \\ a_{1,21} & a_{2,22} \end{bmatrix}$ is a diagonal matrix cu $a_{1,11}$ and $a_{2,22}$ equal 1. The mean of the Minnesota prior distribution for VAR model coefficients are stored in a vectorized form of the two equations of the model: $\bar{\beta}$ = $[0 \; a_{1,11} \; 0 \; 0 \; 0 \; 0 \; 0 \; a_{2,22} \; 0 \; 0]'$, with $P(\beta) \sim N(\bar{\beta}, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ is the prior for covariance matrix. Instead, for the setting of prior variance, the researcher uses some special parameters, called *hyperparameters*, designed to control the dynamic of the VAR system throughout the standard deviation of the parameters. One of the main disadvantages of Minnesota prior is the error covariance matrix Σ from needs to be diagonal and fixed. In order to pick this limit, there will be considered that the two sets of interest parameters come from different distributions, namely β comes from Normal distribution, while Σ comes from an inverse Wishart distribution[1]: $P(\beta|\Sigma) \sim N(\bar{\beta}, \Sigma \times \bar{\Omega})$ and $P(\Sigma) \sim W^{-1}(\Psi, \nu)$. This approach is found in the literature under *natural conjugate prior* or Normal Inverse Wishart prior and unlike the Minnesota prior method permits to the researcher to treats in the same manner the prior variance of lagged variance of both the explained and explanatory variables. As a formal bridge to Structural VAR models, many economic applications suppose some restriction on short or long term horizon of time in regard with response of some variables to shocks in other variable, i.e for example the hypothesis of sticky (or almost sticky) prices or lower bound phenomena in Quantitative Easing framework. From these considerations, the Independent Normal Inverse Wishart prior method involves the separated setting of prior for β and Σ, such as $P(\beta|\Sigma) \sim N(\bar{\beta}, \Omega)$ and $P(\Sigma) \sim W^{-1}(\Psi, \nu)$. The Independent Normal Inverse Wishart prior formulates the following hyperparameters to control how tight are manipulated the coefficients and covariance matrix of VAR model:

- λ_0 is used to control the degree of tightness for the prior on covariance matrix;
- λ_1 is used to control the degree of tightness for the prior on the first lag's coefficients;
- λ_3 is used to control the degree of tightness to zero for the prior on the coefficients of lag higher than 1;
- λ_4 represents the prior of the variance on constant variable.

Within this framework, as λ_1 and $\lambda_4 \to 0$, the related priors are shrunk to 0, while λ_3 increase the prior on the coefficients for lags higher than 1 tends to 0. As compared with standard Minnesota approach, the Independent Normal Inverse Wishart prior method supposes that $\lambda_2 = 1$, which means that coefficients for the lags of dependent variable and the other variables are treated similar [2]. Considering the above notations, Ω is obtained from the Kronecker product of Ψ and $\bar{\Omega}$ that will result in the following multiplication in a matrix form:

$$\begin{pmatrix} \left(\frac{\sigma_1}{\lambda_0}\right)^2 & 0 \\ 0 & \left(\frac{\sigma_2}{\lambda_0}\right)^2 \end{pmatrix} \otimes \begin{pmatrix} (\lambda_0\lambda_4)^2 & 0 & 0 & 0 & 0 \\ 0 & \left(\frac{\lambda_0\lambda_1}{\sigma_1}\right)^2 & 0 & 0 & 0 \\ 0 & 0 & \left(\frac{\lambda_0\lambda_1}{\sigma_2}\right)^2 & 0 & 0 \\ 0 & 0 & 0 & \left(\frac{\lambda_0\lambda_1}{2^{\lambda_3}\sigma_1}\right)^2 & 0 \\ 0 & 0 & 0 & 0 & \left(\frac{\lambda_0\lambda_1}{2^{\lambda_3}\sigma_2}\right)^2 \end{pmatrix} \quad (4)$$

Coming back to the Gibbs sampling algorithm, the main idea behind this procedure is that for a large number of iterations, the samples draws from conditional distribution converge to the real joint and marginal distribution of the underlined data. In the case of VAR system of equations, the following steps are required to be done to implement the Gibbs sampling algorithm in order to compute forecasts, impulse response analysis or variance decompositions. In the following lines we will detail the steps of the Gibbs sampling procedure:

- There are set priors for the VAR coefficients and variance according to multivariate Normal distribution $P(\beta|\Sigma) \sim N(\bar{\beta}, \Omega)$, respectively to inverse Wishart distribution $P(\Sigma) \sim W^{-1}(\Psi, \nu)$.
- According to Hamilton (1991) and considering the conditional posterior distribution for VAR coefficients $\Omega(\beta|\Sigma, Y_t) \sim N(M^*, V^*)$, there are calculated parameters of N :

$$M^*_{(N\times(N\times P+1))\times 1} = \left(\Omega^{-1} + \Sigma^{-1} \otimes X_t'X_t\right)^{-1} \times \left(\Omega^{-1}\bar{\beta} + \Sigma^{-1} \otimes X_t'X_t\hat{\beta}\right) \quad (5)$$

[1] The inverse Wishart distribution it is a generalization of the inverse gamma distribution for the multivariate case.

[2] More exactly the standard Minnesota approach uses λ_2 to control the tightness related to variables less the dependent variable.

$$V^{*}_{(N\times(N\times P+1))\times(N\times(N\times P+1))}=\left(\Omega^{-1}+\Sigma^{-1}\otimes X_t^{'}X_t\right)^{-1} \quad (6)$$

Once M^{*} and V^{*} are determinated, for the sampling of VAR coefficients is used the following procedure:

$$\beta^{1}_{(N\times(N\times P+1))\times 1}=M^{*}_{(N\times(N\times P+1))\times 1}+\left[\tilde{\beta}_{1\times(N\times(N\times P+1))}+(V^{*})^{1/2}_{(N\times(N\times P+1))\times(N\times(N\times P+1))}\right] \quad (7)$$

where β^{1} denotes the first draw and $\tilde{\beta}$ it's a vector sampled from normal distribution. An important remark (advantage) about Bayesian estimation of VAR coefficients is that looking at relation *(5)* we can observe the vector M^{*} result from an average of prior mean $\bar{\beta}$ and estimate of maximum likelihood $\hat{\beta}$ (OLS in this case) weighted with the related variance of the two.

• According to Hamilton (1991) and considering the conditional posterior distribution for VAR variance $\Omega(\Sigma|\beta,Y_t)\sim W^{-1}(\tilde{\Sigma},\nu)$, the scale parameter of W^{-1} is calculated as following:

$$\tilde{\Sigma}=\Psi+\left(Y_t-\beta^{1}X_t\right)^{'}\times\left(Y_t-\beta^{1}X_t\right) \quad (8)$$

where β^{1} is the first draw from previous step and it was shaped as a $(N\times P+1)\times N$ matrix. In order to draw the covariance matrix from $W^{-1}(\Psi,\nu)$, the first step consists in drawing a $Z_{\nu\times n}$ matrix from the multivariate $N(0,\Psi^{-1})$. Then the draw of covariance matrix from W^{-1} is performed according to the formula: $\hat{\Sigma}=\left(\sum_{i=1}^{\nu}Z_iZ_i^{'}\right)^{-1}$

• The steps from 1 to 3 are iterated for *S* times after that are computed the empirical distribution of VAR coefficients and variance from the retained draws $\beta^{1}\dots\beta^{S}$ and $\Sigma^{1}\dots\Sigma^{S}$. In order to decide if *S* is large enough to determine the empirical distributions, there will be used the procedures proposed by Geweke (1991) that will be detailed in the following section.

A more efficient method to compute priors which is mainly indicate for large BSVAR models was proposed by Branbura *et al.* (2007).

$$\begin{bmatrix}Y_t\\ \pi_t\end{bmatrix}=\begin{bmatrix}a_{10}\\ a_{20}\end{bmatrix}+\begin{bmatrix}a_{1,11} & a_{1,12}\\ a_{1,21} & a_{2,22}\end{bmatrix}\begin{bmatrix}Y_{t-1}\\ \pi_{t-1}\end{bmatrix}+\begin{bmatrix}a_{2,11} & a_{2,12}\\ a_{2,21} & a_{2,22}\end{bmatrix}\begin{bmatrix}Y_{t-2}\\ \pi_{t-2}\end{bmatrix}+\begin{bmatrix}\varepsilon_{1t}\\ \epsilon_{2t}\end{bmatrix} \quad (9)$$

Looking backward to the bivariate VAR (2) defined above and there are considered dummy observations to implement the prior. For this purpose we consider a set of artificial data noted by $Y_{D,1}$ and $\Pi_{D,1}$ and the hyperparameters τ , λ_3 and λ_4 that controls for the overall tightness of the prior, while the other two parameters have the same role as in the case of Independent Normal Inverse Wishart prior approach. Given these data, for the implementation of priors on the coefficients for the first lag of y_t and π_t, where $\sigma_{1,2}$ denote the standard deviations from the least squares estimations of the AR parameters, there is proceeding as following:

$$Y_{D,1}=\begin{pmatrix}\left(\frac{1}{T}\right)\sigma_1 & 0\\ 0 & \left(\frac{1}{T}\right)\sigma_2\end{pmatrix} \qquad \Pi_{D,1}=\begin{pmatrix}0 & \left(\frac{1}{T}\right)\sigma_1 & 0 & 0 & 0\\ 0 & 0 & \left(\frac{1}{T}\right)\sigma_2 & 0 & 0\end{pmatrix} \quad (10)$$

Rewriting the VAR model using the artificial data:

and expanding the equation (11) we obtain the following relation:

$$\begin{pmatrix} \left(\frac{1}{\tau}\right)\sigma_1 & 0 \\ 0 & \left(\frac{1}{\tau}\right)\sigma_2 \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{\tau}\right)\sigma_1 a_{1,11} & \left(\frac{1}{\tau}\right)\sigma_1 a_{1,12} \\ \left(\frac{1}{\tau}\right)\sigma_2 a_{1,21} & \left(\frac{1}{\tau}\right)\sigma_2 a_{1,22} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{11}$$

Taking in the first equation $\left(\frac{1}{\tau}\right)\sigma_1 = \left(\frac{1}{\tau}\right)\sigma_1 a_{1,11} + \varepsilon_{1t}$ or $a_{1,11} = 1 - \frac{\tau\varepsilon_{1t}}{\sigma_1}$ and applying expected operator $E(a_{1,11}) = 1 - E\left(\frac{\tau\varepsilon_{1t}}{\sigma_1}\right)$=1, since $E(\varepsilon_{1t}) = 0$ meaning that the dummy variables imply a prior mean of 1 for $a_{1,11}$. Proceeding in the same manner for the variance of estimation, there is obtained relation $a_{1,11} = \frac{\tau^2 var(\varepsilon_{1t})}{\sigma_1^2}$. Analogue, for the second equation we obtain that $a_{1,12} \sim N(0, (\tau^2 var(e_1(2t)))/(\sigma_1^2))$.

In order to get prior for the coefficients on the second lags we use the following matrices:

$$Y_{D,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \Pi_{D,2} = \begin{pmatrix} 0 & 0 & 0 & \left(\frac{1}{\tau}\right)\sigma_1 2^{\lambda_3} & 0 \\ 0 & 0 & 0 & 0 & \left(\frac{1}{\tau}\right)\sigma_2 2^{\lambda_3} \end{pmatrix} \tag{12}$$

Calculating on the same procedure these dummy variables imply a prior mean of 0 for the second la . In order to control the prior on the constants in the model we use the following matrices

$$Y_{D,3} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \Pi_{D,3} = \begin{pmatrix} \frac{1}{\lambda_4} & 0 & 0 & 0 & 0 \\ \frac{1}{\lambda_4} & 0 & 0 & 0 & 0 \end{pmatrix} \tag{13}$$

For setting prior on the error covariance matrix we use the following matrices:

$$Y_{D,4} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \qquad \Pi_{D,4} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{14}$$

The priors are implemented by adding all these dummy observations to the actual data sets: $Y^* = [Y; Y_{D,1}; Y_{D,2}; Y_{D,3,}; Y_{D,4}]$, $\Pi^* = [\Pi; \Pi_{D,1}; \Pi_{D,2}; \Pi_{D,3,}; \Pi_{D,4}]$. For a VAR model with *N* variables and *P* lags, the general form to implement prior dummies proposed by Branbura *et al.* (2007) is:

$$Y_D = \begin{pmatrix} \frac{diag(\chi_1\sigma_1 \dots \chi_N\sigma_N)}{\tau} \\ 0_{Nx(P-1)xN} \\ \dots\dots \\ diag(\sigma_1 \dots \sigma_N) \\ \dots\dots \\ 0_{1xN} \end{pmatrix} \tag{15}$$

$$\Pi_D = \begin{pmatrix} J_P \otimes diag(\sigma_1 \dots \sigma_N) 0_{NPx1} \\ 0_{NxNP} \quad 0_{Nx1} \\ \dots\dots\dots \\ 0_{1xNP} \quad \lambda_3 \end{pmatrix}$$

where χ_i are the prior means for the coefficients on the first lags of the dependent variables and $J_P = diag(1..P)$. If the variables in the VAR have a unit root this information could reflected via a prior the incorporates the belief that coefficients on lags of the dependent variables sum to 1:

$$Y_{D,5} = \begin{pmatrix} \gamma\mu_1 & 0 \\ 0 & \gamma\mu_2 \end{pmatrix} \qquad \Pi_{D,5} = \begin{pmatrix} 0 & \gamma\mu_1 & 0 & \gamma\mu_1 & 0 \\ 0 & 0 & \gamma\mu_2 & 0 & \gamma\mu_2 \end{pmatrix} \tag{16}$$

Where μ_1 is the sample mean of y_t and μ_2 is the sample mean of π_t These dummy observations imply prior means of the form $a_{1,ii} + a_{2,ii}$ =1 and γ controls the tightness of the prior. As $\gamma \to \infty$ the prior is implemented more tightly. Banbura *et al* (2007) underline the fact that these dummy observation for N variable VAR with P lags are given by:

$$Y_D = \frac{\mathrm{diag}(\chi_1 \mu_1 \cdots \chi_N \mu_N)}{\lambda} \qquad \Pi_D = \left(\frac{(1,2..P) \otimes \mathrm{diag}(\chi_1 \mu_1 \cdots \chi_N \mu_N)}{\lambda} \;\; 0_{Nx1} \right) \quad (17)$$

Where $\lambda = \frac{1}{\gamma}$ and $\mu_i, i = 1,\dots,N$ are sample means of each variable included in the VAR

Once the prior are implemented, the next task in SVAR type analysis is to compute specific sensitivity analysis as impulse response function or historical decomposition. Firstly we consider a structural SVAR (2) writing in a scalar form

$$By_t = B_1 y_{t-1} + B_2 y_{t-2} + \varepsilon_t \quad (19)$$

With the reduced form underlying VAR

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + e_t \quad (20)$$

Where $e_t = B^{-1}\varepsilon_t$.

In the following lines we will describe the steps of sign restriction approach to compute impulse response functions. The first hypothesis is that we have ordered the variables that appears in VAR in a recursive way. The VAR residuals, $\hat{e}_t$ are related to the structural residuals as $\hat{e}_t = B^{-1}\hat{\varepsilon}_t$. Having the matrix K that contains the estimated standard deviations of the ε_t on the diagonal and zero in rest, then $\hat{e}_t = B^{-1}KK^{-1}\widehat{\varepsilon_t}_t = T\delta_t$, where $\delta_t = K^{-1}\widehat{\varepsilon_t}_t$ has unit variances. The second hypothesis is that having a square matrix S such as $S'S = SS' = I$ then $\hat{e}_t = TSS'\delta_t = T^*\delta_t^*$. The conclusion is that a new set of estimated shocks δ_t^* have been obtained that have also the property that their covariance matrix is $E(\delta_t^*\delta_t^{*'}) = SE(\delta_t\delta_t{}')S' = I$. In the end it has been obtained a new combination of the shocks δ_t^* that have the same covariance matrix as δ_t but have a different impact upon e_t, hence y_t.

$$\begin{bmatrix} \delta_{1t}^* \\ \delta_{2t}^* \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \delta_{1t} \\ \delta_{2t} \end{bmatrix} \quad (21)$$

Ramirez *et al*. (2010) defined an efficient algorithm to find a A_0 matrix consistent with impulse response of a certain sign consistent with theory in the case of exactly identified models. Considering a VAR (2) model an example

where $var\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix} = \Sigma$, y_t is output growth, π_t is GDP deflator and i_t is the interest rate. The goal is to calculate the impulse response to a specific structural shock, for example a monetary policy shock. An unexpected change in the monetary policy field that determine a positive shock on interest rates will determine a decrease of y_t and π_t. The algorithm contains 5 steps.

- The first step of the algorithm is to define an NxN matrix W from the standard normal distribution.
- Then calculate the matrix S from QR decomposition of W (note that S is orthonormal).
- The third step is to calculate the Choleski decomposition of the current draw of $\Sigma = \widetilde{A}_0'\widetilde{A}_0$.
- The next step is to calculate the candidate A_0 matrix as $A_0 = S\widetilde{A}_0$. Because $S'S = I$ implies that $A_0'A_0 = \Sigma$. The candidate A_0 will have the following form

$A_0 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, where the last row corresponds denotes the interest rate shock. The first check is to see if $a_{31} < 0$ and $a_{32} < 0$ and $a_{33} > 0$ that means a contemporaneous increase in interest rate will lead to a fall in output and inflation. If these are the right sign of contemporaneous relationships among interest variables, the algorithm stops and the candidate matrix A_0 is draw to compute impulse response functions. Otherwise the algorithm will try with other W matrix.

- The step 5 is to repeat the other previous four steps for every retained Gibbs draw.

Additionally to the two previous applications of Bayesian econometrics, this type of methods was also extended to estimate the parameters of state-space models. The benchmark way to estimate the models with unobserved components is the Kalman filter. Consider a standard state-space model, where e_t and v_t are i.i.d shock, $e_t \sim N(0, R)$, $v_t \sim N(0, Q)$ and t = 1,…, T:

$$Y_t = x_t\beta_t + e_t \quad (19)$$
$$\beta_t = \mu + F\beta_{t-1} + v_t \quad (19)$$

The first equation from the state-space model is the measurement equation, while the second one represents the transition equation that describe dynamic of the state variables. The Kalman filter is a recursive procedure that consists in the following steps:

a. Initializing the state variable

In the first step of the procedure are set starting values for the state variable and its variance based on the information sets available upon *t* – 1 moment:

$$\beta_{(t-1|t-1)} = \beta_{00} \qquad (23)$$
$$p_{t-1\backslash t-1} = p_{00}$$

b. Prediction

According to available information there is computed an optimal prediction of dependent variable, where $\eta_{t\backslash t-1}$ and $f_{t\backslash t-1}$ is the prediction error and its variance:

$$\beta_{(t|t-1)} = \mu + F\beta_{(t-1|t-1)} \qquad (24)$$
$$p_{t-1\backslash t-1} = Fp_{t-1\backslash t-1}F' + Q$$
$$y_t = X_t\beta_{(t|t-1)}$$
$$\eta_{t-1\backslash t-1} = Y_t - y_t$$
$$f_{t\backslash t-1} = X_t p_{t\backslash t-1} X_t' + R$$

c. Updating

Last step of the procedure consists in update the previous prediction and its error with the new information set through the so called "Kalman gain" denoted by K_t:

$$K_t = p_{t\backslash t-1}X_t' + f_{t\backslash t-1}^{-1} \qquad (25)$$
$$\beta_{t\backslash t=}\beta_{(t|t-1)} + K_t\eta_{t\backslash t-1}$$
$$p_{t\backslash t} = p_{t\backslash t-1} - K_t X_t p_{t\backslash t-1}$$

In order to implement a Bayesian estimation for the state variable we will call the Carter and Kohn (1994) algorithm. Considering the state variable $\tilde{\beta}_T = [\beta_1, \beta_2, \dots \beta_T]$, the conditional distribution is

$$H(\tilde{\beta}_T) = H(B_T, \tilde{Y}_T)\prod_{t=1}^{T-1} H(\beta_{t\backslash\beta_{t+1}}, \tilde{Y}_t) \qquad (26)$$

Therefore we define above the joint density of the state variable for t=1...T can be factored into the terms $H(B_T, \tilde{Y}_T)$ and $H(\beta_{t\backslash\beta_{t+1}}, \tilde{Y}_t)$ where $\tilde{Y}_t = [Y_1, \dots Y_T]$. As the state space model is linear and Gaussian this implies that $\beta_T \sim N(\beta_{T\backslash T}, p_{T\backslash T})$ (27)

$$\beta_{t\backslash\beta_{t+1}} \sim N\left(\beta_{t\backslash t,B_{t+1}}, p_{t\backslash t,B_{t+1}}\right)$$

The mean and variance in $\beta_T \sim N(\beta_{T\backslash T}, p_{T\backslash T})$ is given by Kalman filter at time t=T. In case of the mean and variance in $N(\beta_{t\backslash t,B_{t+1}}, p_{t\backslash t,B_{t+1}})$ it requires the updating equations derived by Carter and Kohn

$$\beta_{t\backslash t,B_{t+1}} = \beta_{t\backslash t} + p_{t\backslash t}F'(Fp_{t\backslash t}F')^{-1}(\beta_{t+1} - \mu - F\beta_{t\backslash t}) \qquad (28)$$

$$p_{t\backslash t,B_{t+1}} = p_{t\backslash t} + p_{t\backslash t}F'(Fp_{t\backslash t}F' + Q)^{-1}Fp_{t\backslash t} \qquad (29)$$

Empirical evidence showed that shape, sign and magnitude of impulse response functions evolves in time. For this purpose Cogley and Sargent (2002) propose a VAR model with time varying coefficients

$$Y_t = c_t + \sum_{j=1}^{P} B_{j,y}Y_{t-j} + v_t, VAR(v_t) = R \qquad (30)$$
$$\beta_t = \{c_t, B_{1,t} \dots B_{p,t}\}$$
$$\beta_t = \mu + F\beta_{t-1} + e_{t,}\ VAR(e_t) = Q$$

The most important feature of their model showed above as state-space form is that only parameters are allowed to fluctuate while the covariance matrix is fixed.

Calling again the Markov Chain Monte Carlo methods to implement a Bayesian estimation of the state-space form of the VAR model, the Gibbs sampling inference algorithm consists in the following steps.

- The first step is to fix priors for *R* and *Q* and starting values for the Kalman filter. The prior for Q comes from an inverse Wishart p(Q)~IW(Q_0, T_0). We split the initial sample in other two sample. The training sample of size T_0 is used to estimate a standard fixed coefficient VAR using OLS such that $\beta_0 = (X'_{0t}X_{0t})^{-1}[(X]'_{0t}Y_{0t})$, while the coefficient of covariance matrix is by $p_{0\backslash 0} = \Sigma_0 \otimes (X'_{0t}X_{0t})^{-1}$. The scale matrix Q_0 is set equal to $p_{0\backslash 0}xT_0x\tau$, where τ is scaling factor choosen by researchers in order to control the importance of Q_0 The prior degrees of freedom are set equal to T_0.

- The second step consists in sampling $\widetilde{\beta_t}$ conditional on R and Q from its conditional posterior distribution via the Carter and Kohn algorithm.

- The next step is to sample Q from its condition posterior distribution. This, conditioned on $\widetilde{\beta_t}$ is inverse Wishart with scale matrix $\left(\widetilde{\beta_t}^{1} - \tilde{\beta}_{t-1}{}^{1}\right)'\left(\widetilde{\beta_t}^{1} - \tilde{\beta}_{t-1}{}^{1}\right) + Q_0$ and degrees of freedom $T + T_0$ where T is the whole sample's length and $\widetilde{\beta_t}^{1}$ denotes the first draw of the state variable.

- The step 4 is to sample R from the conditional posterior distribution. Conditional on the draw $\widetilde{\beta_t}^{1}$ the posterior of R is inverse Wishart with scale matrix

- The last step is to repeat steps 2 to 4 M times use the last L draws for inference.

2.2 Theoretical framework

This paper aims to analyze the structural degree of synchronization between the Euro Area and the Romanian economies. For this purpose we called a SVAR framework with some certain features. Therefore, in the following lines we will detail the particularities of the approaches that we used to compute impulse response functions to different types of shocks.

The benchmark theoretical framework is a small New-Keynesian model that contains three main equations:

$$Y_t = \alpha_{1Y}Y_{t-1} + \beta_{1Y}E_t(Y_{t+1}) + \gamma_{1i}(i_t - E_t(\pi_{t+1})) + \varepsilon_{vt} \quad (31)$$
$$\pi_t = \alpha_{2\pi}\pi_{t-1} + \beta_{2\pi}E_t(\pi_{t+1}) + \gamma_{2Y}Y_t + \varepsilon_{\pi t} \quad (32)$$
$$i_t = \alpha_{3\pi}i_{t-1} + \beta_{3\pi}E_t(\pi_{t+1}) + \gamma_{3Y}Y_t + \varepsilon_{it} \quad (33)$$

which is also a workhorse for the DSGE models of analysis. In fact, the policy strategy stems in the estimation of the New-Keynesian model under a DSGE fashion. Once the model was estimated, there are set the relationships among variables through the sign of impulse response function. In addition, under the hypothesis of no serial correlation between shocks, the solution $z'_t[Y_t\,\pi_t\,i_t]$ of the maximum likelihood estimates of DSGE parameters can be put in a VAR form. In a strictly manner, the DSGE models can be put in VAR forms under specific form for VAR and under some certain stochastic implication. But in this paper, we will draw only the predictability of DSGE model under a roughly SVAR form as:

$$y_t = z'_{t-1}\gamma_Y + \beta_{Y\pi}\pi_t + \beta_{Yi}i_t + \varepsilon_{vt} \quad (34)$$
$$\pi_t = z'_{t-1}\gamma_\pi + \beta_{\pi Y}Y_t + \beta_{\pi i}i_t + \varepsilon_{\pi t} \quad (35)$$
$$i_t = z'_{t-1}\gamma_i + \beta_{iY}Y_t + \beta_{i\pi}\pi_t + \varepsilon_{it} \quad (36)$$

In the above SVAR system, ε_{yt}, $\varepsilon_{\pi t}$ and ε_{it} denotes the (structural) aggregate demand shock, Cost Push shock (which is a supply side shock), respectively the monetary shock, while the relationships predicted by DSGE literature are:

Table 1. Sign restrictions predicted by DSGE models in a New-Keynesian framework.

Model's Variable/Structural Shock	Aggregate Demand	Cost Push	Monetary
y_t	+	-	-
π_t	+	+	-
i_t	+	+	+

Thus, given the predictions provided by DSGE literature, we will incorporate this information in a SVAR style model with sign restriction. In the current work, we prefer the sign restrictions in place of pure timing restriction or recursive Choleski computation of impulse response function, even that as we explained in previous section our approach is based on exact identification that starts from a Choleski scheme. Now the question arises how to identify the structural shocks given the above framework. Within the literature and applied practices exists several ways to achieve this task, ranging from the standard form of New-Keynesian model to models that incorporate addition equations in order to account for different types of supply shocks. From this view point arises some important questions that we will detail furthermore. Given the main objective of the paper, in the second timeframe we were interested in designing an appropriate benchmark that would allow me further to develop an appropriate framework to fully analyze and explain the main differences between Romania and Euro Area economies in terms of structural convergence. Thus the benchmark model of analysis used in this work represents a preliminary step within a larger project of analysis. To use efficiently this benchmark in order to obtain information about some important issues for this kind of analysis we addressed several limits (*puzzles*) the literature and empirical evidences underlined in regard with SVAR practices. The related *puzzles* are: *multiple models problem* (Fry and Pagan, 2005), *multiple shocks problem* (Fry and Pagan, 2007), *Sims' prices puzzle* (Sims,1992), *Galli's puzzle related to RBC models* (Galli,1999) and *Christiano's evidence on the importance of data form*[3](Christiano *et al.* 2003).

Now we define the model that we used and after will argue the chosen also in regard with the mentioned puzzle:

$$[\ a_{21} \quad a_{22} \quad 0\ ][\ \varepsilon_t^{\pi}\ ] = [\ e_t^{\pi}\ ] \tag{37}$$

where $A_0 = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ comes from Cholesky decomposition of the variance covariance matrix $A_0' A_0 = \Sigma$ This model is known in the literature as a monetary model because it is designed only to produce the response of GDP and inflation to a monetary shock which imply a reaction function from the central bank as response to contemporaneous developments from real and sector prices. Because this specification don't allow for a feedback effect from e_t^i, the model is designed only for the recovery of monetary shocks. For this reason, looking from a technical view point, the above system represents a semi-structural VAR model. Even that theoretical framework it allows for different interaction among structural shock, if we put the theoretical model under a recursive system form, we can identify only the monetary shock. But the use of sign restriction approach in the SVAR may relax the limit of recursive system.

In regard with the previous *puzzle*, we will run the hypothesis from this work, but looking as a preliminary step from a broader analysis. Relating to *multiple models problem*, we called the algorithm of Fry and Pagan (2005), called median target (MT) response. Given the limit of sign restriction methods provided by non-unique values of the rotation matrix's angle $\theta^{(k)}$, we are interesting in those responses that are most common. More exactly, after the obtained impulse response function from Markov Chain Monte Carlo methods, those ones are differentiated with their median and the sum of differences are stored in a vector $\varphi^{(l)}$. Finally, the MT consists in finding that impulse response (*l*) that produces *min*(MT)=$\varphi^{(l)\prime}\varphi^{(l)}$, obtaining in this manner a unique $\theta^{(l)}$. Regarding the *multiple shocks problem*, we used the underlined model mainly for the identification of a monetary shock as a policy objective. But to address this problem, we estimated a Cholesky style model with the same LRAS framework and we compared the results. Regarding the observation of Christiano *et al.* as response to Galli's *puzzle*, we did not change on interest series, allowing thus for a unit root. The other two series are expressed in percents as Fry and Pagan (2007) recommended in the case of sign restriction. The use of Bayesian framework allows testing for importance of rigidities[4]. For example, we can consider that prices don't respond with a lag or two to different shocks and then compare the results with the case of no restrictions. The *Sims' prices puzzle* would be analyzed a bit later based on results.

[3] Levels or differences.
[4] Galli underlined that rigidities (nominal) could be the main trigger of its puzzle.

Since the predictions are offered by DSGE literature, our work consists in looking at the scale and shape of impulse response functions, as well as both the scale and shape of those ones during the time.

3. Data and results

If we look at national statistics we can observe that consumption accounts for a much higher share from aggregate output in Romania as compared with Euro Area. Also during the time, the fluctuation of consumption share within output is higher for Romania. Adding to the previous two mentions the empirical fact of Backus-Kehoe-Kydland *puzzle*, for the main objective of the paper it is worthiness to used an additional specification of the benchmark model in which are used the growth rate of consumption and consumption deflator in place of GDP and GDP deflator.

In this paper we used quarterly data on real GDP, GDP deflator, real consumption[5], consumption deflator and 3M interbank reference rate (ROBOR and EURIBOR) for Romania and Euro Area (EA12) from the EUROSTAT database. The real data on GDP and consumption are expressed in terms of growth rates. All the time series used in this work are seasonally adjusted and adjusted with the number of working days. The period covered by current paper is 2000 Q2 – 2012 Q2. Given the usage of the Bayesian econometrics for involved analysis, the test for unit roots in data shows no significance. The information criteria analysis indicated the use of two lag VAR models

Regarding the prior means for state-space representations of the VAR model there we set $\mu = 0$ and *F*=1, that makes our model a linear Gaussian state space models that allows for a random walk behavior. The training sample used to set the other priors for the state-space representation contains 25 quarters, which means that the Time-Varying Parameters (TVP) VAR model was estimated for 22 quarters, if we take into account the 2 lags of the underlined models. The impulse response horizon is 20 quarters, while the shock was set to one standard deviation.

Even the involved framework is mainly designed to the identification of monetary shocks, we used the sign restriction from table 1 and estimated the IRFs also for shocks that come from the other two structural sources. The use of short-term interest in place of monetary reference rates relaxes the purpose of initial semi-structural VAR if we take into account that in setting quotes for interbank rates, investors incorporates their risk perceptions in regard with supply side risks and in the second timeframe with demand risks. The models were estimated for the version with GDP and GDP deflator in the 3 variables SVAR, as well as for the version with consumption and its deflator.

The computed time varying IRFs for the 12 models are reported in Annex. In the first timeframe, we are interested in analyzing the IRFs for monetary shocks. Obtained results showed that in the case of GDP version (models M1 and M2), while the responses of real growth rate of GDP to a monetary shock translated through an one percent increase in the 3M interbank rates are closed in both cases, the response of GDP deflator to the same shock was around – 1.3 % for Romania, as compared with values ranging between -0.45 % and -0.5 % for EA economy. Instead, in second version of the estimates, the computed IRFs for Romanian economy show a contraction of -0.65 % for consumption growth rate (as compared with around -1.5 % for EA), respectively of -1,9 % for consumption deflator (as compared with around -0.7 % for EA). Except in the case of Romanian economy (consumption version) where the behavior of responses of consumption growth rate and consumption deflator recorded significant changes in the last two years, the others IRFs post low fluctuations in time. On the other side, the *Sims' price puzzle* was present in all the responses of prices indicator to a monetary shock, less the case of model with GDP version data for EA economy. These puzzles were more evident for the models in which we used consumption related data. Also we can observe that for GDP version of the models, the impact of the shock on GDP growth rate and its deflator is more persistent in the case of EA. In the same timeframe, the persistence of the socks in the case of consumption version models is lower than in the GDP version.

On the other side, analysis of interest variables to other types of structural shocks is important especially for two reasons. The first one is focusing on the differences of the responses across two version of data used. Even that for policy purposes, the information provided by the response to the other two types of struc-

[5] Household and NPISH final consumption expenditure.

tural shocks is limited[6], we can observe that responses of Romanian variables to aggregate demand, respectively Cost Push shocks post asymmetric pattern in comparison with IRFs computed for EA case. Secondly, this exercise is important to understand how we can exploit the information contained by such a framework of analysis in regard with structural shocks. For this purpose we estimated a simple static Cholesky identification scheme, for the two versions of data. There were estimated two version of the Cholesky identification scheme, according to the ordering of variables: $z_t'[Y_t \, \pi_t \, i_t]$ and $z_t'[\, \pi_t \, Y_t \, i_t]$. In this manner we addressed the multi shock problem. The results were compared with those obtained from the Bayesian approach with sign restriction and time varying parameters.

4. Conclusions

In this paper we involved a study of structural convergence between Romanian and Euro Zone economies from the view point of synchronization in responses to shocks. For this purpose we called a Bayesian framework in which we estimated a time-varying parameters VAR model. For the identification of structural shocks we started from semi-structural VAR in which we incorporated the standard predictions of DSGE literature for a New-Keynesian model. For several purposes mentioned in the paper, we used two versions of data, replacing GDP and GDP deflator from a standard approach with consumption and its deflator. In this paper we were mainly interested for the response of interest variables to a monetary shock for policy purposes and in a second timeframe for the responses to other types of shocks.

The estimation of response for monetary model showed different structures of economies. The obtained results relieve the fact, for both data sets the structure of Romanian economy is different in comparison with EU economy. The analysis of cost push shock and aggregate demand shock are more likely to be informative and even though the model doesn’t allow identifying other shocks (naïve approach) the structure of economies looked different.

Acknowledgements

This work was co-financed from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013; project number POSDRU/107/1.5/S/77213 „Ph.D. for a career in interdisciplinary economic research at the European standards”.

References

1. Banbura, Marta, Domenico Giannone and Lucrezia Reichlin, 2007, Bayesian VARs with Large Panels, CEPR Discussion Papers 6326, C.E.P.R. Discussion Papers.

2. Benati, Luca and Haroon Mumtaz, 2006, US evolving Macroeconomic Dynamics: A structural investigation., Mimeo European Central Bank.

3. Canova, Fabio, 2007, Methods for Applied Macroeconomic Research, Princeton University Press, Princeton.

4. CARTER, C. K. and R. KOHN, 1994, On Gibbs sampling for state space models, Biometrika 81(3), 541.553.

5. Casella, George and Edward I. George, 1992, Explaining the Gibbs Sampler, The American

6. Cogley, Timothy and Thomas J. Sargent, 2002, Evolving Post-World War II U.S. Inflation Dynamics, NBER Macroeconomics Annual 2001, Volume 16, NBER Chapters, National Bureau of Economic Research, Inc, pp. 331.388.

7. Fry, Renee and Adrian Pagan, 2007, Some Issues in Using Sign Restrictions for Identifying Structural VARs, NCER Working Paper Series 14, National Centre for Econometric Research.

8. Hamilton, J. D., 1994, Time series analysis, Princeton University Press, Princeton.

9. Jacquier, E, N Polson and P Rossi, 2004, Bayesian analysis of stochastic volatility models, Journal of Business and Economic Statistics 12, 371.418.

[6] As we can observe from the reported plots, our models ambiguously meet the sign restrictions from the table, that could means the shocks were not properly identified.

10. Kim, C-J. and C. R. Nelson, 1999, State-space models with regime switching, MIT Press, Cambridge, Massachusetts.

11. Negro, Marco Del and Frank Schorfheide, 2004, Priors from General Equilibrium Models for VARS, International Economic Review 45(2), 643.673.

12. Primiceri, G, 2005, Time varying structural vector autoregressions and monetary policy, The Review of Economic Studies 72(3), 821.852.

13. Ramirez, Juan Rubio, Daniel Waggoner and Tao Zha, 2010, Structural Vector Autoregressions: Theory of Identication and Algorithms for Inference, Review of Economic Studies 77(2), 665.696.

14. Sims, Christopher A and Tao Zha, 1998, Bayesian Methods for Dynamic Multivariate Models, International Economic Review 39(4), 949.68.