



DOCUMENT DE RECHERCHE

EPEE

CENTRE D'ETUDE DES POLITIQUES ECONOMIQUES DE L'UNIVERSITE D'EVRY

Bayesian Analysis of DSGE Models: A Panel Approach

Tareq SADEQ

08 - 03

Bayesian Analysis of DSGE models: A Panel Approach*

Tareq Sadeq[†]
EPEE, Université d'Evry Val d'Essonne

Abstract

This methodological paper in Bayesian econometrics applied to macroeconomic models, presents first the Bayesian estimation of DSGE models, then illustrates the effects of the estimation for a panel of countries, assuming some homogeneity across countries. Under some specifications, the results exhibit more precision about parameters in a panel approach. The gain of precision is important using a panel of small samples of data.

1 Introduction

Bayesian methods in econometrics are increasingly used by researchers and policy-makers. The Bayesian econometrics is simply based on the Baye's rule of probability. This probability helps us to make any inference about parameters or about the model: Test of hypotheses, prediction, model choice,... etc.

This recent approach of econometrics has known many recent expansions. Klein's "*Textbook of Econometrics*" published in 1954, only presents some discussion of the Bayesian theory. Arnold Zellner's book "*An Introduction to Bayesian Inference in Econometrics*" published in 1971 is considered as the first book in Bayesian econometrics. Theil (1978) included a chapter about Bayesian inference and random rational behavior, in which he explained the Baye's theorem and provided some applications. Until the end of the 1980's very few applications of the Bayesian methods were studied. For instance, Zellner published

*I would like to thank my Ph.D supervisor Ferhat Mihoubi for all the useful and helpful comments. I thank also Michel Juillard and Stéphane Adjemian from CEPREMAP for permitting me to participate in some presentations of Dynare group, in addition to all the participants of the EPEE seminar.

[†]tareq.sadeq@univ-evry.fr

an article in 1986 on the Bayesian analysis and prediction using asymmetric loss functions. Poirier (1995) underlies in his book *"Intermediate Statistics and Econometrics: A Comparative Approach."* Bayesian and frequentist methods, but does not discuss models used by economists beyond regression.

The Landon-Lane's (1998) article was the first in evaluating macroeconomic models using bayesian methods, evaluating a Real Business Cycle model using likelihood method, but incorporating parameter uncertainty in the model evaluation through the use of a prior probability distribution. Then, the evaluation of RBC models and general equilibrium models has seen an increasing number of publications. These developments of Bayesian methods in the evaluation of macroeconomic models, notably in the Dynamic Stochastic General Equilibrium models, brought the DYNARE toolbox, which is a group of programs under MATLAB for solving and estimating these models using the Bayesian methods. DYNARE is now one of the most used toolbox by economists and researchers on policy-making.

The Bayesian econometrics, recently developed, takes into account the researcher's belief about parameters. On the contrary, the classical econometrics extracts all the information about parameters from data and does not allow for any prior belief to influence or to change the belief about parameters. Data cannot always give the appropriate information about the parameters for many possible reasons. It may be because of short data samples or because of the complicate structure of the model to estimate. However, using the Bayesian approach, we can give a weight to our belief about the parameters values, or to the theoretical calibrated values of the model's parameters.

In dynamic stochastic general equilibrium models (DSGE), we need to estimate a complicate model with structural parameters as one block. Until a recent time, authors used to calibrate these parameters from data or from previous studies. The reason is that classical econometrics through non-linear estimation methods (Maximum likelihood, GMM, SMM,...etc.) may face identification problems due to the data issues mentioned above¹. As a result, using Maximum likelihood estimation, or other methods depending only on the information given by the data, we can get unlikely parameters values.

In policy-making process, dealing with uncertainty and risk management is the essence of decision-making. Planners should choose the parameters with least risk, this gives us a reason to consider parameters as random variables. Furthermore, when the decision-maker make a

¹These problems appear in DSGE models estimation, where parameters are considered as one block, due to co-movements of variables unexplicated by the data.

forecast of the economy, she needs to consider not only the future path of the economy, but also the distribution of possible outcomes of path of the economy. Then, decision-makers need to reach a judgement for their policy upon probabilities, risks, costs and benefits of the policy choice. The risk-management of the monetary or fiscal policy means dealing with probability distributions. If the decision-maker has a prior belief about the policy choice based on previous observations or constructed by theoretical models, then the risk-management process is an application of Bayesian analysis of the model. The decision-maker once have a look on the posterior distribution of the policy can construct an opinion or a judgement about what policy to adopt.

When decision-makers search for the model that represents the economy, they do not search only for the best model but also for possible models probability distribution that may represent the economy. The best model would be that of the highest probability. The Bayesian approach is both a risk analysis of the model and a strategy analysis for policy directed to achieve economic objectives, for example the objectives of the central bank in price stability and the sustainable economic growth.

When a policy-maker uses the Bayesian choice of policy-making, she uses its own first belief, before any empirical analysis, about the optimal strategy to adopt, with a probability distribution that we call a prior. The question now is how to construct our belief about parameters (or policy-choice). First, the prior covers all the range that a parameter may be defined on. For example, if we know that the parameter takes only positive values, this information can be represented by a Gamma distribution prior. Second, the moments of the prior distribution can be obtained through several ways in order to represent the researcher's belief. The moments can be obtained subjectively from previous outcomes or other studies, or can be obtained objectively through other approaches that we develop later in the core of this paper.

The researcher's belief can then be updated by the observed data distribution, that is known as the likelihood distribution. The fundamental rule of the Bayesian approach is the Baye's rule of conditional probability, that combines the probability distribution reflecting the prior guess and the data distribution. The combination of the prior distribution and the data distribution, will give us a more clear idea about the economic characteristics (parameters, the policy choice, forecasting or the model choice). We can, thus, get a new distribution reflecting a *posteriori* belief. In order to analyze the new distribution, we need to use the posterior ditribution in inference.

Like in classical econometrics, we can take a two-dimension model

applying panel data econometrics. But we do not have to forget that the Bayesian econometrics are probability distribution analyses and not point-estimation like in classical econometrics. The Bayesian approach provides information about not only the parameters estimation, but also about the overall distribution of the parameters conditional to the observables. So, in a panel dimension, instead of point-estimating the individual specific effects, we can estimate the distribution of the individual random effect.

As in classical econometrics, a panel dimension implies an information gain, if we can prove some similarities across the individuals dimension. This feature can be a solution for short-data samples problem. In other words, increasing the dimension of the data can give a higher informative (less flat) likelihood distribution, if the similarities are well specified. The use of a panel dimension is useless, if the similarities are misspecified, since the posterior distribution risks to be with too many modes or to be a flat distribution.

Nowadays, none of the research studying on DSGE models considered a panel-data model, with similarities of parameters across the countries dimension. After scanning the literature of applied Bayesian econometrics on economic models, I found only two interesting articles that deal with panel-data models applied by economists. The first is the article of Koop et al. (2000), that applies Bayesian estimation on an output growth model for a panel of countries. The purpose of this study is to analyze cross-country sources of output growth. The authors specify an informative prior for each parameter associated to each country. The other article is a working-paper of the ECB, by Canova et al (2004) that applies the Bayesian inference on a dynamic panel VAR model for the G-7 countries, assuming interdependence between countries.

Similarities of some parameters in DSGE models can be specified if we apply them to countries located in the same region, or to those who could have some similarities in consumption dynamics or institutional characteristics.

Similarity specification about only a part of parameters allows us to test the heterogeneity of other parameters. However, the heterogeneity in the Bayesian analysis is different from the classical approach, we test the parameters intervals rather than the point-estimates. In the Bayesian analysis, we compare the uncertainty of each parameter, it cannot resume only to the distance between point-estimates.

The priors choice has a big role in the Bayesian estimation process on a Panel data model, since priors must take into account the similarities and differences across countries or individuals. As a result, any misspecification about parameters similarities can affect the posterior distribu-

tion, and thus the risk-management and the decision-making process.

The paper is organized as follows: In section 2 we introduce a general presentation of the Bayesian approach. We focus on priors and their choice process, posterior distributions and their computation for a DSGE model. Section 3 presents Bayesian analysis in panel data DSGE models, and some assumptions about the model and the shocks. Section 4 presents a simple DSGE model in a closed economy, as a Data Generating-Process (DGP). Then the model is estimated on the simulated data in different scenarios, where we compare parameters probability distributions under similarity specification. Finally, we conclude at the end of this paper.

2 A general presentation of the Bayesian approach

In the frequentist point of view, there is a true value for each parameter and we try, on the basis of the information provided by the data, to estimate them. On the other hand, the Bayesian approach considers the parameter as a random variable following a distribution. The Bayesian approach is fundamentally based on the Baye's rule

$$P(\beta | y) = \frac{P(y | \beta) P(\beta)}{P(y)} \quad (1)$$

The distribution, denoted by $P(\beta)$, is called the Prior probability distribution of the parameter β , it reflects the belief of the researcher on that parameter. The conditional distribution of the parameter β , denoted by $P(\beta | y)$, is the updated distribution of the parameter after extracting information from the data, this updated distribution is called the posterior distribution. While the conditional distribution of the data y to the parameter value is just the conditional likelihood function and $P(y)$ denotes the marginal likelihood function. The ratio between the posterior and the prior can tell us how much the data can be informative about the parameter. Since the marginal likelihood function $P(y)$ is constant for all the values of β , the posterior density function $P(\beta | y)$ is proportional to the numerator, the multiplication of the prior and the conditional likelihood, that we'll call the posterior kernel $\check{P}(\beta | y)$. Hereafter, the posterior density function will be approximated by a kernel density function, using sampling methods, like the Metropolis-Hastings algorithm, that combine the prior distribution and the information present in the data given by the likelihood distribution. So, the shape of the distribution of the parameter is updated by the shape of the likelihood function.

$$P(\beta | y) \propto \check{P}(\beta | y) = P(y | \beta) P(\beta) \quad (2)$$

This method works if all the variables are observable in the data, but usually this is not the case in DSGE models, we have unobservable variables such as capital stock, rental rate of capital, shocks values. To estimate the model using the Bayesian method, we need to simulate these unobservable variables. To do so, we need to use a state space representation of the model and to apply the Kalman filter. However, our DSGE model has not only lagged variables but also rational expectations of future variables (backward-forward looking model), that's why we need to transform the model into a state-space representation.

We can re-write the model into a recursive state-space model through different algorithms that exist in literature like King and Watson (1998), Klein (1999), Sims (2002) and the AIM algorithm² using the parameters values, then we can apply the new form of the model to the Kalman filter, in order to simulate the unobservable variables.

2.1 Priors

Priors may represent the beliefs of the researcher about the parameters. The priors are called subjective if the probability distributions are constructed in a subjective way, without using any evaluation based on the data or on previous studies. However, priors may represent the evaluation based on the observations or on previous models evaluations, called objective priors. Some priors may be between these two cases, objective in the forme of the distribution but subjective in its values. When a prior contains enough information on the parameters, it is called an informative prior. Elsewhere, a non-informative prior must minimize its influence on the posterior distribution.

The choice of the prior distribution is a critical point for the Bayesian approach, since the choice of the belief *a priori* about the parameter contributes in the belief *a posteriori*. The choice of the prior distribution is important to say how much the data contributes in updating the information about the parameter's distribution. If we want to let the data speak for themselves about the parameter's distribution, we should choose a non-informative prior. For example, a uniform prior insures that the prior information about a parameter is neutral and does not change the information that we get from observable variables distribution. Non-informative priors are typically selected when information is scarce or when we want to minimize the influence of the prior on the posterior distribution. Non-informative priors are location invariant pri-

²DYNARE Toolbox uses the generalized Schur decomposition of the algorithms of Sims (2002) and Klein (1999). However, Anderson(2006) shows that the AIM is the most accurate and least cost algorithm to transform the model into a state-space model.

ors. This implies that the probability density function is equal for all values in a set Θ :

$$P(\beta_1 \in \Theta) = P(\beta_2 \in \Theta)$$

The Jeffrey approach proposes a method to generate non-informative priors based on Fisher information matrix. The matrix of information is the expected hessian matrix of the probability density function of the data conditional to the parameters. Jeffrey's proposition says that the prior can be calculated in the following way:

$$P(\beta) = (\det [I(\beta)])^{0.5}$$

$$I_{i,j}(\beta) = -E \left[\frac{\partial^2 \log f(y|\beta)}{\partial \beta_i \partial \beta_j} \right]$$

However, if the prior has a big role in the determination of posterior's shape, this means that the *a priori* belief, or the theoretical feature, is an important contribution. Also, in the case of short data samples, priors must be as much precised as possible, since the updating by the likelihood distribution is insufficient to improve our belief about the parameter. The choice of these priors should be in an objective way.

2.2 Posteriors simulation

Often, the posterior distribution $P(\beta | y)$ cannot be calculated analytically, because the marginal likelihood of y is not easy to calculate, it is the sum (integral for a continuous set of parameters) of the product of the prior and the likelihood over all the possible values of the parameter. But, since the marginal likelihood is a constant function of parameters, the posterior distribution has a shape proportional to that of the kernel posterior distribution

$$\check{P}(\beta | y) = P(y | \beta) P(\beta) \tag{3}$$

As a result, we can use the kernel posterior distribution as an approximation of the posterior. The reason for using the kernel as an approximation of the posterior is that the marginal likelihood is complicate to calculate, we need to solve numerically multiple integrals. So, in order to compute the posterior distribution, we could simulate it rather than compute it analytically.

The recent litterature of Bayesian estimation for DSGE models uses Markov Chain Monte-Carlo methods (MCMC), using a Gibbs sampler algorithm or a Metropolis-Hastings algorithm (hereafter called MH). The MCMC methods consider that the parameter sequence follow a Markov Chain process with a transition matrix Q . The advantage of the MCMC methods to sampling methods is that they consider the parameter as a

sequence. The sequence is updated by data at each iteration. In order to stationarize the Markov-Chain, we need to throw out a big number of the first simulations, the rest of the sequence dresses the shape of the posterior distribution. The MH algorithm is time-consuming since the acceptance rate³ of the simulations must be 30-50%; if the acceptance rate is too low the posterior shape will be only on a small range around the mode, on the other hand, if the acceptance rate is too high the posterior's shape will include so far values on the tails of the distribution. Since we throw 50-70% of the replications, we need a huge number of Monte-Carlo simulations. However, we use all the simulations of the Gibbs sampler. But the drawback of this method is in the case when we have a block of parameters, we need to solve multiple integrals which is highly cost, this is the case in DSGE models. This is the reason why many studies use MH algorithm as an MCMC method in DSGE models estimation.

Using the Metropolis-Hastings algorithm, we will draw a large number L of the parameter from the transition probability. In order to get a stationary Markov Chain process, we will throw a large number of the first simulations (n_0). The Metropolis-Hastings algorithm works as follow:

- 1) For $n = 0$, the initial value of the parameter is β_0
- 2) Draw a random number u from a uniform distribution $U[0, 1]$
- 3) Draw β_t from $Q(\beta_n, \beta_t)$, where $Q(\beta_n, \beta_t)$ is the probability to get a value β_t of the parameter sequence, conditional to the parameter value β_n at the precedent iteration.

4) Calculate the Hastings ratio⁴: $H(\beta_n, \beta_t) = \frac{Q(\beta_t, \beta_n)}{Q(\beta_n, \beta_t)}$ and the MH ratio: $MH(\beta_n, \beta_t) = \frac{P(y|\beta_t)P(\beta_t)}{P(y|\beta_n)P(\beta_n)} H(\beta_n, \beta_t)$

If $MH(\beta_n, \beta_t) > u$, then $\beta_{n+1} = \beta_t$, else keep the previous draw $\beta_{n+1} = \beta_n$; $n = n + 1$.

Repeat steps 1 to 4 until we reach the maximum number of iterations L . The iterations from n_0 to L will give the shape of the posterior distribution $P(\beta | y)$.

We can note that the MH algorithm uses acceptance sampling by drawing u between 0 and 1. This is in order to avoid to get sequences on the tails of the posterior distribution.

³The acceptance rate is the part of draws to be used in the posterior's shape construction. It insures that the draws are not issue of draws neither on the tails of the distribution nor concentrated around the mode. See the algorithm steps for details.

⁴Note that the Hastings ratio is equal to one in symmetric Markov transition matrix. In this case, we go back to the initial algorithm, without Hastings development, the Metropolis algorithm.

2.3 The likelihood function of a State-Space model

In order to calculate the posterior kernel, we need to calculate the likelihood, this is not trivial in DSGE models, since they are backward-forward looking and they contain unobserved variables. If the linearized form of the model is under the form:

$$AE_tX_{t+1} + BX_t + CX_{t-1} + D\varepsilon_t = d \quad (4)$$

where A, B, C and D are coefficient matrices, X_t is a vector of variables, ε_t is a vector of shocks and d is a vector of constant coefficients. The problem of these models is that the matrix A may be singular. As a consequence the model cannot be directly written into a recursive form. We need to use the Schur decomposition through algorithms such as AIM, Klein(1999) and Sims(2002). The matrices A and B can be decomposed using the generalized Schur decomposition into QUZ and QMZ , respectively. The matrices Q and Z are supposed to be invertible and the matrices U and M are upper-triangular matrices. A unique stable solution of the model exists if and only if the number of eigenvalues of the matrix Z larger than one (unstable eigenvalues) is equal to the number of forward-looking variables. This last condition is the condition of Blanchard and Kahn (1980) of stability. As a result of the presence of upper-triangular matrices, the model can be written into a recursive state-space model represented in these two blocks of equations:

$$y_t = b_0 + Z_t b_1 + b_2 c_t + u_t \quad (5)$$

$$c_t = Tc_{t-1} + a + \eta_t \quad (6)$$

Equation (5) is the set of space equations or the measurement equations, with y_t are the observable variables, c_t are the unobservable variables, Z_t a set of exogenous variables, b_0, b_1 and b_2 are, respectively, the corresponding coefficients and the serially uncorrelated measurement error is u_t , with a variance-covariance matrix R . Equation (6) is the set of state equations that are recursive equations of the unobservable variables, with a constant transition matrix T , a constant a and a serially uncorrelated error ε_t , with a variance-covariance matrix E .

In order to compute the likelihood distribution of the observed data, we need first to generate the unobserved variables using the Kalman filter. We will expand the discussion about the Kalman filter⁵ and its algorithm using the pooling estimation in a panel approach. The most

⁵See Hamilton (1994) and Harvey (1990)

important output of the Kalman filter is the one-step ahead prediction error of the observed variables

$$v_t = y_t - \hat{y}_{t|t-1}$$

with a variance matrix denoted by F_t , where $\hat{y}_{t|t-1}$ is the prediction of the observed variables using the information available in $t - 1$. Now, we can write the likelihood function:

$$\ln(L(y | c_0, P_0, \beta)) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |F_t| - \sum_{t=1}^T v_t' F_t^{-1} v_t \quad (7)$$

Since the likelihood function is computed, we can write the posterior kernel

$$\check{P}(\beta | y) = P(\beta)L(y | c_0, P_0, \beta) \quad (8)$$

If we cannot compute analytically the shape of the posterior distribution, we need to use the Metropolis-Hastings algorithms explained previously.

3 Pooling Bayesian estimation method

The interesting point of a pooling Bayesian estimation is that we can get more information by the data distribution, the likelihood function of all countries data. Here, we need an assumption on the parameters of the model, they may have a common value across countries. Taking into account the similarity of some parameters across countries increases the contribution in updating our belief about the parameters. In the case of the emerging and transition countries, especially the Eastern Europe countries, where few data are available, this estimation may be more accurate or less dependant on the prior. In addition, it permits to test the heterogeneity between the parameters posteriors across countries.

3.1 Our specification about the model

In DSGE models literature, no one takes into account similarities and differences between countries, in the same model. However, this is an important feature, since we can test the heterogeneity between some parameters and gain more degrees of freedom of the estimation, in comparison with a separate country-by-country estimation. Moreover, we have to take into consideration the correlation of shocks across countries. We will assume that some of the structural parameters are identical across countries. While, some other structural parameters are not restricted to be identical across countries, they are supposed to be heterogenous.

We suppose also that shocks are correlated between countries, since they have important commercial trade between them. This hypothesis

increases the cost to compute the likelihood function and to estimate the model, since the prediction using the Kalman filter must take into account the correlation of shocks between countries. This fact implies a huge amount of covariance parameters to estimate. So, if the variance-covariance matrix of shocks contains a big number of different matrices of variance-covariance, this gives a matrix with a huge number of parameters to estimate:

$$Q = \begin{bmatrix} Q_1 & Q_{12} & \dots & \dots & Q_{1j} & \dots & Q_{1J} \\ Q_{12} & \cdot & & & & & \cdot \\ \cdot & & \cdot & & & & \cdot \\ \cdot & & & & Q_j & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ Q_{1J} & \dots & \dots & & Q_{jJ} & \dots & Q_J \end{bmatrix}$$

This large number of parameters to estimate increases the cost of the estimation, that we will pay in terms of estimation time and estimation accuracy.

In order to take into account the proximity between countries (the correlation between shocks) and the short samples problem, we will suppose that the shocks are composed of two uncorrelated components: one is global for all the region (ε_t^g), while the other (ε_{jt}^l) is local and uncorrelated between countries, with the country index $j = 1, \dots, J$.

$$\varepsilon_{jt} = \varepsilon_t^g + \varepsilon_{jt}^l \quad (9)$$

where each component is a vector of the same number of shocks.

Both the local and the global shocks are normally distributed with variance matrices Q_j^l and Q^g , respectively.

$$\varepsilon_t^c \sim N(0, Q^c)$$

$$\varepsilon_{jt}^l \sim N(0, Q_j^l)$$

As a result, the variance matrix of the pooled vector (ε_t) containing the total shocks (ε_{jt}) can be written as:

$$Q = \begin{bmatrix} Q_1^l + Q^c & Q^c & \cdot & \cdot & \cdot & Q^c \\ Q^c & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & Q_j^l + Q^c & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & Q^c \\ Q^c & \cdot & \cdot & \cdot & \cdot & Q^c \end{bmatrix}_{(JH \times JH)}$$

This matrix contains only $J + 1$ different submatrices.

3.2 Priors

As we mentioned previously, some of the structural parameters are restricted to be identical across countries, so we suppose that they have the same prior distribution and parameter value.

For the unrestricted parameters, prior distributions are selected independently for each country. Priors of the structural parameters, denoted by $P(\beta_j)$, have similar shapes but different values of moments. The priors can be both objective and non-informative, objective priors are taken from other available studies, these priors are related to the structural parameters of the model, while non-informative priors concern mostly the auxiliary parameters that are related to the shocks.

The innovations of the shocks process, $\varepsilon_{j,t}^l$ and ε_t^c are normally distributed, with mean zero and variance Q_j^l and Q^c . The priors of the autoregressive parameters are denoted by $P(\rho_j)$. These priors can be determined using the Jeffrey approach or using uniform priors.

3.3 A multi-country State-Space model

In order to write the likelihood function of the pooled data, we need to build the Kalman filter for a panel state-space model where the error terms are correlated between countries. To get the pooled state-space model for panel data, we will apply the Schur decomposition to each country; then get the state-space model separately, using the priors associated to each country parameters.

$$y_{j,t} = Z_{j,t}b_{1,j} + b_{2,j}c_{j,t} + u_{j,t} \quad (10)$$

$$c_{j,t} = T_j c_{j,t-1} + \eta_{j,t} \quad (11)$$

For the identical parameters across countries, the corresponding parameters in the state-space model are repeated for countries.

We cannot predict the unobservable variables using separate state-space model in the last form, since the shocks are correlated across countries. This correlation must be taken into account in the prediction of the state variables. Now, we can pile up the variables and parameters for all countries taking into account the correlation across countries, given above in the matrix Q . As a result, the pooled state-space model can be written as:

$$y_t^{(JK \times 1)} = Z_t^{(JK \times JM)} S_Z^{(JM \times p)} b_1^{(p \times 1)} + b_2^{(JK \times q)} S_c^{(q \times JH)} c_t^{(JH \times 1)} + u_t^{(JK \times 1)} \quad (12)$$

$$c_t^{(JH \times 1)} = T^{(JH \times JH)} c_{t-1}^{(JH \times 1)} + \eta_t^{(JH \times 1)} \quad (13)$$

Where K is the number of endogenous space variables in the model, M is the number of exogenous variables and H is the number of state variables in the model. M and H contain, both, the heterogenous parameters and the identical parameters corresponding to the variables. S_Z and S_c are selection matrices to pool the heterogenous parameters with the identical parameters, where p, q and r are the number of corresponding parameters, containing one identical parameter and J heterogenous parameters. The transition matrix T is a highly sparse matrix, it contains blocks of transition matrices for countries on the diagonal, the other elements are zero. The matrix T contains some identical parameters.

Now, we will write the algorithm of the Kalman filter, in the pooled model case, with a homoskedastic shock. Given the initial value of the state variables vector $c_0^{(JH \times 1)}$, we will predict the next period state variable by the transition equation:

$$c_{t|t-1} = Tc_{t-1} \quad (14)$$

The variance matrix of the state variable prediction is denoted by P , it can be obtained by the solution of

$$P_{t|t-1} = TP_{t-1}T' + E \quad (15)$$

We can now compare between the prediction of space variables and their observed values, the prediction error is denoted by v_t ,

$$v_t = y_t - y_{t|t-1} \quad (16)$$

where the space variable prediction can be obtained by the space equation (12)

$$y_{t|t-1} = b_2c_{t|t-1} + Z_t b_1 \quad (17)$$

We can also calculate the prediction error variance covariance matrix, denoted by F_t , with dimension $(JK \times JK)$, that is simplified since the model does not contain measurement errors

$$F_t = b_2P_{t|t-1}b_2' + R \quad (18)$$

Finally, we can calculate the state variable c_t , which we can take by the state variable prediction plus the error term of the prediction

$$c_t = c_{t|t-1} + P_{t|t-1}b_2'F_t^{-1}v_t \quad (19)$$

and its variance matrix

$$P_t = P_{t|t-1} + P_{t|t-1}b_2'F_t^{-1}b_2P_{t|t-1}'$$

Once the unobservable variables are simulated and the model is written in a recursive way, we can write the the likelihood function of the data:

$$\begin{aligned} & \ln L(y_1, \dots, y_J \mid \beta_1, \dots, \beta_J, \rho_1, \dots, \rho_J, Q) \\ &= -\frac{JT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |F_t| - \sum_{t=1}^T v_t' F_t^{-1} v_t \end{aligned} \quad (20)$$

3.4 Posteriors simulation

Once the likelihood function is calculated, we can now calculate the kernel function conditional to the other parameters values, for each country independently, because for a hierrarchical model, posteriors are independent if the priors are independent.

$$\begin{aligned} & \check{P}(\beta_j \mid y_1, \dots, y_J, \beta_{j^*}, \rho_1, \dots, \rho_J, Q) \\ &= L(y_1, \dots, y_J \mid \beta_1, \dots, \beta_J, \rho_1, \dots, \rho_J, Q) P(\beta_j) \end{aligned} \quad (21)$$

where β_{j^*} denotes the vector of parameters for the remaining countries.

The posteriors kernel of the shocks autoregressive parameters vector are witten similarly using their priors and the likelihood function

$$\begin{aligned} & \check{P}(\rho_j \mid y_1, \dots, y_J, \beta_1, \dots, \beta_J, \rho_{j^*}, Q) \\ &= L(y_1, \dots, y_J \mid \beta_1, \dots, \beta_J, \rho_1, \dots, \rho_J, Q) P(\rho_j) \end{aligned} \quad (22)$$

Then, using the kernel functions (21)–(22) we can do the Metropolis-Hastings algorithm for each country and simulate the posterior distribution for each country.

3.5 How to test the heterogeneity of parameters?

In order to test the heterogeneity using the Bayesian approach, we can have more than one way. First, we can compare between the modes of the different posterior distributions for each parameter, far distances between modes can indicate heterogeneity of parameters. However, the Bayesian approach considers the parameters as random variables, not as a point. So, in order to analyze the parameters heterogeneity we need to compare, not only between modes of posterior distributions but also between the shape of the distributions. So, the second method that we propose is to compare between the Highest Density Regions (HDR) at a given common level $1 - \alpha$ of the different distributions of the same parameter. The HDR of a posterior distribution, at a level $1 - \alpha$ is

defined as the shortest interval where there is $1 - \alpha$ of the parameter's distribution.

$$HDR = [L_{\alpha_1}; H_{\alpha_2}]$$

with α_1 and α_2 the p-values of the distribution, satisfying $\alpha_1 + \alpha_2 = \alpha$

$$\int_{L_{\alpha_1}}^{H_{\alpha_2}} P(\beta|y) d\beta = 1 - \alpha$$

Once, we have the different HDR's for a differentiated parameter, we can compare between their longitude, larger intervals reflect the fact that the data of that country does not update sufficiently the belief about the parameter.

The third way is using the Bayesian factor, which is defined as the ratio of the posterior probability of a parameter.

$$B_{i,j} = \frac{\check{P}(\beta_i|y_1, \dots, y_J)}{\check{P}(\beta_j|y_1, \dots, y_J)} \quad (23)$$

This way corresponds to testing the null hypothesis that the parameter is identical between two countries, against the alternative that they are different. If the Bayesian factor is different from 1, then the parameters can be considered heterogenous across countries. The Bayesian factor is bilateral (i.e: between two countries), for this fact we must fix a reference country to which we can compare all the other countries.

4 Application on a DSGE model

For the illustration of the methodology in a panel of countries approach, we will apply the Bayesian analysis on a Data Generating Process (DGP) of a simple DSGE model for a panel of J countries. We will study three scenarios of the model: The first assumes that all parameters of the dynamic model are different and will estimate it taking into account this assumption, in the second scenario, some parameters are identical across countries and others are differentiated we will estimate the model taking into account these assumptions. Finally, in the third scenario, we assume an error of specification of the model, where all parameters are differentiated but some are estimated using the assumption that they are identical across countries.

4.1 The model

We use a simple dynamic stochastic general equilibrium in a closed economy, without capital. The model, that is given by Rabanal and Rubio-Ramirez (2003), exhibits nominal rigidity of prices with partial indexation, and nominal rigidity of wages with partial indexation. The shocks

of the economy are: a productivity shock (A), a government expenditure (g) that is $AR(1)$ process with (ε^g) as an innovation, a monetary policy shock (ε^m), and a mark-up shock (λ). The log-linearized model is composed of the following equations:

$$y_t = E_t y_{t+1} - \sigma (i_t - E_t \pi_{t+1}^c + E_t g_{t+1} - g_t) \quad (24)$$

$$y_t = \varepsilon_t^a + (1 - \delta) n_t \quad (25)$$

$$mc_t = w_t + n_t - y_t \quad (26)$$

$$\pi_t = \frac{\omega}{1 + \beta\omega} \pi_{t-1} + \frac{\beta}{1 + \beta\omega} E_t \pi_{t+1} + \kappa_p (mc_t + \lambda_t) \quad (27)$$

$$mrs_t = g_t + \frac{1}{\sigma} y_t + \gamma n_t \quad (28)$$

$$\hat{W}_t - \alpha = \beta E_t \hat{W}_{t+1} + \kappa_w (mrs_t - w_t) \quad (29)$$

$$i_t = \rho_r i_{t-1} + (1 - \rho_r) (\gamma_\pi \pi_t + \gamma_y y_t) + \varepsilon_t^m \quad (30)$$

$$\varepsilon_t^a = (1 - \rho_a) + \rho_a \varepsilon_{t-1}^a + \eta_t^a \quad (31)$$

$$\varepsilon_t^g = (1 - \rho_g) + \rho_g \varepsilon_{t-1}^g + \eta_t^g \quad (32)$$

Equation (24) is an Euler equation with an elasticity of consumption σ , equation (25) is a concave production function with only labor as a production factor, equation (26) is the optimal condition of the cost minimization problem. Equation (27) is a Phillips curve, in a New-Keynesian form⁶, prices are adjusted at each period with a probability θ_p , when prices are not adjusted they are indexed with previous gross inflation rate with a degree of indexation ω between 0 and 1. Equation (28) represents the marginal rate of substitution between labor and consumption, where γ is the inverse elasticity of labor supply with respect to real wages. Equation (29) is the equation of nominal wages growth \hat{W} with respect to marginal rate of substitution, real wages and inflation⁷, wages adjust also with a probability θ_w , when wages are not adjusted they are indexed to past wage with past period inflation indexation. The indexation degree of wages is denoted by α that is between 0 and 1. Equation (30) is a Taylor monetary rule and finally equations (31) and (32) represent the $AR(1)$ process of the shocks.

Since the model is log-linearized around the steady state, the variables of the model above are all equal to zero at the steady state.

⁶ $\kappa_p = \frac{(1-\delta)(1-\theta_p\beta)(1-\theta_p)}{(1+\omega\beta)\theta_p(1+\delta(\bar{\varepsilon}-1))}$ where $\bar{\varepsilon}$ is the steady state elasticity of substitution between differentiated goods.

⁷ $\kappa_w = \frac{(1-\theta_w\beta)(1-\theta_w)}{\theta_w(1+\phi\gamma)}$ where ϕ is the elasticity of substitution between differentiated labor types in the production function.

4.2 The basic model estimation

First, the model has been estimated for one country, without panel dimension, in order to illustrate the Bayesian analysis of this new keynesian models. The model is estimated on the simulated data using the following parameters values:

σ	θ_p	θ_w	α	ω	γ	γ_y	γ_π
0.18	0.66	0.75	0.7	0.7	1.6	0.27	1.4
ρ_a	ρ_g	ρ_r	σ_a	σ_m	σ_g	σ_λ	
0.9	0.9	0.82	0.0096	0.0014	0.04	0.18	

The other parameters are set as in other standard studies $\beta = 0.99$, $\delta = 0.36$, $\bar{\epsilon} = 6$ and $\phi = 6$. Estimating the model on the simulated data for 100 periods, by the methodology explained previously, gives the posterior distributions of the parameters of interest. The estimation shows that the parameters values of the data-generating process are inside the 95% confidence interval constructed from the posterior distribution. The 95% confidence interval is the *highest density region* (HDR) with a 95% level of the posterior distribution. The results obtained are:

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.3505	0.0000	0.6520
ω	0.5	uniform	0.2887	0.7905	0.6363	0.9960
θ_p	0.6	beta	0.1	0.6090	0.5186	0.7414
θ_w	0.7	beta	0.1	0.7265	0.6310	0.8230
γ	1	normal	0.5	1.0559	0.3425	1.6811
γ_π	1.5	normal	0.25	1.4377	1.3146	1.5689
γ_y	0.125	normal	0.25	0.3044	0.1639	0.4356
σ	0.4	gamma	0.2000	0.1737	0.1602	0.1860
ρ_a	0.5	uniform	0.2887	0.9274	0.8938	0.9593
ρ_g	0.5	uniform	0.2887	0.8886	0.8615	0.9159
ρ_r	0.5	uniform	0.2887	0.8273	0.8082	0.8475
σ_{η^a}	0.5	uniform	0.2887	0.0083	0.0054	0.0110
σ_{η^g}	0.5	uniform	0.2887	0.0423	0.0375	0.0470
σ_{ε^m}	0.5	uniform	0.2887	0.0013	0.0012	0.0015
σ_λ	0.5	uniform	0.2887	0.1433	0.0470	0.2339

Table 1: Estimation results for the basic model for one country

The prior choice is due to a prior thought about parameters values. We know *a priori* that $\alpha, \omega, \rho_a, \rho_g,$ and ρ_r between 0 and 1, but we do not

have any idea about the distribution of the parameters between 0 and 1, for this reason we specify uniform priors for these parameters. However, theoretical values about the rigidity parameters give values between 0.4 and 0.7. As a result, we specify a beta distribution between 0 and 1, with a mode near to the theoretical values. The other parameters of elasticity and policy are normally distributed around the theoretical values.

These results in table 1, with the posterior distributions shape in appendix 6.1, give the mean of the posterior distribution, but also the uncertainty about the parameter. Larger HDR or more flat distributions reflect more uncertainty about that parameter. The presence of more than one mode of the posterior increase also the uncertainty about the parameters possible values. The estimation results say that we have more uncertainty about the indexation degree parameters (ω and α), the calvo parameters of prices (θ_p) than the other parameters.

The MCMC diagrams show also if the Metropolis-Hastings draws are stable. This is the case for all the parameters since the draws graphs are stable almost after the first 20,000 draws, the rest 20,000 draws are more stable. In order to verify the presence of other modes, we drawn 3 different sequences from the posterior distribution. The blue line in the MCMC diagrams represents the graph for all the sequences together, while the red line is the graph of the *Within* series (i.e: the difference between the draws of each sequence and the mean of each sequence). Here, to check the presence of more than one mode, we launch three sequences of draws from the posterior parameters. Almost, for all the parameters, the red and the blue lines converge to the same value. As a result, we can conclude that posterior distributions have one mode, except for the Calvo parameter of prices rigidity (θ_p), the indexation degree of wages (α) and the standard error of the mark-up shock (λ), where we observe two modes of the posterior distributions. This states that these parameters still have some uncertainty.

4.3 Estimation of the model in a panel approach

In this section, we discuss the effect of the data pooling on posterior distributions estimation. We do this by simulating data for three countries from the model restricting some parameters to have identical values across countries and differentiating the other parameters. We will assume that the identical parameters are those related to the households, since if countries are located in the same region, households may have many common habits such as wage indexation α , consumption elasticity

σ , and the elasticity of disutility of labor⁸. In order to take into account the shocks correlations across countries, we assume, as explained in the methodology previously, that each shock (ε) is composed of two components: a domestic shock denoted by (ε^l) and a common shock (ε^c). The data are obtained by a stochastic simulation of the model using the following parameters values:

$$\text{Identical parameters values: } \frac{\alpha}{0.7} \mid \frac{\gamma}{1.6} \mid \frac{\sigma}{0.18}$$

$$\begin{array}{l} \text{Country 1: } \frac{\theta_{p1}}{0.66} \mid \frac{\theta_{w1}}{0.75} \mid \frac{\omega_1}{0.7} \mid \frac{\gamma_{y1}}{0.27} \mid \frac{\gamma_{\pi1}}{1.4} \\ \frac{\rho_{a1}}{0.9} \mid \frac{\rho_{g1}}{0.9} \mid \frac{\rho_{r1}}{0.82} \mid \frac{\sigma_{a1}^l}{0.003} \mid \frac{\sigma_{m1}^l}{0.002} \mid \frac{\sigma_{g1}^l}{0.01} \mid \frac{\sigma_{\lambda1}^l}{0.05} \end{array}$$

$$\begin{array}{l} \text{Country 2: } \frac{\theta_{p2}}{0.72} \mid \frac{\theta_{w2}}{0.75} \mid \frac{\omega_2}{0.7} \mid \frac{\gamma_{y2}}{0.25} \mid \frac{\gamma_{\pi2}}{1.35} \\ \frac{\rho_{a2}}{0.9} \mid \frac{\rho_{g2}}{0.9} \mid \frac{\rho_{r2}}{0.75} \mid \frac{\sigma_{a2}^l}{0.004} \mid \frac{\sigma_{m2}^l}{0.002} \mid \frac{\sigma_{g2}^l}{0.015} \mid \frac{\sigma_{\lambda2}^l}{0.18} \end{array}$$

$$\begin{array}{l} \text{Country 3: } \frac{\theta_{p3}}{0.62} \mid \frac{\theta_{w3}}{0.7} \mid \frac{\omega_3}{0.7} \mid \frac{\gamma_{y3}}{0.3} \mid \frac{\gamma_{\pi3}}{1.45} \\ \frac{\rho_{a3}}{0.9} \mid \frac{\rho_{g3}}{0.82} \mid \frac{\rho_{r3}}{0.7} \mid \frac{\sigma_{a3}^l}{0.004} \mid \frac{\sigma_{m3}^l}{0.0025} \mid \frac{\sigma_{g3}^l}{0.02} \mid \frac{\sigma_{\lambda3}^l}{0.06} \end{array}$$

$$\text{Common shocks standard deviation: } \frac{\sigma_a^c}{0.008} \mid \frac{\sigma_m^c}{0.005} \mid \frac{\sigma_g^c}{0.03} \mid \frac{\sigma_\lambda^c}{0.08}$$

The common shocks identification is a problem in calculating the likelihood function, estimating the common shocks standard deviation imposes a big problem of identification. The likelihood function is very flat. In order to avoid this problem we estimate only the total shock standard deviation.

In the estimation process, we give the same priors for parameters across countries. We assume that the parameters concerning households are identical across countries, assuming that households have the same behavior in these countries. The other parameters, even if some of them have similar values, are assumed to be different and have their proper distributions.

⁸This fact can be more illustrated by empirical works. Unfortunately, till now we cannot find estimated structural DSGE models for a panel of countries.

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.6528	0.4386	0.8511
γ	1	normal	0.5	1.5288	0.9556	2.1107
σ	0.4	gamma	0.2000	0.1596	0.1542	0.1650
ω_1	0.5	uniform	0.2887	0.7264	0.6341	0.8272
θ_{p1}	0.6	beta	0.1	0.6609	0.6334	0.6865
θ_{w1}	0.7	beta	0.1	0.7450	0.6792	0.8186
$\gamma_{\pi 1}$	1.5	normal	0.25	1.4543	1.3130	1.5947
γ_{y1}	0.125	normal	0.25	0.3071	0.1608	0.4560
ρ_{a1}	0.5	uniform	0.2887	0.9053	0.8751	0.9312
ρ_{g1}	0.5	uniform	0.2887	0.8720	0.8506	0.8971
ρ_{r1}	0.5	uniform	0.2887	0.8446	0.8337	0.8557
ω_2	0.5	uniform	0.2887	0.6439	0.5423	0.7338
θ_{p2}	0.6	beta	0.1	0.7236	0.7009	0.7532
θ_{w2}	0.7	beta	0.1	0.7502	0.6951	0.8098
$\gamma_{\pi 2}$	1.5	normal	0.25	1.3126	1.2325	1.4013
γ_{y2}	0.125	normal	0.25	0.1781	0.0936	0.2500
ρ_{a2}	0.5	uniform	0.2887	0.9097	0.8875	0.9297
ρ_{g2}	0.5	uniform	0.2887	0.8905	0.8680	0.9137
ρ_{r2}	0.5	uniform	0.2887	0.7693	0.7601	0.7782
ω_3	0.5	uniform	0.2887	0.6278	0.5224	0.7343
θ_{p3}	0.6	beta	0.1	0.6232	0.5985	0.6522
θ_{w3}	0.7	beta	0.1	0.6823	0.5684	0.7689
$\gamma_{\pi 3}$	1.5	normal	0.25	1.4281	1.3261	1.5156
γ_{y3}	0.125	normal	0.25	0.1888	0.1198	0.2606
ρ_{a3}	0.5	uniform	0.2887	0.8969	0.8713	0.9176
ρ_{g3}	0.5	uniform	0.2887	0.7749	0.7436	0.8072
ρ_{r3}	0.5	uniform	0.2887	0.7749	0.7686	0.7801

Table 2a: Estimation results for the model parameters in a panel approach

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
σ_{a1}	0.5	uniform	0.2887	0.0038	0.0029	0.0048
σ_{g1}	0.5	uniform	0.2887	0.0078	0.0052	0.0105
σ_{m1}	0.5	uniform	0.2887	0.0016	0.0013	0.0021
$\sigma_{\lambda1}$	0.5	uniform	0.2887	0.0488	0.0470	0.2339
σ_{a2}	0.5	uniform	0.2887	0.0038	0.0028	0.0047
σ_{g2}	0.5	uniform	0.2887	0.0149	0.0126	0.0172
σ_{m2}	0.5	uniform	0.2887	0.0020	0.0017	0.0022
$\sigma_{\lambda2}$	0.5	uniform	0.2887	0.0598	0.0408	0.0823
σ_{a3}	0.5	uniform	0.2887	0.0042	0.0032	0.0051
σ_{g3}	0.5	uniform	0.2887	0.0165	0.0147	0.0184
σ_{m3}	0.5	uniform	0.2887	0.0027	0.0023	0.0031
$\sigma_{\lambda3}$	0.5	uniform	0.2887	0.0529	0.0396	0.0643

Table 2b: Estimation results for the shocks variance in a panel approach

What we can observe from the estimation results in table 2a and 2b, that we gain more precision about some parameters. First, for the restricted parameters to be identical across countries α , γ and σ , the 95% HDR are shorter than in the estimation for the basic model. But also, some other non-restricted parameters become more precise than in the basic model results, like the Calvo parameter of prices rigidity θ_p for all the countries, the parameter associated to the output gap in the monetary rule γ_y and the autoregressive parameters ρ_a , ρ_g and ρ_r . Even if these parameters are not restricted across countries, we could gain more accuracy about these parameters from the interaction between parameters in a DSGE model.

Looking on the MCMC draws diagram, we see some differences between total draws of three chains of Markov and the Within of the MCMC draws. This means the presence of more than one mode of the posterior distribution for some parameters. After looking at the posteriors shape, in appendix 6.2, we observe that the Calvo rigidity parameters and the indexation degree parameters have more than one mode. However, these differences are small for most of the parameters or converge to the same values as the *Within* of the draws. Even if the posterior distribution has several modes, they are closed to eachother. As a result, pooling the model for three countries increases the precision about parameters.

4.4 Estimation of the model in a panel approach with heterogeneity

4.4.1 Shocks heterogeneity across countries

In this case we assume a shocks heterogeneity in autoregressive parameters ρ and in rigidity parameters of the DGP across countries. The heterogeneity across countries is also represented by reducing the common shocks weight in the total shocks. The simulated data in this case show more heterogeneity than the simulated data section 4.3. The parameters values used in simulation are:

$$\text{Identical parameters values: } \frac{\alpha}{0.7} \mid \frac{\gamma}{1.6} \mid \frac{\sigma}{0.18}$$

$$\begin{aligned} \text{Country 1: } & \frac{\theta_{p1}}{0.66} \mid \frac{\theta_{w1}}{0.75} \mid \frac{\omega_1}{0.7} \mid \frac{\gamma_{y1}}{0.27} \mid \frac{\gamma_{\pi1}}{1.4} \\ & \frac{\rho_{a1}}{0.9} \mid \frac{\rho_{g1}}{0.9} \mid \frac{\rho_{r1}}{0.82} \mid \frac{\sigma_{a1}^l}{0.0066} \mid \frac{\sigma_{m1}^l}{0.007} \mid \frac{\sigma_{g1}^l}{0.03} \mid \frac{\sigma_{\lambda1}^l}{0.12} \end{aligned}$$

$$\begin{aligned} \text{Country 2: } & \frac{\theta_{p2}}{0.72} \mid \frac{\theta_{w2}}{0.75} \mid \frac{\omega_2}{0.65} \mid \frac{\gamma_{y2}}{0.3} \mid \frac{\gamma_{\pi2}}{1.3} \\ & \frac{\rho_{a2}}{0.9} \mid \frac{\rho_{g2}}{0.9} \mid \frac{\rho_{r2}}{0.75} \mid \frac{\sigma_{a2}^l}{0.006} \mid \frac{\sigma_{m2}^l}{0.007} \mid \frac{\sigma_{g2}^l}{0.02} \mid \frac{\sigma_{\lambda2}^l}{0.1} \end{aligned}$$

$$\begin{aligned} \text{Country 3: } & \frac{\theta_{p3}}{0.6} \mid \frac{\theta_{w3}}{0.7} \mid \frac{\omega_3}{0.8} \mid \frac{\gamma_{y3}}{0.3} \mid \frac{\gamma_{\pi3}}{1.3} \\ & \frac{\rho_{a3}}{0.82} \mid \frac{\rho_{g3}}{0.82} \mid \frac{\rho_{r3}}{0.75} \mid \frac{\sigma_{a3}^l}{0.006} \mid \frac{\sigma_{m3}^l}{0.0045} \mid \frac{\sigma_{g3}^l}{0.025} \mid \frac{\sigma_{\lambda3}^l}{0.06} \end{aligned}$$

$$\text{Common shocks standard deviation: } \frac{\sigma_a^c}{0.003} \mid \frac{\sigma_m^c}{0.0005} \mid \frac{\sigma_g^c}{0.01} \mid \frac{\sigma_\lambda^c}{0.06}$$

By assuming the same assumptions on priors and identical parameters as in last case, we obtain the following estimation results:

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.4755	0.1948	0.7726
γ	1	normal	0.5	1.2928	0.4954	2.0137
σ	0.4	gamma	0.2000	0.1698	0.1580	0.1847
ω_1	0.5	uniform	0.2887	0.6848	0.4711	0.8974
θ_{p1}	0.6	beta	0.1	0.6347	0.5562	0.7291
θ_{w1}	0.7	beta	0.1	0.7590	0.6834	0.8464
$\gamma_{\pi 1}$	1.5	normal	0.25	1.5891	1.3826	1.8117
γ_{y1}	0.125	normal	0.25	0.3325	0.1244	0.5416
ρ_{a1}	0.5	uniform	0.2887	0.8782	0.8117	0.9482
ρ_{g1}	0.5	uniform	0.2887	0.8659	0.8367	0.8948
ρ_{r1}	0.5	uniform	0.2887	0.8404	0.8235	0.8582
ω_2	0.5	uniform	0.2887	0.6486	0.5139	0.7995
θ_{p2}	0.6	beta	0.1	0.7151	0.6617	0.7662
θ_{w2}	0.7	beta	0.1	0.7078	0.6132	0.8019
$\gamma_{\pi 2}$	1.5	normal	0.25	1.3338	1.2305	1.4420
γ_{y2}	0.125	normal	0.25	0.3036	0.1437	0.4519
ρ_{a2}	0.5	uniform	0.2887	0.8827	0.8275	0.9495
ρ_{g2}	0.5	uniform	0.2887	0.8817	0.8551	0.9114
ρ_{r2}	0.5	uniform	0.2887	0.7665	0.7454	0.7862
ω_3	0.5	uniform	0.2887	0.6187	0.3978	0.8696
θ_{p3}	0.6	beta	0.1	0.6351	0.5697	0.7036
θ_{w3}	0.7	beta	0.1	0.6286	0.5168	0.7402
$\gamma_{\pi 3}$	1.5	normal	0.25	1.3110	1.1732	1.4413
γ_{y3}	0.125	normal	0.25	0.2469	0.0634	0.4188
ρ_{a3}	0.5	uniform	0.2887	0.6646	0.5324	0.7948
ρ_{g3}	0.5	uniform	0.2887	0.7820	0.7372	0.8302
ρ_{r3}	0.5	uniform	0.2887	0.7595	0.7428	0.7745

Table 3a: Estimation results for the model parameters with shocks heterogeneity

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
σ_{a1}	0.5	uniform	0.2887	0.0078	0.0035	0.0140
σ_{g1}	0.5	uniform	0.2887	0.0287	0.0248	0.0324
σ_{m1}	0.5	uniform	0.2887	0.0067	0.0059	0.0074
$\sigma_{\lambda1}$	0.5	uniform	0.2887	0.1070	0.0336	0.1902
σ_{a2}	0.5	uniform	0.2887	0.0071	0.0041	0.0102
σ_{g2}	0.5	uniform	0.2887	0.0220	0.0190	0.0249
σ_{m2}	0.5	uniform	0.2887	0.0076	0.0067	0.0086
$\sigma_{\lambda2}$	0.5	uniform	0.2887	0.1100	0.0552	0.1659
σ_{a3}	0.5	uniform	0.2887	0.0204	0.0069	0.0345
σ_{g3}	0.5	uniform	0.2887	0.0247	0.0209	0.0282
σ_{m3}	0.5	uniform	0.2887	0.0051	0.0046	0.0057
$\sigma_{\lambda3}$	0.5	uniform	0.2887	0.1660	0.0983	0.2365

Table 3b: Estimation results for the shocks variance with shocks heterogeneity

What we notice from this estimation is that we lose some of the gain in precision obtained by pooling the data for three countries. The loss is notably in the parameter of indexation degrees α and ω . The policy parameters γ_π and γ_y have more risk using this DGP than the DGP of more homogeneous data. However, we can rather conclude that the estimation is still more precise than the basic model with one country. We can also notice that the posterior distributions, in appendix 6.3, of the policy parameters are multi-mode. This fact is also obvious in the MCMC diagram of the Metropolis-Hastings draws. What we can conclude from these results is that pooling the data can bring more precision if the data exhibit sufficient homogeneity in their dynamics. However, if this homogeneity does not exist, the posteriors will be less informative.

4.4.2 Heterogeneity of monetary policy with a misspecification of identical parameters

The purpose of this case demonstration is to examine the consequences of a misspecification assuming identical parameters of monetary policy, while the data are simulated by the model using heterogeneous monetary rule parameters. We use exactly the parameters used in the last case of shocks heterogeneity, with an additional source of heterogeneity using different monetary policy rules. The monetary policy parameters have

the following values:

Parameter	γ_π	γ_y
Country 1	1.4	0.27
Country 2	1.3	0.25
Country 3	1.7	0.4

The estimation results are the following:

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.4308	0.0323	0.7675
γ	1	normal	0.5	1.1073	0.4671	1.7292
σ	0.4	gamma	0.2000	0.1922	0.1714	0.2118
γ_π	1.5	normal	0.25	1.5179	1.3975	1.7215
γ_y	0.125	normal	0.25	0.4091	0.2118	0.5809
ω_1	0.5	uniform	0.2887	0.6744	0.5176	0.8218
θ_{p1}	0.6	beta	0.1	0.6810	0.6051	0.7555
θ_{w1}	0.7	beta	0.1	0.7074	0.6060	0.8009
ρ_{a1}	0.5	uniform	0.2887	0.8901	0.8402	0.9400
ρ_{g1}	0.5	uniform	0.2887	0.8842	0.8619	0.9068
ρ_{r1}	0.5	uniform	0.2887	0.8213	0.7981	0.8441
ω_2	0.5	uniform	0.2887	0.7287	0.5786	0.8806
θ_{p2}	0.6	beta	0.1	0.7322	0.6591	0.8131
θ_{w2}	0.7	beta	0.1	0.6482	0.5093	0.7847
ρ_{a2}	0.5	uniform	0.2887	0.8219	0.7688	0.8762
ρ_{g2}	0.5	uniform	0.2887	0.8569	0.8273	0.8863
ρ_{r2}	0.5	uniform	0.2887	0.7522	0.7213	0.7814
ω_3	0.5	uniform	0.2887	0.6990	0.5404	0.8712
θ_{p3}	0.6	beta	0.1	0.7233	0.6559	0.7927
θ_{w3}	0.7	beta	0.1	0.6671	0.5432	0.7834
ρ_{a3}	0.5	uniform	0.2887	0.8596	0.7835	0.9342
ρ_{g3}	0.5	uniform	0.2887	0.8531	0.8243	0.8820
ρ_{r3}	0.5	uniform	0.2887	0.7323	0.7008	0.7644

Table 4a: Estimation results for the model parameters with misspecification

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
σ_{a1}	0.5	uniform	0.2887	0.0075	0.0036	0.0115
σ_{g1}	0.5	uniform	0.2887	0.0315	0.0275	0.0353
σ_{m1}	0.5	uniform	0.2887	0.0063	0.0055	0.0070
$\sigma_{\lambda 1}$	0.5	uniform	0.2887	0.1312	0.0503	0.2161
σ_{a2}	0.5	uniform	0.2887	0.0081	0.0037	0.0121
σ_{g2}	0.5	uniform	0.2887	0.0230	0.0201	0.0260
σ_{m2}	0.5	uniform	0.2887	0.0073	0.0064	0.0082
$\sigma_{\lambda 2}$	0.5	uniform	0.2887	0.1212	0.0138	0.2102
σ_{a3}	0.5	uniform	0.2887	0.0111	0.0034	0.0207
σ_{g3}	0.5	uniform	0.2887	0.0275	0.0235	0.0310
σ_{m3}	0.5	uniform	0.2887	0.0039	0.0034	0.0044
$\sigma_{\lambda 3}$	0.5	uniform	0.2887	0.4033	0.2045	0.6194

Table 4b: Estimation results for the shocks variance with misspecification

The restriction on policy parameters imply an uncertainty increase for some parameters. The uncertainty about parameters measured by the size of the HDR is larger for the indexation degree of wages (α). The uncertainty increases also for the policy parameters with respect to the policy parameters of country 2 and 3. This is due to the fact that prior's mean is closer to the true value of the policy parameters in country 1 than in countries 2 and 3. We can also notice the presence of more than one mode of the posterior distributions or flat posteriors for the indexation degree parameters.

To conclude this point, a misspecified restriction on parameters of identical parameters across countries, while they are really different increases the uncertainty, according the prior's mean.

4.5 Pooling effect of short data

What we want to show here is the gain of pooling data small samples. Since we are using bayesian estimation, the priors weight is more important than in the previous cases. To illustrate this point we will compare between the estimation of the basic model over a short simulated data sample (40 periods) for one country, and the estimation over the pooled data for three countries⁹. This will clarify the gain from pooling the data in a panel dimension, with short data samples.

⁹We took the first 40 observations of the same simulated data of the previous cases.

4.5.1 The basic model estimation with short data

In this case, we estimated the same model as in section 4.2 over a shorter data sample (40 periods). For a short sample, the contribution of priors in posteriors determination is more important in this estimation than in that obtained in section 4.2. This is actually the contribution of Bayesian econometrics in short data samples problems. For this reason any error of specification of priors will affect the results of the posterior distributions. The results obtained by the estimation of the basic model over 40 periods are the following:

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.0066	0.0000	0.159
ω	0.5	uniform	0.2887	0.6086	0.3162	0.9539
θ_p	0.6	beta	0.1	0.5789	0.4478	0.7073
θ_w	0.7	beta	0.1	0.7507	0.6282	0.8672
γ	1	normal	0.5	1.0897	0.3668	1.8271
γ_π	1.5	normal	0.25	1.5972	1.3775	1.8317
γ_y	0.125	normal	0.25	0.4441	0.2297	0.6829
σ	0.4	gamma	0.2000	0.1565	0.1332	0.1827
ρ_a	0.5	uniform	0.2887	0.9536	0.9273	0.9847
ρ_g	0.5	uniform	0.2887	0.8498	0.8027	0.8968
ρ_r	0.5	uniform	0.2887	0.8456	0.8218	0.8701

Table 5a: Estimation results for the model parameters for short data samples

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
σ_{η^a}	0.5	uniform	0.2887	0.0071	0.0048	0.0094
σ_{η^g}	0.5	uniform	0.2887	0.0436	0.0357	0.0513
σ_{ε^m}	0.5	uniform	0.2887	0.0014	0.0011	0.0017
σ_λ	0.5	uniform	0.2887	0.1360	0.0413	0.2254

Table 5b: Estimation results for the variance shocks for short data samples

The comparison between the posteriors of the basic model estimation for 100 periods and for 40 periods (in appendix 6.1 and 6.5, respectively) can tell us how much the priors contribute in the posteriors determination. For the shorter data sample, the posteriors shapes are closed to the priors shapes. Specially, the parameters of nominal rigidity θ_p and θ_w and the parameter of the elasticity of work in the utility function

γ , they have closed shapes to the prior distributions. Since the priors of the indexation degree parameters (α and ω) are non-informative, the posterior of ω is flat with a 95% HDR over a big part of the prior's region. However, the posterior of α has a smaller standard deviation, but the mean and the HDR of the posterior are too far from the parameter's value used in the data generating process. The parameters of the auto-regressive process, who have non-informative priors, have posteriors with many modes. This is also due to the uniform prior distribution specified to these parameters.

4.5.2 Estimation of the model in a panel approach with short data samples

For this case, we want to show the effect of pooling the data on the posterior distributions. The objective here is to illustrate the gain that can be obtained by pooling the data of some countries which have some similar characteristics. We will use the data simulated by the panel-model of section 4.3, but only for 40 periods.

The estimation results are the following:

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
α	0.5	uniform	0.2887	0.6133	0.2848	0.9558
γ	1	normal	0.5	1.1769	0.4785	1.8686
σ	0.4	gamma	0.2000	0.1580	0.1411	0.1782
ω_1	0.5	uniform	0.2887	0.8488	0.7083	0.9999
θ_{p1}	0.6	beta	0.1	0.6402	0.6026	0.6805
θ_{w1}	0.7	beta	0.1	0.7252	0.6386	0.8141
$\gamma_{\pi 1}$	1.5	normal	0.25	1.4604	1.2757	1.6324
γ_{y1}	0.125	normal	0.25	0.2655	0.0734	0.4398
ρ_{a1}	0.5	uniform	0.2887	0.875	0.8392	0.9386
ρ_{g1}	0.5	uniform	0.2887	0.8578	0.8181	0.9039
ρ_{r1}	0.5	uniform	0.2887	0.8444	0.8169	0.8685
ω_2	0.5	uniform	0.2887	0.6762	0.5314	0.8162
θ_{p2}	0.6	beta	0.1	0.7151	0.6849	0.7446
θ_{w2}	0.7	beta	0.1	0.7066	0.6237	0.7928
$\gamma_{\pi 2}$	1.5	normal	0.25	1.3799	1.2403	1.5145
γ_{y2}	0.125	normal	0.25	0.1548	-0.0161	0.3268
ρ_{a2}	0.5	uniform	0.2887	0.8665	0.8162	0.9155
ρ_{g2}	0.5	uniform	0.2887	0.8669	0.8250	0.9068
ρ_{r2}	0.5	uniform	0.2887	0.7704	0.7337	0.8022
ω_3	0.5	uniform	0.2887	0.6378	0.4261	0.8342
θ_{p3}	0.6	beta	0.1	0.6045	0.5581	0.6564
θ_{w3}	0.7	beta	0.1	0.6670	0.5413	0.7826
$\gamma_{\pi 3}$	1.5	normal	0.25	1.6217	1.3660	1.8402
γ_{y3}	0.125	normal	0.25	0.2664	0.0262	0.5216
ρ_{a3}	0.5	uniform	0.2887	0.8687	0.8073	0.9365
ρ_{g3}	0.5	uniform	0.2887	0.7090	0.6140	0.8082
ρ_{r3}	0.5	uniform	0.2887	0.7814	0.7473	0.8137

Table 6a: Estimation results for the model parameters for short data samples in a panel approach

param.	prior mean	prior	prior stdev	post.mean	95% HDR	
σ_{a1}	0.5	uniform	0.2887	0.0020	0.0000	0.0034
σ_{g1}	0.5	uniform	0.2887	0.0070	0.0025	0.0114
σ_{m1}	0.5	uniform	0.2887	0.0013	0.0005	0.0021
$\sigma_{\lambda1}$	0.5	uniform	0.2887	0.0452	0.0159	0.0711
σ_{a2}	0.5	uniform	0.2887	0.0047	0.0031	0.0063
σ_{g2}	0.5	uniform	0.2887	0.0090	0.0031	0.0063
σ_{m2}	0.5	uniform	0.2887	0.0017	0.0012	0.0023
$\sigma_{\lambda2}$	0.5	uniform	0.2887	0.0584	0.0365	0.0819
σ_{a3}	0.5	uniform	0.2887	0.0055	0.0037	0.0074
σ_{g3}	0.5	uniform	0.2887	0.0181	0.0145	0.0219
σ_{m3}	0.5	uniform	0.2887	0.0026	0.0020	0.0031
$\sigma_{\lambda3}$	0.5	uniform	0.2887	0.0512	0.0247	0.0757

Table 6b: Estimation results for the shocks variance for short data samples in a panel approach

These results show how pooling the data for some countries may update the researcher’s belief on parameters distribution. We notice through the comparison of posteriors relative to priors in appendix 6.5 and appendix 6.6, that the posteriors obtained by the estimation over the pooled data are more informative. The posteriors have also less inaccuracy using the panel approach than the basic model estimation for one country. The results show that the distribution of α includes its value used in data simulation. The distribution of ω is less flat and covers a smaller region than in one-country estimation, that means less inaccuracy. The posterior distributions of the rigidity parameters are also more different from priors than those of one-country estimation. Similarly, the posteriors of the policy parameters are also more informative using the panel approach. Concerning the autoregressive process parameters, the posteriors in a one-country estimation are multi-mode, which is due the non-informative priors specification. However, using the panel approach most of the autoregressive parameters posteriors have one mode.

These results give an interesting conclusion, for countries who have short data samples available, the bayesian estimation in a panel approach can give a better update about parameters distributions than separate estimation. This result is constrained by the presence of some likely dynamics across countries. A heterogeneity across countries could also increase the risk of parameters.

5 Conclusion

The Bayesian analysis in economics has known many recent developments. The contribution of this paper is not only to explain the Bayesian methodology of estimation, but also to apply this analysis in a panel approach. We presented in the core of this paper some illustrations for the panel approach. We presented how the panel approach can update the researcher's belief about parameters possible values assuming the same prior distributions across countries.

The main results of this paper state that accuracy about parameters and about policy choice depends on the form of the prior probability distribution in a Bayesian analysis. Non-informative priors specification for some parameters imply some flat posteriors or highly risk posteriors. In the context of short data samples, the prior choice has a big role in the posterior determination, since the likelihood distribution is not enough informative to update the parameters distribution.

However, under similarity of some dynamic characteristics among a panel of countries, estimating a DSGE model in a panel approach using the pooled data can reduce the importance of the prior distribution. We have seen that the posteriors are more informative in a panel dimension estimation than in one-country dimension. This is due to the use of more available information from other countries data.

In general, passing to a panel dimension reduces the inaccuracy of parameters. It reduces also the bayesian risk of the policy parameters and thus reduces the risk of the policy choice. This result is available if the data have some similar dynamics across countries. Otherwise, if the dynamic characteristics across countries do not satisfy some homogeneity criteria, the posteriors may be less informative since posteriors may be flat or have many modes.

If some parameters are supposed to be identical in the estimation process, these parameters can absorb information from other countries data, if the parameters are really identical. However, if this restriction of identical parameters across countries is a specification error, i.e: parameters have a large heterogeneity across countries, then the posteriors obtained would be either less informative with flat distribution or with many modes.

To conclude this paper, we found that estimating a DSGE model using Bayesian analysis for a panel of countries can be more informative about parameters distribution than one-country dimension. This specification must be done with a big attention about dynamic characteristics similarity. Any error of specification of identical parameters restriction may drive to a non-informative posterior distributions and to some identification problems. We showed in section 4.4.2 how a specifi-

cation of a common monetary policy, while the data are simulated with models containing different monetary policies, increases the uncertainty of the monetary policy parameters choice and also the uncertainty of other parameters.

In a panel approach, researchers are used to deal with random or fixed specific effects in the models, supposing that parameters are identical across countries. This point is interesting to test using the Bayesian analysis. However, in DSGE models, the assumption of identical parameters may be very restrictive and may introduce a specification error. This point will be left for further research in Bayesian analysis.

This paper is a methodology illustration in a panel approach. It gave a more clear vision on the Bayesian analysis under some specifications. For future research, it will be interesting to study the Bayesian estimation of DSGE models on real data of a panel of countries in the same region. Some Central Eastern Europe countries, since they have short data problems, may be interesting to study.

References

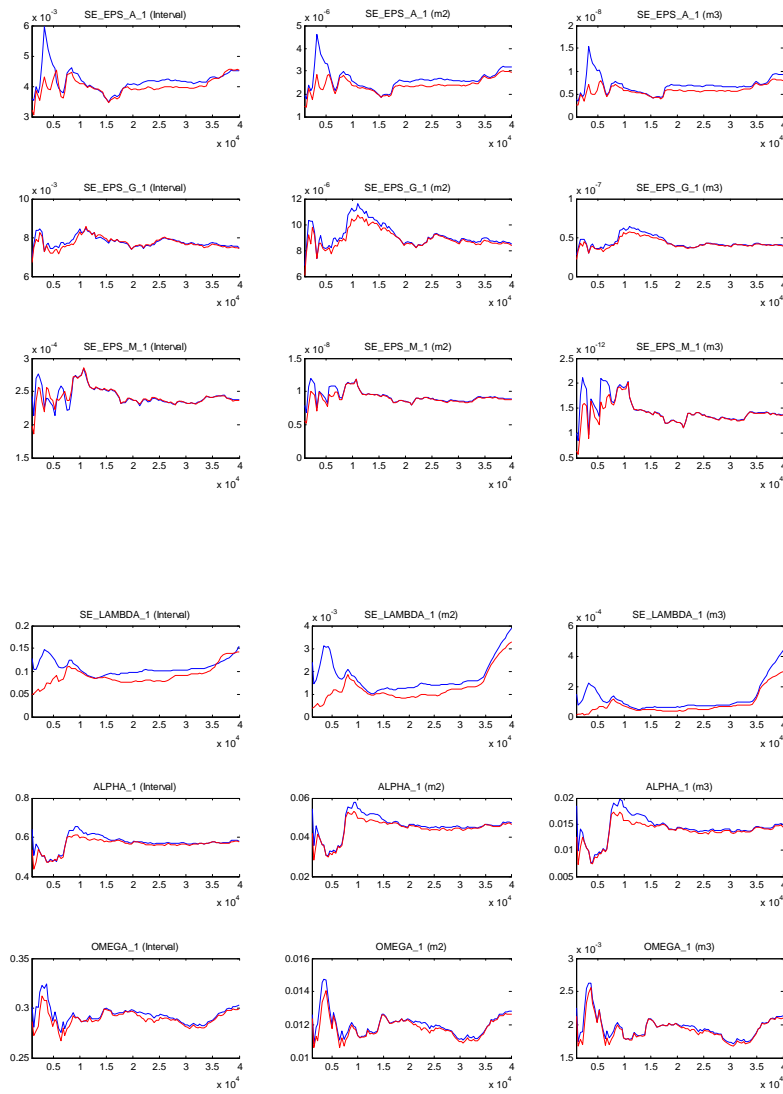
- [1] Alvarez, J. and M. Arellano (2003), "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators", *Econometrica*, vol 71, no. 4, pp 1121-1159.
- [2] An S. and F. Schorfheide (2006), "Bayesian Analysis of DSGE models", working paper no. 06-5, Federal Reserve Bank of Philadelphia.
- [3] Anderson G. (2000), "A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models", Board of Governors of the Federal Reserve System.
- [4] Anderson G. (2006), "Solving Linear Rational Expectations Models: A Horse Race", Board of Governors of the Federal Reserve System.
- [5] Aykroyd, R. G., "Bayesian Estimation for Homogenous and Inhomogenous Gaussian Random Fields", Department of Statistics, University of Leeds.
- [6] Calvo G. (1983) "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics*, vol 12, pp 383-398.
- [7] Canova, F., Ciccarelli, M. and E. Ortega (2004), "Similarities and Convergence in G-7 Cycles".
- [8] Chamberlain, G. (1982), "Panel Data", NBER working paper no. 913
- [9] Chang, Y., Gomes, J. and F. Schorfheide (2002), "Learning by doing", *The American Economic Review*, vol. 92, no. 5, pp 1498-1520.
- [10] Christiano, L.J., Eichenbaum, M. and C. Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", *Journal of Political Economy*, vol. 113, no. 1.
- [11] DeJong, D., Ingram, B and C. Whiteman (2000), "A Bayesian approach to dynamic macroeconomics", *Journal of Econometrics* 98, pp 203 - 223.
- [12] Del Negro M. and F. Schorfheide (2006), "Forming Priors for DSGE Models", working paper no. 2006-16, Federal Reserve Bank of Atlanta.
- [13] Erceg, C., Henderson D. and A. Levin (2000), "Optimal monetary policy with staggered wage and price contracts", *Journal of Monetary Economics*, vol 46, pp 281-313.
- [14] Gali J., Gertler, M. and D. Lopez-Salido (2001), "European inflation dynamics". *European Economic Review* 45, pp 1237-1270.
- [15] Geweke, J. (2005), "*Contemporary Bayesian Econometrics and Statistics*", Chapters 3 & 4, WILEY editions, UK.
- [16] Hamilton, J. (1994), "*Time Series Analysis*", Chapter 13, Princeton University Press.
- [17] Harvey, A. (1989), "*Forecasting, structural time series models and the Kalman filter*", Chapter 3, Cambridge University Press.

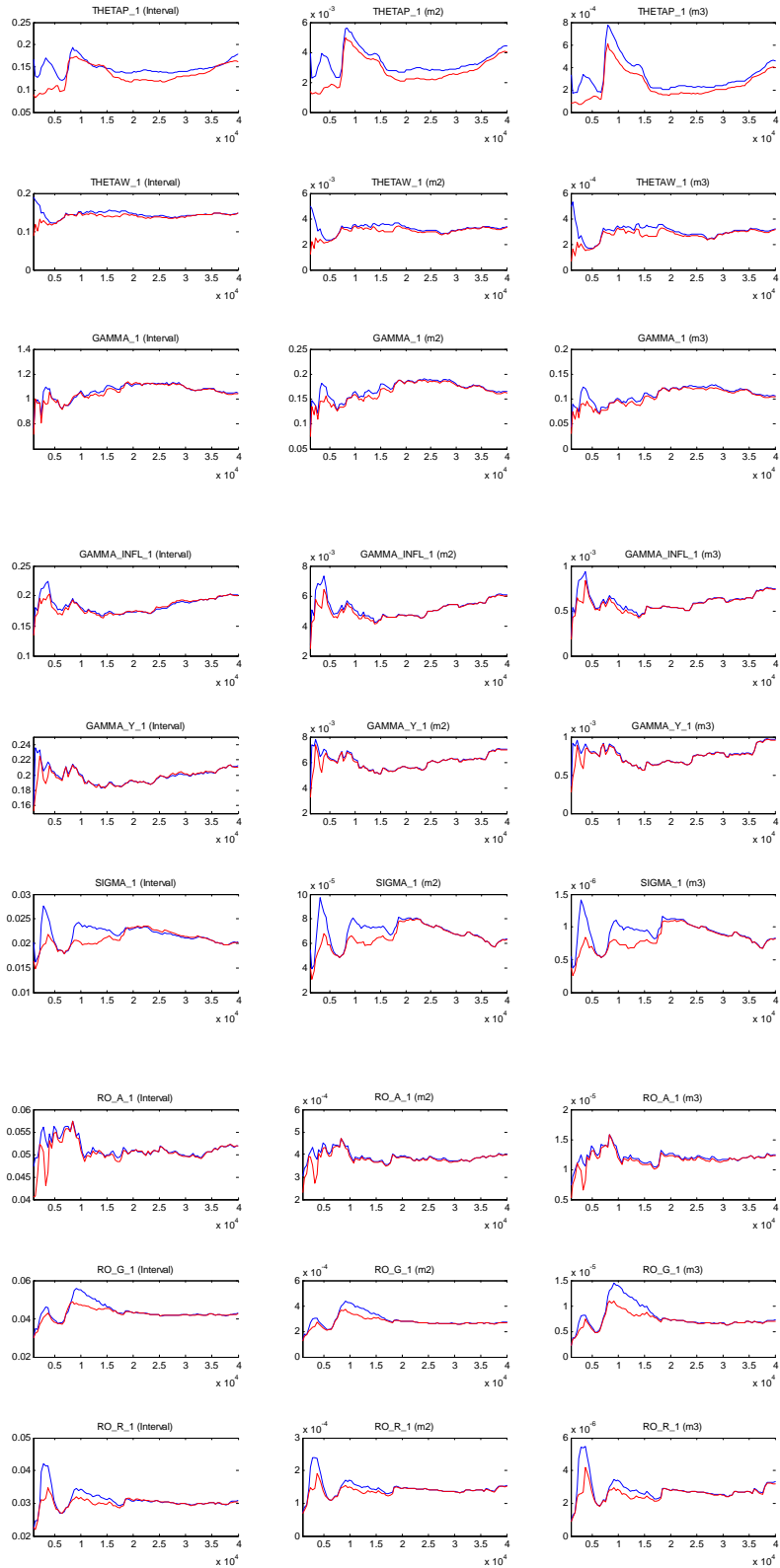
- [18] Hsiao, C. (2002), "*Analysis of Panel Data*", second edition, Chapters 4 & 9, Cambridge University Press.
- [19] Klein, L. (1954), "Textbook of Econometrics", *Revue Economique*, Vol 5, No. 4.
- [20] Koop, G., Osiewalski, J. and M. Steel (2000), "Modeling the sources of Output growth in a panel of countries", *Journal of Business & Economic Statistics*.
- [21] Koop, G. (2003), "*Bayesian Econometrics*", Chapters 7 & 8, WILEY editions, UK.
- [22] Koop, G., Leon-Gonzalez, R. and R. Strachan (2006), "Bayesian Inference in a Cointegrating Panel Data Model", working paper no. 06/2, University of Leicester, UK.
- [23] Landon-Lane, J. (1998), "Bayesian Comparison of Dynamic Macroeconomic Models",
- [24] Peng, Z. and D. Small, "Bayesian Inference for Random Coefficient Dynamic Panel Data Models", Department of Statistics, The Wharton School, University of Pennsylvania.
- [25] Poirier (1995), "*Intermediate Statistics and Econometrics: A Comparative Approach*", David Giles.
- [26] Rabanal P. and J.F. Rubio-Ramirez (2003), "Comparing New Keynesian Models in the Euro Area: A Bayesian Approach", working paper 2003-30, Federal Reserve Bank of Atlanta.
- [27] Smets, F. and R. Wouters (2002b), "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area", working paper no. 171, European Central Bank.
- [28] Sommacampagna, C. and G. Sick (2005), "Estimation of Volatility of Cross Sectional Data: a Kalman Filter Approach".
- [29] Theil, H. (1978), "*Introduction to Econometrics*", Englewood Cliffs, N.J: Prentice-Hall.
- [30] Zellner, A. (1971), "*An Introduction to Bayesian Inference in Econometrics*", Wiley & Sons editions.
- [31] Zellner, A. (2006), "Bayesian Econometrics: Past, Present and Future", working paper no. 0607, University of Chicago.

6 Appendix

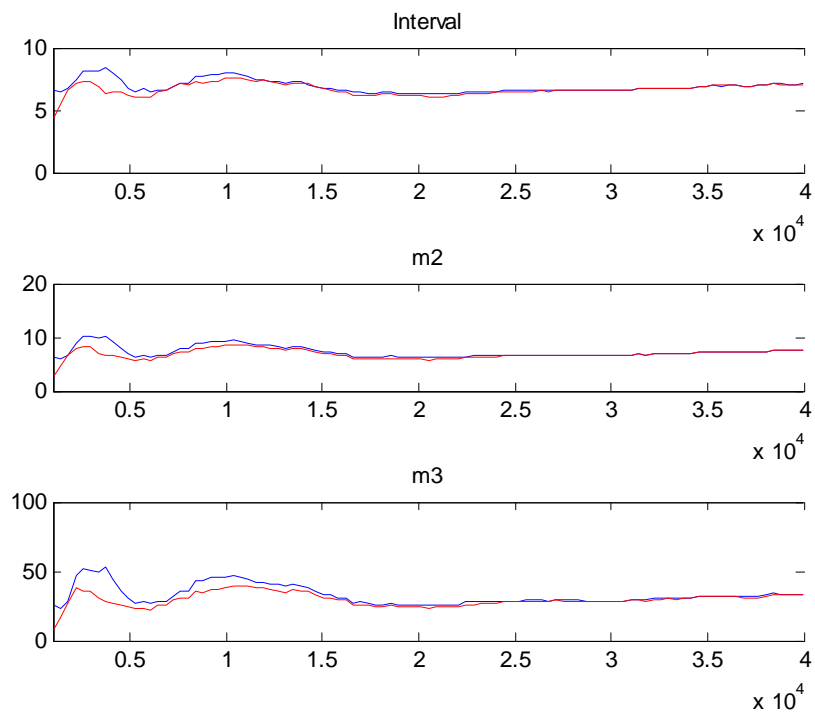
6.1 The basic model estimation results

MCMC univariate diagram of the Metropolis-Hastings draws:

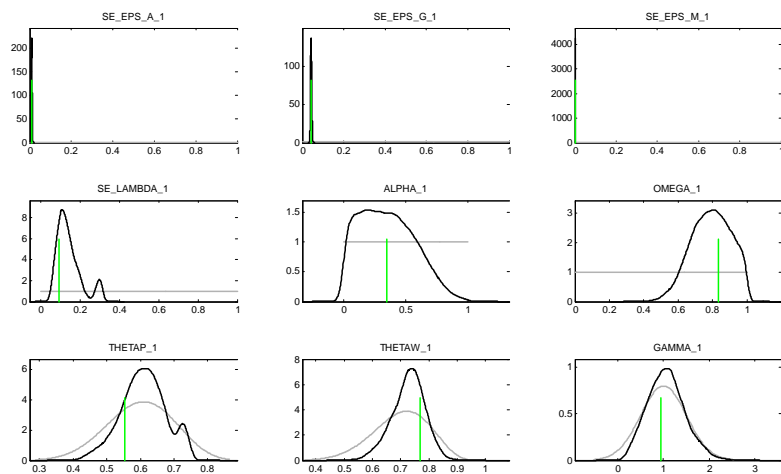


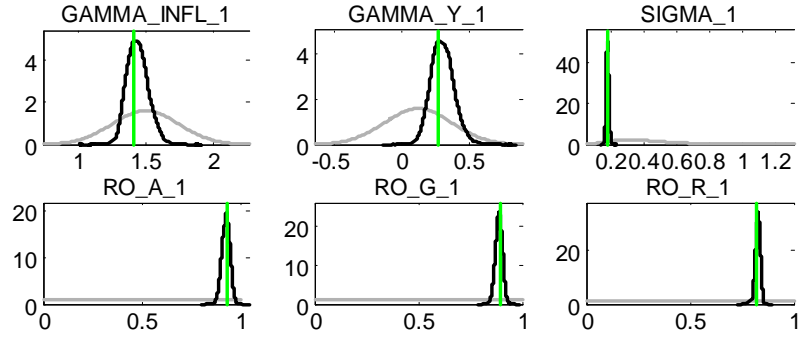


MCMC multivariate diagram of the Metropolis-Hastings draws



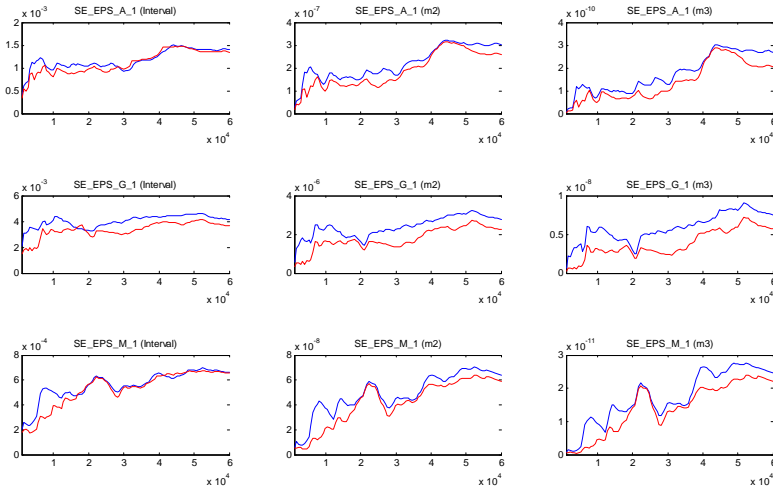
Priors (light lines) and posteriors (dark lines)

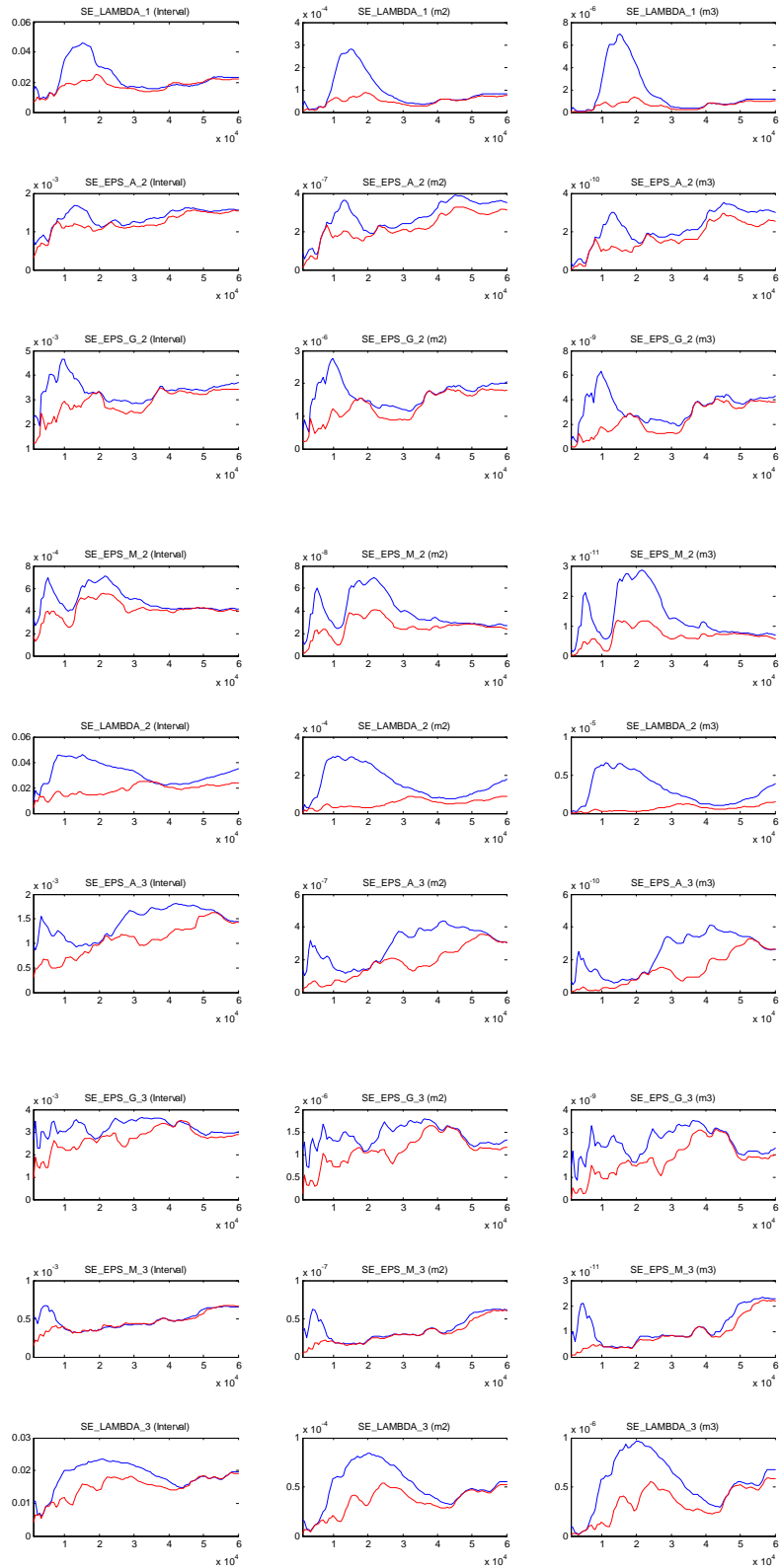


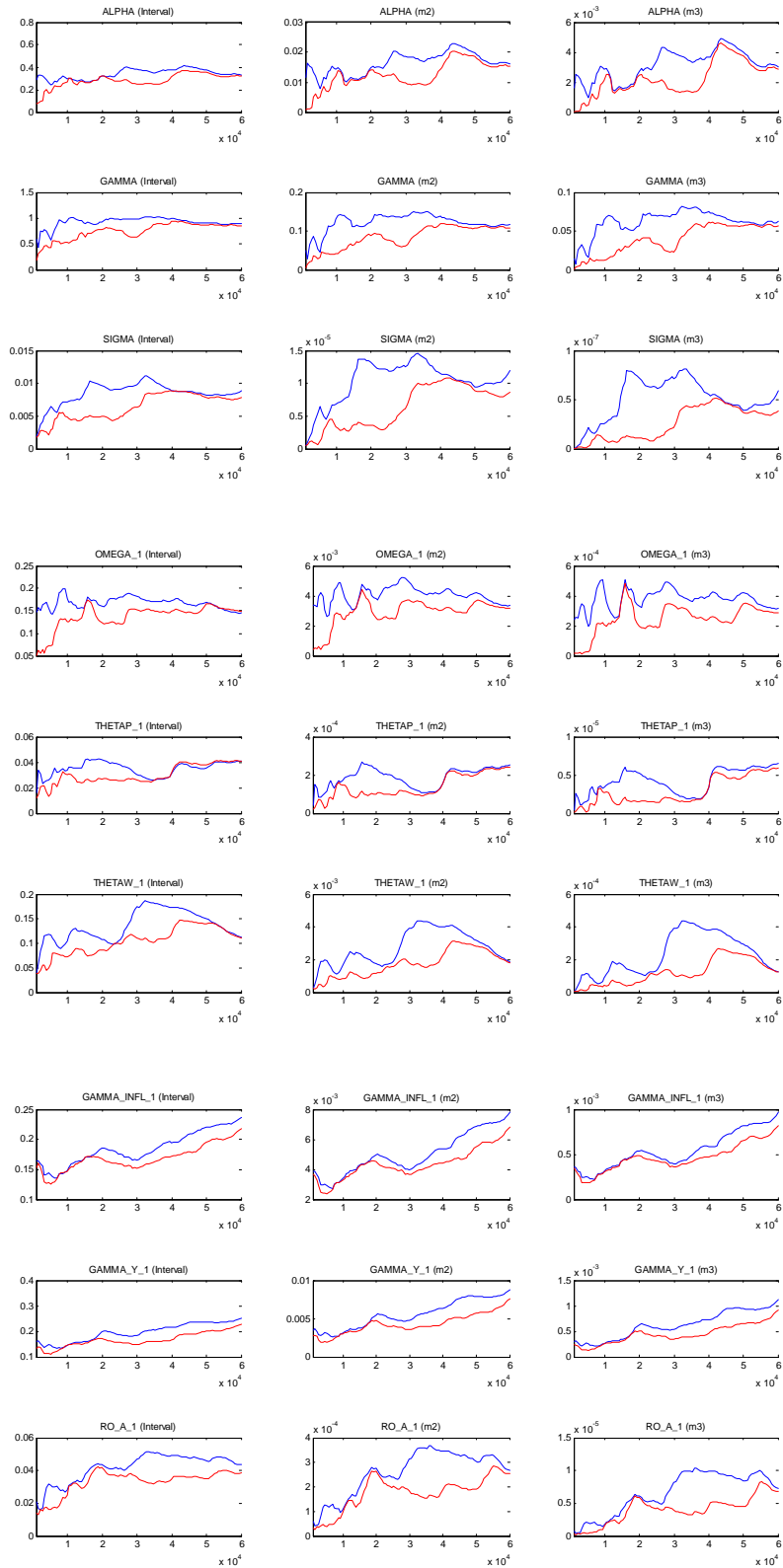


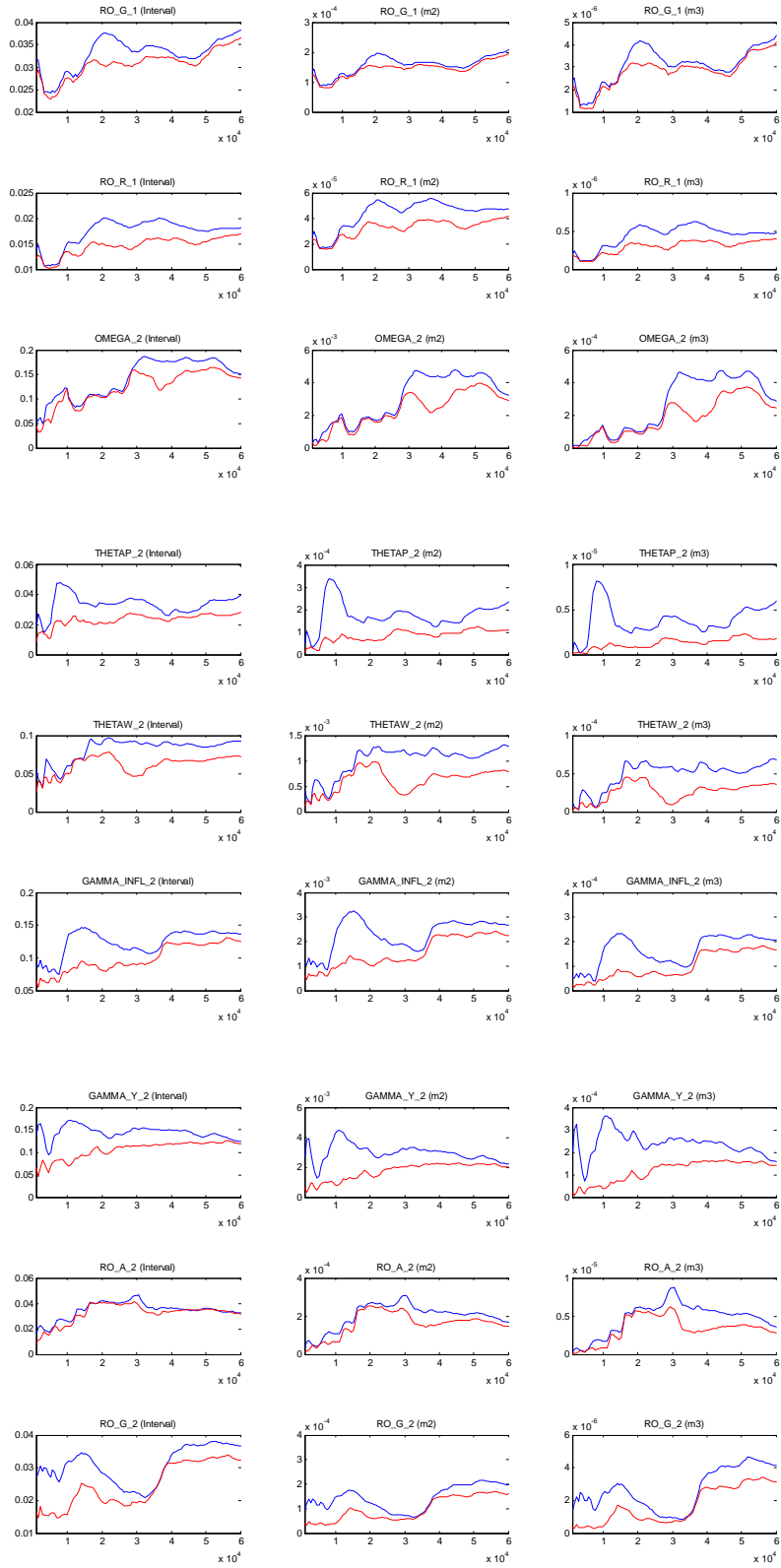
6.2 Estimation results of the model in a Panel approach

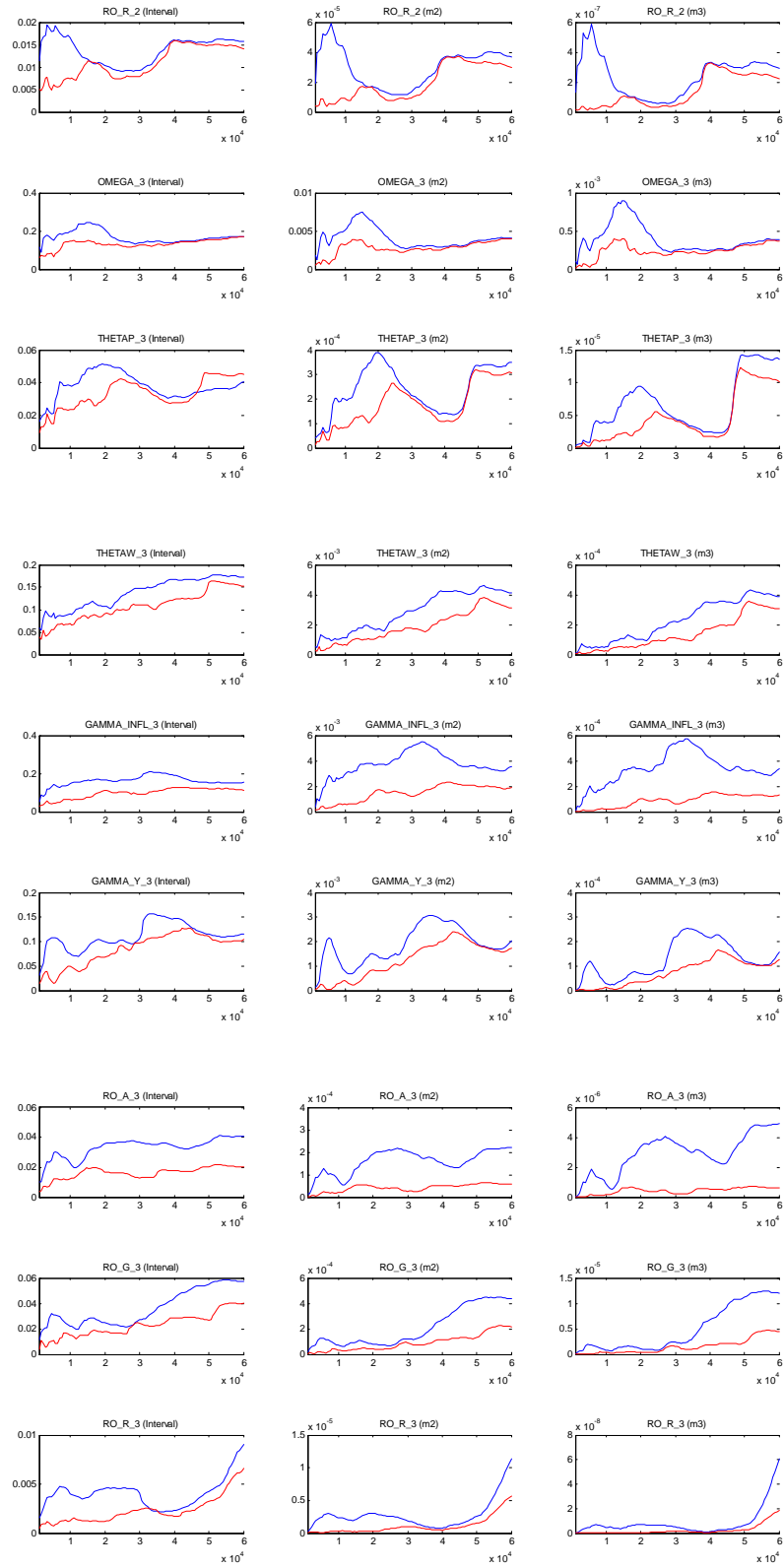
MCMC univariate diagram of the Metropolis-Hastings draws:

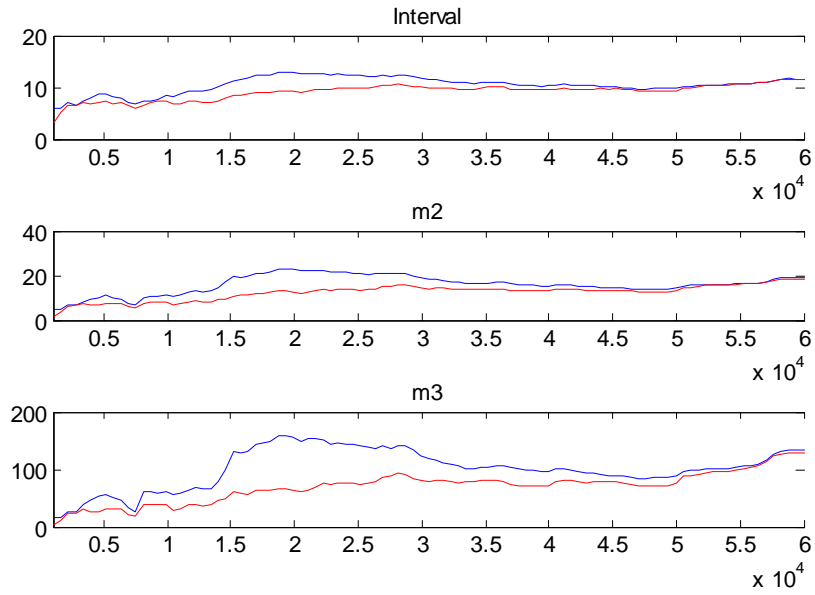




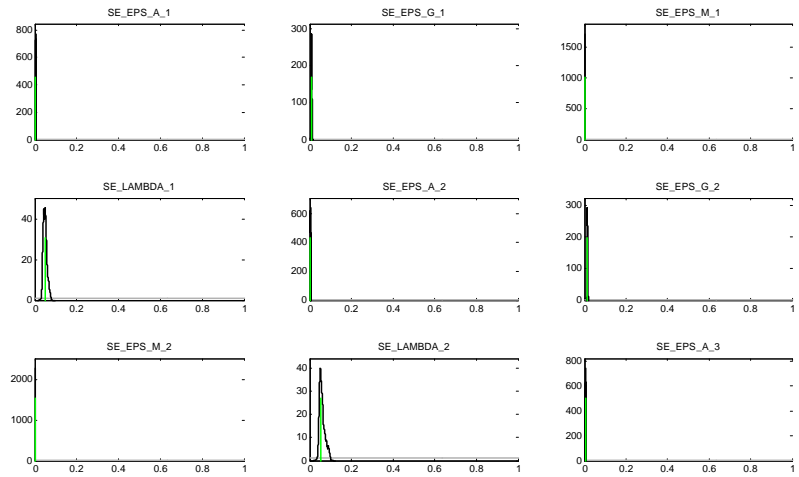


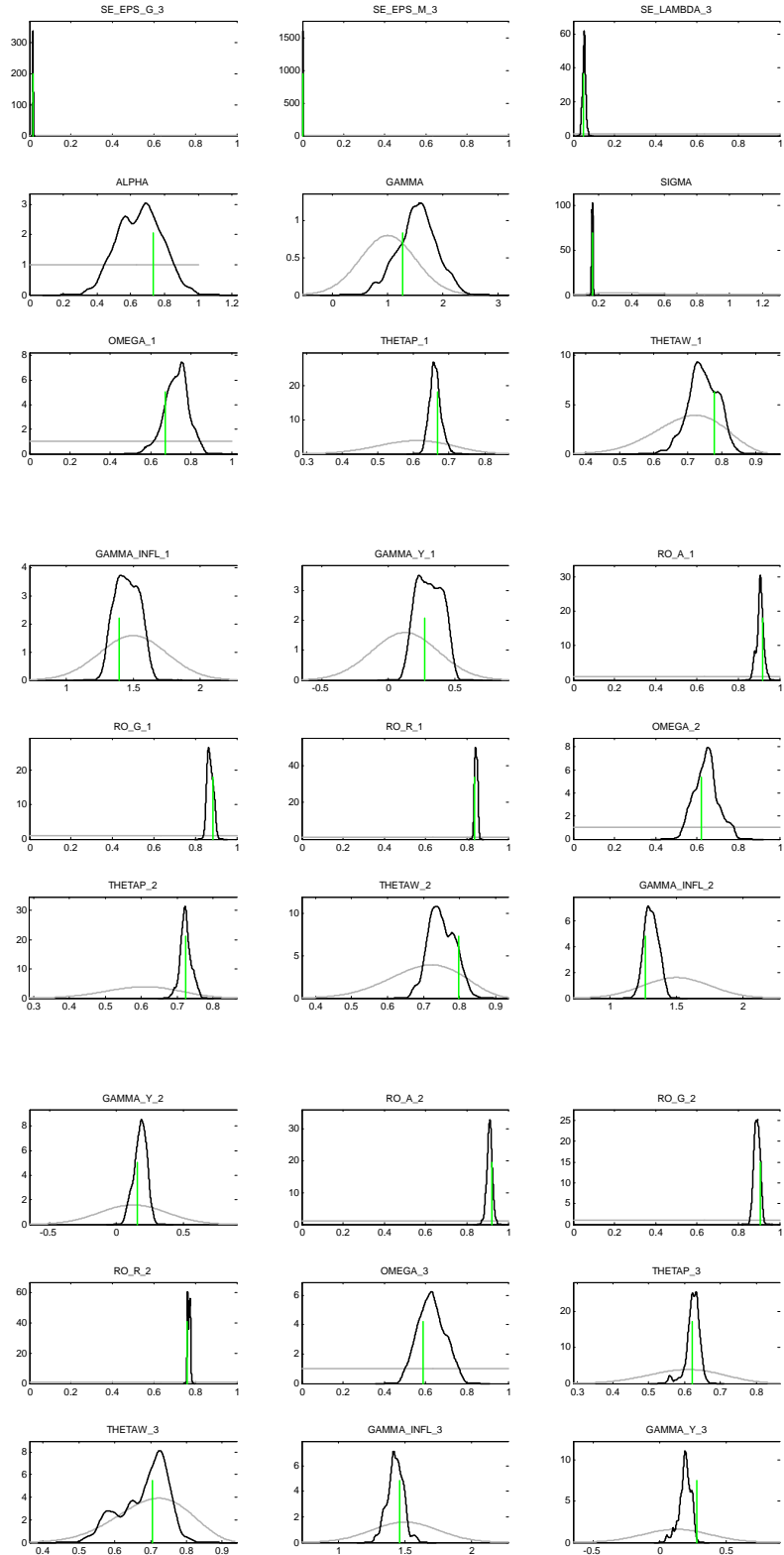






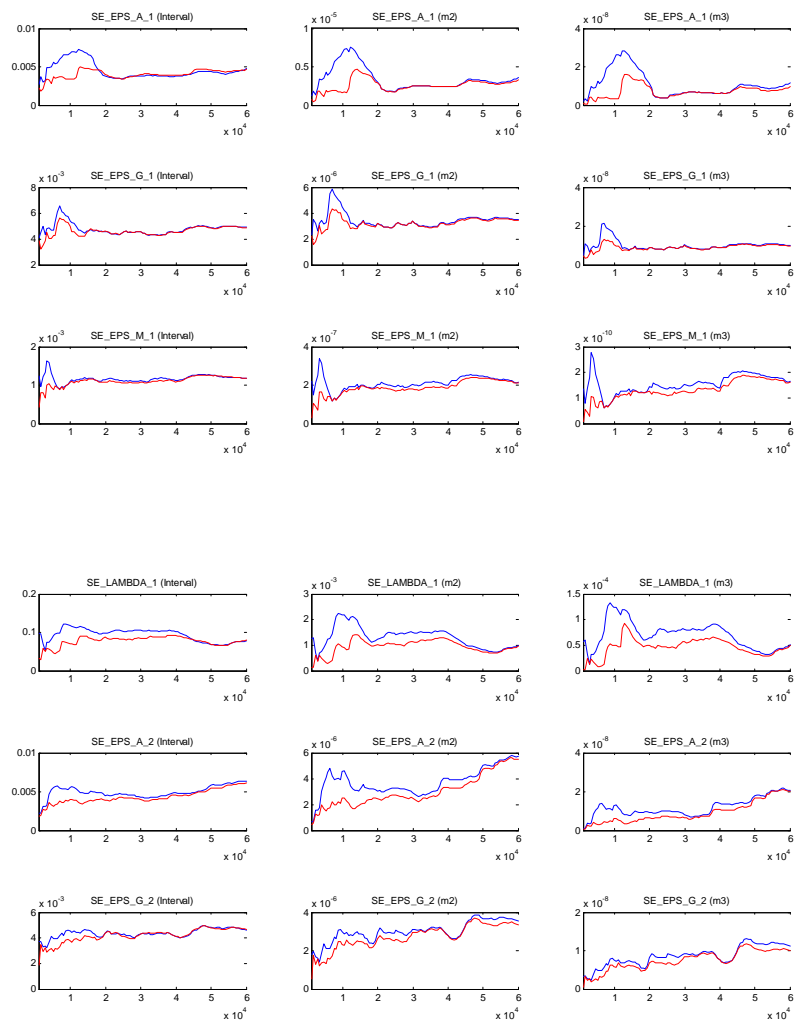
Priors (light lines) and posteriors (dark lines)

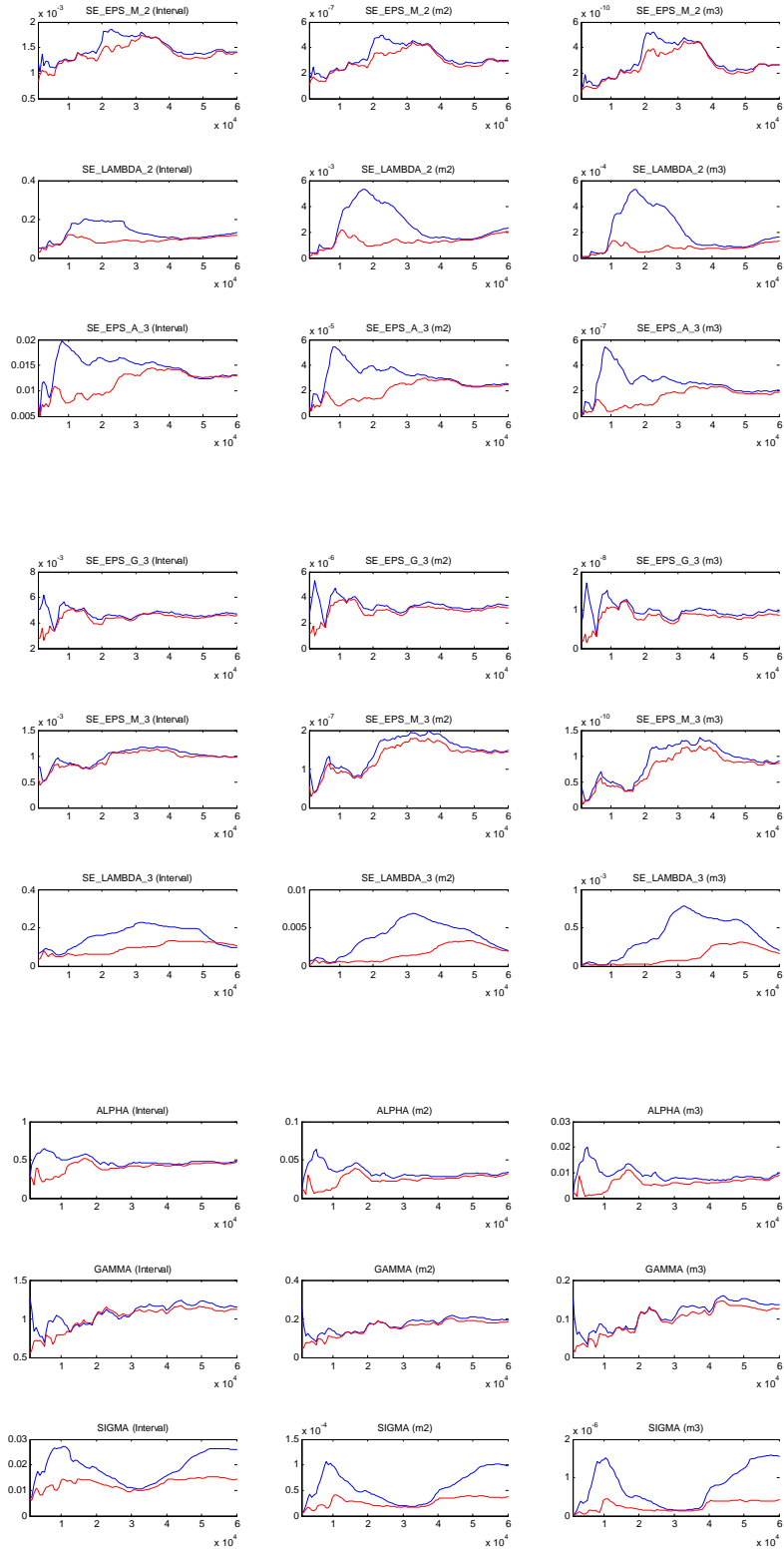


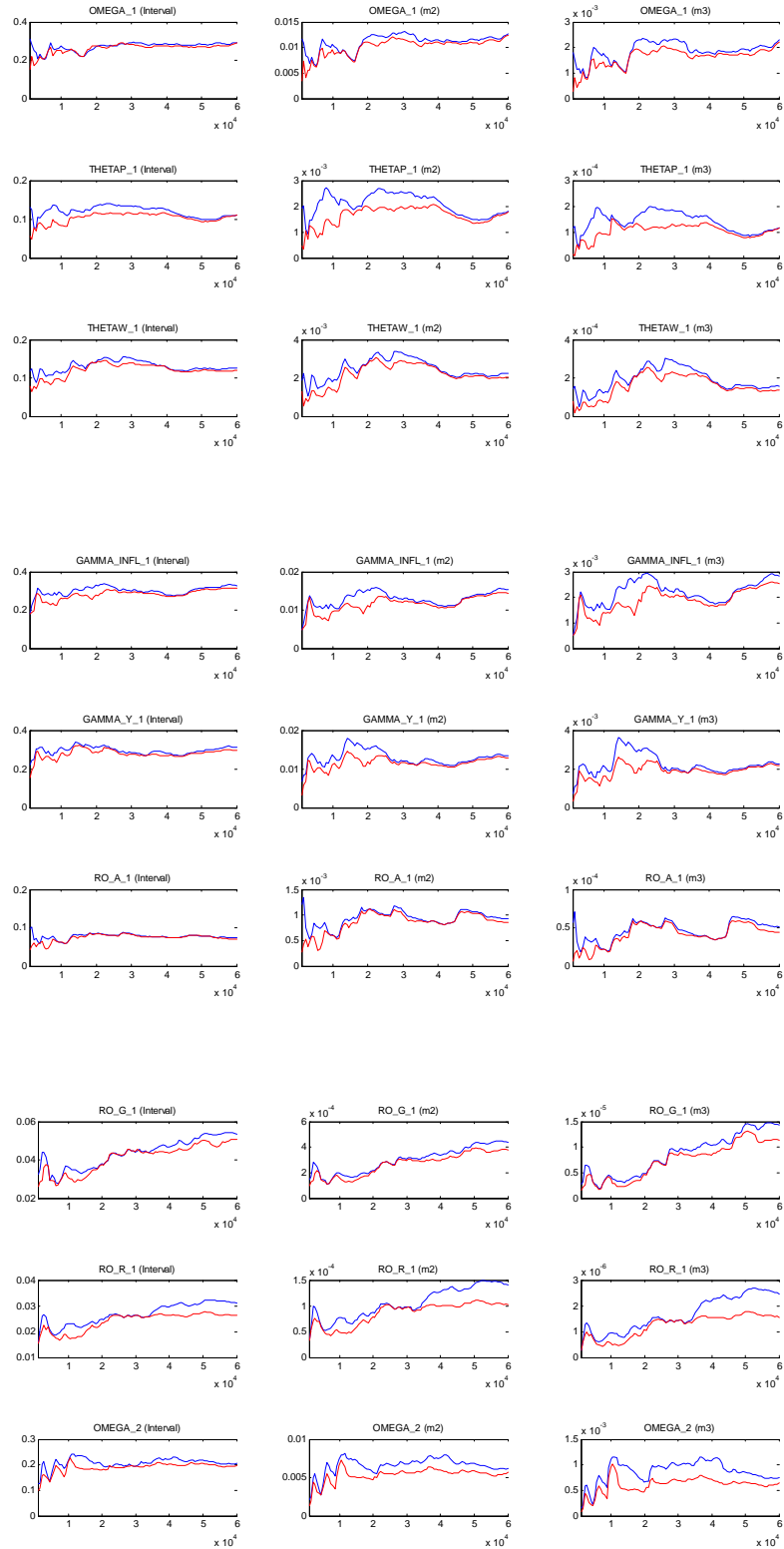


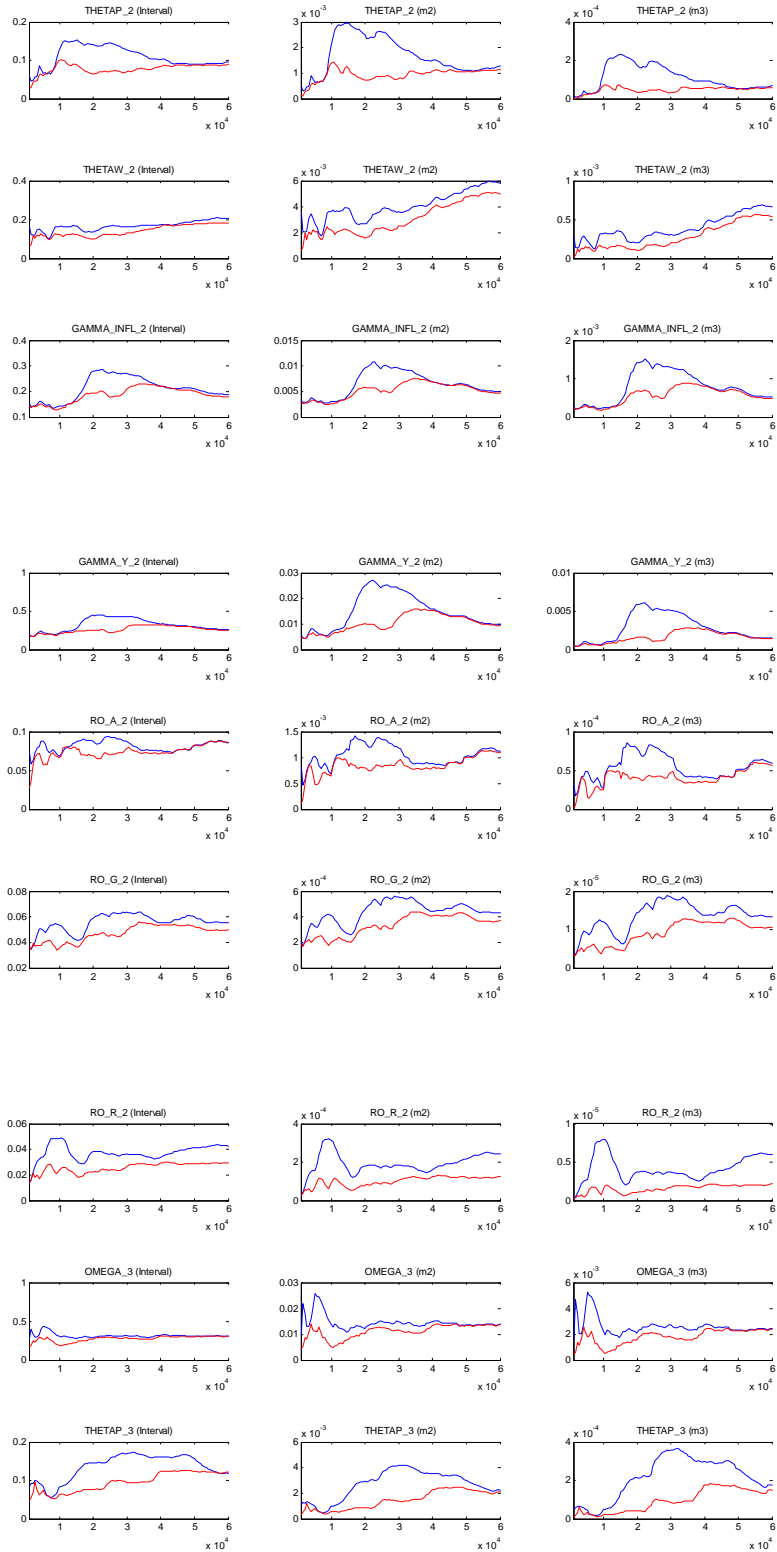
6.3 Estimation results of the model in a panel approach with heterogeneity

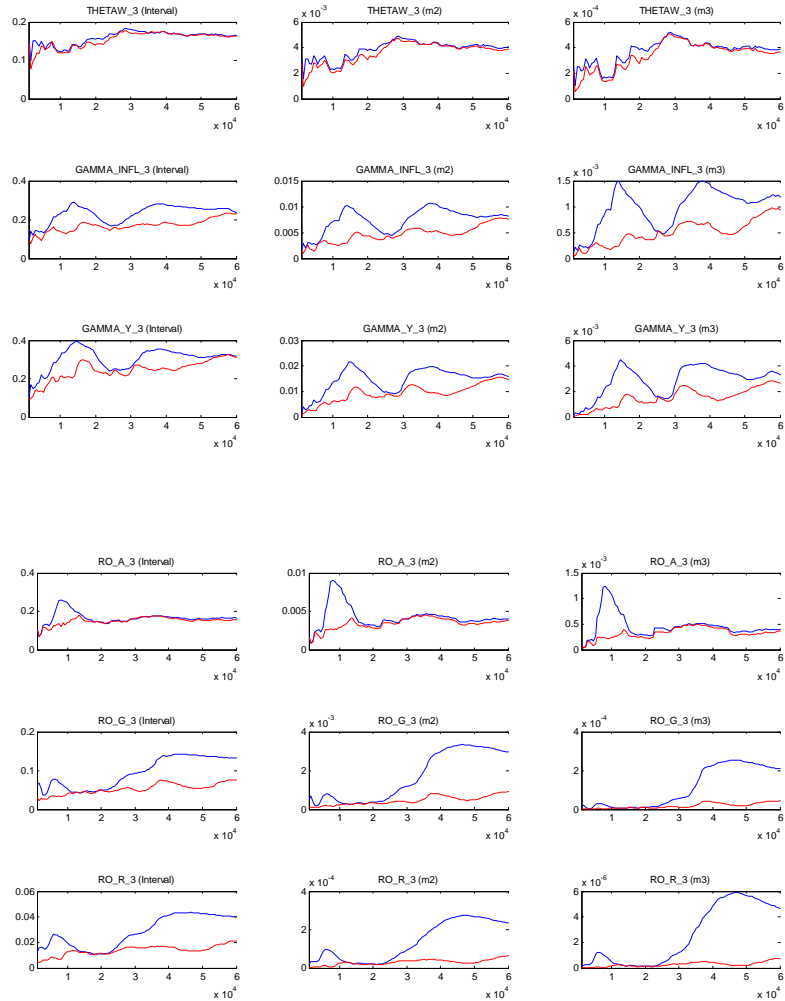
MCMC univariate diagram of the Metropolis-Hastings draws:

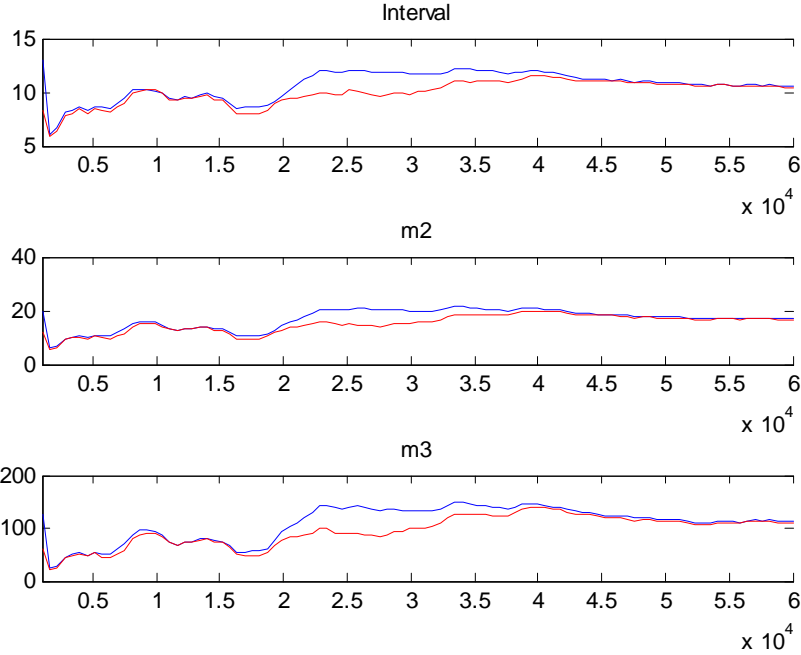




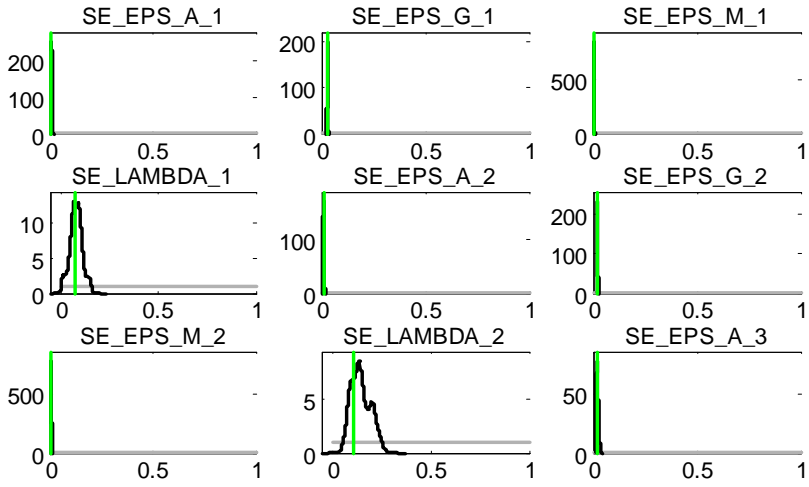


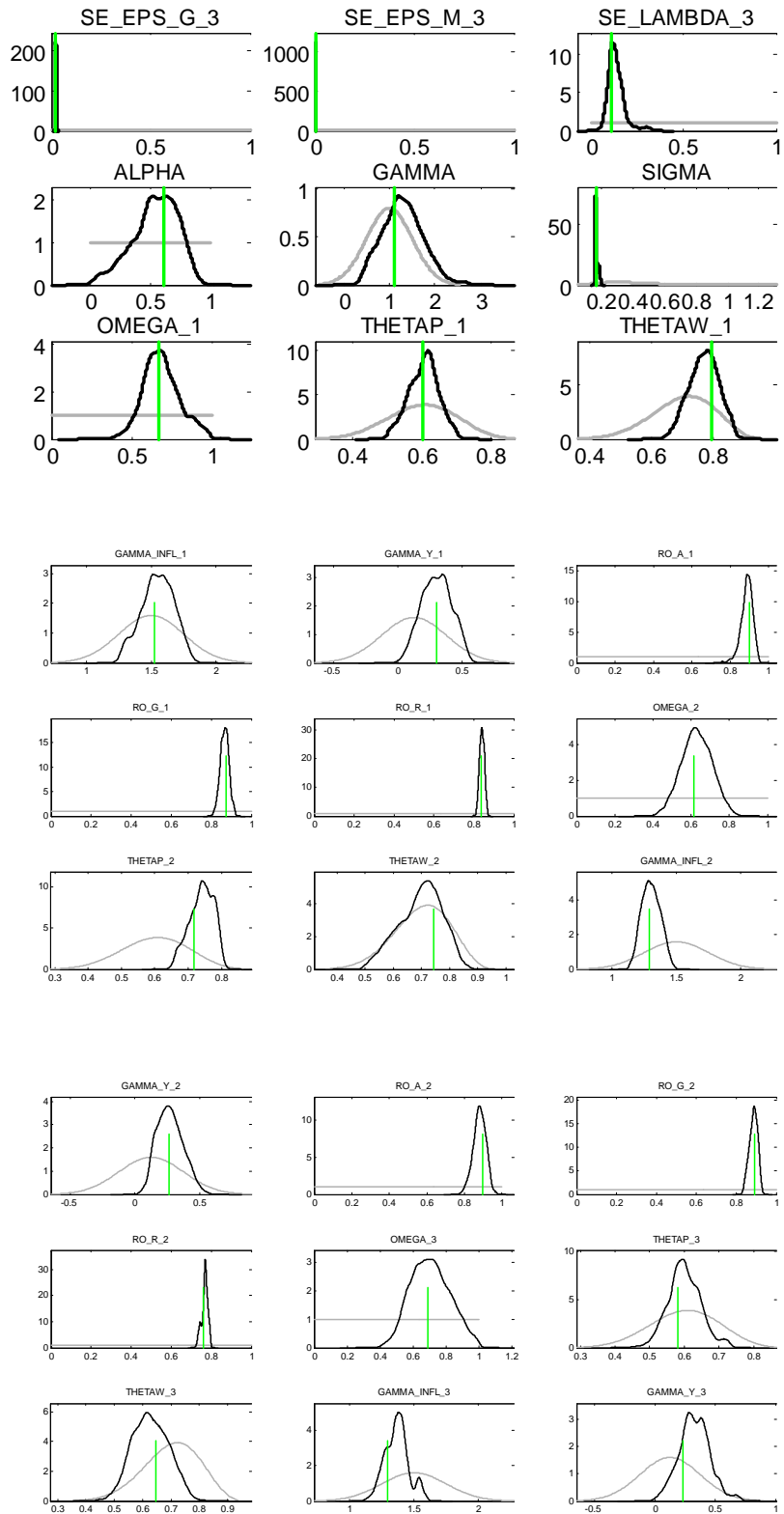






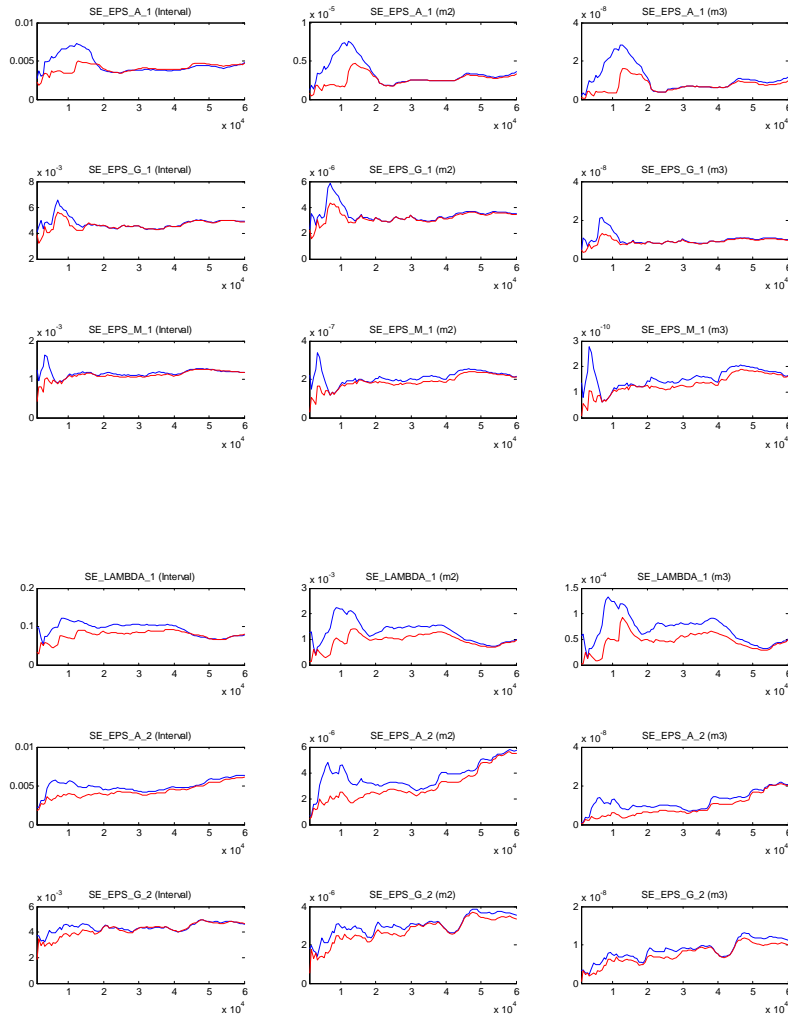
Priors (light lines) and posteriors (dark lines)

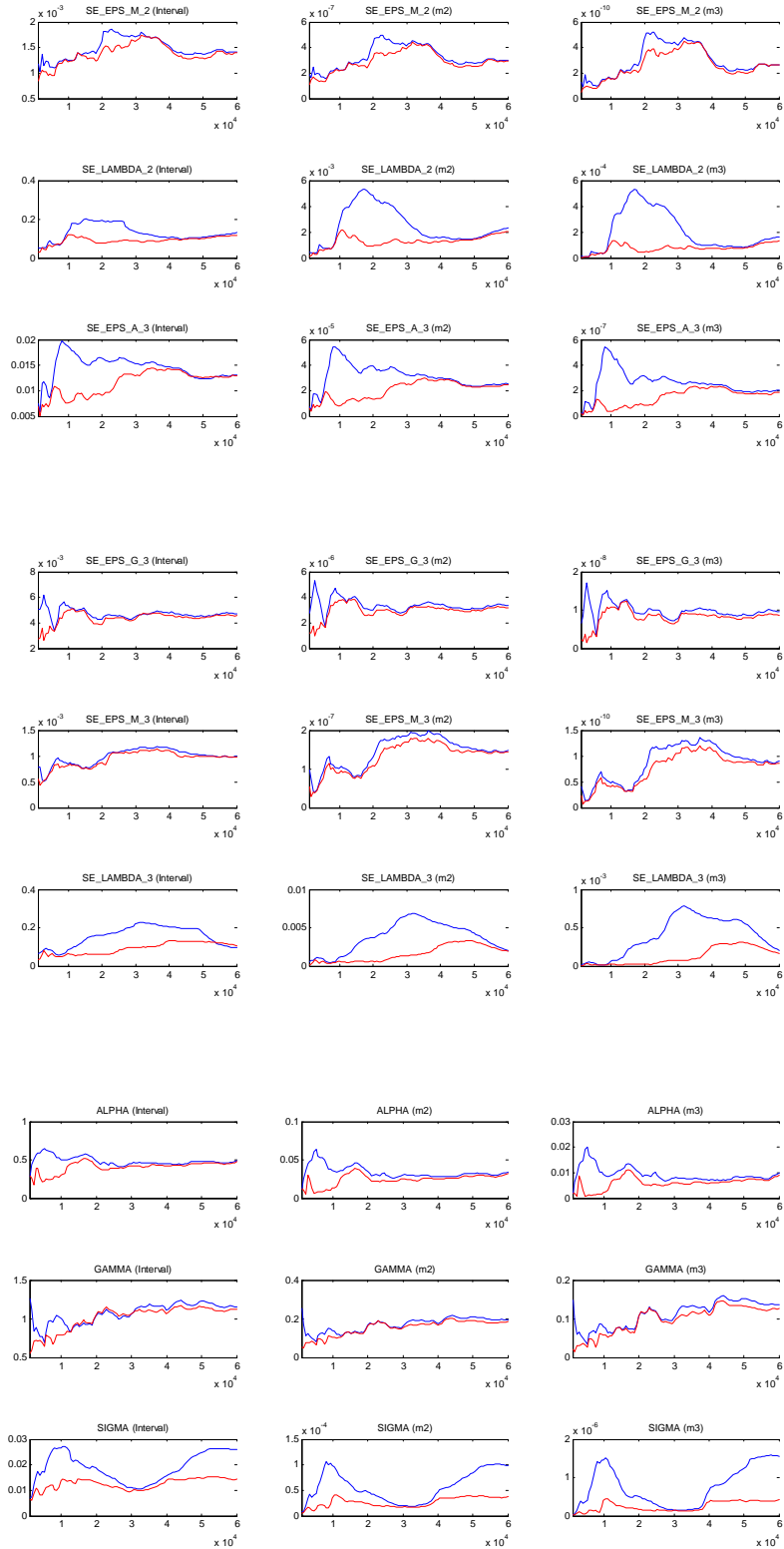


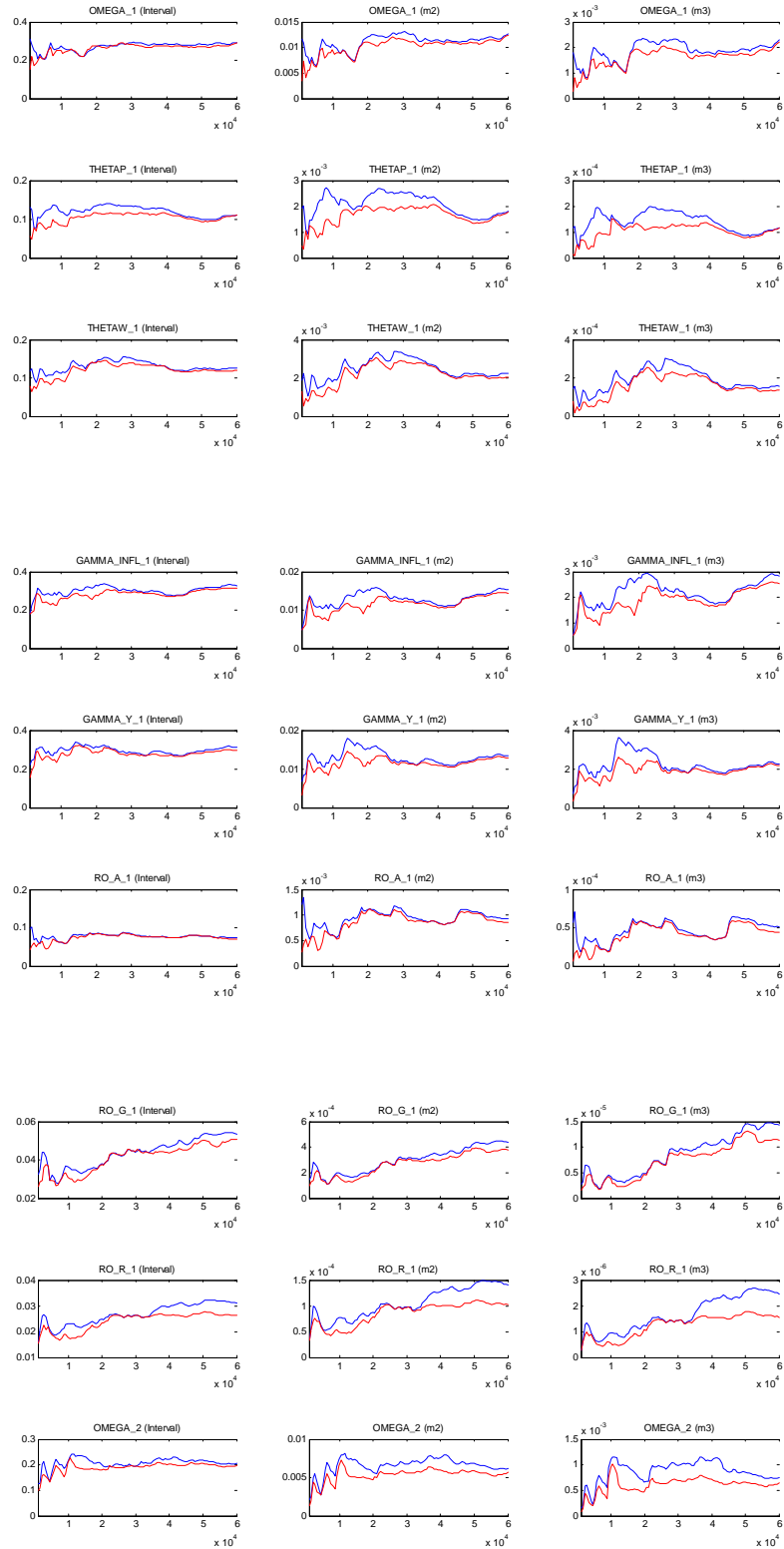


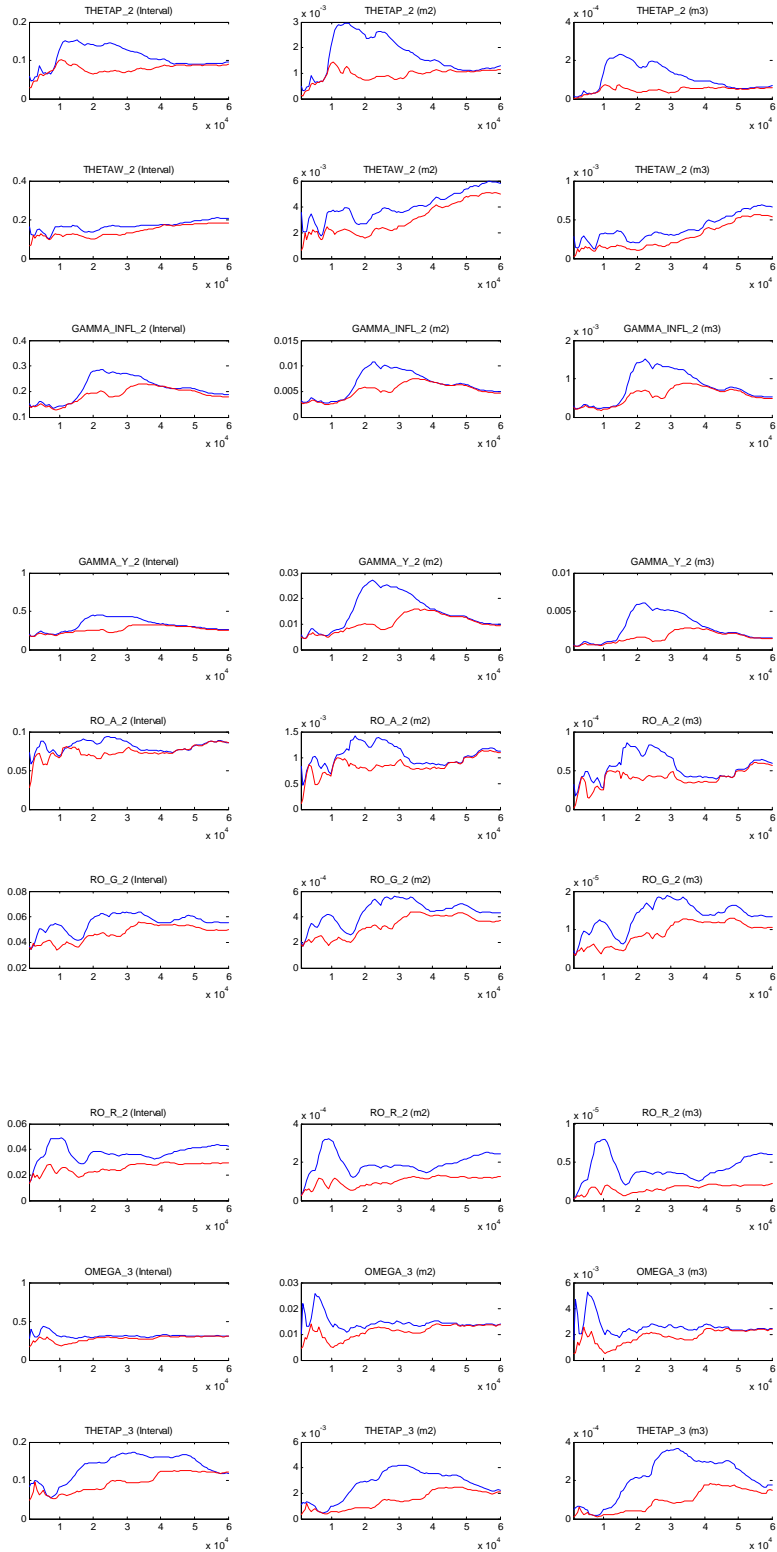
6.4 Estimation results of the model in a panel approach with heterogeneity of the monetary policy with a misspecification of restriction

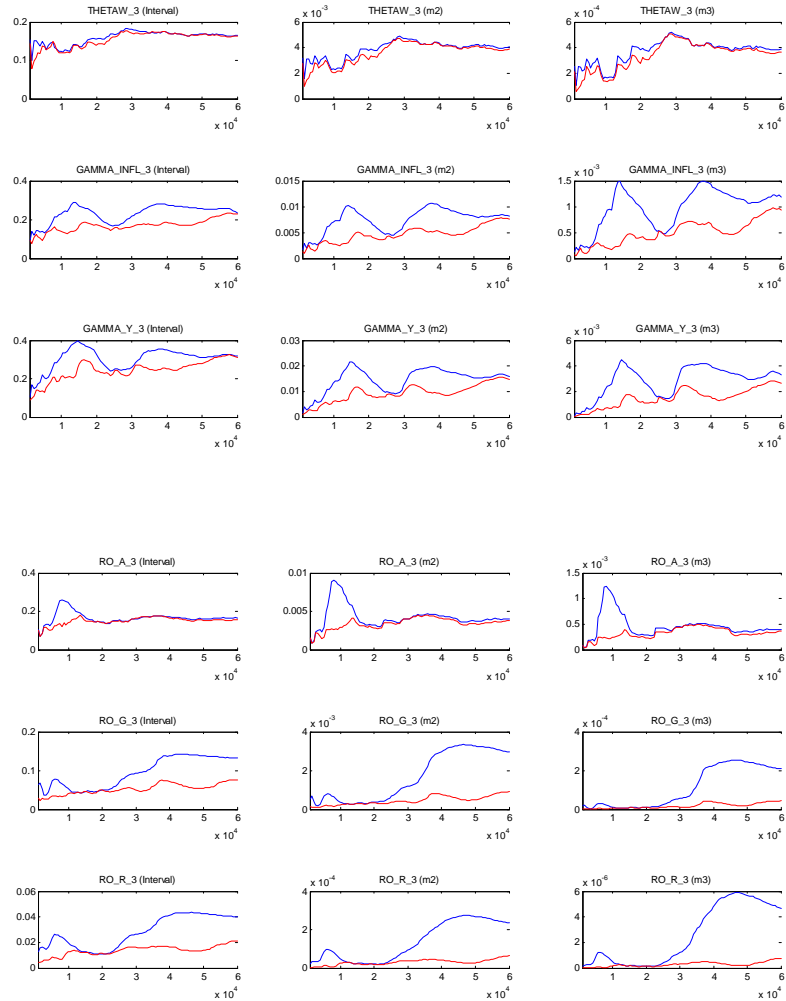
MCMC univariate diagram of the Metropolis-Hastings draws:

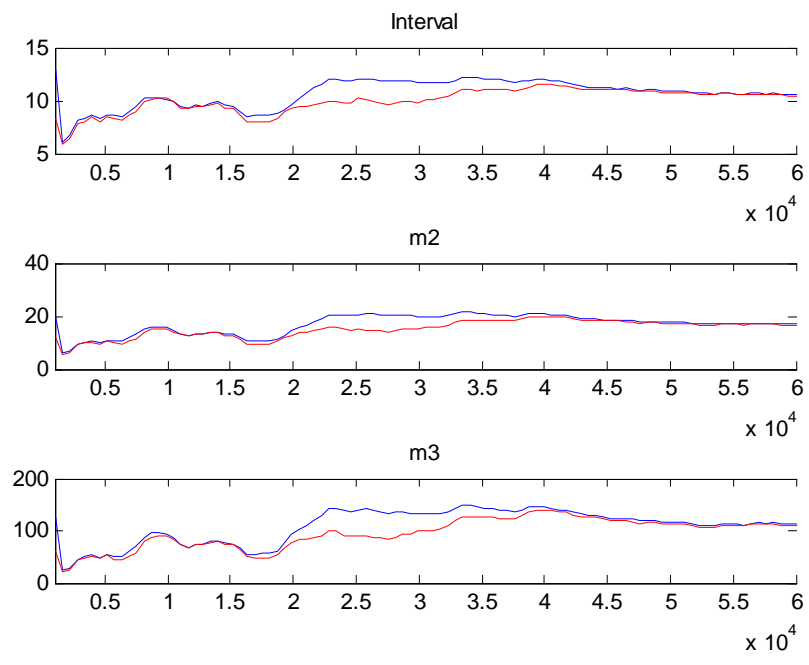




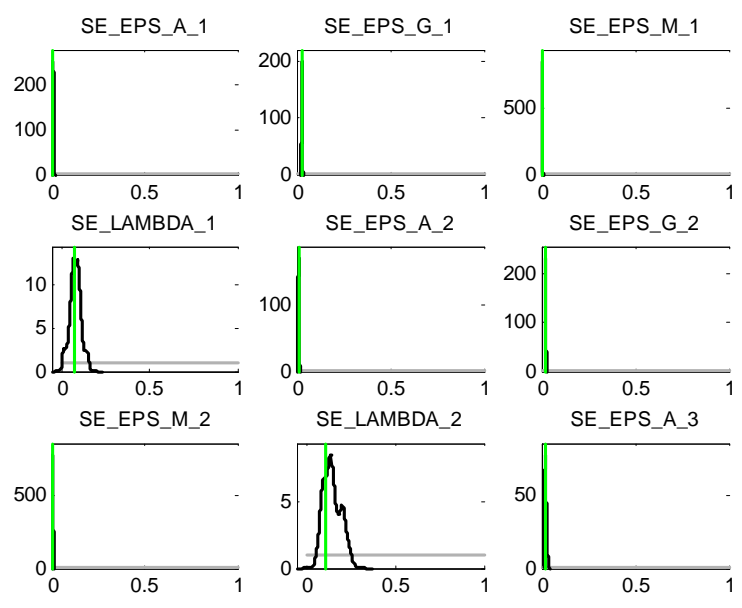


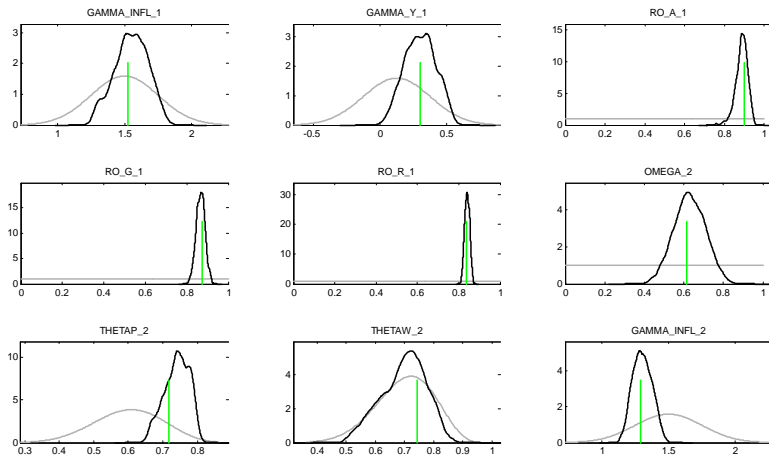
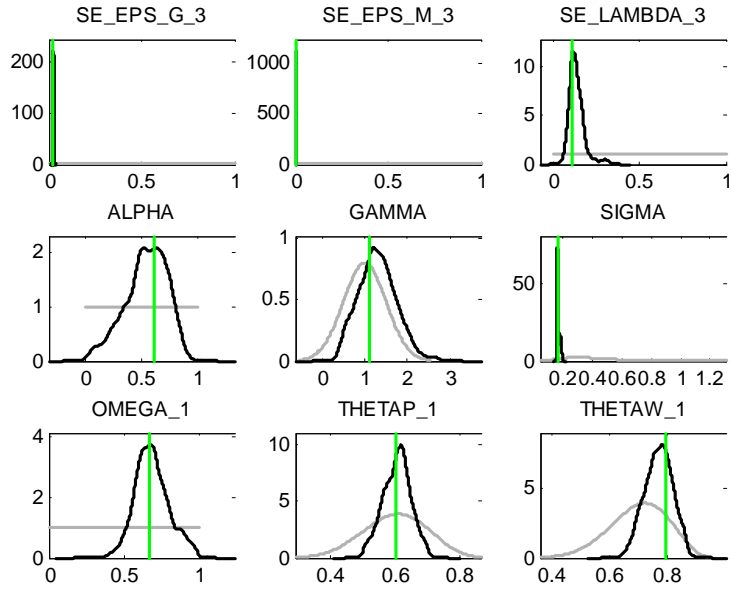


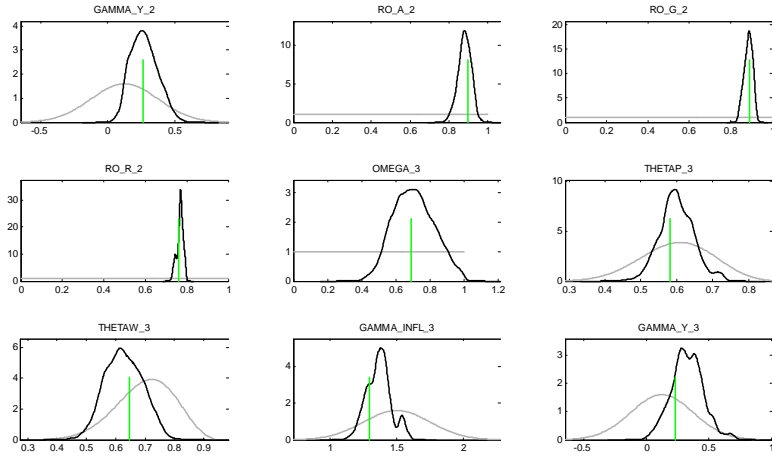




Priors (light lines) and posteriors (dark lines):

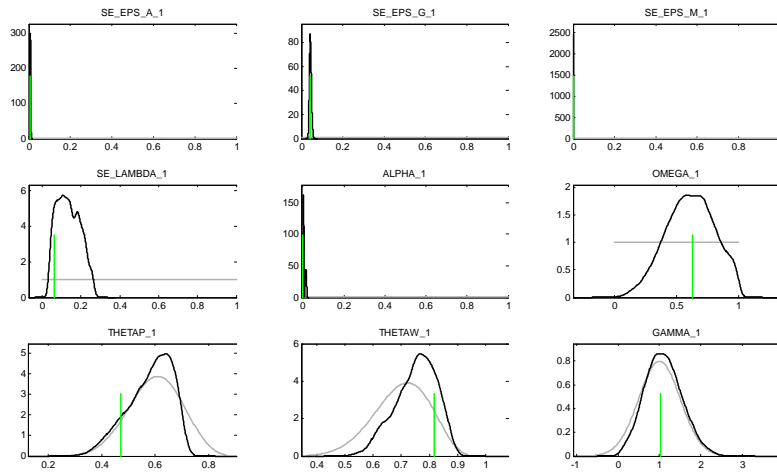


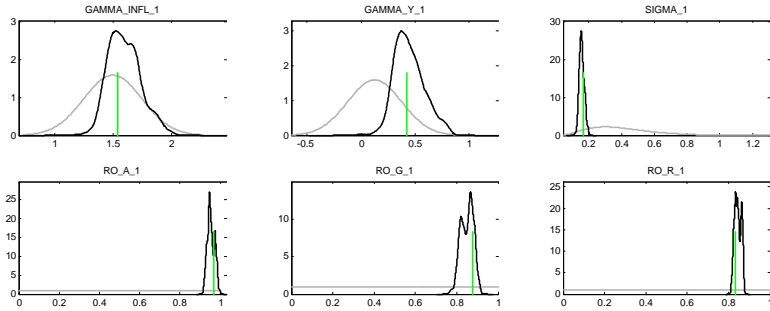




6.5 Estimation results of the basic model for one country with short data sample

Priors (light lines) and posteriors (dark lines):





6.6 Estimation results of the model in a panel approach with short data samples

Priors (light lines) and posteriors (dark lines):

