

Sacht, Stephen

Working Paper

Identification of prior information via moment-matching

Economics Working Paper, No. 2014-04

Provided in Cooperation with:

Christian-Albrechts-University of Kiel, Department of Economics

Suggested Citation: Sacht, Stephen (2014) : Identification of prior information via moment-matching, Economics Working Paper, No. 2014-04

This Version is available at:

<http://hdl.handle.net/10419/90812>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

C | A | U

Christian-Albrechts-Universität zu Kiel

Department of Economics

Economics Working Paper
No 2014-04

identification of prior information via moment-matching

by Stephen Sacht

issn 2193-2476



Identification of Prior Information via Moment-Matching

Stephen Sacht*

January 23, 2014

Abstract

In this paper we apply a sensitivity analysis regarding two types of prior information considered within the Bayesian estimation of a standard hybrid New-Keynesian model. In particular, we shed a light on the impact of micro- and macropriors on the estimation outcome. First, we investigate the impact of the transformation of those model parameters which are bounded to the unit interval, in order to allow for a more diffuse prior distribution. Second, we combine the Moment-Matching (MM; Franke et al. (2012)) and Bayesian technique in order to evaluate macropriors. In this respect we define a two-stage estimation procedure – the so-called Moment-Matching based Bayesian (MoMBay) estimation approach – where we take the point estimates evaluated via MM and consider them as prior mean values of the parameters within Bayesian estimation. We show that while (transformed) micropriors are often used in the literature, applying macropriors evaluated via the MoMBay approach leads to a better fit of the structural model to the data. Furthermore, there is evidence for intrinsic (degree of price indexation) rather than extrinsic (autocorrelation in the shock process) persistence — an observation which stands in contradiction to the results documented in the recent literature.

Keywords: Bayesian estimation; moment-matching estimation; mombay estimation, New-Keynesian model; micropriors; macropriors.

JEL classification: C11, C32, C52, E3.

*Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel & Department of Economics, Christian-Albrechts-University Kiel, Olshausenstrasse 40, 24118 Kiel, Germany. Email: stephen.sacht@ifw-kiel.de or sacht@economics.uni-kiel.de

I thank Reiner Franke, Tae-Seok Jang and Hans-Werner Wohltmann for valuable discussions and comments. This paper is based on Chapter 6 of my doctoral thesis *High-Frequency Analysis and Moment-Matching Estimation of the Baseline New-Keynesian Model* (2014). This working paper version is preliminary.

1 Introduction

It is well known that the estimation outcome within the Bayesian approach is sensitive to the choice of prior knowledge. In the majority of studies each parameter of interest is assumed to be independently distributed a priori. The associated sufficient statistics of these so-called *micropriors* are taken from microeconomic evidence. For example an overwhelming amount of empirical studies report a value of the quarterly discount factor β being equal to 0.99 mostly based on descriptive statistics and without consideration of other (deep) parameters. Lombardi and Nicoletti (2012, pp. 294; *LN* henceforth) claim that for some parameters this kind of information is ambiguous due to different estimation outcomes across several studies (e.g. for the Calvo parameter of price setting) while for others it is simply rare (e.g. for the first and second moments of shock processes). Furthermore, they mention that the property of micropriors to be independent put a restriction on the dynamics of the model variables a priori after a shock occurs. In order to avoid these shortcomings the evolution of macroeconomic variables based on econometric time series analysis, e.g. the moments of observable variables, is considered.

Taking this information into account will lead to the formulation of so-called *macropriors*, which can be then used as prior information within the Bayesian approach. *LN* extract macropriors from IRFs and state that this kind of priors help to clarify how the observables react to a specific shock a priori without explicit consideration of a data-based pre-sample. Similarly, Del Negro and Schorfheide (2008; *DNS* henceforth) formulate macropriors based on the second moments of the data. Furthermore *DNS* (p. 1193) emphasize that the problem of choosing micropriors is given by “the mechanical use of identical prior distribution for alternative model specifications”.

The question arises to what extent the choice of micro- or macropriors has an influence on the estimation outcome. In particular, how sensitive will the estimation results be to a change in the functional form of the micropriors? Is the moment-matching (MM; Franke et al. (2012)) approach able to identify sufficient macropriors, i.e. prior information which leads to higher values of the (logarithmic) marginal likelihood within Bayesian estimation as in the case of micropriors? In order to answer these questions, in this paper we apply a sensitivity analysis regarding the prior information. Then we compare the corresponding outcome (in terms of the goodness of fit) to the one when the standard prior information, which is in line with the recent literature and documented in Castelnovo (2010), is taken into account.

We account mainly for two things. First, we investigate the impact of the transformation of those model parameters which are bounded to the unit interval as discussed in the paper by *LN*. The idea is to allow for a more diffuse prior distribution of the parameters of interest through transformation (e.g. assuming a Gamma instead of a Beta distribution). Hence, we compare the estimation outcome under three alternative sets of micropriors, namely a baseline estimation, which is called 'BR', where the prior information is taken from Castelnovo (2010) and can be found in Table 4 given in the Appendix. Furthermore, two specifications, which differ with respect to the transformation of the price indexation parameter are considered. Second, we shed a light on the investigation of a combination of both MM and Bayesian (maximum likelihood) techniques. Since our parameter estimates obtained via MM account for the second moments of the underlying time-series they can be identified as macropriors for Bayesian estimation. In this respect we define a two-stage estimation procedure: we take the point estimates from the MM estimates for the GI and GM periods, respectively, and consider them as prior mean values of the parameters within Bayesian estimation. For a robustness check, we vary the priors with respect to the standard deviation of the persistence parameters α , χ , and ρ_π in order to allow for a more diffuse prior information.

The remainder of this paper is organized as follows. In the following section we present

the transformation rules for the parameters, which are either bounded below by zero or on the unit interval. In addition, we introduce the two-stage estimation procedure, where the prior mean values for Bayesian estimation are given by the estimates obtained from MM. Finally, we compare all estimation outcomes relative to the ones from Bayesian estimation, where standard prior information from the literature is considered. In section 6.3 we contrast our estimation outcomes for the endo- and exogenous persistence parameters (by applying the two-stage estimation procedure as described above) to the results taken from the literature. Section 6.4 concludes. The Appendix contains the Tables regarding the prior information for the GI and GM periods, respectively. Furthermore, the graphical representations of the posterior densities for all parameters to be estimated across all specifications are presented here.

2 The Baseline New-Keynesian model in Gap Notation

We consider a hybrid version of the New-Keynesian model in gap notation:

$$x_t = \frac{1}{1+\chi} E_t x_{t+1} + \frac{\chi}{1+\chi} x_{t-1} - \tau(\hat{r}_t - E_t \hat{\pi}_{t+1}) + \varepsilon_{x,t} \quad (1)$$

$$\hat{\pi}_t = \frac{\beta}{1+\alpha\beta} E_t \hat{\pi}_{t+1} + \frac{\alpha}{1+\alpha\beta} \hat{\pi}_{t-1} + \kappa x_t + v_{\pi,t} \quad (2)$$

$$\hat{r}_t = \phi_r \hat{r}_{t-1} + (1-\phi_r)(\phi_\pi \hat{\pi}_t + \phi_x x_t) + \varepsilon_{r,t} \quad (3)$$

$$v_{\pi,t} = \rho_\pi v_{\pi,t-1} + \varepsilon_{\pi,t}. \quad (4)$$

The time unit is to be thought of as one quarter. The three (mutually uncorrelated) shocks $\varepsilon_{z,t}$ are normally distributed around zero with variances σ_z^2 ($z = \pi, x, r$). All of the parameters are nonnegative. Specifically, β is the discount factor, κ a composite parameter that depends on the degree of price stickiness and assumptions on the production technology of firms, the coefficient α represents the degree of price indexation ($0 \leq \alpha \leq 1$), and the persistence in the supply shocks is given by the autocorrelation ρ_π ($0 \leq \rho_\pi < 1$). In the IS equation, χ is the representative household's degree of habit formation ($0 \leq \chi \leq 1$) and τ a composite parameter containing its intertemporal elasticity of substitution. In the Taylor rule, ϕ_r determines the degree of interest rate smoothing ($0 \leq \phi_r < 1$), and ϕ_x and ϕ_π are the policy coefficients that measure the central bank's reactions to contemporaneous output and inflation.

The empirical data on which the estimations of the system (1) to (4) are carried out derive from real GDP, the GDP price deflator, and the Federal Funds Rate. Their exogenous trend rates should be moderately flexible and are therefore specified by the convenient Hodrick-Prescott filter (as usual, although debatable, the smoothing parameter is $\lambda=1600$). The total sample period covers the time from 1960 to 2007.

Despite concentrating on trend deviations instead of levels, one has to be aware that there are still great changes over these years in the variances of the three variables and partly also in the entire pattern of their cross covariances. This makes it expedient to subdivide the period into two subsamples, which are commonly referred to as the periods of the Great Inflation (GI) and the Great Moderation (GM). We define the former by the interval 1960:1–1979:2 and the latter by 1982:4–2007:2; the time in between is excluded because of its idiosyncrasy (Bernanke and Mihov (1998)). To give an immediate example for the need of the subdivision, the standard deviation of the annualized inflation gap in GI is 1.41% versus 0.77% in GM; for the output gap it is 1.77% in GI versus 1.15% in GM.

3 The Baseline NKM in Gap Notation

In both experiments we seek to answer the question which kind of prior information is most reliable in order to ensure the best fit of the structural model to the data measured by the (logarithmic) marginal likelihood Υ . We also report the posterior probability $\Pi(BR)$, where the reference model is given by the Bayesian estimation of the model using the standard prior information.

4 Micro- versus Macropriors

4.1 Micropriors: The Impact of Transformation

We follow *LN* (p. 299) and transform the parameters which are either bounded below by zero or on the unit interval. This is especially of interest in such cases presented here, where we find the price indexation parameter $\alpha \approx 0$ based on standard, i.e. non-transformed priors (BR) and $\alpha = \chi = 1, \rho_\pi = \phi_x = \sigma_r = 0$ based on MM in the GI period (cf. estimation *A* provided in Franke et al. (2012)).

For the GM period we find $\alpha > 0, \tau \approx 0$ via BR and $\rho_\pi = 0, \kappa \approx 0$ and even $\chi = \infty$ via MM (cf. estimation *C* provided in Franke et al. (2012)). It goes without saying, that in both cases we choose the MM estimation outcome associated to the lowest value of J . By applying the transformation rules presented below within the Bayesian estimation, we are able to avoid narrow restrictions on those parameters of interest due to a change in their functional form. In this respect transformation means that e.g. a Gamma instead of a Beta distribution is assumed, where the former covers the range from a lower boundary point to unity, while the latter is bounded on the unit interval. In other words, we allow for differences in the micropriors in order to search for the posterior mean values over a broader region of the parameter space relative to the standard prior information we had considered so far. In the end we perform a model comparison exercise.

First, we consider the following relationship:

$$\tilde{k} = f(k) = \frac{1}{1-k} \text{ with } k = \{\chi, \rho_\pi, \phi_r\} \quad (5)$$

where $\tilde{k} = \{\tilde{\chi}, \tilde{\rho}_\pi, \tilde{\phi}_r\}$ represents the transformed counterparts of χ, ρ_π and ϕ_r . The corresponding functional form of the priors is given by $\tilde{k} \sim \Gamma(2, 0.99, 1, \infty)$.¹ Given the posterior mean of \tilde{k} , the value of the original parameters can be easily computed:

$$k = f^{-1}(\tilde{k}) = \frac{\tilde{k} - 1}{\tilde{k}}. \quad (6)$$

Note that, although all three original parameter values are bounded on the unit interval, their transformed counterparts are not. To see this consider both polar cases where the posterior mean values of \tilde{k} are either equal to one or an arbitrary high number (even close to infinity), which is possible under the assumption of the Gamma distribution. According to the transformation rule (6), k becomes then zero or one, respectively.

In contrast τ and κ are bounded below by zero only. The relationship between the original parameters $z = \{\tau, \kappa\}$ and their transformed counterparts $\tilde{z} = \{\tilde{\tau}, \tilde{\kappa}\}$ reads

$$\tilde{z} = f(z) = \frac{1}{z} \quad (7)$$

¹According to this expression the priors $\tilde{\chi}, \tilde{\rho}_\pi$ and $\tilde{\phi}_r$ follow a Gamma distribution with a mean value of 2, a standard deviation of 0.99 and are bounded between unity and infinity.

where $\tilde{z} \sim \Gamma(2, 0.99, 0, \infty)$ holds. Accordingly, the corresponding transformation rule is given by

$$z = f^{-1}(\tilde{z}) = \frac{1}{\tilde{z}}. \quad (8)$$

Suppose that the posterior mean values of \tilde{z} are zero, infinity, one or less than one. After transformation z becomes infinity, zero, one or greater than one, respectively, according to (8).

Regarding the parameter of price indexation α , we distinguish between two specifications based on the transformation rules (6) and (8):

$$\text{T1: } \alpha = \frac{\tilde{\alpha} - 1}{\tilde{\alpha}}, \quad \tilde{\alpha} \sim \Gamma(2, 0.99, 1, \infty) \quad (9)$$

$$\text{T2: } \alpha = \frac{1}{\tilde{\alpha}}, \quad \tilde{\alpha} \sim \Gamma(2, 0.99, 0, \infty) \quad (10)$$

Allowing for two different transformations of α is motivated by our findings in the Franke et al. (2012). In the first case T1 we stick to the economic rationale for the price indexation parameter bounded on the unit interval. In the second case T2 we account for our observation from Franke et al. (2012) where we report evidence for a stronger backward-looking behavior in the New-Keynesian Phillips Curve (NKPC), i.e. $\alpha > 1$. Finally, we are going to investigate, what kind of specification leads to a better fit of the model to the data relative to the other one. Therefore, specification T1 makes use of the transformed micropriors given by (5), (7) and (9) while for specification T2 we consider (5), (7) and (10). Accordingly, we report the posterior point estimates and their corresponding standard deviations, which are computed by applying the transformation rules (6) and (8). Note that we set all prior mean values equal to 2 which (regardless of the type of transformation) leads to a moderate value of the parameters equal to 0.5 a priori. Across both specifications the functional forms of $\phi_\pi, \phi_x, \sigma_\pi, \sigma_x$ and σ_r remain unchanged as motivated by *LN* (p. 300).

The estimation outcomes for T1 and T2 are reported in the third and forth column of the Tables 1 (GI) and 2 (GM). For each period we compare our results to the one in the second column where for the BR case the prior information given by Castelnovo (2010) is used. We find strong evidence for the superiority of BR for the GI period relative to T1 and T2. Even though it can be seen that the marginal likelihood Υ is higher in the BR case, in addition, the posterior probability $\Pi(BR)$ is also strictly in favor of the use of the standard prior information: a value of unity indicates that the probability that the null hypothesis is true — i.e. the reference model BR is superior in describing the data relative to the alternative models T1 and T2 — is given by hundred percent. Furthermore, the value of the loss function J of BR is (slightly) below the ones of T1 and T2. Moreover, the number of moments that missed the confidence intervals of the empirical moments (MCI) is significantly higher for T1 and T2. Note that here we calculate J according to equation (Franke et al. (2012)):

$$\hat{\theta} = \arg \min_{\theta \in \Theta} J(\theta; m_T^{emp}, W) := \arg \min_{\theta \in \Theta} [m(\theta) - m_T^{emp}]' W [m(\theta) - m_T^{emp}]. \quad (11)$$

n_θ denotes the parameters to be estimated, which are considered in a vector $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$, where Θ is the admissible parameter set. Since we are going to match second-moments, i.e. autocovariances, we make their dependence on the particular values of θ explicit by writing $m = m(\theta)$ for them. On the other hand, let T be the length of the sample period and $m_T^{emp} \in \mathbb{R}^{n_m}$ the vector of these empirical moments. The distance between the vectors of the model-generated and sample moments is measured by the quadratic loss function given by 11 that is characterized by an $(n_m \times n_m)$ weighting matrix W (which has to be estimated, too, cf. Franke et al. (2012)). Accordingly, the model is estimated by the set of

GI	Bayesian Prior Information			MM Prior Information			
	BR	T1	T2	V0	V1	V2	V3
α	0.074	0.300	0.137	0.790	0.783	0.610	0.580
	0.000-0.156	0.194-0.700	0.000-0.307	0.625-0.984	0.615-0.983	0.375-0.863	0.324-0.849
κ	0.209	0.202	0.222	0.044	0.044	0.037	0.038
	0.125-0.290	0.141-0.353	0.156-0.395	0.010-0.075	0.010-0.076	0.005-0.066	0.005-0.068
ρ_π	0.570	0.410	0.543	0.052	0.064	0.103	0.135
	0.452-0.693	0.012-0.555	0.230-0.665	0.000-0.111	0.000-0.137	0.000-0.225	0.000-0.286
σ_π	0.694	0.867	0.801	0.771	0.767	0.760	0.753
	0.524-0.864	0.659-1.073	0.592-1.013	0.674-0.871	0.670-0.868	0.653-0.863	0.646-0.866
χ	0.767	0.852	0.853	0.955	0.956	0.948	0.949
	0.689-0.850	0.751-0.894	0.743-0.894	0.908-0.999	0.909-0.999	0.895-1.000	0.896-0.999
τ	0.048	0.135	0.136	0.085	0.085	0.082	0.082
	0.030-0.067	0.105-0.185	0.107-0.189	0.053-0.118	0.052-0.118	0.050-0.115	0.049-0.114
σ_x	0.552	0.617	0.625	0.481	0.481	0.482	0.483
	0.465-0.637	0.504-0.721	0.511-0.735	0.420-0.541	0.418-0.538	0.420-0.538	0.421-0.540
ϕ_π	1.387	1.430	1.433	1.265	1.263	1.262	1.260
	1.124-1.644	1.142-1.707	1.143-1.704	1.147-1.385	1.145-1.384	1.142-1.381	1.141-1.384
ϕ_x	0.759	0.639	0.649	0.211	0.209	0.213	0.211
	0.314-1.193	0.301-0.993	0.285-0.993	0.039-0.371	0.037-0.366	0.043-0.373	0.034-0.365
ϕ_r	0.742	0.663	0.668	0.586	0.584	0.587	0.586
	0.668-0.816	0.546-0.730	0.556-0.735	0.511-0.665	0.507-0.661	0.511-0.665	0.511-0.668
σ_r	0.745	0.786	0.781	0.788	0.787	0.786	0.785
	0.643-0.841	0.675-0.900	0.672-0.892	0.678-0.899	0.676-0.894	0.675-0.898	0.675-0.895
Υ	-341.129	-349.755	-347.690	-337.356	-337.944	-337.205	-337.633
$\Pi(BR)$	0.500	1.000	1.000	0.023	0.039	0.020	0.030
J	213.512	237.934	222.343	99.280	100.291	131.282	134.565
MCI missed	23	32	26	3	4	9	11

Table 1: Bayesian estimation: sensitivity analysis regarding the prior information (GI).

Note: The results for T1 and T2 are computed by applying the formulas (6) and (8) given the transformed posterior point estimates and their corresponding standard deviations. Υ denotes the logarithmic marginal likelihood evaluated via the Laplace approximation. The posterior probability of the null hypothesis (i.e. the reference model BR is the true model) is given by $\Pi(BR) = \frac{\Xi_{BR,i}}{1+\Xi_{BR,i}}$ with $\Xi_{BR,i} = \exp(\Upsilon_{BR} - \Upsilon_i) \cdot \frac{\Lambda_{BR}}{\Lambda_i}$ and $i = \{T1, T2, V0, V1, V2, V3\}$. By construction $\Xi_{BR,BR} = \exp(\Upsilon_{BR} - \Upsilon_{BR}) \cdot \frac{\Lambda_{BR}}{\Lambda_{BR}} = 1$ holds, which leads to $\Pi(BR) = \frac{\Xi_{BR,BR}}{1+\Xi_{BR,BR}} = 0.5$.

parameters $\hat{\theta}$ that minimize this distance over the admissible set Θ . As we turn to a closer look of the point estimates it can be seen that most of them remain almost unchanged across specifications — remarkable differences can be observed regarding α (significantly

greater than zero in T1), τ (nearly three times larger in T1 and T2 compared to BR) and σ_π (larger in T1 and T2 compared to BR).

The results above are confirmed for the GM period. Based on the values for $\Pi(BR)$, the Bayesian estimation of the model using the standard prior information still outperforms the ones using transformed parameters, i.e. the null hypothesis that the reference model BR is the true model can not be rejected. Like for the GI period there are no remarkable differences in the point estimates across specifications with only a few exceptions (see above) — in particular for ρ_π (significantly lower in T1 and T2 compared to BR), while τ remains nearly unchanged. To sum up, in all cases the evidence is in favor of this kind of micropriors which are used in the reference model BR. In other words, the marginal likelihood Υ does not increase when a transformation (i.e. a change in the functional form) of the parameters as in T1 and T2 is applied. This holds independently for the assumption regarding the economic rationale for the price indexation parameter, i.e. if it is assumed to be true (T1: $0 \leq \alpha \leq 1$) or not (T2: $\alpha > 1$).²

4.2 Macropriors: An Investigation via MoMBay

In our second experiment we describe a two-stage estimation approach. At the first stage, the parameters of interest are estimated via MM. Then the resulting point estimates are adopted as prior mean values within the Bayesian estimation. Since these point estimates are the result of the matching of the second moments of the structural model to its empirical counterparts, they can be identified as macropriors. The prior mean values (obtained via MM) can be seen as appropriate starting values which might help to reduce the uncertainty in the researcher’s choice of the prior information. We call this procedure a *Moment-Matching based Bayesian* estimation approach or simply *MoMBay* for short.

In order to judge the reliability of MoMBay, we compare the outcome of this approach to the one when using micropriors reported in BR based on Castelnovo (2010) evaluated via the (logarithmic) marginal likelihood Υ and the posterior probability $\Pi(BR)$. Obviously, we have to determine the functional form of the priors first. The prior mean values are equal to the results of the parameter estimates from the MM estimation A for the GI and the GM period (cf. the corresponding Tables provided in Franke et al. (2012)), respectively. Regarding the prior distributions of the parameters we adopt the one from BR. Due to the fact that as an outcome of the A estimations some parameter values are close to or even hit their boundary points (i.e. $\alpha, \kappa, \rho_\pi, \chi, \phi_x$ and σ_r in the GI case as well as κ, ρ_π, χ and τ in the GM case), within Bayesian estimation the prior mean values are set to 0.1 and 0.9, respectively. For the sake of uniformity all corresponding prior values of the standard deviations are set to 0.08.³ We call this specification V0. Furthermore we consider three additional specifications where we account for more diffuse prior information regarding the extrinsic (ρ_π) and intrinsic (α and χ) persistence parameters. Within specification V1 we increase the prior mean and the standard deviation of ρ_π to 0.2 and 0.15 respectively. Within specification V2 $\rho_\pi \sim \beta(0.2, 0.15)$ together with the intrinsic persistence parameters $\alpha, \chi \sim \beta(0.8, 0.15)$ hold. Finally, in specification V3, $\rho_\pi \sim \beta(0.3, 0.2)$ together with $\alpha, \chi \sim \beta(0.8, 0.15)$ is applied. The functional forms of all other parameters remain the same as specified in V0. An overview of all prior densities are given in the Tables 5 (GI) and 6 (GM) in the Appendix.

²It is indeed the case as we consider T2 as the reference model, the posterior probability ($\Upsilon(T2)$ versus $\Upsilon(T1)$) is given by $\Pi(T2) = 0.887$ for the GI and even $\Pi(T2) = 0.974$ for the GM period. Hence, allowing for a (non-plausible) value of α greater than 1 leads to a better goodness of fit. However, these observations are pure hypothetical since — as we discussed above — both estimations (T1 and T2) are inferior relative to estimation BR.

³This procedure ensures that the inverse Hessian matrix of the second derivatives (evaluated at the posterior mode) will be positive-definite which is a requirement for applying the Random-Walk Metropolis-Hastings algorithm.

GM	Bayesian Prior Information			MM Prior Information			
	BR	T1	T2	V0	V1	V2	V3
α	0.033	0.224	0.057	0.502	0.501	0.285	0.279
	0.000-0.071	0.160-0.366	0.000-0.120	0.364-0.639	0.362-0.641	0.144-0.424	0.137-0.419
κ	0.163	0.155	0.178	0.026	0.026	0.030	0.030
	0.103-0.221	0.119-0.224	0.134-0.267	0.004-0.046	0.004-0.046	0.007-0.052	0.006-0.053
ρ_π	0.389	0.103	0.218	0.030	0.035	0.059	0.066
	0.274-0.510	0.000-0.201	0.000-0.340	0.000-0.064	0.000-0.078	0.001-0.123	0.000-0.143
σ_π	0.517	0.635	0.644	0.528	0.526	0.567	0.563
	0.420-0.611	0.535-0.730	0.512-0.768	0.460-0.593	0.460-0.594	0.486-0.646	0.482-0.644
χ	0.825	0.880	0.881	0.961	0.960	0.943	0.945
	0.759-0.891	0.804-0.913	0.808-0.913	0.917-1.000	0.917-1.000	0.874-0.999	0.879-1.000
τ	0.017	0.081	0.080	0.015	0.015	0.012	0.012
	0.009-0.025	0.066-0.103	0.065-0.100	0.004-0.026	0.003-0.026	0.000-0.021	0.000-0.021
σ_x	0.346	0.437	0.435	0.347	0.347	0.347	0.347
	0.296-0.399	0.366-0.504	0.363-0.504	0.307-0.386	0.309-0.386	0.308-0.386	0.307-0.386
ϕ_π	1.181	1.258	1.243	1.595	1.596	1.596	1.594
	1.001-1.383	0.965-1.510	0.962-1.505	1.465-1.727	1.467-1.729	1.465-1.727	1.459-1.725
ϕ_x	1.014	0.944	0.917	1.038	1.040	1.035	1.037
	0.602-1.419	0.587-1.230	0.558-1.254	0.909-1.166	0.910-1.167	0.912-1.168	0.910-1.166
ϕ_r	0.814	0.731	0.728	0.861	0.861	0.861	0.861
	0.762-0.867	0.659-0.778	0.656-0.774	0.821-0.900	0.822-0.900	0.822-0.898	0.825-0.902
σ_r	0.449	0.537	0.538	0.491	0.491	0.489	0.491
	0.395-0.502	0.466-0.604	0.467-0.601	0.436-0.544	0.438-0.546	0.437-0.543	0.438-0.554
Υ	-299.756	-327.730	-324.108	-287.655	-288.385	-283.164	-283.719
$\Pi(BR)$	0.500	1.000	1.000	0.000	0.000	0.000	0.000
J	170.098	246.827	243.162	154.373	154.873	152.265	153.968
MCI missed	15	25	27	14	13	16	16

Table 2: Bayesian estimation: sensitivity analysis regarding the prior information (GM).

Note: The results for T1 and T2 are computed by applying the formulas (6) and (8) given the transformed posterior point estimates and their corresponding standard deviations. Υ denotes the logarithmic marginal likelihood evaluated via the Laplace approximation. The posterior probability of the null hypothesis (i.e. the reference model BR is the true model) is given by $\Pi(BR) = \frac{\Xi_{BR,i}}{1+\Xi_{BR,i}}$ with $\Xi_{BR,i} = \exp(\Upsilon_{BR} - \Upsilon_i) \cdot \frac{\Lambda_{BR}}{\Lambda_i}$ and $i = \{T1, T2, V0, V1, V2, V3\}$. By construction $\Xi_{BR,BR} = \exp(\Upsilon_{BR} - \Upsilon_{BR}) \cdot \frac{\Lambda_{BR}}{\Lambda_{BR}} = 1$ holds, which leads to $\Pi(BR) = \frac{\Xi_{BR,BR}}{1+\Xi_{BR,BR}} = 0.5$.

This procedure is motivated by the observation that given the BR estimation results, the dynamics in inflation is explained heavily by extrinsic rather than intrinsic persistence, i.e. $\alpha \approx 0$ and $\rho_\pi \gg 0$ hold (while χ is high). In this respect we try to investigate if this

result still holds, when MoMBay is applied or if there is evidence just for the opposite, as reported in the MM case given the marginal likelihood. In other words, we seek to answer the question if, as an outcome of MoMBay, the posterior means of the three persistence parameters tend to depart from their prior mean values (taken from MM; see above) to the left (α and χ) and right (ρ_π) of the posterior distribution. In such a case the results from Bayesian estimation are confirmed under consideration of macro- rather than micropriors - if, of course, the value of the posterior probability will be in favor of MoMBay.

The results for V0 till V3 are reported in the fifth to eighth columns of Table 1 (GI) and 2 (GM). It turns out that the posterior probability $\Pi(BR)$ is in favor for all specifications where macropriors are considered. Independently of the underlying sample period the following observation holds: Given the values of $\Pi(BR)$ the probability that BR is the true model varies between 2 to 3.9 percent in the GI case and be even 0 percent in the GM case. This leads to the conclusion that the specifications V0 to V3 are able to describe the data well. Moreover, applying macro- rather than micropriors within Bayesian estimation results in a slight (GI) and distinct (GM) improvement in the goodness of fit in terms of Υ .⁴

Four observations are also worth mentioning. First, although the values of the marginal likelihoods Υ do not differ strongly across V0 to V3 and especially in contrast to BR (when looking at GI; see the corresponding entries in Table 1), the effect on $\Pi(BR)$ is indeed remarkable. Hence, even a 'small' difference in the marginal likelihood across models (e.g. for GI: $|\Upsilon(BR)| - |\Upsilon(V2)| = |3.924|$) leads to strong changes in the reliability of the null hypothesis measured by $\Pi(BR)$. Second, the goodness of fit of V0 to V3 is also confirmed when looking at the corresponding indices known from the MM approach: higher values of Υ correspond with lower values for J and MCI . This means that the matching of the simulated to the empirical moments is — at least in the GI case — very good when macropriors are considered (given J together with MCI). Third, some of the point estimates of the parameters evaluated via MoMBay differ significantly from the ones of the BR model (see below). However, the point estimates remain unchanged the more we choose relative diffuse priors for the persistence parameters, i.e. as we go from V0 to V3. Exceptions are given by the values for α and ρ_π : The price indexation parameter is estimated to be roughly 30 (78) percent *bigger* in V0 and V1 compared to V2 and V3 in the Great Inflation (Moderation) period. Regarding the shock persistence this previous observation is mirror inverted, since the parameter is estimated to be roughly 51 to 53 percent *lower* in V0 and V1 compared to V2 and V3 in the GI and GM period, respectively. Finally, these observations support the analysis undertaken by *LN* who claim that prior knowledge is sensitive to the use of macropriors, i.e. the moments of observable variables. In the following we will discuss the results for the GI and GM period separately in greater detail.

According to Table 1 specification V0 provides a higher value for Υ (-337.356) relative to BR (-341.129) in the GI case. Although the marginal likelihood for V2 (-337.205) is slightly higher than for V0, we claim that V0 is the most appropriate candidate within MoMBay since the values of J and MCI are the lowest. As we turn to a comparison of the point estimates between BR and V0, differences can be identified. The results for α (0.790) and ρ_π (0.052) reflect our observations from Franke et al. (2012). In this respect using point estimates evaluated via MM as prior information within Bayesian estimation, this leads to distinct evidence for intrinsic rather than extrinsic persistence in the baseline NKM. Hence, concerning the source of persistence, the outcome of the plain MM approach (estimation *A* given in Franke et al. (2012)) is confirmed through MoMBay. More precisely,

⁴Of course, the expressions 'slight' and 'distinct' seem to be a little bit broad in this context. The computation of the posterior probability given a V -specification, i.e. V0 to V3, as being the reference model could help to get things straight. For a clear arrangement we do not apply such kind of analysis, which can be motivated by our observations discussed in section 5.

the Bayesian approach largely confirms the MM prior information, which is approved by the values of Υ and $\Pi(BR)$. Furthermore, the monetary policy parameter ϕ_x as well as the standard deviation of the interest rate shock σ_r are estimated to be significantly greater than zero (as reported in Franke et al. (2012) for GI based on the corresponding estimation A). While the difference in V0 compared to the reference MM estimation A for ϕ_x is moderate (0.211), the value of σ_r is quite high (0.788). Not surprisingly, these results depend heavily on the fact that the macropriors are non-diffuse, i.e. small standard deviations are applied for all parameters within specification V0. Nevertheless, recall that all measurements for the goodness of fit presented here prove that the reliability of V0 over BR in describing the data is unambiguous.

The outcome for the GM case is depicted in Table 2. Here we claim that specification V2 is superior over BR in terms of the log marginal likelihood (-283.164 vs. -299.756) and the values for J (152.265 vs. 170.098). However, it must be emphasized that the matching seems not to be a good one since the values for J are quite high compared to the reference estimation A ($J = 54.1$; cf. Franke et al. (2012)). This can also be seen by looking at the values for MCI : the number of moments missed the confidence interval of the empirical moments is higher for V2 (16) than for BR (15). The lowest values is actually given for the specification V1 (13). A reason for this can probably be found in the steep initial decline in the auto-covariance profile in the inflation gap as discussed in Franke et al. (2012). Concerning the discussion of the point estimates one observation is worth mentioning. As we compare the estimates of α in V2 (0.285) to the reference estimation A (0.816) and BR (0.033), we come to the conclusion that, first, allowing for more diffuse macropriors (as in V2), α tends towards a lower value than 0.5 and second, there is a low degree of (intrinsic, extrinsic *and* inherited) persistence in the inflation gap, even though the standard deviation in the shock σ_π is quite large.⁵ While the other parameter estimates remain nearly the same for V2 in contrast to the reference estimation A (and BR), the low estimated value of α sheds a new light on the issue of backward-looking behavior in the NKPC. In particular, persistence plays a *minor* role in describing the dynamics in the inflation gap as indicated by evidence from Bayesian (MoMBay) as well as MM. This observation is even more surprising as $\Pi(BR)$ is strictly in favor of specification V2. Nevertheless, it is an open question if this result still holds if e.g. the level of inflation instead of the gap is considered. We leave this to further research.

5 Comparison with other Empirical Studies

Next, we seek to compare our results to the most prominent studies from the literature. In *DNS* three different groups of model parameters are considered, which describe the steady-state relationships, exogenous processes and (roughly speaking) the degree of persistence. Prior information on these three groups is either given by micro- or macropriors. In the former case standard priors from the literature are used. In contrast, macropriors (primarily for the parameter of the exogenous processes) are constructed based on a quasi-likelihood function where the second-moments of an auxiliary vector autoregression (VAR) model are taken into account.⁶ A similar procedure is also adopted in *LN*. In addition, the

⁵Note that the prior mean value of α is set from 0.816 to 0.8 across the alternative specifications (see the first row in Table 6 given in the Appendix). As it can be seen by the graphical comparison of the corresponding prior to the posterior distribution in Figure 13 which is provided in the Appendix, the posterior mean value of α indeed tends to the left when applying the Bayesian approach.

⁶A quasi-likelihood function shows similar properties to the standard log likelihood function except for the specification of the parameters' distribution. Instead a quasi-likelihood function consists on a relationship between the variance and the mean of the observations only (see Wedderburn (1974) for a detailed description of this approach). A quasi-likelihood function is most often used in connection to an auxiliary model for the purpose of estimation. The corresponding procedure is known as *Indirect Inference* (II). While

authors introduce an estimation approach where the macropriors are constructed through the minimization of IRFs — which is indeed similar to the matching of second order moments presented here. Following *DNS* (p. 1196), the authors then divide the vector of structural parameters into two blocks of parameters, which represent the steady state relationships among observable variables and the law of motion of the exogenous processes, respectively. These two blocks can be either related or non-related to each other. In both studies the regular algorithm within Bayesian estimation is modified in order to account for the macropriors as discussed above.

We emphasize that MoMBay is located in between both methods presented in *DNS* and *LN*. Obviously, like in *DNS* the second-moments of observables are considered while as in *LN* IRFs are matched. In comparison to *DNS*, we apply our estimation procedure directly (i.e. without use of an auxiliary model) under consideration of the MM approach, which is close to the GMM estimation approach (instead of a specific likelihood function) like in *LN*. Both studies report evidence for differences in the posterior estimates with respect to the underlying prior information. Although, these studies are closely related to our experimental setting presented here, the interpretation of the results must be done with some caution. First, *DNS* investigate a small-scale NKM with sticky prices *and* wages. Furthermore, in their paper they report the outcome of three so-called Low Rigidities, High Rigidities and Agnostic scenarios with respect to micro- and macropriors. Across these scenarios the prior information for the Calvo parameter of price and wage stickiness becomes more diffuse. Note that in our study we consider the baseline NKM with sticky prices only while the slope of the NKPC is characterised by the composite parameter κ . Hence, we do not control *directly* for the Calvo parameter of price stickiness within our estimation. *LN* also work with the baseline NKM but assume a purely forward-looking dynamic IS curve while there exists only a non-autocorrelated (one-off) shock in the hybrid NKPC. However, besides price indexation, persistence is brought into inflation by considering an autocorrelated preference and technology shock. Needless to say, that in both studies the levels (given as demeaned values) rather than the (time-dependent) gaps of the variables are taken into account. Table 3 summarizes the differences of the discussed estimation procedures.

Although there exist structural differences in the setting of our experiments compared to the ones in *DNS* and *LN*, the results help to motivate a fruitful discussion on micro- vs. macropriors. For the sake of clarity (and in order to ensure the comparability with our results), we focus on the outcome for the price indexation parameter (here: α , *LN*: ω , *DNS*: ι_p) and the persistence parameter in the cost-push/mark-up shock (here: ρ_π , *LN*: none, *DNS*: ρ_λ) only. The estimated values for *DNS* are given in Table 6 of their paper, where they discuss the results using micropriors (denoted as \wp_s) and macropriors (denoted as \wp_{QL}) expressed by (Low Rigidities/Agnostic/High Rigidities). In *LN* their Table B2 (micropriors) and Table B3 together with Table B4 (macropriors) given in their Appendix B are primarily of interest.

DNS report significant differences in the estimation results (in terms of posterior mean values and the goodness of fit) with respect to both types of priors. While across all scenarios in the case of micropriors the values of ι_p are low (0.19/0.19/0.20), using macropriors the observation becomes ambiguous (0.47/0.14/0.08). Note here, that the confidence bands are quite wide and include zero except for the Low Rigidities scenario. Hence, at least for the latter it can be stated that there exists evidence for (a moderate degree of) price indexation if macropriors are considered. One explanation for the low value of ι_p in the case of micropriors is given by the associated high values of ρ_λ where (0.86/0.56/0.41)

this approach allows endogenizing the choice of moments (rather than chosen arbitrarily by the researcher within MM), it becomes an open question which kind of auxiliary model will be a good approximation to a DSGE model (*LN* (p. 298, fn. 9)). An overview of the II approach is provided by Carrasco and Florens (2002).

<i>Characteristics</i>	<i>DNS</i>	<i>LN</i>	<i>MoMBay</i>
Moments	(Auto)Covariances	IRFs	(Auto)Covariances
Matching Procedure	Auxiliary VAR, Likelihood Maximization	Minimum Distance	Minimum Distance
Structural Model	Hybrid NKM (sticky wages and prices)	Partly forward-looking NKM (sticky prices only)	Hybrid NKM (sticky prices only)
Sample Period(s)	1982:4–2005:4	1960:1–2001:4	1960:1–1979:2 1982:4–2007:2
Data	Level (demeaned)	Level (demeaned)	Gap (time-dependent)
Price Indexation	ι_p	ω	α
↔ Micro:	0.19 / 0.19 / 0.20	0.77	0.074 / 0.033
↔ Macro:	0.47 / 0.14 / 0.08	0.76	0.790 / 0.285
Persistence Parameter	ρ_λ	—	ρ_π
↔ Micro:	0.86 / 0.56 / 0.41	—	0.570 / 0.389
↔ Macro:	0.89 / 0.88 / 0.68	—	0.052 / 0.059

Table 3: Construction of macropriors.

Note: The results using micro- and macropriors are given in the corresponding rows ‘Micro’ and ‘Macro’, respectively. The estimated values for DNS are given in Table 6 of their paper where they discuss the results using micropriors (denoted as \wp_s) and macropriors (denoted as \wp_{QL}) expressed by so-called (Low Rigidities/Agnostic/High Rigidities) scenarios. For LN the Table B2 (micropriors) and Table B3 together with Table B4 (macropriors) are considered. With respect to the results from MoMBay, the first entry in the row ‘Micro’ denotes the value for GI and the second for GM based on the Bayesian estimation *BR*. The first entry in the row ‘Macro’ denotes the value for GI and the second for GM based on the estimations *V0* and *V2*, respectively.

holds - a result which is confirmed by our study (cf. both BR estimates for the GI and GM period). Interestingly, the corresponding vector in the macroprior case is given by (0.89/0.88/0.68) where the first entry indicates that a high value for ρ_λ coincides with a high value of ι_p . Hence, there exists both intrinsic and extrinsic persistence in inflation where the latter one is more distinctive.

An explanation for this phenomenon is not easy to reveal because in their paper DNS focus on a discussion of price and wage rigidity only. Nevertheless, the authors show that the lowest values in both types of rigidities as well as a quite low level of habit formation (which is denoted by h in their paper) can be found in the Low Rigidities scenario. Since a low degree of price rigidity leads to a high value of the slope of the NKPC, *inherited* persistence plays a major role in explaining the dynamics in inflation.⁷ It follows from that,

⁷Recall from DNS (p. 1195, equation (9)) that under price indexation, the slope of the NKPC is given by $\frac{(1-\zeta_p\beta)(1-\zeta_p)}{\zeta_p(1+\iota_p\beta)}$ where – for a given value of the discount parameter β and price indexation ι_p – this expression increases as price rigidity ζ_p decreases. See also our discussion on the frequency-dependent inherited persistence (FIP) effect in Sacht (2014a,b).

although inherited persistence is important, the persistence in the output gap is not high enough in order to explain the volatility in inflation. In other words, if there is empirical evidence for persistence in inflation while the habit formation parameter is estimated to be low, intrinsic as well as extrinsic inflation persistence must be significantly higher than zero – in order to account for inertia in the corresponding times series. Hence, both values of ι_p and ρ_λ have to be unequal to zero if h is low. Again, these observations must be interpreted with some caution: DNS (p. 1207) notice that probably the time series they use be not informative enough to analyze different scenarios since some posterior estimates mimic their prior values. Our study does not face this kind of problem — with only two exceptions. In particular, for the policy parameters ϕ_π and ϕ_x the corresponding prior distribution does not differ from its posterior counterpart. This holds only in the GM case and across the specifications V0 to V3 (cf. Figures 11 to 14 given in the Appendix). Here it seems that the posterior distribution is driven by the data. Since the policy parameters are not primarily of interest in our study we judged this observation not to be severe. However, further investigations on the influence of non-informative times series within Bayesian estimation are always a necessary task to be undertaken in general. In particular, micropriors might be imposed on ϕ_π and ϕ_x in order to check if the problem remains and if the goodness of fit changes.

LN show that no differences in the estimations for the parameter (ω) of interest can be found. Both in the micro- and macroprior (first and second block) case, price indexation is around a high value of 0.77. Although the cost push shock is characterized by an iid-shock, we already stated above that persistence might come into play through the autoregression parameter in the preference and technology shocks, respectively, which have a (in)direct influence on the inflation dynamics (cf. the Appendix A in LN). The corresponding values are estimated around 0.83 to 0.85 (ρ_a) and 0.89 (ρ_g) across all specifications regarding the prior information. The only changes in the parameter estimates can be found in the values for price rigidity, the elasticity of labor and in the standard deviation of the technology and mark-up shock. The authors argue that this problem might be explained by the misspecification of the model described by a high cross-correlation of the shocks. Under consideration of artificial data they show that this indeed holds with exception of preference shocks. According to our Tables 1 and 2, we see that in our experiment there exist no differences in the standard deviations of the shock at all. However, we do not discard the possibility of misspecification completely since the phenomena might help to explain the changes in the composite parameter κ where the parameter of price rigidity is incorporated. In this respect we call for an explicit consideration of the slope of the NKPC in greater detail — although the number of parameters to be estimated will increase in that case, simply by adding the parameters for various elasticities (cf. the expression(s) for the NKPC in Part One of this thesis).

We sum up that the results presented in DNS and LN mimic those from the literature, where the degree of price indexation is high in the absence of persistence in the cost push (mark-up) shock. If both sources of persistence are considered, the autocorrelation parameter is estimated to be dominant, i.e. being greater than the parameter of price indexation - where the latter can also be insignificant. Moreover, both studies indicate that there exist changes in the parameter estimates if macro- instead of micropriors are applied as prior information. Our experiments (partly) confirm these observations, i.e. we provide evidence for the domination of the persistence parameter if micropriors are considered. The opposite holds under consideration of macropriors, which leads to a better goodness of fit within Bayesian estimation compared to the BR case. Therefore, we claim that the MoMBay approach can be identified as a point of departure for further research.

6 Conclusion

The estimation outcome within the Bayesian approach is sensitive to the choice of prior knowledge. In the majority of studies each parameter of interest is assumed to be independently distributed a priori. The associated sufficient statistics of these so-called micropriors are taken from microeconomic evidence. In contrast, prior information based on econometric time series analysis leads to the formulation of so-called macropriors. We show that while micropriors are often used in the literature, applying MM-based macropriors leads to a better fit of the structural model to the data. Therefore, the advantages of Bayesian (sufficiency) and MM estimation (transparency) can be *unified* within a two-stage estimation procedure called MoMBay, where the macroprior information for Bayesian estimation is evaluated via MM.

The MoMBay approach can be seen as an alternative approach for estimation and evaluation of macroeconomic models in conjuncture with the Bayesian approach. The reason for this is that MoMBay reduces the parameter space from which appropriate starting values (or better: prior mean values) for the optimization within the Bayesian approach are taken. We show that this indeed leads to a change in the estimates of some parameters. Similar observations are also made by Del Negro and Schorfheide (2008) as well as Lombardi and Nicoletti (2012) who also found differences in the posterior values under consideration of macropriors. Problems are stemming from the existence of less or even non-informative time series and misspecification of the model. A deeper investigation on these topics might be fruitful in further research.

7 Appendix

7.1 Prior Densities of the BR Estimations

α	κ	ρ_π	σ_π
$\beta(0.3, 0.2)$	$\Gamma(0.4, 0.1)$	$\beta(0.6, 0.1)$	$I\Gamma(1.0, 8.0)$
χ	τ	—	σ_x
$\beta(0.5, 0.1)$	$\Gamma(0.037, 0.0125)$	—	$I\Gamma(0.25, 2.0)$
ϕ_π	ϕ_x	ϕ_r	σ_r
$N(1.3, 0.2)$	$\Gamma(1.2, 0.8)$	$\beta(0.5, 0.28)$	$I\Gamma(1.0, 8.0)$

Table 4: Prior densities of the BR estimations.

Note: β , N , Γ and $I\Gamma$ denotes a Beta, Normal (Gaussian), Gamma and Inverse Gamma distribution, respectively.

7.2 Prior Densities of the MoMBay Estimations

GI				
	V0	V1	V2	V3
α	$\beta(0.900, 0.080)$	$\beta(0.900, 0.080)$	$\beta(\mathbf{0.800}, \mathbf{0.150})$	$\beta(\mathbf{0.800}, \mathbf{0.150})$
κ	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$
ρ_π	$\beta(0.100, 0.080)$	$\beta(\mathbf{0.200}, \mathbf{0.150})$	$\Gamma(\mathbf{0.200}, \mathbf{0.150})$	$\Gamma(\mathbf{0.300}, \mathbf{0.200})$
σ_π	$I\Gamma(0.614, 0.080)$	$I\Gamma(0.614, 0.080)$	$I\Gamma(0.614, 0.080)$	$I\Gamma(0.614, 0.080)$
χ	$\beta(0.900, 0.080)$	$\beta(0.900, 0.080)$	$\beta(\mathbf{0.800}, \mathbf{0.150})$	$\beta(\mathbf{0.800}, \mathbf{0.150})$
τ	$\Gamma(0.105, 0.080)$	$\Gamma(0.105, 0.080)$	$\Gamma(0.105, 0.080)$	$\Gamma(0.105, 0.080)$
σ_x	$I\Gamma(0.519, 0.080)$	$I\Gamma(0.519, 0.080)$	$I\Gamma(0.519, 0.080)$	$I\Gamma(0.519, 0.080)$
ϕ_π	$N(1.324, 0.08)$	$N(1.324, 0.08)$	$N(1.324, 0.08)$	$N(1.324, 0.08)$
ϕ_x	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$
ϕ_r	$\beta(0.314, 0.080)$	$\beta(0.314, 0.080)$	$\beta(0.314, 0.080)$	$\beta(0.314, 0.080)$
σ_r	$I\Gamma(0.100, 0.080)$	$I\Gamma(0.100, 0.080)$	$I\Gamma(0.100, 0.080)$	$I\Gamma(0.100, 0.080)$

Table 5: Prior densities of the MoMBay estimations (GI).

Note: For specification V0 we adopt the point estimates of the MM estimation A (cf. the corresponding Tables in Franke et al. (2012)) as prior means within Bayesian estimation. If MM parameter estimates hit the boundary points, the corresponding prior means are set to 0.1 or 0.9, respectively. The same prior distributions as in the BR case are considered (cf. Table 4). For the sake of uniformity all corresponding prior values of the standard deviations are set to 0.08. All changes across the specifications V1 to V3 relative to V0 are given in bold type.

GM				
	V0	V1	V2	V3
α	$\beta(0.816, 0.080)$	$\beta(0.816, 0.080)$	$\beta(\mathbf{0.800}, \mathbf{0.150})$	$\beta(\mathbf{0.800}, \mathbf{0.150})$
κ	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$	$\Gamma(0.100, 0.080)$
ρ_π	$\beta(0.100, 0.080)$	$\beta(\mathbf{0.200}, \mathbf{0.150})$	$\beta(\mathbf{0.200}, \mathbf{0.150})$	$\beta(\mathbf{0.300}, \mathbf{0.300})$
σ_π	$I\Gamma(0.200, 0.080)$	$I\Gamma(0.200, 0.080)$	$I\Gamma(0.200, 0.080)$	$I\Gamma(0.200, 0.080)$
χ	$\beta(0.900, 0.080)$	$\beta(0.900, 0.080)$	$\beta(\mathbf{0.800}, \mathbf{0.150})$	$\beta(0.800, 0.150)$
τ	$\Gamma(0.047, 0.080)$	$\Gamma(0.047, 0.080)$	$\Gamma(0.047, 0.080)$	$\Gamma(0.047, 0.080)$
σ_x	$I\Gamma(0.532, 0.080)$	$I\Gamma(0.532, 0.080)$	$I\Gamma(0.532, 0.080)$	$I\Gamma(0.532, 0.080)$
ϕ_π	$N(1.626, 0.08)$	$N(1.626, 0.08)$	$N(1.626, 0.08)$	$N(1.626, 0.08)$
ϕ_x	$\Gamma(1.031, 0.080)$	$\Gamma(1.031, 0.080)$	$\Gamma(1.031, 0.080)$	$\Gamma(1.031, 0.080)$
ϕ_r	$\beta(0.776, 0.080)$	$\beta(0.776, 0.080)$	$\beta(0.776, 0.080)$	$\beta(0.776, 0.080)$
σ_r	$I\Gamma(0.472, 0.080)$	$I\Gamma(0.472, 0.080)$	$I\Gamma(0.472, 0.080)$	$I\Gamma(0.472, 0.080)$

Table 6: Prior densities of the MoMBay estimations (GM).

Note: See Table 5.

7.3 Graphical Results – Prior versus Posterior Densities

Figures 1 to 14 show the graphical results from optimization for the GI and GM period discussed in the sections 4.1 and 4.2. In these standard graphical outputs provided by DYNARE, the dashed/dotted green line represents the posterior mode (calculated from the posterior kernel) *associated* with the posterior distribution (black line; as a result of the Random-Walk Metropolis-Hastings procedure). The red/grey line stands for the prior distribution (collected in the vector $p(\theta)$). It can be stated that all posterior distributions are close to being normal and that the deviations of the posterior means from the modes are negligible low i.e. the dashed/dotted green line is close to the maximum of the posterior distribution. Hence the goodness of the optimization is confirmed: convergence checks by Brooks and Gelman (1998) which are summarized in the uni- and multivariate diagnostics (which are, again, provided upon request) are considered. In the majority of the cases the posterior distributions are not driven by the prior distributions i.e. the shape of the distributions differ and they are not close to each other. Hence, the information from the data is sufficient to ensure that the estimation of the posterior distribution is independent from the priors. The only few exceptions are given by the monetary policy parameters ϕ_π and ϕ_x in the GM case and across the specification V0 (Figure 11) till V3 (14). Here the sample is not informative enough in order to identify a discrepancy between prior and posterior distribution.

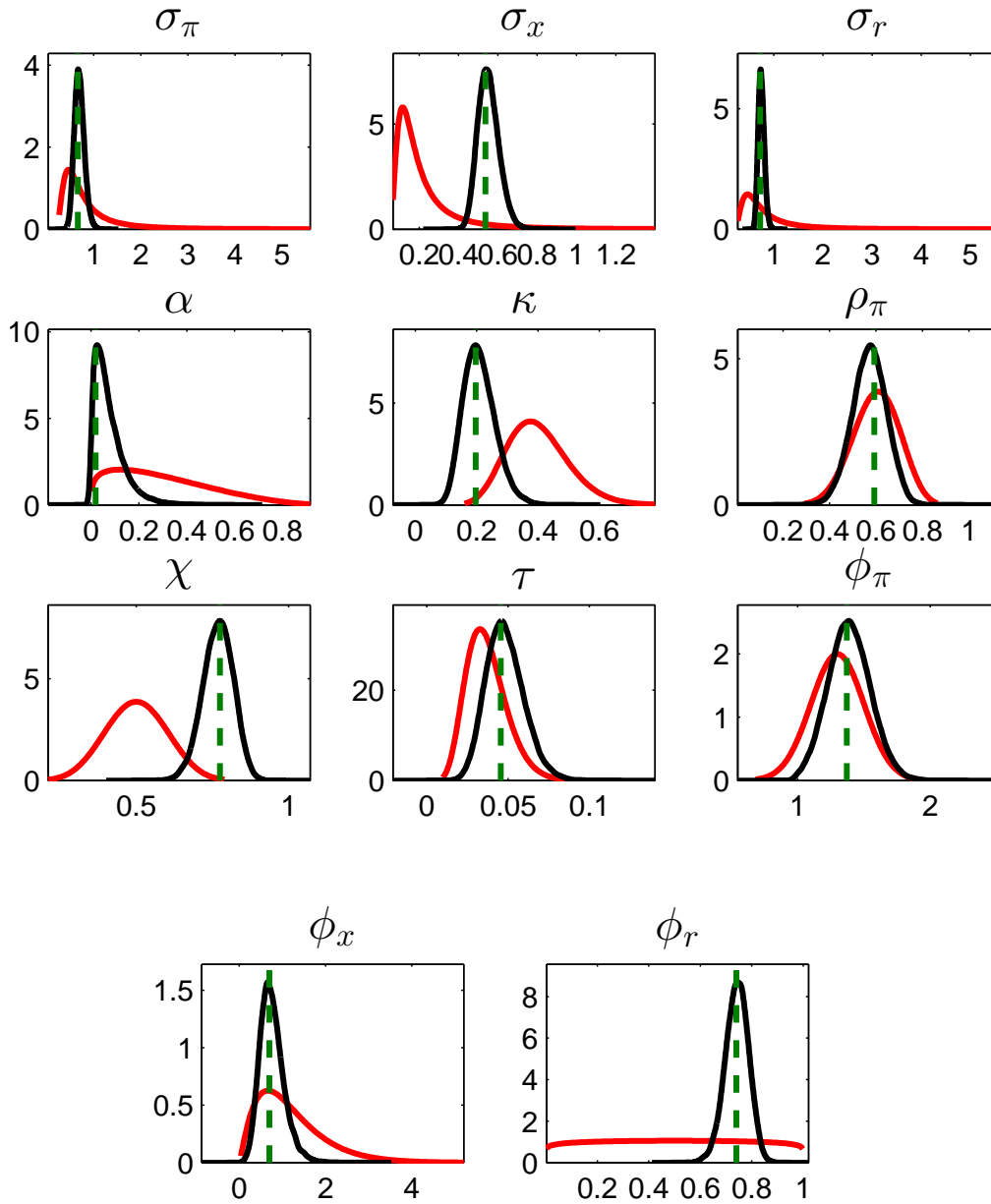


Figure 1: Posterior density estimates for BR (GI).

Note: The dashed/dotted green line represents the posterior mode (calculated from the posterior kernel) *associated* with the posterior distribution (black line; generated by the Random-Walk Metropolis-Hastings procedure). The red/grey line represents the prior distribution (see text for more information regarding the vector of prior information $p(\theta)$). This note equally applies to all of the following Figures.

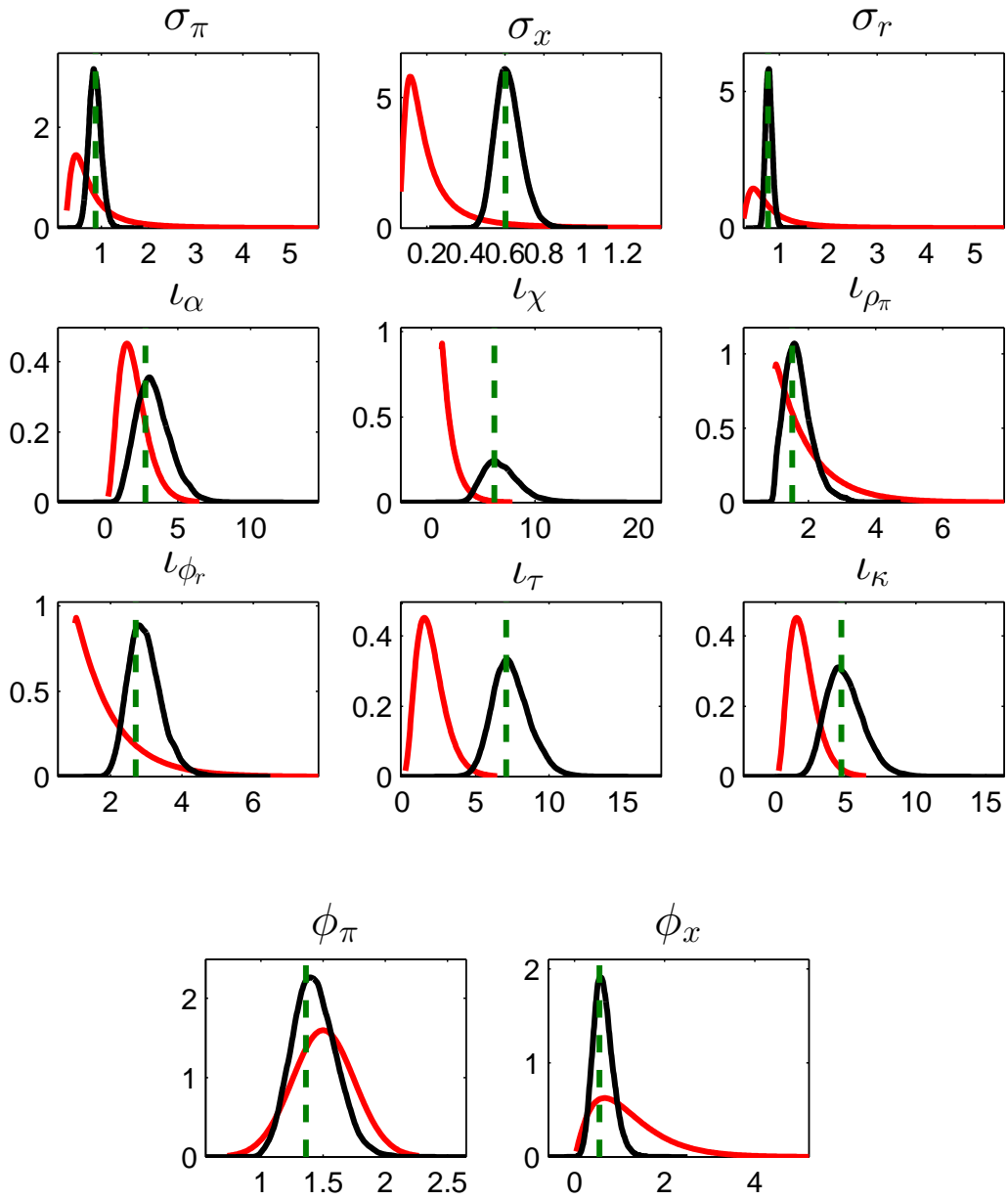


Figure 2: Posterior density estimates for T1 (GI).

Note: See Figure 1. The parameter l_{ii} with $ii \in \{\alpha, \chi, \rho_\pi, \phi_r, \tau, \kappa\}$ denotes the transformed parameters according to the corresponding rules given in section 4.1.

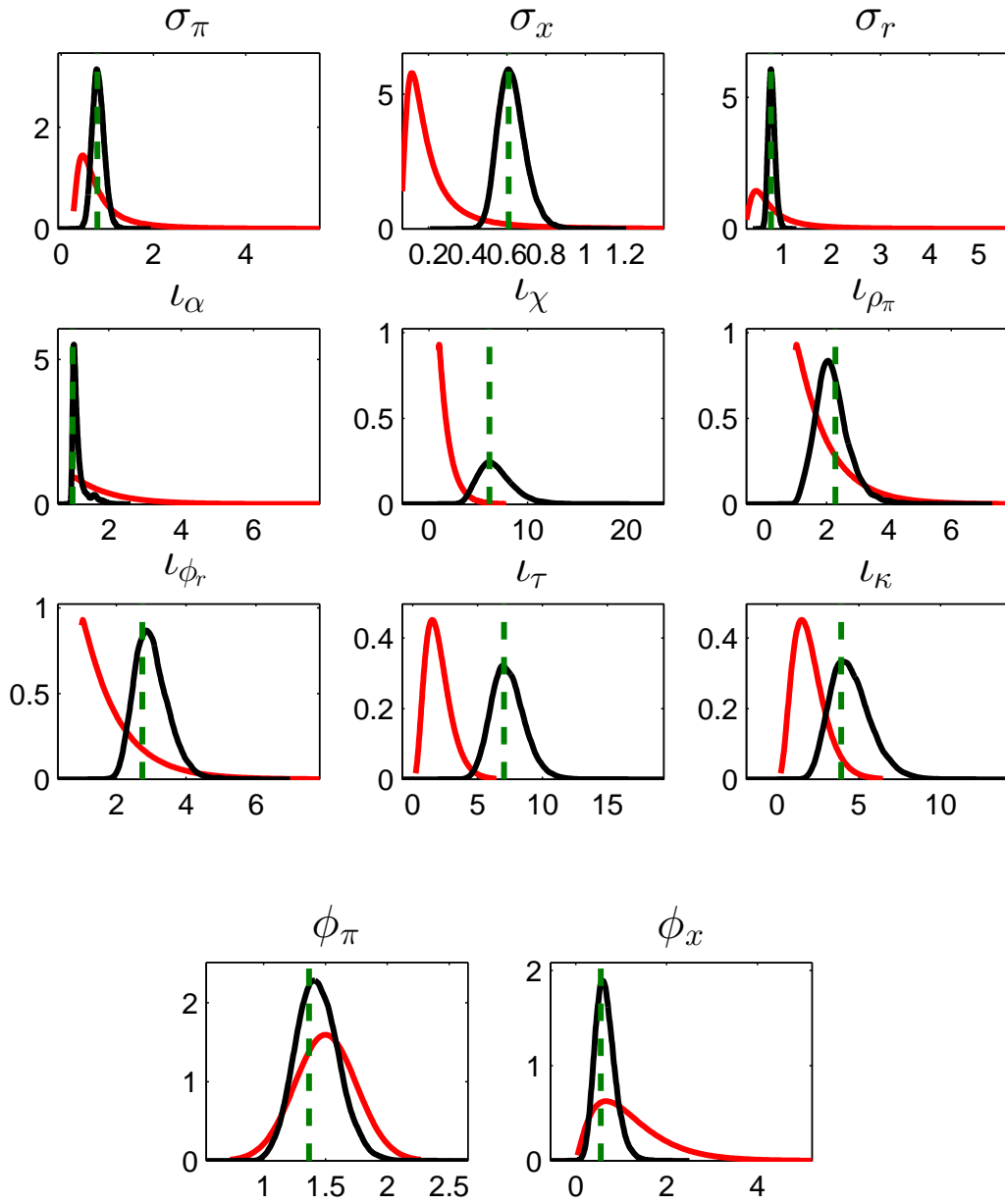


Figure 3: Posterior density estimates for T2 (GI).

Note: See Figure 1. The parameter ι_{ii} with $ii \in \{\alpha, \chi, \rho_\pi, \phi_r, \tau, \kappa\}$ denotes the transformed parameters according to the corresponding rules given in section 4.1.

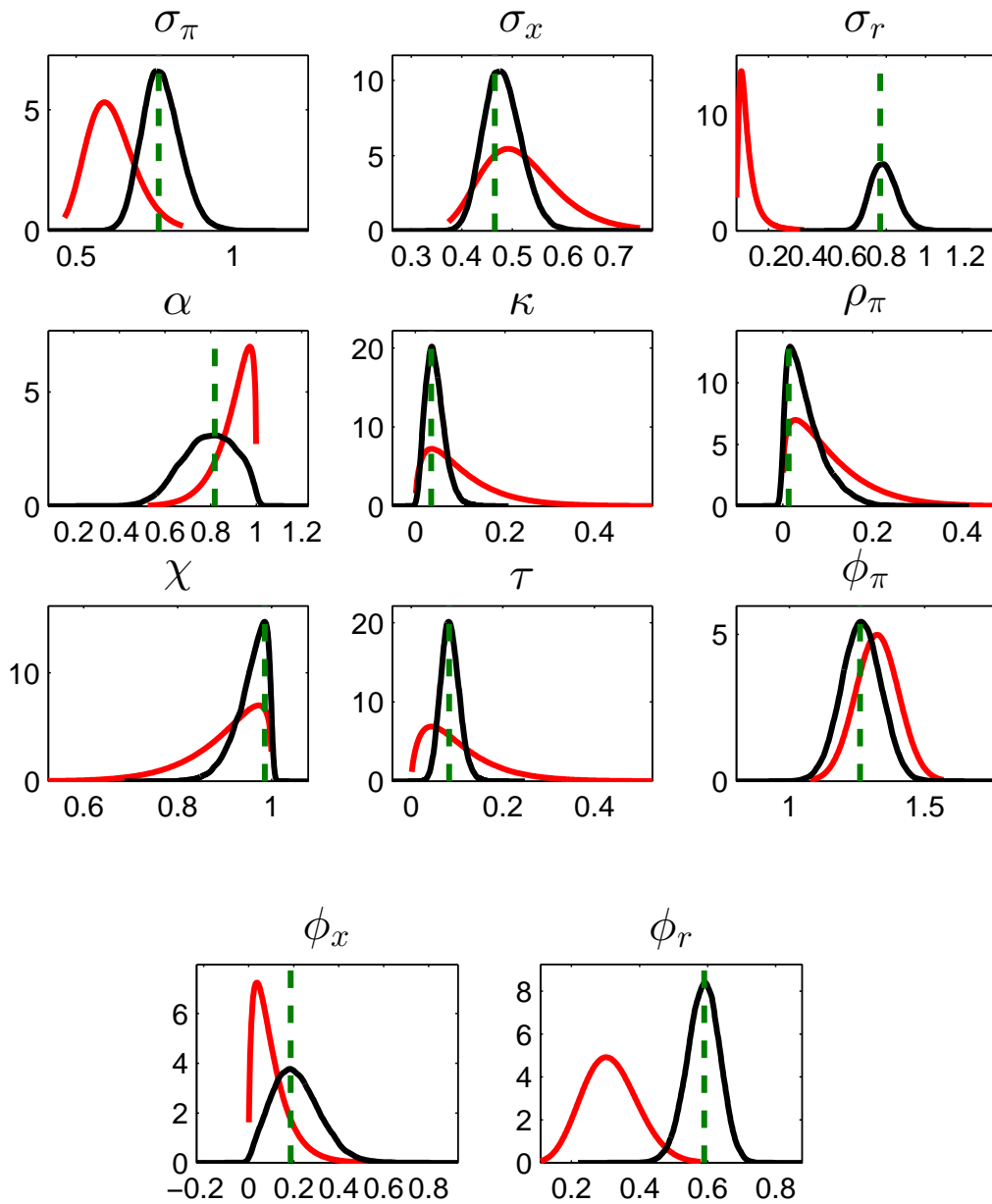


Figure 4: Posterior density estimates for V0 (GI).

Note: See Figure 1.

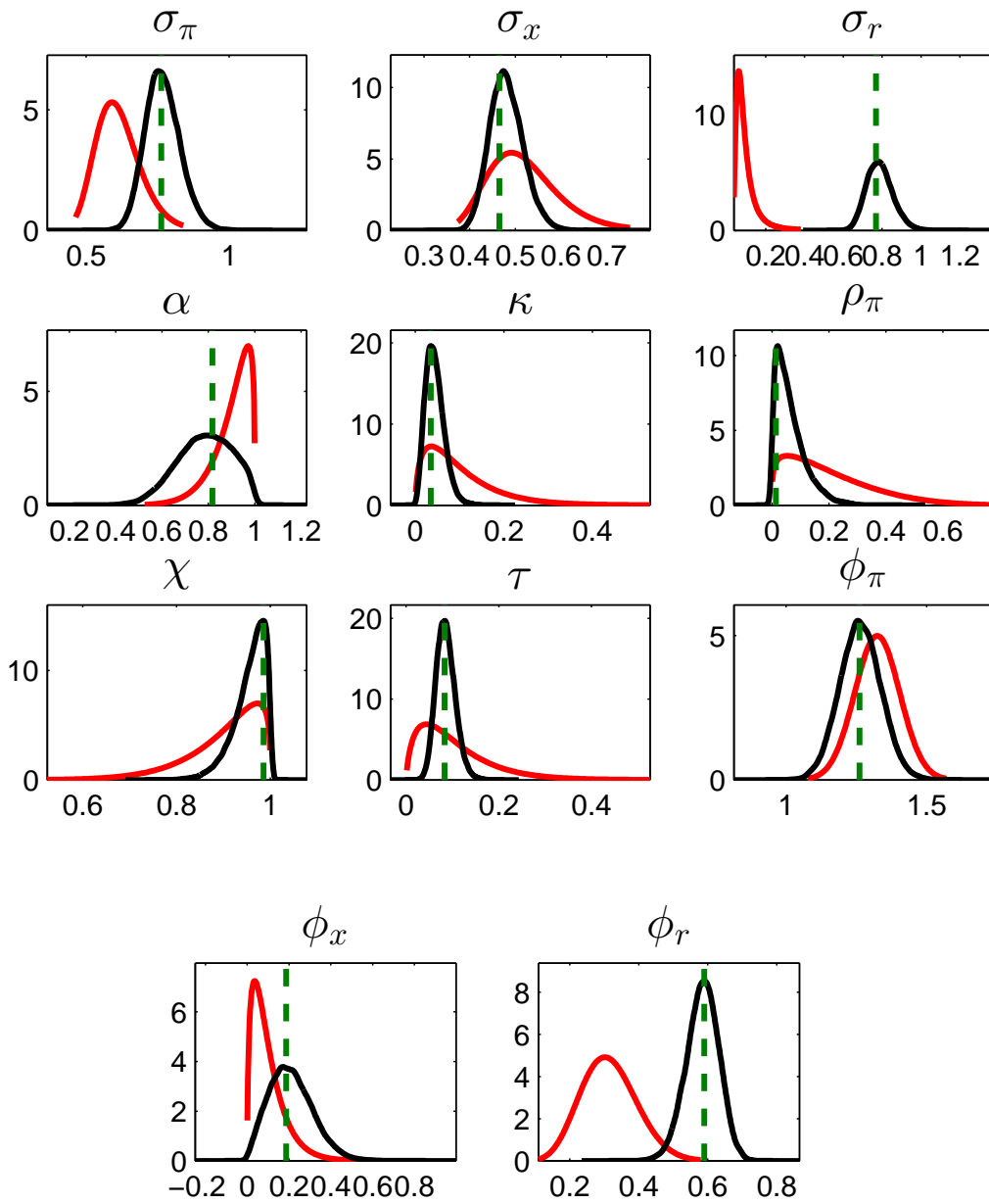


Figure 5: Posterior density estimates for V1 (GI).

Note: See Figure 1.

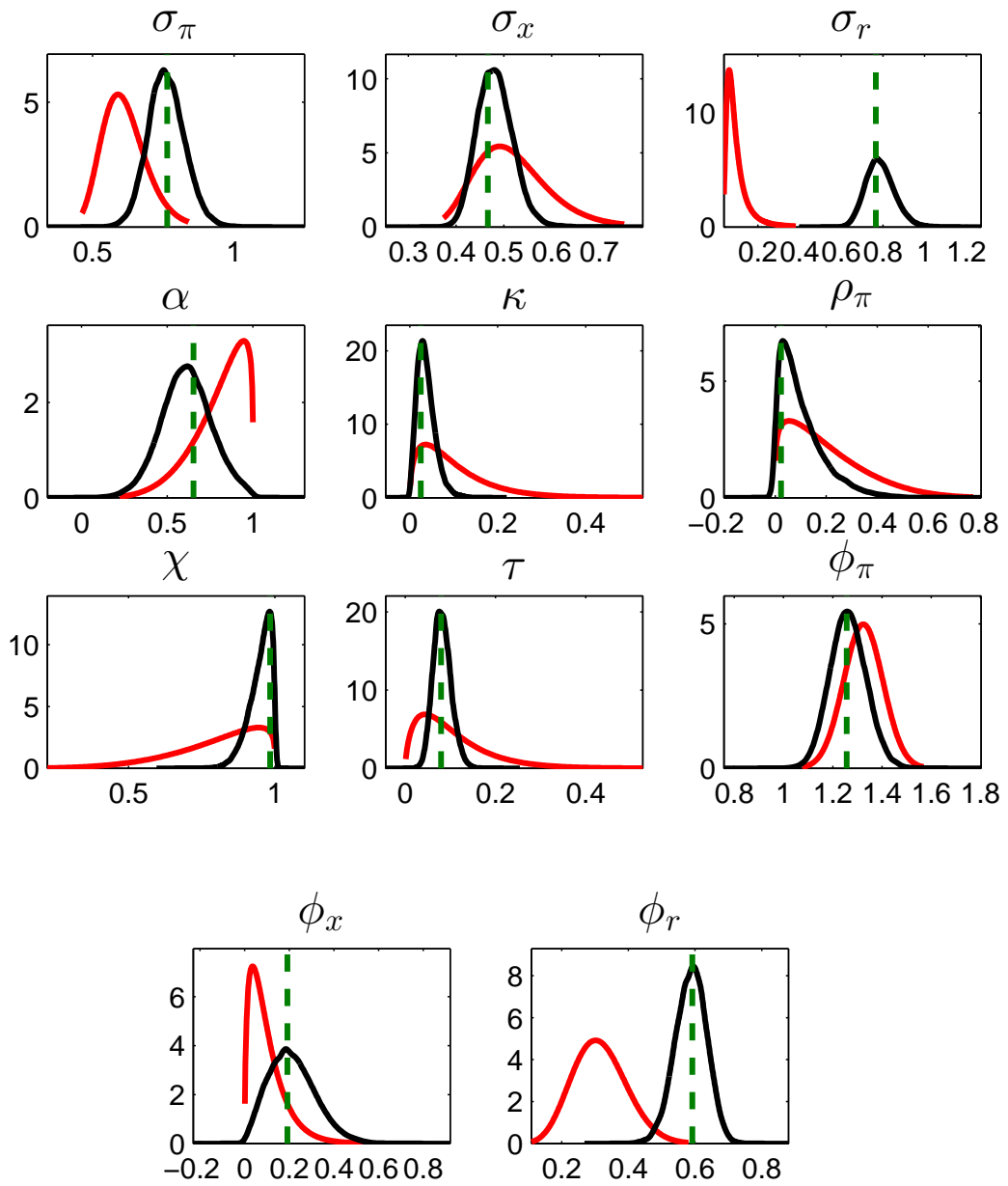


Figure 6: Posterior density estimates for V2 (GI).

Note: See Figure 1.

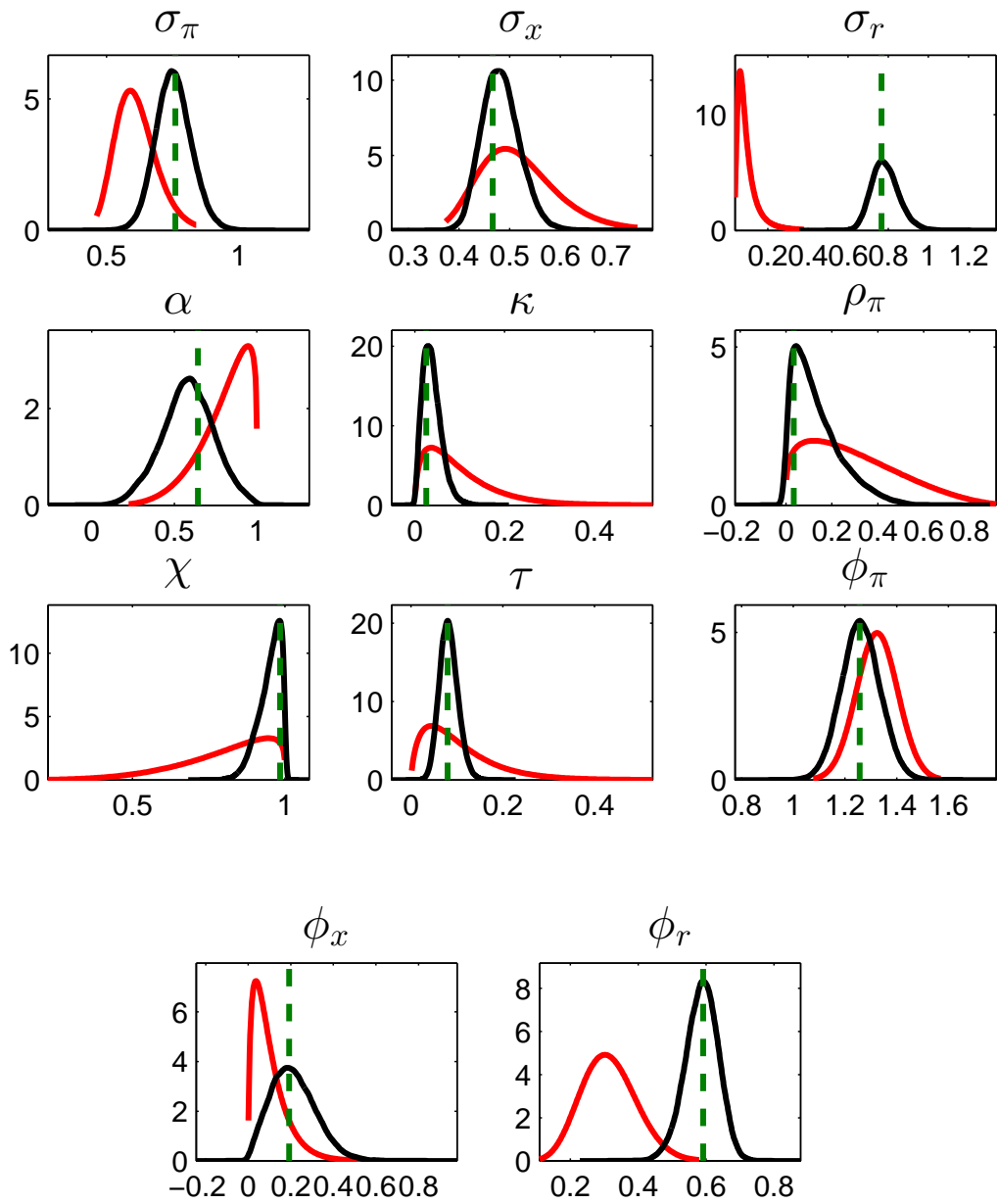


Figure 7: Posterior density estimates for V3 (GI).

Note: See Figure 1.

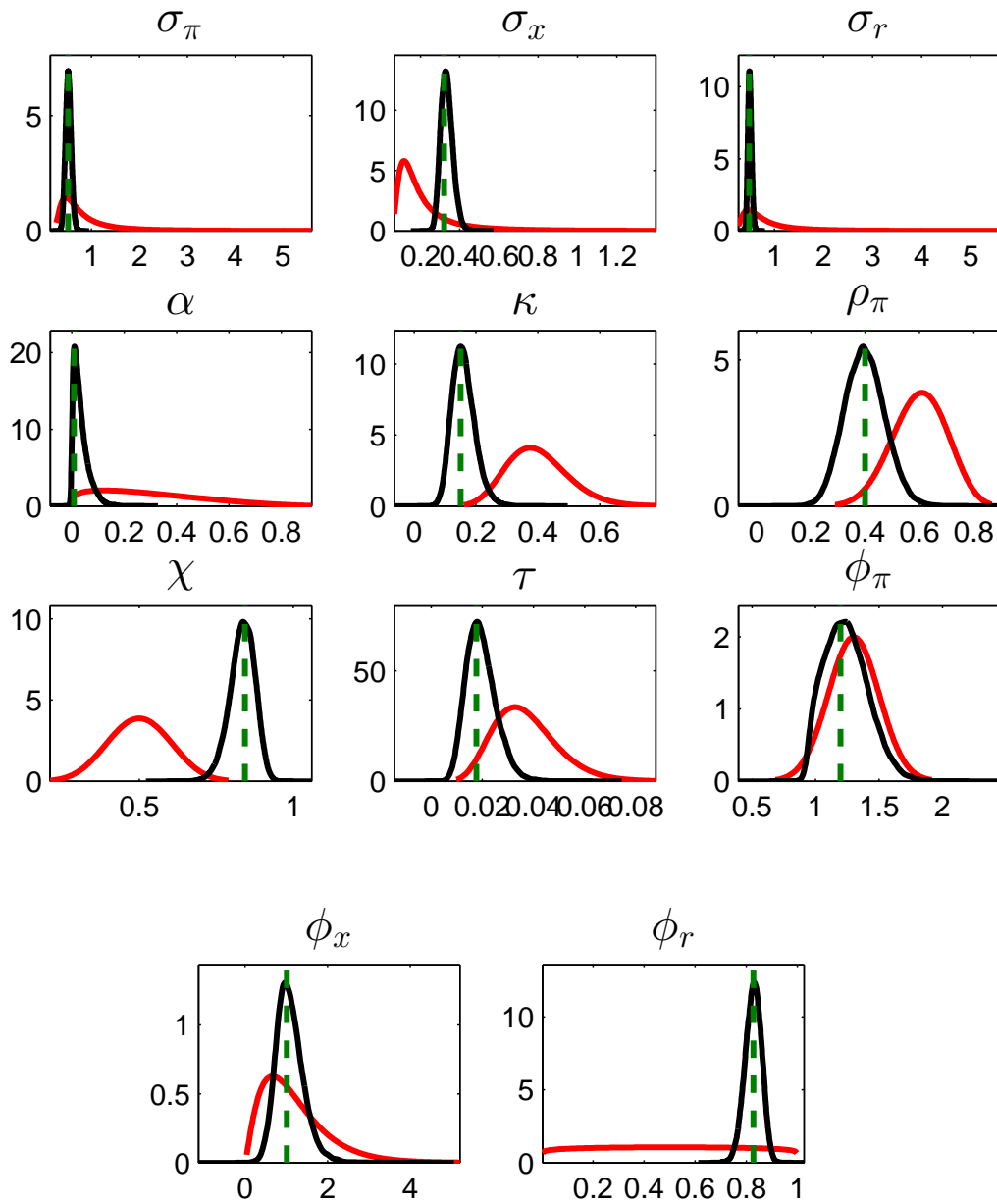


Figure 8: Posterior density estimates for BR (GM).

Note: See Figure 1.

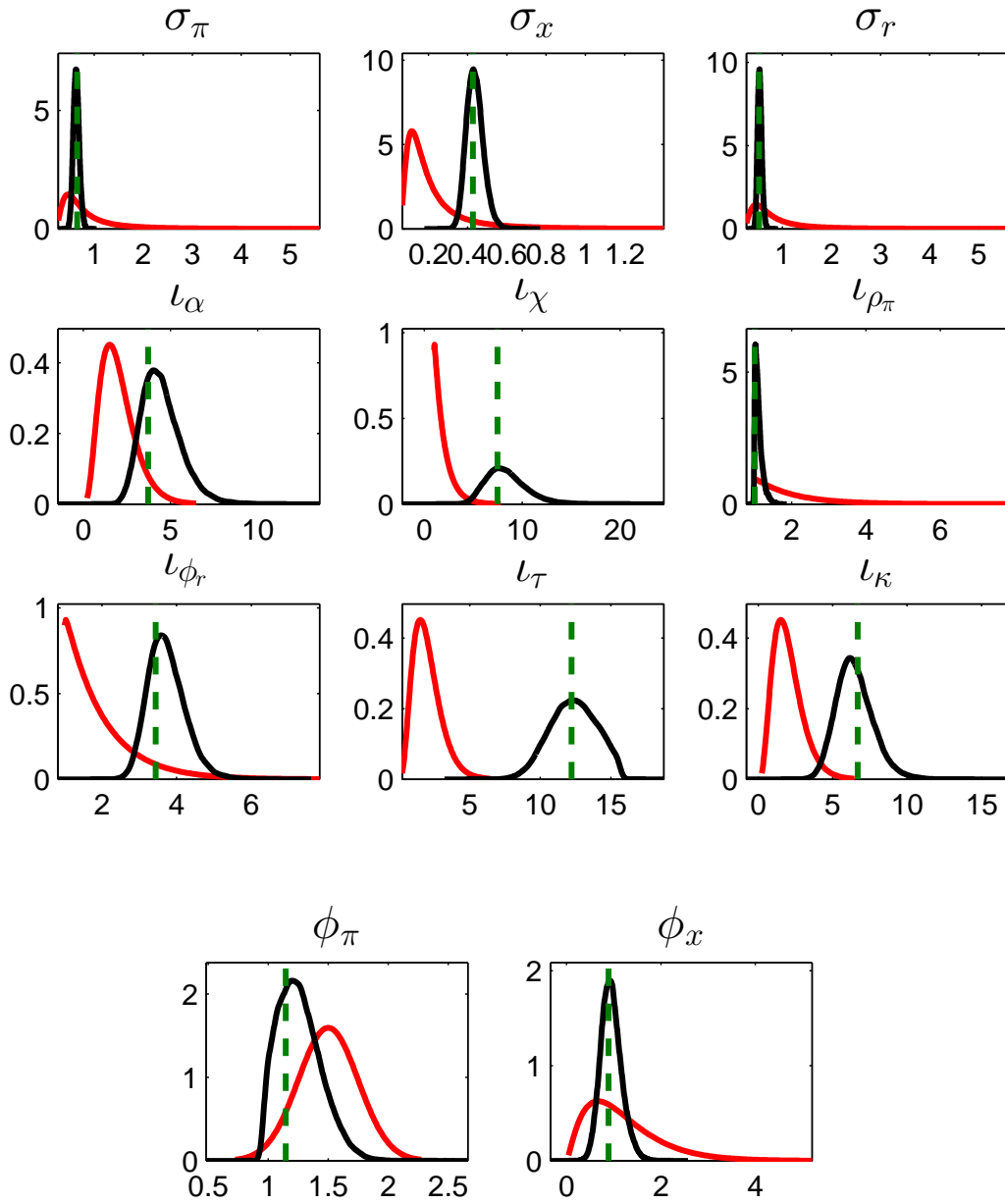


Figure 9: Posterior density estimates for T1 (GM).

Note: See Figure 1. The parameter l_{ii} with $ii \in \{\alpha, \chi, \rho_\pi, \phi_r, \tau, \kappa\}$ denotes the transformed parameters according to the corresponding rules given in section 4.1.

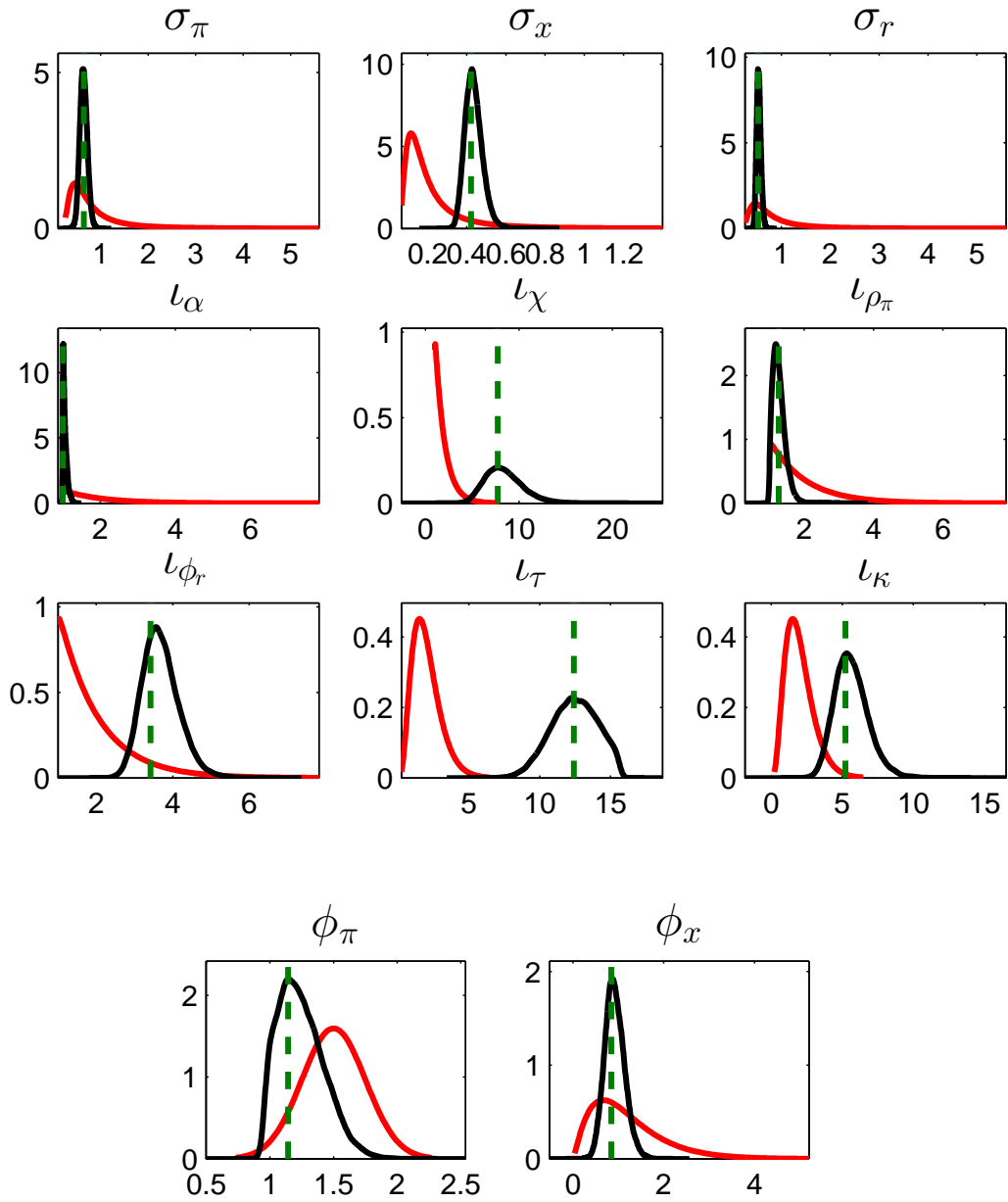


Figure 10: Posterior density estimates for T2 (GM).

Note: See Figure 1. The parameter l_{ii} with $ii \in \{\alpha, \chi, \rho_\pi, \phi_r, \tau, \kappa\}$ denotes the transformed parameters according to the corresponding rules given in section 4.1.

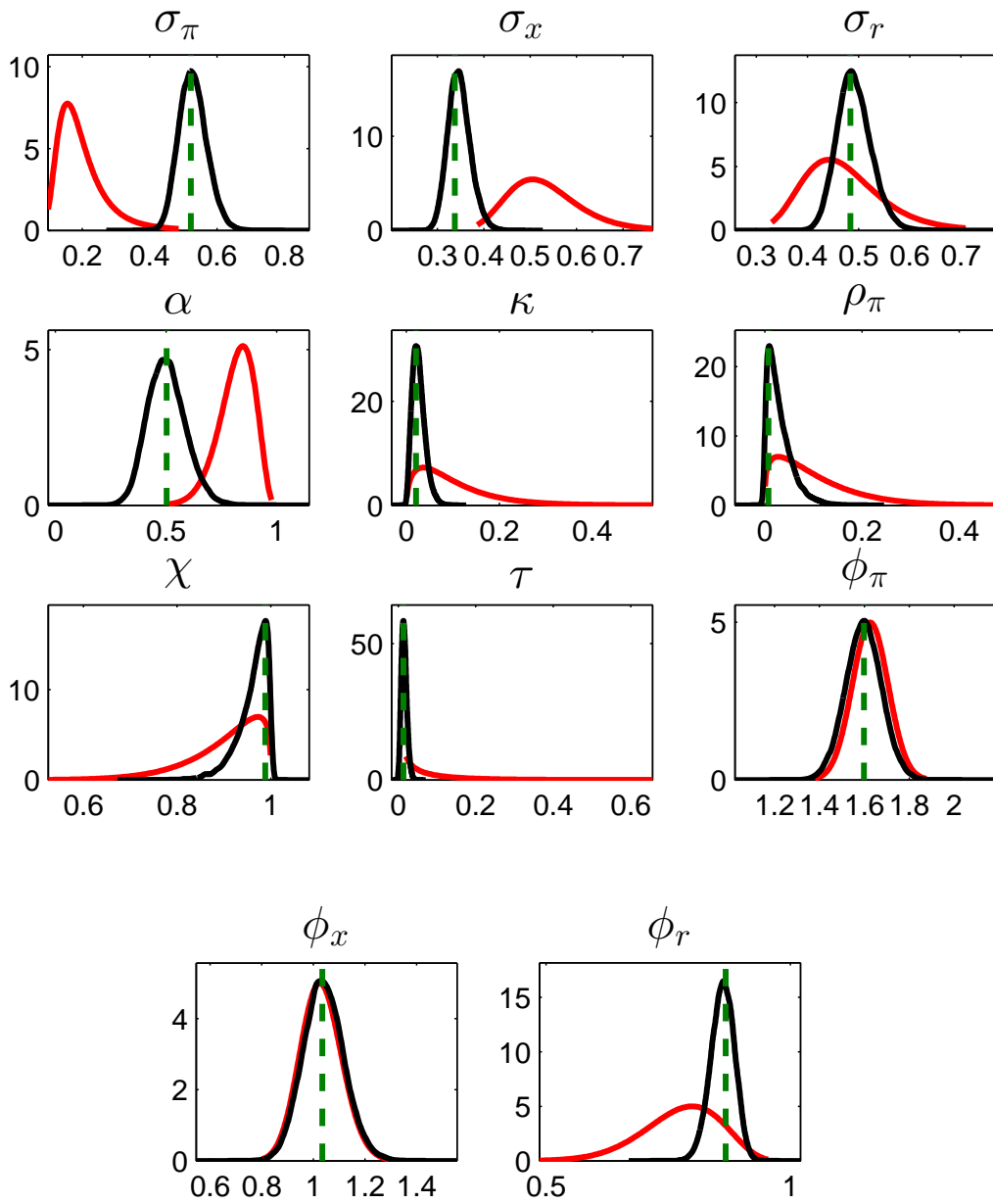


Figure 11: Posterior density estimates for V0 (GM).

Note: See Figure 1.

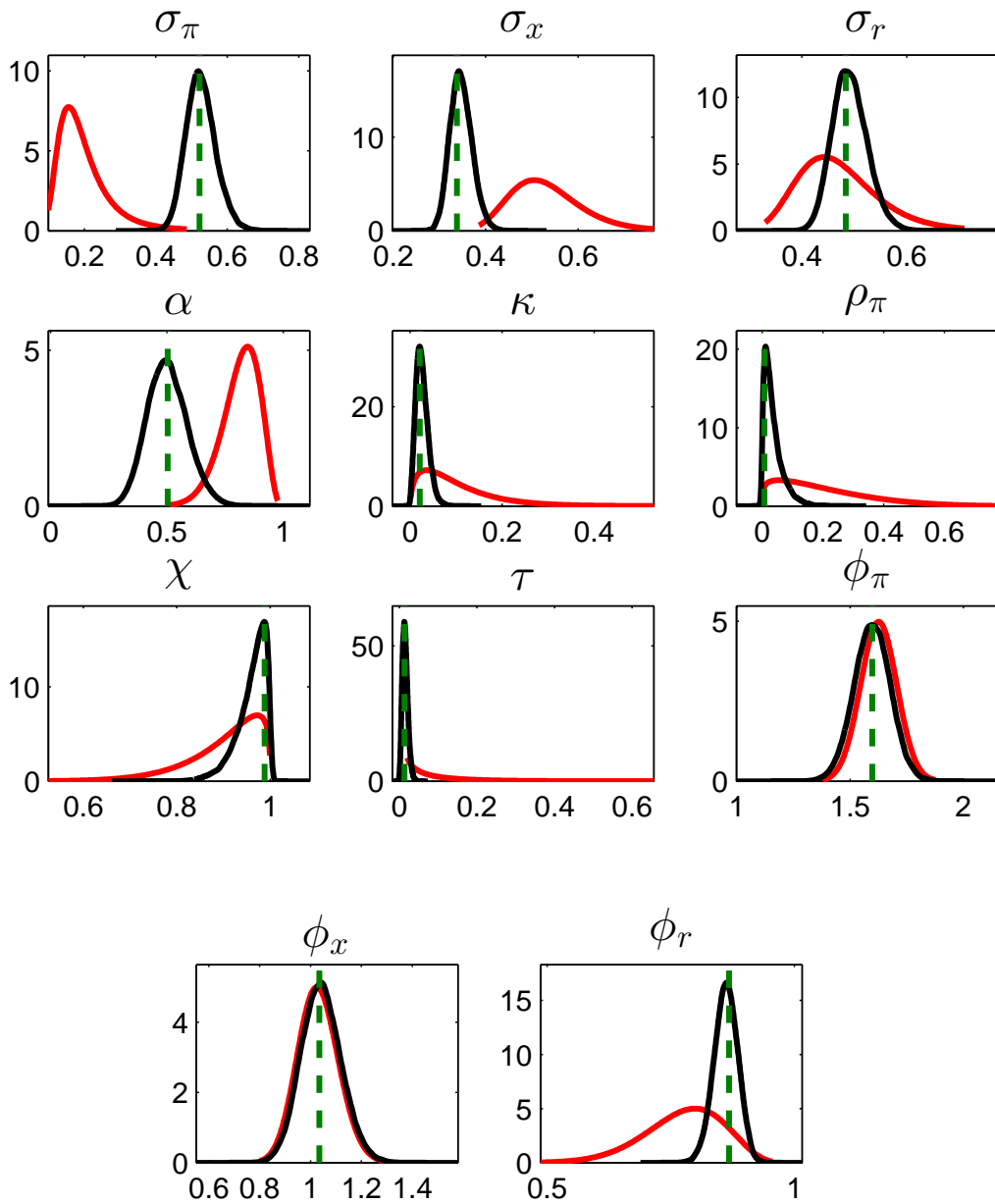


Figure 12: Posterior density estimates for V1 (GM).

Note: See Figure 1.

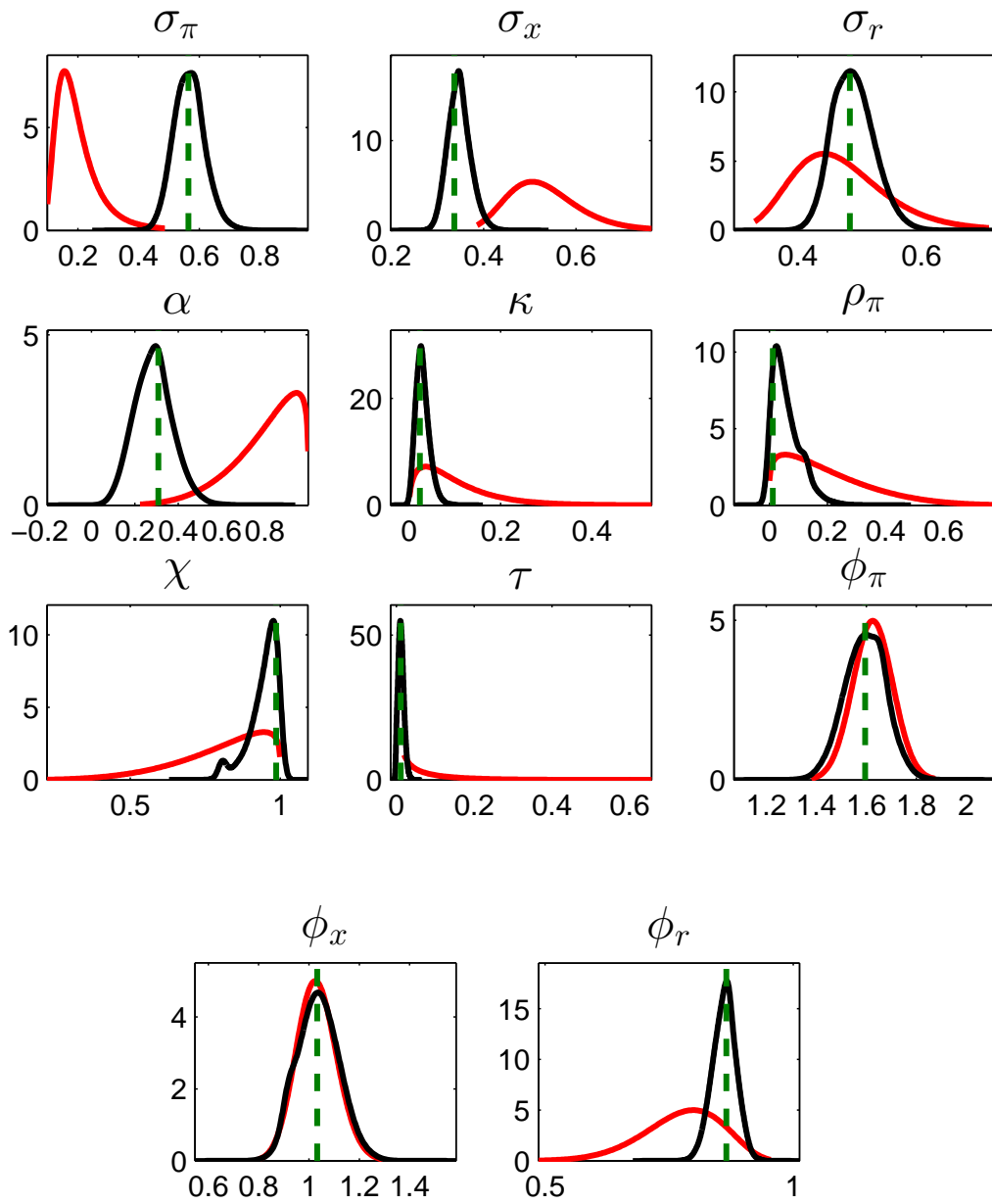


Figure 13: Posterior density estimates for V2 (GM).

Note: See Figure 1.

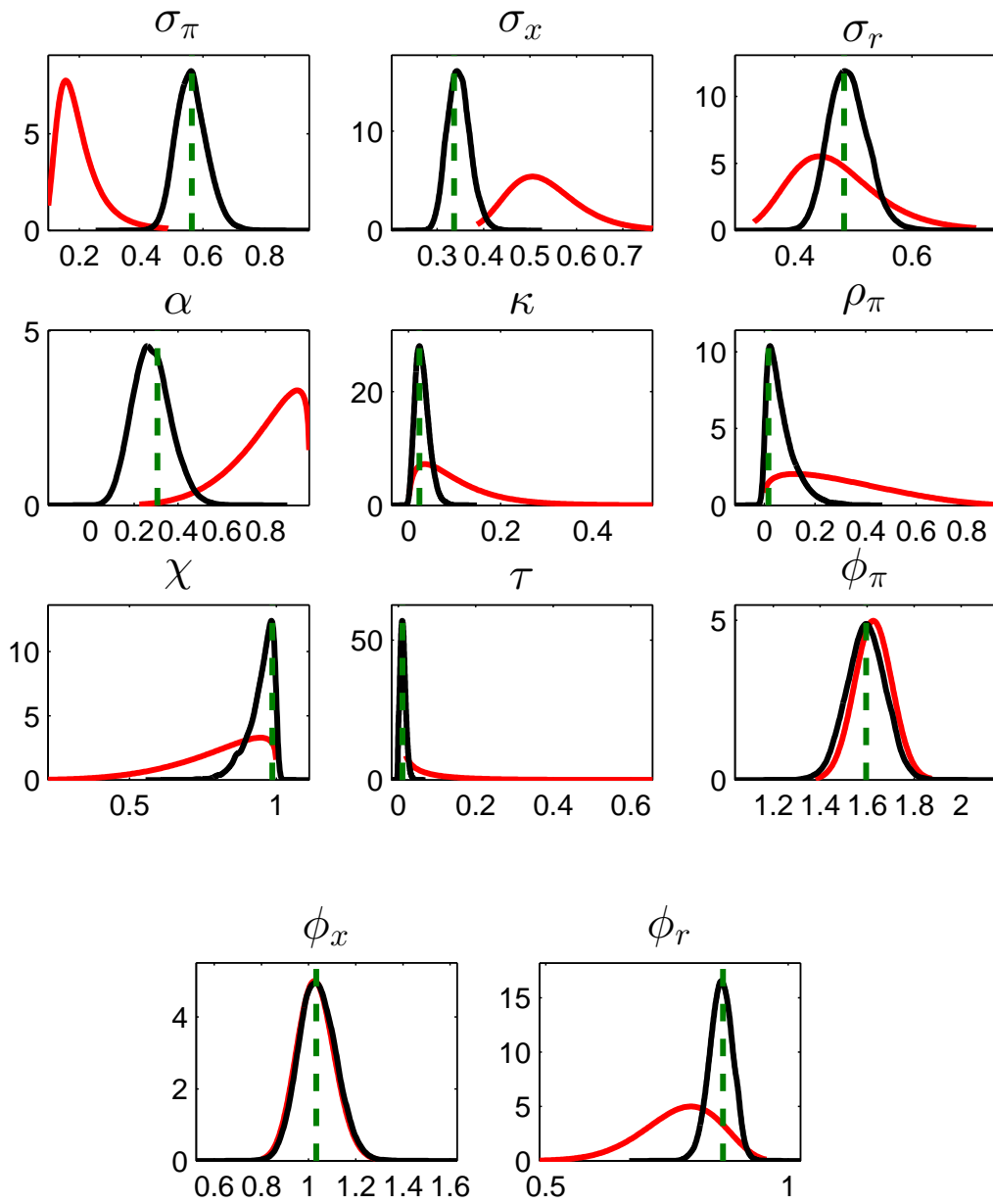


Figure 14: Posterior density estimates for V3 (GM).

Note: See Figure 1.

Literature

BROOKS, S.P. AND GELMAN, A. (1998): General Methods For Monitoring Convergence Of Iterative Simulations. *Journal of Computational and Graphical Statistics*, Vol. 7(4), pp. 434-455.

CARRASCO, M. AND FLORENS, J.-P. (2002): Simulation-Based Method Of Moments And Efficiency. *Journal of Business and Economic Statistics*, Vol. 20(4), pp. 482-492.

CASTELNUOVO, E. (2010): Trend Inflation And Macroeconomic Volatilities In The Post WWII U.S. Economy. *North American Journal of Economics and Finance*, Vol. 21(1), pp. 19-33.

DEL NEGRO, M. AND SCHORFHEIDE, F. (2008): Forming Priors For DSGE Models (And How It Affects The Assessment Of Nominal Rigidities). *Journal of Monetary Economics*, Vol. 55(7), pp. 1191-1208.

FRANKE, R. JANG, T.-S. AND SACT, S. (2012): *Moment Matching versus Bayesian Estimation: Backward-Looking Behaviour in the New-Keynesian Baseline Model*. Economics Working Paper No. 2012-08, Department of Economics, Christian-Albrechts-University of Kiel, September 2012.

LOMBARDI, M.J. AND NICOLETTI, G. (2012): Bayesian Prior Elicitation In DSGE Models: Macro- Vs Micro-priors. *Journal of Economic Dynamics and Control*, Vol. 36(2), pp. 294-313.

SACT, S. (2014a): *Analysis of Various Shocks within the High-Frequency Versions of the Baseline New-Keynesian Model*. Economics Working Paper No. 2014-02, Department of Economics, Christian Albrechts University Kiel, January 2014.

SACT, S. (2014b): *Optimal Monetary Policy Responses and Welfare Analysis*. Economics Working Paper No. 2014-03, Department of Economics, Christian Albrechts University Kiel, January 2014.

WEDDERBURN, R.W.M. (1974): Quasi-Likelihood Functions, Generalized Linear Models, And The Gauss-Newton Method. *Biometrika*, Vol. 61(3), pp. 439-447.