Fundamental shock selection
in DSGE models[*]

Filippo Ferroni
Banque de France
University of Surrey

Stefano Grassi
School of Economics and MaGHiC
University of Kent

Miguel A. León-Ledesma
School of Economics and MaGHiC
University of Kent

April 14, 2015

Abstract

DSGE models are typically estimated assuming the existence of certain structural shocks that drive macroeconomic fluctuations. We analyze the consequences of introducing non-fundamental shocks for the estimation of DSGE model parameters and propose a method to select the structural shocks driving uncertainty. We show that forcing the existence of non-fundamental structural shocks produces a downward bias in the estimated internal persistence of the model. We then show how these distortions can be reduced by allowing the covariance matrix of the structural shocks to be rank deficient using priors for standard deviations whose support includes zero. The method allows us to accurately select fundamental shocks and estimate model parameters with precision. Finally, we revisit the empirical evidence on an industry standard medium-scale DSGE model model and find that government, price, and wage markup shocks are non-fundamental.

Keywords: Reduced rank covariance matrix, DSGE models, stochastic dimension search.

JEL Classification: C10, E27, E32.

1   Introduction

One of the key challenges of modern macroeconomics rests on the identification of the sources of aggregate fluctuations. By specifying a coherent probabilistic structure of economically interpretable endogenous and exogenous processes, DSGE models represent ideal candidates to pin down the shocks driving business cycle fluctuations.[1] A tacit but widespread assumption in the empirical literature on DSGE model estimation is that exogenous disturbances *do* exist in the sense that they capture aggregate economic uncertainty (up to a vector of idiosyncratic measurement errors). Common estimation practice implicitly "imposes" these fundamental shocks by restricting their standard deviation to be non-zero. In a Bayesian context, this assumption is reflected on the prior distributions imposed on the standard deviations of DSGE model shocks (e.g. typically an inverse gamma prior). In classical statistics, standard deviations are re-parameterized by taking logarithmic transformations. In doing so, we rule out boundary solutions and, by construction, structural disturbances always exist.

From an empirical point of view, there is mounting evidence that some of the structural DSGE shocks are unlikely to capture aggregate uncertainty and rather absorb misspecified propagation mechanisms of endogenous variables (see Schorfheide (2013) for an overview). Moreover, it is not infrequent that shocks with dubious structural interpretation are used with the sole purpose of avoiding stochastic singularity and this complicates inference when they turn out to matter, say, for output or inflation fluctuations (see Chari, Kehoe and McGrattan (2008) and Sala, Soderstrom and Trigari (2010)). This is an important question in modern stochastic models of economic fluctuations. Empirically, these models face two challenges. First, unveiling the fundamental innovations that set off fluctuations. Second, identifying the key transmission mechanisms that transform these innovations into business cycles. There is a large literature on the latter. However, because we impose the existence of a set of fundamental shocks, we do not yet understand what are the consequences for inference when estimating a vector of time series with an 'excessive' number of structural disturbances, i.e. what are the consequences for inference when estimating non-fundamental DSGE shocks? This is the first question we tackle in this paper. We then propose a set of easily implementable tools for selecting fundamental shocks entering DSGE models and assess their performance against standard practice.

Non-fundmentalness might arise because some shocks have zero variance or because of the existence of linear combinations among structural innovations. Bar few exceptions (e.g. Cúrdia and Reis (2010)) the vast majority of empirical studies typically postulate orthogonality among innovations and therefore assume a diagonal covariance matrix. Regardless of the shock correlation structure, when taking the model to the data, we want to be able to test, rather then merely postulate, the fundamentalness of structural shocks. I.e., we want to be able to select which innovations are important drivers of aggregate uncertainty. To

---

[1]See Smets and Wouters (2007) and Justiniano, Primiceri and Tambalotti (2010) amongst many others.

be able to estimate the possibly rank deficient covariance matrix of structural shocks, we need to 1) add idiosyncratic measurement errors and 2) abandon standard inverse gamma priors and use distributions (univariate or multivariate) that allow for zero variances (or null eigenvalues).

Using a simple univariate setup, we show analytically that imposing a non-existent exogenous process has deep consequences for inference. E.g. it generates a downward bias in the estimate of the internal persistence of the model. In fully fledged DSGE models, the persistence of model dynamics are controlled not only by the autoregressive parameters, but also by deep parameters capturing real and nominal frictions in the economy. As a result, behavioral parameter estimates can be corrupted as well. With a simulation experiment using a medium-scale model, we quantify the distortions on deep parameter estimates due to incorrect assumptions about the rank of the covariance matrix of structural shock, Σ. In particular, when the econometrician assumes that the rank of Σ is larger than the one of the true DGP, autoregressive parameters and parameters driving price and wage stickiness and indexation are grossly underestimated. The result is that the inclusion of innovations that are not fundamental drivers of macroeconomic uncertainty affects the estimates of the transmission mechanisms. We thus unveil a potential trade-off between the inclusion of a wide set of potential sources of impulses and the correct identification of model parameters that drive propagation.

We then show how these distortions are reduced by considering priors on the structural shocks covariance matrix that allow for rank deficiency. In the context of uncorrelated disturbances, truncated or un-truncated priors can be implemented as long as they attribute non-zero probability to zero standard deviations. We show that proper priors such as normal or exponential distributions have appealing properties since they allow us to recover fundamental and non-fundamental shocks in situations where the true number and combination of fundamental shocks is unknown. In the context of a more general structural shocks covariance matrix, we propose considering the conjugate Metropolis-within-Gibbs sampler proposed in Cúrdia and Reis (2010) adapted for rank deficient covariance matrices [2].

We explore the consequences of our approach in an empirical application and revisit the evidence of a industry standard DSGE model. We estimate the Smets and Wouters (2007) model on seven key quarterly macroeconomic time series, namely, the growth of real output, consumption, investment, and real wages, and hours, inflation, and interest rates. For comparability proposes, we consider the original data span, 1968-2004, with revised data, and only depart from the baseline specification of the model by assuming normal priors on standard deviations and by adding measurement errors.[3] Our findings show that, first, government spending and price markup shocks are non-fundamental for the 1968-2004 sample period and larger samples, and the wage markup shock is non-fundamental for

---

[2] In a rank deficient environment, we consider the singular Inverse Wishart distribution (see Uhlig (1994) and Díaz-García and Gutiérrez-Jáimez (1997)) and the conjugacy results in Díaz-García and Gutiérrez-Jáimez (2006).

[3] We also run estimates with the vintage data and with a sample including more recent years (i.e. up to 2014).

larger estimation spans, i.e. 1968-2009 or onwards. Technology, investment, preference and monetary policy shocks are fundamental. Interestingly, such clustering is very similar to the Chari et al. (2008) classification of structural and non-structural shocks[4]. Second, we show that, when allowing for zero standard deviations, the uncertainty previously coming from these non-fundamental shocks is pushed towards measurement error. This is consistent with Justiniano, Primiceri and Tambalotti (2013), who find that the variability arising from wage markup shocks typically estimated in New Keynesian models is explained mostly by measurement error in wages. Third, the estimated posterior distributions of deep parameters are different when we allow for the possibility of a rank deficient structural shocks matrix. As a consequence, the transmission mechanism of the fundamental shocks is also altered. In particular, those of monetary policy and government spending shocks. Fourth, by means of marginal likelihood comparisons, data prefer versions of the model with a rank deficient shocks structure.

Our methodology is related to the literature on stochastic model specification search in state space models. We draw from Fruhwirth-Schnatter and Wagner (2010) for the selection of structures in unobserved components models or in time varying parameter VAR models as in Belmonte, Koop and Korobilis (2014) or Eisenstat, Chan and Strachan (2014). We build on that literature by proposing to estimate *jointly* the structural parameters and the stochastic specification of the DSGE shock structure. Our paper is also, albeit indirectly, related to the vast literature studying misspecification problems in DSGE model estimation. Invalid cross-equation restrictions (e.g. Ireland (2004), Del Negro and Schorfheide (2009), Inoue, Kuo and Rossi (2014)), parameter instability of various forms (e.g. Fernández-Villaverde and Rubio-Ramírez (2008), Galvao, Giraitis, Kapetanios and Petrova (2015), Canova, Ferroni and Matthes (2015)), incorrect assumptions about shock dynamics (Cúrdia and Reis (2010)), low frequency movements mismatches (e.g. Gorodnichenko and Ng (2010), Ferroni (2011), Canova (2014)), etc., may all plague inference in DSGE models. However, the literature so far is silent on the issue of interest of this paper. We are concerned with *redundant* model-based shocks which can generate distorted estimates and corrupt inference *when forced to be fundamental*.

The remainder paper is organized as follows. Section 2 presents the econometric setup and estimation procedures. Section 3 presents the inference distortions caused by incorrect assumptions about the rank of Σ. Two models are considered: a standard RBC model to convey intuition, and a medium scale DSGE model to measure distortions in models typically used for policy analysis. Section 4 presents the main results of our empirical investigation. Section 5 draws a number of concluding remarks.

---

[4]Quoting Chari et al. (2008), 'We divide these [shocks] into two groups. The *potentially structural shocks* group includes shocks to total factor productivity, investment-specific technology, and monetary policy. The *dubiously structural shocks* group includes shocks to wage markups, price markups, exogenous spending, and risk premia.'. Bar the risk premia shock, we provide additional empirical support to their claim.

2 GENERALIZED PRIORS FOR FUNDAMENTAL DSGE SHOCKS

Consider a DSGE model with (deep) parameters of interest θ. The (control and state) variables of the model, denoted by $\{s_t\}$, are driven by structural shocks with innovations ϵ_t. The model is characterized by a set of equations that define the steady state values s^* and Euler equations that describe the transition dynamics. Linearizing around the steady state gives a system of expectational difference equations that can be solved to yield a solution in the form of difference equations. The linearized solution of a DSGE has the following representation

$$s_{t+1} = A(\theta)s_t + B(\theta)\ \epsilon_{t+1} \tag{1}$$

$$with\ \epsilon_{t+1} \sim N_r(0, \Sigma) \tag{2}$$

where A, B are nonlinear functions of the structural parameters of the model, ϵ_t is a $n \times 1$ vector of the structural (or fundamental) innovations, and s_t is the $n_s \times 1$ vector of endogenous and exogenous states. Σ is a covariance matrix of dimension $n \times n$ whose rank is $r = rank(\Sigma) \leq n$. We denote by $N_r(0, \Sigma)$ the $n \times 1$ multivariate singular Normal distribution with rank r.

If the eigenvalues of $A(\theta)$ are inside the unit circle, the latter structure can be mapped into a $VMA(\infty)$ (see Komunjer and Ng (2011)) representation as follows

$$s_t = (I - A(\theta)L)^{-1}B(\theta)\ \epsilon_t = \Gamma(L;\theta)\ \epsilon_t \tag{3}$$

where L is the lag operator. The mapping from the model based variables to a $n_y \times 1$ vector y_t of observed times series is accomplished through a measurement equation augmented with series specific i.i.d. shocks in order to avoid the possibly stochastically singular model, i.e.

$$y_t = \Lambda\ s_t + e_t = \Phi(L;\theta)\ \epsilon_t + e_t \quad e_t \sim N(0, \Omega) \tag{4}$$

where Λ is a selection matrix , $\Phi(L;\theta) = \Lambda\ \Gamma(L;\theta)$ is full column rank and e_t is a $n_y \times 1$ vector of normal i.i.d. errors. The vector of observables, y_t, that the econometrician uses in estimation is unconstrained and can be of any dimension. It could be larger or equal than the number of structural shocks, i.e. $n_y = n_s \geq n$. Or the observable set can even be larger and include various proxies for the same model-based quantity, i.e. $n_y > n_s > n$, e.g. as in Boivin and Giannoni (2006) or Canova and Ferroni (2011).

To gain more intuition we can rewrite the system as follows, for $j = 1, ..., n_y$

$$y_{j,t} = \Phi_j(L;\theta)\ \epsilon_t + e_{j,t} \tag{5}$$

where Φ_j correspond to the jth raw of $\Phi(L;\theta)$. The fundamental shocks and the measurement shocks are separately identifiable since the former are common and the latter are series specific. Moreover and more importantly, the measurement errors being i.i.d., they cannot explain the cross- and auto-correlation structure of the data, which is entirely determined by the common component, i.e. the DSGE model shocks and its MA structure.

Equation ([5](#)) can be seen as an approximate dynamic factor model where the row vector $\Phi_j(L;\theta)$ represents the factor loadings and ϵ_t the common factors (orthogonal to each other). There is, however, an important difference. While in the factor model we are interested in the number of factors, in this setup we are interested also in the combination of underlying common shocks since they have economic interpretations. This can be accomplished by studying the *null space* of Σ. If we assume Σ diagonal, it is sufficient to check that standard deviations are not zero. If Σ is a symmetric positive definite matrix, the non-zero eigenvalues correspond to the fundamental shocks.

An alternative way to select the number and combination of fundamental DSGE shocks is to compute the marginal likelihood of model specifications with different combinations of structural shocks. However, this can be time consuming because it requires estimating each of the possible models. In models with typically 7 or 8 postulated fundamental shocks, the combinations of models to compare is very large and marginal likelihood comparisons will not be a very useful tool for selection.[5] Our argument is even more persuasive for non-linear models for which the computation of the marginal likelihood is very burdensome. In our approach, the selection of fundamental shocks is done in one step and considering all the observables simultaneously.

Finally, it is important to highlight that, if we have strong priors that a subset of structural shocks do indeed exist, we can postulate inverse gamma priors on the standard deviations of that subset and be more agnostic on others. The shape of priors used are not neutral, in the sense that they are imposing an a priori structure on the sources of business cycle fluctuations. This may be convenient in cases in which we have a strong view about what the sources of uncertainty are. However, the point that we would like to stress is that assuming inverse gamma priors for the standard deviations of shocks that are not fundamental may create severe distortions for inference. The quantitative implications of this are explored in the next subsections. Before that, we need to outline the estimation procedures, priors, and posterior simulators to tackle the estimation of covariance matrices of structural shocks that are rank deficient.

3 Estimation framework

In this section we discuss the prior distributions and posterior samplers used to estimate rank deficient covariance matrices. Since they are different when disturbances are correlated or uncorrelated, we analyze them separately. As mentioned above, our methodology is related to the literature on Bayesian stochastic variable selection in state space models as in Fruhwirth-Schnatter and Wagner (2010). Appendix [A.2](#) briefly summarizes the key ideas of this method and its correspondence with DSGE shock selection.

---

[5]Suppose we have a model with 7 shocks and we sequentially choose between models with 6, 5, 4, 3, and 2 shocks, this would imply estimating 122 models.

3.1 Uncorrelated disturbances, Normal and Exponential priors

Equations (1) and (2) can be rewritten as

$$s_{t+1} = A(\theta)s_t + B(\theta)\ \Sigma^{1/2}\eta_{t+1}$$

$$with\ \eta_{t+1} \sim N(0, I_n)$$

where I_n is the identity matrix and $N(0, I_n)$ is the multivariate normal distribution. While the standard deviation of η_{t+1} is fixed and normalized to one in estimation, the diagonal elements of $\Sigma^{1/2}$ are estimated. We consider classes of prior distributions for the diagonal elements of $\Sigma^{1/2}$ such that the probability of zero is positive, i.e. for $j = 1, ..., n_\epsilon$ we assume that $p(\sigma_j = 0) > 0$. It is important to notice that structural standard deviations are not identified up to sign switch, e.g. $\epsilon_i \sim (0, \sigma_i^2) = \pm\sigma_i\eta \sim (0, 1)$. In other words, the model in the equation with $(-\Sigma^{1/2})(-\eta_{t+1})$ is observationally equivalent to the same model with $\Sigma^{1/2}\eta_{t+1}$. As a consequence, the likelihood function is symmetric around zero along the σ_j dimension and bimodal if the true σ_j is larger than zero. This fact can be exploited to quantify how far the posterior of σ_j is from zero and, in turn, assess fundamentalness. One could also, as it is standard practice, normalize the sign to a positive value and estimate the standard deviations over a non-negative support.

We propose to use the following priors:

1. *Exponential priors*

$$\sigma_j \sim Exp(\lambda_j)$$

   With exponential priors, we fix the sign to be non-negative (but allowing for zero) prior to estimation. In order to assess the fundamentalness of specific shocks, we rely on the confidence sets of the posterior distribution and the statistical distance from zero. Standard Bayesian simulators such as the RW Metropolis-Hastings can be employed to recover the posterior distribution of the parameters.

2. *Normal priors*

$$\sigma_j \sim N(\mu_j, \tau_j^2)$$

   This implies estimating the non-identified standard deviations and fixing the sign after estimation. Accordingly, the prior for structural standard deviations covers the entire real line support. In such a case, the bi-modality of the posterior distribution of the standard deviation implies **existence** of the structural shock in question. Uni-modality (centered on zero) would then imply **non-existence**. In other words, with normal priors, we exploit the information contained in the non-identifiability of the sign. If a shock exists, then the sign should not be identifiable. Standard Bayesian simulators such as the RW Metropolis-Hastings can be employed to recover the posterior distribution with an additional random sign switch of the shocks' standard deviations. Appendix A.3 describes the procedure to introduce the sign switch in an otherwise standard MCMC.

Although this step is not essential, it helps prevent the MCMC chain from getting stuck in one of the two modes.

3.2 Correlated disturbances

Our approach can be extended to situations where some structural disturbances are correlated. In such circumstances, we might be interested in estimating a non-diagonal covariance matrix, Σ, with rank $r = rank(\Sigma) < n$. Such practice might be motivated by the findings in the empirical factor model literature for which aggregate macroeconomic fluctuations are typically explained by an handful of correlated factors. Moreover, Cúrdia and Reis (2010) offer reasons for why arbitrary restrictions on the correlation structure of DSGE model disturbances may be incorrect. It is possible to design an estimation procedure that accounts for both a non-diagonal covariance structure of the data and a rank-deficient covariance matrix.

The estimation procedure combines the ideas of the conjugate-conditional algorithm of Cúrdia and Reis (2010) and the singular generalized inverse Wishart (see Uhlig (1994) and Díaz-García and Gutiérrez-Jáimez (1997)). In this section, we briefly outline the estimation procedure and leave the detailed description of the posterior sampler in Appendix A.4. In particular, the sampling of the parameters needs to be partitioned in two blocks: the covariance matrix of the structural shocks (Σ) and all other parameters (θ). Conditional on a value of θ and on a sequence of states $\{s_t\}_{t=1}^T$, we can derive a sequence of i.i.d. structural shocks from (1) as follows,

$$B(\theta)^+ (s_{t+1} - A(\theta)s_t) = \epsilon_{t+1} \sim N_r(0, \Sigma) \qquad (6)$$

where $B(\theta)^+$ is the left Moore-Penrose generalized inverse of $B(\theta)$. The likelihood of the multivariate normal singular distribution is similar to the non singular one with two main differences. First, the determinant of the covariance matrix is replaced by the product of non-zero eigenvalues. Second, we rescale the sum of the square deviations of the structural shocks by the Moore-Penrose generalized inverse of Σ. Therefore, the likelihood of (6) is given by

$$L(\Sigma|Z) \propto \left(\prod_{k=1}^r \lambda_k \right)^{-\frac{T}{2}} \exp\left(-1/2 \; \text{trace}(\Sigma^+ Z'Z)\right)$$

where λ_k are the non-null eigenvalues of Σ and $Z = (\epsilon_1, ..., \epsilon_T)'$. The latter, combined with a (uninformative or informative) prior gives rise to an expression that is proportional to the kernel of an n-dimension Singular Generalized Inverse Wishart of rank r.[6] We denote this with $W^+(r, \nu, G)$, where $\nu = T - n + 1$ are the degrees of freedom and $G = Z'Z$ the scale matrix. Sampling from this distribution can be accomplished by a sequence of simple steps outlined in Algorithm 3 in Appendix A.4.

---

[6]See the conjugacy results in Díaz-García and Gutiérrez-Jáimez (2006)

With this block structure, we can sequentially sample from Σ given θ and, vice versa, sample θ given Σ. More formally, given r, $\Sigma_{(0)}, \theta_{(0)}$

Algorithm 1

1. Given $\Sigma_{(j-1)}, \theta_{(j-1)}$, we draw a sequence of states, $s_{1:T}^{(j)}$. This distribution is derived from the state space setup and, given linearity and Gaussianity assumptions, it is a normal distribution.

2. Given $\theta_{(j-1)}$ and $s_{1:T}^{(j)}$, draw $\Sigma_{(j)}$ from the n-dimension Singular Generalized Inverse Wishart, $W^+(r, \nu, G^{(j)})$.

3. Given $\Sigma_{(j)}$, draw θ^* form a normal centered in $\theta_{(j-1)}$ and accept the draw with a Metropolis-Hastings probability, i.e. $\min\left\{\frac{L(y_{1:T}|\theta^*, \Sigma_{(j)})p(\theta^*)}{L(y_{1:T}|\theta_{(j)}, \Sigma_{(j)})p(\theta_{(j-1)})}, 1\right\}$.

In steps 1 and 3 the likelihood and the distribution of the states are computed using the Kalman filter recursions. Since the state space is augmented with n_y measurement errors, the covariance matrix of the observables is full rank, hence invertible. Therefore, the Kalman gain, defined as the product of the covariance between states and observables times the inverse of the variance of the observables, can be computed and all the remaining recursions are unaffected.

This algorithm relies on the assumption that the rank of the covariance matrix of structural shocks is known. In applied work, of course, this is not the case. Two approaches can be used to tackle this problem. The first is to run a preliminary test on the data to select the number of common factors that explain a pre-specified portion of the volatility of the observed data. The second is to estimate different specifications with increasing rank dimension from 1 to the number of shocks and select the one that maximizes the marginal likelihood. Once the rank of the covariance matrix is established, redundant or non-fundamental shocks can be obtained by looking a the null space of the posterior distribution of the covariance matrix.

4 Non fundamental DSGE shocks and inference distortions

We now tackle the question of whether the introduction of non-fundamental shocks affects the estimation of parameters governing transmission in the model. We first convey the intuition using a simple model and we then move to an industry-standard New Keynesian DSGE model to quantitatively assess the distortions induced in behavioral parameters and their policy implications.

4.1 Example and experiments with a toy RBC model

We start first by studying the likelihood of the simplest DSGE model, a plain vanilla Real Business Cycle (RBC) model with inelastic labor supply, full capital depreciation, and an

autoregressive process of order one for total factor productivity (TFP) shocks. In this setting, an analytical solution can easily be derived.[7] We obtain the following recursive representation

$$z_{t+1} = \rho z_t + \sigma \epsilon_{t+1}$$

$$k_{t+1} = \alpha k_t + z_t$$

$$y_t = z_t + \alpha k_t + e_t$$

where $\epsilon_{t+1} \sim N(0,1)$ and small case variables indicate the log deviation of the variables from the non stochastic steady state. In particular, k_t is capital per capita, z_t is TFP, y_t is output per capita, and α is the capital share in production. We assume that we observe neither the technology process nor capital. We observe output up to a normal $(0, \omega^2)$ measurement disturbance e_t.

We run a controlled experiment to measure the impact of different priors on the DSGE model shock standard deviations. We simulate artificial data from the RBC model by calibrating structural parameters values to standard values in the literature, i.e. $\alpha = 0.33$, $\rho = 0.70$. We generate data assuming that the technology shock is non fundamental (i.e. standard deviation 0) and fundamental (i.e. standard deviation equal to .05, 0.1, 0.2). We fix the variance of ω to 0.08, i.e. the mean of the range of values of the structural standard deviation. The results obtained in this section are largely invariant to the values of structural and non-structural parameters used to generate data, to the the sample size, and to the location of priors and scale parameters.[8]

We generate 500 data points from the RBC model for each value of σ, and retain the last 100 for inference. We compute and estimate the posterior kernel of σ assuming

- Inverse Gamma Prior with $m = 0.2$ and $SD = 5$,

- Normal Prior with $m = 0.2$ and $SD = 5$,

- Exponential Prior with $m = 5$ and $SD = 5$,

where m stands for the mean, and SD for the standard deviation. While the measures of dispersion are the same across priors, the prior shape and support are different. We first study the posterior kernel of σ conditional on the simulated data and on the other parameters being fixed at their true values. Being a unidimensional problem, we do not have to rely on posterior simulators and we can directly plot the product of the likelihood times the prior (i.e. the posterior kernel) against different values of σ. This allows us to study the behavior of the kernel in the neighborhood of zero. Figure 1 displays the posterior kernel of σ for a range of values of σ ($-0.5 : 0.02 : 0.5$) keeping the remaining parameters fixed at their true values. From the top left panel to the bottom right panel, we present the four cases for the values of the true standard deviation: 0, 0.05, 0.1, 0.2.

---

[7]See Appendix A.1 for details.

[8]Results with different parameterization of the data generating process, scale and location parameters and sample size are available on request.
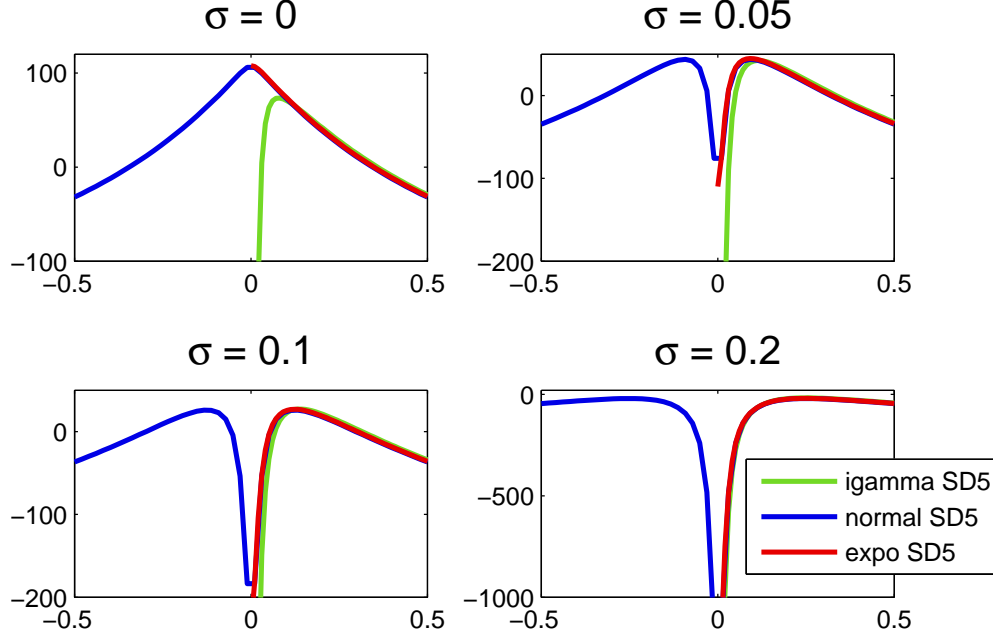
Figure 1: Posterior kernel of σ with inverse gamma IG(0.2,5) (green), normal N(0.2,5) (blue) and exponential Exp(5) (red) prior assuming different true values for σ

A couple of results are worth highlighting. First, when the technology shock has zero variance (i.e. the case of a non-fundamental structural shock), with a normal prior with loose precision the posterior kernel of σ is uni-modal centered on zero and with a tight standard deviation. Similar conclusions apply to the exponential prior, for which the posterior peaks at zero. Hence, the prior information on this parameter does not distort the information of the data likelihood. By assuming an inverse gamma prior, instead, we are forcing the kernel not to explore the region of the parameter space of a null variance and, as a consequence, we are corrupting the information contained in the data. Second, when the technology shock is fundamental the posterior kernel of σ is similar across prior assumptions. Therefore, normal or exponential priors do not seem to create distortions when the shock truly exists.

Conclusions are similar when using posterior simulators to approximate the posterior distributions as shown in figure 2. The posterior moments of the full set of parameters, α, ρ, σ and ω, are computed using the Random Walk Metropolis-Hastings algorithm adapted for the sign switch when assuming normal prior for σ. We postulate a normal prior for α centered in 0.3 with 0.05 standard deviation, a beta distribution for ρ centered in 0.6 with 0.2 standard deviation and an inverse gamma prior for the measurement error centered in 0.2 with a loose standard deviation of 4.

Table 4.1 reports posterior statistics assuming different prior distributions for the standard deviations. While the estimates of σ are very imprecise with an inverse gamma prior, the normal and exponential priors with a sufficiently loose standard deviation allow the MCMC to explore more extensively the parameter space and hence to verify ex-post if the
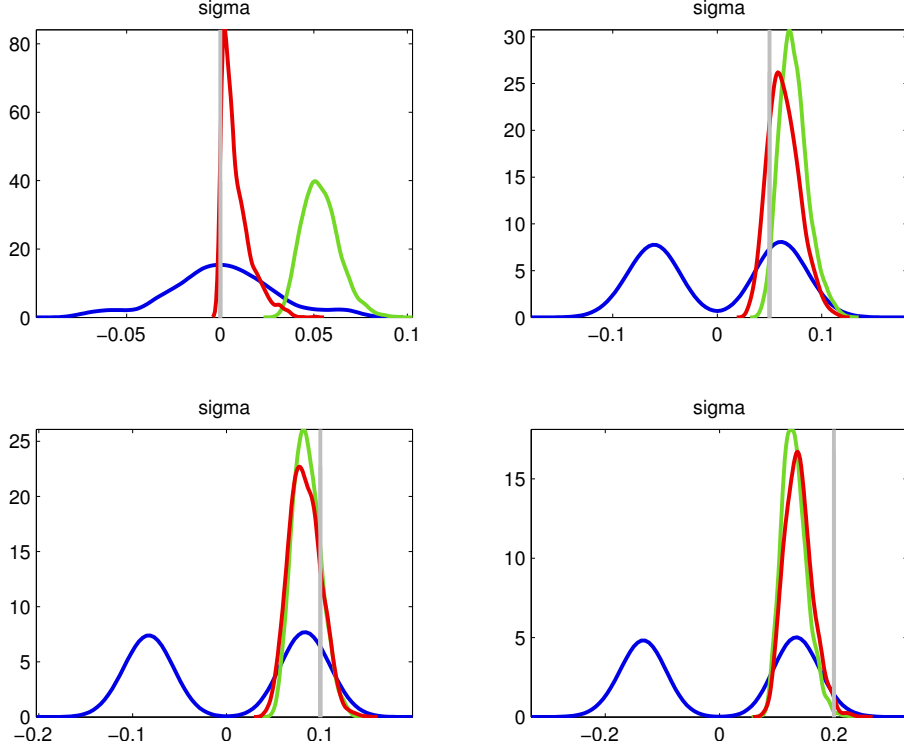
Figure 2: Posterior distribution of σ with inverse inverse gamma IG(0.2,5) (green), normal N(0.2,5) (blue) and exponential Exp(5) (red) prior assuming different true values for σ. RW Metropolis-Hastings simulator is used to approximate the posterior distribution.

structural disturbance exists. It is important to notice that, in this simple example, the structural parameters other than the autoregressive coefficient are unaffected. The extent to which deep structural parameters are influenced by the wrong combination of shocks is explored in the next subsection. However, before that, we wish to explore more in detail the implications of mistaken assumptions about shock existence for the persistence of the observable variable.

In the case of a null standard deviation, the posterior kernel displays a clear trade off between setting the standard deviation close to zero or reducing the persistence of the model dynamics. Since with inverse gamma priors we rule out null standard deviations, the posterior kernel of standard deviations does not include zero and, as a consequence, the structural shock has a dynamic impact on y_t. The only way to tune down the dynamic impact of this shock is to force the autoregressive parameters close to zero. To see this, assume that $\alpha = 0$ and $\omega = 1$. Then, the law of motion for output is given by

$$z_{t+1} = \rho z_t + \sigma \epsilon_{t+1} \quad \epsilon_t \sim N(0,1)$$
$$y_t = z_t + e_t \quad e_t \sim N(0,1)$$

Assume that the true DGP is the one with a null standard deviation, i.e. $\sigma = 0$, we have

θ	$\sigma = 0$ Median [Lower,Upper]	$\sigma = 0.05$ Median [Lower,Upper]	$\sigma = 0.1$ Median [Lower,Upper]	$\sigma = 0.2$ Median [Lower,Upper]
	Inverse Gamma Prior			
α	0.302 [0.218 , 0.382]	0.325 [0.245 , 0.407]	0.328 [0.247 , 0.409]	0.335 [0.256 , 0.416]
ρ	0.219 [0.076 , 0.447]	0.774 [0.584 , 0.905]	0.760 [0.580 , 0.889]	0.889 [0.794 , 0.958]
σ	0.053 [0.039 , 0.071]	0.072 [0.053 , 0.098]	0.085 [0.063 , 0.112]	0.129 [0.097 , 0.172]
ω	0.079 [0.064 , 0.093]	0.087 [0.070 , 0.105]	0.081 [0.061 , 0.101]	0.145 [0.118 , 0.176]
	Normal Prior			
α	0.322 [0.240 , 0.404]	0.327 [0.238 , 0.407]	0.332 [0.243 , 0.408]	0.329 [0.241 , 0.414]
ρ	0.715 [0.319 , 0.966]	0.841 [0.684 , 0.954]	0.735 [0.539 , 0.894]	0.902 [0.811 , 0.966]
σ	0.001 [-0.051 , 0.054]	0.060 [0.040 , 0.088]	0.083 [0.059 , 0.110]	0.134 [0.100 , 0.176]
ω	0.085 [0.061 , 0.101]	0.093 [0.076 , 0.112]	0.082 [0.063 , 0.106]	0.145 [0.117 , 0.178]
	Exponential Prior			
α	0.330 [0.245 , 0.408]	0.330 [0.248 , 0.412]	0.331 [0.247 , 0.412]	0.332 [0.249 , 0.412]
ρ	0.600 [0.238 , 0.913]	0.815 [0.637 , 0.928]	0.765 [0.579 , 0.900]	0.882 [0.780 , 0.955]
σ	0.006 [0.001 , 0.025]	0.062 [0.041 , 0.091]	0.083 [0.058 , 0.111]	0.136 [0.101 , 0.181]
ω	0.091 [0.081 , 0.102]	0.091 [0.074 , 0.110]	0.082 [0.062 , 0.102]	0.143 [0.114 , 0.174]

Table 1: Estimated parameters with normal and inverse gamma priors on standard deviations of structural shocks

that $y_t = e_t$ and the likelihood collapses to

$$\log L(y_T|y^{T-1}; \rho, \sigma = 0) \propto -1/2 \sum_{t=1}^{T} y_t^2$$

While ρ is not identifiable in the true model, the persistence parameter becomes informative in the misspecified model. Suppose we work with a misspecified model in which σ is fixed to a positive value, say $\delta > 0$, which measures the degree of misspecification, i.e. the larger this value the more severe is the misspecification. The likelihood is given by

$$\log L(y_T|y^{T-1}; \rho, \sigma = \delta) \propto -1/2 \sum_{t=1}^{T} \ln(s_t + 1) - 1/2 \sum_{t=1}^{T} \frac{(y_t - z_{t|t-1})^2}{1 + s_t}$$

$$z_{t+1|t} = (1 - k_t)z_{t|t-1} + k_t y_t$$

$$s_{t+1} = k_t + \delta^2$$

$$k_t = \rho \frac{s_t}{1 + s_t}$$

where the recursions are derived from the Kalman filter with $s_1 = \delta^2/(1 - \rho^2)$ and $z_{1|0} = 0$. In order to minimize the information discrepancy between the misspecified model likelihood, i.e. $L(y_T|y^{T-1}; \rho, \sigma = \delta)$, and the true DGP model likelihood, i.e. $L(y_T|y^{T-1}; \rho, \sigma = 0)$, the autoregressive parameter has to go to zero. When $\rho = 0$, we have $k_t = 0$, $z_{t|t-1} = 0$, $s_t = \delta^2$, and the likelihood becomes

$$\log L(y_T|y^{T-1}; \rho = 0, \sigma = \delta) \propto -T/2 \ln(\delta^2 + 1) - \sum_{t=1}^{T} 1/2 \frac{y_t^2}{1 + \delta^2}$$

The fact that the information discrepancy is minimized when $\rho = 0$ can be shown numerically. Figure 3 reports the information discrepancy between the misspecified and the true model likelihood, i.e. $\log L(y_T|y^{T-1};\rho,\sigma=\delta) - \log L(y_T|y^{T-1};\sigma=0)$, for various values of ρ. The closer this value is to zero, the lower is the discrepancy between the misspecified and the true model likelihood. One can see that the larger the persistence of the shock, the larger is the information discrepancy with the true model. Moreover, the larger the degree of misspecification, δ, the steeper the information discrepancy as a function of the persistence becomes.
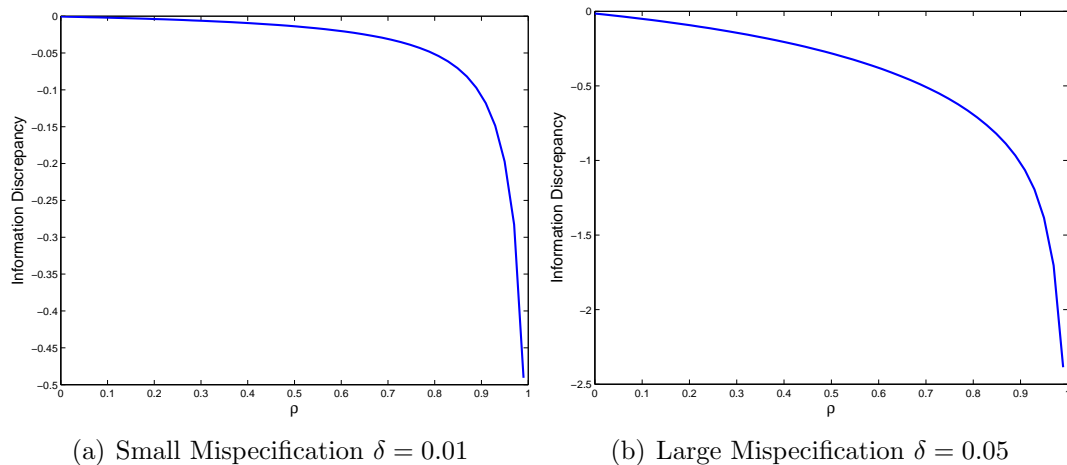

(a) Small Mispecification $\delta = 0.01$ (b) Large Mispecification $\delta = 0.05$

Figure 3: Information discrepancy, $\log L(y_T|y^{T-1};\rho,\sigma=\delta) - \log L(y_T|y^{T-1};\sigma=0)$ for different values of ρ. Left panel $\delta = 0.01$, right panel $\delta = 0.05$.

Hence, postulating the existence of a non-fundamental shock has two consequences. First, it makes the autoregressive parameter informative when it is not in the true DGP. Second, it generates a downward bias in the estimate of the internal persistence. In fully fledged DSGE models, the persistence of model dynamics is controlled not only by the autoregressive parameters, but also by the deep parameters capturing real and nominal frictions in the economy. Therefore, it is likely that those parameters will be affected as well by the incorrect specification of the number and combination of structural shocks. The extent to which deep structural parameters are influenced by the wrong combination of shocks is explored quantitatively in the next subsection.

4.2   Example and Experiments with a medium-scale DSGE model

We now apply the same analysis to the baseline version of the Smets and Wouters (2007) model (henceforth SW). This model is selected because of its widespread use for policy analysis among academics and policymakers, and because it is frequently adopted to study cyclical dynamics and their sources of fluctuations in developed economies. We retain the nominal and real frictions originally present in the model, but we make a number of simplifications, which reduce the computational burden of the experiment, but bear no consequences on the

conclusions we reach. First, we assume that all exogenous processes are stationary. Since we are working with the decision rules of the model, such a simplification involves no loss of generality. Second, we assume that all the shocks are uncorrelated and follow autoregressive processes of order one. Third, since we do not want to have our results driven by identification issues (see Komunjer and Ng (2011) or Iskrev (2010)), we fix a number of parameters and estimate only a subset of them. We then estimate the standard errors, autoregressive parameters, and the parameters driving price and wage indexation and stickiness, habit in consumption, intertemporal elasticity of substitution and the inverse of the elasticity of investment (relative to an increase in the price of installed capital). Table 2 reports the full set of parameter estimates, prior assumptions, and the true values used to generate artificial data, which are taken from the posterior mean estimated in SW.[9]

To study the effect of estimating the model including non-fundamental shocks, we switch off the price markup, the wage markup, and the investment specific shocks and add seven measurement i.i.d. errors (one on each observable) with standard deviation equal 0.08.[10] With this calibration, measurement errors explain on average less then 3% of the volatility of observables. We simulate 1,000 data points and use the last 200 for inference.

We consider seven observable variables: output y_t, consumption c_t, investment i_t, wages w_t, inflation π_t, interest rates r_t, and hours worked h_t. We estimate the model assuming inverse gamma, exponential, and normal priors for the standard deviations and the same priors as in SW for the remaining parameters. We run a 300,000 draws MCMC routine starting from the posterior kernel mode and burn-in the first 200,000 of the chain and keep randomly 1000 for inference. A subset of posterior moments are collected in table 2.

The posterior estimates of structural parameters with inverse gamma priors on standard deviations are inferior to the normal and exponential priors along a number of dimensions. First, and most obvious, with inverse gamma priors we are unable to separate the structural shocks that exist from those that do not. On the contrary, with normal priors the posterior support of the standard deviations of the investment specific shock, and of the wage and price markup shocks include zero. With exponential priors we obtain that, for the non-existing shocks, 90% of the mass of the posterior probability is located between 0 and 0.01. This suggests that, with a sample size comparable to that used in empirical applications, we are able to identify zero and non-zero standard deviations as long as we do not use inverse gamma priors.

Second, the autoregressive parameters estimates of the specification with inverse gamma priors are typically downward biased. As suggested in the previous section, the likelihood has to compensate for the incorrect assumptions about the existence of structural processes by reducing their persistence. In particular, the estimates of ρ_i and ρ_p are heavily downward biased with a relatively tight posterior standard deviation.

---

[9]More details on the model can be found in Appendix A.5, where we report the log-linearized equilibrium conditions.

[10] We also reduced this value to 0.05 and increased it to 0.35 and the main conclusions are unaffected.

Parameters	Normal Prior Median [Lower,Upper]	Igamma Prior Median [Lower,Upper]	Exp Prior Median [Lower,Upper]	true
σ_a	0.513 [0.481 , 0.553]	0.464 [0.420 , 0.510]	0.515 [0.468 , 0.553]	0.450
σ_b	0.250 [0.228 , 0.269]	0.260 [0.233 , 0.290]	0.245 [0.223 , 0.266]	0.240
σ_g	0.519 [0.487 , 0.555]	0.510 [0.466 , 0.554]	0.525 [0.484 , 0.566]	0.520
σ_i	-0.000 [-0.005 , 0.005]	0.068 [0.059 , 0.080]	0.004 [0.000 , 0.010]	0.000
σ_m	0.224 [0.202 , 0.254]	0.246 [0.216 , 0.281]	0.248 [0.215 , 0.269]	0.250
σ_w	-0.000 [-0.001 , 0.001]	0.056 [0.049 , 0.065]	0.000 [0.000 , 0.001]	0.000
σ_p	0.000 [-0.004 , 0.003]	0.060 [0.054 , 0.067]	0.004 [0.000 , 0.008]	0.000
ρ_a	0.956 [0.952 , 0.959]	0.952 [0.940 , 0.960]	0.957 [0.955 , 0.960]	0.958
ρ_b	0.250 [0.213 , 0.282]	0.213 [0.155 , 0.277]	0.271 [0.211 , 0.295]	0.218
ρ_g	0.974 [0.971 , 0.976]	0.973 [0.969 , 0.977]	0.975 [0.974 , 0.977]	0.976
ρ_i	0.786 [0.764 , 0.814]	0.310 [0.285 , 0.335]	0.629 [0.550 , 0.673]	0.710
ρ_r	0.225 [0.189 , 0.252]	0.126 [0.041 , 0.259]	0.208 [0.068 , 0.266]	0.151
ρ_p	0.702 [0.676 , 0.722]	0.069 [0.023 , 0.141]	0.523 [0.491 , 0.556]	0.891
ρ_w	0.839 [0.807 , 0.874]	0.707 [0.679 , 0.736]	0.886 [0.829 , 0.914]	0.968
ϕ	5.688 [5.643 , 5.719]	5.136 [4.918 , 5.261]	5.589 [5.538 , 5.622]	5.744
λ	0.722 [0.708 , 0.734]	0.661 [0.638 , 0.679]	0.704 [0.693 , 0.712]	0.714
ζ_w	0.699 [0.691 , 0.707]	0.595 [0.572 , 0.617]	0.701 [0.695 , 0.706]	0.701
i_w	0.580 [0.558 , 0.602]	0.870 [0.747 , 0.963]	0.576 [0.531 , 0.638]	0.589
i_p	0.174 [0.126 , 0.197]	0.165 [0.075 , 0.240]	0.207 [0.164 , 0.264]	0.240
ζ_p	0.638 [0.627 , 0.648]	0.620 [0.599 , 0.638]	0.642 [0.632 , 0.652]	0.650
r_p	1.929 [1.887 , 1.973]	1.558 [1.448 , 1.683]	1.856 [1.781 , 2.007]	2.045
r_{dy}	0.207 [0.185 , 0.221]	0.194 [0.169 , 0.223]	0.225 [0.210 , 0.242]	0.224
r_y	0.090 [0.084 , 0.101]	0.109 [0.086 , 0.131]	0.136 [0.115 , 0.156]	0.087
ρ	0.799 [0.795 , 0.805]	0.772 [0.746 , 0.801]	0.816 [0.801 , 0.825]	0.808
σ_c	1.329 [1.324 , 1.333]	1.540 [1.530 , 1.549]	1.400 [1.396 , 1.402]	1.380

Table 2: Estimated parameters with various priors on standard deviations of structural shocks

Third, the posterior distributions of deep structural parameters, such as habit in consumption, and price and wage stickiness, appear to be largely influenced by the prior assumptions about the distribution (and hence existence) of structural shocks standard deviations. In particular, the wage stickiness (ζ_w), consumption habit (λ), and price stickiness (ζ_p) parameters tend to be underestimated relative to their true values when using inverse gamma priors. These parameters control the degree of internal persistence of the model. Similar to the previous example, the posterior kernel displays a clear tradeoff between setting the standard deviation close to zero and reducing the persistence of the model dynamics. Since with inverse gamma priors we rule out null values, the posterior kernel of standard deviations does not include zero and, as a consequence, all structural shocks have a dynamic impact on endogenous variables. In order to choke-off the dynamic impact of this shock, the autoregressive parameters are estimated to be close to zero. However, since the dynamic transmission of the shocks is controlled also by deep parameters, the posterior kernel is tilting toward reducing the overall internal persistence of the model, hence inducing a downward bias in the parameters governing wage and price stickiness and indexation. Assuming normal or exponential priors, parameters are estimated without noticeable distortions.

In order to check that the results are not driven by a particular sample, we also designed a Montecarlo experiment where we simulated various samples and estimated the posterior distributions of the parameters using our three different prior distributions, i.e. normal, exponential and inverse gamma. We then considered the bias, measured as the difference between the average posterior mean of different samples and the true value, and we report the results for deep structural parameters in table 3. A positive value means that we are *overestimating* a parameter, and a negative value that we are *underestimating*.

Parameters	Normal	IGamma	Exp	True value
ϕ	-0.002	-1.317	0.008	5.744
λ	0.001	-0.064	0.011	0.714
ζ_w	-0.001	-0.067	0.009	0.701
i_w	0.005	-0.188	0.018	0.589
i_p	0.008	-0.167	-0.008	0.240
ζ_p	0.002	-0.106	0.000	0.650
r_p	-0.058	-0.185	0.008	2.045
r_{dy}	-0.004	-0.044	0.012	0.224
r_y	0.007	0.019	0.015	0.087
ρ	-0.003	-0.043	0.013	0.808
σ_c	-0.007	0.101	0.010	1.380

Table 3: Bias on Montecarlo exercise with 100 different datasets. Bias is measured as the difference between the average posterior mean of different samples and the true value.

Bar a few exceptions, the bias obtained using normal or exponential priors is negligible as the order of magnitude is small. In all cases, the bias using inverse gamma priors is larger than using normal or exponential priors. With inverse gamma priors, we obtain sizable distortions to parameters capturing persistence and others such as the inverse elasticity of investment relative to the price of installed capital.

4.2.1 DSGE model implications: IRFs and Variance Decompositions

Incorrect assumptions about the existence of structural shocks do not only distort parameter estimates, but they have deep consequences for the implications of the model regarding the sources of business cycle fluctuations or the dynamic transmission of structural shocks which are important for policy analysis. Table 4 reports the variance decomposition of output, inflation, wages, and the interest rate in terms of structural shocks under various prior assumptions about the structural standard deviations. Price and wage markup shocks should not explain fluctuations in any of these variables. This is the case for normal and exponential priors. It is not the case, however, for inverse gamma priors where wage and price markup together explain 16 % of the volatility of inflation and 8% of the volatility of wages.

	Variance Decomposition			
	y	π	r	w
Igamma prior on σ_i	0 [0 , 0]	0 [0 , 0]	0[0 , 0]	0[0 , 0]
Igamma prior on σ_w	1 [0 , 1]	9 [6 , 14]	2 [1 , 3]	8[5 , 13]
Igamma prior on σ_p	0 [0 , 0]	7 [5 , 8]	0 [0 , 1]	0[0 , 0]
Normal prior on σ_i	0 [0 , 0]	0 [0 , 0]	0 [0 , 0]	0[0 , 0]
Normal prior on σ_w	0 [0 , 0]	0 [0 , 0]	0 [0 , 0]	0[0 , 0]
Normal prior on σ_p	0 [0 , 0]	0 [0 , 0]	0 [0 , 0]	0[0 , 0]
Exp prior on σ_i	0 [0 , 0]	0 [0 , 0]	0 [0 , 0]	0[0 , 0]
Exp prior on σ_w	0 [0 , 0]	0 [0 , 0]	0 [0 , 0]	0[0 , 0]
Exp prior on σ_p	0 [0 , 0]	0 [0 , 1]	0 [0 , 0]	0[0 , 0]

Table 4: Fraction of the variance of output, inflation, interest rate and wages explained by the non-existing shocks, price (σ_p) and wage (σ_w) markups and the investment specific shock (σ_i). In the true DGP the explained variance should be zero.

Moreover, the transmission of shocks is altered in a substantial way. Figure 4 reports the transmission of monetary policy shocks to output, inflation, and interest rate (top row) and the transmission of a wage markup shock (bottom row). Gray areas (green dashed lines) represent the 90% confidence sets of the response assuming normal (inverse gamma) priors on structural standard deviations and the black solid line the true response.

The responses to a monetary policy shock are qualitatively different under the two settings. Inverse gamma priors tend to produce less persistent dynamic responses. Moreover, on impact, we *overestimate* the reaction of inflation to an interest rate hike and *underestimate* the reaction of output. In this context, disinflation trajectories might result to be less costly in terms of output loss relative to what they truly are.

Even more striking, as would be expected, are the responses to "non-existing" shocks. In the normal prior setup, we obtain statistically insignificant dynamics for all the variables of interest to an increase in the wage markup. Conversely, with inverse gamma priors, output and inflation react strongly and their responses are statistically and economically significant.

In all, the simulation evidence shows that inference and policy conclusions differ substantially when non-existing structural shocks are forced to exist in estimation. Since we do not know ex-ante what are key shocks driving aggregate fluctuations, inverse gamma priors are problematic as they may induce biases in estimated parameters that can be sizeable. It is in this sense that there appears to be a tradeoff between the a-priori inclusion of a large set of sources of macroeconomic uncertainty and the correct identification of the parameters that drive transmission. We have also shown that normal or exponential priors do not suffer from any particular disadvantage when confronted to data that are generated by an
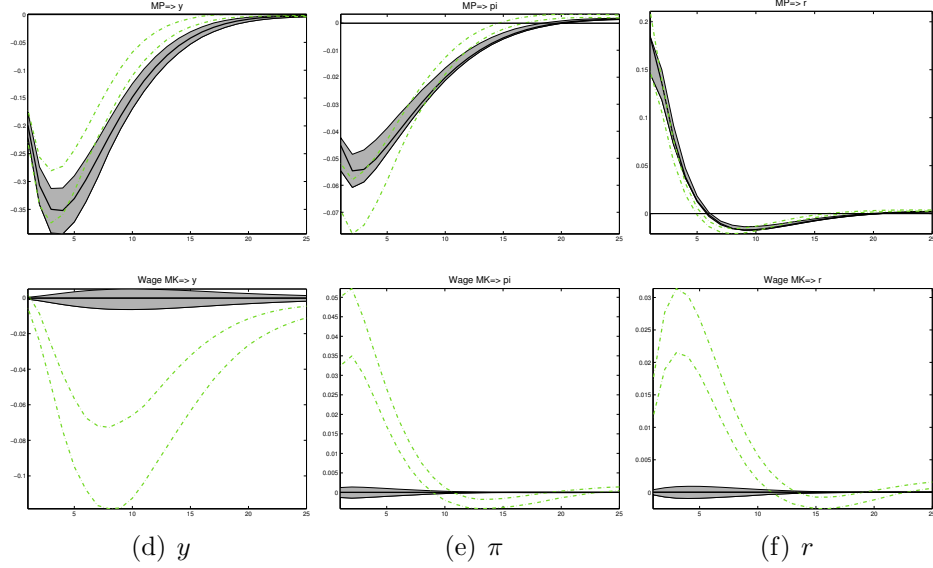
Figure 4: Impulse Response Function of output, inflation and interest rate to a monetary policy shock (top row line) and to a wage markup shock (bottom row) with inverse gamma IG(0.2,5) (green), normal N(0.2,5) (gray shaded area) prior.

unknown number of structural disturbances. If the shock does not exist, the posterior standard deviation distribution is clustered around zero and structural parameters are unbiased. If, instead, the shock exists, the posterior distribution is centered on the true value and the dynamic transmission of shocks is unaffected.

A crucial assumption of our approach is that the vector of observed times series is generated by a combination of structural and non-structural (measurement) shocks. In the absence of measurement errors, the DSGE model with a rank deficient covariance matrix is stochastically singular and, as a consequence, impossible to estimate with likelihood based approaches. The inclusion of measurement error allows us to avoid the stochastic singularity problem. One may argue that measurement error sweeps the rest of the variability of observables that is not explained by fundamental shocks. However, since structural shocks are common factors and measurement errors are variable-specific shocks, when measurement errors capture a larger proportion of the variability of a particular observable, it is precisely indicating that some fundamental shocks may not be true common factors. We will revisit this argument below in the empirical application.

5   EMPIRICAL APPLICATION: FUNDAMENTAL SHOCKS IN THE SW MODEL

We now revisit the empirical evidence on structural DSGE shocks by reconsidering a standard DSGE model using normal and inverse gamma priors using US macroeconomic data. We keep as the benchmark the SW model used in the previous section because it represents a widely used medium scale New Keynesian DSGE model specification. Although there are many more applications of interest, here we focus on the question of whether some of the

standard impulses assumed in the existing literature are truly fundamental and what the consequences of shock selection are for parameter estimates and the transmission of shocks.

While the structural equations of the model are the same as the ones presented in the previous section, we add deterministic growth and measurement equations in order to bridge the model to the observed times series. The SW model is estimated based on seven quarterly macroeconomic time series. The measurement equations for real output, consumption, investment, and real wage growth, hours, inflation, and interest rates are given by:

$$output\ growth = \bar{\gamma} + \Delta y_t + \omega_y e_{y,t}$$
$$consumption\ growth = \bar{\gamma} + \Delta c_t + \omega_c e_{c,t}$$
$$investment\ growth = \bar{\gamma} + \Delta i_t + \omega_i e_{i,t}$$
$$real\ wage\ growth = \bar{\gamma} + \Delta w_t + \omega_w e_{w,t}$$
$$hours = \bar{l} + l_t + \omega_l e_{l,t}$$
$$inflation = \bar{\pi} + \pi_t + \omega_p e_{p,t}$$
$$ffr = \bar{\beta} + R_t + \omega_r e_{r,t}$$
$$e_{x,t} \sim N(0,1)\ \ with\ \ x = y, c, i, w, l, p, r$$

where all variables are measured in percent, $\bar{\pi}$ and $\bar{\beta}$ measure the steady state level of net inflation and short term nominal interest rates, respectively, $\bar{\gamma}$ captures the deterministic long-run growth rate of real variables, and $\bar{l}$ captures the mean of hours. Output growth is measured as the percentage growth rate of Real GDP, consumption growth as the percentage growth rate of personal consumption expenditure deflated by the GDP deflator, and investment growth as the percentage growth rate of Fixed Private Domestic Investment deflated by the GDP deflator. Hourly compensation is divided by the GDP price deflator in order to get the real wage variable. The aggregate real variables are expressed per capita by normalizing by population over 16. Inflation is the first difference of the log of the Implicit Price Deflator of GDP, and the interest rate is the Federal Funds Rate divided by four. For comparability of estimates, we consider the same data span as in SW, 1968-2004, with revised data. However, we also run estimates with the vintage data and with samples including more recent years (i.e. up to 2014). The results differ only marginally and we mention any difference in the text below.

We estimate and fix the same parameters as in SW with one exception. Relative to the original SW specification, we assume that the impact of technology on government spending, ρ_{ga} in their model, is zero so that the government spending process is independent from the technology process. The first specification coincides with the original SW setup, which we call *SW* IGamma. In this specification, we assume that measurement error shocks are zero, i.e. $\omega_x = 0$ for $x = y, c, i, w, l, p, r$, and structural shock standard deviations have an inverse gamma prior. In the second specification, *SW* Normal, we postulate that the structural shock standard deviations are normally distributed with mean zero and standard deviation

20

0.4, and that the measurement error shock standard deviations have an inverse gamma prior with mean 0.1 and standard deviation 2. Given that the results using exponential and normal priors in the previous section were very similar, we report here the results using normal priors only.[11]

5.1 ESTIMATION RESULTS

Table 5 reports the posterior moments of a subset of parameters for specifications *SW* IGamma and *SW* Normal together with prior assumptions.

| Parameters | Posterior Statistics | | Prior Distribution |
| | *SW* IGamma | *SW* Normal | |
	Median [Lower,Upper]	Median [Lower,Upper]	
ι_p	0.30 [0.14 , 0.48]	0.49 [0.25 , 0.72]	B(0.5,0.15)
ϕ	6.09 [4.39 , 7.84]	2.84 [2.00 , 4.21]	N(4,1.5)
ξ_w	0.69 [0.57 , 0.80]	0.79 [0.70 , 0.87]	B(0.5,0.1)
r_y	0.08 [0.05 , 0.12]	0.15 [0.10 , 0.20]	N(0.12,0.05)
ρ_a	0.95 [0.92 , 0.97]	0.95 [0.91 , 0.99]	B(0.5,0.2)
ρ_b	0.23 [0.08 , 0.38]	0.70 [0.51 , 0.85]	B(0.5,0.2)
ρ_g	0.97 [0.96 , 0.99]	0.46 [0.13 , 0.80]	B(0.5,0.2)
ρ_i	0.72 [0.62 , 0.82]	0.74 [0.57 , 0.93]	B(0.5,0.2)
ρ_r	0.16 [0.05 , 0.27]	0.46 [0.18 , 0.85]	B(0.5,0.2)
ρ_p	0.87 [0.80 , 0.95]	0.39 [0.11 , 0.68]	B(0.5,0.2)
ρ_w	0.96 [0.93 , 0.98]	0.98 [0.95 , 1.00]	B(0.5,0.2)
μ_p	0.69 [0.52 , 0.84]	0.58 [0.27 , 0.87]	B(0.5,0.2)
μ_w	0.82 [0.71 , 0.93]	0.61 [0.28 , 0.89]	B(0.5,0.2)
σ_a	0.44 [0.40 , 0.49]	0.34 [0.29 , 0.40]	N(0,0.4)/IG(0.1,2)
σ_b	0.24 [0.20 , 0.28]	0.11 [0.07 , 0.15]	N(0,0.4)/IG(0.1,2)
σ_g	0.57 [0.52 , 0.63]	0.10 [-0.21 , 0.26]	N(0,0.4)/IG(0.1,2)
σ_i	0.41 [0.34 , 0.49]	0.24 [0.09 , 0.40]	N(0,0.4)/IG(0.1,2)
σ_r	0.24 [0.22 , 0.27]	0.15 [0.09 , 0.20]	N(0,0.4)/IG(0.1,2)
σ_p	0.13 [0.10 , 0.15]	-0.06 [-0.14 , 0.13]	N(0,0.4)/IG(0.1,2)
σ_w	0.26 [0.22 , 0.30]	0.03 [0.01 , 0.08]	N(0,0.4)/IG(0.1,2)

Table 5: Subset of estimated parameters with normal and inverse gamma priors on standard deviations of structural shocks. Sample span 1968q1 2003q4

---

[11]While in the previous section the estimates were performed using our own codes, the estimates in this section are performed using the Dynare platform, see Adjemian, Bastani, Karamé, Juillard, Maih, Mihoubi, Perendia, Ratto and Villemot (2011), as the original version of the SW model was estimated in that environment. Since Dynare does not allow for exponential priors, it is more natural to report the results with normal priors for comparability.

There are a number of very relevant results to highlight. First, government shocks and price markup shocks standard deviations are estimated to be 'non-existent', since the posterior support for their standard deviations includes zero. Hence, these two shocks are not fundamental for the observation set used by SW. Technology, investment, preference, wage markup, and monetary policy shocks are instead estimated to be fundamental. The wage markup shock, however, is only marginally so. Moreover, the estimated standard deviations of the fundamental shocks are of the same magnitude in the two specifications, bar for the wage markup shock which is smaller. The wage markup shock actually turns out not to be significant in samples including more recent years. Interestingly, such clustering echoes the classification of structural and non structural shocks proposed by Chari et al. (2008).

Second, except for the parameters of the non fundamental shocks, which are not identifiable, the autoregressive parameters are estimated to be smaller in the *SW* IGamma specification relative to the *SW* Normal specification. This confirms the point made earlier about the tendency for model estimates to reduce internal persistence when assumptions about the fundamentalness of the structural shocks are incorrect.

Third, the posterior distributions of deep parameters appear to be estimated differently in the two setups. For the estimates of ϕ, the second derivative of the investment adjustment cost function, our median estimate is significantly smaller than the 5.58 value found in SW and closer to the value (2.48) available in Christiano, Eichenbaum and Evans (2005).[12] Christiano et al. (2005) estimate a model similar to SW with staggered wage and price contracts, habit formation in preferences for consumption, adjustment costs in investment, and variable capital utilization. The two papers differ in terms of estimation techniques. While SW use full information methods, Christiano et al. (2005) use limited information methods, i.e. by minimizing a measure of the distance between the model and empirical impulse response functions to a monetary policy shock. In a sense, they do not need to impose the existence of (possibly) non-fundamental shocks. Moreover, the SW mode for the price indexation parameter, ι_p, is 0.22 and our estimated parameter is more than twice that value in accordance with the results obtained in the previous sections. Similar conclusions apply for the estimate of ξ_w, the wage stickiness parameter. Overall, estimates of deep parameters change when we allow for the possibility of a rank deficient matrix for the structural shocks.

Fourth, as a consequence of the different structural parameters estimates, the dynamic transmission of structural shocks changes between the two specifications. The effects of policy shocks such as fiscal and monetary policy are different. Figure 5 reports the transmission mechanism of a fiscal (top part) and a monetary (bottom) policy shock for consumption and output growth, inflation, and the interest rate. In the original SW model, fiscal shocks generate an increase in output (not shown here), in inflation and, through the Taylor rule, an interest rate hike. Consumption decreases for the Ricardian motives of the representative

---

[12]Quoting Christiano et al. (2005), "$[1/\phi]$ is the elasticity of investment with respect to a 1 percent temporary increase in the current price of installed capital. Our point estimate implies that this elasticity is equal to 0.40", which roughly coincides with our point estimate.

agent model. While responses are significant in the original SW model, the same model estimated with normal priors on structural standard deviations cannot generate statistically significant dynamics. Similarly, monetary policy shocks seem to generate different dynamics. In the original SW specification, to generate a similar dis-inflationary pattern (panel (e)) we need a much larger hike in the interest rates of 15-20 basis points (panel (f)) as opposed to 0-10 basis points in our specification with normal priors.

(a) govt → cons   (b) govt → infl   (c) govt → int rate

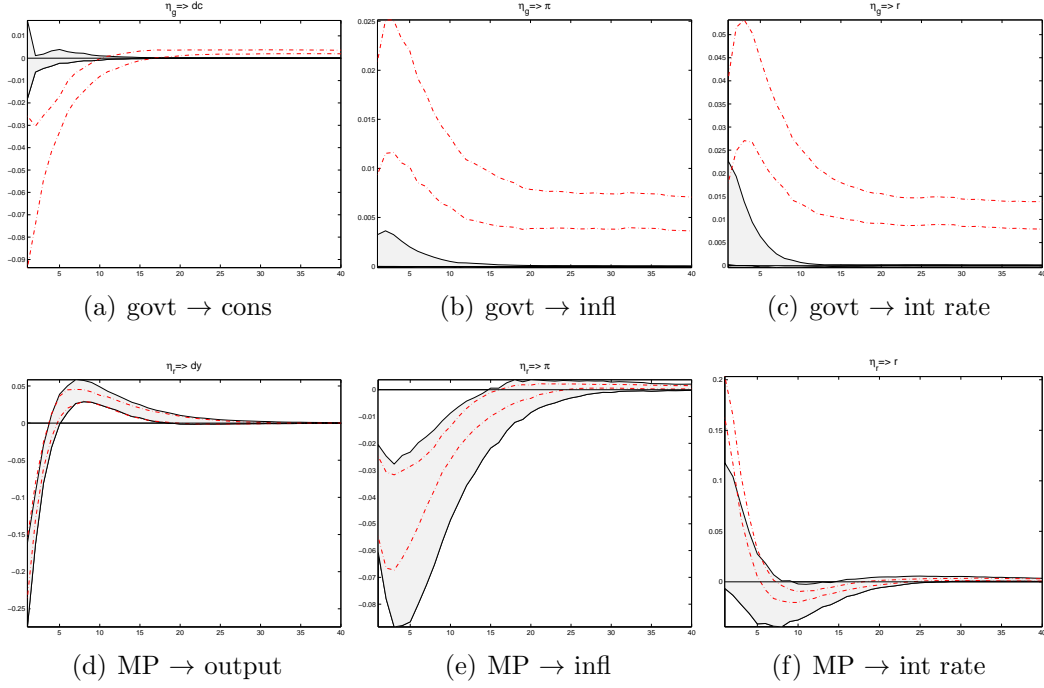(d) MP → output   (e) MP → infl   (f) MP → int rate

Figure 5: Posterior IRF to selected shocks. Gray area the band with normal priors, green dashed lines with inverse gamma. Top panel, responses to a fiscal shock and bottom to a monetary policy shock.

ML	Priors standard deviations structural/measurement		
	SW IG/no meas	SW N/IG	SW N/N
Geweke	-904	-854	-856
Laplace	-904	-854	-857

Table 6: Marginal Log likelihood of different priors specifications on standard deviation of structural and measurement shocks

We might wonder how likely these two polar cases are, i.e. of a full rank covariance matrix of structural shocks (SW IGamma) and of a possibly rank deficient covariance matrix of structural shocks (SW Normal). As a robustness check of these two cases, we estimated and contrasted the marginal likelihood of three different models: SW with inverse gamma priors on structural shocks and no measurement errors (SW IG/no meas), SW with normal priors

for structural shocks and inverse gamma priors for measurement errors (*SW* N/IG), and *SW* with normal priors for both structural shocks and measurement errors (*SW* N/N). To approximate the data marginal likelihood we used both the Laplace approximation around the posterior mode and the Geweke (1999) estimator and results are reported in table 6. The *SW* specifications with normal priors on structural shocks standard deviations are strictly preferred. Between *SW* N/IG and *SW* N/N the differences are very small but the marginal likelihood favors *SW* N/IG. The results, thus, support the idea that the shock structure specified in the SW model is not fundamental.

(a) tech   (b) pref   (c) inv

(d) mp   (e) gov   (f) price markup
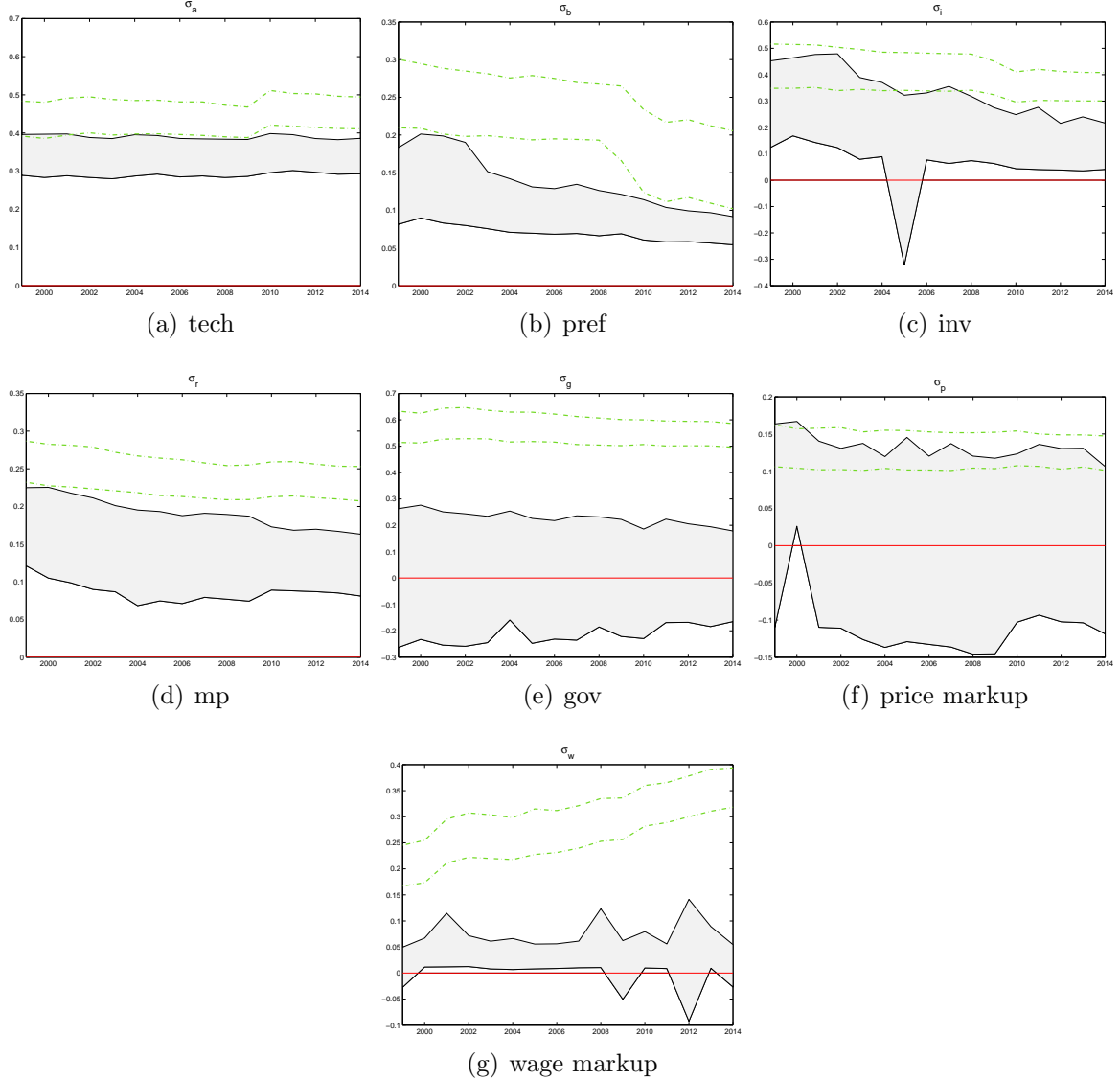
(g) wage markup

Figure 6: Rolling estimates of posterior bands around structural shocks. The gray area the band with normal priors, and the green dashed lines with inverse gamma.

We also extended the SW sample up to recent years in order to verify whether our results were driven by the specific sample in SW. We re-estimated the *SW* IGamma and *SW* Normal models adding one year of data from 2000 up to 2014. Figure 6 reports the estimated bands around the structural standard deviation over rolling subsamples. The results, in terms of

the fundamentalness of structural shocks, also hold if we include more recent data points for estimation. With the exception of wage markup shocks, most of the posterior bands appear to be relatively stable. However, for larger samples, the wage markup shock appears to be non-fundamental. In this case, it is also interesting to observe that, with inverse gamma priors, wage markup shocks appear to be more important if we add recent years. However, using normal priors, we can see that these shocks do not appear to be structural with the addition of more recent data. This finding is consistent with Justiniano et al. (2013). They find that, when using two alternative measures of wages to match the model's wage variable, most of the variability of wages can be explained by measurement error rather than implausibly large fluctuations in the monopoly power of workers.

Finally, we explored further the role of measurement errors in capturing the variability of observables when structural shocks' standard deviations are estimated with normal priors. We compared the variance decomposition of observables in the original SW model and in models where we allow for measurement error and normal priors for structural standard deviations. We also estimated models where we restricted the MCMC algorithm to search the posterior distribution of the standard deviation of the measurement errors in a narrower space, thus forcing the proportion of the variance explained by measurement error to be smaller. The results, available on request, confirm that when forcing on the model potentially non-fundamental shocks, the variance decomposition is heavily distorted. For instance, in the original SW model, 67% of the variance of wage growth is explained by wage markup shocks. When using normal priors, almost 90% of this variability is attributed to the wage measurement error. As we narrow the posterior space for the measurement error, the wage markup shock explains an increasing proportion of the variability of wages. Qualitatively similar conclusions are reached for inflation. For output and hours worked, the fiscal policy shock explains 36% and 14% respectively in the original SW model. In the model with normal priors it explains a negligible proportion. Again, as we narrow the search space for the posterior, the fiscal policy shock resurfaces as an important driver of output and hours fluctuations. In all, as we allow priors for structural (common) shocks to include the zero region, price and wage markup shocks and government spending shocks become insignificant and the variability of observables is pushed towards (variable-specific) measurement errors.[13]

6 Conclusions

One of the key questions in macroeconomics concerns the identification of the main impulses that set off macroeconomic fluctuations, the other key question being the identification of the propagation mechanisms that transform shocks into business cycles. Estimated DSGE models have become the standard methodology to address this question as they provide a coherent and economically interpretable structure. However, the widespread assumption when estimating DSGE models with likelihood methods is that certain exogenous shocks

---

[13]See equation (5) in section 2.

do exist in the sense that they capture macroeconomic uncertainty. In a Bayesian context, this is reflected in the standard practice of using priors for the standard deviation of shocks (typically inverse gamma) whose support does not include zero thus imposing the existence of these shocks and their interpretation as structural or fundamental. Some of these shocks, however, have dubious structural interpretation. We first analyze the consequences of imposing non-fundamental shocks for the estimation of DSGE model parameters and then propose a method that allows us to select the truly fundamental shocks driving macroeconomic uncertainty. The method requires specifying priors that include zero and studying the likelihood of the observable variables in the neighborhood of a null standard deviation.

We show that incorrect assumptions about the rank of the covariance matrix of shocks, Σ, have a non-trivial impact on the remaining estimated parameters and might severely distort structural inference. In particular, postulating the existence of a non-existing exogenous processes generates a substantial downward bias in the estimates of the parameters driving internal persistence of the model. We show this bias analytically for the simplest RBC model with exogenous labor supply and full capital depreciation. Using simulation evidence with a medium scale DSGE model, we also show important biases in the estimated persistence of shocks, and parameters such as wage and price stickiness and indexation. Thus, we unveil a tradeoff between the inclusion of a potentially large number of structural innovations and estimates of the parameters driving propagation.

To prevent this problem, we propose an easily implementable strategy of using normal or exponential priors together with measurement error to avoid stochastic singularity. We also propose a method for cases where the covariance matrix of shocks is not diagonal. Our simulation evidence shows that these priors allow us to select the true fundamental shocks entering the DSGE model and that the remaining parameters are estimated with precision.

We then revisited the evidence on the fundamentalness of structural shocks in the medium-scale New Keynesian model of Smets and Wouters (2007). Our key findings are threefold. First, government spending and price markup shocks are non-fundamental for the 1968-2004 sample and larger samples. The wage markup shock is not fundamental for larger estimation spans, i.e. 1968-2009 or onwards. Technology, investment, preference and monetary policy shocks are found to be fundamental for all samples. Such clustering is very similar to the Chari et al. (2008) classification of structural and non structural shocks. Second, the estimated posterior distributions of deep parameters are different when we allow for the possibility of a rank-deficient matrix of structural shocks. Substantial differences appear in the estimated persistence of shocks, investment adjustment costs parameter, and price and wage indexation and stickiness parameters. And as a consequence, the transmission mechanism of the fundamental shocks, in particular monetary and fiscal shocks, is altered. By means of marginal likelihood comparisons, data prefer versions of the model with a rank deficient structural shocks structure. Third, when estimated using normal priors, most of the variability of observables such as wages is explained by measurement error, i.e. variable-specific

26

rather than common fundamental shocks.

References

Adjemian, S., Bastani, H., Karamé, F., Juillard, M., Maih, J., Mihoubi, F., Perendia, G., Ratto, M. and Villemot, S.: 2011, Dynare: Reference manual version 4, *Dynare Working Papers 1*, CEPREMAP.

Belmonte, M. A. G., Koop, G. and Korobilis, D.: 2014, Hierarchical Shrinkage in Time-Varying Parameter Models, *Journal of Forecasting* **33**, 80–94.

Boivin, J. and Giannoni, M.: 2006, DSGE Models in a Data-Rich Environment, *NBER Technical Working Papers 0332*, National Bureau of Economic Research, Inc.

Canova, F.: 2014, Bridging DSGE model and raw data, *Journal of Monetary Economics* **57**(2), 132–145.

Canova, F. and Ferroni, F.: 2011, Multiple filtering devices for the estimation of cyclical DSGE models, *Quantitative Economics* **2**(1), 73–98.

Canova, F., Ferroni, F. and Matthes, C.: 2015, Approximating time varying structural model with time invariant structures, *Mimeo*, Mimeo.

Chari, V. V., Kehoe, P. J. and McGrattan, E. R.: 2008, New Keynesian Models: Not Yet Useful for Policy Analysis, *American Economic Journal: Macroeconomics* **1**(1), 24–66.

Christiano, L. J., Eichenbaum, M. and Evans, C. L.: 2005, Nominal rigidities and the dynamic effects of a shock to monetary policy, *Journal of Political Economy* **113**(1), 1–45.

Cúrdia, V. and Reis, R.: 2010, Correlated Disturbances and U.S. Business Cycles, *NBER Working Papers 15774*, National Bureau of Economic Research, Inc.

Del Negro, M. and Schorfheide, F.: 2009, Monetary Policy Analysis with Potentially Mis-specified Models, *American Economic Review* **99**(4), 1415–50.

Díaz-García, J. A. and Gutiérrez-Jáimez, R.: 1997, Proof of the conjectures of H. Uhlig on the singular multivaraite beta and the Jacobian of a certain tranformation, *Annals of Statistics* **25**(1), 2018–2033.

Díaz-García, J. A. and Gutiérrez-Jáimez, R.: 2006, Distribution of the generalized inverse of a random matrix and its applications, *Journal of Statistical Planning and Inference* **136**(1), 183 – 192.

Eisenstat, E., Chan, J. C. and Strachan, R. W.: 2014, Stochastic Model Specification Search for Time-Varying Parameter VARs, *CAMA Working Papers 2014-23*, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.
   URL: *http://ideas.repec.org/p/een/camaaa/2014-23.html*

Fernández-Villaverde, J. and Rubio-Ramírez, J. F.: 2008, How Structural Are Structural Parameters?, *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 83–137.

Ferroni, F.: 2011, Trend Agnostic One-Step Estimation of DSGE Models, *The B.E. Journal of Macroeconomics* **11**(1), 1–36.

Frühwirth-Schnatter, S.: 2004, *Efficient Bayesian parameter estimation*, State Space and Unobserved Component Models: Theory and Applications, Cambridge University.

Fruhwirth-Schnatter, S. and Wagner, H.: 2010, Stochastic model specification search for Gaussian and partial non-Gaussian state space models, *Journal of Econometrics* **154**(1), 85–100.

Galvao, A., Giraitis, L., Kapetanios, G. and Petrova, K.: 2015, Local Bayesian estimation and forecasting with a TVP model, *Mimeo*, Mimeo.

Gamerman, D. and Lopes, H. F.: 2006, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Second Edition (Chapman and Hall/CRC Texts in Statistical Science), New York, USA.

Gelfand, A. E., Sahu, S. K. and Carlin, B. P.: 1995, Efficient parameterizations for normal linear mixed models, *Biometrika* **82**, 479 – 488.

George, E. I. and McCulloch, R.: 1993, Variable selection via Gibbs sampling, *Journal of the American Statistical Association* **88**, 881 – 889.

George, E. I. and McCulloch, R.: 1997, Approaches for Bayesian variable selection, *Statistica Sinica* **7**, 339 – 373.

Geweke, J.: 1999, Using simulation methods for bayesian econometric models: inference, development and communication, *Econometric Reviews* **18**(1), 1–73.

Geweke, J.: 2005, *Contemporary Bayesian Econometrics and Statistics*, Wiley, New York, USA.

Gorodnichenko, Y. and Ng, S.: 2010, Estimation of DSGE models when the data are persistent, *Journal of Monetary Economics* **57**(3), 325–340.

Grassi, S. and Proietti, T.: 2014, Characterizing economic trends by bayesian stochastic model specification search, *Computational Statistics and Data Analysis* **71**, 359 – 374.

Harvey, A. C.: 1989, *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

Harvey, A. C.: 2001, Testing in unobserved components models, *Journal of Forecasting* **20**, 1 – 19.

Inoue, A., Kuo, C.-H. and Rossi, B.: 2014, Identifying the Sources of Model Misspecification, *CEPR Discussion Papers 10140*, C.E.P.R. Discussion Papers.

Ireland, P. N.: 2004, A method for taking models to the data, *Journal of Economic Dynamics and Control* **28**(6), 1205–1226.

Iskrev, N.: 2010, Local identification in DSGE models, *Journal of Monetary Economics* **57**(2), 189–202.

Justiniano, A., Primiceri, G. E. and Tambalotti, A.: 2010, Investment Shocks and Business Cycles , *Journal of Monetary Economics* **57**(2), 132–145.

Justiniano, A., Primiceri, G. E. and Tambalotti, A.: 2013, Is There a Trade-Off between Inflation and Output Stabilization?, *American Economic Journal: Macroeconomics* **5**(2), 1–31.

Komunjer, I. and Ng, S.: 2011, Dynamic Identification of Dynamic Stochastic General Equilibrium Models, *Econometrica* **79**(6), 1995–2032.

Proietti, T. and Grassi, S.: 2014, Stochastic trends and seasonality in economic time series: new evidence from bayesian stochastic model specification search., *Empirical Economics* **Forthcoming**.

Sala, L., Soderstrom, U. and Trigari, A.: 2010, The Output Gap, the Labor Wedge, and the Dynamic Behavior of Hours, *Working Paper Series 246*, Sveriges Riksbank (Central Bank of Sweden).

Schorfheide, F.: 2013, Estimation and Evaluation of DSGE Models: Progress and Challenges, *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress*.

Smets, F. and Wouters, R.: 2007, Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach, *American Economic Review* **97(3)**, 586–606.

Uhlig, H.: 1994, On Singular Wishart and Singular Multivariate Beta Distributions, *Annals of Statistics* **22**(1), 395–405.

A   Appendix

A.1   A basic RBC with analytical solution

The representative agent maximizes the following stream of future utility

$$\max E_t \sum_{t=1}^{\infty} \beta^t \ln c_t$$

subject to the following constraints

$$y_t = c_t + k_t$$
$$y_t = z_t k_{t-1}^{\alpha}$$

where y_t is output, c_t consumption, k_t the stock of capital. β is the time discount factor and α is the capital share in production. The system is perturbed by one exogenous disturbance, technology z_t, which follows an AR process

$$\ln z_t = \rho_z \ln z_{t-1} + e_t \quad e_t \sim N(0, \sigma^2)$$

The lagrangian is

$$L = E_0 \sum_{t=0}^{\infty} \beta_t \left[\ln c_t - \lambda_t \left(c_t + k_t - z_t k_{t-1}^{\alpha} \right) \right]$$

The First Order Conditions are

$$1/c_t = \lambda_t$$
$$1/c_t = \beta E_t \left(1/c_{t+1} \alpha z_{t+1} k_t^{\alpha-1} \right)$$

Permanent income model. Guess a solution of the form $c_t = \gamma y_t$, constant saving rate and substitute into the Euler equation.

$$
\begin{aligned}
1 &= \beta E_t \left(\gamma y_t / \gamma y_{t+1} \alpha z_{t+1} k_t^{\alpha-1} \right) \\
&= \beta E_t \left(y_t / y_{t+1} \alpha y_{t+1} / k_t \right) \\
&= \beta E_t \left(\alpha y_t / (y_t - c_t) \right) \\
&= \beta E_t \left(\alpha y_t / (y_t - \gamma y_t) \right) \\
1 &= \frac{\alpha \beta}{1 - \gamma}
\end{aligned}
$$

Hence, $\gamma = 1 - \alpha\beta$. This implies that

$$k_t = (1 - \gamma)y_t = \alpha\beta z_t k_{t-1}^{\alpha}$$

In logs, we can specify a linear state space model in three equations, a law of motion for the exogenous state (z), a law of motion for the endogenous state (k), and the measurement equation (y) as follows

$$\ln k_t = \ln \alpha\beta + \alpha \ln k_{t-1} + \ln z_t$$
$$\ln z_{t+1} = \rho_z \ln z_t + e_{t+1} \quad e_t \sim N(0, \sigma^2)$$
$$\ln y_t = \ln z_t + \alpha \ln k_{t-1} + u_t \quad u_t \sim N(0, \sigma_m^2)$$

At the non stochastic steady state, we have $\ln k = 1/(1-\alpha)\ln\alpha\beta$ and $\ln y = \alpha\ln k$

$$k_t = \alpha k_{t-1} + z_t$$

$$z_{t+1} = \rho_z z_t + e_{t+1}$$

$$y_t = z_t + \alpha k_{t-1} + u_t$$

where small case variables indicate now the log deviation from steady state.

A.2 Stochastic variable selection in state space models

Bayesian stochastic variable selection has a long tradition in Bayesian analysis (see, among others, George and McCulloch (1993, 1997) and the references therein). Recently, this methodology has been extended to state space models and, in particular, to the selection of the unobserved components (level, slope and seasonal cycles) that are the key ingredients in state space modeling (see Frühwirth-Schnatter (2004), Fruhwirth-Schnatter and Wagner (2010), Grassi and Proietti (2014) and Proietti and Grassi (2014)). This approached, called stochastic model selection search (SMSS), hinges on two basic ingredients: the non-centered representation of the unobserved components model and the consequent reparameterization of the variance hyperparameters as regression parameters with unrestricted support.

Consider, for example, modeling a time series $\mathbf{y} = \{y_1, \ldots, y_t\}$ using a local level model, see Harvey (1989) for an introduction:

$$\begin{aligned} y_t &= z_t + u_t & u_t &\sim \mathrm{N}(0, \sigma_u^2) \\ z_t &= z_{t-1} + e_t & e_t &\sim \mathrm{N}(0, \sigma_e^2) \end{aligned} \tag{7}$$

where the latent process z_t follows a random walk starting from unknown initial value μ_0. A typical specification problem arising for this model is to decide if the random walk z_t is time-varying rather than a simple constant. It is well know that testing $\sigma_e^2 = 0$ versus $\sigma_e^2 > 0$ results in a non-regular testing problem, because the null hypothesis lies on the boundary of the parameter space, see Harvey (1989) and Harvey (2001).

A similar specification problem is deciding which components are present in this time series model. For instance, is it necessary to include z_t, that follows a random walk, or should it be removed because z_t is simply a constant? This is another non-regular problem because, again, the null hypothesis can be rephrased as testing $\sigma_e^2 = 0$ versus $\sigma_e^2 > 0$.

The stochastic model specification search methodology proposed by Fruhwirth-Schnatter and Wagner (2010) (FS-W) is based on a reparameterisation of (A.2) with respect to location and scale, known as the non-centred representation. See also Gelfand, Sahu and Carlin 1995, Frühwirth-Schnatter 2004). To give a simple example, the model in equation (A.2) has the

following non-centered representation:

$$y_t = \mu_0 + \sqrt{\theta_1}\tilde{z}_t + u_t, \quad u_t \sim N(0, \sigma_u^2),$$
$$\tilde{z}_t = \tilde{z}_{t-1} + \tilde{e}_t, \qquad \tilde{e}_t \sim N(0, 1), \tag{8}$$
$$\tilde{z}_0 = 0, \quad \theta_1 = \sigma_e^2,$$

where the latent states has been rewritten as follows:

$$z_t = \mu_0 + \sqrt{\theta_1}\tilde{z}_t, \quad t = 1, \ldots, T,$$
$$\tilde{z}_t = \tilde{z}_{t-1} + \tilde{e}_t, \qquad \tilde{e}_t \sim \mathrm{N}(0, 1), \tag{9}$$

where $\tilde{z}_0$ is the starting value of the random walk and $\tilde{z}_t \sim N(0, t)$.

Between the non-centered and centered representation, there exists a one to one relation that can be easily shown using (8) and (9):

$$y_t = \mu_0 + \sqrt{\theta_1}\tilde{z}_t + u_t, \quad u_t \sim \mathrm{N}(0, \sigma_u^2),$$
$$\tilde{z}_t = \tilde{z}_{t-1} + \tilde{e}_t, \qquad \tilde{e}_t \sim \mathrm{N}(0, 1), \tag{10}$$

and rewriting:

$$z_t - z_{t-1} = \sqrt{\theta_1}(\tilde{z}_t - \tilde{z}_{t-1}),$$
$$= \sqrt{\theta_1}\tilde{e}_t = e_t, \quad e_t \sim \mathrm{N}(0, \sigma_e^2). \tag{11}$$

FS-W's key idea is that the non-centered representation is not identified since the model in equation (8) with $(-\sqrt{\theta_1})(-\tilde{z}_t)$ is observationally equivalent to the same model with $(\sqrt{\theta_1})(\tilde{z}_t)$. As a consequence, the likelihood function is symmetric around zero along the $\sqrt{\theta_1}$ dimension and bimodal if the true $\sqrt{\theta_1}$ is larger than zero. This fact can be exploited to quantify how far the posterior of $\sqrt{\theta_1}$ is removed from zero and, in turn, the value of the variance. We stress that the posterior density can also be 0 allowing for boundary conditions. To estimate the model in equation (8) that is equivalent to the model in (A.2) a standard RW-MH algorithm can be used, see Gamerman and Lopes (2006) and Geweke (2005). Finally, we have to underline that this methodology can easily be extended to more complex state space models as shown in FS-W and in Grassi and Proietti (2014) and Proietti and Grassi (2014).

We extend this methodology to linear DSGE models. To show the workings of this extension, consider the basic RBC model in section 4.1 and A.1. In logs, we can specify a linear state space model in three equations, a law of motion for the exogenous state (z), a law of motion for the endogenous state (k), and the measurement equation (y) as follows

$$\ln k_t = \ln \alpha\beta + \alpha \ln k_{t-1} + \ln z_t$$
$$\ln z_{t+1} = \rho_z \ln z_t + e_{t+1} \quad e_{t+1} \sim N(0, \sigma_e^2) \tag{12}$$
$$\ln y_t = \ln z_t + \alpha \ln k_{t-1} + u_t \quad u_t \sim N(0, \sigma_u^2)$$

At the non stochastic steady state, we have $\ln k = 1/(1-\alpha)\ln \alpha\beta$ and $\ln y = \alpha \ln k$ and dropping the ln and lagging the latent state we get:

$$k_{t-1} = \alpha k_{t-2} + z_{t-1}$$
$$z_t = \rho_z z_{t-1} + e_t \tag{13}$$
$$y_t = z_t + \alpha k_{t-1} + u_t$$

where, abusing notation, lower case variables now indicate the log deviation from steady state. Given this, the non-centred state space representation of the model is:

$$y_t = \mu_0 + \sqrt{\theta_1}\tilde{z}_t + \alpha\tilde{k}_{t-1} + u_t \quad u_t \sim N(0, \sigma_u^2)$$
$$\tilde{z}_t = \rho_z\tilde{z}_{t-1} + \tilde{e}_t \quad \tilde{e}_t \sim N(0, 1) \tag{14}$$
$$\tilde{k}_{t-1} = \alpha\tilde{k}_{t-2} + \sqrt{\theta_1}\tilde{z}_{t-1}.$$

This formulation is identified as the following steps show. Define the process

$$z_t = \mu_0 + \sqrt{\theta_1}\tilde{z}_t.$$

Then we have the following formulation:

$$z_t - z_{t-1} = \sqrt{\theta_1}(\tilde{z}_t - \tilde{z}_{t-1})$$
$$= \sqrt{\theta_1}\tilde{e}_t,$$

$$\tilde{k}_{t-1} - \tilde{k}_{t-2} = \alpha\tilde{k}_{t-2} + \sqrt{\theta_1}\tilde{z}_{t-2} - \alpha\tilde{k}_{t-3} + \sqrt{\theta_1}\tilde{z}_{t-3}$$
$$= \alpha(\tilde{k}_{t-2} - \tilde{k}_{t-3}) + \sqrt{\theta_1}(\tilde{z}_{t-1} - \tilde{z}_{t-2})$$
$$= \alpha(\tilde{k}_{t-2} - \tilde{k}_{t-3}) + \sqrt{\theta_1}\tilde{e}_{t-1}.$$

It is clear that $\tilde{k}_{t-1}$ is related to the error term $\sqrt{\theta}\tilde{e}_{t-1}$ as can also be shown in equation (12). The state space formulation of the model then becomes:

$$y_t = \mu_0 + \begin{pmatrix} \sqrt{\theta_1} & \alpha \end{pmatrix} \begin{pmatrix} \tilde{z}_t \\ \tilde{k}_{t-1} \end{pmatrix} + u_t,$$
$$\begin{pmatrix} \tilde{z}_t \\ \tilde{k}_{t-1} \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ \sqrt{\theta_1} & \alpha \end{pmatrix} \begin{pmatrix} \tilde{z}_{t-1} \\ \tilde{k}_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tilde{e}_t.$$

A.3 METROPOLIS HASTINGS MCMC ADJUSTED FOR THE SIGN SWITCH

Here, we explain the steps to adjust the RW Metropolis-Hastings MCMC for a random sign switch. Partition the vector of parameters θ as composed of a column vector of structural standard deviation parameters (σ) and a column vector of the remaining parameters (ϑ), i.e. $\theta = [\sigma', \vartheta']'$. Given an initial value, θ_0, and the information matrix, Ω, from the maximization step, given the size of the jump c, and given a positive sign for the standard deviation, i.e. $W = 1$, for $\ell = 1, .., L$

1. Draw a candidate draw from $\theta^* \sim N(\theta_{\ell-1}, c\Omega)$.

2. Plug it in the DSGE model, $E_t F(x_{t+1}, x_t, x_{t-1}, \epsilon_t; \theta^*) = 0$.

3. Solve the DSGE, and obtain the state space representation

$$s_{t+1} = A(\vartheta^*)s_t + B(\vartheta^*) \Sigma(\sigma^*) W \epsilon_{t+1}$$
$$y_t = \Lambda s_t + e_t.$$

4. Compute the likelihood using the Kalman filter, i.e. $L(y|\theta^*)$.

5. Contrast the kernels of the candidate draw and previous accepted draws.

$$R = \frac{p(\theta^*)L(\theta^*|y)}{p(\theta_{\ell-1})L(\theta_{\ell-1}|y)}$$

6. Keep the draw with certain probability. Draw $u \sim U(0,1)$

$$if \quad R > u, \quad \theta_\ell = \theta^*$$
$$if \quad R \leq u, \quad \theta_\ell = \theta_{\ell-1}$$

7. Switch the sign of the standard deviation with 0.5 probability. Draw form a binomial distribution, $X \sim b(1/2)$, and set the sign of the standard deviation with $W = -1+2X$. Multiply the standard deviation of the structural shocks times W,

$$\sigma_\ell = W\sigma_\ell$$

.

8. Go back to 1.

A.4 Gibbs -Metropolis Hastings MCMC for non-diagonal and rank-deficient matrix

We follow Cúrdia and Reis (2010) and implement a Gibbs-Metropolis algorithm by combining the conjugacy of the singular generalized inverse Wishart of the shock covariance matrix with the RW Metropolis for the structural parameters. In particular, the sampling of the parameters needs to be partitioned in two blocks: the covariance matrix of the structural shocks (Σ) and all other parameters (θ). Conditional on a value of θ and on a sequence of states $\{s_t\}_{t=1}^T$, we can derive a sequence of i.i.d. structural shocks as follows,

$$B(\theta)^+ (s_{t+1} - A(\theta)s_t) = z_{t+1} = \epsilon_{t+1} \sim N_r(0, \Sigma)$$

where $B(\theta)^+$ is the left Moore-Penrose generalized inverse of $B(\theta)$. Conditional on a sequence of states $\{s_t\}_{t=1}^T$, this model can be cast in matrix form as

$$Z = E$$

where $Z = (z_1, ..., z_T)'$ and $p(E|\Sigma) \equiv N_{T,r}(0, \Sigma \otimes I_T)$. As in Díaz-García and Gutiérrez-Jáimez (2006), we denote by $N_{T,r}(0, \Sigma \otimes I_T)$ the $T \times n$ multivariate singular Normal distribution with rank r. The likelihood of the normal singular population can be written as

$$L(\Sigma|Z) \propto \left(\prod_{k=1}^r \lambda_k\right)^{-\frac{T}{2}} \exp\left(-1/2 \operatorname{trace}(\Sigma^+ Z'Z)\right)$$

where λ_k are the non-null eigenvalues of Σ. We consider the following non informative prior density for Σ

$$p(\Sigma) \propto \left(\prod_{k=1}^r \lambda_k\right)^{-\frac{2n-r+1}{2}}$$

34

where λ_k are the non null eigenvalues of Σ. Combining prior and posterior, we obtain the posterior distribution of Σ

$$p(\Sigma|Z) \propto \left(\prod_{k=1}^{r} \lambda_k \right)^{-\frac{T+2n-r+1}{2}} \exp\left(-1/2 \operatorname{trace}(\Sigma^+ Z'Z) \right). \tag{15}$$

This is an n-dimension Singular Generalized Inverse Wishart of rank r with $\nu = T - n + 1$ degrees of freedom and scale matrix $G = Z'Z$, denoted by $W^+(r, \nu, G)$.

We are now in a position to propose the following algorithms. Given r, $\Sigma_{(0)}, \theta_{(0)}$

Algorithm 2

1. Draw $s_{1:T}^{(j)}$ from $p(s_{1:T}^{(j)} \mid y_{1:T}, \ \theta_{(j-1)}, \ \Sigma_{(j-1)})$.
This distribution is derived from the state space and, given linearity assumptions, it is a gaussian normal distribution.

2. Draw $\Sigma_{(j)}$ from $p(\Sigma_{(j)} \mid y_{1:T}, \ \theta_{(j-1)}, \ s_{1:T}^{(j)})$.
This distribution is an n-dimension Singular Generalized Inverse Wishart, $W^+(r, \nu, G^{(j)})$, where $G^{(j)} = Z^{(j)'} Z^{(j)}$, $Z^{(j)} = (z_1^{(j)}, ..., z_T^{(j)})'$ and $z^{(j)} = B(\theta_{(j-1)})^+ \ (s_{t+1}^{(j)} - A(\theta_{(j-1)})s_t^{(j)})$, and degrees of freedom $\nu = T - n + 1$

3. Draw θ^* form a normal centered in $\theta_{(j-1)}$ and accept the draw with a Metropolis-Hastings probability, i.e. $\min \left\{ \frac{L(y_{1:T}|\theta^*, \Sigma_{(j)})p(\theta^*)}{L(y_{1:T}|\theta_{(j)}, \Sigma_{(j)})p(\theta_{(j-1)})}, 1 \right\}$

In step 3 the likelihood is computed using the Kalman filter recursions. Since the state space is augmented with n_y measurement errors, the covariance matrix of the observables is full rank, hence invertible. Therefore, the Kalman gain, defined as the product of the covariance between states and observables times the inverse of the variance of the observables, can be computed and all the remaining recursions are unaffected.

In order to obtain draws at step 2 use the following algorithm:

Algorithm 3 *Singular Inverse Wishart.*
If any U is an n-dimensional Wishart Singular with degrees of freedom ν and scale matrix C, where both U and C are $n \times n$ symmetric positive semi-definite singular matrices of rank r, then we can draw U as follows:

1. *For $C = PLP'$, where PLP' is the non-singular part of the spectral decomposition of C, calculate $PL^{1/2}$.*

2. *Generate $x_1, .., x_\nu$ independently from $N(0, I_r)$.*

3. *$U = \sum_{i=1}^{\nu} W_i W_i'$, where $W_i = Bx_i$*

Then U is drawn from the n-dimensional Wishart Singular with ν degrees of freedom, scale matrix C, and rank r, and U^+ is drawn from the n-dimensional Singular Generalized Inverse Wishart with ν degrees of freedom, scale matrix C, and rank r.

For step 1, we know that $s_{t|T}$ is normally distributed as a singular multivariate normal distribution, i.e. $s_{t|T} \sim N_r(s_{t-1|T}, Q_{t-1|T})$ where $Q_{t-1|T}$ is the covariance of the states. Drawing from this distribution is easy: take the non singular part of the spectral decomposition of $Q_{t-1|T}$, i.e. $Q_{t-1|T} = PLP'$, draw x from a normal $N(0, I_r)$ and $s_{t|T}^{(j)} = PL^{1/2}x$.

A.5 SMETS AND WOUTERS (2007) MODEL

The log linearized equilibrium conditions are summarized as follows

$$y_t = c/y c_t + i/y i_t + r^k k/y z_t + e_t^g$$

$$c_t = c_1 c_{t-1} + (1 - c_1)Ec_{t+1} + c_2(h_t - Eh_{t+1}) - c_3(r_t - E\pi_{t+1} + e_t^b)$$

$$i_t = i_1 i_{t-1} + (1 - i_1)E_t i_{t+1} + i_2 q_t + \epsilon_t^i$$

$$q_t = q_1 Eq_{t+1} + (1 - q_1)Er_{t+1}^k - (r_t - E\pi_{t+1} + e_t^b)$$

$$y_t = \alpha \phi_p k_t + (1 - \alpha)\phi_p h_t + \phi_p \epsilon_t^a$$

$$k_t^s = k_{t-1} + z_t$$

$$z_t = \psi/(1 - \psi)r_t^k$$

$$k_t = k_1 k_{t-1} + (1 - k_1)i_t + k_2 \epsilon_t^i$$

$$mp_t = \alpha(k_t^s - h_t) + e_t^a - w_t$$

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E\pi_{t+1} - \pi_3 mp_t + e_t^p$$

$$r_t^k = -(k_t - h_t) + w_t$$

$$mw_t = w_t - \left(\sigma_n h_t + \frac{1}{1 + \lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})\right)$$

$$w_t = w_1 w_{t-1} + (1 - w_1)E(\pi_{t+1} + w_{t+1}) - w_2 \pi_t + w_3 \pi_{t-1} + mw_t + e_t^w$$

$$R_t = \rho_R R_{t-1} + (1 - \rho_R)(\rho_\pi \pi_t + \rho_y(y_t - y_t^f) + \rho_{\Delta y}\Delta(y_t - y_t^f)) + e_t^r$$

$$+ \; flexible \; economy \; equations$$

where variables with time subscript are variables from the original non linear model expressed in log deviation from the steady state. Flexible output is defined as the level of output that would prevail under flexible prices and wages in the absence of the two mark-up shocks. Seven structural shocks. The model has five AR(1), government, technology, preference, investment specific, monetary policy, and two ARMA(1,1) processes, price and wage markup.

$$c_1 = \lambda/\gamma(1 + \lambda/\gamma)$$

$$c_2 = [(\sigma_c - 1)(W^h h/C)]/[\sigma_c(1 + \lambda/\gamma)]$$

$$c_3 = (1 - \lambda/\gamma)/(1 + \lambda/\gamma)\sigma_c$$

$$k_1 = (1 - \delta)/\gamma$$

$$k_2 = (1 - (1 - \delta)/\gamma)(1 + \beta\gamma^{1-\sigma_c})\gamma^2\phi$$

$$i_1 = 1/(1 + \beta\gamma^{1-\sigma_c})$$

$$i_2 = (1/(1 + \beta\gamma^{1-\sigma_c})\gamma^2\phi$$

$$q_1 = \beta\gamma^{-\sigma_c}(1 - \delta)$$

$$\pi_1 = i_p/(1 + \beta\gamma^{1-\sigma_c}i_p)$$

$$\pi_2 = \beta\gamma^{1-\sigma_c}/(1 + \beta\gamma^{1-\sigma_c}i_p)$$

$$\pi_3 = 1/(1 + \beta\gamma^{1-\sigma_c}i_p)[(1 - \beta\gamma^{1-\sigma_c}\xi_p)(1 - \xi_p)/\xi_p(1 + (\phi_p - 1)\epsilon_p)))]$$

$$w_1 = 1/(1 + \beta\gamma^{1-\sigma_c})$$

$$w_2 = (1 + \beta\gamma^{1-\sigma_c}i_w)/(1 + \beta\gamma^{1-\sigma_c})$$

$$w_3 = i_w/(1 + \beta\gamma^{1-\sigma_c})$$

$$w_4 = 1/(1 + \beta\gamma^{1-\sigma_c})[(1 - \beta\gamma^{1-\sigma_c}\xi_w)(1 - \xi_w)/\xi_w(1 + (\phi_w - 1)\epsilon_w)]$$

$$\overline{\gamma} = 100(\gamma - 1)$$

$$\overline{\pi} = 100(\pi_* - 1)$$

$$\overline{\beta} = ((\pi_*/(\beta * \gamma^{-\sigma_c})) - 1) * 100$$

The coefficients are function of the deep parameters of the model which are summarized and described in table 7.

θ	Description	SW mean or fixed values
γ	slope of the deterministic trend in technology	1.004
δ	depreciation rate	0.025
ε_p	good markets kimball aggregator	10
ε_w	labor markets kimball aggregator	10
λ_w	elasticity of substitution labor	1.5
cg	gov't consumption output share	0.18
β	time discount factor	0.998
ϕ_p	1 plus the share of fixed cost in production	1.61
ϕ	inverse of the elasticity of investment relative to installed capital	5.74
α	capital share	0.19
λ	habit in consumption	0.71
ξ_w	wage stickiness	0.73
ξ_p	price stickiness	0.65
i_w	wage indexation	0.59
i_p	price indexation	0.47
σ_n	elasticity of labor supply	1.92
σ_c	intertemporal elasticity of substitution	1.39
ψ	st. st. elasticity of capital adjustment costs	0.54
ρ_π	monetary policy response to π	2.04
ρ_R	monetary policy autoregressive coeff.	0.81
ρ_y	monetary policy response to y	0.08
$\rho_{\Delta y}$	monetary policy response to y growth	0.22
ρ_a	technology autoregressive coeff.	0.95
ρ_g	gov spending autoregressive coeff.	0.97
ρ_{ga}	cross coefficient tech-gov	0
ρ_b	technology autoregressive coeff.	0.21
ρ_q	technology autoregressive coeff.	0.71
ρ_m	monetary policy autoregressive coeff.	0.15
ρ_p	price markup autoregressive coeff.	0.89
ρ_w	wage markup autoregressive coeff.	0.96
μ_w	wage markup ma coeff.	0
μ_w	wage markup ma coeff.	0
σ_a	sd technology	0.45
σ_g	sd government spending	0.52
σ_b	sd preference	0.25
σ_r	sd monetary policy	0.24
σ_q	sd investment	0
σ_w	sd wage markup	0
σ_p	sd price markup	0

Table 7: Parameters description and numerical values eitehr fixed or obtained from the posterior mean estimated or fixed