Federal Reserve Bank of Chicago

Selecting Primal Innovations in DSGE models

Filippo Ferroni, Stefano Grassi, and Miguel A. León-Ledesma

August 2017

WP 2017-20

Selecting Primal Innovations in DSGE models*

Filippo Ferroni
Chicago FED

Stefano Grassi
Department of Economics and Finance,
University of Rome 'Tor Vergata' and CREATES

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

August 22, 2017

Abstract

DSGE models are typically estimated assuming the existence of certain primal shocks that drive macroeconomic fluctuations. We analyze the consequences of estimating shocks that are "non-existent" and propose a method to select the primal shocks driving macroeconomic uncertainty. Forcing these non-existing shocks in estimation produces a downward bias in the estimated internal persistence of the model. We show how these distortions can be reduced by using priors for standard deviations whose support includes zero. The method allows us to accurately select primal shocks and estimate model parameters with high precision. We revisit the empirical evidence on an industry standard medium-scale DSGE model and find that government and price markup shocks are innovations that do not generate statistically significant dynamics.

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

JEL Classification: C10, E27, E32.

*Ferroni: Chicago FED, 230 LaSalle, Chicago IL. Grassi: University of Rome 'Tor Vergata' and CREATES. León-Ledesma: School of Economics and Macroeconomics, Growth, and History Centre (MaGHiC), Keynes College, University of Kent, Canterbury, CT27NP, UK. An earlier version of this paper circulated with the title, 'Fundamental shock selection in DSGE models'. We wish to thank for helpful discussions Gianni Amisano, Benjamin Born, Fabio Canova, Samuel Hurtado, Alejandro Justiniano, Eric Leeper, Nels Lind, Christian Matthes, Luigi Paciello, Barbara Rossi, Juan Rubio-Ramirez, Luca Sala, Chris Sims, Paolo Surico, Leonardo Melosi, Andrea Tambalotti. The paper also benefited from comments of participants at the 2014 Working Group of Econometric Modeling at the ECB, the 2015 Annual Meeting of the Royal Economic Society, the 2015 Barcelona Summer Forum, the 2015 IAAE conference, and various seminars. Two anonymous referees contributed to improve the empirical part of the paper. Remaining errors are our own. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Chicago or any other person associated with the Federal Reserve System.

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 primal 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, IG henceforth). In classical statistics, standard deviations are re-parameterized by taking logarithmic transformations. In doing so, we rule out boundary solutions and, by construction, primal innovations always exist.[2]

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 et al., 2008 and Sala et al., 2010). This is an important question in modern stochastic models of economic fluctuations. Empirically, these models face two challenges. First, unveiling the primal 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 primal 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-existent DSGE shocks? This is the first question we tackle in this paper.

Some shocks may not be primal because of null 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 than merely postulate, the existence of certain shocks. I.e., we want to be able to select which innovations are important and primal drivers of aggregate uncertainty. To be able to estimate the possibly rank deficient covariance matrix of structural shocks, we need to 1) add idiosyncratic

[1] See Smets and Wouters (2007) and Justiniano et al. (2010) amongst many others.

[2] We use the term "primal" shock throughout the paper to refer to shocks that have an economic interpretation and that affect the dynamics of the model. I.e. they are relevant sources of uncertainty. These shocks could also be termed "structural" or "fundamental", but we prefer to term them primal to avoid confusion with the terminology used in other streams of the literature dealing with very different problems to ours.

measurement errors and 2) abandon standard IG priors and use distributions (univariate or multivariate) that allow for zero variances (or null eigenvalues).

Imposing a non-existent exogenous process has serious consequences for inference. In particular, we argue that it generates a downward bias in the estimate of the internal persistence of the model. In fully fledged DSGE models, the persistence of the model's dynamics is controlled not only by the autoregressive parameters, but also by deep parameters capturing real and nominal frictions in the economy. We try to quantify these distortions in a Montecarlo experiment using a medium-scale model. In particular, when the econometrician assumes that the rank of the covariance matrix of primal shocks, Σ, is larger than the one of the true DGP, autoregressive parameters and parameters driving price and wage stickiness and indexation are grossly underestimated. Thus, we unveil a trade-off between including a wide set of potential sources of impulses (say in order to match more observable variables) and correctly identifying the model parameters that drive propagation. We then show that these distortions are reduced by considering priors on the primal 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. Proper priors such as normal or exponential distributions have appealing properties since they allow us to recover existent and non-existent shocks in situations where the true number and combination of primal shocks is unknown [3].

We explore the consequences of our approach in an empirical application and revisit the evidence of an industry-standard DSGE model. We estimate the Smets and Wouters (2007) model (SW hereafter) on seven key US quarterly macroeconomic time series, namely, the growth rate of real output, consumption, investment, wages, and hours worked, the inflation rate, and the short run interest rate. For comparability proposes, we consider the original data span, 1968-2004, and only depart from the baseline specification of the model by adding measurement errors on each observable and by assuming normal priors on the standard deviation of shocks. We identify as common sources of fluctuations technology shocks, monetary policy surprises, risk-premium, and investment demand shocks. Government spending and price markup shocks do not generate statistically significant dynamics for the variables considered. Wage markup shocks are only marginally significant. The choice of the priors for standard deviations (STDs hereafter) matters for the estimates of the structural parameters of the SW model and, as a consequence, matters when studying the transmission mechanism of primal impulses. Relative to a model estimated without measurement errors, the responses of macro variables to risk-premium shocks are more persistent and hump shaped. When we impose the existence of structural and measurement error shocks (both with inverse gamma priors), we obtain unrealistically strong responses to monetary policy surprises. When we

[3] In the online appendix A.4 we consider a more general structural shocks covariance matrix where innovations can be correlated. We propose considering the conjugate Metropolis-within-Gibbs sampler proposed in Cúrdia and Reis (2010) adapted for rank deficient covariance matrices. In a rank deficient environment, we consider the singular inverse Wishart (IW) distribution (see Uhlig, 1994 and Díaz-García and Gutiérrez-Jáimez, 1997) and the conjugacy results in Díaz-García and Gutiérrez-Jáimez (2006).

are agnostic about the existence of primal and/or measurement errors, the transmission mechanism of monetary policy shocks resembles closely those available from other empirical studies. In this empirical application, we find that normal and inverse gamma priors deliver the same posterior estimates of the parameters when we have the same number of shocks and observables. However, prior specification comparisons in terms of marginal likelihoods are difficult to interpret as the marginal likelihood estimates using inverse gamma priors are not invariant to the prior location. When there are more shocks than observables, inverse gamma priors produce posterior estimates that differ substantially depending on the prior location. It is not straightforward how to select amongst them. Normal priors overcome all these concerns.

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 et al. (2014) or Eisenstat et al. (2016). 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 et al., 2014), parameter instability of various forms (e.g. Fernández-Villaverde and Rubio-Ramírez, 2008, Galvao et al., 2015, Canova et al., 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 exist*.[4]

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 few concluding remarks.

2 Priors for Primal DSGE shocks selection

Consider a DSGE model with (deep) parameters of interest $\mathbf{\Theta}$. The (control and state) variables of the model, denoted by s_t, are driven by primal 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

[4]There is a large literature on shocks "fundamentalness", in particular in the VAR context, see Lippi and Reichlin (1994) and subsequent literature. The concern in that literature focuses on the non-invertibility of the MA representation of the DSGE model. When not invertible, with the available set of observables, it is impossible to recover all the primal or fundamental DSGE shocks using linear projections, i.e. VAR. Although in the text we may refer occasionally to "fundamental" sources of macroeconomic uncertainty, we are clearly not relating to this stream of work.

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:

$$s_{t+1} = \mathrm{A}(\mathbf{\Theta})s_t + \mathrm{B}(\mathbf{\Theta})\epsilon_{t+1} \qquad \epsilon_{t+1} \sim \mathrm{N}_r(0, \Sigma_\epsilon), \tag{1}$$

where A, B are nonlinear functions of the structural parameters of the model $\mathbf{\Theta}$, ϵ_t is a $n \times 1$ vector of the structural innovations, s_t is the $n_s \times 1$ vector of endogenous and exogenous states and Σ_ϵ is a covariance matrix of dimension $n \times n$ whose rank is $r = rank(\Sigma_\epsilon) \leq n$. We denote by $\mathrm{N}_r(0, \Sigma_\epsilon)$ the $n \times 1$ multivariate singular Normal distribution with rank r. If the eigenvalues of $\mathrm{A}(\mathbf{\Theta})$ are inside the unit circle, the latter structure can be mapped into a VMA(∞) (see Komunjer and Ng, 2011). 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, as follows

$$y_{j,t} = \Phi_j(\mathrm{L}; \mathbf{\Theta})\epsilon_t + e_{j,t}, \quad \epsilon_t \sim \mathrm{N}(0, \Sigma_\epsilon), \quad e_t \sim \mathrm{N}(0, \Omega), \quad j = 1, \ldots, n_y, \tag{2}$$

where Φ_j correspond to the jth raw of $\Phi(\mathrm{L}; \mathbf{\Theta})$, the MA polynomial of the DSGE model observable counterparts. The structural 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 their MA structure.

Equation (2) can be seen as an approximate dynamic factor model where the row vector $\Phi_j(\mathrm{L}; \mathbf{\Theta})$ 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_ϵ. If we assume Σ_ϵ diagonal, it is sufficient to check that standard deviations are not zero. If Σ_ϵ is a symmetric positive definite matrix, the non-zero eigenvalues correspond to the primal shocks. An alternative way to select the number and combination of primal DSGE shocks is to compute the marginal likelihood of model specifications with different combinations of shocks. However, this can be time consuming because it requires estimating each of the possible models. In models with typically 7 or 8 postulated primal 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.

Following the literature on Bayesian stochastic variable selection in state space models initiated by

[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.

Fruhwirth-Schnatter and Wagner (2010)[6], we rewrite Equations (1) as follows:

$$\tilde{s}_{t+1} = \mathrm{A}(\boldsymbol{\Theta})\tilde{s}_t + \mathrm{B}(\boldsymbol{\Theta})\Sigma^{1/2}\tilde{\epsilon}_{t+1} \qquad \tilde{\epsilon}_{t+1} \sim \mathrm{N}(0, \mathrm{I}_n),$$

where $\mathrm{N}(0, \mathrm{I}_n)$ is the multivariate normal distribution with unitary variance. While the standard deviation of $\tilde{\epsilon}_{t+1}$ 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 $j = 1, \ldots, n_\epsilon$ we assume that $\mathrm{P}(\sigma_j = 0) > 0$. It is important to notice that structural standard deviations are not identified up to sign switch, e.g. $\epsilon_i \sim \mathrm{N}(0, \sigma_i^2) = \pm\sigma_i\tilde{\epsilon} \sim \mathrm{N}(0, 1)$. In other words, the model in the equation with $(-\Sigma^{1/2})(-\tilde{\epsilon}_{t+1})$ is observationally equivalent to the same model with $\Sigma^{1/2}\tilde{\epsilon}_{t+1}$. 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 the existence of the shock. 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.[7] We propose to use the following priors:

1. *Exponential priors*:

$$\sigma_j \sim \mathrm{Exp}(\lambda_j).$$

   With Exp priors, we fix the sign to be non-negative (but allowing for zero) prior to estimation. In order to assess the existence 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*:

$$\sigma_j \sim \mathrm{N}(\mu_j, \tau_j^2).$$

   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 primal 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[8].

[6] In the online appendix (section A.2) we briefly summarize the key ideas of this method and its correspondence with DSGE shock selection.

[7] This formulation is a simplification of the state space form proposed in FS-W that is derived in online appendix (section A.2). The FS-W formulation is slightly more involved since their main goal is to have a closed form expression for the marginal likelihood which is not our focus here. In the results that follow, we also used their formulations and find no difference in the estimation results.

[8] In the online appendix (section A.2) we describe the procedure to introduce the sign switch in an otherwise standard MCMC routine.

One can be interested in situations where structural disturbances are correlated. In such a case, a singular inverse Wishart prior can be imposed on the variance covariance matrix of the shocks. However, being somehow a non-standard practice on DSGE estimation, we leave the description and the derivation of this specific setup to the online appendix (Section A.4).

3 Inference distortions with non existent DSGE shocks

We now tackle the question of whether the introduction of non-existent shocks affects the estimation of parameters governing the impulse transmissions 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. We derive our conclusions based on intuitive arguments and Montecarlo exercises.

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.[9] We obtain the following recursive representation:

$$\begin{aligned} y_t &= z_t + \alpha k_t + e_t \qquad e_t \sim \mathrm{N}(0, \sigma_e^2), \\ z_{t+1} &= \rho z_t + \sigma \epsilon_{t+1} \qquad \epsilon_{t+1} \sim \mathrm{N}(0,1), \\ k_{t+1} &= \alpha k_t + z_t, \end{aligned} \tag{3}$$

where 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 measurement error, $e_t \sim \mathrm{N}(0, \sigma_e^2)$.

We run a controlled experiment to measure the impact of different priors on the shock standard deviations of the DSGE model. We simulate artificial data from the RBC model by calibrating structural parameters values to standard values in the literature, i.e. $\alpha = 0.33$, $\rho = 0.95$. We generate data assuming that the technology shock standard deviation is zero and positive. We fix the variance of σ_e 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.[10]

We generate 500 data points from the RBC model with the measurement error for each value of σ, and retain the last 100 for inference. We compute and estimate the posterior kernel of σ assuming the following priors where m stands for the mean, and SD for the standard deviation:

- IG(α, β) with $\alpha = 2.0016$ and $\beta = 0.2$ that gives $m = 0.2$ and $SD = 5$;

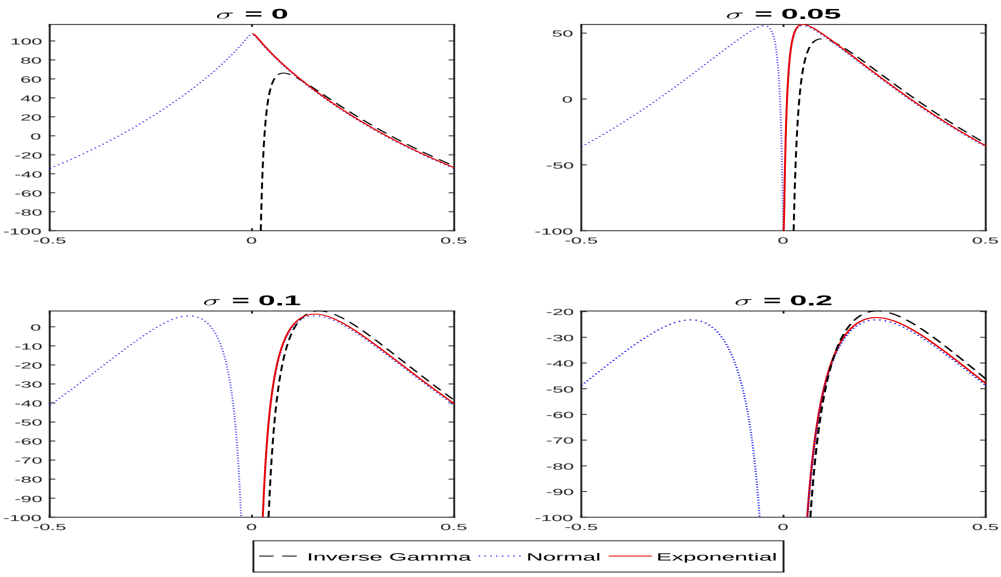
[9] See the online Appendix section A.1 for details.

[10] Results with different parameterization of the data generating process, scale and location parameters and sample size are available on request.

- Normal(μ, σ^2) with $\mu = 0.2$ and $\sigma^2 = 25$ that gives $m = 0.2$ and $SD = 5$;
- Exp(λ) with $\lambda = 0.2$ that gives $m = 5$ and $SD = 5$.

with an abuse of notation we indicate parameter distribution (IG, N and Exp) and in parenthesis mean and standard deviations. While the measures of dispersion are the same across priors, the prior shape and support are different. It is important to notice that, while for the normal and exponential distribution third and higher moment exist, it is not true for the inverse gamma. E.g., with this parametrization, the third and higher moments of the inverse gamma prior do not exist. 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 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 ($[-0.5, 0.5]$) 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. When

Figure 1: *Posterior kernels for different values of σ's. The figure reports the posteriors with: IG(0.2,5) prior (black dashed line); N(0.2, 5) prior (blue dotted line); Exp$(5,5)$ prior (red line). The values considered are $\sigma = \{0, 0.05, 0.10, 0.20\}$.*



the technology shock has zero variance, 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 Exp 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 IG 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 non-zero the posterior kernel of σ is similar across prior assumptions. Therefore, normal or Exp priors do not seem to

Table 1: *Full MCMC estimates of model (3). The model is estimated using the Normal, the IG and the Exp prior for the standard deviations of structural shock. The table reports for all the structural parameters* $\boldsymbol{\Theta}$*: the posterior mode, the lower (5%) and upper (95%) quantile(credible set) and the marginal likelihood (ML) computed with the Laplace approximation and the modified armonic mean (in parenthesis).*

	$\sigma = 0$		
$\boldsymbol{\Theta}$	Mode [Lower,Upper]	Mode [Lower,Upper]	Mode [Lower,Upper]
	IG(**0.05**,5)	Normal(**0.05**,5)	Exp(**5**,5)
α	0.30 [0.21 , 0.37]	0.30 [0.22 , 0.38]	0.30 [0.21 , 0.37]
ρ	0.39 [0.10 , 0.61]	0.50 [0.10 , 0.69]	0.50 [0.09 , 0.68]
σ	0.03 [0.01 , 0.04]	0.00 [-0.05 , 0.04]	0.002 [0.001 , 0.044]
σ_e	0.09 [0.08 , 0.11]	0.09 [0.08 , 0.11]	0.10 [0.08 , 0.10]
ML	88.55 (88.48)	84.47 (84.14)	85.39 (84.38)
	IG(**0.1**,5)	Normal(**0.1**,5)	
α	0.29 [0.21 , 0.37]	0.30 [0.21 , 0.37]	
ρ	0.30 [0.07 , 0.51]	0.50 [0.10 , 0.69]	
σ	0.04 [0.02 , 0.06]	0.00 [-0.04 , 0.04]	—
σ_e	0.09 [0.07 , 0.10]	0.09 [0.08 , 0.11]	
ML	87.36 (87.26)	84.47 (84.16)	
	IG(**0.2**,5)	Normal(**0.2**,5)	
α	0.29 [0.20 , 0.36]	0.29 [0.21 , 0.38]	
ρ	0.20 [0.04 , 0.40]	0.50 [0.10 , 0.69]	
σ	0.06 [0.04 , 0.08]	0.00 [-0.05 , 0.04]	—
σ_e	0.08 [0.06 , 0.10]	0.09 [0.08 , 0.11]	
ML	84.58 (84.49)	84.47 (84.16)	

create distortions when the shock truly exists.

Conclusions are similar when using posterior simulators to approximate the posterior distributions. The posterior moments of the full set of parameters, α, ρ, σ and ω, are computed using the Random Walk Metropolis-Hastings algorithm[11] adapted for the sign switch when assuming normal priors 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 0.5 with 0.2 standard deviation and an IG prior for the measurement error centered in 0.2 with a loose standard deviation of 4.

Table 1 reports posterior statistics assuming different prior distributions and locations for the standard deviations. In particular, from top to bottom, for the inverse gamma prior and the normal, we move the prior location progressively away from zero. We do not report the the results for $\sigma > 0$ since they are similar among priors and close to the true values; they can be consulted on the online appendix (Section B, Additional Tables and Graphs, Table 3).

A few things are worth noting. First, while the estimates of σ are very imprecise with an IG prior,

[11]We launched two chains of 100,000 draws using as a starting point the mode of the posterior kernel.

the Normal and Exp 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 structural disturbance exists.

Second, while invariant for the normal prior, the marginal likelihood estimate depends on the location of the inverse gamma prior. The closer the hyper-parameter controlling the prior mean to zero, the larger the marginal likelihood. Hence, in principle, the marginal likelihood with inverse gamma distribution can be informative about the size of the shock. However, the marginal likelihood with inverse gamma priors is larger than the one obtained in the normal prior setup, which indeed is able to select the true DGP. This feature occurs systematically in short samples and it is not the artifact of one specific realization.[12] While we do not have an analytical explanation for this result, intuitively, it may be related to the third and higher order moments of the probability distribution. For the normal distribution, all moments higher than the first are controlled by the scale parameter (squared variance) only. In the inverse gamma distribution, all moments depend on the shape and scale and, therefore, when changing the prior mean (keeping the dispersion fixed) we are implicitly changing the higher moments of the distribution (if they exist). This makes the comparison across distributions difficult to interpret and care should be taken when expressing a preference for the inverse gamma prior over the normal distribution by means of marginal likelihood.[13] In the empirical section 4 we are going to discuss this issue more extensively.

Third, when the true DGP has $\sigma = 0$ and we impose an inverse gamma prior on σ, we obtain that the autoregressive parameter estimates are heavily downward biased. 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 IG priors we rule out null standard deviations, the posterior kernel of standard deviations does not include zero and, as a consequence, the primal shock has a dynamic impact on y_t. The only way to tone down the dynamic impact of this shock is to force the autoregressive parameter close to zero. To see this, assume that $\alpha = 0$ and $\sigma_e^2 = 1$. Assume that the true DGP is the one with a null standard deviation, i.e. $\sigma = 0$, we have that $y_t = e_t$ and the likelihood collapses to $\log \mathcal{L}(y_T y^{T-1}; \rho, \sigma = 0) \propto -1/2 \sum_{t=1}^{T} y_t^2$. 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

[12] In the online appendix (section A.5), we report the marginal likelihood estimates based on 50 different datasets generated with $\sigma = 0$ assuming either normal or inverse gamma distributions with mean 0.05 and standard deviation 5. On average, there is a 4 to 5 log marginal likelihood difference in favor of the inverse gamma prior over the normal prior. This difference does not vanish with larger samples.

[13] In the online appendix (Section b, Additional Tables and Graphs, Table 2) we provide a decomposition of the (log-) marginal likelihood into a constant, the determinant of the inverse Hessian, the prior, and the likelihood for different locations of the priors. Almost the entire change in the marginal likelihood when using inverse gamma priors with different locations is due to a change in the prior and not in the determinant of the inverse Hessian. Changes in the likelihood are very small.

by:

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

$$z_{t+1|t} = (1-k_t)z_{t|t-1} + k_t y_t, \quad s_{t+1} = k_t + \delta^2, \quad k_t = \rho\frac{s_t}{1+s_t},$$

where the recursions are derived from the Kalman filter with $s_1 = \delta^2/(1-\rho^2)$ and $z_{1|0} = 0$.[14] In order to minimize the information discrepancy between the misspecified model likelihood, i.e. $\mathcal{L}(y_T|y^{T-1};\rho,\sigma=\delta)$, and the true DGP model likelihood, i.e. $\mathcal{L}(y_T|y^{T-1};\rho,\sigma=0)$, 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:

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

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, we expect to see that those parameters will also be affected by the incorrect specification of the number and combination of primal shocks. To quantify these distortions, we consider a baseline DSGE model as presented in Smets and Wouters (2007) (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 an autoregressive process 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 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).[15]

To study the effect of estimating non-existent shocks, we switch off the price markup, the wage markup, and the investment specific shocks and add seven measurement i.i.d. errors (one for each observable) with standard deviation equal 0.08 which we estimate along with the other parameters.[16] With this calibration, measurement errors explain on average less then 3% of the volatility of observables. We simulate 1,000 data

[14]See Durbin and Koopman (2012) for an introduction.

[15]More details on the set of parameter estimates, prior assumptions, parameter description, log-linearized equilibrium conditions and the true values used to generate artificial data, which are taken from the posterior mean estimated in SW, can be found in Appendix A.1.

[16]We also reduced this value to 0.05 and increased it to 0.35 and the main conclusions are unaffected. We discuss a more complete set of robustness checks in section 3.1.

Table 2: *Montecarlo experiment with 100 artificial datasets. The table reports the bias measured as the difference between the average posterior mean (MCMC posterior estimates of the parameter) and the true value. The priors considered are Normal, IG and Exp. The table also reports the true simulated value. ϕ is the inverse of the elasticity of substitution of installed capital, λ habit in consumption, ζ_w wage stickiness, i_w wage indexation, i_p price indexation, ζ_p price stickiness, r_p MP rule response to inflation, r_{dy} MP response to output, r_y MP response to output, ρ interest smoothing, σ_c intertemporal substitution.*

	Structural parameters $\boldsymbol{\Theta}$ Bias										
	ϕ	λ	ζ_w	i_w	i_p	ζ_p	r_p	r_{dy}	r_y	ρ	σ_c
IG Prior	-1.32	-0.06	-0.07	-0.19	-0.17	-0.11	0.19	-0.04	0.02	-0.04	0.10
Normal Prior	0.00	0.00	0.00	0.01	0.01	0.00	-0.06	0.00	0.01	0.00	-0.01
Exp prior	0.01	0.01	0.01	0.02	-0.01	0.00	0.01	0.01	0.02	0.01	0.01
True value	5.74	0.71	0.70	0.59	0.24	0.65	2.05	0.22	0.09	0.81	1.38

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 IG, Exp, and Normal priors for the standard deviations and the same priors as in SW for the remaining parameters. For each of these three specifications, we estimated the standard deviations of measurement errors for which we assumed inverse gamma distributions centered on the true value with a loose precision. 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 1,000 for inference. Converge is checked by means of Brook and Gelman (1998) diagnostics for a subset of estimates [17]

We simulated various samples and estimated the posterior distributions of the parameters using our three different prior distributions, i.e. Normal, Exp and IG. In all samples, normal and exponential priors were able to retrieve primal shocks from non-existent ones. We measured the bias 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 2 (see table 7 in the appendix for the description of the parameters). A positive value means that we are *overestimating* a parameter, and a negative value that we are *underestimating* it. Bar few exceptions, the bias using Normal or Exp priors is negligible as the order of magnitude is small. In all cases, the bias using IG priors is larger than the one using Normal or Exp priors. With IG priors, we obtain sizable distortions in parameters capturing persistence. In particular, price and wage indexation parameters are systematically underestimated.

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 3 reports the variance decomposition of output, inflation, wages, and the interest rate in terms of

[17] In particular, we computed the interval range of the pooled chains and the average interval range within chains and verified when these two lines stabilize and lie one on top of each other. In most of the cases, convergence is achieved within 150,000 draws.

primal shocks under various prior assumptions about their standard deviations. Price and wage markup shocks should not explain fluctuations in any of these variables. This is the case for Normal and Exp priors. It is not the case, however, for IG priors where wage and price markup shocks together explain 16% of the volatility of inflation and 8% of the volatility of wages. Moreover, the transmission of shocks is altered in a

Table 3: *Fraction of the variance of output (y), inflation (π), interest rate (r) and wages (w) explained by the non-existing shocks. In brackets, the 5%-95% confidence sets. The non-existing shocks are: price markup (σ_p), wage markup (σ_w) and investment specific (σ_i) shocks. In the true DGP the explained variances are zero.*

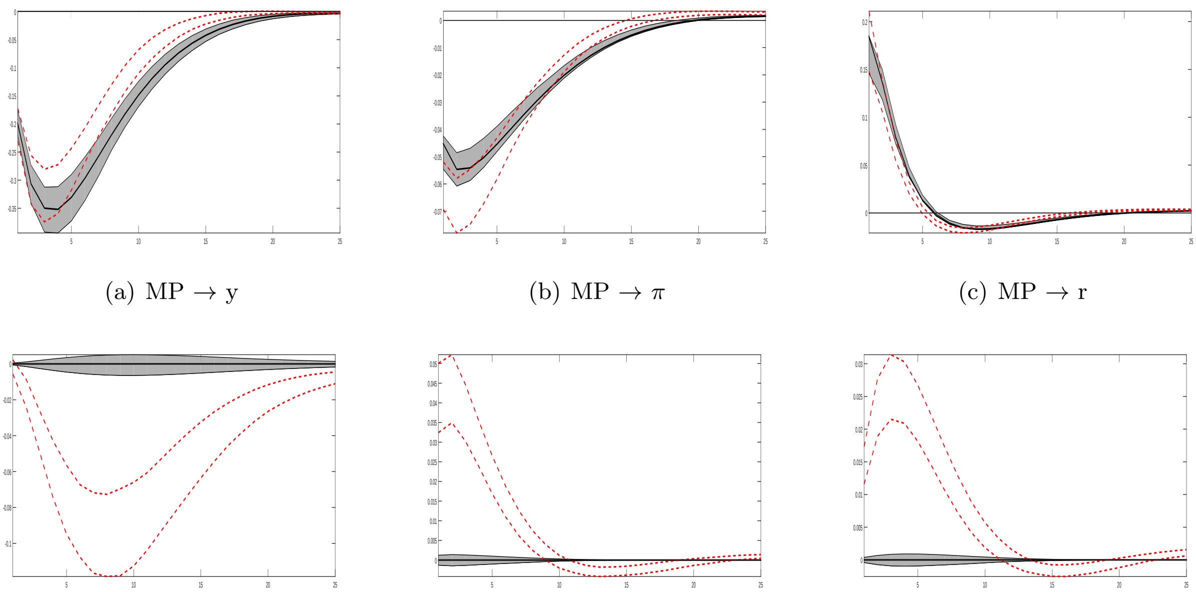
Priors	y	π	r	w
IG prior on σ_i	0 [0,0]	0 [0,0]	0[0,0]	0[0,0]
IG prior on σ_w	1 [0,1]	9[6,14]	2 [1,3]	8[5,13]
IG prior on σ_p	0 [0,0]	7 [5,8]	0 [0,1]	0[0,0]
Normal prior on σ_i	0 [0,0]	0 [0,0]	0 [0,0]	0[0,0]
Normal prior on σ_w	0 [0,0]	0 [0,0]	0 [0,0]	0[0,0]
Normal prior on σ_p	0 [0,0]	0 [0,0]	0 [0,0]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

substantial way. Figure 2 reports the transmission of monetary policy shocks (top row) and wage markup shocks (bottom row) to output, inflation, and the interest rate. Gray areas (red dashed lines) represent the 90% confidence sets of the response assuming Normal (IG) priors on standard deviations and the black solid line the true response. The responses to a monetary policy shock are qualitatively different under the two settings. IG 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. For an interest rate hike of 15-20 basis points, inflation declines more than than it should (black line) and output recovers much faster. In this context, disinflation trajectories might result to be less costly in terms of output loss relative to what they truly are. With normal priors (gray shaded areas) the disinflation trajectories are correctly picked up in terms of size and in terms of speed of adjustment. 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 IG priors, output and inflation react strongly and their responses are statistically and economically significant.

3.1 Further Discussion

The simulation evidence shows that inference and policy conclusions differ substantially when non-existing shocks are forced on the model. Since we do not know ex-ante what are the key shocks driving aggregate

Figure 2: *Impulse response function of: output (y), inflation (π) and interest rate (r) to a monetary policy shock (MP, top row) and to a wage markup shock (Wage MK bottom row) with IG prior (red dashed lines) and Normal prior (gray shaded area). Black line true impulse response function. Results for the Exp prior are not reported for readability reason. They are not different from the Normal prior.*



(a) MP $\rightarrow$ y   (b) MP $\rightarrow$ π (c) MP $\rightarrow$ r

(d) Wage MK $\rightarrow$ y   (e) Wage MK $\rightarrow$ π (f) Wage MK $\rightarrow$ r

fluctuations, IG priors are problematic as they may induce biases in estimated parameters that can be sizable. Normal or Exp priors do not suffer from any particular disadvantage when confronted with data that are generated by an unknown number of primal disturbances. We have tested these priors on different values of the standard deviations of the structural shocks, of the measurement errors (including zeros) and with various sample sizes. Our conclusions are unaffected.

It is important to notice that the autoregressive parameters of non existing shocks are not identifiable. However, since we do not know ex-ante what the non-existing shocks are, we also ignore for which exogenous processes can we identify the persistence coefficient. As a consequence, when estimating the model with the full set of shocks, the likelihood along the dimension of the autoregressive parameter of these non-existing processes is flat and the posterior overlaps with the prior. While this does not cause any problem per se, it makes the estimation inefficient. To cope with this, applied researchers can run a two step estimation procedure. In the first stage, they estimate the full set of exogenous processes' parameters (i.e. standard deviations and autoregressive parameters) and identify the fundamental sources of fluctuations. In the second stage, they can fix the autoregressive parameters of the non-existing shocks obtained in the first stage to avoid the non-identifiability of the persistence coefficient.

A crucial assumption of our approach is that the vector of observed times series is generated by a combination of primal and non-primal (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 primal shocks. However, since primal 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 postulated primal shocks may not be true common factors.

4 Primal innovations and the role of priors for standard deviations

We now study the role played by priors on structural DSGE shocks by reconsidering the estimates of a standard DSGE model using US macroeconomic data. We keep as benchmark the SW model used in the previous section. 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 primal and not just the artifact of avoiding stochastic singularity. Moreover, we try to assess the empirical relevance of the choice of priors for STD for the transmission of DSGE shocks.

While the structural equations of the model are the same as the ones used 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. 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 obtain the real wage variable. The aggregate real variables are expressed in per capita terms using 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. The mapping between observables and model based quantities is accomplished through the following measurement equations

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

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. $e_{x,t}$ are standardized normal i.i.d. measurement error (ME) shocks.

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. Priors for the structural parameters but STD are the same as in SW. We estimate three main specifications (and several variants, see next section):

1. The first specification (IG) coincides with the original SW setup. 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 IG prior with mean 0.1 and standard deviation of 2.
2. In the second specification, we assume an IG prior on both measurement errors and primal shocks. We consider two variants that differ in terms of prior location:
   (a) One variant ($IGIG$), a 'conventional' one, where the mean for the primal shocks is centered in 0.1 and for the measurement error in 0.2.
   (b) One variant ($IGIG_*$) where the prior for measurement error is centered in 0.4 and for primal shocks in 0.2. The latter is chosen based on marginal likelihood considerations, i.e. by selecting the hyper-parameters on a discrete grid of values that maximize the marginal data density.

   To be agnostic on the origin of fluctuations we postulate a loose standard deviation of 10.
3. In the third specification, (NN), we postulate that the standard deviations of primal shocks and measurement errors are normally distributed[18] with a very loose standard deviation of 10 and with location 0.1.[19] So, in principle, we allow for the possibility that some primal and/or measurement shocks are zero.

Table 4 reports the posterior mean and 90% confidence bands for a selection of structural parameters and the STD of shocks.

There are a number of relevant results to highlight. First, government spending (g) shocks and price markup (p) shocks are estimated to be 'non-existent', since the posterior support for their standard deviations includes zero and their posterior distributions are unimodal (not shown here). Technology (a), investment (i), risk-premium (b) or 'liquidity' shocks[20], and monetary policy shocks (r) are instead

[18]Given that the results using Exp and normal priors in the previous section were very similar, we report here the results using normal priors only.
[19]The location parameter for the Normal prior is not very influential. See next section.
[20]As shown in Fisher (2015), the SW risk premium or preference shock can be interpreted as a structural shock to the demand for safe and liquid assets such as short-term Treasury securities.

Table 4: *Posterior median and 90 % confidence bands of a selection of structural parameters in the four specifications: IG, $IGIG_*$, IGIG and NN.*

	NN	$IGIG_*$	$IGIG$	IG
ι_p	0.46 [0.23, 0.71]	0.49 [0.26, 0.72]	0.46 [0.22, 0.69]	0.30 [0.13, 0.48]
ϕ	2.85 [2.00, 4.02]	2.23 [1.18, 3.50]	2.24 [1.14, 3.57]	6.07 [4.35, 7.84]
ξ_w	0.79 [0.70, 0.87]	0.78 [0.68, 0.86]	0.78 [0.69, 0.87]	0.69 [0.57, 0.80]
λ	0.76 [0.64, 0.86]	0.69 [0.55, 0.83]	0.73 [0.60, 0.86]	0.69 [0.62, 0.76]
r_y	0.15 [0.09, 0.20]	0.16 [0.11, 0.22]	0.16 [0.10, 0.21]	0.09 [0.05, 0.12]
ρ_b	0.69 [0.55, 0.82]	0.77 [0.65, 0.87]	0.77 [0.65, 0.87]	0.23 [0.09, 0.38]
ρ_r	0.19 [0.04, 0.37]	0.55 [0.21, 0.91]	0.34 [0.10, 0.90]	0.16 [0.05, 0.27]
σ_a	0.36 [0.31, 0.41]	0.34 [0.28, 0.39]	0.34 [0.30, 0.39]	0.44 [0.40, 0.49]
σ_b	0.11 [0.08, 0.14]	0.10 [0.07, 0.13]	0.09 [0.07, 0.12]	0.24 [0.20, 0.28]
σ_g	-0.01 [-0.21, 0.22]	0.13 [0.07, 0.22]	0.08 [0.04, 0.18]	0.58 [0.53, 0.63]
σ_i	0.22 [0.11, 0.34]	0.17 [0.08, 0.31]	0.11 [0.04, 0.29]	0.40 [0.33, 0.49]
σ_r	0.21 [0.11, 0.25]	0.13 [0.07, 0.19]	0.15 [0.06, 0.22]	0.24 [0.22, 0.27]
σ_p	-0.04 [-0.17, 0.16]	0.10 [0.06, 0.15]	0.08 [0.03, 0.16]	0.13 [0.10, 0.15]
σ_w	0.03 [0.01, 0.10]	0.08 [0.05, 0.13]	0.05 [0.03, 0.09]	0.26 [0.22, 0.30]

estimated to be primal. The wage markup shock, however, is only marginally so.[21] For the $IGIG$ specifications, the estimated STD of government spending and price markup shocks are smaller than in the case with no measurement error (IG) yet still significantly different from zero.[22]

Second, the posterior moments of deep parameters appear to be estimated differently in the three setups. For the estimates of ϕ, the second derivative of the investment adjustment cost function, the mean estimate of NN is significantly smaller than the value in IG (6.07) and larger than the values obtained with the $IGIG$ specifications. This figure is closer to the value (2.48) available in Christiano et al. (2005).[23] 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, and they do not need to specify the whole shock structure. Moreover, the price indexation parameter, ι_p, is 0.30 for the IG model and the NN estimated parameter is substantially larger. Albeit to a smaller degree, similar conclusions apply for the estimate of ξ_w and λ, the wage stickiness and habit parameter respectively. Overall, estimates of deep parameters change when we introduce measurement error shocks

[21]If the model is estimated for rolling sub-samples, the shock appears not to be primal for the majority of the sub-samples. Results available on request.

[22]It is important to note that the fact that, for instance, shocks to government spending are found not to be primal in the NN specification does not imply that government spending is fixed. What this means is that these shocks, within the context of the SW model, are not significant drivers of macroeconomic uncertainty. This could be the case if, for instance, these changes to g were fully expected by agents.

[23]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".

and allow for the possibility of zero STD.

Structural parameter estimates of the specifications where we assume IG on both measurement error and primal shocks are similar to the NN ones, with one important exception. The estimated persistence of the monetary policy shock in the two $IGIG$ specifications are different from the NN (and the IG) specification. From an economic standpoint, this parameter captures the extent to which current monetary policy surprises are predictable using past shocks; small figures are desirable. We find that the persistence of the monetary policy surprises is estimated to be 0.55 in the $IGIG_*$ model and 0.34 in the $IGIG$ model as opposed to values below 0.2 in the NN or IG cases. Besides the fact that the former values have weaker economic interpretation, this parameter is crucial for the dynamic transmission of monetary policy shocks as we will review in the next section.

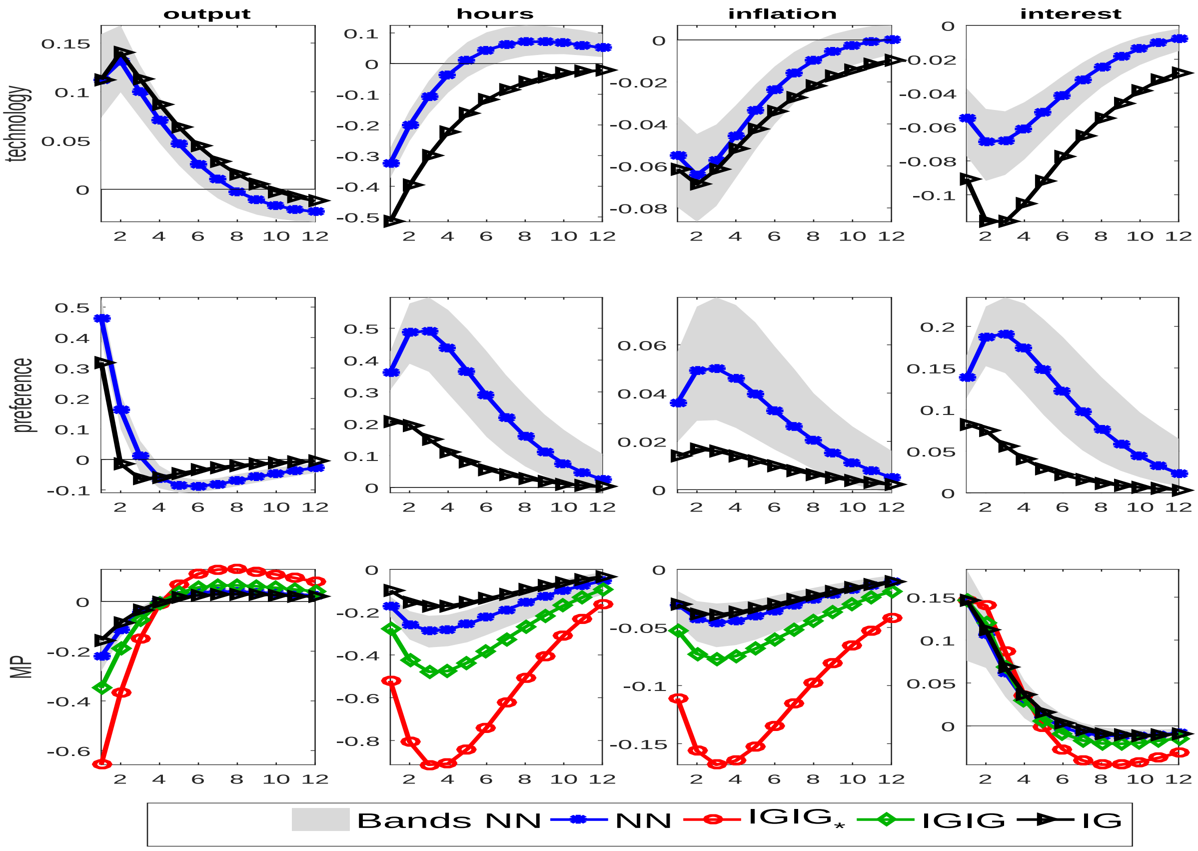
In sum, normal priors identify as common sources of fluctuations technology shocks, monetary policy surprises, risk-premium/liquidity shocks, investment demand shocks, and possibly wage markup shocks. Accordingly, many structural parameters estimates change. This is due partly to the introduction of measurement errors and partly to the prior assumptions about the origin of fluctuations. We investigate this further below.

4.1 The dynamic transmission of the primal sources of fluctuations

To gauge the extent to which the assumptions about the existence of shocks affect the transmission dynamics of the model, we study the impulse response functions of four variables of interest to the identified primal source of fluctuations. We focus our attention on the response of output growth, inflation, interest rate, and hours worked to a supply shock (technology), to a demand shock (risk-premium/liquidity shock), and to a monetary policy surprise. Figure 3 reports the mean estimated impulse response functions under different specifications. The benchmark SW case (IG) is reported with black triangles, the NN specification with blue dots (with 90% gray confidence bands) and the $IGIG_*$ and $IGIG$ specifications in red circles and green diamonds respectively. We have normalized the impact of the shocks across specifications. In particular, a monetary policy impulse generates the same impact increase in the interest rate, a supply impulse the same impact increase in output and a demand impulse the same impact increase in consumption (not shown here). We do not report the response function to demand and supply shocks for the $IGIG$ specifications because they are very similar to the NN one.

There are two important remarks arising from this picture. First, the dynamic transmissions of demand and supply shocks are statistically different between the original SW setup (IG) and the specification where we are agnostic about the origins of fluctuations (NN). While supply shocks generate a negative co-movement between hours and output in both settings, the magnitude is smaller and less persistent in the NN case (see the first row). After a technology shock, in the original SW setup, hours worked need 12 quarters to revert back to steady state, while less then 5 in the NN setup. Demand shocks (second row) generate impulse response functions that are hump-shaped and more persistent with normal priors,

Figure 3: *Transmission mechanism of primal shocks. From top to bottom, dynamic transmission of supply (technology), demand (liquidity preference) and monetary policy shocks to output, hours, inflation and interest rate (left to right). Blue dots NN median estimates and gray areas the 90% confidence sets. Black triangles represent the estimates with IG, green diamonds $IGIG$ and red circles $IGIG^*$.*



with a peak delay of 2/3 quarters.

Second, the prior assumptions about shock standard deviations matter for the estimated transmission dynamics of the model. The third row of figure 3 reports the transmission mechanism of monetary policy surprises under the four different settings. In the original SW setup and the specification where we are agnostic about the sources of fluctuations (NN), an unexpected increase in the federal fund rate of 15 basis points generates a decline in inflation of less then 5 basis points at the peak of the effect and a decline in output of less then 20 basis points. These figures are very close to those obtained in empirical studies using VARs (see Ramey, 2016).

In the setup with inverse gamma priors on structural and measurement errors, the same increase generates a huge response in inflation and output, between twice to more than three times larger.[24] Why

[24]It is also noticeable the large difference between $IGIG$ and $IGIG_*$, which indicates that the choice of prior location matters for these models. This is investigated further in section 4.2 below.

does this difference arise? In the $IGIG$ settings we are forcing all the shocks to exist, both primal and measurement errors. When this happens, the 'residuals' in the interest rate measurement equation are forced to be explained by both a monetary policy shocks and by the interest rate measurement errors. With a small (0.16) ρ_r monetary policy shocks and measurement error are difficult to separately identify as they are both close to be i.i.d. So whenever the likelihood moves toward regions where the measurement error STD is large, it prefers to assign different persistence to these two shocks in order to be able to distinguish them. Since it cannot do this for the measurement error, it does it for the monetary policy shock. Hence the large ρ_r, which generates a statistically large and economically implausible transmission mechanism. Unfortunately, being highly misspecified, this class of models typically enjoy the presence of measurement errors, and large measurement error shocks tend to be associated with high values of the marginal likelihood. As it will become apparent in the next sections, this feature is amplified further by the inverse gamma prior distribution, for which the marginal likelihood favors $IGIG$ variants with large prior mean for the STD of the measurement errors. Hence, when choosing the $IGIG$ specification with the highest marginal likelihood ($IGIG_*$), a researcher would be lead even more astray from a plausible representation of a MP shock.

In all, the choice of the priors for STD matters in practice. When identifying the primal sources of fluctuations and when studying the transmission mechanism of primal shocks. It is in this sense that we uncover a trade-off between imposing a wide set of potential sources of uncertainty and the estimation of the parameters that drive propagation.

4.2 The sensitivity of the posterior to the standard deviation prior location

A noteworthy drawback of using inverse gamma priors on the STD of shocks concerns the sensitivity of the posterior estimates to the prior location. On the contrary, the posterior analysis is invariant to the prior location when assuming normal priors. In other words, even if the parameter estimates driving some impulse response of interest were not too different between inverse gamma and normal priors, posterior estimates are affected by the inverse gamma prior location. This is an added advantage of using normal priors for STDs.

When we consider the SW model with the same the number of shocks and of observables (i.e. 7), the prior distribution of the shocks STD is largely uninfluential for the posterior parameter estimates[25]. All the shocks are needed to explain the stochastic dimension of the observables. And, regardless of the prior assumptions on the STD of the shocks, we obtain statistically indistinguishable posterior estimates. However, in the inverse gamma prior setting, the marginal likelihood varies substantially depending on the prior location, even if the posterior parameter estimates do not. In this context, marginal likelihood comparisons with inverse gamma priors are difficult to interpret.

When we have more shocks than observables (i.e. with measurement errors), the likelihood has some-

[25] Assuming 'reasonable' locations and dispersions.

thing to say about the most likely combination of shocks, i.e. the model specification. As a consequence, it is not longer true that the STD prior distribution does not matter for the posterior analysis. The prior density could favor or penalize the likelihood of a specific model shock configuration, depending on the prior dispersion and location. As a result, it might influence posterior analysis. While this not wrong *per-se*, one should be aware of the extent to which the priors on STDs influence posterior analysis.

Contrary to the inverse gamma case, the normal prior on STD is largely uninfluential for the computation of the posterior parameter distributions and of the marginal likelihood (as long as we have a sufficiently loose precision). It does not really matter if we postulate a priori that the STD of the measurement error on - say - interest rate is close to zero or not when we assume normal priors. The NN setting generates estimates of the posterior distributions and of the marginal data density that are invariant to the prior mean. This appealing property of the normal prior on shocks STD does not carry over for the inverse gamma distribution, where the posterior parameter and the marginal likelihood estimates are very sensitive to the inverse gamma location parameter. In such a case, different configurations of prior means lead to different estimates of the parameters and of the marginal likelihood. This brings us back to square one, as we need a device to select among the prior hyper-parameter.

To show these two points empirically, we run the following exercises. We first considered the case with 7 primal shocks and seven observable variables and no measurement errors. We specified various locations for the prior mean of the STD under the normal and inverse gamma settings. In particular, we assumed that the prior mean for the STD lies in this discrete range of values $[0.08, 0.1, 0.2, 0.4, 0.8, 1, 1.2, 1.5]$ for both the normal prior and the inverse gamma case, and in both setups, we assumed a loose standard deviation equal to 10. For each of these prior specifications (eight for each prior distribution), we estimated the mode of the posterior kernel and computed the log marginal likelihood with the Laplace approximation. Regardless of the prior location, the estimated posterior mode is the same in the normal and inverse gamma setup[26]. Table 5 reports the log marginal likelihood of this exercise and its components. While there are no variations in the marginal likelihood in the normal prior case, there are significant differences in terms of marginal likelihood when inverse gamma priors are used. In other words, since the likelihood for the case with, say, $IG(0.1, 10)$ is the same as the one with $IG(0.4, 10)$, the 13 Log ML difference reported is entirely coming form the prior density rather than from an improvement in the Kalman filter one-step ahead prediction error. The same argument applies when contrasting the marginal likelihood of the N and IG setup. In this sense, marginal likelihood comparisons do not appear to be a useful tool to compare model fit.

When we also introduce measurement errors, the likelihood can express a preference between model specifications (i.e. shocks configuration) which can be favored or penalized by the prior. To assess the extent to which this occurs, we considered different prior locations for measurement errors and for primal shocks, varying on the same discrete range of values. This generated 64 different variants of priors for

[26] To save space, we do not show the results here. They can be consulted on the online appendix in section B, Figure 3-5.

Table 5: *Log marginal likelihood (Laplace approximation) of the SW model (no measurement errors) and its components, under different prior locations for the normal (N) and inverse gamma (IG) case.*

	Prior Mean							
	0.08	0.1	0.2	0.4	0.8	1	1.2	1.5
	Normal prior							
Laplace Approximation	-916.7	-916.7	-916.7	-916.7	-916.7	-916.7	-916.7	-916.7
	Inverse Gamma prior							
Laplace Approximation	-907.6	-904.6	-896.1	-891.1	-898.1	-906.4	-916.9	-935.8
(log) likelihood	-801.1	-801.1	-801.2	-801.4	-802.8	-804.2	-806.4	-811.1
(log) prior	-28.1	-25.1	-16.6	-11.3	-17.0	-23.8	-32.2	-46.6
(log) constant	32.2	32.2	32.2	32.2	32.2	32.2	32.2	32.2
(log) det inverse Hessian	-110.5	-110.5	-110.5	-110.5	-110.6	-110.6	-110.5	-110.3

Table 6: *Log marginal likelihood (Laplace approximation) of the SW model with measurement errors, under different prior locations for the normal (NN) and inverse gamma (IGIG) case.*

	$IGIG$ Prior mean							
STR/ME	0.08	0.1	0.2	0.4	0.8	1	1.2	1.5
0.08	-863	-861	-854	-851	-860	-868	-879	-898
0.1	-862	-869	-852	-850	-859	-868	-879	-897
0.2	-861	-858	-849	-848	-860	-870	-881	-900
0.4	-859	-857	-853	-854	-869	-879	-892	-912
0.8	-877	-876	-877	-877	-894	-907	-921	-943
1	-888	-887	-887	-888	-909	-922	-937	-960
1.2	-900	-900	-900	-904	-925	-939	-953	-976
1.5	-921	-920	-922	-927	-952	-966	-981	-1003

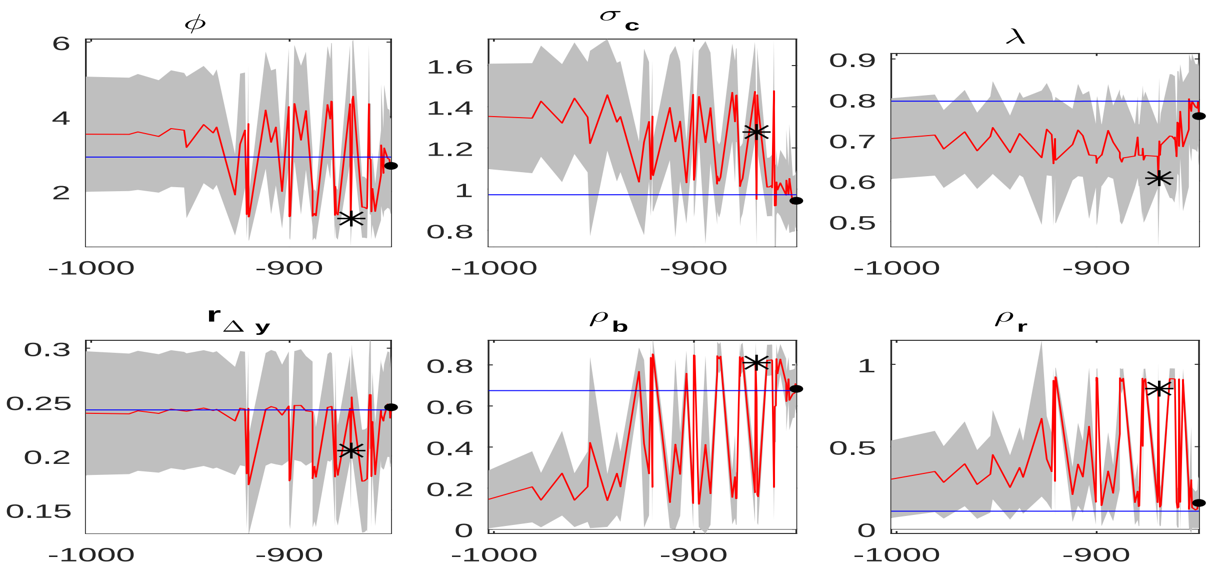
NN
Prior location $me, str = 0, 0.1, ..., 1.5$
-893

each prior distribution assumption, the NN and the $IGIG$. For each of them, we estimated the mode of the posterior kernel and computed the log marginal likelihood with the Laplace approximation.

Table 6 reports the log marginal likelihood of the two specifications, $IGIG$ and NN, under different prior locations. As before, in the NN case with a sufficiently loose prior, the log marginal likelihood (ML) is invariant to the prior mean, meaning that normal priors do not favor nor penalize the likelihood. In the $IGIG$ case there are large swings in the ML, which are the result of the extent to which priors and likelihood are in accordance. Under this prior parametrization, we find a maximum difference of 140 ML points. Even if we focus on more 'reasonable' prior locations, say between 0.1 and 0.2, we still find large changes in the ML. For instance, a variant with $IG(0.1, 10)$ on both measurement and primal shocks has a ML of -869 and a variant with $IG(0.2, 10)$ on both measurement and primal shocks has a ML of -849. The difference is still substantial and, from a Bayesian model selection perspective, it would justify selecting the latter prior location over the former.

In this case, the large swings in marginal likelihood are associated with large variations in parameter estimates. While they turn out to be invariant for the NN case, they are unstable in the $IGIG$ one. Figure 4 reports the mode estimate (red) and the uncertainty (gray) around it for the $IGIG$ variants against the marginal likelihoods so that the mode estimates are ordered from the lowest to the highest marginal likelihood values. We also plot the mode estimates (blue) of the NN variants. Interestingly, even at the left end of the plots (where marginal likelihoods are substantially similar), some parameters exhibit large swings in the mode point estimate, e.g. ϕ, ρ_b and ρ_r. Even if one believes that the point estimate

Figure 4: *Estimated mode selected parameters for NN (blue) and $IGIG$ (red) under different prior locations. Gray bands report the 90% confidence around the mode of specification $IGIG$ (red). On the horizontal axis the marginal likelihood of the $IGIG$ reported in table 6. Black dots represent the estimated variant that maximizes the ML. The asterisk the* $IGIG(0.1, 10)$



analysis does not portray the full picture and wishes to simulate the full posterior distribution, some issues about using the inverse gamma prior remain. This is particularly the case for the autoregressive parameter of the monetary policy shock.[27] And as highlighted in the previous section, this parameter is crucial in order to obtain sensible estimates of the transmission of monetary policy shocks.[28]

One could argue that a finer grid of prior mean values tailored to each individual (primal or mea-

[27]In the online appendix (section A.6), we report a detailed posterior analysis for this parameter, e.g. convergence of the chains and shape of the posterior distribution. The main finding is that with inverse gamma prior on both measurement errors and structural shocks, the posterior distribution of ρ_r is bimodal.

[28]Note that in figure 4 we report, with a black dot, the mode of the estimate of parameters for the $IGIG$ variant that maximizes the marginal likelihood. The mode for ρ_r differs substantially from the mean reported in table 4 because, as mentioned, the posterior is bimodal. As can be seen in figure 2 in the online appendix, the chain flips between 0.2 and 0.8.

surement error) shock might lead to select the vector of shock-specific hyper-parameters that maximizes the marginal likelihood and generates meaningful transmission mechanisms. This could, in principle, be the case. However, we can see two major drawbacks with this approach. First, practical. It would blow up the computing time for estimation. Second, methodological. As Bayesian econometricians, we want to put a reasonable 'distance' between our prior assumptions and the information contained in the data. Such an approach would inevitably blur this distinction.

In all, when we have the same number of shocks as observables, normal and inverse gamma priors deliver the same posterior estimates of the parameters regardless of the prior location. However, marginal likelihood comparisons are difficult to interpret as the marginal likelihood estimates using inverse gamma priors are not invariant to the prior location. When there are more shocks than observables, inverse gamma priors offer a poor platform to generate posterior estimates that are independent of the prior locations. Moreover, it is not straightforward how to select amongst them. Normal priors overcome all these concerns. They offer a direct way of selecting the fundamental drivers of economic fluctuations and are insensitive to the choice of prior locations.

5 Conclusions

One of the key questions in macroeconomics concerns the identification of the primal 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 *do exist* in the sense that they capture macroeconomic uncertainty. We offer reasons for questioning this assumption and for adopting priors that allow us to test, rather than assume, the existence of primal (and non-primal) shocks.

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. Thus, we unveil a tradeoff between the inclusion of a potentially large number of primal innovations and estimates of the parameters driving propagation. To prevent this problem, we propose an easily implementable strategy of using normal (or exponential) priors on the standard deviation of shocks together with measurement errors to avoid stochastic singularity. Our simulation evidence shows that these priors allow us to select the true primal shocks entering the DSGE model and that the remaining parameters are estimated with precision.

We analyzed the evidence on the existence of primal shocks in the medium-scale New Keynesian model of Smets and Wouters (2007). We find that government spending and price markup shocks are innovations that do not generate statistically significant dynamics, with the wage markup shock being only marginally significant. Hence, they are not primal sources of macroeconomic fluctuations in the context of that model. Technology, investment, risk-premium, and monetary policy surprises are all found to be important for the sample and the set of observables considered. By allowing or forcing all these (and measurement errors) shocks to play a role, substantial differences emerge in terms of the estimated parameters and of transmission mechanism of primal shocks.

References

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

Canova, F. (2014). Bridging DSGE Model and Raw Data. *Journal of Monetary Economics*, 57:132–145.

Canova, F., Ferroni, F., and Matthes, C. (2015). Approximating Time Varying Structural Model with Time Invariant Structures. CEPR Discussion Papers N. 10803.

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

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*, 113:1–45.

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 Misspecified Models. *American Economic Review*, 99:1415–1450.

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*, 25:2018–2033.

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*, 136:183–192.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Stace Methods*. Oxford University Press, second edition.

Eisenstat, E., Chan, J. C., and Strachan, R. W. (2016). Stochastic Model Specification Search for Time-Varying Parameter VARs. *Econometrics Review*, 35:1638–1665.

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

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

Fisher, J. D. (2015). On the Structural Interpretation of the Smets-Wouters 'Risk Premium' Shock. *Journal of Money, Credit and Banking*, 47:511–516.

Fruhwirth-Schnatter, S. and Wagner, H. (2010). Stochastic Model Specification Search for Gaussian and Partial non-Gaussian State Space Models. *Journal of Econometrics*, 154:85–100.

Galvao, A., Giraitis, L., Kapetanios, G., and Petrova, K. (2015). Local Bayesian Estimation and Forecasting with a TVP Model. Mimeo, Mimeo.

Gorodnichenko, Y. and Ng, S. (2010). Estimation of DSGE Models when the Data are Persistent. *Journal of Monetary Economics*, 57:325–340.

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*, 28:1205–1226.

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

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

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

Lippi, M. and Reichlin, L. (1994). VAR Analysis, Nonfundamental Representations, Blaschke Matrices. *Journal of Econometrics*, 63:307–325.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. *Handbook of Macroeconomics*, 2:71 – 162.

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. In *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*, 97:586–606.

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

A Appendix

A.1 Smets and Wouters (2007) model

The log linearized equilibrium conditions are summarized as follows:

$$
\begin{aligned}
y_t &= c/y c_t + i/y i_t + r^k k/y z_t + e_t^g \\
c_t &= c_1 c_{t-1} + (1-c_1)Ec_{t+1} + c_2(h_t - Eh_{t+1}) - c_3(r_t - E\pi_{t+1} + e_t^b) \\
i_t &= i_1 i_{t-1} + (1-i_1)E_t i_{t+1} + i_2 q_t + \epsilon_t^i \\
q_t &= q_1 E q_{t+1} + (1-q_1)Er_{t+1}^k - (r_t - E\pi_{t+1} + e_t^b) \\
y_t &= \alpha\phi_p k_t + (1-\alpha)\phi_p h_t + \phi_p \epsilon_t^a \\
k_t^s &= k_{t-1} + z_t \\
z_t &= \psi/(1-\psi) r_t^k \\
k_t &= k_1 k_{t-1} + (1-k_1) i_t + k_2 \epsilon_t^i \\
mp_t &= \alpha(k_t^s - h_t) + e_t^a - w_t \\
\pi_t &= \pi_1 \pi_{t-1} + \pi_2 E\pi_{t+1} - \pi_3 mp_t + e_t^p \\
r_t^k &= -(k_t - h_t) + w_t \\
mw_t &= w_t - \left(\sigma_n h_t + \frac{1}{1+\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})\right) \\
w_t &= w_1 w_{t-1} + (1-w_1)E(\pi_{t+1} + w_{t+1}) - w_2 \pi_t + w_3 \pi_{t-1} + mw_t + e_t^w \\
R_t &= \rho_R R_{t-1} + (1-\rho_R)(\rho_\pi \pi_t + \rho_y(y_t - y_t^f) + \rho_{\Delta y}\Delta(y_t - y_t^f)) + e_t^r \\
&\quad + \ \textit{flexible economy equations}
\end{aligned}
$$

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.

$$
\begin{aligned}
&c_1 = \lambda/\gamma(1+\lambda/\gamma), \quad c_2 = [(\sigma_c - 1)(W^h h/C)]/[\sigma_c(1+\lambda/\gamma)], \quad c_3 = (1-\lambda/\gamma)/(1+\lambda/\gamma)\sigma_c, \quad k_1 = (1-\delta)/\gamma \\
&k_2 = (1-(1-\delta)/\gamma)(1+\beta\gamma^{1-\sigma_c})\gamma^2\phi, \quad i_1 = 1/(1+\beta\gamma^{1\sigma_c}), \quad i_2 = (1/(1+\beta\gamma^{1-\sigma_c})\gamma^2\phi \quad q_1 = \beta\gamma^{-\sigma_c}(1-\delta), \\
&\pi_1 = i_p/(1+\beta\gamma^{1-\sigma_c} i_p), \quad \pi_2 = \beta\gamma^{1-\sigma_c}/(1+\beta\gamma^{1-\sigma_c} i_p), \\
&\pi_3 = 1/(1+\beta\gamma^{1-\sigma_c} i_p)[(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\xi_p(1+(\phi_p - 1)\epsilon_p)))], \quad w_1 = 1/(1+\beta\gamma^{1-\sigma_c}), \\
&w_2 = (1+\beta\gamma^{1-\sigma_c} i_w)/(1+\beta\gamma^{1-\sigma_c}), \quad w_3 = i_w/(1+\beta\gamma^{1-\sigma_c}), \\
&w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/\xi_w(1+(\phi_w - 1)\epsilon_w)], \\
&\overline{\gamma} = 100(\gamma - 1), \quad \overline{\pi} = 100(\pi_* - 1), \quad \overline{\beta} = ((\pi_*/(\beta * \gamma^{\sigma_c})) - 1) * 100
\end{aligned}
$$

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

Table 7: *The table reports the parameter notation (first column), the parameter description (second column) and posterior mean estimated using MCMC or the parameters that are fixed (third column).*

$\mathbf{\Theta}$	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 8: *Posterior median and 90 % confidence bands of the three specifications, IG, IGIG and NN under different prior locations.*

	NN	$IGIG_*$	$IGIG$	IG
ι_w	0.55 [0.31, 0.77]	0.62 [0.39, 0.83]	0.58 [0.35, 0.81]	0.62 [0.41, 0.82]
ι_p	0.46 [0.23, 0.71]	0.49 [0.26, 0.72]	0.46 [0.22, 0.69]	0.30 [0.13, 0.48]
ψ	0.52 [0.33, 0.72]	0.51 [0.31, 0.71]	0.52 [0.32, 0.72]	0.61 [0.44, 0.77]
ϕ_p	1.43 [1.30, 1.57]	1.42 [1.29, 1.56]	1.41 [1.28, 1.55]	1.71 [1.58, 1.84]
α	0.22 [0.18, 0.27]	0.23 [0.18, 0.27]	0.23 [0.19, 0.27]	0.20 [0.16, 0.23]
ϕ	2.85 [2.00, 4.02]	2.23 [1.18, 3.50]	2.24 [1.14, 3.57]	6.07 [4.35, 7.84]
σ_c	1.03 [0.87, 1.23]	1.05 [0.88, 1.25]	1.03 [0.89, 1.19]	1.49 [1.24, 1.75]
σ_l	1.48 [0.38, 2.50]	1.46 [0.33, 2.48]	1.35 [0.25, 2.30]	1.75 [0.84, 2.71]
ξ_w	0.79 [0.70, 0.87]	0.78 [0.68, 0.86]	0.78 [0.69, 0.87]	0.69 [0.57, 0.80]
ξ_p	0.61 [0.52, 0.69]	0.60 [0.52, 0.68]	0.61 [0.52, 0.69]	0.67 [0.58, 0.75]
λ	0.76 [0.64, 0.86]	0.69 [0.55, 0.83]	0.73 [0.60, 0.86]	0.69 [0.62, 0.76]
r_π	1.97 [1.69, 2.27]	2.05 [1.78, 2.35]	2.03 [1.74, 2.33]	2.02 [1.73, 2.32]
r_R	0.83 [0.78, 0.88]	0.82 [0.72, 0.88]	0.83 [0.74, 0.89]	0.81 [0.76, 0.85]
r_y	0.15 [0.09, 0.20]	0.16 [0.11, 0.22]	0.16 [0.10, 0.21]	0.09 [0.05, 0.12]
$r_{\Delta y}$	0.25 [0.20, 0.29]	0.24 [0.19, 0.29]	0.23 [0.18, 0.28]	0.22 [0.18, 0.27]
$\overline{\gamma}$	0.45 [0.40, 0.49]	0.44 [0.40, 0.49]	0.45 [0.40, 0.49]	0.43 [0.40, 0.45]
$\overline{\pi}$	0.67 [0.51, 0.84]	0.68 [0.51, 0.85]	0.67 [0.51, 0.84]	0.83 [0.66, 1.01]
$\overline{\beta}$	0.22 [0.11, 0.34]	0.22 [0.10, 0.34]	0.22 [0.11, 0.35]	0.17 [0.08, 0.28]
$\bar{l}$	0.23 [-0.91, 1.42]	0.08 [-1.11, 1.32]	0.14 [-1.05, 1.40]	0.68 [-1.13, 2.53]
ρ_a	0.94 [0.89, 0.99]	0.94 [0.90, 0.98]	0.94 [0.89, 0.99]	0.95 [0.92, 0.97]
ρ_b	0.69 [0.55, 0.82]	0.77 [0.65, 0.87]	0.77 [0.65, 0.87]	0.23 [0.09, 0.38]
ρ_g	0.48 [0.15, 0.80]	0.45 [0.13, 0.78]	0.47 [0.14, 0.80]	0.98 [0.96, 0.99]
ρ_i	0.74 [0.51, 0.97]	0.77 [0.47, 0.98]	0.74 [0.36, 0.98]	0.72 [0.62, 0.82]
ρ_r	0.19 [0.04, 0.37]	0.55 [0.21, 0.91]	0.34 [0.10, 0.90]	0.16 [0.05, 0.27]
ρ_p	0.37 [0.11, 0.65]	0.35 [0.09, 0.62]	0.39 [0.11, 0.69]	0.87 [0.79, 0.95]
ρ_w	0.98 [0.95, 1.00]	0.97 [0.94, 1.00]	0.97 [0.94, 1.00]	0.96 [0.93, 0.98]
μ_p	0.60 [0.30, 0.88]	0.62 [0.32, 0.89]	0.58 [0.26, 0.88]	0.69 [0.51, 0.84]
μ_w	0.66 [0.32, 0.92]	0.83 [0.69, 0.95]	0.74 [0.51, 0.92]	0.82 [0.71, 0.92]
σ_a	0.35 [0.31, 0.41]	0.34 [0.29, 0.39]	0.35 [0.31, 0.40]	0.44 [0.40, 0.49]
σ_b	0.11 [0.08, 0.15]	0.10 [0.07, 0.13]	0.10 [0.07, 0.14]	0.24 [0.20, 0.28]
σ_g	0.02 [-0.20, 0.22]	0.14 [0.07, 0.22]	0.09 [0.04, 0.19]	0.58 [0.53, 0.63]
σ_i	0.22 [0.09, 0.34]	0.17 [0.08, 0.30]	0.11 [0.04, 0.27]	0.41 [0.34, 0.50]
σ_r	0.21 [0.10, 0.25]	0.12 [0.08, 0.19]	0.21 [0.14, 0.24]	0.24 [0.22, 0.27]
σ_p	0.00 [-0.17, 0.16]	0.10 [0.06, 0.15]	0.11 [0.04, 0.18]	0.13 [0.10, 0.15]
σ_w	0.02 [-0.07, 0.09]	0.08 [0.05, 0.12]	0.05 [0.03, 0.11]	0.26 [0.22, 0.30]

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

The Effects of the Massachusetts Health Reform on Financial Distress **WP-14-01**
Bhashkar Mazumder and Sarah Miller

Can Intangible Capital Explain Cyclical Movements in the Labor Wedge? **WP-14-02**
François Gourio and Leena Rudanko

Early Public Banks **WP-14-03**
William Roberds and François R. Velde

Mandatory Disclosure and Financial Contagion **WP-14-04**
Fernando Alvarez and Gadi Barlevy

The Stock of External Sovereign Debt: Can We Take the Data at 'Face Value'? **WP-14-05**
Daniel A. Dias, Christine Richmond, and Mark L. J. Wright

Interpreting the *Pari Passu* Clause in Sovereign Bond Contracts:
It's All Hebrew (and Aramaic) to Me **WP-14-06**
Mark L. J. Wright

AIG in Hindsight **WP-14-07**
Robert McDonald and Anna Paulson

On the Structural Interpretation of the Smets-Wouters "Risk Premium" Shock **WP-14-08**
Jonas D.M. Fisher

Human Capital Risk, Contract Enforcement, and the Macroeconomy **WP-14-09**
Tom Krebs, Moritz Kuhn, and Mark L. J. Wright

Adverse Selection, Risk Sharing and Business Cycles **WP-14-10**
Marcelo Veracierto

Core and 'Crust': Consumer Prices and the Term Structure of Interest Rates **WP-14-11**
Andrea Ajello, Luca Benzoni, and Olena Chyruk

The Evolution of Comparative Advantage: Measurement and Implications **WP-14-12**
Andrei A. Levchenko and Jing Zhang

Saving Europe?: The Unpleasant Arithmetic of Fiscal Austerity in Integrated Economies **WP-14-13**
Enrique G. Mendoza, Linda L. Tesar, and Jing Zhang

Liquidity Traps and Monetary Policy: Managing a Credit Crunch **WP-14-14**
Francisco Buera and Juan Pablo Nicolini

Working Paper Series *(continued)*

Quantitative Easing in Joseph's Egypt with Keynesian Producers **WP-14-15**
Jeffrey R. Campbell

Constrained Discretion and Central Bank Transparency **WP-14-16**
Francesco Bianchi and Leonardo Melosi

Escaping the Great Recession **WP-14-17**
Francesco Bianchi and Leonardo Melosi

More on Middlemen: Equilibrium Entry and Efficiency in Intermediated Markets **WP-14-18**
Ed Nosal, Yuet-Yee Wong, and Randall Wright

Preventing Bank Runs **WP-14-19**
David Andolfatto, Ed Nosal, and Bruno Sultanum

The Impact of Chicago's Small High School Initiative **WP-14-20**
Lisa Barrow, Diane Whitmore Schanzenbach, and Amy Claessens

Credit Supply and the Housing Boom **WP-14-21**
Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti

The Effect of Vehicle Fuel Economy Standards on Technology Adoption **WP-14-22**
Thomas Klier and Joshua Linn

What Drives Bank Funding Spreads? **WP-14-23**
Thomas B. King and Kurt F. Lewis

Inflation Uncertainty and Disagreement in Bond Risk Premia **WP-14-24**
Stefania D'Amico and Athanasios Orphanides

Access to Refinancing and Mortgage Interest Rates: **WP-14-25**
HARPing on the Importance of Competition
Gene Amromin and Caitlin Kearns

Private Takings **WP-14-26**
Alessandro Marchesiani and Ed Nosal

Momentum Trading, Return Chasing, and Predictable Crashes **WP-14-27**
Benjamin Chabot, Eric Ghysels, and Ravi Jagannathan

Early Life Environment and Racial Inequality in Education and Earnings
in the United States **WP-14-28**
Kenneth Y. Chay, Jonathan Guryan, and Bhashkar Mazumder

Poor (Wo)man's Bootstrap **WP-15-01**
Bo E. Honoré and Luojia Hu

Revisiting the Role of Home Production in Life-Cycle Labor Supply **WP-15-02**
R. Jason Faberman

Working Paper Series *(continued)*

Risk Management for Monetary Policy Near the Zero Lower Bound **WP-15-03**
Charles Evans, Jonas Fisher, François Gourio, and Spencer Krane

Estimating the Intergenerational Elasticity and Rank Association in the US:
Overcoming the Current Limitations of Tax Data **WP-15-04**
Bhashkar Mazumder

External and Public Debt Crises **WP-15-05**
Cristina Arellano, Andrew Atkeson, and Mark Wright

The Value and Risk of Human Capital **WP-15-06**
Luca Benzoni and Olena Chyruk

Simpler Bootstrap Estimation of the Asymptotic Variance of U-statistic Based Estimators **WP-15-07**
Bo E. Honoré and Luojia Hu

Bad Investments and Missed Opportunities?
Postwar Capital Flows to Asia and Latin America **WP-15-08**
Lee E. Ohanian, Paulina Restrepo-Echavarria, and Mark L. J. Wright

Backtesting Systemic Risk Measures During Historical Bank Runs **WP-15-09**
Christian Brownlees, Ben Chabot, Eric Ghysels, and Christopher Kurz

What Does Anticipated Monetary Policy Do? **WP-15-10**
Stefania D'Amico and Thomas B. King

Firm Entry and Macroeconomic Dynamics: A State-level Analysis **WP-16-01**
François Gourio, Todd Messer, and Michael Siemer

Measuring Interest Rate Risk in the Life Insurance Sector: the U.S. and the U.K. **WP-16-02**
Daniel Hartley, Anna Paulson, and Richard J. Rosen

Allocating Effort and Talent in Professional Labor Markets **WP-16-03**
Gadi Barlevy and Derek Neal

The Life Insurance Industry and Systemic Risk: A Bond Market Perspective **WP-16-04**
Anna Paulson and Richard Rosen

Forecasting Economic Activity with Mixed Frequency Bayesian VARs **WP-16-05**
Scott A. Brave, R. Andrew Butters, and Alejandro Justiniano

Optimal Monetary Policy in an Open Emerging Market Economy **WP-16-06**
Tara Iyer

Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis **WP-16-07**
Jeffrey R. Campbell, Jonas D. M. Fisher, Alejandro Justiniano, and Leonardo Melosi

Working Paper Series *(continued)*

Insurance in Human Capital Models with Limited Enforcement **WP-16-08**
Tom Krebs, Moritz Kuhn, and Mark Wright

Accounting for Central Neighborhood Change, 1980-2010 **WP-16-09**
Nathaniel Baum-Snow and Daniel Hartley

The Effect of the Patient Protection and Affordable Care Act Medicaid Expansions
on Financial Wellbeing **WP-16-10**
Luojia Hu, Robert Kaestner, Bhashkar Mazumder, Sarah Miller, and Ashley Wong

The Interplay Between Financial Conditions and Monetary Policy Shock **WP-16-11**
Marco Bassetto, Luca Benzoni, and Trevor Serrao

Tax Credits and the Debt Position of US Households **WP-16-12**
Leslie McGranahan

The Global Diffusion of Ideas **WP-16-13**
Francisco J. Buera and Ezra Oberfield

Signaling Effects of Monetary Policy **WP-16-14**
Leonardo Melosi

Constrained Discretion and Central Bank Transparency **WP-16-15**
Francesco Bianchi and Leonardo Melosi

Escaping the Great Recession **WP-16-16**
Francesco Bianchi and Leonardo Melosi

The Role of Selective High Schools in Equalizing Educational Outcomes:
Heterogeneous Effects by Neighborhood Socioeconomic Status **WP-16-17**
Lisa Barrow, Lauren Sartain, and Marisa de la Torre

Monetary Policy and Durable Goods **WP-16-18**
Robert B. Barsky, Christoph E. Boehm, Christopher L. House, and Miles S. Kimball

Interest Rates or Haircuts?
Prices Versus Quantities in the Market for Collateralized Risky Loans **WP-16-19**
Robert Barsky, Theodore Bogusz, and Matthew Easton

Evidence on the within-industry agglomeration of R&D,
production, and administrative occupations **WP-16-20**
Benjamin Goldman, Thomas Klier, and Thomas Walstrum

Expectation and Duration at the Effective Lower Bound **WP-16-21**
Thomas B. King

Working Paper Series *(continued)*

The Term Structure and Inflation Uncertainty **WP-16-22**
Tomas Breach, Stefania D'Amico, and Athanasios Orphanides

The Federal Reserve's Evolving Monetary Policy Implementation Framework: 1914-1923 **WP-17-01**
Benjamin Chabot

Neighborhood Choices, Neighborhood Effects and Housing Vouchers **WP-17-02**
Morris A. Davis, Jesse Gregory, Daniel A. Hartley, and Kegon T. K. Tan

Wage Shocks and the Technological Substitution of Low-Wage Jobs **WP-17-03**
Daniel Aaronson and Brian J. Phelan

Worker Betas: Five Facts about Systematic Earnings Risk **WP-17-04**
Fatih Guvenen, Sam Schulhofer-Wohl, Jae Song, and Motohiro Yogo

The Decline in Intergenerational Mobility After 1980 **WP-17-05**
Jonathan Davis and Bhashkar Mazumder

Is Inflation Default? The Role of Information in Debt Crises **WP-17-06**
Marco Bassetto and Carlo Galli

Does Physician Pay Affect Procedure Choice and Patient Health?
Evidence from Medicaid C-section Use **WP-17-07**
Diane Alexander

Just What the Nurse Practitioner Ordered:
Independent Prescriptive Authority and Population Mental Health **WP-17-08**
Diane Alexander and Molly Schnell

How do Doctors Respond to Incentives?
Unintended Consequences of Paying Doctors to Reduce Costs **WP-17-09**
Diane Alexander

Closing the Gap: The Impact of the Medicaid Primary Care Rate Increase on Access
and Health Unintended Consequences of Paying Doctors to Reduce Costs **WP-17-10**
Diane Alexander and Molly Schnell

Check Up Before You Check Out: Retail Clinics and Emergency Room Use **WP-17-11**
Diane Alexander, Janet Currie, and Molly Schnell

The Effects of the 1930s HOLC "Redlining" Maps **WP-17-12**
Daniel Aaronson, Daniel Hartley, and Bhashkar Mazumder

Inflation at the Household Level **WP-17-13**
Greg Kaplan and Sam Schulhofer-Wohl

Working Paper Series *(continued)*

The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries **WP-17-14**
Daniel Aaronson, Rajeev Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze

Financialization in Commodity Markets **WP-17-15**
VV Chari and Lawrence J. Christiano

On Interest Rate Policy and Asset Bubbles **WP-17-16**
Franklin Allen, Gadi Barlevy, and Douglas Gale

Estimating the Tax and Credit-Event Risk Components of Credit Spreads **WP-17-17**
Luca Benzoni and Robert S. Goldstein

The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models **WP-17-18**
Sam Schulhofer-Wohl

The Dire Effects of the Lack of Monetary and Fiscal Coordination **WP-17-19**
Francesco Bianchi and Leonardo Melosi

Selecting Primal Innovations in DSGE models **WP-17-20**
Filippo Ferroni, Stefano Grassi, and Miguel A. León-Ledesma