Forming Priors for DSGE Models

(and How It Affects the Assessment of Nominal Rigidities)

Marco Del Negro
Federal Reserve Bank of New York

Frank Schorfheide*
University of Pennsylvania
CEPR and NBER

January 4, 2008

*Correspondence: Marco Del Negro: Research Department, Federal Reserve Bank of New York, 33 Liberty Street, New York NY 10045: marco.delnegro@ny.frb.org. Frank Schorfheide: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu.

We thank Bob King, Simon Potter, Chris Sims, Frank Smets, as well as seminar participants at the Bank of Finland Conference on "Practical Issues in DSGE Modeling at Central Banks", the 7th EABCN Workshop on "Estimation and Empirical Validation of Structural Models for Business Cycle Analysis", the Fall 2006 New York Area Monetary Policy Workshop, the 2007 SED Meetings, the FRB St. Louis, the FRB San Francisco, and the Board of Governors for helpful comments and suggestions. Schorfheide gratefully acknowledges financial support from the Alfred P. Sloan Foundation and the National Science Foundation (Grant SES 0617803). The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

Abstract

The paper discusses prior elicitation for the parameters of dynamic stochastic general equilibrium (DSGE) models, and provides a method for constructing prior distributions for a subset of these parameters from beliefs about the moments of the endogenous variables. The empirical application studies the role of price and wage rigidities in a New Keynesian DSGE model and finds that standard macro time series cannot discriminate among theories that differ in the quantitative importance of nominal frictions.

JEL CLASSIFICATION: C32, E3

KEY WORDS: Bayesian Analysis, DSGE Models, Model Comparisons, Prior Elicitation, Nominal Rigidities

1 Introduction

Bayesian methods are now widely used for the estimation and evaluation of dynamic stochastic general equilibrium (DSGE) models. Of particular interest is the question what endogenous propagation mechanisms to include in the DSGE model to capture the salient features of macroeconomic time series. Several approaches are available in a Bayesian framework: comparison of impulse responses computed from DSGE models and a structural vector autoregression, e.g., Schorfheide (2000) or Del Negro, Schorfheide, Smets, and Wouters (2007); an assessment of how far actual sample moments lie in the tails of prior or posterior predictive distributions from DSGE models, e.g., Canova (1994); a comparison of different DSGE model specifications based on marginal likelihood functions (in-sample fit adjusted for model complexity), e.g. Smets and Wouters (2003, 2007), and Rabanal and Rubio-Ramirez (2005). A comprehensive survey is provided in An and Schorfheide (2007). In all of these approaches prior distributions for the DSGE model parameters play an important role for the analysis. Despite its importance the literature has paid little attention to the systematic elicitation of priors.

The paper makes a methodological and a substantive contribution. First, we provide a framework for constructing priors for different classes of parameters: those determining the steady state, the endogenous propagation of shocks, and the law of motion of exogenous disturbances. As part of this framework, we propose an easily implementable method to elicit prior distributions for DSGE model parameters from beliefs about moments of observable variables. Second, the empirical part of the the paper studies the role of nominal rigidities in a New Keynesian DSGE model with both nominal and real frictions. We find that the macro time series we use – post 1982 U.S. data on output, labor supply, labor share, inflation and interest rates – cannot discriminate among theories that differ in the quantitative importance of these rigidities. Consequently priors play a major role, whence the need to make explicit the information on which priors are based, and to have a transparent method for translating this information into statistical distributions, which is the thrust of our methodological advancement.

Prior distributions either reflect subjective opinions or summarize information derived from data sets not included in the estimation sample. The latter case is essentially equivalent to simplifying the likelihood function for a larger set of observations that would be too complicated to model directly. For instance, when pre-sample information is used to construct a prior, the tacit assumption is that the structure of the economy could have changed prior to the beginning of the estimation sample. Alternatively, priors for parameters that determine labor supply elasticities, mark-ups, frequencies of price changes, and capital adjustment costs are often quantified based on evidence from household or firm-level data sets which makes the specification of a joint likelihood function too cumbersome.[1]

There are three aspects of the prior specification that this paper aims to improve upon. First, researchers typically assume that all DSGE model parameters are independent. This assumption is made for simplicity and has the drawback that the resulting joint distribution assigns non-negligible probability mass to regions of the parameter space where the model is quite unreasonable. Second, since most of the exogenous shock processes are latent, it is difficult to quantify beliefs about their volatilities and autocorrelations. Informally researchers often choose priors that ensure that the model is roughly consistent with the autocovariance patterns observed, for instance, in a pre-sample.[2] Third, after having specified a prior distribution for the parameters of a benchmark model, researchers often use the same prior distribution for alternative model specifications when assessing the relative importance of various model features. But identical parameterizations of the exogenous shock processes potentially generate very different dynamics across model specifications. Hence the use of a common prior for all models can implicitly penalize some specifications and favor others.

We begin by dividing the parameters into three groups, which reflect the information used to construct the prior. Importantly, the placement of the DSGE model parameters into these groups depends on the prior information researchers decide to use. The choices made in this paper are meant to provide guidance but are not meant to be universal. The first group contains the parameters that determine the steady states. In the calibration literature initiated by Kydland and Prescott (1982) these parameters are often pinned down by so-called "great ratios," or other long-run measures such as the average real interest rate. Our method turns error-ridden measures of these magnitudes into a joint prior for the steady state parameters.

The second group includes the taste, technology, and policy parameters governing the DSGE model's endogenous propagation mechanism. For many of these parameters prior information comes from unrelated data sets, e.g. the prior for the labor supply elasticity parameter comes from micro-level studies on labor supply, the one for the price stickiness

[1] As discussed for instance in Chang, Gomes, and Schorfheide (2002) the prior distribution provides a useful device for incorporating micro-level information in the estimation of a aggregate time series model.

[2] The approach of eliciting priors based on beliefs about predictive densities associated with an econometric model dates back at least to Kadane, Dickey, Winkler, Smith, and Peters (1980).

parameters from studies on price changes, *et cetera*. Therefore for this second group we maintain the independence assumption standard in the literature.

The parameters describing the propagation mechanism of exogenous shocks (e.g., autocorrelations, standard deviations) belong to the third group.[3] We propose a method that translates priors about reasonable magnitudes for second moments of observables into a joint prior distribution for these parameters. Such priors may come from pre-sample evidence, for instance, and are assumed to be invariant across different DSGE model specifications. The translation occurs via a quasi-likelihood function of the DSGE model that depends on first and second moments of pre-sample or fictitious observations representing the prior information.

In the remainder of the paper, Section 2 provides a simple example that illustrates that a naive choice of prior distributions can distort Bayesian posterior odds for competing models. Section 3 describes the DSGE model, which is used in Section 4 to present our approach to prior elicitation. The empirical findings are summarized in Section 5 and Section 6 concludes.

2 A Simple Example

The typical choice of priors in DSGE model applications has two related shortcomings: independence across parameters and the mechanical use of the same prior distribution for alternative model specifications. In this section, we present two simple models to illustrate the effect of the current practice on model comparisons. We then show how one can construct an alternative prior in which autoregressive parameters are correlated based on beliefs about predictive distributions. Model $\mathcal{M}_1$ is of the form

$$y_t = \theta + \epsilon_t, \quad \epsilon_t \sim iid\mathcal{N}(0,1) \tag{1}$$

with the following prior distribution for θ: $\theta \sim \mathcal{N}(\underline{\mu}, \lambda^2)$. According to $\mathcal{M}_1$ the y_t's are independent and their marginal distribution is $\mathcal{N}(\underline{\mu}, \lambda^2 + 1)$. Model $\mathcal{M}_2$ allows for serial correlation in y_t:

$$y_t = \theta_1 y_{t-1} + \theta_2 + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1). \tag{2}$$

We will explore two prior distributions for $\mathcal{M}_2$.

[3]Canova (2007) in his discussion of calibration refers to these parameters as "nuisance/auxiliary."

The first prior is motivated as follows. Since both θ in $\mathcal{M}_1$ and θ_2 in $\mathcal{M}_2$ can be interpreted as intercepts of a regression function we use the same prior distribution for the two coefficients and assume that θ_2 is independent of θ_1:

$$\text{Prior } \mathcal{P}_1: \quad \theta_1 \sim \mathcal{U}[0, 1-\xi], \quad \theta_2|\theta_1 \sim \mathcal{N}(\underline{\mu}, \lambda^2). \tag{3}$$

The autoregressive coefficient θ_1 is uniformly distributed on the interval $[0, 1-\xi]$, $\xi > 0$. According to (2) the mean of y_t is given by $\mu = \theta_2/(1-\theta_1)$ and our prior implies that the variance of the mean is increasing in the persistence of the process

$$\mu|\theta_1 \sim \mathcal{N}\left(\underline{\mu}, \frac{\lambda^2}{(1-\theta_1)^2}\right).$$

Alternatively, we could interpret the prior for $\mathcal{M}_1$ as reflecting the belief that the mean of y_t is normally distributed with mean $\underline{\mu}$ and variance λ^2. A straightforward change-of-variable argument then leads to the second prior[4]

$$\text{Prior } \mathcal{P}_2: \quad \theta_1 \sim \mathcal{U}[0, 1], \quad \theta_2|\theta_1 \sim \mathcal{N}\left(\underline{\mu}(1-\theta_1),\ \lambda^2(1-\theta_1)^2\right). \tag{4}$$

To illustrate the effect of the prior distributions on model evaluation in situations in which the sample is not very informative about the parameters, we compute log marginal likelihood ratios for models $\mathcal{M}_1$ and $\mathcal{M}_2$ based on two observations y_1 and y_2. The two panels of Figure 1 depict contour plots of log marginal likelihood ratios of $\mathcal{M}_2(\mathcal{P}_j)$ versus $\mathcal{M}_1$, which can be interpreted as log posterior odds if the prior model odds are one. We chose $\underline{\mu} = 1$ and $\lambda = 2$. Under $\mathcal{P}_1$, a value of the autoregressive parameter close to one implies a diffuse distribution for the mean of y_t and hence a diffuse predictive distribution. Hence, compared to $\mathcal{P}_2$, prior $\mathcal{P}_1$ assigns more mass to realizations of y_1 and y_2 that have the same sign and are large in absolute value. As a consequence, the marginal likelihood for observations that are close to one is smaller for $\mathcal{P}_1$ than for $\mathcal{P}_2$. Moreover, under $\mathcal{P}_1$ these observations would be interpreted as evidence in favor of $\mathcal{M}_1$, whereas they constitute evidence in favor of the autoregressive model $\mathcal{M}_2$ if the second prior is used. In other regions of the sample space, the posterior odds are less sensitive to the choice of prior for $\mathcal{M}_2$. Observations that are large in absolute value and have the identical (opposite) sign, always provide evidence in favor (against) model $\mathcal{M}_2$.

In this simple example it is straightforward to use change-of-variable arguments to transform prior beliefs over the mean and persistence of an autoregressive process into a prior

[4] This prior has been used, for instance, in Schotman and Van Dijk (1991) in the context of unit-root testing.

distribution for θ_1 and θ_2. This transformation ensures that the first two moments of the predictive distribution of y_t are commensurate. In the case if DSGE models, it is often impossible analytically, and hard numerically, to compute the Jacobian terms associated with the change-of-variables due to the non-linearity of the cross-equation restrictions. Hence, in Section 4 we propose alternative methods of constructing prior distributions based on beliefs about the first and second moments of the endogenous variables.

3 The DSGE Model

This section briefly describes the DSGE model to which we apply our methods of constructing prior distributions. We use a medium-scale New Keynesian model with price and wage rigidities, capital accumulation, investment adjustment costs, variable capital utilization, and habit formation. The model is based on work of Smets and Wouters (2003) and Christiano, Eichenbaum, and Evans (2005). The specific version is taken from Del Negro, Schorfheide, Smets, and Wouters (2007), henceforth DSSW. For brevity we only present the log-linearized equilibrium conditions and refer the reader to the above referenced papers for the derivation of these conditions from assumptions on preferences and technologies.

The economy is populated by a continuum of firms that combine capital and labor to produce differentiated intermediate goods. These firms have access to the same Cobb-Douglas production function with capital elasticity α and total factor productivity Z_t. Total factor productivity is assumed to be non-stationary, and its growth rate $z_t = \ln(Z_t/Z_{t-1})$ follows the autoregressive process:

$$z_t = (1-\rho_z)\gamma + \rho_z z_{t-1} + \sigma_z \epsilon_{z,t}. \quad (5)$$

Output, consumption, investment, capital, and the real wage can be detrended by Z_t. In terms of the detrended variables the model has a well-defined steady state. All variables that appear subsequently are expressed as log-deviations from this steady state.

The intermediate goods producers hire labor and rent capital in competitive markets and face identical real wages, w_t, and rental rates for capital, r_t^k. Cost minimization implies that all firms produce with the same capital-labor ratio

$$k_t - L_t = w_t - r_t^k \quad (6)$$

and have marginal costs

$$mc_t = (1-\alpha)w_t + \alpha r_t^k. \quad (7)$$

The intermediate goods producers sell their output to perfectly competitive final good producers, which aggregate the inputs according to a CES function. Profit maximization of the final good producers implies that

$$\widehat{y}_t(j) - \widehat{y}_t = -\left(1 + \frac{1}{\lambda_f e^{\widetilde{\lambda}_{f,t}}}\right)(p_t(j) - p_t). \tag{8}$$

Here $\widehat{y}_t(j) - \widehat{y}_t$ and $p_t(j) - p_t$ are quantity and price for good j relative to quantity and price of the final good. The price p_t of the final good is determined from a zero-profit condition for the final good producers.

We assume that the price elasticity of the intermediate goods is time-varying. Since this price elasticity affects the mark-up that intermediate goods producers can charge over marginal costs, we refer to $\widetilde{\lambda}_{f,t}$ as mark-up shock. Following Calvo (1983), we assume that in every period a fraction of the intermediate goods producers ζ_p is unable to re-optimize their prices. A fraction ι_p of these firms adjust their prices mechanically according to lagged inflation, while the remaining fraction $1 - \iota_p$ adjusts to steady state inflation π^*. All other firms choose prices to maximize the expected discounted sum of future profits, which leads to the following equilibrium relationship, known as New Keynesian Phillips curve:

$$\pi_t = \frac{\beta}{1 + \iota_p \beta} \mathbb{E}_t[\pi_{t+1}] + \frac{\iota_p}{1 + \iota_p \beta} \pi_{t-1} + \frac{(1 - \zeta_p \beta)(1 - \zeta_p)}{\zeta_p (1 + \iota_p \beta)} mc_t + \frac{1}{\zeta_p} \lambda_{f,t}, \tag{9}$$

where π_t is inflation and β is the discount rate.[5] Our assumption on the behavior of firms that are unable to re-optimize their prices implies the absence of price dispersion in the steady state. As a consequence, we obtain a log-linearized aggregate production function of the form

$$\widehat{y}_t = (1 - \alpha) L_t + \alpha k_t. \tag{10}$$

Equations (7), (6), and (10) imply that the labor share lsh_t equals marginal costs in terms of log-deviations: $lsh_t = mc_t$.

There is a continuum of households with identical preferences, which are separable in consumption, leisure, and real money balances. Households' preferences display (internal) habit formation in consumption, that is, period t utility is a function of $\ln(C_t - hC_{t-1})$. Households supply monopolistically differentiated labor services. These services are aggregated according to a CES function that leads to a demand elasticity $1 + 1/\lambda_w$ (see Equation (8)). The composite labor services are then supplied to the intermediate goods producers at real wage w_t. To introduce nominal wage rigidity, we assume that in each period a fraction ζ_w of households is unable to re-optimize their wages. A fraction ι_w of

[5] We used the following re-parameterization: $\lambda_{f,t} = [(1 - \zeta_p \beta)(1 - \zeta_p)\lambda_f / (1 + \lambda_f)(1 + \iota_p \beta)]\widetilde{\lambda}_{f,t}$.

these households adjust their $t-1$ nominal wage by $\pi_{t-1}e^{\gamma}$, where γ represents the average growth rate of the economy, while the remaining fraction $1-\iota_p$ adjusts to steady state wage growth $\pi^* e^{\gamma}$. All other households re-optimize their wages. First-order conditions imply that

$$\begin{aligned}\tilde{w}_t &= \zeta_w \beta I\!E_t\Big[\tilde{w}_{t+1} + \Delta w_{t+1} + \pi_{t+1} + z_{t+1} - \iota_w \pi_{t-1}\Big] \\ &\quad + \frac{1-\zeta_w\beta}{1+\nu_l(1+\lambda_w)/\lambda_w}\left(\nu_l L_t - w_t - \xi_t + \frac{1}{1-\zeta_w\beta}\phi_t\right),\end{aligned} \tag{11}$$

where $\tilde{w}_t$ is the optimal real wage relative to the real wage for aggregate labor services, w_t, and ν_l would be the inverse Frisch labor supply elasticity in a model without wage rigidity ($\zeta_w = 0$) and differentiated labor. Moreover, ϕ_t is a preference shock that affects the intratemporal substitution between consumption and leisure. The real wage paid by intermediate goods producers evolves according to

$$w_t = w_{t-1} - \pi_t - z_t + \iota_w \pi_{t-1} + \frac{1-\zeta_w}{\zeta_w}\tilde{w}_t. \tag{12}$$

Households are able to insure the idiosyncratic wage adjustment shocks with state contingent claims. As a consequence they all share the same marginal utility of consumption ξ_t, which is given by the expression:

$$(e^{\gamma} - h\beta)(e^{\gamma} - h)\xi_t = -(e^{2\gamma} + \beta h^2)c_t + \beta h e^{\gamma} I\!E_t[c_{t+1} + z_{t+1}] + h e^{\gamma}(c_{t-1} - z_t), \tag{13}$$

where c_t is consumption. In addition to state-contingent claims households accumulate three types of assets: one-period nominal bonds that yield the return R_t, capital $\bar{k}_t$, and real money balances. Since preferences for real money balances are assumed to be additively separable and monetary policy is conducted through a nominal interest rate feedback rule, money is block exogenous and we will not use the households' money demand equation in our empirical analysis.

The first order condition with respect to bond holdings delivers the standard Euler equation:

$$\xi_t = I\!E_t[\xi_{t+1}] + R_t - I\!E_t[\pi_{t+1}] - I\!E_t[z_{t+1}]. \tag{14}$$

Capital accumulates according to the following law of motion:

$$\bar{k}_t = (2 - e^{\gamma} - \delta)\big[\bar{k}_{t-1} - z_t\big] + (e^{\gamma} + \delta - 1)i_t, \tag{15}$$

where i_t is investment, δ is the depreciation rate of capital. Investment in our model is subject to adjustment costs, and S'' denotes the second derivative of the investment adjustment cost function at steady state. Optimal investment satisfies the following first-order condition:

$$i_t = \frac{1}{1+\beta}\big[i_{t-1} - z_t\big] + \frac{\beta}{1+\beta}\mathbb{E}_t[i_{t+1} + z_{t+1}] + \frac{1}{(1+\beta)S'' e^{2\gamma}}(\xi_t^k - \xi_t), \tag{16}$$

where ξ_t^k is the value of installed capital and evolves according to:

$$\xi_t^k - \xi_t = \beta e^{-\gamma}(1-\delta)\mathbb{E}_t\big[\xi_{t+1}^k - \xi_{t+1}\big] + \mathbb{E}_t\big[(1-(1-\delta)\beta e^{-\gamma})r_{t+1}^k - (R_t - \pi_{t+1})\big]. \tag{17}$$

Capital utilization u_t in our model is variable and r_t^k in the previous equation represents the rental rate of effective capital $k_t = u_t + \bar{k}_{t-1}$. The optimal degree of utilization is determined by

$$u_t = \frac{r_*^k}{a''} r_t^k. \tag{18}$$

Here a'' is the derivative of the per-unit-of-capital cost function $a(u_t)$ evaluated at the steady state utilization rate. The central bank follows a standard feedback rule:

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2 \widehat{y}_t) + \sigma_R \epsilon_{R,t}. \tag{19}$$

where $\epsilon_{R,t}$ represent policy shocks. The aggregate resource constraint is given by:

$$\widehat{y}_t = (1+g_*)\left[\frac{c_*}{y_*}c_t + \frac{i_*}{y_*}\left(i_t + \frac{r_*^k}{e^{\gamma} - 1 + \delta}u_t\right)\right] + g_t. \tag{20}$$

Here c_*/y_* and i_*/y_* are the steady state consumption-output and investment-output ratios, respectively, and $g_*/(1+g_*)$ corresponds to the government share of aggregate output. The process g_t can be interpreted as exogenous government spending shock. It is assumed that fiscal policy is passive in the sense that the government uses lump-sum taxes to satisfy its period budget constraint.

We find the steady states for the detrended variables and use the method in Sims (2002) to construct a log-linear approximation of the model around the steady state. All subsequent statements about the DSGE model are statements about its log-linear approximation. We collect all the DSGE model parameters in the vector θ, stack the structural shocks in the vector ϵ_t, and derive a state-space representation for our vector of observables y_t, which is composed of

Real output growth (%, annualized)	$400(\ln Y_t - \ln Y_{t-1})$	$=$	$400(\widehat{y}_t - \widehat{y}_{t-1} + z_t)$
Hours (%)	$100 \ln L_t$	$=$	$100(L_t + \ln L^{adj})$
Labor Share (%)	$100 \ln lsh_t$	$=$	$100(L_t + w_t - \widehat{y}_t + \ln lsh_*)$
Inflation (%,annualized)	$400(\ln P_t - \ln P_{t-1})$	$=$	$400(\pi_t + \ln \pi_*)$
Interest Rates (%,annualized)	$400 \ln R_t$	$=$	$400(R_t + \ln R_*)$,

where LS_*, π_*, and R_* are the steady states of the labor share, the inflation rate, and the nominal interest rate, respectively. The parameter L^{adj} captures the units of measured hours. It can be viewed as a re-parameterization of the steady state associated with the time-varying preference parameter ϕ_t that appears in the households' utility function.

4 Forming Priors for DSGE Models

We group the DSGE model parameters into three broad categories. First, we use $\theta_{(ss)}$ to denote parameters that can be easily identified from steady state relationships among observable variables:

$$\theta_{(ss)} = [\alpha, \beta, \gamma, \delta, \lambda_f, \pi_*, g_*, L^{adj}]'.$$

Second, let $\theta_{(exo)}$ denote the parameters that characterize the law of motion of the exogenous processes

$$\theta_{(exo)} = [\rho_z, \sigma_z, \rho_\phi, \sigma_\phi, \rho_{\lambda_f}, \sigma_{\lambda_f}, \rho_g, \sigma_g, \sigma_r]'.$$

Finally, we stack the remaining parameters in the vector $\theta_{(endo)}$:[6]

$$\theta_{(endo)} = [\zeta_p, \iota_p, \zeta_w, \iota_w, \lambda_w, s'', h, a'', \nu_l, \psi_1, \psi_2, \rho_r]'.$$

We will in turn describe our method of forming prior distributions for the parameters in these three blocks. Prior distributions in the context of DSGE model estimation are by and large designed to reflect empirical observations that are excluded from the likelihood function because it would be impractical to specify a more encompassing structural econometric model. Three leading examples of such observations are (i) pre-sample data, e.g., the prior is influenced by pre-1982 observations, whereas the estimation sample is restricted to post-1982 data because of a potential monetary policy change; (ii) the use of data from other countries, e.g., a prior for a DSGE model of the Euro Area is specified based on U.S. data; (iii) the use of observations that are concurrent to the estimation sample but excluded from the likelihood function, e.g., aggregate capital stock data or micro-level data that are informative about labor supply behavior or price rigidities.

[6] Some of the parameters in $\theta_{(endo)}$ do affect steady states. For instance, the habit formation parameter h affects the steady state of the marginal utility of consumption. Nonetheless, we included these parameters in the vector $\theta_{(endo)}$ instead of $\theta_{(ss)}$ because these parameters tend to affect steady states of variables that are difficult to measure in the data.

4.1 Forming Priors for Steady-State Related Parameters

Following the work by Kydland and Prescott (1982), there is a long tradition in the business cycle literature to use long-run averages of macroeconomic time series to infer values for those parameters of the DSGE model that are related to steady states. We use this basic insight to derive a prior for the vector $\theta_{(ss)}$ from beliefs about plausible values of such long-run averages that are based on pre-sample observations or data from other countries. The parameters γ, π_*, and L^{adj} are directly tied to the steady state growth rate of aggregate output, steady state inflation, and the steady state of hours worked. The other parameters in $\theta_{(ss)}$ can be linked to "great ratios" as follows. The steady state labor share and the ratio of consumption and investment relative to output are given by

$$lsh_* = (1-\alpha)(1+\lambda_f), \quad \frac{c_* + i_*}{y_*} = 1/g_*.$$

The Investment-capital and capital-output ratios can be expressed as

$$\frac{i_*}{\bar{k}_*} = e^{-\gamma}(\delta - 1) + 1, \quad \frac{\bar{k}_*}{y_*} = \frac{\alpha}{1+\lambda_f}[\beta^{-1} - e^{-\gamma}(1-\delta)]^{-1}.$$

The drawbacks of specifying a prior distribution directly on the elements of $\theta_{(ss)}$ and assuming independence are twofold: First, choosing the prior means for the elements of $\theta_{(ss)}$ so that they jointly satisfy a set of steady state conditions can be cumbersome in a multidimensional case. Second, and most important, the joint prior potentially assigns substantial mass to parameter combinations that imply unreasonable steady-state relationships.

Instead, we propose to use fictitious measurements of the steady states to construct a prior distribution for $\theta_{(ss)}$ implicitly. Let $\mathcal{S}_D(\theta_{(ss)})$ be a vector-valued function that relates DSGE model parameters and steady states and $\widehat{\mathcal{S}}$ a vector of fictitious measurements

$$\widehat{\mathcal{S}} = \mathcal{S}_D(\theta_{(ss)}) + \eta, \tag{21}$$

where η is a vector of measurement errors. We express (21) in terms of a conditional density (likelihood function) $\mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{\mathcal{S}}\big) = p(\widehat{\mathcal{S}}|\mathcal{S}_D(\theta_{(ss)}))$ and use Bayes theorem in combination with a marginal density $\pi(\theta_{(ss)})$ to generate a conditional distribution that reflects beliefs about steady-state relationships:

$$p(\theta_{ss}|\widehat{\mathcal{S}}) \propto \mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{\mathcal{S}}\big)\pi(\theta_{(ss)}\big). \tag{22}$$

The term $\pi(\theta_{(ss)})$ allows for the possibility that the researcher possesses information on the steady state parameters other than that contained in the term $\mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{\mathcal{S}}\big)$. Two features of this prior are noteworthy. First, the information obtained from the steady states can be overidentifying in the sense that the dimension of $\mathcal{S}_D(\cdot)$ exceeds the dimension of $\theta_{(ss)}$. Second, even if the elements of the vector of measurement errors η are independent, the function $\mathcal{S}_D(\cdot)$ will induce dependence among the elements of $\theta_{(ss)}$. The researcher has to choose the vector $\widehat{\mathcal{S}}$ and make some assumptions about the distribution of the error term η in (21). We will re-visit this issue in Section 5.

4.2 Forming Priors for Shock Related Parameters

We now turn to the specification of prior distributions for the parameters associated with the exogenous shock processes. Whenever the shock processes are latent, it is difficult to quantify such priors. In practice, many researchers informally specify priors for $\theta_{(exo)}$ in an iterative manner. Starting from some initial distribution, one assesses moments of the implied predictive distribution of observable endogenous variables. The prior for $\theta_{(exo)}$ is adjusted until the predictive distribution has the desired properties, for instance, in the sense that it represents prior beliefs formed based on a pre-sample. We provide a formalization of this approach.

To fix ideas, consider a simplified version of the DSGE model presented in Section 3 where we drop capital as factor of production. This means that the parameters δ, s'', and a'' become obsolete. We further impose the following parameter restrictions

$$\alpha = 0,\ \gamma = 0,\ \pi_* = 1,\ g_* = 1,\ \iota_p = 0,\ \iota_w = 0,\ h = 0,\ \psi_1 = 1/\beta,\ \psi_2 = 0,\ \rho_r = 0.$$

We also shut down the government spending and the monetary policy shock for the sake of exposition: $\sigma_g = 0$ and $\sigma_r = 0$. The slopes of the price and wage Phillips curves are denoted by the parameters κ_p and κ_w, respectively, which are functions of the Calvo parameters ζ_p and ζ_w. One can obtain analytical solutions for output, inflation, and the labor share in terms of the structural shocks.

If wages are flexible ($\zeta_w = 0$) but prices are sticky the law of motion of the endogenous variables becomes

$$\begin{aligned}
\widehat{y}_t &= -\frac{\kappa_p \psi_p/\beta}{1-\psi_p\rho_\phi}\sigma_\phi\phi_t - \frac{\psi_p/(\zeta_p\beta)}{1-\psi_p\rho_{\lambda_f}}\sigma_{\lambda_f}\lambda_{f,t} + \frac{\rho_z\psi_p}{1-\psi_p\rho_z}\sigma_z z_t \qquad (23)\\
\pi_t &= \left[1 - \frac{(1+\nu)\kappa_p\psi_p/\beta}{1-\psi_p\rho_\phi}\right]\frac{\kappa_p}{1-\beta\rho_\phi}\sigma_\phi\phi_t \\
&\quad + \left[1 - \frac{(1+\nu)\kappa_p\psi_p/\beta}{1-\psi_p\rho_{\lambda_f}}\right]\frac{1/\zeta_p}{1-\beta\rho_{\lambda_f}}\sigma_{\lambda,f}\lambda_{f,t} + \frac{\kappa_p\psi_p(1+\nu)\rho_z}{(1-\psi_p\rho_z)(1-\beta\rho_z)}\sigma_z z_t \\
lsh_t &= \left[1 - \frac{(1+\nu)\kappa_p\psi_p/\beta}{1-\psi_p\rho_\phi}\right]\sigma_\phi\phi_t - \frac{(1+\nu)\psi_p/(\beta/\zeta_p)}{1-\psi_p\rho_{\lambda_f}}\sigma_{\lambda_f}\lambda_{f,t} + \frac{(1+\nu)\rho_z\psi_p}{1-\psi_p\rho_z}\sigma_z z_t
\end{aligned}$$

where $\psi_p = [1+\kappa_p(1+\nu)/\beta]^{-1}$. If, on the other hand, prices are flexible ($\zeta_p = 0$) and wages are sticky, then the solution changes to

$$
\begin{aligned}
\widehat{y}_t &= -\frac{\psi_w \kappa_w}{\beta(1-\psi_w\rho_\phi)(1-\zeta_w\beta)}\sigma_\phi \phi_t - \frac{\psi_w}{1-\psi_w\rho_{\lambda_f}}\left[1-\rho_{\lambda_f}+\frac{\kappa_w+1-1/\rho_\lambda}{\beta}\right]\sigma_{\lambda_f}\lambda_{f,t} \\
&\quad -\frac{\psi_w}{\rho_{\lambda_f}\beta}\sigma_{\lambda_f}\epsilon_{\lambda,t} + \frac{\psi_w}{\beta(1-\psi_w\rho_z)}\sigma_z z_t \qquad (24) \\
\pi_t &= \frac{\kappa_w}{(1-\beta\rho_\phi)(1-\beta\zeta_w)}\left[1-\frac{\kappa_w\psi_w(1+\nu)}{\beta(1-\psi_w\rho_\phi)}\right]\sigma_\phi\phi_t \\
&\quad +\frac{1}{1-\beta\rho_{\lambda_f}}\left[\beta-\frac{\kappa_w\psi_w(1+\nu)}{1-\psi_w\rho_{\lambda_f}}\right]\left[1-\rho_{\lambda_f}+\frac{\kappa_w+1-1/\rho_\lambda}{\beta}\right]\sigma_{\lambda_f}\lambda_{f,t} \\
&\quad +\frac{1}{\rho_{\lambda_f}}\left[1-\kappa_w\psi_w(1+\nu)/\beta\right]\epsilon_{\lambda_f,t} + \frac{1}{1-\beta\rho_z}\left[(\beta\rho_z-1)+\frac{\kappa_w\psi_w(1+\nu)}{\beta(1-\psi_w\rho_z)}\right]\sigma_z z_t \\
lsh_t &= -\sigma_{\lambda_f}\lambda_{f,t},
\end{aligned}
$$

where $\psi_w = [1+\kappa_w(1+\nu)/\beta]^{-1}$.

Equations (23) and (24) highlight the relationship between the endogenous variables $\widehat{y}_t$, π_t, lsh_t and the exogenous shock processes ϕ_t, $\lambda_{f,t}$, and z_t. Conditional on the parameters $\theta_{(ss)}$ and $\theta_{(endo)}$, priors on $\theta_{(exo)}$ translate into priors on the moments of the joint predictive distribution of output, inflation, and the labor share. Vice versa, one can "invert" Equations (23) and (24) to elicit a prior for $\theta_{(exo)}$ from moments of a predictive distribution for the observables. The latter approach suggests that prior distributions for the shock processes should be specific to a particular model specification.

Consider, for instance, the mark-up disturbance $\lambda_{f,t}$. If prices are assumed to be flexible as in (24), then the volatility and persistence of mark-up process and labor share are identical. Thus, a prior for $\rho_{\lambda,f}$ and $\sigma_{\lambda,f}$ can be formed directly based on views about the stochastic properties of the labor share. Alternatively, in a model in which prices are sticky, the real wage dynamics are also affected by the preference shock ϕ_t and the technology growth process z_t, which implies that it is implausible in the context of the sticky price model to equate beliefs about the persistence and volatility of real wages and the mark-up shock. We now describe a general method that allows us to translate views about the predictive distribution of the observed endogenous variables into a prior distribution for the parameters of the exogenous shock processes.

4.3 Quasi-Likelihood Based Priors

We use a (quasi)-likelihood function to translate prior beliefs about the distribution of observables, represented by a vector of fictitious observations, into a distribution for the model parameters. In principle, one could use the likelihood function of the DSGE for this purpose. We use instead the likelihood function associated with an approximating model, for which there exists a low-dimensional vector of sufficient statistics. The advantage of our approach is that the researcher only needs to specify values for the sufficient statistics on which she has formed beliefs, rather than specifying a full time series of fictitious observations.

Our approximating model is a p'th order VAR of the form

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma), \tag{25}$$

where y_t is an $n \times 1$ vector of observables. Let x_t be the $k \times 1$ vector $[1, y'_{t-1}, \ldots, y'_{t-p}]'$ and re-write the VAR as linear regression model

$$y'_t = x'_t \Phi + u'_t. \tag{26}$$

To relate the DSGE model parameters θ to the VAR parameters Φ, Σ, we assume that the observables have been transformed such that the vector y_t is covariance stationary according to the DSGE model. Let $\Gamma^D_{YY}(\theta)$, $\Gamma^D_{YX}(\theta)$ and $\Gamma^D_{XX}(\theta)$ denote the population autocovariances $I\!\!E^D_\theta[y_t y'_t]$, $I\!\!E^D_\theta[y_t x'_t]$, and $I\!\!E^D_\theta[x_t x'_t]$, respectively, which are calculated from a DSGE model conditional on a particular parameterization θ. We then define a VAR approximation of the DSGE model through the population least-squares regression:

$$\Phi_D(\theta) = [\Gamma^D_{XX}(\theta)]^{-1}\Gamma^D_{XY}(\theta), \quad \Sigma_D(\theta) = \Gamma^D_{YY}(\theta) - \Gamma^D_{YX}(\theta)[\Gamma^D_{XX}(\theta)]^{-1}\Gamma^D_{XY}(\theta). \tag{27}$$

In the multivariate Gaussian linear regression model (26) the sufficient statistics for a set of fictitious observations $\{y^*_t, x^*_t\}_{t=1}^{T^*}$ are given by $\sum y^*_t y^{*'}_t$, $\sum y^*_t x^{*'}_t$, and $\sum x^*_t x^{*'}_t$, which we will write as $T^*\Gamma^*_{YY}$, $T^*\Gamma^*_{YX}$, and $T^*\Gamma^*_{XX}$, respectively. Finally, our prior for the DSGE model parameters is obtained by interpreting the quasi-likelihood function (pre-multiplied by $|\Sigma_D(\theta)|^{-(n+1)/2}$) as a density of θ:

$$\begin{aligned} \mathcal{L}(\theta|\Gamma^*, T^*) &= |\Sigma_D(\theta)|^{-(T^*+n+1)/2} \\ &\quad \times \exp\left\{ -\frac{T^*}{2} tr\left[\Sigma_D(\theta)^{-1}(\Gamma^*_{YY} - 2\Phi_D(\theta)\Gamma^*_{XY} + \Phi'_D(\theta)\Gamma^*_{XX}\Phi_D(\theta) \right] \right\}. \end{aligned} \tag{28}$$

The quasi-likelihood (28), and hence the density of θ, is small at values of θ for which the DSGE model implied autocovariances strongly differ from the Γ^*'s. The parameter T^* captures the precision of our beliefs: The larger T^*, the sharper the peak of $\mathcal{L}(\theta|\Gamma^*, T^*)$.

The most important aspect in the implementation of the prior is the choice of Γ^*, which summarizes the information contained in the dummy observations. Suppose that $p = 0$. Then Γ^* only contains information about the mean and the covariance matrix of y_t and hence the researcher only uses beliefs about location and scale to construct a prior for θ. If $p = 1$ and x_t is composed only of y_{t-1} and the autocovariance matrices in Γ^* are specified in terms of deviations of y_t from its mean, then the prior for θ will indirectly be based on beliefs about the covariance matrix of y_t and first-order autocorrelations. This will be the case considered in the empirical implementation in Section 5.

The numerical values for the Γ^* matrix could be obtained from introspection, calculated from a pre-sample, or based on data from a different country. If Γ^* is directly constructed on the basis of a pre-sample and the lag-length p in the approximating model sufficiently large, then our quasi-likelihood based prior is similar to the use of a training-sample prior for the estimation of the DSGE model.[7] Finally, in our analysis we allow for the possibility that the researcher has additional prior information on the parameters of the exogenous shocks processes $\theta_{(exo)}$. If one adopts a literal interpretation of government spending shocks as shocks to government consumption, for instance, one can measure their persistence and standard deviation. Information from this measures, whenever available, is incorporated in the prior $\pi(\theta_{(exo)})$.

4.4 Putting it All Together

Last, we discuss the specification of the prior distribution for $\theta_{(endo)}$, the DSGE model parameters that control the endogenous propagation mechanism. For many of these parameters the researcher may have beliefs that originate from other sources of information. For instance, views on the degree of price rigidity or labor supply elasticity may arise from micro-level studies on the frequency of price changes or labor supply behavior, respectively. These beliefs are summarized in the informative prior $\pi(\theta_{(endo)})$.[8]

Our overall prior distribution will take the following form

$$\begin{aligned}\text{Prior } \mathcal{P}_{QL}: \quad p(\theta|\widehat{S}, \Gamma^*, T^*) \quad \propto \quad & \pi(\theta_{(ss)})\mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{\mathcal{S}}\big) \times \pi(\theta_{(endo)}) \\ & \times \pi(\theta_{(exo)})\mathcal{L}\big(\underline{\theta}_{(ss)}, \theta_{(exo)}, \underline{\theta}_{(endo)}|\Gamma^*, T^*\big).\end{aligned} \tag{29}$$

[7]Training-sample priors, see for instance Koop (2004), are typically constructed using the likelihood function of the econometric model that is being estimated, rather than a quasi-likelihood of an approximating model as in our approach.

[8]It is apparent from (23) and (24) that the elements of $\theta_{(endo)}$ affect the distribution of the observables in a similar way as the $\theta_{(exo)}$ parameters. This suggests that one could use *a priori* views about the moments of the data to elicit priors on $\theta_{(endo)}$ as well, via the quasi-likelihood function described in Section 4.3. While the implementation of this generalization is straightforward, we do not explore it in this paper.

We will refer to this prior as quasi-likelihood prior, $\mathcal{P}_{QL}$. The $\pi(\cdot)$ terms represent initial distributions for the model parameters that capture information not contained in the information in the fictitious observations. For most elements of $\theta_{(ss)}$ and $\theta_{(exo)}$ we use diffuse densities, mainly to ensure that the resulting prior is proper. Recall that the function $\mathcal{S}_D(\cdot)$ was chosen such that it only depends on the subvector $\theta_{(ss)}$. We fixed the parameters $\theta_{(ss)}$ and $\theta_{(endo)}$ at the values $\underline{\theta}_{(ss)}$ and $\underline{\theta}_{(endo)}$ in the quasi-likelihood function (28). Hence, we can factorize the prior as follows:

$$p(\theta|\widehat{S},\Gamma^*,T^*) = p(\theta_{(ss)}|\widehat{S})p(\theta_{(exo)}|\Gamma^*,T^*,\underline{\theta}_{(ss)},\underline{\theta}_{(endo)})p(\theta_{(endo)}), \tag{30}$$

where

$$p(\theta_{(ss)}|\widehat{S}) = c_1^{-1}\mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{S}\big)\pi(\theta_{(ss)}), \quad c_1 = \int \mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{S}\big)\pi(\theta_{(ss)})d\theta_{(ss)} \tag{31}$$

and

$$\begin{aligned} p(\theta_{(exo)}|\Gamma^*,T^*,\underline{\theta}_{(ss)},\underline{\theta}_{(endo)}) &= c_2^{-1}(\underline{\theta}_{(ss)},\underline{\theta}_{(endo)})\mathcal{L}\big(\underline{\theta}_{(ss)},\theta_{(exo)},\underline{\theta}_{(endo)}|\Gamma^*,T^*\big)\pi(\theta_{(exo)}) \\ c_2(\underline{\theta}_{(ss)},\underline{\theta}_{(endo)}) &= \int \mathcal{L}\big(\underline{\theta}_{(ss)},\theta_{(exo)},\underline{\theta}_{(endo)}|\Gamma^*,T^*\big)\pi(\theta_{(exo)})d\theta_{(exo)}. \end{aligned} \tag{32}$$

The normalization constants c_1 and $c_2(\underline{\theta}_{(ss)},\underline{\theta}_{(endo)})$ have to be computed numerically. Conceptually, it would be desirable not to condition the prior for $\theta_{(exo)}$ on a particular value of $\theta_{(ss)}$ and $\theta_{(endo)}$ and replace (32) by $p(\theta_{(exo)}|\Gamma^*,T^*,\theta_{(ss)},\theta_{(endo)})$. Unfortunately, this modification makes the normalization constant c_2 a function of the parameter vectors $\theta_{(ss)}$ and $\theta_{(endo)}$, which would have to be evaluated by numerical integration for each value that these parameters take in a Markov-Chain Monte-Carlo simulation. In the Bayesian literature, this problem is referred to as a problem of an intractable normalization constant. While there exist computation strategies to deal with such a problem in simple models[9], the computational burden in the context of our specific application is large and lead us to condition on prior mean values $\underline{\theta}_{(ss)}$ and $\underline{\theta}_{(endo)}$.

Due to the nonlinearity of the functions $\mathcal{S}_D(\theta)$, $\Phi_D(\theta)$ and $\Sigma_D(\theta)$ it is not possible to generate draws from the prior distribution of θ directly. In the application in Section 5 we use a random-walk Metropolis algorithm, described in detail for instance in An and Schorfheide (2007), to generate draws from the prior distribution. This algorithm only requires us to be able to numerically evaluate the prior density (30) up to the normalization constant. Based on the output of the Metropolis algorithm, Geweke's (1999) modified harmonic mean estimator can be used to calculate the normalization constants that appear in (31) and (32). The same algorithms can be used to obtain draws from the posterior distribution.

[9] See, for instance, Moeller, Pettitt, Reeves, and Berthelsen (2006).

A few practical considerations are worth mentioning. As long as the $\pi(\cdot)$ functions are integrable, the resulting prior distribution of θ is proper regardless of the choice of $\mathcal{L}\big(\mathcal{S}_D(\theta_{(ss)})|\widehat{\mathcal{S}}\big)$ and $\mathcal{L}\big(\underline{\theta}_{(ss)}, \theta_{(exo)}, \underline{\theta}_{(endo)}|\Gamma^*, T^*\big)$. The curvature of the prior density depends on the particular choice of the quasi-likelihood functions. The curvature of the prior for $\theta_{(ss)}$ will depend on the information included in the vector of steady states $\mathcal{S}_D(\theta_{(ss)})$. Likewise, the curvature of the prior for $\theta_{(exo)}$ will depend on that of the quasi-likelihood. Conditional on $\theta_{(ss)}$ and $\theta_{(endo)}$ it is typically possible to determine the parameters governing the law of motion of the exogenous shocks, $\theta_{(exo)}$, from the autocovariances of order zero and one. Hence in our application the prior density will have enough curvature even for small values of T^* and p for the above mentioned Markov-Chain Monte Carlo methods to work satisfactorily. If one were to rely exclusively on the quasi-likelihood to form a prior on both $\theta_{(exo)}$ and $\theta_{(endo)}$, this may result in a prior that is flat in certain dimensions (see Canova and Sala 2007), which in turn may generate computational issues. Finally, it is important to note that we do not view the quasi-likelihood as a substitute for other sources of prior information if such information is available. We view it as a tool to elicit priors for those parameters for which it is difficult to form a prior directly.

5 Assessing the Role of Nominal Rigidities

In the empirical section we use the framework discussed earlier to investigate the importance of price and wage rigidities. In this investigation we account for the existence of different *a priori* views regarding the importance of nominal rigidities. These prior disagreements reflect either different interpretation of the results of micro studies on rigidities, different assessments on the importance of strategic complementarities, but also more generally deep-rooted convictions regarding the importance of nominal frictions.

In our Bayesian setting, these *a priori* convictions are characterized by the priors on the stickiness parameters ζ_p and ζ_w summarized in Table 1. The priors in the first column of Table 1 are "non-dogmatic:" In principle all these priors put non-zero weight on the entire admissible range, hence with lots of data the likelihood would dominate in all three cases. In practice two of these priors, the *Low* and *High Rigidities*, are quite informative, hence with moderate amount of data they can affect the posterior. The *Low Rigidities* prior is roughly centered at the Bils and Klenow (2004) estimate of price stickyness for ζ_p, which implies an average frequency of price adjustment of between one and two and half quarters, and puts little mass on frequencies above three quarters. This prior assumes that nominal wage rigidity is similarly low. The *High Rigidities* prior is centered at 0.75 for ζ_p and ζ_w, which implies that the frequency of price and wage adjustment is on average four quarters. Most importantly, this prior puts virtually no mass on frequencies below two quarters. The last prior, called *Agnostic* because it is less informative, spans both the regions where the informative priors put most of the mass. We will sometimes refer to these different views of the world as models or specifications. The priors in the remaining column of Table 1 are "dogmatic:" Since either ζ_p (*Flexible Prices*) or ζ_w (*Flexible Wages*) are set to zero, there can be no updating in these dimensions. For the rigidity we do not shut down, we consider three different values corresponding to the *Low*, *High* rigidity, and *Agnostic* case. The "dogmatic" priors are of interest not only because the existing literature (Rabanal and Rubio-Ramirez (2005), Smets and Wouters (2003, 2007)) has focused on these polar cases, but also since they treat price and wage rigidity asymmetrically.

Regardless of how these *a priori* views are formed, an important question is which view of the world best describes the data. In the Bayesian framework, this question is addressed by comparing the posterior odds associated with the different priors. The problem we face is that the priors for the *other* parameters, and in particular for the parameters describing the exogenous processes, can affect this comparison. The standard practice in the Bayesian DSGE model estimation literature is to form a prior for these parameters by 1) assuming independence, and 2) maintaining the same prior across different specifications. We denote this standard prior as $\mathcal{P}_S$. The alternative prior, which we developed in Section 4, is called quasi-likelihood prior and is denoted by $\mathcal{P}_{QL}$.

5.1 Priors: Standard versus Quasi-Likelihood

We start the section by describing in detail the priors we use. We begin by focusing on those parameters ($\theta_{(ss)}$ and $\theta_{(endo)}$) whose priors do not change across specifications, given that the information on which the priors are based (great ratios, micro studies, *et cetera*) is arguably specification-invariant.

The prior for the steady state related parameters has two components: $\pi(\theta_{(ss)})$ and $\mathcal{L}(\mathcal{S}_D(\theta_{(ss)}|\widehat{S}))$. We use $\pi(\theta_{(ss)})$ to represent the prior view that the discount factor is about 0.996, the average annual growth rate of the economy is about 1.65%, the average mark-up is 15%, the inflation rate is 4.3%, and the mean level of ours worked per capita is about 1000 hours per year.[10] Prior standard deviations for these parameters are reported in the last column of Table 2. $\pi(\theta_{(ss)})$ is constant as a function of α, δ, and g_*. As is standard practice in the literature $\pi(\theta_{(ss)})$ is generated as product of marginal densities. We depart from the existing literature by using the quasi-likelihood function $\mathcal{L}(\mathcal{S}_D(\theta_{(ss)})|\widehat{S}))$ to induce an informative distribution for α, δ, and g_*. Table 2 contains fictitious measurements of the labor share, the ratio of the sum of consumption and investment to output, the investment-capital ratio, and the capital-output ratio (see the data section in the appendix for a discussion of how these numbers are computed). The Para (1) entries correspond to $\widehat{S}$, and the Para (2) entries to the standard deviations associated with the η's in (21). The function $\mathcal{L}(\mathcal{S}_D(\theta_{(ss)})|\widehat{S}))$ is constructed by assuming that the elements of η are independently and normally distributed.

We now turn to the priors for $\theta_{(endo)}$. The function $\pi(\theta_{(endo)})$ is a product of marginal densities summarized in the bottom half of Table 2. The priors for the degree of price and wage stickiness are the focus of the model comparison exercise and were discussed in Table 1. The 90% interval for the prior distribution on ν_l implies that the Frisch labor supply elasticity lies between 0.3 and 1.3, reflecting the micro-level estimates at the lower end, and the estimates of Kimball and Shapiro (2003) and Chang and Kim (2006) at the upper end. The distribution for ψ_1 is centered at a value of 2 with a standard deviation of 0.25, conditioning on the view that the central bank responded strongly to inflation movements in the Volcker-Greenspan era. The prior distribution for ψ_2 is approximately centered at 0.2, whereas the smoothing parameter lies in the range from 0.17 to 0.83. The densities for the indexation parameters ι_p and ι_w are nearly uniform over the unit interval. The density for the adjustment cost parameter s'' spans values that Christiano, Eichenbaum, and Evans (2005) find when matching DSGE and VAR impulse response functions. The density for the habit persistence parameter h is centered at 0.7, which is the value used by Boldrin, Christiano, and Fisher (2001). These authors find that $h = 0.7$ enhances the ability of a standard DSGE model to account for key asset market statistics. The density for a'' implies that in response to a 1% increase in the return to capital, utilization rates rise by 0.1 to 0.3%. These numbers are considerably smaller than the one used by Christiano, Eichenbaum, and Evans (2005).[11]

[10] We introduce the following re-parameterizations:

$$r_{(A)} = 400 * (1/\beta - 1), \quad \pi_{(A)} = 400(\pi_* - 1), \quad \gamma_{(A)} = 400\gamma.$$

[11] One can argue that on the last five elements of $\theta_{(endo)}$, namely $\{\iota_p, \iota_w, s'', h, a''\}$, the information from micro studies is fairly limited, and that those parameters are often calibrated to fit the data. Following

The prior for the $\theta_{(exo)}$, the parameters characterizing the exogenous processes (the ρ's and σ's), is summarized in Table 3 under both the standard prior ($\mathcal{P}_S$) and our approach ($\mathcal{P}_{QL}$). Under $\mathcal{P}_S$ this prior is the product of independent marginal distributions: $\pi(\theta_{(exo)}) = \Pi_i \pi_i(\theta_{i,(exo)})$, where i indexes the elements of $\theta_{(exo)}$. These distributions are shown in the left-hand side of Table 3. The prior for ρ_z, which measures the serial correlation of technology growth is centered at 0.4, whereas the priors for the other autocorrelation parameters are centered at 0.75 with a standard deviation of 0.15. These kind of prior settings for the ρ's, which are fairly informative, are standard in the literature. The priors for the σ's are loosely chosen to obtain realistic magnitudes for the implied volatilities and autocorrelation of the endogenous variables under the *Low Rigidities* specification.

Under $\mathcal{P}_{QL}$ the prior for the $\theta_{(exo)}$ parameters is given by the product of two pieces (see expression 32): a standard prior $\pi(\theta_{(exo)})$, which is specified in the right-hand side of Table 3, times the quasi-likelihood function. Unlike in the standard approach, $\pi(\theta_{(exo)})$ is largely uninformative. The prior for all the ρ's is a Beta distribution with mean 0.45 and standard deviation 0.25. This density is almost flat, although gently downward sloping, for most of the $[0, 1)$ interval, but drops sharply as ρ gets very close to one. This is a convenient feature from the computational point of view, as it avoids posterior peaks with ρ stuck at the upper corner of the interval. The prior density for the shock standard deviations is chosen to be proportional to $1/\sigma$. Under $\mathcal{P}_{QL}$ the information in the prior comes from the quasi-likelihood function $\mathcal{L}\big(\underline{\theta}_{(ss)}, \theta_{(exo)}, \underline{\theta}_{(endo)} | \Gamma^*, T^*\big)$, which is constructed based on a VAR with lag length $p = 1$ for demeaned observations on output growth, hours worked, labor share, inflation, and interest rates. To specify Γ^* we are using pre-sample autocovariance matrices of order zero and one. We consider different choices of T^* in the subsequent empirical analysis ($T^* = 4, 6, 10$), but to save space we only show results for $T^* = 6$.

Table 4 compares the mean and 90% intervals for the parameters describing the exogenous processes under the standard ($\mathcal{P}_S$) and ($\mathcal{P}_{QL}$) prior for the different specifications described in Table 1.[12] The point of Table 4 is that many of these numbers, particularly for the autocorrelation parameters, do not look at all that different across specifications. The priors for the σ's are quite different, but it is hard to tell their implications for the models. It would be hard to argue solely on the basis of these numbers that one prior is more or less

this argument, one may as well do this explicitly (using the presample) and include these parameters in the Quasi-Likelihood. We have actually done this in a set of results we do not report because they are similar to those in the paper.

[12]For the "dogmatic" specifications, *Flexible Wages* and *Prices*, we only show the prior under the *Agnostic* rigidity specification to save space.

reasonable than the other. Yet these priors can make substantial differences in terms of the model implications for impulse responses and moments of the observables, as we are going to show. In turn, this will affect model comparisons.

Figure 2 shows four sets of impulse responses. The top set shows the responses to a mark-up shock obtained under the standard prior for the *Low Rigidities* (grey, dash-and-dotted) and the *High Rigidities* (black, solid) specification. The second set shows the same responses under the $\mathcal{P}_{QL}$ prior. It is apparent that the standard prior generates very different implications for the two specifications. In particular, the standard prior implies that for the *High Rigidities* specification mark-up shocks generate what look like implausibly large movements in the observables. Moreover, these shocks generate a large negative correlation between the labor share and inflation, which is also apparent in the simplified model of Section 5.2.[13] From Galí and Gertler (1999) we know that such negative correlation is likely to be counterfactual. The second panel shows that under the $\mathcal{P}_{QL}$ prior the impulse responses for the two specifications, while not identical, are at least of a similar order of magnitude. The responses for the *Low Rigidities* specification do not change dramatically relative to the $\mathcal{P}_S$ prior, which is not surprising given that the standard deviations in the $\mathcal{P}_S$ prior were loosely calibrated on the *Low Rigidities* specification. The last two panels of Figure 2 show the responses to a policy shock under the two specifications. Not surprisingly, these are quite different for the two specifications, as they should be. Under the $\mathcal{P}_{QL}$ prior such differences persist: the $\mathcal{P}_{QL}$ and $\mathcal{P}_S$ are about the same. In summary, the $\mathcal{P}_S$ prior is likely to penalize the *High Rigidities* specification. The $\mathcal{P}_{QL}$ prior appears to remove the penalty, but at the same time maintains the identifying differences between the two models.

Figure 3 shows the implications of using the $\mathcal{P}_S$ versus the $\mathcal{P}_{QL}$ prior for some of the moments of the observables. The Figure shows three sets of plots. The first set depicts the correlation of inflation and the labor share at different lags. The second and third sets depict the autocorrelation of inflation and the labor share, respectively. In each plot the thick dark gray line with crosses represents the statistics as computed from the data (the actual sample used in the estimation, as opposed to that used for the construction of the prior). The black solid and grey dash-and-dotted lines represent the very same statistics computed from the model under the *High Rigidities* and *Low Rigidities* specification, respectively. These statistics are computed by generating parameters from the prior and, conditional on each draw, a size T time series from the model. We repeat this exercise 200,000 times and compute the median and the 90% bands for the statistics: These are the objects shown in

[13]Straightforward algebraic manipulations of (23) reveal that $\partial\pi_t/\partial\lambda_{f,t} > 0$ and $\partial lsh_t/\partial\lambda_{f,t} < 0$.

Figure 3. The statistics are computed using both the $\mathcal{P}_S$ (left column) and the $\mathcal{P}_{QL}$ prior (right column).

From the discussion of Figure 2 it should not be surprising to find that under $\mathcal{P}_S$ the *High Rigidities* specification puts substantial mass on negative correlations between inflation and the labor share (top left chart). The median contemporaneous correlation is -0.4. The *Low Rigidities* specification is somewhat closer to the data. Under the $\mathcal{P}_{QL}$ prior the implications for the two specifications are close to each other.[14] The remaining plots show that for the autocorrelation function of inflation and the labor share the difference between the two specifications is not stark. This is true under both the $\mathcal{P}_S$ and the $\mathcal{P}_{QL}$ prior. Both specifications can roughly match the persistence of inflation and the labor share. If anything, both tend to over-predict the one lag autocorrelation of inflation, which is not very high.

5.2 Posterior Estimates and Model Comparison

The priors in Table 1 capture various hypothesis on the degree of nominal rigidities in the economy. Which one is correct? We address this question by looking at the relative fit of the models, as measured by marginal likelihoods. We also look at the posterior estimates of the rigidity parameters. The role of the quasi-likelihood prior in this exercise is to try levelling the playing field among the different specifications.

Table 5 shows the log marginal likelihoods for different specifications under both the standard prior $\mathcal{P}_S$ (left panel) and the quasi-likelihood prior $\mathcal{P}_{QL}$ (right panel). The comparison among the "non-dogmatic" specifications, on the left column of each panel, is our main focus. The relative fit of the "dogmatic" specifications, *Flexible Wages* and *Prices*, is also informative however as it sheds light on which rigidity may be most needed to describe the data, sticky prices or wages. The model comparison results for the "non-dogmatic" specifications under the $\mathcal{P}_S$ prior are striking, since all models seem to describe the data roughly equally well. The marginal likelihood differences are less than 1.5, which is small,

[14] An interesting feature of the data sample we use is that $Corr(\text{Labor Share}(t), \text{Inflation}(t+k))$ is increasing in k, unlike in the sample used by Galí and Gertler (1999). In other words there seems to be little (unconditional) predictive power of inflation for future labor share, while current labor share predicts future inflation. Moreover the contemporaneous correlation is positive but small. As Figure 3 shows however, this is no evidence against the New-Keynesian model. Even under the *High Rigidities* specification the model is able to roughly replicate this pattern.

in particular if one factors in the numerical approximation error associated with these high-dimensional integrals. Thus, we find that the data cannot discriminate among the *Low* , *High Rigidities*, and *Agnostic* specifications. Under the $\mathcal{P}_{QL}$ prior (right panel) the fit of the *Low Rigidities* specification is clearly worse than that of the other two – mainly because the fit of these specifications improve after levelling the playing field. Yet it is still the case that the *High Rigidities* and *Agnostic* specifications describe the data equally well.

If for all the specifications with similar fit the posterior estimates for the Calvo parameters ζ_p and ζ_w were similar, this finding would not be noteworthy, as there is no posterior disagreement about the magnitude of the nominal rigidities. However, it turns out that the posterior estimates for the Calvo parameters are heavily influenced by the priors and quite different. The left and right panels of Table 6 show the posterior mean and 90% posterior intervals under the standard and quasi-likelihood prior, respectively.[15] The estimates for ζ_p indicate that some degree of price rigidity is needed to describe the data, although precisely how much depends on the prior. These estimates range from about 0.6 for the *Low Rigidities* to about 0.8 for the *High Rigidities* specification. The conclusion that some degree of price rigidity is needed is confirmed by the fact that the fit of the *Flexible Prices* specification is always much worse than that of the corresponding non-dogmatic model (see Table 5).

The assessment of the importance of wage rigidities depends even more on the prior views. Under the *Low Rigidities* specification the posterior estimates of ζ_w are quite low, about 0.25, while under the *High Rigidities* prior they are high, between 0.75 and 0.80. The model comparison results also indicate that the answer to the question "Do we need nominal wage rigidities?" is less robust than in the price rigidity case. The fit of the *Flexible Wages* specification is worse than that of the corresponding non-dogmatic model under *High Rigidities*, both under $\mathcal{P}_S$ and $\mathcal{P}_{QL}$. For the *Low Rigidities* case the fit of the *Flexible Wages* specification is worse than that of the corresponding non-dogmatic model under $\mathcal{P}_S$, but slightly better under $\mathcal{P}_{QL}$. Moreover, the results suggest that, without price rigidity, wage rigidity may not help much to describe the data: In the *Flexible Prices* specification the posterior estimates of ζ_w are always very small, even under the *High Rigidities* specification (the posterior mean is always below 0.25).[16]

[15] In the discussion we focus mainly on ζ_p and ζ_w, but we also tabulate the estimates for the remaining endogenous propagation ($\theta_{(endo)}$) and the exogenous persistence parameters for full disclosure.

[16] We do not show the posterior estimates for the "dogmatic" specifications for brevity. They are available upon request. An additional piece of evidence on the complementarity between wage and price rigidity is that, under the *Agnostic*/$\mathcal{P}_S$ prior, where the posterior of ζ_w spans both the low and the high rigidity region, the posterior correlation between ζ_w and ζ_p is 0.89.

We conjecture that the presence of the labor share among the observables is one of the reasons for the lack of compelling evidence in favor of wage stickiness: In absence of substantial price rigidity, the model with high wage rigidity has a hard time explaining the joint behavior of the labor share, inflation, and output. With low price rigidities movements in the labor share are dominated by mark-up shocks, which tend to generate a counterfactual negative correlation between labor share and inflation. Indeed the evidence in favor of nominal wage rigidities is much stronger whenever the labor share is not among the observables. Table 7 shows the model comparison results and the posterior estimates of ζ_p and ζ_w when the labor share is not among the observables. The estimates of ζ_w are much higher than in Table 6, and the *Low Rigidity* specification is soundly rejected by the data.

In summary, the following two models seem to fit the data equally well: One where price rigidities are moderate and wage rigidity is trivial, and one where both rigidities are high. These two models have strikingly different policy implication, as shown by Figure 4. The Figure plots the impulse responses to a policy shocks for *High Rigidities* (solid black lines) and the *Agnostic* case (grey dash-and-dotted lines) under $\mathcal{P}_{QL}$ (results under $\mathcal{P}_S$ are similar). It is clear that the reduction in output following a decrease in inflation varies substantially between these two models.

6 Conclusions

The choice of priors for DSGE model parameters matters for both posterior estimates and model comparison. Part of this paper's contribution is to provide a framework for eliciting priors for different classes of parameters: those determining the steady state, the endogenous propagation mechanism, and the law of motion of the exogenous disturbances. The main thrust of our approach is to make explicit the information on which priors are based, whether that comes from the pre-sample or other sources, and to provide a systematic approach for translating this information into priors.

The paper uses the approach to investigate the importance of nominal rigidities within a standard New-Keynesian model with several real rigidities. In the profession there are widely diverging views on this subject. One would think that if the macro data spoke very clearly one way or the other, some consensus may eventually emerge as information from the data eventually trumps people's priors. The results in this paper suggest that this is not yet the case. The macro time series we consider – output growth, labor supply, labor share, inflation and interest rates – are not informative enough to discriminate among different theories, in spite of being a natural choice of observables for the question at hand. We indeed show that posterior estimates of the nominal rigidities parameters to some extent mirror the prior views. We also show that the model comparison results are in general not robust to the choice of prior for the parameters describing the exogenous shock processes, as these priors may inadvertently favor one specification relative to another. One promising approach to discriminating among theories is to gather evidence from micro data on the degree of rigidities (see Bils and Klenow 2004, Nakamura and Steinsson 2007) but it remains a challenge to understand how micro-level rigidities aggregate to macro-level rigidities.

References

An, S., Schorfheide, F., 2007. Bayesian Analysis of DSGE Models. Econometric Reviews 26, 113-172.

Bils, M., Klenow, P., 2004. Some Evidence on the Importance of Sticky Prices. Journal of Political Economy 112, 947-985.

Boldrin, M., Christiano, L., Fisher, J., 2001. Habit Persistence, Asset Returns, and the Business Cycle. American Economic Review 91, 149-166.

Calvo, G., 1983. Staggered Prices in a Utility-Maximizing Framework. Journal of Monetary Economics 12, 383-398.

Canova, F., 1994. Statistical Inference in Calibrated Models. Journal of Applied Econometrics 9, S123-144.

Canova, F., 2007. Methods for Applied Macroeconomic Research. Princeton University Press, Princeton, New Jersey.

Canova, F., Sala, L., 2007. Back to Square One: Identification Issues in DSGE Models. Manuscript, Universitat Pompeu Fabra and Universita Bocconi.

Chang, Y., Gomes, J., Schorfheide, F., 2002. Learning-by-Doing as Propagation Mechanism. American Economic Review 92, 1498-1520.

Chang, Y., Kim, S.-B., 2006. From Individual to Aggregate Labor Supply: A Quantitative Analysis based on a Heterogeneous Agent Macroeconomy. International Economic Review 47, 1-27.

Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy 113, 1-45.

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the Fit of New-Keynesian Models. Journal of Business & Economic Statistics 25, 123-162.

Galí, J., Gertler, M., 1999. Inflation Dynamics: A Structural Econometric Analysis. Journal of Monetary Economics 44, 195-222.

Geweke, J., 1999. Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication. Econometric Reviews 18, 1-126.

Kadane, J.B., J.M. Dickey, R.L. Winkler, W.S. Smith, Peters, S.C., 1980. Interactive Elicitation of Opinion for a Normal Linear Model. Journal of the American Statistical Association 75, 845-854.

Koop, G., 2004. Bayesian Econometrics. John Wiley & Sons, Hoboken, New Jersey.

Kydland, F., Prescott, E., 1982. Time to Build and Aggregate Fluctuations. Econometrica 50, 1345-70.

Kimball, M., Shapiro, M., 2003. Labor Supply: Are the Income and Substitution Effects Both Large or Both Small? *Manuscript*, University of Michigan.

Moller, J., Pettitt, A.N., Reeves, R., Berthelsen, K.K., 2006. An Efficient Markov Chain Monte Carlo Method for Distributions with Intractable Normalising Constants. Biometrika 93, 451-458.

Nakamura, E., Steinsson, J., 2007. Five Facts About Prices: A Reevaluation of Menu Costs Models. Manuscript, Columbia University.

Rabanal, P., Rubio-Ramirez, J.F., 2005. Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach. Journal of Monetary Economics 52, 1151-1166.

Schorfheide, F., 2000. Loss-Function Based Evaluation of DSGE Models. Journal of Applied Econometrics 15, 645-670.

Schotman, P., Van Dijk, H., 1991. On Bayesian Routes to Unit Roots. Journal of Applied Econometrics 6, 387-401.

Sims, C., 2002. Solving Rational Expectations Models. Computational Economics 20, 1-20.

Smets, F., Wouters, R., 2003. An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area. Journal of the European Economic Association 1, 1123-75.

Smets, F., Wouters, R., 2007. Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach. American Economic Review 97, 586-606.

A Data

The data set is obtained from Haver Analytics (Haver mnemonics are in italics). We compile observations for the variables that appear in the measurement equation (21). Real output is obtained by dividing the nominal series (*GDP*) by population 16 years and older (*LN16N*), and deflating using the chained-price GDP deflator (*JGDP*). We compute quarter-to-quarter output growth as log difference of real GDP per capita and multiply the growth rates by 100 to convert them into percentages. Our measure of hours worked is computed by taking total hours worked reported in the National Income and Product Accounts (NIPA), which is at annual frequency, and interpolating it using growth rates computed from hours of all persons in the non-farm business sector (*LXNFH*). We divide hours worked by *LN16N* to convert them into per capita terms. We then take the log of the series multiplied by 100 so that all figures can be interpreted as percentage changes in hours worked. The labor share is computed by dividing total compensation of employees (*YCOMP*) obtained from the NIPA by nominal GDP. We then take the log of the labor share multiplied by 100. Inflation rates are defined as log differences of the GDP deflator and converted into annualized percentages. The nominal rate corresponds to the effective Federal Funds Rate (*FFED*), also in percent. We use a pre-sample of observations from 1954:III to 1980:IV to specify the prior distribution. Our estimation sample ranges from 1982:IV to 2005:IV. Annual data on consumption, durable consumption, investment, and capital used to construct the great ratios included in $\hat{S}$ (see Table 2) also come from Haver Analytics (with mnemonics *C*, *CD*, *IDGA*, and *E*, respectively). Since these variables are not included among the observables we use the entire sample (1954-2006) to obtain information on the great ratios. The average labor share measurement included in $\hat{S}$ comes from the pre-sample.

Table 1: Prior Views on the Degree of Price and Wage Rigidity

	"Non-Dogmatic" Priors		"Dogmatic" Priors			
			Flexible Wages		Flexible Prices	
	Mean	(St. Dev.)	Mean	(St. Dev.)	Mean	(St. Dev.)
	Low Rigidities					
ζ_p	.45	(.10)	.45	(.10)	0	
ζ_w	.45	(.10)	0		.45	(.10)
	High Rigidities					
ζ_p	.75	(.10)	.75	(.10)	0	
ζ_w	.75	(.10)	0		.75	(.10)
	Agnostic Prior					
ζ_p	.60	(.20)	.60	(.20)	0	
ζ_w	.60	(.20)	0		.60	(.20)

Notes: Prior standard deviations are in parenthesis whenever the prior is non-degenerate.

Table 2: Prior Distribution for Steady State ($\theta_{(ss)}$) and Endogenous Propagation ($\theta_{(endo)}$) Parameters

Parameter	Domain	Density	Para (1)	Para (2)
		$\pi(\theta_{(ss)})$		
α	[0,1)	Uniform	0.00	1.00
$r_{(A)}$	$\mathbb{R}^+$	Gamma	1.50	1.00
δ	[0,1)	Uniform	0.00	1.00
$\gamma_{(A)}$	$\mathbb{R}^+$	Gamma	1.65	1.00
λ_f	$\mathbb{R}^+$	Gamma	0.15	0.10
$\pi_{(A)}$	$\mathbb{R}$	Normal	4.30	2.50
$g_* - 1$	$\mathbb{R}^+$	Uniform	0.00	∞
L^{adj}	$\mathbb{R}$	Normal	252	10.0
		$\mathcal{L}(\mathcal{S}_D(\theta_{(ss)})\vert\widehat{S})$		
LS_*		Normal	0.57	0.02
$(c_* + i_*)/y_*$		Normal	0.84	0.02
$i_*/\bar{k}_*$		Normal	0.09	0.01
$\bar{k}_*/y_*$		Normal	3.18	0.18
		$\pi(\theta_{(endo)})$		
ζ_p	[0,1)	Beta	see Table 1	
ζ_w	[0,1)	Beta	see Table 1	
ν_l	$\mathbb{R}^+$	Gamma	2.00	0.75
ψ_1	$\mathbb{R}^+$	Gamma	2.00	0.25
ψ_2	$\mathbb{R}^+$	Gamma	0.20	0.10
ρ_r	[0,1)	Beta	0.50	0.20
ι_p	[0,1)	Beta	0.50	0.28
ι_w	[0,1)	Beta	0.50	0.28
s''	$\mathbb{R}^+$	Gamma	4.00	1.50
h	[0,1)	Beta	0.700	0.050
a''	$\mathbb{R}^+$	Gamma	0.20	0.10

Notes: Para (1) and Para (2) correspond to means and standard deviations for the Beta, Gamma, and Normal distributions and to the upper and lower bounds of the support for the Uniform distribution. In case of $\mathcal{L}(\mathcal{S}_D(\theta_{(ss)})|\widehat{S})$ the Para (1) entry can be interpreted as $\widehat{S}$ value and the Para (2) entry as standard deviation of η in Equation (21).

Table 3: Prior for Exogenous Propagation Parameters ($\theta_{(exo)}$)

		$\mathcal{P}_S$			$\mathcal{P}_{QL}$		
Parameter	Domain	Density	Para (1)	Para (2)	Density	Para (1)	Para (2)
ρ_z	[0,1)	Beta	0.40	0.25	Beta	0.45	0.25
ρ_ϕ	[0,1)	Beta	0.75	0.15	Beta	0.45	0.25
ρ_{λ_f}	[0,1)	Beta	0.75	0.15	Beta	0.45	0.25
ρ_g	[0,1)	Beta	0.75	0.15	Beta	0.45	0.25
σ_z	$\mathbb{R}^+$	InvGamma	0.30	4.00		$\propto 1/\sigma_z$	
σ_ϕ	$\mathbb{R}^+$	InvGamma	3.00	4.00		$\propto 1/\sigma_\phi$	
σ_{λ_f}	$\mathbb{R}^+$	InvGamma	0.20	4.00		$\propto 1/\sigma_{\lambda_f}$	
σ_g	$\mathbb{R}^+$	InvGamma	0.50	4.00		$\propto 1/\sigma_g$	
σ_r	$\mathbb{R}^+$	InvGamma	0.20	4.00		$\propto 1/\sigma_r$	
					$\times$	$\dfrac{\mathcal{L}\big(\underline{\theta}_{(ss)}, \theta_{(exo)}, \underline{\theta}_{(endo)} \vert \Gamma^*, T^*\big)}{c_2(\underline{\theta}_{(ss)}, \underline{\theta}_{(endo)})}$	

Notes: Para (1) and Para (2) correspond to means and standard deviations for the Beta distribution and to s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$.

Table 4: Priors for Exogenous Propagation Parameters $\theta_{(exo)}$: Standard vs Quasi-Likelihood

Parameter	$\mathcal{P}_S$	*Agnostic* $\mathcal{P}_{QL}$	*Low Rigidities* $\mathcal{P}_{QL}$	*High Rigidities* $\mathcal{P}_{QL}$	*Flexible Wages* $\mathcal{P}_{QL}$	*Flexible Prices* $\mathcal{P}_{QL}$
ρ_z	0.40	0.47	0.47	0.49	0.42	0.48
	(0.00,0.76)	(0.17,0.75)	(0.16,0.75)	(0.18,0.77)	(0.09,0.72)	(0.19,0.72)
ρ_ϕ	0.75	0.63	0.68	0.51	0.74	0.70
	(0.35,1.00)	(0.31,0.97)	(0.41,0.98)	(0.13,0.89)	(0.49,0.99)	(0.44,0.98)
ρ_{λ_f}	0.75	0.77	0.81	0.61	0.63	0.84
	(0.35,1.00)	(0.56,0.99)	(0.63,0.99)	(0.29,0.95)	(0.24,0.98)	(0.70,0.99)
ρ_g	0.75	0.53	0.52	0.52	0.64	0.50
	(0.35,1.00)	(0.10,0.92)	(0.11,0.93)	(0.11,0.91)	(0.23,1.00)	(0.08,0.89)
σ_z	1.13	1.34	1.40	1.48	1.40	1.48
	(0.48,1.78)	(0.73,1.91)	(0.73,2.01)	(0.73,2.34)	(0.79,2.37)	(0.71,2.21)
σ_ϕ	3.76	6.22	4.89	12.74	3.70	5.41
	(1.59,5.93)	(1.26,11.72)	(1.47,8.21)	(1.43,24.84)	(1.11,7.53)	(1.47,9.44)
σ_{λ_f}	0.25	0.25	0.35	0.21	0.28	0.87
	(0.11,0.40)	(0.10,0.42)	(0.15,0.57)	(0.06,0.37)	(0.11,0.49)	(0.42,1.32)
σ_g	0.63	0.68	0.62	0.67	1.22	0.52
	(0.26,0.99)	(0.00,1.35)	(0.18,1.35)	(0.19,1.27)	(0.22,2.97)	(0.00,1.03)
σ_r	0.50	0.59	0.54	0.56	0.51	0.56
	(0.21,0.79)	(0.23,1.04)	(0.23,0.92)	(0.24,0.91)	(0.00,0.78)	(0.19,1.04)

Notes: The table shows the mean and, in parenthesis, the 90% intervals. Results for the $\mathcal{P}_{QL}$ prior are shown for $T^* = 6$. For *Flexible Wages* and *Prices* we show the prior under the *Agnostic* rigidity specification.

Table 5: Marginal Likelihoods

$\mathcal{P}_S$			$\mathcal{P}_{QL}$		
Non-Dogmatic	Flex. Wages	Flex. Prices	Non-Dogmatic	Flex. Wages	Flex. Prices
Low Rigidities					
-518.83	-526.87	-569.84	-521.96	-519.21	-565.81
Agnostic Prior					
-517.54	-527.20	-568.71	-511.63	-517.58	-563.78
High Rigidities					
-517.32	-527.98	-578.42	-510.11	-519.84	-575.10

Notes: The table shows the log marginal likelihoods for different specifications under both the standard prior $\mathcal{P}_S$ (left panel) and the quasi-likelihood prior $\mathcal{P}_{QL}$ (right panel).

Table 6: Posterior Estimates

Parameter	$\mathcal{P}_S$ *Low Rigidities*	$\mathcal{P}_S$ *Agnostic*	$\mathcal{P}_S$ *High Rigidities*	$\mathcal{P}_{QL}$ *Low Rigidities*	$\mathcal{P}_{QL}$ *Agnostic*	$\mathcal{P}_{QL}$ *High Rigidities*
Rigidities						
ζ_p	0.64	0.78	0.84	0.56	0.65	0.81
	(0.57,0.72)	(0.64,0.88)	(0.80,0.89)	(0.48,0.64)	(0.59,0.72)	(0.75,0.86)
ζ_w	0.26	0.52	0.74	0.24	0.19	0.80
	(0.18,0.34)	(0.16,0.81)	(0.61,0.84)	(0.15,0.32)	(0.10,0.29)	(0.73,0.87)
Other Endogenous Propagation ($\theta_{(endo)}$) Parameters						
ν_l	2.20	0.96	0.78	1.77	1.60	1.75
	(1.22,3.18)	(0.12,1.79)	(0.17,1.37)	(0.90,2.72)	(0.84,2.42)	(0.92,2.58)
ψ_1	2.40	2.32	2.25	2.47	2.48	2.19
	(1.96,2.85)	(1.88,2.77)	(1.84,2.68)	(2.02,2.92)	(2.01,2.92)	(1.80,2.57)
ψ_2	0.04	0.06	0.07	0.06	0.06	0.08
	(0.01,0.07)	(0.02,0.10)	(0.03,0.11)	(0.02,0.10)	(0.02,0.10)	(0.03,0.13)
ρ_r	0.79	0.80	0.81	0.80	0.79	0.81
	(0.75,0.83)	(0.76,0.85)	(0.77,0.85)	(0.75,0.84)	(0.75,0.84)	(0.77,0.85)
ι_p	0.19	0.19	0.20	0.47	0.14	0.08
	(0.00,0.38)	(0.00,0.38)	(0.00,0.41)	(0.10,0.85)	(0.00,0.30)	(0.00,0.18)
ι_w	0.49	0.42	0.36	0.52	0.56	0.42
	(0.08,0.93)	(0.00,0.81)	(0.00,0.71)	(0.12,0.97)	(0.17,1.00)	(0.01,0.77)
s'	8.15	9.76	10.54	8.13	8.58	10.84
	(5.19,11.10)	(5.98,13.18)	(6.91,13.96)	(4.67,11.20)	(5.36,11.71)	(7.40,14.06)
h	0.67	0.80	0.82	0.69	0.70	0.79
	(0.58,0.76)	(0.70,0.89)	(0.77,0.88)	(0.60,0.79)	(0.60,0.79)	(0.72,0.86)
a''	0.25	0.21	0.21	0.26	0.24	0.22
	(0.10,0.40)	(0.05,0.35)	(0.04,0.36)	(0.09,0.42)	(0.09,0.39)	(0.07,0.36)
Exogenous Propagation ($\theta_{(exo)}$) Parameters (ρs only)						
ρ_z	0.13	0.15	0.18	0.16	0.13	0.27
	(0.00,0.25)	(0.00,0.28)	(0.00,0.33)	(0.00,0.30)	(0.00,0.24)	(0.07,0.45)
ρ_ϕ	0.98	0.68	0.50	0.93	0.94	0.27
	(0.95,1.00)	(0.48,1.00)	(0.31,0.69)	(0.86,1.00)	(0.89,1.00)	(0.10,0.45)
ρ_{λ_f}	0.86	0.56	0.41	0.89	0.88	0.68
	(0.77,1.00)	(0.26,0.93)	(0.17,0.65)	(0.82,0.96)	(0.81,0.95)	(0.53,0.85)
ρ_g	0.91	0.96	0.97	0.92	0.92	0.92
	(0.87,0.95)	(0.92,1.00)	(0.92,1.00)	(0.87,0.96)	(0.88,0.96)	(0.89,0.96)

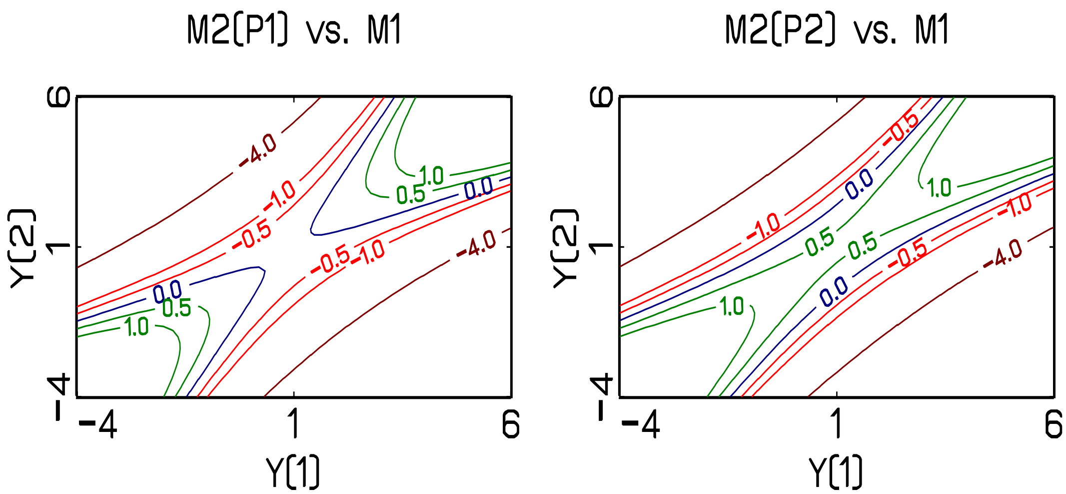
Notes: The table shows the mean and, in parenthesis, the 90% intervals. The left and right panels show the estimates under the standard ($\mathcal{P}_S$) and quasi-likelihood ($\mathcal{P}_{QL}$) prior, respectively.

Table 7: Assessing Nominal Rigidities without the Labor Share Among the Observables

	Low Rigidities	*Agnostic*	*High Rigidities*
ζ_p	0.78	0.84	0.84
	(0.71,0.84)	(0.79,0.89)	(0.79,0.89)
ζ_w	0.73	0.94	0.92
	(0.42,0.89)	(0.91,0.98)	(0.87,0.97)
Marg. Likelihood	-408.45	-387.68	-386.96

Notes: The table shows the mean and, in parenthesis, the 90% intervals for the nominal rigidity parameters under the standard prior for the *Low*, *High Rigidities* and *Agnostic* specifications. The table also shows the log marginal likelihoods associated with each specification.

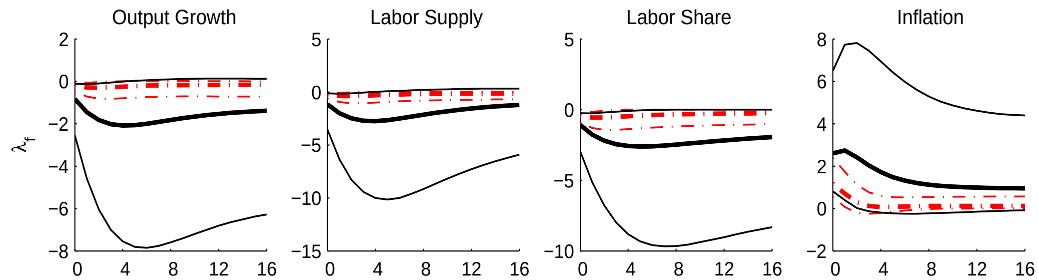
Figure 1: Log Posterior Odds in Favor (Positive Values) of $\mathcal{M}_2$



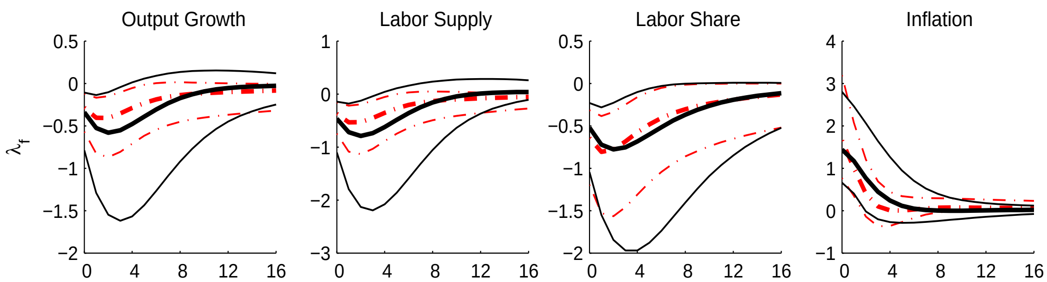
Notes:

Figure 2: A Priori Impulse Response Functions: *Low* versus *High Rigidities*

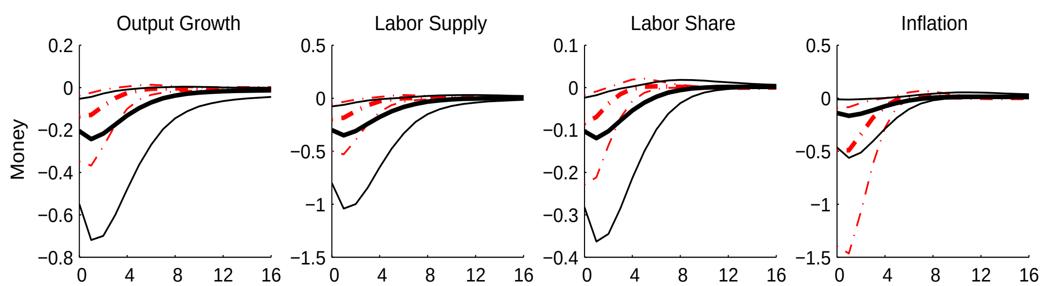
(1) Mark-up Shocks, $\mathcal{P}_S$



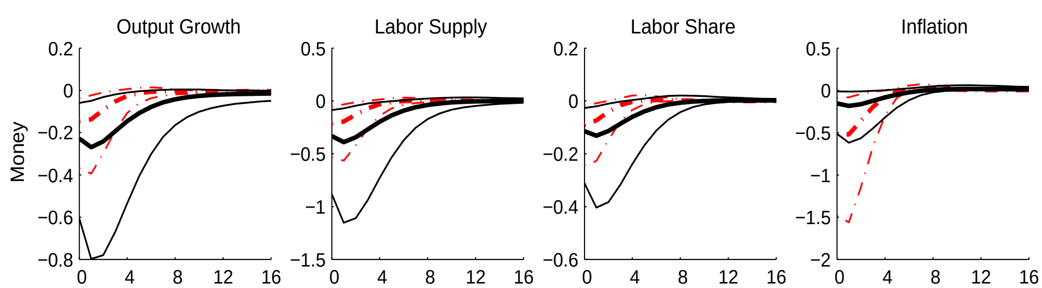
(2) Mark-up Shocks, $\mathcal{P}_{QL}$



(3) Monetary Policy Shocks, $\mathcal{P}_S$

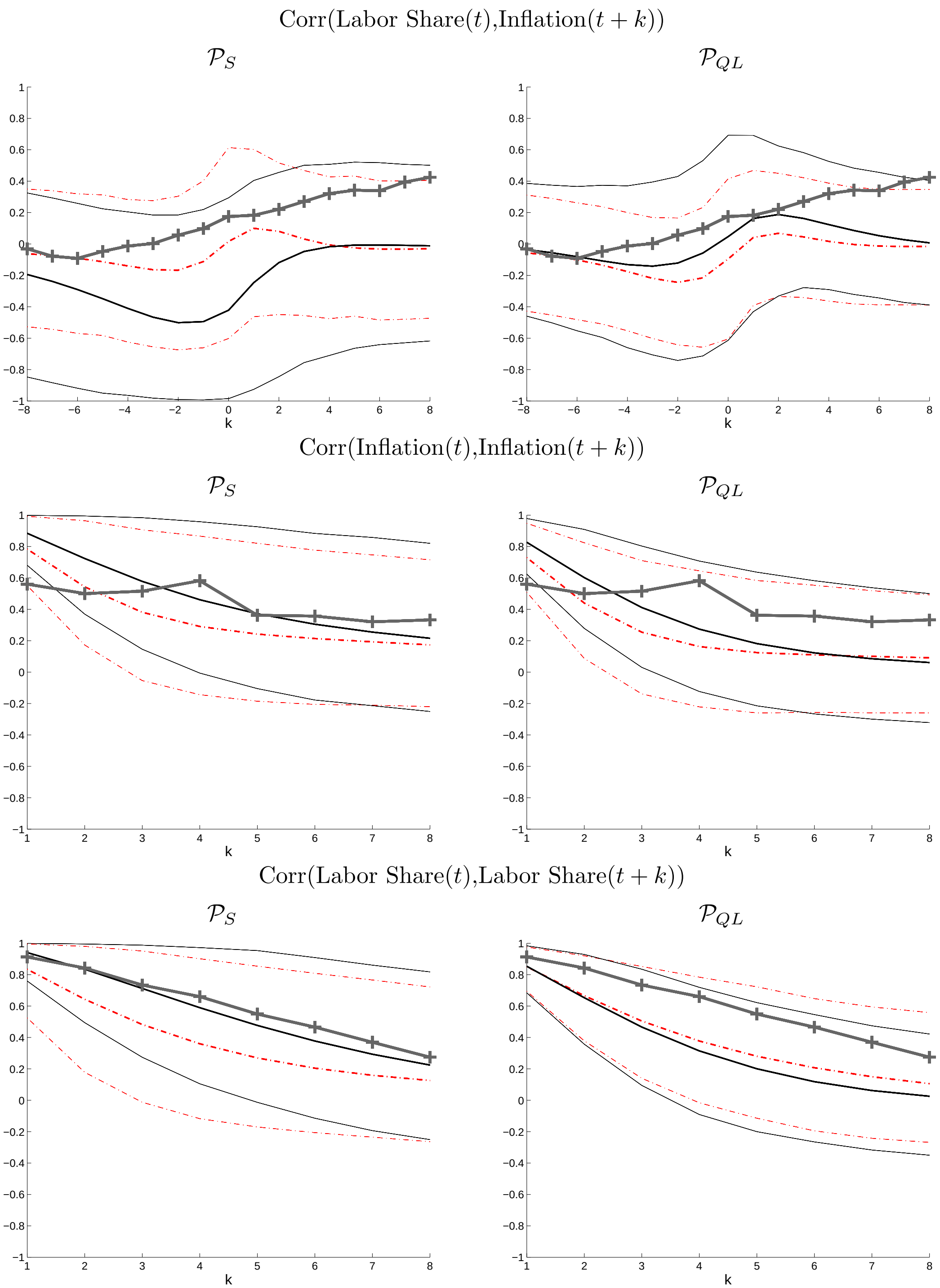


(4) Monetary Policy Shocks, $\mathcal{P}_{QL}$



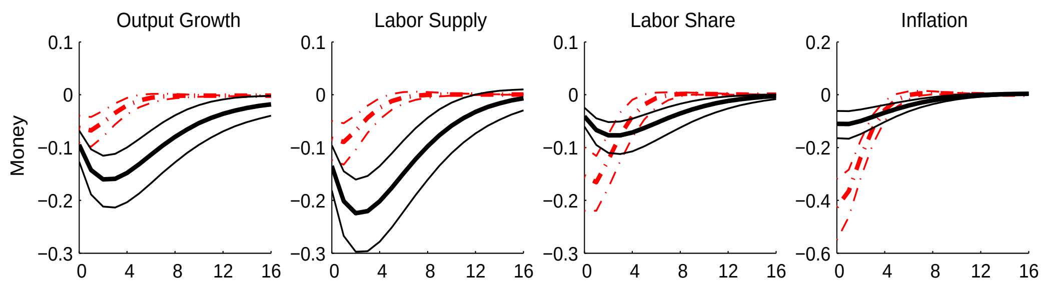
Notes: The black solid and grey dash-and-dotted lines represent the median impulse responses and the 90% bands under the *High Rigidities* and *Low Rigidities* specifications, respectively.

Figure 3: Prior Predictive Moments: *Low* versus *High Rigidities*



Notes: For each plot the thick dark gray line with crosses represents the statistics (Corr(Labor Share(t),Inflation($t+k$)), Corr(Inflation(t),Inflation($t+k$)), Corr(Labor Share(t),Labor Share($t+k$))) as computed from the data. The black solid and grey dash-and-dotted lines represent the statistics computed from the model under the *High Rigidities* and *Low Rigidities* specification, respectively. These statistics are computed by generating parameters from the prior and, conditional on each draw, a size T time series from the model. We repeat this exercise 200,000 times and compute the median and the 90% bands for the statistics. The statistics are computed using both the $\mathcal{P}_S$ (left column) and the $\mathcal{P}_{QL}$ prior (right column).

Figure 4: Implications of Different Assessments of Nominal Rigidities: Posterior Impulse Response Functions to a Monetary Policy Shock based on Prior $\mathcal{P}_{QL}$



Notes: High Rigidities specification IRFs are black solid, *Agnostic* specification IRFs are grey dash-and-dotted.