Monetary Policy Analysis with Potentially Misspecified Models – Technical Appendix

May 1, 2007

1 Introduction

This technical appendix provides detailed derivations for results reported in Del Negro and Schorfheide "Monetary Policy Analysis with Potentially Misspecified Models." We derive the VAR approximation of a linearized DSGE model in Section 2. Section 3 provides details about the DSGE-VAR framework, including the construction of the prior distribution, the likelihood function, the posterior distribution, and a discussion of the various approaches toward policy analysis. Section 4 solves a simplified DSGE model that is used in the paper to explain some of the empirical results obtained with the full DSGE model. Finally, Section 5 describes how we specify the prior distribution.

2 VAR Approximation of DSGE Model

The policy rule can be written in general form as

$$y_{1,t} = x_t'\beta_1(\theta) + y_{2,t}'\beta_2(\theta) + \epsilon_{1,t}, \tag{1}$$

where $y_t = [y_{1,t}, y_{2,t}']'$ and the $k \times 1$ vector $x_t = [y_{t-1}', \ldots, y_{t-p}', 1]'$ is composed of the first p lags of y_t and an intercept. The shock $\epsilon_{1,t}$ corresponds to the monetary policy shock $\sigma_R\widetilde{\epsilon}_{1,t}$ in the DSGE model. The matrices $\beta_1(\theta)$ and $\beta_2(\theta)$ select the appropriate elements of x_t and $y_{2,t}$ to generate the policy rule as a function of θ.

The observed private sector variables are stacked in the vector $y_{2,t}$. We approximate the DSGE model-implied moving average representation of $y_{2,t}$ with a p-th order autoregression,

which we write as

$$y'_{2,t} = x'_t \Psi^*(\theta) + u'_{2,t}. \tag{2}$$

Assuming that the law of motion for $y_{2,t}$ is covariance stationary, we define $\Gamma_{XX}(\theta) = \mathbb{E}^D_\theta[x_t x'_t]$ and $\Gamma_{XY_2}(\theta) = \mathbb{E}^D_\theta[x_t y'_{2,t}]$ and let

$$\Psi^*(\theta) = \Gamma^{-1}_{XX}(\theta)\Gamma_{XY_2}(\theta). \tag{3}$$

The equation for the policy instrument (1) can be rewritten by replacing $y_{2,t}$ with expression (2):

$$y_{1,t} = x'_t\beta_1(\theta) + x'_t\Psi^*(\theta)\beta_2(\theta) + u_{1,t}, \tag{4}$$

where $u_{1,t} = u'_{2,t}\beta_2(\theta) + \sigma_R\widetilde{\epsilon}_{1,t}$. Define $u'_t = [u_{1,t}, u'_{2,t}]$, $B_1(\theta) = [\beta_1(\theta), 0_{k\times(n-1)}]$ and $B_2(\theta) = [\beta_2(\theta), I_{(n-1)\times(n-1)}]$. We obtain a restricted VAR for y_t of the form

$$y'_t = x'_t\Phi^*(\theta) + u'_t, \quad \mathbb{E}[u_t u'_t] = \Sigma^*(\theta) \tag{5}$$

with

$$\begin{aligned}\Phi^*(\theta) &= B_1(\theta) + \Psi^*(\theta)B_2(\theta)\\ \Sigma^*(\theta) &= \Gamma_{YY}(\theta) - \Gamma_{YX}(\theta)\Gamma^{-1}_{XX}(\theta)\Gamma_{XY}(\theta).\end{aligned}$$

Here the population covariance matrices are $\Gamma_{YY}(\theta) = \mathbb{E}^D_\theta[y_t y'_t]$ and $\Gamma_{XY}(\theta) = \Gamma'_{YX}(\theta) = \mathbb{E}^D_\theta[x_t y'_t]$.

Lemma 1 *(i) The VAR coefficient matrix* $\Phi^*(\theta) = \Gamma^{-1}_{XX}(\theta)\Gamma_{XY}(\theta)$. *(ii)* $\mathbb{E}^{VAR}_{\Psi^*(\theta),\Sigma^*(\theta)}[x_t x'_t] = \mathbb{E}^D_\theta[x_t x'_t] = \Gamma_{XX}(\theta)$.

Proof (i) We begin by calculating $\mathbb{E}^D_\theta[x_t y_{1,t}]$:

$$\begin{aligned}\mathbb{E}^D_\theta[x_t y_{1,t}] &= \mathbb{E}^D_\theta[x_t(x'_t\beta_1(\theta) + y_{2,t}\beta_2(\theta) + \epsilon_{1,t})]\\ &= \Gamma_{XX}(\theta)\beta_1(\theta) + \Gamma_{XY_2}(\theta)\beta_2(\theta).\end{aligned}$$

We used the fact that under the DSGE model $\epsilon_{1,t}$ is uncorrelated with past endogenous variables. Using the definitions of $B_1(\theta)$ and $B_2(\theta)$ we deduce

$$\begin{aligned}\mathbb{E}^D_\theta[x_t y_t] &= \mathbb{E}^D_\theta\Big[\Gamma_{XX}(\theta)\beta_1(\theta) + \Gamma_{XY_2}(\theta)\beta_2(\theta),\ \Gamma_{XY_2}(\theta)\Big]\\ &= \Gamma_{XX}(\theta)B_1(\theta) + \Gamma_{XY_2}(\theta)B_2(\theta).\end{aligned}$$

Pre-multiplying this result by $\Gamma^{-1}_{XX}(\theta)$ completes the argument.

(ii) Without loss of generality we can omit the intercept from the subsequent calculations. We present a proof for $p = 2$. The extension to a general p is straightforward. Define $M' = [I_{n\times n}, 0_{n\times n}]$. Moreover, $x_t = [y_{t-1}, y_{t-2}]$. We will write the VAR with parameters $\Phi^*(\theta)$ and $\Sigma^*(\theta)$ in companion form:

$$x_{t+1} = \begin{bmatrix} \Phi^{*'} \\ M' \end{bmatrix} x_t + M u_t.$$

The VAR implied covariance matrix, say Γ^V for brevity, has to satisfy the following relationship

$$\Gamma^V = \begin{bmatrix} \Phi^{*'} \\ M' \end{bmatrix} \Gamma^V \Big[\Phi^*,\, M\Big] + M\Sigma^* M' \tag{6}$$

We will now verify that $\Gamma^V = \Gamma_{XX}$ does satisfy this relationship. Note that

$$\begin{aligned} \Gamma_{XX} &= \begin{bmatrix} \mathbb{E}_\theta^D[y_t y_t'] & \mathbb{E}_\theta^D[y_t y_{t-1}'] \\ \mathbb{E}_\theta^D[y_{t-1} y_t'] & \mathbb{E}_\theta^D[y_t y_t'] \end{bmatrix} \\ \Gamma_{XY} &= \begin{bmatrix} \mathbb{E}_\theta^D[y_{t-1} y_t'] \\ \mathbb{E}_\theta^D[y_{t-2}, y_t'] \end{bmatrix}. \end{aligned} \tag{7}$$

Hence, (6) is satisfied if

$$\Gamma_{XX} = \begin{bmatrix} \Phi^{*'}\Gamma_{XX}\Phi & \Phi^{*'}\Gamma_{XX}M \\ M'\Gamma_{XX}\Phi & M'\Gamma_{XX}M \end{bmatrix} + M\Sigma^* M'.$$

Plugging in the definition of $\Phi^* = \Gamma_{XX}^{-1}\Gamma_{XY}$, the condition can be rewritten as

$$\Gamma_{XX} = \begin{bmatrix} \Gamma_{YX}\Gamma_{XX}^{-1}\Gamma_{XY} & \Gamma_{YX}M \\ M'\Gamma_{XY} & M'\Gamma_{XX}M \end{bmatrix} + M(\Gamma_{YY} - \Gamma_{YX}\Gamma_{XX}^{-1}\Gamma_{XY})M'.$$

Using the expressions for Γ_{XX} and Γ_{XY} in (7) it is straightforward to verify that $\Gamma^V = \Gamma_{XX}$ indeed solves (6), which completes the proof. ■

3 DSGE-VAR Inference

3.1 Prior Distribution

Let $\text{etr}[A] = \exp[-\frac{1}{2}tr[A]]$. The VAR likelihood function is given by

$$\begin{aligned} & p(Y|\Psi, \Sigma, \theta) \\ &= (2\pi)^{-nT/2}|\Sigma|^{-T/2}\text{etr}\Big[\Sigma^{-1}(Y - XB_1(\theta) - X\Psi B_2(\theta))'(Y - XB_1(\theta) - X\Psi B_2(\theta))\Big]. \end{aligned} \tag{8}$$

To capture misspecification, for now only in the conditional mean dynamics, we introduce a discrepancy matrix Ψ^{Δ}, such that $\Psi = \Psi^* + \Psi^{\Delta}$. Now consider the likelihood ratio (omitting the θ-argument of the $B_i(\theta)$ functions)

$$\begin{aligned}
&\ln\left[\frac{p(Y|\Psi^*,\Sigma^*,\theta)}{p(Y|\Psi^*+\Psi^{\Delta},\Sigma^*,\theta)}\right] \\
&= -\frac{1}{2}tr\left[\Sigma^{*-1}(Y-XB_1-X\Psi^*B_2)'(Y-XB_1-X\Psi^*B_2)\right] \\
&\quad +\frac{1}{2}tr\left[\Sigma^{*-1}(Y-XB_1-X\Psi^*B_2-X\Psi^{\Delta}B_2)'(Y-XB_1-X\Psi^*B_2-X\Psi^{\Delta}B_2)\right] \\
&= -\frac{1}{2}tr\left[\Sigma^{*-1}B_2'\Psi^{\Delta'}X'X\Psi^{\Delta}B_2\right] + tr\left[\Sigma^{*-1}B_2'\Psi^{\Delta'}X'(Y-XB_1-X\Psi^*B_2)\right]
\end{aligned}$$

Now suppose we are taking expectations with respect to the probability distribution implied by the DSGE model. Consider the following term

$$\begin{aligned}
\mathbb{E}_{\theta}^{D}\left[X'(Y-XB_1-X\Psi^*B_2)\right] &= \Gamma_{XY}(\theta)-\Gamma_{XX}(B_1+\Psi^*B_2) \\
&= \Gamma_{XY}(\theta)-\Gamma_{XX}\Phi^* \\
&= \Gamma_{XY}(\theta)-\Gamma_{XX}\Gamma_{XX}^{-1}\Gamma_{XY}=0.
\end{aligned}$$

The second equality follows from the definition of Φ^* and the last equality is a consequence of Lemma 1(i). Hence,

$$\mathbb{E}_{\theta}^{D}\left[\ln\left[\frac{p(Y|\Psi^*,\Sigma^*,\theta)}{p(Y|\Psi^*+\Psi^{\Delta},\Sigma^*,\theta)}\right]\right] = -\frac{1}{2}tr\left[\Sigma^{*-1}B_2'(\theta)\Psi^{\Delta'}\Gamma_{XX}(\theta)\Psi^{\Delta}B_2(\theta)\right].$$

We now choose a prior density for Ψ^{Δ} that is proportional to the Kullback-Leibler discrepancy:

$$p(\Psi^{\Delta}|\Sigma^*,\theta) \propto \text{etr}\left[\lambda T\Sigma^{*-1}\left(B_2'\Psi^{\Delta'}\Gamma_{XX}\Psi^{\Delta}B_2\right)\right]. \tag{9}$$

For computational reasons it is convenient to transform this prior into a prior for Ψ. Using standard arguments we deduce that this prior is multivariate normal

$$\Psi|\Sigma^*,\theta \sim \mathcal{N}\left(\Psi^*(\theta),\frac{1}{\lambda T}\left[(B_2(\theta)\Sigma^{*-1}B_2(\theta)')\otimes\Gamma_{XX}(\theta)\right]^{-1}\right). \tag{10}$$

In practice we also have to take potential misspecification of the covariance matrix $\Sigma^*(\theta)$ into account. Hence, we will use the following, slightly modified, prior distribution conditional on θ in the empirical analysis:

$$\begin{aligned}
\Psi|\Sigma,\theta &\sim \mathcal{N}\left(\Psi^*(\theta),\frac{1}{\lambda T}\left[(B_2(\theta)\Sigma^{-1}B_2(\theta)')\otimes\Gamma_{XX}(\theta)\right]^{-1}\right) \qquad (11)\\
\Sigma|\theta &\sim \mathcal{IW}\Big(\lambda T\Sigma^*(\theta),\lambda T-k,n\Big),
\end{aligned}$$

where $\mathcal{IW}$ denotes the inverted Wishart distribution. The latter induces a distribution for the discrepancy $\Sigma^{\Delta} = \Sigma - \Sigma^*$.

We conduct a few additional manipulations that will be useful subsequently. To simplify notation the (θ)-argument of the functions B_1, B_2, Γ_{XY}, Γ_{XX}, and Γ_{YY} is omitted. First, note that the prior density for Σ is given by

$$p(\Sigma|\theta) \propto |\Sigma|^{-\frac{1}{2}(\lambda T-k+n+1)}\text{etr}\left[\Sigma^{-1}\lambda T\Sigma^*(\theta)\right] \tag{12}$$

The prior density for Ψ is of the form

$$\begin{aligned} p(\Psi|\Sigma,\theta) &\propto \text{etr}\left[\lambda T B_2\Sigma^{-1}B_2'(\Psi-\Psi^*)'\Gamma_{XX}(\Psi-\Psi^*)\right] \\ &\propto \text{etr}\left[\lambda T\Sigma^{-1}B_2'\Psi'\Gamma_{XX}\Psi B_2\right] \\ &\quad -2\text{etr}\left[\lambda T\Sigma^{-1}B_2'\Psi'\Gamma_{XX}\Psi^* B_2\right] \end{aligned}$$

It will be convenient for the derivation of the conditional posteriors to rewrite the second term as follows. Recall that

$$y_t' - x_t'B_1 = x_t'\Psi^* B_2 + u_t'$$

Multiplying both sides by x_t and taking expectations under the distribution generated by the DSGE model yields

$$\Gamma_{XY} - \Gamma_{XX}B_1 = \Gamma_{XX}\Psi^* B_2.$$

If we adopt the notation that $\tilde{Y} = Y - XB_1$ and let

$$\begin{aligned} \Gamma_{\tilde{Y}\tilde{Y}} &= \Gamma_{YY} - \Gamma_{YX}B_1 - B_1'\Gamma_{XY} + B_1'\Gamma_{XX}B_1 \\ \Gamma_{X\tilde{Y}} &= \Gamma_{XY} - \Gamma_{XX}B_1. \end{aligned}$$

then we obtain

$$p(\Psi|\Sigma,\theta) \propto \text{etr}\left[\lambda T\Sigma^{-1}\left(B_2'\Psi'\Gamma_{XX}\Psi B_2 - 2B_2'\Psi'\Gamma_{X\tilde{Y}}\right)\right]. \tag{13}$$

Note that (13) implies that we can rewrite the prior mean of Ψ as

$$\Psi^*(\theta) = \bar{\Psi}(\Sigma,\theta) = \Gamma_{XX}^{-1}(\theta)\Gamma_{X\tilde{Y}}(\theta)\Sigma^{-1}B_2'(\theta)[B_2(\theta)\Sigma^{-1}B_2'(\theta)]^{-1}.$$

3.2 Likelihood Function

The likelihood function for the VAR representation was given in (8). It is instructive to factorize the joint distribution of y_t into a marginal distribution of $y_{2,t}$ and a conditional

distribution of $y_{1,t}$ given $y_{2,t}$. Under this factorization, the likelihood function can be expressed as

$$\begin{aligned} & p(Y|\Psi,\Sigma,\theta) \\ \propto\ & |\Sigma_{22}|^{-T/2}\text{etr}\Big[\Sigma^{-1}(Y_2-X\Psi)'(Y_2-X\Psi)\Big] \\ & \times|\Sigma_{11.22}|^{-T/2}\text{etr}\Big[\Sigma_{11.22}^{-1}(Y_1-X\beta_1-X\Psi\beta_2-(Y_2-X\Psi)\Sigma_{22}^{-1}\Sigma_{21})' \\ & \times(Y_1-X\beta_1-X\Psi\beta_2-(Y_2-X\Psi)\Sigma_{22}^{-1}\Sigma_{21})\Big] \end{aligned} \tag{14}$$

where

$$\Sigma_{11.22}=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

The factorization shows that the likelihood function generates an endogeneity correction for the endogenous regressors in the monetary policy rule.

3.3 Posterior

We will now discuss the conditional distributions for the Gibbs Sampler.

Conditional Posterior of Ψ: Combining the prior density (13) with the likelihood function (8) yields

$$\begin{aligned} & p(\Psi|\Sigma,\theta,Y) \\ \propto\ & p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta) \\ \propto\ & \text{etr}\Big[\Sigma^{-1}\lambda T\Big(-2B_2'\Psi'\Gamma_{X\tilde{Y}}+B_2'\Psi'\Gamma_{XX}(\theta)\Psi B_2\Big)+(\tilde{Y}-X\Psi B_2)'(\tilde{Y}-X\Psi B_2)\Big] \\ \propto\ & \text{etr}\Big[\Sigma^{-1}\Big(-2B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y})+B_2'\Psi'(\lambda T\Gamma_{XX}+X'X)\Psi B_2\Big)\Big] \end{aligned} \tag{15}$$

Define

$$\tilde{\Psi}(\Sigma,\theta)=(\lambda T\Gamma_{XX}+X'X)^{-1}(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y})\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}.$$

The previous calculations show that

$$\Psi|\Sigma,\theta,Y\sim\mathcal{N}\Big(\tilde{\Psi}(\Sigma,\theta),\Big[(B_2\Sigma^{-1}B_2')\otimes(\lambda T\Gamma_{XX}+X'X)\Big]^{-1}\Big). \tag{16}$$

Conditional Posterior of Σ: Combining the prior densities (13) and (12) with the likeli-

hood function (8) yields

$$\begin{aligned} p(\Sigma|\Psi,\theta,Y) &\propto p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta)p(\Sigma|\theta) \\ &\propto |\Sigma|^{-\frac{1}{2}((\lambda+1)T-k+n+1)}|(B_2\Sigma^{-1}B_2')^{-1}|^{-\frac{k}{2}} \\ &\quad \text{etr}\Big[\Sigma^{-1}\Big(\lambda T(\Gamma_{\tilde{Y}\tilde{Y}}-\Gamma_{\tilde{Y}X}\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}})+(\tilde{Y}-X\Psi B_2)'(\tilde{Y}-X\Psi B_2)\Big) \\ &\quad +\lambda T(B_2\Sigma^{-1}B_2')(\Psi-\bar{\Psi})'\Gamma_{XX}(\Psi-\bar{\Psi})\Big]. \end{aligned} \tag{17}$$

Using the definition of $\bar{\Psi}$, the last term can be manipulated as follows:

$$\begin{aligned} &\text{etr}\Big[\lambda T B_2\Sigma^{-1}B_2'(\Psi-\bar{\Psi})'\Gamma_{XX}(\Psi-\bar{\Psi})\Big] \\ &= \text{etr}\Big[\lambda T\Sigma^{-1}\Big(B_2'\Psi'\Gamma_{XX}\Psi B_2-2B_2'\Psi'\Gamma_{X\tilde{Y}}\Big) \\ &\quad +\lambda T\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}B_2\Sigma^{-1}\Gamma_{X\tilde{Y}}'\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}}\Big] \end{aligned}$$

Hence,

$$\begin{aligned} p(\Sigma|\Psi,\theta,Y) &\propto |\Sigma|^{-\frac{1}{2}((\lambda+1)T-k+n+1)}|(B_2\Sigma^{-1}B_2')^{-1}|^{-\frac{k}{2}} \\ &\quad \times \text{etr}\Big[\Sigma^{-1}\Big(\lambda T\Gamma_{\tilde{Y}\tilde{Y}}+\tilde{Y}'\tilde{Y}-2B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y}) \\ &\quad +B_2'\Psi'(\lambda T\Gamma_{XX}+X'X)\Psi B_2\Big)\Big] \\ &\quad \times \text{etr}\Big[\lambda T(\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}B_2\Sigma^{-1}-\Sigma^{-1})\Gamma_{X\tilde{Y}}'\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}}\Big]. \end{aligned} \tag{18}$$

If the DSGE model satisfies Equation (1) and the error $u_{1,t}$ is orthogonal to x_t then

$$\Gamma_{X\tilde{Y}} = \Gamma_{XX}\Psi^* B_2$$

and

$$(\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}B_2\Sigma^{-1}-\Sigma^{-1})\Gamma_{X\tilde{Y}}'\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}} = 0. \tag{19}$$

While the conditional posterior distribution of Σ given our prior distribution is not of the $\mathcal{IW}$ form, use an $\mathcal{IW}$ distribution as proposal distribution in a Metropolis-Hastings step. Define

$$\begin{aligned} \tilde{S}(\Psi,\theta) &= \lambda T\Gamma_{\tilde{Y}\tilde{Y}}+\tilde{Y}'\tilde{Y}-(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y})'\Psi B_2-B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y}) \\ &\quad +B_2'\Psi'(\lambda T\Gamma_{XX}+X'X)\Psi B_2 \end{aligned} \tag{20}$$

Our proposal distribution for Σ is

$$\mathcal{IW}(\tilde{S}(\Psi,\theta),(\lambda+1)T,n).$$

Conditional Posterior of θ: The posterior distribution of θ is irregular. Its density is proportional to the joint density of Y, Ψ, Σ, and θ, which we can evaluate numerically since the normalization constants for $p(\Psi|\Sigma,\theta)$ and $p(\Sigma|\theta)$ are readily available.

$$p(\theta|\Psi,\Sigma,Y) \propto p(Y,\Psi,\Sigma,\theta) = p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta)p(\Sigma|\theta)p(\theta). \tag{21}$$

To obtain a proposal density for $p(\theta|\Psi,\Sigma,Y)$ we (i) maximize the posterior density of the DSGE model with respect to θ and (ii) calculate the inverse Hessian at the mode, denoted by $V_{\bar{\theta},DSGE}$. (iii) We then use a random-walk Metropolis step with proposal density

$$\mathcal{N}(\theta_{(s-1)}, cV_{\bar{\theta},DSGE})$$

where $\theta_{(s-1)}$ is the value of θ drawn in iteration $s-1$ of the MCMC algorithm, and c is a scaling factor that can be used to control the rejection rate in the Metropolis step.

3.4 Policy Analysis

In addition to the direct analysis with the DSGE model the paper considers three modes of policy analysis: *backward looking*, *acknowledge misspecification*, *policy-invariant misspecification*.

Backward-looking Analysis: The model developed in the paper takes the form

$$\begin{aligned} y_{1,t} &= x_t'\beta_1(\theta_{(p)}) + x_t'\Psi\beta_2(\theta_{(p)}) + u_{1,t} \\ y_{2,t}' &= x_t'\Psi + u_{2,t}', \end{aligned} \tag{22}$$

where

$$u_{1,t} = u_{2,t}'\beta_2(\theta_{(p)}) + \tilde{\epsilon}_{1,t}\sigma_R. \tag{23}$$

According to the underlying structural model, the one-step-ahead forecast errors $u_{2,t}$ are a function of the monetary policy shock $\tilde{\epsilon}_{1,t}$ and the other structural shocks $\tilde{\epsilon}_{2,t}$. Hence, we express $u_{2,t}$ as

$$u_{2,t}' = \tilde{\epsilon}_{1,t}\sigma_R A_1(\theta_{(p)},\Sigma) + \tilde{\epsilon}_{2,t}' A_2(\theta_{(p)},\Sigma). \tag{24}$$

Combining (23) and (24), we obtain

$$\begin{aligned} u_{1,t} &= \tilde{\epsilon}_{1,t}\sigma_R(1 + A_1\beta_2) + \tilde{\epsilon}_{2,t}' A_2\beta_2 \\ u_{2,t} &= \tilde{\epsilon}_{1,t}\sigma_R A_1 + \tilde{\epsilon}_{2,t}' A_2 \end{aligned}$$

Hence, the partitioned covariance matrix of the reduced-form innovations can be expressed as

$$\begin{aligned}
\Sigma_{11} &= \sigma_R^2(1+\beta_2'A_1')(1+A_1\beta_2)+\beta_2'A_2'A_2\beta_2 \\
\Sigma_{12} &= \sigma_R^2(1+\beta_2'A_1')A_1+\beta_2'A_2'A_2 \\
\Sigma_{22} &= \sigma_R^2A_1'A_1+A_2'A_2.
\end{aligned}$$

We deduce that

$$\begin{aligned}
\sigma_R^2A_1 &= \Sigma_{12}-\beta_2'\Sigma_{22} \\
\sigma_R^2 &= \Sigma_{11}-\beta_2'\Sigma_{22}\beta_2-2(\Sigma_{12}-\beta_2'\Sigma_{22})\beta_2,
\end{aligned}$$

which leads to the following formulas for the effect of the structural shocks on $u_{2,t}'$:

$$A_1 = \left[\Sigma_{11}-\beta_2'\Sigma_{22}\beta_2-2(\Sigma_{12}-\beta_2'\Sigma_{22})\beta_2\right]^{-1}(\Sigma_{12}-\beta_2'\Sigma_{22}) \tag{25}$$

$$A_2'A_2 = \Sigma_{22}-A_1'\left[\Sigma_{11}-\beta_2'\Sigma_{22}\beta_2-2(\Sigma_{12}-\beta_2'\Sigma_{22})\beta_2\right]A_1 \tag{26}$$

Following Sims (1986) we now rewrite our model as structural VAR with interactions between contemporaneous variables on the left-hand-side. Our policy rule is of the form

$$y_{1,t} = x_t'\beta_1+y_{2,t}'\beta_2+\widetilde{\epsilon}_{1,t}\sigma_R. \tag{27}$$

The private sector equations are given by

$$\begin{aligned}
y_{2,t}' &= x_t'\Psi+u_{2,t}' \\
&= x_t'\Psi+\widetilde{\epsilon}_{1,t}\sigma_RA_1+\widetilde{\epsilon}_{2,t}'A_2 \\
&= x_t'\Psi+(y_{1,t}-x_t'\beta_1-y_{2,t}'\beta_2)A_1+\widetilde{\epsilon}_{2,t}'A_2
\end{aligned}$$

Hence,

$$[y_{1,t},y_{2,t}']\begin{bmatrix} I & -A_1 \\ -\beta_2 & I+\beta_2A_1 \end{bmatrix} = x_t'[\beta_1,\Psi-\beta_1A_1]+[\widetilde{\epsilon}_{1,t}\sigma_R,\widetilde{\epsilon}_{2,t}'A_2] \tag{28}$$

In our backward-looking analysis we are only changing the coefficients in the policy rule equation when conducting the policy experiment. The private sector equations will remain unchanged. Mechanically, we implement the analysis with the following computation:

For each value of $\tilde{\theta}_{(p)} \in \Theta_{(p)}$, where $\Theta_{(p)}$ is a grid of policy parameters, and each posterior draw of the triplet (θ,Φ,Σ):

1. Compute β_1, β_2, A_1 and $A_2'A_2$ based on (25) and (26) from the posterior draw of (θ,Φ,Σ).

2. Calculate the unconditional variance of y_t based on the system

$$[y_{1,t}, y'_{2,t}] \begin{bmatrix} I & -A_1 \\ -\beta_2(\tilde{\theta}_{(p)}) & I + \beta_2 A_1 \end{bmatrix} = x'_t[\beta_1(\tilde{\theta}_{(p)}), \Psi - \beta_1 A_1] + [\widetilde{\epsilon}_{1,t}\sigma_R, \widetilde{u}'_{2,t} A_{2,tr}]$$

where $A_{2,tr}$ is the Cholesky factor of $A'_2 A_2$ and $\widetilde{u}_{2,t}$ is a vector of innovations with unit variance that are orthogonal to the monetary policy shock $\widetilde{\epsilon}_{1,t}$.

Acknowledge Misspecification: Starting from the forward-looking analysis we re-introduce the misspecification matrices Ψ^Δ and Σ^Δ into the policy analysis step. Hence, we use

$$\begin{aligned} y_{1,t} &= x'_t\beta_1(\tilde{\theta}_{(p)}) + x'_t\big(\Psi^*(\tilde{\theta}_{(p)}, \theta_{(np)}) + \Psi^\Delta\big)\beta_2(\tilde{\theta}_{(p)}) + u_{1,t} \\ y'_{2,t} &= x'_t\big(\Psi^*(\tilde{\theta}_{(p)}, \theta_{(np)}) + \Psi^\Delta\big) + u'_{2,t}, \end{aligned} \tag{29}$$

and the covariance matrix of u_t is given by $\Sigma^*(\tilde{\theta}_{(p)}, \theta_{(np)}) + \Sigma^\Delta$. For each tuplet $(\tilde{\theta}_{(p)}, \theta_{(np)}, \Psi, \Sigma)$ we can compute A_1 and A_2 according to (25) and (26), let

$$\begin{aligned} u_{1,t} &= \epsilon_{1,t}\sigma_R(1 + A_1\beta_2) + \widetilde{\epsilon}'_{2,t} A_2\beta_2 \\ u_{2,t} &= \epsilon_{1,t}\sigma_R A_1 + \widetilde{\epsilon}'_{2,t} A_2 \end{aligned}$$

and $\sigma_R = 0$ to eliminate the monetary policy shock. In the absence of a firm theory that explains how the discrepancy matrices respond to policy changes, we use the prior distribution to characterize beliefs about post-intervention model misspecification.

Policy-Invariant Misspecification We assume that the estimated discrepancy, in terms of impulse response functions, is policy invariant. For the impulse response functions to be interpretable, it is useful to apply an identification scheme that links them to the structural shocks in the underlying DSGE model. Recall that the monetary policy shock has been identified through an exclusion restriction. However, we still have to identify the matrix A_2 in (24). We follow the approach taken in Del Negro and Schorfheide (2004). Let $A'_{2,tr}A_{2,tr} = A'_2 A_2$ be the Cholesky decomposition of $A'_2 A_2$. The relationship between $A_{2,tr}$ and A_2 is given by $A'_2 = A'_{2,tr}\Omega$, where Ω is an orthonormal matrix that is not identifiable based on the estimates of $\beta(\theta)$, Ψ, and Σ. However, we are able to calculate an initial effect of $\epsilon_{2,t}$ on $y_{2,t}$ based on the DSGE model, denoted by $A_2^D(\theta)$. This matrix can be uniquely decomposed into a lower triangular matrix and an orthonormal matrix:

$$A_2^{D'}(\theta) = A_{2,tr}^{D'}(\theta)\Omega^*(\theta).$$

To identify A_2 above, we combine $A'_{2,tr}$ with $\Omega^*(\theta)$. The calculation is easily implementable in a Markov Chain Monte Carlo analysis. For every draw of $(\theta, \Psi^\Delta, \Sigma^\Delta)$ from their joint posterior distribution we compute $\Omega^*(\theta)$ and A_2.

In order to implement the policy analysis, we use posterior draws of (θ, Ψ, Σ) to create two moving average representations for $y_{2,t}$:

$$\begin{aligned}\sum_{j=0}^{\infty} \widetilde{D}_j^*(\theta) u_{2,t} &= \sum_{j=0}^{\infty} \widetilde{D}_j^*(\theta)\big(A_1^*(\theta)'\tilde{\epsilon}_{1,t} + A_2^*(\theta)'\tilde{\epsilon}_{2,t}\big) \\ \sum_{j=0}^{\infty} \widetilde{D}_j(\Psi) u_{2,t} &= \sum_{j=0}^{\infty} \widetilde{D}_j\big(A_1(\theta,\Sigma)'\tilde{\epsilon}_{1,t} + A_2(\theta,\Sigma)'\tilde{\epsilon}_{2,t}\big)\end{aligned}$$

The first representation is calculated from the VAR approximation of the DSGE model $\Psi^*(\theta)$ and $\Sigma^*(\theta)$. The second representation is obtained from the estimated DSGE-VAR specification. The impulse response function discrepancies (DSGE-VAR($\hat{\lambda}$) versus DSGE-VAR(∞)) are given by

$$IRF_j^{\Delta} = \widetilde{D}_j(\Psi)\big[A_1(\theta,\Sigma)',\ A_2(\theta,\Sigma)'\big] - \widetilde{D}_j^*(\theta)\big[A_1^*(\theta)',\ A_2^*(\theta)'\big].$$

We consider the following post-intervention law of motion for $y_{2,t}$:

$$y_{2,t} = \sum_{j=0}^{\infty}\left[\widetilde{D}_j^*(\tilde{\theta}_{(p)},\theta_{(np)})\big[A_1^*(\tilde{\theta}_{(p)},\theta_{(np)})',\ A_2^*(\tilde{\theta}_{(p)},\theta_{(np)})'\big] + IRF_j^{\Delta}\right]\begin{bmatrix}\tilde{\epsilon}_{1,t}\\ \tilde{\epsilon}_{2,t}\end{bmatrix}. \tag{30}$$

3.5 Measurement Equation

The relationships between the deviations from steady state that appear in the DSGE model description and the observables y_t are given by the following measurement equation:

$$y_{1,t} = r_*^a + 400\gamma + \pi_*^a + 4R_t, \quad y_{2,t} = \begin{bmatrix}\pi_*^a + 4\pi_t \\ \widetilde{y}_t \\ 100\ln(1-\alpha)/(1+\lambda_f) + lsh_t \\ L_t\end{bmatrix}. \tag{31}$$

Here, we have partitioned y_t such that $y_{1,t}$ corresponds to the policymaker's instrument (the interest rate), and $y_{2,t}$ is a vector that includes the remaining four observables. The steady state (net) real interest rate in our model is given by $r_*^a + 400\gamma$. The parameter r_*^a is related to the discount rate β according to $\beta = 1/(1 + r_*^a/400)$. The monetary policy rule can be rewritten in terms of observables as follows:

$$y_{1,t} = (1-\rho_R)[(r_*^a + 400\gamma + \pi_a^*) - \psi_1 \ln \pi_a^*] + y_{1,t-1}\rho_R + y_{2,t}'\begin{bmatrix}(1-\rho_R)\psi_1 \\ 4(1-\rho_R)\psi_2 \\ 0 \\ 0\end{bmatrix} + \sigma_R\epsilon_{1,t}. \tag{32}$$

4 A Simple Example

To gain some intuition for the empirical results reported in the paper and the proposed approaches to policy analysis we consider the following simplified DSGE model:

$$R_t = \psi_1 \pi_t + \psi_2 \widetilde{y}_t + \sigma_R \widetilde{\epsilon}_{1,t} \tag{33}$$

$$\widetilde{y}_t = \mathbb{E}_t[\widetilde{y}_{t+1}] - (R_t - \mathbb{E}_t[\pi_{t+1}]) \tag{34}$$

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa(\widetilde{y}_t - z_t) \tag{35}$$

$$z_t = \rho_z z_{t-1} + \sigma_z \widetilde{\epsilon}_{z,t} \tag{36}$$

To obtain straightforward analytical solutions, we assume that the central bank has historically followed the policy $\psi_1 = 1/\beta$. From (35) we obtain

$$\mathbb{E}_t[\pi_{t+1}] - \frac{1}{\beta}\pi_t = -\frac{\kappa}{\beta}(\widetilde{y}_t - z_t)$$

which can be combined with (33) and (34)

$$(1+\psi_2)\widetilde{y}_t = \mathbb{E}_t[\widetilde{y}_{t+1}] - \frac{\kappa}{\beta}(\widetilde{y}_t - z_t) - \sigma_R \widetilde{\epsilon}_{1,t}.$$

Hence,

$$\widetilde{y}_t = \frac{1}{1+\psi_2+\kappa/\beta}\mathbb{E}_t[\widetilde{y}_{t+1}] + \frac{\kappa/\beta}{1+\psi_2+\kappa/\beta} z_t - \frac{\sigma_R}{1+\psi_2+\kappa/\beta}\widetilde{\epsilon}_{R,t}. \tag{37}$$

Solving (37) forward leads to

$$\widetilde{y}_t = \frac{\kappa/\beta}{1-\rho_z+\psi_2+\kappa/\beta} z_t - \frac{\sigma_R}{1+\psi_2+\kappa/\beta}\widetilde{\epsilon}_{R,t}. \tag{38}$$

Marginal costs in the simple model are given by

$$mc_t = \widetilde{y}_t - z_t = -\frac{1-\rho_z+\psi_2}{1-\rho_z+\psi_2+\kappa/\beta} z_t - \frac{\sigma_R}{1+\psi_2+\kappa/\beta}\widetilde{\epsilon}_{R,t}. \tag{39}$$

According to (35) inflation is the sum of discounted future marginal costs:

$$\pi_t = -\kappa \frac{1-\rho_z+\psi_2}{(1-\rho_z+\psi_2+\kappa/\beta)(1-\beta\rho_z)} z_t - \kappa \frac{\sigma_R}{1+\psi_2+\kappa/\beta}\widetilde{\epsilon}_{R,t} \tag{40}$$

Hence, the law of motion of output and inflation is given by (36), (38), and (40), which corresponds to a restricted moving average representation of $\widetilde{y}_t$ and π_t in terms of $\widetilde{\epsilon}_{1,t}$ and $\widetilde{\epsilon}_{z,t}$. The effects of policy changes, e.g. changes in ψ_2 on the variability of output and inflation can be calculated from (38) and (40). In the absence of a monetary policy shock, the model can be simplified further. We deduce from quasi-differencing of (38) and (40) that

$$\widetilde{y}_t = \rho_z \widetilde{y}_{t-1} + \frac{\kappa/\beta}{1-\rho_z+\psi_2+\kappa/\beta}\sigma_z \widetilde{\epsilon}_{z,t} \tag{41}$$

$$\pi_t = \rho_z \pi_{t-1} - \kappa \frac{1-\rho_z+\psi_2}{(1-\rho_z+\psi_2+\kappa/\beta)(1-\beta\rho_z)}\sigma_z \widetilde{\epsilon}_{z,t} \tag{42}$$

which implies

$$\mathbb{E}_t[\widetilde{y}_{t+1}] = \rho_z \widetilde{y}_t, \quad \mathbb{E}_t[\pi_{t+1}] = \rho_z \pi_t. \tag{43}$$

In general, we regard the cross-equation restrictions of the DSGE model as potentially misspecified and use the DSGE-VAR framework to relax these restrictions. The DSGE-VAR approximates (38) and (40) by a VAR, using lagged output, inflation, and interest rates as right-hand-side variables.

To gain insights into the backward-looking policy analysis, we replace the conditional expectations in (34) and (35) by (43). Hence we can write our model in backward looking form as

$$\begin{aligned} R_t &= \psi_1 \pi_t + \psi_2 \widetilde{y}_t && (44) \\ \widetilde{y}_t &= \rho_z \widetilde{y}_t - (R_t - \rho_z \pi_t) && (45) \\ \pi_t &= \beta \rho_z \pi_t + \kappa(\widetilde{y}_t - z_t) && (46) \\ z_t &= \rho_z z_{t-1} + \sigma_z \widetilde{\epsilon}_{z,t} && (47) \end{aligned}$$

We proceed by manipulating (45):

$$\begin{aligned} \widetilde{y}_t &= \rho_z \widetilde{y}_t - (\psi_1 \pi_t + \psi_2 \widetilde{y}_t - \rho_z \pi_t) \\ &= -\frac{\psi_1 - \rho_z}{1 - \rho_z + \psi_2} \pi_t \end{aligned} \tag{48}$$

Quasi-differencing (46) yields

$$(1 - \beta\rho_z)\pi_t - \kappa \widetilde{y}_t = \rho_z (1 - \beta\rho_z)\pi_{t-1} - \rho_z \kappa \widetilde{y}_{t-1} - \kappa \sigma_z \widetilde{\epsilon}_{z,t}.$$

We can now substitute $\widetilde{y}_t$ and $\widetilde{y}_{t-1}$ and express inflation as an AR(1) process:

$$\pi_t = \rho_z \pi_{t-1} - \kappa \frac{1 - \rho_z + \psi_2}{(1 - \beta\rho_z)(1 - \rho_z + \psi_2) + (\psi_1 \beta - \rho_z \beta)\kappa/\beta} \sigma_z \widetilde{\epsilon}_{z,t}. \tag{49}$$

Notice Equations (42) and (49) are identical for $\psi_1 = 1/\beta$, regardless of ψ_2. The reason is that for $\psi_1 = 1/\beta$ expectations of future output and inflation do not depend on the policy rule parameter ψ_2. Hence, policy analysis in terms of ψ_2 with the backward-looking approximation of the DSGE model will yield the same predictions as the forward-looking analysis. We can see from (49) that an increase in ψ_1 lowers inflation and hence output volatility.

5 Prior Distribution

We are using dogmatic priors for three of the DSGE model parameters: the capital depreciation rate $\delta = 0.025$, growth rate of technology $\gamma = 1.5$, and the steady state mark-up $\lambda_f = 0.3$.

Prior distributions for the remaining DSGE model parameters are summarized in Table 1. All intervals reported below are 90% credible. The distribution for ψ_1 and ψ_2 is approximately centered at Taylor's (1993) values, whereas the smoothing parameter lies in the range from 0.18 to 0.83. The prior mean for the growth-adjusted real interest rate, $r_*^a + \gamma_a$, is 2.5% and annualized steady state inflation ranges from 1 to 7%, which is consistent with pre-1982 long-run historical averages. The prior mean of g^* implies that that the government share of GDP is 20%. According to our prior the habit persistence parameter h lies between 0.6 and 0.8. Boldrin, Christiano, and Fisher (2001) found that a value of 0.7 enhances the ability of a standard DSGE model to account for key asset market statistics. The interval for 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 (2005) at the upper end.

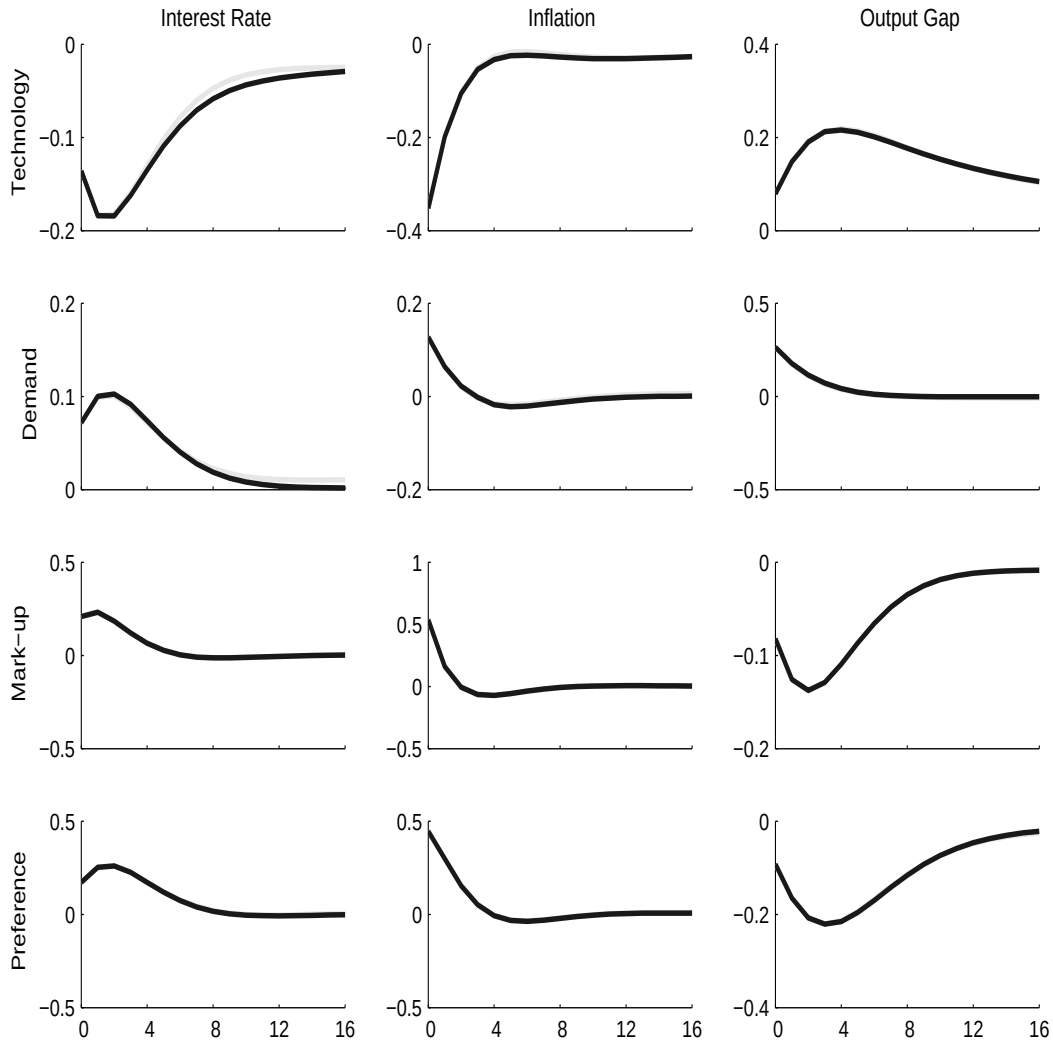
According to the prior for ζ_p, firms re-optimize their prices, on average every 1.5 to 6.5 quarters. This interval encompasses findings in micro-level studies of price adjustments such as Bils and Klenow (2004). The prior for the adjustment cost parameter s'' is consistent with the values that Christiano, Eichenbaum, and Evans (2005) report when matching consumption and investment DSGE impulse response functions, among others, to VAR responses. Our prior for a' implies that in response to a 1% increase in the return to capital, utilization rates rise by 0.15 to 0.4%. These numbers are considerably smaller than the one used by Christiano, Eichenbaum, and Evans (2005). Finally, the priors for the autocorrelations of the exogenous shocks are centered at 0.75 with a standard deviation of 0.1. The priors for the standard deviation parameters are chosen to obtain realistic magnitudes for the implied volatility of the output gap, the labor share, hours worked, inflation, and interest rates.

6 Accuracy of VAR Approximation

To document the accuracy of the finite-order VAR approximation to the state-space representation of the linearized DSGE model the following figure compares DSGE and DSGE-

VAR(∞) impulse responses.

Figure 1: Impulse Responses: DSGE versus DSGE-VAR(∞)



Notes: The figure depicts impulse responses from the DSGE (gray) and the DSGE-VAR(∞) (black) based on the DSGE-VAR($\lambda = 0.75$) posterior estimates of the non-policy parameters $\theta_{(np)}$.

References

Bils, Mark, and Peter Klenow (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, **112**, 947-985.

Boldrin, Michele, Lawrence Christiano, and Jonas Fisher (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, **91**, 149-166.

Chang, Yongsung, and Sun-Bin Kim (2005): "From Individual to Aggregate Labor Supply: A Quantitative Analysis based on Heterogeneous Agent Macroeconomy," *International Economic Review*, forthcoming.

Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, **113**, 1-45.

Del Negro, Marco, and Frank Schorfheide (2004): "Priors from Equilibrium Models for VARs," *International Economic Review*, **45**, 643-673.

Kimball, Miles, and Matthew Shapiro (2003): "Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?" *Manuscript*, University of Michigan, Department of Economics.

Sims, Christopher A. (1986): "Are Forecasting Models Useable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review*, Winter, 2-16.

Table 1: Prior Distributions

Name	Domain	Density	P(1)	P(2)
Policy Rule Parameters				
ψ_1	$I\!R^+$	Gamma	1.500	0.400
ψ_2	$I\!R^+$	Gamma	0.200	0.100
ρ_R	$[0,1)$	Beta	0.500	0.200
Nominal Rigidities				
ζ_p	$[0,1)$	Beta	0.600	0.150
ζ_w	$[0,1)$	Beta	0.600	0.150
Preference Parameters				
h	$[0,1)$	Beta	0.700	0.050
ν_l	$I\!R^+$	Gamma	2.000	0.750
Technology Parameters				
α	$[0,1)$	Beta	0.350	0.050
s''	$I\!R^+$	Gamma	4.000	1.500
a''	$I\!R^+$	Gamma	0.300	0.075
Steady State Parameters				
r_*^a	$I\!R^+$	Gamma	1.000	0.400
π_a^*	$I\!R$	Normal	4.000	2.000
$g^* - 1$	$I\!R^+$	Gamma	0.250	0.100

Table 1: Prior Distributions (Continued)

Name	Domain	Density	P(1)	P(2)
Shock Autocorrelations				
ρ_z	$[0,1)$	Beta	0.750	0.100
ρ_g	$[0,1)$	Beta	0.750	0.100
ρ_{λ_f}	$[0,1)$	Beta	0.750	0.100
ρ_ϕ	$[0,1)$	Beta	0.750	0.100
ρ_μ	$[0,1)$	Beta	0.750	0.100
ρ_b	$[0,1)$	Beta	0.750	0.100
Shock Standard Deviations				
σ_R	$I\!R^+$	InvGamma	0.200	4.000
σ_z	$I\!R^+$	InvGamma	0.300	4.000
σ_g	$I\!R^+$	InvGamma	0.500	4.000
σ_{λ_f}	$I\!R^+$	InvGamma	0.150	4.000
σ_ϕ	$I\!R^+$	InvGamma	3.000	4.000
σ_μ	$I\!R^+$	InvGamma	0.750	4.000
σ_b	$I\!R^+$	InvGamma	0.750	4.000
Hyperparameters for AR(2) Shocks				
ξ_g	$[0,1)$	Beta	0.750	0.100
ω_g	$I\!R^+$	InvGamma	0.500	4.000

Notes: P(1) and P(2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. We are reporting annualized values for π^* and r^* (a-subscript). The following parameters were fixed: $\delta = 0.025$, $\gamma = 1.5$, $\lambda_f = 0.15$, and $\lambda_w = 0.3$. Parameters are assumed to be *a priori* independent. The effective prior is truncated at the boundary of the determinacy region.