Monetary Policy Analysis with Potentially Misspecified Models

Marco Del Negro
Federal Reserve Bank of Atlanta

Frank Schorfheide*
University of Pennsylvania

Preliminary Draft: October 7, 2004

Abstract

This paper proposes a novel method for conducting policy analysis with potentially misspecified dynamic stochastic general equilibrium (DSGE) models and applies it to a New Keynesian DSGE model along the lines of Christiano, Eichenbaum, and Evans (JPE 2004) and Smets and Wouters (JEEA 2003). Specifically, we are studying the effects of coefficient changes in interest-rate feedback rules on the volatility of output growth, inflation, and nominal rates. The paper illustrates the sensitivity of the results to assumptions on the policy invariance of model misspecifications. (JEL: C32)

JEL CLASSIFICATION: C32

KEY WORDS: Bayesian Analysis, DSGE Models, Model Misspecification

---

*Correspondence: M. Del Negro: Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St NE, Atlanta GA 30309-4470. Email: Marco.DelNegro@atl.frb.org. F. Schorfheide: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu. We would like to thank the participants of the 2004 EEA-ESEM session on "Empirical Models for Monetary Policy Analysis" for helpful comments. Schorfheide gratefully acknowledges financial support from the Alfred P. Sloan Foundation. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

1 Introduction

Despite recent successes in improving the empirical performance of dynamic stochastic general equilibrium (DSGE) models, e.g., Smets and Wouters (2003), even large-scale DSGE models suffer to some extent from misspecification (see Del Negro, Schorfheide, Smets, and Wouters 2004). In this paper misspecification means that the DSGE model potentially imposes invalid cross-coefficient restrictions on the moving-average representation of the macroeconomic time series that it aims to explain. As a consequence, one typically observes that the forecasting performance of DSGE models is worse than that of vector autoregressions (VARs) estimated with well-calibrated shrinkage methods. On the other hand, DSGE models have the advantage that one can explicitly assess the effect of policy regime changes on expectation formation and decision rules of private agents. Thus, policy analysis with DSGE models is robust to the Lucas critique and potentially more reliable than conclusions drawn from VARs. This trade-off poses a challenge to policymakers who want to use DSGE models in practice.

Del Negro and Schorfheide (2004a) proposed a framework that combines VARs and DSGE models, extending earlier work by Ingram and Whiteman (1994). In this framework DSGE model restrictions are neither completely ignored as in the unrestricted estimation of VARs, nor are they dogmatically imposed as in the direct estimation of DSGE models. Instead the VAR estimates are tilted toward the restrictions implied by the DSGE model, where the degree of tilting is determined by a Bayesian data-driven procedure that trades off model fit against complexity. Del Negro, Schorfheide, Smets, and Wouters (2004) show that priors arising from the same model used in this paper improve both the in-sample and out-of-sample fit of a VAR and lead to more accurate predictions than those directly obtained from the DSGE model.

In this paper we build upon our earlier work and further develop procedures that are suitable to study the effects of rare regime shifts with potentially misspecified DSGE models. These procedures can be viewed as a Bayesian alternative to the robust control and minimax approaches that recently have been proposed to cope with model misspecification, e.g., Hansen and Sargent (2000) and Onatsky and Stock (2002). One advantage of Bayesian procedures is that the policymaker can learn from existing data about the extent of the DSGE model's misspecification, and consequently adjust her policies. While a companion paper (Del Negro and Schorfheide, 2004b) applies these procedures to a simple three-equation New Keynesian model, the present paper studies policy experiments in the context of a large-scale model with capital accumulation as well as various nominal and real frictions. The DSGE model is based on work by Altig, Christiano, Eichenbaum, and Linde (2002), Christiano, Eichenbaum, and Evans (2004), and Smets and Wouters (2003). We show that conclusions about the effects of changing the response to inflation are to some extent sensitive to assumptions about the policy invariance of observed discrepancies between model and reality.

The paper is organized as follows. The DSGE model is presented in Section 2. In Section 3 we introduce our framework for policy analysis with potentially misspecified models. Section 4 describes the data set, Section 5 discusses our empirical findings, and Section 6 concludes.

2 Model

This section describes the DSGE model, which is a slightly modified version of the DSGE model developed and estimated for the Euro area in Smets and Wouters (2003). In particular, we introduce stochastic trends into the model, so that it can be fitted to unfiltered time series observations. The DSGE model, largely based on the work of Christiano, Eichenbaum, and Evans (2004), contains a large number of nominal and real frictions. Next, we describe each of the agents that populate the model economy and the decision problems they face.

2.1 Final goods producers

The final good Y_t is a composite made of a continuum of intermediate goods $Y_t(i)$, indexed by $i \in [0, 1]$:

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_{f,t}}} di \right]^{1+\lambda_{f,t}}. \tag{1}$$

$\lambda_{f,t} \in (0, \infty)$ follows the exogenous process:

$$\ln \lambda_{f,t} = \ln \lambda_f + \sigma_{\lambda,f} \epsilon_{\lambda,t}, \tag{2}$$

where $\epsilon_{\lambda,t}$ is an exogenous shock with unit variance. The final goods producers are perfectly competitive firms that buy intermediate goods, combine them to the final product Y_t, and resell the final good to consumers. The firms maximize profits

$$P_t Y_t - \int P_t(i) Y_t(i) di$$

subject to (1). Here P_t denotes the price of the final good and $P_t(i)$ is the price of intermediate good i. From their first order conditions and the zero-profit condition we obtain that:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\frac{1+\lambda_{f,t}}{\lambda_{f,t}}} Y_t \quad \text{and} \quad P_t = \left[\int_0^1 P_t(i)^{\frac{1}{\lambda_{f,t}}} di\right]^{\lambda_{f,t}}. \tag{3}$$

2.2 Intermediate goods producers

Good i is made using the technology:

$$Y_t(i) = \max\left\{Z_t^{1-\alpha} K_t(i)^\alpha L_t(i)^{1-\alpha} - Z_t \mathcal{F}, 0\right\}, \tag{4}$$

where the technology shock Z_t (common across all firms) follows a unit root process, and where $\mathcal{F}$ represent fixed costs faced by the firm. We define technology growth $z_t = \log(Z_t/Z_{t-1})$ and assume that z_t follows the autoregressive process:

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

All firms face the same prices for their inputs, labor and capital. Hence profit's maximization implies that the capital/labor ratio is the same for all firms, and equal to:

$$\frac{K_t}{L_t} = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k}., \tag{6}$$

where W_t is the nominal wage and R_t^k is the rental rate of capital. Following Calvo (1983) we assume that in every period a fraction of firms ζ_p is unable to re-optimize their prices $P_t(i)$. These firms adjust their prices mechanically according to

$$P_t(i) = (\pi_{t-1})^{\iota_p} (\pi^*)^{1-\iota_p}, \tag{7}$$

where $\pi_t = P_t/P_{t-1}$ and π^* is the steady state inflation rate of the final good. In our empirical analysis we will restrict ι_p to be either zero or one. Those firms that are able to re-optimize prices choose the price level $\tilde{P}_t(i)$ that solves:

$$\begin{aligned} &\max_{\tilde{P}_t(i)} \mathbb{E}_t \textstyle\sum_{s=0}^{\infty} \zeta_p^s \beta^s \Xi_{t+s}^p \left(\tilde{P}_t(i)\left(\Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}\right) - MC_{t+s}\right) Y_{t+s}(i) \\ &\text{s.t. } Y_{t+s}(i) = \left(\frac{\tilde{P}_t(i)\left(\Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}\right)}{P_{t+s}}\right)^{-\frac{1+\lambda_{f,t}}{\lambda_{f,t}}} Y_{t+s}, \; MC_{t+s} = \frac{\alpha^{-\alpha} W_{t+s}^{1-\alpha} R_{t+s}^{k\ \alpha}}{(1-\alpha)^{(1-\alpha)} Z_{t+s}^{1-\alpha}}. \end{aligned} \tag{8}$$

where $\beta^s \Xi_{t+s}^p$ is today's value of a future dollar for the consumers and MC_t reflects marginal costs. We consider only the symmetric equilibrium where all firms will choose the same $\tilde{P}_t(i)$. Hence from (3) we obtain the following law of motion for the aggregate price level:

$$P_t = [(1-\zeta_p)\tilde{P}_t^{\frac{1}{\lambda_{f,t}}} + \zeta_p (\pi_{t-1}^{\iota_p} \pi_*^{1-\iota_p} P_{t-1})^{\frac{1}{\lambda_{f,t}}}]^{\lambda_{f,t}}. \tag{9}$$

2.3 Labor packers

There is a continuum of households, indexed by $j \in [0,1]$, each supplying a differentiated form of labor, $L(j)$. The "labor packers" are perfectly competitive firms that hire labor from the households and combine it to labor services L_t that are offered to the intermediate goods producers:

$$L_t = \left[\int_0^1 L_t(j)^{\frac{1}{1+\lambda_w}} di\right]^{1+\lambda_w}, \tag{10}$$

where $\lambda_w \in (0,\infty)$. From first-order and zero-profit conditions of the labor packers we obtain the labor demand function and an expression for the price of aggregated labor services L_t:

$$L_t(j) = \left(\frac{W_t(j)}{W_t}\right)^{-\frac{1+\lambda_w}{\lambda_w}} L_t \quad \text{and} \quad W_t = \left[\int_0^1 W_t(j)^{\frac{1}{\lambda_w}} di\right]^{\lambda_w}. \tag{11}$$

2.4 Households

The objective function for household j is given by:

$$I\!E_t \sum_{s=0}^{\infty} \beta^s b_{t+s} \left[\log(C_{t+s}(j) - hC_{t+s-1}(j)) - \frac{\varphi_{t+s}}{1+\nu_l} L_{t+s}(j)^{1+\nu_l} + \frac{\chi}{1-\nu_m}\left(\frac{M_{t+s}(j)}{Z_{t+s}P_{t+s}}\right)^{1-\nu_m}\right] \tag{12}$$

where $C_t(j)$ is consumption, $L_t(j)$ is labor supply, and $M_t(j)$ are money holdings. Household's preferences display habit-presistence. We depart from Smets and Wouters (2003) in assuming separability in the utility function for a reason that will be discussed later. The preference shifters φ_t, which affects the marginal utility of leisure, and b_t, which scales the overall period utility, are exogenous processes common to all households that evolve as:

$$\ln \varphi_t = (1-\rho_\varphi)\ln\varphi + \rho_\varphi \ln\varphi_{t-1} + \sigma_\varphi \epsilon_{\varphi,t}, \tag{13}$$

$$\ln b_t = \rho_b \ln b_{t-1} + \sigma_b \epsilon_{b,t}. \tag{14}$$

Real money balances enter the utility function deflated by the (stochastic) trend growth of the economy, so to make real money demand stationary.

The household's budget constraint written in nominal terms is given by:

$$\begin{aligned} &P_{t+s}C_{t+s}(j) + P_{t+s}I_{t+s}(j) + B_{t+s}(j) + M_{t+s}(j) \le R_{t+s}B_{t+s-1}(j) + M_{t+s-1}(j) \\ &+ \Pi_{t+s} + W_{t+s}(j)L_{t+s}(j) + \left(R^k_{t+s}u_{t+s}(j)\bar{K}_{t+s-1}(j) - P_{t+s}a(u_{t+s}(j))\bar{K}_{t+s-1}(j)\right), \end{aligned} \tag{15}$$

where $I_t(j)$ is investment, $B_t(j)$ is holdings of government bonds, R_t is the gross nominal interest rate paid on government bonds, Π_t is the per-capita profit the household gets from owning firms (households pool their firm shares, and they all receive the same profit), and

$W_t(j)$ is the nominal wage earned by household j. The term within parenthesis represents the return to owning $\bar{K}_t(j)$ units of capital. Households choose the utilization rate of their own capital, $u_t(j)$. Households rent to firms in period t an amount of "effective" capital equal to:

$$K_t(j) = u_t(j)\bar{K}_{t-1}(j), \tag{16}$$

and receive $R_t^k u_t(j)\bar{K}_{t-1}(j)$ in return. They however have to pay a cost of utilization in terms of the consumption good equal to $a(u_t(j))\bar{K}_{t-1}(j)$. Households accumulate capital according to the equation:

$$\bar{K}_t(j) = (1-\delta)\bar{K}_{t-1}(j) + \mu_t \left(1 - S\left(\frac{I_t(j)}{I_{t-1}(j)}\right)\right) I_t(j), \tag{17}$$

where δ is the rate of depreciation, and $S(\cdot)$ is the cost of adjusting investment, with $S'(\cdot) > 0, S''(\cdot) > 0$. The term μ_t is a stochastic disturbance to the price of investment relative to consumption, which follows the exogenous process:

$$\ln \mu_t = (1-\rho_\mu)\ln \mu + \rho_\mu \ln \mu_{t-1} + \sigma_\mu \epsilon_{\mu,t}. \tag{18}$$

The households' wage setting is subject to nominal rigidities á la Calvo (1983). In each period a fraction ζ_w of households is unable to re-adjust wages. For these households, the wage $W_t(j)$ will increase at a geometrically weighted average of the steady state rate increase in wages (equal to steady state inflation π_* times the growth rate of the economy e^γ) and of last period's inflation times last period's productivity $(\pi_{t-1}e^{z_{t-1}})$. The weights are $1-\iota_w$ and ι_w, respectively. Those households that are able to re-optimize their wage solve the problem:

$$\begin{array}{lll} \max_{\tilde{W}_t(j)} & I\!\!E_t \sum_{s=0}^{\infty} (\zeta_w \beta)^s b_{t+s} \left[-\frac{\varphi_{t+s}}{\nu_l + 1} L_{t+s}(j)^{\nu_l + 1} \right] & \\ \text{s.t.} & (15) \text{ for } s = 0, \ldots, \infty, \text{ (11a), and} & (19) \\ & W_{t+s}(j) = \left(\Pi_{l=1}^s (\pi_* e^\gamma)^{1-\iota_w} (\pi_{t+l-1} e^{z^*_{t+l-1}})^{\iota_w} \right) \tilde{W}_t(j). & \end{array}$$

We again consider only the symmetric equilibrium in which all agents solving (19) will choose the same $\tilde{W}_t(j)$. From (11b) it follows that:

$$W_t = [(1-\zeta_w)\tilde{W}_t^{\frac{1}{\lambda_w}} + \zeta_w((\pi_* e^\gamma)^{1-\iota_w}(\pi_{t-1}e^{z^*_{t-1}})^{\iota_w} W_{t-1})^{\frac{1}{\lambda_w}}]^{\lambda_w}. \tag{20}$$

Finally, we assume there is a complete set of state contingent securities in nominal terms, which implies that the Lagrange multiplier $\Xi_t^p(j)$ associated with (15) must be the same for all households in all periods and across all states of nature. This in turn implies that in equilibrium households will make the same choice of consumption, money demand, investment and capital utilization. Since the amount of leisure will differ across households due to the wage rigidity, separability between labor and consumption in the utility function is key for this result.

2.5 Government policies

The central bank follows a nominal interest rate rule by adjusting its instrument in response to deviations of inflation and output from their respective target levels:

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_*}\right)^{\psi_1} \left(\frac{Y_t}{e^{\gamma} Y_{t-1}}\right)^{\psi_2}\right]^{1-\rho_R} \sigma_R e^{\epsilon_{R,t}}, \tag{21}$$

where R^* is the steady state nominal rate, Y_t^s is the target level of output, and the parameter ρ_R determines the degree of interest rate smoothing. In our formulation of the policy rule, the central bank responds to output growth rather than some measure of the output gap.

The government budget constraint is of the form

$$P_t G_t + R_{t-1} B_{t-1} + M_{t-1} = T_t + M_t + B_t, \tag{22}$$

where T_t are nominal lump-sum taxes (or subsidies) that also appear in household's budget constraint. Government spending is given by:

$$G_t = (1 - 1/g_t) Y_t, \tag{23}$$

where g_t follows the process:

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \sigma_g \epsilon_{g,t} \tag{24}$$

2.6 Resource constraint

The aggregate resource constraint:

$$C_t + I_t + a(u_t) \bar{K}_{t-1} = \frac{1}{g_t} Y_t. \tag{25}$$

can be derived by integrating the budget constraint (15) across households, and combining it with the government budget constraint (22) and the zero profit conditions of both labor packers and final good producers.

2.7 Model solution and State-Space Representation

As in Altig, Christiano, Eichenbaum, and Linde (2002) our model economy evolves along stochastic growth path. Output Y_t, consumption C_t, investment I_t, the real wage W_t/P_t, physical capital K_t and effective capital $\bar{K}_t$ all grow at the rate Z_t. Nominal interest rates, inflation, and hours worked are stationary. The model can be rewritten in terms of detrended variables. 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.

Our empirical analysis is based on data on nominal interest rates (annualized), inflation rates (annualized), and quarterly output growth rates. Hence, let $y_t = [R_t^a, \pi_t^a, \Delta \ln Y_t]'$. The relationships between the steady-state deviations $\widetilde{R}_t$, $\widetilde{\pi}_t$, $\widetilde{Y}_t$ and the observables are given by the following measurement equations:

$$\begin{aligned} y_{1,t} &= \ln r_a^* + \ln \pi_a^* + 4\widetilde{R}_t, \\ y_{2,t} &= \begin{bmatrix} \ln \pi_a^* + 4\widetilde{\pi}_t \\ \ln \gamma + \Delta\widetilde{x}_t + \widetilde{z}_t \end{bmatrix}. \end{aligned} \tag{26}$$

Here, $y_{1,t}$ denotes the policymaker's instrument (the interest rate), and $y_{2,t}$ is a vector that includes the remaining two endogenous variables. We collect all the DSGE model parameters in the vector θ and stack the structural shocks in the vector ϵ_t.

3 Setup and Inference

In the subsequent analysis it is assumed that the DSGE model generates a covariance-stationary distribution of the sequence $\{y_t\}$ for all $\theta \in \Theta$. Expectations under this distribution are denoted by $I\!E_\theta^D[\cdot]$. We will derive an (approximate) vector autoregressive representation for the DSGE model and introduce model misspecifications as deviations from this representation. Finally, a prior distribution for these model misspecifications is specified and posterior inference and policy analysis are discussed.

3.1 A VAR Representation of the DSGE Model

Let us rewrite Eq. (21), which describes the policymaker's behavior, in more general form as:

$$y_{1,t} = x_t' M_1 \beta_1(\theta) + y_{2,t}' M_2 \beta_2(\theta) + \epsilon_{1,t}, \tag{27}$$

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 $\sigma_R \epsilon_{R,t}$ in the DSGE model. The matrices M_1 and M_2 select the appropriate elements of x_t and $y_{2,t}$ to generate the policy rule. In our application the vector M_1 selects the intercept and the lagged nominal interest rate and the matrix M_2 extracts inflation, and output growth. The functions $\beta_1(\theta)$ and $\beta_2(\theta)$ can be easily derived from (21) and the measurement equation (26) for R_t.

The remainder of the system for y_t is given by the following reduced form equations:

$$y'_{2,t} = x'_t \Psi^*(\theta) + u'_{2,t}. \tag{28}$$

In general, the VAR representation (28) is not exact if the number of lags p is finite. We define $\Gamma_{XX}(\theta) = I\!E^D_\theta[x_t x'_t]$ and $\Gamma_{XY_2}(\theta) = I\!E^D_\theta[x_t y'_{2,t}]$ and let

$$\Psi^*(\theta) = \Gamma^{-1}_{XX}(\theta)\Gamma_{XY_2}(\theta). \tag{29}$$

Since the system is covariance stationary, the VAR approximation of the autocovariance sequence of $y_{2,t}$ can be made arbitrarily precise by increasing the number of lags p. If in addition, the moving-average (MA) representation of the DSGE model in terms of the structural shocks ϵ_t is invertible, then $u_{2,t}$ can also be expressed as a function of ϵ_t for large p.

The equation for the policy instrument (27) can be rewritten by replacing $y_{2,t}$ with expression (28):

$$y_{1,t} = x'_t M_1 \beta_1(\theta) + x'_t \Psi^*(\theta) M_2 \beta_2(\theta) + u_{1,t}, \tag{30}$$

where $u_{1,t} = u'_{2,t} M_2 \beta_2(\theta) + \epsilon_{1,t}$. Define $u'_t = [u_{1,t}, u'_{2,t}]$, $B_1(\theta) = [M_1 \beta_1(\theta), 0_{k \times (n-1)}]$, $B_2(\theta) = [M_2 \beta_2(\theta), I_{(n-1)\times(n-1)}]$, and let

$$\Phi^*(\theta) = B_1(\theta) + \Psi^*(\theta) B_2(\theta). \tag{31}$$

Hence, we obtain a restricted VAR for y_t

$$y'_t = x'_t \Phi + u'_t, \quad I\!E[u_t u'_t] = \Sigma^*(\theta) \tag{32}$$

with

$$\Phi = \Phi^*(\theta), \quad \Sigma = \Sigma^*(\theta) = \Gamma_{YY}(\theta) - \Gamma_{YX}(\theta)\Gamma^{-1}_{XX}(\theta)\Gamma_{XY}(\theta).$$

Here the population covariance matrices are $\Gamma_{YY}(\theta) = I\!E^D_\theta[y_t y'_t]$ and $\Gamma_{XY}(\theta) = \Gamma'_{YX}(\theta) = I\!E^D_\theta[x_t y'_t]$. The following Lemma will be useful for the subsequent analysis and can be verified by straightforward matrix manipulations. Let $I\!E^{VAR}_{\Psi,\Sigma}[\cdot]$ denote expectations under the probability distribution generated by (32).

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

Since the monetary policy rule (21) in the DSGE model is specified so that it can be exactly reproduced by the VAR, see Eq. (27), $\Phi^*(\theta)$ equals the population least squares coefficients associated with (32), and the covariance matrix of x_t under the DSGE model and its VAR approximation are identical. For the ease of exposition we will subsequently ignore the error made by approximating the state space representation of the DSGE model with the finite-order VAR or, in other words, treat (32) as the structural model that imposes potentially misspecified restrictions on the matrices Φ and Σ.

3.2 Misspecification and Bayesian Inference

We make the following assumptions about misspecification of the DSGE model. There is a vector θ and matrices Ψ^Δ and Σ^Δ such that the data are generated from the VAR in Eq. (32)

$$\Phi = B_1(\theta) + (\Psi^*(\theta) + \Psi^\Delta)B_2(\theta), \quad \Sigma = \Sigma^*(\theta) + \Sigma^\Delta. \tag{33}$$

and there does not exist a $\tilde{\theta} \in \Theta$ such that

$$\Phi = B_1(\tilde{\theta}) + \Psi^*(\tilde{\theta})B_2(\tilde{\theta}), \quad \Sigma = \Sigma^*(\tilde{\theta}).$$

We refer to the resulting specification as DSGE-VAR. Our econometric analysis is casted in a Bayesian framework in which initial beliefs about the DSGE model parameter θ and the model misspecification matrices Ψ^Δ and Σ^Δ are summarized in a prior distribution. We will subsequently motivate this prior distribution with a thought experiment.

For now, we assume that $\Sigma^\Delta = 0$ and condition on the DSGE model parameter vector θ. The goal is to assign low prior density to large values of Ψ^Δ, reflecting the belief that the DSGE model provides a good albeit not perfect approximation of reality. By large, we mean discrepancies that are easily detectable with likelihood ratios. Suppose that a sample of λT observations is generated from (32), where Φ is given by (33). We will construct a prior that has the property that its density is proportional to the expected likelihood ratio of Ψ evaluated at its (misspecified) restricted value $\Psi^*(\theta)$ versus the true value $\Psi = \Psi^*(\theta) + \Psi^\Delta$. Since the likelihood ratio is decreasing in the number of observations λT for fixed Ψ^Δ, the misspecification is re-scaled as follows. Let

$$\Psi^\Delta = \frac{1}{\sqrt{\lambda T}}\widetilde{\Psi}^\Delta.$$

The log-likelihood ratio is

$$\ln\left[\frac{\mathcal{L}(\Psi^*,\Sigma^*,\theta|Y,X)}{\mathcal{L}(\Psi^*+\Psi^\Delta,\Sigma^*,\theta|Y,X)}\right] = -\frac{1}{2}tr\Bigg[\Sigma^{*-1}\Bigg(B_2'\Psi^{*'}X'X\Psi^*B_2 - 2B_2'\Psi^{*'}X'(Y-XB_1)$$
$$-B_2'(\Psi^*+(\lambda T)^{-1/2}\widetilde{\Psi}^\Delta)'X'X(\Psi^*+(\lambda T)^{-1/2}\widetilde{\Psi}^\Delta)B_2$$
$$+2B_2'(\Psi^*+(\lambda T)^{-1/2}\widetilde{\Psi}^\Delta)'X'(Y-XB_1)\Bigg)\Bigg].$$

Here Y denotes the $\lambda T \times n$ matrix with rows y_t' and X_t is the $\lambda T \times k$ matrix with rows x_t'. After replacing Y by $X(B_1+(\Psi^*+\Psi^\Delta)B_2)+U$ the log likelihood ratio simplifies to

$$\ln\left[\frac{\mathcal{L}(\Psi^*,\Sigma^*,\theta|Y,X)}{\mathcal{L}(\Psi^*+\Psi^\Delta,\Sigma^*,\theta|Y,X)}\right] \tag{34}$$
$$= -\frac{1}{2}tr\Bigg[\Sigma^{*-1}\Bigg(B_2'\widetilde{\Psi}^{\Delta'}(\lambda T)^{-1}X'X\widetilde{\Psi}^\Delta B_2 - 2B_2'\widetilde{\Psi}^{\Delta'}(\lambda T)^{-1/2}X'U\Bigg)\Bigg]$$

Taking expectations over X and U using the distribution induced by the data generating process yields (minus) the Kullback-Leibler distance between the data generating process and the DSGE model:

$$I\!E^{VAR}_{\Psi^*+\Psi^\Delta,\Sigma^*}\left[\ln\frac{\mathcal{L}(\Psi^*,\Sigma^*,\theta|Y,X)}{\mathcal{L}(\Psi^*+\Psi^\Delta,\Sigma^*,\theta|Y,X)}\right] = -\frac{1}{2}tr\Bigg[\Sigma^{*-1}\Bigg(B_2'\widetilde{\Psi}^{\Delta'}\Gamma_{XX}\widetilde{\Psi}^\Delta B_2\Bigg)\Bigg]+O((\lambda T)^{-1/2}). \tag{35}$$

Here we have used Lemma 1(ii). The $O((\lambda T)^{-1/2})$ arises because

$$\Phi = B_1(\theta)+\left(\Psi^*(\theta)+\frac{1}{\sqrt{\lambda T}}\widetilde{\Psi}^\Delta\right)B_2(\theta) = \Phi^*(\theta)+\frac{1}{\sqrt{\lambda T}}\widetilde{\Psi}^\Delta B_2(\theta).$$

We now choose a prior density that is proportional ($\propto$) to the Kullback-Leibler discrepancy:

$$p(\widetilde{\Psi}^\Delta|\Sigma^*,\theta) \propto \exp\left\{-\frac{1}{2}tr\Bigg[\Sigma^{*-1}\Bigg(B_2'\widetilde{\Psi}^{\Delta'}\Gamma_{XX}\widetilde{\Psi}^\Delta B_2\Bigg)\Bigg]\right\} \tag{36}$$

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

$$\Psi|\Sigma^*,\theta \sim \mathcal{N}\left(\Psi^*(\theta), \frac{1}{\lambda T}\Big[(B_2(\theta)\Sigma^{*-1}B_2(\theta)')\otimes\Gamma_{XX}(\theta)\Big]^{-1}\right). \tag{37}$$

The hyperparameter λ, which determines the length of the hypothetical sample as a multiple of the actual sample size T, scales the variance of the distribution that generates $\widetilde{\Psi}^\Delta$ and Ψ. If λ is close to zero, the prior variance of the discrepancy $\widetilde{\Psi}^\Delta$ is large. Large values of λ, on the other hand, correspond to small model misspecification and for $\lambda=\infty$ the misspecification disappears.

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:

$$\Psi|\Sigma,\theta \sim \mathcal{N}\left(\Psi^*(\theta), \frac{1}{\lambda T}\left[(B_2(\theta)\Sigma^{-1}B_2(\theta)') \otimes \Gamma_{XX}(\theta)\right]^{-1}\right) \tag{38}$$

$$\Sigma|\theta \sim \mathcal{IW}\left(\lambda T\Sigma^*(\theta), \lambda T - k, n\right), \tag{39}$$

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

The Appendix provides a characterization of the following conditional posterior densities:

$$p(\Psi|\Sigma,\theta,Y), \quad p(\Sigma|\Psi,\theta,Y), \quad \text{and} \quad p(\theta|\Psi,\Sigma,Y).$$

Unfortunately, it is not possible to give a characterization of all conditional distributions in terms of well-known probability distributions. To implement the Gibbs sampler we have to introduce two Metropolis steps that generate draws from the conditional distributions $p(\Sigma|\Psi,\theta,Y)$ and $p(\theta|\Psi,\Sigma,Y)$. The resulting Markov-Chain-Monte-Carlo (MCMC) algorithm is known as Metropolis-within-Gibbs sampler and allows us to generate draws from the joint posterior distribution of θ, Ψ, and Σ. In addition to the posterior distribution of the parameters we are also interested in evaluating marginal data densities of the form

$$p(Y) = \int p(Y|\theta,\Sigma,\Phi)p_{\lambda}(\theta,\Sigma,\Phi)d(\theta,\Sigma,\Phi) \tag{40}$$

for various choices of the hyperparameter λ and restrictions on the parameter space of the DSGE model. Based on the marginal data densities we can compute Bayes factors and posterior probabilities for the various specifications of our model. Under the assumption of equal prior probabilities, ratios of marginal likelihoods can be interpreted as model odds.

3.3 Policy Analysis

At time $t = T$ the policymaker seeks to replace the existing policy rule with one that minimizes the following loss function[1]

$$L(\theta_p,\theta_s,\Psi^{\Delta},\Sigma^{\Delta}) = \min\left\{\mathcal{B}, I\!\!E\left[(1-\delta)\sum_{t=T}^{\infty}\delta^{t-T}(y_t-\bar{y})'\mathcal{W}(y_t-\bar{y})\right]\right\}, \tag{41}$$

[1] We make the simplifying assumption that the public believes the new policy to be in place indefinitely after being announced credibly. This assumption is a short-cut to a more realistic scenario in which there are two types of policy shifts - normal policy making and rare regime shifts (using the terminology of Sims, 1982). In addition we assume that the expectation in (41) is unconditional. The policymaker does not exploit the fact that the public has formed its time $T-1$ expectations based on the $T-1$ policy rule.

where δ is a discount factor, θ is partitioned into policy rule parameters θ_p and taste-and-technology parameters θ_s, and $\mathcal{B}$ is a positive constant that ensures that the loss is bounded even if the VAR coefficients imply non-stationary or explosive behavior of the endogenous variables. The loss function encompasses popular ad-hoc loss functions that penalize inflation, output, and interest rate variability.

The policymaker minimizes the loss $L(\theta_p, \theta_s, \Psi^{\Delta}, \Sigma^{\Delta})$ as a function of the policy parameter θ_p. She has imperfect knowledge about: (i) the policy invariant private sectors' taste and technology parameters θ_s; and (ii) the degree of model misspecification captured by λ, Ψ^{Δ} and Σ^{Δ}. The uncertainty is summarized in the posterior distribution.

We consider four different scenarios for the policy invariance of the misspecification matrices Φ^{Δ} and Σ^{Δ}. Then we calculate the posterior expected loss associated with different policies according to each scenario. If the DSGE model does not suffer from serious misspecification all scenarios collapse to Scenario 1. At this point we have no theory that lets us determine which of the scenarios will provide the most accurate prediction of the policy effects. The goal of the subsequent empirical analysis is to illustrate the sensitivity of policy predictions to assumptions on model misspecification.

Scenario 1: The DSGE model is estimated directly and its potential misspecification is ignored. The policymaker does, however, take the uncertainty with respect to the non-policy parameters into account when calculating the expected loss. This scenario dates back at least to Brainard (1967) and serves as a benchmark. More recent examples in the context of DSGE models include Laforte (2003) and Onatski and Williams (2004).

Scenario 2: The policymaker believes that the sample (hence the posterior) provides no information about potential misspecification after a regime shift has been implemented. This scepticism about the relevance of sample information is shared by the robust control approaches of Hansen and Sargent (2000) and Onatski and Stock (2002). Here, instead of using a minimax argument, our Bayesian policymaker relies on her prior distribution $p(\Psi^{\Delta}, \Sigma^{\Delta}|\theta, \lambda)$ to cope with uncertainty about model misspecification. She still uses the sample to learn about θ_s and λ, however.

Scenario 3: Ψ^{Δ} and Σ^{Δ} are invariant to changes in policy. The sample information is used to learn about the model misspecification via the posterior distribution. Looking forward, the information is used to adjust the policy predictions derived from the DSGE model, $\Psi^*(\tilde{\theta})$ and $\Sigma^*(\tilde{\theta})$. Here $\tilde{\theta}$ denotes the vector of structural parameters that is obtained when θ_p is replaced by a new set of policy parameters $\tilde{\theta}_p$.

Scenario 4: Nature generates a new set of draws from the posterior distribution of Ψ^Δ and Σ^Δ conditional on the post-intervention DSGE model parameters $\tilde{\theta}$ instead of the pre-intervention parameters as in Scenario 3. For small values of λ the conditional posterior distribution of $\Psi^*(\theta) + \Psi^\Delta$ and $\Sigma^*(\theta) + \Sigma^\Delta$ given θ is effectively insensitive to θ. In this case Scenario 4 corresponds to analyzing policy effects with a VAR by simply changing the coefficients in the policy rule.

4 The Data

In our empirical analysis we use observations on interest rates, inflation, and output growth. All data are obtained from Haver analytics (Haver mnemonics are in italics). Real output is obtained by dividing the nominal series (*GDP*) by population 16 years and older (*LF+LH*), and deflating using the chained-price GDP deflator (*JGDP*). Growth rates are computed using log-differences from quarter to quarter, and are in percent. Inflation is computed using log-differences of the GDP delflator, in percent. The nominal rate corresponds to the effective Federal Funds rate (*FFED*), also in percent. The results reported subsequently are based on a sample from 1983:Q3 to 2004:Q1.

5 Empirical Application

Since we estimate the DSGE-VARs based on only three variables we set most shocks equal to zero, except for the technology growth shock $\epsilon_{z,t}$, the monetary policy shock $\epsilon_{R,t}$, and the government spending shock $\epsilon_{g,t}$. Since the model is to some extent able to endogenously generate persistence in real variables, we impose that technology growth shocks are serially uncorrelated, that is, $\rho_z = 0$. In Del Negro, Schorfheide, Smets and Wouters (2004), henceforth DSSW, we did not find evidence in favor of price indexation. Therefore, we let $\iota_p = \iota_w = 0$. Moreover, we set the fixed costs $\mathcal{F} = 0$. Unlike in DSSW, we do not use observations on consumption and investment, which makes it difficult to identify the capital share and the depreciation rate. Therefore, we let $\alpha = 0.25$ and $\delta = 0.025$. Since we are not extracting information from wage and money data we fix the wage-markup parameter $\lambda_w = 0.3$, and the money demand elasticity $\nu_m = 2$. In a log-linear approximation the Calvo parameter is typically not separately identifiable from the price markup parameter λ_f, which we fix at 0.3.

We begin with a direct estimation of the DSGE model using Bayesian techniques described in Schorfheide (2000). Table 1 reports prior mean and standard deviations, as well as posterior means and 90% probability intervals for the structural parameters. The estimates for the inflation and output growth coefficients in the monetary policy rule are 1.43 and 0.36, respectively. Our estimate of the smoothing coefficient is fairly high compared to estimates reported elsewhere in the literature: $\hat{\rho}_r = 0.83$. The Calvo parameters for wages and prices are 0.72, and 0.79, respectively. Thus, agents change their prices on average every 4 quarters. We estimate a large degree of habit persistence, whereas the data seem to be fairly uninformative with respect to the labor supply elasticity μ_l and the cost of capital utilization a".

We proceed by estimating DSGE-VARs for values of λ between 0.25, i.e., large prior variance of the misspecification matrices Ψ^{Δ} and Σ^{Δ}, and 5, i.e., small potential misspecification. The subsequent results are based on $p = 4$ lags. Table 2 describes the posterior of the misspecification parameter λ. The table reports log marginal data densities for the directly estimated DSGE model and DSGE-VARs based on different values of λ. Differences of log marginal densities across model specifications can be interpreted as log posterior odds, under the assumption that the prior odds are equal to one. The odds reported in the last column of Table 2 are relative to $\lambda = 0.75$, which is the specification with the largest marginal data density and, according to this likelihood-based criterion, the best fit. The posterior of λ has an inverted U-shape. There is little variation in the marginal data densities for λ values between 0.5 and 2, whereas values outside of this interval lead to a substantial deterioration in fit. We conclude that over the range of the historical sample the DSGE model is strongly dominated by DSGE-VARs with intermediate values of λ indicating that the structural model is to some extent misspecified and that its policy predictions should be interpreted with care.

Table 3 compares the posterior means of the structural parameters obtained from the estimation of DSGE-VAR specifications for various values of λ. The DSGE-VAR generates Bayesian instrumental variable estimates of the policy rule parameters ψ_1, ψ_2, and ρ_R. For large values of λ the instruments are very similar to the scores of the DSGE models' likelihood function. For values of λ near zero the instruments essentially correspond to lagged values of interest rates, inflation, and output growth. While the estimates of ψ_2 and ρ_R are fairly insensitive to the choice of λ, the estimate of the inflation coefficient rises from 1.43 to 1.99 as the prior variance of the discrepancies Ψ^{Δ} and Σ^{Δ} increases. As shown in Del Negro and Schorfheide (2004a) the estimates of the remaining DSGE model coefficients can be interpreted as minimum distance estimates, in which the estimator of Ψ is projected onto the restricted subspace generated by $\Psi^*(\theta)$.

Based on the parameter estimates we calculate expected policy losses. The loss is based on Eq. (41) where the weighting matrix $\mathcal{W}$ is diagonal with elements $\frac{1}{4}$ (interest rates, annualized), 1 (inflation, annualized), and $\frac{1}{4}$ (output growth, quarter-to-quarter). Our weight on output growth is somewhat larger than in Woodford (2003, Table 6.1) reflecting a larger estimate of κ. Moreover, we place considerable weight on the nominal interest rate, which could be justified by a large interest elasticity of money demand and an important role of real money balances for transactions. The upper bound $\mathcal{B}$ of the loss is set to 20, which is about 5 times larger than the weighted sample variance of the three series. We evaluate the expected loss as a function of ψ_1, the central bank's response to inflation. The other two parameters of the policy rule are set approximately to their respective posterior mean estimates for $\lambda = 0.75$: $\psi_2 = 0.25$, $\rho_R = 0.85$. The results are summarized in Figure 1, which depicts expected loss differentials relative to the benchmark $\psi_1 = 1.8$. Negative differentials indicate an improvement relative to the benchmark.

In Scenario 1 the policymaker calculates the policy loss with the DSGE model, ignoring misspecification. It is well known that as the response to inflation increases, inflation variability drops and the loss decreases. The inference about the misspecification parameter λ in Table 2 casts some doubts on the reliability of DSGE model predictions, however.

In Scenario 2 the policymaker still uses the DSGE model to compute the mean response of the endogenous variables to the change in ψ_1, but recognizes that nature may be injecting noise around these mean responses using the prior distribution. Dark shades of grey in Figure 1 correspond to large values of λ, whereas light shades of grey are associated with small values of the hyperparameter. Not surprisingly for larger values of λ (low misspecification) the shape of the loss does not change relative to Scenario 1. For smaller values of λ (high misspecification) the loss profile becomes flatter. A decomposition of the loss into its three components indicates that for small values of λ the interest rate variability actually rise as the central bank responds more strongly to inflation. However, this rise is off-set by the drop in inflation variability.

In Scenario 3 the policymaker uses sample information to learn about the size of the discrepancies, unlike in the previous scenarios. More specifically, she believes that the historically observed discrepancies Ψ^Δ and Σ^Δ are policy invariant. For $\lambda = 1$ the loss is still a decreasing function of ψ_1, as in the Scenario 1, but for values of λ less than 1 the slope switches sign around $\psi_1 = 2$.

Finally, under Scenario 4 the policymaker again uses sample information to learn about potential model misspecification. Unlike in Scenario 3, the policymaker now asks the question: what is the estimate of the discrepancy if the new policy had been in place during the sample period. For small values of λ inflation and output growth are essentially being forecasted using an unrestricted VAR as changes of agents' decision rules derived from the DSGE models are deemed unreliable. Only the policy equation reflects the change in ψ_1, thereby generating a higher interest rate volatility which leads to the slight positive slope of the loss functions depicted in the fourth panel of Figure 1. However, as λ is increased the loss differentials become more and more similar to the profile calculated under Scenario 1.

At this point we have no theory that lets us determine which of the scenarios will provide the most accurate prediction of the policy effects. We show that the results of the policy analysis to some extent depend on: (i) whether the policymaker relies on the data to assess the degree of misspecifications, i.e., learns about λ; and (ii) the assumption she makes on the process driving the discrepancies between the DSGE model and the data in the aftermath of the policy intervention. According to our analysis, values of ψ_1 less than 1.5 are undesirable regardless of the assumptions about model misspecification and its invariance to policy interventions. Whether ψ_1 should be raised, say to 2.5, is questionable. Under the DSGE model a stronger response to inflation movements leads to a reduction of the expected loss. On the other hand, the DSGE-VAR analysis with the optimal value of λ suggests under Scenarios 3 and 4 that the loss will increase.

6 Conclusion

Current DSGE models are to some extent misspecified, even large-scale models such as the one in Smets and Wouters (2003). While they allow policymakers to assess the effects of rare policy changes on the expectation formation and decision rules of private agents, their fit is typically worse than the fit of alternative econometric models, such as VARs estimated with well-calibrated shrinkage methods. The DSGE-VARs studied in Del Negro and Schorfheide (2004a) and Del Negro, Schorfheide, Smets, and Wouters (2004) provide a framework that allows researchers to account for model misspecification. In this paper we developed techniques to conduct policy analysis with potentially misspecified DSGE models and applied them to a New Keynesian DSGE model with capital accumulation and several real and nominal frictions. We studied the effect of changing the response to inflation under an ad-hoc loss function that penalizes inflation, output growth, and interest rate variability.

We view our framework as an attractive alternative to robust control approaches to model misspecification that deserves to be explored in future research.

References

Altig, David, Lawrence J. Christiano, Martin Eichenbaum, and Jesper Linde (2002): "Technology Shocks and Aggregate Fluctuations," *Manuscript*, Northwestern University.

Brainard, William (1967): "Uncertainty and the Effectiveness of Policy," *American Economic Review*, **57**, 411-25.

Calvo, Guillermo (1983): " Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, **12**, 383-398.

Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans (2004): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, forthcoming.

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

Del Negro, Marco and Frank Schorfheide (2004b): "Policy Predictions if the Model Doesn't Fit," *Manuscript*, University of Pennsylvania.

Del Negro, Marco, Frank Schorfheide, Frank Smets, and Raf Wouters (2004): "On the Fit and Forecasting Performance of New Keynesian Models," *Manuscript*, University of Pennsylvania.

Hansen, Lars P and Thomas J. Sargent (2000): "Wanting Robustness in Macroeconomics," *Manuscript*, New York University, Department of Economics.

Ingram, Beth F. and Charles H. Whiteman (1994): "Supplanting the Minnesota prior – Forecasting macroeconomic time series using real business cycle model priors." *Journal of Monetary Economics*, **34**, 497-510.

Laforte, Jean-Phillipe (2003): "Comparing Monetary Policy Rules in an Estimated General Equilibrium Model of the U.S. Economy," *Manuscript*, Federal Reserve Board of Governors, Washington, DC.

Onatski, Alexei and James H. Stock (2002): "Robust Monetary Policy under Model Uncertainty in a Small Model of the US Economy," *Macroeconomic Dynamics*, **6**, 85-110.

Onatski, Alexei and Noah Williams (2004): "Empirical and Policy Performance of a Forward-Looking Monetary Model," *Manuscript*, Princeton University, Department of Economics.

Schorfheide, Frank (2000): "Loss-Function Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, **15**, 645-70.

Sims, Christopher A. (1982): "Policy Analysis with Econometric Econometric Models," *Brookings Papers on Economic Activity*, **1**, 107-64.

Sims, Christopher A. (2002): "Solving Rational Expectations Models." *Computational Economics*, **20**(1-2), 1-20.

Smets, Frank and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, **1**, 1123-75.

Taylor, John B. (1999): *"Monetary Policy Rules,"* NBER Business Cycles Series, **31**, Chicago.

Woodford, Michael (2003): *"Interest and Prices: Foundations of a Theory of Monetary Policy,"* Princeton University Press, Princeton.

A Implementation of the Posterior Simulation

A.1 Draws from the Posterior

We adopt the notation that $\tilde{Y}(\theta) = Y - XB_1(\theta)$ which leads to the definitions

$$\Gamma_{\tilde{Y}\tilde{Y}} = \Gamma_{YY} - \Gamma_{YX}B_1(\theta) - B_1(\theta)'\Gamma_{XY} - B_1(\theta)'\Gamma_{XX}B_1(\theta), \quad \Gamma_{X\tilde{Y}} = \Gamma_{XY} - \Gamma_{XX}B_1(\theta).$$

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

$$\begin{aligned} & p(Y|\Psi,\Sigma,\theta) \qquad\qquad (A.1) \\ & \propto \quad |\Sigma|^{-T/2}\text{etr}\left[\Sigma^{-1}\Big(Y - X(B_1(\theta)+\Psi B_2(\theta))\Big)'\Big(Y - X(B_1(\theta)+\Psi B_2(\theta))\Big)\right]. \end{aligned}$$

Using Lemma 1(i) we can rewrite the prior mean of Ψ as

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

The prior density for Ψ conditional on Σ is of the form

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

The prior density for Σ is given by

$$p(\Sigma|\theta) \propto |\Sigma|^{-\frac{1}{2}(\lambda T-k+n+1)}\text{etr}\left[\Sigma^{-1}\lambda T\Sigma^*(\theta))\right] \qquad (A.3)$$

To simplify notation the (θ)-argument of the functions B_1, B_2, $\tilde{Y}$, Γ_{XY}, Γ_{XX}, and Γ_{YY} is omitted.

Conditional Posterior of Ψ: Combining the the prior density (A.2) with the likelihood function (A.1) yields

$$\begin{aligned} & p(\Psi|\Sigma,\theta,Y) \\ & \propto \quad p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta) \qquad\qquad (A.4) \\ & \propto \quad \text{etr}\left[\Sigma^{-1}\lambda T\Big(\Gamma_{YY} - 2B_2'\Psi'\Gamma_{X\tilde{Y}} + B_2'\Psi'\Gamma_{XX}(\theta)\Psi B_2\Big) + (\tilde{Y} - X\Psi B_2)'(\tilde{Y} - X\Psi B_2)\right] \\ & \propto \quad \text{etr}\left[\Sigma^{-1}\Big(-2B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}} + X'\tilde{Y}) + B_2'\Psi'(\lambda T\Gamma_{XX} + X'X)\Psi B_2\Big)\right] \end{aligned}$$

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

$$\Psi|\Sigma,\theta,Y \sim \mathcal{N}\left(\tilde{\Psi}(\Sigma,\theta), \left[(B_2\Sigma^{-1}B_2') \otimes (\lambda T\Gamma_{XX} + X'X)\right]^{-1}\right). \qquad (A.5)$$

Conditional Posterior of Σ: Combining the the prior densities (A.2) and (A.3) with the likelihood function (A.1) yields

$$\begin{aligned} p(\Sigma|\Psi,\theta,Y) &\propto p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta)p(\Sigma|\theta) \qquad \text{(A.6)} \\ &\propto |\Sigma|^{-\frac{1}{2}((\lambda+1)T-k+n+1)}|(B_2\Sigma^{-1}B_2')^{-1}|^{-\frac{k}{2}} \\ &\quad \text{etr}\Big[\Sigma^{-1}\Big(\lambda T(\Gamma_{\tilde{Y}\tilde{Y}}-\Gamma_{\tilde{Y}X}\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}})+(\tilde{Y}-X\Psi B_2)'(\tilde{Y}-X\Psi B_2)\Big) \\ &\quad +\lambda T(B_2\Sigma^{-1}B_2')(\Psi-\bar{\Psi})'\Gamma_{XX}(\Psi-\bar{\Psi})\Big]. \end{aligned}$$

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,

$$\begin{aligned} p(\Sigma|\Psi,\theta,Y) &\propto |\Sigma|^{-\frac{1}{2}((\lambda+1)T-k+n+1)}|(B_2\Sigma^{-1}B_2')^{-1}|^{-\frac{k}{2}} \qquad \text{(A.7)} \\ &\quad \times \text{etr}\Big[\Sigma^{-1}\Big(\lambda T\Gamma_{\tilde{Y}\tilde{Y}}+\tilde{Y}'\tilde{Y}-2B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y}) \\ &\quad +B_2'\Psi'(\lambda T\Gamma_{XX}+X'X)\Psi B_2\Big)\Big] \\ &\quad \times \text{etr}\Big[\lambda T(\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}B_2\Sigma^{-1}-\Sigma^{-1})\Gamma_{X\tilde{Y}}'\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}}\Big]. \end{aligned}$$

If the DSGE model satisfies Eq. (27) and the error $u_{1,t}$ is orthogonal to x_t then

$$\Gamma_{X\tilde{Y}} = \Gamma_{XX}\Psi_0(\theta)B_2$$

and

$$(\Sigma^{-1}B_2'(B_2\Sigma^{-1}B_2')^{-1}B_2\Sigma^{-1}-\Sigma^{-1})\Gamma_{X\tilde{Y}}'\Gamma_{XX}^{-1}\Gamma_{X\tilde{Y}} = 0. \qquad \text{(A.8)}$$

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

$$\begin{aligned} \tilde{S}(\Psi,\theta) &= \lambda T\Gamma_{\tilde{Y}\tilde{Y}}+\tilde{Y}'\tilde{Y}-(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y})'\Psi B_2 - B_2'\Psi'(\lambda T\Gamma_{X\tilde{Y}}+X'\tilde{Y}) \qquad \text{(A.9)} \\ &\quad +B_2'\Psi'(\lambda T\Gamma_{XX}+X'X)\Psi B_2 \end{aligned}$$

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.

$$p(\theta|\Psi,\Sigma,Y) \propto p(Y,\Psi,\Sigma,\theta) = p(Y|\Psi,\Sigma,\theta)p(\Psi|\Sigma,\theta)p(\Sigma|\theta)p(\theta). \tag{A.10}$$

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.

Table 1: DSGE Model – Parameter Estimation Results

Parameter	Prior		Posterior		
	Mean	Stdd	Mean	90% Interval	
ψ_1	1.500	0.250	1.433	1.131	1.770
ψ_2	0.125	0.100	0.361	0.102	0.596
ρ_r	0.500	0.200	0.834	0.800	0.869
r^*_a	1.000	1.000	0.577	0.000	1.298
π^*_a	3.000	2.000	4.602	3.085	6.073
γ_a	2.000	1.000	1.945	1.358	2.518
h	0.800	0.100	0.987	0.979	0.997
ν_l	2.000	0.750	2.464	1.131	3.684
ζ_w	0.750	0.100	0.721	0.538	0.957
ζ_p	0.750	0.100	0.794	0.725	0.868
s'	4.000	1.500	6.274	3.734	8.725
$a'\,'$	0.200	0.075	0.225	0.109	0.332
g^*	0.150	0.050	0.131	0.057	0.200
ρ_g	0.800	0.050	0.904	0.867	0.943
σ_z	0.400	2.000	2.086	1.234	2.958
σ_g	0.300	2.000	0.551	0.470	0.634
σ_r	0.200	2.000	0.142	0.121	0.162

Notes: The table reports prior means and standard deviations, and posterior means and 90 percent probability intervals for the estimated parameters. See Section 2 for a definition of the DSGE model's parameters, and Section 4 for a description of the data. We are reporting annualized values for π^*, r^*, and γ (a-subscript). The following parameters were fixed: $\alpha = 0.25$, $\delta = 0.025$, $\iota_p = \iota_w = 0$, $\mathcal{F} = 0$, $\lambda_f = \lambda_w = 0.3$, $\chi = 0$, $\nu_m = 2$, $\rho_z = 0$. All shocks other than ϵ_z, ϵ_R, ϵ_g are equal to zero.

Table 2: Log Marginal Data Densities and Posterior Odds

Specification	$\ln p(Y)$	Post Odds
DSGE Model	-313.58	4E-11
DSGE-VAR, $\lambda = 5.0$	-297.01	6E-04
DSGE-VAR, $\lambda = 2.0$	-293.96	0.012
DSGE-VAR, $\lambda = 1.5$	-292.83	0.039
DSGE-VAR, $\lambda = 1.0$	-290.88	0.270
DSGE-VAR, $\lambda = .75$	-289.58	1.000
DSGE-VAR, $\lambda = .50$	-289.78	0.816
DSGE-VAR, $\lambda = .25$	-298.50	1E-04

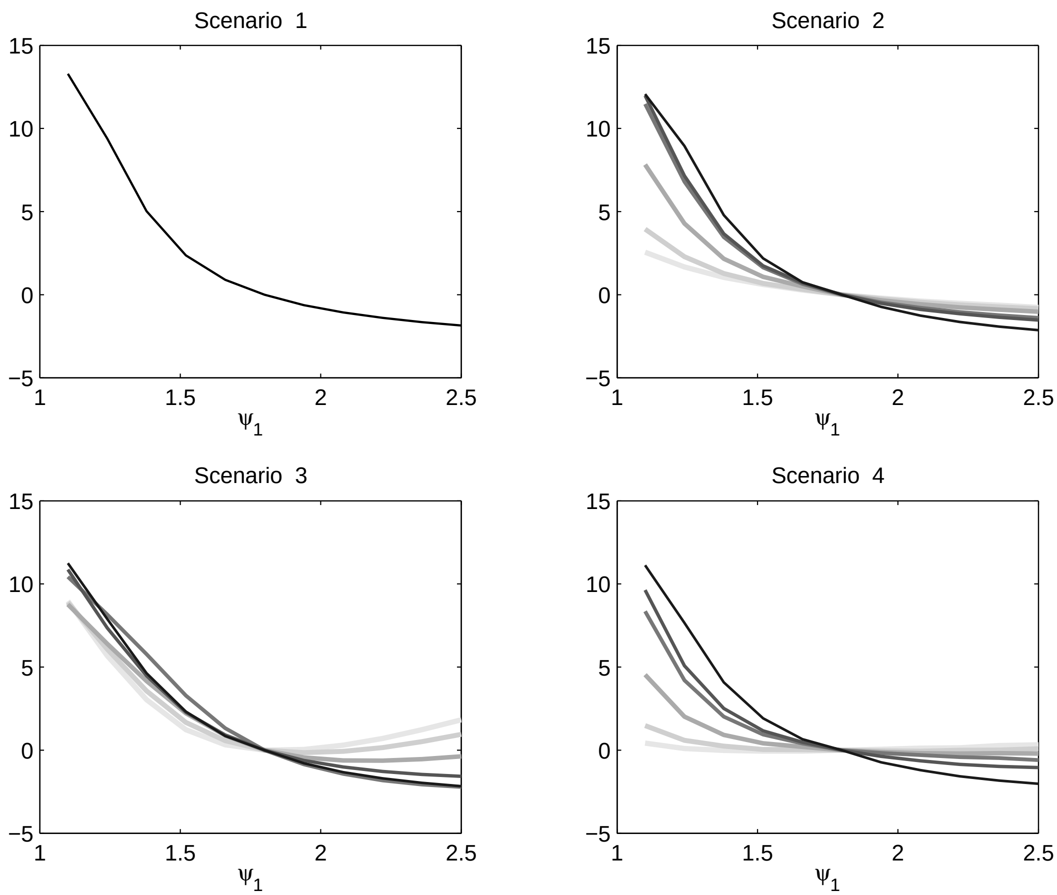
Notes: The marginal data densities are obtained by integrating the likelihood function with respect to the model parameters, weighted by the prior density conditional on λ. The difference of log marginal data densities can be interpreted as log posterior odds under the assumption of that the two specifications have equal prior probabilities.

Table 3: DSGE-VAR – Parameter Estimation Results

Parameter	Prior Mean	Posterior Mean (λ)							
		0.25	0.5	0.75	1.0	1.5	2.0	5.0	DSGE
ψ_1	1.500	1.990	1.788	1.774	1.714	1.824	1.669	1.650	1.433
ψ_2	0.125	0.275	0.278	0.263	0.259	0.285	0.296	0.315	0.361
ρ_r	0.500	0.836	0.845	0.849	0.857	0.861	0.855	0.856	0.834
r^*_a	1.000	0.537	0.378	0.346	0.378	0.498	0.419	0.515	0.577
π^*_a	3.000	3.596	3.392	3.442	3.431	3.782	3.704	3.980	4.602
γ_a	2.000	1.925	1.879	2.081	1.943	2.214	2.044	2.218	1.945
h	0.800	0.944	0.882	0.919	0.970	0.982	0.984	0.987	0.987
ν_l	2.000	2.043	2.161	2.097	2.245	2.326	2.501	2.451	2.464
ζ_w	0.750	0.726	0.728	0.732	0.755	0.727	0.739	0.745	0.721
ζ_p	0.750	0.699	0.618	0.640	0.688	0.739	0.746	0.773	0.794
$a'\,'$	0.200	0.204	0.203	0.220	0.207	0.197	0.214	0.208	0.225
s'	4.000	4.296	4.429	4.503	4.565	4.540	4.500	5.091	6.274
g^*	0.150	0.149	0.158	0.139	0.142	0.141	0.142	0.136	0.131
ρ_g	0.800	0.813	0.823	0.822	0.826	0.823	0.831	0.836	0.904
σ_z	0.400	0.956	0.912	1.033	1.322	1.689	1.837	2.139	2.086
σ_g	0.300	0.303	0.339	0.365	0.369	0.376	0.390	0.424	0.551
σ_r	0.200	0.123	0.129	0.132	0.128	0.132	0.134	0.137	0.142

Notes: The table reports prior and posterior means for the DSGE-VAR as a function of λ and the DSGE model. See Section 2 for a definition of the DSGE model's parameters, and Section 4 for a description of the data. We are reporting annualized values for π^*, r^*, and γ (a-subscript). The following parameters were fixed: $\alpha = 0.25$, $\delta = 0.025$, $\iota_p = \iota_w = 0$, $\mathcal{F} = 0$, $\lambda_f = \lambda_w = 0.3$, $\chi = 0$, $\nu_m = 2$, $\rho_z = 0$. All shocks other than ϵ_z, ϵ_R, ϵ_g are equal to zero.

Figure 1: Expected Policy Loss Differentials



Notes: Mean policy loss differentials relative to baseline policy rule $\psi_1 = 1.8$, $\psi_2 = 0.25$, $\rho_R = 0.85$: DSGE model (black), large values of λ (dark grey), and small values of λ (light grey). Negative differentials signify an improvement relative to baseline rule.