On the Fit and Forecasting Performance of New-Keynesian Models

Marco Del Negro, Frank Schorfheide, Frank Smets, and Raf Wouters*

October, 2004

Abstract

We estimate a large-scale New Keynesian dynamic stochastic general equilibrium (DSGE) model with price and wage stickiness and capital accumulation. Following the approach developed by Del Negro and Schorfheide (2004), we also use the DSGE model to generate a prior distribution for a vector autoregression. We compare the fit and forecasting performance of this DSGE-VAR to that of the directly estimated DSGE model. We find strong evidence that even large scale DSGE models are to some extent misspecified and that relaxing some of their restrictions leads to improved empirical performance.

Preliminary and Incomplete

*Marco Del Negro: Federal Reserve Bank of Atlanta, e-mail: *Marco.DelNegro@atl.frb.org*; Frank Schorfheide: University of Pennsylvania, Department of Economics, e-mail: *schorf@ssc.upenn.edu*; Frank Smets: European Central Bank and CEPR, e-mail: *Frank.Smets@ecb.int*; Raf Wouters: National Bank of Belgium, email: *Rafael.Wouters@nbb.be*. We thank seminar participants at the Atlanta Fed, New York University, Northwestern University, Richmond Fed, the University of Virginia, the workshop on "Empirical Methods and Applications to DSGE Models" at the Cleveland Fed, the $2^{n}d$ Euro Area Business Cycle Network (ECB, Fall 2003), 2004 SED, and 2004 SCE conferences for useful comments. . . . The views expressed in this papers are solely our own and do not necessarily reflect those of the Federal Reserve Bank of Atlanta, the Federal Reserve System, the European Central Bank, or the National Bank of Belgium.

JEL CLASSIFICATION: C11, C32, C53

KEY WORDS: Bayesian Analysis, DSGE Models, Forecasting, Vector Autoregressions

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are not just attractive from a theoretical perspective, but they are also emerging as useful tools for forecasting and quantitative policy analysis in macroeconomics. If some of the models used in the past had poor fit and poor forecasting record (see for instance Schorfheide 2000), a new generation of model in the new-Keynesian tradition appears to be competitive with more heavily parameterized models, like vector autoregressions (VARs), in terms of both in- and out-of-sample fit. Christiano et al. (20004), Altig et al. (2002) and Smets and Wouters (2003a, 2003b) have developed an elaborate DSGE model with capital accumulation as well as various nominal and real frictions. Smets and Wouters (2003a, 2003b) evaluate forecasting performance and Bayesian posterior odds of their model versus a VAR based on (detrended) Euro-area data and U.S. data (Smets and Wouters, 2003b), and show that in the U.S. case the data strongly favor the DSGE model. These findings have had a notable impact on the profession, for they imply a different approach to both empirical macroeconomics and policy making than the one used up to now.

Given the relevance of Smets and Wouters' results, this paper takes a closer look at their model's fit and forecasting performance. We estimate a close variant of their model to the very same variables they use, namely output, consumption, investment, hours worked, nominal wages, prices, and interest rates, except that the data are not detrended. First, we ask whether there is any indication that the model is in disagreement with the data by looking at the model's exogenous driving processes. We find that many of the exogenous driving processes are very persistent, which was well-known, but also that some display a clear trend, in contrast with the stated assumption of stationarity. The exogenous processes have to supply for the model's inability to fit the data in some important dimensions. This heavy reliance on exogenous processes is not welcome, because it brings the model's usefulness for policy analysis into question. If a lot of the action is in the exogenous part of the model and the exogenous processes capture to some extent model misspecification, is it appropriate to assume that the driving processes are policy invariant?

In light of these findings, we ask whether relaxing the DSGE model's cross-equation restrictions, without necessarily going all the way to the unrestricted VAR, can lead to an improvement in in- and out-of-sample fit. In a recent paper, Del Negro and Schorfheide (2004) show that forecasts with a simple three equation New Keynesian DSGE model can be improved by systematically relaxing the DSGE model restrictions. In the framework of Del Negro and Schorfheide the DSGE model is used to generate a prior distribution for the coefficients of a VAR. The prior concentrates most of its probability mass near the restrictions that the DSGE model imposes on the VAR representation and tilts the likelihood estimate of the VAR parameters toward the DSGE model restrictions, without dogmatically imposing them. We will refer to the resulting specification as DSGE-VAR. The framework of Del Negro and Schorfheide can be viewed as an attempt to connect VAR and DSGE model using a hyperparameter – the weight of the DSGE model's prior. The procedure amounts to estimating an unrestricted VAR when the weight of the prior is zero, and the VAR approximation to the DSGE model when the weight goes to infinity.

The accuracy of the VAR approximation of the DSGE model depends on the number of autoregressive lags and the specification of the VAR. In this paper we consider two versions: a VAR in terms of growth rates and a vector error correction model (VECM) that includes the cointegration vectors implied by the DSGE model. We find that with four lags only the VECM is able to approximate the DSGE model reasonably well, in the sense that it produces impulse-response functions that are very close to those of the DSGE model, while it takes ten times as many lags for the VAR in differences to reach the same degree of accuracy.

We assess the empirical performance of the DSGE model based on three criteria. First we construct a posterior distribution for the hyperparameter that controls the weight of the DSGE model prior. Sims (2003) points out that the posterior probabilities of DSGE model versus VAR computed by Smets and Wouters tend to switch between the extremes zero and one, depending on the choice of the data set (U.S. versus Euro-area) and the specification of the VAR prior (Minnesota prior versus training sample prior). In his view these probabilities do not give an accurate reflection of model uncertainty and are largely an artifact of a model space that is too sparse. Following arguments in Gelman, Carlin, Stern, and Rubin (1994), Sims advocates filling the model space by connecting distinct model specifications with continuous parameters and characterize the model uncertainty through the posterior probability distribution of these additional parameters. We create a continuum of models by letting the hyperparameter vary between zero and infinity. We find that for both the VAR and the VECM specification the posterior of the hyperparameter is bell-shaped and peaks for values of the hyperparameter that imply a non-negligible, yet not too large, weight on the DSGE model prior. The best fitting model is therefore equally distant from the VAR or VECM as it is from the DSGE model. This finding suggests, as did the informal inspection of the smoothed exogenous shocks mentioned earlier, that the

DSGE model is to some extent in disagreement with the data.

Second, out-of-sample forecasting statistics are computed for the DSGE-VARs and the DSGE model itself. The DSGE-VAR dominates both the unrestricted VAR and the DSGE model, which again confirms that relaxing the DSGE model restriction helps fitting the data. Third, we compare the impulse responses to structural shocks of the DSGE-VAR and the DSGE model to obtain an indication in which dimension the disagreement between the two specifications is most substantial.

The paper is organized as follows. The DSGE model is presented in Section 2. Section 3 reviews the approach of Del Negro and Schorfheide, and discusses extensions to vector autoregressive models with non-stationary endogenous variables. Section 4 describes the data. Empirical results are presented in Section 5 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 (2003a). 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 $i \in [0,1]$:

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

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

$$(\ln \lambda_{f,t} - \ln \lambda_f) = \rho_{\lambda_f}(\ln \lambda_{f,t-1} - \ln \lambda_f) + \sigma_{\lambda,f}\epsilon_{\lambda,t}, \tag{2}$$

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.[1] The random walk assumption for the technology process 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:

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

[1] Smets and Wouters assume a stationary technology shock that follows an autoregressive process. Their estimate of the autocorrelation coefficient however are very close to the upper boundary of one. We therefore choose to assume a unit root process from the onset.

where $\beta^s \Xi^p_{t+s}$ 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:

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

where $\lambda_w \in (0,\infty)$ is a fixed parameter (Smets and Wouters (2003a) assume that i.i.d. shocks to the degree of labor substitutability are another source of disturbance in the economy). 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:

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

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:

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

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, $A_t(j)$ is the net cash inflow from participating in state-contingent securities, 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 $R^k_t u_t(j)\bar{K}_{t-1}(j)$ 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:

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

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:

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

where $\epsilon_{R,t}$ is the monetary policy shock, R^* is the steady state nominal rate, Y_t^* is the target level of output, and the parameter ρ_R determines the degree of interest rate smoothing. This specification of the Taylor rule is more standard than the one in Smets and Wouters (2003a), who introduce a time-varying inflation objective that varies stochastically according to a random walk. (ADD WHY?) If empirically the random walk inflation target may help, it also bears the implication that inflation is ultimately outside the control of the monetary authorities. This is an undesirable feature since we aim to use this model for policy analysis. We consider two alternative specifications for the target level of output Y_t^* in (21). Under one specification the monetary authorities target the trend level of output: $Y_t^* = Y_t^s$. Under the alternative specification they target the level of output that would have prevailed in absence of nominal rigidities: $Y_t^* = Y_t^f$. The central bank supplies the money demanded by the household to support the desired nominal interest rate.

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 Estimation

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. We collect all the DSGE model parameters in the vector θ, stack the structural shocks in the vector ϵ_t, and derive a state-space representation for the $n \times 1$ vector Δy_t:

$$\Delta y_t = [\Delta \ln Y_t, \Delta \ln C_t, \Delta \ln I_t, \Delta \ln(W_t/P_t), \ln L_t, \pi_t, R_t]',$$

where Δ denotes the temporal difference operator and π_t is the inflation rate. From the state-space representation we can obtain a likelihood function, which we will generically denote by $p_{DSGE}(Y|\theta)$, where Y is the $T \times n$ matrix of available data, with rows $\Delta y_t'$. This likelihood function is combined with a prior density for the structural parameters θ to conduct Bayesian inference. The fit of a DSGE model relative to alternative specifications is assessed using the marginal likelihood (or marginal data density) $p(Y|DSGE)$, defined as

$$p(Y|DSGE) = \int p_{DSGE}(Y|\theta)p(\theta)d\theta. \tag{26}$$

Under the assumption of equal prior probabilities, ratios of marginal likelihoods can be interpreted as model odds.

3 DSGE-VAR as a Toolkit for DSGE Model Evaluation

The DSGE model generates a restricted and potentially misspecified moving average representation for the vector Δy_t. A less restrictive, more flexible MA representation can be obtained, for instance, from a vector autoregression with p-lags:

$$\Delta y_t = \Phi_0 + \Phi_1 \Delta y_{t-1} + \ldots + \Phi_p \Delta y_{t-p} + u_t. \tag{27}$$

We will assume that the vector of reduced-form innovations u_t is normally distributed with mean zero and covariance matrix $\Sigma_u = \Sigma_{tr}\Sigma_{tr}'$, where Σ_{tr} is the lower-triangular Cholesky decomposition matrix of Σ_u. Moreover, we will assume that the reduced-form innovations are a functions of the structural shocks ϵ_t that generate the fluctuations in the DSGE model:

$$u_t = \Sigma_{tr}\Omega\epsilon_t. \tag{28}$$

Ω is an orthonormal matrix with the property $\Omega\Omega' = I$ and does not affect the variance of u_t. While the VAR in Eq. (27) could in principle be estimated using ordinary least squares techniques, macroeconomists typically do not have enough observations to estimate the VAR coefficients accurately if the dimension n of Δy_t is large.

Del Negro and Schorfheide (2004) proposed a Bayesian procedure that tilts the VAR coefficient estimates toward the restrictions implied by the DSGE model. Vice versa, the approach can be interpreted as relaxing the DSGE model restrictions on the moving average representation of Δy_t to overcome potential model misspecification. Loosely speaking, the procedure amounts to adding artificial observations generated from the DSGE model to the actual observations and then estimating the VAR based on this augmented data set. From a Bayesian perspective the artificial observations generate a prior distribution for the VAR coefficients that is centered around the DSGE model restrictions.

The degree to which we allow our coefficient estimates deviate from the DSGE model restrictions depends on the number of dummy observations by which we are augmenting the actual sample, or, in other words, on the weight that we are placing on the DSGE model prior. Let λ denote the ratio of artificial to actual observations. When λ equals zero, the resulting specification is nothing but an unrestricted VAR. As λ increases, the prior will tilt the VAR coefficients toward the cross-coefficient restrictions coming from the DSGE model.

For $\lambda = \infty$ we obtain a finite-order VAR approximation of the DSGE model. Hence, we have a continuum of models indexed by the hyperparameter $\lambda \in [0, \infty]$ of the DSGE prior that has the unrestricted VAR at one extreme and, to the extent that the VAR approximation is accurate, the DSGE model at the other extreme. We will call these models DSGE-VAR(λ), where by "model" here we mean a likelihood function combined with a particular prior distribution, indexed by λ.

The DSGE model presented in Section 2 implies that the set of variables that we consider for our empirical analysis has several common trends. For instance, output, consumption, and investment all grow that the rate Z_t. This suggests that we could obtain a better approximation of the DSGE model if we generate a moving-average representation from the following vector error correction (VECM) specification

$$\Delta y_t = \Phi_0 + \Phi_\beta(\beta' y_{t-1}) + \Phi_1 \Delta y_{t-1} + \ldots + \Phi_p \Delta y_{t-p} + u_t, \tag{29}$$

where the error correction terms are defined as

$$\beta' y_{t-1} = \begin{bmatrix} \ln C_t - \ln Y_t \\ \ln I_t - \ln Y_t \\ \ln(W_t/P_t) - \ln Y_t \end{bmatrix}. \tag{30}$$

We call DSGE-VECM(λ) the VECM with DSGE model prior. DSGE-VECM(λ) is a novelty of this paper, given that in the simple specification considered by Del Negro and Schorfheide (2004) no cointegration relationship was present.

The orthonormal matrix Ω in Eq. (28) is not identifiable since the likelihood function depends only on the covariance matrix

$$\Sigma_u = \Sigma_{tr}\Omega\Omega'\Sigma_{tr}' = \Sigma_{tr}\Sigma_{tr}'.$$

While in principle it is difficult to specify an identification scheme for all the structural shocks if the dimension n of the VAR is large, we are able to do so by using the identification information embodied in the DSGE model. The identification of structural shocks is important for two reasons. First we are able to assess in which dimensions the VAR/VECM is able to approximate the DSGE model, and second, we can evaluate the DSGE model by comparing its impulse response functions to those from the identified DSGE-VAR/VECM.

The remainder of this section reviews the specification of the prior and the characterization of the posterior in some more detail, as well as the identification scheme that determines Ω. A full description can be found in Del Negro and Schorfheide (2004). For the sake of exposition we will only refer to the VAR case: everything applies to the VECM case as well. The last part of this section discusses the accuracy with which DSGE-VARs and DSGE-VECMs can approximate the DSGE model when the prior weight goes to infinity.

3.1 The Likelihood Function

The likelihood for the VAR and the VECM are standard. We assume that the innovations u_t have a multivariate normal distribution $\mathcal{N}(0, \Sigma_u)$ conditional on past observations of Δy_t. Let Y be the $T \times n$ matrix with rows $\Delta y_t'$. Let $k = 1 + np$, X be the $T \times k$ matrix with rows $x_t' = [1, \Delta y_{t-1}', \ldots, \Delta y_{t-p}']$, U be the $T \times n$ matrix with rows u_t', and $\Phi = [\Phi_0, \Phi_1, \ldots, \Phi_p]'$. The VAR can be expressed as $Y = X\Phi + U$ with likelihood function

$$\begin{aligned} & p(Y|\Phi, \Sigma_u) \propto \\ & \quad |\Sigma_u|^{-T/2} \exp\left\{ -\frac{1}{2} tr[\Sigma_u^{-1}(Y'Y - \Phi'X'Y - Y'X\Phi + \Phi'X'X\Phi)] \right\} \end{aligned} \tag{31}$$

conditional on observations $\Delta y_{1-p}, \ldots, \Delta y_0$. For the VECM specification we simply redefine the matrices X and Φ. Let $k = 2 + np$, X be the $T \times k$ matrix with rows $x_t' = [1, (\beta' y_{t-1})', \Delta y_{t-1}', \ldots, \Delta y_{t-p}']$, and $\Phi = [\Phi_0, \Phi_\beta, \Phi_1, \ldots, \Phi_p]'$.

3.2 Generating the DSGE Model Prior

Since the DSGE model depends on the structural parameters θ we have to specify a joint prior distribution for θ and the VAR coefficient matrices Φ and Σ_u. We factorize the joint prior density as follows:

$$p_\lambda(\Phi, \Sigma_u, \theta) = p(\theta) p_\lambda(\Phi, \Sigma_u|\theta) \tag{32}$$

and focus for now on the prior distribution of the VAR parameters given θ and the relative length of the artificial sample λ. The prior density is of the form

$$p_\lambda(\Phi, \Sigma_u|\theta) \propto |\Sigma_u|^{-\lambda T/2} \exp\{-\frac{\lambda T}{2} tr[\Sigma_u^{-1}(\Gamma_{YY}^*(\theta) - 2\Phi'\Gamma_{XY}^*(\theta) + \Phi'\Gamma_{XX}^*(\theta)\Phi)]\}. \tag{33}$$

Here, Γ_{YY}^*, Γ_{XX}^*, and Γ_{XY}^* are the population moments $\mathbb{E}_\theta[\Delta y_t \Delta y_t']$, $\mathbb{E}_\theta[x_t x_t']$, and $\mathbb{E}_\theta[\Delta y_t x_t']$, calculated from the state-space representation of the DSGE model. These terms are scaled by λT can be interpreted as expected values of sample moments of artificial data generated from the DSGE model. We can rewrite (33) as the product of a Inverted-Wishart ($\mathcal{IW}$) density for Σ_u and a Normal ($\mathcal{N}$) density for Φ given Σ_u. These distributions are centered around

$$\Phi^*(\theta) = \Gamma^{*-1}_{XX}(\theta)\Gamma^*_{XY}(\theta) \quad (34)$$

$$\Sigma^*_u(\theta) = \Gamma^*_{YY}(\theta) - \Gamma^*_{YX}(\theta)\Gamma^{*-1}_{XX}(\theta)\Gamma^*_{XY}(\theta) \quad (35)$$

and their variance is a decreasing function of λ. $\Phi^*(\theta)$ and $\Sigma^*_u(\theta)$ are the coefficients of the VAR approximation to the DSGE model.

3.3 The Posterior of the VAR Parameters

The posterior density is proportional to the prior density times the likelihood function. We factorize the posterior distribution into the posterior density of the VAR parameters given the DSGE model parameters and the marginal posterior density of the DSGE model parameters:

$$p_\lambda(\Phi, \Sigma_u, \theta|Y) = p_\lambda(\Phi, \Sigma_u|Y, \theta)p_\lambda(\theta|Y). \quad (36)$$

It is straightforward to show, e.g., Zellner (1971), that the posterior distribution of Φ and Σ is also of the Inverted Wishart – Normal form:

$$\Sigma_u|Y, \theta \sim \mathcal{IW}\Big((\lambda+1)T\tilde{\Sigma}_u(\theta), (1+\lambda)T - k, n\Big) \quad (37)$$

$$\Phi|Y, \Sigma_u, \theta \sim \mathcal{N}\Big(\tilde{\Phi}(\theta), \Sigma_u \otimes (\lambda T\Gamma^*_{XX}(\theta) + X'X)^{-1}\Big), \quad (38)$$

where $\tilde{\Phi}(\theta)$ and $\tilde{\Sigma}_u(\theta)$ are the maximum-likelihood estimates of Φ and Σ_u, respectively, based on both the artificial and the actual sample:

$$\tilde{\Phi}(\theta) = (\lambda T\Gamma^*_{XX}(\theta) + X'X)^{-1}(\lambda T\Gamma^*_{XY} + X'Y) \quad (39)$$

$$\begin{aligned}\tilde{\Sigma}_u(\theta) = \frac{1}{(\lambda+1)T}\Big[&(\lambda T\Gamma^*_{YY}(\theta) + Y'Y) \\ &-(\lambda T\Gamma^*_{YX}(\theta) + Y'X)(\lambda T\Gamma^*_{XX}(\theta) + X'X)^{-1}(\lambda T\Gamma^*_{XY}(\theta) + X'Y)\Big].\end{aligned} \quad (40)$$

Expressions (39) and (40) show that the larger the weight λ of the prior, the closer the posterior of the VAR parameters is to $\Phi^*(\theta)$ and $\Sigma^*_u(\theta)$. The formula for the marginal posterior density of θ and the description of a Markov-Chain-Monte-Carlo algorithm that generates draws from the joint posterior of Φ, Σ_u, and θ are provided in Del Negro and Schorfheide (2004).

3.4 λ as Measure of Fit

Smets and Wouters (2003b) use marginal densities as to compare the fit of the DSGE model to that of VARs. Sims (2003) points out that Bayesian model comparison can lead to misleading conclusions when the model space is too sparse. By "sparse" we mean that the posterior densities associated with the models peak in different regions of an enlarged parameter space that can encompass all the models. In this case small changes in the data set or the prior distribution might move posterior model probabilities from zero to one, and vice versa. Moreover, these (degenerate) probabilities do not give an appropriate characterization of the uncertainty associated with the predictions from these models.

The class of DSGE-VARs "fills" the model space by means of the hyperparameter λ, and hence is able to generate a more insightful characterization of model uncertainty. As pointed out above, λ traces a continuum of models ranging from the unrestricted VAR to the VAR approximation of the DSGE model. We will study the fit of the DSGE model by looking at the posterior distribution of the hyperparameter λ. For computational reasons, we break the continuum into a discrete grid $\Lambda = \{l_1, \dots, l_q\}$. For each element of the grid point (that is, for each data generating process DSGE-VAR(λ)) the likelihood is given by the marginal data density

$$p_\lambda(Y) = \int p_\lambda(Y|\theta)p(\theta)d\theta. \tag{41}$$

For a flat prior on the grid points, the shape of $p_\lambda(Y)$ determines the shape of the posterior.

3.5 Identification

In order to identify structural disturbances in the DSGE-VAR framework we have to determine the orthonormal matrix Ω in Eq. (28). Once identification has been achieved we can study the propagation and relative importance of structural shocks with the DSGE-VAR. A comparison with DSGE model impulse responses can generate important insights in the potential misspecification of the DSGE model.

We construct Ω as follows. The DSGE model itself is identified in the sense that for each value of θ there is a unique matrix $A_0(\theta)$, obtained from the state space representation of the DSGE model, that determines the contemporaneous effect of ϵ_t on Δy_t. Using a QR factorization of $A_0(\theta)$, the initial response of Δy_t to the structural shocks can be can be uniquely decomposed into

$$\left(\frac{\partial y_t}{\partial \epsilon_t'}\right)_{DSGE} = A_0(\theta) = \Sigma_{tr}^*(\theta)\Omega^*(\theta), \tag{42}$$

where $\Sigma^*_{tr}(\theta)$ is lower triangular and $\Omega^*(\theta)$ is orthonormal. According to Equation (27) the initial impact of ϵ_t on the endogenous variables y_t in the VAR is given by

$$\left(\frac{\partial \Delta y_t}{\partial \epsilon_t'}\right)_{VAR} = \Sigma_{tr}\Omega. \tag{43}$$

To identify the DSGE-VAR, we maintain the triangularization of its covariance matrix Σ_u and replace the rotation Ω in Equation (43) with the function $\Omega^*(\theta)$ that appears in (42). The implementation of this identification procedure is straightforward in our framework. Since we are able to generate draws from the joint posterior distribution of Φ, Σ_u, and θ, we can for each draw (i) use Φ to construct a MA representation of Δy_t in terms of the reduced-form shocks u_t, (ii) compute a Cholesky decomposition of Σ_u, and (iii) calculate $\Omega = \Omega^*(\theta)$ to obtain a MA representation in terms of the structural shocks ϵ_t.

3.6 How well is the DSGE Model Approximated?

At the beginning of this section we described DSGE-VAR(λ) (or DSGE-VECM(λ)) as a continuum of specifications with the unrestricted VAR at one extreme and the VAR approximation to the DSGE model at the other. The fact that for $\lambda = \infty$ we only obtain an approximation of the log-linearized DSGE model raises the question why we did not start out from a more general VARMA model that nests the moving average representation of the DSGE model. The answer is twofold. First, VARs have established themselves as popular and powerful tools for empirical research and forecasting in macroeconomics. Second, from a computational perspective the posterior of DSGE-VAR is much easier to analyze than the posterior of a VARMA model with DSGE prior.

The accuracy of the VAR approximation of the DSGE model depends on the number of included autoregressive lags and on the invertibility of the DSGE model's moving average components. The latter is a property of the log-linearized DSGE model whereas the former is typically constrained by the number of observations that are available to initialize VAR lags and to estimate the coefficients. In the remainder of this section we will examine the how well the DSGE-VAR and DSGE-VECM approximates the DSGE model specified in Section 2.

We fix the vector of "deep" parameters θ at the value that maximizes the posterior mode of the DGSE model. We then compare in Figure 1 the impulse responses from the DGSE model with those from DSGE-VAR(∞) and DSGE-VECM(∞) with four lags. These latter impulse responses are obtained using the identification procedure described in the previous section. The solid, dash-and-dotted, and dotted lines represent the impulse responses of the DSGE model, DSGE-VAR(∞), and DSGE-VECM(∞), respectively. The impulse responses to output, consumption, investment, and the real wage are cumulative.

For many of the impulse responses, for example those to monetary policy shock (last row), the approximation is good for both DSGE-VAR and DSGE-VECM. For instance, in terms of cumulative output the maximum difference between the DSGE model's and DSGE-VECM(∞)'s impulse responses over the impulse responses horizon (sixteen quarters) is 5 basis points, versus 8 basis points for DSGE-VAR(∞). For inflation and the interest rate the difference is less than one and 2 basis points, respectively. For other impulse responses however the approximation breaks down, most notably for the responses of investment to shocks to the capital accumulation equation (μ). There, investment reverts to steady state after sixteen periods according to the impulse responses of both the DSGE model and DSGE-VECM(∞), while according to the impulse responses of DSGE-VAR(∞) investment is well above steady state by the end of the horizon. The impulse responses of DSGE-VAR(∞) display a similar behavior in a number of other cases, while those of DSGE-VECM(∞) generally track the DSGE model's well. There are a few instances where the DSGE model's responses and DSGE-VECM(∞)'s are different, such as the responses of cumulative investment to a λ_f (markup) shock. Whenever this is the case, the magnitude of the difference in impulse-responses is still contained, however. The maximum differences between the DSGE model's and DSGE-VECM(∞)'s impulse responses are 64 basis points for cumulative investment, which is small relative to the overall variability of the series. To double check that even these minor differences eventually disappear, we also computed the responses of DSGE-VAR(∞) and DSGE-VECM(∞) with forty lags. Now, the impulse responses of the DSGE model are virtually indistinguishable from those of DSGE-VECM(∞), while the approximation of DSGE-VAR(∞) is about as good as that of DSGE-VECM(∞) with four lags only. This finding suggests that DSGE-VECM(∞) is a parsimonious way to approximate the DSGE model in presence of cointegrating restrictions, and appears to be fairly successful even with a moderate number of lags.[2] Given this finding, the remainder of the paper mainly focuses on results for DSGE-VECM, although we also discuss the results for the DSGE-VAR specification. In estimating and assessing the fit of the DSGE model we condition on the same information set used in DSGE-VECM. Specifically, we are conditioning on x_0, the p initial lags of the endogenous variables, as well as on the initial values

[2] To further investigate the issue of invertibility we randomly generated data from the DSGE model, and then checked whether DSGE-VECM(∞) was able to reproduce the original time series of structural shocks. We find that this is indeed the case, even with four lags.

of the cointegrating vector (Eq. 30).[3]

4 The Data

All data are obtained from Haver analytics (Haver mnemonics are in italics). Real output, consumption, and investment are obtained by dividing the nominal series (*GDP*, *C*, and *I*, respectively) by population 16 years and older (*LF+LH*), and deflating using the chained-price GDP deflator (*JGDP*). The real wage is computed by dividing compensation per hour in the nonfarm business sector (*LXNFC*) by the GDP deflator. Note that compensation per hours includes wages as well as employer contribution. It accounts for both wage and salary workers and proprietors. Labor supply is computed by dividing hours of all persons in the nonfarm business sector (*LXNFH*) by population. Hours of all persons in the nonfarm business sector is an index developed by the BLS that includes the labor supply of both wage and salary workers and proprietors. This measure of labor supply best corresponds to our measure of real wages.[4] All growth rate 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 fed funds rate (*FFED*), also in percent. Data are available from QIII:1954 to QI:2004.

5 Results

5.1 Choosing the DSGE Model Specification

We want to specify the DSGE model so that it has the best shot at fitting and forecasting the variables of interest. We consider the following choices: (i) Price/wage indexation to: (A) steady state inflation/wage growth ($\iota_p = \iota_w = 0$), also known as "static indexation" versus (B) indexation to last period's inflation/wage growth ($\iota_p = \iota_w = 1$), also known as "dynamic indexation"; (ii) Output gap in the Taylor rule (Eq.21) defined using: (A) the trend of output along the balanced growth path ($Y_t^* = Y_t^s$) versus (B) flexible price

[3] This is achieved by running the Kalman filter on the initial observations x_0, and then using the resulting mean and variance for the state as starting values in the estimation on $\Delta y_1 \dots \Delta y_T$. Note that the initial values of the cointegrating vector combined with the sample information $\Delta y_1 \dots \Delta y_T$ implies that we are effectively giving the model information on the values of the cointegrating vector for $t = 1 \dots T$.

[4] Since we use an index as a measure of hours, which enter our specification in level, we need to pin down the average value of the index via an additional free parameter in the DSGE model.

output ($Y_t^* = Y_t^f$); (iii) Fixed costs $\mathcal{F}$ in the production function for the intermediate goods producers in Eq. (4) set to: (A) zero versus (B) endogenously determined to erase steady state profits. Our baseline specification adopts all the (A) choices. Hence, it features indexation to steady state inflation/wage growth ($\iota_p = \iota_w = 0$), steady state output as the target output in the Taylor rule, and $\mathcal{F} = 0$. Table 1 provides a justification for this choice. For each alternative the table shows the differences of the logarithm of the marginal likelihood relative to the baseline and the associated posterior odds.[5] The posterior odds is equal to the exponential of the log-differences in marginal likelihood, and summarizes the odds that the alternative specification is more likely than the baseline given the data we have observed. Table 1 shows that alternatives (i)-B is clearly rejected by the data: the posterior odds are virtually zero. There is evidence that dynamic indexation is favored to static indexation in a single equation framework; that is, when only the Phillips curve is being estimated, as opposed to the entire DSGE model (Eichenbaum and Fisher 2004). This result does not appear to hold in a multiple equation framework such as the one considered here. Alternative (iii)-B (fixed cost $\mathcal{F}$ sets steady-state profits equal to zero) is also heavily rejected. The odds against alternative (ii)-B (flexible-price output target in the Taylor rule) are far less striking. Yet, the odds are still greater than 13/100, and hence favor the baseline model.

5.2 A First Look at the Fit of the DSGE Model

We now discuss the fit of the DGSE model. Figure 2 gives a qualitative impression of how well the DSGE model fits. The figure plots the actual data (dark line), as well as the one-period-ahead forecasts obtained from the Kalman filter (gray line), computed using the vector θ of DGSE model parameters that maximizes the posterior. While later on in this section we will give quantitative measures of fit based on Bayesian marginal likelihood and out-of-sample forecasting accuracy, and draw a comparison with other specifications, Figure 2 provides a first visual diagnostic of the model. The first impression of the model's fit is fairly satisfactory: there appear to be no big misses. In terms of real activity and real wages, the model seem to have a hard time fitting the most volatile periods, such as the mid-seventies. Importantly, the model consistently over-predicts consumption growth in the

[5] All results in this paper that are based on Monte Carlo computations are computed using 110,000 draws and discarding the first 10,000. We checked whether 110,000 draws were sufficient by repeating the estimation procedure a number of times and verifying we would obtain the same results in terms of marginal likelihood.

first part of the sample, and under-predicts consumption growth in the second part, except during the 1990 recession. In terms of nominal variables the model under-predicts inflation in the late seventies, and over-predicts inflation toward the end of the sample. Under (over) -predictions for inflation generally translate into under (over) -predictions for the nominal interest rate.

Table 2 reports on the estimates of the DSGE model parameters. The Table shows the prior mean and standard deviation, as well as the posterior mean and the 90% probability intervals for the estimates.[6] The parameter estimates are generally in line with those of Smets and Wouters (2003b). The model displays a relatively high degree of price and wage stickiness, as measured by the probability that firms (wage setters) cannot change their price (wage) in a given period: the posterior means of ζ_p and ζ_w are .848 and .936, respectively. Smets and Wouters present similarly high estimates of these parameters. Also of particular interest are the parameters describing the persistence of the underlying exogenous processes: ρ_z (technology), ρ_φ (preferences of leisure), ρ_μ (shocks to the capital accumulation equation), b (overall preference shifter), ρ_g (government spending), and ρ_{λ_f} (price markup shocks). The persistence of shocks to the technology growth rate z_t is low. While Smets and Wouters assume that the level of technology Z_t is stationary, and hence obtain a very high degree of persistence, we model technology shocks as a unit root process from the onset. All other processes are quite persistent, particularly those for the government spending shock g_t. However, for most of these shocks the degree of persistence is not nearly as high as that found in Smets and Wouters.

One of the recurrent remarks on the work of Smets and Wouters relates to the persistence of some of the exogenous processes. The near-unit root behavior of many of these processes raises concerns about the fit of the model, because it suggests a lack of a strong endogenous propagation mechanism. Figure 3 plots the Kalman-smoothed time series for the processes z_t (technology), φ_t (preferences of leisure), μ_t (shocks to the capital accumulation equation), b_t (overall preference shifter), g_t (government spending), and $\lambda_{f,t}$ (price markup shocks), computed using the vector θ of DGSE model parameters that maximizes the posterior. Not-surprisingly, given the estimates of the autoregressive coefficients shown in Table 2, many of the exogenous processes are indeed persistent. For instance, leisure preference shocks are positive in the first part of the sample, where total hours are generally lower than average (see Figure 2), and mostly negative in the second part and particularly in the nineties,

[6] A few of the DSGE model's parameters were fixed at the onset: δ, λ_f and λ_w were set at .025, .3 and .3, respectively.

where hours are above average. Shocks to the capital accumulation equation μ_t are also very persistent. Most disturbing of all is the behavior of the government spending process g_t. The process clearly has a downward negative trend that contrasts with the assumption of stationarity stated in Eq. 24. The reason for this negative trend can be traced to the consistent under-prediction of consumption starting in the early eighties, documented in Figure 2. According to the model investment, output, and consumption all grow at the same rate, when measured in nominal terms (or in real terms when deflated by the same price deflator). In the data, consumption has been growing faster than either output or investment. The model's inability to account for this fact may explain the downward path of g_t in Figure 3. The impulse responses (Figure 1) show that government spending shocks have the largest – and opposite – effect on output and consumption, and generally a small impact on investment and the other series ("small" relative to the overall volatility of the series, as can be gauged from other impulse-responses in the same column). Of course, government spending shocks can prop up the model's fit in-sample, but are less helpful out-of-sample, particularly as the forecast horizon increases. Indeed, we will see later that the DSGE model's long run forecasts of consumption are quite inaccurate relative to those of unrestricted VARs. Moreover, we will show that this inaccuracy can be corrected when the DSGE model's cross-equation restrictions are imposed only as priors, as in DSGE-VECM(λ) or DSGE-VAR(λ).

The above results are based on thirty years if data ($T = 120$), starting in QII:1974 and ending in QI:2004. We use thirty years because this the amount of data used in the estimation for the rolling sample forecasting exercise. The findings are qualitatively unchanged when we use the entire sample, from QIII:1955 to QI:2004. Also, we obtain similar results when we estimate the DSGE model without conditioning on the cointegrating vectors (Eq. 30).

5.3 The Fit of the DSGE Model Relative to VARs

One of the important messages from Smets and Wouters (2003a) is that, at least for the Euro area data, new-generation new-Keynesian DSGE models can fit better than VARs. There are two important caveats to their conclusion. First, in- and out-of-sample comparisons are based on linearly detrended data. The second caveat, which we discussed in section 3.4, is that the set of models Smets and Wouters consider is very "sparse". They consider the directly estimated DSGE model, a VAR with a training sample prior, and a VAR with a version of the Minnesota prior. The two VAR specifications do not use the DSGE model restrictions at all. In this comparison, the VAR with Minnesota prior slightly dominates the DSGE model, which is favored over the VAR with training sample prior.

Here we try to address both caveats. We assess the model's fit using data that are not detrended. Perhaps more importantly, we "fill the model space" by considering the fit of DSGE-VECM and DSGE-VAR for a whole range of values of λ, allowing for various degrees of deviation from the DSGE model restriction. As in Smets and Wouters, we use Bayesian marginal data densities to assess in-sample fit.[7]

Table 3 shows the marginal likelihood for DSGE-VECM(λ) and DSGE-VAR(λ) for values of λ ranging from .5 to ∞. The number of lags in the VECM/VAR equals four. Column (1) shows the value of the λ, the weight of the DSGE model prior. Column (2) shows the weight of the prior expressed as a ratio of the number of artificial to the number of total observations ($\frac{\lambda}{1+\lambda}$). For instance, $\lambda = .5$ amounts to artificial observations (λT) being 33 percent of the total ($(1+\lambda)T$). Seen in this perspective, $\lambda = 5$ and $\lambda = \infty$ are not very different, as the former amounts to artificial observations being 83 percent of the total versus (obviously) 100 percent for the latter. Columns (3) and (5) show the marginal likelihood for DSGE-VECM(λ) and DSGE-VAR(λ), respectively, while columns (4) and (6) show the same figures expressed as a deviations from the marginal likelihood of the best-fitting model. The last row shows the marginal likelihood associated with the DSGE model.

The difference between the figures in columns (3) and (5) of the last row of Table 3 show that conditioning on the initial level of the cointegrating vector does not help the fit of the DSGE model. This finding reflects the fact that the cointegrating restrictions imposed by the DSGE model are at odds with the data. When the DSGE model is given the information on the value of the cointegrating vector, it adjusts its forecast for all future period to reflect its belief that in the long run the cointegrating vector should return to zero. This adjustment worsens the forecast accuracy one period ahead, hence the decrease in the marginal likelihood. The decrease in marginal likelihood is common to all values of λ: for any fixed λ_0 the marginal likelihood of DSGE-VAR(λ_0) is always greater to that of DSGE-VECM(λ_0). As we will see in the discussion of the out-of-sample forecasts, the decrease in forecast accuracy is not confined to one quarter ahead forecasts.

Most importantly in terms of the results of Smets and Wouters, we find that the fit

[7]One minor drawback is that we cannot consider the marginal data density for the extreme case $\lambda = 0$, for in that case the prior is flat and the marginal data density is ill-defined. Unlike Smets and Wouters, we do not use training samples to form a proper prior. The DSGE model prior is proper only when $\lambda T \geq k + n$, which is why we do not show results for very small values of λ (say, .1) either.

of the DSGE model is far worse from that of the best fitting DSGE-VECM(λ) (or DSGE-VAR(λ)). The differences in marginal likelihood are so large that the posterior odds of the DSGE model are practically zero. Yet, if the DSGE model itself does not fare well, Table 3 also documents that the DSGE model prior is helpful in terms of fit. The best fitting model is one where the artificial observations are between 50 to 43 percent of the total, depending on whether we are using DSGE-VECM or DSGE-VAR. That is, the results imply that we should use as many artificial observations as there are actual observations. It is interesting to compare this finding with that obtained in Del Negro and Schorfheide (2004), who use a much simpler new-Keynesian model. There, the procedure recommended using half as many artificial observations as there are actual observations.[8]. Furthermore, the Table shows that the posterior distribution of λ is U-shaped, as we would expect. The fit worsens progressively as we use too few or too many artificial observations. It is important to remark that the marginal likelihood for $\lambda = \infty$ is not the same as that for the DSGE model. For DSGE-VECM the difference between the marginal likelihood for $\lambda = \infty$ and for the DSGE model is very small, consistently with the results in the section 3.6. For DSGE-VAR the difference in fit between $\lambda = \infty$ and the DSGE model is much larger. This is due to both the different conditioning information (DSGE-VAR does not use the cointegrating restrictions) and the differences in terms of impulse-responses emphasized in section 3.6. We have also estimated the DSGE model without conditioning on the cointegrating restrictions. We found that the difference narrows, but is still large, not surprisingly given the findings in the section 3.6.

Table 3 is based on thirty years if data ($T = 120$), starting in QII:1974 and ending in QI:2004. The results in Table 3 are remarkably robust: For each date of the rolling sample, from QIV:1985 to QI:2000, the shape of the posterior of λ is qualitatively the same, with the only difference that the peak of the posterior is $\lambda = .75$ for some dates and $\lambda = 1$ for others.

5.4 Out-of-sample Forecast Accuracy

We now discuss the out-sample fit of the DSGE model and compare it to that of unrestricted VECMs with and without DSGE model prior. We also obtained results for the VAR specification, which will discuss later and are available upon request. We assess out-of-sample forecasting accuracy using a rolling sample starting in QIV:1985 and ending in QI:2000, for

[8]Naturally a proper comparison involves applying the model to the same set of data, namely output growth, inflation, and the fed funds rate. We have done that, and the same finding obtains (we do not show these results for lack of space)

a total of fifty-eight periods. At each date of the rolling sample we use the previous 120 observations to estimate the model, and the following twelve quarters to assess forecasting accuracy, which is measured by the root mean squared error (RMSE) of the forecast. For the variables that enter the VECM in growth rates (output, consumption, investment, real wage) and inflation we forecast cumulative changes. For instance, the RMSE of inflation for twelve quarters ahead forecasts measures the error in forecasting cumulative inflation over the next three years (in essence, average inflation), as opposed to inflation exactly three years down the road. At each date, the forecasts from DGSE-VECM(λ) (or DGSE-VAR(λ)) are obtained using the value of λ that maximizes the marginal density, which we call $\hat{\lambda}$. As discussed above, this value hovers between .75 and 1. Note that the choice of $\hat{\lambda}$ is based on an "ex ante" optimality criterion, as opposed to an "ex post" one. Again, when forecasting with the DSGE model we condition on the very same information used by the VECM.

Table 4 shows for each series and for each forecast horizon the RMSE for DGSE-VECM(λ) as well as the percentage improvement (in parenthesis) in RMSE relative to both the DGSE model and the unrestricted VECM. The last three row of the Table show the corresponding figures for the multivariate statistic, a summary measure of joint forecasting performance, which is computed as the converse of the log-determinant of the variance-covariance matrix of forecast errors. For one-step ahead forecasts DGSE-VECM($\hat{\lambda}$) appears to be more accurate than both the DSGE model and the VECM. This is true for the multivariate statistic as well as for each of the individual series, consistently with the likelihood-based in-sample results of the previous section.[9] Relative to the DSGE model, the magnitude of the improvement in the multivariate statistics is largest for medium-run (four to eight quarters) forecasts, but declines in the longer run. For many of the individual series, such as output, consumption, investment, and hours, the improvement increases steadily with the forecast horizons. For forecasts of the real wage beyond one quarter ahead, however, DGSE-VECM($\hat{\lambda}$) actually does worse than the DSGE model, and the discrepancy rises with the forecast horizon. For most of the series other than the real wage, the improvement of DGSE-VECM($\hat{\lambda}$) relative to the DSGE model is fairly large, ranging from 15 to almost 70 percent. The improvements relative to the DSGE model are generally larger than those relative to the unrestrictive VECM. The improvements are substantial for output, consumption, hours and the interest rate, particularly for shorter forecast horizons. For

[9]Separately, we have also plotted the percentage increase in forecasting accuracy of DGSE-VECM(λ) relative to the unrestricted VECM, as measured by the RMSE. Consistently with the results in the previous section we find that for most series and forecasting horizons the increase in forecasting accuracy as a function of λ displays a U-shape, first increasing a then declining as λ goes from zero to infinity.

long run forecasts of investment the unrestricted VECM outperforms DGSE-VECM($\hat{\lambda}$). In summary, we find that DGSE-VECM generally improves relative to the unrestricted VECM, but that for most series and forecast horizons the improvement is even larger relative to the DSGE model. We reach by and large the same conclusions for the VAR specification (results are available upon request). Consistently with the results in Table 3 the VECM specification does worse than the VAR specification in terms of RMSEs for most of the variables with the exception of consumption. For series like investment the reduction in long run forecast accuracy is quite large. In conclusion, the error-correction term implied by the DSGE model appears to do much more harm than good when it comes to both in- and out-of-sample fit.

5.5 Comparing the Propagation of Shocks

We conclude the empirical analysis with a comparison of impulse-response functions calculated from the DSGE-VECM and the DSGE model. Such a comparison can yield insights into the nature of potential DSGE model misspecification. We have demonstrated in Section 3.6 that the discrepancy between the two sets of impulse responses practically vanishes as we increase the weight on the DSGE model prior. However, the results presented in Sections 5.3 and 5.4 show that the best fit of the DSGE-VECM is achieved when the hyperparameter λ is set equal to one, which indicates that the empirical performance of the model can be improved by relaxing its restrictions.

Figure 4 shows the mean impulse-responses for DGSE-VECM($\hat{\lambda}$) (dash-and-dotted lines), the ninety percent bands (dotted lines), and the mean impulse-responses for the DGSE model. The impulse-responses for the DGSE model are computed using the same draws of DSGE model parameters θ that generate the DGSE-VECM($\hat{\lambda}$) impulse-responses. One important feature of the procedure developed in this paper is that it delivers identified impulse responses even for values of λ that are less than infinity. Figure 4 shows that even for values of λ that are not too high ($\lambda = 1$) the identification procedure is fairly successful also for relatively large systems with as many as seven shocks. By successful we mean that the impulse-responses to all seven shocks are economically interpretable, in that they agree with the DSGE model in terms of sign restrictions (see Canova and De Nicoló (...) and Uhlig (...)). We find that several of DGSE-VECM($\hat{\lambda}$)'s impulse-responses are not only qualitatively but also quantitatively in agreement with the DSGE model's. This is the case for the responses to capital adjustment shocks (μ), and mark-up shocks (λ_f). Other impulse responses exhibit discrepancies. The impulse-response to a technology growth shock (*Tech*) is more pronounced in the medium run for output, consumption, investment and hours under DGSE-VECM($\hat{\lambda}$) than under the DSGE model. Also, the response of inflation is more persistent. The effects of the preference shock (φ) on output, consumption, and hours are more persistent according to the VECM specification, which indicates a lack of internal propagation of labor supply shifts. The intertemporal preference shock (b) has a much larger effect on the nominal interest rate in the VECM than it has in the DSGE model. Since b and R are related through the consumption Euler equation, the discrepancy suggests potential misspecification of the consumption-based pricing kernel. According to the DSGE model, output and hours increase immediately in response to a government spending shock and quickly decay monotonically. The VECM specification, on the other hand, predicts delayed, hump-shaped responses of both output and hours that are long-lasting. Moreover, the VECM implies that the government shock is accompanied by a fall in nominal interest rates whereas the DSGE model generates a small rise in R. While both DSGE model and VECM agree on the response of inflation and interest rates to a monetary policy shock, the VECM generates much stronger real effects, suggesting that the DSGE model underestimates the nominal rigidities in the economy.

6 Conclusions

Smets and Wouters (2003a) showed that new-generation New-Keynesian models with real and nominal rigidities can fit as well as unrestricted VARs, and possibly better, and be readily usable for policy analysis. Given the relevance of Smets and Wouters' results, this paper takes a closer look at their model's fit and forecasting performance. We find that in order to fit the data the model needs to rely heavily on the exogenous driving processes. Many of these processes are very persistent, which was well-known, but some also display a clear trend, in contrast with the stated assumption of stationarity. The reliance on exogenous processes brings into question the model's usefulness for policy analysis. In light of these findings, we ask whether we can improve fit by relaxing the DSGE model's cross-equation restrictions along the lines suggested by Del Negro and Schorfheide (2004), without necessarily going all the way to the unrestricted VAR. We find that a vector error correction model with DSGE model prior (VECM-DSGE) can well approximate the DSGE model when the weight of the prior approaches infinity. We also find that VECM-DSGE can outperform the DSGE model in terms of in- and out-of-sample fit when the weight is less than infinity, but still non-negligible. Yet this is only a partial success, because regardless of the weight of the DSGE prior the DSGE-VECM inherits a seemingly counterfactual implication of the DSGE model: that a number of real variables are equally affected by the same stochastic trend. We show that by relaxing this assumption (that is, using DSGE-VAR) we can further improve both fit and forecasting performance. The cost is that DSGE-VAR cannot well approximate the DSGE model with a moderate number of lags.

The conclusions of this paper is that much work lies ahead both in terms of modeling and econometrics. We need to build models that can be successfully taken to non-detrended data – models that fulfill Kydland and Prescott (1982)'s original promise of integrating growth and business cycle theory, so they can at the same time match both growth and business cycle features of the data. On the econometrics side we need to develop approaches that use the DSGE model prior, but down-weight those frequencies where the DSGE model's implications are more at odds with the data, and emphasize those where the DSGE model may be most useful. Progress in either direction may further enhance the use of DSGE models in quantitative policy analysis – the ultimate goal of our research agenda.

References

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

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

Chang, Yongsung, Joao Gomes, and Frank Schorfheide (2002): "Learning-by-Doing as Propagation Mechanism," *American Economic Review*, **92**(5), 1498-1520.

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 (2004): "Priors from General Equilibrium Models for VARs," *International Economic Review*, **45**, 643-673.

Doan, Thomas, Robert Litterman, and Christopher Sims (1984): "Forecasting and Conditional Projections Using Realistic Prior Distributions." *Econometric Reviews*, **3**, 1-100.

Eichenbaum, Fisher, 2004.

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.

Schorfheide, Frank (2000): "Loss Function-Based Evaluation of DSGE Models." *Journal of Applied Econometrics*, **15**, S645-670.

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

Sims, Christopher A. (2003): "Probability Models for Monetary Policy Decisions," *Manuscript*, Princeton University.

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

Smets, Frank and Raf Wouters (2003b): "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach," *Manuscript*, European Central Bank.

Zellner, Arnold (1971): *"Introduction to Bayesian Inference in Econometrics."* John Wiley & Sons, New York.

Table 1: Marginal Likelihood and Posterior Odds of Alternative DSGE Model Specifications

		Log-Differences in Marginal Likelihood wrt Baseline Model	Posterior Odds wrt Baseline Model (in %)
(i)	Dynamic Indexation ($\iota_p = \iota_w = 1$)	-31.49	10^{-12} %
(ii)	Flexible-price Output Target in Eq. (21)	-2.08	12.49 %
(iii)	$\mathcal{F}$ sets steady-state profits $= 0$	-28.87	10^{-11} %

Notes: Posterior odds are computed as the exponential of the log-differences in marginal likelihood between two models, and are expressed in percent. See Section 4 for a description of the data. Results are based on the sample period QII:1974 - QI:2004.

Table 2: DSGE Model's Parameter Estimates

Parameter	Prior Mean	Prior Stdd	Post Mean	90% Lower Band	90% Upper Band
α	0.250	0.100	0.172	0.149	0.195
ζ_p	0.750	0.100	0.848	0.814	0.883
s'	4.000	1.500	5.827	3.238	8.115
h	0.800	0.100	0.793	0.725	0.857
a''	0.200	0.075	0.167	0.067	0.273
ν_l	2.000	0.750	2.204	1.050	3.271
ζ_w	0.750	0.100	0.936	0.913	0.959
r^*	0.500	0.100	0.389	0.270	0.501
ψ_1	1.700	0.100	1.799	1.640	1.944
ψ_2	0.125	0.100	0.065	0.040	0.090
ρ_r	0.800	0.100	0.815	0.775	0.855
π^*	0.650	0.200	1.026	0.807	1.264
γ	0.500	0.250	0.185	0.085	0.276
χ	0.100	0.100	0.100	0.100	0.100
λ_f	0.300	0.500	0.300	0.300	0.300
g^*	0.150	0.050	0.224	0.199	0.253
ρ_z	0.200	0.050	0.218	0.146	0.294
ρ_ϕ	0.800	0.050	0.705	0.625	0.791
ρ_{λ_f}	0.800	0.050	0.518	0.449	0.589
ρ_μ	0.800	0.050	0.884	0.834	0.937
ρ_b	0.800	0.050	0.811	0.743	0.876
ρ_g	0.800	0.050	0.951	0.928	0.975
σ_z	0.400	2.000	0.702	0.625	0.779
σ_ϕ	1.000	2.000	3.450	1.990	4.886
σ_χ	1.000	2.000	0.000	0.000	0.000
σ_{λ_f}	1.000	2.000	0.192	0.168	0.217
σ_μ	1.000	2.000	0.918	0.742	1.077
σ_b	0.200	2.000	0.538	0.439	0.630
σ_g	0.300	2.000	0.406	0.360	0.454
σ_r	0.200	2.000	0.271	0.242	0.300

Notes: See Section 2 for a definition of the DSGE model's parameters, and Section 4 for a description of the data. Results are based on the sample period QII:1974 - QI:2004.

Table 3: Log-Marginal Likelihood for DSGE-VAR(λ)/VECM(λ) and for the DSGE Model.

Weight of the DSGE model prior		DSGE-VECM		DSGE-VAR	
(1)	(2)	(3)	(4)	(5)	(6)
λ	$\frac{\text{Artificial}}{\text{Total}}$	Log-Marginal Likelihood	(Difference wrt Best Model)	Log-Marginal Likelihood	(Difference wrt Best Model)
0.5	0.33	-456.41	(-13.03)	-436.07	(-7.29)
0.75	0.43	-443.45	(-0.07)	-428.78	(0)
1	0.50	-443.38	(0)	-434.89	(-6.11)
1.25	0.56	-450.50	(-7.12)	-437.62	(-8.84)
1.5	0.60	-454.14	(-10.76)	-444.20	(-15.42)
2	0.67	-461.13	(-17.75)	-451.23	(-22.45)
5	0.83	-488.65	(-45.27)	-481.04	(-52.26)
∞	1.00	-529.00	(-85.62)	-501.56	(-72.78)
DSGE		-530.48	(-87.10)	-530.48	(-101.70)

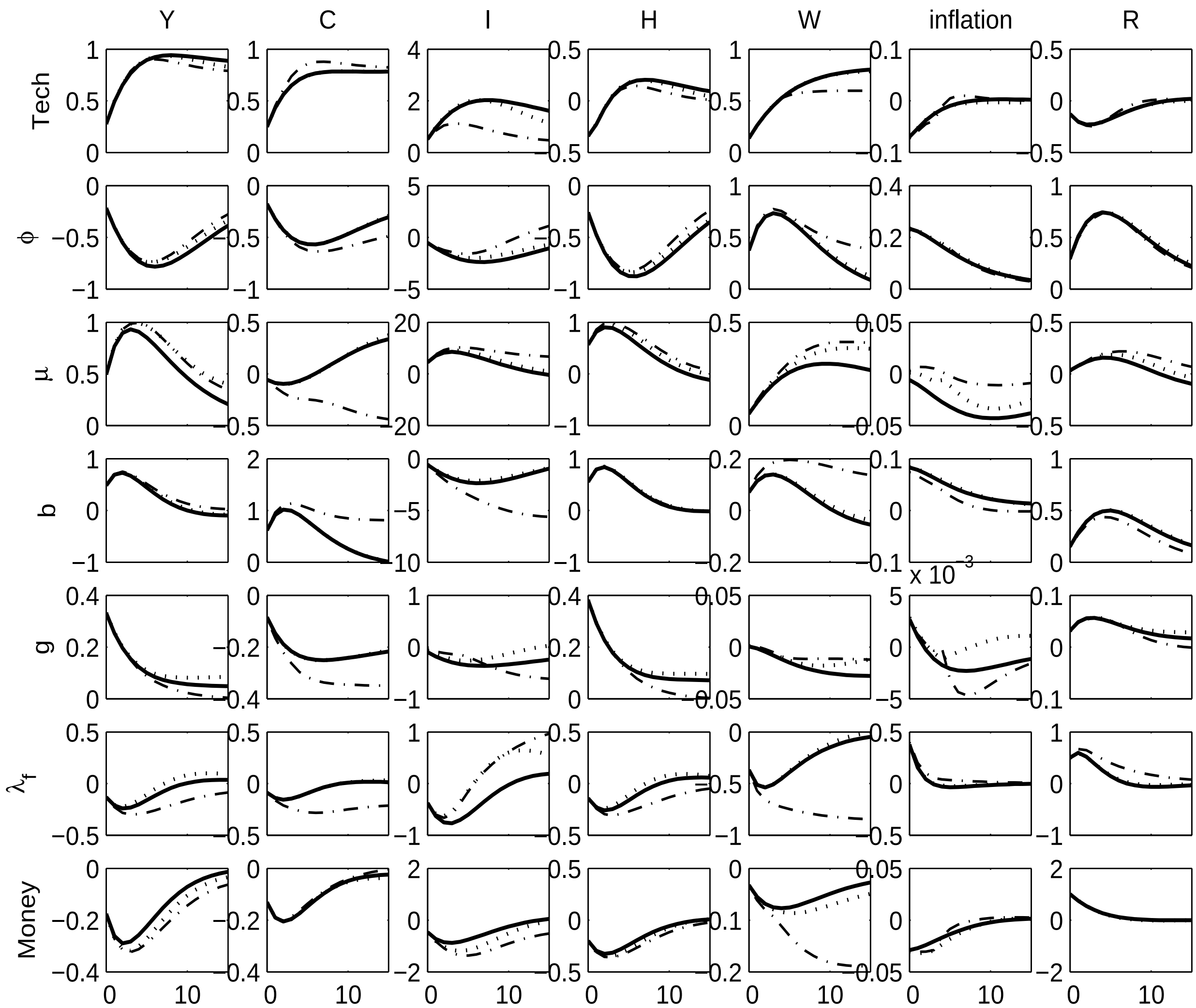
Notes: Column (1) shows the weight of the DSGE model prior λ. Column (2) shows the ratio of artificial over total observations, equal to $\frac{\lambda T}{(1+\lambda)T}$. Columns (3) and (5) show the logarithm of the marginal likelihood for DSGE-VAR(λ) and DSGE-VECM(λ), respectively. Column (4) and (6) shows in parenthesis the difference between the logarithms of the marginal likelihood of DSGE-VAR(λ) and DSGE-VAR($\hat{\lambda}$), and of DSGE-VECM(λ) and DSGE-VECM($\hat{\lambda}$), respectively, where $\hat{\lambda}$ is the value of λ that maximizes the marginal likelihood. See Section 4 for a description of the data. Results are based on the sample period QII:1974 - QI:2004.

Table 4: Out-of-Sample Root Mean Squared Errors: VECM Specification

			forecast horizon					
			1	2	4	6	8	12
Y		DSGE-VECM($\hat{\lambda}$)	0.577	0.909	1.753	2.505	3.141	3.888
	relative to	DSGE	(19.6)	(36.0)	(51.5)	(58.8)	(63.1)	(69.5)
	relative to	VECM	(21.3)	(24.4)	(25.0)	(21.5)	(17.0)	(12.7)
C		DSGE-VECM($\hat{\lambda}$)	0.498	0.767	1.375	1.959	2.450	3.226
	relative to	DSGE	(22.9)	(33.9)	(42.1)	(46.1)	(49.4)	(54.7)
	relative to	VECM	(5.5)	(9.3)	(18.2)	(19.9)	(21.4)	(19.2)
I		DSGE-VECM($\hat{\lambda}$)	3.160	4.955	9.205	13.112	16.520	20.250
	relative to	DSGE	(29.3)	(41.0)	(53.1)	(59.7)	(63.7)	(69.9)
	relative to	VECM	(13.2)	(12.3)	(9.5)	(4.5)	(-2.1)	(-11.8)
H		DSGE-VECM($\hat{\lambda}$)	0.005	0.009	0.019	0.029	0.038	0.050
	relative to	DSGE	(16.3)	(29.9)	(43.7)	(48.5)	(50.2)	(52.5)
	relative to	VECM	(19.5)	(21.1)	(16.4)	(13.6)	(9.4)	(-0.2)
W		DSGE-VECM($\hat{\lambda}$)	0.611	1.022	1.875	2.563	3.162	4.256
	relative to	DSGE	(1.9)	(-0.3)	(-3.9)	(-8.0)	(-12.0)	(-13.2)
	relative to	VECM	(7.5)	(6.4)	(2.5)	(1.5)	(3.6)	(8.8)
Inflation		DSGE-VECM($\hat{\lambda}$)	0.233	0.450	0.833	1.305	1.803	2.820
	relative to	DSGE	(2.4)	(8.6)	(15.5)	(14.1)	(14.0)	(14.5)
	relative to	VECM	(5.2)	(5.6)	(9.0)	(13.4)	(15.3)	(16.5)
R		DSGE-VECM($\hat{\lambda}$)	0.465	0.780	1.288	1.712	2.180	2.596
	relative to	DSGE	(13.1)	(22.4)	(27.4)	(26.5)	(21.2)	(19.4)
	relative to	VECM	(28.3)	(28.8)	(29.2)	(31.4)	(30.1)	(29.1)
Multivariate		DSGE-VECM($\hat{\lambda}$)	1.368	0.939	0.523	0.281	0.117	-0.101
Statistic	relative to	DSGE	(13.8)	(18.2)	(21.7)	(20.7)	(17.8)	(5.0)
	relative to	VECM	(15.0)	(16.3)	(16.0)	(19.7)	(20.4)	(26.5)

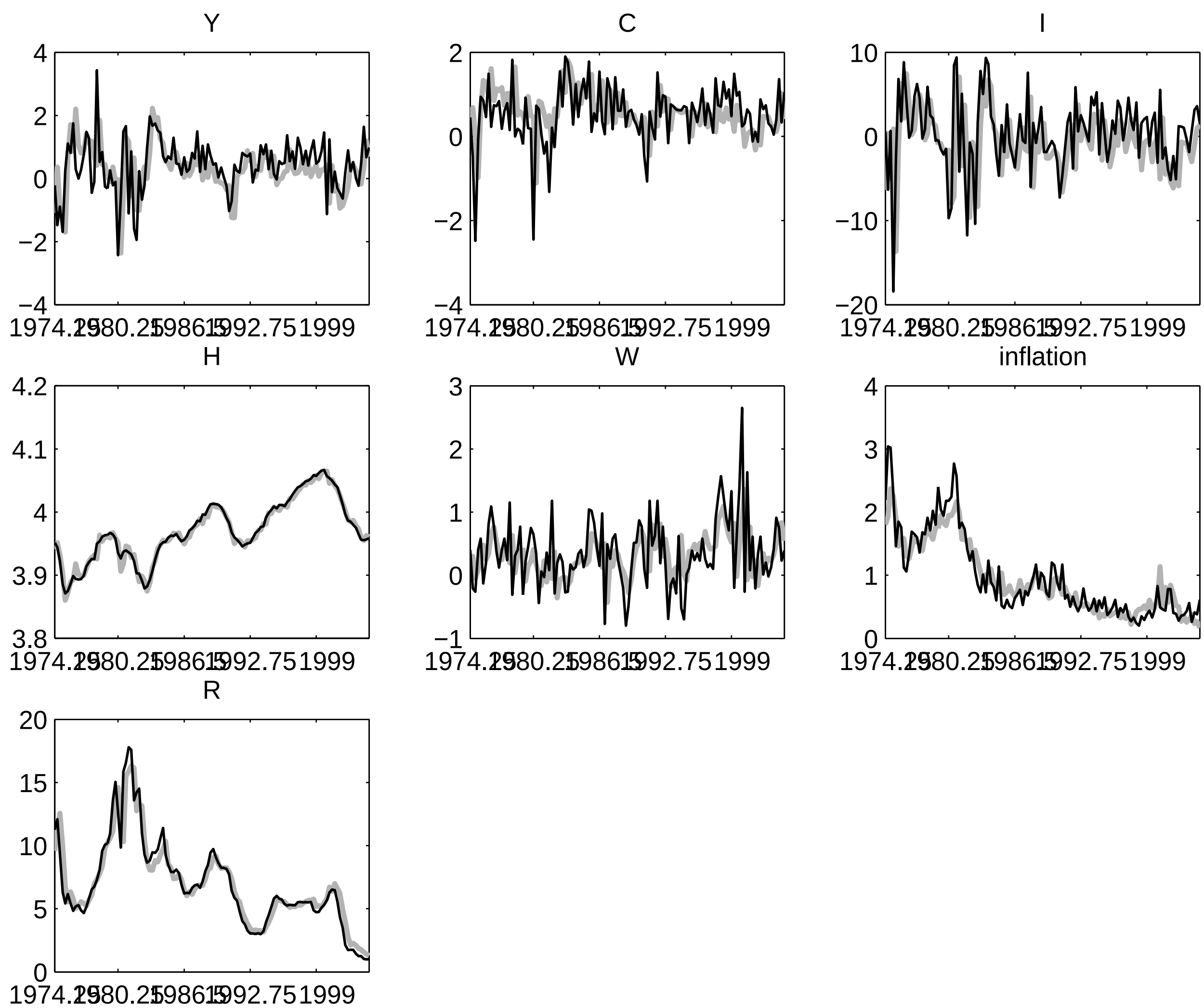
Notes: Results are based on the rolling sample QIV:1985 - QI:2000. At each date of the rolling sample we use the previous 120 observations to estimate the model, and the following twelve quarters to assess forecasting accuracy. For each date we also compute $\hat{\lambda}$, the value of λ that maximizes the marginal likelihood. For each variable, the table shows the root mean squared error (RMSE) of the forecast from DSGE-VECM($\hat{\lambda}$), and in parenthesis the improvement in forecast accuracy relative to the DSGE model and the unrestricted VECM, as measured by the percentage reduction (increase, if negative) in RMSE. The multivariate statistic is computed as the converse of the log-determinant of the variance-covariance matrix of forecast errors. The forecast horizon is measured in quarters. See Section 4 for a description of the data.

Figure 1: How Well Do DSGE-VAR/VECM Approximate the DSGE Model? Identified Impulse Responses



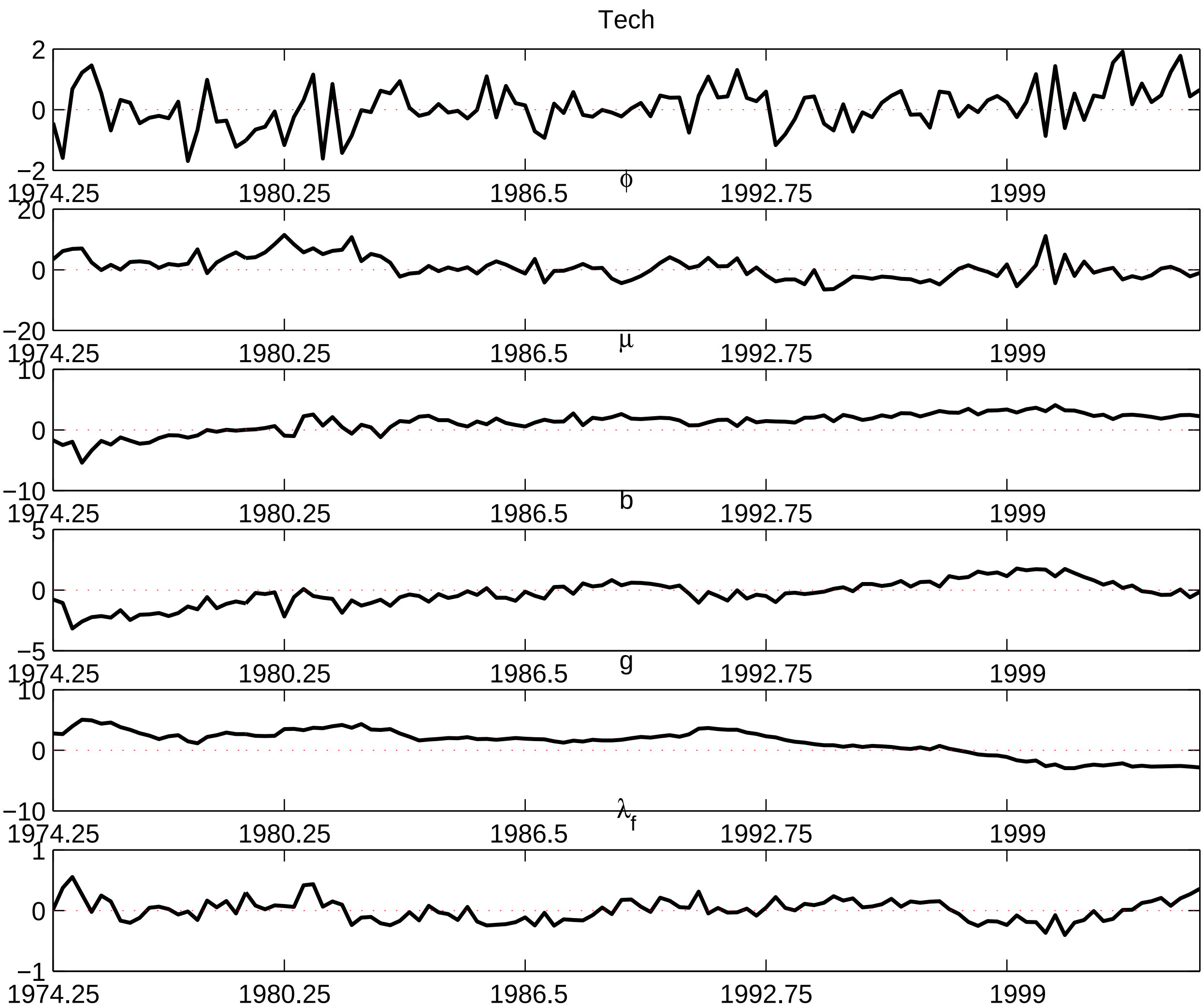
Notes: The solid, dash-and-dotted, and dotted lines represent the impulse responses from one to sixteen quarters ahead of the DSGE model, DSGE-VAR(∞), and DSGE-VECM(∞), respectively, with respect to the following shocks: Tech (technology), φ_t (preferences of leisure), μ_t (shocks to the capital accumulation equation), b_t (overall preference shifter), g_t (government spending), and $\lambda_{f,t}$ (price markup shocks), and Money (monetary policy). All impulse responses are computed using the vector of DSGE model parameters θ that maximizes the posterior of the DGSE model. These impulse responses for DSGE-VAR(∞) and DSGE-VECM(∞) are obtained using the identification procedure described in the section 3.5. The impulse responses to output, consumption, investment, and the real wage are cumulative. Results are based on the sample period QII:1974 - QI:2004. See Section 4 for a description of the data.

Figure 2: In-Sample Fit of the DSGE Model



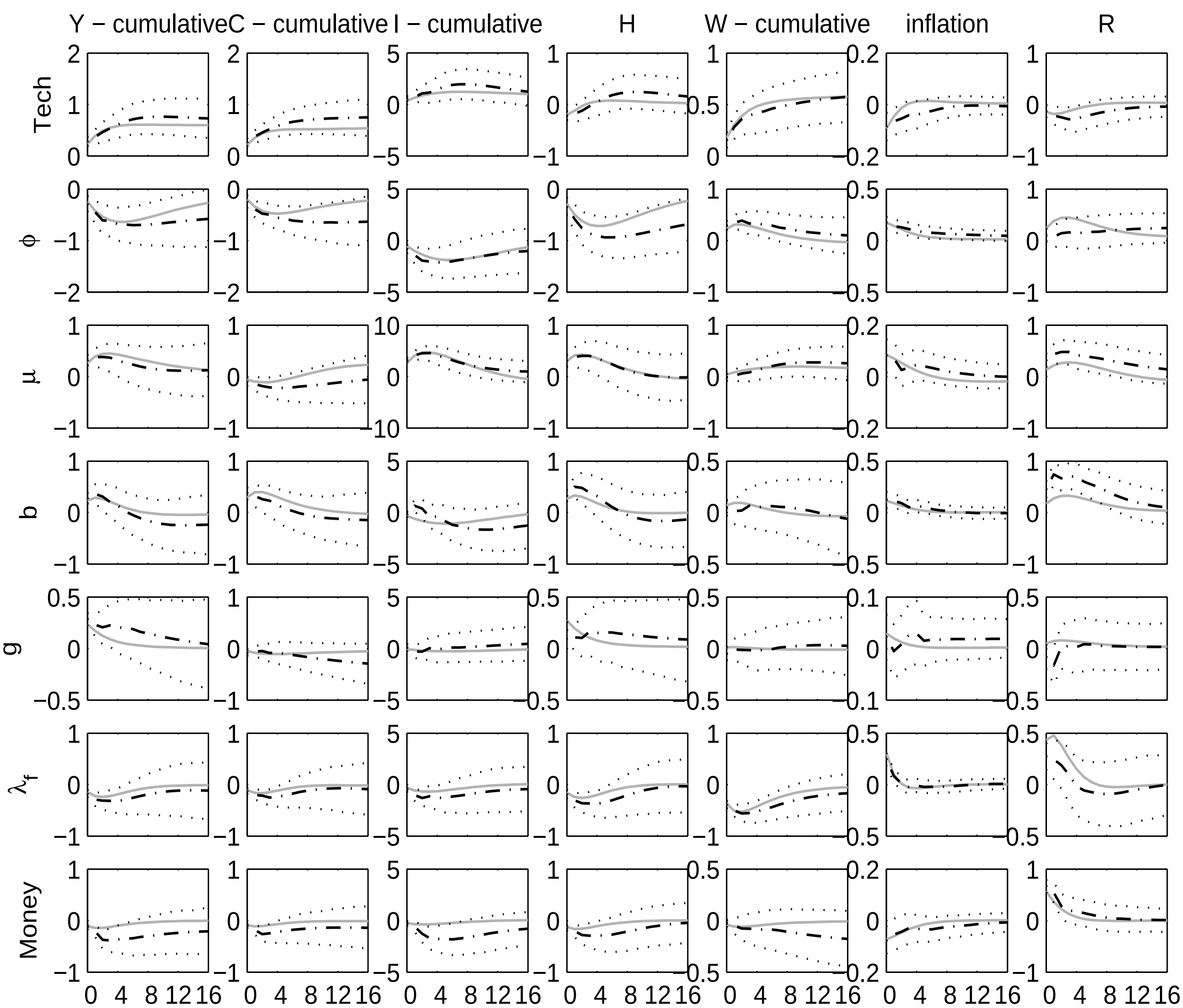
Notes: The figure plots the actual data (dark line), as well as the one-period-ahead forecasts obtained from the Kalman filter (gray line), computed using the vector θ of DGSE model parameters that maximizes the posterior. Results are based on the sample period QII:1974 - QI:2004. See Section 4 for a description of the data.

Figure 3: Exogenous Processes



Notes: The figure plots the Kalman-smoothed time series for the processes z_t (technology), φ_t (preferences of leisure), μ_t (shocks to the capital accumulation equation), b_t (overall preference shifter), g_t (government spending), and $\lambda_{f,t}$ (price markup shocks), computed using the vector θ of DGSE model parameters that maximizes the posterior. Results are based on the sample period QII:1974 - QI:2004. See Section 4 for a description of the data.

Figure 4: Impulse-Responses: DSGE-VECM($\hat{\lambda}$) versus the DSGE model



Notes: The black lines represent the mean impulse-responses (dash-and-dotted lines) of DSGE-VECM($\lambda = 1$) and the associated 90% bands (dotted lines). The gray lines represent the mean impulse-responses (solid lines) of the DSGE model and the associated 90% bands (dotted lines). The impulse-responses are computed with respect to the following shocks: Tech (technology), φ_t (preferences of leisure), μ_t (shocks to the capital accumulation equation), b_t (overall preference shifter), g_t (government spending), and $\lambda_{f,t}$ (price markup shocks), and Money (monetary policy). Results are based on the sample period QII:1974 - QI:2004. See Section 4 for a description of the data.