



EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1110 / NOVEMBER 2009

**PUTTING THE
NEW KEYNESIAN
DSGE MODEL TO THE
REAL-TIME
FORECASTING TEST**

by Marcin Kolasa,
Michał Rubaszek
and Paweł Skrzypczyński



EUROPEAN CENTRAL BANK

EUROSYSTEM



WORKING PAPER SERIES

NO 1110 / NOVEMBER 2009

PUTTING THE NEW KEYNESIAN DSGE MODEL TO THE REAL-TIME FORECASTING TEST¹

by Marcin Kolasa²,
Michał Rubaszek² and Paweł Skrzypczyński²



In 2009 all ECB publications feature a motif taken from the €200 banknote.

This paper can be downloaded without charge from <http://www.ecb.europa.eu> or from the Social Science Research Network electronic library at http://ssrn.com/abstract_id=1499084.

¹ This paper was initiated when Marcin Kolasa and Michał Rubaszek were staying at the ECB. Its earlier versions were presented at the 64th European Meeting of the Econometric Society in Barcelona (August 2009); Macroeconomic Forecasting Conference held by IFO, INSEE and ISAE in Rome (March 2009); the Economic Seminar in the Warsaw School of Economics (February 2009).

We are grateful to participants at these meetings, Michał Brzoza-Brzezina, Andrzej Kocięcki, Frank Schorfheide and Tara Sinclair for helpful comments. The opinions expressed herein are those of the authors and do not necessarily reflect those of the European Central Bank, the National Bank of Poland or the Eurosystem.

² National Bank of Poland, ul. Świętokrzyska 11/21, PL-00-919 Warsaw, Poland (and Warsaw School of Economics); e-E-mail: marcin.kolasa@nbp.pl, michal.rubaszek@nbp.pl and pawel.skrzypczynski@nbp.pl

© European Central Bank, 2009

Address

Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address

Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone

+49 69 1344 0

Website

<http://www.ecb.europa.eu>

Fax

+49 69 1344 6000

All rights reserved.

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the author(s).

The views expressed in this paper do not necessarily reflect those of the European Central Bank.

The statement of purpose for the ECB Working Paper Series is available from the ECB website, <http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html>

ISSN 1725-2806 (online)

CONTENTS

Abstract	4
Non-technical summary	5
1 Introduction	7
2 The DSGE model	9
2.1 Final good producers	9
2.2 Intermediate goods producers	9
2.3 Households	10
2.4 Labor market	11
2.5 Closing the model	12
2.6 Solution and estimation	13
3 BVAR and DSGE-VAR models	13
3.1 Prior specification for BVAR models	14
3.2 Prior specification for DSGE-VAR models	15
4 The Survey of Professional Forecasters	16
5 The data	17
6 Recursive estimates of DSGE and DSGE-VAR parameters	18
7 Forecasts comparison	22
7.1 Mean forecast errors	22
7.2 Root mean squared forecast errors	23
7.3 Root mean squared forecast errors conditional on SPF nowcasts	25
8 Conclusions	26
References	27
European Central Bank Working Paper Series	30

Abstract

Dynamic stochastic general equilibrium models have recently become standard tools for policy-oriented analyses. Nevertheless, their forecasting properties are still barely explored. We fill this gap by comparing the quality of real-time forecasts from a richly-specified DSGE model to those from the Survey of Professional Forecasters, Bayesian VARs and VARs using priors from a DSGE model. We show that the analyzed DSGE model is relatively successful in forecasting the US economy in the period of 1994-2008. Except for short-term forecasts of inflation and interest rates, it is as good as or clearly outperforms BVARs and DSGE-VARs. Compared to the SPF, the DSGE model generates better output forecasts at longer horizons, but less accurate short-term forecasts for interest rates. Conditional on experts' nowcasts, however, the forecasting power of the DSGE turns out to be similar or better than that of the SPF for all the variables and horizons.

Keywords: Forecasting, DSGE, Bayesian VAR, SPF, Real-time data.

JEL Classification: C11, C32, C53, D58, E17.

Non-technical summary

Accurate forecasts of the future path for macroeconomic series such as real GDP, inflation or interest rates are very important information for the business sector, the government or the central bank in their decision-making processes. The problem arises, however, as the number of possible methods that can be used in the process of forecast formulation is large. Generally, these methods can be classified into judgment- and model-based approaches. The former rely on a particular forecaster's skills at interpreting the current economic situation and its future evolution. The latter use formalized econometric models that extrapolate trends from the past into the future.

This paper evaluates the relative accuracy of real-time forecasts formulated on the basis of estimated models and by experts. The main question we pose is whether a richly-specified New Keynesian dynamic stochastic general equilibrium model (DSGE) is able to forecast the US economy better than the Survey of Professional Forecasters (SPF), which we consider to represent the best available judgment-based forecasts. Moreover, we extend our forecasting contest for standard Bayesian vector autoregressive (BVAR) models, which have been widely used as a natural benchmark for estimated DSGE models, as well as for the relatively new tool for policy-oriented analyses, vector autoregressions using priors from a DSGE model (DSGE-VAR). The discussion on forecasting properties of DSGE models can generally be divided into two parts. The first strand of the literature uses latest-available data to compare the accuracy of forecasts from DSGE models to that from VARs or BVARs. These papers indicate that the forecasting performance of DSGE models can be better than that of VARs, which serves as a claim for an increase in the use of DSGE models in forecasting. We argue that for this statement to be persuasive, DSGE models should also perform well in comparison to judgment-based forecasts.

The second strand of the literature addresses this issue by comparing forecasts from DSGE models to experts' judgment. It should be noted that in this kind of analysis it is necessary to use real-time data to ensure that information available to experts and estimated models is comparable. In this article we add to the second strand of the literature by investigating the real-time forecasting properties of the Smets and Wouters (2007) DSGE model, which can be considered to represent a benchmark specification for most DSGE models that are currently used in central banks.

We show that the overall accuracy of real-time forecasts generated by a richly-specified DSGE model is comparable to judgment-based forecasts formulated by the SPF. While the SPF seems to have a clear advantage in short-term forecasts of interest rates, the DSGE model performs significantly better in

forecasting GDP at medium and long horizons. Moreover, except for short-term forecasts of inflation and interest rates, the DSGE model is found to be as good as or clearly outperform the BVAR and DSGE-VAR models in forecasting the US economy. We also demonstrate that the dominance of experts in forecasting nominal variables found in the earlier literature can be attributed to an information advantage of experts, namely their access to current high-frequency data. Conditional on experts' nowcasts, the RMSFEs from the estimated models turn out to be comparable or even smaller than the RMSFEs of the SPF forecasts.

We believe that the above findings contribute to the current discussion on the usefulness of DSGE models in policy oriented analyses. Del Negro et al. (2007) point at an improved time series fit of DSGE models as an important factor behind their increasing use in policy making institutions. We claim that this direction is correct and that DSGE models should be extensively used in forecasting. Furthermore, we propose a method of forecast formulation that involves combining experts' nowcasts with a DSGE model forecasts for the remaining horizons. We also dissuade from adding expert corrections to forecasts generated using this method.

1 Introduction

Accurate forecasts of the future path for macroeconomic series such as real GDP, inflation or interest rates are very important information for the business sector, the government or the central bank in their decision-making processes. The problem arises, however, as the number of possible methods that can be used in the process of forecast formulation is large. Generally, these methods can be classified into judgment- and model-based approaches. The former rely on a particular forecaster's skills at interpreting the current economic situation and its future evolution. The latter use formalized econometric models that extrapolate trends from the past into the future. Our article evaluates the relative accuracy of real-time forecasts formulated on the basis of estimated models and by experts. The main question we pose is whether a richly-specified New Keynesian dynamic stochastic general equilibrium model (DSGE) is able to forecast the US economy better than the Survey of Professional Forecasters (SPF), which we consider to represent the best available judgment-based forecasts. Moreover, we extend our forecasting contest for standard Bayesian vector autoregressive (BVAR) models, which have been widely used as a natural benchmark for estimated DSGE models, as well as for the relatively new tool for policy-oriented analyses, vector autoregressions using priors from a DSGE model (DSGE-VAR).

The discussion on forecasting properties of DSGE models can generally be divided into two parts. The first strand of the literature uses latest-available data to compare the accuracy of forecasts from DSGE models to that from VARs or BVARs. Smets and Wouters (2007), on the basis of quarterly data for the period 1990:1-2004:4, show that a richly-specified DSGE model is able to outperform VAR and BVAR models in forecasting key macroeconomic variables of the US economy, especially if longer horizons are considered. Del Negro, Schorfheide, Smets, and Wouters (2007) develop the DSGE-VAR version of the Smets and Wouters (2003) model and demonstrate that it is able to forecast the US economy better than unrestricted VARs over the evaluation sample of 1985:4-2000:1. In another path-breaking article, Adolfson, Lindé, and Villani (2007) investigate the performance of an open-economy version of the Smets and Wouters (2003) model in forecasting the euro area economy. Using data for the period 1994:1-2002:4 they observe that the accuracy of forecasts from their DSGE model is comparable or even superior to those from VARs and BVARs, both if point forecasts and the whole forecast distributions are considered. All the above articles indicate that the forecasting performance of DSGE models can be better than that of VARs. Consequently, the authors claim that the use of DSGE models in forecasting should increase. We argue that for this statement to be persuasive, DSGE

models should also perform well in comparison to judgment-based forecasts.

The second strand of the literature addresses this issue by comparing forecasts from DSGE models to those formulated by experts. It should be noted that in this kind of analysis it is necessary to use real-time data to ensure that information available to experts and estimated models is comparable. To the best of our knowledge, there are only three studies comparing the forecasting performance of DSGE models with judgment-based forecasts in a real-time context. Rubaszek and Skrzypczynski (2008) demonstrate that for the period 1994:1-2006:2 a small-scale DSGE model is able to better forecast GDP growth in the US than the SPF, while it performs relatively poorly in explaining the future paths of inflation and interest rates. Edge, Kiley, and Laforde (2009) compare forecasts from a large-scale DSGE model to those of the Federal Reserve staff and find that in the evaluation sample of 1996:3-2002:4 the forecast accuracy of the DSGE model is superior for real sector variables, and inferior for inflation and interest rates. Finally, Lees, Matheson, and Smith (2007) analyze the accuracy of forecasts formulated by the Reserve Bank of New Zealand staff relative to those from a small-scale open economy DSGE model and its DSGE-VAR version. On the basis of the evaluation sample of 1998:4-2003:3 the authors find that the DSGE model is relatively successful in forecasting GDP growth, whereas the RBNZ is doing better in forecasting inflation and interest rates. It should be noted that the precision of forecasts from the DSGE and DSGE-VAR models was found to be comparable. The general picture that emerges from these three articles is that DSGE models perform relatively well in forecasting real sector variables, whereas forecasts for nominal variables are less precise than those formulated by experts.

In this article we add to the second strand of the literature by investigating the real-time forecasting properties of the Smets and Wouters (2007) DSGE model, which can be considered to represent a benchmark specification for most DSGE models that are currently used in central banks. Our contribution is threefold. First, we show that this DSGE model outperforms BVAR and DSGE-VAR models in forecasting key US macroeconomic variables. Second, we confirm the finding from the literature that, compared to judgment-based forecasts, DSGE models are relatively good in forecasting GDP growth and relatively bad in forecasting interest rates. Third, we indicate that this feature is due to information advantage of experts: forecasts of nominal variables from the DSGE model, conditional on nowcasts from the SPF, are comparable or even better than forecasts from the SPF.

The rest of the article is structured as follows. The next three sections present methods applied to generate forecasts: the DSGE, BVAR and DSGE-VAR models, and the SPF. In section 5 we describe

the real-time data used in our analysis. Section 6 focuses on parameter estimates and properties of the DSGE and DSGE-VAR models. Section 7 presents the results of the out-of-sample forecast performance analysis. The last section offers conclusions based on the study's main findings.

2 The DSGE model

The DSGE model proposed by Christiano, Eichenbaum, and Evans (2005) and estimated by Smets and Wouters (2003) using Bayesian techniques, is currently considered to be a benchmark richly-specified DSGE model for a closed economy. In this paper we analyze the forecasting performance of the Smets and Wouters (2007) version of this model, which we modify by removing the wage mark-up shock and the wage measurement equation. The reason for this modification is the lack of real-time series for wages in our database. As the model is well documented in the above-referenced articles, here we only summarize its main features.

2.1 Final good producers

The output of the final good Y_t is a composite made of a continuum of intermediate goods $Y_{i,t}$ given implicitly as in Kimball (1995) by:

$$1 = \int_0^1 \Gamma \left(\frac{Y_{i,t}}{Y_t}; \lambda_p, \varepsilon_t^p, \varepsilon_p \right) di, \quad (1)$$

where Γ is a strictly concave and increasing function that satisfies $\Gamma(1) = 1$. The parameter λ_p represents the steady-state price mark-up, ε_p characterizes the curvature of the demand price elasticity,¹ whereas ε_t^p is the disturbance to the price mark-up following an ARMA process $\ln \varepsilon_t^p = \rho_p \ln \varepsilon_{t-1}^p + \eta_t^p - \theta_p \eta_{t-1}^p$, $\eta_t^p \sim NID(0, \sigma_p^2)$. The final good producers minimize the cost $\int_0^1 P_{i,t} Y_{i,t} di$ of producing Y_t , sold at price P_t , by choosing the amount of intermediate inputs, each priced at $P_{i,t}$, subject to constraint (1).

2.2 Intermediate goods producers

Intermediate good i is produced using capital services $K_{i,t}^s$ and labor $L_{i,t}$ as inputs according to the technology:

¹More precisely, ε_p is the percent change in the elasticity of demand due to a one percent change in the relative price of an intermediate good, evaluated in steady state (see Eichenbaum and Fisher, 2007, for an extensive exposition).

$$Y_{i,t} = \varepsilon_t^a (K_{i,t}^s)^\alpha (\gamma^t L_{i,t})^{1-\alpha} - \Phi_t, \quad (2)$$

where ε_t^a is the productivity disturbance that follows an AR process $\ln \varepsilon_t^a = \rho_p \ln \varepsilon_{t-1}^a + \eta_t^p$, $\eta_t^a \sim NID(0, \sigma_a^2)$. The structural parameter γ represents the gross deterministic rate of labor-augmenting technological progress, and fixed costs in production Φ_t are related to the steady-state price mark-up through the zero-profit condition $\Phi_t = (\lambda_p - 1)\bar{Y}_t$, where \bar{Y}_t denotes steady-state output. The marginal cost of production depends thereby on the nominal wage W_t and rental rate on capital R_t^k :

$$MC_t = \alpha^{-\alpha} (1 - \alpha)^{-(1-\alpha)} \gamma^{-(1-\alpha)t} (\varepsilon_t^a)^{-1} (R_t^k)^\alpha (W_t)^{1-\alpha}. \quad (3)$$

As proposed by Calvo (1983), in each period only a fraction $1 - \xi_p$ of randomly selected firms are allowed to optimize their prices, while the remaining firms adjust them mechanically according to:

$$P_{i,t} = P_{i,t-1} (\pi_{t-1})^{\xi_p} (\bar{\pi})^{1-\xi_p}, \quad (4)$$

where $\pi_t = P_t/P_{t-1}$ and $\bar{\pi}$ are the actual and steady-state gross inflation rates of the final good. Those firms that are allowed to re-optimize set their prices at $\tilde{P}_{j,t}$ to maximize the present value of their expected profits:

$$E_t \sum_{s=0}^{\infty} \xi_p^s \Xi_{t,t+s} \left[\tilde{P}_{i,t} \Lambda_{t,t+s}^p - MC_{t+s} \right] Y_{i,t+s}, \quad (5)$$

subject to demand function implied by (1). Here, $\Xi_{t,t+s}$ is the nominal discount factor determined by the relative marginal utility from consumption and $\Lambda_{t,t+s}^p$ stands for price indexation between periods t and $t+s$, which can be calculated recursively using formula (4).

2.3 Households

The household sector consists of a continuum of infinitely-lived households, indexed by j , that choose consumption $C_{j,t}$, hours worked $L_{j,t}$, nominal one-period bond holdings $B_{j,t}$, investment $I_{j,t}$ and capital utilization $Z_{j,t}$ to maximize the following objective function:

$$E_t \sum_{s=0}^{\infty} \beta^s \left[\frac{(C_{j,t+s} - hC_{j,t+s-1})^{1-\sigma_c}}{1 - \sigma_c} \right] \exp \left(\frac{\sigma_c - 1}{1 + \sigma_l} L_{j,t+s}^{1+\sigma_l} \right), \quad (6)$$

subject to the nominal budget constraint:

$$\frac{B_{j,t}}{R_t \varepsilon_t^b} \leq B_{j,t-1} + W_{j,t} L_{j,t} + R_t^k Z_{j,t} K_{j,t-1} + Div_t - P_t (T_t + C_{j,t} + I_{j,t} + \Psi(Z_{j,t}) K_{j,t-1}) \quad (7)$$

and the capital accumulation equation:

$$K_{j,t} = (1 - \delta) K_{j,t-1} + \varepsilon_t^i \left[1 - S \left(\frac{I_{j,t}}{I_{j,t-1}} \right) \right] I_{j,t}. \quad (8)$$

Here, Div_t denotes dividends received from firms and T_t stands for lump-sum taxes net of transfers. The rate of return on assets held by households is a product of the gross interest rate set by the central bank R_t and the risk premium disturbance ε_t^b that follows an AR process $\ln \varepsilon_t^b = \rho_b \ln \varepsilon_{t-1}^b + \eta_t^b$, $\eta_t^b \sim NID(0, \sigma_b^2)$. The structural parameters h , σ_c , σ_l and δ represent external habit formation, the inverse of the intertemporal elasticity of substitution, the inverse of the Frisch elasticity of labor supply and the capital depreciation rate, respectively.

The accumulation of capital $K_{j,t}$ is subject to adjustment costs given by a function S that satisfies $S(\gamma) = 0$, $S'(\gamma) = 0$ and $S''(\gamma) = \varphi$. It also depends on the investment-specific productivity disturbance following an AR process $\ln \varepsilon_t^i = \rho_i \ln \varepsilon_{t-1}^i + \eta_t^i$, $\eta_t^i \sim NID(0, \sigma_i^2)$. The accumulated capital is subsequently transformed into capital services:

$$K_{j,t}^s = Z_{j,t} K_{j,t-1} \quad (9)$$

that are sold to firms. Finally, households have to pay real costs of capital utilization $\Psi(Z_{j,t}) K_{j,t-1}$, where Ψ is an increasing function that satisfies $\Psi(1) = 0$ and $\frac{\Psi''(1)}{\Psi'(1)} = \frac{\psi}{1-\psi}$.

2.4 Labor market

Labor supplied by individual households is bought by perfectly competitive firms, called labor packers, that combine it into aggregate labor. Similarly to the case of final good producers, the aggregation is given implicitly by the Kimball (1995) formula:

$$1 = \int_0^1 \Gamma \left(\frac{L_{j,t}}{L_t}; \lambda_w, \varepsilon_w \right) dj, \quad (10)$$

where λ_w represents the steady-state wage mark-up and ε_w characterizes the curvature of the labor demand elasticity. The aggregated labor is subsequently sold to intermediate goods producers at price



W_t . The labor packers minimize the cost $\int_0^1 W_{j,t} L_{j,t} dj$ of generating L_t , subject to constraint (10).

Wage setting is subject to nominal rigidities á-la Calvo (1983), which means that in each period only a fraction $1 - \xi_w$ of households are allowed to re-optimize their prices. The remaining households adjust their wages mechanically according to:

$$W_{i,t} = W_{i,t-1} \gamma (\pi_{t-1})^{\iota_w} (\bar{\pi})^{1-\iota_w}. \quad (11)$$

Those households that are allowed to re-optimize set their wages at $\tilde{W}_{j,t}$ to maximize the present value of their expected wage income:

$$E_t \sum_{s=0}^{\infty} \xi_w^s \Xi_{t,t+s} \left[\tilde{W}_{j,t} \Lambda_{t,t+s}^w - W_{t+s} \right] L_{j,t+s}, \quad (12)$$

subject to demand function implied by (10). Here, $\Lambda_{t,t+s}^w$ stands for wage indexation between periods t and $t+s$, which can be calculated recursively using formula (11).

2.5 Closing the model

The central bank follows a generalized Taylor rule by adjusting its interest rate in response to deviations of inflation, output and output growth from their target levels:

$$\frac{R_t}{\bar{R}} = \left(\frac{R_t}{\bar{R}} \right)^\rho \left[\left(\frac{\pi_t}{\bar{\pi}} \right)^{r_\pi} \left(\frac{Y_t}{Y^p} \right)^{r_y} \right]^{1-\rho} \left(\frac{Y_t/Y_{t-1}}{Y_t^p/Y_{t-1}^p} \right)^{r_{\Delta y}} \varepsilon_t^r, \quad (13)$$

where \bar{R} denotes the steady-state nominal interest rate, Y_t^p is the potential output and ε_t^r is the monetary policy shock that follows an AR process $\ln \varepsilon_t^r = \rho_r \ln \varepsilon_{t-1}^r + \eta_t^r$, $\eta_t^r \sim NID(0, \sigma_r^2)$.

Government spending G_t is given by an exogenous process:

$$G_t = g_y Y_t \varepsilon_t^g, \quad (14)$$

where g_y denotes the steady-state share of government spending in output and ε_t^g is the government spending disturbance: $\ln \varepsilon_t^g = \rho_g \ln \varepsilon_{t-1}^g + \eta_t^g + \rho_{g\alpha} \eta_t^\alpha$, $\eta_t^g \sim NID(0, \sigma_g^2)$.

The model is closed by the aggregate resource constraint of the following form:

$$Y_t = C_t + I_t + G_t + \Psi(Z_t) K_{t-1}. \quad (15)$$

2.6 Solution and estimation

The empirical implementation of the DSGE model can be described as follows. First, the model is linearized around its steady-state and written as a linear expectation system. The linearized version is described in detail by Smets and Wouters (2007), with the difference that we do not include the wage mark-up disturbance. For given parameter values, such a system can be solved out using standard techniques and transformed into a state-space representation, where the measurement equations relate the model variables to macroeconomic data. Having this representation, the likelihood of the model can be obtained with the Kalman filter.

The structural parameters of the DSGE model are estimated by applying Bayesian techniques. Our assumptions for the priors and five calibrated parameters are identical to those used by Smets and Wouters. For each sample, the posterior mode and the corresponding Hessian matrix are calculated using standard numerical optimization routines. The posterior distribution is approximated using the Metropolis-Hastings algorithm with 25,000 replications, out of which we drop the first 5000.

3 BVAR and DSGE-VAR models

It is well known that a standard DSGE model has a restricted infinite-order VAR representation.² Therefore, VARs have been widely used in the literature as unconstrained benchmarks for evaluating DSGE models. However, because of the large number of parameters and short time series, estimates of unrestricted VAR coefficients are in many cases imprecise and forecasts have large standard errors. As it is common in the literature, we tackle this problem by using a Bayesian approach. We consider two types of priors on VAR coefficients, one atheoretical, the other based on the DSGE model described in the previous section. Henceforth, we will refer to the former case as BVAR and to the latter as DSGE-VAR.

We analyze VAR models:

$$z_t = A_0 + \sum_{i=1}^p A_i z_{t-i} + u_t, \quad (16)$$

where z_t is an n -dimensional vector of observed variables, A_i are matrices of model coefficients, $u_t \sim NID(0, \Sigma_u)$ is the error term, and p denotes the maximum lag order. The model (16) can be expressed

²See Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007) for sufficient conditions regarding the state-space representation of a DSGE model.

in the matrix form as:

$$Z = XA + U, \quad (17)$$

where Z is the $T \times n$ matrix with rows z'_t , X is the $T \times (np+1)$ matrix with rows $x_t = [1, z'_{t-1}, \dots, z'_{t-p}]$, U is the $T \times n$ matrix with rows u'_t , $A = [A_0, A_1, \dots, A_p]'$ and T is the sample size. The likelihood function, conditional on observations x_{1-p}, \dots, x_0 , can be expressed as:

$$f(Z|A, \Sigma_u) \propto |\Sigma_u|^{-T/2} \exp \left\{ -\frac{1}{2} \text{tr} [\Sigma_u^{-1} (Z - XA)'(Z - XA)] \right\}. \quad (18)$$

Below we present the two variants of our prior specification and the resulting form of the posterior distribution for the VAR coefficients A and Σ_u .

3.1 Prior specification for BVAR models

In the case of BVAR models we introduce prior information in line with the method proposed by Litterman (1986) and extended by Sims and Zha (1998). The prior distribution of the VAR parameters consists of three components. The first one is Jeffrey's improper prior. The second component can be described as the likelihood of the form (18) of the VAR model estimated on the basis of T_1 dummy observations Z_1 and X_1 , which are constructed to retain certain features governed by a set of hyperparameters ϑ .³ As in Adolfson, Lindé, and Villani (2007), we depart from the standard random walk prior, centering instead our prior means on the first own lag to zero for real variables expressed in growth rates (output, consumption and investment), and to 0.9 for the remaining variables (hours, inflation and the interest rate). Following Sims (2003), we assume the following values of ϑ . The "overall tightness" ϑ_1 is set to 0.3, the "tightness of the prior on Σ_u " ϑ_2 and the "lag decay" ϑ_3 are set to 1, while the "single stochastic trend" ϑ_4 is set to 5. Since the variables in our dataset (see section 5) are usually considered to be stationary, we do not include the "sum of coefficients" prior. The total number of dummy observations is thereby equal to $T_1 = (p + \vartheta_2)n + 1$. The last component of the prior is equal to the likelihood of the form (18) of the VAR model estimated on the basis of T_2 observations Z_2 and X_2 from the training sample, where we set T_2 to 40 initial quarters.

The resulting conjugate prior is of the Inverse Wishart-Normal form:

³Robertson and Tallman (1999) discuss in detail how to construct the dummy observations.

$$\begin{aligned}\Sigma_u &\sim IW_n\left(\tilde{\Sigma}_u; T_1 + T_2 - (np + 1)\right) \\ A|\Sigma_u &\sim N\left(\tilde{A}; \Sigma_u \otimes \left(\tilde{X}'\tilde{X}\right)^{-1}\right),\end{aligned}\tag{19}$$

where $\tilde{\Sigma}_u$ and \tilde{A} are the OLS estimates of the regression of $\tilde{Z}' = [Z'_1, Z'_2]$ on $\tilde{X}' = [X'_1, X'_2]$.

The posterior distribution coincides with the likelihood function (18) for the VAR model estimated using the actual sample, the training sample and dummy observations. This means that the posterior distribution is

$$\begin{aligned}\Sigma_u|Y, \vartheta &\sim IW_n\left(\hat{\Sigma}_u; T + T_1 + T_2 - (np + 1)\right) \\ A|Y, \Sigma_u, \vartheta &\sim N\left(\hat{A}; \Sigma_u \otimes \left(\bar{X}'\bar{X}\right)^{-1}\right),\end{aligned}\tag{20}$$

where $\bar{Z}' = [Z', \tilde{Z}']$, $\bar{X}' = [X', \tilde{X}']$, $\hat{A} = (\bar{X}'\bar{X})^{-1}(\bar{X}'\bar{Z})$ and $\hat{\Sigma}_u = (\bar{Z} - \bar{X}\hat{A})'(\bar{Z} - \bar{X}\hat{A})$.

3.2 Prior specification for DSGE-VAR models

In this subsection we consider priors for a VAR model that are derived from a DSGE model. We apply the method proposed by Del Negro and Schorfheide (2004), which can be characterized as adding λT artificial observations simulated from the DSGE model to the actual data and estimating the VAR model on the basis of a mixed sample of the artificial and actual observations. The hyperparameter λ denotes the prior tightness so that for $\lambda = 0$ the DSGE-VAR model corresponds to the unrestricted VAR and for $\lambda = \infty$ the DSGE-VAR model becomes the VAR representation of the DSGE model.

A short description of the Del Negro and Schorfheide procedure is as follows. Given the parameters of the DSGE model θ and its state-space representation, it is possible to compute the expected values of artificial data sample moments: $\Gamma_{zz}^* = E_\theta(z_t z_t')$, $\Gamma_{zx}^* = E_\theta(z_t x_t')$, $\Gamma_{xz}^* = E_\theta(x_t z_t')$ and $\Gamma_{xx}^* = E_\theta(x_t x_t')$. The conjugate prior of the VAR coefficients Σ_u and A conditional on θ is of the Inverse Wishart-Normal form:

$$\begin{aligned}\Sigma_u|\theta, \lambda &\sim IW_n\left(\lambda T\left(\Gamma_{zz}^* - \Gamma_{zx}^* \Gamma_{zz}^{*-1} \Gamma_{xz}^*\right); \lambda T - (np + 1)\right) \\ A|\Sigma_u, \theta, \lambda &\sim N\left(\Gamma_{xx}^*{}^{-1} \Gamma_{xz}^*; \Sigma_u \otimes (\lambda T \Gamma_{xx}^*)^{-1}\right).\end{aligned}\tag{21}$$

This means that the posterior distribution of the VAR coefficients is:

$$\begin{aligned}\Sigma_u|Z, \theta, \lambda &\sim IW_n\left((\lambda + 1)T\hat{\Sigma}_u; (\lambda + 1)T - (np + 1)\right) \\ A|Z, \Sigma_u, \theta, \lambda &\sim N\left(\hat{A}; \Sigma_u \otimes (\lambda T \Gamma_{xx}^* + X'X)^{-1}\right),\end{aligned}\tag{22}$$

where $\hat{A} = (\lambda T \Gamma_{xx}^* + X'X)^{-1}(\lambda T \Gamma_{xz}^* + X'Z)$ and $\hat{\Sigma}_u = [(\lambda + 1)T]^{-1}[\lambda T(\Gamma_{zz}^* - \Gamma_{zx}^* \hat{A}) + (Z'Z - Z'X \hat{A})]$. One can notice that the expected value of the posterior distribution of the VAR parameters A is a weighted average of the estimates implied by the expected moments from the DSGE model and the unrestricted OLS estimates, where the weight is determined by the hyperparameter λ .

As in Del Negro and Schorfheide, the prior assumptions given by (21) are complemented with a prior distribution of DSGE model parameters. Following Adjemian, Paries, and Moyen (2008), we also define a prior distribution for the hyperparameter λ , which is assumed to be uniform over the interval $[0,10]$. The VAR coefficients and the parameters related to the DSGE model, including λ , are estimated jointly as the posterior distribution is factorized into the posterior density of the former given the latter and the marginal posterior density of the latter. As in the case of the DSGE model, the posterior distribution of DSGE-VAR parameters is obtained numerically using standard optimization routines and the Metropolis-Hastings algorithm with 25,000 replications, out of which we drop the first 5000.

4 The Survey of Professional Forecasters

The SPF is the oldest quarterly survey of macroeconomic forecasts in the United States. The survey, which was launched and elaborated by the American Statistical Association and the National Bureau of Economic Research in 1968, was taken over by the Federal Reserve Bank of Philadelphia in 1990.⁴ It is carried out at regular three-month intervals and concerns dozens of macroeconomic variables, among

⁴The results of the survey are published quarterly on the Philadelphia Fed website: <http://www.phil.frb.org/econ/spf/index.html>.

them real GDP, the GDP price index and the three-month Treasury bill (TB) rate. In the further part of this paper we focus on the median forecasts of the above-listed variables, which are forecasted by the SPF up to five quarters ahead. It should be noted that the one-step forecasts concern the period when the survey is carried out.

As discussed in more detail by Croushore (2006), the survey's forms are sent at the end of the first month of each quarter, just after the advance release of the national account data for the previous period. The respondents return them in the middle of the next month, i.e. before the data are revised. Nevertheless, the forecasters may use some additional information while formulating their predictions for the US economy, in particular if they monitor leading indicators, business surveys or developments in financial markets. Bearing that in mind, it seems obvious that the SPF has an advantage in forecasting output, prices and especially interest rates in comparison to the estimated models described above, particularly in the one-quarter-ahead horizon. We will address this issue in the second part of our forecasting accuracy investigation.⁵ On the other hand, as pointed out by Edge, Kiley, and Laforde (2009), the DSGE and DSGE-VAR models have an advantage over the SPF in retrospective forecasting of the US economy as these models benefit from the research on what types of models are well fitted to the data. For example, neither the structure of the Smets and Wouters (2007) model nor the priors used in Bayesian estimation were available two decades ago, i.e. in time of forecast formulation by the SPF. Unfortunately, it seems impossible to control our results for this kind of potential biases.

5 The data

The DSGE, BVAR and DSGE-VAR models were estimated on the basis of six key US macroeconomic variables: real GDP, real consumption, real investment, the GDP price index (all expressed as the log difference), log hours worked, and the three-month TB rate. Since the use of the latest-available data in the estimation would give an advantage to the estimated models over the SPF in ex-post forecasts comparisons due to data revisions, we applied the Philadelphia Fed "Real-Time Data Set for Macroeconomists", which is described in more detail by Croushore and Stark (2001). This ensures the comparability of the forecasting errors, as all predictions are formulated on the basis of a similar data set.

⁵Yet another advantage of the SPF, implied by the forecast averaging literature, is that the median forecaster is not the same for each forecasting round, variable and horizon.

The out-of-sample forecast performance is analyzed for horizons ranging from one up to five quarters ahead, whereas the evaluation is based on the data from the period 1994:1-2008:4, called henceforth the evaluation sample. The DSGE, BVAR and DSGE-VAR models were estimated on the set of the recursive samples starting in 1964:2 and ending one quarter before a given vintage date, which is the period of forecast formulation. For instance, the forecasts elaborated in 1994:1 for the period 1994:1-1995:1 were generated using the models estimated on the basis of observations from 1964:2 to 1993:4, using the data available in 1994:1. This procedure is repeated for each quarter from the period 1994:1-2007:4, which gives 56 forecasts for each forecast horizon, model and variable.

6 Recursive estimates of DSGE and DSGE-VAR parameters

Before we move to evaluation of forecasts generated by our competing models, we first briefly discuss the estimation results obtained for the DSGE model. The distribution characteristics of the 1964:2-1993:4 to 1964:2-2008:4 recursive estimates of the posterior median are reported in the right-side columns of Table 1. For convenience, in the left-side columns we also report our assumptions for the priors and five calibrated parameters. They are identical to those used by Smets and Wouters. The only additional parameter we use is the trend population growth rate, which we calibrate at 1.5% per annum and include in the measurement equations for output, consumption and investment.⁶

Comparing our recursive estimates with those obtained by Smets and Wouters we note the following major differences. Averaged across all samples, our results point at lower costs of adjusting investment φ and capacity utilization ψ , less persistent habits h , higher labour supply elasticity σ_L , a higher steady-state price mark-up λ_p and a lower trend growth rate γ . There are also some differences in the characteristics of the shock processes.⁷ The average medians for the remaining parameters fall within the 90% confidence interval obtained by Smets and Wouters. A closer inspection reveals that these discrepancies are almost entirely due to data and sample differences. As we find out by experimenting with the original dataset used by Smets and Wouters, dropping the wage mark-up shock from the model and real wages from the set of observable variables leads only to significantly higher estimates the consumption elasticity σ_c and the price mark-up shock inertia ρ_p .

⁶The reason for this correction is that the real-time series we use are not expressed in per capita terms as in the Smets and Wouters paper. Our results are robust to alternative (reasonable) calibrations of this parameter.

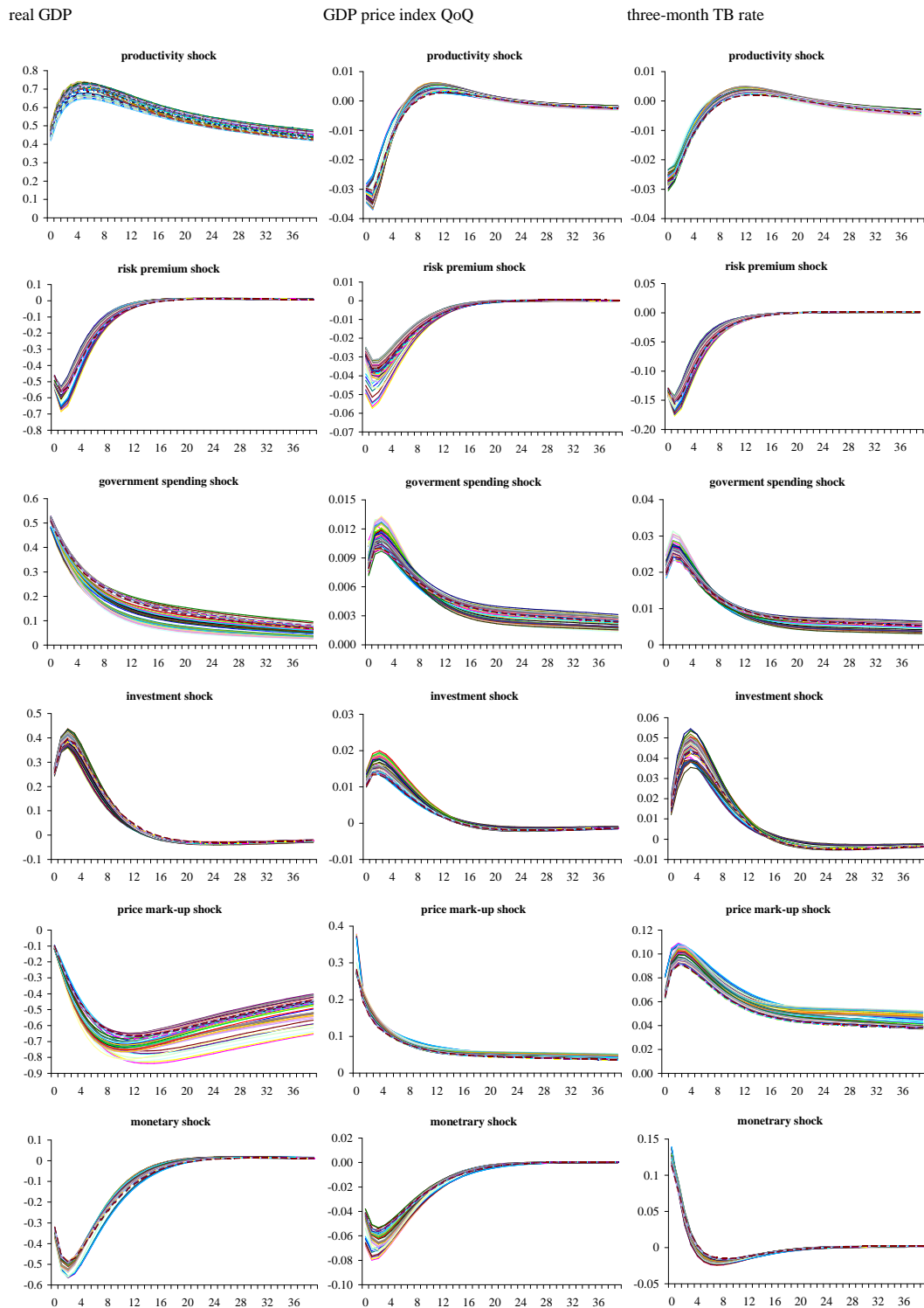
⁷The results reported in Table 1 point at an apparently high estimate of the risk premium volatility. However, as noted by Taylor and Wieland (2009), the risk premium volatility estimated by Smets and Wouters (2007) is actually multiplied with the interest elasticity of consumption.

Table 1: Prior distribution and recursive estimates for model parameters

		Prior distribution			Recursive estimates of posterior median		
		Type	Mean	StDev.	min	av.	max
Estimated parameters							
investment adj. cost	φ	normal	4.00	1.50	2.19	2.70	3.21
consumption elasticity	σ_c	normal	1.50	0.37	1.26	1.46	1.59
consumption habit	h	beta	0.70	0.10	0.45	0.51	0.60
Calvo wages	ξ_w	beta	0.50	0.10	0.56	0.61	0.64
labor supply elasticity	σ_l	normal	2.00	0.75	0.40	0.55	0.72
Calvo prices	ξ_p	beta	0.50	0.10	0.54	0.59	0.66
indexation wages	ι_w	beta	0.50	0.15	0.47	0.51	0.54
indexation prices	ι_p	beta	0.50	0.15	0.19	0.33	0.42
capital util. adj. cost	ψ	beta	0.50	0.15	0.20	0.30	0.35
price mark-up	λ_p	normal	1.25	0.12	1.77	1.84	1.90
int. rate inflation	r_π	normal	1.50	0.25	1.76	1.87	1.97
int. rate smoothing	ρ	beta	0.75	0.10	0.80	0.82	0.84
int. rate output	r_y	normal	0.12	0.05	0.09	0.10	0.12
int. rate output growth	$r_{\Delta y}$	normal	0.12	0.05	0.21	0.23	0.24
steady-state inflation	$\bar{\pi}$	gamma	0.62	0.10	0.63	0.65	0.67
discount factor	$100(\beta^{-1} - 1)$	gamma	0.25	0.10	0.12	0.15	0.19
steady-state hours	\bar{l}	normal	0.00	2.00	1.08	1.80	2.39
trend output p.c. growth	$100(\gamma - 1)$	normal	0.40	0.10	0.26	0.30	0.34
capital share	α	normal	0.30	0.05	0.15	0.16	0.17
Calibrated parameters							
depreciation rate	δ		0.025	0	0.025	0.025	0.025
gov. spending share	g_y		0.18	0	0.18	0.18	0.18
wage mark-up	λ_w		1.50	0	1.50	1.50	1.50
price elasticity curv.	ε_p		10	0	10	10	10
wage elasticity curv.	ε_w		10	0	10	10	10
trend population growth	$100n$		0.37	0	0.37	0.37	0.37
Shock processes							
productivity	ρ_a	beta	0.50	0.20	0.99	0.99	1.00
	σ_a	inv. gam.	0.10	2.00	0.37	0.39	0.41
risk premium	ρ_b	beta	0.50	0.20	0.43	0.56	0.66
	σ_b	inv. gam.	0.10	2.00	0.60	0.79	1.12
investment	ρ_i	beta	0.50	0.20	0.53	0.59	0.62
	σ_i	inv. gam.	0.10	2.00	0.56	0.59	0.62
price mark-up	ρ_p	beta	0.50	0.20	0.97	0.98	0.99
	μ_p	beta	0.50	0.20	0.79	0.84	0.89
	σ_p	inv. gam.	0.10	2.00	0.12	0.14	0.20
monetary	ρ_r	beta	0.50	0.20	0.14	0.23	0.26
	σ_r	inv. gam.	0.10	2.00	0.21	0.23	0.25
gov. spending	ρ_g	beta	0.50	0.20	0.92	0.95	0.96
	σ_g	inv. gam.	0.10	2.00	0.50	0.52	0.54
	ρ_{ga}	beta	0.50	0.20	0.60	0.68	0.76

Notes: For the inverse gamma distribution, the mode and the degrees of freedom are reported.

Figure 1: Recursive impulse response functions



Notes: Recursive IRFs are calculated at the posterior median of parameters for each quarter of the evaluation sample. Each line corresponds to a different quarter of the evaluation sample.

We proceed by evaluating the model's stability and the speed of reversion to the steady-state by looking at the recursive impulse response functions (IRF) to the structural shocks. An informal analysis of Figure 1 shows that the model is relatively stable over the evaluation sample. Importantly, despite the above mentioned differences in the estimates of some of the parameters, the impulse responses turn out to be very similar to those reported in Smets and Wouters.

As regards the impact of individual shocks on the economy, the results are as follows. The productivity shock η_t^a increases output, while also diminishing marginal costs across firms. A respective fall in inflation dominates the output growth effects on the TB rate. It should be noted that the impact of the productivity shock on output is very persistent. The risk premium shock η_t^b leads to a decline in consumption and investment demand, and therefore lowers output. A subsequent fall in inflation is also triggering an interest rate decrease. The government spending shock η_t^g results in higher aggregate output, an increase in inflation and an elevated level of the TB interest rate. The investment shock η_t^i lowers the relative price of investment goods, which raises investment demand and output. A consequent rise in inflation requires interest rate hikes. The price mark-up shock η_t^p raises inflation and suppresses output. The net effect is an increase in the interest rate. Finally, the monetary shock η_t^r entails a rise in the TB rate that results in lower output and inflation.

The recursive estimates of the DSGE-VAR hyperparameter λ are presented in Table 2. According to the results for the DSGE-VAR(1) model, the average recursive estimate of λ is relatively low, standing at 0.32 and indicating that the data give 0.24 probability to the VAR(1) representation of the DSGE model and 0.76 probability to the unrestricted VAR(1). This result should not be surprising as the persistence embedded in our benchmark DSGE model can be approximated by a VAR(1) process only to a limited degree. Increasing the maximum lag p of the DSGE-VAR(p) model leads to a rise in the weight allocated to the DSGE representation of the VAR(p) model. In particular, for $p = 4$ the average posterior estimate of λ amounts to 0.91, indicating 0.48 probability of the DSGE model.

Table 2: Recursive estimates of posterior median for the DSGE-VAR weight parameter

	minimum	average	maximum
DSGE-VAR(1)	0.25	0.32	0.47
DSGE-VAR(2)	0.45	0.58	0.72
DSGE-VAR(3)	0.57	0.72	0.88
DSGE-VAR(4)	0.70	0.91	1.15

Notes: The prior distribution was assumed to be uniform on the interval from 0 to 10.

In the last step, we apply the DSGE, BVAR and DSGE-VAR models to forecasting the US economy.

For this purpose, we generate out-of-sample point forecasts, using the posterior median estimates of each model's parameters. We repeat this procedure for each quarter from the evaluation sample. All calculations are performed with the DYNARE package for MATLAB 7. The results are presented in the next section.

7 Forecasts comparison

Good forecast accuracy is one of the key criteria in the process of model evaluation before it is used in practice. There are two main statistics commonly applied in this subject: the mean forecast error (MFE) and the root mean squared forecast error (RMSFE). While calculating the MFEs and RMSFEs for a model estimated with real-time data the question arises: which vintage should be used to calculate forecast errors. In this paper we present only the results with the “actuals” taken from the last vintage of our sample (2009:1), but we found that the results with different “actuals” are broadly the same.⁸

7.1 Mean forecast errors

We begin our forecasting contest by investigating the MFEs for three key US macroeconomic variables: output growth, inflation and the interest rate. The forecast horizon h is set up to five quarters, i.e. the maximum horizon of the SPF forecasts. Since one-step ahead forecasts ($h = 1$) apply to the period of forecasts formulation, we refer to them as nowcasts.

According to the results presented in Table 3, output growth forecasts are unbiased only in the case of the DSGE model and the SPF. The BVAR and DSGE-VAR models tend to significantly overpredict the future path of GDP for all forecast horizons and for all values of the maximum lag p .⁹ As regards inflation forecasts, all methods perform quite well: the MFEs are not significantly different from zero. Finally, most methods overestimate the future level of the interest rate and the bias is increasing with the forecast horizon h . It can be noted that the BVAR models do slightly better in this area. Overall, the results indicate that, as far as the MFEs are concerned, the SPF and the DSGE dominate the other models in forecasting output, whereas all methods perform comparably well in the case of inflation and the interest rate.

⁸Rubaszek and Skrzypczyński (2008) also find that the general conclusions of forecasts comparison for the latest available and one year after estimation “actuals” are similar.

⁹This somewhat striking result becomes only slightly less pronounced if the models are estimated on the rolling rather than recursive samples. Overall, all our major findings hold in qualitative terms if we use this alternative forecasting setup. The detailed tables, corresponding to those included in the text, are available from the authors upon request.

Table 3: Mean Forecast Errors (MFEs)

h	DSGE	SPF	BVAR				DSGE-VAR			
			$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 1$	$p = 2$	$p = 3$	$p = 4$
<i>Output growth (real GDP, QoQ SAAR)</i>										
1	-0.55**	0.38	-0.87***	-0.92***	-0.75**	-0.73**	-0.93***	-0.98***	-0.85***	-0.99***
2	-0.25	0.17	-1.34***	-1.20***	-0.98***	-0.87**	-1.21***	-1.01***	-0.91**	-0.99***
3	-0.05	0.12	-1.48***	-1.39***	-1.21***	-1.02***	-1.32***	-1.03***	-0.95**	-0.94**
4	0.08	0.02	-1.53***	-1.47***	-1.26***	-1.02***	-1.36***	-1.06***	-0.95**	-0.84**
5	0.07	-0.19	-1.64***	-1.62***	-1.41***	-1.16***	-1.47***	-1.17***	-1.02**	-0.87**
<i>Inflation (GDP price index, QoQ SAAR)</i>										
1	0.16	0.12	0.14	0.10	0.10	0.07	0.17	0.15	0.18	0.17
2	0.06	0.09	0.00	0.00	0.00	-0.03	0.05	0.07	0.11	0.09
3	-0.02	0.01	-0.14	-0.13	-0.08	-0.12	-0.07	-0.04	0.03	0.01
4	-0.04	-0.01	-0.23	-0.21	-0.16	-0.18	-0.15	-0.11	-0.04	-0.04
5	-0.11	-0.06	-0.37	-0.34	-0.29	-0.33	-0.28	-0.23	-0.16	-0.18
<i>Interest rate (three-month TB rate, per annum)</i>										
1	-0.12	-0.07***	0.04	0.02	0.07	0.05	-0.01	-0.03	0.00	-0.05
2	-0.25	-0.19**	-0.04	-0.05	0.05	0.01	-0.10	-0.14	-0.07	-0.17
3	-0.39	-0.33*	-0.18	-0.17	-0.05	-0.11	-0.24	-0.30	-0.22	-0.36
4	-0.52*	-0.47*	-0.34	-0.34	-0.18	-0.25	-0.40	-0.47	-0.37	-0.54*
5	-0.65*	-0.65*	-0.54	-0.54	-0.35	-0.42	-0.59	-0.67*	-0.55	-0.74**

Notes: Symbols ***, ** and * indicate the rejection of the null that the MFE is equal to zero at 1%, 5% and 10% significance levels, respectively. A positive value indicates that on average forecasts are below the actual values. The standard errors are calculated in line with the Newey and West (1987) procedure, where the truncation lag of the modified Bartlett kernel is set using the method proposed by Newey and West (1994).

7.2 Root mean squared forecast errors

We continue our contest by comparing the second moments of the forecast errors. Given the main focus of this paper, we report the levels of the RMSFEs only for the DSGE model, while the remaining numbers in Table 4 are expressed as the ratios to the corresponding RMSFE from the DSGE model. Thus, the values above unity indicate that the DSGE model dominates the alternative method in forecasting a given variable at a forecast horizon h . Moreover, we test whether this difference in the RMSFEs is statistically significant using the HLN-DM test proposed by Harvey, Leybourne, and Newbold (1997).

According to our results, the accuracy of output growth forecasts from the DSGE model is significantly higher than that from the remaining methods at most horizons. In comparison to the SPF, the RMSFEs from the DSGE model are about 20 percent lower for three-, four- and five-quarter ahead forecasts. Moreover, the precision of output growth forecasts from the DSGE model is about 20-30 percent higher than that obtained from the BVAR and DSGE-VAR models. As far as inflation forecasts are concerned, the RMSFEs from the DSGE model, the SPF, and the DSGE-VAR models with the maximum lag of 3 or 4 are comparable, with some evidence in favor of the DSGE-VAR(4) model at short horizons. The low order BVAR and DSGE-VAR models perform significantly worse. Finally,

Table 4: Root Mean Squared Forecast Errors (RMSFEs)

h	DSGE	SPF	BVAR				DSGE-VAR			
			$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 1$	$p = 2$	$p = 3$	$p = 4$
<i>Output growth (real GDP, QoQ SAAR)</i>										
1	1.94	0.98	1.08	1.10	1.09	1.10*	1.10*	1.11	1.13*	1.18**
2	1.99	1.07	1.25**	1.21**	1.20**	1.18**	1.21**	1.18*	1.22**	1.24**
3	1.83	1.21*	1.38**	1.34**	1.31*	1.28	1.32**	1.22	1.27	1.29*
4	1.91	1.18**	1.38**	1.34**	1.29**	1.25*	1.32**	1.21*	1.22*	1.21*
5	2.08	1.17**	1.35**	1.33**	1.28**	1.24*	1.30**	1.21*	1.20*	1.17*
<i>Inflation (GDP price index, QoQ SAAR)</i>										
1	0.96	0.91	1.13***	1.07***	1.01	0.98	1.12***	1.05**	0.97	0.95*
2	0.98	0.95	1.21***	1.11***	1.02	0.98	1.16***	1.06**	0.97	0.92*
3	0.86	1.11	1.35***	1.20***	1.10**	1.06	1.24***	1.11**	1.01	0.96
4	1.02	1.05	1.34***	1.22***	1.13***	1.10**	1.22***	1.11***	1.00	0.96
5	1.11	1.02	1.36***	1.21***	1.12***	1.09	1.23***	1.09***	0.99	0.95
<i>Interest rate (three-month TB rate, per annum)</i>										
1	0.43	0.34***	0.90	0.86	0.92	0.92	0.89	0.87*	0.94	0.93
2	0.80	0.64**	0.90	0.90	0.96	0.95	0.90	0.92	1.00	1.00
3	1.11	0.78	0.94	0.93	0.96	0.95	0.95	0.95	1.00	1.00
4	1.34	0.89	0.99	0.98	0.99	0.98	1.00	1.00	1.02	1.03
5	1.53	1.01	1.04	1.02	1.03	1.02	1.05	1.05	1.07	1.09

Notes: For the DSGE model RMSFEs are reported in levels, whereas for the remaining methods they appear as the ratios to the corresponding RMSFE from the DSGE model. Symbols ***, ** and * indicate the rejection of the null of the HLN-DM test, stating that the RMSFE is not significantly different from the corresponding RMSFE from the DSGE model, at 1%, 5% and 10% significance levels, respectively.

interest rate forecasts formulated by the SPF are substantially better than those generated by the estimated models. The SPF dominance is most evident for one- and two-quarter-ahead forecasts.

In general, these results confirm and extend some findings from the earlier literature surveyed in the introduction. In particular, our results suggest that a richly specified DSGE model is able to outperform BVAR models in forecasting the key US macroeconomic variables. Interestingly, however, the RMSFEs from the DSGE model turned out to be at least as low as those from the DSGE-VAR models for most variables and horizons. This finding can be contrasted with Del Negro, Schorfheide, Smets, and Wouters (2007) who conclude that DSGE-VAR models with an optimally chosen weight of DSGE priors (λ) can perform better than DSGE-VAR models with dogmatic DSGE restrictions. Our results suggest that gains from relaxing these restrictions may actually be more than offset by losses related to the fact that a VAR with a small number of lags is usually a poor approximation to an infinite-order VAR representation of a DSGE model.¹⁰ In other words, even though a DSGE-VAR with optimally selected or, as in our case, estimated λ clearly outperforms its variant with λ set to infinity, it may generate worse forecasts than a state-space representation of the underlying DSGE model.¹¹

¹⁰This point is forcefully made, though in a different context, by Chari, Kehoe, and McGrattan (2008).

¹¹Our results, which are not reported in this article, indicate that the forecast errors from the DSGE model are lower

Finally, we confirm the finding from the literature that, compared to judgment-based forecasts, DSGE models are relatively good in forecasting GDP growth and relatively bad in forecasting interest rates, especially at long and short horizons, respectively. In the next subsection we address this issue by showing that the relative success of the SPF in forecasting the interest rate is due to an information advantage.

7.3 Root mean squared forecast errors conditional on SPF nowcasts

We have already mentioned that one-step ahead forecasts are *de facto* nowcasts as they refer to the period of forecast formulation. This means that, at least in the case of one-step-ahead forecasts, the SPF has an advantage as it can use information unavailable to the estimated models. In particular, the SPF can observe monthly data for CPI inflation, industrial production, retail sales or leading indicators, which might help in nowcasting output growth and inflation. The biggest advantage, however, is in the case of the interest rate as the SPF knows its path up to the middle of the quarter in which forecasts are formulated, whereas the models are estimated with the data ending in the previous quarter. Consequently, it should come as no surprise that the RMSFEs from the SPF at the one quarter horizon are the lowest among all investigated methods, and the superiority of the SPF is most evident in the case of the interest rate. Below we address this issue by comparing the accuracy of forecasts from the estimated models that take into account nowcasts formulated by the SPF.¹²

As indicated in Table 5, in this variant of our forecasting contest the RMSFEs from all methods for the one-quarter ahead horizon are assumed to be the same and equal to those from the SPF. The results for the remaining horizons show that the DSGE model significantly outperforms the other methods in forecasting output growth, with the RMSFEs on average 20 percent lower than those obtained from the SPF, the BVAR and DSGE-VAR models. In the case of inflation, the DSGE and DSGE-VAR models with the maximum lag set to 3 and 4 are characterized by the lowest RMSFEs. The SPF and BVAR(4) are insignificantly less accurate, while the BVAR and low-order DSGE-VAR models are found to be the worst. Finally, the RMSFEs for interest rate forecasts formulated by all methods are comparable.

The results discussed above suggest that the superior performance of experts in forecasting nominal

than those from the VAR representation of the DSGE model even if the maximum lag is set to 5 or 6 quarters.

¹²As the SPF does not forecast average and total hours worked, we applied bridge regressions based on the SPF nowcasts for GDP to generate nowcasts for total hours worked and payroll employment (available SPF nowcasts for the latter start in late 2003). Subsequently, we calculated nowcasts for average hours worked as the ratio of the resulting estimates for total hours worked and payroll employment. The results of these regressions are available upon request.

Table 5: RMSFEs from the models extended for SPF nowcasts

h	DSGE	SPF	BVAR				DSGE-VAR			
			$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 1$	$p = 2$	$p = 3$	$p = 4$
<i>Output growth (real GDP, QoQ SAAR)</i>										
1	1.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	1.91	1.11*	1.10*	1.12**	1.12**	1.13**	1.10*	1.12*	1.15**	1.17**
3	1.88	1.17*	1.27**	1.19	1.16	1.15	1.23*	1.14	1.17	1.20
4	1.83	1.23***	1.40**	1.34**	1.30*	1.27*	1.35**	1.23*	1.24*	1.27*
5	2.09	1.17*	1.34**	1.29**	1.23*	1.18*	1.29**	1.19*	1.16	1.13
<i>Inflation (GDP price index, QoQ SAAR)</i>										
1	0.87	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.92	1.01	1.17***	1.11***	1.05*	1.03	1.13**	1.08**	1.00	0.97
3	0.90	1.06	1.30***	1.16**	1.08	1.06	1.22***	1.08**	1.00	0.96
4	1.00	1.08*	1.33***	1.20***	1.12**	1.10	1.23**	1.11***	1.00	0.97
5	1.11	1.02	1.33**	1.19**	1.10*	1.09	1.22**	1.08**	0.97	0.95*
<i>Interest rate (three-month TB rate, per annum)</i>										
1	0.14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.55	0.92	0.94	0.95	0.98	0.97	0.94	0.96	1.01	0.99
3	0.92	0.94	0.94	0.93	0.97	0.96	0.94	0.95	1.01	1.00
4	1.21	0.99	0.97	0.96	0.98	0.97	0.98	0.97	1.01	1.01
5	1.46	1.06	1.01	0.99	1.00	0.98	1.02	1.01	1.03	1.04

Notes: For the DSGE model RMSFEs are reported in levels, whereas for the remaining methods they appear as the ratios to the corresponding RMSFE from the DSGE model. Symbols ***, ** and * indicate the rejection of the null of the HLN-DM test, stating that the RMSFE is not significantly different from the corresponding RMSFE from the DSGE model, at 1%, 5% and 10% significance levels, respectively.

variables found in the earlier literature can be attributed to their information advantage, related to familiarity with high frequency data. Conditional on the SPF nowcasts, the DSGE model is found to outperform the SPF in forecasting output growth and to generate insignificantly different forecasts of inflation and the interest rate. This indicates that including experts' nowcasts in the process of forecast formulation with an estimated structural model can improve the forecast precision. However, our results also show that forecasts for the remaining horizons should not necessarily be corrected by experts. Finally, our findings also suggest that the current tendency in most central banks and other institutions to intensify the use of DSGE models for policy oriented analyses, including forecasting, is justified.

8 Conclusions

In this paper we have shown that the overall accuracy of real-time forecasts generated by a richly-specified DSGE model is comparable to judgment-based forecasts formulated by the SPF. While the SPF seems to have a clear advantage in short-term forecasts of the interest rate, the DSGE model

performs significantly better in forecasting GDP at medium and long horizons. Moreover, except for short-term forecasts of inflation and the interest rate, the DSGE model has been found to be as good as or clearly outperform the BVAR and DSGE-VAR models in forecasting the US economy. We have also demonstrated that the dominance of experts in forecasting nominal variables found in the earlier literature can be attributed to an information advantage of experts, namely their access to current high-frequency data. Conditional on experts' nowcasts, the RMSFEs from the estimated models turned out to be comparable or even smaller than the RMSFEs of the SPF forecasts.

We believe that the above findings contribute to the current discussion on the usefulness of DSGE models in policy oriented analyses. Del Negro, Schorfheide, Smets, and Wouters (2007) point at an improved time series fit of DSGE models as an important factor behind their increasing use in policy making institutions. We claim that this direction is correct and that DSGE models should be extensively used in forecasting. Furthermore, we propose a method of forecast formulation that involves combining experts' nowcasts with a DSGE model forecasts for the remaining horizons. We also dissuade from adding expert corrections to forecasts generated using this method.

Finally, we would like to emphasize that our results might be dependent on the specification of the DSGE model, the selection of the evaluation sample and the parameterization of priors. We did our best to ensure our choices in these fields are in line with the current state-of-art and best practices. However, more studies for other countries or models would be useful to confirm the main findings of this paper.

References

- ADJEMIAN, S., M. D. PARIÉS, AND S. MOYEN (2008): "Towards a monetary policy evaluation framework," Working Paper Series 942, European Central Bank.
- ADOLFSON, M., J. LINDÉ, AND M. VILLANI (2007): "Forecasting Performance of an Open Economy DSGE Model," *Econometric Reviews*, 26(2-4), 289–328.
- CALVO, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12(3), 383–398.
- CHARI, V., P. J. KEHOE, AND E. R. MCGRATTAN (2008): "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Monetary Economics*, 55(8), 1337–1352.

- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113(1), 1–45.
- CROUSHORE, D. (2006): “An evaluation of inflation forecasts from surveys using real-time data,” Working Papers 06-19, Federal Reserve Bank of Philadelphia.
- CROUSHORE, D., AND T. STARK (2001): “A real-time data set for macroeconomists,” *Journal of Econometrics*, 105(1), 111–130.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2004): “Priors from General Equilibrium Models for VARs,” *International Economic Review*, 45(2), 643–673.
- DEL NEGRO, M., F. SCHORFHEIDE, F. SMETS, AND R. WOUTERS (2007): “On the Fit of New Keynesian Models,” *Journal of Business & Economic Statistics*, 25, 123–143.
- EDGE, R. M., M. T. KILEY, AND J.-P. LAFORTE (2009): “A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model,” Finance and Economics Discussion Series 2009-10, Board of Governors of the Federal Reserve System (U.S.).
- EICHENBAUM, M., AND J. D. FISHER (2007): “Estimating the frequency of price re-optimization in Calvo-style models,” *Journal of Monetary Economics*, 54(7), 2032–2047.
- FERNÁNDEZ-VILLAYERDE, J., J. F. RUBIO-RAMÍREZ, T. J. SARGENT, AND M. W. WATSON (2007): “ABCs (and Ds) of Understanding VARs,” *American Economic Review*, 97(3), 1021–1026.
- HARVEY, D., S. LEYBOURNE, AND P. NEWBOLD (1997): “Testing the equality of prediction mean squared errors,” *International Journal of Forecasting*, 13(2), 281–291.
- KIMBALL, M. S. (1995): “The Quantitative Analytics of the Basic Neomonetarist Model,” *Journal of Money, Credit and Banking*, 27(4), 1241–77.
- LEES, K., T. MATHESON, AND C. SMITH (2007): “Open economy DSGE-VAR forecasting and policy analysis - head to head with the RBNZ published forecasts,” Reserve Bank of New Zealand Discussion Paper Series DP2007/01, Reserve Bank of New Zealand.
- LITTERMAN, R. B. (1986): “Forecasting with Bayesian Vector Autoregressions - Five Years of Experience,” *Journal of Business & Economic Statistics*, 4(1), 25–38.

- NEWKEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–08.
- (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61(4), 631–53.
- ROBERTSON, J. C., AND E. W. TALLMAN (1999): "Vector autoregressions: forecasting and reality," *Economic Review*, (Q1), 4–18.
- RUBASZEK, M., AND P. SKRZYPCZYNSKI (2008): "On the forecasting performance of a small-scale DSGE model," *International Journal of Forecasting*, 24(3), 498–512.
- SIMS, C. A. (2003): "Probability Models for Monetary Policy Decisions," <http://sims.princeton.edu/yftp/ottawa/probmodels.pdf>, Unpublished.
- SIMS, C. A., AND T. ZHA (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–68.
- SMETS, F., AND R. WOUTERS (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175.
- SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.
- TAYLOR, J. B., AND V. WIELAND (2009): "Surprising Comparative Properties of Monetary Models: Results from a New Data Base," NBER Working Papers 14849, National Bureau of Economic Research, Inc.

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (<http://www.ecb.europa.eu>).

- 1077 "The reception of public signals in financial markets – what if central bank communication becomes stale?" by M. Ehrmann and D. Sondermann, August 2009.
- 1078 "On the real effects of private equity investment: evidence from new business creation" by A. Popov and P. Roosenboom, August 2009.
- 1079 "EMU and European government bond market integration" by P. Abad and H. Chuliá, and M. Gómez-Puig, August 2009.
- 1080 "Productivity and job flows: heterogeneity of new hires and continuing jobs in the business cycle" by J. Kilponen and J. Vanhala, August 2009.
- 1081 "Liquidity premia in German government bonds" by J. W. Ejsing and J. Sihvonen, August 2009.
- 1082 "Disagreement among forecasters in G7 countries" by J. Dovern, U. Fritsche and J. Slacalek, August 2009.
- 1083 "Evaluating microfoundations for aggregate price rigidities: evidence from matched firm-level data on product prices and unit labor cost" by M. Carlsson and O. Nordström Skans, August 2009.
- 1084 "How are firms' wages and prices linked: survey evidence in Europe" by M. Druant, S. Fabiani, G. Kezdi, A. Lamo, F. Martins and R. Sabbatini, August 2009.
- 1085 "An empirical study on the decoupling movements between corporate bond and CDS spreads" by I. Alexopoulou, M. Andersson and O. M. Georgescu, August 2009.
- 1086 "Euro area money demand: empirical evidence on the role of equity and labour markets" by G. J. de Bondt, September 2009.
- 1087 "Modelling global trade flows: results from a GVAR model" by M. Bussière, A. Chudik and G. Sestieri, September 2009.
- 1088 "Inflation perceptions and expectations in the euro area: the role of news" by C. Badarinza and M. Buchmann, September 2009.
- 1089 "The effects of monetary policy on unemployment dynamics under model uncertainty: evidence from the US and the euro area" by C. Altavilla and M. Ciccarelli, September 2009.
- 1090 "New Keynesian versus old Keynesian government spending multipliers" by J. F. Cogan, T. Cwik, J. B. Taylor and V. Wieland, September 2009.
- 1091 "Money talks" by M. Hoerova, C. Monnet and T. Temzelides, September 2009.
- 1092 "Inflation and output volatility under asymmetric incomplete information" by G. Carboni and M. Ellison, September 2009.
- 1093 "Determinants of government bond spreads in new EU countries" by I. Alexopoulou, I. Bunda and A. Ferrando, September 2009.
- 1094 "Signals from housing and lending booms" by I. Bunda and M. Ca'Zorzi, September 2009.
- 1095 "Memories of high inflation" by M. Ehrmann and P. Tzamourani, September 2009.

- I096 “The determinants of bank capital structure” by R. Gropp and F. Heider, September 2009.
- I097 “Monetary and fiscal policy aspects of indirect tax changes in a monetary union” by A. Lipińska and L. von Thadden, October 2009.
- I098 “Gauging the effectiveness of quantitative forward guidance: evidence from three inflation targeters” by M. Andersson and B. Hofmann, October 2009.
- I099 “Public and private sector wages interactions in a general equilibrium model” by G. Fernández de Córdoba, J.J. Pérez and J. L. Torres, October 2009.
- I100 “Weak and strong cross section dependence and estimation of large panels” by A. Chudik, M. Hashem Pesaran and E. Tosetti, October 2009.
- I101 “Fiscal variables and bond spreads – evidence from eastern European countries and Turkey” by C. Nickel, P. C. Rother and J. C. Rülke, October 2009.
- I102 “Wage-setting behaviour in France: additional evidence from an ad-hoc survey” by J. Montornés and J.-B. Sauner-Leroy, October 2009.
- I103 “Inter-industry wage differentials: how much does rent sharing matter?” by P. Du Caju, F. Rycx and I. Tojerow, October 2009.
- I104 “Pass-through of external shocks along the pricing chain: a panel estimation approach for the euro area” by B. Landau and F. Skudelny, November 2009.
- I105 “Downward nominal and real wage rigidity: survey evidence from European firms” by J. Babecký, P. Du Caju, T. Kosma, M. Lawless, J. Messina and T. Rõõm, November 2009.
- I106 “The margins of labour cost adjustment: survey evidence from European firms” by J. Babecký, P. Du Caju, T. Kosma, M. Lawless, J. Messina and T. Rõõm, November 2009.
- I107 “Interbank lending, credit risk premia and collateral” by F. Heider and M. Hoerova, November 2009.
- I108 “The role of financial variables in predicting economic activity” by R. Espinoza, F. Fornari and M. J. Lombardi, November 2009.
- I109 “What triggers prolonged inflation regimes? A historical analysis.” by I. Vansteenkiste, November 2009.
- I110 “Putting the New Keynesian DSGE model to the real-time forecasting test” by M. Kolasa, M. Rubaszek and P. Skrzypczyński, November 2009.

