On the forecasting performance of a small-scale DSGE model[1]

Michał Rubaszek*

Paweł Skrzypczyński**

October, 2008

Abstract

Dynamic stochastic general equilibrium (DSGE) models have recently become standard tools for policy analysis. Nevertheless, their forecasting properties are still barely explored. In this article, we address the problem by examining the quality of forecasts from a small-size DSGE model, trivariate vector autoregression (VAR) models and the Philadelphia Fed Survey of Professional Forecasters (SPF) for the key U.S. economic variables: the three-month Treasury bill yield, the GDP growth rate and GDP price index inflation. The *ex post* forecast errors are evaluated on the basis of the data from the period 1994-2006. We apply the Philadelphia Fed "Real-Time Data Set for Macroeconomists" to ensure that the information available to the SPF was comparable to the data used in estimating the DSGE and VAR models.

Overall, the results are mixed. It appears that when comparing the root mean squared errors for some forecast horizons, the DSGE model outperforms the other methods in forecasting the GDP growth rate. However, this characteristic turned out to be statistically insignificant. In principle, most of the SPF's forecasts of GDP price index inflation and the short-term interest rate are better than those from the DSGE and VAR models.

Keywords: Bayesian inference; real-time data; DSGE model, Survey of Professional Forecasters; Vector Autoregression models.

JEL Classification: C11, C32, C53, E12, E17.

---

[1] Published in the International Journal of Forecasting 24 (2008): 498-512.

* National Bank of Poland, Warsaw School of Economics, e-mail: Michal.Rubaszek@nbp.pl (corresponding author)

** National Bank of Poland, e-mail: Pawel.Skrzypczynski@nbp.pl.

1. Introduction

Forecasting inflation, output and interest rates in the United States is one of the crucial tasks for many foreign financial institutions and other economic entities. The reason is that the ability to predict the future state of the U.S. economy accurately facilitates a decision-taking process. For instance, a precise forecast of short-term interest rates would be useful information for an investment fund while setting the duration of its bond portfolio. Similarly, many foreign central banks would like to know more about the future economic situation in the United States while setting the level of domestic interest rates. As a result, the question arises as to which method is the most appropriate in forecasting the U.S. economy. We explore this issue by comparing the forecast performance of a small-scale dynamic stochastic general equilibrium (DSGE) model, the Survey of Professional Forecasters (SPF) and vector autoregression (VAR) models.

DSGE models have recently become standard tools for quantitative policy analysis in macroeconomics, mainly due to their characteristics such as micro-foundations or explicitly modeled expectations. Nevertheless, as stated by Smets & Wouters (2004), these models are only rarely applied to forecasting on the basis of the argument that they perform poorly in this field. In contrast, Del Negro et al. (2005) claim that due to the improved time series fit of these models, their role in forecasting should increase. Although a few central banks have recently decided to use DSGE models for projecting inflation, the discussion about the application of DSGE models for macroeconomic forecasting is still open, especially as the documentation of their out-of-sample performance is still scarce. We consider this issue to be of special importance, in particular as the growing use of DSGE models requires an answer to the question about their ability to forecast future economic developments.

We believe that the SPF, which is the oldest regularly conducted survey of macroeconomic forecasts, represents a plausible approximation of market expectations about the future economic situation in the United States. Therefore, we deem the SPF to represent a very good reference point in evaluating forecasts from an estimated model.

Since the publications of Sims (1980) and Litterman (1986), VAR and Bayesian VAR (BVAR) models have been widely applied in macroeconomics, both for policy analysis and forecasting. Moreover, since infinite order VARs constitute unconstrained versions of DSGE models, they have also been extensively used as benchmarks for evaluating the performance of these micro-founded models. This was done by comparing impulse response functions, as in Christiano et al. (2005), or forecast errors, as in Smets & Wouters (2004). Better forecast accuracy of a DSGE model than that of the parallel VAR would justify constraints given by the economic theory. Bearing this in mind, we also investigate whether the quality of forecasts from our DSGE model is superior to those from the corresponding VAR and BVAR.

As stated by Croushore & Stark (2001a), while comparing the *ex post* forecast performance of an estimated model and the SPF on the basis of the latest-available data, the researcher is favoring his model for two reasons. First, he knows the *ex post* realizations of the data and thereby has a richer data set for building the model. We are aware of the fact that the professional forecasters could not use estimated DSGE models in the mid-1980s because these models were not well developed then. Second, the latest-available data may substantially differ from those disposable to the professional forecasters due to revisions. We address this issue by applying the Philadelphia Fed "Real-Time Data Set for Macroeconomists", described in more detail by Croushore & Stark (2001a). This controls that the information available to the SPF is exactly the same as the data applied to estimate our DSGE, VAR and BVAR models.

The contribution of the paper is twofold. First of all, we extend knowledge about the forecasting properties of small-scale DSGE models. Secondly, we believe that this is the first study that compares forecasts from a DSGE model with those from the SPF in a real-time environment.[2]

The paper is organized as follows. In section 2 we examine the literature that discusses the forecasting performance of DSGE and VAR models, as well as the SPF. Section 3 outlines the methods applied to generate forecasts. In section 4 we describe the real-time data used for estimating the DSGE, VAR and BVAR models. Section 5 focuses on the parameter estimates and properties of the proposed DSGE model. Finally, section 6 presents the results of the out-of-sample forecast performance analysis. We conclude in the last section.

2. Literature review

The number of articles evaluating the forecasting properties of DSGE models in a real-time environment is small. According to our best knowledge, the only such analysis, which was elaborated by Edge et al. (2006), compares forecasts from a random walk, VARs and a richly-specified DSGE model to Federal Reserve (FRB) staff projections. The considered variables are the growth rates of GDP and real consumption, as well as GDP price index and

[2] Edge et al. (2006) have independently performed a similar analysis. The difference is that they use a medium-size DSGE model and that the forecast performance evaluation period is shorter.

PCE inflation. The authors find that FRB staff is the best inflation forecaster, while the DSGE, VAR and BVAR models dominate in forecasting the GDP growth rate. It should be noted, however, that the forecast evaluation period of 1996-2000 is relatively short, and thereby the results might not be representative.

There are a number of articles that compare the quality of forecasts from DSGE and VAR models. However, these analyses are generally not carried out in a real-time context. The two most notable examples are the papers by Smets & Wouters (2004) and Del Negro et al. (2005).[3] The former article illustrates how a medium-scale DSGE model for the euro area, introduced by Smets & Wouters (2003), can be applied for macroeconomic projections and economic analyses. The authors compare the root mean squared errors (RMSE) of forecasts from the DSGE model to those from the corresponding VARs for seven macroeconomic variables, among them output growth, inflation and nominal interest rates, and find that the DSGE model moderately outperforms the VARs. The latter article develops a DSGE-VAR model, which can be characterized as a BVAR with priors deriving from a DSGE model, and applies this new concept to the previously mentioned Smets & Wouters (2003) framework. Afterwards, on the basis of the rolling sample from 1985-2000, the authors compare forecasts for seven variables describing the U.S. economy, and report that the RMSEs of the forecasts from the DSGE and VAR models are generally comparable, but the forecast accuracy of both these models is inferior to that of the DSGE-VAR model.

A number of papers compare the quality of real-time forecasts from the SPF to those from VARs or univariate atheoretical models such as ARIMAs. For example, Croushore (2006) focuses on forecasts of GDP price index inflation in the United States over the period 1971-2004. The author tests and accepts the hypothesis that the SPF forecasts are unbiased, except for the sub-sample period 1971-1981, when the U.S. economy was affected by the oil shocks. Subsequently, he compares the forecast errors from the SPF and a simple ARIMA model estimated on the basis of the real-time data, and concludes that there is little evidence that the univariate model can outperform the professional forecasters. As stated by the author, the evidence in favor of ARIMA models in the earlier studies was derived from using the latest-available data rather than the real-time ones. Clark & McCracken (2006) present a more extensive study comparing real-time forecasts of the SPF, FRB staff and numerous atheoretical models. They report that the performance of VARs and univariate models in forecasting the GDP growth rate, GDP price index and CPI inflation, and the Treasury bill

[3] Similar analyses and results are presented also e.g. by Korenok & Swanson (2005) or Adolfson et al. (2007).

yield is roughly the same. They also find that forecast errors from VARs are significantly higher than those from the SPF, especially for the one-quarter forecast horizon. Finally, the SPF appeared to be more successful than FRB staff in forecasting the GDP growth rate, but less successful in the case of GDP price index and CPI inflation.

The general picture that emerges from the above studies is that in a real-time context the SPF can better forecast the economy than atheoretical models such as VARs or ARIMAs. Furthermore, if the forecast performance is evaluated on the basis of the latest-available data, DSGE models behave comparable or even superior to VARs. The question arises whether DSGE models can beat the SPF in forecasting the U.S. economy, if the real-time data are used. Providing an answer is the main purpose of this article.

3. Models

This section presents the three methods that are applied to forecast the key macroeconomic variables of the U.S. economy. We start with an extensive description of the structure of a small-scale DSGE model. Then, we give a brief outline of the Survey of Professional Forecasters. Finally, we focus on trivariate VAR and BVAR models for output, inflation and short-term interest rates.

3.1. DSGE model

The model economy is populated by three groups of agents: households that optimize their lifetime utility, firms that maximize profits and monetary authorities that, according to the law, care for price and output stability. The model consists of three core equations: a dynamic IS curve, a forward looking Phillips curve and a monetary policy rule, which determine the path of output, prices and short-term nominal interest rates. The system is put in motion by three structural shocks. The first, productivity, shock affects the level of production technology. The second, demand, shock impacts households' decisions concerning consumption and savings. The third, monetary, shock derives from monetary authorities' decisions.

3.1.1. Firms

Production of a homogenous final good is divided into two stages. First, firms indexed by $k \in [0,1]$, operating under monopolistic competitive conditions, are producing differentiated intermediate goods Y_t^k, which are then sold at price P_t^k. Then, intermediate goods are transformed into the final good by perfectly competitive firms, which have access to constant returns to scale technology of Dixit and Stiglitz (1977):

$$Y_t = \left[\int_0^1 \left(Y_t^k \right)^{\frac{\theta-1}{\theta}} dk \right]^{\frac{\theta}{\theta-1}}, \qquad (1)$$

where $\theta > 1$ is the elasticity of intra-temporal substitution. Final good producers minimize the cost of elaborating their output Y_t by deciding on the amount of each differentiated intermediate good they purchase, taking its price as given. The minimal cost is hence equal to:

$$P_t = \left[\int_0^1 \left(P_t^k \right)^{1-\theta} dk \right]^{\frac{1}{1-\theta}}, \qquad (2)$$

and constitutes the consumer price of the final good. The optimal decisions of final good producers also determine the demand for *k*-th intermediate good as:

$$Y_t^k = \left(\frac{P_t^k}{P_t} \right)^{-\theta} Y_t. \qquad (3)$$

Each differentiated good is produced by one firm that uses L_t^k units of labor as the only input. The total output Y_t^k is given by the production function with constant returns to scale:

$$Y_t^k = A_t \varepsilon_t^S L_t^k - \frac{Y_t}{\theta}, \qquad (4)$$

where Y_t / θ represents fixed costs guarantying that in equilibrium profits are null, A_t is a deterministic trend in technology $\ln(A_t) = gt$, and ε_t^S is a supply shock following an AR(1) process $\varepsilon_t^S = \left(1 - \rho^S\right)\bar{\varepsilon}^S + \rho^S \varepsilon_{t-1}^S + \eta_t^S$ with η_t^S being an IID white noise disturbance. The first-order condition for cost minimization implies that the nominal marginal cost per unit of output is equal to:

$$MC_t^N = \frac{W_t}{A_t \varepsilon_t^S}, \qquad (5)$$

and the instantaneous profits are given by:

$$D_t^k = \left(P_t^k - MC_t^N\right)\left(\frac{P_t^k}{P_t}\right)^{-\theta} Y_t - \frac{P_t Y_t}{\theta}. \qquad (6)$$

Intermediate goods producers are assumed to operate in a sticky-price environment introduced in line with the staggered contract framework proposed by Calvo (1983). In each period the representative firm is allowed to set the price of its output at a desirable level with probability $(1-\xi)$. In the other case, the price is automatically adjusted by a steady-state inflation rate ($\overline{\Pi}$) and a fraction δ of the last period's excessive inflation rate.[4] Thus, if firm k has not re-optimized the price of its output since period t, then the price in period $t+s$ is equal to:

$$P_{t+s}^k = P_t^k \left(\frac{P_{t+s-1}}{P_{t-1}\overline{\Pi}^s}\right)^{\delta} \left(\overline{\Pi}\right)^s. \qquad (7)$$

Producers that are allowed to re-optimize their price are maximizing the present value of their discounted intertemporal profits:

$$\max_{P_t^k} E_t \left\{ \sum_{s=0}^{\infty} \xi^s Q_{t,t+s} D_{t+s}^k \right\}, \qquad (8)$$

where $Q_{t,t+s}$ is a stochastic discount factor. Substituting equations (6) and (7) into equation (8) yields the following optimization problem:

$$\max_{P_t^k} E_t \left\{ \sum_{s=0}^{\infty} \xi^s Q_{t,t+s} \left[\left(P_t^k \left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\delta} \overline{\Pi}^s - MC_{t+s}^N \right) \left(\frac{P_t^k \left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\delta} \overline{\Pi}^s}{P_{t+s}} \right)^{-\theta} - \frac{P_{t+s}}{\theta} \right] Y_{t+s} \right\}, \qquad (9)$$

for which the first-order condition yields:

[4] Similar indexation patterns were introduced by Smets & Wouters (2003) or Christiano et al. (2005).

$$E_t\left\{\sum_{s=0}^{\infty}\xi^s Q_{t,t+s} Y_{t+s}^k\left(\tilde{P}_t^k\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\delta}\overline{\Pi}^s-\frac{\theta}{\theta-1}MC_{t+s}^N\right)\right\}=0, \tag{10}$$

where $\tilde{P}_t^k$ is the price that maximizes the expected value of future dividends. According to equation (10), the chosen price depends positively on current and expected future marginal costs. Finally, according to the definition of the aggregate price index given by equation (2), the price level is equal to:

$$P_t=\left[\xi\left(P_{t-1}\left(\frac{P_{t-1}}{P_{t-2}}\right)^{\delta}\overline{\Pi}\right)^{1-\theta}+(1-\xi)\left(\tilde{P}_t^k\right)^{1-\theta}\right]^{\frac{1}{1-\theta}}. \tag{11}$$

3.1.2. Households

The model economy is populated by a continuum of homogenous households indexed by $i\in[0,1]$. In each period, a typical household maximizes its lifetime utility function:

$$E_t\sum_{s=0}^{\infty}\beta^{t+s}U_{t+s}^i\left(H_{t+s}^i;L_{t+s}^i;\varepsilon_{t+s}^D\right), \tag{12}$$

where $\beta<1$ is a time-invariant discount factor. The utility function U_t^i is an increasing function of instantaneous consumption of the representative household with respect to a fraction λ of aggregate past consumption adjusted for the growth rate of technology g, called henceforth habit:

$$H_t^i=C_t^i-\lambda(1+g)C_{t-1}, \tag{13}$$

and a decreasing function of labor supplied by the typical household L_t^i:

$$U_t^i=E_t\sum_{s=0}^{\infty}\beta^s\varepsilon_{t+s}^D\left\{\frac{\left(H_{t+s}^i/A_{t+s}\right)^{1-\sigma}}{1-\sigma}-v_L\frac{\left(L_{t+s}^j\right)^{1+\varphi}}{1+\varphi}\right\}. \tag{14}$$

The coefficient σ is the inverse of the intertemporal elasticity of substitution, φ is the inverse of labor supply elasticity with respect to real wages, and ε_t^D is a demand shock that is assumed to be an AR(1) process $\varepsilon_t^D = (1-\rho^D)\bar{\varepsilon}^D + \rho^D \varepsilon_{t-1}^D + \eta_t^D$, where η_t^D is an IID white noise disturbance.

In each period, the representative household receives nominal remuneration for the work effort $W_t L_t^i$, dividends from owned firms D_t^i and repayment of funds previously invested in one-period bonds B_{t-1}^i. The money is spent on consumption $P_t C_t^i$ or invested in debt securities, which are sold at a discount rate $1/R_t$. Consequently, the budget constraint is of the following form:

$$\frac{B_t^i}{R_t} + P_t C_t^i = B_{t-1}^i + W_t L_t^i + D_t^i . \qquad (15)$$

In order to maximize the intertemporal utility function (14) subject to the budget constraint (15) the typical household must take two decisions. First, it must choose how much money should be spent on current consumption and how much should be invested in bonds. The solution of this problem leads to the specification of the following dynamic IS curve:

$$\left(\frac{H_t^i}{A_t}\right)^{-\sigma} = \beta E_t \left\{ \frac{\varepsilon_{t+1}^D}{\varepsilon_t^D} \left(\frac{R_t}{\Pi_{t+1}}\right) \frac{A_t}{A_{t+1}} \left(\frac{H_{t+1}^i}{A_{t+1}}\right)^{-\sigma} \right\}, \qquad (16)$$

where $\Pi_t = P_t / P_{t-1}$ is the inflation rate. Second, the typical household must decide how much time it is willing to spend at work. On the one hand, higher labor intensity increases its revenue form remuneration, but on the other, it lessens the amount of its leisure time. The outcome of the optimization yields the labor supply curve:

$$\frac{W_t}{P_t A_t} = v_L \left(\frac{H_t^i}{A_t}\right)^{\sigma} \left(L_t^i\right)^{\varphi} . \qquad (17)$$

3.1.3. Monetary authorities and market clearing condition

The central bank is supposed to be obliged by law to minimize the variation of inflation and output. The short-term nominal interest rate is hence adjusted, as in Taylor (1993), in response to changes in these two variables. Following Rudebusch (2002) and Orphanides & Williams (2002), we extended Taylor's original specification by introducing interest rate smoothing into the monetary policy reaction function:

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}}\right)^{\gamma} \left[\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{(\gamma_\pi)} \left(\frac{Y_t}{Y_{t-1}(1+g)}\right)^{(\gamma_{\Delta y})}\right]^{(1-\gamma)} e^{(\eta_t^M)}, \tag{18}$$

where η^M is a monetary shock following an IID white noise process.

The model is closed by the following clearing condition on the final good market:

$$Y_t = C_t, \tag{19}$$

stating that aggregate supply is equal to aggregate demand.

3.2. 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. It is carried out in regular three-month intervals and concerns dozens of macroeconomic variables, among them output, inflation and short-term interest rates. Since each of anonymous respondents to the survey is producing regular economic forecasts as part of his or her professional activity, it was named as the Survey of Professional Forecasters. The results of the survey are published quarterly on the Philadelphia Fed website.[5]

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 GDP data for the previous period, and 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 U.S. 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 above-described estimated models, particularly in the one-quarter-ahead horizon.

This paper focuses on the median forecasts of the three following variables: the quarterly GDP growth rate, GDP price index inflation and the three-month Treasury bill yield. All these variables are forecasted by the SPF up to four quarters ahead, where the one-step

[5] See: http://www.phil.frb.org/econ/spf/index.html.

forecasts concern the period when the survey is carried out. For instance, the forecasts for the period 1994:1-1994:4 come from the survey from the first quarter of 1994.

3.3. VAR model

Since infinite order VARs constitute unconstrained representations of DSGE models, they have been widely used as benchmarks for evaluating the forecasting performance of these micro-funded models (see Smets & Wouters (2004), Del Negro et al. (2005) or Adolfson et al. (2007)). In this regard, we also analyze VAR models and its Bayesian counterparts for the vector of the three U.S. macroeconomic variables explained by the DSGE model introduced in the previous subsection, namely the short-term interest rate, inflation and output growth. For the vector of these variables $\mathbf{X}_t$ we estimate the following VAR:

$$\mathbf{X}_t = \mathbf{\Gamma}_0 + \sum_{i=1}^{P} \mathbf{\Gamma}_i \mathbf{X}_{t-i} + \boldsymbol{\varepsilon}_t , \tag{20}$$

where $\mathbf{\Gamma}_i$ for $i = 0,1,...,P$ are matrices of coefficients, P is a lag order and $\boldsymbol{\varepsilon}_t$ is a vector of residuals. In the case of the VAR model, we chose the optimal lag order on the basis of the final prediction error criterion, where the maximum available lag was set to five. For the BVAR model, we fixed the maximum lag at four quarters and assumed the Minnesota prior distribution for the model's coefficients (see Litterman, 1986). The relevant hyperparameters for overall tightness, cross-equation tightness and harmonic lag decay were taken, as suggested by LeSage (1999), to be 0.2, 0.1 and 1, respectively.

4. The data

We focus on the following key macroeconomic variables for the U.S. economy: the three-month Treasury bill yield and the quarterly growth rates of seasonally adjusted real GDP and GDP price index. These variables represent the models' short-term interest rate, output growth and inflation, respectively. Due to the fact that some of the above time series are subject to revisions, the use of the latest-available data would lead to favoring the investigated estimated models upon the SPF for the reasons discussed in previous sections. We tackle this problem by using the real-time data, which increases the comparability of the forecasting errors, as all predictions are formulated on the basis of the same data set.

Henceforth, by the term "real-time data" we understand values of macroeconomic time series available to a researcher on a given date in the past, which we refer to as a "vintage".

In this analysis, the real-time data for GDP and the GDP price index are taken from the Philadelphia Fed Real-Time Data Set for Macroeconomists.[6] Since the vintages are chosen to be the middle day of each quarter, they include the advance release of national account data for the previous period. This means that the real-time data match up exactly with the information available to the SPF. Concerning the three-month Treasury bill yield, the time series are not revised over time and thereby the latest-available data are the same as the real-time ones.[7]

The out-of-sample forecast performance is analyzed for horizons ranging from one up to four quarters ahead, whereas the evaluation is based on the data from the period 1994:1-2006:2, called henceforth the "evaluation sample". The DSGE, VAR and BVAR models are estimated on the basis of the most recent 60 quarterly observations for a given vintage date, which is the period of forecast formulation. For instance, the forecasts elaborated in the first quarter of 1994 for the period 1994:1-1994:4 are generated using the models estimated on the basis of observations from 1979:1 to 1993:4 given by the vintage of 1994:1. The next set of models is estimated with the real-time data from 1979:2-1994:1 given by the vintage of 1994:2, and these models are applied to forecast the U.S. economy for the period 1994:2-1995:1. This procedure is repeated for each quarter from the period 1994:1-2005:3, which means that we calculate 47 forecasts for each forecast horizon, model and variable.

5. DSGE estimation results

The empirical implementation of the DSGE model is done in several steps. First, the model is log-linearized around its steady-state (see Appendix 1) and written as a linear expectation system. Second, the system is solved out using standard techniques and transformed into a state-space representation, where the measurement equation relates the model's variables to the observable ones, namely the three-month Treasury bill yield, GDP growth rate and GDP price index inflation. The state-space representation includes eleven structural parameters of the DSGE model, three parameters that describe a diagonal matrix of shocks variance and three parameters that pin down steady-state values of output growth, inflation and nominal interest rate to the data.

[6] The data are available on the Philadelphia Fed web page: http://www.phil.frb.org/econ/forecast/reaindex.html.
[7] The data are available on the Fed web page: http://www.federalreserve.gov/releases/.

Before the estimation, we fixed the value of the discount factor β at 0.995 and, as in Irealand (2004), we relaxed the model's assumption that $\overline{R} = (1+g)\overline{\Pi}/\beta$, which is not confirmed by historical observations. Instead, we regard $\overline{R}$ as the additional parameter to be estimated. The inference about the remaining parameters is drawn by applying Bayesian techniques. The left-side columns of Table 1 show our assumptions for the prior distribution of the estimated parameters, which generally correspond to those in Ireland (2004), Smets & Wouters (2004), Del Negro et al. (2005) or Adolfson et al. (2007).

The posterior density function, which is the product of the likelihood function and the prior density, is maximized with respect to the state-space model parameters. The posterior mode and the corresponding Hessian matrix are calculated using standard numerical optimization routines. The distribution characteristics of the recursive estimates of the posterior mode are reported in the right-side columns of Table 1. According to the results, all coefficients are reasonable across the evaluation sample, and their values are close to those found in the above-mentioned articles. In brief, the model displays low inflation inertia, a moderate degree of price stickiness, standard reaction of monetary policy to changes in output and inflation, and high persistence of the supply shock.

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 three structural shocks. Figure 1 shows that the positive demand shock leads to a hump-shaped increase in output, which in turn produces an upward pressure on wages, inflation and interest rates. The supply shock increases output, while also diminishing marginal costs across the firms. A respective fall in inflation dominates the output growth effects on interest rates. The impact of the supply shock on the economy is very persistent, except for the initial six observations of the evaluation sample. The monetary shock entails a rise in nominal interest rates that results in lower output and inflation. Finally, it can be noticed that the informal analysis of the recursive IRFs shows the relative stability of the model across the evaluation sample, except for the few initial observations.

In the last step, we apply the model to forecasting the U.S. economy. For this purpose, we approximate the posterior distribution of the structural parameters through Monte-Carlo Markov-Chain sampling methods. Then, for each of the 25,000 draws, we compute the out-of-sample values for the observable variables and take the median of these forecasts for further analysis. We repeat this procedure for each quarter from the evaluation sample. All calculations are performed with DYNARE package for MATLAB.[8]

6. Forecast performance evaluation

This section reports the out-of-sample forecast performance of the SPF, VAR, BVAR and DSGE model for the short-term interest rate, output growth and inflation at horizons up to four quarters. Since the analysis is conducted in a real-time data environment, while calculating the forecast errors, it must be decided what to use as the "actuals" for the forecasted variables. We evaluate the quality of the forecasts in two variants. First, we take the latest-available data set, i.e. from the vintage of 2006:3, as the actuals. The graphical comparison of the forecasts to this kind of actuals is presented in Figure 2. Second, we evaluate the forecasts with the real-time data available one year after the date of the vintage used in estimation. We label the former case as the "latest-available", and the latter as the "one year after estimation".

We begin by examining whether the forecasts are biased. In this regard, for each method, variable and forecast horizon we regress the "actuals" (X_t) on the forecasts (X_t^F):

$$X_t = \alpha_0 + \alpha_1 X_t^F + \varepsilon_t \, , \tag{21}$$

and test the null hypothesis that the constant term is null ($\alpha_0 = 0$) and the slope coefficient is unity ($\alpha_1 = 1$), which if accepted, indicates that the forecast is unbiased. We apply the Wald Chi-squared test corrected for heteroskedasticity and autocorrelation of the residuals. The adequate covariance matrix is estimated in line with the Newey & West (1987) procedure: we use the modified Bartlett kernel, where the truncation lag is dependent on the number of observations as proposed by Newey & West (1994). Since the results for the "latest-available" and the "one year after estimation" actuals are broadly the same, in Table 2 we report only the main findings for the former ones. At the 5% significance level, the short-term interest rate forecasts are unbiased, except for the one-quarter-ahead forecast from the SPF. The forecasts of output growth and inflation are imprecise: the relevant coefficients of determination are low and barely exceed 10%. Moreover, for the output growth forecasts, the unbiasedness hypothesis cannot be rejected only in the case of the DSGE model at the two and four-quarter horizons, and the SPF at the one-quarter horizon. Inflation forecasts are biased in all cases.

[8] The Dynare package is available on the website: http://www.cepremap.cnrs.fr/dynare.

We proceed by comparing the mean errors (ME), the mean absolute errors (MAE) and the root mean squared errors (RMSE) of the forecasts. According to the results, which are reported in Table 3, the MAEs and RMSEs of the forecasts for the short-term interest rate are the lowest for the SPF and the highest for the VAR. The superiority of the SPF over the remaining methods is evident, especially for the one-quarter-ahead forecast, which should not be surprising taking into account that the professional forecasters know about interest rate movements of the first half of the quarter for which the forecast is elaborated. It should be emphasized, however, that for all forecast horizons, the SPF forecasts are inferior to the VAR and BVAR ones in terms of the MEs. For the output growth forecasts, the best performer is the DSGE model, which is characterized by the lowest MAEs and RMSEs at the three- and four-quarter horizons, and by the MEs that are the closest to zero. With regard to the inflation forecasts, we find that the SPF forecasts are characterized by the lowest RMSEs among the competing methods at all horizons. Overall, the results indicate that the SPF performance is the best in the case of forecasts for the short-term interest rate and inflation, while the DSGE forecasts are the most plausible in the case of output.

While the RMSEs are widely used for evaluating the forecast performance of a given method, these statistics do not indicate whether one method is statistically better than another. We tackle this issue by employing the Harvey-Leybourne-Newbold (1997) modification of the Diebold-Mariano (1995) test for the null hypothesis of equal forecast accuracy from two competing models. The results, which are reported in Table 4, show that the short-term interest rate forecasts from the SPF are significantly better than those from the remaining methods at the horizons up to three quarters. For the output growth forecasts, we find that both the SPF and the DSGE model perform better than the VAR for the "latest-available" actuals. However, the hypothesis that the DSGE model significantly dominates the SPF and BVAR in forecasting output was not confirmed by the data. The comparison of the inflation forecasts RMSEs shows that the SPF generally outperforms the VAR and DSGE model at up to the three-quarter horizons.

7. Conclusions

In the paper we have compared the quality of the forecasts from the DSGE, VAR and BVAR models, as well as from the SPF in the case of the three key U.S. macroeconomic variables. Since the analysis was carried out in a real-time context, we controlled the comparability of the information available to the SPF with the time series used to estimate coefficients of the analyzed models. We believe that this is the first study that compares the forecast errors from a DSGE model with those from the SPF in a real-time environment.

The results show that the short-term interest rate forecasts were unbiased in the case of all methods and horizons, except for the one-quarter-ahead forecast from the SPF. In contrast, all inflation forecasts turned out to be biased. Moreover, the forecasts of inflation and output came out to be imprecise: the relevant coefficients of determination were always below 10%. In the case of the output growth forecasts, the only unbiased predictions were generated by the DSGE model. Comparison of the RMSEs showed that the DSGE model performs relatively well in forecasting output growth, while the SPF is superior to the estimated models in forecasting the short-term interest rate and inflation. The Harvey-Leybourne-Newbold test of the null hypothesis about equal forecast accuracy showed that in some cases forecasts of inflation and the short-term interest rate from the SPF are significantly better than those from the DSGE, VAR and BVAR models.

The general picture that emerges from the above analysis is that the proposed DSGE model is not able to significantly outperform the SPF in forecasting output growth, inflation nor interest rates in the United States. We found, however, that the DSGE model generates the forecasts which are very close in accuracy to the SPF predictions. Moreover, the DSGE model was found to perform comparable to the trivariate VAR and BVAR models. Clearly, additional research is required to document the out-of-sample performance of DSGE models: the structure of the DSGE model presented in this article is relatively simple and hence forecasting properties of a more complex DSGE model could be studied. Furthermore, the forecast accuracy of DSGE models could be compared to a larger group of methods than the SPF and VAR models.

References

Adolfson, M., Linde, J. & Villani, M. (2007). Forecasting Performance of an Open Economy DSGE Model. *Econometric Review*, 26, 289-328.

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

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

Clark, T. & McCracken, M. (2006). *Forecasting with Small Macroeconomic VARs in the Presence of Instabilities*. Federal Reserve Bank of Kansas City Working Paper 06-09.

Croushore, D. (2006). *An Evaluation of Inflation Forecasts From Surveys Using Real-Time Data*. Federal Reserve Bank of Philadelphia Working Paper 06-19.

Croushore, D. & Stark, T. (2001a). A Real-Time Data Set for Macroeconomists. *Journal of Econometrics*, 105, 111-130.

Croushore, D. & Stark, T. (2001b). *Forecasting with a Real-Time Data Set for Macroeconomists.* Federal Reserve Bank of Philadelphia Working Paper 01-10.

Del Negro, M., Schorfheide, F., Smets, F. & Wouters, R. (2005). *On the Fit and Forecasting Performance of New Keynesian Models.* CEPR Discussion Paper Series 4848.

Diebold, F. X. & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263.

Dixit, A. & Stiglitz, J. (1977). Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67, 297-308.

Edge, R., Kiley, M. & Laforte, J. (2006). *A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model.* http://www.eabcn.org/workshops/zurich_2006/documents/Laforte.pdf.

Harvey, D., Leybourne, S. & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), 281-291.

Ireland, P. (2004). *Technology Shocks in the New Keynesian Model.* NBER Working Paper 10309.

Korenok, O. & Swanson, N. (2005). The Incremental Predictive Information Associated with Using Theoretical New Keynesian DSGE Models vs. Simple Linear Econometric Models. *Oxford Bulletin of Economics & Statistics*, 67(S1), 905-930.

LeSage J. (1999). *Applied Econometrics Using MATLAB*. University of Toledo.

Litterman, R. (1986). Forecasting with Bayesian Vector Autoregressions – Five Years of Experience. *Journal of Business and Economic Statistics*, 4(1), 25-38.

Newey, W. K. & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703-708.

Newey, W. K. & West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *The Review of Economic Studies*, 61(4), 631-653.

Orphanides, A. & Williams, J. (2002). Robust monetary policy rules with unknown natural rates. *Brookings Papers on Economic Activity*, 2, 63-118.

Rudebusch, G. (2002). Term structure evidence on interest-rate smoothing and monetary policy inertia. *Journal of Monetary Economics*, 49, 1116-1186.

Sims, Ch. (1980). Macroeconomics and Reality. *Econometrica*, 48, 1-48.

Smets, F. & Wouters, R. (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association*, 1(5), 1123-1175.

Smets, F. & Wouters, R. (2004). Forecasting With a Bayesian DSGE Model. An Application to the Euro Area. *Journal of Common Market Studies*, 42(4), 841-867.

Taylor, J. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Series on Public Policy*, 39, 195-214.

Appendix 1. *Log-linearized version of the DSGE model*

In the steady-state of the model, all intermediate goods-producing firms set their price at the same level, output is equal across the firms and profits are null. The value of bond holdings is also equal to zero. For the reason that in the equilibrium dynamics of some variables depend on the level of technology A_t or price index P_t, we introduce the following de-trended variables $y_t = Y_t / A_t$, $h_t = H_t / A_t$ and $mc_t = MC_t^N / P_t$. In the absence of structural shocks, the economy converges to the stationary steady-sate given by $\bar{y}$, $\bar{h}$, $\overline{mc}$, $\bar{L}$, $\bar{\Pi}$ and $\bar{R}$.

Fluctuations of the economy around this equilibrium are specified by a log-linearized version of the model, in which deviations of a variable x_t from its steady-state value are represented by $\hat{x}_t = \log(x_t/\bar{x})$. The list of equations is as follows. Habit formation is given by:

$$(1-\lambda)\hat{h}_t = \hat{y}_t - \lambda \hat{y}_{t-1},$$

where the Euler equation for habits is:

$$\hat{h}_t = -\frac{1}{\sigma}\left[\hat{R}_t - \hat{\Pi}_{t+1} + \hat{\varepsilon}_{t+1}^D - \hat{\varepsilon}_t^D\right] + \hat{h}_{t+1}.$$

The specification of the augmented forward looking Phillips curve is:

$$\hat{\Pi}_t = \frac{\delta}{1+\beta\delta}\hat{\Pi}_{t-1} + \frac{\beta}{1+\beta\delta}\hat{\Pi}_{t+1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\beta\delta)\xi} m\hat{c}_t,$$

where real marginal costs are equal to:

$$m\hat{c}_t = \sigma \hat{h}_t + \varphi \hat{L}_t - \hat{\varepsilon}_t^S,$$

and labor supply is given by:

$$\hat{L}_t = \hat{y}_t - \hat{\varepsilon}_t^S.$$

Subsequently, a log-linear approximation of the monetary policy reaction function yields:

$$\hat{R}_t = \gamma \hat{R}_{t-1} + (1-\gamma)\left(\gamma_\pi \hat{\Pi}_t + \gamma_{\Delta y}\left(\hat{y}_t - \hat{y}_{t-1}\right)\right) + \eta_t^M.$$

The two last equations specify the law of motion for the demand and supply shocks:

$$\hat{\varepsilon}_t^i = \rho^i \hat{\varepsilon}_{t-1}^i + \eta_t^i,$$

where $i = \{D, S\}$.

Appendix 2. *Tables and figures*

Table 1. Prior distribution and recursive estimates of structural parameters

Parameter		*Prior distribution*			*Recursive mode of posterior*			
		Type	*Mean*	*Std/df*	*min*	*med.*	*mean*	*max*
Habit formation	λ	beta	0.70	0.10	0.45	0.56	0.54	0.65
Elasticity of substitution	σ	normal	1.00	0.38	0.65	0.98	0.97	1.22
Labor supply elasticity	φ	normal	2.00	0.75	1.83	1.97	1.97	2.14
Indexation prices	δ	beta	0.75	0.15	0.25	0.33	0.33	0.43
Calvo prices	ξ	beta	0.75	0.15	0.68	0.80	0.78	0.85
Interest rate smoothing	γ	beta	0.80	0.10	0.69	0.76	0.76	0.84
Inflation response	γ_π	normal	1.70	0.10	1.62	1.74	1.73	1.78
Output growth response	$\gamma_{\Delta y}$	normal	0.15	0.05	0.19	0.22	0.22	0.24
Steady-state inflation		normal	0.50	0.15	0.64	0.72	0.71	0.82
Steady-state output growth		normal	0.75	0.15	0.51	0.71	0.69	0.79
Steady-state nominal int. rate		normal	1.50	0.15	1.29	1.40	1.38	1.44
Supply shock persistence	ρ^S	beta	0.85	0.10	0.94	0.99	0.98	0.99
Demand shock persistence	ρ^D	beta	0.85	0.10	0.64	0.89	0.87	0.93
Supply shock STD	σ^S	inv. gamma	0.40	2	0.75	1.06	1.11	1.70
Demand shock STD	σ^D	inv. gamma	0.20	2	0.85	1.43	1.51	2.76
Monetary shock STD	σ^M	inv. gamma	0.10	2	0.10	1.12	1.14	0.23

Note: The recursive posterior modes of the model's parameters are estimated over fifteen-year window for each quarter of the evaluation sample.

Table 2. Test of forecast unbiasedness – "latest-available" data set case

h	*Short-term interest rate*				*Output growth*				*Inflation*			
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	*p-val.*	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	*p-val.*	$\hat{\alpha}_0$	$\hat{\alpha}_1$	R^2	*p-val.*
SPF												
1	0.002 (0.030)	0.987 (0.007)	0.997	***0.004***	1.206 (0.635)	0.774 (0.189)	0.189	*0.103*	1.488 (0.518)	0.284 (0.216)	0.033	***0.002***
2	0.007 (0.131)	0.964 (0.031)	0.946	*0.174*	2.924 (1.026)	0.103 (0.301)	0.001	***0.011***	2.196 (0.582)	-0.057 (0.247)	0.001	***0.000***
4	0.036 (0.566)	0.909 (0.127)	0.656	*0.337*	4.882 (1.190)	-0.577 (0.405)	0.032	***0.000***	3.379 (0.713)	-0.560 (0.290)	0.090	***0.000***
VAR												
1	0.232 (0.243)	0.969 (0.055)	0.912	*0.497*	3.780 (1.002)	-0.146 (0.228)	0.005	***0.000***	2.008 (0.404)	0.030 (0.177)	0.001	***0.000***
2	0.735 (0.574)	0.845 (0.128)	0.705	*0.440*	2.605 (1.112)	0.178 (0.281)	0.009	***0.006***	2.119 (0.402)	-0.022 (0.200)	0.000	***0.000***
4	1.304 (0.897)	0.654 (0.211)	0.377	*0.257*	1.761 (1.610)	0.376 (0.373)	0.024	***0.009***	2.644 (0.382)	-0.270 (0.179)	0.065	***0.000***
BVAR												
1	0.040 (0.154)	0.999 (0.041)	0.943	*0.901*	2.751 (1.723)	0.144 (0.424)	0.003	***0.015***	1.613 (0.327)	0.242 (0.148)	0.047	***0.000***
2	0.176 (0.349)	0.957 (0.092)	0.807	*0.879*	2.941 (1.390)	0.077 (0.354)	0.001	***0.010***	1.785 (0.340)	0.153 (0.170)	0.019	***0.000***
4	0.494 (0.738)	0.820 (0.189)	0.503	*0.550*	2.765 (1.619)	0.123 (0.409)	0.002	***0.025***	2.161 (0.306)	-0.016 (0.146)	0.000	***0.000***
DSGE												
1	-0.495 (0.206)	1.116 (0.056)	0.916	*0.055*	1.892 (1.099)	0.415 (0.270)	0.038	***0.036***	1.727 (0.360)	0.170 (0.137)	0.027	***0.000***
2	-0.689 (0.478)	1.134 (0.121)	0.776	*0.297*	2.926 (1.583)	0.090 (0.462)	0.001	*0.138*	2.275 (0.459)	-0.091 (0.174)	0.005	***0.000***
4	-0.369 (1.170)	0.983 (0.268)	0.439	*0.420*	1.614 (1.956)	0.552 (0.715)	0.010	*0.390*	3.361 (0.654)	-0.511 (0.251)	0.099	***0.000***

Notes: Table presents the coefficient estimates of regression (21). The figures in parentheses denote the corrected standard errors. The *p*-values relate to the test of the null hypothesis that the forecast is unbiased, where bold figures indicate the rejection at the 5% significance level.

Table 3. Out-of-sample forecast evaluation

h	*Short-term interest rate*				*Output growth*				*Inflation*			
	SPF	*VAR*	*BVAR*	*DSGE*	*SPF*	*VAR*	*BVAR*	*DSGE*	*SPF*	*VAR*	*BVAR*	*DSGE*
ME – "latest-available" data set case												
1	-0.046	0.119	**0.035**	-0.048	0.605	-0.269	-0.293	**-0.045**	**0.042**	0.229	0.202	0.091
2	-0.136	0.172	**0.013**	-0.159	0.392	-0.205	-0.321	**0.024**	**0.000**	0.217	0.160	-0.092
3	-0.234	**0.071**	-0.109	-0.299	0.387	-0.657	-0.518	**0.183**	**-0.077**	0.238	0.173	-0.205
4	-0.343	**-0.025**	-0.236	-0.442	0.350	-0.666	-0.507	**0.306**	**-0.106**	0.222	0.142	-0.282
ME – "one year after estimation" data set case												
1	-0.046	0.119	**0.035**	-0.048	0.703	-0.172	-0.195	**0.053**	-0.086	0.101	0.074	**-0.037**
2	-0.136	0.172	**0.013**	-0.159	0.548	**-0.049**	-0.165	0.180	-0.159	0.059	**0.001**	-0.251
3	-0.234	**0.071**	-0.109	-0.299	0.554	-0.490	-0.351	**0.350**	-0.253	0.062	**-0.002**	-0.380
4	-0.343	**-0.025**	-0.236	-0.442	0.317	-0.699	-0.540	**0.273**	-0.364	**-0.036**	-0.117	-0.541
MAE – "latest-available" data set case												
1	**0.075**	0.355	0.284	0.401	**1.499**	1.937	1.722	1.683	0.640	0.828	0.733	0.756
2	**0.285**	0.671	0.554	0.620	1.731	1.706	1.700	**1.644**	**0.738**	0.844	0.772	0.831
3	**0.563**	0.879	0.744	0.797	1.767	1.662	1.683	**1.516**	**0.773**	0.923	0.821	0.878
4	**0.820**	1.062	0.934	0.988	1.780	1.768	1.789	**1.550**	**0.881**	1.023	0.938	0.980
MAE – "one year after estimation" data set case												
1	**0.075**	0.355	0.284	0.401	**1.472**	1.742	1.608	1.685	**0.561**	0.732	0.653	0.734
2	**0.285**	0.671	0.554	0.620	1.574	1.639	**1.522**	1.700	**0.677**	0.806	0.693	0.809
3	**0.563**	0.879	0.744	0.797	1.667	1.522	**1.503**	1.539	0.723	0.837	**0.712**	0.819
4	**0.820**	1.062	0.934	0.988	1.579	1.545	1.530	**1.520**	**0.861**	0.926	0.892	0.950
RMSE – "latest-available" data set case												
1	**0.099**	0.516	0.405	0.518	**1.910**	2.311	2.127	2.029	**0.820**	1.025	0.921	0.963
2	**0.419**	0.964	0.742	0.830	2.140	2.192	2.161	**2.090**	**0.896**	1.066	0.976	1.018
3	**0.740**	1.193	0.986	1.094	2.250	2.096	2.114	**1.978**	**0.922**	1.120	0.983	1.010
4	**1.056**	1.439	1.240	1.338	2.226	2.151	2.165	**2.024**	**1.027**	1.259	1.118	1.107
RMSE – "one year after estimation" data set case												
1	**0.099**	0.516	0.405	0.518	**1.878**	2.116	1.975	1.985	**0.722**	0.923	0.827	0.896
2	**0.419**	0.964	0.742	0.830	2.032	2.021	**1.995**	2.101	**0.867**	1.003	0.897	0.971
3	**0.740**	1.193	0.986	1.094	2.086	1.927	**1.878**	1.909	0.880	0.998	**0.879**	0.999
4	**1.056**	1.439	1.240	1.338	1.932	1.910	1.897	**1.791**	**1.042**	1.151	1.067	1.148

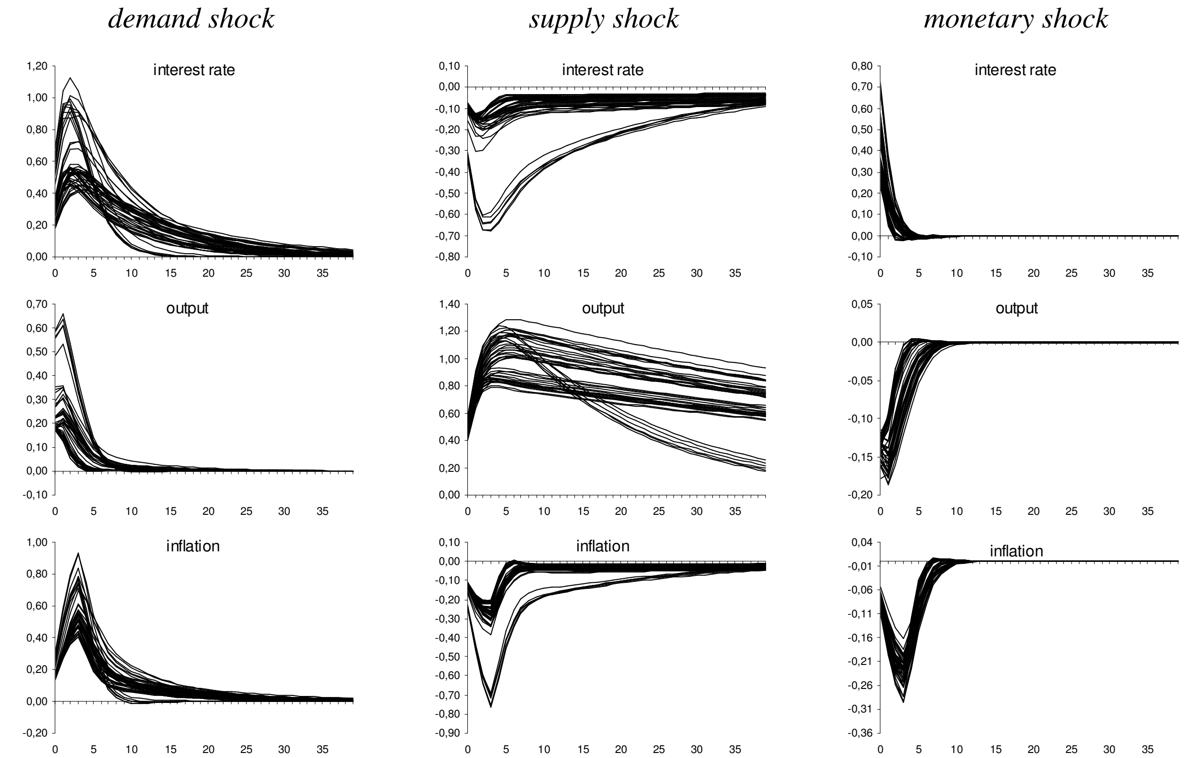
Note: Bold figures are the minimum absolute values for the MEs, MAEs and RMSEs.

Table 4. Test for equal forecast accuracy

h	*The "latest-available" data set*						*The "one year after estimation" data set*					
	Short-term interest rate		*Output growth*		*Inflation*		*Short-term interest rate*		*Output growth*		*Inflation*	
	HLN	*p-val.*	*HLN*	*p-val.*	*HLN*	*p-val.*	*HLN*	*p-val.*	*HLN*	*p-val.*	*HLN*	*p-val.*
DSGE vs. SPF												
1	**3.84**	***0.000***	0.78	*0.441*	1.81	*0.077*	**3.84**	***0.000***	0.52	*0.606*	**2.04**	***0.047***
2	**2.65**	***0.011***	-0.35	*0.727*	**2.51**	***0.016***	**2.65**	***0.011***	0.47	*0.640*	**2.05**	***0.047***
3	1.65	*0.105*	-1.68	*0.100*	**2.61**	***0.012***	1.65	*0.105*	-1.08	*0.286*	**4.24**	***0.000***
4	1.36	*0.180*	-1.41	*0.164*	1.50	*0.140*	1.36	*0.180*	-1.02	*0.314*	1.48	*0.145*
DSGE vs. VAR												
1	0.02	*0.981*	**-2.09**	***0.042***	-0.78	*0.438*	0.02	*0.981*	-1.12	*0.268*	-0.37	*0.717*
2	-0.57	*0.573*	-0.47	*0.639*	-0.69	*0.491*	-0.57	*0.573*	0.43	*0.670*	-0.60	*0.552*
3	-0.39	*0.699*	-0.49	*0.629*	-1.13	*0.265*	-0.39	*0.699*	-0.08	*0.938*	0.02	*0.987*
4	-0.38	*0.705*	-0.87	*0.387*	-1.12	*0.270*	-0.38	*0.705*	-0.62	*0.540*	-0.03	*0.975*
DSGE vs. BVAR												
1	2.01	*0.051*	-1.07	*0.288*	0.46	*0.647*	2.01	*0.051*	0.12	*0.903*	0.83	*0.409*
2	0.75	*0.457*	-0.49	*0.628*	0.48	*0.634*	0.75	*0.457*	0.90	*0.374*	1.04	*0.304*
3	0.65	*0.517*	-0.73	*0.472*	0.26	*0.797*	0.65	*0.517*	0.17	*0.865*	1.47	*0.148*
4	0.56	*0.581*	-0.94	*0.353*	-0.08	*0.939*	0.56	*0.581*	-0.58	*0.564*	0.73	*0.469*
BVAR vs. SPF												
1	**3.60**	***0.001***	1.31	*0.195*	1.48	*0.145*	**3.60**	***0.001***	0.49	*0.630*	1.71	*0.093*
2	**2.70**	***0.010***	0.12	*0.908*	0.81	*0.423*	**2.70**	***0.010***	-0.22	*0.826*	0.44	*0.665*
3	**2.52**	***0.015***	-0.57	*0.570*	0.52	*0.608*	**2.52**	***0.015***	-0.88	*0.381*	-0.01	*0.996*
4	1.73	*0.090*	-0.28	*0.777*	0.70	*0.490*	1.73	*0.090*	-0.13	*0.894*	0.28	*0.778*
BVAR vs. VAR												
1	-1.82	*0.075*	-1.49	*0.143*	-1.80	*0.078*	-1.82	*0.075*	-1.53	*0.133*	**-2.22**	***0.032***
2	-1.23	*0.226*	-0.20	*0.845*	-1.13	*0.264*	-1.23	*0.226*	-0.17	*0.865*	-1.88	*0.067*
3	-1.37	*0.177*	0.26	*0.795*	-1.61	*0.114*	-1.37	*0.177*	-1.11	*0.274*	-1.73	*0.091*
4	-1.38	*0.174*	0.30	*0.765*	-1.88	*0.067*	-1.38	*0.174*	-0.19	*0.849*	-1.39	*0.170*
VAR vs. SPF												
1	**3.13**	***0.003***	**2.03**	***0.048***	**2.97**	***0.005***	**3.13**	***0.003***	1.07	*0.289*	**3.58**	***0.001***
2	**2.01**	***0.050***	0.25	*0.805*	1.86	*0.069*	**2.01**	***0.050***	-0.05	*0.958*	**2.22**	***0.032***
3	**2.03**	***0.048***	-0.56	*0.581*	1.82	*0.075*	**2.03**	***0.048***	-0.62	*0.539*	1.99	*0.053*
4	1.78	*0.082*	-0.33	*0.743*	1.69	*0.098*	1.78	*0.082*	-0.08	*0.940*	1.41	*0.166*

Notes: A positive value of the HLN statistic informs that the RMSE of A is higher than that of B, where A and B stand for SPF, VAR, BVAR or DSGE. For a positive HLN value, the rejection of the null means that B is superior in forecasting to A. The bold figures indicate the rejection of the null at the 5% significance level.

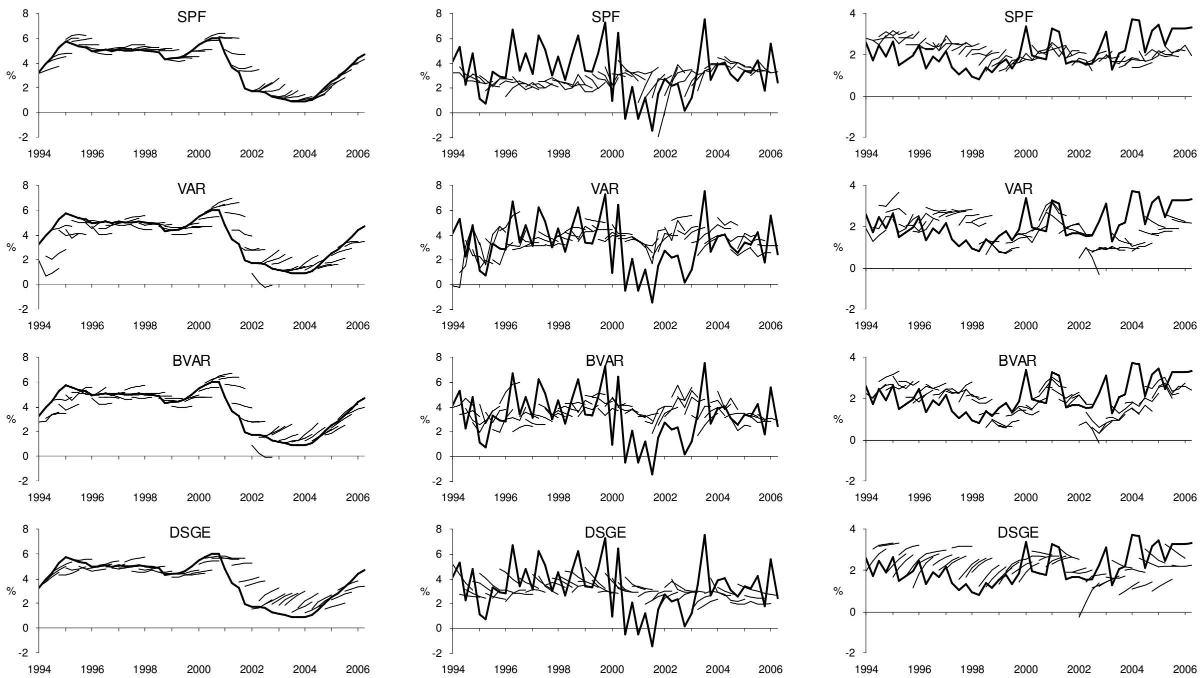
Figure 1. Recursive impulse response functions



Notes: Recursive impulse response functions are calculated at the posterior mean for each quarter of the evaluation sample. The interest rate refers to the three-month Treasury bill yield in annual terms, output denotes the level of GDP and inflation is the annual growth rate of the GDP price index.

Figure 2. Forecasts and the “latest-available” actuals

Three-month TB yield (p.a.)	*GDP (QoQ SAAR)*	*GDP price index (QoQ SAAR)*



Note: Recursive forecasts are calculated at the posterior median for each quarter of the evaluation sample.