Open Econ Rev (2014) 25:171–193
DOI 10.1007/s11079-013-9304-6

RESEARCH ARTICLE

How Useful are DSGE Macroeconomic Models for Forecasting?

Michael Wickens

Published online: 23 January 2014

Abstract A review of the literature shows that forecasts from DSGE models are not more accurate than either times series models or official forecasts, but neither are they any worse. Further, all three types of forecast failed to predict the recession that started in 2007 and continued to forecast poorly even after the recession was known to have begun. The aim of this paper is to investigate why these results occur by examining the structure of the solution of DSGE models and compare this with pure time series models. The main factor seems to be the dynamic structure of DSGE models. Their backward-looking dynamics gives them a similar forecasting structure to time series models and their forward-looking dynamics, which consists of expected values of future exogenous variables, is difficult to forecast accurately. This suggests that DSGE models should not be tested through their forecasting ability.

Keywords DSGE models · Forecasting · VAR models

JEL C5 · E1

1 Introduction

Increasingly, DSGE models are being used by central banks not only for policy analysis, but also for forecasting. A number of papers, including some by central banks, have reported the forecasting performance of DSGE models and compared this with that of pure time series models. Reviewing these results, we find that forecasts from DSGE models are not more accurate than either times series models or official forecasts, but neither are they any worse. We also find that all three types of forecast

M. Wickens (✉)
Cardiff Business School, Cardiff, UK
e-mail: mike.wickens@york.ac.uk

M. Wickens
The University of York, York, UK

failed to predict the recession that started in 2007 and continued to forecast poorly even after the recession was known to have begun.

The aim of this paper is to examine why these results occurred and whether there are any broader lessons to learned. We investigate this is by examining the structure of the solution of DSGE models and compare this with pure time series models. We show that there are three key elements to the answer. First, the solution to a DSGE model consists of both backward-looking and a forward-looking dynamics. The forward-looking terms are expected values of future exogenous variables. Being able to accurately forecast these is crucial to the overall forecasting performance of DSGE models. As they are exogenous, however, we have no theory for them. We might use announcements as our forecasts of these exogenous variables. Alternatively, we might use forecasts obtained from a backward-looking time series model of the exogenous variables. This would be equivalent to including this model for the exogenous variables as part of our solution procedure for the DSGE model. The solution would then be a backward-looking time series model of the full data set. The only way that this solution differs from a conventional pure time series model is that it incorporates the restrictions implied by the structural DSGE model. It follows that the DSGE model would only out-forecast a pure time series model if these restrictions were valid. Our finding that the forecasts from DSGE models are no better (or worse) than those from pure time series forecasts suggests that including the restrictions adds little but, at the same time, they don't grossly violate the data either.

These results also have an interesting implication for testing macroeconomic models. If the accuracy of the model's forecasts depends heavily on those of the current and future exogenous variables whose generating process is not part of our theory, then we may reject a theory when it is not the theory that is at fault. In effect, we have a joint hypothesis: the theory and the data generating processes of the exogenous variables, and not a simple hypothesis consisting just of the theory of interest. It may, therefore, be best to test a DSGE model not through its forecasting properties - i.e. out-of-sample - but solely within-sample where the data are known. Even within-sample the DSGE models will, in general, involve the expected future values of the exogenous variables and, although in-sample the exogenous variables would be known, it would not be correct to use these actual values as the forecast values. Consequently, in-sample testing still involves forming forecasts. The key difference from out-of- sample forecasts is that the lagged endogenous variables in the solution are known.

The paper is set out as follows. In Section 2 we examine the theoretical implications of using DSGE models for forecasting, drawing as examples on the standard neoclassical growth model and a New Keynesian inflation model. In Section 3 we review the forecasting performance of official forecasts and time series models. In Section 4 we compare the forecasting records of several DSGE models—especially those used by official agencies—with pure time series models and official forecasts. Our conclusions are presented in Section 4.

2 Theoretical Issues in Using DSGE Models for Forecasting

We illustrate the issues in using DSGE models for forecasting by considering two well-known models: the standard neoclassical growth model which is the basis of the real business cycle (RBC) model and the New Keynesian model.

2.1 RBC Model

The representative economic agent is assumed to maximize

$$E_t \sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\sigma}}{1-\sigma}$$

subject to

$$\begin{aligned} Y_t &= C_t + I_t \\ Y_t &= A_t K_t^{\alpha} L_t^{1-\alpha} \\ \Delta K_{t+1} &= I_t - \delta K_t \\ L_t &= (1+n)^t L_0 \\ A_t &= (1+\mu)^t Vt \\ \ln V_t &= \upsilon_t, \Delta \upsilon_t = e_t \sim i.i.d\left(0, \omega^2\right) \end{aligned}$$

where Y is output, C is consumption, I is investment, K is capital, L is labour which grows at the rate n, A is technical progress which grows in steady-state at the rate μ, and V is a permanent shock to technical progress.

It can be shown—the full details are in Wickens (2012)—that the log-linearized solution is

$$\begin{aligned} E_t \Delta \ln c_{t+1} &= -\left(\eta + \frac{\delta+\theta}{\sigma}\right)(1-\alpha) E_t \ln k_{t+1} + \left(\eta + \frac{\delta+\theta}{\sigma}\right) \upsilon_t \\ \ln k_{t+1} &= -[\theta + \eta(\sigma-2)] \ln c_t + (1+\theta+\sigma\eta) \ln k_t + \frac{\theta+\delta+(\sigma-\alpha)\eta}{\alpha} \upsilon_t. \end{aligned}$$

This can be written in matrix form as

$$\begin{bmatrix} 1 & 0 \\ \left(\eta + \frac{\delta+\theta}{\sigma}\right)(1-\alpha) & 1 \end{bmatrix} \begin{bmatrix} \ln k_{t+1} \\ E_t \ln c_{t+1} \end{bmatrix} = \begin{bmatrix} 1+\theta+\sigma\eta & -[\theta+\eta(\sigma-2)] \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln k_t \\ \ln c_t \end{bmatrix}$$

$$+ \begin{bmatrix} \dfrac{\theta+\delta(\sigma-\alpha)\eta}{\alpha} \\ \dfrac{\eta+\delta+\theta}{\sigma} \end{bmatrix} \upsilon_t$$

The model can be rewritten as

$$\begin{bmatrix} x_{t+1} \\ E_t y_{t+1} \end{bmatrix} = A \begin{bmatrix} x_t \\ y_t \end{bmatrix} + C z_t \qquad (1)$$

where x_t, in general, is a vector of predetermined variables (stocks), y_t is a vector of "jump" variables (flows or asset prices) and z_t is a vector of exogenous variables (including policy variables) and structural disturbances. Here, $x_t = k_t$, $y_t = c_t$ and $z_t = \upsilon_t$. Denoting the canonical decomposition of A as $A = Q\Gamma Q^{-1}$, the solution has the form of a forward-looking VARX

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = M \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + N\sum\nolimits_{s=0}^{\infty} \Gamma_{yy}^{-s} P E_t z_{t+s} + J z_{t-1} + K\xi_t$$
$$\xi_t = x_t - E_{t-1} x_t.$$

In order to forecast x_t and y_t it is therefore necessary to forecast the exogenous variables z_{t+s}, $s\geq 0$. Hence the forecasting performance of a DSGE model depends on having good forecasts of the exogenous variables. If the exogenous variables are policy variables then we might be able to replace $E_t z_{t+s}$ with credible policy announcements. Or, alternatively, there might be a policy rule, such as a Taylor rule, that is used to determine these variables, in which case these variables are no longer exogenous and we would include the policy rule in the DSGE model. More generally, having no theory for the exogenous variables - otherwise they would be part of the DSGE model - we would need to use a pure time series model to forecast them. Thus testing a DSGE model by its forecasting performance may result in a rejection of the model just because the forecasts of the future exogenous variables are poor.

The solution also shows that ξ_t, shocks to the predetermined endogenous variables, are a further source of forecast error. However, in contrast to forecasting errors to the exogenous variables, these structural shocks are transitory and will disappear at a speed that depends on the internal dynamics of the model. The forecasting performance described later indicates that such shocks are not a major cause of persistent forecast error.

Suppose that the exogenous variables may be represented by the VAR

$$z_{t+1} = R z_t + \varepsilon_{t+1}$$

when $E_t z_{t+s} = R^s z_t$. The solution is then the VARX

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = M \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + H z_t + J z_{t-1} + K\xi_t,$$

Where $H = N(\sum_{s=0}^{\infty} \Gamma_{yy}^{-s} P R^s)$, which is a purely backward-looking model. We can also express the complete data set as the VAR(1)

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} I_x & 0 & -H_x \\ 0 & I_y & -H_y \\ 0 & 0 & I_z \end{bmatrix} \begin{bmatrix} M_{xx} & M_{xy} & J_x \\ M_{yx} & M_{yy} & J_y \\ 0 & 0 & R \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} I_x & 0 & -H_x \\ 0 & I_y & -H_y \\ 0 & 0 & I_z \end{bmatrix} \begin{bmatrix} K_{xx} & K_{xy} & 0 \\ K_{yx} & K_{yy} & 0 \\ 0 & 0 & I_z \end{bmatrix} \begin{bmatrix} \xi_{xt} \\ \xi_{yt} \\ \varepsilon_t \end{bmatrix}$$

or

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = F \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + G \begin{bmatrix} \xi_{xt} \\ \xi_{yt} \\ \varepsilon_t \end{bmatrix}$$

where I_x, I_y and I_z are identity matrices of size equal to the length of the associated vectors x_t, y_t and z_t, and M, K and ξ_t are partitioned conformably with x and y.

This solution has a number of implications. The difference between this solution and a pure time series VAR is that this solution has coefficient restrictions arising from the

DSGE model whereas the pure time series version has no restrictions. This solution also shows that the internal dynamics of a DSGE model may be replicated by an unrestricted VAR. Thus an unrestricted VAR may be expected to provide at least as accurate forecasts as a DSGE model, especially if the coefficient restrictions are incorrect. Following our earlier observations on the problems of testing a DSGE model by its forecasting performance, we note that the alternative is to test the model in-sample by testing these restrictions. This may be carried out using classical statistical inference or using the method of indirect inference in which VAR estimates based on actual data are compared with those based on data simulated from the solution to the DSGE model.

To complete the solution to the RBC model, we note that as $\Delta v_t = e_t$ we have $E_t v_{t+s} = v_t$ (*or* $R=I$), and so the solution is

$$\begin{bmatrix} \Delta \ln k_t \\ \Delta \ln c_t \end{bmatrix} = \begin{bmatrix} \Gamma_{xx} & 0 \\ Q_{yx}\Gamma_{xx} & 0 \end{bmatrix} \begin{bmatrix} \Delta \ln k_{t-1} \\ \Delta \ln c_{t-1} \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} He_t + \begin{bmatrix} C_{xx} \\ Q_{yx}C_{xx} \end{bmatrix} e_{t-1}$$

where $H = Q_{yy}(I - \Gamma_{yy}^{-1})P$. The solution can therefore be written as a VARMA in changes in x_t and y_t

$$\begin{aligned} X_t &= AX_{t-1} + u_t \\ u_t &= Be_t + Ce_{t-1}. \end{aligned}$$

where $X_t = (x_t, y_t)'$ and A, B and C are appropriately redefined.

2.2 New Keynesian Model

The basic New Keynesian model—written in familiar notation and redefining variables accordingly—is

$$\begin{aligned} \pi_t &= \phi + \beta E_t \pi_{t+1} + \gamma x_t + e_{\pi t} \\ x_t &= E_t x_{t+1} - \alpha(R_t - E_t \pi_{t+1} - \theta) + e_{xt} \end{aligned}$$

where π is inflation, x is the output gap and R is the nominal interest rate, $e_{\pi t}$ and e_{xt} are independent, zero mean iid processes and $\varphi = (1-\beta)\pi^*$, where π^* is target inflation.

2.2.1 Discretionary Monetary Policy

In this case R is exogenous. The solution to the model may be obtained by writing it in the form of Eq. (1) and proceeding as before. Alternatively, we can eliminate x_t from the model to obtain the reduced-form dynamic equation for π_t which is

$$\begin{aligned} \pi_t - (1 + \beta + \alpha\gamma)E_t\pi_{t+1} + \beta E_t \pi_{t+2} &= \alpha\gamma\theta + z_t \\ z_t &= -\alpha\gamma R_t + e_{\pi t} + \gamma e_{xt}. \end{aligned}$$

The long-run solution is therefore $\pi_t = R_t - \theta$. The short-run dynamic behaviour of π_t is obtained by rewriting this as

$$\beta L^{-2} f(L)\pi_t = \alpha\gamma\theta + z_t.$$

where L is the lag operator and $f(L)$ defines the characteristic equation

$$\begin{aligned} f(L) &= 1 - \frac{1+\beta+\alpha\gamma}{\beta}L + \frac{1}{\beta}L^2 \\ &= \left(\frac{1}{\beta}\right)(L-\lambda_1)(L-\lambda_2) \\ &= -\left(\frac{1}{\beta}\right)\lambda_1 L\left(1-\lambda_1^{-1}L\right)\left(1-\lambda_2 L^{-1}\right) \\ &= 0. \end{aligned}$$

As f(1)=-$\alpha\gamma/\beta$ <0 there is a saddlepath solution with one root greater than unity (the stable root λ_1) and one root less than unity (the unstable root λ_2). The solution is therefore

$$\begin{aligned} \pi_t &= -\frac{\alpha\gamma\theta}{\beta\lambda_1(1-\lambda_2)} + \frac{1}{\lambda_1}\pi_{t-1} - \frac{1}{\beta\lambda_1}\sum\nolimits_{s=0}^{\infty}\lambda_2^s L^{-s} z_{t-1} \\ &= -\frac{\alpha\gamma\theta}{\beta\lambda_1(1-\lambda_2)} + \frac{1}{\lambda_1}\pi_{t-1} + \frac{\alpha\gamma}{\beta\lambda_1}\left[R_{t-1} + \sum\nolimits_{s=0}^{\infty}\lambda_2^{s+1}E_t R_{t+s}\right] \\ &\quad -\frac{\lambda_2}{\beta\lambda_1}[e_{\pi t} + \gamma e_{xt}] - \frac{1}{\beta\lambda_1}\left[e_{\pi,t-1} + \gamma e_{x,t-1}\right]. \end{aligned}$$

Hence, inflation depends on past, current and future expected interest rates.

The solution for x_t is

$$\begin{aligned} \Delta x_t &= -\frac{\alpha\theta}{\lambda_1}\left(1-\lambda_1-\frac{\gamma}{\beta(1-\lambda_2)}\right) + \frac{1}{\lambda_1}\Delta x_{t-1} \\ &\quad -\frac{\alpha}{\lambda_1}\left(1+\frac{\alpha\gamma}{\beta}\right)R_{t-1} + \alpha\left(1-\frac{\alpha\gamma\lambda_2}{\beta\lambda_1}\right)R_t - \frac{\alpha^2\gamma}{\beta\lambda_1}\sum\nolimits_{s=1}^{\infty}\lambda_2^{s+1}E_t R_{t+s} \\ &\quad +\frac{\alpha\lambda_2}{\beta\lambda_1}[e_{\pi t} + \gamma e_{xt}] + \frac{\alpha}{\beta\lambda_1}e_{\pi,t-1} - \left(1-\frac{\lambda}{\beta\lambda_1}\right)e_{x,t-1} + \frac{1}{\lambda_1}e_{x,t-2.} \end{aligned}$$

Putting the solutions for π_t and x_t together gives a VARMAX(1,2) in π_t and Δx_t with lagged, current and future expected values of the exogenous variable R_t.

Under credible forward guidance the future interest rates in these solutions are announced in advance by the central bank. Without forward guidance they must be determined somehow: for example, by forward rates or other forecasts.

2.2.2 Rules-Based Monetary Policy

We now assume that the monetary authority uses a Taylor Rule and we consider the following model where, for convenience, we have suppressed the intercept terms and we assume that the disturbances are iid variables with zero means

$$\begin{aligned} \pi_t &= \beta E_t\pi_{t+1} + \gamma x_t + e_{\pi t} \\ x_t &= E_t x_{t+1} - \alpha(R_t - E_t\pi_{t+1}) + e_{xt} \\ R_t &= \lambda\pi_t + \eta x_t + e_{Rt} \end{aligned}$$

It can be shown that the solution for $\lambda>1$ is

$$\begin{bmatrix} \pi_t \\ x_t \\ R_t \end{bmatrix} = \frac{1}{1+\alpha(\eta+\lambda\gamma)}\begin{bmatrix} 1+\alpha\eta & \gamma & -\alpha\gamma \\ -\alpha\lambda & 1 & -\alpha \\ -\lambda & \lambda\gamma+\eta & 1 \end{bmatrix}\begin{bmatrix} e_{\pi t} \\ e_{xt} \\ e_{Rt} \end{bmatrix}$$

It may be noted that this does not involve β and that a slightly different solution is obtained if $\lambda<1$. The solution is backward-looking. Once again forecasts of inflation and the output gap from this model should therefore be little different from those from a pure time series model of the three variables.

The solutions of the New Keynesian model show their dependence on the specification of the model. The solution under discretionary monetary policy requires forecasts of the future interest rate while that under rules-based policy is purely backward looking. Forecasting is therefore easier under rules-based policy than under discretion when an entirely different approach is needed. Under a rules-based policy a pure time series model should, in theory, forecast similarly to a DSGE model. But under discretionary policy forecasting from a DSGE model is more problematical, and the forecasts are more likely to differ from those from a pure time series model.

3 The Forecasting Record of Official Forecasts and Time Series Models

In the next two sections we review the findings from various studies, drawing both from published and unpublished material. Rather than contribute further results to this literature, our purpose is to see whether any general conclusions may be drawn from this evidence on the properties of forecasts from DSGE and pure time series models, and whether the theoretical results of the previous section can be of help in explaining these findings.

Most central banks and fiscal authorities use a variety of models in constructing their official forecasts. These include time series models, structural macroeconometric models and small DSGE models. Among the official agencies that use DSGE models as part of their forecasting round are the US Federal Reserve, the Bank of England (the Bank is now also using CGE models), the New Zealand Reserve Bank and the Riksbank. Their published official forecasts are not simply those from a DSGE model but include additional discretionary input not revealed to the public. A very helpful analysis of the forecasting performance of several central banks that is drawn on in our discussion is that of Wieland and Wolters (2013). We also refer to several studies undertaken by the central banks themselves which compare official forecasts with forecasts from time series models and DSGE models. We focus mainly on the forecasting performance surrounding the recent recession that started in 2007.

Before analyzing these forecasts it is important to point out that forecasting was not the original objective in constructing DSGE models. This was primarily to describe the past behaviour of the economy. For example, the aim in the first DSGE models (the real business cycle model) was to see whether they could explain the business cycle. This was an exercise in within-sample analysis not out-of-sample analysis, as in forecasting. The other principal uses of DSGE models are policy analysis in which the effects of altering policy instruments is examined (as unlike standard macroeconometric models and pure time series models, DSGE models are not subject to the Lucas Critique), and within-sample shock decomposition which can be used to explain the causes of past behaviour. The value of both of these uses depends on the appropriateness of the model's specification for the problem at hand.

Recent interesting examples of the latter are the decomposition of the shocks in the Smets and Wouters (2007) model of the United States to investigate the causes of the recent recession, see Wieland and Wolters (2013), and Gali et al. (2012) who have used a modified version of the Smets-Wouters model that incorporates unemployment to analyze why the recovery has been so slow.

3.1 Official Forecasts

3.1.1 US Fed

Edge et al. (2009) have studied the performance of the "Greenbook" forecasts of the US Fed. and compared these with forecasts from the FRB/US macroeconometric model and a simple time series model, an AR(2), over the period 1996–2002. An example of their findings is reported in Table 1 which gives the root mean square forecast errors for growth, inflation and the Fed. funds rate of the Greenbook forecasts and those from the FRB/US model both relative to those from an AR(2).

For growth the AR(2) has the smaller RMSE. For inflation the FRB/US model has a smaller RMSE than the AR(2); the Greenbook forecasts are the best of all 1 year-ahead, but not for one quarter ahead. Perhaps not surprisingly, as the Fed. sets it, the Greenbook forecasts for the Fed funds rate are superior over a 2-year forecasting horizon.

3.1.2 Riksbank

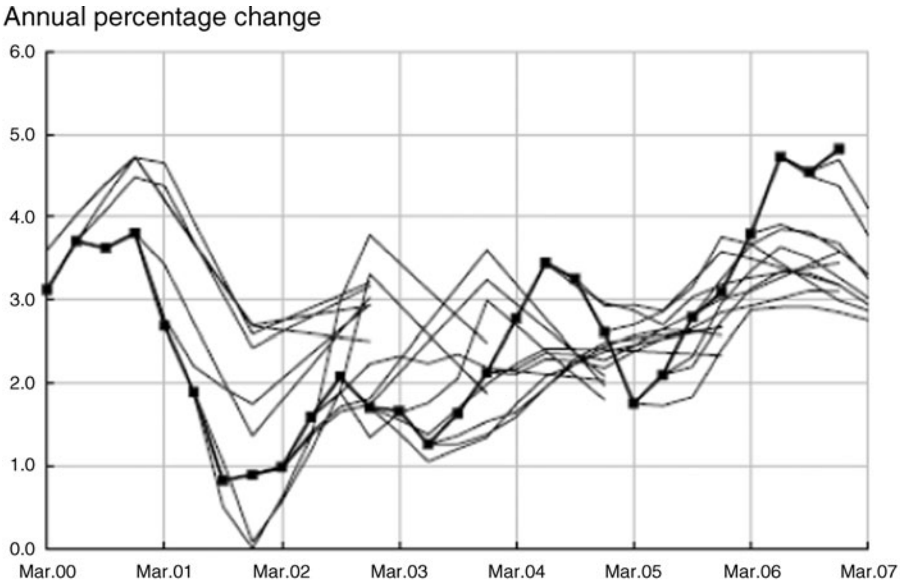
The Riksbank's official forecasts of GDP over the period 2000–2007 are given by Andersson et al. (2007). Their comparison of the official forecasts made each quarter with the outcomes is shown in Fig. 1. In general, the forecasts tended not to pick up fluctuations in GDP. Sweden had an output downturn in 2001. The start of this downturn was clearly missed but, once observed, it was picked up afterwards by the lagged dynamic structure of their forecasting procedure.

3.1.3 Bank of England

Unlike the US Fed., the mandate of the Bank of England is to focus solely on controlling inflation. Its forecasting record on inflation is shown in Fig. 2. The upper panel shows the path of inflation up to August 2011 and the Monetary Policy

Table 1 US: RMSE relative to an AR(2) 1996–2002

Model	1Q	4Q	8Q
Growth			
Greenbook	1.153	1.189	1.104
FRB/US	1.066	1.158	1.138
Inflation			
Greenbook	1.063	0.701	0.934
FRB/US	0.941	0.918	0.865
FFR			
Greenbook	0.743	0.888	0.983
FRB/US	0.743	0.888	0.983



Note. The curve marked with squares is made up of each quarter's first available outcome and the other curves represent the Riksbank's forecasts at each forecasting round.

Fig. 1 The Riksbank's forecasts of GDP 2000–2007

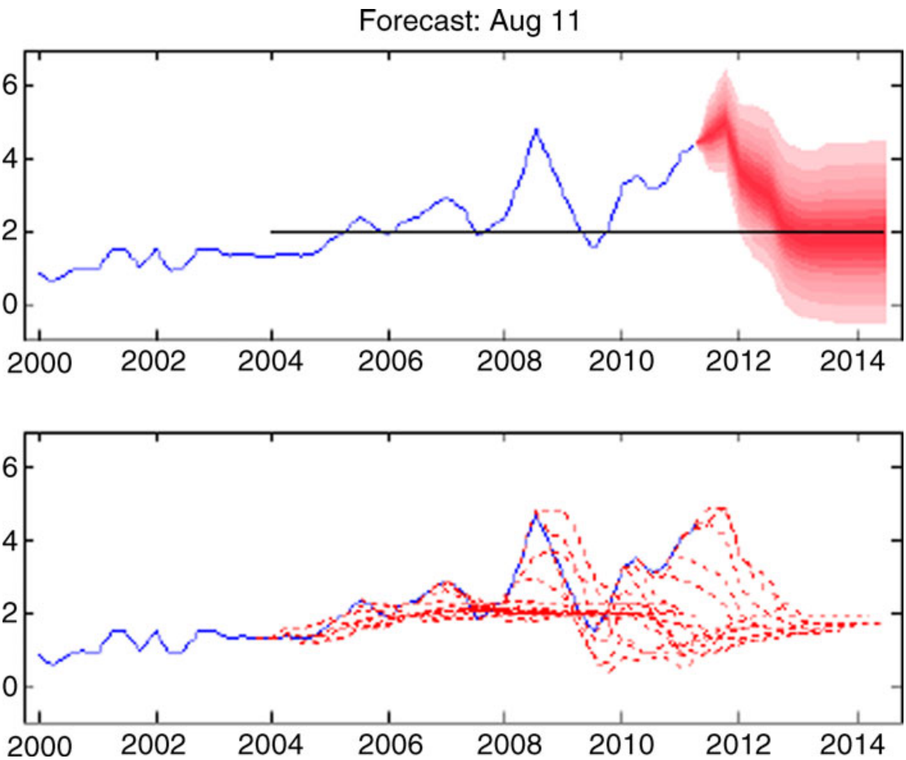


Fig. 2 MPC forecasts of inflation

Committee's assessment afterwards. Starting in 2004, the lower panel compares successive inflation forecasts for the next 3 years of the MPC with actual inflation. Except for 2009 the MPC has consistently under-estimated inflation. Nonetheless, the MPC has kept bank rate at 0.5 % since 2008.

The Bank of England does not appear to have carried out a comparison of the official forecasts of inflation with forecasts based on simple time series models over the period that includes the recent recession. Groen and Price (2009) compared the Bank of England's forecasts of inflation with forecasts from pure time series models for the period 1997Q3-2006Q2, which is prior to the financial crisis, and concluded that the two were similar. For the period 2008Q1-2010Q4, Barnett et al. (2012) have examined the forecasting performance for inflation of 24 time series models with time-varying parameters to capture structural change. This additional feature might be expected to improve the time series forecasts. Figure 3a shows one step ahead forecasts and Fig. 3b shows 4 step ahead forecasts. The bold line with the crosses is actual inflation. Again, none of these models accurately forecast inflation. Most fail to pick up the turning points and strongly mean revert. The three models that do pick up turning points also incorporate additional information and so are not true time series models. If anything, however, their forecasts are even worse as they also hugely over-estimate the fluctuations in inflation.

3.2 Time Series Forecasts

Following Nelson (1972, 1981), who found that a univariate ARIMA model forecast better than the FMP macroeconometric model, it has been widely accepted that pure

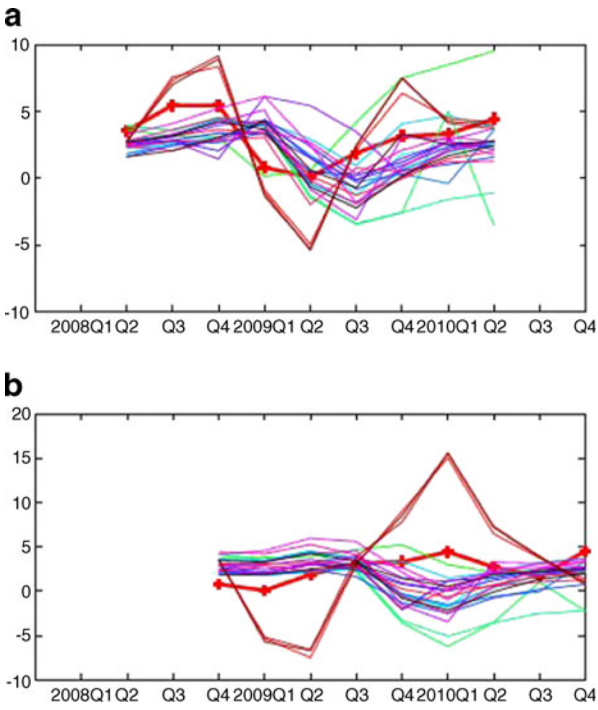


Fig. 3 **a** One step ahead forecasts of UK inflation **b** Four step ahead forecasts of UK inflation

time series models often provide better forecasts than macroeconometric models. Multivariate time series forecasts are often based on a VAR. We note that omitting a variable from a VAR just adds to length of the lag structure of the VAR and creates a moving average error structure, making the model a VARMA. Christofferson and Diebold (1997) have shown that over long horizons an unrestricted VAR forecasts just as well as a VAR that takes account of any cointegration present in the VAR. As many terms in a VAR are insignificant, and including insignificant terms tends to worsen forecasting performance, Doan and Litterman (1984) found that using a Bayesian VAR, which shrank poorly determined coefficients via the "Minnesota prior", improved VAR forecasts.

Wieland and Wolters (2013) compare the forecasting performance of an AR(4) and a BVAR(4). Their results for the US are shown in Fig. 4. The growth forecasts consistently return growth to its long-run level and so miss the depth of the recession; the inflation forecasts flatline as do the unemployment forecasts; together the growth and inflation forecasts explain the over-estimate of interest rate forecasts.

Figure 5 compares the official forecasts of the Riksbank with those from an AR for the 2001 recession in Sweden. Both miss the recession, the AR forecasts (the dotted lines) more than the official forecasts. Once more, therefore, the VAR forecasts are flatlining more than the data and more than the official forecasts which are also relatively flat.

Castle et al. (2011), Clements and Hendry (1998) and Hendry (2005) have found that the most useful way of improving forecasts is to "robustify" them against structural breaks by taking account of shifts in the mean. This is especially helpful if there is cointegration. For the error correction model

$$\Delta y_t = \beta \Delta x_t - (1-\alpha)\left[y_{t-1} - \mu - \theta x_{t-1}\right] + e_t$$

this entails robustifying the model through an intercept adjustment to μ which also affects the rate of growth. In this way permanent deviations from the new path of y_t are

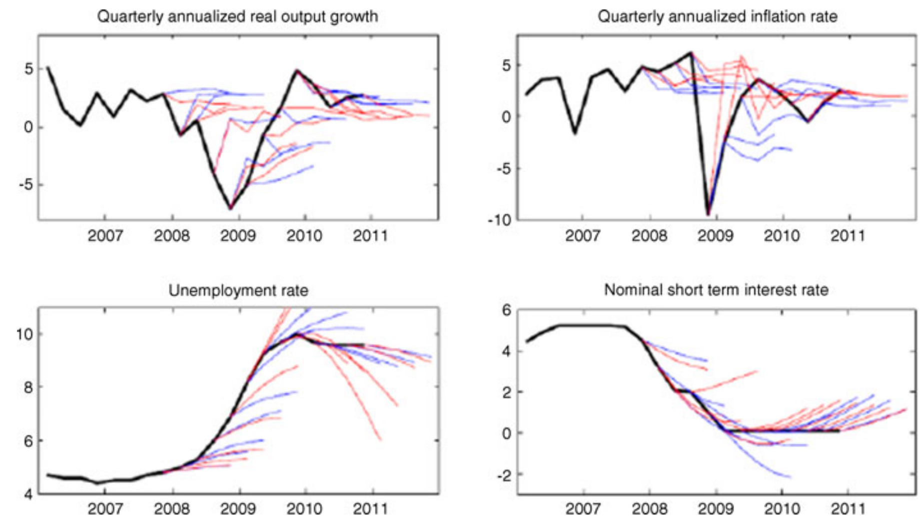
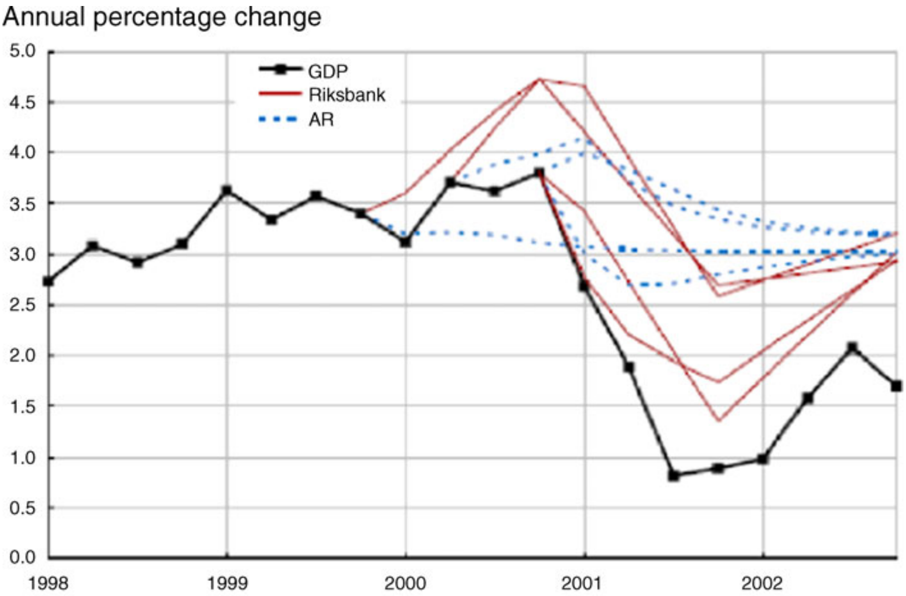


Fig. 4 AR and BVAR forecasts of US growth *Blue line*: BVAR(4) forecast; *red line*: AR(4) forecast



Note. The curve marked with squares is made up of each quarter's first available outcome and the other solid curves represent the Riksbank's forecast at each forecasting round. Dotted lines represent the corresponding AR forecasts.

Fig. 5 Comparison of the Riksbank's forecasts with an AR model

corrected. Rebasing forecasts by using the latest value of the lagged variables also helps.

3.3 Data Problems

The accuracy of forecasts also depend on using accurate data. Macroeconomic data is usually estimated initially and these estimates are then revised. Forecasters seeking timely forecasts therefore use data that is being changed. For example, over the period 1998q2-2004q1, 9 of the 19 official preliminary estimates of UK growth were more than one standard deviation away from the final estimate, while only 4 out of 18 of the previous quarter's growth rates were more than one standard deviation away from the final estimate. The assumption of no change in the growth rate gave estimates that were twice as accurate as the worst interim estimate. In other words, over this period, using last period's growth rate gave the best estimate of the current growth rate, see House of Lords (2004).

Often the current value of a variable may exist and so must be estimated. This is known as the problem of nowcasting. Time series models are usually used to estimate these current values.

4 The Forecasting Record of DSGE Models

The focus of our attention is on the forecasting record of DSGE models used by the central banks in New Zealand and the United States. We also examine forecasts made by the IMF and those using the Smets-Wouters model. We draw upon Wieland and Wolters (2013) for some of this information.

4.1 Reserve Bank of New Zealand

The RBNZ have estimated a following small New Keynesian open-economy DSGE model for New Zealand over the period 1990Q1-2005Q4. See Lees et al. (2007) for details. Table 2 compares the mean square forecast errors (MSFE) of the DSGE model and those from and an unrestricted VAR (UNR), a Bayesian VAR with a Minnesota prior (MVAR) and a VAR constructed from data simulated from the DSGE model (DVAR) with the official forecasts over the period 1998q4-2003q3. The forecasts are made in real time. Gains over the official forecasts are positive and losses are negative.

The results show that the MVAR forecasts are best across all forecast horizons for all variables. The DSGE model was considerably worse for growth and inflation but performed better for the exchange rate than the MVAR at longer horizons; its interest rate forecasts were only marginally worse.

4.2 IMF New Keynesian Model of the US

Wieland and Wolters (2013) report the forecasts of the IMF's small model for the United States for the period 2008–2011. It is basically a New Keynesian model with a

Table 2 Real time forecasts by the RBNZ 1998q4-2003q3

Percentage gain (loss) in MSFE over the real-time RBNZ forecasts								
h	UNR	MVAR	DVAR	DSGE	UNR	MVAR	DVAR	DSGE
GDP growth					Inflation			
1	−33.7	5.6	19.6	6.4	−19.0	0.9	−10.0	−40.1
2	12.8	23.0	36.0	18.3	−49.2	−4.7	−26.3	−83.4
2	17.0	30.2	29.1	18.3	−43.1	4.0	−9.0	−75.0
4	8.5	42.5*	29.0	28.1	−17.2	15.5	2.1	−78.4
5	−10.3	43.2*	3.7	25.2*	−9.7	21.6*	10.1	−86.3
6	−22.2	42.2	−16.3	17.0	4.0	23.9*	13.8	−110.2
7	−27.4	47.6	−26.3	2.2	10.0	24.9	15.7	−106.8
8	−20.2	43.0*	−24.4	−20.5	12.5	22.9	13.3	−114.3
Interest rate					Exchange rate			
1	−289.5	−74.9	−165.9	−144.7	−24.9	4.7	−10.0	−3.8
2	−274.2	−21.0	−47.1	−64.0	−32.9	−0.6	−22.0	−21.8
3	−119.0	5.7	13.2	−7.6	−14.8	1.4	−0.4	1.5
4	−31.1	33.8*	14.9	24.8	−14.4	0.7	4.6	15.9
5	−23.6	43.3**	23.5	26.0	−21.0	−9.9	−9.6	4.4
6	−47.4	39.6*	32.3	29.8	−21.8	−18.9	−24.2	−8.2
7	−86.8	33.8	33.6	17.7	−19.6	−26.7	−23.3	−6.7
8	−139.6	23.9	33.3	20.5	−22.6	−38.7	−30.3	−10.4

Taylor rule and, in addition, an unemployment equation. The forecasts and the outcomes for quarterly (but annualized) and annual rates of growth—the rate of change in the output gap—are shown in Fig. 6. Clearly the DSGE model misses the recession as it is attempting to eliminate the output gap

4.3 Riksbank

Adolfson et al. (2007) compare the forecasting performance of a small open-economy DSGE model for Sweden with those from a BVAR and the official Riksbank forecasts over the period 1999q1-2005q4. The forecasts of GDP growth, inflation and the rate of interest in Figs. 7, 8 and 9 are updated each period. The top panel is the official forecast, the second panel is the DSGE model forecasts and the bottom panel is the BVAR forecasts.

All three forecasts miss the growth downturn starting in 2000 and are generally too optimistic after the downturn. The DSGE and BVAR forecasts also miss the upturn that preceded the downturn. The forecasts of the DGSE and BVAR models are similar throughout, as they are for inflation, but less similar for the interest rate. Prior to 2004 the official forecasts tend to under-predict inflation whereas the DSGE and BVAR forecasts tend to over-estimate inflation. Both models tend to miss turning points in the rate of inflation. The official forecasts consistently over-estimate the interest rate throughout as does the DSGE model, but the BVAR, by flatlining, misses increases in the interest rate on the downside and decreases on the upside.

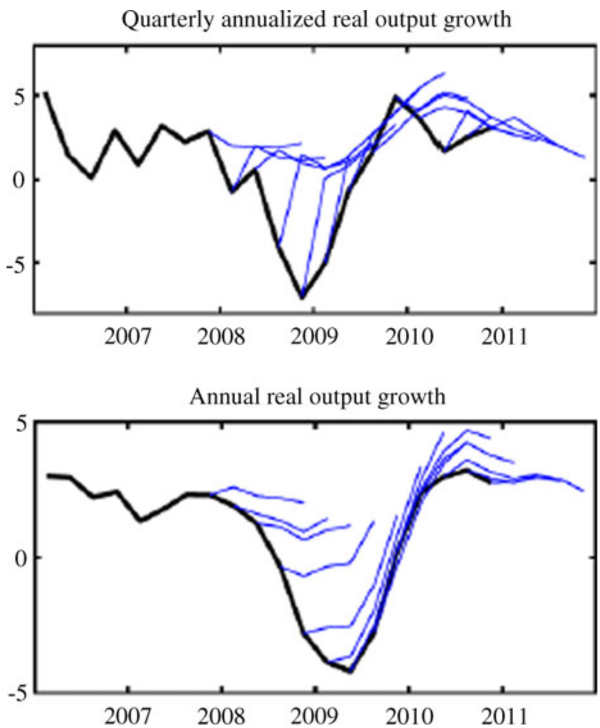


Fig. 6 Forecasts of US quarterly and annual growth based on the IMF's NK model

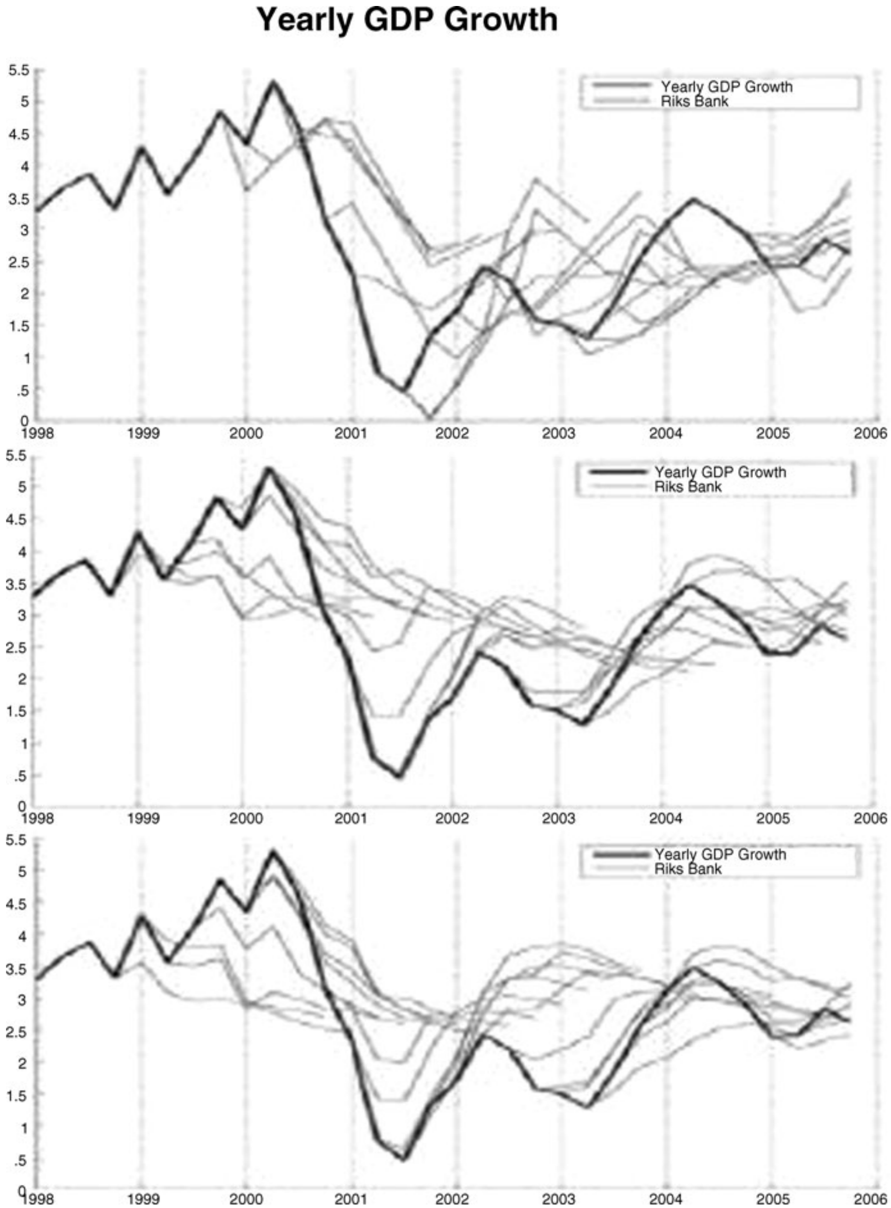


Fig. 7 GDP growth

4.4 US Federal Reserve

Edge et al. (2009) report forecasts of the US economy for the period 1996.9–2004.11, comparing its RMSE with forecasts from an AR(2). The results for US growth and inflation are reported in Table 3 together with forecasts from a VAR and a BVAR over horizons from one to eight quarters. Table 4 reports the mean biases. Tables 5 and 6 report the RMSEs and mean biases for the Fed funds rate.

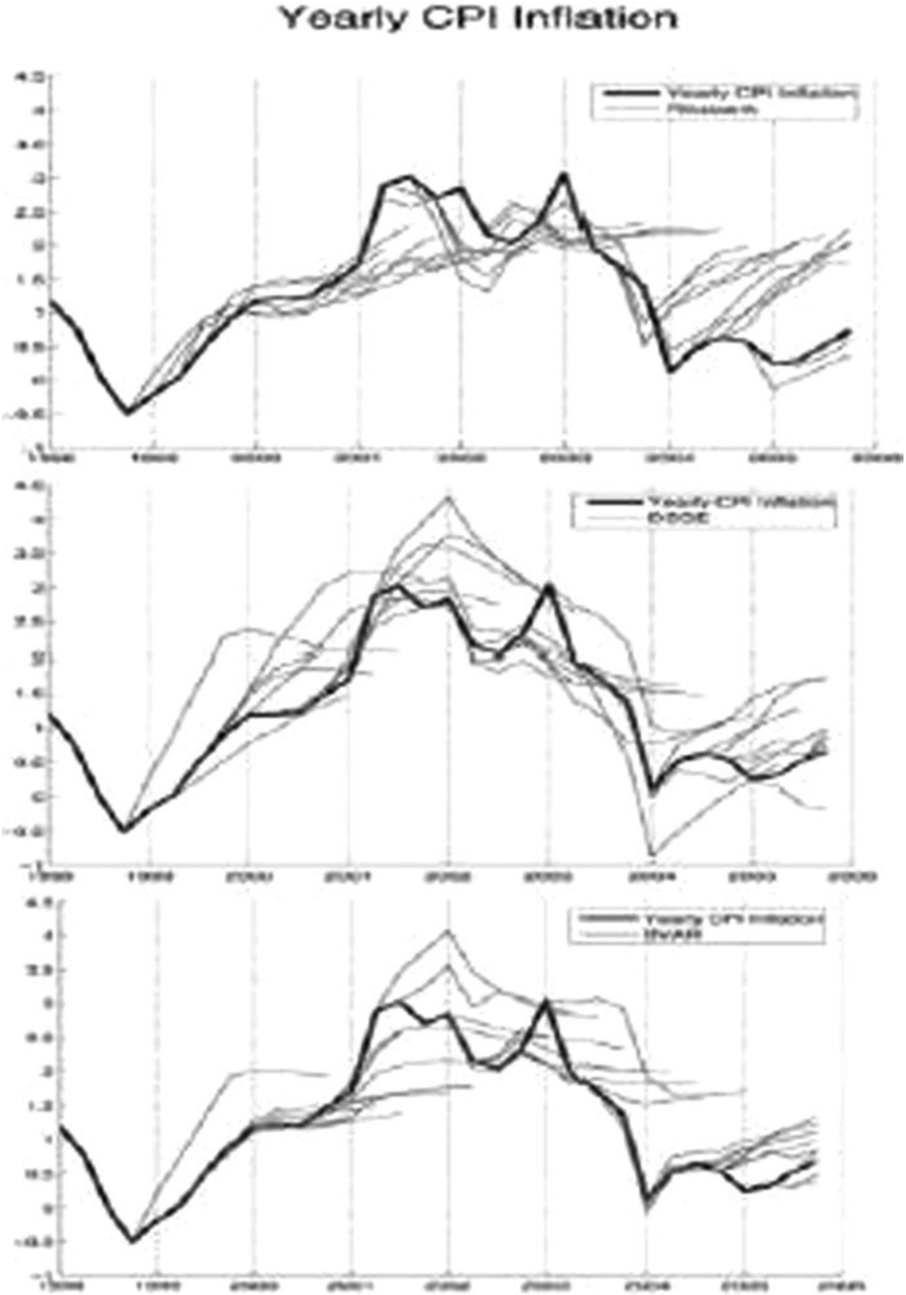


Fig. 8 CPI inflation

For the forecasts for growth the DSGE model has a slightly smaller RMSE than the AR(2) model, and the AR(2) model has slightly smaller RMSEs than either the VAR and BVAR. The DSGE model also has smaller mean bias except at the 2-year horizon. Again, except at a 2-year horizon, the AR(2) forecasts inflation better than all three

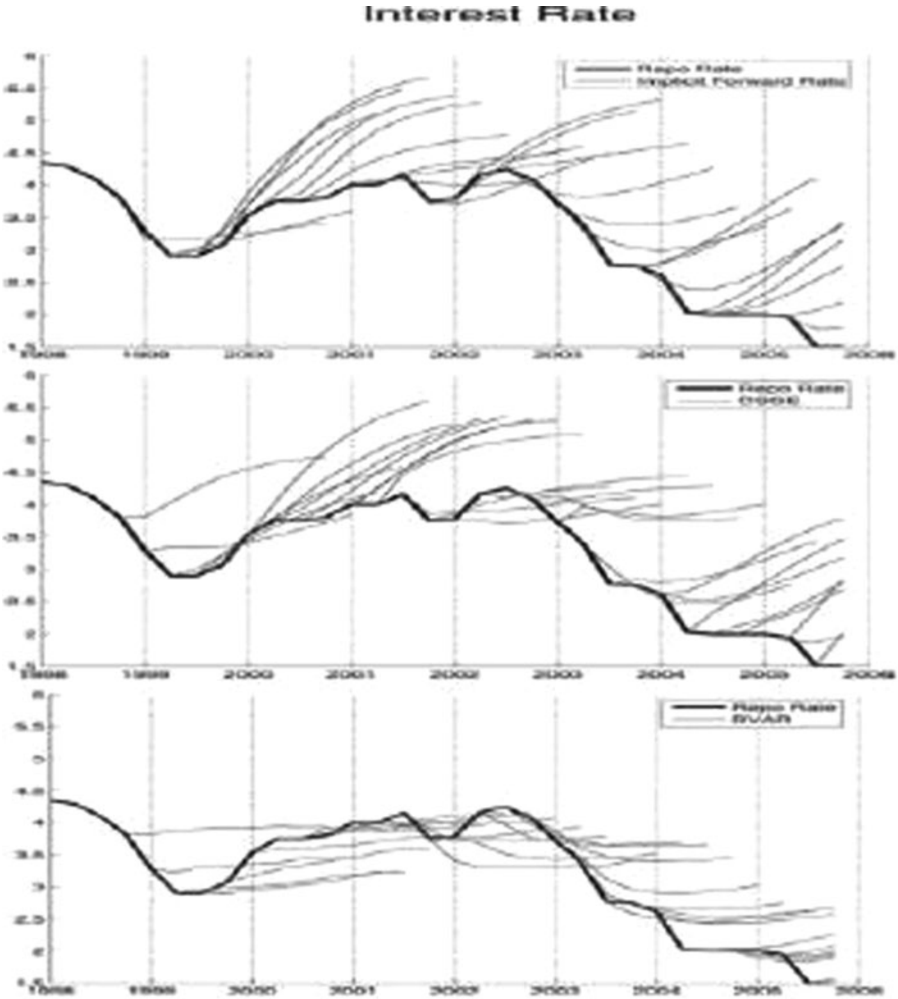


Fig. 9 Interest rate

models. The mean bias is, however, smallest for the BVAR model. In terms of the RMSE, apart from the 2-year horizon, the DSGE model forecasts the Fed. funds rate worst and has the worst mean biases.

4.5 Smets-Wouters US Model

The DSGE model of the US of Smets and Wouters (2007) raised the bar on the specification of DSGE models by introducing New Keynesian features related to price and wage dynamics and has set the standard for many of the official DSGE models that have been developed since. It is also widely used as a test-bed for DSGE models. We consider two examples of such studies that examine its forecasting properties. Although the dynamic specification of the Smets-Wouters model is much richer than typical RBC models, which should help its forecasting ability, it has some unsatisfactory features.

Table 3 Relative RMSEs for forecasts of growth and inflation from various US Fed. Models

Model	1Q	2Q	3Q	4Q	8Q
Real GDP growth					
AR(2)	0.470	0.521	0.497	0.547	0.551
Relative RMSE					
DSGE/Edo	0.953	0.920	0.948	0.917	0.976
VAR(1)	1.125	1.052	1.121	1.018	1.073
BVAR(2)	1.096	1.031	1.071	1.002	1.078
GDP price inflation					
AR(2)	0.276	0.258	0.243	0.281	0.288
Relative RMSE					
DSGE B	1.064	1.065	1.061	1.038	0.925
VAR(1)	1.139	1.175	1.222	1.231	1.258
BVAR(2)	1.088	1.137	1.150	1.192	1.177

For example, Le et al. (2011) find that the model is rejected against more general alternative specifications of price and wage dynamics. Another unsatisfactory element of the model is that many of the structural disturbances are highly autocorrelated also suggesting dynamic misspecification.

Wieland and Wolters (2013) report forecasts for the US from the Smets-Wouters model for the period 2008–2011. These are shown in Fig. 10. The same pattern emerges as before: the forecasts miss the recession and then the model tries to return the economy to its steady-state too quickly.

Gurkaynak, Kisacikoglu and Rossi (2013) compare the out-of-sample forecasting ability of the Smets-Wouters model those of several reduced-form time series models using data up to 2007. They conclude that simple AR models forecast output growth best over short horizons and DSGE models are more accurate at long horizons, and that this is reversed when forecasting inflation. Another finding is that low-dimension AR

Table 4 Mean biases for forecasts of growth and inflation from various US Fed. Models

Model	1Q	2Q	3Q	4Q	8Q
Real GDP growth					
AR(2)	−0.037	−0.116	−0.056	−0.140	0.175
DSGE/Edo	0.024	−0.050	−0.011	−0.104	−0.165
VAR(1)	−0.074	−0.091	−0.021	−0.088	−0.122
BVAR(2)	−0.040	−0.088	−0.021	−0.094	−0.124
GDP price inflation					
AR(2)	0.075	0.087	0.104	0.126	0.092
DSGE/Edo	0.072	0.094	0.108	0.125	0.079
VAR(1)	0.044	0.023	0.024	0.026	−0.038
BVAR(2)	0.032	0.020	0.016	0.021	−0.052

Table 5 Relative RMSEs for forecasts of the Fed. funds rate from various US Fed. Models

Nominal funds rate					
AR(2)	0.170	0.260	0.339	0.426	0.618
Relative RMSE					
DSGE/Edo	1.224	1.208	1.153	1.092	0.941
VAR(1)	0.986	1.068	1.116	1.144	1.248
BVAR(2)	1.006	1.054	1.084	1.104	1.058

and VAR models forecast more accurately than large-scale Bayesian VAR models. These are very similar to the findings of the studies reported above.

4.6 Various Unofficial DSGE Models for the US 2000–2002

We complete our examination of the forecasting performance of DSGE models by examining the record of a number of unofficial DSGE models over the period 2000–2002. These results are from Wolters (2013) and are for US growth, inflation and the Fed. funds rate. The forecasts from these models together with their fan-charts, the Greenbook forecasts (white line) and the outcomes are shown in Fig. 11.

The results show that none of the forecasts are close the actual outcomes and no model consistently out-forecasts another for all variables, for all sub-periods or for all forecast horizons. The Greenbook forecasts are consistently better for output growth but the models give better forecasts of inflation than the Greenbook.

5 Explaining the Evidence

The forecasting record of all of the models, whether a DSGE model or a time series model, show a consistent pattern: having a tendency to flatline, they all fail to anticipate turning points, especially recessions, and they all try to return the economy to the steady-state too fast. The official forecasts have the same defects. In broad terms, DSGE models do not appear to forecast less accurately than pure time series models, but neither do they forecast any better. There is, however, some evidence to suggest that for some variables (eg. output) DSGE models forecast long-run outcomes better than pure time series models but produce worse short-run forecasts, while for other variables (inflation) this is reversed. The aim of this paper is to examine why these results occur.

Table 6 Mean biases for forecasts of the Fed. funds rate from various US Fed. MODELS

Nominal rate fund					
AR(2)	0.050	0.089	0.129	0.178	0.309
DSGE/Edo	0.086	0.152	0.202	0.251	0.336
VAR(1)	0.048	0.087	0.127	0.177	0.282
BVAR(2)	0.057	0.099	0.138	0.185	0.255

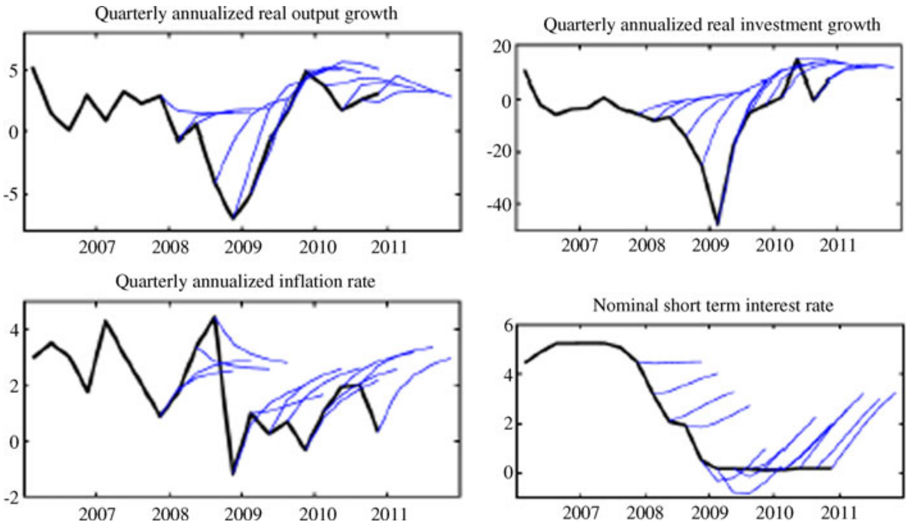
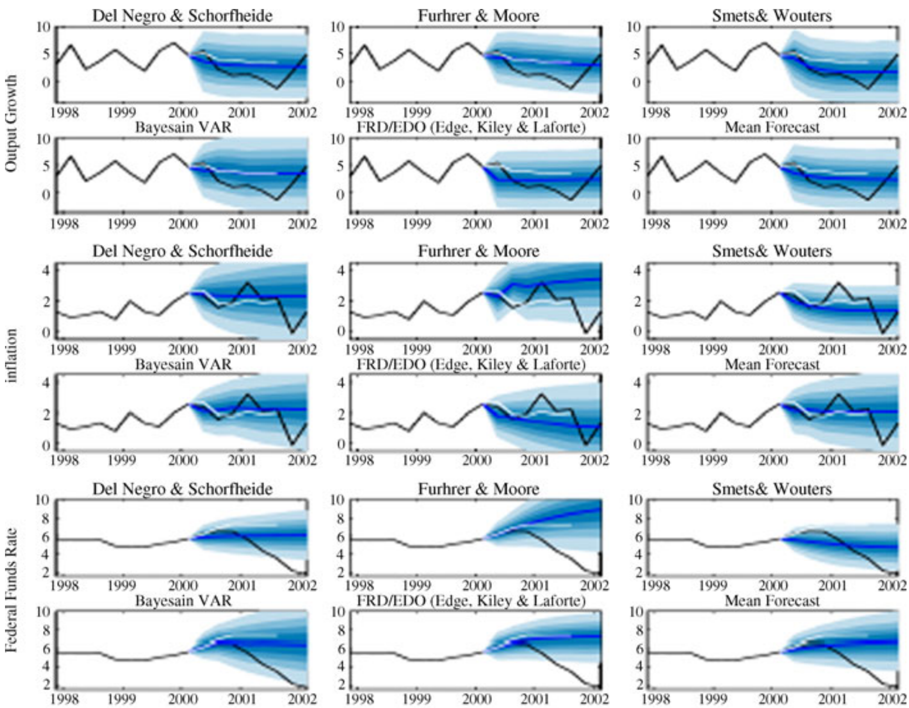


Fig. 10 Forecasts of the US economy from the Smets-Wouters model

We have suggested that one explanation for forecast inaccuracy by DSGE models may lie in the nature of the solution to a DSGE model. The solution has three features of potential relevance. First, in general, the solution involves the presence of expected



Note: the black line show real-time data until the forecast start and revised data afterwards: the shaded areas show 90% 70%, 50% and 30% confidence bands; the line in the middle of the confidence bands shows the mean forecast for horizons 0 to 7; the short white line the Greenbook forecast for horizons 0 to 5. Mean Forecast is the averge of the four model forecasts.

Fig. 11 Forecasts for the US from unofficial DSGE models and the Greenbook forecasts future values of exogenous variables—the forward-looking dynamics—and so the forecasting performance of a DSGE model may well depend critically on the accuracy of the forecasts of future exogenous variables. As they are exogenous, we have no theory about them. If they are policy variables, then we might have official announcements about their future values as in forward guidance. If these announcements are credible, we might use these as our forecasts. The alternative in both cases is probably to use time series models to forecast future exogenous variables. The findings of the studies we have reported have, however, shown that such forecasts are unlikely to be very accurate, especially when there are sharp changes in the exogenous variables. An important reason why the DSGE models miss turning points is that the exogenous variables are not well-forecasted. The failure to forecast changes in the exogenous variables probably accounts for the tendency of DSGE models to flatline and give persistent forecast errors. Not surprisingly, when the economy is growing steadily, and there are no turning points, it is much easier to forecast the exogenous variables. All of these forecasts then perform much better.

Related conclusions about the role of exogenous variables in macroeconometric modelling were made some years ago by Adelman and Adelman (1959). They asked what features of a macroeconometric model could cause business cycle-like behaviour. They found that the internal dynamics of the model produced cycles of far too small an amplitude, and that the disturbances were too small to produce plausible cycles. They therefore concluded that business cycles must be caused by fluctuations in the exogenous variables. Howrey (1971) came to a similar conclusion.

A second feature of the solutions of DSGE models is their lag structure—the backward-looking dynamics. These are likely to be similar to those estimated from a time series model in the same variables which is why DSGE models forecast as badly as time series models, and vice-versa. The tendency of DSGE models to forecast a faster return to the long-run following a recession than has actually occurred may be partly due to misspecified dynamics, possibly because the lags are longer than specified, or because the estimates of the coefficients on the lags both in the structural model and in the time series model used to forecast the exogenous variables are biased downwards. For example, we know from Shaman and Stine (1988) that for an AR(1) model with an intercept and an autoregressive parameter ρ the OLS estimator has a downward bias to the estimate of ρ of $-(1+3\rho)/T$ where T is the sample size. This implies a faster mean reversion than is correct. Incorrect restrictions imposed by the DSGE model could also cause such biases. Shocks to the structural equations do not appear to be a cause of persistent forecast error as they tend to be dominated by the backward-looking dynamics observed in DSGE models. An exception is when structural disturbances are highly serially correlated as in the Smets-Wouters model.

Whether or not variables are better forecast from a DSGE model than a pure time series model in the long or short run will depend on the persistence of these variables. A time series model may be expected to produce accurate forecasts of variables that are highly persistent but not of variables that are close to being serially uncorrelated. A DSGE model may be expected to perform similarly except when it is difficult to obtain accurate forecasts of key exogenous variables. For example, long-run forecasting tends to be more accurate when the data are strongly trending or are constant and there are no breaks in the processes generating the exogenous variables. Long-run forecasting may also be accurate when shocks are temporary and serially uncorrelated as the shocks average out in the long run. Conversely, temporary shocks may be expected to reduce the accuracy of short-run forecasts.

These conclusions have potentially important implications for how to test macroeconomic models. Models are usually tested from their predictions. But if the accuracy of a DSGE model's forecasts depends heavily on the accuracy of the forecasts of the current and future exogenous variables, whose generating process is not part of our theory, then we may reject a theory when it is not the theory that is at fault. In effect, we have a joint hypothesis: the theory and the exogenous generating process, and not a simple hypothesis consisting just of the theory of interest. It may, therefore, be best to test a macroeconomic model not out-of-sample through its forecasting properties, but solely within-sample. Even within-sample the DSGE models will, in general, involve the expected future values of the exogenous variables and, although in-sample the exogenous variables would be known, it would not be correct to use these actual values as the forecast values. Consequently, in-sample testing still involves forming forecasts. The key difference from out-of- sample forecasts is that the lagged endogenous variables in the solution are known.

References

Adelman FL, Adelman I (1959) The dynamic properties of the Klein-Goldberger model. Econometrica 27: 596–625

Adolfson M, Andersson MK, Linde J, Villani M, Vredin A (2007) Modern forecasting models in action: Improving analyses at central banks. Int J Central Banking 3:111–144

Andersson MK, Karlsson G, Svensson J (2007) The Riksbank's forecasting performance. Sveriges Riksbank WP218

Barnett A, Mumtaz H, Theodoridis K (2012) Forecasting UK GDP growth, inflation and interest rates under structural change: a comparison of models with time-varying parameters. Bank of England Working Paper No. 450

Castle JL, Fawcett NWP, Hendry DF (2011) Forecasting with equilibrium-correction models during structural breaks. J Econometrics 158:25–36

Christofferson PF, Diebold FX (1997) Cointegration and long-horizon forecasting. NBER WP217

Clements MP, Hendry DF (1998) Intercept corrections and structural change. J App Econom 11:475–494

Doan T, Litterman R (1984) Forecasting and conditional rejection using realistic prior distributions. Econometric Reviews 3(1):1–100

Edge RM, Kiley MT, Laforte J-P (2009) A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model. Finance and Economics Discussion Series, Federal Reserve Board

Gali J, Smets F, Wouters R (2012) Slow recoveries: a structural interpretation. J Money Credit Banking 44:9–30

Groen JJJ, Kapetanios G, Price S (2009) A real time evaluation of Bank of England forecasts of inflation and growth. Int J Forecasting 25:74–80

Gurkaynak RS, Kisacikoglu B, Rossi B (2013) Do DSGE models forecast more accurately out-of-sample than VAR models? CEPR Discussion Paper DP9576

Hendry DF (2005) Robustifying forecasts from equilibrium-correction models. Journal of Econometrics 135: 399–426

House of Lords (2004) Report of the Select Committee of Economic Affairs on "The State of the Economy and Economic Policy, Long Term and Short Term"

Howrey EP (1971) Stochastic properties of the Klein-Goldberger model. Econometrica 39:73–87

Lees K, Matheson T, Smith C (2007) Open economy DSGE-VAR forecasting and policy analysis: Head to head with the RBNZ published forecasts. RBNZ DP2007/01

Le VPM, Meenagh D, Minford APL, Wickens MR (2011) How much nominal rigidity is there in the US economy? Testing a new Keynesian DSGE model using indirect inference. Journal of Economic and Dynamic Control 35(12):2078–2104

Nelson CR (1972) The prediction performance of the FRB-MIT-Penn model of the US economy. American Economic Review 62:902–917

Nelson CR (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attntion to measurement of the business cycle. Journal of Monetary Economics 7:151–171

Shaman P, Stine RA (1988) The bias of autoregressive coe¢ cient estimators. Journal of the American Statistical Association 83:842–848

Smets F, Wouters R (2007) Shocks and frictions in US business cycles: a bayesian DSGE approach. American Economic Review 97:586–606

Wieland V, Wolters MH (2013) Forecasting and policy making. In: Elliott G, Timmerman A (eds.) Handbook of economic forecasting, Vol. 2. Elsevier, 5:238–237

Wickens MR (2012) Macroeconomic theory: a dynamic general equilibrium approach (second edition). Princeton University Press, Princeton

Wolters M (2013) Evaluating point and density forecasts of DSGE models. J Appl Econom (forthcoming)