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

Michał Rubaszek[1] Paweł Skrzypczyński[2]

[1]National Bank of Poland & Warsaw School of Economics

[2]National Bank of Poland

1st Macroeconomic Forecasting Conference, March 2009

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

Outline

Table of Contents

Motivation and Contribution

Motivation:

- Widespread use of DSGE models for quantitative policy analysis in macroeconomics
- Increasing role of DSGE models in forecasting
- Scarce documentation of DSGE models out-of-sample performance

Contribution:

- Extend knowledge about the forecasting properties of small-scale DSGE models
- The first study that compares forecasts from a DSGE model with those from the SPF in a real-time environment

The form of the analysis

We use four methods:

- Dynamic Stochastic General Equilibrium model (DSGE)
- Vector Autoregression model (VAR)
- Bayesian VAR (BVAR)
- Survey of Professional Forecasters (SPF)

to compare up to 4-quarter real time forecasts for key US variables:

- GDP growth rate
- GDP price index inflation
- Three-month Treasury bill rate

Table of Contents

Literature review

An analysis the most similar to ours

Edge, Kiley & Laforte (2006)

- Use of real-time data
- Comparison of forecasts from a random walk, VARs and a richly-specified DSGE model to Federal Reserve (FRB) staff projections
- FRB staff was found to be the best inflation forecaster
- DSGE, VAR and BVAR models dominate FRB staff in forecasting the GDP growth rate
- The forecast evaluation period (1996-2000) is short and thereby the results might not be representative

Literature review

An article comparing the quality of forecasts from DSGE and VAR models, but not using real-time data

Smets & Wouters (2004)

- A medium-scale DSGE model for the euro area (introduced by Smets & Wouters, 2003) is compared to VARs
- Comparison on the basis of the RMSEs for 7 macroeconomic variables
- DSGE found to moderately outperform VARs

Literature review

An article comparing the quality of real-time forecasts from the SPF to those from VARs or ARIMAs

Clark & McCracken (2006)

- Extensive study comparing real-time forecasts of the SPF, FRB staff and numerous atheoretical models
- Forecast accuracy of VARs and univariate models is roughly the same
- Forecast errors from VARs are significantly higher than those from the SPF, especially for one-quarter forecast horizon
- The SPF appears 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.

Literature review

A general picture

- In a real-time context the SPF can better forecast the economy than atheoretical models such as VARs or ARIMAs
- If the forecast performance is evaluated on the basis of the latest-available data, DSGE models behave comparable or even superior to VARs

A question and the purpose of our article

- Can a DSGE model beat the SPF in forecasting the U.S. economy, if the real-time data are used

Table of Contents

DSGE model

- Standard small-scale New-Keynesian model for output, inflation and interest rates augmented for:
  - consumption habits
  - inflation persistence
  - interest rate smoothing
- The system is put in motion by three shocks: supply, demand and monetary
- Observable variables are: GDP growth rate, GDP price index inflation and 3-month Treasury bill rate
- Application of Bayesian inference to estimate the structural parameters

(B)VAR model

- Infinite order VARs - unconstrained representations of DSGE models
- Good benchmarks for evaluating the performance of DSGE models
- We analyze (B)VAR models for the vector of the three U.S. variables (interest rates, inflation and output)
- The optimal lag order of the VAR model chosen using the final prediction error criterion
- For the BVAR model - the maximum lag set at 4 quarters with the Minnesota prior distribution

Survey of Professional Forecasters

- The SPF - the oldest quarterly survey of macroeconomic forecasts in the US (published on the Philadelphia Fed website)
- Carried out in regular three-month intervals and concerns dozens of macroeconomic variables (among them output, inflation and interest rates)
- Forecasts are formulated up to four quarters ahead
- Information available to the SPF concerning NA data is the same as the real-time data we use to estimate DSGE and VAR models[2]

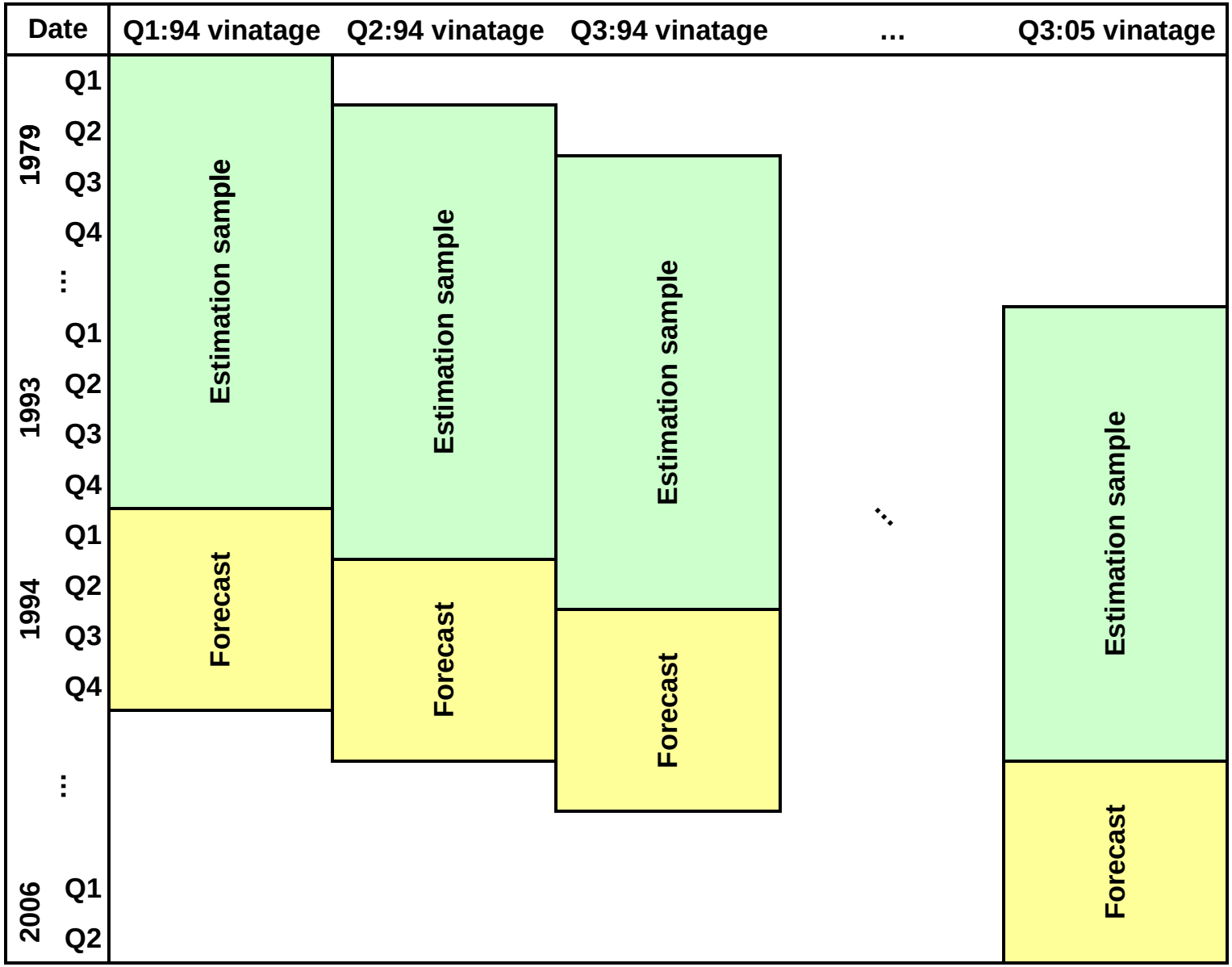
[2] However, the professional forecasters have some an advantage over the models as they can use other, higher frequency data while formulating their forecasts. This is especially evident for interest rate forecasts.

Table of Contents

The data

- Variables of our interest: 3-month Treasury bill yield, GDP (QoQ SAAR), GDP price index (QoQ SAAR)
- We use the real-time data from the Philadelphia Fed Real-Time Data Set for Macroeconomists

The sample


Date
Q1:94 vinatage
Q2:94 vinatage
Q3:94 vinatage
...
Q3:05 vinatage
1979
1993
1994
2006
Q1
Q2
Q3
Q4
Estimation sample
Forecast

The choice of "actuals"

We evaluate the quality of the forecasts in two variants of the actuals:

- with the latest-available data set (from the vintage of 2006:3)
- with the real-time data available one year after the estimation vintage

Here, we present only the result for the "latest-available" case.[3]

[3] In the article we show, however, that the general findings are broadly the same.

Forecasts and the "latest-available" actuals

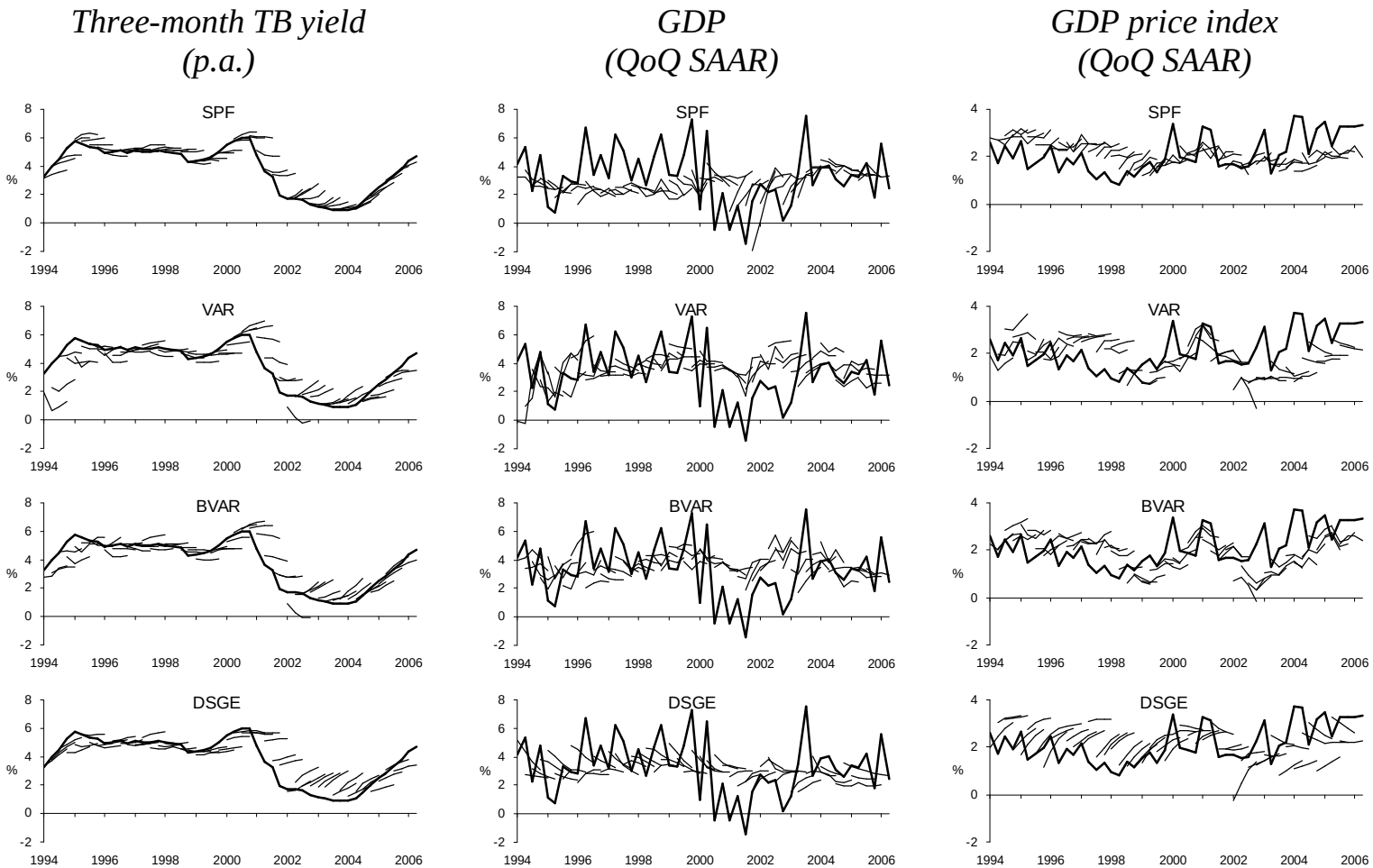


Table of Contents

A comparison of forecast errors

h	*Short-term interest rate*				*Output growth*				*Inflation*			
	SPF	*VAR*	*BVAR*	*DSGE*	*SPF*	*VAR*	*BVAR*	*DSGE*	*SPF*	*VAR*	*BVAR*	*DSGE*
Mean Error												
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
Mean Absolute Error												
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
Root Mean Squared Error												
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

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

A test of equal forecast accuracy

We proceed 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.

h	*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						**DSGE vs. VAR**					
1	**3.84**	***0.000***	0.78	*0.441*	1.81	*0.077*	0.02	*0.981*	**-2.09**	***0.042***	-0.78	*0.438*
2	**2.65**	***0.011***	-0.35	*0.727*	**2.51**	***0.016***	-0.57	*0.573*	-0.47	*0.639*	-0.69	*0.491*
3	1.65	*0.105*	-1.68	*0.100*	**2.61**	***0.012***	-0.39	*0.699*	-0.49	*0.629*	-1.13	*0.265*
4	1.36	*0.180*	-1.41	*0.164*	1.50	*0.140*	-0.38	*0.705*	-0.87	*0.387*	-1.12	*0.270*
	DSGE vs. BVAR						**BVAR vs. SPF**					
1	2.01	*0.051*	-1.07	*0.288*	0.46	*0.647*	**3.60**	***0.001***	1.31	*0.195*	1.48	*0.145*
2	0.75	*0.457*	-0.49	*0.628*	0.48	*0.634*	**2.70**	***0.010***	0.12	*0.908*	0.81	*0.423*
3	0.65	*0.517*	-0.73	*0.472*	0.26	*0.797*	**2.52**	***0.015***	-0.57	*0.570*	0.52	*0.608*
4	0.56	*0.581*	-0.94	*0.353*	-0.08	*0.939*	1.73	*0.090*	-0.28	*0.777*	0.70	*0.490*
	BVAR vs. VAR						**VAR vs. SPF**					
1	-1.82	*0.075*	-1.49	*0.143*	-1.80	*0.078*	**3.13**	***0.003***	**2.03**	***0.048***	**2.97**	***0.005***
2	-1.23	*0.226*	-0.20	*0.845*	-1.13	*0.264*	**2.01**	***0.050***	0.25	*0.805*	1.86	*0.069*
3	-1.37	*0.177*	0.26	*0.795*	-1.61	*0.114*	**2.03**	***0.048***	-0.56	*0.581*	1.82	*0.075*
4	-1.38	*0.174*	0.30	*0.765*	-1.88	*0.067*	1.78	*0.082*	-0.33	*0.743*	1.69	*0.098*

Notes: A positive value of the HLN statistic informs that the RMSE of A is higher than that of B. The bold figures indicate the rejection of the null at the 5% significance level.

Summary

The general picture that emerges from the above analysis is that:

- the proposed DSGE model was not able to significantly outperform the SPF in forecasting output growth, inflation nor interest rates in the United States
- the DSGE model was found to perform comparable or even better to the trivariate VAR and BVAR models

Further research

Clearly, additional research is required to document the out-of-sample performance of DSGE models:

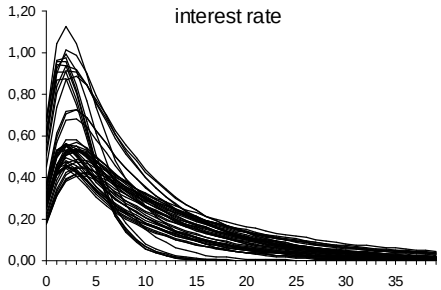
- the structure of the DSGE model presented in this article is relatively simple and hence forecasting properties of a more complex DSGE model could be studied
- the forecast accuracy of DSGE models could be compared to a larger group of methods than the SPF and VAR models

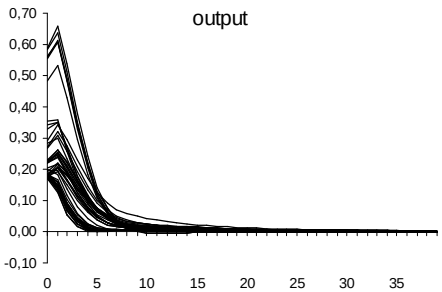
We believe, however, that our analysis constitutes a step in extending our knowledge about the forecasting properties of DSGE models

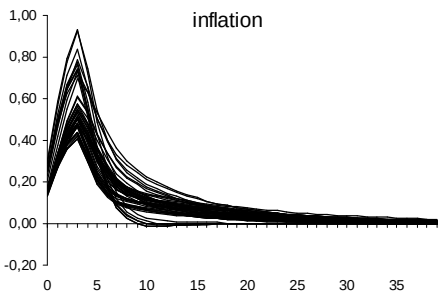
Thank you for your attention

DSGE model - recursive IRFs

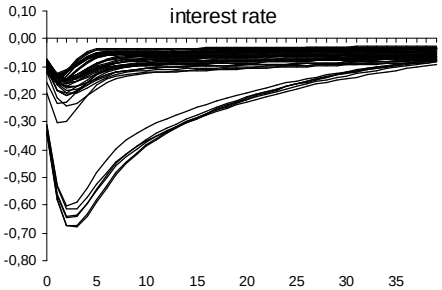
demand shock

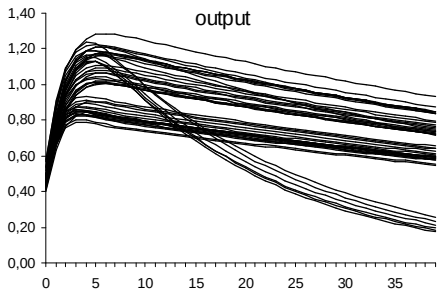


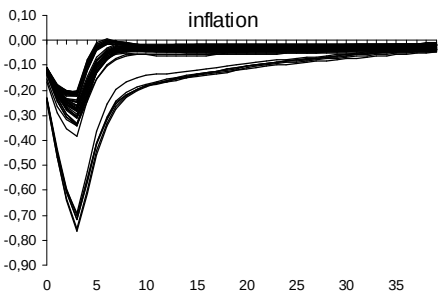




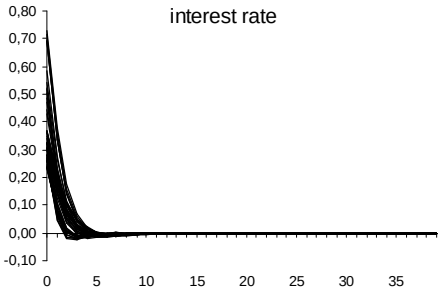
supply shock

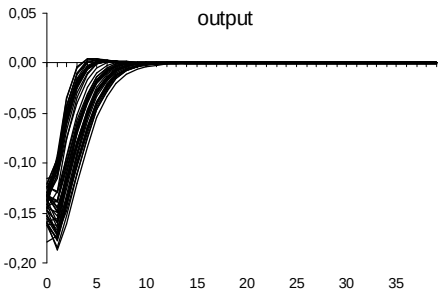


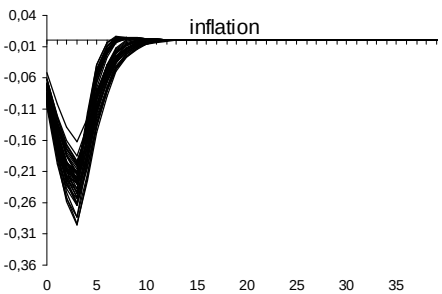




monetary shock







Test of unbiasedness

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

$$X_t = \alpha_0 + \alpha_1 X_t^F + \varepsilon_t$$

and test the null hypothesis of unbiasedness: $\alpha_0 = 0 \wedge \alpha_1 = 1$ using the Wald Chi-squared test corrected for heteroskedasticity and autocorrelation.

Test of unbiasedness

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: Bold figures indicate the rejection at the 5% significance level.