HOUSE OF FINANCE
Goethe-Universität Frankfurt

POLICY PLATFORM | Policy Letter

Macroeconomic Model Comparisons and Forecast Competition

Volker Wieland
Institute for Monetary and Financial Stability

Maik Wolters
Goethe-University Frankfurt

February 2012

Macroeconomic Model Comparisons and Forecast Competition[1]

Volker Wieland and Maik Wolters

The failure of researchers and professional forecasters to predict the ”Great Recession” of 2008 and 2009 has generated much criticism regarding the state of economic forecasting and macroeconomic modeling. Distinguished economists – among them Nobel Prize winner Paul Krugman – have blamed developments in macroeconomic modeling over the last 30 years and particularly the usage of dynamic stochastic general equilibrium (DSGE) models for this failure.

Key policy makers take a more pragmatic view, namely that there is no alternative to the usage of simplified models, but that the development of complementary tools to improve the robustness of policy decisions is required. For example, ECB President Jean-Claude Trichet recently expressed these needs very clearly:

> *“We need macroeconomic and financial models to discipline and structure our judgemental analysis. How should such models evolve? The key lesson I would draw from our experience is the danger of relying on a single tool, methodology or paradigm. Policymakers need to have input from various theoretical perspectives and from a range of empirical approaches. Open debate and a diversity of views must be cultivated – admittedly not always an easy task in an institution such as a central bank. We do not need to throw out our DSGE and asset-pricing models: rather we need to develop complementary tools to improve the robustness of our overall framework”.*[2]

Against this background, we present a new paper (Wieland et al., 2012), in which we propose a comparative approach to macroeconomic policy analysis that is open to competing modeling paradigms. We have developed a data base of macroeconomic models that enables a systematic comparative approach to macroeconomic modelling, with the objective of identifying policy recommendations that are robust to model uncertainty. This comparative approach enables individual researchers to conduct model comparisons easily, frequently, at low cost and on a large scale.

The macroeconomic model data base is available for downloading at www.macromodelbase.com and includes over 50 models. We have included models that are used at policy institutions like the IMF, the ECB or the Fed and in academia. The data base includes models of the U.S. economy, the Euro area economy and several multi-country models. Some of the models are fairly small and focus on explaining output, inflation and interest rate dynamics. Many others are of medium scale and cover many key macroeconomic aggregates.

Some models in the data base are fairly large in scale such as the Federal Reserve’s FRB-US model, the model of the G7 economies of John Taylor or the ECB’s area-wide model. Most of the models can be classified as New Keynesian models because they incorporate rational expectations, imperfect competition and wage or price rigidities. Many of these New

[1] This article was first published on voxeu.org on February 13, 2012.
[2] See Jean-Claude Trichet, „Reflections on the nature of monetary policy non-standard measures and finance theory“, speech given on the occasion of the ECB Central Banking Conference Frankfurt, 18 November 2010.

Keynesian models fully incorporate recent advances in terms of microeconomic foundations. However, some models that assign little role to forward-looking behavior by economic agents are included in the data base as well.

This data base can be used to compare the implications of specific economic policies across models, but it can also serve as a testing ground for new models. New modeling approaches may offer more sophisticated explanations of the sources of the financial crisis and carry the promise of improved forecasting performance. This promise should be put to a test rather than presumed. Wieland and Wolters (2011) show how to implement such model competition. We start by analyzing whether existing DSGE models pass a test of fulfilling necessary minimum requirements to be considered for business cycle analysis. Based on the new macroeconomic model database, we compute forecasts for five different rational expectation models and compare them to forecasts from professionals as collected in the Survey of Professional Forecasters (SPF).

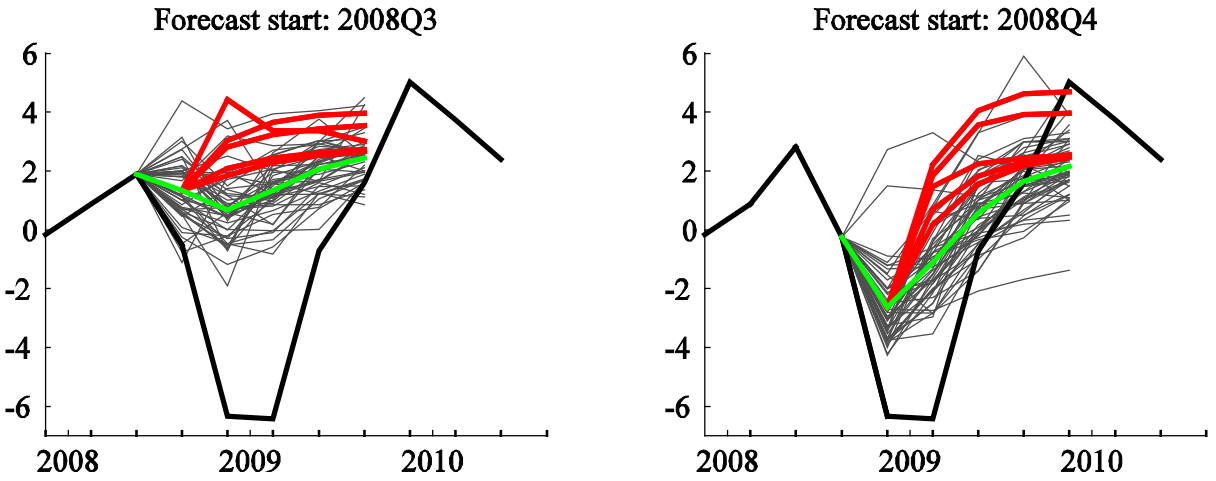
In recent years, researchers such as Smets & Wouters (2004), Adolfson et al. (2007) and Edge et al. (2010) have reported encouraging findings regarding the forecasting performance of DSGE models. However, the existing papers are based on samples with long periods of average volatility and can therefore not address specifically how well DSGE model-based forecasts perform during recessions and recoveries. Therefore, we analyse the forecasting performance of models and experts around the five most recent NBER defined recessions. Turning points pose the greatest challenge for economic forecasters, are of most importance for policy makers and can help us to understand current limitations of economic forecasting especially with respect to the recent financial crisis.

In Wieland and Wolters (2011) we use two small micro-founded New Keynesian models, two medium-size state-of-the-art New Keynesian business cycle models – often referred to as DSGE (dynamic-stochastic general equilibrium) models – and, for comparison purposes, an earlier-generation New-Keynesian model (also with rational expectations and nominal rigidities but less strict microeconomic foundations) and a Bayesian VAR model. For each forecast we *reestimate all five models using exactly the data as it was available for professional forecasters when they submitted their forecasts to the SPF*. Using these historical data vintages is crucial to ensure comparability to historical forecasts by professionals. We compute successive quarter-by-quarters forecasts up to five quarters ahead for all models.

Predicting the recession of 2008-2009

Figure 1 shows forecasts for annualized quarterly real output growth for the recent financial crisis. The black line shows real-time data until the forecast starting point and revised data afterwards. The grey lines show forecasts collected in the SPF and the green line shows their mean. Model forecasts are shown in red. While data for real GDP becomes available with a lag of one quarter, professional forecasters can use within quarter information from data series with a higher frequency. In contrast the models can process only quarterly data. To put the models on an equal footing in terms of information with the forecasts of experts, we condition their forecasts on the mean estimate of the current state of the economy from the SPF.

Figure 1



Notes: Solid black line shows annulized quarterly output growth (real-time data vintage until forecast starting point and revised data afterwards), grey lines show forecasts from the SPF, green line shows mean forecast from the SPF, red lines show model forecasts conditional on the mean nowcast from the SPF.

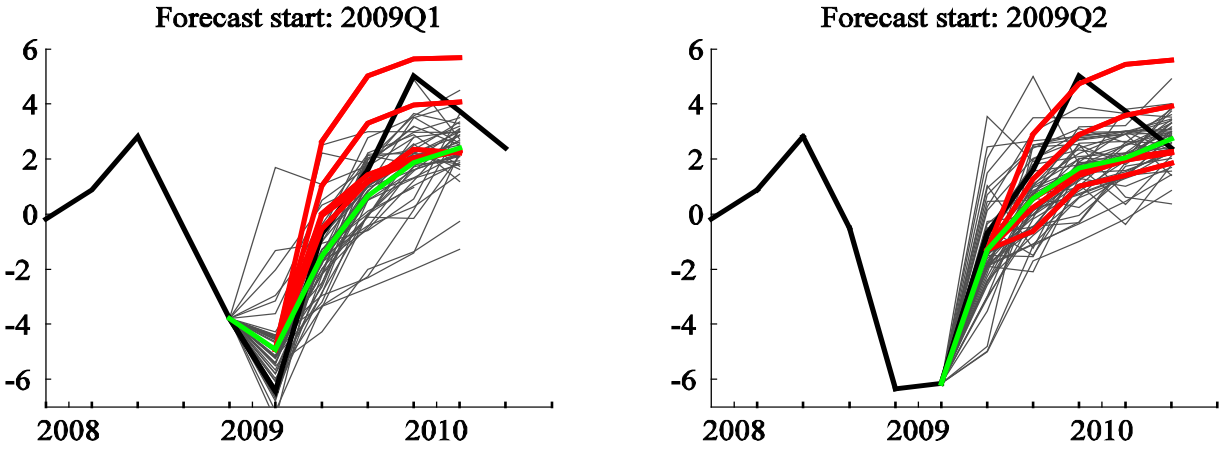
The forecasts shown in the left graph start in the third quarter 2008 and have been computed before the collapse of Lehman brothers. It is apparent that all professional forecasters failed to foresee the downturn. The mean SPF forecast indicates a slowdown of growth in the fourth quarter of 2008 followed by a return to higher growth in the first quarter of 2009. The model-based forecasts would not have performed any better and predict even higher growth rates than most professional forecasters. The figure on the right shows that in the fourth quarter of 2008, following the Lehman debacle, professional forecasters drastically revised their assessments of the current state of the economy downwards. Still, growth turned out to be even much lower than estimated. Professional forecasters as well as model forecasts wrongly predicted that the trough had already been reached. While the models predict positive growth rates one quarter ahead, some of the professional forecasters were somewhat more pessimistic. The model-based predictions and the professional forecasters are, however, far from predicting an extreme downturn of as much as minus six percent output growth.

Given this failure to predict the recession and its length and depth, the wide-spread criticism of the state of economic forecasting before and during the financial crisis applies to business forecasting experts as well as modern and older macroeconomic models. Professional forecasters, who are able to use information from hundreds of data series, including information about financial market conditions and all kinds of different forecasting tools and thus have clear advantage over purely model-based forecasts, were not able to predict the great recession either. Thus, there is no reason to single out DSGE models and favour more traditional Keynesian-style models that may still be more popular among business experts. In particular, Paul Krugman's proposal to rely on such models for policy analysis in the financial crisis and disregard three decades of economic research is mis-placed.

Is there any hope left for economic forecasting and the use of modern structural models in this endeavour?

Figure 2 shows professional and model-based forecasts starting in the first and the second quarter of 2009. Professional forecasters continued to revise their estimated nowcast downwards for the first quarter of 2009 and predict an increase of growth rates afterwards. Interestingly, from the first quarter of 2009 onwards, the model-based forecasts perform quite well in predicting the recovery of the U.S. economy. Three-quarters-ahead, model-based forecasts dominate expert forecasts in several cases.

Figure 2



Comparing the forecasting accuracy of professional and model-based forecasts

We obtain a number of interesting findings with regard to the relative accuracy of model-based and professional forecasts measured in terms of root mean squared prediction errors (RMSE). In Wieland and Wolters (2011) we report detailed results for all the different recessions. To illustrate the main results table 1 shows the RMSEs of the different models for all the different recessions together:

Table 1: Root Mean Squared Prediction Errors

Horizon	NK1	NK2	RE	DSGE1	DSGE2	BVAR	Mean	SPF
0	3.74	3.87	3.47	3.07	3.04	3.31	3.19	1.90
1	3.90	3.97	4.03	3.81	3.62	3.60	3.64	2.90
2	3.67	3.72	4.28	3.93	3.74	4.17	3.78	3.28
3	3.88	3.89	4.43	4.10	3.85	4.30	3.96	3.71
4	3.77	3.77	4.23	3.98	3.80	4.22	3.85	3.60

Notes: NK1: standard New Keynesian model (Del Negro & Schorfheide, 2004); NK2: as NK1, but with additional shock processes; RE: older generation rational expectation model (Fuhrer, 1997), DSGE1: medium scale model (Smets & Wouters, 2007); DSGE2: large scale model (Edge, Kiley & Laforte 2008), BVAR: Bayesian VAR; Mean: mean forecast of the previous six models; SPF: mean forecast of the Survey of Professional Forecasters.

The model forecasts are on average less accurate than the mean SPF forecasts. Of course, taking the mean of all forecasts collected in the SPF can increase the forecasting accuracy compared to individual forecasts. Looking at individual forecasts from the SPF, we observe that the precision of the different model forecasts is well in line with the precision range of forecasts from professionals. The difference between the RMSEs of model and expert forecasts decreases with the forecast horizon. Structural models are therefore suitable for medium-term forecasts while expert forecasts incorporate additional information that helps improve near-term forecasts or nowcasts. However, for practical policy usage, medium term horizon forecasts might be of much more interest due to long transmission lags.

Computing the mean forecast of all six models, as denoted by the column “Mean”, we obtain a robust forecast that is close to the accuracy of the forecast from the best model. Conditioning the model forecasts on the nowcast of professional forecasters (reported in the paper) can further increase the accuracy of model-based forecasts. Overall, model-based

forecasts still exhibit somewhat greater errors than expert forecasts, but this difference is surprisingly small considering that the models only take into account few economic variables and incorporate theoretical restrictions that are essential for evaluations of the impact of alternative policies but often considered a hindrance for effective forecasting.

Conclusion

Both – models and professionals – failed to predict the financial crisis. At the current state of knowledge about macroeconomics and the limitations to use all this knowledge in simplified models, large recessions might simply be difficult to forecast at all. The question is whether approaches that have been proposed by some economists as alternatives or successors can improve on this important aspect of business cycle analysis. While recently the development of modern DSGE models with financial sectors has taken off, it remains to be seen whether these models perform better in explaining or predicting the financial crisis. Furthermore, other modeling approaches such as agent-based models or behavioural models apparently have not yet reached a state of development that allows the conduct of the type of model competition described above. We hope this will change in the future.

Wieland et al. (2012) presents the data base and a computational platform to compare policy implications across models. We have found that the implications of a certain policy can differ substantially for different modeling approaches. In Wieland and Wolters (2011) we have extended this comparative approach to forecasting. By comparing the forecasts from different models, we can hedge against outliers and find predictions that are robust across several models. Our macroeconomic model database provides a testing ground for macroeconomists to compare new models to a large range of existing benchmarks. We thus provide the tools for a comparison with established benchmarks and current forecasting practice as documented in the SPF. It is important to base discussions about competing modeling approaches on a solid basis. In this paper we have shown how such a comparison of different models can be pursued.

References

Adolfson, M., M. K. Andersson, J. Linde, M. Villani, A. Vredin (2005). Modern forecasting models in action: Improving macroeconomic analyses at central banks. *Sveriges Riksbank Working Paper* No. 190.

Del Negro, M., F. Schorfheide (2004). Priors from general equilibrium models for VARs. *International Economic Review* 45(2):643–673.

Edge, R. M., M. T. Kiley, J.P. Laforte (2008). Natural rate measures in an estimated DSGE model of the U.S. economy. *Journal of Economic Dynamics and Control* 32:2512–2535.

Edge, R. M., M. T. Kiley, J.-P. Laforte (2010). A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced form models, and a DSGE model. *Journal of Applied Econometrics* 25(4): 720-754.

Smets, F. R. Wouters (2004). Forecasting with a Bayesian DSGE model: An application to the euro area. *Journal of Common Market Studies* 42(4):841–867.

Smets F., R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *The American Economic Review* 97(3):586–606.

Taylor, J. and V. Wieland (2011). Surprising Comparative Properties of Monetary Models: Results from a New Monetary Model Base, *Review of Economics and Statistics*, forthcoming.

Wieland, V., T. Cwik, G. J. Müller, S. Schmidt, M. H. Wolters (2012). A new comparative approach to macroeconomic modeling and policy analysis. *Manuscript, Center for Financial Studies*, Frankfurt.

Wieland, V., M. H. Wolters (2011). The Diversity of Forecasts from Macroeconomic Models of the U.S. Economy, *Economic Theory* 47(2-3):247-292 (also available as CEPR Discussion Paper 7870).