Model comparison and robustness: A proposal for policy analysis after the financial crisis*

Volker Wieland
Goethe University of Frankfurt and IMFS

This Version: September 9, 2011

* Funding support from European Community grant MONFISPOL under grant agreement SSH-CT-2009-225149 is gratefully acknowledged. Excellent research assistance was provided by Elena Afanasyeva, Matthias Burgert and Sebastian Schmidt. Helpful comments by Robert Boyer and the participants of the Cournot Centre for Economic Studies' conference, 'What's Right with Macroeconomics?', held 2-3 December 2010 in Paris, were greatly appreciated. The usual disclaimer applies.

1 Introduction

In the aftermath of the financial crisis, the state of macroeconomic modelling and the use of macroeconomic models in policy analysis has come under heavy criticism. Media and other commentators have criticized macroeconomists for failing to predict the great recession of 2008–09, or at least failing to provide adequate warning of the risk of such a recession. Practitioners have attributed this failure to academic and central bank researchers' love of a particular modelling paradigm. They blame so-called dynamic stochastic general equilibrium (DSGE) models for misdirecting the attention of policymakers. Indeed, even some well-known academics-cum-bloggers have published scathing commentaries on the current state of macroeconomic modelling. On 3 March 2009, Willem Buiter wrote on the Financial Times blog,[1] '... the typical graduate macroeconomics and monetary economics training received at Anglo-American universities during the past 30 years or so, may have set back by decades serious investigations of aggregate economic behavior and economic policy-relevant understanding'. This view was echoed by Nobel Laureate Paul Krugman on 11 June 2009 in the weekly *The Economist*, 'Most work in macroeconomics in the past 30 years has been useless at best and harmful at worst'.

Against this background, this chapter aims to develop a more constructive proposal for how to use macroeconomic modelling – whether state-of-the-art or 1970s vintage – in practical policy design. It is written in the vein of the 1992 call for a pluralistic and rigorous economics by leading economists. The undersigned – among them Nobel Laureates Paul Samuelson and Franco Modigliani – were concerned with 'the threat to economic science posed by intellectual monopoly' and pleaded for 'a new spirit of pluralism in economics, involving critical conversation and tolerant communication between different approaches.'[2] It is in that spirit that I propose a systematic comparative approach to macroeconomic modelling with the objective of identifying policy recommendations that are robust to model uncertainty.[3] This approach is open to a wide variety of modelling paradigms.

Scientific rigor demands a level-playing field on which models can compete. Rather than using rhetoric to dismiss competing approaches, emphasis should be placed on empirical benchmarks that need to be satisfied by the models in order to stay in the race. For example, macroeconomic models used for monetary policy could be required to be estimated to fit the empirical dynamics of key time series, such as output, inflation and nominal interest rates. Models should be able to provide answers to policy makers' typical questions, such as what is the effect of an unanticipated increase (or decrease) in the central bank's operating target for the money market rate, or of an unanticipated

[1] "The unfortunate uselessness of most 'state of the art' academic monetary economics": http://blogs.ft.com/maverecon/2009/03/the-unfortunate-uselessness-of-most-state-of-the-art-academic-monetary-economics/#axzz1SqddzN1g

[2] See the advertisement section of the *American Economic Review* – AEA Papers and Proceedings, May 1992.

[3] See Taylor and Wieland (2011) and Wieland et al. (2009) for an implementation of this model comparison approach.

temporary increase (or decrease) in government spending or transfers? Another common concern is the degree of output, inflation and interest rate volatility and persistence predicted by the models under different policy rules. New modelling approaches may offer more sophisticated explanations of the sources of the great recession of 2008–09 and carry the promise of improved forecasting performance. This promise should be put to a test rather than presumed. Estimated models could be compared along their relative real-time forecasting performance, in particular during periods of great change, such as recessions and recoveries. An example of such a model competition is given by Volker Wieland and Maik Wolters (2011).

Macroeconomic data, however, are unlikely to provide sufficient testing grounds for selecting a single, preferred model for policy purposes. Instead, policy recommendations should be made robust to model uncertainty. In particular, the robustness of policy prescriptions can be improved by introducing them in multiple, competing models and comparing performance across models according to established target criteria. Policy makers continue to require models and are aware of the need for robustness. Most recently, European Central Bank President Jean-Claude Trichet 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.*[4]

Macroeconomic model comparison projects have already helped produce some very influential insights for practical policy making. For example, John B. Taylor (1993a) credits the comparison project organized by the Brookings Institution and summarized in Ralph Bryant et al. (1993) as the crucial testing ground for what later became known as the Taylor rule for monetary policy. More recently, the International Monetary Fund (IMF) organized a large-scale model comparison exercise in order to evaluate the likely consequences of temporary fiscal stimulus measures (see Coenen et al., 2011). Such model comparisons have been, nevertheless, infrequent and costly, because they require

[4] See Jean-Claude Trichet, 'Reflections on the nature of monetary policy non-standard measures and finance theory', speech given at the ECB Central Banking Conference in Frankfurt on 18 November 2010.

the input of many teams of researchers and multiple meetings to obtain a limited set of comparative findings.

To remedy this situation, Volker Wieland et al. (2009) have implemented a new approach to model comparison that enables individual researchers to conduct such comparisons easily, frequently, at low cost and on a large scale. This methodology comes with a model archive that includes many well-known empirically estimated models of business-cycle dynamics. These models can be used to evaluate the performance of macroeconomic stabilization policies. A computational platform that allows straightforward comparisons of models' implications using MATLAB and DYNARE[5] software is available online to download.[6] Researchers can easily include new models in the database and compare the effects of novel extensions to established benchmarks.

Our proposal is to use this comparative approach systematically in order to organize a pluralistic, yet rigorous and productive communication between competing modelling paradigms in macroeconomics. So far, the model database contains small-, medium- and large-scale macroeconomic models of different vintages and methodological traditions. The first release of November 2010 covers 38 models including many state-of-the-art New Keynesian DSGE models, but also earlier vintage New-Keynesian models with rational expectations and nominal rigidities, as well as some models that offer a more traditional Keynesian-style perspective on macroeconomic fluctuations with largely backward-looking dynamics. The model database and the computational platform for model comparison provide a level-playing field that is open to new entrants. Going forward, we propose to cover as many competing modelling paradigms as possible, so as to compare models' empirical implications in a systematic fashion, and to search for policy prescriptions that are robust along relevant dimensions of model uncertainty.

The next section shortly describes the comparison methodology, and gives an overview of the models available in the database. It also outlines a list of competing modelling paradigms that promise improvements in our understanding of macroeconomic dynamics. In future work, they should be compared to those approaches that have received the most attention in recent years. Section 3 gives an example of a demanding test of comparative model performance, namely the real-time forecasting evaluation of a range of models relative to experts in the last five US recessions conducted by Wieland and Wolters (2010). Section 4 reviews findings from recent comparative studies regarding the impact of fiscal stimulus packages and reports estimates of the impact of government transfers in selected models. Section 5 concludes.

[5] See Juillard (1996).
[6] See http://www.macromodelbase.com.

2 Model comparison

The six older comparison projects reported in Bryant et al. (1988), Bryant et al. (1989), Klein (1991), Bryant et al. (1993), Taylor (1999), Hughes-Hallett and Wallis (2004), as well as the recent IMF exercise by Coenen et al. (2011) have all involved multiple teams of researchers, each team working only with one or a small subset of available models. The approach by Wieland et al. (2009), on the other hand, is meant to provide users easy access to the complete set of models considered in a comparison exercise. Furthermore, users should find it fairly straightforward to integrate their own models. To this end, Wieland et al. (2009) present a formal exposition of their comparative methodology. Taylor and Wieland (2011) use the model database to compare three well-known models of the US economy and analyse the robustness of simple monetary policy rules.

A general class of nonlinear dynamic stochastic macroeconomic models is augmented with a space of common comparable variables, parameters and shocks. Augmenting models in this manner is a necessary precondition for a systematic comparison of particular model characteristics. Given a space of common variables and parameters, one can define common policy rules as model input, and produce comparable objects as model output. These objects are also defined in terms of common variables, parameters and shocks. Examples for such objects are impulse response functions, autocorrelation functions and unconditional distributions of key macroeconomic aggregates.

The space of common variables, parameters and policy rules comprises only a subset of each model's variables, parameters and equations. Most model-specific equations remain unchanged. Only the model-specific policy rules are replaced with common policy rules that express policy variables as functions of common variables and parameters. Nevertheless, a new set of definitional equations needs to be added to each model. These definitional equations define the common variables in terms of model-specific variables. Once each model is augmented with the appropriate definitional equations and the common policy rules, it is ready for comparative exercises. For a formal exposition of the procedure for integrating new models, see Wieland et al. (2009). Several examples are carried out step-by-step in that paper. A detailed documentation of the augmented model files is also provided.

2.1 A model database

In the following, I give a brief overview of the model archive that is available with the comparison software. This database includes many well-known, empirically-estimated macroeconomic models that may be used for quantitative analysis of monetary and fiscal stabilization policies. It contains estimated and calibrated models of the US economy and the euro area. There are also a number of small open-economy models of countries such as Canada, Chile and Brazil. Finally, it also includes

several multi-country models that cover industrialized economies. The models, made available in the first release as of November 2010, are listed in Table 1.

Most models could be classified as New Keynesian, because they incorporate rational expectations, imperfect competition and wage or price rigidities. A subset of the models could be characterized as monetary business cycle models where all behavioural equations are derived in a completely consistent manner from the optimization problems of representative households and firms. Many authors use the term 'dynamic stochastic general equilibrium' (DSGE) model to refer to this particular class of model. Thus, the database offers interesting opportunities for comparing policy implications of this class of model to a broader set of empirically estimated, dynamic, stochastic, economy-wide macro models.

While most of the models assume that market participants form rational, forward-looking expectations, we have also included some models that assume little or no forward-looking behaviour.[7] Comparative analysis of these classes of models will be useful to evaluate recently voiced criticisms that the newer models have been rendered invalid by the global financial crisis.

The models are grouped into five categories in Table 1. The first category includes small, calibrated versions of the basic New-Keynesian DSGE model. These models concentrate on explaining output, inflation and interest-rate dynamics. Some of them are calibrated to US data. The model taken from Clarida et al. (2002) is a two-country version of this type of model.

The second category covers estimated models of the US economy. It includes small models of output, inflation and interest-rate dynamics, such as Fuhrer and Moore (1995) and Rudebusch and Svensson (1999). Other models are of medium scale, such as Orphanides and Wieland (1998) or the well-known models of Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007), which fully incorporate recent advances in terms of microeconomic foundations. The database includes the version of the Christiano–Eichenbaum–Evans model estimated by Altig et al. (2005), because this version contains other economic shocks in addition to the monetary policy shock studied by Christiano et al. (2005).[8] We have also included an additional version of the Altig et al. (2005) model used in Taylor and Wieland (2011) that omits the cost channel of monetary policy.[9] The largest model of the US economy in the database is the Federal Reserve's FRB–US model of Reifschneider et al. (1999). We

[7] For example, the models of Rudebusch and Svensson (1999) and Orphanides (2003) are essentially structural VAR models with some restrictions on the coefficients. The ECB's Area-Wide Model is a medium-sized structural model with a relatively limited role for forward-looking behaviour compared to the other structural rational expectations models in the database.
[8] Because of complications in programming the informational timing assumptions regarding expectations in this model in DYNARE, two versions are included: one for simulating the consequences of a monetary policy shock, and the other for simulating the consequences of the other economic shocks in the model.
[9] This version was created in Taylor and Wieland (2011) to evaluate the effect of this assumption in comparing the Altig et al. (2005) model with the model of Smets and Wouters (2007), which features no such cost channel.

have included a linearized version of this model with rational expectations that was previously used in Levin et al. (2003), as well as two more recent versions from 2008, one with rational expectations and one with adaptive expectations based on a reduced form vector-autoregression (VAR). Federal Reserve economists Rochelle Edge, Michael Kiley and Jean-Philippe Laforte (2010) have developed a new two-sector DSGE model of the US economy that is also included in the database; a version of this model is estimated in Wieland and Wolters (2010).

In addition, there are a number of smaller estimated models of the US economy that offer new insights into the role of financial frictions in economic fluctuations. Ioan Carabenciov et al. (2008), for example, augment a simple backward-looking model with a measure of financial linkages and frictions. Ferre De Graeve (2008) and Ian Christensen and Ali Dib (2008) introduce constraints on firms' financing following Ben Bernanke et al. (1999) in a fully-fledged estimated DSGE model of the US economy. These models provide an endogenous account of firms' external finance premium over the business cycle. Matteo Iacoviello (2005) includes the housing sector in a DSGE model. The model of N. Gregory Mankiw and Ricardo Reis (2007) deviates from the assumption of rational expectations and allows for rational inattention. This mechanism introduces a role for outdated expectations (or informational frictions) in business-cycle dynamics. All these extensions of the standard DSGE framework were accomplished before the financial crisis and could potentially be helpful in rendering DSGE models more useful in explaining developments during the crisis.

Table 1: MODELS CURRENTLY AVAILABLE IN THE DATABASE

(NOVEMBER 2010)

1. SMALL CALIBRATED MODELS		
1.1	NK RW97	Rotemberg and Woodford (1997)
1.2	NK LWW03	Levin et al. (2003)
1.3	NK CGG99	Clarida et al. (1999)
1.4	NK CGG02	Clarida et al. (2002)
1.5	NK MCN99cr	McCallum and Nelson (1999), (Calvo–Rotemberg model)
1.6	NK IR04	Ireland (2004)
1.7	NK BGG99	Bernanke et al. (1999)
1.8	NK GM05	Gali and Monacelli (2005)
2. ESTIMATED US MODELS		
2.1	US FM95	Fuhrer and Moore (1995)
2.2	US OW98	Orphanides and Wieland (1998) equivalent to MSR model in Levin et al. (2003)
2.3	US FRB03	Federal Reserve Board model linearized as in Levin et al. (2003)
2.4	US FRB08	2008 linearized version of Federal Reserve Board model
2.5	US FRB08mx	2008 linearized version of FRB model (mixed expectations)
2.6	US SW07	Smets and Wouters (2007)
2.7	US ACELm	Altig et al. (2005) (monetary policy shock)
	US ACELt	Altig et al. (2005) (technology shocks)
	US ACELswm	no cost channel as in Taylor and Wieland (2011) (mon. pol. shock)
	US ACELswt	no cost channel as in Taylor and Wieland (2011) (tech. shocks)
2.8	US NFED08	based on Edge et al. (2007), version used for estimation in Wieland and Wolters (2010)
2.9	US RS99	Rudebusch and Svensson (1999)
2.10	US OR03	Orphanides (2003)
2.11	US PM08	IMF projection model US, Carabenciov et al. (2008)
2.12	US PM08fl	IMF projection model US (financial linkages), Carabenciov et al. (2008)
2.13	US DG08	DeGraeve (2008)
2.14	US CD08	Christensen and Dib (2008)
2.15	US IAC05	Iacoviello (2005)
2.16	US MR07	Mankiw and Reis (2007)

3. ESTIMATED EURO AREA MODELS

3.1	EA CW05ta	Coenen and Wieland (2005) (Taylor-staggered contracts)
3.2	EA CW05fm	Coenen and Wieland (2005) (Fuhrer–Moore staggered contracts)
3.3	EA AWM05	ECB's area-wide model linearized as in Dieppe et al. (2005)
3.4	EA SW03	Smets and Wouters (2003)
3.5	EA SR07	Sveriges Riksbank euro area model of Adolfson et al. (2007)
3.6	EA QUEST3	QUEST III, model by DG-ECFIN EC, Ratto et al. (2009)

4. ESTIMATED/CALIBRATED MULTI-COUNTRY MODELS

4.1	G7 TAY93	Taylor (1993b) model of G7 economies
4.2	G3 CW03	Coenen and Wieland (2002) model of USA, Euro Area and Japan
4.3	EACZ GEM03	Laxton and Pesenti (2003) model calibrated to Euro Area and Czech Republic
4.4	G2 SIGMA08	Federal Reserve's SIGMA model from Erceg et al. (2008) calibrated to the US economy and a symmetric twin
4.5	EAUS NAWM08	Coenen et al. (2008), New Area Wide model of Euro Area and USA

5. ESTIMATED MODELS OF SMALL OPEN ECONOMIES

5.1	CL MS07	Medina and Soto (2007), model of the Chilean economy
5.2	CA ToTEM10	ToTEM model of Canada, based on Murchison and Rennison (2006), 2010 vintage
5.3	BRA SAMBA08	Gouvea et al. (2008), model of the Brazilian economy

==========

The third category in Table 1 covers estimated models of the euro-area economy. Four of those models have been used in a recent study of robust monetary policy design for the euro area by Keith Kuester and Volker Wieland (2010): the medium-scale model of Smets and Wouters (2003), two small models by Coenen and Wieland (2005) that differ by the type of staggered contracts that induce inflation rigidity, and a linearized version of the Area-Wide Model that was used at the European Central Bank (ECB) for forecasting purposes. The latter was recently replaced by a new DSGE model. In addition, we have included an estimated DSGE model of the euro area recently developed at the Sveriges Riksbank, Adolfson et al. (2007), and at the European Commission, Ratto et al. (2009). The latter model was developed with a particular focus on the analysis of euro-area fiscal policy.

The fourth category includes estimated and calibrated models of two or more economies. Currently, the largest model in the database is the estimated model of the G7 economies of Taylor (1993b). The estimated model of Coenen and Wieland (2002) with rational expectations and price rigidities aims to explain inflation, output and interest-rate dynamics and spill-over effects between the United States, the euro area and Japan. The model of Laxton and Pesenti (2003) is a two-country model with extensive microeconomic foundations calibrated to the economies of the euro area and the Czech Republic. The Federal Reserve's SIGMA model is similarly rich in microeconomic foundations. The parameters in the two-country version of this model from Erceg et al. (2008) are calibrated to the US economy and a symmetric twin. Finally, there is a two-country calibrated version of the ECB's new area-wide DSGE model as presented by Coenen et al. (2008). This model also covers the US economy.

The fifth category of models covers small open-economy DSGE models of Canada, Chile and Brazil. In addition to openness to trade and capital flows, these models also consider particular economic features of the respective countries, such as the important role that a natural resources sector might play in the economy.

In sum, the current breadth of model coverage allows for a variety of interesting comparison exercises, for example, between earlier vintage and more recent Keynesian-style models of business cycle dynamics for a given country; cross-country comparisons between the United States, the euro area and some small open economies; or comparisons between standard New-Keynesian DSGE models and DSGE models that also include some financial or informational frictions.

2.2 A proposal for extending the coverage of competing modelling paradigms

In the aftermath of the financial crisis, the DSGE modelling approach has come under heavy criticism. Many critics have argued that models of this type that were in use prior to the crisis did not incorporate realistic treatments of banking, and therefore failed to account for the macroeconomic risks resulting from a fragile financial sector. Other critics have suggested that the crucial flaw of the DSGE approach is of a more fundamental nature. Many of them question the central assumption of homogenous, rational expectations. They point out that in practice, economic agents are imperfectly informed, they are engaged in a learning process, and they often disagree about likely future developments. Others go further, calling into question the basic microeconomic assumption of rational optimizing behaviour by households and firms.

Policy makers are keen to have modelling frameworks at their disposal that address these criticisms. Their interest in the matter is well exemplified by ECB President Trichet who requests the following steps:

> *…we need to better integrate the crucial role played by the financial system into our macroeconomic models, ... we may need to consider a richer characterisation of expectation formation,.... We need to deal better with heterogeneity across agents and the interaction among those heterogeneous agents, (and) we need to entertain alternative motivations for economic choices.*[10]

The following paragraphs highlight some recent studies that explore these different directions.

Financial sector risks

Proponents of the DSGE approach have been hard at work to provide more explicit modelling of financial intermediation and risks by extending the standard DSGE framework. As a minimum, such models should include a financial sector where banks are exposed to risk and where the functioning of the banking sector affects the real economy. Recent contributions along these lines include Goodfriend and McCallum (2007), Gertler et al. (2007), De Fiore et al. (2009), DeWalque et al. (2010), Christiano et al. (2010), Gerali et al. (2010), Angeloni and Faia (2010), Meh and Moran (2010), Nolan and Thoenissen (2009), Dib (2010), Gertler and Karadi (2009) and Gertler and Kiyotaki (2009).

All these contributions examine the interaction of financial risk, business cycle dynamics and monetary policy. They differ in how banking and financial intermediation are modelled, and in the focus of the particular policy application. Some of them investigate the implications of banking and financial intermediation on business cycle fluctuations in a fully-fledged DSGE model. From the perspective of the modellers, an important question to be investigated in the future is whether such extensions offer a satisfactory explanation of the financial crisis. One would hope that such an explanation would reveal not only the sources that caused the crisis in the form of particular economic shocks, but also the propagation mechanisms that would help modellers to predict the development of such a crisis in the future.

Learning and diverse beliefs

Households and firms in complex, modern DSGE models are assumed to have access to forecasts that are equivalent to the expectation calculated under complete knowledge about the

[10] See Jean-Claude Trichet, 'Reflections on the nature of monetary policy non-standard measures and finance theory', speech given at the ECB Central Banking Conference in Frankfurt on 18 November 2010.

structural features of the model economy. Households and firms typically are assumed to share homogeneous expectations regarding future developments of key macroeconomic variables. Expectations play a crucial role in determining the dynamics of these models and the policy recommendations derived from them. Expectations, of course, also appear to play a very important role in actual real-world markets and economies. Thus, the debate among modellers should not be about the importance of expectations in macroeconomics, but rather about the sensitivity of the business cycle and policy implications derived under the homogeneous rational expectations assumption to alternative specifications of market participants' beliefs.

A number of different approaches to modelling less-than-fully-rational expectations and belief diversity have been proposed in the economic literature. A first step away from rational expectation is adaptive learning. It implies that market participants re-estimate simple reduced-form models of the variables to be forecasted and update the parameter estimates of these forecasting models once they obtain new data. An example of such a learning process is recursive least squares. Adaptive learning has been discussed in macroeconomics for more than two decades.[11] The expectations obtained from adaptive learning, however, are typically homogeneous across market participants in these models. Some examples of recent investigations of the implications of adaptive learning for macroeconomic dynamics as well as monetary and fiscal policy are Orphanides and Williams (2006), Slobodyan and Wouters (2008) and Wieland (2009). Nevertheless, medium- to large-scale DSGE models used at central banks and other policy institutions are typically simulated under rational expectations.

It would be of interest to conduct a systematic comparison of DSGE models with rational expectations versus DSGE models with adaptive learning in order to evaluate whether adaptive learning plays an important role in interpreting the period leading up to the global financial crisis or the reaction of market participants during the crisis. Furthermore, if adaptive learning better characterizes real-world market participants' process of expectations formation, models with adaptive learning might also perform better in forecasting exercises.

Expectation heterogeneity, however, has so far been largely ignored in structural macroeconomic models used for policy analysis. While empirical studies have documented a substantial degree of heterogeneity of professional forecasts[12], theoretical research has emphasized that expectational heterogeneity itself can be an important propagation mechanism for economic fluctuations and a driving force for asset price dynamics. Theories of heterogeneous expectations and endogenous fluctuations have been advanced, for example, by Kurz (1994a, 1994b, 1996, 1997a,1997b, 2009),

[11] A well-known textbook that provides a comprehensive framework for adaptive learning is Evans and Honjapohja (2001).
[12] see Kurz et al. (2003); Kurz et al. (2005); Giordani and Söderlind (2003); Capistran and Timmermann (2009); and Wieland and Wolters (2010).

Brock and Hommes (1998), Kurz et al. (2005), Chiarella et al. (2007), Branch and McGough (2011), Branch and Evans (2011), and De Grauwe (2011).

Since belief diversity can cause economic volatility, macroeconomic policy analysis cannot ignore the diversity of expectations among households, firms and policy makers themselves. While in homogenous models such volatility would be attributed to other shock processes, models with heterogeneous expectations offer a possibility to disentangle which fraction of economic volatility can be attributed to heterogeneous expectations and which fraction is explained by other economic shocks. Some of the above-mentioned studies explore the impact of diverse beliefs in small New Keynesian models. It would be of great interest to introduce such models into the macroeconomic model database and conduct a systematic comparison of models with homogenous and heterogeneous beliefs. Possibly, variations in the diversity of beliefs – the degree of optimism and pessimism among market participants – may have played an important role in the asset price boom before the crisis, and its subsequent collapse. If these variations in diversity act as a propagation mechanism and were themselves to some extent predictable, then models with diverse beliefs might stand a chance to deliver a better forecasting performance before such recessions than standard DSGE models.

Deviations from strictly optimizing behaviour

In the last two decades, several significant strands of literature have developed that investigate fundamental deviations from strictly optimizing behaviour by economic agents and consider alternative motivations for economic choices and decisions. Behavioural economics brings lessons from psychology to bear on the analysis of economic decision making. Contributions to this strand of the literature have argued that empirical failures of the classical paradigm of fully rational behaviour may be resolved by introducing particular psychological motivations for economic behaviour (see Diamond and Vartiainen, 2007; and Akerlof and Shiller, 2010). With regard to policy analysis with structural macroeconomic models, an important question is which behavioural macroeconomic models are best suited to be considered as competitors of standard DSGE models, or the new DSGE models with detailed banking sector and financial frictions. Sometimes behavioural approaches are mentioned in support for more traditional Keynesian-style models with backward-looking dynamics. The richness of the behavioural economics approach would suggest, however, that a new line of structural macroeconomic models should emerge from this literature.

Another large body of literature is known under the term ‘agent-based modelling’ and crosses the borders between engineering, physics and economics.[13] Agent-based modelling is the computational study of economic processes modelled as dynamic systems of interacting agents. Here, ‘agent’ refers

[13] For a recent survey of agent-based models in economics, see Tesfatsion and Judd (2006).

broadly to a bundle of data and behavioural methods representing an entity that constitutes part of a computationally constructed world. Instead of the fully optimizing rational decision makers in standard DSGE models, these agents can range from active data-gathering decision makers with sophisticated learning capabilities to passive world features with no cognitive function. Researchers have built computational 'laboratories' with thousands or even millions of such agents. These laboratories have been used to investigate whether agent-based modelling can replicate some empirical regularities in financial, goods and labour markets, as well as other areas of economics. Another aim is to test certain government regulations and policies in terms of the simulation outcomes they would generate in such models. A recent contribution that describes a model of the euro area economy is Deissenberg et al. (2008). With regard to policy analysis with structural macroeconomic models, an important question is how agent-based models can be used to deliver answers to the type of questions policy makers typically ask of DSGE models. For example, what are the models' predictions for growth and inflation over the coming year? What would be the effect of an increase in the central bank's interest rate or of an unexpected increase in fiscal transfers, such as a tax rebate? A comparison of agent-based and DSGE models with regard to such questions would be tremendously useful for practical macroeconomic policy analysis.

3 A recent model comparison: Forecasting US recessions

In general, macroeconomic models used for policy analysis in a particular economy ought to be empirically estimated or calibrated to fit macroeconomic time series of that economy. A more demanding test, however, would be to evaluate the real-time forecasting performance of such models. Recently, Wieland and Wolters (2010) have conducted such a forecasting exercise with six different models of the US economy. They have investigated the accuracy and heterogeneity of output growth and inflation forecasts during the current and the four preceding NBER[14]-dated US recessions. Model forecasts were compared to professional forecasts from the Federal Reserve's Greenbook and the Survey of Professional Forecasters (SPF).[15] Importantly, the model parameters and model forecasts were derived from historical data vintages so as to ensure comparability to historical forecasts by professionals. The comparison was conducted for successive quarter-by-quarter forecasts up to four quarters into the future. Arguably, the periods around recessions and recoveries posed the greatest challenge for economic forecasters.

[14] National Bureau of Economic Research

[15] The SPF is conducted quarterly and contains responses from 30 to 50 professional forecasters. It was initiated in 1968 by the American Statistical Association and the NBER and has been administered by the Federal Reserve Board of Philadelphia since 1990. The Greenbook is not a survey. It contains a single forecast produced by the staff of the Board of Governors of the Federal Reserve System in Washington DC, which becomes publicly available within a five-year lag.

Wieland and Wolters (2010) ('WW' hereafter) considered six macroeconomic models: three small-scale New Keynesian models that differ in terms of structural assumptions, a non-structural Bayesian VAR model, and two medium-scale New Keynesian DSGE models of the type currently used by leading central banks. Two of the small-scale models were variants of the New-Keynesian models 1.1 and 1.2 in Table 1, estimated with US data. They were estimated by Marco Del Negro and Frank Schorfheide (2004) and Wieland and Wolters (2011), respectively, and are denoted by the acronyms *NK-DS* and *NK-WW*. The third small-scale model was a variant of model 2.1. in Table 1, developed by Fuhrer (1997). It is denoted by *NK-Fu* while the VAR model is referred to as *BVAR-WW*. The three small-scale, New Keynesian models and the Bayesian VAR were estimated to fit three macroeconomic time series: real gross domestic product (GDP) growth, inflation measured by the GDP deflator, and the federal funds rate.

The first medium-scale model is the well-known DSGE model estimated by Frank Smets and Rafael Wouters (2007) (model 2.6. in Table 1), which itself is a version of the DSGE model developed in Christiano et al. (2005). It is referred to as the *CEE-SW* model in the forecasting exercise. It is estimated with seven variables, including consumption, investment, wages and hours worked. The largest model in the forecasting exercise is a version of the Federal Reserve's new DSGE model estimated by Edge et al. (2010) (model 2.8. in Table 1). It is denoted by *FRB-EDO* in the forecast evaluation. This model accounts for the breakdown in durables versus non-durables and services consumption, residential versus business investment, and the related deflators. It is estimated on 11 macroeconomic data series.

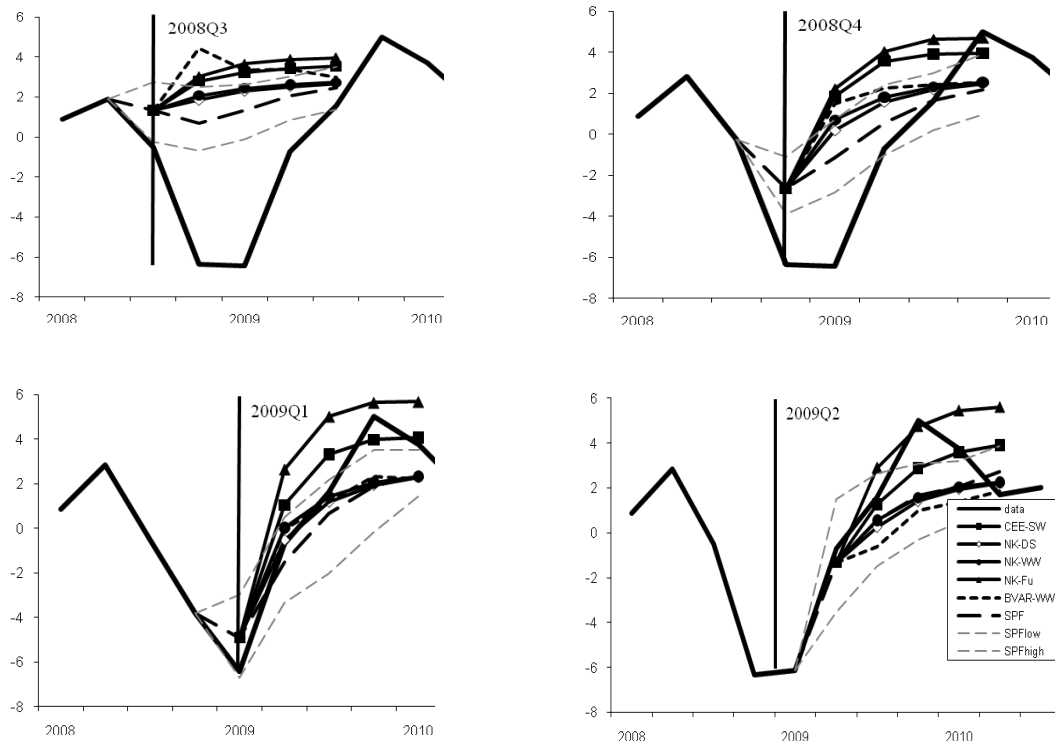
3.1 Forecasting the 2008-09 recession: Models versus experts

To render model-based forecasts comparable to historical SPF and Greenbook forecasts, Wieland and Wolters (WW) have put them on a similar footing in terms of the data vintage used for parameter estimation and initial conditions. Thus, WW have created a large, real-time data set that contains all the historical quarterly vintages of the 11 time series used in the largest model. For every quarter, they re-estimate all the model parameters on the basis of the data vintage that was available at that exact point in time. Using this parameterization, they compute an estimate of the current state of the economy – the so-called *nowcast* – and forecasts for one to four quarters into the future. Then, they assess the forecasting precision relative to the revised data that became available during the subsequent quarters for the dates to which the forecasts apply.

The model-based forecasts only use quarterly data vintages, where the most recent data entries concern the quarter preceding the quarter in which the forecast was made. In practice, however, there are many data series that are available on a monthly, weekly or daily frequency that can be used to improve

current-quarter estimates of GDP. Examples are industrial production, sales, unemployment, money, opinion surveys, interest rates and other financial prices. These data can be used to improve nowcasts; the Federal Reserve staff and many professional forecasters make use of them. Methods for using higher frequency data systematically in combination with quarterly structural macroeconomic models in conjunctural analysis are available (see Giannone et al., 2009). To illustrate the impact of the timeliness of the nowcast on model-based forecasts, WW compare model forecasts initiated with both types of nowcasts.

Figure 1: Real output growth forecasts during the 2007–09 recession



Note: The mean SPF nowcast forms the starting point for model-based forecasts regarding future quarters.

The four panels in Figure 1 replicate the individual model forecasts from WW for the 2008-09 recession that are initialized with the mean SPF nowcast. Each panel displays model forecasts relative to the mean SPF forecast (dash-dotted line) and the actual data (solid line) that has become available so far. In addition, I have included a measure of the central tendency of SPF forecasts for comparative purposes. It is indicated by the grey dashed-lines labelled SPFlow and SPFhigh. This measure of the central tendency is computed in the same manner as the Federal Reserve computes central tendencies of FOMC forecasts, that is by omitting the three greatest outliers on the high and the low side. The top left panel shows forecasts made in the third quarter of 2008. The top right panel then reports forecasts from the fourth quarter of 2008, and the two lower panels from the first two quarters of 2009.

As shown in the top left panel, professional forecasters, on average, failed to foresee the downturn as late as in the third quarter of 2008. The central tendency of professional forecasts, however, anticipated somewhat less growth than the model forecasts. The mean SPF forecast indicates a slowdown in the fourth quarter followed by a return to higher growth in the first quarter of 2009. The model-based forecasts based on the data vintage of the third quarter of 2008 do not perform any better.

Following the Lehman debacle, professional forecasters drastically revised their assessments of the current situation downwards, and continued to do so in the first quarter of 2009.Interestingly, from 2009:Q1 onwards, the model-based forecasts perform quite well in predicting the recovery of the US economy. From that point onwards, several of the models deliver predictions that are very similar to the mean SPF forecast, and match up with the subsequent data releases surprisingly well. The 2009:Q1 forecasts for the second and third quarter of 2009 – implied by the *CEE-SW* and *NK-WW* models – already look fairly accurate relative to the subsequent data releases. The central tendency of SPF forecasts indicates a somewhat more pessimistic outlook regarding the speed of recovery than the models. However, the above-mentioned two models came closer to the actual data for the following quarters.

3.2 The relative accuracy of model-based and expert forecasts

For the purpose of a systematic evaluation of forecast accuracy, WW compute the root mean squared errors (RMSE) of the nowcast and forecasts from one to four quarters ahead for each model during the five recessions. The typical recession sample covers the period from four quarters prior to the trough, determined by the NBER Business Cycle Dating Committee, to four quarters after the trough.[16] The Greenbook nowcast is used as the initial condition for the model-based forecasts, except in the latest recession where the mean SPF nowcast is applied. Models are re-estimated every quarter, and forecasts are computed for horizons of one-to-four quarters into the future. Table 2 reports the associated root mean squared errors of output growth and inflation forecasts for the different recession episodes from WW. It compares the accuracy of the individual model forecasts to the mean model forecast (the average of the six models), the mean SPF forecast and the Greenbook forecast.

[16] Exceptions are the 1980 and 2008–09 recessions. In the first case, they start only 2 quarters prior to the trough because of data availability. In the second case, the trough is not yet determined. They start in 2007Q4 (peak) and end in 2009Q3.

Table 2: RMSEs of Output Growth Forecasts Initialized with Expert Nowcasts

Sample / Horizon	NK-DS	NK-WW	CEE-SW	FRB-EDO	NK-Fu	BVAR-WW	Mean	GB	SPF
1980:1 - 1981:3									
0	5.05	5.05	5.05	5.05	5.05	5.05	5.05	5.05	–
1	8.14	8.13	6.33	6.06	7.18	6.69	5.83	6.65	–
2	6.34	6.36	4.80	5.60	6.48	6.48	4.83	5.54	–
3	5.50	5.74	5.20	5.37	6.49	7.74	5.20	6.11	–
4	5.56	5.75	4.23	4.24	4.12	5.50	4.05	5.32	–
1981:4 - 1983:4									
0	2.42	2.42	2.42	2.42	2.42	2.42	2.42	2.42	2.14
1	4.28	4.50	3.74	3.27	3.80	3.23	3.54	3.58	3.88
2	3.99	4.05	4.22	4.09	3.98	4.09	3.86	3.93	4.11
3	4.14	4.23	4.05	4.52	4.64	4.87	4.25	3.91	4.41
4	4.08	4.11	4.07	4.67	4.73	4.89	4.28	3.84	4.02
1990:1 - 1992:1									
0	1.27	1.27	1.27	1.27	1.27	1.27	1.27	1.27	1.12
1	2.64	2.87	3.22	1.70	3.11	2.00	2.47	2.09	1.45
2	2.95	3.04	3.80	1.92	3.68	2.28	2.82	2.34	2.06
3	3.08	3.13	3.78	2.42	3.67	2.55	2.94	2.31	2.54
4	2.71	2.76	3.65	2.16	3.48	2.29	2.69	2.18	2.37
2000:4 - 2002:4									
0	2.28	2.28	2.28	2.28	2.28	2.28	2.28	2.28	2.22
1	2.17	2.15	2.31	2.84	2.06	2.48	2.23	2.20	2.30
2	2.09	2.10	2.11	2.61	2.35	1.98	2.11	2.34	2.21
3	2.74	2.72	2.68	2.98	2.51	2.66	2.65	2.76	2.65
4	2.25	2.26	2.08	2.40	2.24	2.30	2.19	2.18	2.13
2007:4 - 2009:3									
0	1.94	1.94	1.94	–	1.94	1.94	1.94	–	1.94
1	3.74	3.90	4.24	–	4.54	4.85	4.21	–	3.30
2	4.52	4.62	4.94	–	5.48	5.10	4.89	–	4.11
3	5.05	5.11	5.39	–	5.83	5.27	5.32	–	4.80
4	5.50	5.52	5.86	–	6.07	5.57	5.70	–	5.39

Model forecasts perform differently. There is no single model that dominates the others in all recessions. The *CEE-SW* model performs best in the 1980–81 recession. It even beats the Greenbook forecast in this recession, though the Greenbook forecast is unusually far off the mark in this period compared to later ones. The *NK-DS* and *CEE-SW* models perform best among the model forecasts in the 1981–82 recession, while the Greenbook forecast performs best overall in this case. In the 1990–92 period, the *FRB-EDO* model and the Bayesian VAR deliver the best predictions among the models. For the short horizon, the mean SPF forecast is best, but the Greenbook dominates the other forecasts over three to four quarters. In the period around the 2001 recession, the *NK-Fu* model and the *CEE-SW* model dominate the others over a horizon of three to four quarters. In the 2008–09 period, so far, the *NK-DS* and *NK-WW* models appear to perform best among the models. Of course, as shown in Figure 1, it is only during the recovery phase that the models appear to have some predictive power. WW, however, obtain very similar findings for the four preceding recessions.

Interestingly, the mean model forecast outperforms the Greenbook forecast in the 1980 and 2001 recessions. The mean model forecast also compares well to the mean SPF forecast in the 1981–82 and 2001 recessions. The Greenbook forecasts still perform best in the 1981–82 and 1990–91 recessions,

while the mean SPF forecast appears to be the most accurate in the ongoing recession, for which no Greenbook data and forecasts are publicly available.

The forecast comparison gives no cause for much shoulder-slapping among modellers nor among professional forecasters. Both tend to miss the onset of recessions. This is not only true in the 2008–09 recession, but also in the four preceding ones. Thus, there is no reason for expert forecasters, who tend to rely more often on traditional Keynesian-style models with backward-looking dynamics, to point fingers at DSGE modellers for supposedly having too much influence on central banks. Experts and models exhibit some predictive power during the recovery phase. They predict a return to mean, and the speed of return predicted seems to be reasonably accurate once the recovery has actually started. Some encouragement for modelling efforts, however, can be drawn from the finding that mean model forecasts perform well at horizons of three to four quarters and sometimes dominate Greenbook or mean SPF forecasts.

Given these findings, it does not seem to be appropriate to utterly dismiss state-of-the-art New Keynesian DSGE models in favour of those Keynesian-style models that only use theory from more than 30 years ago, as suggested by the Buiter and Krugman commentaries cited in the introduction of the chapter. Nevertheless, it appears urgent to investigate whether any of the innovations discussed in section 2.2. – such as more thorough modelling of the financial sector, the inclusion of learning and diverse beliefs, or behavioural and agent-based modelling – can deliver on the promise of improved forecasting power.

4 Investigating policy robustness

Model competitions in terms of empirical fit or predictive power will certainly help narrow down the field of models relevant for policy analysis. The preceding forecasting exercise suggests, however, that such competitions are not likely to deliver a unique preferred model for policy purposes. As recognized by policy makers such as ECB President Trichet, multiple models need to be used as tools for making policy recommendations robust to model uncertainty. There exist recent examples. Several model comparison studies have been conducted to investigate the likely consequences of temporary fiscal stimulus.

Importantly, in January 2009, Christina Romer, then Chair of the US President's Council of Economic Advisers, and Jared Bernstein, Chief Economist of the Office of the Vice-President, used macroeconomic models in a report on the likely impact of a large-scale stimulus package proposed by the Obama Administration. Soon after, the US Congress approved 787 billion US dollars in additional spending, transfers and tax reductions with the 2009 *American Recovery and Reinvestment Act*

(ARRA). The ARRA extended over five years, with much of the additional spending occurring in the first two years. Many other economies around the world also announced fiscal stimulus measures. In Europe, the EU initiated the *European Economic Recovery Plan* (EERP), while national governments announced their own fiscal stimuli. Among them, the German government launched two *Konjunkturpakete* in a row. The European stimulus packages were to be concentrated on 2009 and 2010 only.

4.1 Recent comparative evaluations of fiscal stimulus

The literature on fiscal stimulus has expanded very quickly. Here, I only focus on a few contributions that used multiple structural macroeconomic models to evaluate the likely impact of such measures, with an eye towards robustness to model uncertainty. Romer and Bernstein (2009), for example, provide numerical estimates of the impact of an increase in government spending and government transfers, respectively, on GDP and employment by averaging over models of the Federal Reserve and private-sector business consultancies. They estimate that an increase in government purchases of 1 per cent of GDP would induce an increase in real GDP of 1.6 per cent.[17] Thus, one or more of the models they use exhibits a text-book Keynesian multiplier effect.

The text-book multiplier follows from the national accounts' spending identity when combined with the Keynesian consumption function. An increase in government spending boosts aggregate spending, and thereby aggregate output and after-tax household income. Consumption is assumed to increase with current after-tax income. Consequently, a debt-financed increase in government spending boosts total spending (and therefore total GDP) more than one for one. Details of the individual model simulations behind this average effect have not been made available for the Romer–Bernstein study, however the authors clarify that interest rates were assumed to remain constant for five years in the model simulations. On that basis, they project that a package similar in magnitude to the eventual ARRA legislation would boost US GDP by 3.6 per cent.

Shortly after the ARRA had passed the House and Senate, John F. Cogan et al. (2009) (later published as Cogan et al., 2010) evaluated its likely impact on US GDP using empirically-estimated New Keynesian models, such as the models of Taylor (1993b) and Smets and Wouters (2007), that is model *4.1 G7-TAY93* and *2.6 US-SW07* in Table 1. In these models, government purchases multipliers are typically smaller than one. They exhibit significant crowding-out of private consumption and private investment following an increase in government purchases. Consumption declines because forward-looking households expect increased government debt to be paid off at a later stage with higher taxes. This negative wealth effect induces additional saving and reduced consumption earlier on. Private

[17] See Romer and Bernstein (2009), Appendix 1, p. 12. This paper was written during the transition period in early January before Romer was sworn in as Chair of the Council of Economic Advisers.

investment declines because increased government debt puts upward pressure on interest rates. The expectation of higher interest rates and lower wealth in the future, in turn reduces private investment already in the near term. As a consequence, estimates of the GDP effects of ARRA legislation obtained with the model of Smets and Wouters (2007) are only one-sixth as large as those of Romer and Bernstein (2009).

While the original Smets–Wouters model only contains forward-looking 'permanent-income' consumers, Cogan et al. (2010) also estimated a variant of this model that includes households whose consumption is determined by their current after-tax income, as prescribed by the Keynesian consumption function. The empirically estimated share of these 'rule-of-thumb' households is 26.5 per cent. At this scale, the presence of rule-of-thumb consumers only has a small impact on the multiplier of the ARRA government purchases. It remains well below one in this model. In addition, Cogan et al. (2010) investigate the interaction of monetary and fiscal policy when monetary policy is constrained at the zero-interest-rate floor. If the central bank's desired policy rate is negative – and thus below the actual rate of zero – it will not respond to an increase in GDP by raising the policy rate as in normal times. Consequently, the crowding-out effect of an increase in government purchases would be lessened. Cogan et al. (2010) consider simulations with one and two years of constant interest rates, as well as simulations where the time spent at the zero lower bound is endogenous and projected from the trough of the recession in the first quarter of 2009 onwards. Though the GDP impact of ARRA purchases increases it remains far below the Romer–Bernstein estimate of 3.6 per cent by the end of 2010.

The euro area stimulus measures were summarized and evaluated in Cwik and Wieland (2011)[18] using five different structural models of the euro area economy based on Fagan et al. (2005), Smets and Wouters (2003), Ratto et al. (2009), Laxton and Pesenti (2003) and Taylor (1993b), respectively, models *3.3 EA-AWM05, 3.4 EA-SW03, 3.6 EA-QUEST3, 4.3 EACZ-GEM03* and *4.1 G7 TAY93* from Table 1. The ECB's Area-Wide model described in Fagan et al. (2005) largely ignores forward-looking motives for private decision-making and provides a traditional Keynesian perspective. The other four models are of the New Keynesian variety with forward-looking households and firms. Smets and Wouters' (2003) model is a euro area version of the medium-sized DSGE model of Christiano et al. (2005). The *EA-QUEST3* model is an estimated DSGE model developed for fiscal policy analysis at the European Commission by Ratto et al. (2009). This model also accounts for rule-of-thumb consumers. Their share is estimated at 35 per cent, not too far from the estimate obtained with US data by Cogan et al. (2010). The *EACZ-GEM03* model is a calibrated, two-country DSGE model of the euro area and Czech Republic developed by IMF researchers. Together with Taylor's

[18] See also the earlier working paper version, Cwik and Wieland (2009).

multi-country model, it can account for the possible diversion of fiscal stimulus towards import demand.

Cwik and Wieland (2011) confirm the differential assessment of traditional Keynesian and New Keynesian models concerning the size of the government purchases multiplier emphasized by Cogan et al. (2010) relative to Romer and Bernstein (2009). In their baseline scenario, the New Keynesian models of the euro area provide no support for a traditional Keynesian multiplier effect of government purchases. Crowding-out of consumption, investment and net exports dominates. The ECB's area-wide model, however, supports a strong impact of government spending on GDP that is substantially greater than one for one. The boom is nevertheless followed by a bust. Thus, the cumulative effect of government on private spending eventually also turns negative in that model. More importantly, models with backward-looking dynamics may not be as well suited for the analysis of major policy measures as the New Keynesian models that account for the interaction of policy announcements and private-sector expectations.

Overall, the euro area stimulus package was much smaller in magnitude than the US package, and more concentrated on 2009 and 2010. The findings in Cwik and Wieland (2011) suggest that such a shorter and sharper increase in government spending induces less crowding-out than the ARRA package, which includes significant additional spending from 2011 onwards. Cwik and Wieland (2011) also discuss some factors that may have played a role in the recession of 2008–09: namely implementation lags, the zero-interest-rate floor and the share of rule-of-thumb consumers. Time lags arise because of the steps needed to move from a timely announcement to the actual implementation of government spending plans. Such implementation lags lead to more crowding out and may even cause an initial contraction. If interest rates are anticipated to remain constant due to zero-bound effects for two years, that represents the complete period of fiscal stimulus. Cwik and Wieland document a small crowding-in effect in some of the DSGE models. For the multiplier to be greater than one, however, it is important that the two-year constant rate window is already anticipated as of the first quarter of 2009.

A number of studies have used other structural macroeconomic models to assess the impact of different fiscal policy tools. Typically they focus on a single model. An interesting extension of the EU-Quest model by Werner Roeger and Jan in't Veld (2009) includes a third type of household that is credit constrained. Their benchmark calibration apparently has 40 per cent liquidity-constrained households and another 20 per cent credit-constrained households, which would be too high relative to the survey evidence available regarding the share of such households during the financial crisis. The IMF's new preferred model for fiscal policy analysis – the so-called GIMF model – has been used by Charles Freedman et al. (2010) to analyse the consequences of different fiscal measures. An

innovative element of this model is that it features overlapping generations of households with finite horizons. As shown by Taylor (2010b), the effects of longer-lasting or permanent fiscal stimuli in the GIMF are very close to the effects reported by Cogan et al. (2010) for New Keynesian DSGE models. A short-term government spending shock in GIMF has a multiplier of unity under normal circumstances. Unfortunately, the GIMF model is calibrated and not estimated with state-of-the-art methods to fit US or euro-area data. It would be very useful to see how it fares in estimation relative to the estimated models I have used.

Recently, a very commendable model comparison study was carried out by 17 researchers from the IMF, the Organisation for Economic Co-operation and Development (OECD), the ECB, the Federal Reserve and the European Commission in Coenen et al. (2011). It covers seven structural models used at policy institutions, including GIMF, the modified version of EU-Quest with additional constrained households, the Fed's SIGMA and FRB-US models, the OECD Fiscal Model, the Bank of Canada-GEM model and the ECB's New Area-Wide Model. For comparative purposes, they also consider the Smets-Wouters model and the version with rule-of-thumb consumer estimated by Cogan et al. (2010) (CCTW) and they simulate the near-permanent fiscal expansion as well as the ARRA spending plan investigated by CCTW. In both, cases the outcomes under the CCTW model fall well inside the range of outcomes obtained with the other 7 policy models. Thus, they corroborate the robustness of the evaluation of the likely impact of ARRA spending by CCTW. However, Coenen et al emphasise that a counterfactual on-off increase in spending restricted to two years of anticipated constant interest rates would have delivered greater stimulative effects, also in the CCTW model. Such a shorter stimulus is closer to the euro area stimulus evaluated by Cwik and Wieland (2011). However, Coenen et al (2011) neglect the possibility of implementation lags investigated by Cwik and Wieland (2011). Furthermore, several of their models assume shares of 40 to 50 per cent of rule-of-thumb households that are much higher than the 26.5 per cent share estimated by CCTW. I will address the question whether such a higher share is more likely in the recent recession or not further below.

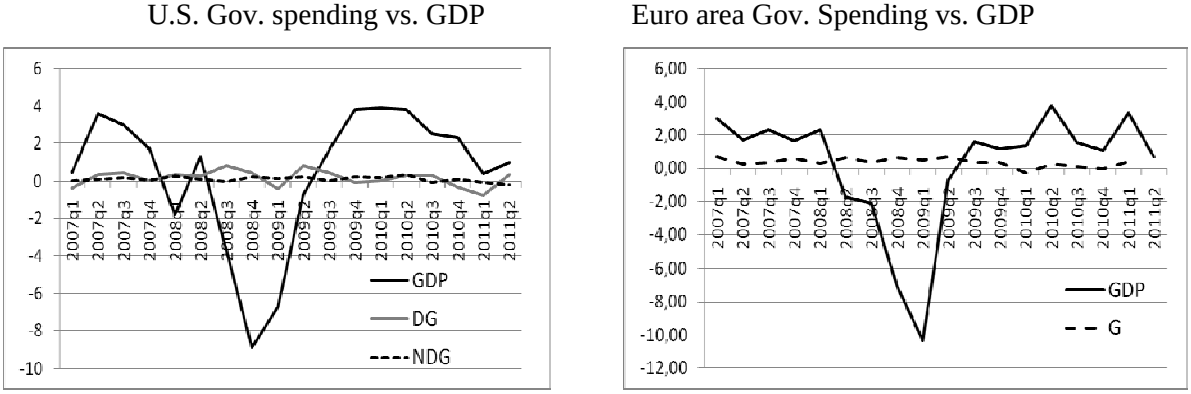
Further comparison of the findings in CCTW and Cwik and Wieland with those by Coenen et al. (2011) would be very useful. Unfortunately, however, Coenen et al. (2011) use a traditional model comparison approach whereby separate teams of researchers conduct a specific set of experiments, each team in their own model, and report outcomes. It would be very useful if the policy institutions represented by these research teams would choose to create a platform for model comparison as in Wieland et al. (2009), or add their models to this new model data base. Such a platform would also render their model simulations directly replicable and transparent to researchers outside those teams and institutions. Replicability is a basic scientific standard that ensures that correct comparisons can be made and policy recommendations properly scrutinized. Software for replicating the Coenen et al (2011) model simulations is made available on the American Economic Journal: Macroeconomics

website. Unfortunately it is based on TROLL, a software tool that is not easily available for individual researchers outside central banks.

4.2 Government purchases versus government transfers

The preceding review of the literature focused primarily on the likely size of the government purchases multiplier. Recently, however, doubts have surfaced as to whether the 2009 ARRA legislation in the United States did achieve the announced increase in government consumption and infrastructure investment – that is, the announced multiplicand. Using new data from the Bureau of Economic Analysis (BEA) and considering developments at the federal, state and local levels, Cogan and Taylor (2010) find that the government purchases multiplicand through the 2nd quarter of 2010 is only 2 per cent of the total spending announced by the ARRA. This increase in purchases occurred mainly at the federal level, while state and local governments used the substantial grants they received under the ARRA to reduce borrowing and increase transfer payments rather than purchases.

Figure 2: Government spending contributions to GDP growth: 2007–10



Notes: GDP: GDP growth; NDG: non-defence spending contribution to GDP growth; DG: defence spending contribution; G: government spending contribution to growth. Sources: BEA and ECB. Vintage: August 2011.

The Cogan and Taylor (2010) finding seems to explain why the contribution of government spending to GDP growth in the national accounts has remained rather flat. As shown in the left panel of Figure 2, non-defence spending (dashed line) varied little over the recession and recovery from 2008 to 2010. There is no strong upward spike in its contribution to GDP growth visible in 2009 or 2010. Interestingly, the contribution of government spending to GDP growth in the euro area also remained fairly flat throughout the recession and recovery as is apparent from the right panel of Figure 2 (dashed line). It seems difficult to make a case for a crucial role of government spending in stimulating growth based on the inspection of this chart. It would be very useful if the European Commission and national

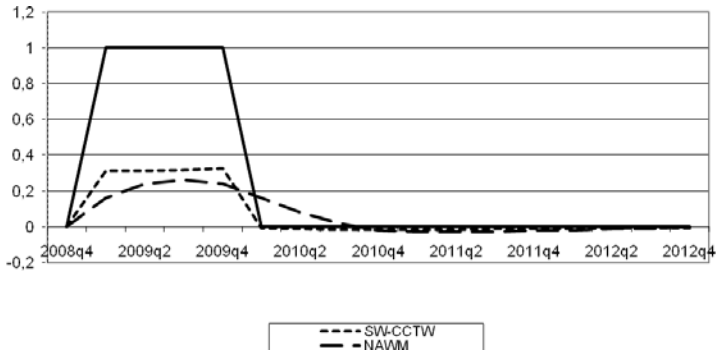
euro area governments would similarly publish information on the actual spending pattern in relation to the announced measures. This information is crucial for making appropriate ex-post evaluations of their effectiveness.

In light of these findings, assessments of the impact of the government stimulus packages in 2009 in the USA and euro area should perhaps focus more on the likely effect of government transfers on GDP growth. Romer and Bernstein (2009) estimated, based on their models, that additional permanent government transfers and tax cuts planned by the US administration would increase GDP by 2010 by one for one.[19] Given, of course, that the overall amount of the ARRA was limited, the announced changes in transfers and taxes were primarily temporary in nature. The effect of such temporary measures depends crucially on the importance of different motives for consumer behaviour. Traditional Keynesian-style models may predict a positive impact of temporary transfers on GDP, because consumption is modelled as a function of current after-tax income. Forward-looking, permanent-income consumers would instead see through the temporary increase and expect little or no boost to permanent income, because future tax increases may be needed to pay off the government debt incurred. Thus, their consumption would not change at all. For this reason, the DSGE models of Smets and Wouters (2003) and (2007), as well as the new DSGE model of Fed researchers Edge et al. (2010), predict that a temporary increase in government transfers, tax cuts or tax rebates have no effect on GDP.

As discussed earlier, some of the empirically-estimated New Keynesian DSGE models allow for the presence of rule-of-thumb consumers. In Cogan et al. (2010), the estimated share is 26.5 per cent, similar to other estimates available in the literature. In order to illustrate the impact of government transfers and tax cuts, I simulated an increase in government transfers in that model of 1 per cent of GDP for the duration of one year (solid black line in Figure 3). Figure 3 also shows that GDP (dotted line) then increases by about 30 basis points for a year in the CCTW model. In this simulation, interest rates are set according to Taylor's rule. I have also considered interest-rate accommodation for one or two years due to zero bound effects. The resulting increase in the GDP effect of transfers is rather small, however.

[19] See Appendix 1 of Romer and Bernstein (2009): multipliers for different types of spending.

Figure 3: The GDP impact of a temporary increase in government transfers of 1 % of GDP



Notes: SW-CCTW: DSGE model with rule-of-thumb households estimated in Cogan et al. (2010) with US data; NAWM: calibrated two-country version of ECB's New-Area-Wide model taken from Coenen et al. (2008); transfers: temporary increase of transfers in the US economy.

For comparison, I also include a simulation of the same temporary increase in government transfers in the calibrated two-country version of the ECB's New-Area-Wide model taken from Coenen et al. (2008). An estimated, single-economy, euro-area version of this model has recently replaced the AWM model in ECB policy analysis. This estimated model is also used in the Coenen et al. (2011) study. In the calibrated two-country model, the share of rule-of-thumb agents is set at 25 per cent. Transfers, however, are assumed to be unevenly distributed, in per-capita terms, over the two types of households. The rule-of-thumb households are favoured at a ratio of 3 to 1. As indicated by the dashed line in Figure 3, the impact on GDP of a one-per-cent increase in transfers is similar, though slightly smaller, than the estimate obtained with the model of Cogan et al. (2010).

In sum, this exercise suggests that the effects of the temporary increase in government transfers and tax cuts implied by the ARRA may be significantly smaller than expected by Romer and Bernstein. The two models I have considered suggest an effect between zero and 30 basis points on GDP per one per cent of GDP increase in transfers. A possible concern is that the share of rule-of-thumb households increased during the course of the financial crisis. The argument goes as follows. A standard justification for hard-wiring rule-of-thumb households in macro models is to capture borrowing constraints. Households that desire to borrow but are credit constrained would increase consumption along with increases in current disposable income. The number of such households might then have increased during the financial crisis, because banks were more reluctant to extend credit. Alternatively, it is also possible that the share of consumers who wanted to borrow declined during the recession. In particular, households that expect a lasting reduction in life-time income, because of less promising job opportunities, asset losses, sustained unemployment, or higher taxes, may decide to save more. In this manner, some of those households that were borrowing-constrained before may now want to save rather than spend any additional income they might receive from the government.

The tax rebate offered by the Bush administration in spring 2008 and similar tax rebates or credits by the Obama administration offered in the context of the ARRA in spring 2009 were the focus of a recent survey by Claudia Sahm and Slemrod (2010) that may shed some light on the direction of this effect. They write that 25 per cent of households reported that the one-time economic stimulus payment in 2008 led them mostly to increase their spending. In 2009, 13 per cent of households reported that they had mostly spent the extra pay from the lower withholding. This finding, taken together with the above model-based analysis, may help to explain the behaviour of aggregate consumption and income. As pointed out by Taylor (2009) and (2010a), the rebate payments are directly apparent as upward spikes in aggregate disposable income in May and June 2008 and 2009, while aggregate consumption growth in those periods is relatively smooth and flat.

5 Conclusions

In this chapter, I have presented a proposal for a comparative approach to macroeconomic policy analysis that is open to competing modelling paradigms. I have reviewed recent work on building a macroeconomic model archive and platform that make it much easier to conduct extensive model comparisons. In addition, I have pointed towards a range of competing modelling approaches that should be the focus of a systematic model comparison exercise in the future.

To illustrate the use of empirical benchmarks in a model competition, I have reviewed findings from a recent model comparison by Wieland and Wolters (2010) in terms of forecasting performance during US recessions and recoveries. This comparison has indicated that models and experts tend to miss the onset of recessions. Both, models and experts have some forecasting power during recoveries. Interestingly, several of the state of the art models performed better than many experts over a horizon of three to four quarters. Thus, there is no reason for forecasting professional using time series methods or traditional Keynesian-style models to dismiss modern DSGE models.

Model comparison can be a very valuable tool for increasing the robustness of policy recommendations. To illustrate what is meant by policy robustness, I have reviewed recent findings regarding the effectiveness of temporary fiscal stimulus measures from a range of models and studies. I found that a range of model comparisons suggests a significantly smaller impact of the government spending planned under the American Recovery and Reinvestment of 2009 as initially projected by Administration economists in Romer and Bernstein (2009). From these findings I conclude that arguments in favour of fiscal stimulus that are based on the supposed multiplicative effect of government purchases should be viewed with substantial scepticism. Furthermore, given the uncertainty about the appropriate macroeconomic model of the economy, policy analysis needs to take

into account a range of models, including New-Keynesian DSGE models rather than relying only on more traditional Keynesian-style models as in Romer and Bernstein (2009).

References

Adolfson, M., S. Lasen, J. Lind, and M. Villani (2007), 'Bayesian estimation of an open economy DSGE model with incomplete pass-through', *Journal of International Economics*, 72, pp. 481–511.

Akerlof, G., and R. J. Shiller (eds) (2010), *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*, Princeton, N.J.: Princeton University Press.

Altig, D.E., L.J. Christiano, M. Eichenbaum and J. Lind (2005), 'Firm-specific capital, nominal rigidities and the business cycle', CEPR Discussion Papers 4858.

Angeloni, I., and E. Faia (2010), 'Capital regulation and monetary policy with fragile banks'. Manuscript, University of Frankfurt

Bernanke, B., M. Gertler and S. Gilchrist (1999), 'The financial accelerator in a quantitative business cycle framework', in J. B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics Volume 1C*, Amsterdam: Elsevier Science, North-Holland.

Branch, W.A. and G.W. Evans (2011), 'Monetary policy with heterogeneous expectations', *Economic Theory*, forthcoming.

Branch, W.A. and B. McGough (2011), 'Business cycle amplification with heterogeneous expectations', *Economic Theory*, forthcoming.

Brock, W. and C. Hommes (1998), 'Heterogeneous beliefs and routes to chaos in a simple asset pricing model', *Journal of Economic Dynamics and Control*, 22, pp. 1235–74.

Bryant, R., D. Currie, J. Frenkel, P. Masson, and R. Portes (eds) (1989), *Macroeconomic Policies in an Interdependent World*, Washington, D.C.: The Brookings Institution.

Bryant, R., D.W. Henderson, G. Holtham, P. Hooper, and S.A. Symansky, (eds) (1988), *Empirical Macroeconomics for Interdependent Economies*, Washington, D.C.: The Brookings Institution.

Bryant, R., P. Hooper, and C. Mann, (eds) (1993), *Evaluating Policy Regimes: New Research in Empirical Macroeconomics*, Washington, D.C.: The Brookings Institution.

Capistran, C. and A. Timmermann (2009), 'Disagreement and biases in inflation expectations', *Journal of Money, Credit, and Banking*, 41, pp. 365–96.

Carabenciov, I., I. Ermolaev, C. Freedman, M. Juillard, O. Kamenik, D. Korshunov and D. Laxton (2008), 'A small quarterly projection model of the US economy', IMF Working Paper 08/278.

Chiarella, C., R. Dieci and X.-Z. He (2007), 'Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework', *Journal of Economic Behavior and Organization*, 62, pp. 408–27.

Christensen, I. and A. Dib, (2008), 'The financial accelerator in an estimated New Keynesian model', *Review of Economic Dynamics*, 11, pp.155–78.

Christiano, L., R. Motto and M. Rostagno (2010), 'Financial factors in economic fluctuations', ECB Working Paper 1192.

Christiano, L.J., M. Eichenbaum and C.L. Evans (2005), 'Nominal rigidities and the dynamic effects of a shock to monetary policy', *Journal of Political Economy*, 113(1), pp. 1–45.

Clarida, R., J. Gal and M. Gertler (1999), 'The science of monetary policy: A New Keynesian perspective', *Journal of Economic Literature*, 37(4), pp. 1661–707.

Clarida, R., J. Gal and M. Gertler (2002), 'A simple framework for international monetary policy analysis', *Journal of Monetary Economics*, 49, pp. 879–904.

Claudia Sahm, M.S. and J. Slemrod (2010), 'Check in the mail or more in the paycheck: Does the effectiveness of fiscal stimulus depend on how it is delivered?', NBER Working Paper 16246.

Coenen, G., C.J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Lind, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt and J. in't Veld (2011), 'Effects of fiscal stimulus in structural models',, forthcoming in, *American Economic Journal: Macroeconomics*.

Coenen, G., P. McAdam and R. Straub (2008), 'Tax reform and labour-market performance in the euro area: A simulation-based analysis using the new area-wide model, *Journal of Economic Dynamics & Control*, 32 (8), pp. 2543–83.

Coenen, G. and V. Wieland (2002), 'Inflation dynamics and international linkages: A model of the United States, the euro area and Japan', ECB Working Paper Series 181.

Coenen, G. and V. Wieland (2005), 'A small estimated euro area model with rational expectations and nominal rigidities' *European Economic Review*, 49, pp.1081–104.

Cogan, J.F., T. Cwik, J.B. Taylor and V. Wieland (2009), 'New Keynesian versus old Keynesian government spending multipliers', NBER Working Paper Series No. 14782.

Cogan, J. F., T. Cwik, J.B. Taylor and V. Wieland (2010), 'New Keynesian versus old Keynesian government spending multipliers', *Journal of Economic Dynamics and Control*, 34, pp. 281–95.

Cogan, J.F. and J.B. Taylor (2010), 'What the government purchases multiplier actually multiplied in the 2009 stimulus package', NBER Working Paper 16505.

Cwik, T. and Wieland, V. (2009). Keynesian government spending multipliers and spillovers in the euro area. CEPR Discussion Paper 7389.

Cwik, T. and V. Wieland (2011), 'Keynesian government spending multipliers and spillovers in the euro area, *Economic Policy*, 26, pp. 493-549.

De Fiore, F., P. Teles, and O. Tristani (2009), 'Monetary policy and the financing of firms', ECB Working Paper 1123.

De Grauwe, P. (2011), 'Animal spirits and monetary policy', *Economic Theory*, forthcoming.

DeWalque, G., O. Pierrard, and A. Rouabah (2010), 'Financial (in)stability, supervision and liquidity injections: A dynamic general equilibrium approach', *Economic Journal*, Royal Economic Society, 120 (549), pp. 1234–61.

De Graeve, F. (2008), 'The external finance premium and the macroeconomy: US post-WWII evidence', *Journal of Economic Dynamics and Control*, 32, pp. 3415–440.

Deissenberg, C., H. Dawid, and S. van der Hoog (2008), 'Eurace: A massively parallel agent-based model of the European economy', *Applied Mathematics and Computation*, 204, pp. 541–52.

Del Negro, M. and F. Schorfheide (2004), 'Priors from general equilibrium models for VARs', *International Economic Review*, 45 (2), pp. 643–73.

Diamond, P. and H. Vartiainen, (eds) (2007), *Behavioral Economics and Its Applications*, Princeton, NJ: Princeton University Press.

Dib, A. (2010), 'Banks, credit market frictions, and business cycles', Bank of Canada, Working Paper 2010-24.

Dieppe, A., K. Kuester and P. McAdam (2005), 'Optimal monetary policy rules for the euro area: An analysis using the area wide model', *Journal of Common Market Studies*, 43 (3), pp. 507–37.

Edge, R.M., M.T. Kiley and J.-P. Laforte (2007), 'Documentation of the Research and Statistics Division's estimated DSGE model of the U.S. economy: 2006 version', Finance and Economics Discussion Series 07-53, Board of Governors of the Federal Reserve System.

Edge, R.M., M.T. Kiley and 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*, forthcoming.

Erceg, C.J., L. Guerrieri and C. Gust (2008), 'Trade adjustment and the composition of trade', *Journal of Economic Dynamics & Control*, 32, pp. 2622–650.

Evans, G. and Honjapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Fagan, G., J. Henry and R. Mestre (2005), 'An area-wide model for the euro area',. *Economic Modelling*, 22, pp. 39–59.

Freedman, C., M. Kumhof, D. Laxton, D. Muir and S. Mursula (2010), 'Global effects of fiscal stimulus during the crisis', *Journal of Monetary Economics*, 57, pp. 506–26.

Fuhrer, J.C. (1997), 'Inflation/output variance trade-offs and optimal monetary policy', *Journal of Money, Credit and Banking*, 29 (2), pp. 214–34.

Fuhrer, J.C. and G. Moore (1995), 'Inflation persistence', *The Quarterly Journal of Economics*, 110 (1), pp. 127–59.

Gerali, A., S. Neri, L. Sessa and F. Signoretti (2010), 'Credit and banking in a DSGE model of the euro area', *Journal of Money, Credit and Banking*, 42 (6), supplement.

Gertler, M., S. Gilchrist and F. Natalucci (2007), 'External constraints on monetary policy and the financial accelerator', *Journal of Money, Credit and Banking*, 39 (2-3), pp. 295–330.

Gertler, M. and P. Karadi (2009), 'A model of unconventional monetary policy' mimeo.

Gertler, M. and N. Kiyotaki (2009), 'Financial intermediation and credit policy in business cycle analysis', (in preparation for the *Handbook of Monetary Economics*).

Giannone, D., F. Monti and L. Reichlin (2009), 'Incorporating conjunctural analysis in structural models', in V. Wieland (ed.), *The science and practice of monetary policy today*, Berlin: Springer Science, pp. 41–57.

Giordani, P. and Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, 47:1037–1059.

Goodfriend, M. and B.T. McCallum (2007), 'Banking and interest rates in monetary policy analysis: A quantitative exploration', *Journal of Monetary Economics*, 54 (5), pp. 1480–507.

Gouvea, S., A. Minella, , R. Santos and N. Souza-Sobrinho, (2008), 'Samba: Stochastic analytical model with a Bayesian approach', Manuscript.

Hughes-Hallett, A. and K.F. Wallis (eds) (2004), 'EMU Macroeconomic Model Comparison Exercise for the Euroconference 7-8 June 2002', Economic Modelling 21 (5).

Iacoviello, M. (2005), 'House prices, borrowing constraints, and monetary policy in the business cycle', *The American Economic Review*, 95 (3), pp. 739–64.

Juillard, M. (1996). Dynare: A program for the resolution and simulation of dynamic models with forward variables through the use of a relaxation algorithm. CEPREMAP working paper, 9602.

Klein, L. (ed.) (1991), *Comparative Performance of U.S. Econometric Models*, Oxford, England, Oxford University Press.

Krugman, Paul (2009), "Paul Krugman's London lectures. Dismal Science. The Nobel laureate speaks on the crisis in the economy and in economics", *The Economist*, 11 June.

Kuester, K. (2010), 'Insurance policies for monetary policy in the euro area', *Journal of the European Economic Assocation*, 8, pp. 872–912.

Kurz, M. (1994a), 'On rational belief equilibria', *Economic Theory*, 4, pp. 859–76.

Kurz, M. (1994b), 'On the structure and diversity of rational beliefs', *Economic Theory*, 4, pp. 877–900.

Kurz, M. (1996), 'Rational beliefs and endogenous uncertainty: an introduction', *Economic Theory*, 8, pp. 383–97.

Kurz, M. (1997a), 'Endogenous economic fluctuations and rational beliefs: A general perspective', in M. Kurz (ed.), *Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs*, Springer Series in Economic Theory, (6), Berlin, Heidelberg, New York: Springer Verlag.

Kurz, M. (ed.) (1997b), *Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs*, Springer Series in Economic Theory, (6), Berlin, Heidelberg, New York: Springer Verlag.

Kurz, M. (2009), 'Rational diverse beliefs and market volatility', in T. Hens and K. Schenk-Hoppe (eds), *Handbook of financial markets: dynamics and evolution*, Amsterdam: North Holland.

Kurz, M., H. Jin and M. Motolese (2003), *Knowledge, Information and Expectations in Modern Macroeconomics: Essays In Honor of Edmund S. Phelps*, Chapter 10: 'Endogenous Fluctuations and the Role of Monetary Policy', Princeton, N.J.: Princeton University Press, pp. 188 –227.

Kurz, M., H. Jin and M. Motolese (2005), 'The role of expectations in economic fluctuations and the efficacy of monetary policy', *Journal of Economic Dynamics & Control*, 29, pp. 2017–65.

Laxton, D. and P. Pesenti (2003), 'Monetary rule for small, open, emerging economies', *Journal of Monetary Economics*, 50, pp. 1109–146.

Levin, A., V. Wieland and J.C. Williams (2003), 'The performance of forecast-based monetary policy rules under model uncertainty', *The American Economic Review*, 93 (3), 622–45.

Mankiw, N.G. and R. Reis (2007), 'Sticky information in general equilibrium', *Journal of the European Economic Association*, 5 (2-3), pp. 603–13.

McCallum, B. and E. Nelson (1999), 'Performance of operational policy rules in an estimated semi-classical structural model', in J.B. Taylor (ed.), *Monetary Policy Rules*, Chicago: University of Chicago Press.

Medina, J.P. and C. Soto (2007), 'The Chilean business cycles through the lens of a stochastic general equilibrium model', Central Bank of Chile Working Paper 457.

Meh, C.A. and K. Moran (2010), 'The role of bank capital in the propagation of shocks', *Journal of Economic Dynamics and Control*, 34 (3), pp. 555–76.

Murchison, S. and A. Rennison (2006), 'ToTEM: The Bank of Canada's new quarterly projection model', Bank of Canada Technical Report No. 97.

Nolan, C. and C. Thoenissen (2009), 'Financial shocks and the US business cycle', *Journal of Monetary Economics*, 56, pp. 596–604.

Orphanides, A. (2003), 'The quest for prosperity without inflation', *Journal of Monetary Economics*, 50, pp. 633–63.

Orphanides, A. and V. Wieland (1998), 'Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero', Finance and Economics Discussion Series 98–35, Board of Governors of the Federal Reserve System.

Orphanides, A. and J.C. Williams (2006), 'Monetary policy under imperfect knowledge', *Journal of the European Economic Association*, 4, pp. 366–375.

Ratto, M., W. Roeger and J. in 't Veld (2009), 'QUEST III: An estimated open-economy DSGE model of the euro area with fiscal and monetary policy', *Economic Modelling*, 26 (1), pp. 222–33.

Reifschneider, D., R. Tetlow and J.C. Williams (1999), 'Aggregate disturbances, monetary policy and the macroeconomy: The FRB/US perspective', *Federal Reserve Bulletin*, 85 (1), pp. 1–19.

Roeger,W. and J. in't Veld (2009), 'Fiscal policy with credit constrained households', Economic Papers 357, European Commission.

Romer, C. and J. Bernstein (2009), 'The job impact of the American Recovery and Reinvestment Plan', 8 January.

Rotemberg, J.J. and M. Woodford (1997), 'An optimization-based econometric framework for the evaluation of monetary policy', *NBER Macroeconomics Annual*, 12, pp. 297–346.

Rudebusch, G.D. and L.E.O. Svensson (1999), 'Policy rules for inflation targeting', in J.B. Taylor (ed.), *Monetary Policy Rules*, Chicago: University of Chicago Press.

Slobodyan, S. and R. Wouters (2008), 'Estimating a medium-scale DSGE model with expectations based on small forecasting models', mimeo.

Smets, F. and R. Wouters (2003), 'An estimated dynamic stochastic general equilibrium model of the euro area', *Journal of the European Economic Association*, 1 (5), pp. 1123–75.

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

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

Taylor, J.B. (1993b), *Macroeconomic Policy in a World Economy*, New York: W.W. Norton.Taylor, J.B. (1999), *Monetary Policy Rules*, Chicago: The University of Chicago Press.

Taylor, J.B. (2009), 'The lack of an empirical rationale for a revival of discretionary fiscal policy', *American Economic Review, Papers and Proceedings*, 99 (2), pp. 550–55.

Taylor, J. (2010a), 'Getting back on track: Policy lessons from the financial crisis', *Federal Reserve Bank of St. Louis Review*, 92, pp. 165–76.

Taylor, J.B. (2010b), 'Comment on "global effects of fiscal stimulus during the crisis"', *Journal of Monetary Economics*, 57, pp. 527–30.

Taylor, J.B. and V. Wieland (2011), 'Surprising comparative properties of monetary models: Results from a new data base', forthcoming in *The Review of Economics and Statistics*.

Tesfatsion, L. and Judd, K. L., editors (2006). *Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics*. North Holland.

Wieland, V. (2009), 'Learning, endogenous indexation and disinflation in the New-Keynesian model', in C. Walsh and K. Schmidt-Hebbel (eds), *Monetary Policy under Uncertainty and Learning*, Santiago: Central Bank of Chile.

Wieland, V., T. Cwik, G.J. Müller, S. Schmidt and M. Wolters (2009), 'A new comparative

approach to macroeconomic modelling and policy analysis', Manuscript, Center for Financial Studies, Frankfurt.

Wieland, V. and M. Wolters (2010), 'The diversity of forecasts from macroeconomic models of the U.S. economy', *Economic Theory*,47, pp. 247-292.