Essays on
Business Cycle Models, Forecasting and Monetary Policy

Inaugural-Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Wirtschaftswissenschaften der
Johann Wolfgang Goethe-Universität
Frankfurt am Main

JOHANN WOLFGANG GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

vorgelegt von
Maik Hendrik Wolters
aus Hannover
Juni 2010

Erstgutachter:	Prof. Volker Wieland, Ph.D.
Zweitgutachter:	Prof. Stefan Gerlach, Ph.D.
Tag der Promotion:	4. Oktober 2010

List of Original Working Papers

This thesis consists of the following working papers:

1.) Volker Wieland, Tobias Cwik, Gernot J. Müller, Sebastian Schmidt and Maik H. Wolters (2009):
A New Comparative Approach to Macroeconomic Modeling and Policy Analysis
Manuscript, Center for Financial Studies, Frankfurt.

2.) Volker Wieland and Maik H. Wolters (2010):
The Diversity of Forecasts from Macroeconomic Models of the U.S. Economy[1]
CEPR Discussion Paper No. 7870 and CFS Working Paper No. 2010/08.

3.) Maik H. Wolters (2010):
Forecasting under Model Uncertainty
Working Paper, Goethe University Frankfurt.

4.) Tobias Cwik, Gernot J. Müller and Maik H. Wolters (2008):
Does Trade Integration Alter Monetary Policy Transmission?
CFS Working Paper No. 2008/29.

5.) Maik H. Wolters (2009):
Estimating Monetary Policy Reaction Functions Using Quantile Regressions
Working Paper, Goethe University Frankfurt.

[1]This paper is forthcoming in *Economic Theory*.

Abstract

This dissertation introduces in chapter 1 a new comparative approach to model-based research and policy analysis by constructing an archive of business cycle models. It includes many well-known models used in academia and at policy institutions. A computational platform is created that allows straightforward comparisons of models' implications for monetary and fiscal stabilization policies. Chapter 2 applies business cycle models to forecasting. Several New Keynesian models are estimated on historical U.S. data vintages and forecasts are computed for the five most recent recessions. The extent of forecast heterogeneity for models and professional forecasts is analysed. Chapter 3 extends the forecasting analysis to a long sample and to the evaluation of density forecasts. Weighted forecasts are computed using a variety of weighting schemes. The accuracy of forecasts is evaluated and compared to professional forecasts and forecasts from nonstructural time series methods. Chapter 4 adds a new feature to existing business cycle models. Specifically, a medium-scale New Keynesian model is constructed that allows for strategic complementarities in price-setting. The role of trade integration for monetary policy transmission is explored. A new dimension of the exchange rate channel is highlighted by which monetary policy directly impacts domestic inflation. Chapter 5 tests whether simple symmetric monetary policy rules used in most business cycle models are a sufficient description of reality. I use quantile regressions to estimate policy parameters and find asymmetric reactions to inflation, the output gap and past interest rates.

Contents

Introduction **1**

Zusammenfassung **15**

1 A New Comparative Approach to Macroeconomic Modeling and Policy Analysis **29**
1.1 Introduction . . . 30
1.2 A general approach to model comparison . . . 32
1.2.1 Augmenting models for the purpose of comparison . . . 32
1.2.2 Conducting a comparison . . . 38
1.3 A data base of macroeconomic models . . . 40
1.4 Comparing monetary and fiscal policies across models: an example . . . 44
1.5 Conclusion . . . 49
References . . . 50
A.1 A detailed overview of available models . . . 54
A.1.1 Small calibrated models . . . 54
A.1.2 Estimated U.S. models . . . 57
A.1.3 Estimated Euro area models . . . 62
A.1.4 Estimated/calibrated multi-country models . . . 67

2 The Diversity of Forecasts from Macroeconomic Models of the U.S. Economy **73**
2.1 Introduction . . . 74
2.2 Forecasting models . . . 76
2.3 Forecasting methodology . . . 79
2.4 An illustration: forecasting the 2001 recession . . . 82
2.5 Model-based versus expert nowcasts and the 2008/09 recession . . . 86
2.6 The relative accuracy of model-based and expert forecasts . . . 89
2.7 The heterogeneity of model-based and expert forecasts . . . 94
2.8 Conclusions . . . 98
References . . . 102
A.1 The macroeconomic models used to compute forecasts . . . 107
A.2 The quarterly vintage database . . . 118

3 Forecasting under Model Uncertainty **121**
3.1 Introduction . . . 121
3.2 Forecasting models . . . 124
3.3 A real-time dataset . . . 127
3.4 Forecasting methodology . . . 129
3.5 Forecast evaluation . . . 135
3.6 Model averaging . . . 139
3.7 Forecast evaluation of combined forecasts . . . 142
3.8 Density forecast evaluation . . . 146
3.9 Conclusion . . . 149

References . . . 151
A.1 Additional results . . . 155

4 Does Trade Integration Alter Monetary Policy Transmission? 161
4.1 Introduction . . . 161
4.2 Model . . . 164
4.2.1 Final goods firms . . . 165
4.2.2 Intermediate good firms . . . 167
4.2.3 Households . . . 168
4.2.4 Monetary policy . . . 169
4.2.5 Model solution . . . 169
4.2.6 The exchange rate channel revisited . . . 170
4.3 Estimation . . . 171
4.3.1 Empirical impulse response functions . . . 171
4.3.2 Estimation of general equilibrium model . . . 172
4.3.3 Parametric setup . . . 174
4.3.4 Results . . . 175
4.4 The role of openness in monetary policy transmission . . . 177
4.4.1 The role of openness . . . 177
4.4.2 Implications for monetary policy . . . 179
4.5 Conclusion . . . 181
References . . . 183
A.1 The New Keynesian Phillips curve . . . 187
A.1.1 Pricing problem of LCP-firm . . . 187
A.1.2 Pricing problem of PCP-firm . . . 188
A.1.3 New Keynesian Phillips Curve . . . 190
A.2 Data . . . 191

5 Estimating Monetary Policy Reaction Functions Using Quantile Regressions 193
5.1 Introduction . . . 193
5.2 Data . . . 195
5.3 Least squares regressions . . . 196
5.3.1 Specification tests . . . 197
5.3.2 Least squares estimation results . . . 198
5.4 Quantile regression . . . 199
5.4.1 Inverse quantile regression . . . 200
5.4.2 Moving blocks bootstrap . . . 201
5.5 Estimation results . . . 202
5.5.1 Robustness . . . 205
5.6 Decomposing deviations from policy rules . . . 206
5.7 Conclusion . . . 209
References . . . 211

List of Figures

1.1 Negative monetary policy shock 45
1.2 Autocorrelation functions 47
1.3 Fiscal policy shock 48

2.1 Real GDP growth forecast at the start of the 2001 recession 83
2.2 Real GDP growth forecast for the 2001 recession 85
2.3 Real output growth forecast during the 2001 recession. 87
2.4 Real output growth forecast during the 2007-2009 recession. 88
2.5 Standard deviations of output growth forecasts: experts (solid) and models (dashed) . 94
2.6 Standard deviations of inflation forecasts: experts (solid) and models (dashed) 95

3.1 Structural forecasts: data vintage May 12, 2000 133
3.2 Forecast errors and output growth rates 134
3.3 Weighted structural forecasts: data vintage May 12, 2000 143
3.4 Evaluation of structural density forecasts: 1984 - 2000 147
3.5 Evaluation of structural density forecasts: 1984 - 2000 148

4.1 Effects of a monetary policy shock 173
4.2 Demand function for intermediate goods 176
4.3 Impulse responses to a monetary policy shock 178
4.4 Openness and pass-through for the U.S. 180

5.1 Federal funds rate, inflation forecasts and output gap nowcasts 196
5.2 Estimated coefficients ($\alpha_i = 0$) 202
5.3 Estimated coefficients ($\alpha_i \neq 0$) 204
5.4 Fed funds rate, policy rule, quantiles and deviation decomposition ($\alpha_i = 0$) 207
5.5 Fed funds rate, policy rule, quantiles and deviation decomposition ($\alpha_i \neq 0$) 209

List of Tables

1.1 Model-specific variables, parameters, shocks and equations . . . 33
1.2 Comparable common variables, parameters, shocks and equations . . . 34
1.3 Comparable common variables . . . 35
1.4 Model 1 - The hybrid model of Clarida, Galí and Gertler (1999) (NK_CGG99) . . . 36
1.5 Model 2 - The New Keynesian model of Woodford (2003) (NK_RW97) . . . 37
1.6 Models currently available in the data base . . . 43
1.7 Policy rules . . . 44

2.1 Model overview . . . 77
2.2 State space representation and model equations . . . 80
2.3 RMSEs of output growth forecasts . . . 90
2.4 RMSEs of inflation forecasts . . . 91
2.5 RMSEs of output growth forecasts initialized with expert nowcasts . . . 92
2.6 RMSEs of inflation forecasts initialized with expert nowcasts . . . 93
2.7 RMSE of best, worst, and average output growth forecaster from survey and models . 96
2.8 Best, worst, and average inflation forecaster from survey and models . . . 97

3.1 Model overview . . . 127
3.2 Greenbook RMSE and relative RMSE of model forecasts: 1984-2000 . . . 136
3.3 Greenbook RMSE and relative RMSE of weighted model forecasts: 1984-2000 . . . 145
3.4 Percentage of periods alternative forecast better than Greenbook: 1984-2000 . . . 155
3.5 Percentage of periods weighted forecast better than Greenbook: 1984-2000 . . . 156
3.6 Combination weights for data vintage May 12, 2000: output growth . . . 157
3.7 Combination weights for data vintage May 12, 2000: inflation . . . 158
3.8 Combination weights for data vintage May 12, 2000: federal funds rate . . . 159

4.1 Estimated parameter values of DSGE model . . . 175
4.2 Monetary policy trade-off . . . 181

5.1 p-values of subsample stability tests . . . 197
5.2 p-values of tests for exogeneity . . . 198
5.3 Estimated policy reaction parameters . . . 198

Introduction

Many macroeconomic models that attempt to explain the behavior of the main economic variables over the business cycle have been developed in recent years. Model builders include not only academics but also researchers at central banks, treasuries and international organizations. If one model were to be found to dominate all others in terms of theoretical appeal and empirical fit, this model could be used exclusively to develop policy recommendations. Yet, there is no agreement on a best approach to macroeconomic modeling.
Theory based dynamic stochastic general equilibrium (DSGE) models that are consistently derived from microeconomic optimization problems of households and firms have become the workhorse of modern monetary economics. However, critics argue sharply against using DSGE models and suggest to go back to earlier generation models. While several competing models describe historical data of key aggregates reasonably well, based on different theoretical approaches macroeconomic models have a different economic structure with different implications for policy analysis. To derive reliable policy recommendations from macroeconomic models one needs to compare the findings from several models to establish "robustness" of policy recommendations. Such an approach is recommended by McCallum (1988), McCallum (1999), Blanchard and Fischer (1989), Taylor (1999) and many others. Comparing empirical predictions of different models is difficult and rare, and evaluating the performance of different policies across many models typically is work intensive and costly. The six comparison projects reported in Bryant et al. (1988), Bryant et al. (1989), Klein (1991), Bryant et al. (1993), Taylor (1999) and Hughes-Hallett and Wallis (2004) have involved multiple teams of researchers, each team working only with one or a small subset of available models. While these initiatives have helped produce some very influential insights such as the Taylor rule,[2] the range of systematic, comparative findings has remained limited.

This dissertation provides a new comparative approach to model-based research that enables individual researchers to conduct model comparisons easily, frequently, at low cost and on a large scale. Using this approach an archive of business cycle models is built that includes many well-known empirically estimated models that may be used for quantitative analysis of monetary

[2] Taylor (1993a) credits the comparison project summarized in Bryant et al. (1993) as the crucial testing ground for what later became known as the Taylor rule.

and fiscal stabilization policies. Building on this comparative approach this dissertation includes two applications that compare the predictive ability of several macroeconomic models. Finally, two chapters analyse specific aspects of macroeconomic models. One studies the role of trade integration for monetary policy transmission in a New Keynesian model and the other provides an empirical test that shows whether simple linear monetary policy rules used in most business cycle models are a sufficient description of reality.
Chapter 1 introduces a computational platform and an archive of business cycle models. This archive allows the simulation of macroeconomic models and model comparison based on statistics like impulse response functions and autocorrelation functions. While the models in the archive are based on the model parameters that are provided in the original references of the specific models, in chapter 2 several models are estimated on historical U.S. data vintages. Forecasts are computed for the five most recent recessions as defined by the NBER. The extent of forecasting heterogeneity is analysed and compared to forecasts from the Survey of Professional Forecasters (SPF). Chapter 3 extends the forecasting evaluation analysis to a longer sample and to density forecasts. It is shown that combining forecasts from several models can increase the accuracy of forecasts. Chapter 4 shows an example how to build a macroeconomic model. A medium-scale two country New Keynesian model is developed that allows for strategic complementarities in price-setting. The role of trade integration together with strategic price complementarities for monetary policy transmission is analysed. Finally, chapter 5 studies simple monetary policy rules of the type usually assumed in macroeconomic models. Monetary policy rules are estimated without any restrictive model assumptions about other economic dynamics. I use quantile regressions to estimate policy parameters over the whole conditional distribution of the interest rate. I find that simple symmetric rules might be too restrictive to reflect actual monetary policy.

Chapter 1, which is joined work with Volker Wieland, Tobias Cwik, Gernot J. Müller and Sebastian Schmidt introduces a database of macroeconomic business cycle models. It enables individual researchers to conduct systematic model comparisons and policy evaluations. A general class of nonlinear dynamic stochastic macroeconomic models is augmented with a space of common comparable variables, parameters and shocks to allow for a systematic comparison of particular model characteristics. On this basis, common policy rules can be defined and their implications can be compared across models. Comparison is based on objects such as impulse response functions, autocorrelation functions and unconditional distributions of key macroeconomic aggregates.
The database 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 (cf. Clarida et al. (1999), Rotemberg and Woodford (1997), Fuhrer and Moore (1995), McCallum and Nelson (1999), Coenen and Wieland (2005), etc). Many others are of medium scale and cover many key macroeconomic aggregates (cf. Christiano et al. (2005), Coenen et al. (2004), Smets and Wouters (2003, 2007)). Some models in the data base are fairly large in scale such as the Federal Reserve's FRB-US model of Reifschneider et al. (1999), the model of the G7 economies of Taylor (1993b) or the ECB's area-wide model of Dieppe et al. (2005). 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 Keynesian models fully incorporate recent advances in terms of microeconomic foundations. Well-known examples of this class of models are Christiano et al. (2005), Smets and Wouters (2003, 2007), Laxton and Pesenti (2003) and Adolfson et al. (2007). However, some models that assign little role to forward-looking behavior by economic agents (cf. the ECB's area-wide model) or none at all (cf. Rudebusch and Svensson (1999) and Orphanides (2003)) are included into the database as well.

The model database is augmented with a computational platform. It allows users to solve structural models and conduct comparative analysis. Comparisons of impulse response functions of common variables in response to common shocks, or of autocorrelation functions of common variables in response to model-specific shocks, or of unconditional distributions of common variables are generated. It can also be used to conduct a systematic investigation of policy rules across models. This platform accepts specific economic policy rules as common comparable input for multiple economic models. It generates as output a comparison across models of statistics describing characteristics of the main macroeconomic variables, which are predicted to result from these policies according to different economic models. The platform admits nonlinear as well as linear models and allows for perturbation-based approximation of nonlinear models with forward-looking variables.
The comparative approach to modeling and policy analysis is illustrated with several examples. Impulse responses to monetary and fiscal policy shocks are compared under alternative monetary policy rules, and the predictions of different models and different policies for inflation and output persistence are investigated. Important differences of the monetary policy transmission mechanism are found between small and large models. Different modeling philosophies like New Keynesian models with microeconomic foundations and larger disaggregated models with less strict theoretical foundations appear to be the reasons for differences in the magnitude and length of responses of core variables like output and inflation to fiscal and monetary policy shocks.
New models may easily be introduced into the model database and compared to established benchmarks thereby fostering a comparative rather than insular approach to model building. Wide application of this approach could help improve the replicability of quantitative macroeconomic analysis, reduce the danger of circular developments in model-based research and strengthen the robustness of policy recommendations.

Chapter 2 extends the comparative approach to macroeconomic modeling in an important direction. The models in the model archive introduced in chapter 1 have been implemented with parameters as estimated or calibrated by the respective original authors. Therefore, the model database provides no measure that shows which model is best suited to analyse a specific data sample. In chapter 2 three small and two large New Keynesian DSGE models are linked to a common dataset. Model parameters are estimated and forecasts are computed. The accuracy of forecasts can be viewed as a measure of data fit. It shows to what extent business cycle dynamics can be explained by these models. Specifically, the accuracy of output growth and inflation forecasts is analysed. Besides evaluating forecasts, the focus of chapter 2 is to quantify the heterogeneity of model forecasts and compare them to survey forecasts in order to learn more about the extent, dynamics and sources of forecast heterogeneity.
Recent empirical studies have documented substantial variations in the accuracy and heterogeneity of expert forecasts of GDP and inflation (see Kurz et al. (2003, 2005), Giordani and Söderlind (2003), Kurz (2009) and Capistran and Timmermann (2009)). At the same time, 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 (c.f. Kurz (1994a,b, 1996, 1997a,b, 2009), Brock and Hommes (1998), Kurz et al. (2005), Chiarella et al. (2007), Branch and McGough (2010), Branch and Evans (2010) and de Grauwe (2010)).
Forecast heterogeneity arises for several reasons. First of all, forecasters need a forecast-generating framework. Such a framework may be a fully developed economic structure, a non-structural collection of statistical relationships or a simple rule-of-thumb. The particular modeling assumptions embedded in this forecasting framework represent an important source of belief heterogeneity. Another source of heterogeneity is the information used by the forecaster. Information sets may differ

in terms of the number of economic aggregates or prices for which the forecasters collect data and the timeliness of the data vintage. The data is needed to estimate the state of the economy and the parameters of the forecasting framework.

While expert forecasts are published in various surveys, the underlying modeling assumptions, information sets and parameter estimates are not publicly available. Instead, in this chapter different macroeconomic models of the U.S. economy are used to generate output and inflation forecasts. The precision and diversity of expert forecasts from the SPF and the Federal Reserve's Greenbook are used as benchmarks for comparison. This comparison is conducted for successive quarter-by-quarter forecasts up to four quarters into the future during the five most recent recessions of the U.S. economy as dated by the NBER. Periods around recessions pose the greatest challenge for economic forecasters, and arguably expectational heterogeneity may itself play a role in these shifts in economic activity.

The mean model forecast comes surprisingly close to the mean of all forecasts collected in the SPF and to Greenbook forecasts in terms of accuracy even though the models only make use of a small number of data series. Model forecasts compare particularly well to professional forecasts at a horizon of three to four quarters and during recoveries. The extent of forecast heterogeneity is similar for model and professional forecasts but varies substantially over time. Of course, the models used by professional forecasters may differ from the models used in this chapter. While the particular reasons for diversity in professional forecasts are not observable, the diversity in model forecasts can be traced to different modeling assumptions, information sets and parameter estimates. These three sources of disagreement are found to be sufficient to generate an extent of heterogeneity that is similar to the heterogeneity observed among expert forecasts. Furthermore, the recursive updating of model parameter estimates with incoming data induces dynamics in model forecast heterogeneity. Expert forecast diversity even exhibits roughly similar variations. Thus, the findings of chapter 2 can be taken as an indication that much of the observed time variation in forecast heterogeneity may be explained by disagreement about appropriate modeling assumptions and differences in parameter estimates rather than irrationality of particular forecasters. This belief diversity itself may be a source of volatility. Of course, the models used in this chapter would attribute such volatility to shocks or other propagation mechanisms rather than endogenous heterogeneity in beliefs. Models with heterogenous expectations provide an avenue for distinguishing this source of economic fluctuations from other candidate propagation mechanisms.

Chapter 3 extends the analysis of chapter 2 to a longer evaluation sample from 1984 to 2000 including periods of high and low volatility. While the focus of chapter 2 is the evaluation of forecast accuracy around business cycle turning points, the diversity of forecasts and how these are linked to business cycles theories with hetereogenous expectations, chapter 3 is a pure forecast evaluation exercise.

In recent years, researchers such as Smets and Wouters (2004), Adolfson et al. (2005), Smets and Wouters (2007), Christoffel et al. (2008), Del Negro et al. (2007) and Wang (2009) have reported encouraging findings regarding the forecasting performance of state-of-the-art structural models. By contrast, the failure of researchers and professional forecasters to predict the "Great Recession" of 2008 and 2009 has generated much public criticism regarding the state of economic forecasting and macroeconomic modeling. Against this background, analysing the forecasting performance of structural models provides new insights. Specifically, I investigate the accuracy of point and density forecasts of four DSGE models for output growth, inflation and the interest rate. All of the models have been used in chapter 2 as well. Using structural models facilitates an economically meaningful interpretation of the forecasts. However, a thorough assessment of different structural models including a comparison to forecasts from sophisticated time series models and to professional forecasts for a long sample has not been undertaken yet. Recent comparison studies of state of the art forecasting methods have been restricted to nonstructural econometric methods (c.f. Stock and Watson, 2002; Bernanke and Boivin, 2003; Forni et al., 2003; Marcellino et al., 2003; Faust and Wright, 2009; Hsiao and Wan, 2010).

I use the same sample and real-time dataset as Faust and Wright (2009) who assess the forecasting accuracy of eleven nonstructural models. Therefore, the DSGE forecasts are directly comparable to these nonstructural forecasts. The dataset is perfectly synchronized with the Greenbook and thus the results can also be compared to a best practice benchmark given by the Greenbook projections of the Federal Reserve.

The considered models cover to some extent the range of closed-economy DSGE models used in academia and at policy institutions. The model parameters are reestimated for the historical data vintages using maximum likelihood or Bayesian estimation. Given this estimate, I compute a nowcast and forecasts up to five quarters into the future that take into account information that was actually available at the forecast start. The evaluation results confirm the reasonable forecast accuracy of DSGE models found in the above mentioned studies. The forecast quality of the structural models is in particular competitive to the Greenbook projections for medium term horizons. Point forecasts of some models are comparable to the forecast accuracy of atheoretical forecasting methods that can process large data sets. Especially the model by Smets and Wouters (2007) yields relatively precise forecasts. Structural forecasts perform quite well during normal times, but they are not able to detect large recessions and turning points due to their weak internal propagation meachanism.

The forecasting literature using nonstructural models has found that combining several forecasts from different models can increase the forecast accuracy (Timmermann, 2006). Chapter 3 confirms this finding for structural models. I consider several simple and sophisticated model averaging schemes to compute weighted forecasts. A simple mean of model forecasts is more accurate than forecasts from individual models and is hard to beat by other forecast weighting methods.

While point forecasts are interesting, economists are concerned about the uncertainty surrounding these. Therefore, I derive density forecasts for the DSGE models that take into account parameter uncertainty and uncertainty about economic shocks expected in the future. I find that all the model forecasts overestimate actual uncertainty. A reason might be the tight restrictions imposed on the data. If the data rejects these restrictions, large shocks are needed to fit the models to the data resulting in high shock uncertainty (see also Gerard and Nimark, 2008).

Chapter 2 and 3 are applications of existing macroeconomic models. Chapter 4, which is joined work with Tobias Cwik and Gernot J. Müller is an example how to construct a macroeconomic model of the type contained in the model database introduced in chapter 1 and estimated in chapters 2 and 3.

The surprisingly good forecasting results in the previous two chapters are obtained while abstracting from external trade altogether. Taken at face value, this suggests that trade integration, or openness, plays no important role for business cycle dynamics of large open economies. There is, however, a secular trend in trade integration, suggesting that economies are becoming considerably more open over time. In the U.S., imports, as a fraction of GDP, have risen from about 6 percent in 1973 to 16 percent to date. In fact, as this trend has been accelerating over the last decade, some observers have identified increasing trade integration as an important manifestation of globalization. In this chapter, we investigate more systematically the role of trade integration for monetary policy transmission. We assess how increasing openness alters quantitatively the effects of monetary policy shocks on domestic inflation and domestic absorption.

We develop a New Keynesian DSGE model featuring two symmetric countries and several frictions which recent business cycle research has found to be important in accounting for several macroeconometric observations. In addition, following Gust et al. (2006), Sbordone (2007) and Guerrieri et al. (2008), we assume a fairly general aggregation technology for final goods. It induces strategic complementarities in price setting for firms not only with respect to domestic, but also with respect to foreign competitors. Hence, the domestic currency price charged by foreign competitors enters the decision problem of domestic firms and eventually the New Keynesian Phillips curve. As a result, a new dimension of the exchange rate channel emerges. Traditionally, monetary policy is thought to directly impact CPI-inflation and to indirectly impact domestic inflation via the exchange rate, where the latter effect comes about through changes in demand induced by 'expenditure-switching'. With strategic price-setting complementarities, changes in the exchange rate, which alter the domestic currency prices charged by foreign competitors, directly impact domestic inflation. We analyse this new dimension of the exchange rate channel, by which monetary policy gains direct leverage over domestic inflation. We find that the importance of this effect increases with i) the extent of strategic complementarities in price-setting; ii) the openness of an economy and iii) the amount of exchange rate pass-through.

In order to quantify the effects of openness on monetary transmission, we estimate, in a first step, a VAR on U.S. time series relative to an aggregate of industrialized countries. We identify monetary policy shocks by imposing an identification scheme which is consistent with our theoretical model and trace out the transmission mechanism through impulse response functions. Having used maximum likelihood and Bayesian estimation in chapters 2 and 3, in this chapter a third method is used to pin down key parameters of the model: we find parameter values of the DSGE model by matching its impulse responses to those obtained from the VAR. This method is advantageous in the context of analysing monetary policy transmission as it only accounts for the effects of a monetary policy shocks and does not require the model to explain other aspects of business cycle dynamics. We find that the estimated model is generally able to mimic the empirical response functions quite closely. In a second step, we compare the effects of a monetary policy shock in the estimated model to counterfactual scenarios with different import shares. We find for all scenarios that limited exchange rate pass-through prevents the new dimension of the exchange rate channel from having strong quantitative effects. If we repeat our experiment while assuming higher exchange rate pass-through, the effects of monetary policy shocks become considerably stronger.

Finally, turning to the implications for monetary policy, we stress that while increasing openness could, in principle, improve the trade-off faced by monetary policy, such a development is likely to be prevented by low exchange rate pass-through. At current trends, it appears that while trade integration is on the rise, exchange rate pass-through is declining as far as major industrialized countries are concerned. We conclude that while policy makers should keep a close eye on the joint development of openness and exchange rate pass-through, future research may investigate possible causes underlying these trends.

While chapter 4 has focused on the impact of monetary policy shocks, i.e. the exogenous part of monetary policy, chapter 5 studies the endogenous part of monetary policy. Systematic monetary policy reactions are typically modeled in business cycle models of the previous chapters with a simple monetary policy rule of the type developed by Taylor (1993a). The forecast evaluation of chapter 3 has shown that the interest rate forecasts from a Bayesian VAR are much better than those from DSGE models. While the policy rule implicit in the VAR includes four lags of the interest rate, output growth and inflation, the policy rules in DSGE models are typically restricted to include only contemporaneous inflation, an output gap and one lag of the interest rate. Chapter 5 tests whether simple symmetric policy rules of this kind are a realistic description of actual monetary policy. In doing so, I estimate monetary policy rules while being agnostic about other economic dynamics, i.e.

the estimation does without specifying a complete macroeconomic model.
In reality the Federal Reserve does not follow a policy rule mechanically: "The monetary policy of the Federal Reserve has involved varying degrees of rule- and discretionary-based modes of operation over time," (Greenspan, 1997). This raises the question how the Federal Open Market Committee (FOMC) responds to inflation and the output gap during periods that cannot be described accurately by a policy rule. Except anecdotal descriptions of some episodes (e.g. Taylor, 1993a; Poole, 2006) there appears to be a lack of studies that analyze deviations from Taylor's rule systematically and quantitatively.
In addition to changes between discretionary and rule-based policy regimes, economic theory provides several reasons for deviating at least at times from a linear policy rule framework. First, asymmetric central bank preferences can lead in an otherwise linear model to a nonlinear policy reaction function (Gerlach, 2000; Surico, 2007; Cukierman and Muscatelli, 2008). A nonlinear policy rule can be optimal when the central bank has a quadratic loss function, but the economy is nonlinear (Schaling, 1999; Dolado et al., 2005). Even in a linear economy with symmetric central bank preferences an asymmetric policy rule can be optimal if there is uncertainty about specific model parameters (c.f. Meyer et al. (2001) and Tillmann (2010)). Finally, when interest rates approach the zero lower bound, responses to inflation might increase to avoid the possibility of deflation (Orphanides and Wieland, 2000; Kato and Nishiyama, 2005; Sugo and Teranishi, 2005; Adam and Billi, 2006). Despite these concerns in the empirical literature estimation of linear policy rules prevails with only few exceptions.
Policy rule parameters estimated with least squares methods characterize the conditional mean of the interest rate. Thus, during deviations of the interest rate from a linear policy rule the Federal Reserve sets the interest rate not at its conditional expected value, but at some other part of its conditional distribution. Chevapatrakul et al. (2009) estimate interest rate reactions at various points of its conditional distribution. I extend their work to real-time data, a recent IV quantile method and a gradual adjustment of interest rates. Using real-time data is crucial as the output gap was perceived by the Federal Reserve to be negative in real-time for almost the whole time between 1970 and 1990. I use real-time inflation forecasts from the Greenbook that are at times quite different from ex post realized inflation rates. Using Hausman tests I find significant endogeneity of inflation forecasts and output gap nowcasts and therefore use in addition to quantile regression (QR) inverse quantile regression (IQR) proposed by Chernozhukov and Hansen (2005) to compute consistent parameter estimates.
The results indicate that policy parameters fluctuate significantly over the conditional distribution of the federal funds rate. These deviations from the parameter estimates at the conditional mean of the interest rate are systematic: inflation reactions and the interest rate smoothing parameter increase and output gap responses decrease over the conditional distribution of the interest rate. This indicates that the FOMC has sought to stabilize inflation more and output less when setting the interest rate higher than implied by the estimated policy rule and vice versa. Thus, a fraction of deviations from an estimated linear policy rule are possibly not caused by policy shocks, but by systematic changes in the policy parameters or an asymmetric policy rule.
Having analyzed how the Federal Reserve sets interest rates when deviating from the conditional mean it is of interest whether these deviations are related to the business cycle. I find that the Fed reacted more to the output gap during recessions than during expansions. This leads to lower interest rates during recessions than implied by a simple symmetric policy rule. A recession avoidance preference of the FOMC found by Cukierman and Muscatelli (2008) is thus confirmed.

This dissertation studies macroeconomic models from different perspectives and yields several interesting new insights. First, differences in macroeconomic modeling are important. Different models can have significantly different implications for policy analysis. This should be taken into account in model based research on monetary and fiscal stabilization policies. An appropriate framework for comparative studies is introduced. Second, the heterogeneity of forecasts generated from different models is similar to heterogeneity found in surveys of professional forecasts. The findings can be taken as an indication that much of the observed time variation in forecast heterogeneity may be explained by disagreement about appropriate modeling assumptions rather than irrationality of particular forecasters. Including heterogenous beliefs in macroeconomic models is an important task as it may be a source of economic volatility. Third, stylized DSGE models yield surprisingly accurate point forecasts despite their reliance on very few observable data series. Future work is needed to demonstrate the structural interpretation of forecasts. Fourth, current generation DSGE models overestimate actual uncertainty. Developing models with stronger propagation mechanisms can potentially improve the accurace of density forecasts. Fifth, combining forecasts from several structural models can increase the forecast accuracy. Therefore, it is useful to consider several forecasting models in applied work. Sixth, strategic price complementarities can lead to a new dimension of the exchange rate channel of monetary policy transmission that can improve the trade-off faced by monetary policy. Its empirical relevance depends on the degree of trade integration, exchange rate pass-through and strategic price complementarities. Seventh, simple symmetric monetary policy rules are an insufficient description of actual monetary policy. Asymmetric reactions to inflation, the output gap and past interest rates have been detected. Future macroeconomic models should include more realistic monetary policy rules.

References

Adam, K., Billi, R. M., 2006. Optimal monetary policy under commitment with a zero bound on nominal interest rates. Journal of Money, Credit, and Banking 38(7), 1877–1905.

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

Adolfson, M., Laséen, S., Lindé, J., Villani, M., 2007. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Journal of International Economics 72(2), 481–511.

Bernanke, B. S., Boivin, J., 2003. Monetary policy in a data-rich environment. Journal of Monetary Economics 50(3), 525–546.

Blanchard, O., Fischer, S., 1989. Lectures on Macroeconomics. The MIT Press.

Branch, W. A., Evans, G. W., 2010. Monetary policy with heterogeneous expectations. Economic Theory, forthcoming.

Branch, W. A., McGough, B., 2010. Business cycle amplification with heterogeneous expectations. Economic Theory, forthcoming.

Brock, W., Hommes, C., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22, 1235–1274.

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

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

Bryant, R. C., Hooper, P., Mann, C., 1993. Design and implementation of the empirical simulations. In: Evaluating Policy Regimes: New Research in Empirical Macroeconomics. The Brookings Institution, Washington DC.

Capistran, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. Journal of Money, Credit, and Banking 41, 365–396.

Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. Econometrica 73(1), 245–261.

Chevapatrakul, T., Kim, T.-H., Mizen, P., 2009. The Taylor principle and monetary policy approaching a zero bound on nominal rates: Quantile regression results for the United States and Japan. Journal of Money, Credit and Banking 41(8), 1705–1723.

Chiarella, C., Dieci, R., He, X.-Z., 2007. Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework. Journal of Economic Behavior and Organization 62, 408–427.

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

Christoffel, K., Coenen, G., Warne, A., 2008. The New Area-Wide Model of the euro area - a micro-founded open-economy model for forecasting and policy analysis, European Central Bank Working Paper 944.

Clarida, R., Galí, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. Journal of Economic Literature 37(4), 1661–1707.

Coenen, G., Orphanides, A., Wieland, V., 2004. Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero. Berkeley Electronic Press: Advances in Macroeconomics 4(1), 1–23.

Coenen, G., Wieland, V., 2005. A small estimated euro area model with rational expectations and nominal rigidities. European Economic Review 49, 1081–1104.

Cukierman, A., Muscatelli, A., 2008. Nonlinear Taylor rules and asymmetric preferences in central banking: Evidence from the United Kingdom and the United States. The B.E. Journal of Macroeconomics 8(1).

de Grauwe, P., 2010. Animal spirits and monetary policy. Economic Theory, forthcoming.

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of New Keynesian models. Journal of Business and Economic Statistics 25(2), 123–143.

Dieppe, A., Kuester, K., McAdam, P., 2005. Optimal monetary policy rules for the euro area: An analysis using the Area Wide Model. Journal of Common Market Studies 43 (3), 507–5372.

Dolado, J. J., Maria-Dolores, R., Naveira, M., 2005. Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve. European Economic Review 49, 485–503.

Faust, J., Wright, J. H., 2009. Comparing Greenbook and reduced form forecasts using a large realtime dataset. Journal of Business and Economic Statistics 27(4), 468–479.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the euro area? Journal of Monetary Economics 50, 1243–1255.

Fuhrer, J. C., Moore, G., 1995. Inflation persistence. The Quarterly Journal of Economics 110(1), 127–159.

Gerard, H., Nimark, K., 2008. Combing multivariate density forecasts using predictive criteria, Research Discussion Paper 2008-2, Reserve Bank of Australia.

Gerlach, S., 2000. Asymmetric policy reactions and inflation, working paper, Bank for International Settlements.

Giordani, P., Söderlind, P., 2003. Inflation forecast uncertainty. European Economic Review 47, 1037–1059.

Greenspan, A., September 1997. Rules vs. discretionary monetary policy, speech at the 15th Anniversary Conference of the Center for Economic Policy Research at Stanford University, Stanford, California.

Guerrieri, L., Gust, C., López-Salido, D., 2008. International competition and inflation: A New Keynesian perspective, International Finance Discussion Papers, 918.

Gust, C., Leduc, S., Vigfusson, R. J., 2006. Trade integration, competition, and the decline in exchange-rate pass-through, International Finance Discussion Papers, Number 864, Board of Governors of the Federal Reserve System.

Hsiao, C., Wan, S. K., 2010. Is there an optimal forecast combination?, Working Paper University of Southern California.

Hughes-Hallett, A., Wallis, K. F. (Eds.), 2004. EMU macroeconomic model comparison excercise for the Euroconference 7-8 June 2002. Economic Modelling 21(5).

Kato, R., Nishiyama, S.-I., 2005. Optimal monetary policy when interest rates are bounded at zero. Journal of Economic Dynamics & Control 29, 97–133.

Klein, L. (Ed.), 1991. Comparative Performance of U.S. Econometric Models. Oxford, Eng.: Oxford University Press.

Kurz, M., 1994a. On rational belief equilibria. Economic Theory 4, 859–876.

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

Kurz, M., 1996. Rational beliefs and endogenous uncertainty: an introduction. Economic Theory 8, 383–397.

Kurz, M., 1997a. Endogenous economic fluctuations and rational beliefs: A general perspective. In: Kurz, M. (Ed.), Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M. (Ed.), 1997b. Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M., 2009. Rational diverse beliefs and market volatility. In: Hens, T., Schenk-Hoppe, K. (Eds.), Handbook of financial markets: dynamics and evolution. North Holland.

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

Kurz, M., Jin, H., Motolese, M., 2005. The role of expectations in economic fluctuations and the efficacy of monetary policy. Journal of Economic Dynamics & Control 29, 2017–2065.

Laxton, D., Pesenti, P., 2003. Monetary rule for small, open, emerging economies. Journal of Monetary Economics 50, 1109–1146.

Marcellino, M., Stock, J., Watson, M., 2003. Macroeconomic forecasting in the euro area: country-specific versus area-wide information. European Economic Review 47, 1–18.

McCallum, B., 1988. Robustness properties of a rule for monetary policy. Carnegie-Rochester Conference Series on Public Policy 29, 173–204.

McCallum, B., 1999. Issues in the design of monetary policy rules. In: Taylor, J. B., Woodford, M. (Eds.), Handbook of Macroeconomics. Amsterdam: Elsevier Science, North-Holland.

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

Meyer, L. H., Swanson, E. T., Wieland, V., 2001. Nairu uncertainty and nonlinear policy rules. American Economic Review 91(2), 226–231.

Orphanides, A., 2003. The quest for prosperity without inflation. Journal of Monetary Economics 50, 633–663.

Orphanides, A., Wieland, V., 2000. Efficient monetary policy design near price stability. Journal of the Japanese and International Economies 14, 327–365.

Poole, W., August 2006. Understanding the Fed, speech at the Dyer County Chamber of Commerce Annual Membership Luncheon, Dyersburg, Tenn.

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

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

Rudebusch, G. D., Svensson, L. E. O., 1999. Policy rules for inflation targeting. In: Taylor, J. B. (Ed.), Monetary Policy Rules. Chicago: University of Chicago Press.

Sbordone, A. M., 2007. Globalization and inflation dynamics: the impact of increased competition, Federal Reserve Bank of New York.

Schaling, E., 1999. The nonlinear phillips curve and inflation forecast targeting, Bank of England Working Paper No. 98.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the euro area. Journal of the European Economic Association. 1 (5), 1123–1175.

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

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

Stock, J., Watson, M., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.

Sugo, T., Teranishi, Y., 2005. The optimal monetary policy rule under the non-negativity constraint on nominal interest rates. Economics Letters 89, 95–100.

Surico, P., 2007. The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. Journal of Economic Dynamics & Control 31, 305–324.

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

Taylor, J. B., 1993b. Macroeconomic Policy in a World Economy. W.W. Norton, New York, online edition available on: http://www.stanford.edu/ johntayl/MacroPolicyWorld.htm.

Taylor, J. B., 1999. Monetary Policy Rules. The University of Chicago Press.

Tillmann, P., 2010. Parameter uncertainty and non-linear monetary policy rules. Macroeconomic Dynamics, forthcoming.

Timmermann, A., 2006. Forecast combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Amsterdam: North Holland, pp. 135–196.

Wang, M.-C., 2009. Comparing the DSGE model with the factor model: An out-of-sample forecasting experiment. Journal of Forecasting 28(2), 167–182.

Zusammenfassung

In den letzten Jahren sind viele makroökonomische Modelle zur Erkärung der Dynamik wichtiger ökonomischer Variablen entwickelt worden. Dynamische stochastische allgemeine Gleichgewichtsmodelle, die konsequent von mikroökonomischen Optimierungsproblemen privater Haushalte und Unternehmen hergeleitet werden, sind zum Hauptanalyseinstrument der modernen Konjunkturtheorie geworden. Diese Modelle werden nicht nur von Forschern an Universitäten, sondern auch von Ökonomen an Zentralbanken, Ministerien und internationalen Politikorganisationen wie dem internationalen Währungsfonds oder der Europäischen Kommission entwickelt und verwendet. Gäbe es einen breiten Konsenz über ein Referenzmodell, das gegenüber anderen Ansätzen hinsichtlich seiner theoretischen Fundierung und empirischen Validierung zu bevorzugen wäre, so könnte dieses für einheitliche Politikanalysen und -empfehlungen verwendet werden. Bisher ist jedoch kein Konsens hinsichtlich der makroökonomischen Modellierung absehbar.
Während viele Forscher mikroökonomisch fundierte Neu-Keynesianische Modelle bevorzugen, gibt es scharfe Kritiker dieser Modelle. Sie empfehlen, zu traditionelleren Modellierungsformen zurückzukehren. Auch die empirische Validierung von konkurrierenden Modellierungsansätzen führt nicht weiter, da verschiedenste Modelle historische Daten der wichtigsten makroökonomischen Variablen ähnlich gut erklären. Basierend auf diesen verschiedenen makroökonomischen Theorien führt die Struktur konkurrierender Konjunkturmodelle zu unterschiedlichen - und möglicherweise gegensätzlichen - Handlungsempfehlungen für politische Entscheidungsträger.
Auf Grund der Uneinigkeit über geeignete Modelle für die Analyse von Geldpolitik, Fiskalpolitik und Finanzstabilitätspolitik können vergleichende Studien entscheidend zu einer besseren makroökonomischen Modellierung und fundierteren Politikanalysen beitragen. Ein Vergleich der Erkenntnisse aus verschiedenen Modellsimulationen kann zu verlässlichen Handlungsempfehlungen für die Politik führen. Solche Empfehlungen sind somit robust in Bezug auf Modellunsicherheit. Vergleichende Analysen können vermeiden, dass eine Politik verfolgt wird, die in einzelnen Modellen Erfolg verspricht, in der Realität jedoch nicht funktioniert. Ein solcher Ansatz wird von McCallum (1988), McCallum (1999), Blanchard und Fischer (1989), Taylor (1999) und vielen anderen empfohlen. Ein Vergleich von empirischen Implikationen verschiedener Modelle ist jedoch schwierig. Die Bewertung von unterschiedlichen Politikszenarien über verschiedene Modelle hinweg ist arbeitsintensiv und kostspielig und daher selten. Es gibt nur sechs größere Vergleichsprojekte, an denen jeweils mehrere

Forscherteams beteiligt waren: Bryant et al. (1988), Bryant et al. (1989), Klein (1991), Bryant et al. (1993), Taylor (1999) und Hughes-Hallett und Wallis (2004). Jedes Team arbeitete dabei nur mit einem oder wenigen Modellen. Während diese Initiativen zu einigen einflussreichen Erkenntnissen wie der Taylor-Regel[3] geführt haben, ist das Spektrum systematischer, vergleichender Forschungserkenntnisse sehr begrenzt geblieben.

Diese Dissertation stellt einen neuen Ansatz für die vergleichende modellbasierte Forschung vor. Dieser Ansatz ermöglicht es Ökononomen erstmals, Modellvergleiche einfach und in großem Stil durchzuführen. Eine Datenbank mit makroökonomischen Modellen wird erstellt, die viele bekannte empirische Modelle für die quantitative Analyse monetärer und fiskalischer Stabilisierungspolitik enthält. Aufbauend auf diesem vergleichenden Ansatz der modellbasierten ökonomischen Forschung umfasst diese Dissertationen zwei Anwendungen makroökonomischer Modelle zur Berechnung und Evaluierung makroökonomischer Prognosen. Daran anschließend greifen zwei Kapitel einzelne Aspekte der makroökonomischen Modellierung auf: In einem Kapitel wird die Auswirkung gestiegener Handelsintegration auf die Transmission der Geldpolitik analysiert und in einem zweiten Kapitel teste ich empirisch, ob einfache lineare geldpolitische Regeln - wie sie in den meisten makroökonomischen Modellen verwendet werden - eine hinreichende Beschreibung der Realität darstellen.
Kapitel 1 stellt eine Datenbank makroökonomischer Modelle und einen Ansatz zum Modellvergleich vor. Dieser Ansatz ist in ein Programm implementiert, dass die Simulation von Modellen und den Vergleich von Modellen anhand von Statistiken wie Impulsantwortfolgen und Autokorrelationsfunktionen ermöglicht. Während die Modelle der Datenbank mit den Modellparametern implementiert sind, die von den jeweiligen Autoren verwendet wurden, werden in Kapitel 2 mehrere Modelle anhand eines gemeinsamen Datensatzes verglichen. Prognosen werden für die fünf jüngsten von der NBER definierten U.S. Rezessionen berechnet. Die Heterogenität von Vorhersagen wird analysiert und ein Vergleich zu Prognosen des Survey of Professional Forecasters (SPF) durchgeführt. Kapitel 3 dehnt die Prognoseevaluation auf eine längere Datenstichprobe und die Evaluation der Prognoseverteilungen aus. Es wird gezeigt, dass die Kombination von Prognosen mehrerer Modelle die Präzision der Vorhersagen erhöhen kann. Kapitel 4 stellt ein Beispiel dafür da, wie ein makroökomisches Modell konzipiert werden kann: Ein Neu-Keynesianisches Modell zweier großer offener Volkswirtschaften mit strategischen Komplementaritäten in der Preissetzung von Firmen wird entwickelt und die Auswirkung einer erhöhten Handelsintegration auf die Transmission von geldpolitischen Impulsen untersucht. Kapitel 5 analysiert einfache geldpolitische Regeln, die in makroökonomischen Modellen zur Beschreibung der Geldpolitik eingesetzt werden. Geldpolitische Regeln werden losgelöst von einschränkenden Modellannahmen über andere wirtschaftliche Zusammenhänge empirisch geschätzt. Politikparameter werden über die gesamte bedingte Verteilung des Leitzinses mittels Quantilregressionen geschätzt. Es zeigt sich, dass einfache symmetrische Regeln zu restriktiv sind, um die tatsächliche Geldpolitik widerzuspiegeln.

Kapitel 1 der Dissertation ist eine gemeinsame Arbeit mit Volker Wieland, Tobias Cwik, Gernot J. Müller und Sebastian Schmidt. Eine Datenbank gesamtwirtschaftlicher Konjunkturmodelle wird vorgestellt. Diese Datenbank ermöglicht Forschern, systematische Vergleiche von Modellen und deren Politikimplikationen durchzuführen. Eine allgemeine Klasse von nichtlinearen dynamischen stochastischen makroökonomischen Modellen wird durch gemeinsame vergleichbare Variablen,

[3]Taylor (1993a) nennt das Projekt von Bryant et al. (1993) als entscheidenden Test für seine einfache geldpolitische Regel, die als Taylor-Regel bekannt wurde.

Parameter und Schocks ergänzt. Auf dieser Grundlage können gemeinsame geldpolitische und fiskalische Regeln definiert und deren Auswirkungen in verschiedenen Modellen verglichen werden. Solche Vergleiche basieren auf Statistiken wie Impulsantwortfolgen, Autokorrelationsfunktionen und Varianzen der wichtigsten makroökonomischen Aggregate.

Die Datenbank umfasst Modelle der US-Wirtschaft, der Eurozone und mehrere Mehrländer-Modelle. Einige der Modelle sind recht klein und konzentrieren sich auf die Erklärung von Wirtschaftswachstum, Inflation und Zinsen (vgl. Clarida et al. (1999), Rotemberg und Woodford (1997), Fuhrer und Moore (1995), McCallum und Nelson (1999), Coenen und Wieland (2005), etc). Viele Modelle sind mittlerer Größe und decken viele wichtige makroökonomische Aggregate ab (vgl. Christiano et al. (2005), Coenen et al. (2004), Smets und Wouters (2003, 2007)). Einige Modelle in der Datenbank sind relativ groß wie beispielsweise das FRB-US-Modell der amerkanischen Notenbank von Reifschneider et al. (1999), das Modell der G7-Volkswirtschaften von Taylor (1993b) oder das Modell der Eurozone der Europäischen Zentralbank (Dieppe et al., 2005). Die meisten Modelle können als Neu-Keynesianische Modelle klassifiziert werden, da sie rationale Erwartungen, unvollkommenen Wettbewerb und Lohn- oder Preisrigiditäten enthalten. Bekannte Beispiele dieser von mikroökonomischen Optimierungsproblemen hergeleiteteten Neu-Keynesianischen Modelle sind Christiano et al. (2005), Smets und Wouters (2003, 2007), Laxton und Pesenti (2003) und Adolfson et al. (2007). Die Datenbank enthält allerdings auch Modelle, die nicht mikroökonomisch fundiert sind und in denen Erwartungen von Agenten nur eine geringe (vgl. das EZB Modell der Eurozone von (Dieppe et al., 2005)) oder gar keine Rolle spielen (z.B. Rudebusch und Svensson (1999) und Orphanides (2003)).

Zu der Modelldatenbank gehört ein Programm, mit dem Modelle gelöst, simuliert und vergleichende Analysen durchgeführt werden können. Vergleiche basieren dabei auf drei Statistiken: Erstens auf Impulsantwortfolgen in Reaktion auf gemeinsame Schocks von Variablen, die in allen Modellen gleich definiert sind, zweitens auf Autokorrelationsfunktionen gemeinsamer Variablen in Reaktion auf modellspezifische Schocks oder drittens auf der Varianz gemeinsamer Variablen. Das Programm kann auch verwendet werden, um systematisch die Auswirkung von Politikregeln in unterschiedlichen Modellen zu untersuchen. Dabei werden geldpolitische oder fiskalische Regeln basierend auf gemeinsamen Variablen der Modelle definiert und das Programm implementiert diese in die verschiedenen Modellen. Unterschiedliche Implikationen dieser Regeln in den verschiedenen Modellen können anhand von Statistiken über makroökonomische Variablen, die in allen Modellen existieren, analysiert werden. Das Programm kann für lineare und nichtlineare Modelle genutzt werden, wobei Perturbationsmethoden verwendet werden, um Lösungen nichtlinearer Modelle linear zu approximieren.

Das Kapitel enthält einige Beispiele zur vergleichenden Modell- und Politikanalyse. Impulsantwortfolgen auf monetäre und fiskalische Schocks von mehreren Modellen werden für verschiedene geldpolitische Regeln verglichen und so Unterschiede in der implizierten Inflations- und Output-Persistenz untersucht. Wir finden deutliche Unterschiede des geldpolitischen Transmissionsmechanismus zwischen kleinen und großen Modellen. Verschiedene Modellierungsphilosophien wie einerseits Neu-Keynesianische Modelle mit mikroökonomischen Grundlagen und andererseits größere disaggregierte Modelle mit einer weniger strikten theoretischen Fundierung scheinen Gründe für die Unterschiede in der Größe und Länge der Impulsantwortfolgen zentraler Variablen wie Produktion und Inflation auf fiskalische und geldpolitische Schocks zu sein.

Neue Modelle können problemlos in die Modelldatenbank aufgenommen und mit etablierten Modellen verglichen werden. Der vorgestellte vergleichende Ansatz kann dazu beitragen, die Reproduzierbarkeit von quantitativen makroökonomischen Analysen zu verbessern. Er kann so die Gefahr zirkulärer Entwicklungen in der modellbasierten Forschung maßgeblich verringern und zu robusteren und fundierteren Handlungsempfehlungen für die Politik führen.

Kapitel 2 - eine gemeinsame Arbeit mit Volker Wieland - erweitert den komparativen Ansatz makroökonomischer Modellierung in eine wichtige Richtung. Die Modelle der Datenbank in Kapitel 1 sind mit den Parametern implementiert, die von den jeweiligen Originalautoren geschätzt oder kalibriert wurden. Daher enthält die Modelldatenbank keinerlei Maß oder Statistik, die zeigt, welches Modell zur Beschreibung und Analyse der Daten einer bestimmten Stichprobe am besten geeignet ist. In Kapitel 2 werden daher drei einfache und zwei detailiertere Neu-Keynesianische Modelle mit einem gemeinsamen Datensatz verknüpft. Modellparameter werden geschätzt und darauf aufbauend Prognosen der wirtschaftlichen Aktivität und der Inflation berechnet. Die Prognosegenauigkeit kann als ein Maß gesehen werden, das zeigt, wie gut ein Modell die Daten beschreibt. Neben der Evaluation von makroökonomischen Prognosen wird die Heterogenität von Modellvorhersagen quantifiziert und mit Prognosen von profesionellen Makroökonomen verglichen, um mehr über das Ausmaß, die Dynamik und die Gründe für die Heterogenität von Prognosen zu erfahren.
Empirische Studien haben erhebliche Unterschiede in der Genauigkeit und der Heterogenität von Prognosen professioneller Makroökonomen für das BIP und die Inflation dokumentiert (siehe Kurz et al. (2003, 2005), Giordani und Söderlind (2003), Kurz (2009) und Capistran und Timmermann (2009)). In theoretischen makroökonomischen Modellen kann die Heterogenität von Erwartungen hinsichtlich zukünftiger Realisierungen makroökonomischer Variablen selbt zu einer Verstärkung konjunktureller Schwankungen führen und eine treibende Kraft für die Dynamik an Finanzmärkten sein (siehe Kurz (1994a,b, 1996, 1997a,b, 2009), Brock und Hommes (1998), Kurz et al. (2005), Chiarella et al. (2007), Branch und McGough (2010), Branch und Evans (2010) und de Grauwe (2010)).
Es gibt mehrere Gründe für die Heterogenität von Prognosen. Zunächst brauchen Ökonomen ein Prognoseinstrument. Dies kann ein theoriebasiertes umfangreiches makroökonomisches Modell, ein nicht strukturelles ökonometrisches Verfahren zur Erfassung statistischer Beziehungen oder eine einfache Daumenregel sein. Die unterschiedlichen Modellierungsannahmen dieser Prognoseinstrumente führen zu unterschiedlichen Prognosen. Eine weitere Quellen der Heterogenität stellen die Unterschiede in den Daten dar, die Ökonomen verwenden. Die Stichproben können sich in Bezug auf die Anzahl der volkswirtschaftlichen Aggregate und Preise und hinssichtlich der Frequenz und der Länge unterscheiden, für die der Prognostiker Daten sammelt. Die Daten werden benötigt, um den aktuellen Zustand der Volkswirtschaft zu erfassen und Modellparameter ökonometrisch zu schätzen.
Expertenprognosen werden in verschiedenen Umfragen veröffentlicht. Die zugrundeliegenden Modellannahmen, Daten und geschätzten Parameter sind aber unbekannt. Deshalb verwenden wir in diesem Kapitel verschiedene makroökonomische Modelle der US-Wirtschaft, um Wirtschaftswachstums- und Inflationsprognosen zu generieren. Die Präzision und die Heterogenität der Expertenprognosen aus dem SPF und dem Greenbook der amerikanischen Zentralbank werden als Benchmarks zum Vergleich herangezogen. Dieser Vergleich wird für Prognosen für die letzten fünf von der NBER definierten Rezessionen der US-Wirtschaft durchgeführt und es werden Quartalsprognosen für bis zu vier Quartale in die Zukunft berechnet. Rezessionen sind die größte Herausforderung für makroökonomische Prognostiker und die Heterogenität dieser Prognosen selbst kann die Tiefe und Länge einer Rezesssion beeinflussen.
Zusätzlich zu den einzelnen Modellprognosen berechnen wir den Mittelwert dieser Prognosen. Dessen Präzision liegt erstaunlich nah am Mittelwert der SPF-Prognosen und der Greenbook-Prognosen. Das ist bemerkenswert, da die Modelle den Informationsgehalt einer nur geringen Anzahl von Datenreihen nutzen. Die Genauigkeit von Modellprognosen im Vergleich zu Expertenprognosen ist insbesondere bei einem mittleren Prognosehorizont von drei bis vier Quartalen in der Zukunft gut.

Das Ausmaß der Prognoseheterogenität von Modellen und Expertenprognosen ist ähnlich und variiert erheblich im Laufe der Zeit. Natürlich sind die Modelle, die in diesem Kapitel verwendet werden, nicht die Modelle, die von professionellen Prognostikern genutzt wurden. Während die Gründe für die Heterogenität von Expertenprognosen nicht beobachtbar sind, kann die Heterogenität der Modellvorhersagen auf unterschiedliche Modellannahmen, Daten und Modellparameter zurückgeführt werden. Diese drei Unterschiede reichen aus, um Unterschiede in den Modellprognosen zu generieren, die den Unterschieden in den Expertenprognosen ähneln. Die rekursive Aktualisierung der Modellschätzungen mit historischen Daten führt zu einer ähnlichen Dynamik der Prognoseheterogenität der Modelle wie der Prognosen aus dem SPF. Die Ergebnisse dieses Kapitels deuten darauf hin, dass ein Großteil der beobachteten Unterschiede und deren Schwankungen über die Zeit durch Unterschiede in den Prognoseinstrumenten und Parameterschätzungen und nicht durch irrationales Verhalten erklärt werden können. Diese Heterogenenität selbst kann wiederum ein Grund für ökonomische Dynamik sein. Die makroökonomischen Modelle, die in der vorliegenden Dissertation verwendet werden, würden diese Volatilität allerdings ökonomischen Schocks und andereren Modellmechanismen zuschreiben. Daher ist es wichtig, in der Zukunft heterogene Erwartungen in makroökonomische Modelle einzubeziehen, um diese Quelle konjunktureller Schwankungen von anderen Einflüssen unterscheiden zu können.

Kapitel 3 erweitert die Analyse von Kapitel 2. Die Daten zur Evaluation von Vorhersagen sind nicht mehr auf Rezessionen beschränkt, sondern reichen durchgehend von 1984 bis 2000. Während der Schwerpunkt von Kapitel 2 auf der Auswertung der Prognosegüte um konjunkturelle Wendepunkte herum und der Analyse von Prognoseheterogenität liegt, besteht Kapitel 3 aus einer detaillierten Prognoseevaluation.
In den letzten Jahren haben viele Ökonomen herausgefunden, dass Neu-Keynesianische Modelle zu relativ genauen Prognosen führen (vgl. Smets und Wouters (2004), Adolfson et al. (2005), Smets und Wouters (2007), Christoffel et al. (2008), Del Negro et al. (2007) und Wang (2009). Im Gegensatz dazu ist es Ökonomen nicht gelungen, die große Rezession von 2008 und 2009 zu prognostizieren, was zu starker öffentlicher Kritik am Stand makroökonomischer Vorhersagemethoden und Modelle geführt hat. Vor diesem Hintergrund kann die Analyse der Prognosequalität struktureller makroökonomischer Modelle wichtige neue Erkenntnisse liefern. In diesem Kapitel untersuche ich die Genauigkeit von Punkt- und Dichteprognosen von vier dynamischen stochastischen allgemeinen Gleichgewichtsmodellen für das Wirtschaftswachstum, die Inflation und den Leitzins. Alle Modelle wurden ebenfalls in Kapitel 2 verwendet. Die Nutzung theoriebasierter Modelle ermöglicht eine sinnvolle Interpretation der Prognosen. Meines Wissens ist dies die erste ausführliche Analyse der Prognosegüte mehrerer struktureller Modelle inklusive eines Vergleichs mit der Prognosegüte nichtstruktureller Zeitreihenmodelle und Expertenprognosen. Studien zum Vergleich aktueller Prognoseverfahren haben sich bisher auf nichtstrukturelle ökonometrische Methoden beschränkt (Stock und Watson, 2002; Bernanke und Boivin, 2003; Forni et al., 2003; Marcellino et al., 2003; Faust und Wright, 2009; Hsiao und Wan, 2010).
Ich nutze die gleichen Daten wie Faust und Wright (2009), welche die Prognosegüte elf nichtstruktureller Modelle untersuchen. Dies ermöglicht einen direkten Vergleich der Vorhersagen struktureller und nichtstruktureller Modelle. Auf den gleichen Daten basieren die Greenbookprognosen der amerikanischen Notenbank, so dass ein Vergleich der Ergebnisse auch hier möglich ist. Die betrachteten Modelle repräsentieren zu einem gewissen Grad die Bandbreite der in der Wissensschaft und an Zentralbanken verwendeten Neu-Keynesianischen Modelle. Die Modellparameter werden rekursiv für die jeweiligen historischen Daten mit Maximum Likelihood oder Bayesianischer Schätzung aktualisiert. Darauf aufbauend werden aktuelles BIP, Inflation und Zins geschätzt und

Prognosen für bis zu fünf Quartale in die Zukunft berechnet. Es werden zu jedem Zeitpunkt nur Informationen verwendet, wie sie für Ökonomen in der Vergangenheit tatsächlich vorlagen. Die Evaluationsergebnisse bestätigen die relativ genauen Vorhersagen, die von den oben genannten Forschern für einzelne Neu-Keynesianische Modelle gefunden wurden. Die Prognosequalität der Modelle steigt relativ zur Prognosequalität der Greenbookprognosen mit dem Prognosehorizont. Theoriebasierte Modelle sind also insbesondere für Prognosen der mittleren Frist geeignet. Die Genauigkeit der Punktprognosen einiger Modelle ist vergleichbar mit der Genauigkeit atheoretischer Prognoseverfahren, die Informationen großer Datenmengen berücksichtigen können. Insbesondere das Modell von Smets und Wouters (2007) liefert relativ präzise Vorhersagen. Strukturelle Modelle führen zu relativ akkuraten Prognosen in normalen Zeiten. Sie sind allerdings nicht in der Lage, große Rezessionen und konjunkturelle Wendepunkte vorherzusagen. Ein Grund hierfür liegt darin, dass die Modelle sehr stilisiert sind und nur wenig endogene Volatilität erzeugen. Ein großer Teil der Dynamik wird von exogenen Schockprozessen und nicht von der theoriebasierten endogenen Struktur der Modelle erfasst.

Die Literatur über Prognosen nichtstruktureller Modelle kommt zu dem Schluss, dass die Kombination mehrerer Prognosen verschiedener Modelle die Prognosegenauigkeit erhöhen kann (siehe z.B. den Überblicksartikel von Timmermann, 2006). Die Ergebnisse aus Kapitel 3 zeigen, dass das auch für strukturelle Modelle gilt. Verschiedene einfache und komplizierte Verfahren zur Gewichtung mehrere Prognosen werden getestet und es zeigt sich, dass ein einfacher Durchschnitt der einzelnen Modellvorhersagen genauer ist als die Vorhersagen der einzelnen Modelle und kaum von fundierteren Gewichtungsmethoden zu schlagen ist.

Während Punktprognosen einige Anhaltspunkte für zu erwartende Entwicklungen geben, sind Ökonomen insbesondere an der Unsicherheit der Prognosen interessiert. Daher berechne ich Dichteprognosen, die Parameterunsicherheit und Unsicherheit über zukünftige exogene Schocks berücksichtigen. Die Evaluationsmethoden von Diebold et al. (1998) und Diebold et al. (1999) zeigen, dass die hier verwendeten Modelle die tatsächliche Unsicherheit überschätzen. Ein Grund dafür liegt in den starken Restriktionen, die die Modelle für die Daten implizieren. Werden diese empirisch abgelehnt, so sind starke Schwankungen exogener Schocks notwendig, damit die Modelle die Daten überhaupt widerspiegeln können (siehe auch Gerard und Nimark, 2008).

In den Kapiteln 2 und 3 wurden bereits bestehende makroökonomische Modelle verwendet. Kapitel 4, das in gemeinsamer Arbeit mit Tobias Cwik und Gernot J. Müller enstanden ist, ist ein Beispiel dafür, wie ein makroökonomisches Modell ähnlich der Modelle der in Kapitel 1 eingeführten Modelldatenbank konstruiert werden kann.

Die guten Prognoseergebnisse der beiden vorangegangenen Kapitel wurden ohne Berücksichtigung des Außenhandels unter Verwendung von Modellen der geschlossenen Volkswirtschaft erzielt. Das könnte bedeuten, dass der Außenhandel keine wichtige Rolle für die Konjunkturanalyse großer offener Volkswirtschaften wie den Vereinigten Staaten spielt. Es gibt jedoch in den letzten Jahren einen Anstieg des Welthandels, was darauf hindeutet, dass der Außenhandel eine größere Rolle als früher spielt. In den USA ist der Anteil der Importe am BIP von rund 6 Prozent 1973 auf bis zuletzt 16 Prozent gestiegen. Die Beschleunigung dieses Trends in den letzten zehn Jahren hat dazu geführt, dass Ökonomen die verstärkte Handelsintegration als einen der wichtigsten Bestandteile der Globalisierung ansehen. In diesem Kapitel wird die Auswirkung eines Anstiegs des Außenhandels auf den geldpolitischen Transmissionsmechanismus untersucht.

Wir entwickeln ein Neu-Keynesianisches Modell mit zwei symmetrischen Ländern und vielen nominalen und realen Friktionen, die wichtig für die Erklärung der Daten durch das Modell sind. Darüber hinaus übernehmen wir einen sehr allgemeinen Aggregator für Endprodukte wie er in Gust

et al. (2006), Sbordone (2007) und Guerrieri et al. (2008) verwendet wurde. Dieser Aggregator führt zu strategischen Komplementaritäten in der Preisbildung der Unternehmen. Dabei berücksichtigen Unternehmen nicht nur Preise inländischer, sondern auch ausländischer Wettbewerber. Der Preis von Gütern ausländischer Unternehmen in inländischer Währung beeinflusst somit die Entscheidungen inländischer Unternehmen und letzendlich die Neu-Keynesianische Phillips-Kurve. Dadurch ensteht eine neue Dimension des Wechselkurskanals der geldpolitischen Transmission. Der traditionelle Wechselkurskanal besteht aus einer unmittelbaren Auswirkung von Leitzinsänderungen und damit verbundenen Wechselkursänderungen auf den Verbraucherpreisindex, der Importe enthält. Außerdem wirkt sich die Geldpolitik indirekt über den Wechselkurs und die dadurch implizierte Veränderung der Nachfrage nach heimischen relativ zu importierten Gütern auf die inländische Inflation aus. Durch strategische Komplementaritäten in der Preissetzung führen durch Wechselkursänderungen induzierte Änderungen der Preise importierter Güter zu einer Änderung der Preise heimischer Güter und damit der inländischen Inflation in der gleichen Richtung. Wir analysieren diese neue Dimension des Wechselkurskanals, durch den die Geldpolitik verstärkten direkten Einfluss auf die inländische Inflation hat. Es zeigt sich, dass die Bedeutung dieses Kanals durch drei Quellen beeinflusst wird. Sie steigt mit i) dem Ausmaß der strategischen Komplementaritäten in der Preissetzung, ii) der Offenheit einer Volkswirtschaft und iii) der Höhe der Durchlässigkeit von Wechselkursänderungen auf Importpreise (Pass-Through).
Um die Auswirkung des Außenhandels auf die Transmission der Geldpolitik zu quantifizieren, schätzen wir als erstes ein Vektorautoregressionsmodell (VAR) von US-Zeitreihen relativ zu aggregierten Zeitreihen der anderen wichtigsten Industrieländer. Geldpolitische Schocks werden konsistent mit den Annahmen in dem theoretischen Modell identifiziert. Impulsantwortfolgen quantifizieren die geldpolitische Transmission. Während in den Kapiteln 2 und 3 Maximum Likelihood und Bayesianische Schätzungsmethoden verwendet werden, benutzen wir in diesem Kapitel ein drittes Verfahren, um die Modellparameter zu schätzen: Parameterwerte werden hier so gewählt, dass Impulsantwortfolgen des Modells und des VARs möglichst wenig voneinander abweichen. Dieses Verfahren ist vorteilhaft im Zusammenhang mit der Analyse der Transmission geldpolitischer Schocks, da es nur die Auswirkungen geldpolitischer Schocks berücksichtigt und es Erkärungen anderer Aspekte der makroökonomischen Dynamik nicht erforderlich macht. In einem zweiten Schritt vergleichen wir die Effekte eines geldpolitischen Schocks in dem geschätzten Modell mit Simulationen für unterschiedliche Offenheitsgrade der Volkswirtschaft. Wir finden heraus, dass in allen Szenarien eine starke Auswirkung des neuen Wechselkurskanals durch die sehr begrenzte Durchlässigkeit des Wechselkurses auf Importpreise verhindert wird. Wir wiederholen die Simulationen mit einer höheren Durchlässigkeit des Wechselkurses auf Importpreise, was zu einer stärkeren Wirkung der neuen Dimension des Wechselkurskanals führt.
In Hinsicht auf Implikationen für die Geldpolitik ist festzuhalten, dass eine Erhöhung der Offenheit einer Volkswirtschaft potenziell den Zielkonflikt der Geldpolitik entschärfen kann, aber eine solche Entwicklung durch die geringe Durchlässigkeit des Wechselkurses verhindert wird. Während die Handelsintegration steigt, zeigen empirische Untersuchungen, dass die Durchlässigkeit von Wechselkursen rückläufig ist. Politische Entscheidungsträger sollten also die weitere Entwicklung verfolgen und zukünftige Forschung könnte mögliche Ursachen für diese Entwicklungen untersuchen.

Während in Kapitel 4 die Auswirkung geldpolitischer Schocks - d.h. die Auswirkung des exogenen Teils der Geldpolitik - analysiert wird, behandelt Kapitel 5 den endogenen Teil der Geldpolitik. In den Konjunkturmodellen der vorangegangenen Kapitel werden einfache systematische geldpolitische Reaktionsfunktionen ähnlich der von Taylor (1993a) entwickelten Regel verwendet. Die Auswertung der Prognosen in Kapitel 3 hat gezeigt, dass Zinsprognosen eines Bayesiansischen VARs viel akkurater als Zinspronosen Neu-Keynesianischer Modelle sind. Während die Gleichung für den Zins im VAR den Zins, das Wirtschaftswachstum und die Inflationsrate der vier vorherigen Quartale umfasst, beschränken sich die geldpolitischen Regeln in strukturierten Modellen in der Regel auf die aktuelle Quartalsinflation, die aktuelle Produktionslücke und den Zins im vorherigen Quartal. Kapitel 5 testet, ob diese einfachen symmetrischen Politikregeln eine realistische Beschreibung der tatsächlichen Geldpolitik sind. Die Schätzung geldpolitischer Regeln erfolgt dabei losgelöst von der restriktiven Modellierung weiterer ökonomischer Zusammenhänge.

In der Realität folgt die amerkanische Notenbank nicht mechanisch einer Regel. „Die Geldpolitik der US-Notenbank basiert auf variirenden Anteilen regelbasierter und diskretionärer Entscheidungen", so Greenspan (1997). Das wirft die Frage auf, wie der Offenmarktausschuss der US-Notenbank in Phasen, die nicht gut durch eine Politikregel beschrieben werden können, auf Inflation und die Produktionslücke reagiert. Außer Beschreibungen einzelner Episoden (z.B. Taylor, 1993a; Poole, 2006) gibt es keine Studien, die systematisch und quantitativ Abweichungen der Geldpolitik von der Taylor-Regel analysieren.

Auch die makroökonomische Theorie liefert mehrere Gründe, warum die Geldpolitik zumindest zeitweise von linearen Reaktionsfunktionen abweicht. Asymmetrische Präferenzen der Zentralbanker können in einem ansonsten linearen Modell zu einer nichtlinearen geldpolitischen Regel führen (Gerlach, 2000; Surico, 2007; Cukierman und Muscatelli, 2008). Eine nichtlineare Politikregel kann optimal sein, wenn die Zentralbank in einem nichtlinearen Modell eine quadratische Verlustfunktion hat (Schaling, 1999; Dolado et al., 2005). Selbst in einem linearen Modell mit symmetrischen Präferenzen auf Seiten der Zentralbank kann eine asymmetrische Politikregel optimal sein, wenn Unsicherheit über bestimmte Modellparameter besteht (Meyer et al., 2001; Tillmann, 2010). Wenn der Nominalzins nicht unter null fallen kann, reagiert die Zentralbank möglicherweise stärker auf die Inflation, um die Gefahr einer Deflation zu vermindern (Orphanides und Wieland, 2000; Kato und Nishiyama, 2005; Sugo und Teranishi, 2005; Adam und Billi, 2006). Trotz dieser Argumente werden in der empirischen Literatur und in den meisten Modellen überwiegend lineare Politikregeln verwendet.

Mit der Methode der kleinsten Quadrate geschätzte geldpolitische Regeln charakterisieren den bedingten Mittelwert des Zinssatzes. Weicht die Zentralbank vom Zinssatz ab, der durch eine geschätze Regel impliziert wird, so setzt sie den Zins nicht am bedingten Erwartungswert, sondern in einem anderen Bereich der bedingten Verteilung des Zinses. Chevapatrakul et al. (2009) schätzen Zinsreaktionen an verschiedenen Punkten der bedingten Verteilung des Zinses. Ich erweitere ihre Arbeit in mehrfacher Hinsicht. Ich benutze Echtzeitdaten, ein aktuelles Instrumentenvariablen-Quantilregressions-Verfahren und modelliere die empirisch beobachtete Zinsglättung. Echtzeitdaten zu verwenden ist entscheidend, da zur jeweiligen Zinsentscheidung fast über die ganze Zeit von 1970 bis 1990 von der US-Notenbank eine negative Produktionslücke berechnet wurde. Ich verwende Echtzeitinflationsprognosen aus dem Greenbook, da diese sich zeitweise stark von ex post realisierten Inflationsraten unterscheiden. Der Hausman-Test zeigt, dass Inflationsprognosen und Produktionslücken endogen sind. Daher verwende ich zusätzlich zur Quantilsregression (QR) die inverse Quantilsregression (IQR) von Chernozhukov und Hansen (2005), um konsistent geschätzte Parameter zu berechnen.

Die Ergebnisse zeigen, dass die Parameter der geldpolitischen Regel erheblich über die bedingte Verteilung des Leitzinses variieren. Diese Abweichungen von der Parameterschätzung am bedingten Mittelwert des Zinssatzes sind systematisch: Inflationsreaktionen und Zinsglättungsparameter steigen über die bedingte Verteilung des Zinses an, während der Reaktionsparameter der Produktionslücke sinkt. Das deutet darauf hin, dass die US-Notenbank - wenn der Zins höher war, als von einer Zinsregel impliziert - versucht hat, Inflation stärker als das Wirtschaftswachstum zu stabilisieren und

umgekehrt, wenn der Zins niedriger war. Somit wird ein Teil der Abweichungen von einer geschätzten linearen geldpolitischen Reaktionsfunktion nicht von exogenenen Politikschocks, sondern von systematischen asymmetrischen Politikreaktionen verursacht.
Neben der Analyse, wie die US-Notenbank den Leitzins während Abweichungen von einer Zinsreaktionsfunktion setzt, ist es interessant herauszufinden, ob ein zeitlicher Zusammenhang dieser Abweichungen zum Verlauf von Konjunkturzyklen besteht. Es zeigt sich, dass die Zentralbank während einer Rezession mehr auf die Produktionslücke reagiert, als während eines Aufschwungs. Dies führt zu niedrigeren Zinsen während einer Rezession, als wenn die Zentralbank einer einfachen symmetrischen Politikregel gefolgt wäre. Dies bestätigt die von Cukierman und Miscatelli (2008) gefundene Rezessions-Vermeidungspräferenz der US-Notenbank.

Die vorliegende Dissertation analysiert makroökonomische Modelle aus unterschiedlichen Perspektiven und liefert einige interessante neue Erkenntnisse.
Erstens: Unterschiede in der makroökonomischen Modellierung haben wichtige Implikationen. Verschiedene Modelle können erheblich unterschiedliche Implikationen für politische Handlungsempfehlungen haben. Das sollte in der modellbasierten Forschung zur geldpolitischen und fiskalischen Stabilitätspolitik berücksichtigt werden. Ein geeignetes Analyseinstrumentarium für entsprechende vergleichende Studien wird vorgestellt.
Zweitens: Die Prognoseheterogenität, die durch verschiedene Neu-Keynesianische Modelle generiert wird, entspricht ungefähr der Heterogenität von Expertenprognosen. Die Ergebnisse können als Indiz dafür angesehen werden, dass die beobachtbaren Unterschiede von Prognosen durch die Uneinigkeit über die richtige Modellierung - und weniger durch die Irrationalität einzelner Vorhersagen - erklärt werden können. Heterogenene Erwartungen in makroökonomischen Modellen zu berücksichtigen, ist eine wichtige Aufgabe für die Zukunft, da heterogene Erwartungen selbst zu endogener Volatilität führen können.
Drittens: Stilisierte dynamische stochastische allgemeine Gleichgewichtsmodelle liefern überraschend genaue Prognosen, obwohl sie den Informationsgehalt nur weniger Zeitreihen nutzen. Weitere Arbeiten sind erforderlich, welche die strukturelle Interpretierbarkeit dieser Prognosen demonstrieren.
Viertens: Aktuelle Neu-Keynesianische Modelle überschätzen die tatsächlich zu erwartende Unsicherheit. Die Entwicklung von Modellen mit stärkeren Multiplikatoreffekten und einem größeren Grad an endogen generierter Volatilität kann die Qualität von Dichteprognosen in der Zukunft verbessern.
Fünftens: Eine Kombination von Vorhersagen mehrerer struktureller Modelle kann die Prognosegüte erhöhen. Für zukünftige Anwendungen ist es daher sinnvoll, mehrere Prognosemodelle gleichzeitig zur Berechnung von Vorhersagen zu nutzen.
Sechstens: Strategische Komplementaritäten in der Preissetzung können zu einer neuen Dimension des Wechselkurskanals der geldpolitischen Transmission führen. Das kann den Zielkonflikt der Geldpolitik entschärfen. Die empirische Relevanz des Kanals hängt von der Handelsintegration, der Durchlässigkeit von Wechselkursschwankungen auf Importpreise und der Stärke von strategischen Preiskomplementaritäten ab.
Siebtens: Einfache, symmetrische geldpolitische Regeln bieten eine unzureichende Beschreibung der tatsächlichen Geldpolitik, da asymmetrische empirische Reaktionen auf die Inflation, die Produktionslücke und den Zins gefunden wurden. Zukünftige makroökonomische Modelle sollten diese realistischere Darstellung der Geldpolitik beinhalten.

References

Adam, K., Billi, R. M., 2006. Optimal monetary policy under commitment with a zero bound on nominal interest rates. Journal of Money, Credit, and Banking 38(7), 1877–1905.

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

Adolfson, M., Laséen, S., Lindé, J., Villani, M., 2007. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Journal of International Economics 72(2), 481–511.

Bernanke, B. S., Boivin, J., 2003. Monetary policy in a data-rich environment. Journal of Monetary Economics 50(3), 525–546.

Blanchard, O., Fischer, S., 1989. Lectures on Macroeconomics. The MIT Press.

Branch, W. A., Evans, G. W., 2010. Monetary policy with heterogeneous expectations. Economic Theory forthcoming.

Branch, W. A., McGough, B., 2010. Business cycle amplification with heterogeneous expectations. Economic Theory forthcoming.

Brock, W., Hommes, C., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22, 1235–1274.

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

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

Bryant, R. C., Hooper, P., Mann, C., 1993. Design and implementation of the empirical simulations. In: Evaluating Policy Regimes: New Research in Empirical Macroeconomics. The Brookings Institution, Washington DC.

Capistran, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. Journal of Money, Credit, and Banking 41, 365–396.

Chernozhukov, V., Hansen, C., 2005. An iv model of quantile treatment effects. Econometrica 73(1), 245–261.

Chevapatrakul, T., Kim, T.-H., Mizen, P., 2009. The taylor principle and monetary policy approaching a zero bound on nominal rates: Quantile regression results for the united states and japan. Journal of Money, Credit and Banking 41(8), 1705–1723.

Chiarella, C., Dieci, R., He, X.-Z., 2007. Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework. Journal of Economic Behavior and Organization 62, 408–427.

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

Christoffel, K., Coenen, G., Warne, A., 2008. The New Area-Wide Model of the euro area - a micro-founded open-economy model for forecasting and policy analysis, european Central Bank Working Paper 944.

Clarida, R., Galí, J., Gertler, M., 1999. The science of monetary policy: A new keynesian perspective. Journal of Economic Literature 37(4), 1661–1707.

Coenen, G., Orphanides, A., Wieland, V., 2004. Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero. Berkeley Electronic Press: Advances in Macroeconomics 4(1), 1–23.

Coenen, G., Wieland, V., 2005. A small estimated euro area model with rational expectations and nominal rigidities. European Economic Review 49, 1081–1104.

Cukierman, A., Muscatelli, A., 2008. Nonlinear taylor rules and asymmetric preferences in central banking: Evidence from the united kingdom and the united states. The B.E. Journal of Macroeconomics 8(1).

de Grauwe, P., 2010. Animal spirits and monetary policy. Economic Theory forthcoming.

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of New Keynesian models. Journal of Business and Economic Statistics 25(2), 123–143.

Diebold, F. X., Gunther, T. A., Tay, A. S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39(4), 863–883.

Diebold, F. X., Hahn, J., Tay, A. S., 1999. Multivariate density forecast evaluation and calibration in financial risk management: High-frequency returns on foreign exchange. Review of Economics and Statistics 81(4), 661–673.

Dieppe, A., Kuester, K., McAdam, P., 2005. Optimal monetary policy rules for the euro area: An analysis using the area wide model. Journal of Common Market Studies 43 (3), 507–5372.

Dolado, J. J., Maria-Dolores, R., Naveira, M., 2005. Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the phillips curve. European Economic Review 49, 485–503.

Faust, J., Wright, J. H., 2009. Comparing Greenbook and reduced form forecasts using a large realtime dataset. Journal of Business and Economic Statistics 27(4), 468–479.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the Euro area? Journal of Monetary Economics 50, 1243–1255.

Fuhrer, J. C., Moore, G., 1995. Inflation persistence. The Quarterly Journal of Economics 110(1), 127–159.

Gerard, H., Nimark, K., 2008. Combing multivariate density forecasts using predictive criteria, research Discussion Paper 2008-2, Reserve Bank of Australia.

Gerlach, S., 2000. Asymmetric policy reactions and inflation, working paper, Bank for International Settlements.

Giordani, P., Söderlind, P., 2003. Inflation forecast uncertainty. European Economic Review 47, 1037–1059.

Greenspan, A., September 1997. Rules vs. discretionary monetary policy, speech at the 15th Anniversary Conference of the Center for Economic Policy Research at Stanford University, Stanford, California.

Guerrieri, L., Gust, C., López-Salido, D., 2008. International competition and inflation: A new keynesian perspective, international Finance Discussion Papers, 918.

Gust, C., Leduc, S., Vigfusson, R. J., 2006. Trade integration, competition, and the decline in exchange-rate pass-through., international Finance Discussion Papers, Number 864, Board of Governors of the Federal Reserve System.

Hsiao, C., Wan, S. K., 2010. Is there an optimal forecast combination?, working Paper University of Southern California.

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

Kato, R., Nishiyama, S.-I., 2005. Optimal monetary policy when interest rates are bounded at zero. Journal of Economic Dynamics & Control 29, 97–133.

Klein, L. (Ed.), 1991. Comparative Performance of U.S. Econometric Models. Oxford, Eng.: Oxford University Press.

Kurz, M., 1994a. On rational belief equilibria. Economic Theory 4, 859–876.

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

Kurz, M., 1996. Rational beliefs and endogenous uncertainty: an introduction. Economic Theory 8, 383–397.

Kurz, M., 1997a. Endogenous economic fluctuations and rational beliefs: A general perspective. In: Kurz, M. (Ed.), Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M. (Ed.), 1997b. Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M., 2009. Rational diverse beliefs and market volatility. In: Hens, T., Schenk-Hoppe, K. (Eds.), Handbook of financial markets: dynamics and evolution. North Holland.

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

Kurz, M., Jin, H., Motolese, M., 2005. The role of expectations in economic fluctuations and the efficacy of monetary policy. Journal of Economic Dynamics & Control 29, 2017–2065.

Laxton, D., Pesenti, P., 2003. Monetary rule for small, open, emerging economies. Journal of Monetary Economics 50, 1109–1146.

Marcellino, M., Stock, J., Watson, M., 2003. Macroeconomic forecasting in the Euro area: Country-specific versus area-wide information. European Economic Review 47, 1–18.

McCallum, B., 1988. Robustness properties of a rule for monetary policy. Carnegie-Rochester Conference Series on Public Policy 29, 173–204.

McCallum, B., 1999. Issues in the design of monetary policy rules. In: Taylor, J. B., Woodford, M. (Eds.), Handbook of Macroeconomics. Amsterdam: Elsevier Science, North-Holland.

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

Meyer, L. H., Swanson, E. T., Wieland, V., 2001. Nairu uncertainty and nonlinear policy rules. American Economic Review 91(2), 226–231.

Orphanides, A., 2003. The quest for prosperity without inflation. Journal of Monetary Economics 50, 633–663.

Orphanides, A., Wieland, V., 2000. Efficient monetary policy design near price stability. Journal of the Japanese and International Economies 14, 327–365.

Poole, W., August 2006. Understanding the fed, speech at the Dyer County Chamber of Commerce Annual Membership Luncheon, Dyersburg, Tenn.

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

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

Rudebusch, G. D., Svensson, L. E. O., 1999. Policy rules for inflation targeting. In: Taylor, J. B. (Ed.), Monetary Policy Rules. Chicago: University of Chicago Press.

Sbordone, A. M., 2007. Globalization and inflation dynamics: the impact of increased competition, federal Reserve Bank of New York.

Schaling, E., 1999. The nonlinear phillips curve and inflation forecast targeting, bank of England Working Paper No. 98.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the euro area. Journal of the European Economic Association. 1 (5), 1123–1175.

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

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

Stock, J., Watson, M., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.

Sugo, T., Teranishi, Y., 2005. The optimal monetary policy rule under the non-negativity constraint on nominal interest rates. Economics Letters 89, 95–100.

Surico, P., 2007. The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. Journal of Economic Dynamics & Control 31, 305–324.

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

Taylor, J. B., 1993b. Macroeconomic Policy in a World Economy. W.W. Norton, New York, online Edition available on: http://www.stanford.edu/ johntayl/MacroPolicyWorld.htm.

Taylor, J. B., 1999. Monetary Policy Rules. The University of Chicago Press.

Tillmann, P., 2010. Parameter uncertainty and non-linear monetary policy rules. Macroeconomic Dynamics forthcoming, working Paper, University Giessen.

Timmermann, A., 2006. Forecast combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Amsterdam: North Holland, pp. 135–196.

Wang, M.-C., 2009. Comparing the DSGE model with the factor model: An out-of-sample forecasting experiment. Journal of Forecasting 28(2), 167–182.

Chapter 1

A New Comparative Approach to Macroeconomic Modeling and Policy Analysis

(with Volker Wieland, Tobias Cwik Gernot J. Müller and Sebastian Schmidt)

Abstract Macroeconomic model comparison projects have helped produce some very influential insights such as the Taylor rule. However, they have been infrequent and costly, because they require the input of many teams of researchers and multiple meetings to obtain a limited set of comparative findings. This chapter provides a new comparative approach to model-based research and policy analysis that enables individual researchers to conduct model comparisons easily, frequently, at low cost and on a large scale. Using this approach a model archive is built that includes many well-known empirically estimated models that may be used for quantitative analysis of monetary and fiscal stabilization policies. A computational platform is created that allows straightforward comparisons of models' implications. Its application is illustrated by comparing different monetary and fiscal policies across selected models. Researchers can easily include new models in the data base and compare the effects of novel extensions to established benchmarks thereby fostering a comparative instead of insular approach to model development.

Keywords: macroeconomic models, model uncertainty, policy rules, robustness, monetary policy, fiscal policy, model comparison.

JEL-Codes: E52, E58, E62, F41

1.1 Introduction

According to Lucas (1980) the objective of macroeconomic model building is

> " *to provide fully articulated, artificial economic systems that can serve as laboratories in which policies that would be prohibitively expensive to experiment with in actual economies can be tested out at much lower cost. [...] Our task as I see it [...] is to write a FORTRAN program that will accept specific economic policy rules as 'input' and will generate as 'output' statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies.*"

Since then, many new models that aim to explain the behavior of the main aggregates of the world's economies have been developed. Model builders include not only academics but also researchers at many central banks, treasuries and international organizations. Not surprisingly, the models differ in terms of economic structure, estimation methodology and parameter estimates.

If one model were to be found to dominate all others in terms of theoretical appeal and empirical fit, this model could be used exclusively to develop policy recommendations. Or, if no model's structure is considered completely satisfactory from a theoretical perspective, and if many of the competing models describe historical data of key aggregates reasonably well, one could use these models to establish "robustness" of policy recommendations.[1] Yet, systematic comparisons of the empirical implications of a large variety of available models are rare. Evaluating the performance of different policies across many models typically is work intensive and costly. The six comparison projects reported in Bryant et al. (1988), Bryant et al. (1989), Klein (1991), Bryant et al. (1993), Taylor (1999) and Hughes-Hallett and Wallis (2004) have involved multiple teams of researchers, each team working only with one or a small subset of available models. While these initiatives have helped produce some very influential insights such as the Taylor rule,[2] the range of systematic, comparative findings has remained limited.

This chapter provides a new comparative approach to model-based research and policy analysis that enables individual researchers to conduct systematic model comparisons and policy evaluations easily and at low cost. Following this approach it is straightforward to include new models and compare their empirical and policy implications to a large number of established benchmarks.

We start by presenting a formal exposition of our approach to model comparison. 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 pre-condition

[1]Such an approach is recommended by McCallum (1988), McCallum (1999), Blanchard and Fischer (1989), Taylor (1999) and many others. McCallum (1999), for example, proposes *" to search for a policy rule that possesses robustness in the sense of yielding reasonably desirable outcomes in policy simulation experiments in a wide variety of models."* Taylor and Wieland (2009) follow this recommendation and investigate the policy implications of three well-known models of the U.S. economy available in this data base.

[2]Taylor (1993a) credits the comparison project summarized in Bryant et al. (1993) as the crucial testing ground for what later became known as the Taylor rule.

for a systematic comparison of particular model characteristics. On this basis, common policy rules can be defined as model input. Then we derive comparable objects that may be produced 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. An illustrative example with two well-known small New Keynesian models is provided.

Next, we give a brief overview of the model archive that we have built. This data base includes many well-known empirically-estimated macroeconomic models that may be used for quantitative analysis of monetary and fiscal stabilization policies. These are models of the U.S. and Euro area economies and several multi-country models. Some of the models are fairly small and focus on explaining output, inflation and interest rate dynamics (cf. Clarida et al (1999), Rotemberg and Woodford (1997), Fuhrer and Moore (1995), McCallum and Nelson (1999), Coenen and Wieland (2005), etc.). Others are of medium scale and cover many key macroeconomic aggregates (cf. Christiano, Eichenbaum and Evans (2005), Coenen et al. (2004), Smets and Wouters (2003, 2007)). Some models in the data base are fairly large in scale such as the Federal Reserve's FRB-US model of Reifschneider et al. (1999), the model of the G7 economies of Taylor (1993b) or the ECB's Area-wide model of Dieppe et al. (2005). 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 Keynesian models fully incorporate recent advances in terms of microeconomic foundations. Well-known examples of this class are models by Christiano et al. (2005), Smets and Wouters (2003, 2007), Laxton and Pesenti (2003) and Adolfson et al. (2007). In addition, we have included models that assign little role to forward-looking behavior by economic agents (cf. the ECB's Area-wide model) or none at all (cf. Rudebusch and Svensson (1999) and Orphanides (2003)).

We have created a computational platform that implements our approach to model comparison. It allows users to solve structural models and conduct comparative analysis. Comparisons of impulse response functions of common variables in response to common shocks, or of autocorrelation functions of common variables in response to model-specific shocks, or of unconditional distributions of common variables are generated. It can also be used to conduct a systematic investigation of policy rules across models. Paraphrasing Lucas (1980), *we have completed the task of writing a program that will accept specific economic policy rules as common input for multiple economic models and will generate as output a comparison of statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies according to different economic models*. The platform admits nonlinear as well as linear models and allows for perturbation-based approximation of nonlinear models with forward-looking variables.[3] New models may easily be introduced and compared to established benchmarks thereby fostering a comparative rather than insular approach to model building.

[3]This software is written for MATLAB and utilizes DYNARE software for model solution. For further information on DYNARE see Juillard (2001) and Juillard (1996).

Finally, the comparative approach to modeling and policy analysis is illustrated with several examples. We compare monetary and fiscal policy shocks under alternative monetary policy rules, and investigate the predictions of different models and different policies for inflation and output persistence. A detailed description of the models included in the data base is provided in the Appendix A.1, respectively.

1.2 A general approach to model comparison

Macroeconomic models differ in terms of modeling assumptions. They may include different economic concepts and therefore different variables and parameters; they may use different policy rules; and invariably they tend to use different notation and definitions of the same key macroeconomic aggregates. As a consequence, model output is not directly comparable. In the following, we describe formally how to augment any model in a way that renders comparison of policy implications across models straightforward, while keeping the number of necessary modifications of the original models at a minimum.

1.2.1 Augmenting models for the purpose of comparison

We start by introducing the notation for a general nonlinear macroeconomic model of the economy. The letter m is used to refer to a specific model considered in the comparison. Thus, $m = (1, 2, 3, ..., M)$ will appear as a superscript on any variables or parameters that are part of this model.[4] These variables or parameters need not be comparable across models nor follow particular naming conventions across models. Our notation regarding the vectors model-specific variables, parameters, and shocks is summarized in Table 1.1.

We distinguish two types of model equations, policy rules, which we denote by $g_m(.)$, and the other equations and identities that make up the rest of the model, that we denote by $f_m(.)$. The two types of equations together determine the endogenous model variables, which are denoted by the vector x_t^m. The model variables are functions of each other, of model-specific shocks, $(\epsilon_t^m \eta_t^m)$, and of model parameters $(\beta^m \gamma^m)$. A particular model m may then be defined as follows:

$$E_t[g_m(x_t^m, x_{t+1}^m, x_{t-1}^m, \eta_t^m, \gamma^m)] = 0 \tag{1.1}$$

$$E_t[f_m(x_t^m, x_{t+1}^m, x_{t-1}^m, \epsilon_t^m, \beta^m)] = 0 \tag{1.2}$$

The superscript m refers to the original version of the respective model as supplied by the developers. The model may include current values, lags and the expectation of leads of endogenous variables. In

[4]In the computational implementation m may be associated with a particular list of model names rather than a list of numbers.

Table 1.1: Model-specific variables, parameters, shocks and equations

Notation	Description
x_t^m	endogenous variables in model *m*
$x_t^{m,g}$	policy variables in model *m* (also included in x_t^m)
η_t^m	policy shocks in model *m*
ϵ_t^m	other economic shocks in model *m*
$g_m(.)$	policy rules in model *m*
$f_m(.)$	other model equations in model *m*
γ^m	policy rule parameters in model *m*
β^m	other economic parameters in model *m*
Σ^m	covariance matrix of shocks in model *m*

equations (1.1) and (1.2) the lead- and lag-lengths are set to unity. This assumption is for notational convenience only and should not be understood as a restriction on the type of model that is admitted.[5] The model may also include innovations that are random variables with zero mean and covariance matrix, Σ^m:

$$E([\eta_t^m \epsilon_t^m]') = 0 \tag{1.3}$$

$$E([\eta_t^{m\prime} \epsilon_t^{m\prime}]'[\eta_t^{m\prime} \epsilon_t^{m\prime}]) = \Sigma^m = \begin{pmatrix} \Sigma_\eta^m & \Sigma_{\eta\epsilon}^m \\ \Sigma_{\eta\epsilon}^m & \Sigma_\epsilon^m \end{pmatrix} \tag{1.4}$$

In the following we refer to innovations interchangeably as shocks. Some model authors instead differentiate between serially correlated economic shocks that are themselves functions of random innovations. This practice does not prevent us from including such models in a comparison. The serially correlated economic shocks of these authors would appear as elements of the vector of endogenous variables x_t^m and only their innovations would appear as shocks in our notation. Equation (1.4) distinguishes the covariance matrices of policy shocks and other economic shocks as Σ_η^m and Σ_ϵ^m. The correlation of policy shocks and other shocks is typically assumed to be zero, $\Sigma_{\eta\epsilon}^m = 0$.

If one wants to compare the implications of different models, it is necessary to define a limited set of comparable variables, shocks and parameters that will be in common to all models considered in the comparison exercise. It is then possible to express policies in terms of particular parameters, variables and policy shocks that are identical across models, and study the consequences of these policies for a set of endogenous variables that are defined in a comparable manner across models. Our notation for common endogenous variables, policy instruments, policy shocks, policy rules and parameters is introduced in Table 1.2.

[5]The software implementation does not restrict the lead- and lag-lengths of participating models.

Table 1.2: Comparable common variables, parameters, shocks and equations

Notation	Description
z_t	common variables in all models
z_t^g	common policy variables in all models (also included in z_t)
η_t	common policy shocks in all models
$g(.)$	common policy rules
γ	common policy rule parameters

Any model that is meant to be included in a comparison first has to be augmented with common variables, parameters and shocks. Augmenting the model implies adding equations. These additional equations serve to define the common variables in terms of model specific variables. We denote these definitional equations or identities by $h_m(.)$. By their nature they are model-specific. A further step is to replace the original model-specific policy rules with the common policy rules. All the other equations, variables, parameters and shocks may be preserved in the original notation of the model developers. As a consequence, the augmented model consists of three components: (i) the common policy rules, $g(.)$, expressed in terms of common variables, z_t, policy shocks, η_t, and policy rule parameters, γ; (ii) the model-specific definitions of common variables in terms of original model-specific endogenous variables, $h_m(.)$, with parameters θ^m; (iii) the original set of model-specific equations $f_m(.)$ that determine the endogenous variables. Thus, the augmented model may be represented as follows:

$$E_t[g(z_t, z_{t+1}, z_{t-1}, \eta_t, \gamma)] = 0 \tag{1.5}$$

$$E_t[h_m(z_t, x_t^m, x_{t+1}^m, x_{t-1}^m, \theta^m)] = 0 \tag{1.6}$$

$$E_t[f_m(x_t^m, x_{t+1}^m, x_{t-1}^m, \epsilon_t^m, \beta^m)] = 0 \tag{1.7}$$

Models augmented in this manner can be used in comparison exercises. For example, it is possible to compare the implications of a particular policy rule for the dynamic properties of those endogenous variables that are defined in a comparable manner across models. An advantage of this approach is that it requires only a limited set of common elements. With regard to the remainder of the model the original notation used by model authors can be left unchanged, in particular the variable names and definitions of endogenous variables, x_t^m, the other economic shocks ϵ_t^m, the equations $f_m(.)$ with model parameters β^m and the covariance matrix of shocks Σ_ϵ^m. The covariance matrix of policy shocks Σ_η may be treated as an element of the vector of policy parameters or constrained to zero.
The essential step in introducing a new model in a comparison exercise is to define the common variables in terms of model-specific variables. It involves setting up the additional equations, $h_m(.)$, and determining the definitional parameters, θ^m. We illustrate this process with an example.

A simple example

The vector of common variables, z_t, is assumed to contain six variables that are meant to be comparable across models:

$$z_t = [\ i_t^z \quad g_t^z \quad \pi_t^z \quad p_t^z \quad y_t^z \quad q_t^z\ ]' \tag{1.8}$$

These variables are characterized in Table 1.3. They are expressed in percentage deviations from steady state values, because the example applies to linear models. The monetary policy instrument is

Table 1.3: Comparable common variables

Notation	Description
i_t^z	annualized quarterly money market rate
g_t^z	discretionary government spending (share in GDP)
π_t^z	year-on-year rate of inflation
p_t^z	annualized quarter-to-quarter rate of inflation
y_t^z	quarterly real GDP
q_t^z	quarterly output gap (dev. from flex-price level)

the annualized short-term money market rate in quarter t denoted by i_t^z. The fiscal policy instrument is discretionary government spending expressed in terms of its share in GDP and denoted by g_t^z. Economic outcomes are measured with regard to inflation, real output and the output gap. π_t^z denotes the year-on-year rate of inflation, while p_t^z refers to the annualized quarter-to-quarter rate of inflation. y_t^z is quarterly real GDP. q_t^z refers to the output gap defined as the difference between actual output and the level of output that would be realized if the price level were flexible.[6]

Next, we define common monetary and fiscal policy rules. The monetary rule serves to determine the nominal interest rate, i_t^z. It includes a systematic response to output and inflation, defined in comparable terms, as well as a monetary policy shock. The fiscal rule determines discretionary government spending, g_t^z. It is simply defined as the product of a random innovation and a policy parameter:

$$i_t^z = \gamma_i i_{t-1}^z + \gamma_p p_t^z + \gamma_q q_t^z + \eta_t^i \tag{1.9}$$

$$g_t^z = \gamma_g \eta_t^g \tag{1.10}$$

The common policy shocks and parameters are denoted by:

$$\eta_t = [\ \eta_t^i \quad \eta_t^g\ ] \tag{1.11}$$

$$\gamma = [\ \gamma_i \quad \gamma_p \quad \gamma_q \quad \gamma_g\ ] \tag{1.12}$$

[6] The latter concept of potential output is used in whichever way a particular model defines it. Another interesting exercise would be to compare different concepts of potential output and output gaps across models by introducing additional common variables.

Having defined common variables, shocks and policy parameter, we proceed to consider two simple New Keynesian models for conducting a model comparison, $m = \{1, 2\}$. One model is taken from Clarida et al. (1999), ($m = 1$ refers to the model name NK_CGG99), while the other one is from Woodford (2003) and based on Rotemberg and Woodford (1997), ($m = 2$ refers to NK_RW97). These are well-known benchmarks in the literature. We present the original model equations as published by the authors and then show how to augment them appropriately for a comparison exercise. This step may seem trivial in the case of such simple models, but it is nevertheless important in order to avoid a case of comparing apples and oranges.

Table 1.4: Model 1 - The hybrid model of Clarida, Galí and Gertler (1999) (NK_CGG99)

Description	Equations and Definitions
Original Model	
variables	$x_t^1 = [\ i_t \quad x_t \quad \pi_t\ ]', \quad x_t^{1,g} = [i_t]$
shocks	$\epsilon_t^1 = [\ g_t \quad u_t\ ]'$
parameters	$\beta_1 = [\ \varphi \quad \theta \quad \phi\ ]', \quad \gamma_1 \quad = [\ \alpha \quad \gamma_\pi \quad \gamma_x\ ]'$
model equations	
$g_1(.)$	$i_t = \alpha + \gamma_\pi(\pi_t - \bar{\pi}) + \gamma_x x_t$
$f_1(.)$	$x_t = -\varphi(i_t - E_t\pi_{t+1}) + \theta x_{t-1} + (1-\theta)E_t x_{t+1} + g_t$
...	$\pi_t = \lambda x_t + \phi\pi_{t-1} + (1-\phi)\beta E_t\pi_{t+1} + u_t$
Augmented Model	
$z_t, \eta_t, \gamma, g(.)$	as defined by equations (1.8-1.12).
$f_1(.)$	as defined above in original model.
$h_1(z_t, x_t^1, E_t x_{t+1}^1, x_{t-1}^1, \theta^1)$	$i_t^z = 4i_t$
...	$\pi_t^z = \pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}$
...	$p_t^z = 4\pi_t$
...	$q_t^z = x_t$

The Clarida et al. (1999) model is presented in Table 1.4. The model in the authors' notation consists of three equations: (i) a Phillips curve relating quarterly inflation, π_t, to inflation expectations, past inflation, the output gap, x_t, and a cost-push shock, u_t; (ii) an IS equation relating the current output gap to past and expected future gaps, the expected real interest rate, $i_t - E_t\pi_{t+1}$, and a demand shock, g_t; (iii) and a policy rule relating the quarterly interest rate to inflation and the output gap.[7] Clarida et al. (1999) call it the hybrid model because it involves forward- and backward-looking elements in

[7] These are equations 6.1, 6.2 and 7.1 in Clarida et al. (1999) respectively.

the Phillips and IS curves.

In the augmented version of the model the original policy rule is replaced with the common rule, equation (1.7). The other equations from the original model, $f_m(.) = f_1(.)$, remain unchanged. The additional equations in the augmented model, $h_m(.,\theta^m) = h_1(.,\theta^1)$, provide the appropriate definitions of common comparable variables in terms of model-specific variables.[8]

Table 1.5: Model 2 - The New Keynesian model of Woodford (2003) (NK_RW97)

Description	Equations and Definitions
Original Model	
variables	$x_t^2 = [\ \hat{i}_t \quad \pi_t \quad x_t \quad \hat{r}_t^n \quad g_t \quad u_t \quad y_t \quad y_t^n\ ]', \quad x_t^{2,g} = [\hat{i}_t]$
shocks	$\epsilon_t^2 = [\ \epsilon_{u,t}\ ] \quad \eta_t^{2,g} = [\epsilon_{g,t}]$
parameters	$\beta^2 = [\ \beta \quad \kappa \quad \sigma \quad \rho_g \quad \rho_u \quad \omega\ ]'\ ,\ \gamma_2 = [\ \phi_\pi \quad \phi_x \quad \bar{\pi} \quad \bar{x}\ ]'$
model equations	
$g_2(.)$	$\hat{i}_t = \bar{i}_t + \phi_\pi(\pi_t - \bar{\pi}) + \frac{\phi_x}{4}(x_t - \bar{x})$
$f_2(.)$	$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + u_t$
...	$x_t = E_t x_{t+1} - \sigma(\hat{i}_t - E_t \pi_{t+1} - \hat{r}_t^n)$
...	$\hat{r}_t^n = \sigma^{-1}[(g_t - y_t^n) - E_t(g_{t+1} - y_{t+1}^n)]$
...	$g_t = \rho_g g_{t-1} + \epsilon_{g,t}$
...	$u_t = \rho_u u_{t-1} + \epsilon_{u,t}$
...	$y_t = x_t + y_t^n$
...	$y_t^n = \frac{\sigma^{-1}}{\omega + \sigma^{-1}} g_t$
Augmented Model	
$z_t, \eta_t, \gamma, g(.)$	as defined by equations (1.8-1.12).
$f_2(.)$	as defined above in original model.
$h_2(z_t, x_t^2, , E_t x_{t+1}^2, x_{t-1}^2 \theta^2)$	$i_t^z = 4\hat{i}_t$
...	$g_t^z = \epsilon_{g,t}$
...	$\pi_t^z = \pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}$
...	$p_t^z = 4\pi_t$
...	$y_t^z = y_t$
...	$q_t^z = x_t$

[8] This model is defined in terms of the output gap relative to a variable called flexible price output without further information on the determination of said variable. Thus, a comparable definition of the level of output is not available in this model. Therefore, this model remains silent on the time series characteristics of the level of output, y_t^z, in the comparison exercise. It is important that a systematic approach to model comparison identifies such cases so as to avoid comparing apples and oranges. Furthermore, the model does not explicitly include government spending. Therefore, it also remains silent with regard to the common variable g_t^z.

The Rotemberg and Woodford (1997) model is presented in Table 1.5. For simplicity, the linearized version is used. Of course, the nonlinear version could similarly be augmented for comparison purposes following the approach outlined in this chapter. There are some interesting differences to the hybrid model of Clarida et al. (1999). The Rotemberg-Woodford model does not exhibit endogenous persistence due to the inclusion of lagged inflation and output in the Phillips and IS curves. Instead, however, it allows for persistence in the exogenous shocks. Furthermore, it includes government spending, the natural real interest rate and the natural level of output explicitly. The model in the notation of Woodford (2003) consists of eight equations:[9] (i) a policy rule determining the nominal interest rate, $\hat{i}_t$; (ii) a purely forward-looking Phillips curve equation that determines quarterly inflation, π_t; (iii) a forward-looking IS equation determining the quarterly output gap x_t; (iv) a definition of the natural rate of interest, $\hat{r}_t^n$; (v, vi) definitions of serially correlated government spending dynamics, g_t, and cost-push shocks u_t with random innovations,[10] $\epsilon_{g,t}$ and $\epsilon_{u,t}$; (vii, viii) and definitions of output, y_t, and the natural level of output, y_t^n.

1.2.2 Conducting a comparison

Given models augmented with common policy rules and comparable variables it is possible to conduct a proper comparison. It requires solving the augmented models, constructing appropriate objects for comparison, and defining a metric that quantifies the differences of interest.

Model solution

A solution to the general nonlinear model is obtained by solving out the expectations of future variables conditional on the available information. This step requires an assumption of how expectations are formed. So far, we have used the statistical expectation that is appropriate for models with rational expectations. Solution methods for linear and nonlinear models with rational expectations are available and implemented in the computational platform provided with the working paper version of this chapter.[11] Most of the models in the data base assume rational expectations. However, other assumptions regarding expectations formation can also be admitted.[12] Existence and uniqueness of equilibrium also need to be checked in the solution step.[13] The solution of the structural nonlinear model may then be expressed in terms of the following nonlinear reduced-form equations:

$$z_t = k_z(z_{t-1}, x_{t-1}^m, \eta_t, \epsilon_t^m, \kappa_z) \tag{1.13}$$

$$x_t^m = k_x(z_{t-1}, x_{t-1}^m, \eta_t, \epsilon_t^m, \kappa_x) \tag{1.14}$$

[9]See Woodford (2003), page 246-247, equations 1.12-1.14, 2.2-2.4.

[10]In the quantitative analysis we rely on estimates of the autoregressive parameters in the shock processes provided by Adam and Billi (2006), while we obtained the structural parameters from Woodford (2003).

[11]The software is available on www.macromodelbase.com.

[12]Examples would be the introduction of adaptive learning in the Smets and Wouters (2007) model by Slobodyan and Wouters (2007), or a version of the FRB-US model with VAR-based expectations instead of rational expectations.

[13]In linear models the Blanchard-Kahn conditions provide the necessary information.

(κ_z, κ_x) denote the reduced-form parameters, which are complex functions of the structural parameters, β^m, the policy parameters, γ, and the covariance matrix Σ^m.

Solutions of the nonlinear model can be obtained using numerical methods, for example, perturbation-based methods (Collard and Juillard, 2001) or by linearizing it around a deterministic steady state and using the methods of Uhlig (1995) (generalized eigenvalue-eigenvector problem), Klein (2000) (generalized Shur decomposition), Sims (2001) (QZ decomposition), Christiano (2002) (undetermined coefficients) and many others.

In the remainder of this section we consider the first-order approximation to the reduced form solution of the augmented nonlinear model and show how it may be used to obtain particular objects for comparison defined in terms of comparable variables. The first-order that is linear approximation to the nonlinear solution (or the linear solution to originally linear models as in the preceding example) is given by:

$$\begin{pmatrix} z_t \\ x_t^m \end{pmatrix} = K_m(\gamma) \begin{pmatrix} z_{t-1} \\ x_{t-1}^m \end{pmatrix} + D_m(\gamma) \begin{pmatrix} \eta_t \\ \epsilon_t^m \end{pmatrix} \tag{1.15}$$

where the reduced-form matrices $K_m(\gamma)$ and $D_m(\gamma)$ are complicated functions of the structural parameters including the policy parameters, γ. We denote the dependence on the other (model specific) parameters β^m with the subscript m.

With the linear reduced form in hand one can derive particular objects for comparison, for example, the dynamic response of a particular common variable (an element of z) to a policy shock conditional on a certain policy rule. Impulse response functions describe the isolated effect of a single shock on the dynamic system holding everything else constant. Formally the impulse response functions in period $t + j$ to the common monetary policy shock η_t^i are defined as:

$$IR_{t+j}^m(\gamma; \eta^i) = \begin{pmatrix} E[z_{t+j}|z_{t-1}, x_{t-1}^m, I_t] - E[z_{t+j}|z_{t-1}, x_{t-1}^m] \\ E[x_{t+j}^m|z_{t-1}, x_{t-1}^m, I_t] - E[x_{t+j}^m|z_{t-1}, x_{t-1}^m] \end{pmatrix} = K_m(\gamma)^j D_m(\gamma) I_t \tag{1.16}$$

where I_t is a vector of zeros that is augmented with a single entry equal to the size of the common policy shock, for which the impulse response is computed. Using the ordering from equation (1.8) and setting $I_t(1) = -0.01$ the sixth entry of $IR_{t+j}^1(\gamma; \eta^i)$ gives the impulse response of the output gap in the first model (NK_CGG99) to a surprise interest rate reduction of 1 percent. Similarly, the sixth entry of $IR_{t+j}^2(\gamma; \eta^i)$ gives the impulse response of the output gap in the second model (NK_RW97) to the same type of shock.

It is then straightforward to compare the impulse responses of common variables to common shocks across models and policy rules. Such a comparison provides interesting insights into the transmission channels of monetary policy. We define a metric s that measures the distance between two or more models for a given characteristic of economic time series like an impulse response function. For example, the difference in the cumulative sum of the response of the output gap to a monetary policy shock of -1 percent for the models NK_CGG99 $(m = 1)$ and NK_RW97 $(m = 2)$ is given by the sixth entry of:

$$s(\gamma, z) = \sum_{j=0}^{\infty}(IR^1_{t+j}(\gamma; \eta^i; z) - IR^2_{t+j}(\gamma; \eta^i; z)). \tag{1.17}$$

The index z is meant as a reminder that we can only compare the entries in the impulse response vector for the common variables, but not the model specific variables. For the two models we get $s(\gamma, 6) = -0.0399$ under the Taylor rule, that is when the policy parameters γ imply an inflation reaction coefficient of 1.5, an output gap reaction of 0.5 and no interest rate smoothing.
Other possible characteristics for comparison are unconditional variances and serial correlation functions. The unconditional contemporaneous covariance matrix V_0^m for $([z\ x^m]')$ is given by:

$$V_0^m = \sum_{j=0}^{\infty} {K_m}^j D_m \Sigma^m {D_m}' {K_m}^{j\prime} \tag{1.18}$$

The variance is defined by the implicit expression $V_0^m = K_m V_0^m {K_m}' + D_m \Sigma^m {D_m}'$ and is solved for with an algorithm for Lyapunov equations. Given V_0^m the autocovariance matrices of $([z\ x^m]')$ are readily computed using the relationship:

$$V_j^m = {K_m}^j V_0^m \tag{1.19}$$

Again, we can compute objects for comparison between models in terms of the unconditional variance or the serial correlations and cross-correlations of common variables. Then, suitable metrics for measuring the distance between two or more models may be calculated. For example, the absolute difference of the unconditional variance for the two models given by:

$$\omega = |V_0^1(z) - V_0^2(z)| \tag{1.20}$$

The sixth entry on the diagonal of ω constitutes the difference of the unconditional variance of the output gaps of the two simple New Keynesian models considered.Its value is given by $\omega(6, 6) = 10.7919$.
It is straightforward to construct other metrics that measure the differences between the models. In section 1.4 of this chapter, for example, we will also study autocorrelation functions of comparable variables in different models of the U.S. economy.

1.3 A data base of macroeconomic models

Implementing the approach to model comparison outlined in the preceding section on a broader scale requires an archive of benchmark models. Individual researchers may then expand this model data base by introducing new models and conducting comparative analysis. The data base that we have created includes many well-known empirically-estimated macroeconomic models. The models implemented are summarized in Table 1.6. A more detailed overview of each model is provided in Appendix A.1. The data base may easily be expanded. A description of the model comparison software is available in appendix of the working paper version of this chapter.[14] It also includes an explanation how to incorporate new models in the data base and augment them with comparable variables.

Currently, the data base includes estimated and calibrated models of the U.S. economy and the Euro area, as well as several multi-country models. Most but not all models could be classified as New Keynesian because they incorporate rational expectations, imperfect competition and wage or price rigidities. All models are dynamic, stochastic, general equilibrium (DSGE) models if the term general equilibrium is taken to refer to economy-wide models compared to models of a particular sector of the economy. Only a subset of the models could be characterized as monetary business cycle models where all behavioral equations are derived in a completely consistent manner from the optimization problems of representative households and firms. Many authors use the term DSGE model to refer to this particular class of models. Thus, our data base offers interesting opportunities for comparing policy implications of this class of models 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 which assume little or no forward-looking behavior.[15] In our view, comparative analysis of these classes of models will be useful to evaluate recently voiced criticisms that the new models are rendered invalid by the experience of the world financial crisis.

The models are grouped in four categories in Table 1.6. The first category includes small, calibrated versions of the basic New Keynesian model such as the two models discussed in section 1.2. These models concentrate on explaining output, inflation and interest rate dynamics. Some of them are calibrated to U.S. data. The model taken from Clarida et al. (2002) is a two-country version of the basic New Keynesian model.

The second category covers estimated models of the U.S. economy. It includes small models of output, inflation and interest rate dynamics such as Fuhrer and Moore (1995a) and Rudebusch and Svenson (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) that fully incorporate recent advances in terms of microeconomic foundations. The data base includes the version of Christiano, Eichenbaum and Evans model estimated by Altig, Christiano, Eichenbaum and Linde (2004) because it contains other economic shocks in addition to the monetary policy shock studied by Christiano et al (2005). Because of complications in programming the informational timing assumptions on expectations in this model in DYNARE, two versions are included, one version for simulating the consequences of the monetary policy shock and the other version for simulating the

[14] The appendix is available in the working paper version of this chapter on www.macromodelbase.com.

[15] For example, the models of Rudebusch and Svensson (1999) and Orphanides (2003) are essentially structural VAR models with some restrictions on some of the coefficients. The ECB's Area-Wide Model is a medium-size structural model but with a relatively limited role for forward-looking behavior compared to the other structural, rational expectations models in the data base.

consequences of the other economic shocks in the model. Furthermore, we have included an additional version of the Altig et al (2004) model used in Taylor and Wieland (2009) that omits the cost-channel of monetary policy.[16] The largest model of the U.S economy in the data base is the Federal Reserve's FRB-US model of Reifschneider et al. (1999). We have included a linearized version of this model with rational expectations that was previously used in Levin et al (2003).

The third category in Table 1.6 covers estimated models of the Euro area economy. Four of these models have been used in a recent study of robust monetary policy design for the Euro area by Kuester and 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 inducing inflation rigidity, and a linearized version of the Area-wide Model used at the ECB for forecasting purposes. In addition, we have included an estimated DSGE model of the Euro area recently developed at the Sveriges Riksbank.

The fourth category includes estimated and calibrated models of two or more economies. Currently, the largest model in the data base is the estimated model of the G7 economies of Taylor (1993b). The estimated model of Coenen and Wieland (2003) with rational expectations and price rigidities aims to explain inflation, output and interest rate dynamics and spill-over effects between the U.S., 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 U.S. economy and a symmetric twin.

[16] This version was created in Taylor and Wieland (2009) to evaluate the effect of this assumption in comparing the Altig et al (2004) model with the model of Smets and Wouters (2007) that features no such cost channel.

Table 1.6: Models currently available in the data base

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_MCN99pb	McCallum and Nelson (1999), (P-bar model)
2. ESTIMATED US MODELS		
2.1	US_FM95	Fuhrer and Moore (1995a)
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_SW07	Smets and Wouters (2007)
2.5	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 (2009) (mon. pol. shock)
	US_ACELswt	no cost channel as in Taylor and Wieland (2009) (tech. shocks)
2.6	US_RS99	Rudebusch and Svensson (1999)
2.7	US_OR03	Orphanides (2003)
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)
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 U.S, Euro area and Japan
4.3	EACZ_GEM03	Laxton and Pesenti (2003) model calibrated to Euro area and Czech republic
4.4	G2_SIGMA08	The Federal Reserve's SIGMA model from Erceg et al. (2008) calibrated to the U.S. economy and a symmetric twin.

1.4 Comparing monetary and fiscal policies across models: an example

We have created a computational platform that renders comparisons of impulse response functions of common variables in response to common shocks, comparisons of autocorrelation functions of common variables in response to model-specific shocks and systematic investigations of policy rules across models straightforward. This result may be described by paraphrasing Lucas (1980) as follows: *we have completed the task of writing a program (in MATLAB instead of FORTRAN) that will accept specific economic policy rules as common comparable input for multiple economic models and will generate as output a comparison across models of statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies according to different economic models*. The computational platform utilizes DYNARE software for model solution. New models may easily be introduced and compared to established benchmarks thereby fostering a comparative rather than insular approach to model building.

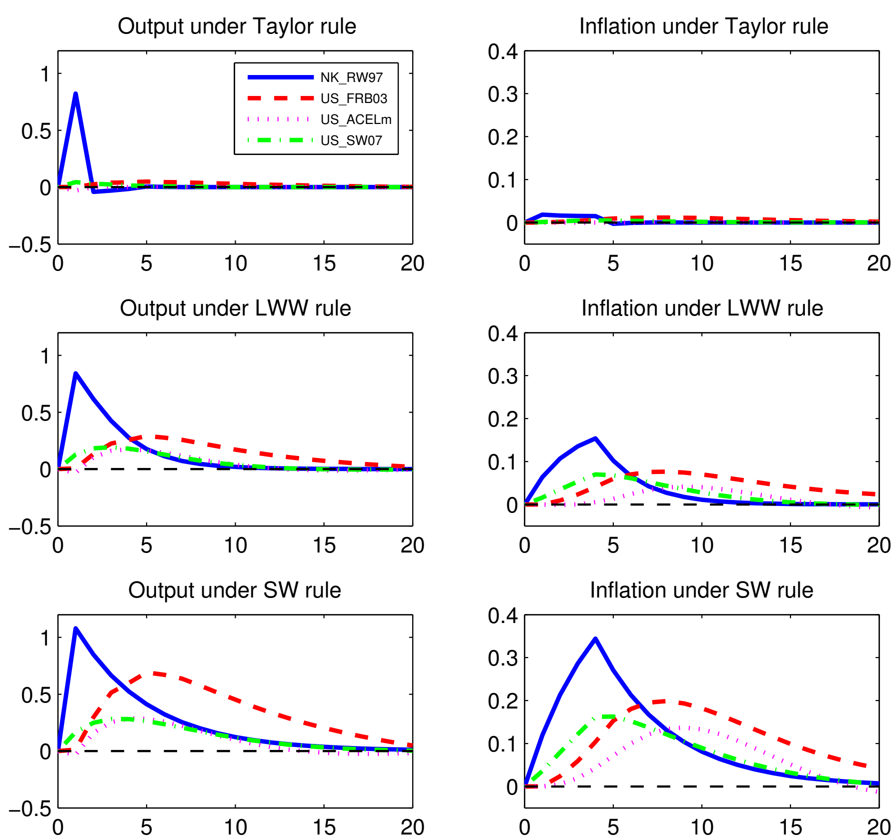
The software implementation and model database discussed in the preceding section contain a generalized interest rate rule that allows for much richer specifications than equation (1.9). For the comparison exercise in this chapter, we consider five parameterizations of this generalized rule that are taken from Taylor (1993a), Levin et al. (2003), Smets and Wouters (2007), Christiano, Eichenbaum and Evans (2005) and Gerdesmeier and Roffia (2004), respectively. The specific formulas are shown in Table 1.7.

Table 1.7: Policy rules

Taylor (1993a):	$i_t^z = \sum_{j=0}^{3} 0.38 p_{t-j}^z + 0.50 q_t^z + \eta_t^i$
Levin et al. (2003):	$i_t^z = 0.76 i_{t-1}^z + \sum_{j=0}^{3} 0.15 p_{t-j}^z + 1.18 q_t^z - 0.97 q_{t-1}^z + \eta_t^i$
Smets and Wouters (2007):	$i_t^z = 0.81 i_{t-1}^z + 0.39 p_t^z + 0.97 q_t^z - 0.90 q_{t-1}^z + \eta_t^i$
Christiano et al. (2005)	$i_t^z = 0.8 i_{t-1}^z + 0.3 E_t p_{t+1}^z + 0.08 q_t^z + \eta_t^i$
Gerdesmeier and Roffia (2004):	$i_t^z = 0.66 i_{t-1}^z + \sum_{j=0}^{3} 0.17 p_{t-j}^z + 0.10 q_t^z + \eta_t^i$

The first rule in the table, that is the simple monetary policy rule of Taylor (1993a) is well-known beyond academic economics and central banks for the following reasons. In the 1990s it became widely known that this rule described Federal Reserve interest rate decisions since 1987 surprisingly well. More recently, the large deviation of Federal Reserve policy from this rule between 2002 and 2006 has been cited as the source of cheap money fueling a housing bubble in the United States that ultimately triggered the world financial crisis. Perhaps little known is that Taylor (1993b) credits the comparison exercise of Bryant et al (1993) as the crucial testing ground that helped select this particular simple rule. Variations of the rule, motivated either by empirical estimation or model performance, abound in the literature. For comparison, we consider a rule originally estimated with U.S. data by Orphanides and Wieland (1998) and simulated in five models of the U.S. economy by Levin et al. (2003) (LWW). Their choice of models is included in our data base. The LWW rule allows for interest-rate smoothing and includes the lag of the output gap in addition to current inflation and the output gap that make up the Taylor rule. Smets and Wouters (2007) (SW) have estimated the same type of rule with interest smoothing, current inflation, current and past output gaps using Bayesian techniques together with the other structural parameters of their model. Christiano, Eichenbaum and Evans (2005) consider a different policy rule that they attribute to Clarida, Gali and Gertler (1999). Their rule includes a response to the forecast of inflation rather than current inflation. It has also been studied in Taylor and Wieland (2009). Furthermore, we add a rule estimated with Euro area data. This rule is due to Gerdesmeier and Roffia (2004) and has been simulated in Kuester and Wieland (2010) in four models of the Euro area economy that are also included in our data base.

Figure 1.1: Negative monetary policy shock



Finally, the comparative approach to macroeconomic modeling and policy analysis is applied with several examples. We compare monetary and fiscal policy shocks under alternative monetary policy rules and investigate the predictions of different models and different policies for inflation and output persistence. Figure 1.1 reports on the effect of a monetary policy shock on output and inflation in four different models of the U.S. economy under the Taylor rule, the LWW rule and the SW rule. The models considered are the calibrated New Keynesian model of Rotemberg and Woodford (1997) from Table 1.5 (NK_RW97, solid line), the Federal Reserve's FRB-US model from Levin et al. (2003) (US_FRB03, dashed line), the model of Smets and Wouters (2007) (US_SW07, dashed-dotted line) and the model of Altig et al. (2005) (US_ACELm, dotted line). The particular shock considered is a one time unexpected reduction of the nominal interest rate of 1 percentage point. Following the initial shock the nominal interest rate path corresponds to the prescriptions of the policy rule. Three rules are compared by the three rows of panels in Figure 1.1, the Taylor, LWW and SW rules.

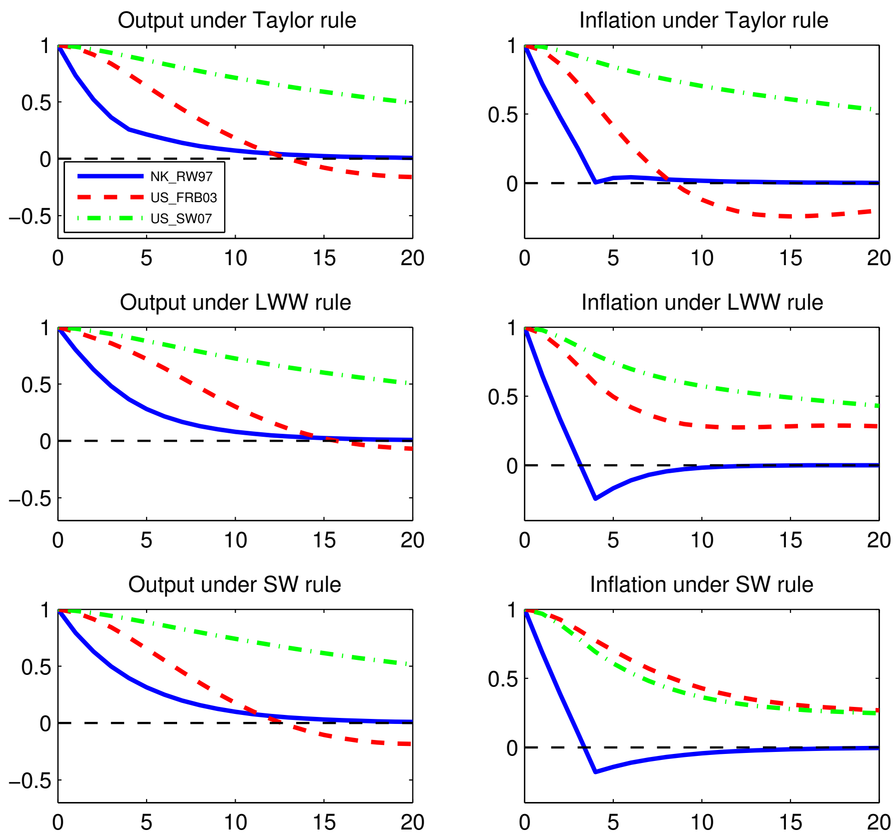
The simulation results exhibit the following findings regarding the transmission of a monetary policy shock. All four models exhibit nominal rigidities and therefore indicate that a monetary shock affects real output as indicated by the left column of panels. Under the Taylor rule, the effect on output is short-lived. In three of the four models the effect is also very small. The exception is the simple calibrated New Keynesian model (NK_RW97) which indicates a sharp large but temporary boost to output. Under the LWW and SW rules the effect on real output in the US_SW07, US_ACELm and FRB_03 models builds up over time. The reason for the larger and longer-lasting effect on real output lies in the persistent effect of the shock on interest rates due to the near-unity reaction coefficient on the lagged interest rate in these two rules. In NK_RW97 the effect on real output remains sharp and large but also peters out more slowly. An interesting difference between FRB_03 and the other models is that the peak effect of the monetary shock on real output in FRB_03 is reached only in the second year but in the first year in the other models. Thus, the models that incorporate recent advances in microeconomic foundations contradict conventional policy maker wisdom regarding "long" policy lags of more than one year. The reason for this finding is that these models give more room to the possibility of forward-looking and optimizing behavior by households and firms. The effects of the monetary shock on real output in the two estimated DSGE models with microeconomic foundations are almost identical as already noted by Taylor and Wieland (2009).

The effects of a monetary policy shock on inflation (second column of panels) are more drawn out with the peak effect occurring later than the peak in output, typically in the second or third year after the initial shock. Again, the results from the calibrated simple New Keynesian model (NK_RW97) appear too extreme relative to the findings from the empirically-estimated models.

Figure 1.2 reports the autocorrelation functions of output and inflation under the Taylor, LWW and SW monetary policy rules. These time series characteristics are derived assuming that shocks are drawn from the empirical distribution of structural shocks of these models. Only the variance of the monetary policy shock is set to zero. The Altig et al. (2005) model is omitted from the comparison because the two non-monetary shocks in that model explain only a relatively small part of the empirical output and inflation volatility in the U.S. economy (see Taylor and Wieland (2009)). The small calibrated New Keynesian model (NK_RW97) exhibits the lowest degree of output and inflation persistence among the three remaining models whichever policy rule is considered. As discussed in section 1.2 this model does not allow for lagged terms of inflation and output in the New Keynesian IS and Phillips curves.

Only, the exogenous shocks exhibit persistence in that model. The Federal Reserve's estimated model of the U.S. economy, however, implies a larger degree of output and inflation persistence. Thus, better empirical fit is obtained by allowing for a richer set of dynamics and adjustment costs that imply the appearance of one or more lags of endogenous variables in key behavioral equations.

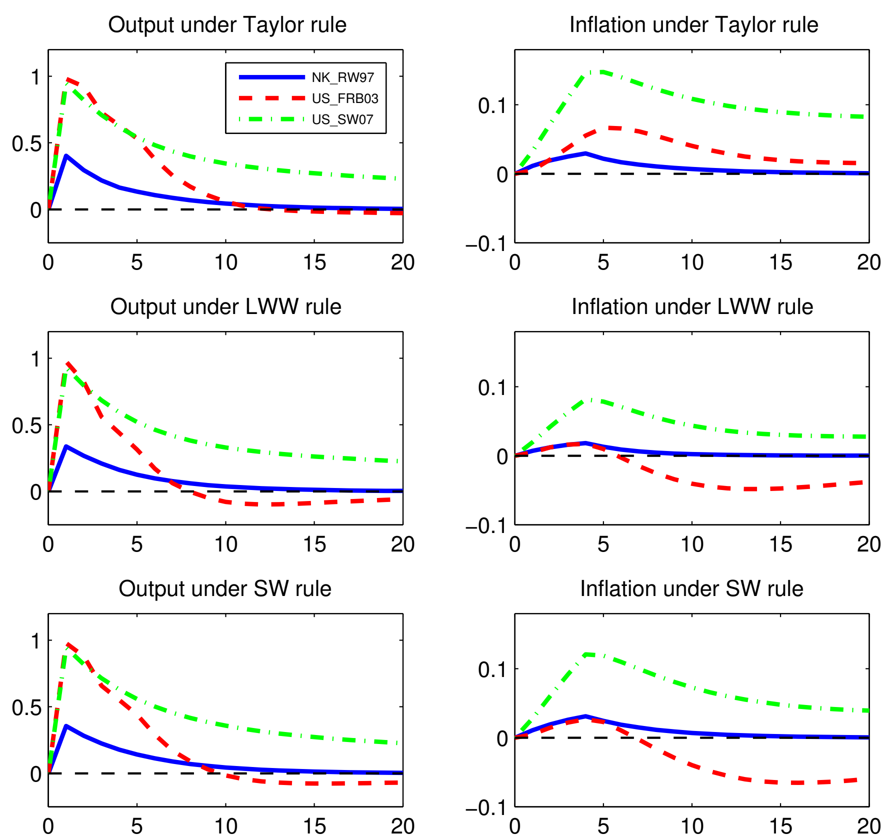
Figure 1.2: Autocorrelation functions



A rather surprising finding is that the estimated DSGE Model of Smets and Wouters (2007) exhibits the highest degree of output persistence under all three policy rules, even under the SW rule that is estimated along with the model. One might have expected that this model with microeconomic foundations would lie somewhere in between the small calibrated model of Rotemberg and Woodford (1997) and the FRB-US model. Much criticism of models such as the Federal Reserve's model was that they introduce too many adjustment costs and therefore too much endogenous persistence. Given our findings one might therefore suspect that Smets and Wouters (2007) have built in too much persistence in their model, a criticism recently voiced by Chari et al. (2009). It would be of interest to further investigate the sources of persistence in this model in future work.

Next, we turn to an evaluation of the consequences of a government spending shock of 1 percent of GDP in the three models. The fiscal policy rule for discretionary government spending is defined as in section 1.2 by equation (1.10) with a coefficient γ_g of unity. The estimated degree of persistence of such a shock to government spending differs in each model. Its implications for output and inflation are shown in Figure 1.3. In all three models, the initial shock causes output to increase in the same quarter, followed by a slow drawn-out decline over subsequent years. This profile holds under all monetary policies considered. The magnitude of the effect is rather similar for the monetary rules considered, but differs a lot across models. The impact effect is smallest in the small New Keynesian model around 0.4 percent of output, compared to about 1 percent of output in the other two models. Thus, private consumption and investment are crowded out immediately in the small model. In the other two models, private consumption and investment also decline from the start but more slowly. Somewhat surprisingly, output declines faster and inflation increases less in the US_FRB03 model than in the US_SW07 model.

Figure 1.3: Fiscal policy shock



Comparative evaluations of the consequences of fiscal policy and the robustness of policy recommendations for fiscal stimulus are particularly urgent given the amount of resources to such measures recently. Cogan et al. (2010) provide a first assessment of the American Recovery and Reinvestment of 2009. Their analysis based the Smets and Wouters (2007) model and the Taylor (1993b) model from this data base suggests that the estimates of fiscal multipliers implied by government advisers (cf. Romer and Bernstein, January 8, 2009) are far too optimistic and not robust to model uncertainty. The simulation in Figure 1.3 suggests that an evaluation using the US_FRB03 model with rational expectation would result in similar conclusions, while the NK_RW97 model would provide an even more pessimistic assessment. Interestingly, the US_FRB03 considers different components of government spending such as federal versus state expenditures and government consumption versus government investment. The shock simulated here is spread across all components according to their steady-state shares in total government spending. Further studies evaluating the non-linear timing and anticipation effects of such fiscal stimulus packages highlighted by Cogan et al. (2010) would also be of interest.

1.5 Conclusion

This chapter provides a new comparative approach to model-based research and policy analysis that enables individual researchers to conduct model comparisons easily, frequently, at low cost and on a large scale. Using this approach a model archive is built that includes many well-known empirically estimated models that may be used for quantitative analysis of monetary and fiscal stabilization policies. A computational platform is created that allows straightforward comparisons of models' implications. Its application is illustrated by comparing different monetary and fiscal policies across selected models. Researchers can easily include new models in the data base and compare the effects of novel extensions to established benchmarks thereby fostering a comparative instead of insular approach to model development. Wide application of this approach could help dramatically improve the replicability of quantitative macroeconomic analysis, reduce the danger of circular developments in model-based research and strengthen the robustness of policy recommendations.

References

Adolfson, M., Laséen, S., Lindé, J., Villani, M., 2007. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Journal of International Economics 72, 481–511.

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

Blanchard, O., Fischer, S., 1989. Lectures on Macroeconomics. The MIT Press.

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

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

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

Calvo, G., 1983. Staggered prices in a utility maximizing framework. Journal of Monetary Economics 12, 383–398.

Chari, V. V., Kehoe, P., McGrattan, E., 2009. New Keynesian models: Not yet useful for policy analysis. American Economic Journal: Macroeconomics 1(1), 242–266.

Christiano, L. J., 2002. Solving dynamic equilibrium models by a method of undetermined coefficients. Computational Economics 20(1-2), 21–55.

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

Clarida, R., Galí, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. Journal of Economic Literature 37(4), 1661–1707.

Clarida, R., Galí, J., Gertler, M., 2002. A simple framework for international monetary policy analysis. Journal of Monetary Economics 49, 879–904.

Coenen, G., Orphanides, A., Wieland, V., 2004. Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero. Berkeley Electronic Press: Advances in Macroeconomics 4(1), 1–23.

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

Coenen, G., Wieland, V., 2005. A small estimated euro area model with rational expectations and nominal rigidities. European Economic Review 49, 1081–1104.

Cogan, J. F., Cwik, T., Taylor, J. B., Wieland, V., 2010. New keynesian versus old keynesian government spending multipliers. Journal of Economic Dynamics and Control 34(3), 281–295, nBER Working Paper 14782.

Collard, F., Juillard, M., 2001. Accuracy of stochastic perturbation methods: The case of asset pricing models. Journal of Economic Dynamics & Control 25, 979–999.

Dieppe, A., Kuester, K., McAdam, P., 2005. Optimal monetary policy rules for the euro area: An analysis using the Area Wide Model. Journal of Common Market Studies 43 (3), 507–5372.

Erceg, C. J., Guerrieri, L., Gust, C., 2008. Trade adjustment and the composition of trade. Journal of Economic Dynamics & Control 32, 2622–2650.

Fagan, G., Henry, J., Mestre, R., 2005. An area-wide model for the euro area. Economic Modelling 22, 39–59.

Fuhrer, J. C., Moore, G., 1995a. Inflation persistence. The Quarterly Journal of Economics 110(1), 127–159.

Fuhrer, J. C., Moore, G., 1995b. Monetary policy trade-offs and the correlation between nominal interest rates and real output. The American Economic Review 85(1), 219–239.

Galí, J., Monacelli, T., 2005. Monetary policy and exchange rate volatility in a small open economy. Review of Economic Studies 72, 707–734.

Gerdesmeier, D., Roffia, B., 2004. Empirical estimates of reaction functions for the euro area. Swiss Journal of Economics and Statistics 140(1), 37–66.

Hughes-Hallett, A., Wallis, K. F. (Eds.), 2004. EMU macroeconomic model comparison excercise for the Euroconference 7-8 June 2002. Economic Modelling 21(5).

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.

Juillard, M., 2001. Dynare: A program for the simulation of rational expectation models, Computing in Economics and Finance 213.

Kimball, M., 1995. The quantitative analytics of the basic neomonetarist model. Journal of Money, Credit and Banking 27(4), 1241–1277.

Klein, L. (Ed.), 1991. Comparative Performance of U.S. Econometric Models. Oxford, Eng.: Oxford University Press.

Klein, P., 2000. Using the generalized Schur form to solve a multivariate linear rational expectations model. Journal of Economic Dynamics and Control 24(10), 1405–1423.

Kuester, K., Wieland, V., 2010. Insurance policies for monetary policy in the euro area. Journal of the European Economic Association 8(4), 872–912.

Laxton, D., Pesenti, P., 2003. Monetary rule for small, open, emerging economies. Journal of Monetary Economics 50, 1109–1146.

Levin, A., Wieland, V., Williams, J. C., 2003. The performance of forecast-based monetary policy rules under model uncertainty. The American Economic Review 93(3), 622–645.

Lucas, R., 1980. Methods and problems in business cycle theory. Journal of Money, Credit and Banking 12(4), 696–715.

McCallum, B., 1988. Robustness properties of a rule for monetary policy. Carnegie-Rochester Conference Series on Public Policy 29, 173–204.

McCallum, B., 1999. Issues in the design of monetary policy rules. In: Taylor, J. B., Woodford, M. (Eds.), Handbook of Macroeconomics. Amsterdam: Elsevier Science, North-Holland.

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

Orphanides, A., 2003. The quest for prosperity without inflation. Journal of Monetary Economics 50, 633–663.

Orphanides, A., Wieland, V., 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.

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

Romer, C., Bernstein, J., January 8, 2009. The job impact of the american recovery and reinvestment plan.

Rotemberg, J. J., 1982. Sticky prices in the United States. Journal of Political Economy 90 (4), 1187–1211.

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

Rudebusch, G. D., Svensson, L. E. O., 1999. Policy rules for inflation targeting. In: Taylor, J. B. (Ed.), Monetary Policy Rules. Chicago: University of Chicago Press.

Sims, C., 2001. Solving linear rational expectations models. Journal of Computational Economics 20(1-2), 1–20.

Slobodyan, S., Wouters, R., 2007. Learning in an estimated medium-scale DSGE model, Working Paper.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the euro area. Journal of the European Economic Association. 1 (5), 1123–1175.

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

Taylor, J. B., 1980. Aggregate dynamics and staggered contracts. Journal of Political Economy 88(1), 1–24.

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

Taylor, J. B., 1993b. Macroeconomic Policy in a World Economy. W.W. Norton, New York, online edition available on: http://www.stanford.edu/ johntayl/MacroPolicyWorld.htm.

Taylor, J. B., 1999. Monetary Policy Rules. The University of Chicago Press.

Taylor, J. B., Wieland, V., 2009. Surprising comparative properties of monetary models: Results from a new data base, NBER Working Paper 14849.

Tinsley, P. A., 1993. Fitting both data and theories: Polynomial adjustment costs and error-correction decision rules, FEDS Working Paper 93-21.

Uhlig, H., 1995. A toolkit for analyzing nonlinear dynamic stochastic models easily, Discussion Paper 97, Tilburg University, Center for Economic Research.

Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.

Yellen, J. L., 2007. John taylor's contributions to monetary theory and policy, speech at the Federal Reserve Bank of Dallas Conference: "John Taylor's Contributions to Monetary Theory and Policy".

Yun, T., 1996. Nominal price rigidity, money supply endogeneity, and business cycles. Journal of Monetary Economics 37, 345–370.

Appendix

A.1 A detailed overview of available models

This section describes the structure and the most important features of the different models in the macro model data base.
Most models assume that expectations of future realizations of model variables such as for example future exchange rates, prices, interest rates, wages and income are formed in a model-consistent, rational manner. A few models assume backward-looking expectations formation, in particular the models from Rudebusch and Svensson (1999) and Orphanides (2003). Most, but not all models are linear, or linear approximations of nonlinear models. In this case the variables appear as percentage deviations from their steady state values. There are many differences in model structure, in terms of size, in terms of countries covered, or the extent of microeconomic foundations considered.

A.1.1 Small calibrated models

NK_RW97: Rotemberg and Woodford (1997)

The model and the estimation strategy is discussed in detail in Rotemberg and Woodford (1997). The equations of this model can be derived from the behavior of optimizing agents. The expectational IS equation and the policy rule together can be viewed as determining aggregate demand, while the New Keynesian Phillips curve equation determines aggregate supply. The Phillips curve equation can be obtained as a log-linear approximation to the first-order condition of optimizing firms with either Calvo-style staggered price contracts (Yun, 1996) or convex costs of price adjustment (Rotemberg, 1982). The IS equation can be obtained as a log-linear approximation of the representative household's first-order equation in a model in which consumption, leisure, and real money balances are each additively separable in the utility function, and total consumption demand (private and government consumption) is equal to aggregate output.

- Aggregate Demand: Standard New Keynesian IS curve.
- Aggregate Supply: Standard New Keynesian Phillips curve.
- The Foreign Sector: no foreign sector
- Microeconomic foundation: yes
- Shocks: A cost-push shock following an AR(1) process, the common monetary policy shock, a government spending shock representing the common fiscal policy shock.
- Calibration/Estimation: Rotemberg and Woodford (1997) match the empirical impulse response functions to a monetary policy shock in a VAR (detrended real GDP, inflation, funds rate) and

the empirical variances with the variances and the theoretical impulse responses from the model to all three shocks. Quarterly U.S. data for the period 1980:1–1995:2 is used. The estimated parameters are taken from Woodford (2003) Table 6.1. However, we do not have information on the calibration of the shock processes. Hence, we employ the estimation results from Adam and Billi (2006) for the NK_RW97 shock specifications.

NK_LWW03: Levin et al. (2003)

This model is used for comparison in the robustness analysis of monetary policy rules by Levin et al. (2003). Its structure is similar to the NK_RW97 model presented above, but without explicit treatment of government spending.

- Aggregate Demand: Standard New Keynesian IS curve.
- Aggregate Supply: Standard New Keynesian Phillips curve.
- Shocks: A cost-push shock, a shock to the real interest rate, the common monetary policy shock and a fiscal policy shock that we add to the IS equation.
- Calibration/Estimation: In calibrating the model, the parameter values of Woodford (2003) adjusted for annualized variables as in Levin et al. (2003) are used.

NK_CGG99: Clarida et al. (1999), hybrid model

The model is similar to NK_RW97 but it features a hybrid Phillips curve with endogenous persistence in inflation. Also, government spending is not treated explicitly. The model and its implications for monetary policy are discussed in detail in Clarida et al. (1999) from page 1691 onwards.

- Aggregate Demand: Hybrid New Keynesian IS curve.
- Aggregate Supply: Hybrid New Keynesian Phillips curve.
- Shocks: The common monetary policy shock, cost-push shock, demand shock representing the common fiscal policy shock.
- Calibration/Estimation: Expected inflation enters the Phillips curve with a weight of 0.52 and lagged inflation with a weight of 0.48. In the IS curve the expected output gap has a weight of 0.56 and the lagged output gap has a weight of 0.44. All other parameters are the same as in the baseline model.

NK_CGG02: Clarida et al. (2002), 2-country model

Clarida et al. (2002) derive a small-scale, two-country, sticky-price model to analyse optimal monetary policy. The two countries are symmetric in size, preferences and technology.

- Aggregate Demand: Households maximize their lifetime utility, where the utility function is separable in consumption and leisure, subject to an intertemporal budget constraint. They own the firms, are a monopolistically competitive supplier of labor to the intermediate firms and additionally hold their financial wealth in the form of one-period, state-contingent bonds, which can be traded both domestically and internationally.

- Aggregate Supply: Domestic production takes place in two stages. First there is a continuum of intermediate goods firms, each producing a differentiated material input under monopolistic competition using a production function that is linear in labor input and includes an exogenous technology parameter. They set nominal prices on a staggered basis à la Calvo and receive a subsidy in percent of their wage bill to achieve an undistorted steady state. Final goods producers then combine these inputs into output, which they sell to households under perfect competition. Wages are perfectly flexible. Thus, all workers will charge the same wage and work the same amount of hours. Clarida et al. (2002) introduce an exogenous time-varying elasticity of labor demand to vary the wage-mark-up over time. The system of equations is collapsed into an IS equation and a Phillips curve, which determine the output gap and inflation, conditional on the path of the nominal interest rate both for the domestic and the foreign economy.

- Foreign Sector: Producer currency pricing is assumed so that the Law of one price holds for the final consumption good and the CPI based real exchange rate is unity. Together with the assumption of complete markets this ensures that the consumption levels are equal in both countries at any point in time.

- Shocks: Cost push shock, the common monetary policy shock and the common fiscal policy shock which is added to the IS- equation.

- Calibration/Estimation: We take the parametrization of the small open economy model in Galí and Monacelli (2005) to calibrate the model. Galí and Monacelli (2005) calibrate the stochastic properties of the exogenous driving forces by fitting AR(1) processes to log labor productivity in Canada, which is their proxy for the domestic country, and log U.S. GDP, which they use as proxy for world output. The sample period comprises 1963:1–2002:4.

NK_MCN99: McCallum and Nelson (1999)

The model in McCallum and Nelson (1999) is used to monitor the performance of operational monetary policy rules. Two distinct variants of the model are used, mainly differing in the choice of the aggregate supply setup. In the first setup, aggregate supply is based on a standard Calvo-Rotemberg (NK_MCN99cr) specification of the Phillips curve where inflation is linked to expected inflation and the output gap. In the second setup of the model, the authors introduce the so-called P-bar price adjustment (NK_MCN99pb) where price changes occur in order to gradually eliminate deviations of actual from market clearing values of output.

- Aggregate Demand: Standard New Keynesian IS and LM curve.
- Aggregate Supply: Two setups: (i) Standard New Keynesian Phillips curve (NK_MCN99cr), (ii) P-bar price adjustment (NK_MCN99pb).
- Shocks: Shock to the IS curve which follows an AR(1) process, the innovation of which represents the common fiscal policy shock, shock to the LM curve, investment shock and shock on capacity output.
- Calibration/Estimation: The model equations are estimated individually by ordinary least squares and instrumental variable estimation for U.S. data. The sample period comprises 1955-1996.

A.1.2 Estimated U.S. models

US_FM95: Fuhrer and Moore (1995a)

The model is described in Fuhrer and Moore (1995a) and Fuhrer and Moore (1995b). We employ the parametrization used in Levin et al. (2003). Fuhrer and Moore introduce a new wage contracting model where agents care about relative real wages in order to match the strong inflation persistence observed in U.S. data.

- Aggregate Demand: The US_FM95 model represents aggregate spending by a single reduced-form equation corresponding to an IS curve. The current output gap depends on its lagged values over the past two quarters and the lagged value of the long-term real interest rate, which is defined as a weighted average of ex-ante short-term real interest rates with a duration of 40 quarters.
- Aggregate Supply: The aggregate price level is a constant mark-up (normalized to one) over the aggregate wage rate. The aggregate wage dynamics are determined by overlapping wage contracts. In particular, the aggregate wage is defined to be the weighted average of current and three lagged values of the contract wage rate. The real contract wage, that is the contract wage deflated by the aggregate wage, is determined as a weighted average of expected real contract wages, adjusted for the expected average output gap over the life of the contract. This specification yields a hybrid Phillips curve that depends additionally on current and past demand and expectations about future demand.
- Shocks: Originally the model contains an ad hoc supply shock and a monetary policy shock. We add an ad hoc demand shock to the IS equation representing the common fiscal policy shock.
- Calibration/Estimation: Full-information maximum likelihood estimation on U.S. data from 1966–1994.

- Replication: We replicated the impulse response functions for annualized quarterly inflation and the output gap to a 100 basis point innovation to the federal funds rate in Figure 2 of Levin et al. (2003).

US_OW98: FRB Monetary Studies, Orphanides and Wieland (1998)

This is a small open economy model described in Orphanides and Wieland (1998) and used to investigate the consequences of the zero bound on nominal interest rates.

- Aggregate Demand: The US_OW98 model disaggregates real spending into five components: private consumption, fixed investment, inventory investment, net exports, and government purchases. The aggregate demand components exhibit partial adjustment to their respective equilibrium levels, measured as shares of potential GDP. Partial adjustments reflect habit persistence. Equilibrium consumption and fixed investment are functions of permanent income (discounted at 10 percent) and depend on the long-term real rate. The long-term nominal interest rate is an average of expected future nominal short-term rates. The long-term real rate is determined by the Fisher equation. Inventory investment depends on three lags of output. Government spending is an AR(1) process.

- Aggregate Supply: The structure is similar to the US_FM95 model. In US_FM95 and US_OW98, the aggregate price level is a constant mark-up over the aggregate wage rate.

- Foreign Sector: Net exports depend on domestic output, world output, the real exchange rate and lagged net exports. The exchange rate is determined by an UIP condition.

- Shocks: Five demand shocks including the common fiscal policy shock in the government spending equation, an ad hoc cost push shock to the nominal wage contracts and the common monetary policy shock.

- Calibration/Estimation: The model is estimated for the period 1980–1996 using U.S. data. The demand block is estimated via IV-estimation equation-by-equation. For the supply side simulation-based indirect inference methods are used.

- Replication: We replicated the impulse response functions for annualized quarterly inflation and the output gap to a 100 basis point innovation to the federal funds rate in Figure 2 of Levin et al. (2003).

US_FRB03: FRB-US model

The FRB model is a large-scale model of the U.S. economy with a relatively detailed representation of the supply side of the economy. The version US_FRB03 was linearized by Levin et al. (2003).

- Aggregate Demand: Real spending is divided into five components: private consumption, fixed investment, inventory investment, net exports and government purchases. The broad components are disaggregated further i.e. spending on fixed investment is separated into equipment, nonresidential structures and residential construction. Government spending is divided into six sub-components, each of which follows a simple reduced-form equation that includes a counter-cyclical term. The specification of most non-trade private spending equations follows the generalized adjustment cost model due to Tinsley (1993).

- Aggregate Supply: Potential output is modeled as a function of the labor force, crude energy use, and a composite capital stock, using a three-factor Cobb-Douglas production technology. The equilibrium output price is a mark-up over a weighted average of the productivity-adjusted wage rate and the domestic energy price. The specification of the wage and price dynamics follows the generalized adjustment cost framework used in the aggregate demand block. Wage inflation depends on lagged wage inflation over the previous three quarters, as well as expected future growth in prices and productivity, and a weighted average of expected future unemployment rates. Price inflation depends on its own lagged values over the past two quarters, as well as expected future changes in equilibrium prices and expected future unemployment rates. In addition, both wages and prices error-correct to their respective equilibrium levels. A vertical long-run Phillips curve is imposed in estimation. The model contains a detailed accounting of various categories of income, taxes, and stocks, an explicit treatment of labor markets, and endogenous determination of potential output. Long-run equilibrium in the model is of the stock-flow type; the income tax rate and real exchange rate risk premium adjust over time to bring government and foreign debt-to-GDP ratios back to specified (constant) levels.

- Foreign Sector: The full model includes detailed treatments of foreign variables. Twelve sectors (countries or regions) are modeled, which encompass the entire global economy. In the model used in the modelbase the full set of equations describing the foreign countries is replaced by two reduced form equations for foreign output and prices, to reduce computational cost.

- Shocks: The model exhibits a large range of shocks to which we add the common monetary policy shock and a fiscal shock that equally affects all three components of federal government spending such that a unit demand shock affects output by 1 percent.

- Replication: We replicated the impulse response functions for annualized quarterly inflation and the output gap to a 100 basis point innovation to the federal funds rate in Figure 2 of Levin et al. (2003).

US_SW07: Smets and Wouters (2007)

Smets and Wouters (2007) develop a medium-scale closed economy DSGE-Model and estimate it for the U.S. with Bayesian techniques. The model features a deterministic growth rate driven by labor-augmenting technological progress, so that the data do not need to be detrended before estimation.

- Aggregate Demand: Households maximize their lifetime utility, where the utility function is nonseparable in consumption and leisure, subject to an intertemporal budget constraint. Smets and Wouters (2007) include external habit formation to make the consumption response in the model more persistent. Households own firms, rent capital services to firms and decide how much capital to accumulate given certain capital adjustment costs. They additionally hold their financial wealth in the form of one-period, state-contingent bonds. Exogenous spending follows a first-order autoregressive process with an iid-normal error term and is also affected by the productivity shock.

- Aggregate Supply: The final goods, which are produced under perfect competition, are used for consumption and investment by the households and by the government. The final goods producer maximizes profits subject to a Kimball (1995) aggregator of intermediate goods, which introduces monopolistic competition in the market for intermediate goods and features a non constant elasticity of substitution between different intermediate goods, which depends on their relative price. A continuum of intermediate firms produce differentiated goods using a production function with Cobb-Douglas technology and fixed costs and sell these goods to the final-good sector. They decide on labor and capital inputs, and set prices according to the Calvo model. Labor is differentiated by a union using the Kimball aggregator, too, so that there is some monopoly power over wages, which results in an explicit wage equation. Labor packers buy the labor from the unions and resell it to the intermediate goods producer in a perfectly competitive environment. Sticky wages à la Calvo are additionally assumed. The Calvo model in both wage and price setting is augmented by the assumption that prices that cannot be set freely are partially indexed to past inflation rates.

- Shocks: A total factor productivity shock, a risk premium shock, an investment-specific technology shock, a wage and a price mark-up shock and two policy shocks: the common fiscal policy shock entering the government spending equation and the common monetary policy shock.

- Calibration/Estimation: The model is estimated for the U.S. with Bayesian techniques for the period 1966:1–2004:4 using seven key macroeconomic variables: real GDP, consumption, investment, the GDP deflator, real wages, employment and the nominal short-term interest rate.

- Replication: We replicated the impulse response functions to a positive one standard deviation monetary policy shock in Figure 6 of Smets and Wouters (2007). The variables include output, hours, quarterly inflation and the interest rate.

US_ACEL: CEE/ACEL by Altig et al. (2005)

The purpose of the authors is to build a model with optimizing agents that can account for the observed inertia in inflation and persistence in output (Christiano et al., 2005). In the version by Altig et al. (2005) firm-specific capital is introduced to get a Calvo parameter consistent with the microeconomic evidence of price re-optimizations on average once every 1.5 quarters. The modelbase contains four different specifications of the CEE/ACEL model, labeled by m = monetary policy shock, t = technology shock and sw = SW assumptions, i.e. no cost channel and no timing constraints as in Taylor and Wieland (2009).

- Aggregate Demand: The representative household's utility is separable in consumption and leisure and allows for habit formation in consumption. Expected-lifetime utility is maximized, choosing optimal consumption and investment, as well as the amount of capital services supplied to the intermediate firms (homogenous capital model) and portfolio decisions. Investment adjustment costs are introduced. Furthermore, the household determines the wage rate for its monopolistically supplied differentiated labor services whenever it receives a Calvo signal. In those periods, in which it does not receive a signal, the wage is increased by the lagged inflation rate augmented by the steady state growth rate of a combination of the neutral technology shock and the shock to capital embodied technology. Labor services are sold to a competitive firm that aggregates the differentiated services and supplies the resulting aggregated labor to the intermediate goods firms.
  In the firm-specific capital model, the capital stock is owned by the firms.

- Aggregate Supply: The final consumption good is produced under perfect competition using differentiated intermediate goods as inputs. Each intermediate good is producted by a monopolist employing capital (which is firm-specific in one variant of the model) and labor services. The production function is augmented by a technology shock. Capital is pre-determined. Hence, if capital is firm-specific, marginal costs depend positively on the firm's output level. Furthermore, it is assumed that the monopolistic firms have to pay the wage bill in advance which requires borrowing from a financial intermediary. Nominal frictions are introduced in the form of Calvo sticky prices. Non-reoptimizing firms index their prices to previous periods inflation.

- Shocks: The common monetary policy shock, a neutral technology shock, an investment specific technology shock and the common fiscal policy shock which is added to the resource constraint.

- Calibration/Estimation: The model has been estimated by matching the empirical impulse response functions to a monetary policy shock in a ten variable VAR with the theoretical impulse responses from the model to a monetary policy shock. Quarterly U.S. data from 1959:2–2001:4 is used.

- Replication: Using the US_ACELm model we replicated the impulse response functions for annualized quarterly inflation, output, annualized quarterly money growth and the annualized quarterly interest rate to a one standard deviation monetary policy shock.

US_RS99: Rudebusch and Svensson (1999)

Rudebusch and Svensson (1999) set up a simple linear model of the U.S. economy which is used to examine the performance of different policy rules taking into account an inflation targeting monetary policy regime. The model equations are backward looking.

- Aggregate Demand: An IS curve relates the output gap to its own lags and the difference between the average federal funds rate and the average inflation rate over the current and three preceding quarters.
- Aggregate Supply: Phillips curve of the accelerationist form.
- Shocks: A cost-push shock and a demand shock which represents the common fiscal policy shock.
- Calibration/Estimation: The model equations are estimated individually by ordinary least squares for U.S. data. The sample period comprises 1961:1-1996:2.

US_OR03:Orphanides (2003)

Orphanides (2003) conducts a counterfactual analysis based on the historical experience of the United States economy to give an example of the difficulties in identifying robust policy strategies. The counterfactual analysis gives an insight how inflation and the output gap would have evolved from the 1960s to the 1990s if the Federal Reserve had actually followed two distinct activist monetary policy rules taking into account the difference between realistic and non-realistic assumptions on the availability of information on the output gap.

- Aggregate Demand: The demand side of the structural model of the economy is represented by an IS equation which relates the output gap to its own lags, lags of inflation and the federal funds rate.
- Aggregate Supply: The supply side is represented by an accelerationist form of the Phillips curve with an adaptive representation of inflation expectations.
- Shocks: A cost-push shock and a demand shock representing the common fiscal policy shock.
- Calibration/estimation: The aggregate demand and aggregate supply equation are estimated in a setup that can be interpreted as a mildly restricted structural vector autoregression (VAR) of up to four lags estimated using quarterly data from from 1960 to 1993.

A.1.3 Estimated Euro area models

EA_CW05: Coenen and Wieland (2005)

Coenen and Wieland (2005) develop a small-scale macroeconomic model for various staggered pricing schemes. We use a version with the nominal contract specification of Taylor (1980), labeled

EA_CW05ta, and a version with the relative real wage contract specification of Fuhrer and Moore (1995a), labeled EA_CW05fm.

- Aggregate Demand: The aggregate demand equation is backward looking: two lags of aggregate demand (should account for habit persistence in consumption, adjustment costs and accelerator effects in investment) and one lag of the long-term interest rate (allows for a transmission lag of monetary policy). The long-term nominal interest rate is an average of expected future nominal short-term rates. The long-term real rate is determined by the Fisher equation.
- Aggregate Supply: As in US_FM95 and US_OW98.
- Shocks: The common monetary policy shock, the common fiscal policy shock represented by a demand shock entering the aggregate demand equation and a contract wage shock.
- Calibration/Estimation: The model has been estimated on data from the ECB Area Wide Model data set from 1974:1–1998:4. The contract wage specifications have been estimated by a limited information indirect inference technique while the IS equation has been estimated by means of the GMM.
- Replication: We replicated the impulse response functions of annual inflation and the output gap to a 100bps temporary unanticipated rise in the nominal short term rate in the upper panel of Figure 7 of Kuester and Wieland (2010) for both versions of the model.

EA_AWM05: Area Wide model linearized by Dieppe, Kuester and McAdam (2005)

The model is described in Fagan et al. (2005). It was one of the first models to treat the Euro area as a single economy. In the modelbase we use the linearized version from Dieppe, Kuester and McAdam (2005) that is also used in Kuester and Wieland (2010). The EA_AWM05 is an open economy model of the Euro area. Expectation formation is largely backward-looking. Activity is demand-determined in the short-run but supply determined in the long-run with employment having converged to a level consistent with the exogenously given level of equilibrium unemployment. Stock-flow adjustments are accounted for, e.g., the inclusion of a wealth term in consumption.

- Aggregate Demand: Demand is disaggregated into private consumption, government consumption, investment, variation of inventories, exports, and imports. The term structure (12-year bond) is forward-looking. Private consumption is specified as a function of households' real disposable income and wealth, where the latter consists of net foreign assets, public debt and the capital stock. The change in the log of the investment/output ratio depends on the real interest rate, the real GDP/capital stock ratio and the lagged investment/output ratio. The authors stress that this investment equation represents the key channel through which interest rates affect aggregate demand. Government consumption is treated as exogenous.

- Aggregate Supply: Output follows a whole economy production function. Short-run employment dynamics are driven by output growth and real wages. The deflator for real GDP at factor costs, which according to Fagan et al. (2005) is the key price index of the model, is a function of unit labor costs, import prices, the output gap and inflation expectations. The growth rate of wages depends on consumer price inflation, productivity and the unemployment gap, defined as the deviation of the current unemployment rate from the NAIRU.
- Foreign Sector: Besides extra-area flows, exports and imports also include intra-area flows. World GDP and world GDP deflator are treated as exogenous variables. The exchange rate is a forward-looking variable determined by uncovered interest rate parity.
- Shocks: Employment shock, factor cost-push shock, private consumption cost-push shock, gross investment cost-push shock, gross investment shock, exports cost-push shock, imports cost-push shock, private consumption shock, term structure shock, common fiscal policy shock and common monetary policy shock.
- Calibration/Estimation: Estimation on Euro area data equation by equation from 1970:1–1997:4, whereas the estimation period of some equations starts later, but not later than 1980:1.
- Replication: We replicated the impulse response functions of annual inflation and the output gap to a 100bps temporary unanticipated rise in the nominal short term rate in the upper panel of Figure 7 of Kuester and Wieland (2010).

EA_SW03: Smets and Wouters (2003)

The EA_SW03 model of Smets and Wouters (2003) is a medium-scale closed economy DSGE model with various frictions and estimated for the Euro area with Bayesian techniques.

- Aggregate Demand: Households maximize their lifetime utility, where the utility function is separable in consumption, leisure and real money balances, subject to an intertemporal budget constraint. Smets and Wouters (2003) include external habit formation to make the consumption response in the model more persistent. Households own firms, rent capital services to firms and decide how much capital to accumulate given certain capital adjustment costs. They additionally hold their financial wealth in the form of cash balances and one-period, state-contingent bonds. Exogenous spending is introduced by a first-order autoregressive process with an iid-normal error term.
- Aggregate Supply: The final goods, which are produced under perfect competition, are used for consumption and investment by the households and by the government. The final goods producer maximizes profits subject to a Dixit-Stiglitz aggregator of intermediate goods, which introduces monopolistic competition in the market for intermediate goods and features a constant elasticity of substitution between individual, intermediate goods. A continuum of intermediate

firms produce differentiated goods using a production function with Cobb-Douglas technology and fixed costs and sell these goods to the final-goods sector. They decide on labor and capital inputs, and set prices according to the Calvo model. Labor is differentiated over households using the Dixit-Stiglitz aggregator, too, so that there is some monopoly power over wages, which results in an explicit wage equation. Sticky wages à la Calvo are additionally assumed. The Calvo model in both wage and price setting is augmented by the assumption that prices that cannot be freely set, are partially indexed to past inflation rates.

- Shocks: Ten orthogonal structural shocks are introduced in the model. Three preference shocks in the utility function: a general shock to preferences, a shock to labor supply and a money demand shock. Two technology shocks: an AR(1) process with an iid shock to the investment cost function and a productivity shock to the production function. Three cost push-shocks: shocks to the wage and price mark-up, which are iid around a constant and a shock to the required rate of return on equity investment. And finally two monetary policy shocks: a persistent shock to the inflation objective and a temporary common monetary policy shock. In addition, the common fiscal policy shock is added in the form of a government spending shock. Since government spending is expressed in output units, we set the coefficient which scales the shock to unity to achieve a shock size of one percent of GDP.

- Calibration/Estimation: The model is estimated using Bayesian techniques on quarterly Euro area data. The data set used is comprised of seven key macroeconomic variables consisting of real GDP, real consumption, real investment, the GDP deflator, real wages, employment and the nominal interest rate over the period 1970:1–1999:4.

- Replication: We replicated the impulse response functions of annual inflation and the output gap to a 100bps temporary unanticipated rise in the nominal short term rate in the upper panel of Figure 7 of Kuester and Wieland (2010).

EA_SR07: Euro Area Model of Sveriges Riksbank, Adolfson et al. (2007)

Adolfson et al. (2007) develop an open economy DSGE model and estimate it for the Euro area using Bayesian estimation techniques. They analyse the importance of several rigidities and shocks to match the dynamics of an open economy.

- Aggregate Demand: Households maximize lifetime utility subject to a standard budget constraint. Preferences are separable in consumption, labor and real cash holdings. Persistent preference shocks to consumption and labor supply are added to the representative utility function. Internal habit formation is imposed with respect to consumption. Aggregate consumption is specified as a CES function, being composed of domestically produced as well as imported consumption goods. Households rent capital to firms. Capital services can be increased via investment and via an increase in the capital utilization rate, where both options are involved with

costs. Total investment in the domestic economy is represented by a CES aggregate consisting of domestic and imported investment goods. Households are assumed to be able to save through acquiring domestic bonds and foreign bonds in addition to holding cash and accumulating physical capital. A premium on foreign bond holdings assures the existence of a well-defined steady state. Households monopolistically supply a differentiated labor service. Wage stickiness is introduced in the form of the Calvo model augmented by partial indexation.
Government consumption of the final domestic good is financed via taxes on capital income, labor income, consumption and payroll. Any surplus or deficit is assumed to be carried over as a lump-sum transfer to households.

- Aggregate Supply: The final good is produced via a CES aggregator using a continuum of differentiated intermediate goods as inputs. The production of intermediate goods requires homogeneous labor and capital services as inputs and is affected by a unit-root technology shock representing world productivity as well as a domestic technology shock. Fixed costs are imposed such that profits are zero in steady state. Due to working capital, (a fraction of) the wage bill has to be financed in advance of the production process. Price stickiness of intermediate goods is modeled as in the Calvo (1983) model. In addition, partial indexation to the contemporaneous inflation target of the central bank and the previous periods inflation rate is included for those firms that do not receive a Calvo signal in a given period. This results in a hybrid new Keynesian Phillips curve.

- Foreign Sector: Importing firms are assumed to buy a homogeneous good in the world market and differentiate it to sell it in the domestic market. Similarly, exporting firms buy the homogeneous final consumption good produced in the domestic economy and differentiate it to sell it abroad. Specifically, the differentiated investment and consumption import goods are aggregated in a second step via a CES function, respectively. The same applies to the export goods. Calvo pricing is also assumed for the import and export sector, allowing for incomplete exchange rate pass-through in the short run. The foreign economy is described by an identified VAR model for foreign prices, foreign output and the foreign interest rate.

- Shocks: Unit root technology shock, stationary technology shock, investment specific technology shock, asymmetric technology shock, consumption preference shock, labor supply shock, risk premium shock, domestic mark-up shock, imported consumption mark-up shock, imported investment mark-up shock, export mark-up shock, inflation target shock, the common monetary policy shock, shocks to the four different tax rates and a government spending shock which represents the common fiscal policy shock and which we have adjusted so that we achieve a shock size of one percent of GDP.

- Calibration/Estimation: The model is estimated using Bayesian estimation techniques for the Euro area using quarterly data from 1970:1–2002:4 in order to match the dynamics of 15 selected variables. According to the authors, they calibrated those parameters that should be weakly identified by the 15 variables used for estimation.

- Replication: We replicated the impulse response functions for annualized quarterly inflation, output, employment and the annualized interest rate to a one standard deviation monetary policy shock in Figure 3 of Adolfson et al. (2007).

A.1.4 Estimated/calibrated multi-country models

G7_TAY93: Taylor (1993b) G7 countries

Taylor (1993b) describes an estimated international macroeconomic framework for policy analysis in the G7 countries: U.S., Canada, France, Germany, Italy, Japan and the U.K. The model consists of 98 equations and a number of identities. This model was the first to demonstrate that it is possible to construct, estimate, and simulate large-scale models for real-world policy analysis (Yellen, 2007). Taylor (1993b) argues that a multicountry model is appropriate for the evaluation of policy questions like the appropriate mix of fiscal and monetary policy or the choice of an exchange rate policy.

- Aggregate Demand: The IS components are more disaggregated than in the US_OW98 model. For example, spending on fixed investment is separated into three components: equipment, non-residential structures, and residential construction. The specification of these equations is very similar to that of the more aggregated equations in the US_OW98 model. The aggregate demand components exhibit partial adjustment to their respective equilibrium levels. In G7_TAY93, imports follow partial adjustment to an equilibrium level that depends on U.S. income and the relative price of imports, while exports display partial adjustment to an equilibrium level that depends on foreign output and the relative price of exports. Uncovered interest rate parity determines each bilateral exchange rate (up to a time-varying risk premium); e.g., the expected one-period-ahead percent change in the DM/U.S.$ exchange rate equals the current difference between U.S. and German short-term interest rates.

- Aggregate Supply: The aggregate wage rate is determined by overlapping wage contracts. In particular, the aggregate wage is defined to be the weighted average of current and three lagged values of the contract wage rate. In contrast to the US_FM95 model and the US_OW98 model, G7_TAY93 follows the specification in Taylor (1980), where the current nominal contract wage is determined as a weighted average of expected nominal contract wages, adjusted for the expected state of the economy over the life of the contract. This implies less persistence of inflation than in the US_FM95 and the US_OW98 model. The aggregate price level is not set as a constant mark-up over the aggregate wage rate as in US_FM95 and US_OW98. Prices are set as a mark-up over wage costs and imported input costs. This mark-up varies and prices adjust slowly to changes in costs. Prices follow a backward-looking error-correction specification. Current output price inflation depends positively on its own lagged value, on current wage inflation, and on lagged import price inflation, and responds negatively (with a coefficient of -0.2) to the lagged percent deviation of the actual price level from equilibrium. Import prices adjust slowly (error-correction form) to an equilibrium level equal to a constant mark-up over a weighted average of foreign prices converted to dollars. This partial adjustment of import and output prices imposes somewhat more persistence to output price inflation than would result from staggered nominal wages alone.

- Foreign Sector: G7_TAY93 features estimated equations for demand components and wages and prices for the other G7 countries at about the level of aggregation of the U.S. sector. Financial capital is mobile across countries.

- Shocks: Interest rate parity shock, term structure shock, durable consumption shock, non-durable consumption shock, services consumption shock, total consumption shock, aggregate consumption shocks for Germany and Italy, for the other countries disaggregated, nonresidential equipment investment shock, nonresidential structures investment shock, residential investment shock, inventory investment shock, fixed investment shock, inventory investment shock, real export shock, real import shock, contract wage shock, cost-push shock, import price shock, export price shock, fiscal policy shock, where we have adjusted the size of the fiscal policy shock for the U.S. - the common fiscal shock - so that a unit shock represents a 1 percent of GNP shock and a monetary policy shock where again the common modelbase monetary policy shock enters the monetary policy rule for the U.S..

- Calibration/Estimation: The model is estimated with single equation methods on G7 data from 1971–1986.

- Replication: We replicated the impulse response functions for annualized quarterly inflation and the output gap to a 100 basis point innovation to the federal funds rate in Figure 2 of Levin et al. (2003).

G3_CW03: Coenen, Wieland (2002, 2003) G3 countries

In this model different kinds of nominal rigidities are considered in order to match inflation and output dynamics in the U.S., the Euro area and Japan. Staggered contracts by Taylor (1980) explain best inflation dynamics in the Euro area and Japan and staggered contracts by Fuhrer and Moore (1995a) explain best U.S. inflation dynamics. The authors evaluate the role of the exchange rate for monetary policy and find little gain from direct policy response to exchange rates.

- Aggregate Demand: The open-economy aggregate demand equation relates output to the lagged ex-ante long-term real interest rate and the trade-weighted real exchange rate and additional lags of the output gap. The demand equation is very similar to the G7_TAY93 model without any sectoral disaggregation. Lagged output terms are supposed to account for habit persistence in consumption as well as adjustment costs and accelerator effects in investment. The lagged interest rate allows for lags in the transmission of monetary policy. The exchange rate influences net exports and thus enters the aggregate demand equation. The long term nominal interest rate is an average of expected future nominal short-term rates. The long-term real interest rate is determined by the Fisher equation.

- Aggregate Supply: For the U.S., relative real wage staggered contracts by Fuhrer and Moore (1995a) are used (see the US_FM95 model for a detailed exposition). For the Euro area and Japan the nominal wage contracts by Taylor (1980) are used. Note that Taylor contracts, with a maximum contract length exceeding two quarters, result in Phillips curves that explicitly include lagged inflation and lagged output gaps. Thus, the critique that with Taylor contracts inflation persistence is solely driven by output persistence (Fuhrer and Moore, 1995a) is mitigated.

- Foreign Sector: All three countries are modeled explicitly. The modelbase rule replaces monetary policy for the U.S.. For the Euro area and Japan the original interest rules remain. Foreign output does not affect domestic output directly, but indirectly via the exchange rate in the demand equation. The bilateral exchange rates are determined by UIP conditions.

- Shocks: Contract wage shock, demand shock, where the demand shock to the U.S. economy represents the common fiscal policy shock which is adjusted such that a unit shock has the size of one percent of aggregate demand, the common monetary policy shock which is added for the U.S..

- Calibration/Estimation: Euro area data, (fixed GDP weights at PPP rates from the ECB Area-wide model database), U.S. data and Japanese data. For the U.S. and Japan OECD's output gap estimates are used. For the Euro area log-linear trends are used to derive potential output. The estimation is robust to different output gap estimations. Demand block: GMM estimation where lagged values of output, inflation, interest rates, and real exchange rates are used as instruments. Supply side: simulation-based indirect inference methods. Estimation period: U.S. 1980:1–1998:4, Euro area 1980:1–1998:4 and Japan 1980:1–1997:1.

- Replication: We replicated the impulse response functions to 0.5 percentage points demand shocks in the United States, the Euro Area und Japan plotted in Figure 3 of Coenen, Wieland (2002). Variables include the output gap, annual inflation and the short-term nominal interest rate of the United States, the Euro Area and Japan.

EACZ_GEM03: IMF model of Euro Area and Czech Republic, Laxton and Pesenti (2003)

The model is a variant of the IMF's Global Economy Model (GEM) and consists of a small and a large open economy. The authors study the effectiveness of Taylor rules and inflation-forecast-based rules in stabilizing variability in output and inflation. They check if policy rules designed for large and relatively closed economies can be adopted by small, trade-dependent countries with less developed financial markets and strong movements in productivity and relative prices and destabilizing exposure to volatile capital flows. In contrast to Laxton and Pesenti (2003) we focus on the results for the large open economy (Euro area) rather than the small open economy (Czech Republic).

- Aggregate Demand: Infinitely lived optimizing households; government spending falls exclusively on nontradable goods, both final and intermediate. Households face a transaction cost if they take a position in the foreign bond market.

- Aggregate Supply: Monopolistic intermediate goods firms produce nontradeable goods and tradable goods. It exists a distribution sector consisting of perfectly competitive firms. They purchase tradable intermediate goods worldwide (at the producer price) and distribute them to firms producing the final good (at the consumer price). Perfectly competitive final good firms (Dixit-Stiglitz aggregator) use nontradable and tradeable goods and imports as inputs. Households are monopolistic suppliers of labor and wage contracts are subject to adjustment costs. Households own domestic firms, nonreproducable resources and the domestic capital stock. Markets for land and capital are competitive. Capital accumulation is subject to adjustment costs. Labor, capital and land are immobile internationally. Households trade a short-term nominal bond, denominated in foreign currency. All firms exhibit local currency pricing, thus exchange rate pass-through is low.

- Shocks: Risk premium shock, productivity shock, shock to the investment depreciation rate, shock to the marginal utility of consumption, government absorption shock where the one affecting the large foreign economy represents the common fiscal policy shock, shock to the marginal disutility of labor, preference shifter. We add the common monetary policy shock to the policy rule of the large economy.

- Calibration/Estimation: Calibrated to fit measures of macro-variability of the Euro area (1970:1–2000:4) and Czech Republic (1993:1–2001:4).

- Notes: Due to the symmetric setup of the model, we use the same policy rule in both countries.

- Replication: We replicated the standard deviations of annual inflation, the output gap and the first difference of the interest rate under the optimal Taylor rule implied by the loss function specification 2 of Laxton and Pesenti (2003) as listed in the second row of Table 4 in their paper.

G2_SIGMA08: FRB-SIGMA by Erceg et al. (2008)

The SIGMA model is a medium-scale, open-economy, DSGE model calibrated for the U.S. economy. Erceg et al. (2008) in particular take account of the expenditure composition of U.S. trade and analyse the implications for the reactions of trade to shocks compared to standard model specifications.

- Aggregate Demand: There are two types of households: households that maximize a utility function separable in consumption, with external habit formation and a preference shock, leisure and real money balances, subject to an intertemporal budget constraint (forward-looking households) and the remainder that simply consume after-tax disposable income (hand-to-mouth households). Households consume, own the firms and accumulate capital, which they rent to the intermediate goods producers. Erceg et al. (2008) introduce investment adjustment costs à la Christiano et al. (2005), where it is costly for the households to change the level of gross investment. Households also choose optimal portfolios of financial assets, which include domestic money balances, government bonds, state-contingent domestic bonds and a non-state contingent foreign bond. It is assumed that households in the home country pay an intermediation cost when purchasing foreign bonds, which ensures the stationarity of net foreign assets. Households rent their labor in a monopolistic market to firms, where forward-looking households set their nominal wage in Calvo-style staggered contracts analogous to the price contracts and hand-to-mouth households simply set their wage each period equal to the average wage of the forward-looking households.

- Aggregate Supply: Intermediate-goods producers have an identical CES production function and rent capital and labor from competitive factor markets. They sell their goods to final goods producers under monopolistic competition and set prices in Calvo-style staggered contracts. Firms, that don't get a signal to optimize their price in the current period, mechanically adjust their price based on lagged aggregate inflation. Final good producers in the domestic and foreign market assemble the domestic and foreign intermediate goods into a single composite good by a CES production function of the Dixit-Stiglitz form and sell the final good to households in their country. Erceg et al. (2008) introduce quadratic import adjustment costs into the final goods aggregator, which are zero in steady state. It is costly for a firm to change its share of imports in a final good relative to their lagged aggregate shares. Thus, the import share of consumption or investment goods is relatively unresponsive in the short-run to changes in the relative price of imported goods even while allowing the level of imports to jump costlessly in response to changes in overall consumption or investment demand. Government purchases are assumed to be a constant fraction of output. Government revenue consists of income from capital taxes (net of the depreciation write off), seignorage income and revenue from lump-sum taxes (net of transfers). The government issues bonds to finance the difference between government revenue and expenditure. Lump-sum taxes are adjusted both in response to deviations of the government debt/GDP ratio from a target level and to the change in that ratio.

- Foreign Sector: Local currency pricing is assumed. Intermediate goods producers price their product separately in the home and foreign market leading to an incomplete exchange rate pass-through. Erceg et al. (2008) point out, that empirically imports and exports in the U.S. are heavily concentrated, with about 75 percent in capital goods and consumer durables, but the production share of capital goods and consumer durables is very low. To account for this fact in the two-country model they allow the import share in the final good aggregator for investment goods to be higher than the import share in the final good aggregator for consumption goods.

- Shocks: Since we have no information about the variances of the shock terms, we set all shock variances equal to zero. The government spending shock of the home country represents the common fiscal policy shock. The common monetary policy shock is added for the home country.

- Calibration/Estimation: The model is calibrated at a quarterly frequency. Parameters of the original monetary policy rule are estimated using U.S. data from 1983:1–2003:4.

- Replication: We replicated the impulse response functions for real exports, real imports and the exchange rate to a foreign investment demand sock represented by a decline in the foreign capital income tax rate as plotted in Figure 3 (disaggregated trade case) of Erceg et al. (2008).

Chapter 2

The Diversity of Forecasts from Macroeconomic Models of the U.S. Economy

(with Volker Wieland)

Abstract This chapter investigates the accuracy and heterogeneity of output growth and inflation forecasts during the current and the four preceding NBER-dated U.S. recessions. We generate forecasts from six different models of the U.S. economy and compare them to professional forecasts from the Federal Reserve's Greenbook and the Survey of Professional Forecasters (SPF). The model parameters and model forecasts are derived from historical data vintages so as to ensure comparability to historical forecasts by professionals. The mean model forecast comes surprisingly close to the mean SPF and Greenbook forecasts in terms of accuracy even though the models only make use of a small number of data series. Model forecasts compare particularly well to professional forecasts at a horizon of three to four quarters and during recoveries. The extent of forecast heterogeneity is similar for model and professional forecasts but varies substantially over time. Thus, forecast heterogeneity constitutes a potentially important source of economic fluctuations. While the particular reasons for diversity in professional forecasts are not observable, the diversity in model forecasts can be traced to different modeling assumptions, information sets and parameter estimates.

Keywords: forecasting, business cycles, heterogenous beliefs, forecast distribution, model uncertainty, Bayesian estimation

JEL-Codes: C53, D84, E31, E32, E37

2.1 Introduction

Recent empirical studies have documented substantial variations in the accuracy and heterogeneity of expert forecasts[1] of GDP and inflation (see Kurz, Jin and Motolese (2003, 2005), Giordani and Söderlind (2003), Kurz (2009) and Capistran and Timmermann (2009)). At the same time, 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, in Kurz (1994a, 1994b, 1996, 1997a,1997b, 2008), Brock and Hommes (1998), Kurz et al. (2005), Chiarella et al. (2007), Branch and McGough (2010), Branch and Evans (2010) and de Grauwe (2010).

Forecast heterogeneity arises for several reasons. First of all, forecasters need a forecast-generating framework. Such a framework may be a fully developed economic structure, a non-structural collection of statistical relationships or a simple rule-of-thumb. The particular modeling assumptions embedded in this forecasting framework represent an important source of belief heterogeneity. Another source of heterogeneity is the information used by the forecaster. Information sets may differ in terms of the number of economic aggregates or prices for which the forecasters collect data and the timeliness of the data vintage. The data is needed to estimate the state of the economy and the parameters of the forecasting framework.

While expert forecasts are published in various surveys, the underlying modeling assumptions, information sets and parameter estimates are not publicly available. Instead, this chapter uses six different macroeconomic models of the U.S. economy to generate output and inflation forecasts and investigate the impact of modeling assumptions, information sets and parameter estimates on forecast precision and heterogeneity.[2] The precision and diversity of expert forecasts from the Survey of Professional Forecasters (SPF) and the Federal Reserve's Greenbook are used as benchmark for comparison.[3] This comparison is conducted for successive quarter-by-quarter forecasts up to four quarters into the future during the five most recent recessions of the U.S. economy as dated by the NBER. We focus on periods around recessions because downturns and recoveries pose the greatest challenge for economic forecasters, and arguably expectational heterogeneity may itself play a role in these shifts in economic activity.

Among the six macroeconomic models considered in this chapter are 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 dynamic stochastic general equilibrium (DSGE) models of

[1] Expert forecasts are available via surveys such as Bluechip Economic Indicators by Aspen Publishers or the Survey of Professional Forecasters by the Federal Reserve Bank at Philadelphia.

[2] We draw on a recent research initiative that aims to build a database of macroeconomic models and offers a new comparative approach to model building and the search for macroeconomic policies that are robust under model uncertainty (see Taylor and Wieland (2009) and Wieland et al. (2009)).

[3] The SPF is conducted quarterly and contains responses by 30 to 50 professional forecasters. It was initiated in 1968 by the American Statistical Association and the NBER and is administered by the FRB 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 and becomes publicly available with a five-year lag.

the type currently used by leading central banks. The four small models are estimated to fit three macroeconomic time series: real GDP growth, inflation measured by the GDP deflator and the federal funds rate. The two medium-scale models are estimated with data for 7 and 11 variables, respectively. These variables include consumption, investment, wages and hours worked. The largest model even accounts for the breakdown in durables versus non-durables and services consumption, residential versus business investment, and the related deflators. We consider each of the six macroeconomic models as a reasonable forecast-generator. Such models are used at central banks and similar models may also be used by professionals in the private sector. Although the five structural models all embody the popular modeling assumption of homogenous rational expectations, they can be used together to generate an estimate of forecast heterogeneity due to differences in other modeling assumptions, information sets and parameter estimates. The properties of these models are discussed in more detail in the next section.

To render model-based forecasts comparable to historical SPF and Greenbook forecasts, we have to put them on a similar footing in terms of the data vintage used for parameter estimation and initial conditions. Thus, we 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. Every quarter we 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 we 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, we assess forecast precision relative to the revised data that became available during the subsequent quarters for the dates to which the forecasts apply. This assessment is obtained for periods surrounding recessions of the U.S. economy in 2008/09, 2001, 1990/91, 1981/82 and 1980. Forecasts are generated starting 4 quarters prior to the trough determined by the NBER Business Cycle Dating Committee up to 4 quarters after the trough.[4]

The approach taken in this chapter breaks new ground in several respects. First, to our knowledge there exists no comparable assessment of the forecasting accuracy of multiple structural macroeconomic models with historical data vintages. Real-time forecasts of non-structural time series models have been compared recently by Faust and Wright (2009) and in earlier work by Bernanke and Boivin (2003). Edge et al. (2010) have provided an assessment of the real-time forecasting performance of a single structural model. Furthermore, this chapter is the first attempt to quantify the heterogeneity of model forecasts and compare them to survey forecasts in order to learn more about the extent, dynamics and sources of forecast heterogeneity.

We obtain a number of interesting findings with regard to the relative accuracy of model-based and professional forecasts as well as the extent and dynamics of forecast diversity. The mean model forecast comes surprisingly close to the mean SPF and Greenbook forecasts in terms of accuracy even though the models only make use of a small number of data series. Model forecasts compare particularly well to professional forecasts at a horizon of three to four quarters and during recoveries.

[4]Exceptions are the 1980 and 2008/9 recessions. In the first case, we start only 2 quarters prior to the trough because some data is not available for earlier vintages. In the second case, the trough is not yet determined. We start in 2007Q4 and end in 2009Q3.

The extent of forecast heterogeneity is similar for model and professional forecasts but varies substantially over time. This variation itself may constitute a potentially important source of economic fluctuations. While the particular reasons for diversity in professional forecasts are not observable, the diversity in model forecasts can be traced to different modeling assumptions, information sets and parameter estimates. Of course, the models used by professional forecasters may differ from our models. Furthermore, New Keynesian DSGE models have only been developed in the last decade and would not have been available to forecasters in earlier recessions. However, non-structural VAR models such as the Bayesian VAR were already in use in the 1980s and the model of Fuhrer (1997) is a good example of the type of structural models with rational expectations that have been used since the early 1990s. Even if most private sector forecasters still favor traditional structural models over the New Keynesian DSGE models with microeconomic foundations preferred by academia and central banks, the two types of models exhibit some similar reduced-form relationships such as price and wage-inflation Phillips curves and aggregate demand equations with a mixture of backward- and forward-looking components. Thus, our findings can be taken as an indication that much of the observed time variation in forecast heterogeneity may be explained by disagreement about appropriate modeling assumptions and differences in parameter estimates rather than irrationality of particular forecasters.

The remainder of this chapter proceeds as follows. Section 2.2 summarizes the most important features of the different macroeconomic models that we use to compute forecasts. Section 2.3 describes the estimation and forecasting methodology. Section 2.4 provides an illustrative example by forecasting the 2001 recession. The difference between model-based and professional nowcasts and their impact on forecasting performance in the current recession are demonstrated in section 2.5. Section 2.6 provides a comparison of forecast accuracy of model and professional forecasts. The extent and dynamics of forecast heterogeneity is studied systematically in section 2.7. Section 2.8 summarizes our findings and concludes.

2.2 Forecasting models

In total, we consider six different models of the U.S. economy. One of the models is a simple vector autoregression model (VAR) that incorporates no behavioral interpretations of parameters or equations. The other five models are structural representations of the U.S. economy. Table 2.1 summarizes the most important model features, while Appendix A1 provides a complete description of the model equations.

The VAR model is estimated with four lags of output growth, inflation and the federal funds rate. It is well-known that unrestricted VARs are heavily over-parameterized and therefore not very useful for forecasting purposes. As proposed by Doan et al. (1984) we use a Bayesian approach with so-called Minnesota prior to shrink the parameters towards zero and render the VAR model more effective in forecasting. It is referred to as the *BVAR-WW model* in the following. The extension *WW* is meant to

Table 2.1: Model overview

Name/Reference	Short Name	Type	Observable Variables	Original Authors' Sample
Bayesian VAR estimated in this chapter	BVAR-WW	Bayesian VAR with 4 lags and Minnesota priors	3: output growth, inflation, interest rate	
Fuhrer (1997)	NK-Fu	small-scale closed economy New Keynesian model with relative real wage contracts and backward looking IS curve	3: output growth, inflation, interest rate	1966Q1-1994Q1
Del Negro and Schorfheide (2004)	NK-DS	standard 3-equation New Keynesian model with Calvo contracts and forward looking IS-equation	3: output growth, inflation, interest rate	1955Q3-2001Q3
New Keynesian Model estimated in this chapter	NK-WW	standard 3-equation New Keynesian model with mark-up and preference shocks	3: output growth, inflation, interest rate	
Christiano et al. (2005) as estimated in Smets and Wouters (2007)	CEE-SW	medium-scale closed economy DSGE-model of the type used by policy institutions	7: output growth, consumption growth, investment growth, inflation, wages, hours, interest rate	1966Q1-2004Q4
Edge et al. (2008)	FRB-EDO	medium-scale closed economy DSGE-model developed at he Federal Reserve. Two sectors with different technology growth rates	11: output growth, inflation, interest rate, consumption of non-durables and services, consumption of durables, residential investment, business investment, hours, wages, inflation for consumer nondurables and services, inflation for consumer durables	1984Q1-2004Q4

indicate that we have estimated this model without reference to an earlier parameterization by other authors. Nevertheless, such models have been used in forecasting by many practitioners at least since the early 1980s, that is throughout all the recessions studied in our forecast comparison.

The structural models we have chosen reflect the developments in macroeconomic modeling in the last two decades. The model of Fuhrer (1997) is a good example of the New Keynesian models that were developed in the 1980s and early 1990s.[5] While academics still focused primarily on developing the microeconomic foundations of real business cycle theory, these models became quite popular among central bank researchers and practitioners. They took into account adaptive and forward-looking behavior of market participants, real effects of monetary policy and output and inflation persistence. Fuhrer (1997) used maximum likelihood estimation to parameterize the model and we follow the same approach in re-estimating this model in the present chapter. It is referred to as the *NK-Fu model* in our analysis.

The New Keynesian model laid out by Rotemberg and Woodford (1997) and Goodfriend and King (1997) and developed in detail in Woodford (2003) and Walsh (2003) accounts more systematically

[5]These models combined rational expectations and nominal rigidities as in the seminal paper of Taylor (1979). For other examples see the model comparison projects of Bryant et al. (1988), Bryant et al. (1989), Klein (1991), and Bryant et al. (1993). Taylor (1993) already presented an estimated multi-country model of the G-7 economies of this type.

for microeconomic foundations in terms of the optimizing and forward-looking behavior of representative households and firms. Such a framework is particularly useful for quantifying likely market responses to changes in macroeconomic policies as emphasized in the famous Lucas critique. The New Keynesian model also incorporates restrictions in terms of monopolistic competition and price rigidity that imply important interactions between nominal and real economic variables. It has quickly become the principal workhorse model of monetary economics in the last decade.[6] Key model variables are output, inflation and interest rates just as in the *BVAR-WW* and *NK-Fu* models, but the microeconomic foundations imply additional restrictions on the reduced-form VAR representation of this model. We consider two empirical implementations. The first specification is taken from Del Negro and Schorfheide (2004). They use a Bayesian estimation methodology to fit the model to output, inflation and interest rate data. In the following, it is referred to as the *NK-DS model*. The second specification differs in terms of the modeling assumptions regarding the particular economic shocks that are the source of fluctuations. It is also estimated with Bayesian methods and termed the *NK-WW model*.

Christiano, Eichenbaum and Evans (2005) extended the New Keynesian DSGE modeling approach and showed how to build medium-scale models that can fit a significant number of important empirical regularities of the U.S. economy. To this end, they introduced additional dimensions for optimizing behavior as well as additional economic frictions. Such medium-scale models include physical capital in the production function and account for endogenous capital formation. Labor supply is modeled explicitly. Nominal frictions include sticky prices and wages and inflation and wage indexation. Real frictions include consumption habit formation, investment adjustment costs and variable capital utilization. Smets and Wouters (2003, 2007) extended and estimated the model of Christiano, Eichenbaum and Evans with Bayesian methods to fit key macroeconomic series. We generate forecasts from a version of this model estimated with Bayesian methods and refer to it as the *CEE-SW model* in the following. DSGE modeling has rapidly gained in popularity and many central banks have estimated larger and more sophisticated DSGE models for their respective countries. The fifth structural model in our forecasting pool is a version of the new DSGE model developed at the Federal Reserve by Edge et al. (2008). Following these authors we refer to it as the *FRB-EDO model*.

The two medium-scale models are fit to 7 and 11 economic time series, respectively. The CEE-SW model is estimated with data on real GDP growth, inflation as measured by the GDP deflator, the federal funds rate, wages, hours worked, consumption and investment. The FRB-EDO model allows for further disaggregation. It features two production sectors, which differ in their pace of technological progress. This structure can capture the different growth rates and relative prices observed in the data. Accordingly, the expenditure side is disaggregated as well. It is divided into business investment and three categories of household expenditure: consumption of non-durables and services, investment in durable goods and residential investment. The data used in estimation covers output growth, inflation, the federal funds rate, consumption of non-durables and services, consumption of durables, residential

[6]For recent discussions of the application of the New Keynesian approach in practical monetary policy see Wieland (2009).

investment, business investment, hours, wages, inflation for consumer nondurables and services and inflation for consumer durables.

2.3 Forecasting methodology

This section demonstrates how the forecasts are computed. Three aspects are best distinguished and discussed separately: model specification and solution, parameter estimation, and the sequence of steps necessary to generate quarter-by-quarter forecasts.

Model specification and solution. The simple New Keynesian model estimated by Del Negro and Schorfheide (2004) serves as a good example. It is a log-linearized approximation of the original nonlinear model consisting of three equations: a New Keynesian IS equation that is derived from the household's intertemporal first-order condition, a New Keynesian Phillips curve that is implied by the price-setting problem of the firm under monopolistic competition and price rigidity, and the central bank's interest rate rule:

$$x_t = E_t x_{t+1} - \tau^{-1}(R_t - E_t \pi_{t+1}) + (1-\rho_g) g_t + \rho_z \tau^{-1} z_t \tag{2.1}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (x_t - g_t) \tag{2.2}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2 x_t) + \epsilon_{R,t} \tag{2.3}$$

The notation of equations, variables and parameters is the same as in Del Negro and Schorfheide (2004). Variables are defined as percentage deviations from their steady state level. x_t denotes output, π_t inflation and R_t the federal funds rate. g_t is a government spending shock and z_t a technology shock. Both shocks follow an AR(1) process (not shown). The monetary policy shock $\epsilon_{R,t}$ is iid-normally distributed. $(\beta, \tau, \gamma, r^*, \pi^*, \kappa, \rho_R, \psi_1, \psi_2)$ represent model parameters that need to be estimated. The vector of parameters also includes the AR parameters (ρ_g, ρ_z) governing the dynamics of economic shocks and the standard deviations of the associated innovations, $(\sigma_R, \sigma_g, \sigma_z)$.

The model is connected with the available data by adding measurement equations that link the model variables to observable quarterly output growth, quarterly inflation, and the quarterly federal funds rate:

$$YGR_t = \ln \gamma + \Delta x_t + z_t \tag{2.4}$$

$$INFL_t = \ln \pi^* + \pi_t \tag{2.5}$$

$$INT_t = \ln r^* + \ln \pi^* + R_t. \tag{2.6}$$

YGR_t denotes the first difference of the log of GDP, $INFL_t$ the first difference of the log GDP deflator, and INT_t the quarterly federal funds rate. The system of linear expectational difference equations that comprises model and measurement equations is then solved using a conventional solution method such as the technique of Blanchard and Kahn and the state space representation of the system is derived:

$$y_t^{obs} = y(\theta) + \lambda + y_t^s, \tag{2.7}$$

$$y_t = g_y(\theta)y_{t-1} + g_u(\theta)u_t, \tag{2.8}$$

$$E(u_t u_t') = Q(\theta), \tag{2.9}$$

Here, the first equation summarizes the measurement equations, the second equation constitutes the transition equation and the third equation denotes the variance-covariance matrix Q. θ refers to the vector of structural parameters. These include the shock variances, so that Q also depends on elements of θ. A state space representation of this form is derived for each forecasting model and the notation in equations (7), (8) and (9) is general enough to apply to all the structural models considered. As an example, Table 2.2 shows how to link the variables and parameters in the state space representation to those in the Del Negro & Schorfheide model.

Table 2.2: State space representation and model equations

structural parameters	$\theta = (\beta, \tau, \rho_g, \rho_z, \gamma, r^*, \pi^*, \kappa, \rho_R, \psi_1, \psi_2, \sigma_R, \sigma_g, \sigma_z)$
observable variables	$y_t^{obs} = [\ YGR_t \quad INFL_t \quad INT_t\ ]'$
steady state	$y(\theta) = [\ 0 \quad \ln \pi^* \quad \ln r^* + \ln \pi^*\ ]'$
deterministic trend	$\lambda = [\ \ln \gamma \quad 0 \quad 0\ ]'$
subset of endogenous variables	$y_t^s = [\ \Delta x_t + z_t \quad \pi_t \quad R_t\ ]'$
endogenous variables	$y_t = [\ x_t \quad R_t \quad \pi_t \quad g_t \quad z_t\ ]'$
shocks	$u_t = [\ \epsilon_{R,t} \quad \epsilon_{z,t} \quad \epsilon_{g,t}\ ]'$

The observable variables y_t^{obs} that are defined by the measurement equations are functions of the stationary steady state $y(\theta)$, of a subset of the endogenous variables expressed in deviations from steady state, y_t^s, and of the deterministic trend λ. The transition equation comprises the decision rules. Its parameters are given by the two solution matrices g_y and g_u which are nonlinear functions of the structural parameters θ. Thus, the transition equations relate the endogenous variables y_t to lags of themselves and the vector of exogenous shocks u_t. Since, the measurement equations include the deterministic growth path that is driven by labor-augmenting technological progress no separate de-trending of the data is necessary.

Model Estimation. Whenever possible, we estimate the models using the same techniques as the original authors. The model by Fuhrer (1997) is estimated using maximum likelihood techniques while the NK-DS, CEE-SW and FRB-EDO models are estimated using a Bayesian methodology. We also use Bayesian methods to estimate the NK-WW and BVAR-WW models. Maximum likelihood estimation maximizes the likelihood of the model, while Bayesian estimation combines the likelihood with prior beliefs obtained from economic theory, microeconomic data or previous macro studies. An extensive survey of the methodology is presented in An and Schorfheide (2007).

Because the reduced-form coefficients of the state-space representations are nonlinear functions of the structural parameters, θ, the calculation of the likelihood is not straightforward. The Kalman filter is applied to the state space representation to set up the likelihood function (see e.g. Hamilton, 1994, chapter 13.4).[7] Since the models considered here are stationary we can initialize the Kalman Filter using the unconditional distribution of the state variables. Combining the likelihood with the priors yields the log posterior kernel $ln\mathcal{L}(\theta|Y^T) + lnp(\theta)$ that is maximized over θ using numerical methods so as to obtain the posterior mode. We use the posterior mode to generate point forecasts. As a robustness check we compared point forecasts obtained from the posterior mean and posterior mode in several cases. To this end, we simulated the posterior distribution using the Metropolis-Hastings-Algorithm. Since the two alternative point forecasts were quite similar we relied on the posterior mode for forecast generation in the remainder of our analysis so as to keep the computational burden resulting from the large number of model re-estimations manageable.

In estimating the Bayesian VAR we follow Doan et al. (1984) and use the so-called Minnesota prior to avoid over-parameterization. This prior implies shrinking the parameters towards zero by assuming that the price level, real output and the interest rate follow independent random walks. All parameters are assumed to be normally distributed with mean zero. The prior variance of the parameters decreases with the lag length.

Forecasting. For a given date, we estimate each of the models on the basis of the most recent data vintage that would have been available at that time. Thus, data vintages are identical across models and change quarter-by-quarter as in real time. The information sets differ across models only if the models use different variables. Forecasts may also differ due to different estimation methods and different modeling assumptions. While the information set for the three small models and the Bayesian VAR is comprised of three time series, the information set of the CEE-SW model contains seven time series and the information set of our variant of the FRB-EDO model contains eleven time series. The particular time series and the sources for the real-time data set are described in Appendix A2.

We re-estimate the models quarter-by-quarter with every arrival of a new data vintage. Thus, the newly estimated model specification uses parameter estimates $\hat{\theta}_t$ that are based on the information set I_t which contains the most recent data vintage available in quarter t. Of course, data on real GDP, the components of GDP and the associated deflators become available with a time lag and is not part of the current quarter t information set. Current quarter estimates of economic growth and inflation are obtained using $t-1$ observations of those variables. The current quarter estimate is typically referred

[7] We consider only unique stable solutions. If the Blanchard-Kahn conditions are violated we set the likelihood equal to zero.

to as the *nowcast*, that is the "forecast" at a horizon of zero quarters. The model forecasts for horizons $h \in (0, 1, 2, 3, 4)$ are computed under the assumption that expected future shocks are equal to zero, $E[u_{t+h}|I_t] = 0$. They are generated by iterating over the following equation:

$$E[y_{t+h}^{obs}|I_t] = y(\hat{\theta}_t) + \hat{\lambda}_t + g_y(\hat{\theta}_t)^{h+1} y_{t-1}. \tag{2.10}$$

A hat on the structural parameters θ and the subscript t denote that they are estimated on the basis of the information set at time t, I_t, which contains the most recent releases of economic aggregates through quarter $t-1$. Recall also that the reduced form solution matrices g_y are functions of these estimates and change over time as new data vintages become available.
It is instructive to summarize the different steps needed to generate diverse model forecasts:

1. Model setup: create a model file with the model equations and add measurement equations that link the model to observable time series.

2. Solution: solve the model and write it in state space form.

3. Data update: update the data with the current data vintage.

4. Prior: add a prior distribution of the model parameters if necessary.

5. Estimation: estimate the structural parameters by maximizing the likelihood or the posterior kernel.

6. Forecast: compute forecasts by iterating over the solution matrices setting the expected value of future shocks to zero.

7. Repeat steps 3 to 6 quarter-by-quarter for the time-period of interest.

8. Repeat steps 1 to 7 for different models while extending the information set with additional variables as required by the respective model.

2.4 An illustration: forecasting the 2001 recession

Next, we illustrate the real-time forecasting process with an example focusing on the 2001 recession in the United States. According to the NBER Business Cycle Dating Committee a peak in economic activity in March 2001 was followed by a trough in November 2001.
Figure 2.1 shows real output growth forecasts that were obtained on the basis of data available in the first quarter of 2001. The vertical line serves to indicate the current quarter. The nowcasts in 2001:Q1, of course, differ from the actual 2001:Q1 data that is released subsequently. The solid line in Figure 2.1 reports the actual data on annualized quarter on quarter output growth. This time series consists of the data vintage 2001:Q1 until the starting point of the nowcast/forecast in the fourth quarter of 2000 and revised data from 2001:Q1 onwards. The revised GDP data is drawn from later data vintages.

GDP data is first released about one month after the end of the quarter to which the data refers, the so-called advance release. These data are then revised several times at the occasion of the preliminary release, final release, annual revisions and benchmark revisions. We follow Faust and Wright (2009) and use the data point in the vintage that was released two quarters after the quarter to which the data refer to as revised data. For example, revised data for 2001:Q1 is obtained by selecting the entry for 2001:Q1 from the data vintage released in 2001:Q3. Revised data for 2001:Q2 is obtained by using the entry for 2001:Q2 from the data vintage released in 2001:Q4, and so on. Hence, we do not attempt to forecast annual and benchmark revisions, because the models cannot predict changes in data definitions. The revised data against which we judge the accuracy of forecasts will typically correspond to the final NIPA release.

Three different forecasts are reported in Figure 2.1. The model-based forecast depicted by the dashed-dotted line is derived from the CEE-SW model. It is compared to the Fed's Greenbook forecast (dashed line) and the mean SPF forecast (dotted line). The SPF is a quarterly survey of professional macroeconomic forecasters conducted by the Federal Reserve Bank of Philadelphia. Typically, 30 to 50 respondents report projections of several key macroeconomic variables.[8] Since these experts

[8]Other surveys include Bluechip Economic Indicators, the Michigan Survey of Consumer Attitudes and Behavior and the Livingston Survey. Livingston and Bluechip are surveys of professionals like the SPF. Bluechip is not available free of charge. The Livingston survey is only conducted semi-annually. The Michigan survey reports assessments of 1000 to 3000 households. Mankiw et al. (2004) compare inflation expectations from these different surveys: median inflation expectations are relatively accurate and similar for the different surveys. Histograms show substantial disagreement; especially among consumers. There are extreme outliers that show up in long tails of the forecast distribution. Disagreement varies dramatically over time but similarly for consumers and professionals. Mishkin (2004) is sceptical of household surveys and notes that households have no incentive to compute detailed forecasts to answer survey questions about their expectations. Given the long tail in forecast distributions, he questions whether respondents with extreme expectations behave in a way consistent with these expectations. Professional forecasters, who make their living in this business, will put serious effort in computing a good forecast.

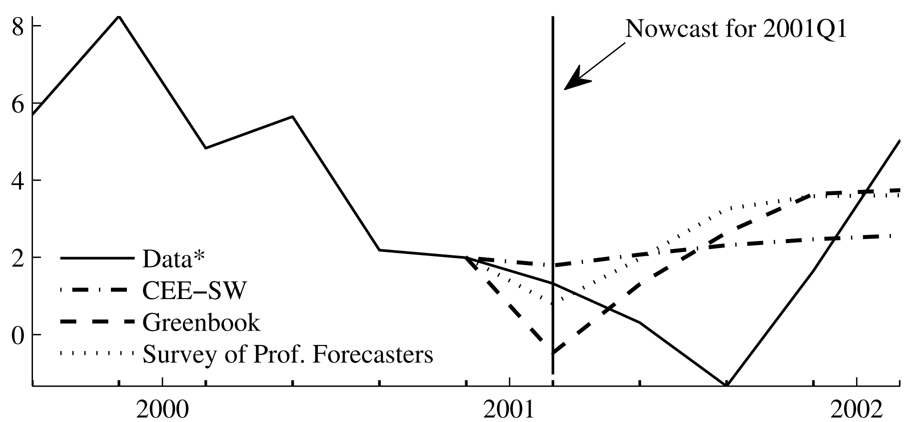


Figure 2.1: Real GDP growth forecast at the start of the 2001 recession

Notes: NBER defined peak: 2001Q1, NBER defined trough: 2001Q4.

*) The solid line shows data vintage 2001Q1 until 2000Q4 and revised data afterwards.

tend to earn their living in the forecasting business and may be expected to put serious effort in the production of the forecast, we consider it a reasonable benchmark for comparison with our model forecasts. Of course, it is well known that there exist incentives not to report the best possible forecast in such a survey.[9] For this reason, we also consider the Greenbook forecast prepared by the staff of the Board of Governors of the Federal Reserve System for the Federal Open Market Committee.[10]

All three forecasts imply a reduction in output growth in 2001:Q1, the current quarter, followed by a re-bound in subsequent quarters. The CEE-SW model only projects a slight decline in the growth rate compared to the larger declines implied by mean SPF forecast and the Greenbook. However, in this particular quarter the Greenbook nowcast of negative growth is far too pessimistic and the least accurate among the three nowcasts. As to the subseqent quarters, all three forecasts turned out to be mistaken in predicting an immediate re-bound starting in 2001:Q2. The economy deteriorated in the second and third quarter of 2001. The lowest quarterly output growth rate was reached in 2001:Q3, after which the economy recovered.

Successive forecasts throughout the course of the 2001 recession are shown in Figure 2.2. The left-hand-side column of panels in Figure 2.2 compares the real-time forecasts generated with the CEE-SW model (solid line with square markers) to the Greenbook (dashed line) and SPF (dotted line) forecasts and the actual data (solid line). The top-left panel replicates Figure 2.1 with the 2001:Q1 forecasts. Moving down the columns the data vintages and forecasts are shifted forward quarter-by-quarter. The second left-hand-side panel indicates that the Greenbook and SPF nowcasts in 2001:Q2 were much closer to the actual decline in GDP growth than the CEE-SW model's nowcast. In 2001:Q3 the CEE-SW nowcast and forecasts for subsequent quarters are very similar to the Greenbook and SPF forecast. In 2001:Q4 the CEE-SW nowcast and forecasts clearly dominate the two expert forecasts in terms of accuracy. At that point the Greenbook and mean SPF forecast implied a deepening of the recession. The revised data shows that instead a recovery took place as predicted by the model forecast. In 2002:Q1 the model nowcast is again more accurate. Also, the forecast for the third quarter is right on target although at the expense of overshooting in the next two quarters.

The panels in the right-hand-side column of Figure 2.2 provide a comparison of the quarter-by-quarter forecasts generated from the six different macroeconomic models. The CEE-SW forecast is shown together with the forecasts from the NK-DS, NK-WW, NK-Fu, BVAR-WW and FRB-EDO models. The solid line again indicates the data that is used as benchmark for assessing the accuracy of the model forecasts. The model forecasts generally fail to forecast the downturn in the U.S. economy from the first to the third quarter of 2001. However, the mean SPF and Greenbook forecasts also largely miss the downturn. The model forecasts, however, perform relatively well with regard to the recovery, once the trough in 2001:Q3 has been reached. Model forecasts are quite heterogeneous with the extent of heterogeneity varying over time. Forecast differences narrow in 2001:Q2 and 2001:Q3

[9] Forecasters have incentives to publish a forecast close to the consensus (Scharfstein and Stein, 1990; Lamont, 2002) as well as to publish a distinct forecast (Laster et al., 1999).

[10] Greenbook projections are prepared by the Federal Reserve's staff before each FOMC meeting and have been found to dominate forecasts from other professional forecasters in terms of forecasting accuracy (Romer and Romer, 2000; Sims, 2002; Bernanke and Boivin, 2003). They are made public with a five-year lag.

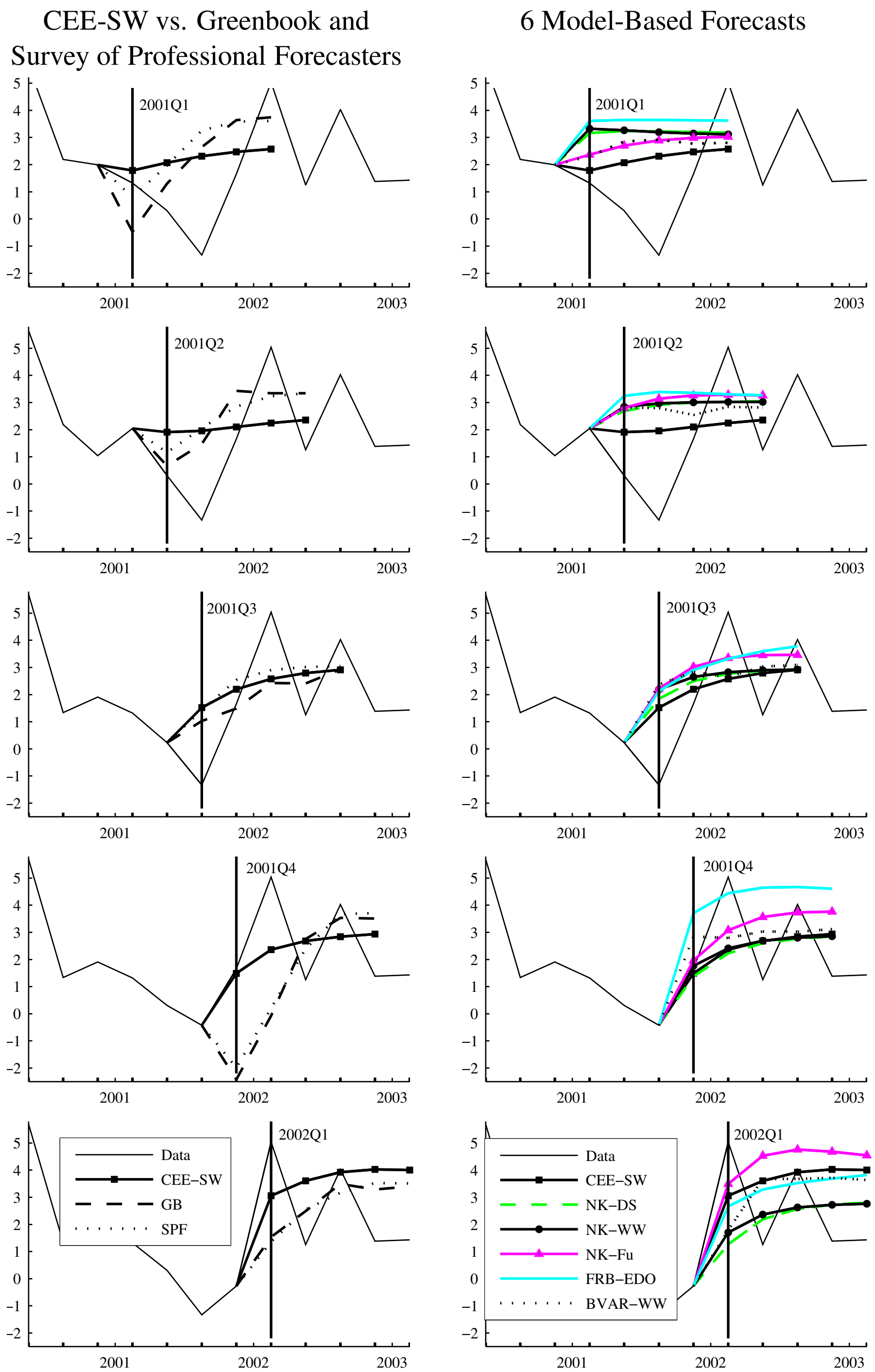


Figure 2.2: Real GDP growth forecast for the 2001 recession

Notes: NBER defined peak: 2001Q1, NBER defined trough: 2001Q4.

and widen again in 2001:Q4 and 2002:Q1.

2.5 Model-based versus expert nowcasts and the 2008/09 recession

The model-based forecasts shown in Figures 2.1 and 2.2 only use quarterly data vintages where the most recent data entries concern the quarter preceding the quarter in which the forecast is 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. This data can be used to improve nowcasts and the Federal Reserve staff and many professional forecasters certainly make use of it. The use of higher-frequency data may well be the main reason for better nowcasts by the Greenbook and Survey of Professional Forecasters compared to our six models.

In principle, there exist methods for using higher frequency data in combination with quarterly structural macroeconomic models. For example, Giannone et al. (2009) show how to incorporate such conjunctural analysis systematically in structural models. Employing such methods, however, is beyond the scope of this chapter. Instead, we approximate the use of higher-frequency information in quarterly model nowcasts simply by using Greenbook and mean SPF nowcasts to initialize model forecasts for future quarters.

The difference between using model versus expert nowcasts as initial conditions for model-based forecasts is illustrated in Figure 2.3. The top panel in Figure 2.3 partly replicates the second right-hand-side panel in Figure 2.2. It shows the 2001:Q2 forecasts from the CEE-SW, FRB-EDO, NK-WW and BVAR-WW models in comparison to the Greenbook forecast (dashed line) and the revised data (solid line). As discussed previously, the Greenbook nowcasts in 2001:Q2 came much closer to capturing the beginning of the downturn than the model nowcasts. Clearly, by that time it had become apparent to the Federal Reserve staff that the economy was deteriorating perhaps because of evidence obtained from higher-frequency data. The models miss this early evidence of the downturn as they are only using quarterly data concerning 2001:Q1.

The lower panel of Figure 2.3 displays the effect of using the Greenbook nowcast as the basis for the model forecasts. As a consequence, the model forecasts differ much less from each other than in the upper panel. The one-quarter-ahead model forecasts are more optimistic than the Greenbook. The two quarter-ahead forecasts from the models, however, are somewhat below the Greenbook and a bit closer to the eventual realization of output growth.

Altogether, we investigate and compare successive forecasts throughout the five most recent recessions on the U.S. economy in this manner. Of course, at the current juncture it is of particular interest to investigate the accuracy and diversity of forecasts in the on-going recession. In 2008 and 2009 public criticism of economic forecasters for failing to predict the downturn that is now often referred to as "The Great Recession" has been very pronounced. Figure 2.4 provides a perspective on successive model forecasts relative to the mean SPF forecast (dash-dotted line) and the actual data (solid line) that

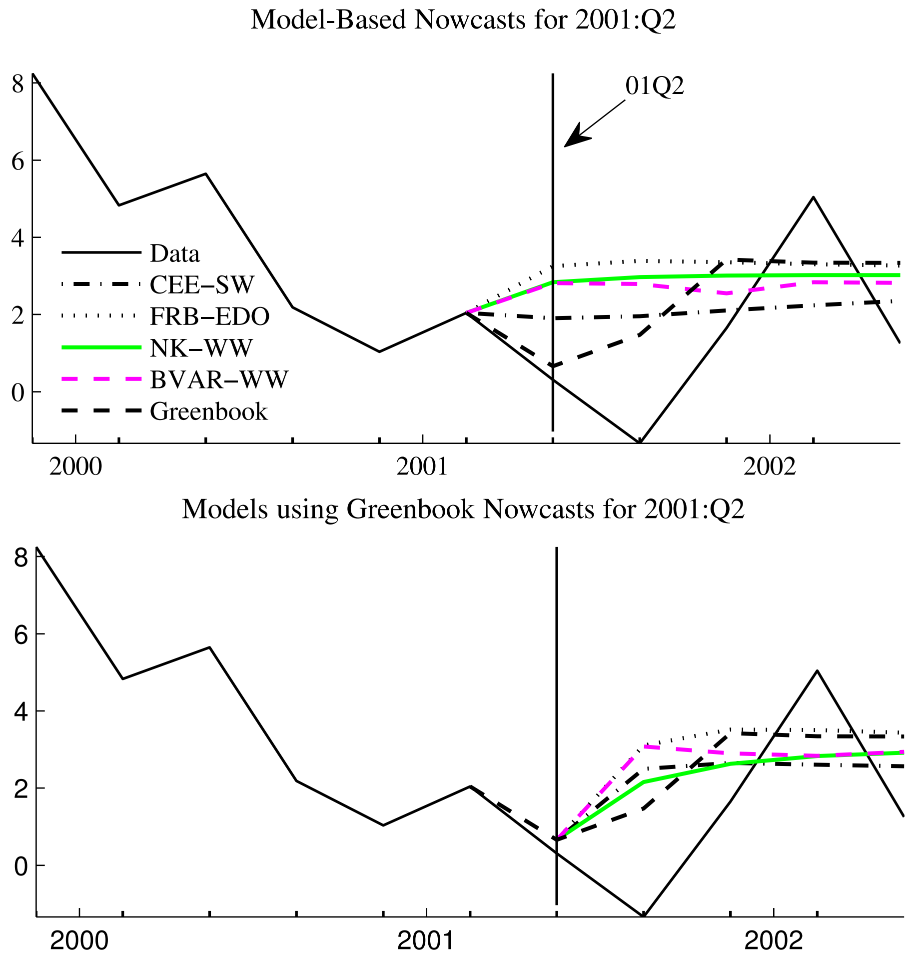


Figure 2.3: Real output growth forecast during the 2001 recession.

Notes: NBER defined peak: 2001:Q1, NBER defined trough: 2001:Q4. In the upper panel the model-generated nowcast based on the information set with information on $t-1$ aggregates is used. In the lower panel the Greenbook nowcast forms the starting point for model-based forecasts regarding future quarters.

has become available so far. The top row of panels shows forecasts made in the third quarter of 2008. Lower rows report subsequent forecasts quarter-by-quarter as new data vintages become available. In the panels of the left-hand-side column model-based nowcasts are generated from the most recent quarterly data vintage. In the right column, instead, mean SPF nowcasts are used to initialize the model forecasts.

As is apparent from the top left panel, professional forecasters, on average, failed to foresee the downturn as late as in the third quarter of 2008. The mean SPF forecast indicates a slowdown in the fourth quarter followed by a return to higher growth in the first quarter of 2009. Not surprisingly, this misdiagnosis has generated much public criticism. The model-based forecasts we generate based on the data vintage of 2008:Q3 would not have performed any better. In fact, they do not indicate any impending decline in economic activity. In the fourth quarter of 2008, however, the mean SPF nowcast

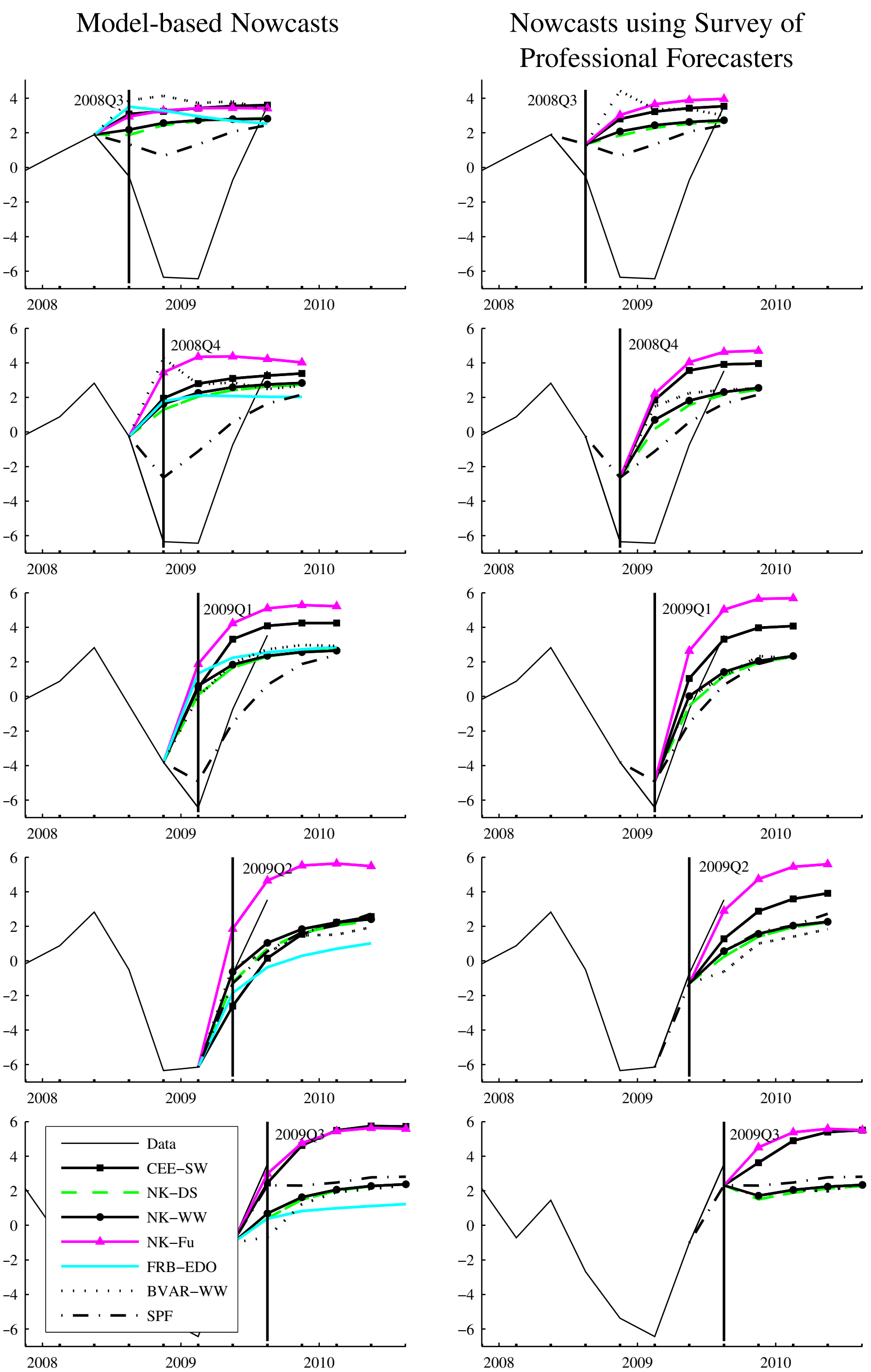


Figure 2.4: Real output growth forecast during the 2007-2009 recession.

Notes: NBER defined peak: 2007Q4. In the left-hand-side panels the model-generated nowcast based on the information set with information on $t-1$ aggregates is used. In the right-hand-side panels the mean SPF nowcast forms the starting point for model-based forecasts regarding future quarters.

and the model-based nowcast diverge dramatically. Following the Lehman debacle professional forecasters drastically revised their assessments downwards, and continued to do so in the first quarter of 2009.

Interestingly, from 2009:Q2 onwards the model-based forecasts perform quite well in predicting the recovery of the U.S. 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. An inspection of the right-hand-side panels suggests that initializing the model forecasts with the mean SPF nowcasts further strenghtens the models performance during the recovery phase. In this case, the 2009:Q1 forecast for the second and third quarter of 2009 that is implied by the CEE-SW, NK-WW and FRB-EDO models already looks surprisingly accurate relative to the data releases that have become available so far.

2.6 The relative accuracy of model-based and expert forecasts

For a systematic evaluation of forecast accuracy we 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. Our typical recession sample covers the period from 4 quarters prior to the trough determined by the NBER Business Cycle Dating Committee to 4 quarters after the trough.[11] The accuracy of the individual model forecasts is compared to the mean model forecast, that is the average of the six models, the mean SPF forecast and the Greenbook forecast. The RMSE for model m at forecasting horizon h given a recession sample that starts in period p and ends in period q is given by:

$$RMSE_m^h = \sqrt{\sum_{t=p}^{q} \left(E[y_{t+h}^{obs}|I_t^m] - y_{t+h}^{obs}\right)^2 / (q-p+1)}, \tag{2.11}$$

where I_t^m denotes the information set of a specific model m at time t. I_t^m includes the model equations and the data vintage for period t. y_{t+h}^{obs} denotes the data realizations h periods ahead.

Our findings are reported in Table 2.3. In most cases the model forecasts are on average less accurate than the Greenbook and mean SPF forecasts. Sometimes the best forecast is given by the Greenbook but at other times by the mean SPF forecast. 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 nowcasts and near-term forecasts. An exception is the 2001 recession during which the quality of all forecasts is very similar. Root mean squared errors are lower during the 1990-91 recession and the 2001 recession than during the other recessions.

Among the structural models there is none that consistently outperforms the others. During a specific recession, the best forecasts at different horizons may also come from different models. Nevertheless,

[11] Exceptions are the 1980 and 2008/9 recessions. In the first case, we start only 2 quarters prior to the trough because of data availability. In the second case, the trough is not yet determined. We start in 2007Q4 (peak) and end in 2009Q3.

Table 2.3: RMSEs of output growth forecasts

Sample / Horizon	NK-DS	NK-WW	CEE-SW	FRB-EDO	NK-Fu	BVAR-WW	Mean	GB	SPF
1980:1 - 1981:3									
0	7.19	7.12	6.42	5.64	6.88	6.46	5.13	5.05	–
1	7.28	7.20	5.59	5.95	6.78	7.63	5.59	6.65	–
2	5.56	5.67	5.24	5.77	7.43	8.69	5.70	5.54	–
3	5.50	5.67	4.33	4.92	5.62	6.28	4.56	6.11	–
4	5.43	5.57	4.45	4.39	5.56	7.33	4.84	5.32	–
1981:4 - 1983:4									
0	5.54	5.68	2.89	3.23	3.69	3.68	3.68	2.42	2.14
1	5.14	5.25	3.69	4.32	3.96	3.98	4.02	3.58	3.88
2	4.09	4.16	4.06	4.59	4.84	5.72	4.31	3.93	4.11
3	4.16	4.22	4.15	4.53	5.10	5.74	4.45	3.91	4.41
4	4.09	4.12	4.02	4.56	4.66	5.74	4.33	3.84	4.02
1990:1 - 1992:1									
0	2.82	3.01	3.22	1.80	2.92	1.76	2.50	1.27	1.12
1	3.15	3.22	3.94	2.06	3.79	2.24	2.98	2.09	1.45
2	3.08	3.13	4.00	2.15	3.84	2.38	2.99	2.34	2.06
3	3.13	3.14	3.90	2.38	3.81	2.56	3.03	2.31	2.54
4	2.79	2.78	3.56	2.30	3.73	2.32	2.80	2.18	2.37
2000:4 - 2002:4									
0	2.32	2.33	1.94	2.43	2.30	2.63	2.22	2.28	2.22
1	2.22	2.24	2.19	2.49	2.64	2.28	2.25	2.20	2.30
2	2.23	2.21	2.29	2.61	2.54	2.35	2.29	2.34	2.21
3	2.69	2.67	2.74	2.82	2.74	2.71	2.67	2.76	2.65
4	2.24	2.25	2.08	2.58	2.17	2.12	2.19	2.18	2.13
2007:4 - 2009:3									
0	3.58	3.75	3.78	4.05	4.37	4.42	3.91	–	1.94
1	4.36	4.43	4.81	4.72	5.18	4.95	4.69	–	3.30
2	4.78	4.83	4.89	4.85	5.36	5.05	4.94	–	4.11
3	5.20	5.21	5.35	5.13	5.66	5.29	5.29	–	4.80
4	5.56	5.55	5.85	5.29	5.91	5.61	5.62	–	5.39

a detailed comparison reveals some systematic differences. The CEE-SW model and the FRB/EDO model deliver fairly good forecasts in four out of five recessions. Several times, they yield the most accurate forecasts. In those cases where they are less precise than other models, the differences to the most accurate forecast are small. Both models have a rich economic structure and consider more observable data series than the other models. At the same time the parameterization is tight enough to yield accurate forecasts. The BVAR-WW model forecasts quite accurately in the 1990-91 and the 2001 recession, but more poorly in the other three recessions. Output growth in the 1990 and 2001 recession was less volatile. Perhaps, the lag structure of the Bayesian VAR is more appropriate during normal times and minor recessions. In more volatile times, sharp spikes in output fluctuations continue to feed through to forecasts for several quarters due to the lags included in the model. This results in less accurate forecasts.

The NK-DS and NK-WW models perform quite well during the most recent three recessions, but more poorly in the first two recessions. These models rely on three time series only. Persistence in output fluctuations arises primarily due to ad-hoc AR(1) shock processes. It is less pronounced than in the BVAR-WW model with four lags of endogenous variables. In these models a sharp spike in real GDP growth has a short but strong effect on the forecast. Finally, the NK-Fu model performs worse than the NK-DS and NK-WW models in most of the recessions. This model does not allow ad-hoc

Table 2.4: RMSEs of inflation forecasts

Sample / Horizon	NK-DS	NK-WW	CEE-SW	FRB-EDO	NK-Fu	BVAR-WW	Mean	GB	SPF
1980:1 - 1981:3									
0	1.77	1.76	2.05	2.64	2.04	2.67	1.90	1.67	1.52
1	1.92	1.90	2.52	3.55	2.76	2.18	2.19	1.25	1.81
2	1.59	1.38	2.05	2.57	2.20	1.75	1.45	1.66	1.92
3	2.89	2.32	2.36	3.34	2.96	3.88	2.53	1.77	2.23
4	3.07	2.29	2.51	3.79	2.83	3.97	2.58	2.21	2.56
1981:4 - 1983:4									
0	1.90	1.76	1.69	1.37	2.41	1.49	1.58	1.12	1.13
1	2.71	2.24	1.98	1.47	2.16	2.24	1.98	1.32	1.76
2	2.63	1.99	1.89	1.29	1.81	2.13	1.70	1.26	1.68
3	2.85	2.01	2.10	1.31	2.07	2.31	1.80	1.07	1.95
4	2.87	1.95	2.26	1.22	1.61	2.46	1.67	1.48	2.06
1990:1 - 1992:1									
0	1.21	1.16	1.07	1.21	1.80	1.05	1.15	0.73	1.09
1	1.76	1.64	1.29	1.20	2.03	1.16	1.43	0.84	0.98
2	1.69	1.76	1.35	1.33	1.15	1.07	1.25	0.95	1.01
3	1.30	1.76	1.53	0.91	0.81	0.95	1.01	1.06	1.19
4	1.69	1.87	1.71	1.39	1.65	1.37	1.40	1.02	1.19
2000:4 - 2002:4									
0	1.08	1.05	1.04	1.27	1.17	0.90	0.98	0.56	0.70
1	1.18	1.15	1.12	1.43	1.26	0.92	1.07	0.87	0.87
2	1.35	1.38	1.16	1.50	1.48	1.11	1.19	0.70	0.92
3	1.42	1.49	1.21	1.75	1.63	1.16	1.28	0.75	0.93
4	1.45	1.59	1.07	1.64	1.83	1.30	1.27	0.78	0.98
2007:4 - 2009:3									
0	2.06	1.96	1.69	2.19	1.61	1.58	1.69	–	1.11
1	1.53	1.51	1.14	1.83	1.52	1.21	1.23	–	1.03
2	1.56	1.54	1.23	1.95	1.61	1.31	1.31	–	1.10
3	1.86	1.82	1.36	1.77	1.99	1.60	1.61	–	1.24
4	1.60	1.74	1.38	1.64	1.78	1.48	1.40	–	1.40

persistence via AR(1) shock processes. Shocks are assumed i.i.d. and output and inflation persistence can only arise from lags of output and inflation in the IS-curve and the overlapping wage structure. These dynamics may not be be sufficient to yield precise output growth forecasts.

The mean model forecast shown in the seventh column of the table averages the six model forecasts. It performs very well. Most of the time it turns out to be fairly close to the best individual model forecast in terms of root mean squared error.

In addition, we have investigated the accuracy of inflation forecasts. Table 2.4 reports the associated root mean squared errors of nowcasts and forecasts for the five recession episodes. Again, the root-mean-squared errors at horizons from zero to four quarters into the future are recorded separately. The Federal Reserve's Greenbook forecast for inflation is almost always more accurate than the other forecasts including the mean forecast from the Survey of Professional Forecasters. Perhaps, the better performance of the Greenbook forecast reflects an informational advantage regarding the inflationary consequences of Federal Reserve policies and future policy intentions.

Interestingly, the quality of the mean model forecast of inflation is quite similar to the mean SPF forecast. As in the case of output growth it is difficult to draw general conclusions about how differences in models influence the forecasting results. The BVAR-WW yields very good forecasts for the three latest recessions, but performs worse for the two recessions in the 1980s. The reason might be that the

Table 2.5: 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

BVAR-WW has a high a number of lags relative to the other models which may be more useful during less volatile times than during the 1980s disinflation. The CEE-SW model delivers one of the best inflation forecasts in several recessions and never one of the worst forecasts. In contrast to our findings for output growth, the FRB-EDO medium-scale model does not always perform as well as CEE-SW in inflation forecasting. It delivers very good inflation forecasts in two of the five recessions, but is among the most inaccurate for the others. The NK-WW model performs better than the fairly similar NK-DS model, because the additional mark-up shocks appear to better capture inflation dynamics. Finally, the NK-Fu model yields less satisfactory inflation forecasts. Perhaps, the overlapping wage contracts help the model capture the output-inflation tradeoff apparent in the 1980s recession but may induce more rigidity than required to match inflation dynamics in more recent recessions. The mean model forecast of inflation comes quite close to the best individual model forecast most of the time.

As discussed in the preceding section, the quality of a forecast for the future very much depends on how accurate the assessment of the current state of the economy is that forms the starting point for the forecast. The model forecasts lack information on specific events that have happened in the current quarter such as the failure of Lehman in the fall of 2008 nor do they make use of higher-frequency data that becomes available during the quarter ahead of quarterly GDP releases. Expert forecasts may take into account both types of information. Therefore, we check if the superior forecast performance

Table 2.6: RMSEs of inflation 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	1.67	1.67	1.67	1.67	1.67	1.67	1.67	1.67	1.52
1	2.73	2.59	2.57	2.76	2.97	2.94	2.59	1.25	1.81
2	2.89	2.56	2.49	2.53	2.76	3.33	2.59	1.66	1.92
3	2.70	1.86	1.98	1.39	1.48	2.71	1.73	1.77	2.23
4	4.02	2.92	2.54	3.00	3.15	4.94	3.22	2.21	2.56
1981:4 - 1983:4									
0	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.13
1	2.31	2.06	1.97	1.72	2.15	1.71	1.86	1.32	1.76
2	2.53	2.05	2.04	1.58	2.46	1.61	1.92	1.26	1.68
3	2.53	1.91	2.02	1.16	2.32	1.67	1.79	1.07	1.95
4	2.78	2.01	2.25	1.41	2.36	1.66	1.87	1.48	2.06
1990:1 - 1992:1									
0	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	1.09
1	1.03	1.10	1.01	0.94	1.77	0.93	1.03	0.84	0.98
2	1.42	1.58	1.36	0.81	1.61	1.04	1.23	0.95	1.01
3	1.49	1.77	1.63	1.11	0.89	0.93	1.20	1.06	1.19
4	1.31	1.70	1.62	1.34	0.87	1.07	1.16	1.02	1.19
2000:4 - 2002:4									
0	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.70
1	0.92	0.95	0.90	0.97	1.13	0.76	0.85	0.87	0.87
2	1.33	1.38	1.18	1.24	1.61	1.04	1.23	0.70	0.92
3	1.29	1.41	1.18	1.48	1.68	1.02	1.25	0.75	0.93
4	1.53	1.65	1.17	1.68	2.02	1.35	1.45	0.78	0.98
2007:4 - 2009:3									
0	1.11	1.11	1.11	–	1.11	1.11	1.11	–	1.11
1	1.15	1.19	1.00	–	1.48	1.11	1.10	–	1.03
2	1.28	1.37	1.17	–	1.56	1.22	1.28	–	1.10
3	1.50	1.61	1.30	–	1.87	1.49	1.51	–	1.24
4	1.69	1.81	1.39	–	1.92	1.59	1.65	–	1.40

of the expert forecasts is due to the same informational advantage that induces better nowcasts. As in the preceding section, we simply use the Greenbook nowcast (and for the latest recession the mean SPF nowcast) as initial conditions for the model-based forecasts. On this basis, we re-estimate the models and compute forecasts for horizons of one to four quarters into the future. Tables 2.5 and 2.6 report the associated root mean squared errors of output growth and inflation forecasts for the different recession episodes.

The GDP growth forecasts improve for most models and horizons when the expert nowcast is added to the models' information sets. An exception is the recession of 1980, probably because the Greenbook nowcasts were not very good during this period. The mean model forecast now even 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 1981-82 and 1990-91 recessions, while the mean SPF forecast still appears to be the most accurate in the ongoing recession, for which no Greenbook data and forecasts are publicly available.

With regard to forecasts of inflation, the addition of the expert nowcast to the information set of the model does not improve model-based forecasts quite as much as in the case of GDP forecasts. Also, the Greenbook forecast performance tends to remain superior to the model forecasts. Thus, one might speculate that the Federal Reserve staffs advantage in forecasting inflation is driven either by modeling assumptions or information regarding the FOMC's objectives and future policies.

2.7 The heterogeneity of model-based and expert forecasts

The model-based forecasts of output growth in the 2001 and 2008/09 recessions shown in Figures 2.1 to 2.4 indicate a substantial degree of heterogeneity that varies over time during these episodes. In this section, we document the extent and dynamics of forecast heterogeneity somewhat more systematically. To quantify forecast heterogeneity we compute the standard deviation of the cross section of individual forecasts for each horizon at any point in time. This standard deviation is defined as follows:

$$\sigma_t = \sqrt{\sum_{m=1}^{M} \left(E[y^{obs}_{t+h}|I^m_t] - \frac{1}{M} \sum_{m=1}^{M} E[y^{obs}_{t+h}|I^m_t] \right)^2 /(M-1)}, \tag{2.12}$$

where I^m_t denotes the information set of a specific model m at time t and M denotes the number of models used to forecast.

As a benchmark for comparison, we compute the same measure of forecast diversity for the cross section of individual expert forecasts from the Survey of Professional Forecasters. We only take into account forecasters who contributed at least four forecasts during one of the recessions. As a result of this selection, the number of individual forecasts taken from the SPF ranges from 9 to over 50, compared to the 6 individual model forecasts.

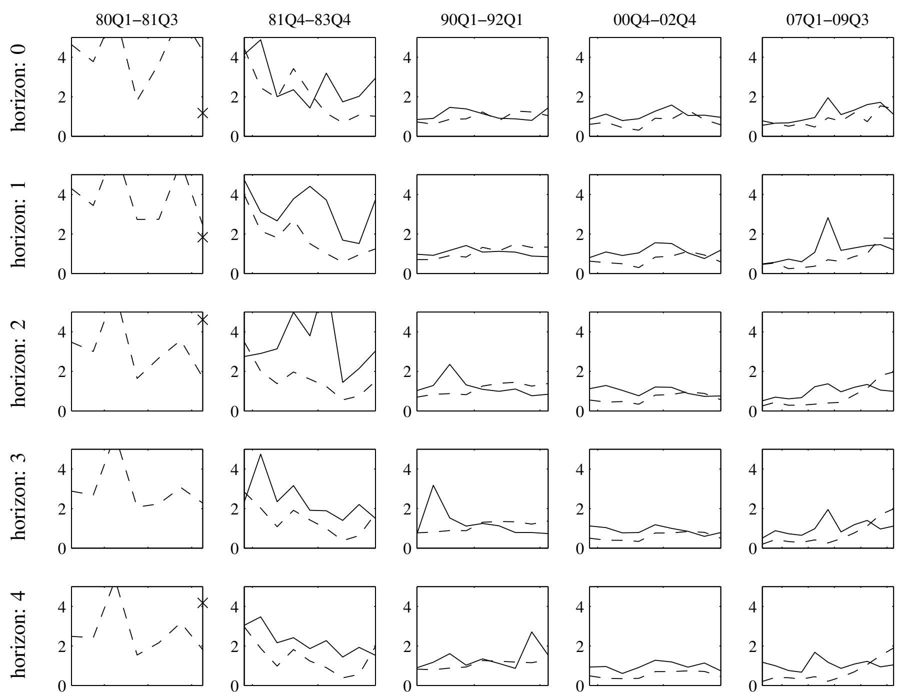


Figure 2.5: Standard deviations of output growth forecasts: experts (solid) and models (dashed)

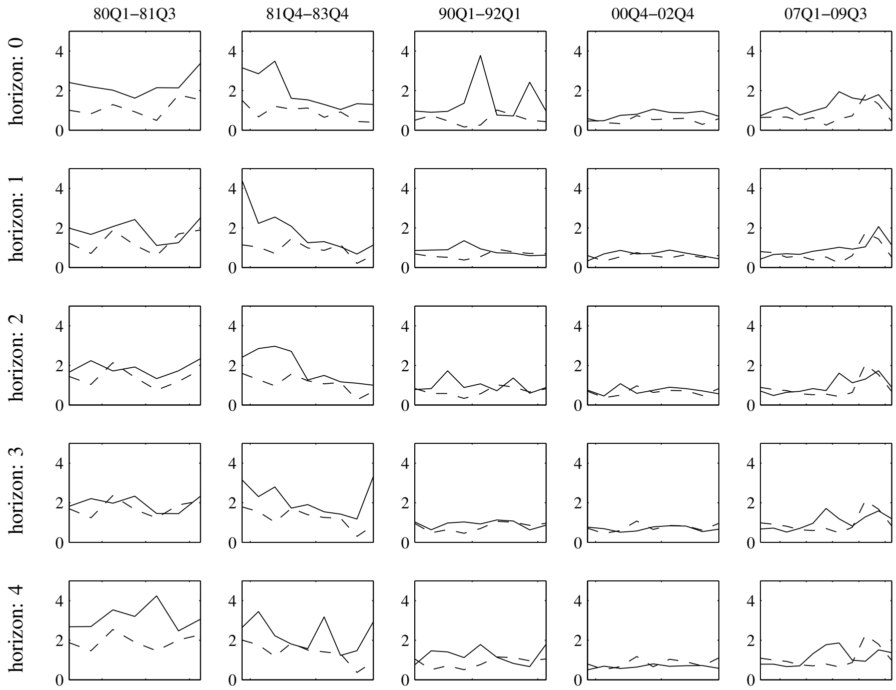


Figure 2.6: Standard deviations of inflation forecasts: experts (solid) and models (dashed)

Figures 2.5 and 2.6 display the standard deviations of model-based forecasts (dashed line) and professional forecasts (solid line). The rows show the different forecast horizons and the columns the different recessions. The dashed line indicates the diversity of model forecasts while the solid line measures the diversity of survey forecasts. Output growth forecasts of the SPF start in 1981Q3 which is marked with an x.

The extent of heterogeneity of GDP growth and inflation forecasts is roughly in the same range for model-based and expert forecasts, although it is somewhat lower for the models relative to the experts. The latter finding might be attributed to the much smaller number of individual model forecasts. The diversity of forecasts among the six models provides an indication of the extent of disagreement that may arise from different modeling assumptions, information sets and estimation methods. Since experts are faced with those same choices in developing their forecasting frameworks, the observed extent of heterogeneity in expert forecasts need not attributed to irrationality on behalf of individual forecasters.

We conduct some robustness checks to find out whether the heterogeneity measured by the standard deviation is strongly influenced by outliers. To this end, we compute the range between the 0.166 and 0.833 quantile for model-based and professional forecasts, that is we drop the highest and the lowest model forecast, compute the range between the second highest and second lowest forecast and compare to the same measure obtained from expert forecasts. The results confirm the finding that the models generate a similar degree of diversity as observed in the Survey of Professional Forecasters.

In addition, it is apparent from Figures 2.5 and 2.6 that the extent of forecast heterogeneity varies substantially over time. For example, diversity in output growth forecasts is most pronounced in the

1980s recessions and much smaller in the 1990-91 and 2001 recessions. It increases again in the 2008/09 recession. At several occasions model-based and survey forecasts of GDP growth exhibit similar dynamics. Examples are the decline in the diversity of three- to four-quarter ahead forecasts over the course of the 1981-82 recession (last two panels in the second column), or the increase in diversity in the middle of the 2000-2002 period (fourth column of panels). Also, heterogeneity increases throughout the latter part of the 2008/09 recession for model as well as expert forecasts. Of course, we also observe some spikes in disagreement among forecasters in the SPF that do not appear in the model-based forecasts. Examples are found in the GDP growth forecasts in 1990 and 2008. Such occasional spikes are not too surprising given that the SPF contains some extreme outliers. Rather, the co-movement visible in several episodes constitutes the more interesting finding, in our view.

Table 2.7: RMSE of best, worst, and average output growth forecaster from survey and models

	Horizons:	0	1	2	3	4
1980:1 - 1981:3						
min RMSE	Survey / Models	– / 5.64	– / 5.59	– / 5.24	– / 4.33	– / 4.39
max RMSE	Survey / Models	– / 7.19	– / 7.63	– / 8.69	– / 6.28	– / 7.33
average RMSE	Survey / Models	– / 6.62	– / 6.74	– / 6.39	– / 5.39	– / 5.46
1981:4 - 1983:4						
min RMSE	Survey / Models	1.15 / 2.89	2.37 / 3.69	1.40 / 4.06	2.30 / 4.15	2.26 / 4.02
max RMSE	Survey / Models	10.33 / 5.68	15.12 / 5.25	18.91 / 5.72	9.77 / 5.74	10.22 / 5.74
average RMSE	Survey / Models	3.30 / 4.12	4.95 / 4.39	4.93 / 4.58	4.73 / 4.65	4.28 / 4.53
1990:1 - 1992:1						
min RMSE	Survey / Models	0.69 / 1.76	0.63 / 2.06	0.86 / 2.15	0.97 / 2.38	0.08 / 2.30
max RMSE	Survey / Models	2.36 / 3.22	2.74 / 3.94	4.67 / 4.00	5.23 / 3.90	8.54 / 3.73
average RMSE	Survey / Models	1.54 / 2.59	1.69 / 3.07	1.88 / 3.09	1.88 / 3.15	2.01 / 2.91
2000:4 - 2002:4						
min RMSE	Survey / Models	1.34 / 1.94	0.82 / 2.19	1.33 / 2.21	1.76 / 2.67	0.94 / 2.08
max RMSE	Survey / Models	4.72 / 2.63	3.49 / 2.64	4.22 / 2.61	3.76 / 2.82	3.10 / 2.58
average RMSE	Survey / Models	2.38 / 2.33	2.44 / 2.34	2.37 / 2.37	2.73 / 2.73	2.22 / 2.24
2007:4 - 2009:4						
min RMSE	Survey / Models	1.06 / 3.58	0.56 / 4.36	0.46 / 4.78	0.68 / 5.13	1.36 / 5.29
max RMSE	Survey / Models	12.95/ 4.42	12.03 / 5.18	7.77 / 5.36	9.28 / 5.66	7.70 / 5.91
average RMSE	Survey / Models	5.62 / 3.99	4.60 / 4.74	2.78 / 4.96	4.84 / 5.31	4.98 / 5.63

Another aspect of heterogeneity concerns the range of accuracy of forecasts by individual forecasters. Some forecasters perform consistently better than average while others tend to make greater errors on average. Thus, we also compare the accuracy range among expert forecasters to the range among individual model forecasts. To this end, we compute the root mean squared error of the forecasts made by individual participants in the SPF for the different recession samples.

Table 2.8: Best, worst, and average inflation forecaster from survey and models

	Horizons:	0	1	2	3	4
1980:1 - 1981:3						
min RMSE	Survey / Models	0.35 / 1.76	1.12 / 1.90	0.60 / 1.38	0.30 / 2.32	1.84 / 2.29
max RMSE	Survey / Models	5.81 / 2.67	4.92 / 3.55	4.50 / 2.57	4.46 / 3.88	8.49 / 3.97
average RMSE	Survey / Models	1.90 / 2.15	2.19 / 2.47	2.16 / 1.92	2.71 / 2.96	3.36 / 3.08
1981:4 - 1983:4						
min RMSE	Survey / Models	0.70 / 1.37	0.58 / 1.47	0.82 / 1.29	1.38 / 1.31	0.82 / 1.22
max RMSE	Survey / Models	6.52 / 2.41	9.36 / 2.71	6.42 / 2.63	9.58 / 2.85	6.56 / 2.87
average RMSE	Survey / Models	1.94 / 1.77	2.38 / 2.13	2.41 / 1.96	2.67 / 2.11	2.73 / 2.06
1990:1 - 1992:1						
min RMSE	Survey / Models	0.63 / 1.05	0.51 / 1.16	0.50 / 1.07	0.41 / 0.81	0.38 / 1.37
max RMSE	Survey / Models	8.40 / 1.80	2.27 / 2.03	2.98 / 1.76	2.35 / 1.76	2.46 / 1.87
average RMSE	Survey / Models	1.63 / 1.25	1.19 / 1.52	1.25 / 1.39	1.30 / 1.21	1.35 / 1.61
2000:4 - 2002:4						
min RMSE	Survey / Models	0.36 / 0.90	0.21 / 0.92	0.44 / 1.11	0.41 / 1.16	0.31 / 1.07
max RMSE	Survey / Models	2.50 / 1.27	1.83 / 1.43	2.73 / 1.50	2.18 / 1.75	1.85 / 1.83
average RMSE	Survey / Models	0.92 / 1.08	1.00 / 1.18	1.07 / 1.33	1.03 / 1.44	1.08 / 1.48
2007:4 - 2009:4						
min RMSE	Survey / Models	0.77 / 1.58	0.42 / 1.14	0.75 / 1.23	0.56 / 1.36	0.55 / 1.38
max RMSE	Survey / Models	6.00 / 2.19	2.52 / 1.83	4.21 / 1.95	4.31 / 1.99	4.99 / 1.78
average RMSE	Survey / Models	1.63 / 1.85	1.23 / 1.46	1.43 / 1.53	1.46 / 1.73	1.61 / 1.60

Table 2.7 reports the worst, best and average RMSE of the individual expert forecasters during the five recession episodes. We only take into account those forecasters who contribute at least four forecasts for one of the recessions, otherwise a very low RMSE can be achieved by forecasting only during times of little volatility. The average RMSE for output growth forecasts of survey participants and the six models lies in a similar range, with the 1990-91 recession being an exception. During this

recession the model forecasts are on average of worse quality than the forecasts of survey participants. The range of forecast accuracies is much wider in the SPF than among the six models. The SPF has some extreme outliers. The worst RMSE is as high as 18.91 in the 1981-82 recession for a forecast horizon of two quarters. The highest model RMSE of 8.69 is generated by the BVAR-WW model in the 1980 recession for a forecast horizon of two quarters. With few exceptions the maximal RMSE is higher among survey participants than among the models and the minimal RMSE is lower among survey participants than among models. The lowest survey RMSE is as low as 0.08 for a four-quarter horizon in the 1990-91 recession. The lowest RMSE among the models is the nowcast of output growth in the 1990's recession with 1.76 and is also produced by the BVAR-WW model.
Table 2.8 reports the same statistics for the inflation forecasts. The average RMSE from the survey participants is always close to the average RMSE from the models. The best survey forecaster always performs better than the best model forecast. The worst survey forecast is with only one exception worse than the worst model forecast. The best survey RMSE is achieved for the 2001 recession for forecasting horizon of one quarter with a RMSE of 0.21. The best model RMSEs are given by 0.81 for the 1990-91 recession at a horizon of three quarters produced by the NK-Fu model and by 0.82 for the 2001 recession nowcast produced by the FRB-EDO model. We checked whether including the Greenbook or Survey nowcast in the information set for model-based forecasts changes these statistics. The models' minimal, maximal, and average RMSEs decrease by a small amount.

2.8 Conclusions

In recent years, researchers such as Smets and Wouters (2004), Adolfson et al. (2005), Smets and Wouters (2007), Christoffel et al. (2008), Del Negro et al. (2007) and Wang (2009) have reported encouraging findings regarding the forecasting performance of state-of-the art structural models. By contrast, the failure of researchers and professional forecasters to predict the "Great Recession" of 2008 and 2009 has generated much public criticism regarding the state of economic forecasting and macroeconomic modeling. Against this background, our analysis of the forecasting performance of models and experts during recessions provides several new insights.

The relative accuracy of model versus expert forecasts

First, we depart from the above-mentioned studies by using the real-time data vintages that were available in the past as the basis for evaluating forecasts of structural macroeconomic models. In doing so, we follow Faust and Wright (2009) who have shown that forecasts from non-structural models using ex-post revised data have uniformly smaller RMSEs than their counterparts estimated on real-time data. Thus, a comparison of structural model forecasts with historical expert forecasts has to be conducted on the basis of the real-time data vintages that could have been used by these experts at the time.[12]

[12]Faust and Wright (2009) find that the relative performance of non-structural models is less affected by using ex-post

Our focus on forecasting performance during recessions helps reveal that both, model and expert forecasts, tend to miss downturns. Interestingly, however, the model-based forecasts can do quite well during the recovery phase, sometimes even better than the Greenbook or mean-professional forecasts. Some model forecasts also predict the speed of recovery from the "Great Recession" surprisingly well. Model-based forecasts, in particular the mean model forecast,[13] compare quite well to the Greenbook and mean SPF forecasts, especially at a horizon of three to four quarters into the future. 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.

Professional forecasters typically make use of extensive survey information and higher-frequency indicators that help improve the estimate of current GDP prior to the first GDP release from the Bureau of Economic Analysis. Thus, it is not surprising if their forecasts detect recessions a little earlier than model forecasts. However, model forecasts could be combined with such higher-frequency information (e.g. Giannone, Monti and Reichlin (2009)). To approximate the effect of efficient now-casting we also conduct our comparisons between model-based and professional forecasts by starting from the professional nowcast. As a result, the gap between the two types of forecasts is further reduced.

Comparing model and expert forecast heterogeneity

We also quantify the extent of heterogeneity by means of the standard deviation across individual expert and model forecasts for a given forecasting horizon. The six model forecasts exhibit a broadly similar extent of forecast heterogeneity as the Survey of Professional Forecasters. The degree of forecast heterogeneity can change substantially over time. The standard deviations of model and professional forecasts vary over the course of the particular recession episodes that we examine as well as between different episodes. In some episodes the dynamics of forecast diversity derived from the two types of forecasts are quite similar.

In addition, we compare the forecast quality of different forecasters and models. In other words, we contrast the best, worst and average forecaster among models and professionals. This range is much greater among the professionals in the SPF than among the different models. In other words, some professional forecasters are consistently worse than the worst model, while some others perform consistently better than the best model. Thus, the range of accuracy of individual model forecasts does not approach the range observed in the Survey of Professional Forecasters.

How can the comparison of expert and model forecast heterogeneity be interpreted? Of course, some of the models considered were not available to professional forecasters during the earlier recession

revised data. Whether this is also true for structural model still needs to be investigated.

[13]Our mean model forecast combines five structural models with a non-structural Bayesian VAR model. In light of the finding by Del Negro et al. (2007) that a 'hybrid' model which contains priors from a DSGE model and has otherwise a VAR structure performs better than either a structural DSGE model or a non-structural VAR this combination should be expected to improve forecast performance.

episodes. For example, state-of-the-art medium-scale DSGE models such as the CEE-SW and FRB-EDO models only became available in time for the recession of 2008/2009. Non-structural VAR models, however, have been used during all the episodes that we consider and the model of Fuhrer (1997) is representative of the New Keynesian structural models that were already in use in the late 1980s and early 1990s. Furthermore, the reduced-form three-equation VAR implied by the linearized New Keynesian models with microeconomic foundations (NK-DS and NK-WW) is not that different from the reduced-form VAR's implied by the earlier generation of New Keynesian models. The microeconomic foundations simply imply additional cross-equation restrictions.

We interpret the comparison of the extent and dynamics of heterogeneity of model and expert forecasts as follows: while we can only speculate about the sources of disagreement among expert forecasters, the extent of disagreement among our six model forecasts can be traced to differences in modeling assumptions, different data coverage and different estimation methods. These three sources of disagreement are found to be sufficient to generate an extent of heterogeneity that is similar to the heterogeneity observed among expert forecasts. Furthermore, the recursive updating of model parameter estimates with incoming data induces dynamics in model forecast heterogeneity. In several episodes, expert forecast diversity even exhibits roughly similar variations. As a consequence of these findings, we would argue that it is not necessary to take recourse to irrational behavior or perverse incentives in order to explain the dynamics of expert forecast diversity.[14] Rather, this diversity may largely be due to model uncertainty and belief updating in a world where the length of useful data series is limited by structural breaks.[15]

On one side, our findings are encouraging in terms of the accuracy of forecasts derived from currently available structural macroeconomic models relative to expert forecasts from surveys. On the other side, our findings underscore the importance of research on models with heterogenous expectations. Using models with homogenous rational expectations for real-world forecasting, we estimate a significant range of forecast diversity that arises from different beliefs about appropriate modeling assumptions, estimation techniques and parameter estimates. This belief diversity itself may be a source of volatility. Of course, our models would attribute such volatility to shocks or other propagation mechanisms rather than endogenous heterogeneity in beliefs. Models with heterogenous expectations provide an avenue for distinguishing this source of economic fluctuations from other candidate propagation mechanisms.

Clearly, this is an important area for research on macroeconomic modeling. One direction for progress is suggested by the theory of rational beliefs (see Kurz, 2009, for a detailed introduction into the the-

[14] Notwithstanding forecasters may face incentives to publish a forecast close to the consensus (Scharfstein and Stein, 1990; Lamont, 2002) or a very distinct forecast (Laster et al., 1999).

[15] Others have documented the strong time variation of disagreement among survey forecasts. For example, Mankiw et al. (2004) have investigated disagreement in inflation surveys. Engelberg et al. (2009) and Clements (2010) investigate the properties of SPF forecasts, the extent of heterogeneity and the cross-sectional histograms of survey forecasts. Similar in spirit to our analysis, Williams (2004) used multiple non-structural time series model to quantify the extent of inflation forecast heterogeneity due to model uncertainty. He concludes that model uncertainty provides an intuitively more appealing description of the observed diversity of inflation expectations than staggered information updating as suggested by Mankiw and Reis (2007).

ory of rational beliefs). Our set of models might be interpretable as beliefs in such a context. The theory of rational beliefs assumes people optimize given the limited knowledge they have and may make mistakes. They know that it is impossible to ever learn the true structural relationships and probability laws because structural breaks limit the length of useful data series. Diversity arises when market participants have different beliefs about the true data generating process and therefore estimate different models to forecast macroeconomic variables. Diverse beliefs are rational if they are consistent with the empirical distribution. The papers by Kurz and Motolese (2010), Guo et al. (2010) and Nielsen (2010) apply the theory of rational beliefs. Branch and McGough (2010), Branch and Evans (2010) and de Grauwe (2010) provide another avenue for studying heterogeneity of beliefs by modeling agents with cognitive limitations that generate boundedly rational forecasting rules. The latter two papers impose heterogenous expectations directly into a New Keynesian model. Instead of having rational expectations agents use small forecasting models. An interesting area for future research would be to estimate such models with heterogeneous expectations and compare the importance of belief diversity as a source of economic fluctuations relative to the propagation mechanisms considered by the homogenous rational expectations models in this chapter.

References

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

An, S., Schorfheide, F., 2007. Bayesian analysis of DSGE models. Econometric Reviews 26(2-4), 113–172.

Bernanke, B. S., Boivin, J., 2003. Monetary policy in a data-rich environment. Journal of Monetary Economics 50(3), 525–546.

Branch, W. A., Evans, G. W., 2010. Monetary policy with heterogeneous expectations. Economic Theory, forthcoming.

Branch, W. A., McGough, B., 2010. Business cycle amplification with heterogeneous expectations. Economic Theory, forthcoming.

Brock, W., Hommes, C., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22, 1235–1274.

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

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

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

Capistran, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. Journal of Money, Credit, and Banking 41, 365–396.

Chiarella, C., Dieci, R., He, X.-Z., 2007. Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework. Journal of Economic Behavior and Organization 62, 408–427.

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

Christoffel, K., Coenen, G., Warne, A., 2008. The New Area-Wide Model of the euro area - a micro-founded open-economy model for forecasting and policy analysis, European Central Bank Working Paper 944.

Clements, M. P., 2010. Explanations of the inconsistencies in survey respondents' forecasts. European Economic Review 54(4), 536–549.

de Grauwe, P., 2010. Animal spirits and monetary policy. Economic Theory, forthcoming.

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

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of New Keynesian models. Journal of Business and Economic Statistics 25(2), 123–143.

Doan, T., Litterman, R., Sims, C., 1984. Forecasting and conditional projection using realistic prior distributions. Econometric Reviews 3, 1 – 100.

Edge, R. M., Kiley, M. T., Laforte, J.-P., 2007. Documentation of the research and statistics divisions estimated DSGE model of the U.S. economy: 2006 version, Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.: 2007-53.

Edge, R. M., Kiley, M. T., Laforte, J.-P., 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., Kiley, M. T., Laforte, J.-P., 2010. A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced form models, and a DSGE model. Journal of Applied Econometrics forthcoming.

Engelberg, J., Manski, C. F., Williams, J., 2009. Assessing the temporal variation of macroeconomic forecasts by a panel of changing composition, Northwestern University Working Paper.

Faust, J., Wright, J. H., 2009. Comparing Greenbook and reduced form forecasts using a large realtime dataset. Journal of Business and Economic Statistics 27(4), 468–479.

Francis, N., Ramey, V. A., 1995. Measures of per capita hours and their implications for the technology-hours debate, NBER Working Paper 11694.

Fuhrer, J. C., 1997. Inflation/output variance trade-offs and optimal monetary policy. Journal of Money, Credit and Banking 29(2), 214–234.

Fuhrer, J. C., Moore, G., 1995a. Inflation persistence. The Quarterly Journal of Economics 110(1), 127–159.

Fuhrer, J. C., Moore, G., 1995b. Monetary policy trade-offs and the correlation between nominal interest rates and real output. The American Economic Review 85(1), 219–239.

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

Giordani, P., Söderlind, P., 2003. Inflation forecast uncertainty. European Economic Review 47, 1037–1059.

Goodfriend, M., King, R. G., 1997. The New Neoclassical Synthesis and the role of monetary policy. In: Bernanke, B. S., Rotemberg, J. J. (Eds.), National Bureau of Economic Research Macroeconomics Annual 1997. MIT Press, Cambridge, MA.

Guo, W. C., Wang, F. Y., Wu, H. M., 2010. Financial leverage and market volatility with diverse beliefs. Economic Theory, forthcoming.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Kimball, M., 1995. The quantitative analytics of the basic monetarist model. Journal of Money, Credit and Banking 27(4), 1241–1277.

Klein, L. (Ed.), 1991. Comparative Performance of U.S. Econometric Models. Oxford, Eng.: Oxford University Press.

Kurz, M., 1994a. On rational belief equilibria. Economic Theory 4, 859–876.

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

Kurz, M., 1996. Rational beliefs and endogenous uncertainty: an introduction. Economic Theory 8, 383–397.

Kurz, M., 1997a. Endogenous economic fluctuations and rational beliefs: A general perspective. In: Kurz, M. (Ed.), Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M. (Ed.), 1997b. Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M., 2009. Rational diverse beliefs and market volatility. In: Hens, T., Schenk-Hoppe, K. (Eds.), Handbook of financial markets: dynamics and evolution. North Holland.

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

Kurz, M., Jin, H., Motolese, M., 2005. The role of expectations in economic fluctuations and the efficacy of monetary policy. Journal of Economic Dynamics & Control 29, 2017–2065.

Kurz, M., Motolese, M., 2010. Diverse beliefs and time variability of risk premia. Economic Theory, forthcoming.

Lamont, O. A., 2002. Macroeconomic forecasts and microeconomic forecasters. Journal of Economic Behavior and Organization 48(3), 265–280.

Laster, D., Bennett, P., Geoum, I. S., 1999. Rational bias in macroeconomic forecasts. Quarterly Journal of Economics 114(1), 293–318.

Mankiw, N. G., Reis, R., 2007. Sticky information in general equilibrium. Journal of the European Economic Association 5 (2-3), 603–613.

Mankiw, N. G., Reis, R., Wolfers, J., 2004. Disagreement about inflation expectations. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2003. Vol. 18. Cambridge, Mass.: MIT Press, pp. 209–248.

Mishkin, F., 2004. Discussion of 'Disagreement about inflation expectations' by N. Gregory Mankiw, Ricardo Reis, and Justin Wolfers. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2003. Vol. 18. Cambridge, Mass.: MIT Press, pp. 257–268.

Nielsen, C. K., 2010. Price stabilizing, pareto improving policies. Economic Theory, forthcoming.

Romer, C. D., Romer, D. H., 2000. Federal reserve information and the behavior of interest rates. American Economic Review 90, 429–457.

Rotemberg, J. J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy, in B. Bernanke and J. Rotemberg, (eds.), NBER Macroeconomics Annual, The MIT Press.

Scharfstein, D. S., Stein, J. C., 1990. Herd behavior and investment. American Economic Review 80(3), 465–479.

Sims, C. A., 2002. The role of models and probabilities in the monetary policy process. Brookings Papers on Economic Activity 2, 1–40.

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

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

Taylor, J. B., 1979. Estimation and control of a macroeconomic model with rational expectations. Econometrica 47 (5), 1267–1286.

Taylor, J. B., 1993. Macroeconomic Policy in a World Economy. W.W. Norton, New York.

Taylor, J. B., Wieland, V., 2009. Surprising comparative properties of monetary models: Results from a new data base, NBER Working Paper 14849.

Walsh, C. E., 2003. Monetary theory and policy. The MIT Press, Cambridge.

Wang, M.-C., 2009. Comparing the DSGE model with the factor model: An out-of-sample forecasting experiment. Journal of Forecasting 28(2), 167–182.

Wieland, V. (Ed.), 2009. The Science and Practice of Monetary Policy Today. Springer Science.

Wieland, V., Cwik, T., Müller, G. J., Schmidt, S., Wolters, M., 2009. A new comparative approach to macroeconomic modeling and policy analysis, Manuscript, Center for Financial Studies, Frankfurt.

Williams, J. C., 2004. Discussion of 'Disagreement about inflation expectations' by N. Gregory Mankiw, Ricardo Reis, and Justin Wolfers. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2003. Vol. 18. Cambridge, Mass.: MIT Press, pp. 257–268.

Woodford, M., 2003. Interest and prices: Foundations of a theory of monetary policy. Princeton University Press, Princeton.

Appendix

A.1 The macroeconomic models used to compute forecasts

This appendix provides a description of the six macroeconomic models that are used in this chapter to generate forecasts. In the case of the NK-Fu, NK-DS, CEE-SW and FRB/EDO models our notation follows exactly the notation in the model authors' original articles.

BVAR-WW Model: Non-structural VAR models have been available to forecasters for decades and are still being used by practitioners today. Such a VAR is a more general description of the data than the DSGE models as it imposes little restrictions on the data generating process. All variables are treated symmetrically and therefore the VAR incorporates no behavioral interpretations of parameters or equations. We estimate such a VAR on output growth, inflation and the federal funds rate using Bayesian methods. Each of the variables is regressed on a constant, four lagged values of the variable itself and four lagged values of the other two variables. It is well known that unrestricted VARs are heavily overparameterized. To improve forecast performance it is important to shrink the parameter space in some manner. We follow Doan et al. (1984) and use the so-called Minnesota prior to avoid over-parameterization. This prior implies shrinking the parameters towards zero by assuming that the price level, real output and the interest rate follow independent random walks. All parameters are assumed to be normally distributed with mean zero. The variance around these zero priors decreases with lag-length. The rationale for this assumption is that short lags contain more information about the dependent variables than long lags.

NK-Fu Model: The model of Fuhrer (1997) is a good example of the New Keynesian models that were developed in the 1980s and early 1990s.[16] While academics still focused primarily on developing the microeconomic foundations of real business cycle theory, these models became quite popular among central bank researchers and practitioners. They took into account adaptive and forward-looking behavior of market participants, real effects of monetary policy and output and inflation persistence. The model of Fuhrer (1997) exhibits a high degree of inertia with respect to aggregate demand which is determined by the following IS-curve:

$$\tilde{y}_t = a_0 + a_1\tilde{y}_{t-1} + a_2\tilde{y}_{t-2} + a_\rho\rho_{t-1} + \epsilon_{y,t}, \tag{2.13}$$

$\tilde{y}_t$ denotes the output gap, which is computed as the deviation from the log-linear trend. ρ_t denotes the long-term real interest rate and $\epsilon_{y,t}$ a demand shock. The long-term real interest rate is determined by an intertemporal arbitrage condition that equalizes the expected holding-period yields on government

[16] For other examples see the model comparison projects of Bryant et al. (1988), Bryant et al. (1989), Klein (1991), and Bryant et al. (1993).

bonds and real long-term bonds:

$$\rho_t - D\left[E_t(\rho_{t+1}) - \rho_t\right] = f_t - E_t(\pi_{t+1}). \tag{2.14}$$

f_t denotes the federal funds rate, π_t the quarterly inflation rate and D is a constant approximation for Macaulay's duration that is set equal to 10 years.

The short-run aggregate supply nexus between output and inflation is importantly influenced by overlapping wage contracts. Fuhrer assumes that wage contracts that remain in effect for one to four quarters are negotiated relative to the real wage implied by those set in the recent past and those that are expected to be negotiated in the near future (see Fuhrer and Moore, 1995a,b). ν_t denotes an index of wage contracts that are currently in effect:

$$\nu_t = \sum_{i=0}^{3} \omega_i (x_{t-i} - p_{t-i}), \tag{2.15}$$

where x_t denotes the log wage contract negotiated in period t and p_t the log price level. The weights ω_i are the proportions of the outstanding contracts and sum to one. The weights decrease for contracts negotiated in earlier periods. The current nominal wage contract is determined such that the current real wage contract equals the average real contract wage index expected to prevail over the life of the contract. Additionally, it adjusted for expected excess demand conditions as reflected in current and expected future output gaps:

$$x_t - p_t = \sum_{i=0}^{3} \omega_i (\nu_{t+i} + \gamma \tilde{y}_{t+i}) + \epsilon_{p,t}. \tag{2.16}$$

$\epsilon_{p,t}$ is a cost-push shock. The aggregate log wage index is a weighted average of the log of wage contracts. The aggregate price level is a constant mark-up (normalized to zero) over the aggregate wage rate. Inflation dynamics depend on current, past and expected future demand. The model is quite successful in matching the strong inflation persistence observed in U.S. data. Inflation is given by an average of changes in the log nominal wage contracts:

$$\pi_t = \sum_{i=0}^{3} \omega_i (x_{t-i} - x_{t-i-1}). \tag{2.17}$$

The model is closed with a monetary policy reaction function. The Fed is assumed to set the federal funds rate with respect to a constant equilibrium value, the lagged funds rate, inflation, lagged inflation, the output gap and the change in the output gap. Deviations from the reaction function are interpreted as monetary policy shocks:

$$f_t = \alpha_0 + \alpha_{f1} f_{t-1} + \alpha_{\pi 0} \pi_t + \alpha_{\pi 1} \pi_{t-1} + \alpha_{\Delta y} (\tilde{y}_t - \tilde{y}_{t-1}) + \alpha_y \tilde{y}_t + \epsilon_{f,t}. \tag{2.18}$$

Contrary to the other structural models considered in this chapter, Fuhrer allows for the possibility of contemporaneously correlated structural shocks. The variance-covariance matrix is estimated together with the parameters of the model.

NK-DS Model: The model by Del Negro and Schorfheide (2004) is an example of small-scale New Keynesian models with microeconomic foundations in the vein of Rotemberg and Woodford (1997) and Goodfriend and King (1997). A representative household derives utility from consumption relative to a habit stock that depends on the level of technology. Hours worked reduce the household's utility and real money balances increase it. The utility function is additively separable. Utility is maximized over an infinite lifetime subject to the household's budget constraint. The household earns income from different sources: wage income from supplying perfectly elastic labor services to firms, interest rate payments from bond holdings and profits from the firms. It pays lump-sum taxes. Utility maximization implies an Euler equation. Linearizing this equation and imposing market clearing (output equals consumption and government spending) yields the New Keynesian forward-looking IS-equation:

$$x_t = E_t x_{t+1} - \tau^{-1}(R_t - E_t \pi_{t+1}) + (1 - \rho_g) g_t + \rho_z \tau^{-1} z_t, \tag{2.19}$$

x_t denotes output, π_t inflation and R_t the federal funds rate. τ is the risk aversion parameter of the household. All variables are defined in percentage deviations from steady state. g_t and z_t are government spending and technology shock processes. Both shocks follow AR(1) processes (not shown) with parameters ρ_g and ρ_z. The government consumes a fraction of output which fluctuates exogenously according to the shock process: ξ_t denotes the fraction of output consumed by the government and the shock is defined as $g_t = 1/(1 - \xi_t)$. The government issues bonds that can be bought by households and it collects lump-sum taxes to finance its expenditures.

The production sector consists of a continuum of monopolistically competitive firms that are owned by the households. They face demand curves that can be derived from a Dixit-Stiglitz final good aggregator. Nominal rigidities are modelled via quadratic price adjustment costs. Firms pay these costs in form of an output loss when they desire to set a price in deviation from the level implied by steady-state inflation. The production function is linear in labor. Labor is hired from the households. Total factor productivity follows a unit root process. Thus, it induces a stochastic trend into the model. As a result, output fluctuates around the steady-state growth rate. Firms maximize the present value of expected profits over an infinite horizon. The optimality condition implies that prices are set as a fixed mark-up over marginal cost. Linearizing this first order condition leads to the following New Keynesian forward-looking Phillips curve:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (x_t - g_t), \tag{2.20}$$

where β is the household's discount factor and κ is a function of the price adjustment cost parameter and the elasticity of demand. Inflation is a function of marginal cost which can be substituted with the output gap. The model is closed with a monetary policy rule. The rule assumes that the central bank sets the current interest rate as a function of current inflation, the output gap, and the previous interest rate choice:

$$R_t = \rho_R R_{t-1} + (1 - \rho_R)(\psi_1 \pi_t + \psi_2 x_t) + \epsilon_{R,t}. \tag{2.21}$$

The monetary policy shock, $\epsilon_{R,t}$, is assumed iid-normally distributed. ρ_R indicates the degree of interest rate smoothing and ψ_1 and ψ_2 capture the policy response to inflation and output gaps. The IS equation and the policy rule together represent the aggregate demand side, while the Phillips curve captures fluctuations in aggregate supply.

NK-WW model: The NK-WW model generalizes the NK-DS model in terms of the economic shocks considered. To allow for richer output and inflation dynamics we add serially correlated preference and mark-up shock processes χ_t and Φ_t. Both shocks follow AR(1) processes with parameters ρ_χ and ρ_Φ. The preference shock enters the consumption term in the utility function and appears in the New Keynesian IS-equation:

$$x_t = E_t x_{t+1} - \tau^{-1}(R_t - E_t \pi_{t+1}) + (1 - \rho_g) g_t + \rho_z \tau^{-1} z_t + \tau^{-1}(1 - \rho_\chi)\chi_t, \tag{2.22}$$

Both shocks enter the New Keynesian Phillips curve. The mark-up shock has a direct effect on inflation. The preference shock influences marginal costs and thereby also inflation determination:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \left[x_t - g_t + \tau^{-1}(\Phi_t - \chi_t) \right]. \tag{2.23}$$

The monetary policy rule is the same as in the NK-DS model.

CEE-SW Model: Building on the above-mentioned micro-founded New Keynesian model Christiano, Eichenbaum and Evans (2005) developed the first medium-scale New Keynesian DSGE model that can fit a significant number of important empirical regularities of the U.S. economy (NBER working paper 2001). Smets and Wouters (2003, 2007) extended this model and estimated it with Bayesian methods. The CEE-SW model contains a large number of frictions and structural shocks. Physical capital is included in the production function and capital formation is endogenous. Labor supply is modeled explicitly. Nominal frictions include sticky prices and wages and inflation and wage indexation. Real frictions include consumption habit formation, investment adjustment costs and variable capital utilization. Utility is nonseparable in consumption and leisure. There exist fixed costs in production and the Dixit-Stiglitz aggregator is replaced with the aggregator by Kimball (1995) which implies a non-constant elasticity of demand. The model contains seven structural shocks and is fit to seven time series. Among the shocks are, total factor productivity, risk premium, investment-specific technology, wage mark-up, price mark-up, government spending and monetary policy shocks. All

shock processes are serially correlated. In the following we describe each of the linearized equations of the model following the notation in Smets and Wouters (2007).
The resource constraint is given by:

$$y_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g, \tag{2.24}$$

where output y_t is the sum of consumption, c_t, and investment, i_t, weighted with their steady state ratios to output (c_y and i_y), the capital-utilization cost which depends on the capital utilization rate, z_t, and an exogenous government spending shock ϵ_t^g. ϵ_t^g follows an AR(1) process and is also affected by the technology shock. z_y equals $R_*^k k_y$, where k_y is the ratio of capital to output in steady state and R_*^k is the rental rate of capital in steady state. Combining the households' first order conditions for consumption and bond holdings yields the consumption Euler equation

$$c_t = c_1 c_{t-1} + (1 - c_1)E_t(c_{t+1}) + c_2(l_t - E_t(l_{t+1})) - c_3(r_t - E_t(\pi_{t+1}) + \epsilon_t^b. \tag{2.25}$$

The parameters are $c_1 = (\lambda/\gamma)/(1+\lambda/\gamma)$, $c_2 = [(\sigma_c - 1)(W_*^h L_*/C_*)]/[(\sigma_c(1+\lambda/\gamma)]$ and $c_3 = (1-\lambda/\gamma)/[(1+\lambda/\gamma)\sigma_c]$. λ governs the degree of habit formation, γ is the labor augmented steady growth rate, σ_c the inverse of the intertemporal elasticity of substition and parameters with a $*$ subscript denote steady state values. ϵ_t^b denotes an AR(1) shock process on the premium over the central bank controlled interest rate. Consumption is a weighted average of past and expected consumption due to habit formation. The consumption Euler equation depends on hours worked, l_t, because of the nonseparability of utility. When consumption and hours are complements ($\sigma_c > 1$), consumption increases with current hours and decreases with expected hours next period. The real interest rate and the shock term affect aggregate demand by inducing intertemporal substitution in consumption.
The investment Euler equation is given by

$$i_t = i_1 i_{t-1} + (1 - i_1)E_t(i_{t+1}) + i_2 q_t + \epsilon_t^i, \tag{2.26}$$

where $i_1 = 1/(1 + \beta\gamma^{1-\sigma_c})$ and $i_2 = [1/(1 + \beta\gamma^{1-\sigma_c})\gamma^2\phi]$. β denotes the discount factor, ϕ the elasticity of the capital adjustment cost function, q_t Tobin's Q and ϵ_t^i an investment specific technology shock that follows an AR(1) process. Current investment is a weighted average of past and expected future investment due to the existence of capital adjustment costs. It is positively related to the real value of the existing capital stock. This dependence decreases with the elasticity of the capital adjustment cost function. The arbitrage equation for the real value of the capital stock is:

$$q_t = q_1 E_t(q_{t+1}) + (1 - q_1)E_t(r_{t+1}^k) - (r_t - E_t(\pi_{t+1}) + \epsilon_t^b), \tag{2.27}$$

where $q_1 = \beta\gamma^{-\sigma_c}(1 - \delta)$. r_t^k denotes the real rental rate of capital and δ the depreciation rate of capital. The real value of the existing capital stock is a positive function of its expected value next period and the rental rate on capital and a negative function of the real interest rate and the external

finance premium.
The production process is assumed to be determined by a Cobb-Douglas production function with fixed costs:

$$y_t = \phi_p(\alpha k_t^s + (1-\alpha)l_t + \epsilon_t^a). \tag{2.28}$$

k_t^s denotes effective capital (physical capital adjusted for the capital utilization rate), ϵ_t^a a neutral productivity shock that follows an AR(1) process and ϕ_p is one plus the share of fixed costs in production. Output is produced using capital and labour and is boosted by technology shocks. Capital used in production depends on the capital utilization rate and the physical capital stock of the previous period as new capital becomes effective with a lag of one quarter:

$$k_t^s = k_{t-1} + z_t. \tag{2.29}$$

Household income from renting capital services to firms depends on r_t^k and changing capital utilization is costly so that the capital utilization rate depends positively on the rental rate of capital:

$$z_t = (1-\psi)/\psi r_t^k, \tag{2.30}$$

where $\psi \in [0,1]$ is a positive function of the elasticity of the capital utilization adjustment cost function. The law of motion for physical capital is given by:

$$k_t = k_1 k_{t-1} + (1-k_1)i_t + k_2\epsilon_t^i, \tag{2.31}$$

where $k_1 = (1-\delta)/\gamma$ and $k_2 = (1-(1-\delta)/\gamma)(1+\beta\gamma^{1-\sigma_c})\gamma^2\phi$. The price mark-up μ_t^p equals the difference between the marginal product of labor and the real wage w_t:

$$\mu_t^p = \alpha(k_t^s - l_t) + \epsilon_t^a - w_t. \tag{2.32}$$

Monopolistic competition, Calvo-style price contracts, and indexation of prices that are not free to be chosen optimally combine to yield the following Phillips curve:

$$\pi_t = \pi_1\pi_{t-1} + \pi_2 E_t(\pi_{t+1}) - \pi_3\mu_t^p + \epsilon_t^p, \tag{2.33}$$

with $\pi_1 = \iota_p/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, $\pi_2 = \beta\gamma^{1-\sigma_c}/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, and $\pi_3 = 1/\ (1+\beta\ \gamma^{1-\sigma_c}\ \iota_p)$ $[(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\xi_p((\phi_p-1)\epsilon_p+1)]$. This Phillips curve contains not only a forward-looking but also a backward-looking inflation term because of price indexation. Firms that cannot adjust prices optimally either index their price to the lagged inflation rate or to the steady-state inflation rate. Note, this indexation assumption ensures also that the long-run Phillips curve is vertical. ξ_p denotes the Calvo parameter, ι_p governs the degree of backward indexation, ϵ_p determines the curvature of the Kimball (1995) aggregator. The Kimball aggregator complementarity effects enhance the price rigidity resulting from Calvo-style contracts. The mark-up shock ϵ_t^p follows an ARMA(1,1) process.

A monopolistic labor market yields the condition that the wage mark-up μ_t^w equals the real wage minus the marginal rate of substition mrs_t:

$$\mu_t^w = w_t - mrs_t = w_t - (\sigma_l l_t + \frac{1}{1-\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})), \tag{2.34}$$

with σ_l being the Frisch elasticity of labor supply. The wage Phillips curve ist given by:

$$w_t = w_1 w_{t-1} + (1-w_1)(E_t(w_{t+1}) + E_t(\pi_{t+1})) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \epsilon_t^w, \tag{2.35}$$

where $w_1 = 1/(1+\beta\gamma^{1-\sigma_c})$, $w_2 = (1+\beta\gamma^{1-\sigma_c}\iota_w)/((1+\beta\gamma^{1-\sigma_c}))$, $w_3 = \iota_w/(1+\beta\gamma^{1-\sigma_c})$, and $w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/(\xi_w((\phi_w-1)\epsilon_w+1))]$. The parameter definition is analogous to the price Phillips curve.
Setting $\xi_p = 0$, $\xi_w = 0$, $\epsilon_t^p = 0$ and $\epsilon_t^w = 0$ one obtains the efficient flexible price and flexible wage allocation. The output gap x_t is defined as the log difference between output and flexible price output just like in the small-scale New Keynesian models above.
The monetary policy rule reacts to inflation, the output gap and the change in the output gap and incorporates partial adjustment:

$$r_t = \rho r_{t-1} + (1-\rho)(r_\pi \pi_t + r_x x_t) + r_{\Delta x_t}(x_t - x_{t-1}) + \epsilon_t^r. \tag{2.36}$$

ϵ_t^r is a monetary policy shock that follows an AR(1) process.

FRB-EDO Model: The model by Edge et al. (2008) is a more disaggregated DSGE model that was developed at the Board of Governors of the Federal Reserve System. It features two production sectors, which differ in their pace of technological progress. This structure can capture the different growth rates and relative prices observed in the data. Accordingly, the expenditure side is disaggregated as well. It is divided into business investment and three categories of household expenditure: consumption of non-durables and services, investment in durable goods and residential investment. The model is able to capture different cyclical properties in these four expenditure categories. It includes 14 structural shocks: technology shocks, price and wage mark-up shocks, preference shocks, capital efficiency shocks, an external spending shock and a monetary policy shock. The model is estimated to fit eleven empirical time series: output growth, inflation, the federal funds rate, consumption of non-durables and services, consumption of durables, residential investment, business investment, hours, wages, inflation for consumer nondurables and services and inflation for consumer durables. We estimate a variant of the FRB-EDO model that is built as close to the documentation of (Edge et al., 2007) as possible. While the aggregate dynamics implied by our version of the model do not exactly match the figures in the authors' documentation, they come reasonably close to that.

In the following we describe the main equations of the model. There are two types of intermediate-good producing firms that differ with respect to the rate of technological progress in their production technology. Production depends on technology, utilized non-residential capital and labor. Non-residential capital is rented from capital owners and labor is hired from households. The first sector is called the business and institutions sector and most of its output is used for consumption. The sector is therefore denoted by cbi. The technology of the second sector grows at a faster rate. This sector is called the business sector and the produced goods are used for capital accumulation. It is therefore denoted by kb.

The intermediate-goods producing firms' cost minimization problems with respect to labor and utilized non-residential capital lead to the following optimal factor input conditions:

$$L_t^s = (1-\alpha)\widetilde{X}_t^s \frac{\widetilde{MC}_t^s}{\widetilde{W}_t^s}, \qquad \text{for } s = cbi, kb \tag{2.37}$$

$$\frac{\widetilde{K}_t^{u,nr,s}}{\Gamma_t^{x,kb}} = \alpha \widetilde{X}_t^s \frac{\widetilde{MC}_t^s}{\widetilde{R}_t^{nr,s}}, \qquad \text{for } s = cbi, kb \tag{2.38}$$

L_t^s is the labor input, $\widetilde{X}_t^s$ are the produced goods, $\widetilde{MC}_t^s$ are marginal costs, $\widetilde{W}_t^s$ is the nominal wage rate, $\widetilde{K}_t^{u,nr,s}$ is the amount of utilized non-residential capital, $\Gamma_t^{x,kb}$ is the growth rate of output in the kb sector, $\widetilde{R}_t^{nr,s}$ is the aggregate nominal rental rate on non-residential capital and α denotes the capital share in the production function. A tilde on a variable denotes stationarized variables.

The stationarized production function is given by:

$$\widetilde{X}_t^s = (L_t^s)^{(1-\alpha)} \left(\frac{\widetilde{K}_t^{u,nr,s}}{\Gamma_t^{x,kb}} \right)^{\alpha}, \qquad \text{for } s = cbi, kb \tag{2.39}$$

The intermediate-goods firms face monopolistic competition. Thus, they are able to set prices that maximize the present value of profits in the infinite future. When maximizing profits the firms have to take into account the demand for their goods. This demand function is derived from perfectly competitive final good firms that use a Dixit-Stiglitz aggregation function. Furthermore, price adjustment is constrained by a quadratic adjustment cost function. Adjustment costs are paid in the form of an output loss when the price adjustment exceeds an average of the steady state inflation rate and last period's inflation rate. The Phillips curve is given by:

$$\begin{aligned} \Theta_t^{x,s} \widetilde{MC}_t^s \widetilde{X}_t^s = {} & (\Theta_t^{x,s} - 1)\widetilde{P}_t^s \widetilde{X}_t^s \\ & + 100\chi^p (\Pi_t^{p,s} - \eta^p \Pi_{t-1}^{p,s} - (1-\eta^p)\Pi_*^{p,s})\Pi_t^{p,s} \widetilde{X}_t^s \widetilde{P}_t^s \\ & - \beta E_t \left(\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} 100\chi^p (\Pi_{t+1}^{p,s} - \eta^p \Pi_t^{p,s} - (1-\eta^p)\Pi_*^{p,s})\Pi_{t+1}^{p,s} \widetilde{P}_{t+1}^s \widetilde{X}_{t+1}^s \right), \end{aligned} \tag{2.40}$$

where $s = cbi, kb$. $\Theta_t^{x,s}$ is the stochastic elasticity of substitution between differentiated intermediate goods and governs shocks to the price mark-up over marginal cost. $\Pi_t^{p,s}$ is the inflation rate and $\Pi_*^{p,s}$

is the steady state inflation rate. $\widetilde{P}_t^s$ is the price level relative to the cbi sector ($\widetilde{P}_t^{cbi}$ is equal to 1). $\widetilde{\Lambda}_t^{cnn}$ denotes the marginal utility of the consumption good. The parameter χ^p reflects the size of adjustment costs in re-setting prices. η^p determines the relative importance of lagged inflation and steady state inflation in the adjustment cost function and β is the household's discount factor.

There are three different types of capital owners who invest in goods, transform these into the three different capital stocks and rent them to households and firms. Goods from the fast growing sector (kb) are transformed into non-residential capital or consumer-durable capital. Goods from the slow growing sector (cbi) are transformed into residential capital stock or directly used for household consumption. Capital evolution depends on a quadratic investment adjustment cost that is paid via a capital loss if current investment differs from investment in the previous period adjusted by the growth rate of the respective sector production. In addition there are stochastic capital efficiency shocks. The first-order condition of the non-residential capital owners with respect to the capital stock is given by:

$$\widetilde{Q}_t^{nr} = \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \frac{1}{\Gamma_{t+1}^{x,kb}} \left(\widetilde{R}_{t+1}^{nr} + (1-\delta^{nr}) \widetilde{Q}_{t+1}^{nr} \right) \right\}, \tag{2.41}$$

where $\widetilde{Q}_t^{nr}$ is the price of installed non-residential capital, $\widetilde{R}_t^{nr}$ is the nominal rental rate on non-residential capital and δ^{nr} is the depreciation rate. The first order condition with respect to investment in non-residential capital is given by:

$$\begin{aligned} \widetilde{P}_t^{kb} &= \widetilde{Q}_t^{nr} \left[A_t^{nr} - 100\chi^{nr} \left(\frac{\widetilde{E}_t^{nr} - \widetilde{E}_{t-1}^{nr}}{\widetilde{K}_t^{nr}} \Gamma_t^{x,kb} \right) \right] \\ &+ \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \widetilde{Q}_{t+1}^{nr} 100\chi^{nr} \left(\frac{\widetilde{E}_{t+1}^{nr} - \widetilde{E}_t^{nr}}{\widetilde{K}_{t+1}^{nr}} \Gamma_{t+1}^{x,kb} \right) \right\}. \end{aligned} \tag{2.42}$$

A_t^{nr} is a capital efficiency shock, χ^{nr} is an investment adjustment cost parameter, $\widetilde{E}_t^{nr}$ denotes expenditure on goods used for non-residential investment and $\widetilde{K}_t^{nr}$ is the non-residential capital stock. Other conditions that include the capital accumulation equation and the market clearing condition for non-residential capital used in the production process in both sectors are given by:

$$\widetilde{R}_t^{nr,s} = \frac{\widetilde{R}_t^{nr}}{U_t^s}, \qquad \text{for } = cbi, kb \tag{2.43}$$

$$U_t^s = \left(\frac{1}{\kappa} \frac{\widetilde{R}_t^{nr,s}}{\widetilde{P}_t^{kb}} \right)^{\frac{1}{\psi}}, \qquad \text{for } = cbi, kb \tag{2.44}$$

$$\widetilde{K}_{t+1}^{nr} = (1-\delta^{nr}) \frac{\widetilde{K}_t^{nr}}{\Gamma_t^{x,kb}} + A_t^{nr} \widetilde{E}_t^{nr} - \frac{100\chi^{nr}}{2} \left(\frac{\widetilde{E}_t^{nr} - \widetilde{E}_{t-1}^{nr}}{\widetilde{K}_t^{nr}} \Gamma_t^{x,kb} \right)^2 \frac{\widetilde{K}_t^{nr}}{\Gamma_t^{x,kb}} \tag{2.45}$$

$$\widetilde{K}_t^{nr} = \widetilde{K}_t^{nr,cbi} + \widetilde{K}_t^{nr,kb}. \tag{2.46}$$

U_t^s is the capital utilization rate, κ is a scaling parameter for the cost of changing the capacity utilization rate and ψ is the elasticity of the capacity utilization cost. $\widetilde{R}_t^{nr,cbi}$ and $\widetilde{R}_t^{nr,kb}$ denote the nominal rental rate on non-residential capital used in the cbi and kb sector denoted by $\widetilde{K}_t^{nr,cbi}$ and $\widetilde{K}_t^{nr,kb}$, respectively.

The first order conditions for the consumer durable capital owners and residential capital owners are similar. As these types of capital are not used in the production process, there are only three first order conditions for each capital owner. The only difference between the two types of capital is that the consumer durable capital good is produced in the fast growing (kb) sector and the residential capital good is produced in the slow growing (cbi) sector:

$$
\begin{aligned}
\widetilde{Q}_t^{cd} &= \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \frac{1}{\Gamma_{t+1}^{x,kb}} \left(\widetilde{R}_{t+1}^{cd} + (1-\delta^{cd}) \widetilde{Q}_{t+1}^{cd} \right) \right\} && (2.47) \\
\widetilde{P}_t^{kb} &= \widetilde{Q}_t^{cd} \left[A_t^{cd} - 100\chi^{cd} \left(\frac{\widetilde{E}_t^{cd} - \widetilde{E}_{t-1}^{cd}}{\widetilde{K}_t^{cd}} \Gamma_t^{x,kb} \right) \right] && (2.48) \\
&+ \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \widetilde{Q}_{t+1}^{cd} 100\chi^{cd} \left(\frac{\widetilde{E}_{t+1}^{cd} - \widetilde{E}_t^{cd}}{\widetilde{K}_{t+1}^{cd}} \Gamma_{t+1}^{x,kb} \right) \right\} \\
\widetilde{K}_{t+1}^{cd} &= (1-\delta^{cd}) \frac{\widetilde{K}_t^{cd}}{\Gamma_t^{x,kb}} + A_t^{cd} \widetilde{E}_t^{cd} - \frac{100\chi^{cd}}{2} \left(\frac{\widetilde{E}_t^{cd} - \widetilde{E}_{t-1}^{cd}}{\widetilde{K}_t^{cd}} \Gamma_t^{x,kb} \right)^2 \frac{\widetilde{K}_t^{cd}}{\Gamma_t^{x,kb}} && (2.49)
\end{aligned}
$$

and

$$
\begin{aligned}
\widetilde{Q}_t^{r} &= \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \frac{1}{\Gamma_{t+1}^{x,cbi}} \left(\widetilde{R}_{t+1}^{r} + (1-\delta^{r}) \widetilde{Q}_{t+1}^{r} \right) \right\} && (2.50) \\
\widetilde{P}_t^{cbi} &= \widetilde{Q}_t^{r} \left[A_t^{r} - 100\chi^{r} \left(\frac{\widetilde{E}_t^{r} - \widetilde{E}_{t-1}^{r}}{\widetilde{K}_t^{r}} \Gamma_t^{x,cbi} \right) \right] && (2.51) \\
&+ \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} \widetilde{Q}_{t+1}^{r} 100\chi^{r} \left(\frac{\widetilde{E}_{t+1}^{r} - \widetilde{E}_t^{r}}{\widetilde{K}_{t+1}^{r}} \Gamma_{t+1}^{x,cbi} \right) \right\} \\
\widetilde{K}_{t+1}^{r} &= (1-\delta^{r}) \frac{\widetilde{K}_t^{r}}{\Gamma_t^{x,cbi}} + A_t^{r} \widetilde{E}_t^{r} - \frac{100\chi^{r}}{2} \left(\frac{\widetilde{E}_t^{r} - \widetilde{E}_{t-1}^{r}}{\widetilde{K}_t^{r}} \Gamma_t^{x,cbi} \right)^2 \frac{\widetilde{K}_t^{r}}{\Gamma_t^{x,cbi}} && (2.52)
\end{aligned}
$$

The variable definitions are the same as for non-residential capital (nr) and the capital type is denoted by cd for consumer durable capital and r for residential capital.

A representative household derives utility from consumer non-durable goods and non-housing services, $\widetilde{E}_t^{cnn}$, the flow of services from consumer-durable capital, $\widetilde{K}_t^{cd}$, the flow of services from residential capital, $\widetilde{K}_t^{r}$ and leisure implicitly defined by hours worked in the two sectors, $L_t^{cbi} + L_t^{kb}$. Utility is influenced by a habit stock of each component scaled by the parameters h^{cnn}, h^{cd} and h^{r}. There are stochastic preference shocks to the different components denoted by Ξ_t^{cnn}, Ξ_t^{cd}, Ξ_t^{r} and Ξ_t^{l}. Households maximize utility and are monopolistic suppliers of labor. The household's budget constraint incorporates wage income, capital income, expenditure on consumption, rental payments on durable capital and residential capital, wage setting adjustment costs (depend on the parameter χ^w and the lagged and steady-state wage inflation rate) and costs in altering the composition of labor supply. Utility maximization and wage setting are constrained by the household's budget and the demand curve for the household's differentiated labor. The household's first order conditions are given by:

$$\widetilde{\Lambda}_t^{cnn} = \beta R_t E_t \left\{ \widetilde{\Lambda}_t^{cnn} \frac{1}{\Pi_{t+1}^{p,cbi} \Gamma_{t+1}^{x,cbi}} \right\} \tag{2.53}$$

$$\widetilde{\Lambda}_t^{cnn} = \widetilde{\Lambda}_t^{cd} \frac{1}{\widetilde{R}_t^{cd}} \tag{2.54}$$

$$\widetilde{\Lambda}_t^{cnn} = \widetilde{\Lambda}_t^{r} \frac{1}{\widetilde{R}_t^{r}} \tag{2.55}$$

$$\widetilde{\Lambda}_t^{cnn} = \varsigma^{cnn} \frac{\Xi_t^{cnn}}{\widetilde{E}_t^{cnn} - (h^{cnn}/\Gamma_t^{x,cbi}) \widetilde{E}_{t-1}^{cnn}} - \beta \varsigma^{cnn} E_t \left\{ \frac{(h^{cnn}/\Gamma_{t+1}^{x,cbi}) \Xi_{t+1}^{cnn}}{\widetilde{E}_t^{cnn} - (h^{cnn}/\Gamma_{t+1}^{x,cbi}) \widetilde{E}_t^{cnn}} \right\} \tag{2.56}$$

$$\frac{\widetilde{\Lambda}_t^{cd}}{\Gamma_t^{x,kb}} = \varsigma^{cd} \frac{\Xi_t^{cd}}{\widetilde{K}_t^{cd} - (h^{cd}/\Gamma_{t-1}^{x,kb}) \widetilde{K}_{t-1}^{cd}} - \beta \varsigma^{cd} E_t \left\{ \frac{(h^{cd}/\Gamma_t^{x,kb}) \Xi_{t+1}^{cd}}{\widetilde{K}_{t+1}^{cd} - (h^{cd}/\Gamma_t^{x,kb}) \widetilde{K}_t^{cd}} \right\} \tag{2.57}$$

$$\frac{\widetilde{\Lambda}_t^{r}}{\Gamma_t^{x,cbi}} = \varsigma^{r} \frac{\Xi_t^{r}}{\widetilde{K}_t^{r} - (h^{r}/\Gamma_{t-1}^{x,cbi}) \widetilde{K}_{t-1}^{r}} - \beta \varsigma^{r} E_t \left\{ \frac{(h^{r}/\Gamma_t^{x,cbi}) \Xi_{t+1}^{r}}{\widetilde{K}_{t+1}^{r} - (h^{r}/\Gamma_t^{x,cbi}) \widetilde{K}_t^{r}} \right\}, \tag{2.58}$$

where ς^{cnn}, ς^{cd}, ς^{r} and ς^{l} are scale parameters that tie down the ratios between the household's consumption components. $\widetilde{\Lambda}_t^{cnn}$, $\widetilde{\Lambda}_t^{cd}$ and $\widetilde{\Lambda}_t^{r}$ denote marginal utility of the different goods and R_t denotes the nominal interest rate.

The household's labor-supply decisions imply the following wage Phillips curves:

$$\begin{aligned}
\Theta_t^l \quad & \frac{\Lambda_t^{l,cbi}}{\widetilde{\Lambda}_t^{cnn}} L_t^{cbi} && (2.59)\\
= \quad & (\Theta_t^l - 1) \widetilde{W}_t^{cbi} L_t^{cbi} \\
- \quad & \Theta_t^l 100 \chi^l \left(\frac{L_*^{cbi}}{L_*^{cbi} + L_*^{kb}} \widetilde{W}_t^{cbi} + \frac{L_*^{kb}}{L_*^{cbi} + L_*^{kb}} \widetilde{W}_t^{kb} \right) \left(\frac{L_t^{cbi}}{L_t^{kb}} - \eta^l \frac{L_{t-1}^{cbi}}{L_{t-1}^{kb}} - (1 - \eta^l) \frac{L_*^{cbi}}{L_*^{kb}} \right) \\
+ \quad & 100 \chi^\omega \left(\Pi_t^{\omega,cbi} - \eta^\omega \Pi_{t-1}^{\omega,cbi} - (1 - \eta^\omega) \Pi_*^{\omega,cbi} \right) \Pi_t^{\omega,cbi} \widetilde{W}_t^{cbi} L_t^{cbi} \\
- \quad & \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} 100 \chi^\omega \left(\Pi_{t+1}^{\omega,cbi} - \eta^\omega \Pi_t^{\omega,cbi} - (1 - \eta^\omega) \Pi_*^{\omega,cbi} \right) \Pi_{t+1}^{\omega,cbi} \widetilde{W}_{t+1}^{cbi} L_{t+1}^{cbi} \right\}
\end{aligned}$$

and

$$\begin{aligned}
\Theta_t^l & \frac{\Lambda_t^{l,kb}}{\widetilde{\Lambda}_t^{cnn}} L_t^{kb} \qquad (2.60)\\
= & (\Theta_t^l - 1)\widetilde{W}_t^{kb} L_t^{kb} \\
+ & \Theta_t^l 100\chi^l \left(\frac{L_*^{cbi}}{L_*^{cbi} + L_*^{kb}} \widetilde{W}_t^{cbi} + \frac{L_*^{kb}}{L_*^{cbi} + L_*^{kb}} \widetilde{W}_t^{kb} \right) \left(\frac{L_t^{cbi}}{L_t^{kb}} - \eta^l \frac{L_{t-1}^{cbi}}{L_{t-1}^{kb}} - (1-\eta^l) \frac{L_*^{cbi}}{L_*^{kb}} \right) \\
+ & 100\chi^\omega \left(\Pi_t^{\omega,kb} - \eta^\omega \Pi_{t-1}^{\omega,kb} - (1-\eta^\omega)\Pi_*^{\omega,kb} \right) \Pi_t^{\omega,kb} \widetilde{W}_t^{kb} L_t^{kb} \\
- & \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} 100\chi^\omega \left(\Pi_{t+1}^{\omega,kb} - \eta^\omega \Pi_t^{\omega,kb} - (1-\eta^\omega)\Pi_*^{\omega,kb} \right) \Pi_{t+1}^{\omega,kb} \widetilde{W}_{t+1}^{kb} L_{t+1}^{kb} \right\}.
\end{aligned}$$

Θ_t^l denotes the elasticity of substitution between the differentiated labor inputs into production, $\Lambda_t^{l,s}$ denotes the marginal disutility of supplying labor in the two sectors, $\widetilde{W}_t^s$ denotes the nominal wage rates and $\Pi_t^{\omega,s}$ denotes the wage inflation rates. The parameter χ^l reflects the size of adjustment costs of altering the labor supply and χ^ω the size of adjustment costs in re-setting wages. η^l determines the importance of the lagged sectoral mix of labor relative to its steady state value in the labor composition adjustment costs. η^ω determines the importance of the lagged wage inflation rate relative to its steady state value in the wage adjustment cost function.

Additionally, there are market clearing conditions and some definitional equations, for example, regarding GDP growth H_t^{gdp} and GDP deflator inflation $\Pi_t^{p,gdp}$. Finally the model is closed with a monetary policy reaction function. The nominal interest rate R_t is adjusted gradually to the central bank's target interest rate $\bar{R}_t$:

$$R_t = (R_{t-1})^{\phi^r} (\bar{R}_t)^{(1-\phi^r)} \exp[\epsilon_t^r] \qquad (2.61)$$

$$\bar{R}_t = \left(\Pi_t^{p,gdp}/\Pi_*^{p,gdp}\right)^{\phi^{\pi,gdp}} \left(\Delta\Pi_t^{p,gdp}\right)^{\phi^{\Delta\pi,gdp}} \left(H_t^{gdp}/H_*^{gdp}\right)^{\phi^{h,gdp}} \left(\Delta H_t^{gdp}\right)^{\phi^{\Delta h,gdp}} R_*. \qquad (2.62)$$

ϵ_t^r is a monetary policy shock. ϕ^r, $\phi^{\pi,gdp}$, $\phi^{\Delta\pi,gdp}$, $\phi^{h,gdp}$ and $\phi^{\Delta h,gdp}$ denote policy response parameters and R_* the steady state interest rate.

A.2 The quarterly vintage database

This appendix describes the data series and the data sources for the quarterly data vintages that form the basis of the quarterly real-time re-estimation of macroeconomic models over the business cycle in this chapter.

All models are estimated using quarterly real-time data for real output, the output deflator and the effective federal funds rate. For the Christiano-Eichenbaum-Evans/Smets-Wouters model we use in addition real-time data for consumption, investment, hours and wages. The estimation of the model Edge et al. (2007) additionally requires data for consumption of non-durable goods and services, consumption of durable goods, residential investment, nonresidential investment, hours, wages, inflation for consumer nondurable goods and services and inflation for consumer durable goods. All time series are obtained from the Federal Reserve Bank of St. Louis' Alfred database except for hours and wages. For the 1980s and 1990s recessions we use data on aggregate weekly hours and employee compensation per hour from Faust and Wright (2009). For the 2001 and 2009 recessions we use the average weekly hours and the hourly compensation time series as in Smets and Wouters (2007) which we obtain from the Alfred database.

Consumption, investment and wages are expressed in real terms through division with the output deflator. Inflation is computed as the first difference of the log output deflator. The interest rate is expressed on a quarterly basis. Output, consumption and investment are expressed per capita by division with the civilian noninstitutional population over 16. For the 1980s and 1990s we obtain annual realtime population data from the Statistical Abstract of the United States.[17] We assume a constant population growth rate within one year to construct quarterly data. For the 2001 and 2009 recessions quarterly real-time population data is available from the Alfred database.

For the 1980s and 1990s recessions we compute hours per capita by dividing aggregate hours with civilian employment (16 years and older). Realtime employment data is obtained from the Alfred database. The hours per capita series is also influenced by low frequency movements in government employment, schooling and the aging of the population that cannot be captured by the macroeconomic models. Thus, we we follow Francis and Ramey (1995) and remove these trends by computing deviations of the hours per capita series using the HP filter with a weight of 16000 (compared to the standard weight of 1600 used for business-cycle frequency de-trending). The real-time character of the data is not affected by this procedure. For the 2001 and 2009 recessions average weekly hours are multiplied with the civilian employment (16 years and older) as in Smets and Wouters (2007) to take into account the limited coverage of the nonfarm business sector compared to GDP. Finally, this hours series is expressed per capita by dividing with the population over 16.

Output, consumption, investment, wages and hours are expressed in 100 times the logarithm. Growth rates are computed as the first difference of output, consumption, investment and wages. For the FRB/EDO model we use nominal time series except for output. Inflation of nondurables and services prices and durable consumer goods prices is computed by dividing the relevant nominal and real time series.

In the forecasting exercises, per capita output growth forecasts are converted into aggregate forecasts by assuming that the average quarterly population growth of the last two years holds in the future. All data and forecasts of output growth and inflation coincide with the definition of official annualized quarterly series as we remove rounding errors of the log expressions used for the estimation of the models.

[17] Scanned documents are available as .pdf files on http://www.census.gov/prod/www/abs/statab.html

Chapter 3

Forecasting under Model Uncertainty

Abstract This chapter investigates the accuracy of point and density forecasts of four dynamic stochastic general equilibrium (DSGE) models for output growth, inflation and the interest rate. The model parameters are estimated and forecasts are derived successively from historical U.S. data vintages synchronized with the Fed's Greenbook projections. In addition, I compute weighted forecasts using simple combination schemes as well as likelihood based methods. While forecasts from structural models fail to forecast large recessions and booms, they are quite accurate during normal times. Model forecasts compare particularly well to nonstructural forecasts and to Greenbook projections for horizons of three quarters ahead and higher. Weighted forecasts are more precise than forecasts from single models. A simple average of forecasts yields an accuracy comparable to the one obtained with state of the art time series methods that can incorporate large datasets. Comparing density forecasts of DSGE models with the actual distribution of observations shows that the models overestimate uncertainty around point forecasts.

Keywords: forecasting, model uncertainty, density forecasts, business cycle models

JEL-Codes: C53, E31, E32, E37

3.1 Introduction

For a long time business cycle models with microeconomic foundations have been calibrated and used for policy simulations while atheoretical time series methods have been used to forecast macroeconomic variables. Recently, several researchers have shown that estimated DSGE models can generate forecasts of reasonable accuracy (Smets and Wouters, 2004; Adolfson et al., 2005; Smets and Wouters, 2007; Edge et al., 2009; Wang, 2009; Christoffel et al., 2010). While these studies analyse only one model at a time, Wieland and Wolters (2010) compute forecasts from several theory based models for the five most recent U.S. recessions. The advantage of using structural models is that an economically meaningful interpretation of the forecasts can be given. While the forecasting accuracy of structural models is interesting on its own, it is also a test to which extent this class of models explains real world business cycle dynamics. A thorough assessment of different structural models including a comparison to forecasts from sophisticated time series models and to professional forecasts has not been undertaken yet. Recent comparison studies of state of the art forecasting methods have been restricted to nonstructural econometric methods (c.f. Stock and Watson, 2002; Bernanke and Boivin, 2003; Forni et al., 2003; Marcellino et al., 2003; Faust and Wright, 2009; Hsiao and Wan, 2010).

In this chapter, I carry out a detailed assessment of the forecasting accuracy of a suite of structural models. I use the same sample and real-time dataset as Faust and Wright (2009) who assess the forecasting accuracy of eleven nonstructural models. Therefore, my results are directly comparable to the forecasts from these models. The dataset is perfectly synchronized with the Greenbook and thus the results can also be compared to a best practice benchmark given by the Greenbook projections. The Greenbook projections are computed by the Federal Reserve's staff before each FOMC meeting and have been found to dominate forecasts from other professional forecasters in terms of forecasting accuracy (Romer and Romer, 2000; Sims, 2002; Bernanke and Boivin, 2003). The dataset includes data vintages for 145 FOMC meetings between March 1980 and December 2000.

I consider models that cover to some extent the range of closed-economy DSGE models used in academia and at policy institutions. The first model is a purely forward looking small-scale New Keynesian model with sticky prices that is analysed in detail in Woodford (2003). The second model by Fuhrer (1997) has a backward looking demand side, while the Phillips curve is derived from overlapping wage contracts. The third model is a medium-scale New Keynesian model as developed in Christiano et al. (2005). I use the estimated version by Smets and Wouters (2007). The fourth model is a version of the DSGE model by Edge et al. (2007) that features two production sectors with different technology growth rates and is itself an extension of the Christiano, Eichenbaum & Evans model. To determine how much of the forecasting accuracy of these four models is due to the theoretical foundations and what can be attributed to the parsimonious parametrization of these stylized models, I also consider a Bayesian VAR. It is a datadriven nonstructural counterpart to the four DSGE models with a comparably strict parametrization.

The parameters of the models are reestimated on three to eleven time series - as proposed by the original authors - for historical data vintages. Given this estimate, I compute a nowcast and forecasts up to five quarters into the future that take into account information that was actually available at the forecast start. Forecast precision is assessed relative to the revised data that became available during the subsequent quarters of the dates to which the forecasts apply.

Good forecasts are in general based on good forecasting methods and an accurate assessment of the current state of the economy. The Fed's great efforts to evaluate the current state of the economy are reflected in the accuracy of the Greenbook nowcasts. Sims (2002) suggests that this accurate data basis is a main reason for the precise Greenbook projections. The Fed's nowcasts exploit high frequency time series with more recent data than quarterly time series. In principle, there are methods available that allow the use of such data in combination with structural macroeconomic models (see Giannone et al., 2009). Employing such methods is beyond the scope of this chapter. To approximate the effect of using more information in nowcasting, I investigate the effect of using Greenbook nowcasts as a starting point for model-based forecasts by appending them to the actually available data. Thus, the potential informational advantage of the Fed about the current state of the economy is eliminated and a proper comparison of model forecasts with Greenbook projections is possible.

Timmermann (2006) surveys model averaging methods and finds that weighted forecasts from several nonstructural models outperform forecasts from individual models. Combining several models provides a hedge against model uncertainty when it is not possible to identify a single model that consistently dominates the forecasting accuracy of other models. Therefore, in addition to the individual model forecasts, I consider several simple and sophisticated model averaging schemes to compute weighted forecasts. For example, Gerard and Nimark (2008) and Bache et al. (2009) take into account forecasting uncertainty due to model uncertainty by combining forecasts from VARs and a single DSGE model. This chapter is an extension of their approach to a suite of theory based business cycle models.

The evaluation results of the point forecasts confirm the reasonable forecasting accuracy of DSGE models found in the above mentioned studies. The forecasting quality of the structural models is in particular competitive to the Greenbook projections for medium term horizons. For output growth, several models outperform the Greenbook projections and have an accuracy comparable to the best nonstructural models. Large scale models perform better than small scale models. However, quarterly output growth has little persistence and is thus difficult to forecast in general. Only one of the DSGE models gives more accurate forecasts than a simple univariate autoregressive process. The Greenbook inflation forecast is more accurate than all model forecasts. For the interest rate projections, the structural models perform worse than a Bayesian VAR probably due to the very simple monetery policy rules imposed in the models. The forecasts from the model by Smets and Wouters (2007) are in many cases more precise than forecasts from the other models. The model has a rich economic structure and is estimated on more variables than the standard New Keynesian models. Yet the parameterization is tight enough to yield accurate forecasts.

I find that weighted forecasts have a higher accuracy than forecasts from individual models. Combined forecasts based on simple weighting schemes that give significant weight to several models are superior to likelihood based weighting schemes that turn out to identify a single model rather than giving weight to several models. The forecasts of a simple average of the forecasts of all models are in many cases most accurate and otherwise only marginally less accurate than weighted forecasts from more sophisticated weighting methods.

While point forecasts are interesting, economists are concerned about the uncertainty surrounding these. Therefore, I derive density forecasts for the DSGE models that take into account parameter uncertainty and uncertainty about economic shocks in the future. I find that all the model forecasts overestimate actual uncertainty, i.e. density forecasts are very wide when compared with the actual distribution of data. A reason might be the tight restrictions imposed on the data. If the data rejects these restrictions, large shocks are needed to fit the models to the data resulting in high shock uncertainty (see also Gerard and Nimark, 2008). In a second step, I take into account model uncertainty and compute combined density forecasts using the same model averaging methods as for the point forecasts. This is similar to Gerard and Nimark (2008) who combine density forecasts of a DSGE model, a FAVAR model and a Bayesian VAR. Given the bad performance of individual models' density forecasts, it comes at no surprise that combined denstity forecasts overestimate uncertainty as well.

The remainder of this chapter proceeds as follows. Section 3.2 outlines the different macroeconomic models that are used to compute forecasts. Section 3.3 gives an overview of the dataset. Section 3.4 describes the estimation and forecasting methodology. Section 3.5 evaluates point forecasts from the individual models and compares them to Greenbook projections and nonstructural forecasts. Section 3.6 describes several model combination schemes. Section 3.7 provides a comparison of the accuracy of weighted forecasts, individual forecasts, Greenbook projections and nonstructural forecasts. Section 3.8 evaluates density forecasts of individual models and weighted models. Section 3.9 summarizes the findings and concludes.

3.2 Forecasting models

I consider five different models of the U.S. economy. Four are structural New Keynesian macroeconomic models and one model is a Bayesian VAR. The latter is representative of simple vector autoregression models that are often used to summarize macroeconomic dynamics without imposing strong theoretical restrictions. It is thus the unrestricted counterpart of the three variables output growth, inflation and the federal funds rate that are common to the four structural models. The models are chosen to broadly reflect the variety of DSGE models used in academia and at policy institutions.[1] I briefly describe the main features of the models. All models have been applied in Wieland and Wolters (2010) to compute point forecasts during the last five U.S. recessions.

Small New Keynesian Model estimated by Del Negro & Schorfheide (DS) The New Keynesian model is described, e.g., in Goodfriend and King (1997) and Rotemberg and Woodford (1997). It is often referenced to be the workhorse model in modern monetary economics and a comprehensive

[1]A comparison to large scale econometric models in the tradition of the Cowles Commission is unfortunately more burdensome. Fair (2007) compares the forecasting accuracy of a large econometric model to a DSGE model by Del Negro et al. (2007).

analysis is presented in the monograph of Woodford (2003). The model consists of three main equations: an IS curve, a monetary policy rule and a Phillips curve. The expectational IS curve can be derived from the behavior of optimizing and forward looking representative households that have rational expectations. Together with a monetary policy rule, it determines aggregate demand. The New Keynesian Phillips curve determines aggregate supply and can be derived from monopolistic firms that face sticky prices. Del Negro and Schorfheide (2004) use Bayesian estimation to fit the model to output growth, inflation and interest rate data. The methodology is reviewed in An and Schorfheide (2007). Wang (2009) shows that the small number of frictions is sufficient to provide reasonable output growth and inflation forecasts.

Small Model with Overlapping Wage Contracts by Fuhrer & Moore (FM) This is a small scale model of the U.S. economy described in Fuhrer (1997). It differs from the New Keynesian model with respect to the degree of forward lookingness and the specification of sticky prices. Aggregate demand is determined by a reduced form backward looking IS curve together with a monetary policy rule. Aggregate supply is modelled via overlapping wage contracts: agents care about real wage contracts relative to those negotiated in the recent past and those that are expected to be negotiated in the near future (see Fuhrer and Moore, 1995a,b). The aggregate price level is a constant mark-up over the aggregate wage rate. The resulting Phillips curve depends on current and past demand and expectations about future demand. Fuhrer (1997) uses maximum likelihood estimation to parameterize the model. In contrast to all other models in this chapter, variables are not defined in percentage deviations from the steady state. While a measurement equation is needed to link output growth via a trend growth rate to the data, inflation and the interest rate are directly defined in the model equations as in the data.

Medium Scale Model by Smets & Wouters (SW) The small New Keynesian model has been extended by Christiano et al. (2005) to fit a high fraction of U.S. business cycle dynamics. It is a closed economy model that incorporates physical capital in the production function and capital formation is endogenized. Labor supply is modelled explicitly. Nominal frictions include sticky prices and wages as well as inflation and wage indexation. Real frictions include consumption habit formation, investment adjustment costs and variable capital utilization. Smets and Wouters (2007) added nonseparable utility and fixed costs in production. They replaced the Dixit-Stiglitz aggregator with the aggregator by Kimball (1995) which leads to a non-constant elasticity of demand. The model includes equations for consumption, investment, price and wage setting as well as several identities. Smets and Wouters (2007) used Bayesian estimation with a complete set of structural shocks to fit the model to seven U.S. time series.

Medium Scale Model by Edge, Kiley & Laforte (FRB/EDO) The so-called FRB/EDO model by Edge et al. (2008) has been developed at the Federal Reserve and also builds on the work by Christiano et al. (2005). It features two production sectors, which differ with respect to the pace of technological progress. This structure can capture the different growth rates and relative prices observed in the data. Accordingly, the expenditure side is disaggregated as well. It is divided into business investment and three categories of household expenditure: consumption of non-durables and services, investment in durable goods and residential investment. The model is able to capture different cyclical properties in these four expenditure categories. As in the Smets & Wouters model all behavioral equations are derived in a completely consistent manner from the optimization problems of representative households and firms. The model is documented in Edge et al. (2007).[2] To estimate the model using Bayesian techniques, 14 structural shocks are added to the equations and the model is estimated on eleven time series.

Bayesian VAR (BVAR) In addition to the four structural models, I estimate a VAR on output growth, inflation and the federal funds rate using four lags. The VAR is a more general description of the data than the DSGE models as it imposes little restrictions on the data generating process. All variables are treated symmetrically and therefore the VAR incorporates no behavioral interpretations of parameters or equations. Unrestricted VARs are heavily overparametrized and therefore not suitable for forecasting. I therefore use a Minnesota prior (see Doan et al., 1984) to shrink the parameters towards zero. The Minnesota prior assumes that the vector of time series is well-described as a collection of independent random walks. I use growth rates or stationary time series and therefore put a prior assumption of a zero coefficient on the first lag of the dependent variable instead of a one. Therefore, all parameters are assumed to be normally distributed with mean zero. The prior variance of the parameters decreases with the lag length. While larger Bayesian VARs and specifications with the level of output and prices can potentially increase the accuracy of foreacsts, I estimate a version that uses the same variable definitions as the DSGE models. This can be helpful to disentangle the importance of theoretical foundations and a parsimonious parametrization for accurate forecasts.

Table 3.1 summarizes the most important features of the four structural models and the Bayesian VAR. The number of equations refers to all equations in a model taking into account shock processes, measurement equations and identities. For example the standard New Keynesian model consists of 3 structural equations, 2 shock processes (+1 iid shock) and 3 measurement equations. It is apparent that the size of the models differs a lot from each other. Furthermore, the number of estimated parameters per equation are different. The FRB/EDO model includes about one parameter per equation implying high cross equation restrictions. The authors added measurement errors to the model to fit it to 11 time series. The Fuhrer & Moore model in contrast has two parameters per equation. The number of parameters in the Bayesian VAR can vary from 3 shock variances to 39 parameters depending on the significance of the four lags of each variable in each of the three equations. The method of estimating the structural parameters also varies across the models: I adapt the methodology used by the original authors and use maximum likelihood estimation for the Fuhrer & Moore model while Bayesian estimation is used to estimate the other models.[3]

[2] My version is not able to replicate the figures in the documentation exactly, but is reasonably close.

[3] To be sure, I approximate maximum likelihood estimation by defining wide uniform priors for all parameters and use

Table 3.1: Model overview

Type	Eq.	Par.	Est. Par.	Observable Variables	Reference
Small-scale microfounded forward looking New Keynesian Model	8	13	13	3: output growth, inflation, interest rate	Del Negro and Schorfheide (2004)
Small-scale model with overlapping real wage contracts and a backward looking IS curve	10	20	19	3: output growth, inflation, interest rate	Fuhrer (1997)
Medium-scale DSGE-model with many nominal and real frictions as used by policy institutions	27	42	37	7: output growth, consumption growth, investment growth, inflation, wages, hours, interest rate	Smets and Wouters (2007)
Large-scale DSGE-model developed at the Federal Reserve. Two production sectors with different technology growth rates. The demand side is disaggregated into four categories	59	71	51	11: output growth, inflation, interest rate, consumption of non-durables and services, consumption of durables, residential investment, business investment, hours, wages, inflation for consumer nondurables and services, inflation for consumer durables	Edge et al. (2008)
Bayesian VAR with 4 lags; Minnesota priors	3	3-39	39	3: output growth, inflation, interest rate	Doan et al. (1984)

Notes: Type: short classification of the models according to the main modelling assumptions; Eq.: number of equations including shock processes, measurement equations and identities, but excluding variable definitions and flexible price allocations; Par.: total number of parameters in the model file excluding all auxiliary parameters; Est. Par.: exact number of estimated parameters including shock variances and covariances; Observable Variables: the number and names of the observable variables; Reference: original reference that is closest to the implemented version in this chapter.

For the priors, I use the ones in the original research referenced in Table 3.1. Except for the model by Fuhrer & Moore, variables are defined in percentage deviations from steady state and thus measurement equations that include an output growth trend and the steady state of inflation, the interest rate and other observables are needed to link the equations to the data. The FRB/EDO model is implemented nonlinearly and I derive a first order approximation of the solution. All other models are linearized.

3.3 A real-time dataset

I use the real-time dataset described in Faust and Wright (2009).[4] The dataset is prepared by the Federal Reserve staff to compute the Greenbook forecasts. The data is perfectly synchronized with the Greenbook and contains historical samples, i.e. data vintages, of 109 variables as observed at the time the Greenbook was published. In addition, it contains nowcasts and forecasts up to five quarters for all variables. The dataset contains data vintages for 145 FOMC meetings from March 1980

then the same Bayesian estimation algorithm as for the other models. Therefore, exactly the same statistics are derived for all models which is important for the computation of weighted forecasts in section 3.6.

[4]The dataset can be downloaded from the website of Jon Faust: http://e105.org/faustj/papgbts.php?d=n. A detailed data appendix is available on the same website.

to December 2000, while the different data series start in 1960.[5] While some of the nonstructural forecasting models considered in Faust and Wright (2009) can process as many data series as available, the structural models considered in this chapter use only a small subset of the available time series varying from three to eleven variables to estimate the different models. Still some variables for the FRB/EDO model are not available in the data set. Therefore I add the necessary real-time data series from the Federal Reserve Bank of St. Louis' Alfred database and also the accordant nowcasts from the Greenbook. To each data vintage I add only observations that would have been available at the Greenbook publication date.

There is a trade-off between using a long sample to get precise parameter estimates and for leaving out a fraction of past data that might contain structural breaks. Therefore, I use a moving window of the latest eighty quarterly observations of each data vintage to estimate the models. Aside from structural breaks the high inflation periods of the 70's and 80's influence the estimated inflation steady state which can bias the inflation forecasts of the late 80's and the 90's. Therefore a window of 80 observations gives at least the chance of a diminishing effect on the forecasts. The first sample for the FOMC meeting of March 1980 starts in 1960Q1 and ends in 1979Q4, the second sample for the FOMC meeting of April 1980 starts in 1960Q2 and ends in 1980Q1, and this goes on until the last sample for the FOMC meeting of December 2000 that starts in 1980Q4 and ends in 2000Q3.

I forecast annualized quarterly real output growth as measured by the GNP/GDP real growth rate, annualized quarterly inflation as measured by the GNP/GDP deflator and the federal funds rate. GDP data is first released about one month after the end of the quarter to which the data refer, the so-called advance release. These data series are then revised several times at the occasion of the preliminary release, final release, annual revisions and benchmark revisions. I follow Faust and Wright (2009) and use actual realized data as recorded in the data vintage that was released two quarters after the quarter to which the data refer to evaluate the forecasting accuracy. For example, revised data for 1999Q1 is obtained by selecting the entry for 1999Q1 from the data vintage released in 1999Q3. Hence, I do not attempt to forecast annual and benchmark revisions, because the models cannot predict changes in data definitions. The revised data against which the accuracy of forecasts is judged will typically correspond to the final NIPA release.

While the models by Del Negro & Schorfheide, Fuhrer & Moore and the Bayesian VAR are estimated on the three key variables - output growth, inflation and the federal funds rate - the other two models are fit to seven and eleven time series, respectively. The Smets & Wouters model is estimated on the three key variables and a wage time series, hours worked, consumption and investment. The FRB/EDO model is estimated on eleven empirical time series: output growth, inflation, the federal funds fate, consumption of non-durables and services, consumption of durables, residential investment, business investment, hours, wages, inflation for consumer nondurables and services and

[5] The dataset ends in 2000 because Greenbook data remains confidential for 5 years after the forecast date. I don't update the data for the additional years that are now available to make the forecasting results directly comparable to Faust and Wright (2009).

inflation for consumer durables.[6]

3.4 Forecasting methodology

Computing recursive forecasts using structural models and real-time data vintages requires a sequence of steps that are explained in the following. First, the models need to be specified, solved and linked to the empirical data. Second, the data needs to be updated to the current vintage and parameters have to be estimated. Third, density and point forecasts are computed.

Model specification and solution. Each of the models consists of a number of linear or nonlinear equations that determine the dynamics of the endogenous variables. A number of structural shocks is included in each model. Any of the models $m = 1, ..., 4$ can be written as follows:

$$E_t \left[f_m(y_t^m, y_{t+1}^m, y_{t-1}^m, \epsilon_t^m, \beta^m) \right] = 0 \tag{3.1}$$

$$E(\epsilon_t^m) = 0 \tag{3.2}$$

$$E(\epsilon_t^m \epsilon_t^{m\prime}) = \Sigma_\epsilon^m, \tag{3.3}$$

where $E_t \left[f_m(.) \right]$ is a system of expectational difference equations, y_t^m is a vector of endogenous variables, ϵ_t^m a vector of exogenous stochastic shocks, β^m a vector of parameters and Σ_ϵ^m is the variance-covariance matrix of the exogenous shocks. The parameters and the variance-covariance matrix are either calibrated or estimated or a mixture of both.

A subset of the endogenous variables consists of empirically observable variables $y_t^{m,obs}$. If variables in the models are defined in percentage deviations from steady state then there is a subset of the equations that are so-called measurement equations $f_m^{obs}(.)$. These link the observable variables to the other endogenous variables through the inclusion of steady state values or steady state growth rates. Another possibility is that the observable variables are directly included in the general equations of a model. The latter is the case in the Fuhrer & Moore model. Inflation and the interest rate are included in the model as they appear in the data and are not redefined as deviations from steady states. For the FRB/EDO model, it is assumed that not all observable variables are measured exactly and therefore a set of nonstructural measurement shocks is added to the measurement equations.

[6] Output is in real terms available in the data set and growth rates can be computed directly. Consumption, investment and wages are expressed in real terms as defined in the models through division with the output deflator. Growth rates are computed afterwards. Inflation is computed as the first difference of the log output deflator. The nominal interest rate is expressed on a quarterly basis. I compute hours per capita by dividing aggregate hours with civilian employment (16 years and older). The hours per capita series includes low frequent movements in government employment, schooling and the aging of the population that cannot be captured by the models. I remove these following Francis and Ramey (1995) by computing deviations of the hours per capita series from its low frequent HP-filtered trend with a parameter of 16000. The realtime characteristic of the data remains unaffected by this procedure. For the FRB/EDO model nominal time series except for output growth are used. Growth rates are computed for consumption of non-durables and services, consumption of durables, residential investment and business investment. Inflation of nondurables and services and inflation of durable goods is computed by dividing the accordant nominal and real time series and calculating log first differences.

The system of equations is solved using a conventional solution method for rational expectations models such as the technique of Blanchard and Kahn. In the case of the FRB/EDO model a first order approximation of the solution is derived. The other models are already linearized before solving them.[7] Given the solution, the following state space representation of the system is derived:

$$y_t^{m,obs} = \Gamma^m \bar{y}^m + \Gamma^m y_t^m + \epsilon_t^{m,obs}, \tag{3.4}$$

$$y_t^m = g_y^m(\beta^m) y_{t-1}^m + g_\epsilon^m(\beta^m)\epsilon_t^m, \tag{3.5}$$

$$E(\epsilon_t^m \epsilon_t^{m\prime}) = \Sigma_\epsilon^m \tag{3.6}$$

The first equation summarizes the measurement equations and shows the link between observable variables and the endogenous model variables via steady state values or deterministic trends $\bar{y}^m$. The matrix Γ^m might include lots of zero entries as not all variables are directly linked to observables. The measurement errors $\epsilon_t^{m,obs}$ are a subset of the shocks ϵ_t^m. The second equation constitutes the transition equations including the solution matrices g_y^m and g_ϵ^m that both are nonlinear functions of the structural parameters β^m. The transition equations relate the endogenous variables to their own lags and the vector of exogenous shocks. The third equation denotes the variance-covariance matrix Σ_ϵ^m.

Estimation. Having solved the model and linked to the data, one needs to update the data before estimating the model. I use for each forecast the 80 most recent observations of the respective historical data vintage that was available at the time of the forecast start. Estimating DSGE models using Bayesian estimation has become a popular approach due to the combination of economic theory which is imposed on the priors and data fit summed up in the posterior estimates. A survey of the methodology is presented in An and Schorfheide (2007). Therefore, I only give a short overview of the algorithm. I approximate maximum likelihood estimation with Bayesian estimation with wide uniform priors, so that exactly the same estimation algorithm is used. Due to the nonlinearity in β^m the calculation of the likelihood is not straightforward. The Kalman filter is applied to the state space representation to set up the likelihood function (see e.g. Hamilton, 1994, chapter 13.4).[8] Since the models considered are stationary, one can initialize the Kalman Filter using the unconditional distribution of the state variables. Combining the likelihood with the priors yields the log posterior kernel $ln\mathcal{L}(\beta^m|y_1^{m,obs},...,y_t^{m,obs}) + lnp(\beta^m)$ that is maximized over β^m using numerical methods to compute the posterior mode. The posterior distribution of the parameters is a complicated nonlinear function of the structural parameters. The Metropolis-Hastings algorithm offers an efficient method to derive the posterior distribution via simulation. Details are provided for example in Schorfheide (2000). I compute 500,000 draws from the Metropolis-Hastings algorithm and use the first 25,000 of these to calibrate the scale such that an acceptance ratio of 0.3 is achieved. Another 25,000 draws are disregarded as a burn in sample. The models are reestimated for the first data vintage of each year.

[7] I use the solution procedure of the Dynare software package. See www.dynare.org and Juillard (1996) for a description.

[8] I consider only unique stable solutions. If the Blanchard-Kahn conditions are violated I set the likelihood equal to zero.

Reestimating the models for all 145 available data vintages would be computationally too intensive. Finally, the mean parameters can be computed from the posterior distribution of β^m.

Forecast computation. Having estimated the different models, forecasts for the horizons $h \in (0, 1, 2, 3, 4, 5)$ are derived. First, a density forecast is computed and afterwards a point forecast is calculated as the mean of the density forecast. For each parameter a large number of values are drawn from the parameter's posterior distribution. For a random draw s a projection of the observable variables is derived by iterating over the solution matrix $g_y^m(\hat{\beta}^{m,s})$. At each iteration i in addition a vector of shocks $\epsilon_i^{m,s}$ is drawn from a mean zero normal distribution where the variance is itself a random draw from the posterior distribution of the variance-covariance matrix:

$$
\begin{aligned}
y_{t+h}^{s,m,obs} &= \Gamma^m \hat{\bar{y}}^{m,s} + \Gamma^m g_y^m(\hat{\beta}^{m,s})^{h+1} y_{t-1}^m + \Gamma^m \sum_{i=0}^{h} g_\epsilon^m(\hat{\beta}^{m,s})^{(h+1-i)} \epsilon_i^{m,s} && (3.7)\\
\epsilon_i^{m,s} &\sim N(0, \hat{\Sigma}_\epsilon^{m,s}), && (3.8)
\end{aligned}
$$

where a hat on the structural parameters $\beta^{m,s}$, the variance covariance matrix $\Sigma_\epsilon^{m,s}$ and the steady state values of observable variables $\bar{y}^{m,s}$ denotes that they are estimated. The reduced form solution matrices g_y^m and g_ϵ^m are functions of the estimated parameters and change over time as the models are reestimated. The procedure is repeated 10,000 times ($s = 1, ..., 10,000$) and finally the forecast density is given by the ordered set of forecast draws $y_{t+h}^{s,m,obs}$. The point forecast is given by the mean of the forecast density.

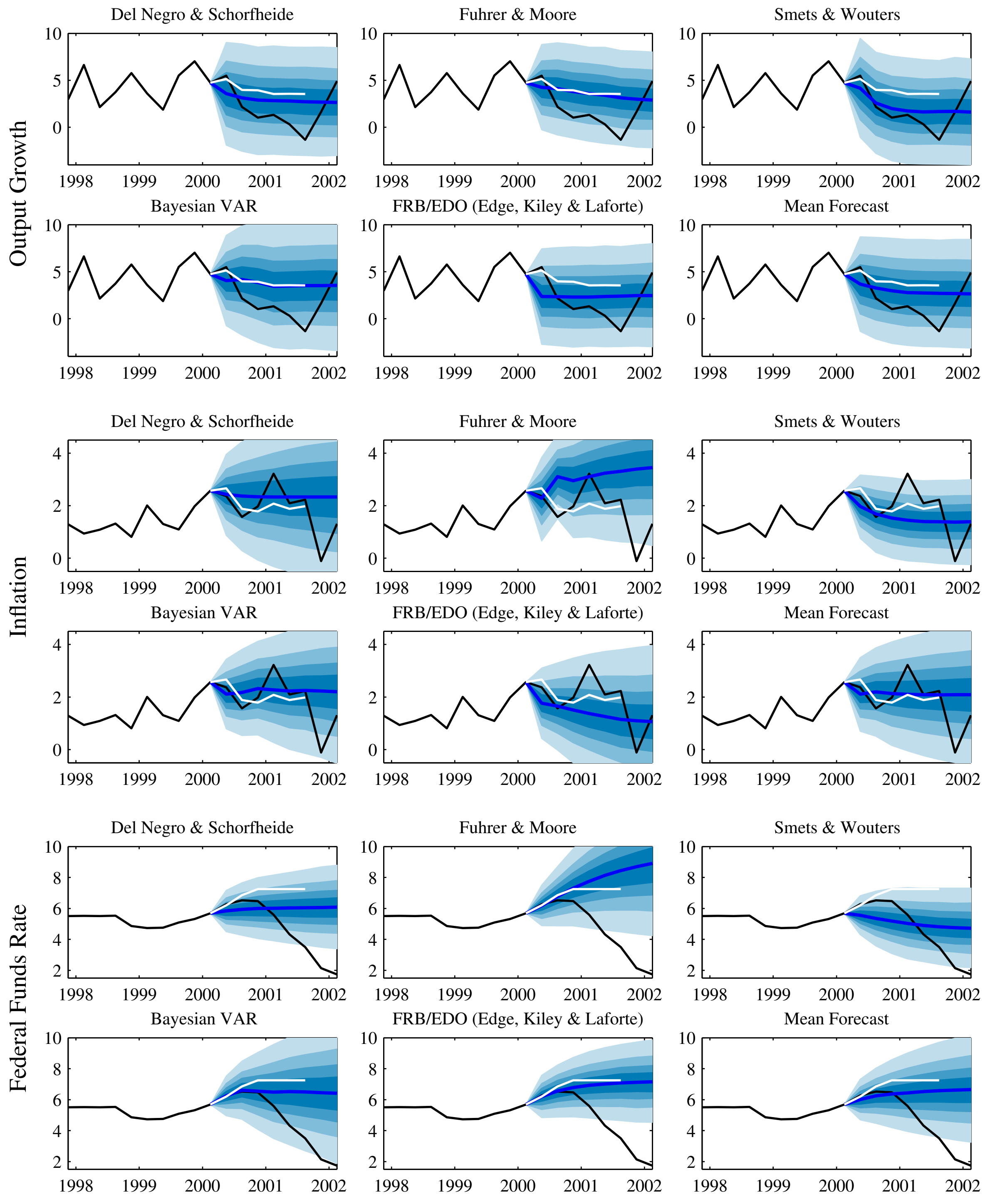
The different steps to compute forecasts are:

1. Model specification: set up a file with the model equations and add measurement equations that link the model to the empirical time series.
2. Solution: solve the model and express it in state space form.
3. Data update: update the data with the current vintage.
4. Estimation: reestimate the model for the first data vintage of each year. Otherwise, use the posterior distribution of the parameters from previous estimation. Add a prior distribution of the model parameters. Estimate the structural parameters by maximizing the posterior kernel. Afterwards simulate the posterior distribution of the parameters using the Metropolis-Hastings algorithm.
5. Density forecast: compute forecast draws by iterating over the solution matrices for different parameter values drawn from the posterior distribution. At each iteration draw a vector of shocks from a mean zero normal distribution with the variance itself being a draw from the posterior distribution. The forecast density is given by the ordered forecast draws.
6. Mean forecast: compute the mean of the forecast density to get the point forecast.
7. Repeat steps 3 to 6 for all data vintages.

8. Repeat steps 1 to 7 for different models, possibly extending the information set by additional variables as required by the respective model.

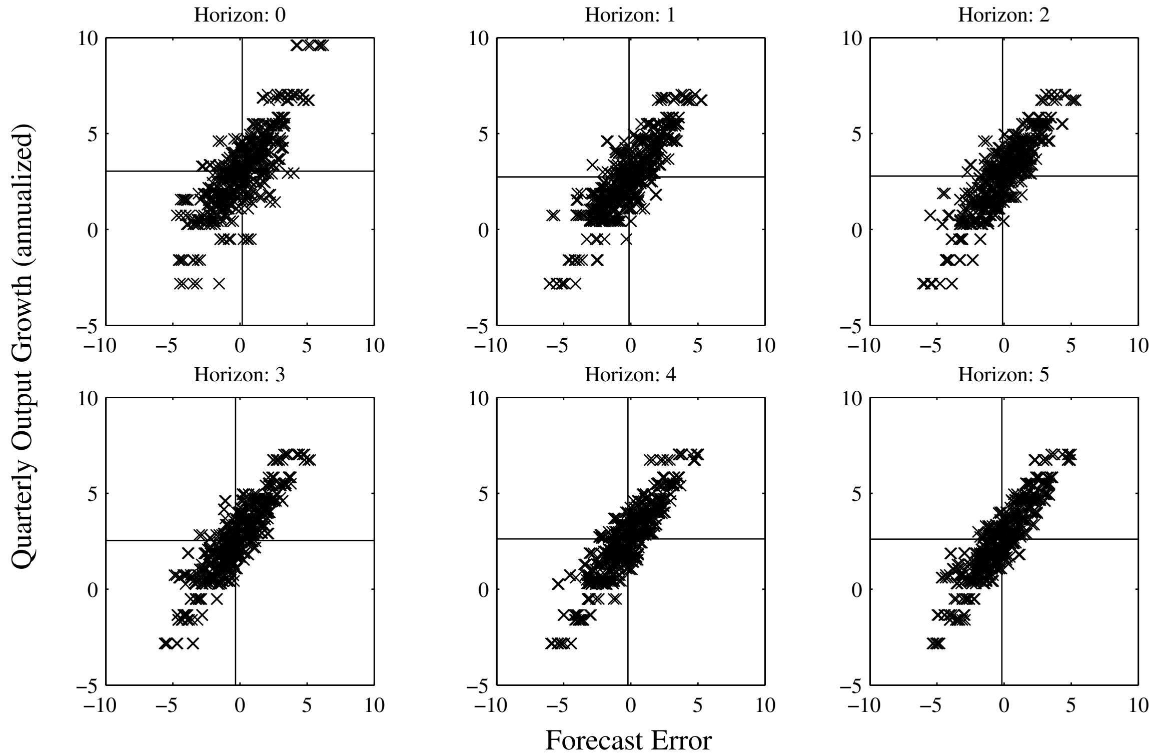
Figure 3.1 shows as an example forecasts for output growth, inflation and the federal funds rate derived from data vintage May 12, 2000. The black line shows real-time data until the forecast start and revised data afterwards. I plot the 0.05, 0.15, 0.25, 0.35 and 0.65, 0.75, 0.85 and 0.95 percentiles to graphically represent the density forecasts. The different shades therefore show for 90%, 70%, 50% and 30% probability bands. The line in the middle of the confidence bands shows the mean forecast for each model. The short white line shows the correspondent Greenbook projections. Data is available until the first quarter of 2000. The current state of the economy in the second quarter of 2000 is estimated using the different models. The economy was in a boom in early 2000 and the models broadly predict the return to average growth rates over the next quarters. They are not able to predict the 2001 recession that has been defined by the NBER to take place between the first and the fourth quarter of 2001. Inflation is predicted by the Del Negro & Schorfheide model and the Bayesian VAR to stay on a similar level as in the first quarter of 2000. The Fuhrer & Moore model predicts an increase of the inflation rate. The FRB/EDO and the Smets & Wouters models are able to predict the inflation decrease in the third quarter of 2000. None of the models is able to predict the short inflation increase in the first quarter of 2001. The interest rate is forecast to increase by the Fuhrer & Moore model, the FRB/EDO model and the Bayesian VAR. It is predicted to stay constant by the Del Negro & Schorfheide model and to decrease by the Smets & Wouters model. The average of the five forecasts predicts the interest rate path quite precisely until the end of the year. The decrease in the federal funds rate beginning in 2001 is not captured by the forecasts. This is consistent with the output growth forecasts that miss the recession in 2001 that is in turn a reason for the interest rate cuts.

Figure 3.1: Structural forecasts: data vintage May 12, 2000



Notes: the black line shows real-time data until the forecast start and revised data afterwards; the shaded areas show 90% 70%, 50% and 30% probability bands; the line in the middle of the probability bands shows the mean forecast for horizons 0 to 7; the short white line shows the Greenbook forecast for horizons 0 to 5. The graph labeled "Mean Forecast" shows the average of the four model forecasts and the Bayesian VAR forecast.

Figure 3.2: Forecast errors and output growth rates



Notes: the figure shows observed output growth rates and the corresponding forecast errors of the four DSGE models and the Bayesian VAR for different forecasting horizons. The horizontal lines show the mean output growth rate and the vertical line the mean forecast errors of all models for each horizon.

I plot a figure like this for the forecasts derived from each data vintage. Unfortunately, it is not possible to show all these figures in this chapter. However, screening over all the forecasts for the different historical data vintages reveals some notable observations. Structural models and the Bayesian VAR are well suited to forecast during normal times. Given small or average exogenous shocks the models give a good view about how the economy will return back to steady state. In contrast, large recessions or booms and the respective turning points are impossible to forecast with these models. Figure 3.2 plots the forecast errors (outcome minus forecast) of all models on the horizontal axis and the correponding realized output growth rate on the vertical axis. A clear positive relation is visible. When output growth is highly negative the models are not able to forecast such a sharp downturn and thus the forecast errors are negative. The models require large exogenous shocks to capture large deviations from the balanced growth path and the steady state inflation and interest rate. This is due to the weak internal propagation mechanism of the models. Therefore for a given shock all the models including the Bayesian VAR predict a quick return back to the steady state growth rate. Even if one of the models would imply more persistence, it is unlikely to capture the length of recessions accurately as these are rare events with few data points so that their implied persistence cannot be captured precisely when estimating a model. Each recessions might be caused by different exogenous reasons and therefore there is no information in previous data samples that can be used to forecast the length of such a recession in the future. While the point forecasts cannot predict a recession, the possibility that a large deviation from steady state values occurs is captured by the wide confidence bands. Once the turning point of a recession has been reached, all models predict the economic recovery back to the balanced growth path well. Recoveries in this data sample are quick with little persistence just like the internal propagation mechanism of the models used in this chapter.

3.5 Forecast evaluation

Table 3.2 reports the root mean squared prediction errors (RMSE) for output growth, inflation and interest rate forecasts from the Greenbook, the four structural models, the Bayesian VAR and the respective best and worst performing nonstructural model considered by Faust and Wright (2009). The first column gives the RMSE for the Greenbook and all other columns report the RMSE of the specific models relative to the Greenbook RMSE. Values less than one show that a model forecast is more accurate than the corresponding Greenbook projection. The last two columns report the relative RMSEs of the most and the least accurate nonstructural forecasting model from Faust and Wright (2009) for each horizon.

The first six rows in each table show forecasts based on the available data at the starting point of the forecast. The current state of the economy is not available in the data and therefore needs to be forecast. This nowcast is labeled as a forecast for horizon zero. As the data becomes available with a lag of one quarter, the results are labeled as "jump off -1". 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, opinion surveys, interest rates and other financial prices. This data can be used to improve nowcasts and the Federal Reserve staff and many professional forecasters certainly make use of it. To approximate the effect of using more information in nowcasting, I investigate the effect of using Greenbook nowcasts as a starting point for model-based forecasts regarding future quarters. The results are shown in the last five rows of each table and are labeled as "jump off 0".

I follow Faust and Wright (2009) in leaving out the period from 1980-1983 from the evaluation as this period was very volatile and might bias the assessment of forecasting accuracy for the whole sample. Therefore, the results start in 1984 so that the RMSEs for output growth and inflation are directly comparable to Table 2 in Faust and Wright (2009). The reported RMSEs are thus based on 122 forecasts from 1984 to 2000. I evaluate whether the difference of Greenbook RMSEs and model RMSEs is statistically significant based on the Diebold-Mariano statistic (Diebold and Mariano, 1995) using a symmetric loss function. Asymptotic p-values are computed using Newey-West standard errors with a lag-length of 10, covering a bit more than a year, to account for serial correlation of forecast errors.

Table 3.2: Greenbook RMSE and relative RMSE of model forecasts: 1984-2000

(a) Output growth

horizon	GB	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1				
0	1.75	1.20	1.13	1.24	1.21	1.11	1.09	1.39
1	2.12	**0.95**	1.05	**0.91**	**0.91**	**0.97**	**0.86**	1.20
2	2.01	1.06	1.10	**0.93**	1.00	**0.96**	**0.95**	1.15
3	2.15	**0.99**	1.09	**0.86·**	**0.95**	**0.97**	**0.94**	1.12
4	2.08	1.01	1.05	**0.89**	**0.94**	**0.94**	**0.99**	1.11
5	2.08	1.02	1.05	**0.90**	**0.99**	1.00	**0.97**	1.09
				jump off 0				
1	2.12	**0.95**	1.03	**0.93**	**0.94**	**0.94**	**0.84**	1.07
2	2.01	1.06	1.13	**0.94**	1.00	**0.97**	**0.90**	1.12
3	2.15	1.00	1.12	**0.87**	**0.97**	**0.96**	**0.95**	1.18
4	2.08	1.01	1.08	**0.88**	**0.97**	**0.97**	**0.96**	1.09
5	2.08	1.03	1.06	**0.89·**	1.01	**0.99**	**0.98**	1.11

(b) Inflation

horizon	GB	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1				
0	0.69	1.52●	1.86●	1.48●	1.65●	1.47●	1.34●	1.63●
1	0.79	1.59●	1.80●	1.44●	1.50●	1.45●	1.22●	1.86●
2	0.81	1.38●	1.57●	1.29●	1.59●	1.30●	1.15·	1.92●
3	0.93	1.17•	1.42•	1.20•	1.50●	1.14	1.03	1.84●
4	0.89	1.28•	1.80●	1.29●	1.46●	1.35•	1.08	2.11●
5	1.14	1.24•	1.62●	1.24●	1.33●	1.30	**0.99**	1.83●
				jump off 0				
1	0.79	1.24•	1.61●	1.15•	1.17•	1.25●	1.20●	1.58●
2	0.81	1.25●	1.50●	1.18·	1.16·	1.25●	1.18	1.69●
3	0.93	1.24●	1.27•	1.22●	1.27•	1.15·	1.04	1.66●
4	0.89	1.19●	1.51●	1.20●	1.26•	1.19	1.05	1.91●
5	1.14	1.15•	1.47•	1.21●	1.14	1.19	**0.97**	1.77●

(c) Federal Funds Rate

horizon	GB	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1				
0	0.11	5.91●	4.84●	4.63●	5.98●	3.57●	-	-
1	0.49	2.13●	1.88●	1.89●	2.39●	1.55●	-	-
2	0.90	1.49•	1.46●	1.37•	1.75●	1.18	-	-
3	1.25	1.19	1.25·	1.10	1.53●	1.01	-	-
4	1.60	1.05	1.22	**0.97**	1.40●	**0.96**	-	-
5	1.90	**0.97**	1.23·	**0.87**	1.29•	**0.92**	-	-
				jump off 0				
1	0.49	1.37•	1.30●	1.19·	1.66●	1.06	-	-
2	0.90	1.18	1.08	1.07	1.53●	**0.96**	-	-
3	1.25	1.02	1.01	**0.95**	1.45●	**0.90**	-	-
4	1.60	**0.95**	1.03	**0.89**	1.38●	**0.88**	-	-
5	1.90	**0.90**	1.08	**0.83**	1.31●	**0.86**	-	-

Notes: GB: Greenbook; DS: Del Negro & Schorfheide; FM: Fuhrer & Moore; SW: Smets & Wouters; EDO: FRB/EDO Model by Edge, Kiley & Laforte; BVAR: Bayesian VAR; Best FW: Best performing atheoretical model for the specific horizon considered by Faust & Wright; Worst FW: Worst performing atheoretical model for the specific horizon considered by Faust & Wright. The first column shows the forecast horizon. The second column shows the RMSE for the Greenbook. The other columns show RMSEs of alternative models relative to the Greenbook. Values less than one are in bold and show that a forecast is more accurate than the one by the Greenbook. The symbols ●, •, ·, indicate that the relative RMSE is significantly different from one at the 1, 5, or 10% level, respectively.

The results for inflation, output growth and the federal funds rate are very different. For output growth the Greenbook nowcast is more precise than the model nowcasts. This was expected as the Fed can exploit more information about the current state of the economy. However, this precise estimate of the current state of the economy does not translate into a superior forecasting performance at higher horizons. The SW, EDO and BVAR models' forecasts dominate the Greenbook forecast from horizon 1 onwards. The DS model yields a similar forecasting accuracy as the Greenbook. Only the FM model is slightly less accurate than the Greenbook forecast for all horizons. If I include the Greenbook nowcast in the information set used to compute forecasts the results hardly change as quarterly output growth is not very persistent. Viewing the Greenbook as a best practice benchmark, one could be tempted to judge the forecasting ability of the structural models as very good. However, one should keep in mind that quarterly output growth has little persistence and thus is difficult to forecast in general. The reported RMSEs in Faust and Wright (2009) show that none of their nonstructural forecasting methods is more accurate than an univariate autoregressive forecast.[9] I find that only the SW model's forecasts are more precise than an autoregressive forecast from horizon 2 onwards. The forecasting accuracy of the EDO and BVAR model is similar to the autoregressive forecast and the DS and FM forecasts are less precise. In addition, none of the models RMSEs differs statistically significant from the Greenbook RMSE with the SW model's forecasts for horizon 3 being the only exception. The difference in the forecasting accuracy of the models can be traced to the different modelling assumptions. The SW and EDO model have a richer economic structure than the DS and FM model. The BVAR also performs very good as the higher number of lags compared to the other models can catch important business cycle dynamics. Despite this richer structure the SW, EDO and BVAR models are tightly enough parametrized to yield precise forecasts.

The Greenbook inflation forecasts are more accurate than all structural as well as all nonstructural inflation forecasts. The structural forecasts have an accuracy in line with the accuracy range of the nonstructural forecasts. None of models reaches the forecasting quality of the best nonstructural forecasts. Among the DSGE models the DS and SW model show a good forecasting performance. They achieve a forecast of similar accuracy as the BVAR. The EDO model forecasts are somehow less precise and the FM forecasts are relatively imprecise. The forecasting accuracy relative to the Greenbook forecasts improves with increasing horizons for all models. When I add the Greenbook nowcast to the information set of the models, the forecasting accuracy increases, but does not reach the quality of the Greenbook forecasts. While it is not possible to forecast inflation with DSGE models as precise as the Fed does, the forecasts are reasonable: with the exception of the FM model they are as good or better than a simple autoregressive forecast from horizon 3 onwards and for all horizons for the jump of 0 scenario.

The Greenbook projections are conditioned on a hypothetical path of policy. This hypothetical federal

[9] Faust and Wright (2009) consider two types of autoregressive forecasts. First, a recursive autoregression, where the h-period ahead forecast is constructed by recursively iterating the one-step ahead forecast forward. Second, they use a direct forecast from the autoregression by regressing h-period ahead output growth values on the autoregressive process. For both types they use four lags and get a similar forecasting accuracy.

funds rate is not meant to be a forecast. Nevertheless, viewing it as a forecast its accuracy for short horizons is extremely high. Therefore, the Fed might have conditioned the projections on a policy path that is likely to be implemented in the future and it is reasonable to view this as a forecasting benchmark. Faust and Wright (2009) did not compute interest rate forecasts, so that I cannot compare the structural forecasts to forecasts from their time series models. Due to the Greenbook's extremely high accuracy in the short term, the structural forecasts do much worse than the Greenbook for horizons 0 to 3. For medium term forecasts, however, the forecasting accuracy of the DS, SW and BVAR models dominates the Greenbook path. For short forecasting horizons it is apparent that the BVAR forecasts have a much higher accuracy than the DSGE forecasts. The monetary policy rules in the DSGE models include only few variables and might be too simple. In contrast, the policy rule implicit in the BVAR contains four lags of the interest rate, output growth and the inflation rate. Among the DSGE models the EDO forecasts are very imprecise as they underestimate the level of the interest rate many times. Taking the Greenbook nowcast as given, the forecasting accuracy of the models relative to the Greenbook increases. The results might be sensitive to the hypothetical policy path characteristic of the Greenbook projection. If the Fed's staff would compute an unconditional best forecast for the federal funds rate it might as well dominate the model forecasts for all horizons. For future work it is interesting to condition model forecasts on the Greenbook federal funds path to compute output growth and inflation forecasts. This might change the results significantly.

Del Negro and Schorfheide (2004) propose to use DSGE models as priors for VARs. They show that the forecasting accuracy of these so-called DSGE-VARs improves relative to a VAR and partly to a BVAR with Minnesota priors. They advocate to use DSGE-VARs for forecasting until structural models are available that have the same forecasting performance. The forecasting results in Table 3.2 show that at least the SW models' forecasting performance for output growth, inflation and the interest rate is already good enough to be considered for forecasting exercises on its own.

Faust and Wright (2009) present a table showing the percentage of forecast periods in which the time series model forecasts are more accurate than the Greenbook. This metric is not as sensitive to outliers as the RMSEs. I compute accordant numbers for the structural forecasts which are shown in Table 3.4 in the Appendix. A value higher than 50% indicates that the specific forecast was more accurate than the Greenbook forecast for more than half of the sample. The results are similar to the RMSE results: the Greenbook output growth nowcast dominates the model nowcasts. For the other horizons the model forecasts for output growth are as good as the Greenbook forecasts or even better. For inflation the Greenbook forecasts are more accurate than all model forecasts. The interest rate path of the Greenbook is more precise than model forecasts for short horizons, but model forecasts do as well as the Greenbook for medium forecasts with the EDO model being an exception.

3.6 Model averaging

Density forecasts are useful to show uncertainty around point forecasts. Having estimated the posterior parameter distribution of a certain model, it is straightforward to compute density forecasts that include various sources of uncertainty. One computes forecasts for a large number of draws from the models' posterior parameter distribution to take into account parameter uncertainty. Uncertainty about future realizations of shocks is incorporated by repeatedly drawing from their estimated distribution. However, the largest source of uncertainty - model uncertainty - is ignored. Using only one model to forecast is equivalent to a subjective prior of the forecaster that the specific model is the best representation of the unknown true data generating process. Gerard and Nimark (2008) take into account model uncertainty by combining forecasts from a Bayesian VAR, a FAVAR and a DSGE model. I extend their work to combining forecasts from four DSGE models and an unconstrained Bayesian VAR. Computing weighted forecasts is interesting for a second reason: the results in the empirical literature on forecast combination show that combining multiple forecasts increases the forecasting accuracy. Unless one can identify a single model with superior forecasting performance, forecast combinations are useful for diversification reasons as one does not have to rely on the forecast of a single model. I consider several methods to combine forecasts from the set of models: likelihood based weights, relative performance weights based on past RMSEs, a least squares estimator of weights, and non-parametric combination schemes (mean forecast, median forecast and weights based on model ranks reflecting past RMSEs). While many of these methods have been applied to nonstructural forecasts (see Timmermann, 2006, for a survey) there are to my knowledge no applications to a suite of structural models. From a theoretical point of view likelihood based weights or weights estimated by least squares are appealing. In practice, these estimated weights have the disadvantage that they introduce estimation errors. In the applied literature simple combination schemes like equal-weighting of all models have widely been found to perform better than theoretically optimal combination methods (see e.g. Hsiao and Wan, 2010, for the disconnect of Monte Carlo simulation results and empirical results).

Let I_t^m be the information set of model m at time t including the model equations, parameter estimates and the observable time series of the accordant data vintage. A combined point forecast of models $m = 1, ..., M$ for horizon h denoted as $E[y_{t+h}^{obs}|I_t^1, ..., I_t^M, \omega_{1,h}, ..., \omega_{M,h}]$ can be written as the weighted sum of individual density forecasts $p[y_{t+h}^{obs}|I_t^m]$ with assigend weights $\omega_{m,h}$ divided by the number of draws S:

$$E[y_{t+h}^{obs}|I_t^1, ..., I_t^M, \omega_{1,h}, ..., \omega_{M,h}] = \frac{1}{S}\sum_{m=1}^{M} \omega_{m,h} p[y_{t+h}^{obs}|I_t^m]. \tag{3.9}$$

I take 10,000 draws from each individual forecast and order them in ascending order to get the density forecast for each model. Afterwards I weight each of the 10,000 draws for each model with the specific model weights to compute 10,000 draws of the combined forecast. This is the weighted or averaged density forecast. The weighted point forecast is computed as the mean of the 10000 draws of the weighted forecast. In the following, I discuss various methods how to choose the weights $\omega_{m,h}$.

A natural way to weight different models in a Bayesian context is to use Bayesian Model Averaging. The marginal likelihood $ML(y_T^{obs}|m)$ - with T denoting all observations of a specific historical data sample observed in period t - is computed for each model $m = 1, ..., M$ and posterior probability weights are given by:

$$\omega_m = ML(m|y_T^{obs}) = \frac{ML(y_T^{obs}|m)}{\sum_{m=1}^{M} ML(y_T^{obs}|m)}, \tag{3.10}$$

where a flat prior belief about model m being the true model is used so that no prior beliefs show up in the formula. This weighting scheme is based on the fit of a model to the observed time series. Unfortunately posterior probability weights are not comparable for models that are estimated on a different number of time series. A second problem of the posterior probability weights is that over-parameterized models that have an extreme good in-sample fit, but a bad out-of-sample forecasting accuracy are assigned high weights. To circumvent these problems Gerard and Nimark (2008) use an out-of-sample weighting scheme based on predictive likelihoods as proposed by Eklund and Karlsson (2007) and Andersson and Karlsson (2007).

Predictive Likelihood (PL) The available data is split into a training sample used to estimate the models and a hold-out sample used to evaluate each model's forecasting performance. The forecasting performance is measured by the predictive likelihood, i.e. the marginal likelihood of the hold-out sample conditional on a specific model. I follow the approach suggested by Andersson and Karlsson (2007) and used by Gerard and Nimark (2008) to compute a series of small hold-out sample predictive likelihoods for each horizon. Equation (3.11) shows how to compute the predictive likelihood PL of model m for horizon h:

$$PL_h^m = ML(y_{holdout}^{obs}|y_{training}^{obs}) = \prod_{t=l}^{T-h} ML(y_{t+h}^{obs}|y_t^{obs}). \tag{3.11}$$

Starting with an initial trainings sample of length l, one computes the marginal likelihood for horizon h using the hold-out sample. The training sample is expanded by one observation to $l + 1$ and a second maginal likelihood is computed for the hold-out sample that is one observation shorter than the previous one. This continues until the trainings sample has increased to lenght $T - h$ and the hold-out sample has shrinked to length h. To make the results comparable among models, only the three common variables output growth, inflation and the interest rate are considered for the computation of the predictive likelihood. Finally, the predictive likelihood weights are computed by replacing the marginal likelihood in equation (3.10) with the predictive likelihood:

$$\omega_{m,h} = \frac{PL_h^m}{\sum_{m=1}^{M} PL_h^m}. \tag{3.12}$$

The predictive likelihood weighting scheme allows for different weights to be assigned to a given model at different forecast horizons.

Ordinary Least Squares Weights (OLS) In model averaging applications of time series models it is common to assume a linear-in-weights model and estimate combination weights by ordinary least squares (see Timmermann, 2006). I use the forecasts from previous vintages for each model and the accordant data realizations to regress the realizations y_{t+h}^{obs} on the forecasts $E[y_{t+h}^{obs}|I_t^m]$ from the different models via constrained OLS separately for each variable:

$$y_{t+h}^{obs} = \omega_{1,h} E[y_{t+h}^{obs}|I_t^1] + ... + \omega_{M,h} E[y_{t+h}^{obs}|I_t^M] + \epsilon_{t+h}, \qquad s.t. \sum_{m=1}^{M} \omega_{m,h} = 1. \tag{3.13}$$

The resulting parameter estimates $\omega_{1,h}, ..., \omega_{M,h}$ are the combination weights. Therefore, the combination weights differ for different horizons and also for the three different variables. I omit an intercept term and restrict the weights to sum to one so that the weights can be interpreted as the fractions the specific models contribute to the weighted forecast. It also ensures that the combined forecast lies inside the range of the individual forecasts.

RMSE based weights (RMSE) There are several ways to compute simple relative performance weights. I consider here weightings based on RMSEs of past forecasts and weights based on the relative past forecast accuracy by ranking the accuracy of the different models. For the prior case RMSE based weights can be computed by taking forecasts from previous vintages and compute the RMSE for each model. The weights are then calculated by taking the inverse relative RMSE performance:

$$\omega_{m,h} = \frac{(1/RMSE_h^m)}{\sum_{m=1}^{M} (1/RMSE_h^m)}. \tag{3.14}$$

Rank based weights (Rank) A second possibility to compute relative performance weights is to assign ranks R from 1 to M according to the past forecasting accuracy measured by the RMSEs. This method is similar to the RMSE based weights while being more robust to outliers. The performance rank based weights are computed as follows:

$$\omega_{m,h} = \frac{(1/R_h^m)}{\sum_{m=1}^{M} (1/R_h^m)}. \tag{3.15}$$

Both methods can assign different weights to forecasts of different variables and the different forecasting horizons.

Mean Forecast (Mean) The simplest method to compute a weighted forecast is to give equal weight to each model and simply compute the mean forecast of all models. From a theoretical point of view this approach is not preferable as the weights are purely subjective prior weights implicitly given by the choice of models. However, it has often been found that simple weighting schemes perform well (see e.g. Hsiao and Wan, 2010). A reason is that they give weight to several models instead of choosing one optimal model and are thus robust.

Median Forecast (Median) Another possibility is choose the median of different model forecasts. I compute the median forecast for each of the ordered draws of all models. This gives the density of the median forecast which is used to compute the mean of all these draws as a point forecast. The approach is similar to taking the mean forecast, but is more robust to outliers. The medians from the ordered forecast draws need not to come from the same model for different slices of the ordered forecast draws. By counting the fraction that the median forecast is generated by a specific model one can compute pseudo weights of the different model forecasts that show the contribution of each model to the final point forecast.

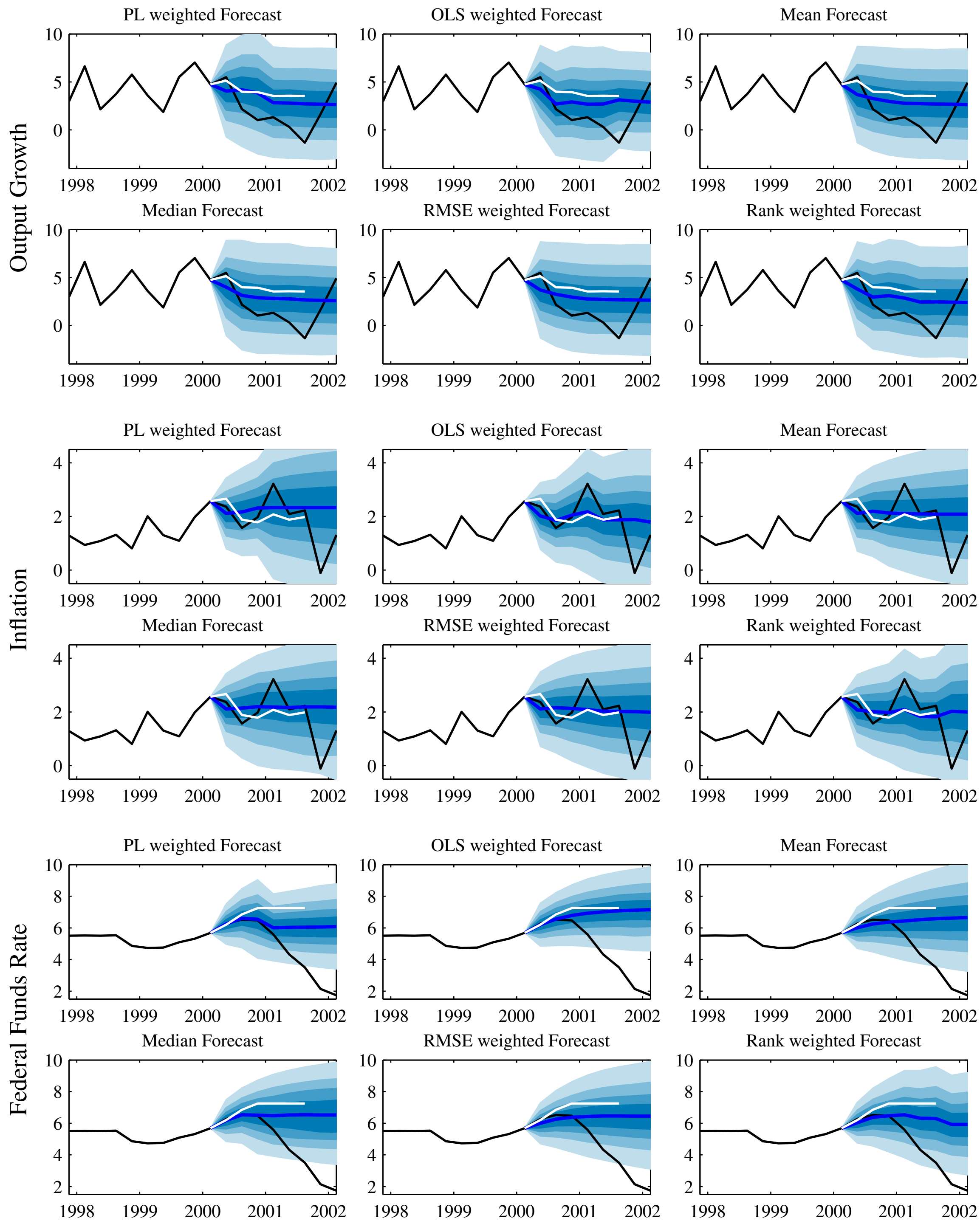
Figure 3.3 shows as an example weighted forecasts computed for the data vintage of May 12, 2000. In comparison with the individual forecasts in Figure 3.1 the forecasts are more robust as no outliers are visible. All methods predict a slightly lower output growth path than the Greenbook and a slight decrease of inflation in the current quarter. Afterwards inflation is predicted to remain about constant. For the interest rate forecasts all models predict an increase in the interest rate for the next three to four quarters. Afterwards the interest rate is predicted to remain at roughly six percent. Only the weighted forecasts based on the predictive likelihood and on ranked past forecasting performance predict a slight interest rate decrease.

3.7 Forecast evaluation of combined forecasts

In Table 3.3, I report the RMSEs for output growth, inflation and interest rate forecasts from the Greenbook, and RMSEs of the six weighted forecasts relative to the Greenbook RMSE. The second last column shows for comparison the relative RMSEs of the best single model as reported in Table 3.2 and the last column shows the relative RMSEs of the best nonstructural model for each horizon as computed by Faust and Wright (2009).

For output growth, inflation and the federal funds rate, it is apparent that the weighted forecasts have in general an accuracy higher than forecasts from most single models. For output growth the Greenbook nowcast is slightly better than all other forecasts, but for all other horizons the weighted model forecasts dominate the Greenbook forecast. The PL weighting scheme is an exception with a forecasting quality not better, but still comparable to the Greenbook. Taking the Greenbook nowcast as given does not translate into more accurate forecasts due to the low persistence of output growth data.

Figure 3.3: Weighted structural forecasts: data vintage May 12, 2000



Notes: the black line shows real-time data until the forecast start and revised data afterwards; the shaded areas show 90% 70%, 50% and 30% probability bands; the line in the middle of the probability bands shows the mean forecast for horizons 0 to 7; the short white line shows the Greenbook forecast for horizons 0 to 5.

For horizons two and above most weighted forecasts dominate RMSEs of a simple autoregressive forecast as reported in Faust and Wright (2009). In contrast, in the case of single model forecasts only the Smets & Wouters model is able to beat the autoregressive forecast. All the differences in output growth forecasting accuracy are statistically insignificant, with the Rank weighted horizon 3 forecast being the only exception.
For the inflation forecast, weighted forecasts increase the forecasting accuracy compared to most single model forecasts. However, the performance of the Greenbook forecasts is still the best. The weighting schemes can roughly be devided into two groups: the PL and OLS weighted forecasts are less precise than the Median, Mean, RMSE and Rank weighted forecasts. The simple Mean forecast is most accurate. Especially for the medium term forecast it improves upon the best single model forecast. For medium term horizons it is only slightly worse than the Greenbook forecast and the best nonstructural forecast. The forecasting accuracy relative to the Greenbook increases with increasing horizons for all weighting schemes. This shows that structural forecasts are especially useful for medium term forecasts. An univariate autoregressive forecast is less precise than the weighted forecasts from horizon 2 onwards. Appending the Greenbook nowcast to the information set of the forecasting models increases the forecasting performance of all weighting methods and the Mean forecast becomes as precise as the best nonstructural forecast. For the jump of 0 scenario all weighted forecasts are more accurate than an univariate autoregressive forecast.
The interest rate forecast results for individual models showed that the Bayesian VAR model performed better than all other models at least for short horizons. Nevertheless, combining this forecast with other less accurate forecasts even improves the forecasting quality: the Mean, RMSE and Rank weighted forecasts are more accurate than the forecasts from the Bayesian VAR. While the Greenbook interest rate path is significantly more accurate for horizons 0 to 2, the Mean, RMSE and Rank weighted forecasts are more precise for horizons 3 to 5. The relative forecasting accuracy improves with increasing horizons for all weighting schemes. Taking the Greenbook nowcast as given, the accuracy of all weighting schemes increases due to the high persistence of the interest rate. The Mean forecast is as precise as the Greenbook policy path for horizon 1 and dominates it for all other horizons.

Overall it turns out that model combination methods that give weight to several models perform well. Likelihood based weighting methods are preferable in theory, but do not work as well in practice. Tables 3.6 to 3.8 in the Appendix report as an example model weights for forecasts derived from data vintage May 12, 2000. Wieland and Wolters (2010) report RMSEs for structural forecasts for five different recessions and find that there is no model that consistently outperforms other models. This shows that the forecasting performance of different models relative to each other varies over time. Therefore, it is important to choose an average of several models to hedge against inaccurate forecasts of individual models. Differences in predictive likelihoods of different models are so high that at most times all weight is given to a single model. Combining several models gives a more robust forecast

Table 3.3: Greenbook RMSE and relative RMSE of weighted model forecasts: 1984-2000

(a) Output growth

horizon	GB	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
					jump off -1				
0	1.75	1.17	1.05	1.07	1.06	1.06	1.04	1.11	1.09
1	2.12	**0.93**	**0.90**	**0.89**	**0.86**	**0.86**	**0.87**	**0.91**	**0.86**
2	2.01	1.06	**0.93**	**0.92**	**0.92**	**0.91**	**0.90**	**0.93**	**0.95**
3	2.15	**0.99**	**0.88**	**0.91**	**0.90**	**0.89**	**0.85**•	**0.86**·	**0.94**
4	2.08	1.00	**0.92**	**0.90**	**0.89**	**0.89**	**0.87**	**0.89**	**0.99**
5	2.08	1.02	**0.92**	**0.91**	**0.92**	**0.92**	**0.90**	**0.90**	**0.97**
					jump off 0				
1	2.12	**0.96**	**0.90**	**0.85**	**0.85**	**0.85**	**0.85**	**0.93**	**0.84**
2	2.01	1.01	**0.94**	**0.93**	**0.91**	**0.91**·	**0.90**•	**0.94**	**0.90**
3	2.15	1.02	**0.94**	**0.92**	**0.90**	**0.90**	**0.91**	**0.87**	**0.95**
4	2.08	1.02	**0.93**	**0.92**	**0.90**	**0.90**	**0.89**	**0.88**	**0.96**
5	2.08	1.03	**0.98**	**0.92**	**0.92**	**0.92**	**0.95**	**0.89**·	**0.98**

(b) Inflation

horizon	GB	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
					jump off -1				
0	0.69	1.52●	1.60●	1.45●	1.44●	1.44●	1.45●	1.47●	1.34
1	0.79	1.58●	1.54●	1.47●	1.43●	1.44●	1.47●	1.44●	1.22
2	0.81	1.37●	1.42●	1.25•	1.23•	1.23•	1.25•	1.29●	1.15
3	0.93	1.17•	1.20•	1.10	1.06	1.07	1.11	1.14	1.03
4	0.89	1.28•	1.32•	1.20·	1.15	1.17	1.20·	1.28•	1.08
5	1.14	1.24•	1.21	1.19·	1.11	1.12	1.16	1.24•	**0.99**
					jump off 0				
1	0.79	1.23•	1.25·	1.16•	1.18•	1.17•	1.17•	1.15•	1.20●
2	0.81	1.24•	1.27●	1.19•	1.16·	1.16·	1.17•	1.16·	1.18
3	0.93	1.23●	1.29●	1.15•	1.09·	1.09·	1.11·	1.15·	1.04
4	0.89	1.18●	1.18·	1.10	1.07	1.07	1.14·	1.19	1.05
5	1.14	1.15•	1.17•	1.12·	1.06	1.06	1.09	1.14	**0.97**

(c) Federal Funds Rate

horizon	GB	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
					jump off -1				
0	0.11	5.95●	4.45●	3.77●	3.56●	3.49●	3.42●	3.57●	-
1	0.49	2.14●	2.13●	1.65●	1.47●	1.47●	1.45●	1.55●	-
2	0.90	1.49•	1.54•	1.22·	1.14	1.14	1.14	1.18	-
3	1.25	1.19	1.33•	1.01	**0.99**	**0.99**	1.00	1.01	-
4	1.60	1.05	1.26•	**0.95**	**0.94**	**0.94**	**0.97**	**0.96**	-
5	1.90	**0.97**	1.19•	**0.91**	**0.92**	**0.91**	**0.91**	**0.87**	-
					jump off 0				
1	0.49	1.37•	1.63●	1.08	1.01	1.02	1.07	1.06	-
2	0.90	1.18	1.49•	**0.99**	**0.93**	**0.93**	**0.97**	**0.96**	-
3	1.25	1.02	1.29·	**0.89**	**0.86**	**0.87**	**0.94**	**0.90**	-
4	1.60	**0.95**	1.23·	**0.88**	**0.87**	**0.87**	**0.92**	**0.88**	-
5	1.90	**0.90**	1.19·	**0.86**	**0.86**	**0.86**	**0.89**	**0.86**	-

Notes: PL: Predictive Likelihood; OLS: Ordinay Least Squares; Median: Median forecast; Mean: Mean forecast; RMSE: weighted by inverse RMSE; Rank: weighted by inverse ranks; best M: best single model forecast; Best FW: Best performing atheoretical model for the specific horizon considered by Faust & Wright; The first column shows the forecast horizon. The second column shows the RMSE for the Greenbook. The other columns show RMSE of alternative forecasts relative to the Greenbook. Values less than one are in bold and show that a forecast is more accurate than the one by the Greenbook. The symbols ●, •, ·, indicate that the relative RMSE is significantly different from one at the 1, 5, or 10% level, respectively.

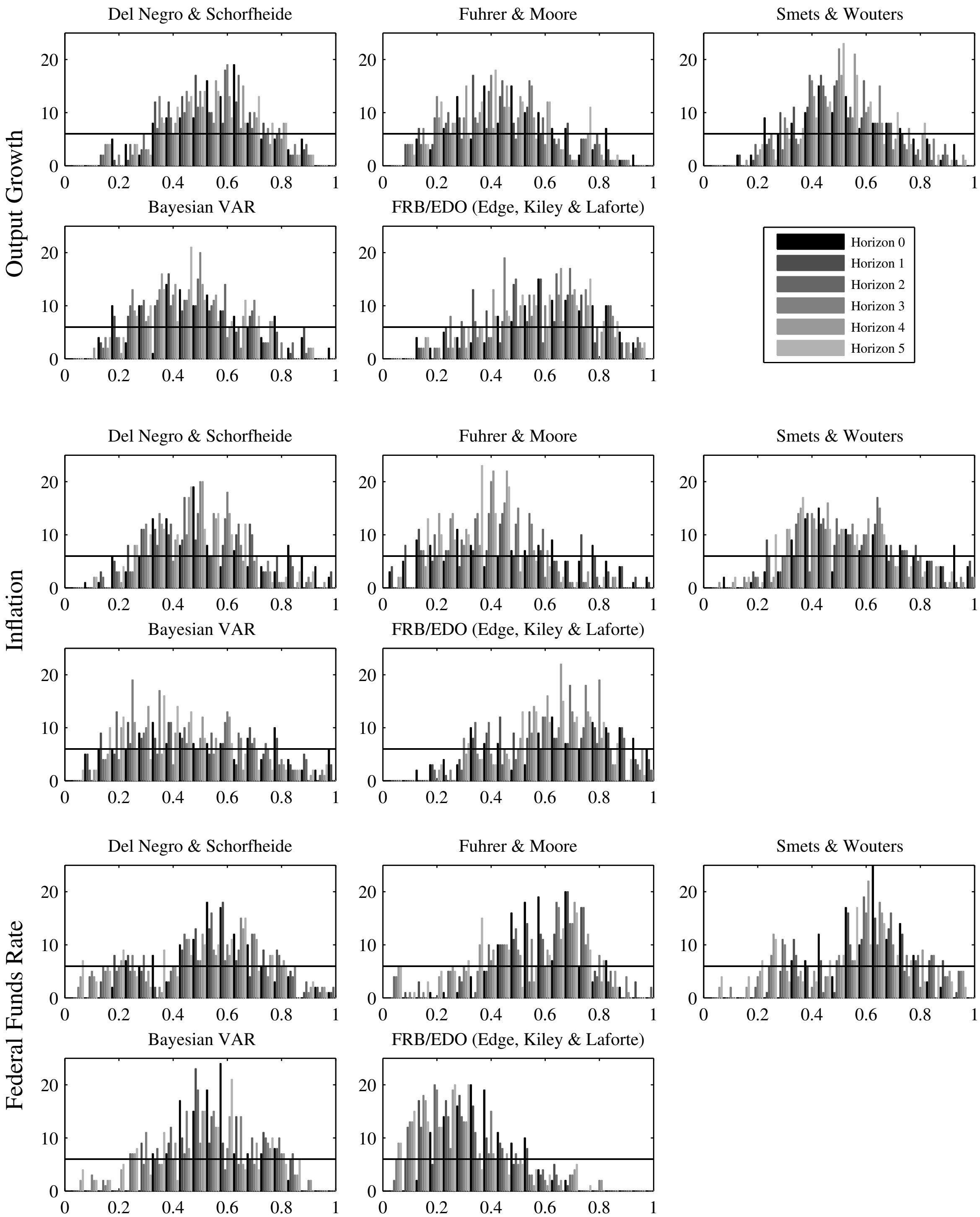
as it prevents against choosing an outlier that produces high forecast errors. Also estimated weights by least squares do not perform as good as simpler combination schemes: restricting the weights to sum to one leads to estimation problems so that in many cases weight is given only to one model. The Median forecast works quite well as it ensures that outliers are not chosen. The best forecasting performance is achieved by the Mean forecast and the RMSE and Rank based weighted forecasts. However, the RMSE weights deviate only slightly from the Mean forecast. The Rank weights take past forecasting performance more into account: this increases the accuracy of the output growth forecast, but does not improve on the Mean forecast for inflation and the interest rate. Therefore, at this stage, one can conclude that a simple Mean forecast is the preferable method. It is very easy to compute as one needs no forecasts and realization from earlier data vintages to calculate model weights and it yields precise forecasts that are quite robust to outliers. Table 3.5 shows the percentage of forecast periods in which the weighted forecasts are more accurate than the Greenbook projections. The results of this robust statistic are very similar to the RMSE results.

To sum up the point forecast evaluation, the forecasts of the Smets & Wouters model are good. The accuracy of forecasts that give considerable weight to several forecasts is as high as the Smets & Wouters forecast and in most cases even better. The accuracy of the Mean forecast is comparable to nonstructural forecasting methods that can process large data sets. All forecasts based on structural models are especially suited to compute medium term forecasts.

3.8 Density forecast evaluation

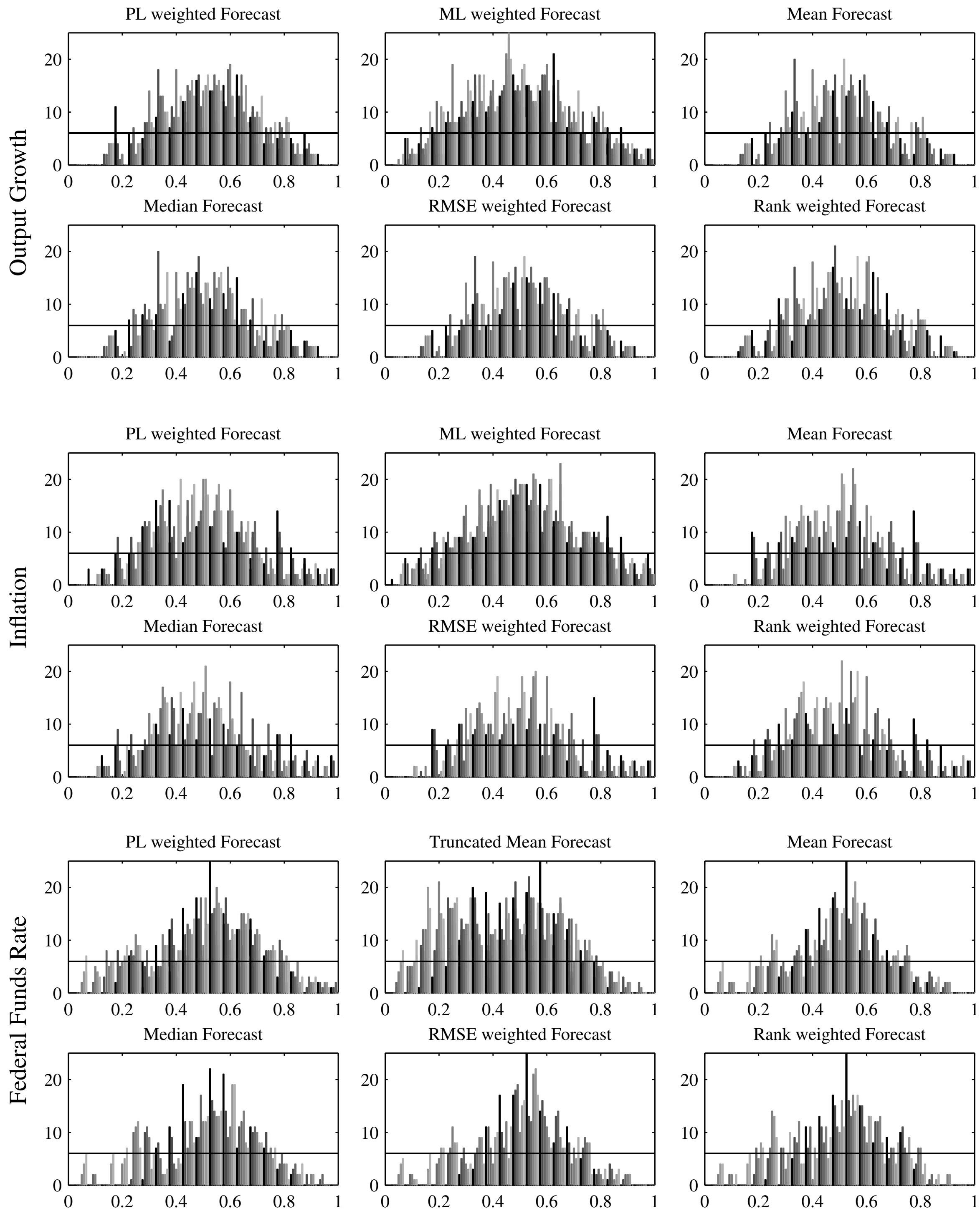
Assuming a symmetric loss function, the accuracy of point forecasts can be easily compared by computing RMSEs. Evaluating density forecasts is less straightforward. The true density is never observed. Still one can compare the distribution of observed data with density forecasts to check whether the forecasts provide a realistic description of actual uncertainty. I use the following evaluation procedure: I split up the density forecasts into probability bands that each cover 5% of the probability mass. This is similar to disaggregating the fan charts plotted in Figures 3.1 and 3.3 further into smaller confidence bands. For each data realization I can check into which of the 20 probability bands of the accordant density forecast it falls. Doing this for all realization and the corresponding density forecasts, 5% of the realizations should be contained in each of the probability bands. Otherwise the density forecasts are not a good characterization of the distribution of the data realizations. In general, if one divides density forecasts into probability bands of equal coverage, data realisations should be uniformly distributed across all probability bands. This is the approach outlined in Diebold et al. (1998) and Diebold et al. (1999). More formally, it is based on the relationship between the data generating process and the sequence of density forecasts via probability integral transforms of the observed data with respect to the density forecasts. The probability integral transform (PIT) is

Figure 3.4: Evaluation of structural density forecasts: 1984 - 2000



Notes: The figures show the distribution of realized data points on the density forecasts. The density forecasts are represented as probability bands each covering 5% of the density. The bars show how many of the realized observations fall in each of the probability bands. If the density forecast is an accurate description of actual uncertainty, than about six of the 122 observations should fall in each probability band.

Figure 3.5: Evaluation of structural density forecasts: 1984 - 2000



Notes: The figures show the distribution of realized data points on the density forecasts. The density forecasts are represented as probability bands each covering 5% of the density. The bars show how many of the realized observations fall in each of the probability bands. If the density forecast is an accurate description of actual uncertainty, than about six of the 122 observations should fall in each probability band.

the cumulative density function corresponding to the sequence of n density forecasts $\{p_t(y_{t+h}^{obs})\}_{t=1}^{n}$ evaluated at the corresponding observed data points $\{y_{t+h}^{obs}\}_{t=1}^{n}$:

$$z_t = \int_{-\infty}^{y_{t+h}^{obs}} p_t(u)du, \qquad \text{for} \quad t = 1, ..., n. \tag{3.16}$$

The PIT is the probability implied by the density forecast that a realized data point would be equal or less than what is actually observed. If the sequence of density forecasts is an accurate description of actual uncertainty, the sequence of PITs, $\{z_t\}_{t=1}^{n}$, should be distributed uniformly between zero and one. Figures 3.4 and 3.5 presents a visual assessment of the distribution of realized data points on the sequence of PITs that is represented as a histogram of 20 probability bands each covering 5%. There are $n = 122$ forecasts, so that there should be about 6 observations in each of the probability bands if the density forecasts are accurate. This is represented by the horizontal line. The bars shaded in different colors reflect PITs for the different forecasting horizons.

The peak in the middle of the histograms of the output growth forecasts shows that these overestimate actual uncertainty. The histograms for inflation are closer to a uniform distribution, especially for the inflation nowcast. There is only a slight peak in the middle of the distributions and the histogramms for some models cover the entire distribution including the tails. Higher horizon forecasts overestimate actual inflation uncertainty. The density forecasts are imprecise for the federal funds rate. The tails are not covered, especially for short horizons, and thus uncertainty is overestimated by the density forecasts. Gerard and Nimark (2008) give a plausible reason for the overestimation of actual uncertainty by DSGE models. The models impose tight restrictions on the data. If the data rejects these restrictions, large shocks are needed to fit the models to the data resulting in high shock uncertainty. As all individual model forecasts overestimate actual uncertainty it is not possible that the weighted forecasts yield a more realistic assesment of uncertainty. Therefore, the averaged density forecasts overestimate uncertainty as well.[10]

3.9 Conclusion

During the last decade theory based DSGE models that are consistently derived from microeconomic optimization problems of households and firms have become the workhorse of modern monetary economics. Despite their stylized nature and their reliance on few equations they are widely used in academics as well as at policy institutions. Computing out of sample forecasts is an ultimate test of the ability of this class of models to explain business cycles. In this chapter, I have assessed the accuracy of point and density forecasts of four DSGE models using real-time data. While point forecasts are surprisingly precise, density forecasts have been shown to overestimate actual uncertainty.

[10] In principle, there are tests available to formally check for a uniform distribution (Berkowitz, 2001). Unfortunately, the results have to be treated with high caution (see Elder et al., 2005; Gerard and Nimark, 2008). As the visual assessement has already shown clear evidence against a uniform distribution of the PITs, I do not use additional formal tests.

Point forecasts of some models are comparable to the forecasting accuracy of atheoretical forecasting methods that can process large data sets. Especially the model by Smets and Wouters (2007) yields relatively precise inflation, output growth and interest rate forecasts. Combining several forecasts can increase the forecasting accuracy. Combination methods that give significant weight to several models are preferable over methods that aim to identify a single best model. The accuracy of a simple mean of model forecasts is hard to beat by other forecast weighting methods. DSGE based forecasts perform particularly well for medium term forecasts in comparison with Greenbook projections and nonstructural forecasts. Structural forecasts perform quite well during normal times, but they are not able to detect large recessions and turning points due to their weak internal propagation meachanism. Large shocks are needed to fit the models to volatile periods of the sample. This is also the reason for their wide confidence bands.

References

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

An, S., Schorfheide, F., 2007. Bayesian analysis of DSGE models. Econometric Reviews 26(2-4), 113–172.

Andersson, M. K., Karlsson, S., 2007. Bayesian forecast combination for VAR models, Sveriges Riksbank Working Paper No 216.

Bache, I. W., Jore, A. S., Mitchell, J., Vahey, S. P., 2009. Combining VAR and DSGE forecast densities, Norges Bank Working paper 2009/23.

Berkowitz, J., 2001. Testing density forecasts, with applications to risk management. Journal of Business and Economic Statistics 19(4), 465–474.

Bernanke, B. S., Boivin, J., 2003. Monetary policy in a data-rich environment. Journal of Monetary Economics 50(3), 525–546.

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

Christoffel, K., Warne, A., Coenen, G., 2010. Forecasting with DSGE models, ECB Working Paper No. 1185.

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

Del Negro, M., Schorfheide, F., Smets, F. R., Wouters, R., 2007. On the fit of new Keynesian models. Journal of Business and Economic Statistics 25, 123–143.

Diebold, F. X., Gunther, T. A., Tay, A. S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39(4), 863–883.

Diebold, F. X., Hahn, J., Tay, A. S., 1999. Multivariate density forecast evaluation and calibration in financial risk management: High-frequency returns on foreign exchange. Review of Economics and Statistics 81(4), 661–673.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Doan, T., Litterman, R., Sims, C., 1984. Forecasting and conditional projection using realistic prior distributions. Econometric Reviews 3, 1 – 100.

Edge, R. M., Kiley, M. T., Laforte, J.-P., 2007. Documentation of the research and statistics divisions estimated DSGE model of the U.S. economy: 2006 version, Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.: 2007-53.

Edge, R. M., Kiley, M. T., Laforte, J.-P., 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., Kiley, M. T., Laforte, J.-P., 2009. A comparison of forecast performance between federal reserve staff forecasts, simple reduced form models and a DSGE model, Working Paper.

Eklund, J., Karlsson, S., 2007. Forecast combination and model averaging using predictive measures. Econometric Reviews 26(2-4), 329–363.

Elder, R., Kapetanios, G., Taylor, T., Yates, T., 2005. Assessing the MPC's fan charts. Bank of England Quarterly Bulletin Autumn, 326–345.

Fair, R. C., 2007. Evaluating inflation targeting using a macroeconometric model. Economics: The Open-Access, Open-Assessment E-Journal 8.

Faust, J., Wright, J. H., 2009. Comparing Greenbook and reduced form forecasts using a large realtime dataset. Journal of Business and Economic Statistics 27(4), 468–479.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the Euro area? Journal of Monetary Economics 50, 1243–1255.

Francis, N., Ramey, V. A., 1995. Measures of per capita hours and their implications for the technology-hours debate, nBER Working Paper 11694.

Fuhrer, J. C., 1997. Inflation/output variance trade-offs and optimal monetary policy. Journal of Money, Credit and Banking 29(2), 214–234.

Fuhrer, J. C., Moore, G., 1995a. Inflation persistence. The Quarterly Journal of Economics 110(1), 127–159.

Fuhrer, J. C., Moore, G., 1995b. Monetary policy trade-offs and the correlation between nominal interest rates and real output. The American Economic Review 85(1), 219–239.

Gerard, H., Nimark, K., 2008. Combing multivariate density forecasts using predictive criteria, Research Discussion Paper 2008-2, Reserve Bank of Australia.

Giannone, D., Monti, F., Reichlin, L., 2009. Incorporating conjunctural analysis in structural models. In: Wieland, V. (Ed.), The Science and Practice of Monetary Policy Today. Springer Science, pp. 41–57.

Goodfriend, M., King, R. G., 1997. The new neoclassical synthesis and the role of monetary policy. In: Bernanke, B. S., Rotemberg, J. J. (Eds.), National Bureau of Economic Research Macroeconomics Annual 1997. MIT Press, Cambridge, MA.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Hsiao, C., Wan, S. K., 2010. Is there an optimal forecast combination?, Working Paper University of Southern California.

Kimball, M., 1995. The quantitative analytics of the basic monetarist model. Journal of Money, Credit and Banking 27(4), 1241–1277.

Marcellino, M., Stock, J., Watson, M., 2003. Macroeconomic forecasting in the Euro area: Country-specific versus area-wide information. European Economic Review 47, 1–18.

Romer, C. D., Romer, D. H., 2000. Federal reserve information and the behavior of interest rates. American Economic Review 90, 429–457.

Rotemberg, J. J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy, in B. Bernanke and J. Rotemberg, (eds.), NBER Macroeconomics Annual, The MIT Press.

Schorfheide, F., 2000. Loss function-based evaluation of DSGE models. Journal of Applied Econometrics 15, 645–670.

Sims, C. A., 2002. The role of models and probabilities in the monetary policy process. Brookings Papers on Economic Activity 2, 1–40.

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

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

Stock, J., Watson, M., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.

Timmermann, A., 2006. Forecast combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Amsterdam: North Holland, pp. 135–196.

Wang, M.-C., 2009. Comparing the DSGE model with the factor model: An out-of-sample forecasting experiment. Journal of Forecasting 28(2), 167–182.

Wieland, V., Wolters, M. H., 2010. The diversity of forecasts from macroeconomic models of the U.S. economy. Economic Theory, forthcoming.

Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.

Appendix

A.1 Additional results

Table 3.4: Percentage of periods alternative forecast better than Greenbook: 1984-2000

(a) Output Growth

horizon	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1			
0	29	34	37	32	38	43	30
1	**52**	45	48	48	**51**	**60**	39
2	48	47	**53**	49	**53**	**58**	37
3	47	43	**59**	**51**	**51**	**57**	42
4	44	45	**52**	48	48	**54**	36
5	45	43	**60**	49	42	**52**	43
				jump off 0			
1	43	**51**	49	48	50	**59**	40
2	48	49	**57**	43	**53**	**55**	41
3	48	47	**55**	48	**52**	**57**	38
4	46	47	**53**	42	**52**	**57**	39
5	43	44	**55**	43	47	49	43

(b) Inflation

horizon	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1			
0	41	30	41	29	38	37	25
1	29	31	44	38	35	40	21
2	41	38	36	35	39	43	25
3	44	36	33	32	40	44	17
4	43	30	36	31	34	43	11
5	37	31	38	35	35	46	16
				jump off 0			
1	36	35	36	43	36	41	30
2	37	32	40	45	38	40	21
3	42	43	37	38	48	43	20
4	37	26	33	36	38	43	18
5	38	31	31	50	33	48	15

(c) Federal Funds Rate

horizon	DS	FM	SW	EDO	BVAR	best FW	worst FW
				jump off -1			
0	8	13	6	4	13	-	-
1	28	27	22	11	25	-	-
2	45	33	32	18	38	-	-
3	50	34	39	23	45	-	-
4	**56**	31	45	30	48	-	-
5	**60**	34	50	29	**56**	-	-
				jump off 0			
1	33	31	29	23	38	-	-
2	41	35	39	27	50	-	-
3	46	42	48	27	**53**	-	-
4	48	40	**53**	29	**57**	-	-
5	**53**	42	**54**	24	**59**	-	-

Notes: GB: Greenbook; DS: Del Negro & Schorfheide; FM: Fuhrer & Moore; SW: Smets & Wouters; EDO: FRB/EDO Model by Edge, Kiley & Laforte; BVAR: Bayesian VAR; Best FW: Best performing atheoretical model for the specific horizon considered by Faust & Wright; Worst FW: Worst performing atheoretical model for the specific horizon considered by Faust & Wright. The first column shows the forecast horizon. The other columns show the percentage of forecast periods in which forecast errors of specific models are smaller in absolute value than the Greenbook forecast error. Entries greater than 50 percent indicate that the alternative forecast is better more than half the time and are in bold.

Table 3.5: Percentage of periods weighted forecast better than Greenbook: 1984-2000

(a) Output Growth

horizon	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
				jump off -1				
0	36	43	36	40	40	39	38	43
1	**52**	**55**	**55**	**55**	**56**	**55**	**52**	**60**
2	45	**55**	**54**	**57**	**57**	**56**	**53**	**58**
3	47	**57**	**55**	**57**	**58**	**63**	**59**	**57**
4	44	49	**60**	**54**	**54**	**65**	**52**	**54**
5	45	49	**54**	**56**	**55**	**56**	**60**	**52**
				jump off 0				
1	44	**53**	**54**	**57**	**57**	**56**	**51**	**59**
2	46	**54**	**62**	**58**	**61**	**53**	**57**	**55**
3	44	**53**	**55**	**55**	**55**	**56**	**55**	**57**
4	46	**54**	**55**	**53**	**53**	**53**	**53**	**57**
5	43	49	**53**	**53**	**53**	**53**	**55**	49

(b) Inflation

horizon	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
				jump off -1				
0	39	32	42	37	38	40	41	37
1	33	34	33	38	38	34	44	40
2	41	40	46	43	44	46	41	43
3	44	43	45	49	48	46	44	44
4	43	42	43	44	45	43	43	43
5	37	38	39	43	44	41	38	46
				jump off 0				
1	38	40	37	37	38	39	43	41
2	38	39	41	43	43	45	45	40
3	42	37	43	47	46	50	48	43
4	37	38	39	44	43	42	38	43
5	38	43	35	40	42	43	50	48

(c) Federal Funds Rate

horizon	PL	OLS	Median	Mean	RMSE	Rank	best M	best FW
				jump off -1				
0	10	9	13	12	12	13	13	-
1	29	14	29	29	32	31	28	-
2	43	29	42	41	38	40	45	-
3	50	37	48	**51**	**54**	50	49	-
4	**56**	34	**57**	**56**	**57**	**56**	**56**	-
5	**60**	33	**58**	**61**	**61**	**60**	**60**	-
				jump off 0				
1	31	23	36	38	37	40	38	-
2	43	29	45	48	45	**53**	50	-
3	45	38	**55**	**58**	**57**	**51**	50	-
4	48	38	**59**	**56**	**57**	**54**	**57**	-
5	**53**	33	**60**	**63**	**62**	**53**	**59**	-

Notes: PL: Predictive Likelihood; OLS: Ordinay Least Squares; Median: Median forecast; Mean: Mean forecast; RMSE: weighted by inverse RMSE; Rank: weighted by inverse ranks; best M: best single model forecast; Best FW: Best performing atheoretical model for the specific horizon considered by Faust & Wright; The first column shows the forecast horizon. The other columns show the percentage of forecast periods in which forecast errors of specific models are smaller in absolute value than the Greenbook forecast error. Entries greater than 50 percent indicate that the alternative forecast is better more than half the time and are in bold.

Table 3.6: Combination weights for data vintage May 12, 2000: output growth

model	PL	OLS	Median	Mean	RMSE	Rank
			horizon 0			
DS	0.00	0.00	0.01	0.20	0.19	0.09
FM	0.00	1.00	0.33	0.20	0.21	0.22
SW	0.00	0.00	0.33	0.20	0.19	0.11
EDO	0.00	0.00	0.00	0.20	0.19	0.15
BVAR	1.00	0.00	0.32	0.20	0.22	0.44
			horizon 1			
DS	0.00	0.00	0.98	0.20	0.19	0.11
FM	0.00	0.00	0.00	0.20	0.18	0.09
SW	0.00	0.42	0.00	0.20	0.21	0.44
EDO	0.00	0.45	0.02	0.20	0.21	0.22
BVAR	1.00	0.12	0.00	0.20	0.21	0.15
			horizon 2			
DS	0.00	0.00	0.93	0.20	0.19	0.11
FM	0.00	0.00	0.02	0.20	0.18	0.09
SW	0.00	0.19	0.00	0.20	0.21	0.22
EDO	0.00	0.44	0.05	0.20	0.21	0.15
BVAR	1.00	0.37	0.00	0.20	0.21	0.44
			horizon 3			
DS	1.00	0.00	0.78	0.20	0.19	0.11
FM	0.00	0.00	0.06	0.20	0.18	0.09
SW	0.00	0.19	0.00	0.20	0.21	0.44
EDO	0.00	0.42	0.10	0.20	0.21	0.15
BVAR	0.00	0.38	0.06	0.20	0.21	0.22
			horizon 4			
DS	1.00	0.00	0.75	0.20	0.19	0.09
FM	0.00	0.00	0.09	0.20	0.19	0.11
SW	0.00	0.28	0.00	0.20	0.21	0.44
EDO	0.00	0.37	0.12	0.20	0.20	0.15
BVAR	0.00	0.35	0.04	0.20	0.21	0.22
			horizon 5			
DS	1.00	0.00	0.53	0.20	0.19	0.09
FM	0.00	1.00	0.26	0.20	0.20	0.15
SW	0.00	0.00	0.00	0.20	0.21	0.44
EDO	0.00	0.00	0.15	0.20	0.19	0.11
BVAR	0.00	0.00	0.06	0.20	0.21	0.22

Notes: PL: Predictive Likelihood; OLS: Ordinay Least Squares; Median: Median forecast; Mean: Mean forecast; RMSE: weighted by inverse RMSE; Rank: weighted by inverse ranks; DS: Del Negro & Schorfheide; FM: Fuhrer & Moore; SW: Smets & Wouters; EDO: FRB/EDO Model by Edge, Kiley & Laforte; BVAR: Bayesian VAR; The first column shows the model name and the rows show the weight of each model for the different combination schemes. For each horizon, the five model weights sum up to 1.

Table 3.7: Combination weights for data vintage May 12, 2000: inflation

model	PL	OLS	Median	Mean	RMSE	Rank
			horizon 0			
DS	0.00	0.00	0.00	0.20	0.22	0.15
FM	0.00	0.00	0.11	0.20	0.16	0.09
SW	0.00	0.62	0.05	0.20	0.23	0.44
EDO	0.00	0.00	0.00	0.20	0.18	0.11
BVAR	1.00	0.38	0.84	0.20	0.22	0.22
			horizon 1			
DS	0.00	0.00	0.21	0.20	0.20	0.15
FM	0.00	0.00	0.00	0.20	0.17	0.09
SW	0.00	0.49	0.03	0.20	0.23	0.44
EDO	0.00	0.14	0.00	0.20	0.19	0.11
BVAR	1.00	0.37	0.76	0.20	0.22	0.22
			horizon 2			
DS	0.00	0.00	0.50	0.20	0.20	0.15
FM	0.00	0.30	0.00	0.20	0.19	0.11
SW	0.00	0.35	0.07	0.20	0.22	0.44
EDO	0.00	0.23	0.00	0.20	0.17	0.09
BVAR	1.00	0.11	0.44	0.20	0.22	0.22
			horizon 3			
DS	1.00	0.25	0.44	0.20	0.24	0.44
FM	0.00	0.35	0.00	0.20	0.17	0.09
SW	0.00	0.00	0.10	0.20	0.22	0.22
EDO	0.00	0.39	0.00	0.20	0.17	0.11
BVAR	0.00	0.00	0.46	0.20	0.20	0.15
			horizon 4			
DS	1.00	0.00	0.36	0.20	0.22	0.22
FM	0.00	0.31	0.00	0.20	0.16	0.09
SW	0.00	0.16	0.11	0.20	0.23	0.44
EDO	0.00	0.54	0.00	0.20	0.20	0.15
BVAR	0.00	0.00	0.52	0.20	0.19	0.11
			horizon 5			
DS	1.00	0.00	0.33	0.20	0.22	0.22
FM	0.00	0.33	0.00	0.20	0.16	0.09
SW	0.00	0.15	0.13	0.20	0.23	0.44
EDO	0.00	0.52	0.00	0.20	0.20	0.15
BVAR	0.00	0.00	0.54	0.20	0.18	0.11

Notes: PL: Predictive Likelihood; OLS: Ordinay Least Squares; Median: Median forecast; Mean: Mean forecast; RMSE: weighted by inverse RMSE; Rank: weighted by inverse ranks; DS: Del Negro & Schorfheide; FM: Fuhrer & Moore; SW: Smets & Wouters; EDO: FRB/EDO Model by Edge, Kiley & Laforte; BVAR: Bayesian VAR; The first column shows the model name and the rows show the weight of each model for the different combination schemes. For each horizon, the five model weights sum up to 1.

Table 3.8: Combination weights for data vintage May 12, 2000: federal funds rate

model	PL	OLS	Median	Mean	RMSE	Rank
			horizon 0			
DS	0.00	0.00	0.00	0.20	0.18	0.11
FM	0.00	0.00	0.00	0.20	0.21	0.15
SW	0.00	0.00	0.00	0.20	0.22	0.22
EDO	0.00	1.00	1.00	0.20	0.14	0.09
BVAR	1.00	0.00	0.00	0.20	0.25	0.44
			horizon 1			
DS	0.00	0.00	0.00	0.20	0.18	0.11
FM	0.00	0.00	0.00	0.20	0.23	0.22
SW	0.00	0.00	0.00	0.20	0.20	0.15
EDO	0.00	1.00	1.00	0.20	0.14	0.09
BVAR	1.00	0.00	0.00	0.20	0.24	0.44
			horizon 2			
DS	0.00	0.00	0.03	0.20	0.19	0.11
FM	0.00	0.00	0.00	0.20	0.22	0.22
SW	0.00	0.00	0.00	0.20	0.20	0.15
EDO	0.00	1.00	0.54	0.20	0.15	0.09
BVAR	1.00	0.00	0.43	0.20	0.25	0.44
			horizon 3			
DS	1.00	0.00	0.12	0.20	0.19	0.11
FM	0.00	0.00	0.00	0.20	0.20	0.22
SW	0.00	0.00	0.00	0.20	0.20	0.15
EDO	0.00	1.00	0.38	0.20	0.16	0.09
BVAR	0.00	0.00	0.50	0.20	0.24	0.44
			horizon 4			
DS	1.00	0.00	0.16	0.20	0.21	0.15
FM	0.00	0.00	0.00	0.20	0.18	0.11
SW	0.00	0.00	0.00	0.20	0.21	0.22
EDO	0.00	1.00	0.38	0.20	0.16	0.09
BVAR	0.00	0.00	0.46	0.20	0.23	0.44
			horizon 5			
DS	1.00	0.00	0.22	0.20	0.21	0.15
FM	0.00	0.00	0.00	0.20	0.17	0.09
SW	0.00	0.00	0.00	0.20	0.22	0.22
EDO	0.00	1.00	0.38	0.20	0.17	0.11
BVAR	0.00	0.00	0.40	0.20	0.23	0.44

Notes: PL: Predictive Likelihood; OLS: Ordinay Least Squares; Median: Median forecast; Mean: Mean forecast; RMSE: weighted by inverse RMSE; Rank: weighted by inverse ranks; DS: Del Negro & Schorfheide; FM: Fuhrer & Moore; SW: Smets & Wouters; EDO: FRB/EDO Model by Edge, Kiley & Laforte; BVAR: Bayesian VAR; The first column shows the model name and the rows show the weight of each model for the different combination schemes. For each horizon, the five model weights sum up to 1.

Chapter 4

Does Trade Integration Alter Monetary Policy Transmission?

(with Tobias Cwik and Gernot J. Müller)

Abstract This chapter explores the role of trade integration—or openness—for monetary policy transmission in a medium-scale New Keynesian model. Allowing for strategic complementarities in price-setting, we highlight a new dimension of the exchange rate channel by which monetary policy directly impacts domestic inflation. Although the strength of this effect increases with economic openness, it also requires that import prices respond to exchange rate changes. In this case domestic producers find it optimal to adjust their prices to exchange rate changes which alter the domestic currency price of their foreign competitors. We pin down key parameters of the model by matching impulse responses obtained from a vector autoregression on U.S. time series relative to an aggregate of industrialized countries. While we find evidence for strong complementarities, exchange rate pass-through is limited. Openness has therefore little bearing on monetary transmission in the estimated model.

Keywords: monetary policy transmission, open economy, trade integration, exchange rate channel, strategic complementarity, exchange rate pass-through

JEL-Codes: F41, F42, E52

4.1 Introduction

Recent research on the monetary transmission mechanism has focused on the quantitative performance of dynamic stochastic general equilibrium (DSGE) models. Specifically, interest has centered on their ability to account for the dynamic effects of monetary policy shocks as apparent from estimated vector autoregression (VAR) models. In a seminal study, Christiano et al. (2005) show that a medium scale New Keynesian model mimics quite closely the VAR-responses to a monetary policy shock of as many as nine variables. This result is obtained while abstracting from external trade altogether. Taken at face value, it suggests that trade integration, or openness, plays no important role for monetary policy transmission—at least as far as a large open economy such as the U.S. is concerned.[1]

There is, however, a secular trend in trade integration, suggesting that economies are becoming considerably more open over time. In the U.S., imports, as a fraction of GDP, have risen from about 6 percent in 1973 to 16 percent to date. In fact, as this trend has been accelerating over the last decade, some observers have identified increasing trade integration as an important manifestation of globalization.[2] In this chapter, we investigate more systematically the role of trade integration for monetary policy transmission, where we measure trade integration by the import-to-GDP ratio. Specifically, we assess how increasing openness alters quantitatively the effects of monetary policy shocks on domestic (i.e. producer price) inflation and domestic absorption. We focus on these variables, because they are well defined in closed economy models as well.

Taking an analytical perspective, earlier work by Clarida et al. (2001) and Galí and Monacelli (2005) has stressed the similarity between open and closed economy versions of the New Keynesian baseline model. In fact, apart from being a source of additional shocks, 'openness' merely alters some of the reduced-form coefficients of the canonical representation of the model which is, in fact, shown to be isomorphic in closed and open economies. More recently, Erceg et al. (2007) have shown that the difference between closed and open economies in this class of models hinges on the relative size of the intertemporal elasticity of substitution and the trade price elasticity. Moreover, these authors argue that—for reasonable calibrations—increasing openness is unlikely to alter the transmission of domestic shocks, monetary policy shocks inclusive, in a quantitatively important way.

However, taking up the question within the New Keynesian baseline model twists the analysis towards finding no effect of openness. A key assumption underlying the derivation of the New Keynesian Phillips curve and, hence, its isomorphism in closed and open economies, is that the demand functions faced by intermediate goods firms are characterized by a constant elasticity of substitution. This, in turn, implies that the desired markup is independent of the price of competitors, i.e. there are no strategic complementarities in price setting. Such complementarities arise under a more general formulation of the demand functions, or, rather, the underlying aggregation technology. In this case, the isomorphism of the New Keynesian Phillips curve in closed and open economies breaks down. Intuitively, strategic complementarities arise not only with respect to domestic, but also with respect

[1] Other studies which employ this approach find similarly satisfactory results for variants of the New Keynesian model. Rotemberg and Woodford (1997), Amato and Laubach (2003), Bovin and Giannoni (2006) and Meier and Müller (2006) are examples. These studies also assume counterfactually closed economy models. Clearly, other studies have explored the empirical performance of open economy DSGE models; yet these studies have typically not been particularly concerned with monetary transmission, see, e.g., Lubik and Schorfheide (2006) and Adolfson et al. (2007).

[2] The consequences of globalization for monetary policy are widely discussed both in academia and among policy makers. Most commentators, taking a fairly general perspective, have argued that globalization does not fundamentally affect the central bank's ability to control the economy, see, e.g., Mishkin (2007) and Bernanke (2007). Changes brought about by globalization may nevertheless require, as Yellen (2006) puts it, "some recalibration of policy responses".

to foreign competitors. Hence, the domestic currency price charged by foreign competitors enters the decision problem of domestic firms and eventually the New Keynesian Phillips curve. Recently, Guerrieri et al. (2008) have highlighted the importance of this mechanism in accounting for inflation dynamics.[3]

In this chapter, we take price-setting complementarities into account when exploring the role of openness for monetary transmission. As a result, a new dimension of the exchange rate channel emerges. Traditionally, monetary policy is thought to directly impact CPI-inflation and to indirectly impact domestic inflation via the exchange rate, where the latter effect comes about through changes in demand induced by 'expenditure-switching'. With strategic price-setting complementarities, changes in the exchange rate, which alter the domestic currency prices charged by foreign competitors, directly impact domestic inflation. The importance of this effect increases with i) the extent of strategic complementarities in price-setting; ii) the openness of an economy and iii) the amount of exchange rate pass-through.

Our analysis is based on a medium-scale two-country DSGE model. It features an aggregation technology for the production of final goods which gives rise to strategic complementarities in price-setting; in addition, the aggregation technology determines trade integration by giving unequal weight to domestically produced and imported intermediate goods. The model also features a number of frictions which the literature has found to increase the empirical success of this class of models; notably, we allow exchange rate pass-through to be limited in the short-run. Overall, the model structure is rich enough to provide a quantitatively realistic account of the monetary transmission mechanism such as to allow us to study the quantitative implications of trade integration on monetary transmission.

As a benchmark, we compute impulse responses to a monetary policy shock within a VAR model estimated on quarterly time series data for the U.S. relative to an aggregate of industrialized countries. In addition to standard 'closed-economy' variables, the VAR model also includes CPI-inflation as well as U.S. net exports. We treat the impulse responses as a characterization of the actual monetary transmission mechanism and estimate the structural parameters of the DSGE model employing the minimum distance estimation strategy suggested by Rotemberg and Woodford (1997) and Christiano et al. (2005). To avoid identification problems we fix several parameter values prior to the estimation, most notably the degree of openness which we assume to be 12 percent, i.e. the average import-GDP-ratio of the U.S. in our sample. We estimate the values of nine parameters and find that the estimated model is able to replicate the VAR evidence fairly well for plausible parameter values. Three estimates are particularly noteworthy: a low value for the trade price elasticity, strong complementarities in price-setting and limited exchange rate pass-through.

In order to explore the role of openness, we compute the effects of a monetary policy shock in an economy that is approximately closed and an economy where imports account for 40 percent of GDP. Relative to the baseline economy, there is hardly any difference in the responses of domestic inflation and absorption in these counterfactual economies. Two reasons are key for this result. First, the es-

[3] Specifically, they estimate the resulting variant of the New Keynesian Phillips curve on the basis of single equation techniques. Importantly, in contrast to our analysis, they assume that all firms engage in local currency pricing.

timated value for the trade price elasticity is close to intertemporal elasticity of substitution, which, according to the results reported by Erceg et al. (2007), prevents openness from altering the dynamics of the New Keynesian baseline model. Second, as exchange rate pass-through is limited, the exchange rate channel is prevented from operating in a quantitatively important way. We find, however, that strategic complementarities in price-setting would, in principle, constitute an important channel through which openness impacts monetary transmission. Specifically, if we increase the exchange rate pass-through from an estimated value of 12 percent to 40 percent, openness has sizeable effects. In this case, moving from the closed to the very open economy increases the effects of a monetary policy shock on domestic inflation by some 25 percent. As an implication for monetary policy, we stress that the joint evolution of trade integration as well as exchange rate pass-through should be monitored closely.

The remainder of this chapter is organized as follows. In section 4.2 we introduce the details of the model economy. Section 4.3 presents time series evidence from the estimated VAR model and discusses the estimation of the DSGE model. In section 4.4, we take a closer look at the role of trade integration for monetary transmission. Section 4.5 concludes.

4.2 Model

In this section we develop a two-country DSGE model to study monetary policy transmission in open economies. Most of the model features are standard and familiar from so-called medium scale DSGE models as put forward, for instance, in Christiano et al. (2005) or Smets and Wouters (2005) in a closed economy context.[4] There is a representative household in each country owning the capital stock which is rented together with labor services to intermediate goods producers on a period-by-period basis. Adjusting the level of investment is costly. International financial markets are assumed to be complete.

We assume that in each country there is a continuum of intermediate good producers operating under monopolistic competition and being constrained in price setting à la Calvo. A fraction of these firms invoices exports in their own currency. Using common terminology, these firms are engaging in ‘producer currency pricing’, or ‘PCP’ for short. The remaining firms are engaging in ‘local currency pricing’, or ‘LCP’, by invoicing domestic sales and exports in the currency of domestic and foreign buyers, respectively. A key aspect of monetary transmission in open economies is the extent of exchange rate pass-through. In our setup it will be smaller, the more pervasive LCP for any given degree of price rigidity.[5]

In each country final goods firms combine domestic and imported intermediate goods to provide

[4]In setting up the model we also draw on earlier work by Chari et al. (2002), Kollmann (2002), Galí and Monacelli (2005) and Corsetti and Pesenti (2005), among others.

[5]See Bergin (2006) for a similar formulation, Betts and Devereux (1996, 2000) for early contributions and Obstfeld and Rogoff (2000) for a critical discussion. Note that in the present model nominal rigidities are critical for limiting the extent of exchange rate pass-through. Corsetti and Dedola (2005) and Gust et al. (2006), in contrast, provide real models of limited exchange rate pass-through.

households with final goods used for consumption and investment purposes. The aggregation technology employed by final goods firms may imply unequal weights of domestic and imported intermediates in the production of final goods—thereby determining the degree of openness. In addition, the aggregation technology induces demand functions for intermediate goods which are characterized by a non-constant price elasticity of substitution (NCES). Such an aggregation technology has recently been advocated by Gust et al. (2006), and Guerrieri et al. (2008) in an open economy context. Importantly, it induces strategic complementarities in price-setting among intermediate good firms not only with respect to domestic, but also with respect to foreign competitors.[6]

In the following we give a formal exposition of the model, discussing in turn the problems of final goods firms, intermediate good firms, and the representative household. We close the model with a feedback rule to characterize monetary policy. As both countries are symmetric, of equal size, and have isomorphic structures, we focus on the domestic economy, i.e. on the 'home' country. When necessary we refer to foreign variables by means of a star superscript.

4.2.1 Final goods firms

Final goods are composites of intermediate goods produced by a continuum of monopolistic competitive firms in both countries. We use $j \in [0,1]$ to index intermediate good firms as well as their products and prices. Final goods firms operate under perfect competition and purchase domestically produced intermediate goods, $A_t(j)$, as well as imported intermediate goods, $B_t(j)$. Final goods, F_t are not traded across countries, but are used for domestic consumption, C_t, investment, I_t, and government spending, G_t. In each period, market clearing requires that $F_t = C_t + X_t + G_t$.

Letting $P_t^A(j)$ denote the domestic price of a domestically produced intermediate good and $P_t^B(j)$ the domestic price of an imported intermediate good, the problem of the representative final goods firm is to produce F_t while minimizing expenditures given by

$$\int_0^1 P_t^A(j)A_t(j)dj + \int_0^1 P_t^B(j)B_t(j)dj \tag{4.1}$$

subject to

$$\left[V_{Dt}^{\frac{\sigma-1}{\sigma}} + V_{Mt}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} - \left[\frac{1}{(1+\eta)\upsilon} - 1\right] = 1, \tag{4.2}$$

where V_{Dt} and V_{Mt} are defined as follows

$$V_{Dt} = \int_0^1 \omega^{\frac{\sigma}{\sigma-1}} \frac{1}{(1+\eta)\upsilon}\left[\frac{(1+\eta)}{\omega}\frac{A_t(j)}{F_t} - \eta\right]^{\upsilon} dj, \tag{4.3}$$

[6]The original closed economy formulation goes back to Dotsey and King (2005) or, more generally, to Kimball (1995). Sbordone (2007) uses a similar technology when discussing the consequences of firm entry for the slope of the New Keynesian Phillips curve. While Gust et al. (2006) and Guerrieri et al. (2008) focus on pass-through and inflation dynamics, respectively, we explore the implications for monetary transmission.

$$V_{Mt} = \int_0^1 (1-\omega)^{\frac{\sigma}{\sigma-1}} \frac{1}{(1+\eta)\upsilon} \left[\frac{(1+\eta)}{(1-\omega)} \frac{B_t(j)}{F_t} - \eta \right]^{\upsilon} dj. \tag{4.4}$$

Our aggregation technology given by (4.2), (4.3) and (4.4) follows Gust et al. (2006) closely. A few remarks concerning key parameters are in order. The trade price elasticity, i.e. the elasticity which measures the extent of substitution from goods produced at home to those produced abroad for a given change in relative prices, is a key parameter for the international transmission mechanism. In our setup it is a function of several parameters and given by

$$\tilde{\sigma} = \frac{-\sigma}{(\sigma(\upsilon-1)-\upsilon)(1+\eta)}. \tag{4.5}$$

The elasticity of substitution between goods produced within the same country is generally time varying. In steady state it is constant and given by

$$\epsilon = \frac{1}{1-\upsilon} \frac{1}{1+\eta}. \tag{4.6}$$

The parameter η plays a crucial role for both elasticities. It provides a measure of how strongly our setup deviates from the special case where the elasticity of substitution is constant (CES), which is nested in our model for $\eta = 0$. Finally, the parameter ω measures the weight of domestically produced goods in final goods in steady state. $1-\omega$ measures the fraction of imports in final goods in steady state and thus corresponds to the import-GDP-ratio.

Optimization behavior of domestic and foreign final goods firms gives rise to demand functions for domestically produced intermediate goods

$$A_t(j) = \frac{\omega}{1+\eta} \left[\left(\frac{P_t^A(j)}{P_t^A} \right)^{\frac{1}{\upsilon-1}} \left(\frac{P_t^A}{\Gamma_t} \right)^{\frac{\sigma}{\sigma(\upsilon-1)-\upsilon}} + \eta \right] F_t, \tag{4.7}$$

$$A_t^*(j) = \frac{1-\omega}{1+\eta} \left[\left(\frac{P_t^{A*}(j)}{P_t^{A*}} \right)^{\frac{1}{\upsilon-1}} \left(\frac{P_t^{A*}}{\Gamma_t^*} \right)^{\frac{\sigma}{\sigma(\upsilon-1)-\upsilon}} + \eta \right] F_t^*, \tag{4.8}$$

where Γ_t is a price index defined below. Global demand for a generic good j is then given by

$$Y_t(j) = A_t(j) + A_t^*(j). \tag{4.9}$$

Note that the demand function includes a linear term if $\eta \neq 0$. As a result, price elasticities of demand and the desired markup of intermediate goods firms will be time-varying, or, in other words, price-setting behavior at the level of intermediate goods firms is characterized by strategic complementarities.

The optimization problem of final goods firm implicitly defines price indices. For further reference, it is useful to explicitly distinguish between the prices charged by LCP and PCP-firms. Therefore, let $P_t^{A,PCP}(j)$ and $P_t^{A,LCP}(j)$ denote the domestic price charged by a domestic intermediate goods firm engaged in PCP and LCP, respectively. Letting $\alpha \in [0,1]$ be the fraction of LCP-firms and $(1-\alpha)$

the fraction of PCP-firms, the domestic producer price index P_t^A and the import prices index P_t^B are given by the following expressions:

$$P_t^A = \left(\int_0^\alpha P_t^{A,LCP}(j)^{\frac{\upsilon}{\upsilon-1}} dj + \int_\alpha^1 P_t^{A,PCP}(j)^{\frac{\upsilon}{\upsilon-1}} dj\right)^{\frac{\upsilon-1}{\upsilon}}, \tag{4.10}$$

$$P_t^B = \left(\int_0^\alpha P_t^{B,LCP}(j)^{\frac{\upsilon}{\upsilon-1}} dj + \int_\alpha^1 P_t^{B,PCP}(j)^{\frac{\upsilon}{\upsilon-1}} dj\right)^{\frac{\upsilon-1}{\upsilon}}. \tag{4.11}$$

The price index for final goods is given by

$$\begin{aligned} P_t &= \frac{1}{1+\eta}\Gamma_t + \frac{\eta}{1+\eta}\omega\left(\int_0^\alpha P_t^{A,LCP}(j)dj + \int_\alpha^1 P_t^{A,PCP}(j)dj\right) \\ &+ \frac{\eta}{1+\eta}(1-\omega)\left(\int_0^\alpha P_t^{B,LCP}(j)dj + \int_\alpha^1 P_t^{B,PCP}(j)dj\right), \end{aligned} \tag{4.12}$$

where

$$\Gamma_t = \left[\omega(P_t^A)^{\frac{(\sigma-1)\upsilon}{\sigma(\upsilon-1)-\upsilon}} + (1-\omega)(P_t^B)^{\frac{(\sigma-1)\upsilon}{\sigma(\upsilon-1)-\upsilon}}\right]^{\frac{\sigma(\upsilon-1)-\upsilon}{(\sigma-1)\upsilon}}. \tag{4.13}$$

Finally, letting S_t denote the nominal exchange rate and assuming that the law of one price holds for PCP-firms, we obtain the following relationships:

$$P_t^{B,PCP}(j) = S_t P_t^{B,PCP*}(j); \quad P_t^{A,PCP}(j) = S_t P_t^{A,PCP*}(j). \tag{4.14}$$

4.2.2 Intermediate good firms

The production of intermediate goods, $Y_t(j)$, is governed by a Cobb-Douglas production function

$$Y_t(j) = K_t(j)^\theta H_t(j)^{1-\theta}, \tag{4.15}$$

where $H_t(j)$ and $K_t(j)$ denote labor and capital employed by firm j. Letting W_t and R_t denote the nominal wage rate and the rental rate of capital, respectively, minimizing costs implies for (nominal) marginal costs

$$MC_t(j) = \frac{W_t H_t(j)}{(1-\theta)Y_t(j)} = \frac{R_t K_t(j)}{\theta Y_t(j)}. \tag{4.16}$$

We assume that price setting is constrained exogenously by a discrete time version of the mechanism suggested by Calvo (1983). Each firm has the opportunity to change its price with a given probability $1-\xi$. Moreover, we assume that when a firm has the opportunity to do so, it sets the new price in order to maximize the expected discounted value of net profits before the realization of shocks in a given period.[7] Firms that do not reoptimize in a certain period index their price to last period's producer price inflation, where the degree of indexation is given by the parameter $\kappa \in [0,1]$.

[7] In other words, period t prices are set conditional on the information period $t-1$, see Christiano et al. (2005).

In setting the new price $P_t^{A,PCP}(j)$, the problem of a generic PCP-firm is given by

$$\max \sum_{k=0}^{\infty} \xi^k E_{t-1} \left(\frac{Q_{t,t+k} Y_{t+k}(j)}{P_{t+k}} \left[P_t^{A,PCP}(j) \prod_{s=1}^{k} \left(\Pi_{t+s-1}^A \right)^{\kappa} - MC_{t+k} \right] \right), \tag{4.17}$$

subject to the demand function (4.9), the production function (4.15) and the optimality condition on factor inputs (4.16).[8] $\Pi_t^A = P_t^A / P_{t-1}^A$ denotes domestic inflation. Profits are discounted with the stochastic discount factor, $Q_{t,t+1}$, implicitly defined below.

The pricing problem of a generic LCP-firm is subject to the same constraints as those of the PCP-firm. It sets two distinct prices for the domestic and foreign market. The domestic price $P_t^{A,LCP}(j)$ is set to solve

$$\max \sum_{k=0}^{\infty} \xi^k E_{t-1} \frac{Q_{t,t+k} A_{t+k}(j)}{P_{t+k}} \left[P_t^{A,LCP}(j) \prod_{s=1}^{k} \left(\Pi_{t+s-1}^A \right)^{\kappa} - MC_{t+k} \right], \tag{4.18}$$

subject to the demand function (4.7), while $P_t^{A,LCP*}(j)$ is set to solve

$$\max \sum_{k=0}^{\infty} \xi^k E_{t-1} \frac{Q_{t,t+k} A_{t+k}^*(j)}{P_{t+k}} \left[S_{t+k} P_t^{A,LCP*}(j) \prod_{s=1}^{k} \left(\Pi_{t+s-1}^B \right)^{\kappa} - MC_{t+k} \right] \tag{4.19}$$

subject to the demand function (4.8).

4.2.3 Households

A representative household allocates consumption expenditures intertemporally on final goods and supplies labor, H_t, to intermediate good firms. The preferences of the household are given by

$$\sum_{t=0}^{\infty} \beta^t \frac{[(C_t - bC_{t-1})^{\mu} (1 - H_t)^{1-\mu}]^{1-\gamma}}{1-\gamma}, \tag{4.20}$$

where β is a time discount factor and $b \in [0, 1)$ measures the extent of consumption habits. The parameters γ and μ are positive constants characterizing preferences.

Households own the domestic capital stock, K_t, which is internationally immobile as are labor services. As in Christiano et al. (2005) it may be costly to adjust the level of investment, I_t. Specifically, the law of motion for capital is given by

$$K_{t+1} = (1 - \delta) K_t + [1 - \Psi(I_t / I_{t-1})] I_t, \tag{4.21}$$

where δ denotes the depreciation rate; restricting $\Psi(1) = \Psi'(1) = 0$ and $\Psi''(1) = \chi > 0$ ensures that the steady state capital stock is independent of investment adjustment costs captured by χ.

A complete set of state-contingent securities is traded at an international level. Letting Ξ_{t+1} denote the period $t+1$ payoff of the portfolio held at the end of period t, the gross short-term nominal interest

[8] In our formulation we implicitly assume that demand for intermediate good j is met at all times.

rate, $(1 + i_t)$, is implicitly defined by $(1 + i_t)^{-1} = E_t Q_{t,t+1}$, while the budget constraint reads as follow

$$W_t H_t + R_t K_t + \Upsilon_t + T_t - P_t (C_t + X_t) = E_t \{Q_{t,t+1} \Xi_{t+1}\} - \Xi_t. \tag{4.22}$$

Υ_t denotes nominal profits earned by monopolistic firms and transferred to households and T_t denotes lump-sum taxes. We assume that government spending is financed entirely through lump-sum taxes: $T_t = P_t G_t$.

We assume that the household decides on consumption and investment expenditures in period t before period-t uncertainty is revealed. Subject to this additional constraint as well as to (4.21) and (4.22), the household maximizes the expected value of (4.20).

4.2.4 Monetary policy

To close the model, we assume that monetary policy is characterized by an interest rate feedback rule as in Clarida et al. (2000). Specifically, we assume for the interest rate

$$i_t = \rho i_{t-1} + (1 - \rho) \left(i + \beta^{-1} \phi_\pi \left(\Pi_t^A - \Pi^A \right) + (4F\beta)^{-1} \phi_y (F_t - F) \right) + \nu_t, \tag{4.23}$$

where letters without time subscript refer to steady state values. The parameter $\rho \in [0, 1]$ captures interest rate smoothing, ϕ_π captures the long-run adjustment of the interest rate to producer price inflation and ϕ_y captures stabilization of domestic absorption.[9] Finally, ν_t represents a zero-mean shock to the short-term interest rate not accounted for by the systematic feedback rule. It thus represents a monetary policy shock.

4.2.5 Model solution

We solve the model numerically by applying standard techniques. Specifically, we use (4.23) together with the linearized first order conditions and constraints of the firms' and household problem as well as their foreign counterparts to determine the equilibrium allocation near the deterministic and symmetric steady state. We use the approximate solution of the model to investigate the effects of monetary policy shocks on the economy. To simplify the analysis, we focus on country differences, i.e. the behavior of a domestic variable relative to its foreign counterpart. Before discussing our strategy to assign parameter values, we briefly turn to the implications of strategic price-setting complementarities for the exchange rate channel of monetary policy transmission.

[9] We assume that monetary policy responds to domestic inflation and absorption, because under this assumption we can identify monetary policy shocks in our VAR model in a way which is consistent with our theoretical model. Note also that in open economy models focusing on domestic inflation rather than CPI-inflation is often preferable from a welfare point of view, see Galí and Monacelli (2005). In addition, our formulation of the interest rate rule (4.23) is meant to facilitate a comparison of the parameter values ϕ_π and ϕ_y to those obtained in the empirical literature on interest rate rules where inflation and interest rate are typically annualized.

4.2.6 The exchange rate channel revisited

Strategic complementarities in price-setting may alter monetary policy transmission in open economies by adding a new dimension to the exchange rate channel. Traditionally, two dimensions of the exchange rate channel have been distinguished (see, for instance, Svensson, 2000). First, under sticky prices, nominal exchange rate changes translate into real exchange rate changes that in turn induce an expenditure switching effect. As a result, exchange rate changes alter the demand for domestic goods and thus affect domestic producer prices. Note that in this case, the exchange rate impacts only indirectly—via demand—on domestic inflation. Second, nominal exchange rate changes feed directly into the prices of imported goods and hence into CPI-inflation. Both effects, however depend on the extent of exchange rate pass-through. If import prices are insulated from exchange rate movements, the exchange rate channel is failing to operate along both dimensions.

Strategic price-setting complementarities add a new dimension to the exchange rate channel. In order to show this formally, we focus on the case where exchange rate pass-through is complete ($\alpha = 0$) and derive a variant of the New Keynesian Phillips curve as an approximation of the intermediate goods firms' price setting problem around a deterministic, zero inflation steady state:

$$E_{t-1}\pi_t = \beta E_{t-1}\pi_{t+1} + \lambda(1-\Psi)E_{t-1}mc_t + \lambda\Psi(1-\omega)\frac{2\omega\tilde{\sigma}}{\epsilon}E_{t-1}q_t, \qquad (4.24)$$

where π_t denotes percentage points of domestic inflation, mc_t measures the percentage deviation of marginal costs from steady state and q_t denotes percentage deviation of the relative price of imports expressed in domestic currency. The coefficient $\lambda = (1-\beta\xi)(1-\xi)\xi^{-1}$ is familiar from the New Keynesian baseline model and provides a measure for the pass-through of marginal costs onto inflation. The coefficient Ψ depends on the extent of strategic complementarities in price-setting and other structural parameters of the model: $\Psi = -1\eta\epsilon(\epsilon(1-\eta)-1)^{-1}$.[10]

The relationship (4.24) governs the dynamics of domestic inflation. Note that if $\eta = 0$, we have $\Psi = 0$ and the term q_t disappears from the Phillips curve. In fact, in this case the Phillips curve takes the form which is well-known from the closed-economy New Keynesian baseline model. Clarida et al. (2001) and Galí and Monacelli (2005) have stressed this isomorphism, i.e. the fact that the form of the Phillips curve for the open economy corresponds to that of the closed economy. This case is nested in our model.

Turning to the case where such complementarities are present ($\eta < 0 \rightarrow \Psi > 0$), we observe that the relative price of imports directly matters for domestic inflation. Consider, for instance, a decrease in the domestic currency price of imports resulting from an exchange rate appreciation. In this case, given strategic price-setting complementarities, domestic producers will find it optimal to lower their prices, because the price charged by foreign competitors is reduced: domestic inflation falls. In addition to the coefficient Ψ, two more parameters govern the strength of this effect. First, the larger

[10] Expression (4.24) abstracts from indexation. In appendix A.1 we derive the New Keynesian Phillips curve considering the general case $\alpha \in [0, 1]$. Guerrieri et al. (2008) provide a derivation under the assumption that $\alpha = 1$.

the trade price elasticity relative to the elasticity of substitution across domestically produced goods ($\tilde{\sigma}/\epsilon$), the stronger the impact of import prices on domestic inflation. Second, the impact will also be stronger, the more open an economy. This follows from imports making up for a larger fraction of the final goods basket, measured by $1 - \omega$.

As a consequence, monetary policy may *directly* impact *domestic* inflation via the exchange rate. A monetary contraction which appreciates the nominal exchange rate and lowers the price of imports reduces domestic inflation. This adds a new dimension to the exchange rate channel, which is not present in models without price-setting complementarities. Its importance, however, depends on the extent of exchange rate pass-through in addition to the parameters discussed above. If import prices are unresponsive to exchange rate changes, the exchange rate channel fails to operate. In order to gauge its importance, we need to quantify the extent of exchange rate pass-through along with other key parameters of the model.

4.3 Estimation

Our model is agnostic as regards the sources of business cycle fluctuations and only allows for monetary policy shocks. Accordingly, by bringing the model to the data, we isolate fluctuations in actual time series which can be attributed to monetary policy shocks. Specifically, we focus on the empirical impulse response functions obtained from a VAR estimated on U.S. time series relative to an aggregate of industrialized countries. We use these statistics to pin down the values of key parameters of the model. Such a limited information approach enables our DSGE model to provide an empirically plausible account of the monetary transmission mechanism.[11]

4.3.1 Empirical impulse response functions

We estimate the VAR on quarterly time series data for the period 1973–2006. We focus on relative variables, i.e. the difference of a variable in the U.S. and its counterpart for an aggregate of industrialized countries, which is meant to proxy for the rest of the world ('ROW' for short), see also Clarida and Gali (1994) and Rogers (1999). Specifically, we consider the log of relative consumption, the log of relative investment, the difference in domestic inflation rates (computed on the basis of the GDP deflator), the difference in short term interest rates, the difference in CPI-inflation rates as well as real net exports for the U.S., where real net exports are defined as the log difference in deflated exports and imports.[12] Letting Y_t denote the vector of endogenous variables, we estimate the structural VAR model

$$A(L)Y_t = \varepsilon_t, \tag{4.25}$$

[11] A natural alternative is to estimate the model using full information techniques. This would require to take a stand of all possible sources of business cycle fluctuations, which we can avoid for the purpose of the present study.

[12] We treat CPI-inflation as the empirical counterpart of the DSGE model's inflation rate for final goods. A detailed description of the data is given in appendix A.2. We remove a constant linear trend from consumption and investment before computing relative variables.

where $A(L) = \sum_{i=0}^{4} A_i L^i, LY_t = Y_{t-1}$ and $E(\varepsilon_t \varepsilon_t') = I$.

In order to identify (relative) monetary policy shocks, we assume that A_0 is lower triangular, i.e. we impose the recursive identification scheme which is frequently employed to study the effects of monetary policy shocks, see Kim (2001) for an open economy context. We attach a structural interpretation only to the innovation in relative short-term interest rates. Hence, what matters for identification is how the other variables in Y_t are ordered relative to this variable, see Christiano et al. (1999). We order relative consumption, relative investment as well as the differential of domestic inflation before and the differential of CPI-inflation and net exports after the short-term interest rate differential. The implied identification assumptions are consistent with our DSGE model: consumption, investment and domestic inflation are predetermined relative to monetary policy shocks, while consumer (i.e. final goods) prices and real net exports are free to adjust immediately. As in the theoretical model, we are allowing monetary policy to adjust the interest rate contemporaneously to changes in domestic inflation and domestic absorption.[13]

Figure 4.1 displays the impulse responses to a monetary policy shock, i.e. an increase by 100 basis points in the U.S. short rate relative to the aggregate of industrialized countries. The solid line shows the point estimate, while the shaded area measures 90 percent confidence bounds obtained from bootstrap sampling. The upper row shows the responses of consumption and investment in relative terms; for both we find a protracted and hump-shaped decline. While consumption falls by roughly 0.3 percent, investment falls by about 1.25 percent, with the maximum effect occurring between three and six quarters after the shock.

Domestic inflation responds somewhat sluggishly; the maximum decline of about 8 basis points is observed five quarters after the shock. According to our point estimate, it takes another 3 to 4 years for inflation to return to its pre-shock level. The shock to the interest rate differential is mildly persistent, with the short rate returning to its pre-shock level after about one year. The response of CPI-inflation is remarkably close to that of domestic inflation, both from a quantitative and a qualitative point of view. Finally, U.S. net exports display a hump-shaped increase with the maximum effect of about 0.2 percent occurring after about a year.

4.3.2 Estimation of general equilibrium model

The second step of the analysis consists in matching empirical and theoretical impulse responses in order to obtain estimates for the parameters of the DSGE model. This approach has gained popularity in closed economy studies of monetary policy transmission following the pioneering work of Rotemberg and Woodford (1997) and Christiano et al. (2005).

[13]Alternative approaches to identify monetary policy shocks in open economy frameworks consider on monetary aggregates and non-recursive identification schemes, see Eichenbaum and Evans (1995), Cushman and Zha (1997) and Kim and Roubini (2000). More recently, Faust and Rogers (2003) and Scholl and Uhlig (2008) use sign restrictions to achieve identification. These studies have typically been concerned with the behavior of the exchange rate in the face of monetary policy shocks and on the importance of the latter to account for fluctuations in the former. In the present chapter, we are not taking up these issues. Instead, we use the VAR responses as a key statistic to pin down parameter values of our DSGE model.

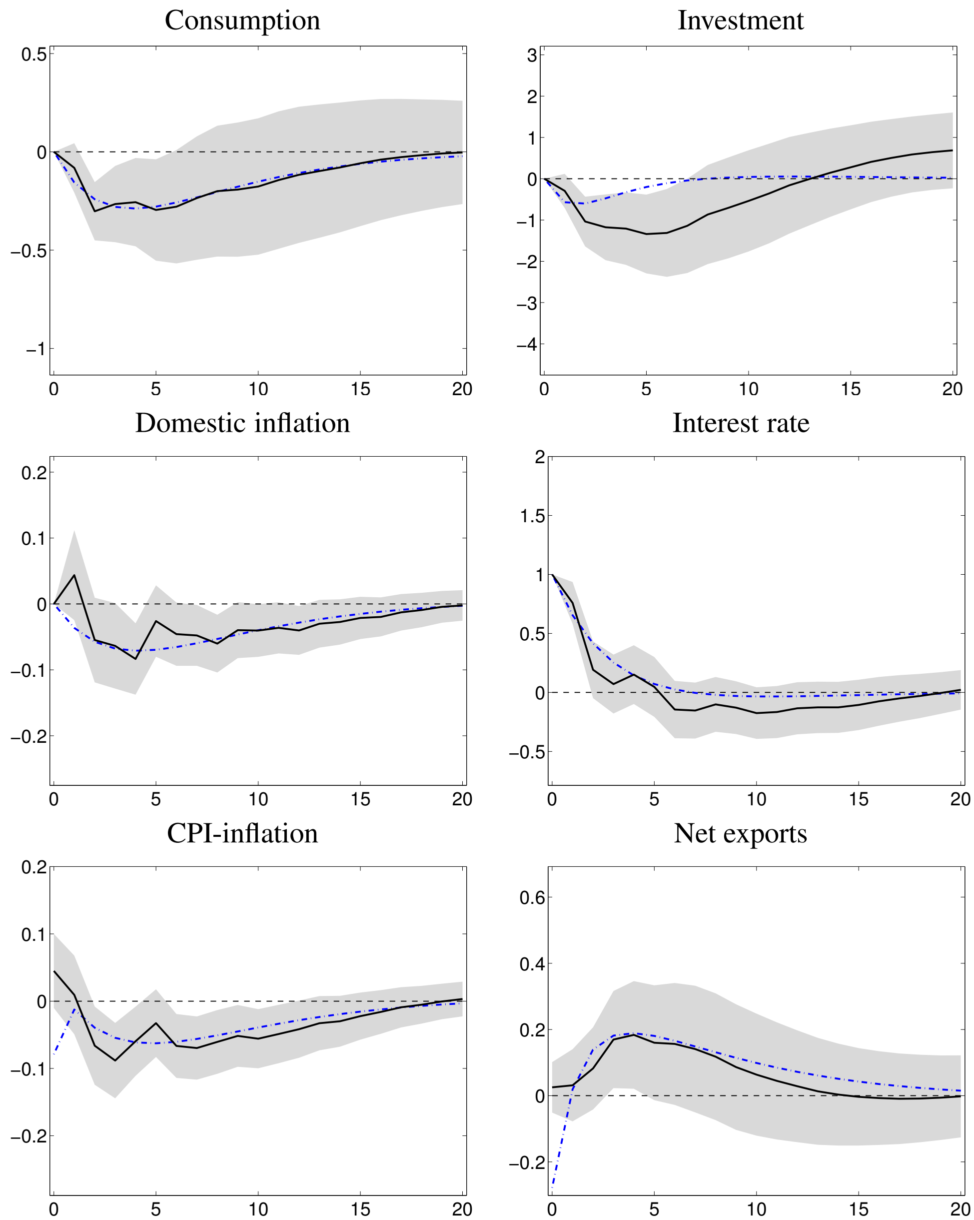


Figure 4.1: Effects of a monetary policy shock

Notes: Shock and responses are in relative terms (U.S. vs. ROW), except for net exports which is the log difference of U.S. exports and imports. Solid line: point estimate; shaded areas: bootstrapped 90 percent confidence intervals; dashed-dotted line: responses of estimated DSGE model; Vertical axes: percent, except for inflation and interest rate (percentage points). Horizontal axes: quarters.

To illustrate this approach, define IR^e to be the empirical impulse response function characterizing the data. The model itself assigns to each admissible vector of structural parameters θ a theoretical impulse response function $IR = IR(\theta)$. We obtain an estimate for the parameter vector of interest, $\widehat{\theta}$, by minimizing the weighted distance between empirical and theoretical impulse response functions, i.e., IR^e and IR:

$$\widehat{\theta} = \arg\min \left(IR^e - IR(\theta)\right)' W \left(IR^e - IR(\theta)\right), \tag{4.26}$$

where W represents a diagonal matrix whose diagonal entries are the reciprocal values of the variance of the empirical impulse responses. Using this weighting matrix ensures that the theoretical impulse responses are made to be as close to the empirical ones as possible, in terms of point-wise standard deviations. Regarding the length of the impulse response functions, we consider 20 quarters starting from the second quarter as most variables return to their steady state within 5 years.

The relationship between structural parameters and the implied impulse response functions is non-linear; we therefore obtain theoretical impulse response functions by applying standard numerical techniques. Note that our procedure only admits saddle path stable solution and thus rules out by construction any parameterization of the model which would give rise to equilibrium indeterminacy. Standard errors for $\widehat{\theta}$ are computed using the following expression for the asymptotic variance of our estimator, taken from Wooldridge (2002):

$$\widehat{Avar}\left(\widehat{\theta}\right) = \left(G'WG\right)^{-1}\left(G'W\widehat{\Sigma}WG\right)\left(G'WG\right)^{-1}. \tag{4.27}$$

where $G = \nabla_{\theta} IR$ represents the Jacobian of the impulse response function generated from the model and $\widehat{\Sigma}$ denotes the variance matrix of the impulse responses obtained from bootstrap sampling.

4.3.3 Parametric setup

In practice, given the number of the structural parameters, it is not possible to identify all of them simultaneously. We therefore fix those parameters prior to the estimation which are either given by first moments of the data or are fairly uncontroversial.

First we set $\omega = 0.88$ which implies an import-to-GDP ratio of 12 percent, the average value for the U.S. in our sample period. Moreover, we set, as, for instance, in Backus et al. (1994) $\beta = 0.99$, $\gamma = 2$ and $\mu = 0.34$ as well as $\theta = 0.36$ and $\delta = 0.025$. In addition, we assume that government spending accounts for 20 percent of GDP, close to the average in our sample period. Regarding price rigidities, we set $\xi = 0.75$, which implies an average duration of prices of one year which is broadly in line with the evidence discussed in Nakamura and Steinsson (2008). We set υ such that the markup earned by intermediate goods firms in steady state is 20 percent.

We are thus left with nine parameters for which we seek to obtain estimates by solving (4.26). We estimate a value for the trade price elasticity, $\tilde{\sigma}$, by adjusting σ according to the relationship (4.5). In addition, we pin down values for the parameters measuring investment adjustment costs, χ, price indexation, κ, habits, b, as well as for those parameters which specify the interest rate feedback rule:

Table 4.1: Estimated parameter values of DSGE model

Parameter	Description	
$\tilde{\sigma}$	Trade price elasticity	0.48 (0.71)
χ	Investment adjustment costs	1.11 (0.75)
κ	Price indexation	1.00 $(-)$
ϕ_π	Inflation coefficient in policy rule	1.00 (0.50)
ϕ_y	Output coefficient in policy rule	0.02 (0.14)
ρ	Interest rate smoothing	0.67 (0.09)
b	Habits	0.89 (0.05)
α	Share of firms with local currency pricing	0.89 (0.15)
η	NCES-parameter	-10.37 (14.30)

Notes: Parameter estimates obtained from matching DSGE and VAR impulse response functions; standard errors are reported in parentheses. Those parameter values which have been estimated to be at their theoretical bounds have been assumed to take this value prior to estimation; in this case no standard error is reported.

ϕ_π, ϕ_y and ρ. Two additional parameters, which are of particular importance for the international monetary transmission mechanism are α, measuring the fraction of LCP-firms and η which is directly related to the degree of strategic price-setting complementarities.

4.3.4 Results

Table 4.1 provides the estimation results. We find plausible point estimates and fairly narrow confidence bounds implied by the standard errors reported in parentheses. The estimated trade price elasticity is below the values often used or found in the literature. Yet several recent studies suggest that a low trade price elasticity may help to account for a larger set of macroeconometric observations, see Lubik and Schorfheide (2006), Kollmann (2006) and de Walque et al. (2005). Also χ, the parameter capturing investment adjustment costs is somewhat below the value reported in Christiano et al. (2005). This is likely to be the result of the aggregation function of final goods, see the discussion in Backus et al. (1994).

In line with earlier research we also find full indexation of prices, see, for instance, Meier and Müller (2006). Regarding monetary policy we find parameter values which imply a fairly loose monetary stance. Note, however, that our solution procedure rules out equilibrium indeterminacy. The degree of interest rate smoothing is in line with previous findings in the literature, see, for instance, Clarida et al. (2000) for the U.S. We find a considerable amount of habits in consumption, somewhat above the values reported in Smets and Wouters (2005) both for the Euro area and the U.S.

For the share of firms engaged in LCP we find a value somewhere between 80 and 99 percent reported

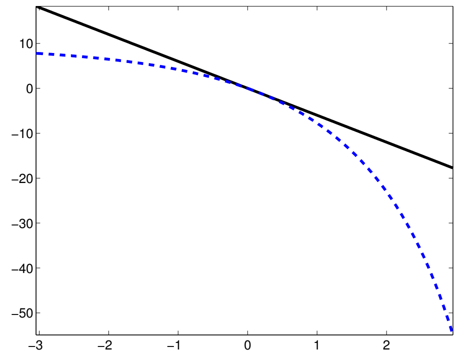


Figure 4.2: Demand function for intermediate goods

Notes: Solid line: CES case ($\eta = 0$); dashed-dotted line: NCES case ($\eta = -11.1$); vertical axes: relative demand in percent; horizontal axes: relative price in percent.

by Campa and Goldberg (2005) and Bergin (2006), respectively for the U.S. Finally, the estimate for the parameter η provides a measure for the curvature of our demand functions. Our estimate is somewhat higher than the values assumed by Gust et al. (2006) and Guerrieri et al. (2008), but close to the value assumed by Smets and Wouters (2007) in a closed economy context.

In order to assess the implication of our estimate for η, we display in Figure 4.2 the percentage change in demand for a generic good (vertical axis) resulting from a percentage change in its relative price (horizontal axis). The dashed line shows the implied demand function for our estimate of η, while the solid line displays the results for $\eta = 0$ implying a constant elasticity of substitution (CES). Relative to the CES case, our estimate implies strongly curved demand functions. As a result, if the relative price increases, demand falls more than proportionally, while, if the relative price falls, demand increases less than proportionally. This induces strategic complementarities in price-setting, which, ceteris paribus, provides firms with an incentive to adjust prices so as to avoid large deviations from the domestic currency price charged by domestic and foreign competitors.

Given the estimated parameter values, we compute the impulse responses of the model and compare them to those obtained from the VAR model. The dashed-dotted lines in the panels of Figure 4.1 show that the model responses track the empirical responses quite closely. All the responses are within the confidence bounds of the VAR responses, except for the impact response of CPI-inflation and net exports. Also the theoretical response of investment is somewhat less pronounced than its empirical counterpart. The response of the consumption differential, as well as those of domestic inflation and the interest rate are matched particularly closely. Overall, we conclude that the DSGE model—if evaluated at the point estimates—provides a quantitatively satisfactory account of the monetary transmission mechanism as apparent for the estimated VAR model.

4.4 The role of openness in monetary policy transmission

In this section we take up the question which motivates our investigation: does trade integration play a quantitatively important role for the transmission of monetary policy? Given that the estimated DSGE model provides a structural and quantitatively realistic account of the monetary transmission mechanism, it is well suited for counterfactual experiments which allow us to quantify the role of openness. We will also briefly explore some implications for monetary policy.

4.4.1 The role of openness

Several quantitative studies have demonstrated that it is possible to account for the actual transmission mechanism while abstracting from foreign trade altogether, see Christiano et al. (2005). At the same time, economies are bound to become more open as a result of increasing trade integration. While the average import share for the U.S. over the period 1973–2006 has been about 12 percent, it has been increasing secularly: from about 6 percent at the beginning of the sample to about 16 percent at the end of the sample. Interestingly, the trend seems to have been accelerating over the last 10 years or so. Against this background, we compare monetary transmission in the estimated model where imports account for 12 percent to two counterfactual scenarios: an approximately closed economy with imports accounting for less than 0.01 percent and a very open economy with imports accounting for 40 percent of final goods.

Figure 4.3 displays impulse responses of domestic inflation (upper row) and domestic absorption (lower row) to a domestic monetary policy shock, i.e. an exogenous increase in the nominal interest rate by 100 basis points. The responses in the left column are computed using the estimated DSGE model where all parameters, except for ω, are kept at their (estimated) baseline values, notably α measuring the fraction LCP-firms. The dashed lines show the responses for the baseline case where imports account for 12 percent of GDP, while solid lines show the responses for the 'closed' economy; the dashed-dotted line shows the responses for the high-openness scenario. Recall that we focus on domestic inflation and absorption, because these variables are well defined in closed-economy models as well.[14] A comparison of the responses reveals that openness matters very little for the transmission of monetary policy shocks in the estimated model (left column).

In a first step to interpret this results, recall that Clarida et al. (2001) and Galí and Monacelli (2005) have shown that there exists an isomorphic representation of the baseline New Keynesian model for closed and open economies. Specifically, the dynamic 'IS-curve' and the New Keynesian Phillips curve have the same structure. Relaxing the closed economy assumption induces only changes in the parameters governing the pass-through of marginal costs onto domestic inflation and the interest elasticity of demand, i.e. it alters only 'slope' coefficients.[15] More specifically, Erceg et al. (2007)

[14] The behavior of CPI inflation and output displays dynamics similar to domestic inflation and absorption, respectively. An exception is the impact period where changes in the nominal exchange rate and net exports dominate the behavior of domestic variables, because the latter are predetermined.

[15] Actually, for certain parameterizations even the difference in the slope coefficients disappears such that 'openness' is

show that the difference between closed and open economies in this class of models can be attributed to the effects of a single composite parameter: the weighted average of the intertemporal elasticity of substitution and the trade price elasticity. As openness determines the relative weights, an increase in openness will alter the dynamic behavior of the economy strongly only if the trade price elasticity differs considerably from the intertemporal elasticity of substitution.

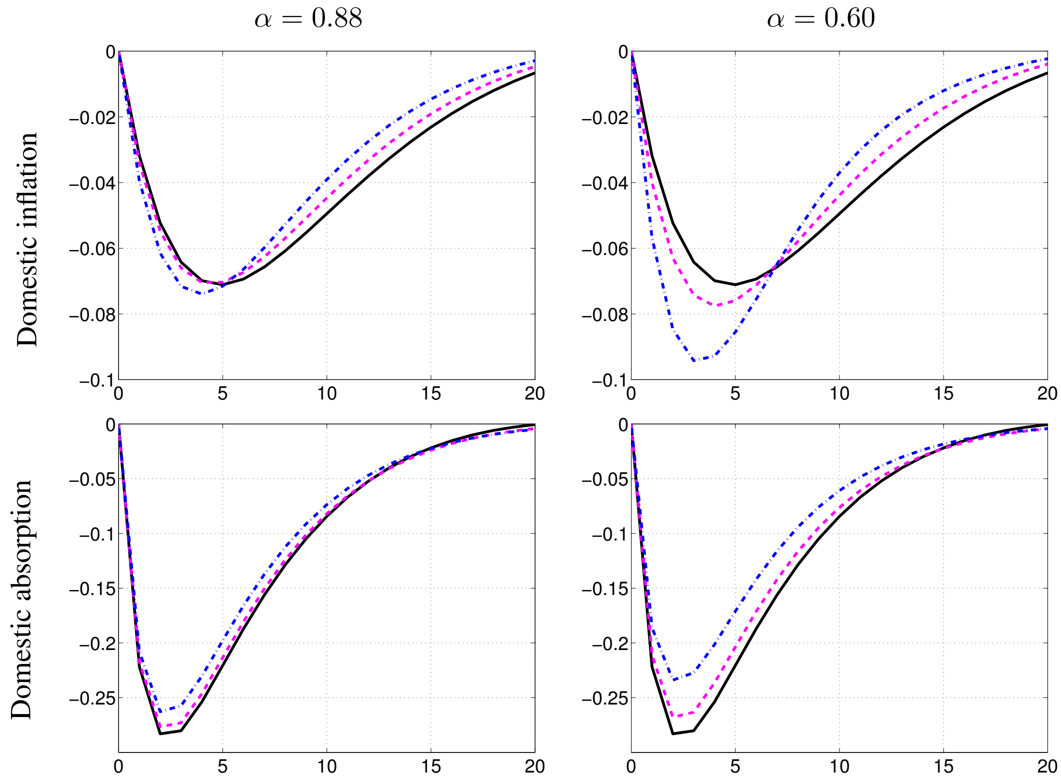


Figure 4.3: Impulse responses to a monetary policy shock

Notes: Shock is exogenous increase in domestic nominal interest rate by 100 basis points; lines show response of domestic variables. Solid line displays responses for zero import share; dashed line: 12 percent import share; dashed-dotted line: 40 percent; all parameter values are kept at the values used or obtained in the estimation of the model.

This result is useful in interpreting our finding. Abstracting from habit formation, our choice of parameter values for μ and γ implies a value for the intertemporal elasticity of substitution for consumption of about $3/4$ which is in the middle of the range of the values discussed in the literature. Our estimate for the trade price elasticity suggests a value which is only slightly lower. It thus appears that because the trade price elasticity and the intertemporal elasticity of substitution are of similar magnitude, openness plays a very limited role in the monetary transmission mechanism.[16]

However, we have so far drawn on a discussion of the New Keynesian baseline model where strategic price-setting complementarities are absent, while we stressed a new dimension of the exchange rate channel emerging under such complementarities, see section 4.2.6. Specifically, in this case openness

merely a source of additional shocks.

[16]In fact, when we increase the trade price elasticity, we find openness to impact more strongly on monetary transmission.

is likely to alter monetary transmission mechanism as it provides monetary policy with direct leverage on domestic inflation. Yet this effect is not evident in the response of domestic inflation displayed in Figure 4.3—despite our estimate for η which suggests strong complementarities.

Yet openness and complementarities are not sufficient for this effect to be present. As stressed above, a third condition is a fair amount of exchange rate pass-through. To see this, consider a monetary contraction: only if the resulting appreciation is reflected in foreign competitors charging lower *domestic currency* prices, will domestic firms find it optimal to lower their prices as well. In this case, there will be downward pressure on domestic inflation due to strategic complementarities, in addition to downward pressure resulting from muted demand and marginal costs.

In principle, this dimension of the exchange rate channel can be quite powerful from a quantitative point of view. This is illustrated in the upper right panel of Figure 4.3, which displays the impulse responses of domestic inflation for the different degrees of openness, assuming a higher degree of exchange rate pass-through: we lower the value of α from our estimate of 0.88 to 0.6. In this case, increasing openness induces a much quicker and stronger fall in domestic inflation. In the open economy (40 percent imports, dashed-dotted line) the response peaks after 3 quarters rather than after 5 quarters in the closed economy. Moreover, the strength of the response increases by some 25 percent.[17]

The lower panels of Figure 4.3 display the response of domestic absorption for all three openness scenarios, both for $\alpha = 0.88$ (left panel) and $\alpha = 0.6$ (right panel). Generally, domestic absorption falls less in response to the monetary policy shock in the more open economy. The effect of openness, however, is considerably more pronounced if the fraction of LCP-firms is lower, i.e. if exchange rate pass-through is higher. To understand this result, recall that while a monetary policy shock is an exogenous increase in the nominal interest rate, what matters for the dynamic adjustment of domestic absorption is the ex ante real interest rate. Its response depends on the dynamics of CPI-inflation which, in turn, will vary with the degree of openness. On impact, CPI-inflation falls more strongly than domestic inflation, because of the exchange rate appreciation. Yet as the exchange rate overshoots, subsequent changes in the exchange rate tend to raise CPI-inflation relative to domestic inflation—thereby dampening the rise in the real rate. Hence, the fall in domestic absorption is less pronounced in more open economies. Again, this effect is stronger, the more pervasive exchange rate pass-through.

4.4.2 Implications for monetary policy

Assuming strategic complementarities in price setting, monetary policy gains better control over domestic inflation as trade integration increases, at least in principle. A necessary condition is that import prices are not completely isolated from exchange rate movements. Yet our estimates suggest that exchange rate pass-through is fairly limited. Moreover, several recent studies suggest that exchange rate pass-through has been declining over the last one or two decades. Figure 4.4 provides

[17] Interestingly, Erceg et al. (2007) also discuss results for the NCES case. However, they still find that the role of openness (for the transmission of technology shocks) is limited which is likely to be the result of assuming that all firms engage in LCP.

suggestive evidence for recent trends both in trade integration and exchange rate pass-through in the U.S. The left panel displays the import-to-GDP ratio over the period 1973–2006. The right panel displays a reduced-form recursive estimate of exchange rate pass-through for the same period.[18] Our results, suggesting a decline in pass-through over the last 10-15 years, are broadly in line with those obtained in the literature, see, for instance, Marazzi et al. (2005) and Ihrig et al. (2006).

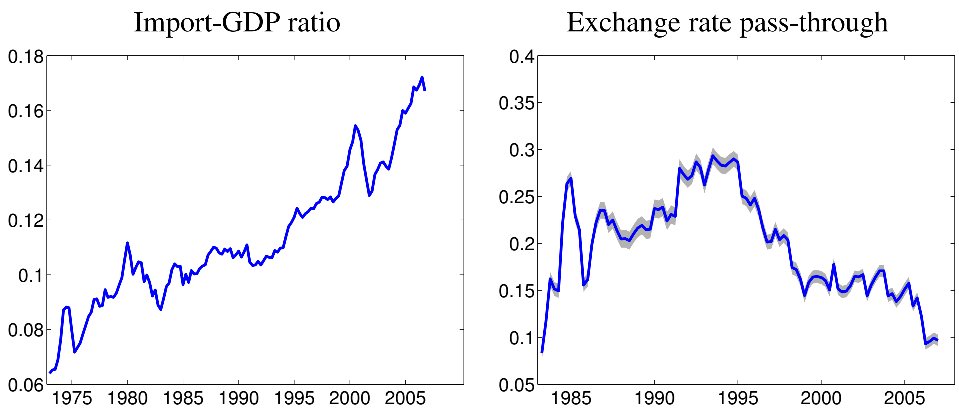


Figure 4.4: Openness and pass-through for the U.S.

Notes: Left panel displays import-GDP ratio; right panel displays reduced form estimate of exchange rate pass-through for 10 year rolling window recursive estimates, shaded area displays two-standard error confidence bounds.

Hence, it appears that although openness is on the rise, pass-through will continue to decline, if current trends prevail. This observation has important implications for monetary policy. To assess this more formally, we compute, as a measure for the trade-off faced by monetary policy, the cumulative reduction in domestic absorption relative to the cumulative reduction in domestic inflation for the first year after a monetary policy shock.[19] Again we consider counterfactual scenarios and compare it to our baseline case: an economy which is approximately closed and an economy where imports account for 40 percent. First, we keep pass-through low (at the value implied by our estimate of $\alpha = 0.88$), but allow, in a last experiment, for higher pass-through by lowering α to 0.6.

Table 4.2 reports the results, which confirm our earlier findings. As a result of strategic price-setting complementarities, monetary policy has *direct* leverage on domestic inflation, which operates irrespectively of a contraction in demand. The more open the economy, the stronger this effect appears. At

[18] As it is not possible to obtain rolling window estimates based on the structural estimation approach employed above, we resort to reduced form estimates. Specifically, similar to Gust et al. (2006) we regress recursively, using a 10 year rolling window, the log-differenced relative import price (measured as the nominal price of non-commodity imports of goods and services divided by the CPI-Index) on the log-differenced real effective exchange rate and a constant.

[19] To be precise about the trade-off faced by monetary policy, it would be necessary to specify an objective for monetary policy. Assuming that monetary policy aims at stabilizing both domestic inflation and the output gap, one may argue that there is no real trade-off in the present model: if both monetary authorities stabilize domestic inflation perfectly, they are likely to stabilize the output gaps as well. However, this is only true in the absence of cost-push shocks, which are typically found to be an important source of business cycle fluctuations, see Smets and Wouters (2007). While our model is agnostic about the sources of business cycle fluctuations, our measure for the monetary policy trade-off might provide some idea of how much reduction in domestic demand is necessary in order to engineer a certain reduction in domestic inflation. Our measure is thus related to the sacrifice ratio, except that we do not consider a permanent reduction in inflation.

Table 4.2: Monetary policy trade-off

$1-\omega$	α	
0.00	0.88	4.8
0.12	0.88	4.5
0.40	0.88	3.9
0.40	0.60	2.6

Notes: Right column measures cumulative reduction in domestic absorption relative to domestic inflation for the first year after monetary policy shock.

the same time, domestic absorption falls by less, because the monetary contraction implies a smaller increase in the real interest rate. Both effects tend to improve our trade-off measure. Yet from a quantitative point of view, this improvement is contained if pass-through is limited—as becomes apparent from the results of the fourth experiment (last row) where pass-through is increased to counterfactually high levels.

It thus appears that, as long as exchange rate pass-through remains limited, increasing trade openness has little bearing on the monetary transmission mechanism and the trade-off faced by monetary policy.[20] As a matter of fact, current trends suggest that while trade integration is increasing, pass-through is decreasing. Yet it is conceivable that both phenomena are interwined at a fundamental level. While the present framework has allowed us to study isolated the effects of features, it seems worthwhile to explore the possibility of a joint cause for both trends in future research.[21]

4.5 Conclusion

In this chapter we explore the role of trade integration for monetary policy transmission. First, we develop a New Keynesian DSGE model featuring two symmetric countries and several frictions which recent business cycle research has found to be important in accounting for several macroeconometric observations. In addition, following Gust et al. (2006), Sbordone (2007) and Guerrieri et al. (2008), we assume a fairly general aggregation technology for final goods. It induces strategic complementarities in price-setting with respect to domestic and foreign competitors such that domestic firms will find it optimal to adjust their prices in response to exchange rate changes which alter the domestic currency price of imports—a new dimension of the exchange rate channel by which monetary policy gains direct leverage over domestic inflation.

In order to quantify the effects of openness on monetary transmission, we estimate, in a first step, a

[20]Erceg et al. (2007) simulate the reduction of the inflation target incorporated in an interest rate feedback rule using the SIGMA model of the FED. They compute the sacrifice ratio for different degrees of openness finding no important role for the latter. Note, however, that while they assume strategic complementarities in price-stetting, they also assume LCP such that import prices are isolated from exchange rate changes in the short-run.

[21]Dornbusch (1987) argues that the extent of exchange rate pass-through and goods market integration are jointly determined. Gust et al. (2006) also link trade integration and exchange rate pass-through in a framework with strategic complementarities. However, they abstract from nominal rigidities.

VAR on U.S. time series relative to an aggregate of industrialized countries. We identify monetary policy shocks by imposing an identification scheme which is consistent with our theoretical model and trace out the transmission mechanism through impulse response functions. In a second step, we find parameter values of the DSGE model by matching its impulse responses to those obtained from the VAR. We find that the estimated model is generally able to mimic the empirical response functions quite closely. Importantly, for the model to do so, we require a low value for the trade price elasticity and the exchange rate pass-through, but strong complementarities in price-setting.

In a third step, we compare the effects of a monetary policy shock in the estimated model where imports account for 12 percent of final goods to two alternative scenarios: an economy which is approximately closed and one in which imports account for 40 percent. We find the effects on domestic inflation and absorption to be almost identical. Closer inspection reveals two reasons underlying this finding. First, the estimated value of the trade price elasticity is close to the intertemporal elasticity of substitution. In this case, openness has been shown to induce little change in the New Keynesian baseline model, see Erceg et al. (2007). Second, as regards the new dimension of the exchange rate channel, we find that limited exchange rate pass-through prevents it from having strong quantitative effects. If we repeat our experiment while assuming higher exchange rate pass-through, the effects of monetary policy shocks become considerably stronger.

Finally, turning to the implications for monetary policy, we stress that while increasing openness could, in principle, improve the trade-off faced by monetary policy, such a development is likely to be prevented by low exchange rate pass-through. At current trends, it appears that while trade integration, or openness, is on the rise, exchange rate pass-through is declining as far as major industrialized countries are concerned. We conclude that while policy makers should keep a close eye on the joint development of openness and exchange rate pass-through, future research may investigate possible causes underlying these trends.

References

Adolfson, M., Laséen, S., Lindé, J., Villani, M., 2007. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Journal of International Economics 72(2), 481–511.

Amato, J. D., Laubach, T., 2003. Estimation and control of an optimization-based model with sticky prices and wages. Journal of Economic Dynamics and Control 27(7), 1181–1215.

Backus, D., Kehoe, P., Kydland, F., March 1994. Dynamics of the trade balance and the terms of trade: the J-curve? American Economic Review 84 (1), 84–103.

Bergin, P. R., 2006. How well can the new open economy macroeconomics explain the exchange rate and current account? Journal of International Money and Finance 25(5), 675–701.

Bernanke, B. S., 2007. Globalization and monetary policy, speech at the Fourth Economic Summit, Stanford Institute for Economic Policy Research.

Betts, C., Devereux, M. B., 1996. The exchange rate in a model of pricing-to-market. European Economic Review 40(3-5), 1007–1021.

Betts, C., Devereux, M. B., 2000. Exchange rate dynamics in a model of pricing-to-market. Journal of International Economics 50(1), 215–244.

Bovin, J., Giannoni, M., 2006. Has monetary policy become more effective? The Review of Economics and Statistics 88(3), 445–462.

Calvo, G., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12(3), 383–398.

Campa, J. M., Goldberg, L. S., 2005. Exchange rate pass-through into import prices. The Review of Economics and Statistics 87(4), 679–690.

Chari, V. V., Kehoe, P. J., McGrattan, E. R., 2002. Can sticky price models generate volatile and persistent real exchange rates? Review of Economics Studies 69(3), 533–563.

Christiano, L. J., Eichenbaum, M., Evans, C. L., 1999. Monetary policy shocks: What have we learned and to what end? In: Taylor, J. B. (Ed.), Handbook of Macroeconomics. Elsevier B.V., pp. 319–347.

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

Clarida, R., Galí, J., Gertler, M., 2001. Optimal monetary policy in open versus closed economies: an integrated approach. American Economic Review 91(2), 248–252.

Clarida, R., Gali, J., 1994. Sources of real exchange-rate fluctuations: How important are nominal shocks? Carnegie-Rochester Conference Series on Public Policy 41, 1–56.

Clarida, R., Galí, J., Gertler, M., February 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. Quarterly Journal of Economics 115(1), 147–180.

Corsetti, G., Dedola, L., 2005. A macroeconomic model of international price discrimination. Journal of International Economics 67(1), 129–155.

Corsetti, G., Pesenti, P., 2005. International dimensions of optimal monetary policy. Journal of Monetary Economics 52(2), 281–305.

Cushman, D. O., Zha, T., 1997. Identifying monetary policy in a small open economy under flexible exchange rates. Journal of Monetary Economics 39(3), 433–448.

de Walque, G., Smets, F., Wouters, R., 2005. An estimated two-country DSGE model for the Euro area and the US economy, mimeo.

Dornbusch, R., 1987. Exchange rates and prices. American Economic Review 77(1), 93–106.

Dotsey, M., King, R. G., 2005. Implications of state-dependent pricing for dynamic macroeconomic models. Journal of Monetary Economics 52(1), 213–242.

Eichenbaum, M., Evans, C. L., 1995. Some empirical evidence on the effects of shocks to monetary policy on exchange rates. The Quarterly Journal of Economics 110(4), 975–1009.

Erceg, C., Gust, C., López-Salido, D., 2007. The transmission of domestic shocks in open economies, mimeo.

Fagan, G., Henry, J., Mestre, R., 2001. An area-wide model (AWM) for the euro area, ECB Working Paper 42.

Faust, J., Rogers, J. H., 2003. Monetary policy's role in exchange rate behavior. Journal of Monetary Economics 50(7), 1403–1424.

Galí, J., Monacelli, T., 2005. Monetary policy and exchange rate volatility in a small open economy. Review of Economic Studies 72(3), 707–734.

Guerrieri, L., Gust, C., López-Salido, D., 2008. International competition and inflation: A New Keynesian perspective, International Finance Discussion Papers, 918.

Gust, C., Leduc, S., Vigfusson, R. J., 2006. Trade integration, competition, and the decline in exchange-rate pass-through., International Finance Discussion Papers, Number 864, Board of Governors of the Federal Reserve System.

Ihrig, J. E., Marazzi, M., Rothenberg, A. D., 2006. Exchange-rate pass-through in the G-7 countries, international Finance Discussion Papers 851.

IMF, 2007. World economic outlook database.

Kim, S., 2001. International transmission of U.S. monetary policy shocks: Evidence from VAR's. Journal of Monetary Economics 48(2), 339–372.

Kim, S., Roubini, N., 2000. Exchange rate anomalies in the industrial countries: A solution with a structural VAR approach. Journal of Monetary Economics 45(3), 561–586.

Kimball, M., 1995. The quantitative analytics of the basic monetarist model. Journal of Money, Credit and Banking 27(4), 1241–1277.

Kollmann, R., 2002. Monetary policy rules in the open economy: Effects on welfare and business cycles. Journal of Monetary Economics 49(5), 989–1015.

Kollmann, R., 2006. International portfolio equilibrium and the current account. CEPR Discussion Paper 5512.

Lubik, T., Schorfheide, F., 2006. A bayesian look at new open economy macroeconomics. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2005, Volume 20. MIT Press, Cambridge MA, pp. 313–366.

Marazzi, M., Sheets, N., Vigfusson, R., Faust, J., Gagnon, J., Marquez, J., Martin, R., Reeve, T., Rogers, J., 2005. Exchange rate pass-through to U.S. import prices: Some new evidence, International Finance Discussion Papers 833.

Meier, A., Müller, G. J., 2006. Fleshing out the monetary transmission mechanism: Output composition and the role of financial frictions. Journal of Money, Credit and Banking 38, 2099–2134.

Mishkin, F. S., 2007. Globalization, macroeconomic performance, and monetary policy (speech). BIS Review 108.

Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models, mimeo Columbia University.

Obstfeld, M., Rogoff, K., 2000. New directions for stochastic open economy models. Journal of International Economics 50(1), 117–153.

OECD, 2007. SourceOECD. http://www.sourceoecd.org/.

Rogers, J. H., 1999. Monetary shocks and real exchange rates. Journal of International Economics 49(2), 269–288.

Rotemberg, J. R., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy. In: Taylor, J. B. (Ed.), NBER Macroeconomics Annual. Cambridge, MA.: MIT Press, pp. 297–346.

Sbordone, A. M., 2007. Globalization and inflation dynamics: the impact of increased competition, Federal Reserve Bank of New York.

Scholl, A., Uhlig, H., 2008. New evidence from the puzzles: Results from agnostic identification on monetary policy and exchange rates, *Journal of International Economics*, forthcoming.

Smets, F., Wouters, R., 2005. Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach. Journal of Applied Econometrics 20(2), 161–183.

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

Svensson, L. E., 2000. Open-economy inflation targeting. Journal of International Economics 50(1), 155–183.

Yellen, J. L., 2006. Monetary policy in a global environment, speech at The Euro and the Dollar in a Globalized Economy Conference, U.C. Santa Cruz, Stanta Cruz, CA.

Appendix

A.1 The New Keynesian Phillips curve

In the following, we go through the main steps of deriving the New Keynesian Phillips curve equation (4.24). We split the derivation into 3 parts. In part one we solve the pricing problem of a generic intermediate good LCP-firm in the domestic market (eq. 4.18). Part 2 solves the pricing problem of a generic intermediate good PCP-firm in the domestic market (eq. 4.17). In part 3 we bring the first parts together using the first order approximation of the definition of the producer price index.

A.1.1 Pricing problem of LCP-firm

Defining $I_{t+k} = \prod_{s=1}^{k} (\Pi^A_{t+s-1})^{\kappa}$ and maximizing equation (4.18) subject to the demand function (4.7), we derive the following first order condition

$$E_{t-1} \sum_{k=0}^{\infty} \xi^k Q_{t,t+k} \left(P_{t+k}\right)^{-1} I_{t+k} \left[1 - \left(1 - \frac{MC_{t+k}}{I_{t+k} P_t^{A,LCP}(j)}\right) \epsilon_{t+k}(j)\right] A_{t+k}(j) = 0, \tag{4.28}$$

where the elasticity of demand for good j in the domestic market is

$$\epsilon_{t+k}(j) = \frac{1}{1-\upsilon} \left[1 + \eta \left(\frac{P_t^{A,LCP}(j) I_{t+k}}{P^A_{t+k}}\right)^{\frac{1}{1-\upsilon}} \left(\frac{P^A_{t+k}}{\Gamma_{t+k}}\right)^{\frac{-\sigma}{\sigma(\upsilon-1)-\upsilon}}\right]^{-1}. \tag{4.29}$$

Rewriting equation (4.28) using the definition of real marginal cost $MC_t^R = \frac{MC_t}{P_t^A}$, defining the contract price as $P_t^{AQ,LCP}(j) = \frac{P_t^{A,LCP}(j)}{P_t^A}$ and linearizing gives

$$\begin{aligned} E_{t-1}\left[\widehat{P}_t^{AQ,LCP}(j)\right] &= \sum_{k=1}^{\infty} (\beta\xi)^k E_{t-1}\left[\widehat{\Pi}^A_{t+s} - \kappa \widehat{\Pi}^A_{t+s-1}\right] \\ &+ (1-\beta\xi) \sum_{k=0}^{\infty} (\beta\xi)^k E_{t-1}\left[\widehat{MC}^R_{t+k} - \frac{1}{\epsilon - 1}\widehat{\epsilon}_{t+k}(j)\right]. \end{aligned}$$

In the above equation all variables are expressed in log-deviations from steady-state. Log-linearizing the elasticity of demand for good j equation (4.29), with $\Gamma_t^Q = \frac{\Gamma_t}{P_t^A}$, we get

$$\widehat{\epsilon}_{t+k}(j) = -\eta\epsilon \left(\widehat{P}_t^{AQ,LCP}(j) - \sum_{k=1}^{\infty} \left(\widehat{\Pi}^A_{t+s} - \kappa \widehat{\Pi}^A_{t+s-1}\right)\right) + \eta\tilde{\sigma}\widehat{\Gamma}^Q_{t+k}. \tag{4.30}$$

Substituting this expression for the demand elasticity in the first order condition, we have

$$E_{t-1}\left[\widehat{P}_t^{AQ,LCP}(j)\right] = \sum_{k=1}^{\infty}(\beta\xi)^k E_{t-1}\left[\widehat{\Pi}_{t+s}^A - \kappa\widehat{\Pi}_{t+s-1}^A\right]$$
$$+\left(\frac{1-\beta\xi}{1-\frac{\eta\epsilon}{\epsilon-1}}\right)\sum_{k=0}^{\infty}(\beta\xi)^k E_{t-1}\left[\widehat{MC}_{t+k}^R - \frac{\eta\epsilon}{\epsilon-1}\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}_{t+k}^Q\right].$$

Using the definition of the steady state markup $\mu = \frac{\epsilon}{\epsilon-1}$ and the definition of $\Psi = \frac{-\eta\mu}{1-\eta\mu}$, this expression after quasi-differencing can be written as

$$E_{t-1}\left[\widehat{P}_t^{AQ,LCP}(j) - \beta\xi\widehat{P}_{t+1}^{AQ,LCP}(j)\right] = \beta\xi E_{t-1}\left(\widehat{\Pi}_{t+1}^A - \kappa\widehat{\Pi}_t^A\right)$$
$$+(1-\beta\xi)E_{t-1}\left[(1-\Psi)\widehat{MC}_t^R + \Psi\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}_t^Q\right].$$

The log-linearized version of the competitive price index equation (4.13) in the domestic country implies that

$$\widehat{\Gamma}_t^Q = (1-\omega)\widehat{q}_t, \tag{4.31}$$

where $q_t = \frac{P_t^B}{P_t^A}$ is the relative import price in domestic currency. Using this to substitute for the relative competitive price index above we get

$$E_{t-1}\left[\widehat{P}_t^{AQ,LCP}(j) - \beta\xi\widehat{P}_{t+1}^{AQ,LCP}(j)\right] = \beta\xi E_{t-1}\left(\widehat{\Pi}_{t+1}^A - \kappa\widehat{\Pi}_t^A\right)$$
$$+(1-\beta\xi)E_{t-1}\left[(1-\Psi)\widehat{MC}_t^R + \Psi\frac{\tilde{\sigma}}{\epsilon}(1-\omega)\widehat{q}_t\right].$$

A.1.2 Pricing problem of PCP-firm

We can derive a similar expression for the PCP-firms. Maximizing equation (4.17) subject to the demand function (4.9), we derive the following first order condition:

$$E_{t-1}\sum_{k=0}^{\infty}\xi^k Q_{t,t+k}\left(P_{t+k}\right)^{-1} I_{t+k}\left[Y_{t+k} - \left(1 - \frac{MC_{t+k}}{I_{t+k}P_t^{A,PCP}(j)}\right)\left(\epsilon_{t+k}^H(j)A_{t+k}(j) + \epsilon_{t+k}^F(j)A_{t+k}^*(j)\right)\right] = 0,$$

where the elasticity of demand for good j in the domestic market is similar to the LCP-firms problem

$$\epsilon_{t+k}^H(j) = \frac{1}{1-\upsilon}\left[1 + \eta\left(\frac{P_t^{A,PCP}(j)I_{t+k}}{P_{t+k}^A}\right)^{\frac{1}{1-\upsilon}}\left(\frac{P_{t+k}^A}{\Gamma_{t+k}}\right)^{\frac{-\sigma}{\sigma(\upsilon-1)-\upsilon}}\right]^{-1}, \tag{4.32}$$

and the elasticity of demand for good j in the foreign market is given by

$$\epsilon^F_{t+k}(j) = \frac{1}{1-\upsilon}\left[1+\eta\left(\frac{P^{A,PCP}_t(j)I_{t+k}}{S_{t+k}P^{A*}_{t+k}}\right)^{\frac{1}{1-\upsilon}}\left(\frac{P^{A*}_{t+k}}{\Gamma^*_{t+k}}\right)^{\frac{-\sigma}{\sigma(\upsilon-1)-\upsilon}}\right]^{-1}. \tag{4.33}$$

Linearizing the first order condition of the firms problem using $P^{AQ,PCP}_t(j) = \frac{P^{A,PCP}_t(j)}{P^A_t}$ gives

$$\begin{aligned}
&E_{t-1}\left[\widehat{P}^{AQ,PCP}_t(j)\right] = \sum_{k=1}^{\infty}(\beta\xi)^k E_{t-1}\left[\widehat{\Pi}^A_{t+s} - \kappa\widehat{\Pi}^A_{t+s-1}\right] \\
&+(1-\beta\xi)\sum_{k=0}^{\infty}(\beta\xi)^k E_{t-1}\left[\widehat{MC}^R_{t+k} - \frac{1}{\epsilon-1}\omega\widehat{\epsilon}^H_{t+k}(j) - \frac{1}{\epsilon-1}(1-\omega)\widehat{\epsilon}^F_{t+k}(j)\right].
\end{aligned}$$

Linearizing both demand elasticities defining $\Gamma^{Q*}_t = \frac{\Gamma^*_t}{P^{A*}_t}$ and the law-of-one-price gap as $q^{A*}_t = \frac{S_t P^{A*}_t}{P^A_t}$ gives

$$\begin{aligned}
\widehat{\epsilon}^H_{t+k}(j) &= -\eta\epsilon\left(\widehat{P}^{AQ,PCP}_t(j) - \sum_{k=1}^{\infty}\left(\widehat{\Pi}^A_{t+s} - \kappa\widehat{\Pi}^A_{t+s-1}\right)\right) + \eta\tilde{\sigma}\widehat{\Gamma}^Q_{t+k}, \\
\widehat{\epsilon}^F_{t+k}(j) &= -\eta\epsilon\left(\widehat{P}^{AQ,PCP}_t(j) - \sum_{k=1}^{\infty}\left(\widehat{\Pi}^A_{t+s} - \kappa\widehat{\Pi}^A_{t+s-1}\right) - \widehat{q}^{A*}_{t+k}\right) + \eta\tilde{\sigma}\widehat{\Gamma}^{Q*}_{t+k}.
\end{aligned}$$

Substituting the demand elasticities into the first order condition and simplifying yields

$$\begin{aligned}
&E_{t-1}\left[\widehat{P}^{AQ,PCP}_t(j)\right] = \sum_{k=1}^{\infty}(\beta\xi)^k E_{t-1}\left[\widehat{\Pi}^A_{t+s} - \kappa\widehat{\Pi}^A_{t+s-1}\right] \\
&+(1-\beta\xi)\sum_{k=0}^{\infty}(\beta\xi)^k E_{t-1}\left[(1-\Psi)\widehat{MC}^R_{t+k} + \Psi\omega\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}^Q_{t+k} + \Psi(1-\omega)\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}^{Q*}_{t+k} + \Psi(1-\omega)\widehat{q}^{A*}_{t+k}\right].
\end{aligned}$$

After quasi-differencing, the expression can be rewritten as

$$\begin{aligned}
&E_{t-1}\left[\widehat{P}^{AQ,PCP}_t(j) - \beta\xi\widehat{P}^{AQ,PCP}_{t+1}(j)\right] = \beta\xi E_{t-1}\left(\widehat{\Pi}^A_{t+1} - \kappa\widehat{\Pi}^A_t\right) \\
&+(1-\beta\xi)E_{t-1}\left[(1-\Psi)\widehat{MC}^R_t + \Psi\omega\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}^Q_t + \Psi(1-\omega)\frac{\tilde{\sigma}}{\epsilon}\widehat{\Gamma}^{Q*}_t + \Psi(1-\omega)\widehat{q}^{A*}_t\right].
\end{aligned}$$

One can linearize the competitive price index in the foreign country analogously to the one in the home country defining the relative export price in foreign currency as $q^{B*}_t = \frac{P^{A*}_t}{P^{B*}_t}$:

$$\widehat{\Gamma}^{Q*}_t = -\omega\widehat{q}^{B*}_t \tag{4.34}$$

Using this expression and equation (4.31) to substitute for the relative competitive price indices above we get

$$
\begin{aligned}
&E_{t-1}\left[\widehat{P}_t^{AQ,PCP}(j)-\beta\xi\widehat{P}_{t+1}^{AQ,PCP}(j)\right]=\beta\xi E_{t-1}\left(\widehat{\Pi}_{t+1}^{A}-\kappa\widehat{\Pi}_t^{A}\right)\\
&+(1-\beta\xi)E_{t-1}\left[(1-\Psi)\widehat{MC}_t^{R}+\Psi(1-\omega)\omega\frac{\tilde{\sigma}}{\epsilon}\widehat{q}_t^{B}-\Psi(1-\omega)\omega\frac{\tilde{\sigma}}{\epsilon}\widehat{q}_t^{B*}+\Psi(1-\omega)\widehat{q}_t^{A*}\right].
\end{aligned}
$$

A.1.3 New Keynesian Phillips Curve

The log-linearized version of the producer price index, equation (4.10), reads as

$$
\alpha\widehat{P}_t^{AQ,LCP}(j)+(1-\alpha)\widehat{P}_t^{AQ,PCP}(j)=\frac{\xi}{1-\xi}\left(\widehat{\Pi}_t^{A}-\kappa\widehat{\Pi}_{t-1}^{A}\right). \tag{4.35}
$$

Using the final equations in the two subsections above to substitute for the contract prices of LCP- and PCP-firms one finally obtains a general formulation for the New Keynesian Phillips curve:

$$
\begin{aligned}
&E_{t-1}\left(\widehat{\Pi}_t^{A}-\kappa\widehat{\Pi}_{t-1}^{A}\right)=\beta E_{t-1}\left(\widehat{\Pi}_{t+1}^{A}-\kappa\widehat{\Pi}_t^{A}\right)\\
&+\lambda E_{t-1}(1-\Psi)\widehat{MC}_t^{R}\\
&+\Psi\left((1-\omega)(\alpha+(1-\alpha)\omega)\frac{\tilde{\sigma}}{\epsilon}\widehat{q}_t^{B}-(1-\omega)\omega(1-\alpha)\frac{\tilde{\sigma}}{\epsilon}\widehat{q}_t^{B*}+(1-\omega)(1-\alpha)\widehat{q}_t^{A*}\right)
\end{aligned}
$$

with $\lambda=(1-\beta\xi)(1-\xi)\xi^{-1}$.

The special cases with $\alpha=0$ and $\eta=0$ are discussed in section 4.2.6. Here we briefly discuss the case of incomplete pass-through ($0<\alpha<1$) and strategic complementarities in price setting ($\eta<0$). In addition to the closed economy Phillips curve or the open economy Phillips curve without strategic complementarities three additional terms show up: $\widehat{q}_t^{B}$, $\widehat{q}_t^{B*}$ and $\widehat{q}_t^{A*}$. We discuss the underlying economics in turn focusing on a monetary contraction which appreciates the nominal exchange rate.
A reduction of the relative import price $\widehat{q}_t^{B}$, induces domestic LCP firms to reduce their prices as their demand elasticity increases with a decrease of the import price index relative to the domestic price index. Domestic PCP-firms react in a similar way; in addition they adjust their price to changes in the relative export prices.
Following a nominal appreciation, the relative export price of PCP-firms expressed in foreign currency, $\widehat{q}_t^{B*}$, increases. Recall that PCP-firms can adjust export prices only through adjustments in domestic prices which are then translated via the law of one price into foreign currency. Hence, the increase in the export price, puts downward pressure on (domestic currency) price of PCP-firms.
Following a nominal appreciation, the export prices PCP-firms increase relative to the export prices of LCP-firms—in foreign currency terms. This is captured by a decrease in $\widehat{q}_t^{A*}$. As the PCP-firms can adjust their export price only by adjusting their domestic price, this puts additional downward pressure on domestic prices of PCP-firms.

All these effects become stronger with the degree of strategic price-setting complementarities η and the import share $1-\omega$. As stressed in the main text, the effects also depend on the degree of exchange rate pass-through. Note that if there are only LCP-firms ($\alpha = 1$), the last two terms in the New Keynesian Phillips curve drop out and only real marginal cost and the relative import price govern the domestic inflation dynamics. Yet, in this case import prices do not directly respond to exchange rate changes.

A.2 Data

Our data are obtained from the OECD Economic Outlook database, see OECD (2007). The ROW aggregate comprises data for Canada, the U.K., Japan and the Euro area. We use data for private consumption (volume), private fixed investment (excl. stockbuilding, volume), and the deflator for private consumption and the deflator for GDP. The latter series are used to construct the CPI-inflation and domestic inflation, respectively.

To construct a measure for net exports of the U.S., we deflate exports (exports of goods and services, value, local currency) and imports (imports of goods and services, value, local currency) with their deflators (export or import price goods and services, local currency) and compute the log-difference of both series. Measures for the short term interest rates are also obtained from the Economic Outlook database (interest rate, short-term) except for the Euro area. In this case we draw on data (STN) from the Area-Wide Model database of the ECB, see Fagan et al. (2001).

To compute the ROW series, we calculate quarterly growth rates and aggregate these series on the basis of GDP weights (PPP-adjusted, year 2000), based on data from the IMF (2007). To obtain levels, we cumulate aggregated growth rates.

Chapter 5

Estimating Monetary Policy Reaction Functions Using Quantile Regressions

Abstract Monetary policy rule parameters are usually estimated at the mean of the interest rate distribution conditional on inflation and an output gap. This is an incomplete description of monetary policy reactions when the parameters are not uniform over the conditional distribution of the interest rate. I use quantile regressions to estimate parameters over the whole conditional distribution of the federal funds rate. Inverse quantile regressions are applied to deal with endogeneity. Real-time data of inflation forecasts and the output gap are used. I find significant and systematic variations of parameters over the conditional distribution of the interest rate.

Keywords: monetary policy rules, IV quantile regression, real-time data

JEL-Codes: C14, E52, E58

5.1 Introduction

Policy rules of the form proposed by Taylor (1993) to understand the interest rate setting of the Federal Open Market Committee (FOMC) in the late 1980s and early 1990s have been used as a tool to study historical monetary policy decisions. Although estimated versions describe monetary policy in the U.S. quite well, in reality the Federal Reserve does not follow a policy rule mechanically: "The monetary policy of the Federal Reserve has involved varying degrees of rule- and discretionary-based modes of operation over time," (Greenspan, 1997). This raises the question how the FOMC responds to inflation and the output gap during periods that cannot be described accurately by a policy rule. Except anecdotal descriptions of some episodes (e.g. Taylor, 1993; Poole, 2006) there appears to be a lack of studies that analyze deviations from Taylor's rule systematically and quantitatively.

In addition to changes between discretionary and rule-based policy regimes, economic theory provides several reasons for deviating at least at times from a linear policy rule framework. First, asymmetric central bank preferences can lead in an otherwise linear model to a nonlinear policy reaction function (Gerlach, 2000; Surico, 2007; Cukierman and Muscatelli, 2008). A nonlinear policy rule can be optimal when the central bank has a quadratic loss function, but the economy is nonlinear (Schaling, 1999; Dolado et al., 2005). Even in a linear economy with symmetric central bank preferences an asymmetric policy rule can be optimal if there is uncertainty about specific model parameters: Meyer et al. (2001) analyse uncertainty regarding the NAIRU and Tillmann (2010) studies optimal policy with uncertainty about the slope of the Phillips curve. Finally, when interest rates approach the zero lower bound, responses to inflation might increase to avoid the possibility of deflation (Orphanides and Wieland, 2000; Kato and Nishiyama, 2005; Sugo and Teranishi, 2005; Adam and Billi, 2006). Despite these concerns in the empirical literature estimation of linear policy rules prevails with only few exceptions.

Estimated policy rule parameters characterize the conditional mean of the interest rate. Thus, during deviations of the interest rate from a linear policy rule the Federal Reserve sets the interest rate not at its conditional expected value, but at some other part of its conditional distribution. Chevapatrakul et al. (2009) estimate interest rate reactions at various points of its conditional distribution. I extend their work to real-time data, a recent IV quantile method and a gradual adjustment of interest rates. Using real-time data is crucial as the output gap was perceived by the Federal Reserve to be negative in real-time for almost the whole time between 1970 and 1990. I use real-time inflation forecasts from the Greenbook that are at times quite different from ex post realized inflation rates. Using Hausman tests I find significant endogeneity of inflation forecasts and output gap nowcasts and therefore use in addition to quantile regression (QR) inverse quantile regression (IQR) proposed by Chernozhukov and Hansen (2005) to compute consistent parameter estimates. I find that allowing for a structural change in the output gap coefficient in 1979 the remaining parameters are stable for the period 1969 through 2002 confirming the breakpoint test results of Orphanides (2004).

The results indicate that policy parameters fluctuate significantly over the conditional distribution of the federal funds rate. These deviations from the parameter estimates at the conditional mean of the interest rate are systematic: inflation reactions and the interest rate smoothing parameter increase and output gap responses decrease over the conditional distribution of the interest rate. The results are robust to variations in the sample. They indicate that the FOMC has sought to stabilize inflation more and output less when setting the interest rate higher than implied by the estimated policy rule and vice versa. Thus, a fraction of deviations from an estimated linear policy rule are possibly not caused by policy shocks, but by systematic changes in the policy parameters or an asymmetric policy rule.

Having analyzed how the Federal Reserve sets interest rates when deviating from the conditional mean it is of interest whether these deviations are related to the business cycle. I estimate for each observation at which quantile of its conditional distribution the interest rate is located. Knowing the parameters at the mean and at the estimated quantile for each observation of the sample one can

decompose overall deviations of the federal funds rate from a linear policy rule into differences in the inflation reaction, the output gap reaction, the reaction to the lagged interest rate and differences in the constant. I find anticyclical deviations of monetary policy from a linear policy rule with respect to the output gap response for the Volcker-Greenspan era. Together with a decreasing output gap parameter over the conditional distribution of the interest rate one can conclude that the Fed reacted more to the output gap during recessions than during expansions. This leads to lower interest rates than implied by a linear policy rule during recessions. A recession avoidance preference of the FOMC found by Cukierman and Muscatelli (2008) is thus confirmed.
The remainder of this chapter is organized as follows: Section 5.2 presents the real-time dataset. Section 5.3 presents estimation results for standard methods. Afterwards, section 5.4 gives an overview on quantile regression methods. In section 5.5 the quantile regression results are presented and discussed. Section 5.6 links parameter variations to the business cycle. Finally, section 5.7 concludes.

5.2 Data

I use real-time data from 1969 through 2003 that were available at the Federal Reserve at the time of policy decisions.[1] For expected inflation I compute year-on-year inflation forecasts four quarters ahead of the policy decisions using four successive quarter-on-quarter forecasts of the GDP/GNP deflator computed by Federal Reserve staff for the Greenbook.[2] Data sources for output gap nowcasts as used by the Federal Reserve are described by Orphanides (2004) in detail. From 1969 until 1976 output gap estimates were computed by the Council of Economic Advisors. Afterwards the Federal Reserve staff started to compute an own output gap series. The output gap estimates by the Fed were not officially published in the Greenbook, but were used to prepare projections of other variables included in the Greenbook. Finally, the interest rate is measured as the annual effective yield of the federal funds rate.
An important aspect of the analysis is that the different data series correspond exactly to the information available at the dates of the specific FOMC meetings. I use observations of as many FOMC meetings as possible to describe U.S. monetary policy with high accurracy. Therefore, the frequency of the observations is not equally spaced and varies over the sample: data from 1969 to 1971 is annual, the observations for 1972 and 1973 are seminannual, data until 1987 is quarterly and for most years of the remaining sample there is data available for eight FOMC meetings per year. In addition, I create quarterly spaced data for robustness checks. A plot of the data is shown in Figure 5.1. It is noticeable

[1] Greenbook data remains confidential for some years, so I cannot use data after 2003.

[2] To be sure, these forecasts need not to coinicide with the forecasts of the FOMC members. Orphanides and Wieland (2008) use the forecasts of the FOMC members from the semiannual Humphrey-Hawkins Reports to estimate monetary policy rules. I stick to the staff's forecast as the higher frequency of the data is useful to get precise estimates using quantile regression methods. Orphanides (2001) notes that the Greenbook forecast are an useful approximation for the forecast of the FOMC.

that the Fed perceived the output gap to be negative in real-time for large parts of the sample.

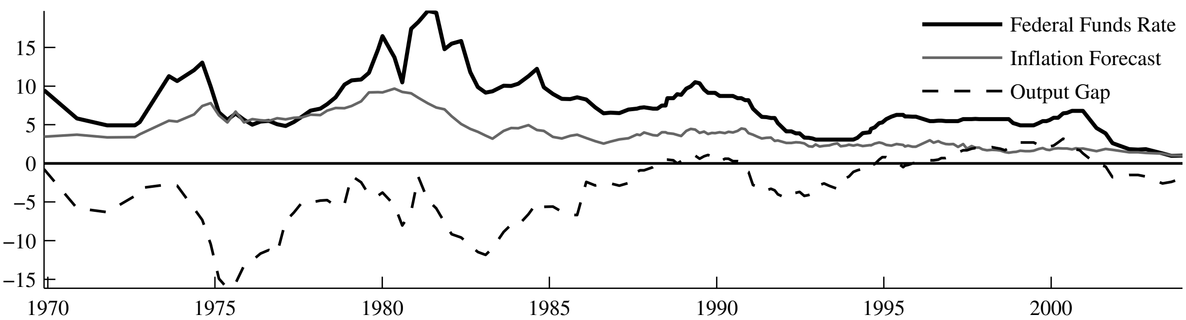


Figure 5.1: Federal funds rate, inflation forecasts and output gap nowcasts
Notes: Inflation forecasts reflect percentage year-over-year changes in the GDP/GNP deflator. Output gap nowcasts measure deviations of real output from potential output in percent. The interest rate is the annual effective yield of federal funds rate.

5.3 Least squares regressions

I estimate a monetary policy rule of the form:

$$i_t = \rho i_{t-1} + (1-\rho)(i^* + \beta(\pi_{t+4|t} - \pi^*) + \gamma y_t) + \epsilon_t, \tag{5.1}$$

where i_t is the nominal short term interest rate, i^* is the targeted nominal rate, $\pi_{t+4|t}$ is a four-quarter-ahead inflation forecast, π^* is the inflation target, y_t is the output gap and ϵ_t is a policy shock. ρ, β and γ are policy parameters. Thus, the federal funds rate responds systematically to deviations of the inflation forecast from a target and to the output gap. The interest rate is adjusted gradually to its target. Orphanides (2001) shows that forward-looking policy rules provide a better description of U.S. monetary policy than backward-looking rules in the sense that they do not violate the Taylor principle when being estimated with real-time data.

The nominal interest rate target can be decomposed into the targeted real interest rate and the inflation target: $i^* = r^* + \pi^*$. To use linear estimation techniques equation (5.1) is rewritten:

$$i_t = \alpha_0 + \alpha_i i_{t-1} + \alpha_\pi \pi_{t+4|t} + \alpha_y y_t + \epsilon_t, \tag{5.2}$$

where $\alpha_0 = (1-\rho)(r^* + (1-\beta)\pi^*)$, $\alpha_i = \rho$, $\alpha_\pi = (1-\rho)\beta$ and $\alpha_y = (1-\rho)\gamma$. Parameters can be estimated at the conditional expected value of the federal funds rate with standard methods like ordinary least squares (OLS) or two-stage least squares (TSLS) to handle endogeneity problems:

$$E(i_t | i_{t-1}, \pi_{t+4|t}, y_t) = \alpha_0 + \alpha_i i_{t-1} + \alpha_\pi \pi_{t+4|t} + \alpha_y y_t. \tag{5.3}$$

5.3.1 Specification tests

Clarida et al. (2000) find using revised data differences in policy rule parameters prior to Paul Volcker's appointment as Fed chairman and afterwards. Orphanides (2004) found using a real-time dataset similar to the one used in this study a more activist policy response to the output gap prior to 1979 than afterwards, but no change in the inflation response. I estimate equation (5.3) and examine restrictions on the constancy of specific parameters to decide on an appropriate specification. Inflation forecasts and output gap nowcasts might be endogenous and therefore specification tests are repeated using TSLS. For the results using TSLS I use lags up to four quarters of the federal funds rate, inflation and the output gap as instruments as in Clarida et al. (2000) and Orphanides (2001). These lagged variables are predetermined and are thus appropriate instruments for the inflation forecast and the output gap nowcast.

Table 5.1: p-values of subsample stability tests

	OLS		TSLS	
Parameters	all data	quarterly data	all data	quarterly data
All	0.12	0.14	0.05	0.06
α_0	0.17	0.15	0.09	0.09
α_π	0.09	0.08	0.03	0.03
α_y	0.02	0.01	0.01	0.01
α_i	0.10	0.11	0.20	0.19
α_0 (α_y varies)	0.67	0.72	0.91	0.78
α_π (α_y varies)	0.95	0.94	0.38	0.43
α_i (α_y varies)	0.81	0.90	0.34	0.49

Notes: The entries show p-values of parameter stability tests across the subsamples 1969:4-1979:2 and 1979:3-2003:4. Test results are shown for all available FOMC meetings and for quarterly data. Row 1 examines the null hypothesis of joint constancy of all parameters. Rows 2-5 test the null hypothesis that the specific parameter shown is constant, under the assumption that remaining parameters are constant. Rows 6-8 test the null hypothesis that the specific parameter shown is constant when α_y is allowed to vary and remaining parameters are constant.

Table 5.1 shows that the null hypothesis of no structural break cannot be rejected. However, as the p-values in the case of the TSLS estimates are close to rejection I investigate if there is a structural break in specific parameters. The hypothesis of no structural breaks in the constant and the interest rate smoothing parameters are accepted, while the evidence is mixed for the inflation parameter. Constancy of the output gap response parameter is rejected in all cases. Allowing this parameter to vary, the null hypothesis of no structural break in all the other parameters is accepted. Based on this, I estimate policy rules over the period 1969:4-2003:4, allowing for a structural change of α_y in 1979:3.

Policy rule estimates using revised data of inflation and the output gap have relied on instrumental variable methods, (see, e.g., Clarida et al., 1998). In contrast, the literature using real-time data has not used instrumental variable methods as inflation forecasts and output gap nowcasts are prepared before the FOMC meetings and are not revised afterwards. However, forecasts might be based on fairly accurate expectations about the policy actions of the FOMC and still a simultaneity problem with the interest rate can arise. I compute Hausman tests to detect possible endogeneity problems:

Table 5.2: p-values of tests for exogeneity

	$\alpha_i = 0$ all data	quarterly data	$\alpha_i \neq 0$ all data	quarterly data
1969:4 - 2003:4	0.00	0.01	0.00	0.00
1969:4 - 1979:2	0.51	0.51	0.45	0.45
1979:3 - 2003:4	0.00	0.00	0.00	0.00
α_y varies	0.00	0.00	0.00	0.00

Notes: The entries show p-values of Hausman tests of the null hypothesis of no endogeneity. Specifications with and without interest rate smoothing are estimated. Rows 1-3 show results for different subsamples. Row 4 shows p-values for the whole sample when the output gap reaction α_y is allowed to change in 1979:3.

The tests results indicate that except for the pre-Volcker subsample endogeneity of inflation expectations and the output gap cannot be rejected at high significance levels. I therefore present results for standard methods and instrumental variable counterparts.

5.3.2 Least squares estimation results

Table 5.3 shows the estimated policy reaction parameters at the conditional mean of the federal funds rate. Results typically found in the real-time policy rule literature are confirmed: the Taylor principle is fulfilled over the whole sample. The reaction to the output gap is high for the first part of the sample while it is close to zero and partly insignificant in the second part. The high inflation of the 1970's might have been caused by the high reaction to the output gap that was perceived to be highly negative in real-time. Interest rate smoothing parameters are high and significant.

Table 5.3: Estimated policy reaction parameters

	$\alpha_i = 0$ OLS	TSLS	$\alpha_i \neq 0$ OLS	TSLS
α_0	1.78	1.38	0.04	0.10
	(0.68)	(0.78)	(0.22)	(0.23)
α_π	1.60	1.72	0.49	0.41
	(0.22)	(0.27)	(0.09)	(0.11)
$\alpha_y : 1969:4-1979:2$	0.44	0.48	0.17	0.14
	(0.12)	(0.14)	(0.04)	(0.05)
$\alpha_y : 1979:3-2003:4$	-0.02	0.00	0.09	0.06
	(0.12)	(0.13)	(0.04)	(0.04)
α_i	-	-	0.78	0.81
			(0.04)	(0.05)

Notes: The entries show estimated parameters together with bootstrapped standard errors in brackets. The estimated equation is $i_t = \alpha_0 + \alpha_i i_{t-1} + \alpha_\pi \pi_{t+4|t} + (\alpha_{y,1} + D\alpha_{y,2})y_t + \epsilon_t$, D is a dummy variable that equals zero until 1979:2 and one afterwards. The output gap coefficients are computed as follows: $\alpha_y = \alpha_{y,1}$ until 1979:2 and $\alpha_y = \alpha_{y,1} + D\alpha_{y,2}$ afterwards.

The estimation results impose the untested restriction that the parameters are the same across the quantiles of the conditional distribution of the federal funds rate. The restriction of parameter constancy across quantiles is testable by estimating equation (5.2) at different quantiles and checking for significant differences in policy reaction parameters at different parts of the conditional distribution of the interest rate.

5.4 Quantile regression

Quantiles are values that divide a distribution such that a given proportion of observations is located below the quantile. The τ^{th} conditional quantile is the value $q_\tau(i_t|i_{t-1}, \pi_{t+4|t}, y_t)$ such that the probability that the conditional interest rate will be less than $q_\tau(i_t|i_{t-1}, \pi_{t+4|t}, y_t)$ is τ and the probability that it will be more than $q_\tau(i_t|i_{t-1}, \pi_{t+4|t}, y_t)$ is $1-\tau$:

$$\int_{-\infty}^{q_\tau(i_t|i_{t-1},\pi_{t+4|t},y_t)} f_{i_t|i_{t-1},\pi_{t+4|t},y_t}(x|i_{t-1}, \pi_{t+4|t}, y_t)dx = \tau, \quad \tau\in(0,1) \tag{5.4}$$

where $f(.|.)$ is a conditional density function. The policy rule at quantile τ can accordingly be written as:

$$q_\tau(i_t|i_{t-1}, \pi_{t+4|t}, y_t) = \alpha_0(\tau) + \alpha_i(\tau)i_{t-1} + \alpha_\pi(\tau)\pi_{t+4|t} + \alpha_y(\tau)y_t. \tag{5.5}$$

Estimating policy parameters at different quantiles instead of the mean can be done with quantile regressions as introduced by Koenker and Basset (1978). Estimating this equation for all $\tau \in (0,1)$ yields a set of parameters for each value of τ and characterizes the entire conditional distribution of the federal funds rate. While preserving the linear policy rule framework, quantile regression imposes no functional form constraints on parameter values over the conditional distribution of the interest rate.

As in the case of least squares, parameters estimated using quantile regression are biased when regressors are correlated with the error term. A two-stage least absolute deviations estimator has been developed by Amemiya (1982) and Powell (1983) and has been extended to quantile regression by Chen and Portnoy (1996). The first stage equals the standard two-stage least squares procedure of regressing the endogenous variables on the exogenous variables and additional instruments. The second stage estimates obtained by quantile regression yield the parameters $\hat{\alpha}_i(\tau)$, $\hat{\alpha}_0(\tau)$, $\hat{\alpha}_\pi(\tau)$ and $\hat{\alpha}_y(\tau)$. However, Chernozhukov and Hansen (2001) show that these estimates are only unbiased if changes in the endogenous variables do not affect the scale or shape of the distribution of the dependent variables, but only shift its location. This assumption is restrictive and excludes interesting cases. It is not fulfilled when estimating policy rules: if inflation decreases and thus interest rates decrease, the shape of the conditional distribution of the interest rate is altered as zero remains the lower bound of the interest rate.

Chernozhukov and Hansen (2001) developed inverse quantile regression that generates consistent estimates without restrictive assumptions.[3] They derive the following moment condition as the main identifying restriction of IQR:

$$P(Y \leq q_\tau(D, X)|X, Z) = \tau, \tag{5.6}$$

where $P(.|.)$ denotes the conditional probability, Y denotes the dependent variable i_t, D a vector of endogenous variables $\pi_{t+4|t}$ and y_t, X a vector of exogenous variables including a constant and i_{t-1} and Z a vector of instrument variables. This equation is similar to the definition of conditional quantiles given above except for conditioning on additional instrument variables. The main assumption for this moment condition is fulfilled if rank invariance holds: it requires that the expected ranking of observations by the level of the interest rate does not change with variations in the covariates. If for example inflation rises, the level of the interest rate would rise for all observations exposed to the change in inflation. Hence, it is likely that the ranking of these observations is not altered by the change in inflation.[4,5]

5.4.1 Inverse quantile regression

IQR transforms equation (5.6) into its sample analogue. The moment condition is equivalent to the statement that 0 is the τ^{th} quantile of the random variable $Y - q_\tau(D, X)$ conditional on (X, Z).[6] Therefore, one needs to find parameters of the function $q_\tau(D, X)$ such that zero is the solution to the quantile regression problem, in which one regresses the error term $Y - q_\tau(D, X)$ on any function of (X, Z). Let $\lambda_D = [\alpha_\pi \;\; \alpha_y]'$ denote the parameters of the endogenous variables and $\lambda_X = [\alpha_0 \;\; \alpha_i]'$ denote a vector of parameters of the exogenous variables and Λ a set of possible values for λ_D. Write the conditional quantile as a linear function: $q_\tau(Y|D, X) = D'\lambda_D(\tau) + X'\lambda_X(\tau)$. The following algorithm implements IQR:[7]

1. First stage regression: regress the endogenous variables on the exogenous variables and additional instruments using OLS. This yields fitted values $\hat{D}$.

[3]Alternatively, one could use a control function approach as in Lee (2004). Results are likely to be similar to IQR. However, using IQR retains the simple structure of Taylor type rules. This facilitates the interpretation of the results. For a comparison of the two approaches see Chernozhukov and Hansen (2005).

[4]A weaker similiarity condition together with some other assumptions discussed in detail in Chernozhukov and Hansen (2001) is sufficient, too. Similarity requires that the distribution of the error term has to be equal for all values of each endogenous variable, holding everything else constant. Rank invariance is a stricter, but in the context of policy rule estimation also more intuitive condition than similarity.

[5]An additional advantage of IQR is that it allows for measurement errors in the instruments. This will be the case in policy rule estimation using real-time data for the instruments as the data is revised later on. However, even using revised data will include measurement errors. Orphanides (2001) notes that mismeasurement is solved for many macroeconomic variables only slowly through redefinitions and rebenchmarks, but most likely never completely. Additionally, the output gap is an unobservable variable in practice and thus the output gap itself is an estimate.

[6]A simple example for unconditional quantiles may help to illustrate this equivalence: consider a sample $Y = \{2, 5, 6, 9, 10\}$ and the quantile at $\tau = 0.4$ that is computed to be $q_{0.4} = 5$. Now compute $Y - q_{0.4} = \{-3, 0, 1, 4, 5\}$. It is clear that 0 is the 0.4 quantile of this expression.

[7]The dependence of the parameters on the quantile τ is omitted in the following equations to keep the notation simple.

2. Second stage regression: estimate for all $\lambda_D \in \Lambda$:

$$[\hat{\lambda}_X(\lambda_D)\ \ \hat{\lambda}_Z(\lambda_D)]' = \arg\min_{\{\lambda_X,\lambda_Z\}} \frac{1}{T}\sum_{t=1}^{T} \varphi_\tau(Y_t - D_t'\lambda_D - X_t'\lambda_X - \hat{D}_t'\lambda_Z), \tag{5.7}$$

where $\varphi_\tau(u) = \tau - 1(u < 0)u$ is the asymmetric least absolute deviation loss function from standard quantile regression (see e.g. Koenker and Basset, 1978) and λ_Z are additional parameters on $\hat{D}$.

3. Inverse step: find $\hat{\lambda}_D$ by minimizing an Euclidian norm of $\hat{\lambda}_Z(\lambda_D)$ over $\lambda_D \in \Lambda$:

$$\hat{\lambda}_D = \arg\min_{\{\lambda_D \in \Lambda\}} \sqrt{\hat{\lambda}_Z(\lambda_D)'\hat{\lambda}_Z(\lambda_D)} \tag{5.8}$$

This minimization ensures that $Y - q_\tau(D, X)$ does not depend on $\hat{D}$ anymore which is the above mentioned function of (X, Z).

Chernozhukov and Hansen (2001) call this procedure the inverse quantile regression as the method is inverse to conventional quantile regression: first, one estimates $\hat{\lambda}_Z(\lambda_D)$ and $\hat{\lambda}_X(\lambda_D)$ by quantile regression for all $\lambda_D \in \Lambda$. The inverse step (5.8) yields the final estimates $\hat{\lambda}_D$, $\hat{\lambda}_Z(\hat{\lambda}_D)$ and $\hat{\lambda}_X(\hat{\lambda}_D)$. The procedure is made operational through numerical minimization methods combined with standard quantile regression estimates. Through increasing τ from 0.01 to 0.99 one traces partial effects over the entire distribution of i_t conditional on i_{t-1}, $\pi_{t+4|t}$ and y_t including all the cases when the central bank deviates from a policy rule estimated at its conditional mean.
Throughout this study stationarity of all variables used in the regressions is assumed. It is reasonable to assume stationarity of the output gap. Using standard Dickey-Fuller tests Clarida et al. (1998) find that the federal funds rate and inflation are at the border between being I(0) and I(1). They proceed to estimate with an I(0) assumption under the argument that the Dickey-Fuller test lacks power in small samples.

5.4.2 Moving blocks bootstrap

Fitzenberger (1997) presents moving blocks bootstrap (MBB) as an estimator for standard errors in quantile regression that is robust to heteroskedasticity and autocorrelation of unknown forms. The MBB is modified in this study for usage with IQR. Following Clarida et al. (1998) the autocorrelation considered is limited to one year. For each bootstrap blocks of the variables are drawn randomly from the whole sample. This includes the dependent variable, the endogenous variables, the exogenous variables and the instruments. For each of the 1000 bootstraps the IQR estimates are computed. Finally, standard errors of the coefficients are computed as the standard deviation of the 1000 estimates of $\alpha_i(\tau)$, $\alpha_0(\tau)$, $\alpha_\pi(\tau)$ and $\alpha_y(\tau)$, respectively.

5.5 Estimation results

Figure 5.2 shows the estimated coefficients of the inflation forecast, the output gap and the constant when restricting α_i to zero. The varying solid black lines show the QR and IQR coefficients over the conditional distribution of the federal funds rate denoted by the quantiles $\tau \in (0, 1)$ on the x-axis. The shaded areas show 95% confidence bands. OLS and TSLS coefficients together with 95% confidence intervals are denoted by straight horizontal lines. The coefficients vary for both the QR and IQR estimates significantly over the conditional distribution of the federal funds rate except for the output gap coefficient in the first subsample.[8] The deviations of the parameter estimates from the OLS and TSLS coefficients reflect persistent deviations of the federal funds rate from a policy rule estimated at the mean. The systematic variations show that at least parts of the deviations from the policy rule are beyond unsystematic policy shocks. The QR and IQR estimation results have qualitative similar patterns over the distribution.

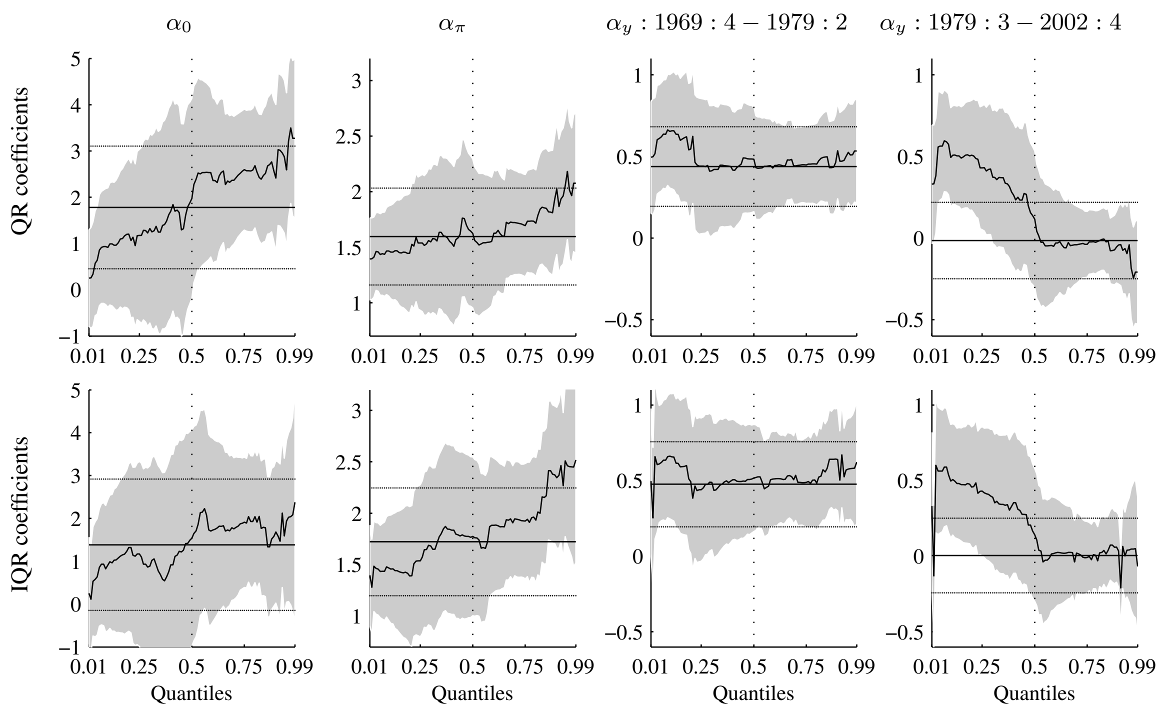


Figure 5.2: Estimated coefficients ($\alpha_i = 0$)

Notes: The solid line in row 1 presents QR estimates and in row 2 IQR estimates of: $i_t = \alpha_0(\tau) + \alpha_\pi(\tau)\pi_{t+4|t} + (\alpha_{y,1}(\tau) + D\alpha_{y,2}(\tau))y_t + \epsilon_t$ for $\tau \in (0, 1)$. See Table 5.3 for a description of the dummy variable D. Shaded areas denote 95% confidence bands of 1000 bootstraps. Solid straight horizontal lines show OLS estimates in row 1 and TSLS estimates in row 2 together with 95% confidence bands.

The estimation results show that the Federal Reserve responded systematically to inflation. The

[8] The significance occurs in two aspects: first, the QR and IQR point estimates lie outside of the OLS and TSLS confidence bands at the lower or upper quantiles. Second, the QR and IQR point estimates of the upper quantiles lie outside the confidence bands of the QR and IQR estimates at the lower quantiles and vice versa.

IQR inflation coefficient is significantly different from zero and increases from 1.5 to 2 (QR) and 2.5 (IQR), respectively, over the distribution satisfying the Taylor principle over the whole distribution. An evaluation of the Taylor principle over the distribution of the interest rate is the focus of Chevapatrakul et al. (2009). The estimation results confirm their finding that the Taylor principle is not violated over the whole conditional distribution of the federal funds rate using real-time instead of revised data and a different IV quantile estimation method. The upper part of the distribution covers periods where the interest rate has been set higher than the least squares policy rule estimates suggest and the lower part periods where it has been set lower. Therefore, the inflation response is stronger when the interest rate is set higher than on average and lower when the interest rate is set lower than on average. While the QR and IQR inflation coefficients are similar at the lower border of the distribution the IQR coefficient increases faster over the range of quantiles than the QR coefficient. This is reflected in the coefficients at the conditional mean: the TSLS inflation coefficient is higher than the OLS inflation coefficient.

The response to the output gap is higher in the first part of the sample than in the second part. In the first part of the sample the output gap response is significant and close to the estimated coefficients at the mean of 0.45. The estimates of the second subsample show that the output gap is significantly different from zero only for the lower range of the distribution. The Fed therefore did not always respond countercyclically to the output gap. The output gap reactions decrease significantly over the conditional distribution from 0.5 to about 0. The output gap coefficients are different from the ones estimated by Chevapatrakul et al. (2009). They find an output gap coefficient that varies between 0.3 and 1 and that does not show a clear decreasing pattern. Their mean estimate is close to 0.5 while I find a mean estimate close to zero. The interest rate reaction to the output gap is weaker when the interest rate is set above an estimated policy rule and stronger when the interest rate is set below an estimated policy rule. The IQR output gap coefficient is over almost the entire distribution higher than the TSLS estimate showing that conventional methods presumably underestimate the output response of the Fed.

The constant shows high variations over the conditional distribution of the federal funds rate, but also wide confidence bands. It increases from 0 to 3.5 (QR) and 2.5 (IQR), respectively, deviating largely from estimated parameters at the mean. The constant includes variations in the natural real interest rate and the inflation target, but also includes variations in the inflation coefficient: $\alpha_0 = r^* + (1 - \alpha_\pi)\pi^*$. While an estimate of α_π is known, the targeted interest r^* and inflation rate π^* are not identified separately. As the constant and the inflation coefficient are negatively related when assuming a positive inflation target, but the graphs show an increase of both coefficients over the range of quantiles, one can infer that there is a substantial degree of variation in the natural interest rate, the inflation target or both.

Figure 5.3 shows the estimated coefficients of the inflation forecast, the output gap, the constant and an interest rate smoothing term for the whole conditional distribution of the federal funds rate when allowing for a gradual adjustment of interest rates. As in the case without interest rate smoothing it is apparent that uniform coefficients of standard estimations of linear monetary reaction function are an incomplete description of monetary policy. All QR and IQR parameters estimates vary significantly over the conditional distribution of the federal funds rate and support important nonlinearities over the conditional distribution of FOMC policy reactions. Although policy rules with an interest rate smoothing term show a high fit in general, the estimation results show that this is misleading and in fact high deviations from policy reaction parameters at the conditional mean of the interest rate appear. QR and IQR estimation results show similar patterns over the range of quantiles while variations of IQR coefficients are less smooth than variations of the QR coefficients.

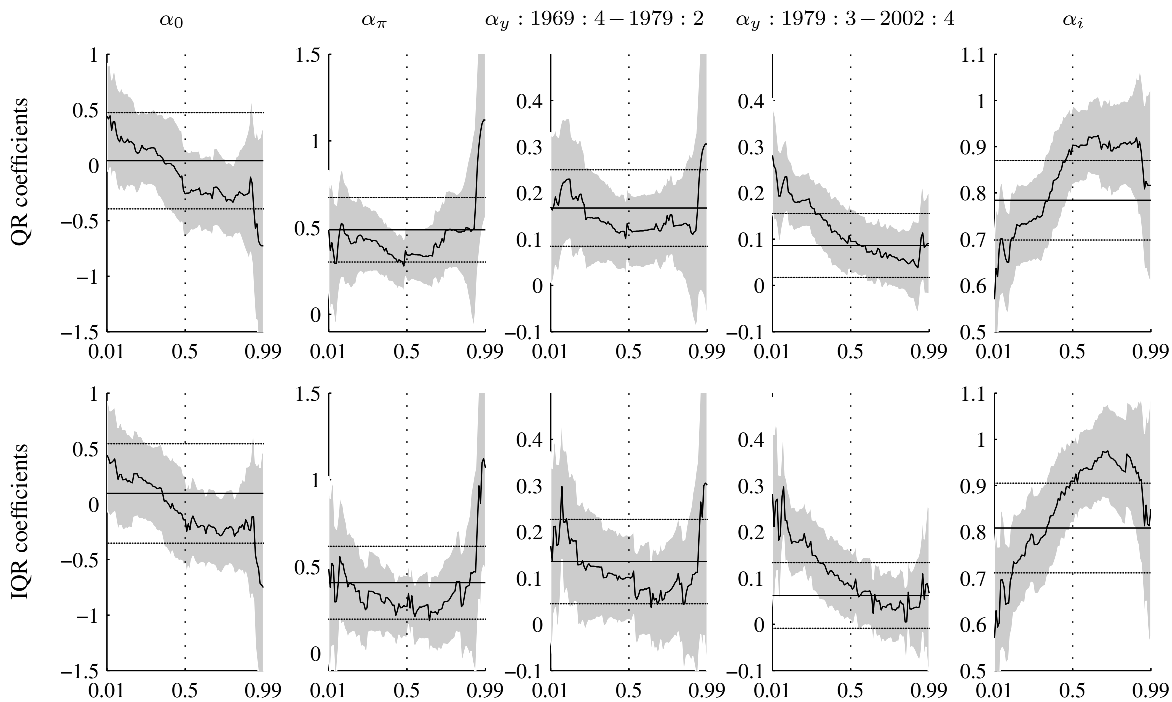


Figure 5.3: Estimated coefficients ($\alpha_i \neq 0$)

Notes: see Figure 5.2 for a description of the different graphs. The estimated equation is $i_t = \alpha_i(\tau)i_{t-1} + \alpha_0(\tau) + \alpha_\pi(\tau)\pi_{t+4|t} + (\alpha_{y,1}(\tau) + D\alpha_{y,2}(\tau))y_t + \epsilon_t$, for $\tau \in (0, 1)$.

The inflation response is significantly different from zero except for small outlier regions. Combining the inflation parameter and the smoothing parameter one can compute that the structural inflation response $\beta = \alpha_\pi/(1 - \alpha_i)$ is satisfying the Taylor principle over the entire distribution. The inflation coefficient is slightly below the mean estimates of 0.5 (OLS) and 0.4 (TSLS) between the 0.01 and the 0.75 quantile and increases strongly in the upper range of the distribution to 1.2. The median inflation coefficient is below the OLS/TSLS estimates.

The response to the output gap is decreasing over the distribution in both subsamples. The decrease is more pronounced in the second subsample from values around 0.25 to 0.05. In the first subsample the decrease ranges from values around 0.2 to 0.05 with an upward kink to 0.3 for estimates at the highest quantiles. The decrease of the output gap coefficient in the second subsample is highly significant. In both subsamples the instrumental variable estimates show that the output gap response is significant only for the lower 50% of the conditional distribution.

The interest rate smoothing parameter shows sizeable variations over the range of quantiles. With a mean estimate around 0.8 it increases from 0.6 to almost 1 at the 0.75 quantile and decreases thereafter slightly. The parameter is significantly different from zero over the whole distribution suggesting that interest rate smoothing is a prevalent characteristic of monetary policy of the Federal Reserve. The narrow confidence bands until the 0.75 quantile show that the parameter increase is highly significant. The median interest rate smoothing parameters is significantly higher than the OLS/TSLS estimate.

Finally, the constant shows a large decline over the distribution from 0.5. to -0.5 with a mean estimate slightly above 0. The confidence bands are wide and the constant is nowhere significantly different from 0. The constant can be written as $\alpha_0 = (1 - \alpha_i)r^* + (1 - \alpha_i - \alpha_\pi)\pi^*$ which shows that a large part of the decrease of α_0 is due to the increase of α_i. The sharp decrease at the highest quantiles reflects the high increase of α_π in this region of the distribution.

In summary, the estimation results for both specifications suggest that the Federal Reserve responded more aggressive to inflation and less to the output gap during upward deviations from a monetary policy reaction function estimated at the mean and the other way around during downward deviations. For the first part of the sample variations in the output gap response are limited especially in the case without a gradual adjustment of interest rates. The regression constant includes sizeable variations of the natural real interest rate and/or the inflation target over the conditional distribution of the federal funds rate. For the specification with a gradual adjustment of the federal funds rate the interest rate smoothing parameter amplifies the higher weight of inflation relative to the output gap during upward deviations from a policy rule. During downward deviations the lower smoothing parameter diminishes the relatively low inflation reaction further. It also dampens the more active output stabilizing policy compared to estimates at the mean as the structural coefficients $\beta(\tau)$ and $\gamma(\tau)$ are computed by division of $\alpha_\pi(\tau)$ and $\alpha_y(\tau)$ by $1 - \alpha_i(\tau)$. Systematic deviations from policy rule parameters estimated at the mean are strong even when taking into account interest rate smoothing as they overcompensate in this case the decrease of the constant over the conditional distribution of the federal funds rate.

5.5.1 Robustness

To ensure robustness of the results I repeat the estimations for quarterly spaced data, for the subsamples 1969:4-1979:2, 1979:3-2002:4, 1983:1-2002:4 and in addition for the whole sample abstracting from the structural break of the output gap imposed in the previous section.[9] The subsamples starting in 1979 and in 1983 are widely used in the literature on policy rules (see e.g. Clarida et al., 2000).

[9] I refer to the estimates from the previous section as the baseline case in the following.

Repeating regressions of the baseline specification with quarterly data yields similar results to the baseline results. In the case of no interest rate smoothing the increase in the inflation response over the conditional distribution of the interest rate is even more pronounced while the decrease of the output gap coefficient after 1979 is only visible between the 0.01 and the 0.25 quantile. The latter shows that it is important to use all available observations as one would otherwise capture an important feature of U.S. monetary policy not so clearly. In the case with interest rate smoothing the results are hardly distinguishable from the baseline estimation results. Estimation results for the different subsamples confirm the findings of the baseline case: an increase in the inflation coefficient, a decrease in the output gap coefficient for the Volcker-Greenspan era and a constant output gap coefficient for the pre-Volcker era. In the case without interest rate smoothing the regression constant increases, while it decreases when interest rate smoothing is allowed. The interest rate smoothing parameter increases in all subsamples. Especially the results for the sample starting in 1979 and in 1983 are close to the baseline results. The data with the highest frequency originate from this period. Therefore, the baseline results are not driven by the high inflation period of the 70's. However, the findings are not for all subsamples significant as the smaller number of observations leads to wide confidence bands. Results using all available data and quarterly data are similar while the confidence bands of the latter are wider.

5.6 Decomposing deviations from policy rules

The strong variation of policy coefficients raises the question if these are connected to expansions and recessions. For example, central bankers might be more averse to the danger of running into a recession than to accepting higher inflation during an expansion (Blinder, 1998). Thus, if the probability of a recession rises they might favor to decrease the interest rate by reacting more to the output gap compared to other times (Cukierman and Muscatelli, 2008). I estimate at which part of its conditional distribution the federal funds rate is set at each point of the sample. First, I compute for each observation fitted values of the interest rate at all quantiles using the parameters from IQR for all $\tau \in (0, 1)$. I then choose the quantile τ_t that minimizes the absolute difference of the fitted value and the actual value of the federal funds rate in period t.[10] In this way one generates a time series of quantiles τ_t that shows the path of the position of the federal funds rate on its conditional distribution.[11] Using this information one can decompose the deviations of the federal funds rate from an estimated policy rule into differences in the reactions to the covariates as follows:

$$i_t - \hat{i}_t \approx [\hat{\alpha}_0(\tau_t) - \hat{\alpha}_0] + [\hat{\alpha}_\pi(\tau_t) - \hat{\alpha}_\pi]\pi_{t+4|t} + [\hat{\alpha}_y(\tau_t) - \hat{\alpha}_y]y_t \tag{5.9}$$

[10] I find that this minimization problem is well behaved and features a unique minimum.

[11] I check robustness of the results using probit, logit and nonparametric estimation methods to estimate realized quantiles. Probit and logit estimates give similar results to the ones reported here. Nonparametric regression yields by trend similar results though showing some high frequency jumps of the estimated quantiles that might be caused by the low number of observations.

For example the second term on the right side shows how much the central bank's reaction to expected inflation deviates at time t from the reaction implied by the policy rule.[12,13]

Figure 5.4 shows the federal funds rate, the policy rule without interest rate smoothing estimated in section 5.5, estimated quantiles and a decomposition of deviations.[14] Row 2 shows the series of estimated quantiles which is linked closely to the least squares error term shown in row 3. Row 4 shows that deviations of the IQR constant from the TSLS constant are negligible. Major deviations from the policy rule are due to persistent deviations in the inflation response shown in row 5 and the output gap response in row 6.

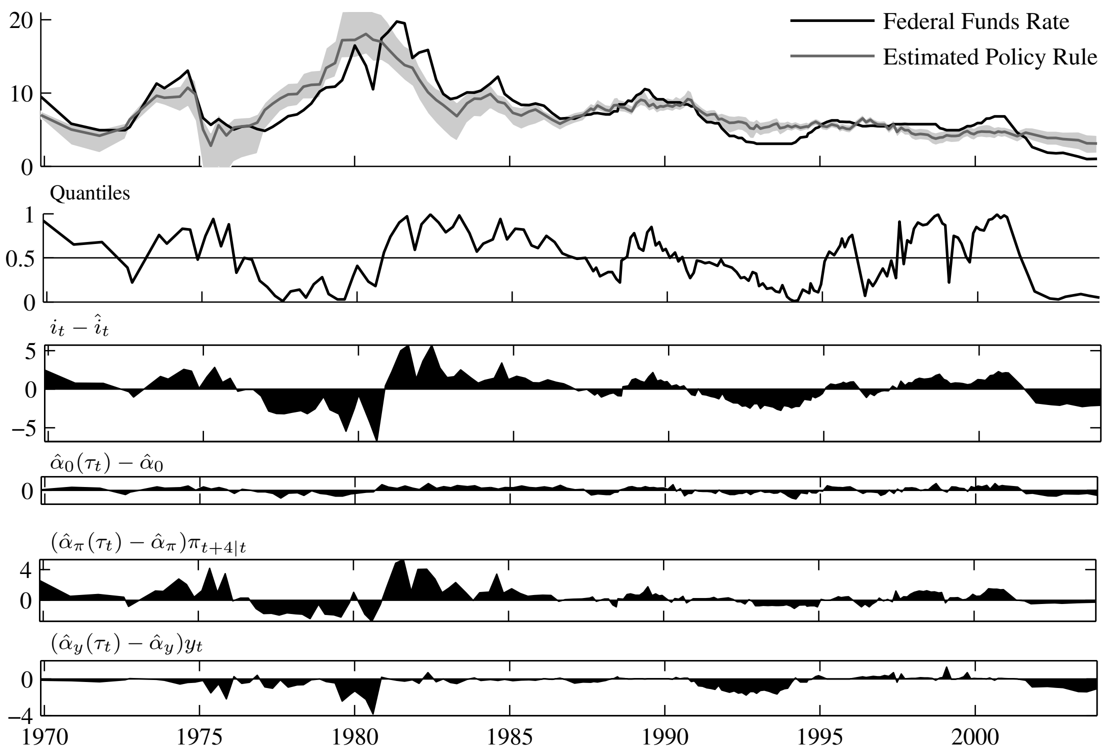


Figure 5.4: Fed funds rate, policy rule, quantiles and deviation decomposition ($\alpha_i = 0$)

Notes: Row 1 shows the federal funds rate and fitted values of the estimated policy rule using TSLS together with a 95% confidence band. Row 2 shows a series of estimated quantiles τ_t. Row 3 shows the difference between the policy rule and the federal funds rate. Rows 4-6 show the difference between estimated policy reactions and implied reactions by the policy rule. Summing up values from rows 4-6 yields row 3.

Differences between the estimated output gap responses and the response implied by the policy rule

[12]The major advantage of the methodology used here in comparison to logit and nonparametric approaches is that the estimated terms of the right side sum up almost exactly to the overall deviations on the left side. This is not the case when switching to other methods for estimating the quantile series. A disadvantage is that policy shocks do not show up anymore, but are absorbed in the variations of the parameters.

[13]The methodology is easily expanded to analyze deviations of the federal funds rate from benchmark policy rules. Deviations from Taylor's rule can be for example decomposed as follows: $i_t - i_t^{Taylor} = [\hat{\alpha}_0(\tau_t) - 1] + [\hat{\alpha}_\pi(\tau_t) - 1.5]\pi_{t+4|t} + [\hat{\alpha}_y(\tau_t) - 0.5]y_t$.

[14]I report only results for IQR and TSLS estimates here as they are close to the QR and OLS results.

are negative for large parts of the sample reflecting the finding from Figure 5.3 that the IQR coefficients are for large parts of the conditional interest rate distribution higher than the TSLS estimates. I compute correlations of the overall deviations of the interest rate from the policy rule estimated at the mean to the real-time output gap series. Overall deviations are negatively correlated with the business cycle for the period 1969:4-1979:2 (correlation coefficient: -0.35, p-value: 0.07), not correlated for the period 1979:3 - 2002:4 (correlation coefficient: 0.04, p-value: 0.63), but positively correlated for the post-Volcker period 1983:3 - 2002:4 (correlation coefficient: 0.34, p-value: 0.00). Thus, the Federal Reserve deviated from the policy responses proposed by a simple linear policy rule procyclically for the pre-Volcker period and anticyclically for the post-Volcker period. One can check further if these anticyclicality is due to deviations from a linear policy rule with respect to the inflation or the output gap reaction. There is no clear correlation between deviations in the inflation response and the business cycle. Deviations in the output gap response are uncorrelated with the business cycle during the pre-Volcker period (correlation coefficient: -0.01, p-value: 0.96), but positively correlated for the period 1979:3 - 2002:4 (correlation coefficient: 0.18, p-value: 0.03) and also for the period 1983:3 - 2002:4 (correlation coefficient: 0.42, p-value: 0.00). Thus, Federal Reserve policy responses to the output gap deviate anticyclically from a linear policy rule for the Volcker-Greenspan era. This anticyclicality together with a decreasing output gap coefficient over the conditional distribution of the interest rate implies a recession avoidance preference for the 1980 - 2002 period. The central bank reacted more to the output gap during recessions leading to a lower interest rate setting than proposed by a linear policy rule. This confirms the recession avoidance preference of the Federal Reserve found by Cukierman and Muscatelli (2008) for the Greenspan period. They estimate an interest rate rule with smooth-transisition models for inflation deviations from a target and the output gap to capture nonlinearities in the reaction to these two variables. Gerlach (2000) and Surico (2007) also find that the Federal Reserve responded more strongly to recessions than to expansions, but only between 1960 and 1980 and not afterwards. Gerlach (2000) uses a nonlinear policy reaction function and a HP-filtered output gap, while Surico (2007) uses the CBO output gap and squared inflation and output gap terms in a linear policy rule. The differences to my results might be due to the different methodological approach and the usage of real-time data in this study.

The graphs reflect the anticyclicality for important episodes of monetary policy: for example during the downturn of the early nineties due to FOMC concerns about "financial headwinds" (Poole, 2006) the output gap response is high. As the real-time output gap is negative for most of the time (see Figure 5.1) this high output gap reaction brings about an anticyclical decrease in the interest rate.

Figure 5.5 shows the same decomposition for the case with interest rate smoothing. Even though differences between the federal funds rate and the fitted values from the policy rule are hardly visible in row 1 of the graph, the series of quantiles in row 2 shows that deviations from the policy rule are persistent during some periods and row 3 shows that these even take values between -4% and 5% during the reserve targeting period in the early 1980's. The Fed deviates in its reactions to inflation, the lagged interest rate and during some periods in the reaction to the output gap from the estimated policy rule. Overall deviations from the policy rule and deviations in the inflation response from the linear rule are uncorrelated to the real-time output gap. Deviations in the output gap response from the linear policy rule are negatively correlated for the period 1969:4-1979:2 (correlation coefficient: -0.63, p-value: 0.00) and positively correlated for the period 1979:3 - 2002:4 (correlation coefficient: 0.28, p-value: 0.00) and also for the period 1983:3 - 2002:4 (correlation coefficient: 0.29, p-value: 0.00). Thus, the Federal Reserve's output gap response deviated procyclically from the one suggested by a linear policy rule for the pre-Volcker period and anticyclically for the post-Volcker period. The latter confirms the result from the case without interest rate smoothing and the recession avoidance preference found by Cukierman and Muscatelli (2008). One can conclude that even though the deviations from a policy rule are small when allowing for a gradual adjustment of interest rates, quantile regression is still useful as it allows a more precise description of monetary policy that is otherwise hidden behind the high degree of interest rate smoothing.

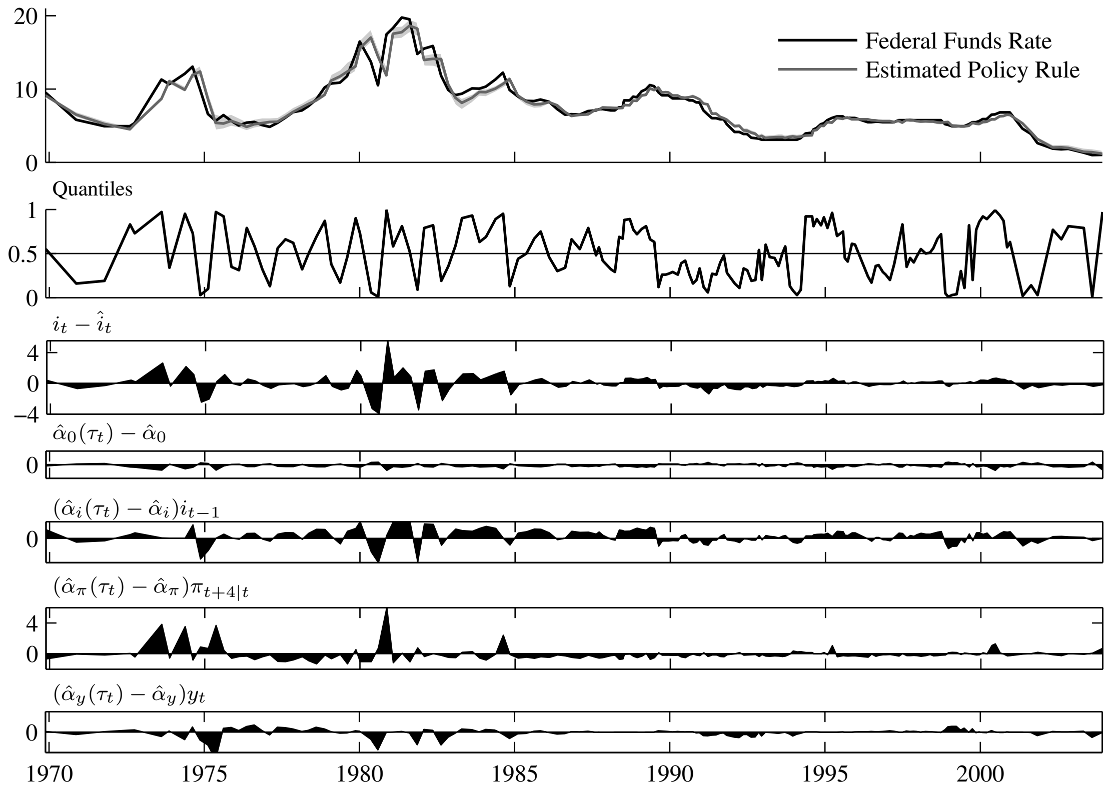


Figure 5.5: Fed funds rate, policy rule, quantiles and deviation decomposition ($\alpha_i \neq 0$)

Notes: see Figure 5.4 for a description of the different graphs.

5.7 Conclusion

Using quantile regressions to estimate monetary policy rules appears to be useful: without including additional variables, one obtains more detailed estimates than with standard estimation techniques without violating the robustness property of simple rules. Deviations of the federal funds rate from standard policy rule estimates are caused to a large extent by systematic changes in the inflation and output gap reaction parameters and the interest rate smoothing parameter over the conditional distribution of the federal funds rate rather than by policy shocks. Inflation reactions increase and output gap responses decrease over the conditional distribution of the interest rate. Allowing for a gradual adjustment of interest rates pretends a high fit of an estimated policy rule, while quantile regression reveals systematic and significant movements of monetary policy reaction coefficients over the conditional distribution of the federal funds rate. Estimating at which part of its conditional distribution the interest rate is located for each observation of the sample shows that deviations of the output gap response from a linear policy rule are procyclical for the pre-Volcker period and anitcyclical for the Volcker-Greenspan era. The anticyclical output gap response together with a decreasing output gap coefficient over the conditional distribution of the interest rate for the second part of the sample implies at least a mild recession avoidance preference of the Federal Reserve for the period 1980 - 2003.

References

Adam, K., Billi, R. M., 2006. Optimal monetary policy under commitment with a zero bound on nominal interest rates. Journal of Money, Credit, and Banking 38(7), 1877–1905.

Amemiya, T., 1982. Two stage least absolute deviations estimators. Econometrica 50, 689–711.

Blinder, A. S., 1998. Central Banking in Theory and Practice. Cambridge, MA: MIT Press.

Chen, L.-A., Portnoy, S., 1996. Two-stage regresson quantiles and two-stage trimmed least squares estimators for structural equation models. Communications in Statistics. Theory and Methods 25(5), 1005–1032.

Chernozhukov, V., Hansen, C., 2001. An IV model of quantile treatment effects, Massachusetts Institute of Technology, Department of Economics, Working Paper 02-06.

Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. Econometrica 73(1), 245–261.

Chevapatrakul, T., Kim, T.-H., Mizen, P., 2009. The Taylor principle and monetary policy approaching a zero bound on nominal rates: Quantile regression results for the United States and Japan. Journal of Money, Credit and Banking 41(8), 1705–1723.

Clarida, R., Galí, J., Gertler, M., 1998. Monetary policy rules in practice: Some international evidence. European Economic Review 42, 1003–1067.

Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. Quarterly Journal of Economics 115(1), 147–180.

Cukierman, A., Muscatelli, A., 2008. Nonlinear taylor rules and asymmetric preferences in central banking: Evidence from the United Kingdom and the United States. The B.E. Journal of Macroeconomics 8(1).

Dolado, J. J., Maria-Dolores, R., Naveira, M., 2005. Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve. European Economic Review 49, 485–503.

Fitzenberger, B., 1997. The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. Journal of Econometrics 82, 235–287.

Gerlach, S., 2000. Asymmetric policy reactions and inflation, working paper, Bank for International Settlements.

Greenspan, A., September 1997. Rules vs. discretionary monetary policy, speech at the 15th Anniversary Conference of the Center for Economic Policy Research at Stanford University, Stanford, California.

Kato, R., Nishiyama, S.-I., 2005. Optimal monetary policy when interest rates are bounded at zero. Journal of Economic Dynamics & Control 29, 97–133.

Koenker, R., Basset, G. W., 1978. Regression quantiles. Econometrica 46(1), 33–50.

Lee, S., 2004. Endogeneity in quantile regression models: A control function approach. CeMMAP Working Paper, University College London 08/04.

Meyer, L. H., Swanson, E. T., Wieland, V., 2001. Nairu uncertainty and nonlinear policy rules. American Economic Review 91(2), 226–231.

Orphanides, A., 2001. Monetary policy rules based on real-time data. American Economic Review 91, 964–985.

Orphanides, A., 2004. Monetary policy rules, macroeconomic stability, and inflation: A view from the trenches. Journal of Money, Credit and Banking 36(2), 151–175.

Orphanides, A., Wieland, V., 2000. Efficient monetary policy design near price stability. Journal of the Japanese and International Economies 14, 327–365.

Orphanides, A., Wieland, V., 2008. Economic projections and rules of thumb for monetary policy. Federal Reserve Bank of St. Louis Review 90(4), 307–324.

Poole, W., August 2006. Understanding the Fed, speech at the Dyer County Chamber of Commerce Annual Membership Luncheon, Dyersburg, Tenn.

Powell, J. L., 1983. The asymptotic normality of two-stage least absolute deviations estimators. Econometrica 51(5), 1569–1575.

Schaling, E., 1999. The nonlinear Phillips curve and inflation forecast targeting, Bank of England Working Paper No. 98.

Sugo, T., Teranishi, Y., 2005. The optimal monetary policy rule under the non-negativity constraint on nominal interest rates. Economics Letters 89, 95–100.

Surico, P., 2007. The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. Journal of Economic Dynamics & Control 31, 305–324.

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

Tillmann, P., 2010. Parameter uncertainty and non-linear monetary policy rules. Macroeconomic Dynamics, forthcoming.

Curriculum Vitae

	Maik Hendrik Wolters
Home Address	Würzburger Straße 34 60385 Frankfurt am Main
Date of Birth	28 December 1982
Citizenship	German

Education

Oct 2006 – Jun 2010	Ph.D. Program in Economics at Goethe University Frankfurt
Oct 2005 – Mar 2007	M.Sc. in Quantitative Economics, Goethe University Frankfurt
Aug 2004 – Jul 2005	M.A. in International Business, ESC Rennes School of Business, France
Oct 2002 – Jul 2004	Undergraduate studies in Economics (Pre-Deploma), Bielefeld University
Jul 2002	Abitur at Humboldt-Gymnasium, Bad Pyrmont