MONFISPOL - Deliverable 5.1.3

Part 1
The Diversity of Forecasts from Macro-economic Models of the U.S. Economy

Part 2
Surprising Comparative Properties of Monetary Models: Results from a New Model Database

Part 3
The New Keynesian Approach to Dynamic General Equilibrium Modeling: Models, Methods and Macroeconomic Policy Evaluation

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

Volker Wieland
Goethe University Frankfurt, CEPR and CFS†

Maik H. Wolters
Goethe University Frankfurt

May 20, 2010

Abstract

This paper 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

*We are grateful to Mordecai Kurz, two anonymous referees and the participants at the Stanford Institute for Theoretical Economics Workshop on Diverse Beliefs and the MONFISPOL Networks Stresa workshop for very helpful comments on this paper. Wieland acknowledges funding support from European Community grant MONFISPOL under grant agreement SSH-CT-2009-225149. The usual disclaimer applies.

†Corresponding Author: Volker Wieland; Mailing address: Goethe University of Frankfurt, Grueneburgplatz 1, Goethe Universitaet, House of Finance, 60323 Frankfurt am Main, Germany, E-mail: wieland@wiwi.uni-frankfurt.de. Other Author: Maik Wolters; Mailing address: Grueneburgplatz 1, House of Finance, 60323 Frankfurt am Main, Germany.

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 (2011), Branch and Evans (2011) and De Grauwe (2011).

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 paper 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 paper are three small-scale New-Keynesian

[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.

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 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 paper 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 paper 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,

[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.

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. 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 paper proceeds as follows. Section 2 summarizes the most important features of the different macroeconomic models that we use to compute forecasts. Section 3 describes the estimation and forecasting methodology. Section 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 5. Section 6 provides a comparison of forecast accuracy of model and professional forecasts. The extent and dynamics of forecast heterogeneity is studied systematically in section 7. Section 8 summarizes our findings and concludes.

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 equa-

tions. The other five models are structural representations of the U.S. economy. Table 1 summarizes the most important model features, while appendix A1 provides a complete description of the model equations.

Table 1: Model Overview

Name/Reference	Short Name	Type	Observable Variables	Original Authors' Sample
Bayesian VAR estimated in this paper	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 paper	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

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 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

[5]These models combined rational expectations and nominal rigidities as in the seminal paper of Taylor (1979). For other

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 paper. 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 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

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.

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

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

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:

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

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (x_t - g_t) \tag{2}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2 x_t) + \varepsilon_{R,t} \tag{3}$$

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 $\varepsilon_{R,t}$ 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:

$$YGR_t = \ln\gamma + \Delta x_t + z_t \tag{4}$$

$$INFL_t = \ln\pi^* + \pi_t \tag{5}$$

$$INT_t = \ln r^* + \ln\pi^* + R_t. \tag{6}$$

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:

$$y_t^{obs} = y(\theta) + \lambda + y_t^s, \tag{7}$$

$$y_t = g_y(\theta) y_{t-1} + g_u(\theta) u_t, \tag{8}$$

$$E(u_t u_t') = Q(\theta), \tag{9}$$

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 shows how to link the variables and parameters in the state space representation to those in the Del Negro & Schorfheide model.

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.

Table 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	$u_t = [\ \varepsilon_{R,t} \quad \varepsilon_{z,t} \quad \varepsilon_{g,t}\ ]'$

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 $ln\mathscr{L}(\theta|Y^T) + lnp(\theta)$ 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.

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

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 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:

$$E[y^{obs}_{t+h}|I_t] = y(\hat{\theta}_t) + \hat{\lambda}_t + g_y(\hat{\theta}_t)^{h+1} y_{t-1}. \tag{10}$$

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.

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 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 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

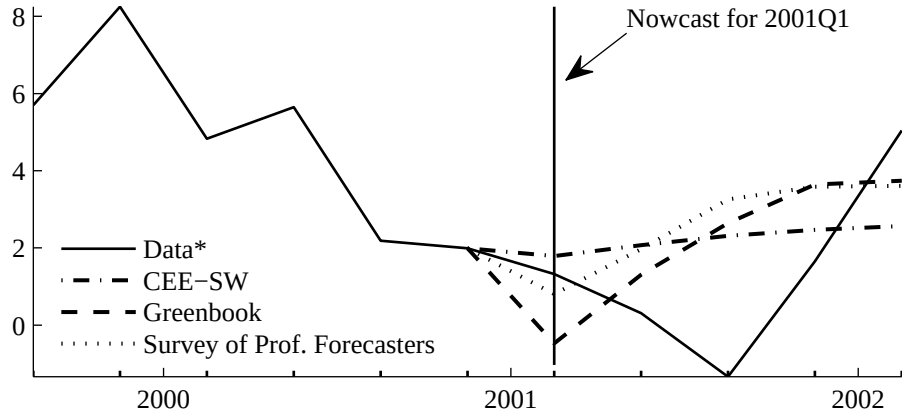


Figure 1: Real GDP Growth Forecast at the Start of the 2001 Recession (NBER defined peak: 2001Q1, NBER defined trough: 2001Q4).

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

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 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 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

[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.

[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.

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. The left-hand-side column of panels in Figure 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 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 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 and widen again in 2001:Q4 and 2002:Q1.

5 Model-Based versus Expert Nowcasts and the 2008/09 Recession

The model-based forecasts shown in Figures 1 and 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.

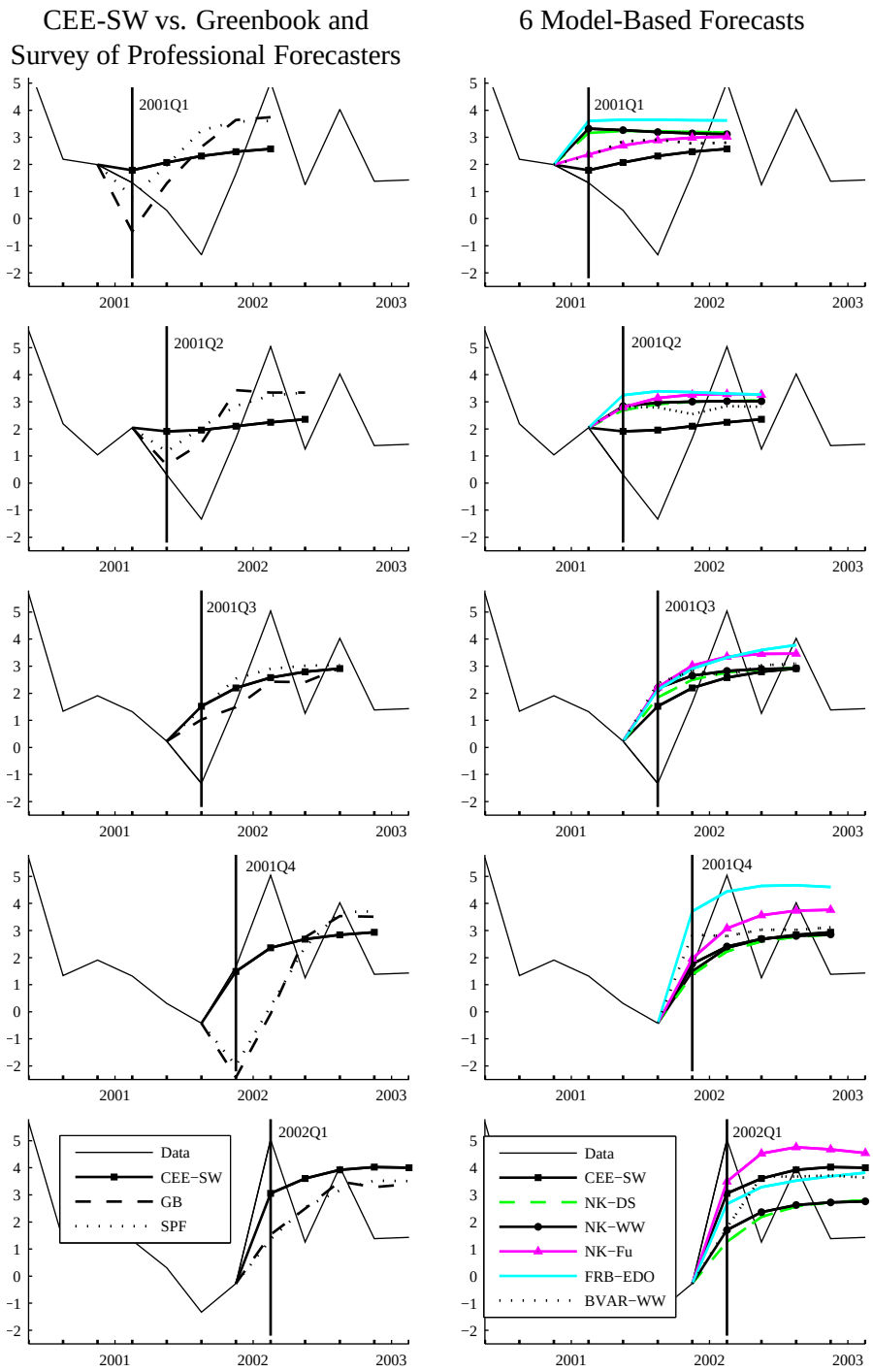


Figure 2: Real GDP Growth Forecasts for the 2001 Recession (NBER defined peak: 2001Q1, NBER defined trough: 2001Q4)

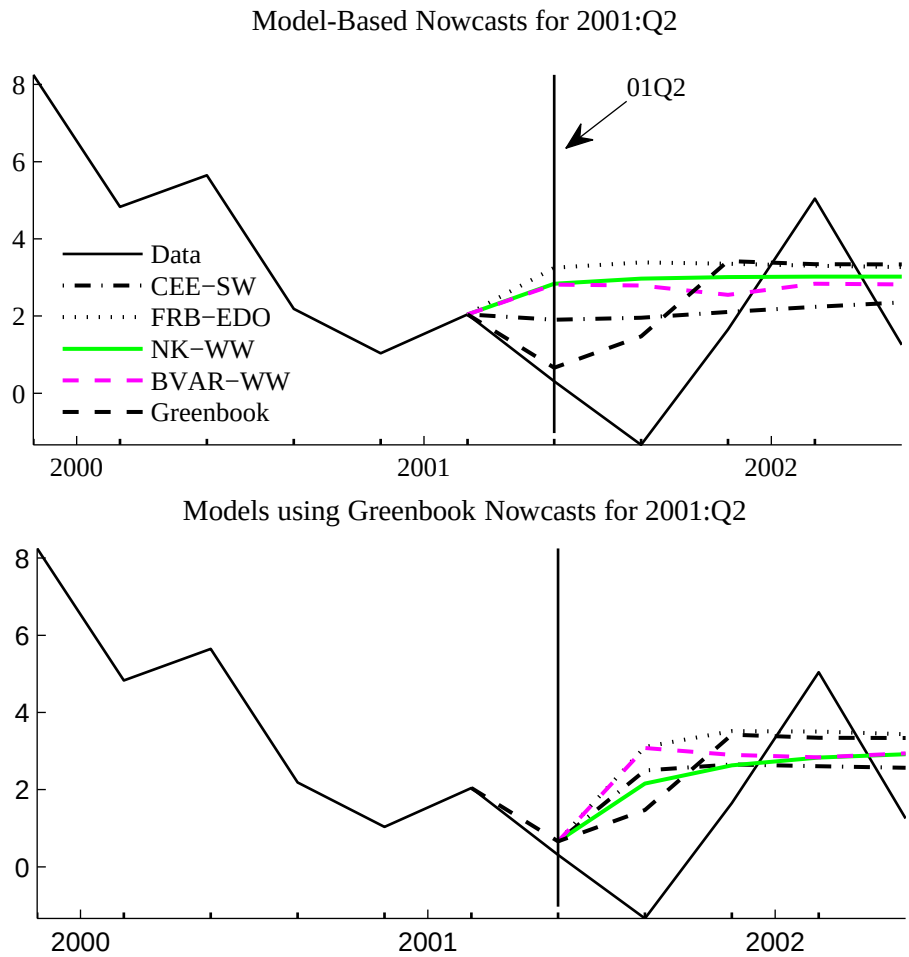


Figure 3: Real Output Growth forecast during the 2001 recession (NBER defined peak: 2001:Q1, NBER defined trough: 2001:Q4).

Notes: 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.

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 paper. 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 3. The top panel in Figure 3 partly replicates the second right-hand-side panel in Figure 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 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 4 provides a perspective on successive model forecasts relative to the mean SPF forecast (dash-dotted line) and the actual data (solid line) that 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 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

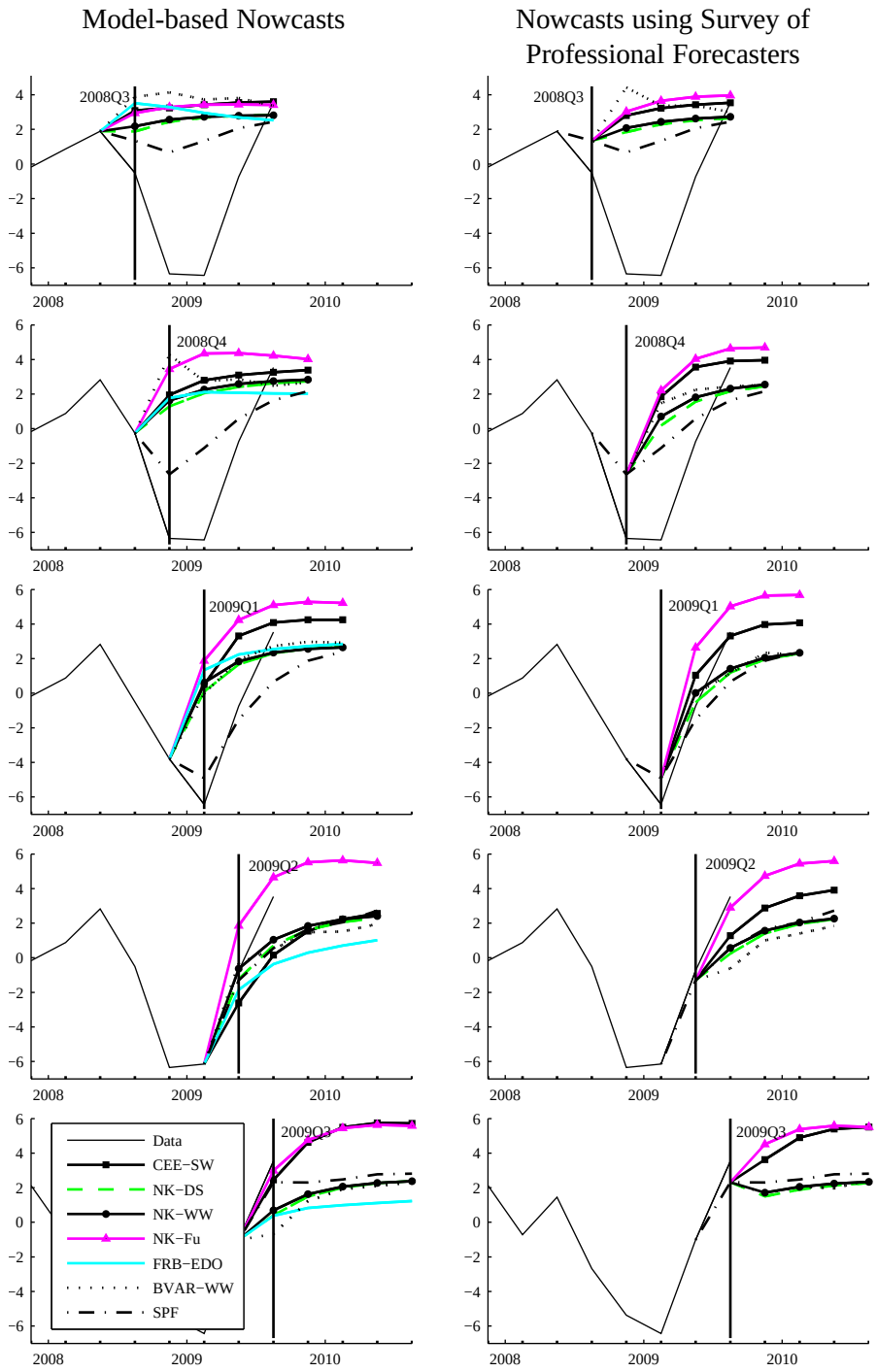


Figure 4: Real Output Growth forecast during the 2007-2009 recession (NBER defined peak: 2007Q4).

Notes: 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.

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.

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:

$$RMSE_m^h = \sqrt{\sum_{t=p}^{q} \left(E[y_{t+h}^{obs}|I_t^m] - y_{t+h}^{obs}\right)^2 / (q-p+1)}, \tag{11}$$

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 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, 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

[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 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

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 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

Table 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

model forecast in terms of root mean squared error.

In addition, we have investigated the accuracy of inflation forecasts. Table 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 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

Table 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

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 shock 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 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

Table 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

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 5 and 6 report the associated root mean squared errors of output growth and inflation forecasts for the different recession episodes.

The GDP growth forecast 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.

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 1 to 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:

$$\sigma_t = \sqrt{\sum_{m=1}^{M} \left(E[y_{t+h}^{obs}|I_t^m] - \frac{1}{M}\sum_{m=1}^{M} E[y_{t+h}^{obs}|I_t^m] \right)^2 /(M-1)}, \tag{12}$$

where I_t^m 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.

Figures 5 and 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 5 and 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.

Another aspect of heterogeneity concerns the range of accuracy of forecasts by individual fore-

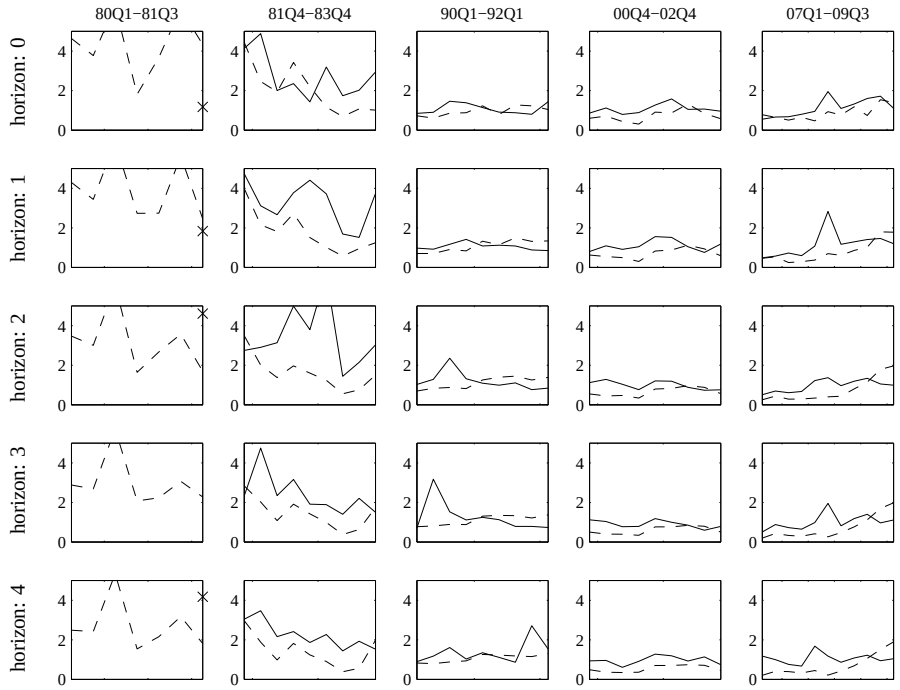


Figure 5: Standard Deviations of Output Growth Forecasts: Experts (solid) and Models (dashed)

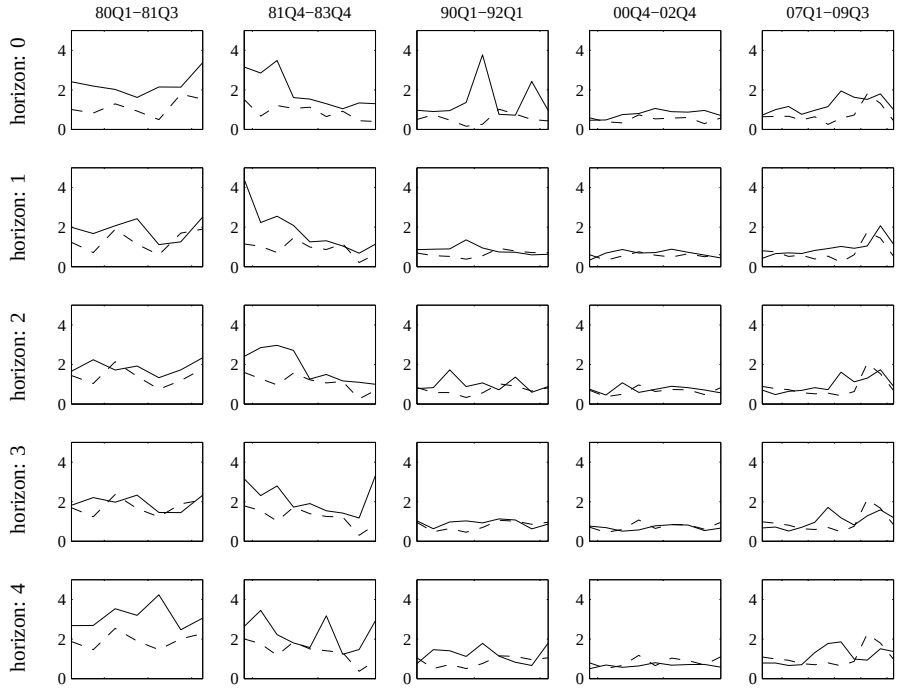


Figure 6: Standard Deviations of Inflation Forecasts: Survey (Black) and Models (Red)

casters. 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 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 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

Table 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

Table 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

statistics. The models' minimal, maximal, and average RMSE decrease by a small amount.

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]

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 nowcasting 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

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

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 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

[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).

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 theory 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 (2011), Guo et al. (2011) and Nielsen (2011) in this issue apply the theory of rational beliefs. Branch and McGough (2011), Branch and Evans (2011) and De Grauwe (2011) 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 our paper.

[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).

References

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

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

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

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

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

Brock, W. and 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., and Portes, R., editors (1989). *Macroeconomic Policies in an Interdependent World*. Washington, D.C.: The Brookings Institution.

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

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

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

Chiarella, C., Dieci, R., and 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., and 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., and 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. (2011). Animal spirits and monetary policy. *Economic Theory*, this volume.

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

Del Negro, M., Schorfheide, F., Smets, F., and Wouters, R. (2007). On the fit of New Keynesian models. *Journal of Business and Economic Statistics*, 25(2):123143.

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

Edge, R. M., Kiley, M. T., and 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., and 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., and 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., and Williams, J. (2009). Assessing the temporal variation of macroeconomic forecasts by a panel of changing composition. Northwestern University Working Paper.

Faust, J. and 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. and 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. and Moore, G. (1995a). Inflation persistence. *The Quarterly Journal of Economics*, 110(1):127–159.

Fuhrer, J. C. and 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., and Reichlin, L. (2009). Incorporating conjunctural analysis in structural models. In Wieland, V., editor, *The science and practice of monetary policy today*, pages 41–57. Springer Science.

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

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

Guo, W. C., Wang, F. Y., and Wu, H. M. (2011). Financial leverage and market volatility with diverse beliefs. *Economic Theory*, this volume.

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., editor (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., editor, *Endogenous Economic Fluctuations: Studies in the Theory of Rational Beliefs*. Springer Series in Economic Theory, No. 6, Springer Verlag.

Kurz, M., editor (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. and Schenk-Hoppe, K., editors, *Handbook of financial markets: dynamics and evolution*. North Holland.

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

Kurz, M., Jin, H., and Motolese, M. (2005). The role of expectations in economic fluctuations and

the efficacy of monetary policy. *Journal of Economic Dynamics & Control*, 29:20172065.

Kurz, M. and Motolese, M. (2011). Diverse beliefs and time variability of risk premia. *Economic Theory*, this volume.

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

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

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

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

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

Nielsen, C. K. (2011). Price stabilizing, pareto improving policies. *Economic Theory*, this volume.

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

Rotemberg, J. J. and 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. and 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. and 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. and 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. and 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., editor (2009). *The Science and Practice of Monetary Policy Today*. Springer Science.

Wieland, V., Cwik, T., Müller, G. J., Schmidt, S., and 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. and Rogoff, K., editors, *NBER Macroeconomics Annual 2003*, volume 18, pages 257–268. Cambridge, Mass.: MIT Press.

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

Appendix A1: The Macroeconomic Models Used to Compute Forecasts

This appendix provides a description of the six macroeconomic models that are used in this paper 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:

$$\tilde{y}_t = a_0 + a_1\tilde{y}_{t-1} + a_2\tilde{y}_{t-2} + a_\rho \rho_{t-1} + \varepsilon_{y,t}, \tag{13}$$

$\tilde{y}_t$ 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 $\varepsilon_{y,t}$ 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:

$$\rho_t - D\left[E_t(\rho_{t+1}) - \rho_t\right] = f_t - E_t(\pi_{t+1}). \tag{14}$$

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). v_t denotes an index of wage contracts that are currently in effect:

$$v_t = \sum_{i=0}^{3} \omega_i (x_{t-i} - p_{t-i}), \tag{15}$$

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:

$$x_t - p_t = \sum_{i=0}^{3} \omega_i (v_{t+i} + \gamma \tilde{y}_{t+i}) + \varepsilon_{p,t}. \tag{16}$$

$\varepsilon_{p,t}$ 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:

$$\pi_t = \sum_{i=0}^{3} \omega_i (x_{t-i} - x_{t-i-1}). \tag{17}$$

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:

$$f_t = \alpha_0 + \alpha_{f1} f_{t-1} + \alpha_{\pi 0} \pi_t + \alpha_{\pi 1} \pi_{t-1} + \alpha_{\Delta y} (\tilde{y}_t - \tilde{y}_{t-1}) + \alpha_y \tilde{y}_t + \varepsilon_{f,t}. \tag{18}$$

Contrary to the other structural models considered in this paper, 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:

$$x_t = E_t x_{t+1} - \tau^{-1}(R_t - E_t \pi_{t+1}) + (1-\rho_g) g_t + \rho_z \tau^{-1} z_t, \tag{19}$$

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:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(x_t - g_t), \tag{20}$$

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:

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2 x_t) + \varepsilon_{R,t}. \tag{21}$$

The monetary policy shock, $\varepsilon_{R,t}$, 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:

$$x_t = E_t x_{t+1} - \tau^{-1}(R_t - E_t \pi_{t+1}) + (1-\rho_g) g_t + \rho_z \tau^{-1} z_t + \tau^{-1}(1-\rho_\chi)\chi_t, \tag{22}$$

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:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \left[x_t - g_t + \tau^{-1}(\Phi_t - \chi_t) \right]. \tag{23}$$

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 (dynamic stochastic general equilibrium) 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:

$$y_t = c_y c_t + i_y i_t + z_y z_t + \varepsilon_t^g, \tag{24}$$

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

$$c_t = c_1 c_{t-1} + (1-c_1)E_t(c_{t+1}) + c_2(l_t - E_t(l_{t+1})) - c_3(r_t - E_t(\pi_{t+1}) + \varepsilon_t^b. \tag{25}$$

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

$$i_t = i_1 i_{t-1} + (1-i_1)E_t(i_{t+1}) + i_2 q_t + \varepsilon_t^i, \tag{26}$$

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:

$$q_t = q_1 E_t(q_{t+1}) + (1-q_1)E_t(r^k_{t+1}) - (r_t - E_t(\pi_{t+1}) + \varepsilon^b_t), \tag{27}$$

where $q_1 = \beta\gamma^{-\sigma_c}(1-\delta)$. r^k_t 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:

$$y_t = \phi_p(\alpha k^s_t + (1-\alpha)l_t + \varepsilon^a_t). \tag{28}$$

k^s_t denotes effective capital (physical capital adjusted for the capital utilization rate), ε^a_t 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:

$$k^s_t = k_{t-1} + z_t. \tag{29}$$

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

$$z_t = (1-\psi)/\psi r^k_t, \tag{30}$$

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

$$k_t = k_1 k_{t-1} + (1-k_1)i_t + k_2\varepsilon^i_t, \tag{31}$$

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^p_t equals the difference between the marginal product of labor and the real wage w_t:

$$\mu^p_t = \alpha(k^s_t - l_t) + \varepsilon^a_t - w_t. \tag{32}$$

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:

$$\pi_t = \pi_1\pi_{t-1} + \pi_2\pi_{t+1} - \pi_3\mu^p_t + \varepsilon^p_t, \tag{33}$$

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 $\pi_3 = 1/\ (1+\beta\ \gamma^{1-\sigma_c}\ \iota_p)\ [(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\xi_p((\phi_p-1)\varepsilon_p+1)]$. 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:

$$\mu_t^w = w_t - mrs_t = w_t - (\sigma_l l_t + \frac{1}{1-\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})), \tag{34}$$

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

$$w_t = w_1 w_{t-1} + (1-w_1)(E_t(w_{t+1}) + E_t(\pi_{t+1})) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \varepsilon_t^w, \tag{35}$$

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 $w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/(\xi_w((\phi_w-1)\varepsilon_w+1))]$. The parameter definition is analogous to the price Phillips curve.

Setting $\xi_p = 0$, $\xi_w = 0$, $\varepsilon_t^p = 0$ and $\varepsilon_t^w = 0$ 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:

$$r_t = \rho r_{t-1} + (1-\rho)(r_\pi\pi_t + r_x x_t) + r_{\Delta x_t}(x_t - x_{t-1}) + \varepsilon_t^r. \tag{36}$$

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:

$$L_t^s = (1-\alpha)\widetilde{X}_t^s \frac{\widetilde{MC}_t^s}{\widetilde{W}_t^s}, \qquad \text{for } s = cbi, kb \tag{37}$$

$$\frac{\widetilde{K}_t^{u,nr,s}}{\Gamma_t^{x,kb}} = \alpha \widetilde{X}_t^s \frac{\widetilde{MC}_t^s}{\widetilde{R}_t^{nr,s}}, \qquad \text{for } s = cbi, kb \tag{38}$$

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:

$$\widetilde{X}_t^s = (L_t^s)^{(1-\alpha)} \left(\frac{\widetilde{K}_t^{u,nr,s}}{\Gamma_t^{x,kb}} \right)^{\alpha}, \qquad \text{for } s = cbi, kb \tag{39}$$

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:

$$\begin{aligned}\Theta_t^{x,s}\widetilde{MC}_t^s\widetilde{X}_t^s = (\Theta_t^{x,s}-1)\widetilde{P}_t^s\widetilde{X}_t^s \\ +100\chi^p(\Pi_t^{p,s}-\eta^p\Pi_{t-1}^{p,s}-(1-\eta^p)\Pi_*^{p,s})\Pi_t^{p,s}\widetilde{X}_t^s\widetilde{P}_t^s \\ -\beta E_t\left(\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}100\chi^p(\Pi_{t+1}^{p,s}-\eta^p\Pi_t^{p,s}-(1-\eta^p)\Pi_*^{p,s})\Pi_{t+1}^{p,s}\widetilde{P}_{t+1}^s\widetilde{X}_{t+1}^s\right), \quad (40)\end{aligned}$$

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:

$$\widetilde{Q}_t^{nr} = \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\frac{1}{\Gamma_{t+1}^{x,kb}}\left(\widetilde{R}_{t+1}^{nr}+(1-\delta^{nr})\widetilde{Q}_{t+1}^{nr}\right)\right\}, \quad (41)$$

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:

$$\begin{aligned}\widetilde{P}_t^{kb} &= \widetilde{Q}_t^{nr}\left[A_t^{nr}-100\chi^{nr}\left(\frac{\widetilde{E}_t^{nr}-\widetilde{E}_{t-1}^{nr}}{\widetilde{K}_t^{nr}}\Gamma_t^{x,kb}\right)\right] \quad (42)\\ &+ \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\widetilde{Q}_{t+1}^{nr}100\chi^{nr}\left(\frac{\widetilde{E}_{t+1}^{nr}-\widetilde{E}_t^{nr}}{\widetilde{K}_{t+1}^{nr}}\Gamma_{t+1}^{x,kb}\right)\right\}.\end{aligned}$$

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:

$$\widetilde{R}_t^{nr,s} = \frac{\widetilde{R}_t^{nr}}{U_t^s}, \qquad \text{for } = cbi, kb \tag{43}$$

$$U_t^s = \left(\frac{1}{\kappa}\frac{\widetilde{R}_t^{nr,s}}{\widetilde{P}_t^{kb}}\right)^{\frac{1}{\psi}}, \qquad \text{for } = cbi, kb \tag{44}$$

$$\widetilde{K}_{t+1}^{nr} = (1-\delta^{nr})\frac{\widetilde{K}_t^{nr}}{\Gamma_t^{x,kb}} + A_t^{nr}\widetilde{E}_t^{nr} - \frac{100\chi^{nr}}{2}\left(\frac{\widetilde{E}_t^{nr} - \widetilde{E}_{t-1}^{nr}}{\widetilde{K}_t^{nr}}\Gamma_t^{x,kb}\right)^2 \frac{\widetilde{K}_t^{nr}}{\Gamma_t^{x,kb}} \tag{45}$$

$$\widetilde{K}_t^{nr} = \widetilde{K}_t^{nr,cbi} + \widetilde{K}_t^{nr,kb}. \tag{46}$$

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:

$$\widetilde{Q}_t^{cd} = \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\frac{1}{\Gamma_{t+1}^{x,kb}}\left(\widetilde{R}_{t+1}^{cd} + (1-\delta^{cd})\widetilde{Q}_{t+1}^{cd}\right)\right\} \tag{47}$$

$$\begin{aligned}\widetilde{P}_t^{kb} &= \widetilde{Q}_t^{cd}\left[A_t^{cd} - 100\chi^{cd}\left(\frac{\widetilde{E}_t^{cd} - \widetilde{E}_{t-1}^{cd}}{\widetilde{K}_t^{cd}}\Gamma_t^{x,kb}\right)\right] \\ &+ \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\widetilde{Q}_{t+1}^{cd}100\chi^{cd}\left(\frac{\widetilde{E}_{t+1}^{cd} - \widetilde{E}_t^{cd}}{\widetilde{K}_{t+1}^{cd}}\Gamma_{t+1}^{x,kb}\right)\right\}\end{aligned} \tag{48}$$

$$\widetilde{K}_{t+1}^{cd} = (1-\delta^{cd})\frac{\widetilde{K}_t^{cd}}{\Gamma_t^{x,kb}} + A_t^{cd}\widetilde{E}_t^{cd} - \frac{100\chi^{cd}}{2}\left(\frac{\widetilde{E}_t^{cd} - \widetilde{E}_{t-1}^{cd}}{\widetilde{K}_t^{cd}}\Gamma_t^{x,kb}\right)^2 \frac{\widetilde{K}_t^{cd}}{\Gamma_t^{x,kb}} \tag{49}$$

and

$$\widetilde{Q}_t^{r} = \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\frac{1}{\Gamma_{t+1}^{x,cbi}}\left(\widetilde{R}_{t+1}^{r} + (1-\delta^{r})\widetilde{Q}_{t+1}^{r}\right)\right\} \tag{50}$$

$$\begin{aligned}\widetilde{P}_t^{cbi} &= \widetilde{Q}_t^{r}\left[A_t^{r} - 100\chi^{r}\left(\frac{\widetilde{E}_t^{r} - \widetilde{E}_{t-1}^{r}}{\widetilde{K}_t^{r}}\Gamma_t^{x,cbi}\right)\right] \\ &+ \beta E_t\left\{\frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}}\widetilde{Q}_{t+1}^{r}100\chi^{r}\left(\frac{\widetilde{E}_{t+1}^{r} - \widetilde{E}_t^{r}}{\widetilde{K}_{t+1}^{r}}\Gamma_{t+1}^{x,cbi}\right)\right\}\end{aligned} \tag{51}$$

$$\widetilde{K}_{t+1}^{r} = (1-\delta^{r})\frac{\widetilde{K}_t^{r}}{\Gamma_t^{x,cbi}} + A_t^{r}\widetilde{E}_t^{r} - \frac{100\chi^{r}}{2}\left(\frac{\widetilde{E}_t^{r} - \widetilde{E}_{t-1}^{r}}{\widetilde{K}_t^{r}}\Gamma_t^{x,cbi}\right)^2 \frac{\widetilde{K}_t^{r}}{\Gamma_t^{x,cbi}} \tag{52}$$

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:

$$\widetilde{\Lambda}_t^{cnn} = \beta R_t E_t \left\{ \widetilde{\Lambda}_t^{cnn} \frac{1}{\Pi_{t+1}^{p,cbi} \Gamma_{t+1}^{x,cbi}} \right\} \tag{53}$$

$$\widetilde{\Lambda}_t^{cnn} = \widetilde{\Lambda}_t^{cd} \frac{1}{\widetilde{R}_t^{cd}} \tag{54}$$

$$\widetilde{\Lambda}_t^{cnn} = \widetilde{\Lambda}_t^{r} \frac{1}{\widetilde{R}_t^{r}} \tag{55}$$

$$\widetilde{\Lambda}_t^{cnn} = \varsigma^{cnn} \frac{\Xi_t^{cnn}}{\widetilde{E}_t^{cnn} - (h^{cnn}/\Gamma_t^{x,cbi})\widetilde{E}_{t-1}^{cnn}} - \beta \varsigma^{cnn} E_t \left\{ \frac{(h^{cnn}/\Gamma_{t+1}^{x,cbi})\Xi_{t+1}^{cnn}}{\widetilde{E}_t^{cnn} - (h^{cnn}/\Gamma_{t+1}^{x,cbi})\widetilde{E}_t^{cnn}} \right\} \tag{56}$$

$$\frac{\widetilde{\Lambda}_t^{cd}}{\Gamma_t^{x,kb}} = \varsigma^{cd} \frac{\Xi_t^{cd}}{\widetilde{K}_t^{cd} - (h^{cd}/\Gamma_{t-1}^{x,kb})\widetilde{K}_{t-1}^{cd}} - \beta \varsigma^{cd} E_t \left\{ \frac{(h^{cd}/\Gamma_t^{x,kb})\Xi_{t+1}^{cd}}{\widetilde{K}_{t+1}^{cd} - (h^{cd}/\Gamma_t^{x,kb})\widetilde{K}_t^{cd}} \right\} \tag{57}$$

$$\frac{\widetilde{\Lambda}_t^{r}}{\Gamma_t^{x,cbi}} = \varsigma^{r} \frac{\Xi_t^{r}}{\widetilde{K}_t^{r} - (h^{r}/\Gamma_{t-1}^{x,cbi})\widetilde{K}_{t-1}^{r}} - \beta \varsigma^{r} E_t \left\{ \frac{(h^{r}/\Gamma_t^{x,cbi})\Xi_{t+1}^{r}}{\widetilde{K}_{t+1}^{r} - (h^{r}/\Gamma_t^{x,cbi})\widetilde{K}_t^{r}} \right\}, \tag{58}$$

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:

$$
\begin{aligned}
\Theta_t^l \quad & \frac{\Lambda_t^{l,cbi}}{\widetilde{\Lambda}_t^{cnn}} L_t^{cbi} \\
= \quad & (\Theta_t^l - 1)\widetilde{W}_t^{cbi} L_t^{cbi} \\
- \quad & \Theta_t^l 100\chi^l \left(\frac{L_*^{cbi}}{L_*^{cbi}+L_*^{kb}} \widetilde{W}_t^{cbi} + \frac{L_*^{kb}}{L_*^{cbi}+L_*^{kb}} \widetilde{W}_t^{kb} \right) \left(\frac{L_t^{cbi}}{L_t^{kb}} - \eta^l \frac{L_{t-1}^{cbi}}{L_{t-1}^{kb}} - (1-\eta^l)\frac{L_*^{cbi}}{L_*^{kb}} \right) \\
+ \quad & 100\chi^\omega \left(\Pi_t^{\omega,cbi} - \eta^\omega \Pi_{t-1}^{\omega,cbi} - (1-\eta^\omega)\Pi_*^{\omega,cbi} \right) \Pi_t^{\omega,cbi} \widetilde{W}_t^{cbi} L_t^{cbi} \\
- \quad & \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} 100\chi^\omega \left(\Pi_{t+1}^{\omega,cbi} - \eta^\omega \Pi_t^{\omega,cbi} - (1-\eta^\omega)\Pi_*^{\omega,cbi} \right) \Pi_{t+1}^{\omega,cbi} \widetilde{W}_{t+1}^{cbi} L_{t+1}^{cbi} \right\}
\end{aligned}
\tag{59}
$$

and

$$
\begin{aligned}
\Theta_t^l \quad & \frac{\Lambda_t^{l,kb}}{\widetilde{\Lambda}_t^{cnn}} L_t^{kb} \\
= \quad & (\Theta_t^l - 1)\widetilde{W}_t^{kb} L_t^{kb} \\
+ \quad & \Theta_t^l 100\chi^l \left(\frac{L_*^{cbi}}{L_*^{cbi}+L_*^{kb}} \widetilde{W}_t^{cbi} + \frac{L_*^{kb}}{L_*^{cbi}+L_*^{kb}} \widetilde{W}_t^{kb} \right) \left(\frac{L_t^{cbi}}{L_t^{kb}} - \eta^l \frac{L_{t-1}^{cbi}}{L_{t-1}^{kb}} - (1-\eta^l)\frac{L_*^{cbi}}{L_*^{kb}} \right) \\
+ \quad & 100\chi^\omega \left(\Pi_t^{\omega,kb} - \eta^\omega \Pi_{t-1}^{\omega,kb} - (1-\eta^\omega)\Pi_*^{\omega,kb} \right) \Pi_t^{\omega,kb} \widetilde{W}_t^{kb} L_t^{kb} \\
- \quad & \beta E_t \left\{ \frac{\widetilde{\Lambda}_{t+1}^{cnn}}{\widetilde{\Lambda}_t^{cnn}} 100\chi^\omega \left(\Pi_{t+1}^{\omega,kb} - \eta^\omega \Pi_t^{\omega,kb} - (1-\eta^\omega)\Pi_*^{\omega,kb} \right) \Pi_{t+1}^{\omega,kb} \widetilde{W}_{t+1}^{kb} L_{t+1}^{kb} \right\}.
\end{aligned}
\tag{60}
$$

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 $\bar{R}_t$:

$$
R_t = (R_{t-1})^{\phi^r} (\bar{R}_t)^{(1-\phi^r)} \exp[\varepsilon_t^r] \tag{61}
$$

$$
\begin{aligned}
\bar{R}_t = \quad & \left(\Pi_t^{p,gdp} / \Pi_*^{p,gdp} \right)^{\phi^{\pi,gdp}} \left(\Delta \Pi_t^{p,gdp} \right)^{\phi^{\Delta\pi,gdp}} \\
& \left(H_t^{gdp} / H_*^{gdp} \right)^{\phi^{h,gdp}} \left(\Delta H_t^{gdp} \right)^{\phi^{\Delta h,gdp}} R_*.
\end{aligned}
\tag{62}
$$

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.

Appendix A2: 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 paper.

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

[17] Scanned documents are available as .pdf files on http://www.census.gov/prod/www/abs/statab.html

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.

Surprising Comparative Properties of Monetary Models: Results from a New Model Database

John B. Taylor and Volker Wieland*

This Version: September 2, 2010

Abstract

In this paper we investigate the comparative properties of empirically-estimated monetary models of the U.S. economy. We make use of a new databaseof models designed for such investigations. We focus on three representative models: the Christiano, Eichenbaum, Evans (2005) model, the Smets and Wouters (2007) model, and the Taylor (1993a) model. Although the three models differ in terms of structure, estimation method, sample period, and data vintage, we find surprisingly similar economic impacts of unanticipated changes in the federal funds rate. However, the optimal monetary policy rules are different in the different models. Simple model-specific policy rules that include the lagged interest rate, inflation and current and lagged output gaps are not robust. Some degree of robustness can be recovered by using rules without interest-rate smoothing or with GDP growth deviations from trend in place of the output gap. However, improvement vis-à-vis other models, comes at the cost of significant performance deterioration in the original model. Model averaging offers a much more effective strategy for improving the robustness of policy rules.

* John B. Taylor is Mary and Robert Raymond Professor of Economics at Stanford University and George P. Shultz Senior Fellow in Economics at Stanford's Hoover Institution. Volker Wieland is Professor for Monetary Theory and Policy at Goethe University of Frankfurt. Wieland is grateful for financial support by a Willem Duisenberg Research Fellowship at the European Central Bank while this research project was initiated. Excellent research assistance was provided by Elena Afanasyeva, Tobias Cwik and Maik Wolters from Goethe University Frankfurt. Wieland acknowledges research assistance funding from European Community grant MONFISPOL under grant agreement SSH-CT-2009-225149. Comments by Mark Watson, Frank Smets, three anonymous referees and session participants at the AEA Meetings 2009, the German Economic Association Monetary Committee Meeting 2009, and the NBER Summer Institute 2009 Monetary Economics Group also proved very helpful. All remaining errors are our own.
Contact: Taylor: Herbert Hoover Memorial Building, Stanford University, Stanford, CA 94305, U.S.A.
Wieland: House of Finance, Goethe University of Frankfurt, Grueneburgplatz 1, D-60323 Frankfurt am Main, Germany, wieland@wiwi.uni-frankfurt.de.

Ever since the 1970s revolution in macroeconomics, monetary economists have been building quantitative models that incorporate the fundamental ideas of the Lucas critique, time inconsistency, and forward-looking expectations, in order to evaluate monetary policy more effectively. The common characteristic of these monetary models, compared with earlier models, is the combination of rational expectations, staggered price and wage setting, and policy rules, all of which have proved essential to policy evaluation.

Over the years the number of monetary models with these characteristics has grown rapidly as the ideas have been applied in more countries, as researchers have endeavoured to improve on existing models by building new ones, and as more data shed light on the monetary transmission process. The last decade, in particular, has witnessed a surge of macroeconomic model building as researchers have further developed the microeconomic foundations of monetary models and applied new estimation methods. In our view it is important for research progress to document and compare these models and assess the value of model improvements in terms of the objectives of monetary policy evaluation. Keeping track of the different models is also important for monetary policy in practice because by checking the robustness of policy in different models one can better assess policy.

With these model comparison and robustness goals in mind we have recently created a new "monetary model database," an interactive collection of models that can be simulated, optimized, and compared. The monetary model database can be used for model comparison projects and policy robustness exercises. Perhaps because of the large number of models and the time and cost of bringing modellers together, there have not been many model comparison projects and robustness exercises in recent years. In fact the most recent policy robustness exercise, which we both participated in, occurred 10 years ago as part of an NBER

conference.[1] Our monetary model base provides a new platform that makes model comparison much easier than in the past and allows individual researchers easy access to a wide variety of macroeconomic models and a standard set of relevant benchmarks.[2] We hope in particular that many central banks will participate and benefit from this effort as a means of getting feedback on model development efforts.

This paper investigates the implications of three well-known models included in the model database for monetary policy in the U.S. economy. The first model, which is a multi-country model of the G-7 economies built more than fifteen years ago, has been used extensively in the earlier model comparison projects. It is described in detail in Taylor (1993a). The other two models are the best known representatives of the most recent generation of empirically estimated new Keynesian models, the Christiano, Eichenbaum and Evans (2005) model of the United States and the Smets and Wouters (2007) model of the United States.

The latter two models incorporate the most recent methodological advances in terms of modelling the implications of optimizing behavior of households and firms. They also utilize new estimation methods. The Christiano, Eichenbaum and Evans (2005) model is estimated to fit the dynamic responses of key macroeconomic variables to a monetary policy shock identified with a structural vector autoregression. The Smets and Wouters (2007) model is estimated with Bayesian methods to fit the dynamic properties of a range of key variables in response to a full set of shocks.

---

[1] The results are reported in the conference volume, *Monetary Policy Rules,* Taylor (1999). Several of the models in this earlier comparison and robustness exercise are also included in our new monetary model database, including Rotemberg-Woodford (1999), McCallum and Nelson (1999), and Taylor (1993).

[2] See the Appendix of this paper for the current list of 35 models and Wieland, Cwik, Müller, Schmidt and Wolters (2009) for a detailed exposition of the platform for model comparison. The model base includes small calibrated text-book-style models, estimated medium- and large-scale models of the U.S. and euro area economies, and some estimated open-economy and multi-country models. Software and models are available for download from http://www.macromodelbase.com. This platform relies on the DYNARE software for model solution and may be used with Matlab. For further information on DYNARE see Collard and Juillard (2001) and Juillard (1996) and http://www.cepremap.cnrs.fr/dynare/.

First, we examine and compare the monetary transmission process in each model by studying the impact of monetary policy shocks in each model. Second, we calculate and compare the optimal monetary policy rules within a certain simple class for each of the models. Third, we evaluate the robustness of these policy rules by examining their effects in each of the other models relative to the rule that would be optimal for the respective model.

The model comparison and robustness analysis reveals some surprising results. Even though the two more recent models differ from the Taylor (1993a) model in terms of economic structure, estimation method, data sample and data vintage, they imply almost identical estimates of the response of U.S. GDP to an unexpected change in the federal funds rate, that is, to a monetary policy shock. This result is particular surprising in light of earlier findings by Levin, Wieland and Williams (1999, 2003) indicating that a number of models built after Taylor (1993a) exhibit quite different estimates of the impact of a monetary policy shock and the monetary transmission mechanism.[3] We also compare the dynamic responses to other shocks. Interestingly, the impact of the main financial shock, that is the risk premium shock, on U.S. GDP is also quite similar in the Smets and Wouters (2008) and the Taylor (1993) model. This finding is of interest in light of the dramatic increase in risk premia observed since the start of the financial crisis in August 2007.[4] Differences emerge with regard to the consequences of other demand and supply shocks.

The analysis of optimized simple interest rate rules reveals further interesting similarities and differences across the three models. All three models prefer rules that include the lagged interest rate in addition to inflation deviations from target and output deviations from potential. The two more recent new Keynesian models favour the inclusion of the growth rate of output gaps.

---

[3] For example, the model of Fuhrer and Moore (1995) and the Federal Reserve's FRB/US model of Reifschneider et al (1999), both exhibited longer-lasting effects of policy shocks on U.S. GDP that peak several quarters later than in Taylor (1993a). See Levin, Wieland and Williams (1999, 2003) for a comparison.

[4] As noted by Smets and Wouters (2007) the risk premium shock represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households and has similar effects as so-called net-worth shocks in models with an explicit financial sector such as Bernanke et al (1999).

The robustness exercise, however, delivers more nuanced results. Model-specific rules with interest rate smoothing and output gaps are not robust. Some degree of robustness can be recovered by focusing on 2-parameter rules with inflation and the output gap, or 3-parameter rules with interest-rate smoothing, inflation and the deviation of output growth from trend instead of output gap growth. This increase in robustness vis-à-vis other models comes at the cost of significant performance deterioration in the original model. Fortunately, however, model comparison offers an avenue for improving over the robustness properties of model-specific rules. Rules that are optimized with respect to the average loss across multiple models achieve very good robustness properties at much lower cost.

1. Brief Description of the Models

Taylor (1993a)

This is an econometrically-estimated rational expectations model fit to data from the G7 economies for the period 1971:1 to 1986:4. All our simulations focus on the United States. The model was built to evaluate monetary policy rules and was used in the original design of the Taylor rule. It has also been part of several model comparison exercises including Bryant et al (1985), Klein (1991), Bryant et al (1993) and Taylor (1999). Shiller (1991) compared this model to the "old Keynesian" models of the pre rational expectations era, and he found that there were large differences in the impact of monetary policy due largely to the assumptions of rational expectations and more structural models of wage and price stickiness.

To model wage and price stickiness Taylor (1993a) used the staggered wage and price setting approach rather than ad hoc lags of prices or wages which characterized the older pre-rational expectations models. However, because the Taylor (1993a) model was empirically estimated it used neither the simple example of constant-length four-quarter contracts presented in Taylor (1980) nor the geometrically-distributed contract weights proposed by Calvo (1983). Rather it lets the weights have a general distribution which is empirically

estimated using aggregate wage data in the different countries. In Japan some synchronization is allowed for.

The financial sector is based on several "no-arbitrage" conditions for the term structure of interest rates and the exchange rate. Expectations of future interest rates affect consumption and investment, and exchange rates affect net exports. Slow adjustment of consumption and investment is explained by adjustment costs such as habit formation or accelerator dynamics. A core principle of this model is that after a monetary shock the economy returns to a growth trend, which is assumed to be exogenous to monetary policy as in the classical dichotomy.

Most of the equations of the model were estimated with Hansen's instrumental variables estimation method, with the exception of the staggered wage setting equations which were estimated with maximum likelihood.

Christiano, Eichenbaum, Evans* (*2005)

Many of the equations in the model of Christiano, Eichenbaum and Evans (CEE 2005 in the following) exhibit similarities to the equations in the Taylor model, but they are explicitly-derived log-linear approximations of the first-order conditions of optimizing representative firms and households. Their model also assumes staggered contracts but with Calvo weights and backward-looking indexation in those periods when prices and wages are not set optimally. Long-run growth and short-run fluctuations are modelled jointly rather than separately as in Taylor's model. Thus, the CEE (2005) model explicitly accounts for labor supply dynamics as well as the interaction of investment demand, capital accumulation and utilization. Furthermore, their model includes a cost-channel of monetary policy. Firms must borrow working capital to finance their wage bill. Thus, monetary policy rates have an immediate impact on firms' profitability.

The CEE (2005) model was estimated for the U.S. economy over the period 1959:2-2001:4 by matching the impulse response function to the monetary shock in a structural VAR. An important assumption of the VAR that carries over to the model is that monetary policy innovations affect the interest rate in the *same* quarter, but other variables, including output and inflation, only by the *following* quarter.

The monetary policy innovation represents the single, exogenous economic shock in the original CEE model. However, additional shocks can be incorporated in the structural model and the variance of such shocks may be estimated using the same methodology. The additional shocks would first be identified in the structural VAR. Then, the parameters of the structural model including innovation variances would be re-estimated by matching the impulse response functions implied by the model with their empirical counterparts from the VAR. Altig, Christiano, Eichenbaum and Linde (2004), (ACEL 2004 in the following), follow this approach and identify two additional shocks – a neutral and an investment-specific technology shock. These shocks exhibit serial correlation and have permanent effects on the level of productivity. Together with the monetary policy shock they account for about 50% of the variation in output. The impulse response function for the monetary policy shock in ACEL (2004) is almost identical to CEE (2005). Therefore, we will use the ACEL (2004) parameterization of the CEE model for the computational analysis in our paper. A drawback of this model is that it does not yet provide a complete characterization of the observed output and inflation volatility.

The CEE model, which was initially circulated in 2001, represented the first medium-sized, estimated example of the new generation of New-Keynesian dynamic stochastic general equilibrium models explicitly derived from optimizing behavior of representative households and firms.[5] It stimulated the development of similar optimization-based models for many other countries once Smets and Wouters (2003) showed how to make use of new

[5] The paper was published in 2001 as NBER Working Paper 8403.

advances in Bayesian techniques (see e.g. Geweke (1999) and Schorfheide (2000)) in estimating such models.

Smets and Wouters (2007)

The model of the U.S. economy estimated by Smets and Wouters (2007) (SW 2007 in the following) with U.S. data from 1966:1 to 2004:4 may be viewed as an extended version of the CEE/ACEL model. The SW model contains a greater set of macroeconomic shocks and aims to fully explain the variation in key variables, such as aggregate output and its components as well as inflation, wages and interest rates. They use a Bayesian estimation methodology that allows the use of priors on model parameters informed from theory and literature. The posterior distributions then incorporate the information in the available macroeconomic data. Whenever the data does not help in pinpointing parameter values very precisely, theoretical priors dominate. Such priors can in some cases be based on evidence from microeconomic studies. The Bayesian estimation methodology has quickly been popularized and widely applied by researchers in central banks and academia. It has been implemented for use with the DYNARE software that we also utilize in our model base.

Smets and Wouters (2007) modify some of the structural assumptions embodied in the CEE/ACEL model. In the long-run, the SW model is consistent with a balanced steady-state growth path driven by deterministic labor-augmenting technological progress. While the CEE model assumes wages and prices are indexed to last period's inflation rate in the absence of a Calvo-style signal, the SW model allows firms to index to a weighted average of lagged and steady-state inflation. Furthermore, SW drop two more assumptions that have important short-run implications in the CEE/ACEL model. First, they do not impose the delayed effect of monetary policy on other variables that CEE built into the structural model so as to match the constraints required by the structural VAR to identify monetary policy shock. Second, SW (2007) do not require firms to borrow working capital to pay the wage bill. Thus, the so-called

cost channel is absent from the model. Smets and Wouters note that they did not find this channel necessary for fitting the dynamics in U.S. data. In our simulations, we will also investigate the implications of adopting the SW assumptions of no cost channel and no timing constraints on monetary policy shocks in the original CEE/ACEL model.

2. Shocks to Monetary Policy as Deviations from Two Policy Rules

We first use the model database to assess the extent of differences between models regarding the transmission of monetary policy to output and inflation. To this end we compare the effect of monetary policy shocks in the three models. A monetary policy shock is defined as a surprise deviation from systematic policy behavior which is characterized by interest rate policy rules.

In our comparison, we focus on two estimated rules used by SW 2007 and CEE 2005 respectively to characterize systematic central bank policy. Smets and Wouters estimate the coefficients of this interest rate rule along with the other equations in their model. We refer to it as the SW rule in the remainder of the paper. They call it a generalized Taylor rule, because it includes the lagged federal funds rate, the lagged output gap, and a serially correlated policy shock, in addition to the current inflation rate and output gap that appear in the original Taylor (1993b) rule. The SW rule implies the following setting for the federal funds rate, i_t:

(1) $$i_t = 0.81 i_{t-1} + 0.39\pi_t + 0.97 y_t - 0.90 y_{t-1} + \varepsilon_t^i, \text{ where } \varepsilon_t^i = 0.15\varepsilon_{t-1}^i + \eta_t^i$$

Here, π_t refers to the annualized, quarterly inflation rate and y_t to the output gap.[6] In the SW and CEE model the gap measure used in the policy rule is defined as the difference between

[6] Note, the response coefficients differ from the values reported in SW 2007. In equation (1), interest and inflation rates are annualized, while SW used quarterly rates. The original specification in SW 2007 corresponds to $i_t^q = (1-0.81)(2.04\pi_t^q + 0.09 y_t) + 0.22\Delta y_t + 0.81 i_{t-1}^q + \varepsilon_t^i$, where the superscript q refers to quarterly rates.

the actual output level and the level that would be realized if prices adjust flexibly to macroeconomic shocks, the so-called flex-price output level.[7] In the Taylor model (and the original Taylor rule) the output gap is defined as difference between actual output and long-run potential output as measured by the trend. The policy shock is denoted by ε_t^i and follows a first-order autoregressive process with an IID normal error term, η_t^i. As a result of serial correlation and the inclusion of the lagged interest rate in the reaction function, an IID innovation will have a persistent effect on nominal interest rates and due to price rigidity also on real rates and aggregate output.

CEE (2005) define the central bank's policy rule in terms of a reaction function for the growth rate of money.[8] They identify monetary policy shocks in a structural VAR as orthogonal innovations to the interest rate reaction function. Then, they estimate the parameters of the structural model including the parameters of the money growth rule by matching the impulse response in the structural model and the VAR. In addition, they contrast their findings under the money growth rule with the effect of a policy shock under an extended Taylor rule for the federal funds rate:[9]

(2) $$i_t = 0.80 i_{t-1} + 0.3 E_t \pi_{t+1} + 0.08 y_t + \varepsilon_t^i$$

Just like the SW rule it incorporates partial adjustment to the lagged federal funds rate. However, it is forward-looking and responds to the expected inflation rate for the upcoming quarter. The coefficient on the output gap is much smaller than in the SW rule and it does not include the lag of the output gap. The policy shock is IID. In the following we refer to this rule as the CEE rule.

---

[7] Smets and Wouters set wage and price markup shocks equal to zero in the derivation of the flex-price output measure used to define their output gap.

[8] CEE (2005) and ACEL(2004) model monetary policy in terms of innovations to the growth-rate of money that they denote by μ_t: $\mu_t = \mu + \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} \ldots$

[9] Note, we use annualized interest and inflation rates and transcribe the CEE rule accordingly. In CEE 2005 they define their rule as: $i_t^q = (1-0.80)(1.5 E_t \pi_{t+1}^q + 0.1 y_t) + 0.8 i_{t-1}^q + \varepsilon_t^i$. CEE (2005) attribute this estimated rule to Clarida et al (1999). However, the coefficients reported in Clarida et al (1999) are different. Their rule corresponds to $i_t = (1-0.79)(2.15 E_t \pi_{t+1} + 0.93 y_t) + 0.79 i_{t-1} + \varepsilon_t^i$.

3. Monetary Policy Shocks in Three Monetary Models of the U.S. Economy

We compare the consequences of a monetary policy shock in the Taylor, SW and CEE/ACEL models to shed light on their implications for the transmission of Federal Reserve interest rate decisions to aggregate output and inflation. In particular, we want to find out to what extent the current-generation DSGE models, CEE/ACEL (2004) and SW (2007), imply quantitatively different effects of monetary policy than the model by Taylor (1993a). Since the models differ in terms of economic structure and parameter estimates are obtained for different data series, estimation periods and data vintages, we would expect to obtain quantitatively different assessments of the monetary transmission mechanism.

Figure 1 reports the consequences of a 1 percentage point shock to the federal funds rate for nominal interest rates, output and inflation. The panels on the left-hand side refer to the outcomes when the Federal Reserve sets interest rates following the initial shock according to the prescriptions of the SW rule, while the right-hand-side panels refer to the outcome under the CEE rule. Each panel shows the findings from four model simulations. The dark solid line refers to the SW model, the light solid line to the TAYLOR model, the dashed line to the CEE/ACEL model and the dotted line to the CEE/ACEL model with SW assumptions.[10]

Surprisingly, the effect of the policy shock on real output and inflation given a common policy rule is very similar in the four models. For example, under the SW rule the nominal interest rate increases on impact by 0.8 to 1 percentage points and then returns slowly to steady state, real output falls over three to four quarters to a trough of about -0.35 percent

[10] The CEE/ACEL model with SW Assumptions implies the following modifications: We remove the timing constraints that were imposed on the structural model by the authors so that it coincides with the identification restrictions on the VAR that they used to obtain impulse responses for the monetary policy shock. Furthermore we remove the constraint from the ACEL model that requires firms to finance the wage bill by borrowing cash in advance from a financial intermediary. As a result of this constraint the interest rate has a direct effect on firms' costs.

before returning to steady-state, and inflation declines more slowly with a trough of about 20 basis points roughly 2 to 3 quarters later than output.

The quantitative implications for real output in the Taylor (1993) and SW (2007) models are almost identical. The outcome under the CEE/ACEL model initially differs slightly from the other two models. In the quarter of the shock we observe a tiny increase in output, while inflation does not react at all. From the second quarter onwards output declines to the same extent as in the other two models but the profile is shifted roughly one quarter into the future. The decline in inflation is similarly delayed. Once we implement the CEE/ACEL model with the SW assumptions of no timing constraint on policy and no cost channel, the timing of output and inflation dynamics is more similar to the other two models.

The outcome of a monetary policy shock given the Fed follows the CEE rule is shown in the right-hand-side panels of **Figure 1**. Again, the magnitude of the effect of the policy shock on real output and inflation is almost identical in the Taylor model, the SW model and the ACEL/CEE model, particularly when the latter model is implemented with the SW assumptions. Furthermore, the reduction in output is very similar to the case when the Fed follows the SW rule. The decline in inflation is a bit smaller.

The original Lucas critique stated that a change in the systematic component of policy can have important implications for the dynamics of macroeconomic variables. Thus, it is not surprising that the output and inflation effects of monetary policy shocks change if we consider a wider set of monetary policy rules. For example, in the case of the original Taylor (1993b) rule an IID policy shock would influence the nominal interest rate only for one period, because the Taylor rule does not include the lagged interest rate. We have investigated the real output effects of a monetary policy shock with different response coefficients (for example, a four times smaller response to output), different inflation measures (such as year-on-year inflation) and different rules such as the original Taylor rule or the benchmark rules considered in Levin, Wieland and Williams (2003) and Kuester and Wieland (2010).

Different rules have quite different implications for the real consequences of monetary policy shocks. However, the Taylor model, the SW model and the CEE model continue to imply surprisingly similar dynamics of aggregate real output and inflation in response to a policy shock for a given, common policy rule.

The finding that the two best-known models of the recent generation of new Keynesian models provide very similar estimates of the impact of a policy shock on U.S. real GDP as the model of Taylor (1993a) is particularly surprising in light of earlier comparison projects. For example, the comparison in Levin, Wieland and Williams (1999) and (2003) indicated that models built and estimated after Taylor (1993a) such as the model of Fuhrer and Moore (1995) or the Federal Reserve's FRB/US model of Reifschneider, Tetlow and Williams (1999) provided different assessments of the U.S. monetary transmission mechanism. In particular, these models suggested that the impact of monetary policy shocks on real output would be longer-lasting and reach its peak more than a year after the initial impulse. This view is often considered conventional wisdom among practitioners. The model data base associated with this paper also allows users to replicate the impulse response function comparison in the Fuhrer and Moore and FRB/US models.

So far we have focused on the overall effect of the policy shock on output and inflation. Now we turn to the effects on other macroeconomic variables. **Figure 2** illustrates some additional common aspects of the transmission mechanism in the three models of the U.S. economy, while **Figure 3** highlights interesting differences. Monetary policy is assumed to follow the SW rule after the policy shock.[11] The real interest rate increases almost to the same extent in all three models as shown in panel 2a. As a result, aggregate consumption and aggregate investment decline. The decline in consumption is smaller in the Taylor model than in the other two models, while the decline in investment is much greater. The quantitative comparison of the dynamics of GDP components, however, is hampered by the fact that the

[11] Similar figures for the case of the CEE rule are provided in the online appendix.

models use different deflators in generating real consumption and investment series.[12] Another similarity regarding monetary policy transmission in the three models is that real wages decline along with aggregate demand following the monetary policy shock.

The three models also exhibit some interesting differences regarding monetary policy transmission. For example, panels a. and b. in **Figure 3** indicate that only the Taylor model accounts for international feedback effects. As a result of the policy shock the US dollar appreciates temporarily in real trade-weighted terms. Exports and imports, both, decline. However, the fall in imports is much greater than in exports and as a result net exports increase. The strong decline in imports occurs due to the domestic demand effect that figures very importantly in the U.S. import demand equation. The resulting increase in net exports partly offsets the impact of the large negative decline in investment demand on aggregate output in the Taylor model. Furthermore, panels c. through f. in **Figure 3** illustrate that only the SW and CEE models account for the effects of the policy shock on labor supply, capital stock, the rental rate of capital and capital utilization. All four measures decline in response to the monetary shock. This explanation of supply-side dynamics is missing from the Taylor model.

4. Other shocks and their implications for policy design

Unexpected changes in monetary policy are of interest in order to identify aspects of the monetary transmission mechanism. When it comes to the question of policy design, however, the standard recommendation is to avoid policy surprises since they only generate additional output and inflation volatility. Instead optimal and robust policy design focuses on the proper choice of the variables and the magnitude of the response coefficients in the policy rule that

[12] While the Taylor model simulates the components of GDP in real terms, the simulations in the SW and CEE models concern the nominal components divided by the GDP deflator. It is not possible to make the series directly comparable because none of the models accounts for the consumption and investment deflators separately from the GDP deflator.

characterizes the systematic component of monetary policy. The policy rule is then designed to stabilize output and inflation in the event of shocks emanating from other sectors of the economy. In this respect, it is of interest to review and compare the potential sources of economic shocks in the three models under consideration.

In light of the recent financial crisis, we start by comparing the effect of particular financial shocks. Only the Taylor and SW models contain such shocks. **Figure 4** illustrates the effect of an increase in the term premium by 1 percentage point on real output and inflation in the Taylor and SW models. The initial impact of these shocks on real output is almost identical in the two models and lies between -0.22 and -0.24 percent of output. This finding is particularly surprising since the shocks are estimated quite differently in the two models. In the Taylor model the term premium shock is estimated from the term structure equation directly using data on short- and long-term interest rates, that is, the federal funds rate versus 10-year US treasuries. In the SW model the risk premium shock is estimated from the consumption and investment equation. It assumes the term structure relation implicitly but uses no data on long-term rates. In earlier work on the euro area, Smets and Wouters (2003) included instead a consumption demand or preference shock. This shock is omitted in their model of the U.S. economy to keep the number of shocks in line with the number of observed variables. SW emphasize that the premium shock represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households and has similar effects as so-called net-worth shocks in models with an explicit financial sector such as Bernanke et al (1999).[13]

[13] In the model file available from the AER website along with the SW (2007) paper the shock is multiplied with minus the consumption elasticity. This is consistent with figure 2 of that paper, where the shock appears as a "demand" shock, i.e. an increase has a positive effect on output. It is not consistent with equation (2) in SW (2007) that identifies the shock as a risk premium shock (i.e. an increase has a negative effect). We have modified the model file consistent with the notation as risk premium shock in equation (2) in SW (2007). In addition, we have checked that re-estimating the SW model with the shock entering the consumption Euler equation as defined by equation (2) in their paper does not have an important effect on the parameter estimates.

Figure 5 provides a comparison of what could be termed "demand" or "spending" shocks in the three models. These are shocks that push output and inflation in the same direction. The Taylor model contains many such shocks. Panels a. and b. show the effects of shocks to non-durables consumption, equipment investment, inventory investment, government spending and import demand on the output gap and inflation. The SW model contains two shocks of this type, an exogenous spending shock that comprises government spending as well as net exports and an investment-specific technology shock. The ACEL model contains an investment-specific technology shock that initially lowers inflation but then raises it. It has stronger long-term effects than the investment-specific technology shock in SW (2007).

Figure 6 compares supply shocks in the three models, i.e. shocks that push output and inflation in opposite directions. The Taylor model has a number of such shocks, in particular innovations to the contract wage equations, the final goods price equation, import prices and export prices. The SW model contains price mark-up and wage mark-up shocks that are somewhat similar to the contract wage and aggregate price shocks in the Taylor model. Only the SW and the ACEL models include neutral technology shocks. In the ACEL model these shocks have a long-term effect on productivity growth, while their effect on productivity growth in the SW model is temporary.

Comparing the three models, it is important to keep in mind that only the Taylor and SW model aim to fully explain the variation in the macroeconomic variables included in the model as an outcome of exogenous shocks and endogenous propagation. The ACEL model only aims to explain that part of the variation that is caused by the three shocks in the structural VAR that was used to identify them. **Figures 5** and **6** indicate that the investment-specific and neutral technology shocks in the ACEL model have negligible effects on inflation. Consequently, the ACEL model omits most sources of inflation volatility outside of policy shocks and is of limited usefulness for designing monetary policy rules. With this

caution in mind, we will nevertheless explore the implications of the ACEL model for policy design together with the other two models.

5. Optimal simple policy rules in the Taylor, CEE/ACEL and SW models

The first question on policy design, that we address concerns the models´ recommendations for the optimal policy response to a small number of variables in a simple interest rate rule. We start by considering rules that incorporate a policy response to two variables, that is, the current year-on-year inflation rate and the output gap as in the original Taylor (1993b) rule:

(3) $$i_t = \alpha\pi_t + \beta_0 y_t$$

In the SW and ACEL models, the output gap y is defined as the deviation of actual output from the level of output that would be realized if the price level were fully flexible. This flexible-price output varies in response to some of the economic shocks. We use the same definition of flexible price output as in Smets and Wouters (2007). In the Taylor model the gap is calculated relative to a measure of potential that grows at an exogenous rate.

In a second step, we extend the rule to include the lagged nominal interest rate as in Levin, Wieland and Williams (1999, 2003):

(4) $$i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$$

Then, we also include the lagged output gap as in the estimated rule in the Smets and Wouters (2007) model:

(5) $$i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$$

We choose the response coefficients of the rules (i.e. ρ, α, β_0, β_1) in each of the models by minimizing a loss function L that includes the unconditional variances of inflation, the output gap and the change of the nominal interest rate:

$$L = Var(\pi) + \lambda_y Var(y) + \lambda_{\Delta i} Var(\Delta i) \tag{6}$$

This form of loss function has been used extensively in earlier analyses, including the above-mentioned model comparison studies. With $\lambda_{\Delta i}$=0, it corresponds to the unconditional expectation of a second-order approximation of household utility in a small New-Keynesian model derived from microeconomic foundations as shown in Rotemberg and Woodford (1999). The magnitude of the implied value of λ_y is very sensitive to the particular specification of overlapping nominal contracts: random-duration "Calvo-style" contracts imply a very low value on the order of 0.01, whereas fixed-duration "Taylor-style" contracts imply a value near unity (Erceg and Levin (2001)). For this reason, we consider values of λ_y=(0, 0.5,1). In addition, we assign a positive weight to interest volatility and consider values of $\lambda_{\Delta i}$ =(0.5,1). It is intended to capture central banks' well-known tendency to smooth interest rates and to avoid extreme values of optimized response coefficients that would be very far from empirical observation and regularly violate the non-negativity constraint on nominal interest rates (Woodford 1999).

The optimized response coefficients are shown in **Table 1**. It reports results for two-, three- and four-parameter rules in the Taylor, SW and CEE/ACEL models. The central bank's objective is assumed to assign a weight of unity to inflation and interest rate volatility

and either a weight of zero or unity to output gap volatility.[14] First, with regard to two-parameter rules all three models prescribe a large response coefficient on inflation and a small coefficient on the output gap, if the output gap does not appear in the loss function. If the output gap receives equal weight in the loss function then the optimal coefficient on output increases but remains quite a bit below the response to inflation. The coefficient on inflation declines in the SW and CEE/ACEL models but increases in the Taylor model when output appears in the loss function.

For three-parameter rules the optimized value of the coefficient on the lagged nominal interest rate is near unity. This property applies in all three models and with different values of the objective function weights except for one case that is discussed below. The coefficients on inflation are much smaller than in the two-parameter rules but they typically remain positive.

In the ACEL model the loss function is very flat. There appear to be multiple local optima and the global optimum we identify has very extreme coefficients in the case of the three-parameter rule with a positive weight on output gap volatility in the loss function.[15] As noted earlier, a weakness of the ACEL model is that it only contains two technology shocks that explain little of the variation of inflation and output gaps but have permanent effects on the growth of steady state output. The ACEL model contains no short-run demand and supply shocks as do the other two models. For this reason the model may not be considered suitable in its current form for an evaluation of the role of interest rate rules in stabilization policy. Nevertheless, we continue to replicate the analysis conducted in the other two models also in the ACEL model throughout this paper.[16]

---

[14] Additional findings for a weight of 0.5 on the unconditional variance of the change of the nominal interest rate are reported in the additional appendix available online. Further sensitivity studies for intermediate weights have been conducted and are available from the authors upon request.

[15] A local optimum at less extreme values is observed for $\rho = 0.01$, $\alpha = 2.9$, $\beta_0 = 0.5$.

[16] Following the suggestion of an anonymous referee, we have investigated whether the SW model exhibits similar properties as the ACEL model if the number of shocks is reduced to the investment-specific and the neutral technology-shock as in ACEL. We find that the response coefficients on inflation and the output gap in the two-parameter and three-parameter rules increase in absolute terms. However, the three parameter rule in the

Next, we turn to the rules with four parameters that include the lagged output gap in addition to current output, inflation and the lagged interest rate. The coefficients on the lagged interest rate typically remain near unity. Interestingly, the coefficient on the lagged output gap, that is β_1, in the CEE/ACEL and SW models is almost equal to $-\beta_0$, the coefficient on the current output gap. Thus, the CEE/ACEL and SW models appear to desire a policy response to the growth rate of the output gap rather than its level. In fact, restricting $\beta_1=-\beta_0$ and re-optimizing the response coefficients in these models implies a coefficient of 1.65 in the SW and 2.0 in the ACEL model, respectively. Changes in the other response coefficients are very limited. By contrast, in the Taylor model, which uses trend output as a measure of potential, the optimal coefficients on current and lagged output gaps are both positive.

Different findings between the Taylor model and the SW and CEE/ACEL models may be due to different definitions of potential output. The flex-price output level used as a measure of potential in the SW and CEE/ACEL models exhibits substantial variation due to economic shocks and its growth rate may deviate substantially from trend growth.[17] Thus, simply differencing the output gap in our policy rule does not eliminate the effect of different concepts of potential output on the optimized response coefficients. Instead, we proceed to evaluate the performance of a fourth class of rules that respond to the deviation of actual GDP growth from trend (or steady-state) growth, denoted by Δy_t :

(7) $$i_t = \rho i_{t-1} + \alpha \pi_t + \beta_\Delta \Delta y_t$$

In this manner, potential output growth is defined similarly across the three models. Researchers such as Orphanides (2003) have recommended such rules as a way to reduce the impact of central bank misperceptions about the level of potential output on interest rate

SW model does not take the extreme coefficient values observed in the ACEL model, nor do we observe multiple local optima as in the ACEL model. We make these findings available along with other material in an additional appendix that is available online.

[17] A number of recent contributions have emphasized the differences between flex-price measures of potential and more traditional views on the trending components of real activity (see Palmqvist (2007), Basu and Fernald (2009) and Gupta (2009)).

setting.[18] The last 3 rows in **Table 1** report the optimal coefficients of the 3-parameter policy rule with deviations of actual from steady-state output growth.

If output gap variability does not appear in the loss function, (λ_y=0), the optimal coefficient on output growth, $\beta_{\Delta y}$, is very close to zero, just as in the 3-parameter rules with the output gap, If output variability receives a weight of unity in the loss function, the optimal interest rate rule responds positively to output growth, at least in the Taylor and SW models. In the ACEL models it is near zero. Thus, in the SW and ACEL models, it matters quite a lot whether the rule uses the deviation of actual GDP growth from trend growth or from flexible-price output growth.

Table 2 reports on the relative stabilization performance with two-, three- and four-parameter rules. Two different measures are reported, the percentage increase in loss and, in parentheses, the absolute increase in loss when one reduces the number of parameters (and therefore variables) in the policy rule starting from the case of four-parameter rules. In the following, we will focus on the absolute loss differences because the percentage differences tend to give misleading signals.

The particular measure of the increase in absolute loss that is shown is the implied inflation variability premium proposed by Kuester and Wieland (2010) (referred to as the IIP in the following). This measure translates a particular increase in absolute loss into the increase in the standard deviation of inflation (in percentage point terms) that would raise the loss to the same extent keeping all else equal (i.e. for a constant output or interest volatility). The advantage of this measure is that it is easily interpreted in practical terms and therefore provides a clear signal of those properties of interest rate rules that are of economic importance.

[18] See also Schmitt-Grohe and Uribe (2004) and Burriel, Fernández-Villaverde and Rubio-Ramírez (2009). Beck and Wieland (2008) have instead proposed a non-linear cross-checking mechanism that would correct the prescriptions from an output gap-based rule whenever there is statistical evidence of distorted policy outcomes, but take advantage of gap estimates in normal times.

To give an example, consider the number in the fourth row and third column of **Table 2** in parentheses. Its value is 2.14 and it implies the following: if the Taylor model represents the U.S. economy and the central bank considers using the optimized two-parameter rule instead of an optimized three-parameter rule, and if the central bank's loss-function assigns equal weight to output and inflation, the resulting increase in loss (due to higher inflation, output and interest volatility) is equivalent to an increase in the standard deviation of inflation of 2.14 percentage points all else equal. This difference is economically important. Although, it is the largest IIP reported in the table the associated percentage increase of 98.8% is only the fourth-largest in the table. The third-largest percentage increase in the table is 229%. It is associated with a switch from the three-parameter to the two-parameter rule in the ACEL model when the central bank's loss function assigns zero weight to output volatility. However, the associated IIP of 0.04 is tiny. Thus, the particular switch in rule is economically irrelevant in spite of the large percentage increase in loss. In this case, the reason is that the ACEL model only contains two shocks that cause little inflation volatility and very small losses.

The findings in **Table 2** suggest that there is little additional benefit from including the lagged output gap in the rule. Dropping the lagged output gap from the rule barely increases the central bank's loss. The associated IIP's in the first column of Table 2 lie between 0.001 and 0.47. However, it appears very beneficial to include the lagged interest rate in the rule. Dropping the lagged interest rate from the rule and moving from three to two response parameters implies an economically significant increase in the central bank's loss function, in particular in the SW and Taylor models, where it is equivalent to an increase in the standard deviation of inflation by 1 and 2 percentage points, respectively, (3rd column in Table 2). Among three-parameter rules, the rule with the output gap performs better than the rule with growth rate of output (in deviation from trend growth) across all three models. As shown in the middle column of Table 2 the IIP's relative to the four-parameter rule are uniformly

greater for the growth rate than the gap version. They are particularly large in the Taylor model. However, the growth-rate version of the three-parameter rule still performs better than the 2-parameter rule with inflation and the output gap.

6. Robustness

What if the model used by the central bank in designing a policy rule is not a good representation of the economy and one of the other two models provides a much better representation of the U.S. economy? In other words, how robust are model-specific optimized policy rules with respect to the range of model uncertainty reflected in the three models considered in this paper? **Table 3** provides answers to these questions. Robustness is measured in the following manner. The rule optimized for model X is implemented in model Y. The resulting loss in model Y is compared to the loss that would be realized under the rule with the same number of parameters that has been optimized for that particular model. The difference is expressed in terms of IIP only.

The findings in **Table 3** show that from the perspective of a central bank that aims to minimize inflation and interest rate volatility but assigns no weight to output volatility ($\lambda_y=0$), all four classes of policy rules are quite robust. Typically, a rule optimized in one of the models performs quite well in any of the other model compared to the best possible rule with the same number of parameters in that model.

Unfortunately, the preceding conclusion is almost completely reversed when one takes the perspective of a policy maker who cares equally about output and inflation volatility, that is when $\lambda_y=1$. In this case, only the 2-parameter rules remain fairly robust. The lack of robustness is most pronounced for 3- and 4-parameter rules that use output gaps. While these rules offer substantial performance improvements when the true model is known, performance can deteriorate markedly if the economy is better approximated by another

model. For example, using the 4-parameter rule that is optimal in the SW model instead in the Taylor model, implies an IIP of 2.71. Alternatively, the 4-parameter rule optimized for the Taylor model implies an IIP of 7.18 in the SW model and generates multiple equilibria in the ACEL model.

Similar problems arise with regard to 3-parameter rules that use the output gap, even if the CEE/ACEL model is excluded from the robustness analysis because of its odd behaviour under such rules as discussed earlier. As shown in the second column of **Table 3**, the rule optimized in the Taylor model implies an IIP of 5.41 in the SW model, while the rule optimized for the SW model delivers an IIP of 3.20 in the Taylor model. Replacing the output gap in the 3-parameter rules with the deviation of output growth from its trend improves their robustness properties at the cost of substantial performance deterioration in the true model as shown previously in **Table 2**. However, the IIP's are not negligible and remain near or above unity in three cases, two of which concern the rule optimized in the ACEL model.

Only the rules with two parameters that respond to inflation and the current output gap deliver a fairly robust stabilization performance across the three models. The IIP's are always substantially below unity and often near zero. Thus, a policymaker with a strong preference for robustness against model uncertainty might prefer to choose an optimized two-parameter rule that responds to inflation and the output gap but not the lagged interest rate.

Unfortunately, such rules perform quite a bit worse than rules with interest-rate smoothing when it is known which of the models best captures the true dynamics in the economy. To quantify this loss, we re-evaluate robustness with respect to the best 4-parameter rule when the model is known, rather than the best rule of the same class. With respect to this benchmark 2-parameter rules exhibit IIP's of 2.64 (SW rule in Taylor model) and 1.53 (Taylor rule in SW model), respectively. Thus, they remain more robust than 3- and 4-parameter rules with output gaps. However, 3-parameter rules that replace the output gap with the deviation of actual GDP growth from trend perform slightly better from this perspective as

long as the ACEL model is excluded from the comparison. They exhibit IIP's of 2.28 (SW rule in Taylor model) and 1.21 (Taylor rule in SW model), respectively, when compared to the 4-parameter rule optimized for the correct model.

Using the model database, however, it is possible to produce policy recommendations that are more robust than those based on a single model. For example, one may optimize a particular policy rule with respect to multiple models by minimizing the average loss across models. This approach has been proposed by Levin, Wieland and Williams (2003) and Brock, Durlauf and West (2003), among others. In this case, the response coefficients of the rules, (ρ, α, β_0, β_1, β_Δ), are chosen to minimize the average loss across the three models:

$$\sum_{m \in \mathrm{M}} \frac{1}{3} L_m = \sum_{m=\mathrm{M}} \frac{1}{3} (Var(\pi_m) + \lambda_y Var(y_m) + \lambda_{\Delta i} Var(\Delta i_m)) \tag{8}$$

Here, the subscript *m* refers to a particular element of *M=(TAYLOR, SW, ACEL)* – the set of available models. We focus on the performance of such rules in those cases where model-specific rules were not robust, that is when the central bank assigns similar weights to output and inflation in the loss function. The parameter values for the model averaging rules are reported in **Table 4**. The 2-parameter rules remain fairly similar to the model-specific optimization because those were already quite robust. The interest-smoothing coefficient for 3- and 4-parameter rules now lies very close to unity, in between the values that are optimal in the SW and the TAYLOR model. The response to inflation is small but positive ranging from 0.2 to 0.4 depending on whether the rules include current and lagged output gaps or the deviation of output growth from trend. Response coefficients on the current output gap, output gap growth or output growth deviations from trend vary between 0.2 and 0.8.

As shown in **Table 5,** model averaging generally improves the robustness of all four classes of simple policy rules that we have evaluated. Again, the numerical values reported in different cells of the table refer to the increase in the loss function – expressed in terms of inflation variability premia (IIP) – when a rule optimized in model X is used in model Y and

evaluated relative to the same type of rule optimized in model Y. By this measure 2-parameter rules that respond to inflation and the output gap are the rules that are most robust to model uncertainty. The robustness properties of rules with interest rate smoothing that respond to inflation and output growth deviations from trend are slightly worse. However, this ordering can be reversed if the 4-parameter rule optimized in the correct model is used as benchmark (IIP values in parentheses) and the ACEL model is dropped from the comparison. More importantly, model averaging helps to identify rules with interest rate smoothing and a response to output gaps that are fairly robust to model uncertainty, while regaining much of the improvement in stabilization performance promised by such rules in the absence of model uncertainty.

We note that model averaging mirrors Bayesian decision-making with equal prior beliefs. Kuester and Wieland (2010) compare Bayesian decision-making with worst-case analysis and ambiguity aversion, which combines both objectives, in an application that deals with monetary policy modelling in the euro area. They also explore the impact of learning on posteriors and Bayesian objectives over time.

7. Conclusions and Extensions

The preceding comparison of the Taylor (1993a) model with the two well-known examples from the current generation of new Keynesian models of the U.S. economy by Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007) indicates a surprising similarity of the monetary transmission mechanism. The empirical, model-based assessment of the impact of an unanticipated change in the federal funds rate on real U.S. GDP has not changed in 14 years that lie in between the publication of these models. This finding is encouraging for policy makers that want to rely on such models. It differs from earlier comparison projects

which showed that models built later in the 1990s such as the FRB/US model suggested that the impact of policy shocks on real output was much more drawn out over time. Conventional wisdom on the lags of monetary policy decisions may therefore need to be revised.

The robustness analysis of simple policy rules with the three models reveals less similarity than the comparative assessment of the transmission mechanism. If the central bank has the task of stabilizing both output and inflation, then an optimal rule derived in one of our models is not robust in the other models. By sacrificing optimality in each model one can identify some policy rules that are fairly robust, in particular, 2-parameter rules that respond to inflation and the output gap and 3-parameter rules that include interest rate smoothing but replace the output gap with the deviation of GDP growth from trend.

We also find that model averaging substantially improves the robustness properties of policy rules. Hence, using a model database, such as the one described in this paper, one cab derive policy rules that are more robust to model uncertainty than those obtained with a single preferred model.

Our findings also suggest at least two important extensions focusing on (1) the implications of utility-based loss functions and (2) a wider range of macroeconomic models.

Utility Based Loss Functions

We selected the loss function in equation (6) becuase it has been used extensively in the past and because it corresponds to the unconditional expectation of a second-order approximation of household utility in a small New-Keynesian model derived from microeconomic foundations. However, if the loss function is interpreted as a measure of utility, then its parameters (λ_y,$\lambda_{\Delta i}$) are model-dependent (as we noted previously) and the list of variables appearing in the loss function must be expanded. For example, if wage rigidities are present in addition to price rigidities, not only price but also wage fluctuations will affect household

utility. Onatski and Williams (2004) derive the following quadratic approximation of the unconditional expectation of household utility in the model of Smets and Wouters (2003):

(9) $$L_{OW2004} = E[\pi_t^2 + 0.21K_{t-1}^2 - 0.51\pi_t\pi_{t-1} + 0.24(w_t + \pi_t)(w_t - w_{t-1})]$$

Here w_t refers to the real wage and K_{t-1} to the lagged capital stock. To illustrate how such a loss function would affect our results, we optimized the four types of simple policy rules with respect to this utility-based loss in the SW model augmented with the variance of the change of the interest rate. and summarize them here. Interestingly, the optimized 2-, 3- and 4-parameter rules have fairly similar welfare implications under the Onatski-Williams approximation of household utility in the Smets-Wouters model with a maximum difference of 1.17 in IIP terms. We also evaluate the robustness of rules optimized with respect to the simpler loss function defined by equation (6) under this new loss function. Again, model-specific 2-parameter rules with inflation and the output gap, and 3-parameter rules with interest-rate smoothing, inflation and output growth deviations from trend remain fairly robust, but not the other model-specific rules. More details about these results are available in the online/web appendix.

Robustness to other macroeconomic models

While we have focused on three models of the U.S. economy, the new monetary model database offers the possibility of comparing many other empirically estimated models. With regard to future research, it would be of great interest to investigate the robustness of monetary policy rules in models that offer a more detailed treatment of the financial sector. As an illustration we extended our model comparison and robustness analysis to include the model of De Graeve (2008). De Graeve introduces a financial intermediary, capital goods producers and entrepreneurs as in Bernanke et al (1999) in a medium-size DSGE model of the same type as the CEE and SW models we have considered. His model, which he estimates with Bayesian methods, generates an endogenous external finance premium that is impacted

by a variety of economic shocks. Interestingly, we find that the GDP response to a monetary policy shock in the De Graeve (DG) model remains very close to the impulse responses in the Taylor, SW and CEE/ACEL models reported in **Figure 1**. The robustness of optimized model-specific rules, however, deteriorates further once we include the DG model. Especially 2-parameter rules optimized in the DG model perform badly in the Taylor and SW models. However, model-averaging rules remain very robust. In fact, they need not be changed. Including the DG model in the model-averaging loss function defined by equation (8) has only a marginal effect on the optimal response coefficients in the policy rules. More information about these results is available in the online/web appendix.

References

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani (2007). "Bayesian estimation of an open economy DSGE model with incomplete pass-through." *Journal of International Economics*, 72: 481–511.

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani (2008). "Evaluating An Estimated New Keynesian Small Open Economy Model." *Journal of Economic Dynamics and Control* 32(8): 2690-2721.

Altig, David, Lawrence, Christiano, Martin Eichenbaum, and Jesper Linde. (2004). "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." Working Paper Series, WP-05-01, Federal Reserve Bank of Chicago.

Basu, Susanto and John G. Fernald. (2009). "What do we know and not know about potential output?" Working Paper Series, WP-2009-05, Federal Reserve Bank of San Francisco.

Beck, Günter and Volker Wieland (2008), "Central Bank Misperceptions and the Role of Money in Interest Rate Rules", *Journal of Monetary Economics*, 55 (S1).

Bernanke, Ben, Mark Gertler and Simon Gilchrist, (1999). "The Financial Accelerator in a Quantitative Business Cycles Framework." In John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics Volume 1C*. Amsterdam: Elsevier Science, North-Holland.

Brayton, Flint and Thomas Laubach (2008). "Documentation of Linearized FRB/US." Unpublished manuscript, Board of Governors of the Federal Reserve System, Division of Research and Statistics.

Brock, William, Steven Durlauf, and Kenneth West, (2003), "Policy Evaluation in Uncertain Environments", Brookings Papers on Economic Activity, 1:2003, 235-302.

Bryant, Ralph, David Currie, Jacob Frenkel, Paul Masson, and Richard Portes, eds. (1989). *Macroeconomic Policies in an Interdependent World.* Washington, D.C.: The Brookings Institution.

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

Burriel, Pablo, Jesús Fernández-Villaverde, and Juan F. Rubio-Ramírez. (2009). "MEDEA: A DSGE Model for the Spanish Economy," FEDEA Working Papers, WP-2009-17, FEDEA.

Calvo, Guillermo (1983). "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12: 383–98.

Carabenciov, Ioan, Igor Ermolaev, Charles Freedman, Michel Juillard, Ondra Kamenik, Dmitry Korshunov, and Douglas Laxton (2008). "A Small Quarterly Projection Model of the US Economy." IMF Working Paper 08/278.

Christensen, Ian and Ali Dib (2008). "The Financial Accelerator in an Estimated New Keynesian model," Review of Economic Dynamics 11, pp. 155-178.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113, 1: 1–45.

Clarida, Richard, Jordi Gali and Mark Gertler (1999). "The Science of Monetary Policy: A New-Keynesian Perspective." *Journal of Economic Literature* 37 (December): 1661–1707.

Clarida, R., J. Galí, and M. Gertler (2002). "A Simple Framework for International Monetary Policy Analysis." *Journal of Monetary Economics* 49: 879-904.

Coenen, Günter and Volker Wieland (2002). "Inflation Dynamics and International Linkages: A Model of the United States, the Euro Area and Japan." ECB Working Paper Series 181, 2002.

Coenen, Günter and Volker Wieland (2005). "A Small Estimated Euro Area Model with Rational Expectations and Nominal Rigidities." *European Economic Review*, 49: 1081–1104.

Coenen, Günter, Andrew Levin, and Volker Wieland (2005). "Data Uncertainty and the Role of Money as an Information Variable for Monetary Policy." *European Economic Review*, 49: 975–1006.

Coenen, Günter, Peter Mcadam, and Roland Straub (2008). "Tax Reform and Labour-Market Performance in the Euro Area: A Simulation-Based Analysis Using the New Area-Wide Model." *Journal of economic dynamics and control*, Vol. 32, No. 8: 2543-2583.

Collard, Fabrice and Michael Juillard (2001). "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models." *Journal of Economic Dynamics and Control* 25: 979–999.

De Graeve, Ferre (2008). "The External Finance Premium and the Macroeconomy: US Post-WWII Evidence", *Journal of Economic Dynamics and Control* 32: 3415-3440.

Dieppe, Alistair, Keith Kuester, and Peter McAdam (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.

Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte (2007). "Documentation of the Research and Statistics Division's Estimated DSGE Model of the U.S. Economy: 2006 Version." Finance and Economics Discussion Series Working Paper 07-53, Board of Governors of the Federal Reserve System.

Erceg, Christopher J., and Andrew Levin (2006). "Optimal Monetary Policy with Durable Consumption Goods," *Journal of Monetary Economics*, 53(7): 1341-1359.

Erceg, Christopher J., Luca Guerrieri, and Christopher Gust (2008). "Trade Adjustment and the Composition of Trade." *Journal of Economic Dynamics & Control*, 32: 2622–2650.

Fuhrer, Jeffrey and George Moore (1995). "Inflation Persistence." *Quarterly Journal of Economics* 110, 1: 127–159.

Galí, Jordi, and Tommaso Monacelli (2005). "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies* 72: 707-734.

Geweke, John (1999). "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication." *Econometric Reviews* 18: 1-126.

Gupta, Abhishek (2009), "A Forecasting Metric for Evaluating DSGE Models." Mimeo, Johns Hopkins University.

Iacoviello, Matteo (2005). "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle." *The American Economic Review* 95: 739-764.

Ireland, Peter N. (2004). "Money's Role in the Monetary Business Cycle." *Journal of Money, Credit and Banking*, Vol. 36: 969-983.

Juillard, Michel. (1996), "Dynare: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables Through the Use of a Relaxation Algorithm." CEPREMAP working paper, 9602.

Klein, Lawrence, ed. (1991). *Comparative Performance of U.S. Econometric Models.* Oxford, Eng.: Oxford University Press.

Kuester, Keith and Volker Wieland (2010), "Insurance Policies for Monetary Policy in the Euro Area." *Journal of the European Economic Association* 8, 4: 872-912.

Laxton, Douglas and Paolo Pesenti (2003). "Monetary Rules for Small, Open, Emerging Economies." *Journal of Monetary Economics*, 50:1109–1146.

Levin, Andrew, Volker Wieland, and John C. Williams. (1999), "Robustness of Simple Monetary Policy Rules under Model Uncertainty," in John B. Taylor, ed., *Monetary Policy Rules*. Chicago: University of Chicago Press, 1999.

Levin, Andrew, Volker Wieland, and John. C. Williams. (2003), "The Performance of Forecast-Based Monetary Policy Rules under Model Uncertainty." *American Economic Review* 93, 3: 622–645.

Levin, Andrew, Alexei Onatski, John C. Williams, and Noah M. Williams. (2006), "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," *NBER Macroeconomics Annual 2005, Volume 20.* NBER, Cambridge, MA.

Lucas, Robert (1976). "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

McCallum, Bennett, and Edward Nelson. (1999), "Performance of Operational Policy Rules in an Estimated Semi-Classical Structural Model," in John B. Taylor, ed., *Monetary Policy Rules*. Chicago: University of Chicago Press, 1999.

Medina, Juan P. and Claudio Soto (2007). "The Chilean Business Cycles through the Lens of a Stochastic General Equilibrium Model." Working Paper 457, Central Bank of Chile.

Onatski, Alexei and Noah M. Williams. (2004), "Empirical and Policy Performance of a Forward-Looking Monetary Model", Working Paper, University of Wisconsin.

Orphanides, Athanasios (2003a). "Historical Monetary Policy Analysis and the Taylor Rule", *Journal of Monetary Economics*, 50, 5: 983-1022.

Orphanides, Athanasios (2003b). "The Quest for Prosperity without Inflation." *Journal of Monetary Economics*, 50: 633–663.

Orphanides, Athanasios and Volker Wieland (1998). "Price Stability and Monetary Policy Effectiveness when Nominal Interest Rates are Bounded at Zero." Finance and Economics Discussion Series 98-35, Board of Governors of the Federal Reserve System.

Ratto, Marco, Werner Roeger, and Jan in ’t Veld (2009). "QUEST III: An Estimated Open-Economy DSGE Model of the Euro Area with Fiscal and Monetary Policy." *Economic Modelling* 26(1): 222-233.

Reifschneider, David, Tetlow, Robert and John C. Williams. (1999), "Aggregate Disturbances, Monetary Policy, and the Macroeconomy: The FRB/US Perspective." *Federal Reserve Bulletin,* 85, 1: 1–19.

Rotemberg, Julio J. and Michael Woodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual* 12: 297-346.

Rotemberg, Julio and Michael Woodford. (1999), "Interest-Rate Rules in an Estimated Sticky Price Model," in John B. Taylor, ed., *Monetary Policy Rules*. Chicago: University of Chicago Press, 1999.

Rudebusch, Glenn D. and Lars E. O. Svensson (1999). "Policy Rules for Inflation Targeting." In J. B. Taylor, editor, *Monetary Policy Rules*. Chicago: University of Chicago Press.

Schmitt-Grohe, Stephanie and Martin Uribe. (2004), "Optimal Operational Monetary Policy in the Christiano-Eichenbaum-Evans Model of the U.S. Business Cycle," NBER Working Papers 10724, National Bureau of Economic Research, Inc.

Schorfheide, Frank (2000). "Loss Function Based Evaluation of DSGE Models." *Journal of Applied Econometrics* 15, 6: 645-670.

Smets, Frank and Raf Wouters. (2003), "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1, 5: 1123–1175.

Smets, Frank and Raf Wouters. (2007). "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97, 3: 506–606.

Shiller, Robert J. (1991). "Comment." In *Comparative Performance of U.S. Econometric Models,* edited by Lawrence Klein. Oxford, Eng.: Oxford University Press.

Taylor, John, B. (1993a). *Macroeconomic Policy in a World Economy*. New York: Norton, 1993.

Taylor, John, B. (1993b). "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy,* 39: 195-214.

Taylor, John, B. (1999). ed. *Monetary Policy Rules*. Chicago: University of Chicago Press, 1999.

Taylor, John, B. (1980). "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88: 1–24.

Wieland, Volker, Tobias Cwik, Gernot Mueller, Sebastian Schmidt and Maik Wolters (2009). "A New Comparative Approach to Macroeconomic Modelling and Policy Analysis" Manuscript, Center for Financial Studies, Frankfurt.

Woodford, Michael (1999), “Optimal Monetary Policy Inertia”, Manchester School, Supplement, 67:1-35.

Table 1
Optimal Simple Policy Rules[1)]

Rules: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1} + \beta_\Delta \Delta y_t$

Rule /Model	**Loss (λ_y = 0):** $Var(\pi) + Var(\Delta i)$					**Loss (λ_y=1):** $Var(\pi) + Var(y) + Var(\Delta i)$				
	ρ	A	β_0	β_1	β_Δ	ρ	α	β_0	β_1	β_Δ
	2 Parameters (Gap)[2)]					$\alpha\pi_t + \beta_0 y_t$				
TAYLOR		2.54	0.19				3.00	0.52		
SW		2.33	-0.10				2.04	0.26		
CEE/ACEL		4.45	0.28				2.57	0.45		
	3 Parameters (Gap)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$				
TAYLOR	0.98	0.37	0.09			0.98	0.21	0.53		
SW	1.06	0.49	0.01			1.13	0.012	0.015		
CEE/ACEL	0.97	0.99	0.02			2.84	7.85	-2.12		
	4 Parameters (Gaps)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$				
TAYLOR	0.98	0.37	0.07	0.02		0.96	0.18	0.41	0.19	
SW	1.06	0.46	-0.03	0.03		1.07	0.16	1.63	-1.62	
CEE/ACEL	1.01	1.11	0.18	-0.18		1.04	0.51	2.24	-2.30	
	3 Parameters (Growth)[3)]					$\rho i_{t-1} + \alpha\pi_t + \beta_\Delta \Delta y_t$				
TAYLOR	1.01	0.52			0.07	1.13	0.40			0.68
SW	1.03	0.48			-0.01	1.01	0.20			1.04
CEE/ACEL	1.02	1.07			-.002	0	3.71			.002

Notes:

[1)] The loss function includes the variance of inflation and the variance of the first-difference of nominal interest rates with a weight of unity, $\lambda_{\Delta i}$=1. λ_y denotes the weight on the variance of the output gap.

[2)] In the Taylor model the output gap denotes the difference between actual and trend output. In the SW and ACEL models it is the difference to the level realized under flexible prices given current macroeconomic shocks.

[3)] The output growth measure Δy_t is defined relative to steady-state/trend output growth in all three models.

Table 2
Increase in Loss when Reducing the Number of Parameters in the Rule
Percentage Increase (Increase in IIP)[1)]

Models	**Loss (λ_y = 0):** $Var(\pi) + Var(\Delta i)$		
	4 versus 3 Parameters (Gaps)	4 Parameters (Gaps) vs 3 Par. (Growth)	3 versus 2 Parameters (Gaps)
TAYLOR	0.12% (0.001)	13.5% (0.10)	278% (1.38)
SW	0.22% (0.001)	1.40% (0.01)	316% (0.78)
CEE/ACEL	5.10% (0.001)	10.0% (0.003)	229% (0.04)
	Loss(λ_y = 1): $Var(\pi) + Var(y) + Var(\Delta i)$		
TAYLOR	1.81% (0.07)	67.1% (1.61)	98.8% (2.14)
SW	10.6% (0.47)	18.1% (0.76)	25.6% (1.17)
CEE/ACEL	14.4% (0.11)	36.7% (0.22)	9.67% (0.11)

Notes:
[1)]The values in parentheses measure the increase in absolute loss in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

Table 3
Robustness of Policy Rules
Increase in IIP[1)] when a rule optimized in model X is used in model Y
and evaluated relative to the same type of rule optimized in model Y

Loss (λ_y=0): $Var(\pi)+Var(\Delta i)$				
IIP if evaluated in	Rules[3)] optimized in TAYLOR Model			
Model:	2 Parameters	3 Par. (Gap)	3 Par. (Growth)	4 Par. (Gaps)
SW	0.37	0.83	0.01	0.90
ACEL	0.03	0.12	0.01	0.14
	Rules optimized in SW Model			
TAYLOR	0.27	0.13	0.03	0.15
ACEL	0.15	0.02	0.01	0.02
	Rules optimized in ACEL Model			
SW	0.54	0.11	0.10	0.09
TAYLOR	0.76	0.27	0.25	0.34
Loss(λ_y=1): $Var(\pi)+Var(y)+Var(\Delta i)$				
IIP if evaluated in:	Rules optimized in TAYLOR Model			
Model:	2 Parameters	3 Par. (Gap)	3 Par. (Growth)	4 Par. (Gaps)
SW	0.17	5.41	0.66	7.18
ACEL	0.001	M.E.[3)]	0.31	M.E.[3)]
	Rules optimized in SW Model			
TAYLOR	0.86	3.20	1.05	2.71
ACEL	0.03	0.21	0.44	0.13
	Rules optimized in ACEL Model			
SW	0.07	108	1.69	0.53
TAYLOR	0.12	24.9	1.40	3.85

Notes:

[1)] The values in this table concern the increase in absolute loss in model Y under a rule optimized for model X relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

[2)] Rules: 2 Parameters: $i_t = \alpha\pi_t + \beta_0 y_t$, 3 Parameters (Gap): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$;

3 Parameters(Growth),: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_A \Delta y_t$; 4 Parameters (Gaps): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$.

[3)]**M.E.** refers to indeterminacy and the existence of multiple self-fulfilling equilibria.

Table 4
Optimized Model-Averaging Rules

Objective: Min $\sum_{m \in M} \frac{1}{3}(Var(\pi_m) + Var(y_m) + Var(\Delta i_m))$

Rules: $i_t = \rho i_{t-1} + \alpha \pi_t + \beta_0 y_t + \beta_1 y_{t-1} + \beta_\Delta \Delta y_t$

Set of equally-weighted models: $M = \{SW, TAYLOR, ACEL\}$	ρ	α	β_0	β_1	β_Δ
2 Parameters Rule (Gap)		2.75	0.52		
3 Parameters Rule (Gap)	1.05	0.41	0.23		
3 Parameters Rule (Growth)	1.09	0.20			0.76
4 Parameters rule (Gap)	1.06	0.19	0.67	-0.59	

Table 5
Robustness of Model-Averaging Policy Rules
Increase in IIP[1)] when a rule optimized in model X is used in model Y
and evaluated relative to the same type of rule optimized in model Y

Loss(λ_y=1): $Var(\pi) + Var(y) + Var(\Delta i)$				
IIP if evaluated in	2 Parameters	3 Par. (Gap)	3 Par. (Growth)	4 Par. (Gaps)
SW	0.11 (1.50) [2)]	1.02	0.13 (0.84) [2)]	0.47
TAYLOR	0.03 (2.18) [2)]	0.56	0.19 (1.71) [2)]	1.28
ACEL	0.00 (0.17) [2)]	0.27	0.40 (0.44) [2)]	0.12

Notes:
[1)] The values in this table concern the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.
[2)] The values in parenthesis refer to the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a 4-parameter rule optimized in model Y.

Figure 1
The Effect of a Policy Shock on Interest Rates, Output and Inflation
1 Percentage Point Increase in the Nominal Policy Rate

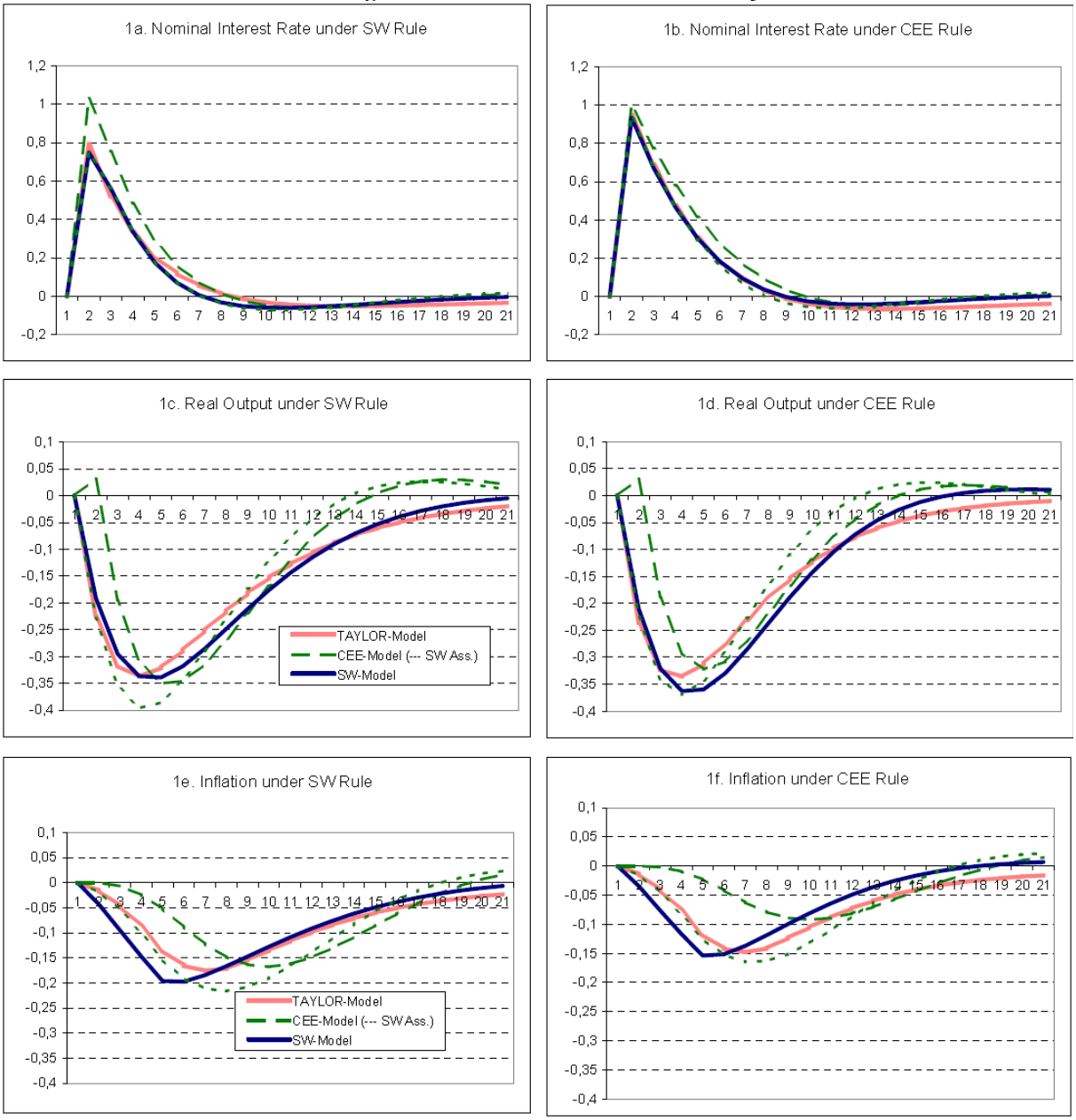


Figure 2
Common Aspects of the Transmission Mechanism in the Three Models (SW Rule)

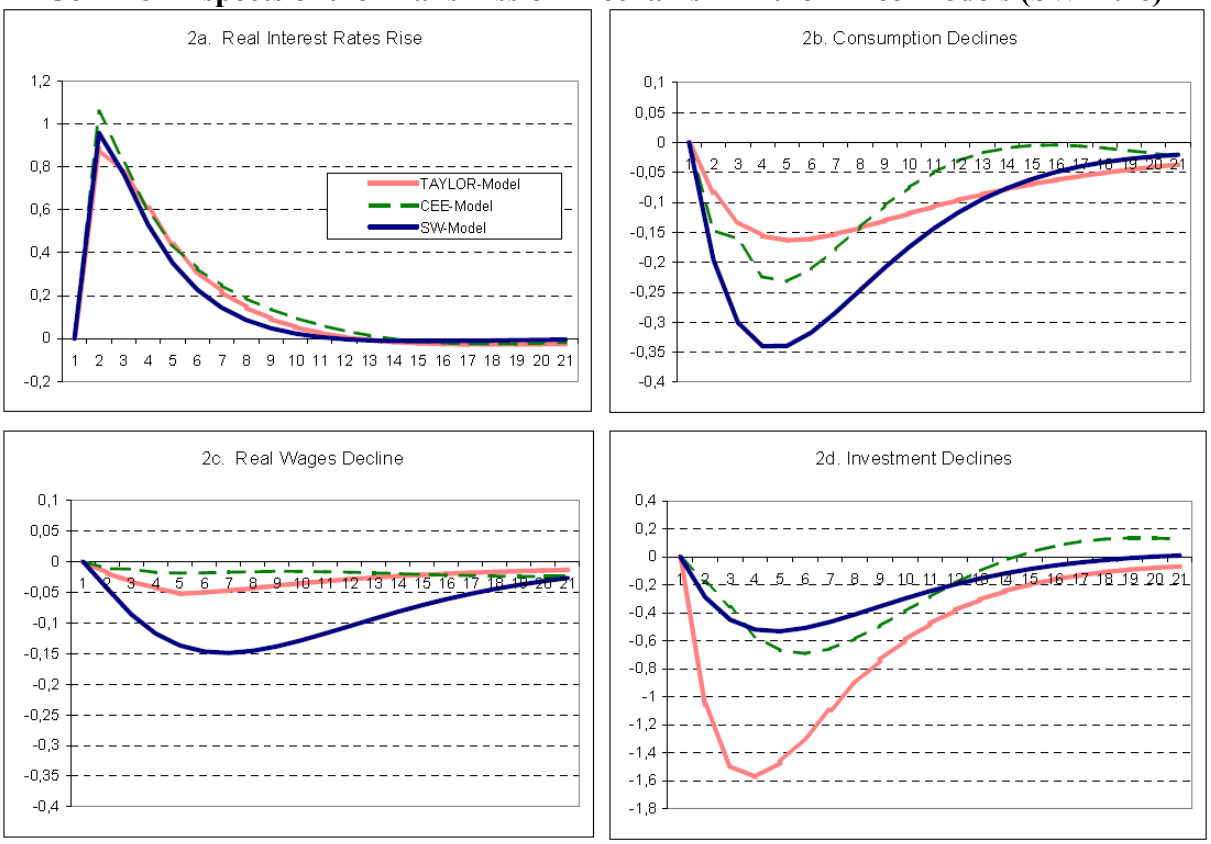
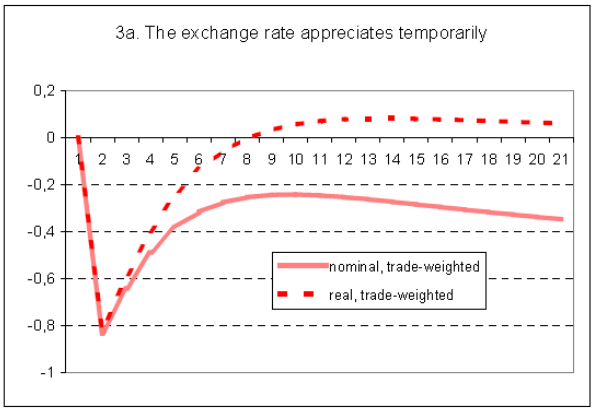
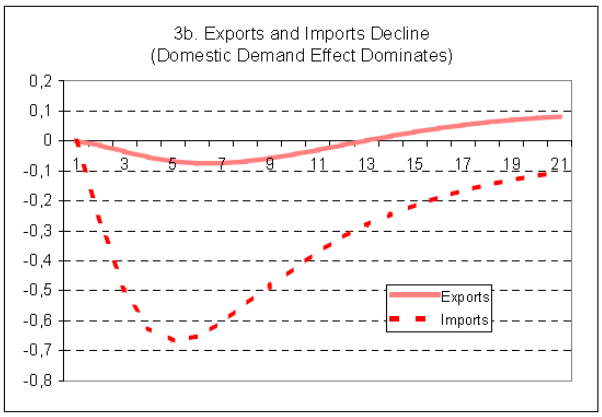


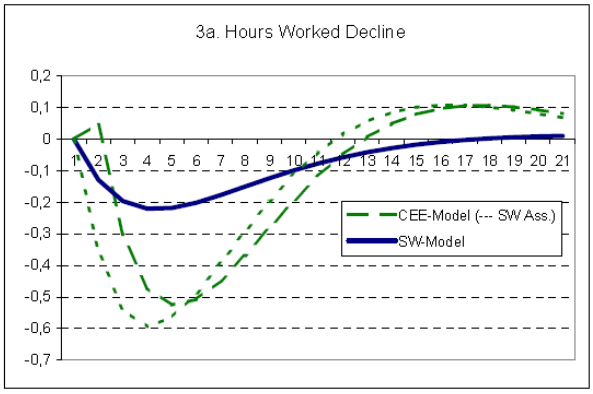
Figure 3
Differences in the Transmission Mechanism in the Three Models (SW Rule)

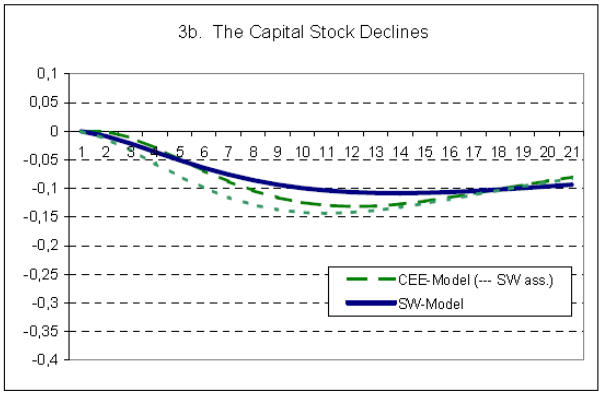
Only the TAYLOR Model Accounts for International Feedback

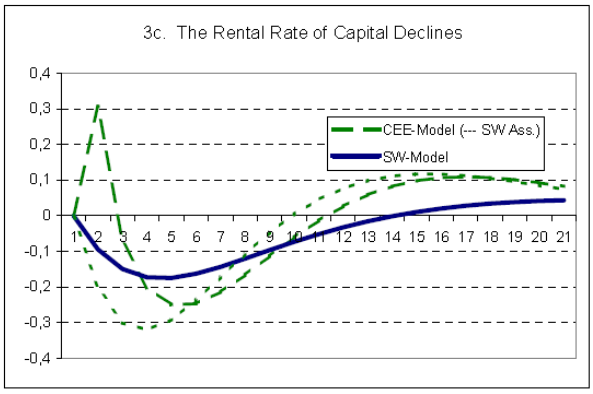




Only the SW and CEE Models Account for Labor Supply, Capital Stock and Capital Utilization







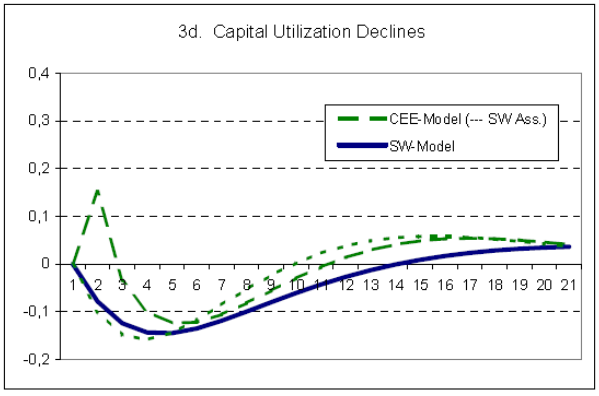
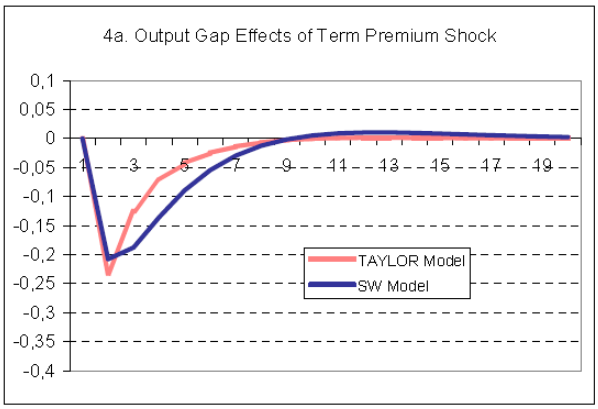


Figure 4
Term Premium Shock in the Taylor and SW Models (SW Rule)
1 Percentage Point Increase in the Term Premium



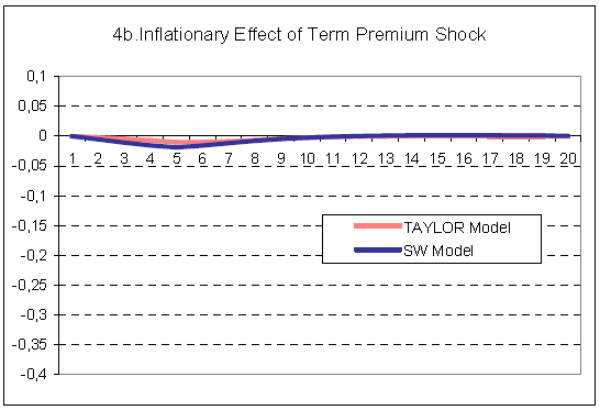


Figure 5
"Demand" Shocks in the Taylor, SW and CEE Models (SW Rule)
1 Percent Increase in the Relevant Variables

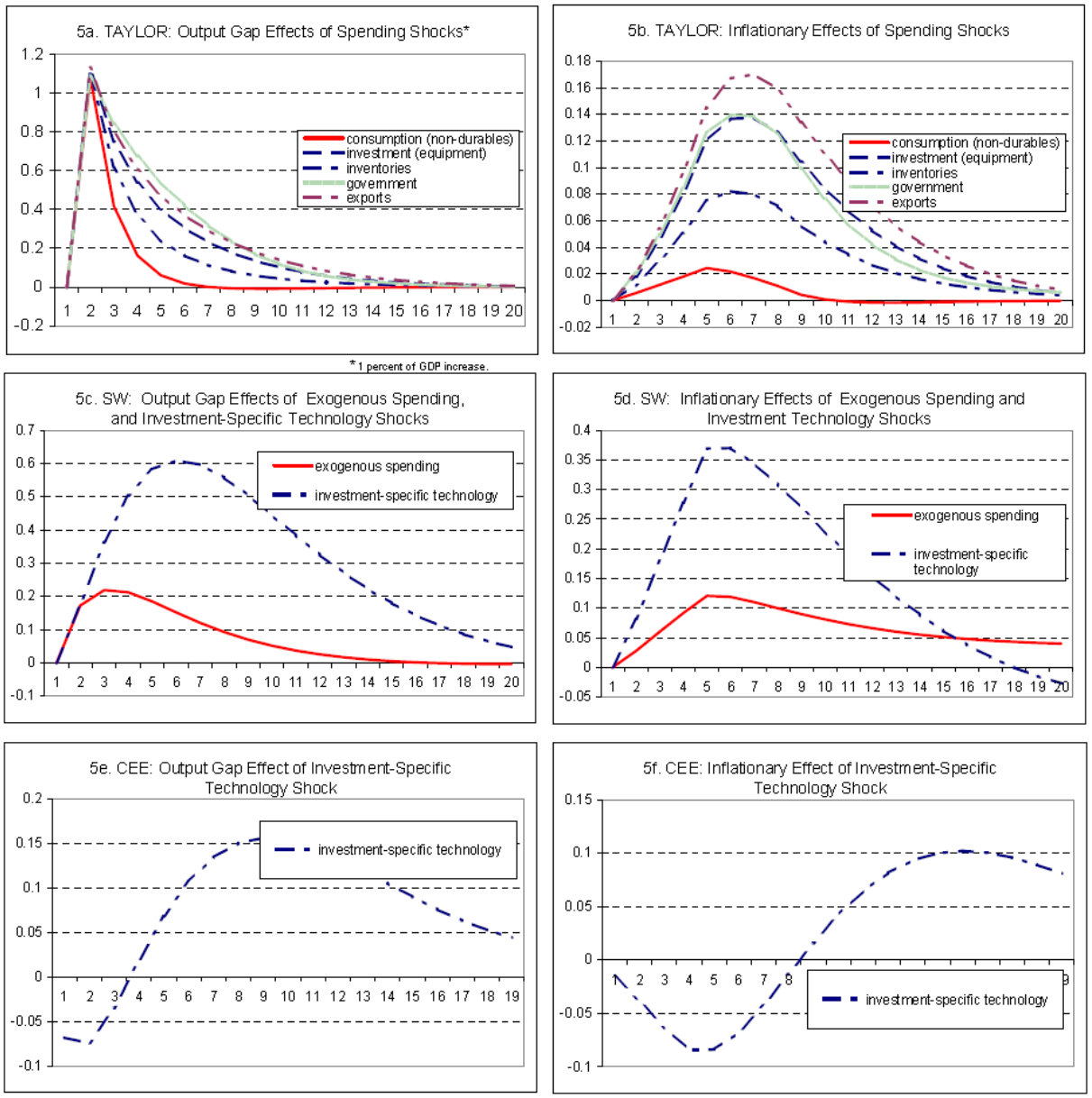
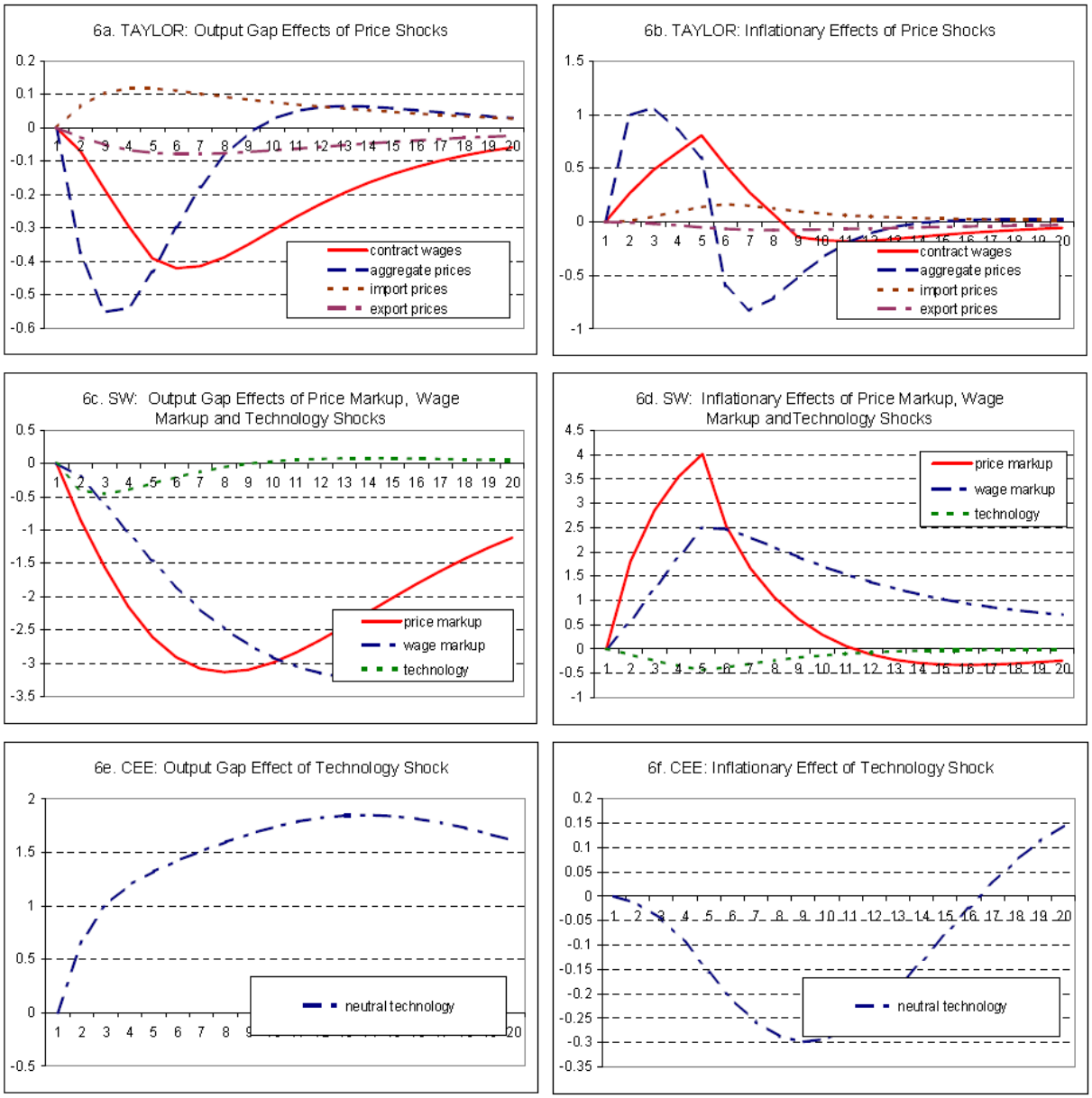


Figure 6
Short-Run and Long-Run "Supply" Shocks in Taylor, SW and CEE Models (SW Rule)
1 Percent Increase in the Relevant Variables



Appendix 35 Models Included in the Model Base as of August 2010[19]

1. Small Calibrated Models

1.1 Rotemberg, Woodford (1997)	NK_RW97
1.2 Levin, Wieland, Williams (2003)	NK_LWW03
1.3 Clarida, Gali, Gertler (1999)	NK_CGG99
1.4 Clarida, Gali, Gertler 2-Country (2002)	NK_CGG02
1.5 McCallum, Nelson (1999)	NK_MCN99
1.6 Ireland (2004)	NK_IR04
1.7 Bernanke, Gertler, Gilchrist (1999)	NK_BGG99
1.8 Gali, Monacelli (2005)	NK_GM05

2. Estimated US Models

2.1 Fuhrer, Moore (1995)	US_FM95
2.2 Orphanides, Wieland (1998)	US_OW98
2.3 FRB-US model linearized as in Levin, Wieland, Williams (2003)	US_FRB03
2.4 FRB-US model 08 linearized by Brayton and Laubach (2008)	US_FRB08
2.5 FRB-US model 08 mixed expectations, linearized by Laubach (2008)	US_FRB08mx
2.6 Smets, Wouters (2007)	US_SW07
2.7 CEE/ACEL Altig, Christiano, Eichenbaum, Linde (2004)	US_ACELm
(m=monetary policy shock, t=technology shock, sw=SW	US_ACELt
assumptions = no cost channel, no timing constraints)	US_ACELswm
	US_ACELswt
2.8 New Fed US Model by Edge, Kiley, Laforte (2007)	US_NFED08
2.9 Rudebusch, Svensson (1999)	US_RS99
2.10 Orphanides (2003b)	US_OR03
2.11 IMF projection model by Carabenciov et al. (2008)	US_PM08
2.12 De Graeve (2008)	US_DG08
2.13 Christensen, Dib (2008)	US_CD08
2.14 Iacoviello (2005)	US_IAC05

3. Estimated Euro Area Models

3.1 Coenen, Wieland (2005) (ta: Taylor-staggered contracts)	EA_CW05ta
3.2 Coenen, Wieland (2005) (fm: Fuhrer-Moore staggered contracts)	EA_CW05fm
3.3 ECB Area Wide model linearized as in Dieppe et al. (2005)	EA_AWM05
3.4 Smets, Wouters (2003)	EA_SW03
3.5. Euro Area Model of Sveriges Riksbank (Adolfson et al. 2007)	EA_SR07
3.6. Euro Area Model of the DG-ECFIN EU (Ratto et al. 2009)	EA_QUEST3
3.7. ECB New-Area Wide Model of Coenen, McAdam, Straub (2008)	EA_NAWM08

[19] See Wieland, Cwik, Müller, Schmidt and Wolters (2009) for a detailed exposition of the platform for model comparison. Software and models are available for download from http://www.macromodelbase.com.

4. Estimated Small Open-Economy Models (other countries)

4.1. RAMSES Model of Sveriges Riskbank, Adolfson et al.(2008b)	SW_ALLV08
4.2 Model of the Chilean economy by Medina, Soto (2007)	CL_MS07

5. Estimated/Calibrated Multi-Country Models

5.1 Taylor (1993a) model of G7 economies	G7_TAY93
5.2 Coenen,Wieland (2002, 2003) G3 economies	G3_CW03
5.3 IMF model of euro area & CZrep by Laxton, Pesenti (2003)	EACZ_GEM03
5.4 FRB-SIGMA model by Erceg, Gust, Guerrieri (2008)	G2_SIGMA08

Web-Appendix: Additional Sensitivity Studies

This appendix provides material on additional sensitivity studies conducted by the authors that are not included in the printed text of the paper. These findings are summarized at different points throughout the paper. They are organized in the appendix in the same order as they are mentioned in the text under the headings of the relevant sections of the main text.

A-1. Sensitivity studies for "3. Monetary policy shocks in three monetary models of the U.S. economy"

Figures A-1 to A-3 report common aspects and differences regarding the monetary transmission mechanism in the Taylor, SW and CEE/ACEL models under the policy rule used in Christiano, Eichenbaum and Evans (2005) It is referred to as the CEE rule and defined by equation (2) in the main text. The findings should be compared to Figures 2 and 3 in the main text that are obtained under the SW rule (equation 1).

Figure A-1
Common Aspects of the Transmission Mechanism in the Three Models (CEE Rule)

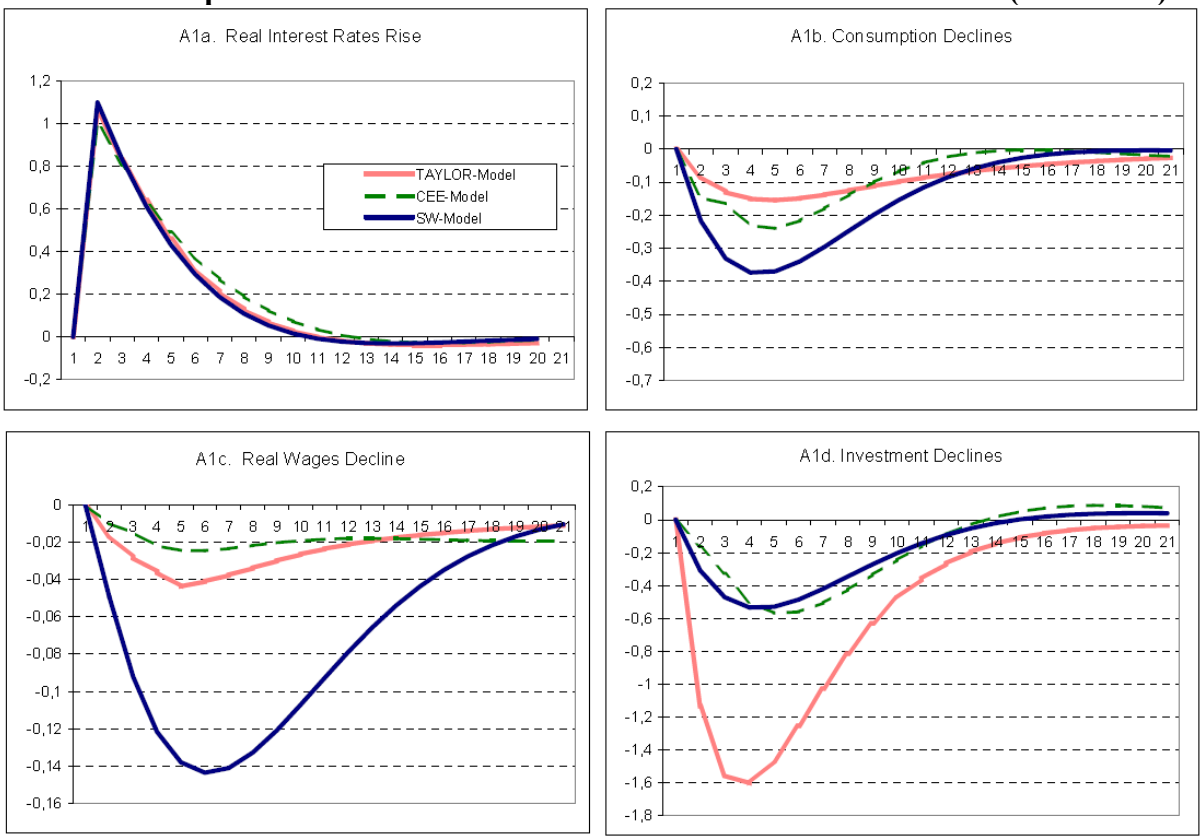


Figure A-2
Only the TAYLOR Model Accounts for International Feedback (CEE Rule)

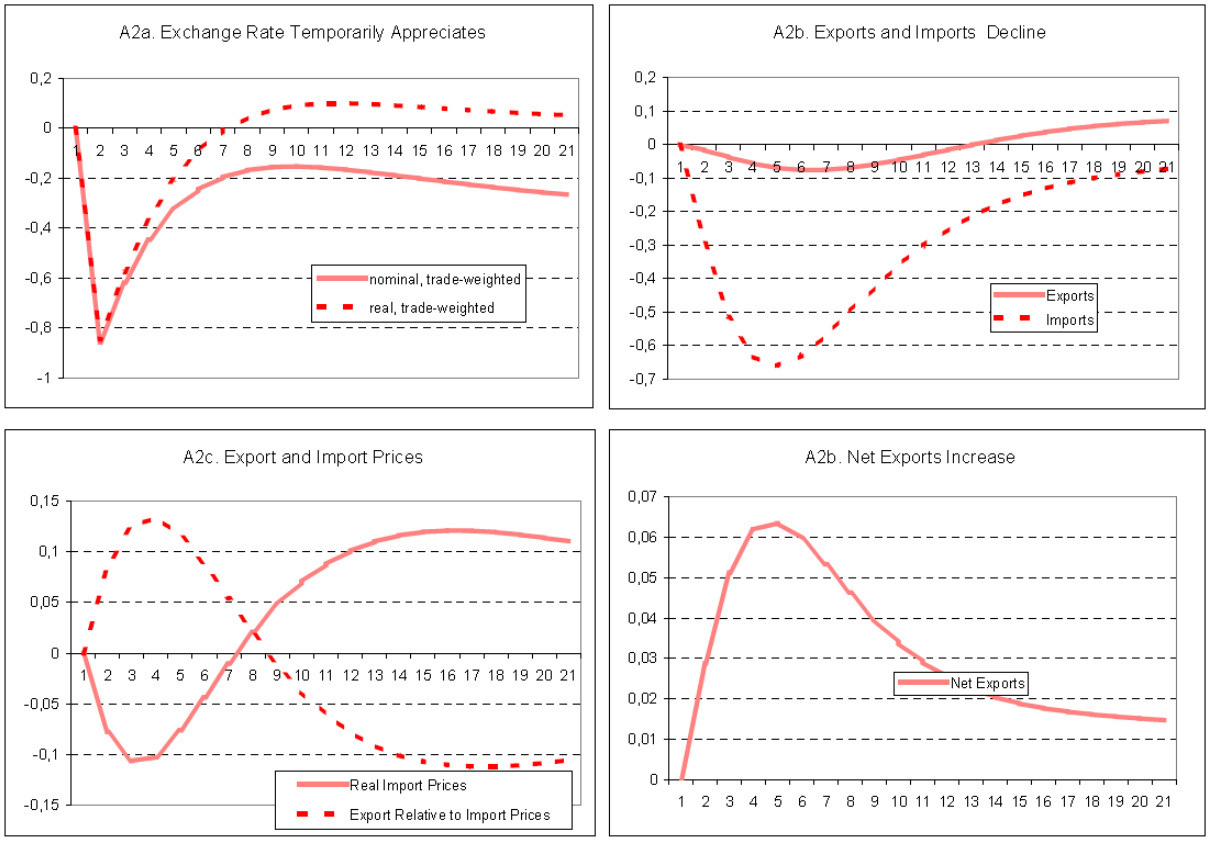
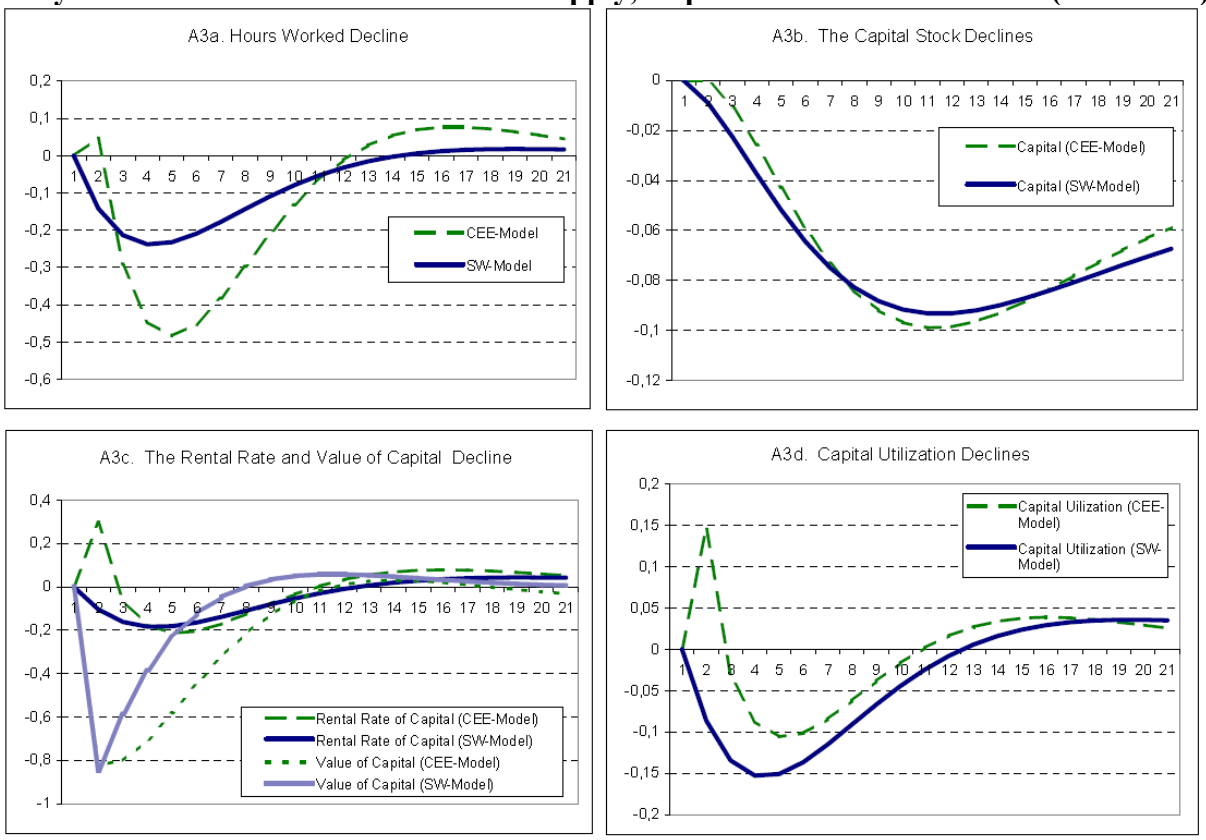


Figure A-3
Only SW and CEE Account for Labor Supply, Capital Stock and Utilization (CEE Rule)



A-2. Sensitivity studies for "5. Optimal simple policy rules in the Taylor, CEE/ACEL and SW models"

Tables A-1 and A-2 provide information on a sensitivity study with a smaller weight of $\lambda_{\Delta i}$=0.5 on the standard deviation of changes in the short-term nominal interest rate. Further sensitivity studies (not reported) were conducted with respect to a weight of 0.5 on the output gap, with respect to the definition of the output gap relative to steady-state output rather than flexible-price output (SW and CEE/ACEL model), and with respect to a version of the CEE/ACEL model with the SW assumptions of no cost-channel and no exogenous delay of the impact of policy. The conclusions discussed in the main text remained the same.

Table A-1
Optimized 2-, 3- and 4-Parameter Rules

$$i_t = \rho i_{t-1} + \alpha \pi_t + \beta_0 y_t + \beta_1 y_{t-1}$$

Rule /Model	**Loss[1] (λ_y = 0,$\lambda_{\Delta i}$=0.5):** $Var(\pi) + 0.5Var(\Delta i)$				**Loss (λ_y=1, $\lambda_{\Delta i}$=0.5):** $Var(\pi) + Var(y) + 0.5Var(\Delta i)$			
	ρ	α	β_0	β_1	ρ	α	β_0	β_1
	2 Parameters (Gap)[2]				$\alpha\pi_t + \beta_0 y_t$			
TAYLOR		3.00	0.22			3.46	0.76	
SW		2.81	-0.12			2.15	0.30	
CEE/ACEL		5.91	0.27			2.87	0.49	
	3 Parameters (Gap)				$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$			
TAYLOR	09.7	1.44	0.02		0.97	0.27	0.76	
SW	1.05	0.71	0.01		1.12	0.012	0.015	
CEE/ACEL	0.97	0.99	0.02		2.14	8.29	-1.99	
					(0.01)[3]	(2.90)[3]	(0.50)[3]	
	4 Parameters (Gaps)				$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$			
TAYLOR	0.98	0.51	0.09	0.02	0.95	0.24	0.60	0.26
SW	1.05	0.65	-0.04	0.05	1.07	0.21	2.22	-2.21
CEE/ACEL	1.01	1.86	0.15	-0.15	1.01	0.75	3.11	-3.18

Notes:
[1] The loss function includes the variance of inflation and the variance of the first-difference of nominal interest rates with a weight of unity, $\lambda_{\Delta i}$=1. λ_y denotes the weight on the variance of the output gap.
[2] In the Taylor model the output gap denotes the difference between actual and trend output. In the SW and ACEL models it is the difference to the level realized under flexible prices given current macroeconomic shocks.
[3] The ACEL model, which is has only two shocks, exhibits only very small values of the loss function and multiple local optima. For example, a local optimum with much smaller parameter values is displayed in parentheses.

Table A-2
Increase in Loss when Reducing the Number of Parameters in the Rule
Percentage Increase (Increase in IIP)[1)]

Models	**Loss (λ_y = 0 $\lambda_{\Delta i}$=0.5):** $Var(\pi) + 0.5Var(\Delta i)$	
	4 versus 3 Parameters (Gaps)	3 versus 2 Parameters (Gaps)
TAYLOR	0.10% (0.001)	210% (1.04)
SW	0.26% (0.001)	253% (0.62)
CEE/ACEL	3.51% (0.001)	217% (0.03)
	Loss(λ_y =1 $\lambda_{\Delta i}$=0.5): $Var(\pi) + Var(y) + 0.5Var(\Delta i)$	
TAYLOR	1.78% (0.06)	88.0% (1.81)
SW	12.4% (0.53)	23.7% (1.10)
CEE/ACEL	13.7% (0.10)	11.18% (0.12)

Notes:
[1)]The values in parentheses measure the increase in absolute loss in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

Because of the odd properties exhibited by the ACEL model in the policy optimization (specifically the existence of multiple local optima) we investigated whether a small number of shocks would induce similar properties in the policy optimization with the SW model. The ACEL model has a monetary policy shock (which is turned of in the optimization), and investment-specific technology shock and a neutral technology shock. The SW model also has an investment-specific and a neutral technology shock in addition to other shocks. Thus, we have shut down these other shocks in the SW model and re-optimized the policy rules in the SW model with only the two technology shocks. The findings are reported in Table A-3 in comparison to the other rules reported previously in Table 1 in the paper. The weight on interest volatility, $\lambda_{\Delta i}$, is set to unity as in the main text of the paper. The optimized rules in the SW model change, but they do not approach the extreme values found in the ACEL model, nor does the computation exhibit multiple local optima. Other features of the ACEL model must be at the root of these findings. Thus, we keep the ACEL in the optimization and robustness exercise.

Table A-3
Optimal Simple Policy Rules[1)]
with Shocks in SW turned off to compare to ACEL
Rules: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$

Rule /Model	**Loss (λ_y = 0):** $Var(\pi) + Var(\Delta i)$				**Loss (λ_y=1):** $Var(\pi) + Var(y) + Var(\Delta i)$			
	ρ	α	β_0	β_1	ρ	α	β_0	β_1
	2 Parameters (Gap)[2)]				$\alpha\pi_t + \beta_0 y_t$			
SW		2.33	-0.10			2.04	0.26	
CEE/ACEL		4.45	0.28			2.57	0.45	
SW/2 shocks		5.07	-0.29			5.12	1.08	
	3 Parameters (Gap)				$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$			
SW	1.06	0.49	0.01		1.13	0.012	0.015	
CEE/ACEL	0.97	0.99	0.02		2.84	7.85	-2.12	
SW/2 shocks	1.07	0.89	0.03		1.02	0.35	0.87	
	4 Parameters (Gaps)				$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$			
SW	1.06	0.46	-0.03	0.03	1.07	0.16	1.63	-1.62
CEE/ACEL	1.01	1.11	0.18	-0.18	1.04	0.51	2.24	-2.30
SW/2 shocks	1.08	0.96	0.11	-0.10	1.02	0.42	1.09	-0.29

Notes:
[1)] The loss function includes the variance of inflation and the variance of the first-difference of nominal interest rates with a weight of unity, $\lambda_{\Delta i}$=1. λ_y denotes the weight on the variance of the output gap.
[2)] In the Taylor model the output gap denotes the difference between actual and trend output. In the SW and ACEL models it is the difference to the level realized under flexible prices given current macroeconomic shocks.

A-3. Sensitivity studies for "6. Robustness"

Table A-3 provide information on a sensitivity study with a smaller weight of $\lambda_{\Delta i}$=0.5 on the standard deviation of changes in the short-term nominal interest rate. The findings should be compared to Table 3 in the main text of the paper. The conclusions discussed in the main text remain the same.

Table A-4
Robustness of Policy Rules
Increase in IIP[1)] when a rule optimized in model X is used in model Y and evaluated relative to a rule optimized in model Y.

Loss (λ_y=0, $\lambda_{\Delta i}$=0.5): $Var(\pi) + 0.5Var(\Delta i)$			
IIP if evaluated in Model:	Rules[3)] optimized in TAYLOR Model		
	2 Parameters	3 Parameters (Gap)	4 Parameters (Gaps)
SW ACEL	0.34 0.02	0.70 0.11	0.77 0.12
	Rules optimized in SW Model		
TAYLOR ACEL	0.19 0.11	0.12 0.01	0.77 0.12
	Rules optimized in ACEL Model		
SW TAYLOR	0.47 0.82	0.10 0.28	0.12 0.46
Loss(λ_y=1,$\lambda_{\Delta i}$=0.5): $Var(\pi) + Var(y) + 0.5Var(\Delta i)$			
IIP if evaluated in: Model:	Rules optimized in TAYLOR Model		
	2 Parameters	3 Parameters (Gap)	4 Parameters (Gaps)
SW ACEL	0.18 0.02	6.04 M.E.[3)]	7.63 M.E.[3)]
	Rules optimized in SW Model		
TAYLOR ACEL	1.09 0.02	3.36 0.23	2.39 0.13
	Rules optimized in ACEL Model		
SW TAYLOR	0.08 0.27	55.4 24.4	0.54 3.57

Notes:

[1)] The values in this table concern the increase in absolute loss in model Y under a rule optimized for model X relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

[2)] Rules: 2 Parameters: $i_t = \alpha\pi_t + \beta_0 y_t$, 3 Parameters (Gap): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$;

3 Parameters(Growth),: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_A \Delta y_t$; 4 Parameters (Gaps): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$.

[3)]**M.E.** refers to indeterminacy and the existence of multiple self-fulfilling equilibria.

A-4. Sensitivity studies for "7. Conclusions and extensions / Utility Based Loss Functions"

We have computed optimized policy rules for the following quadratic approximation of the unconditional expectation of consumer welfare in the model of Smets and Wouters (2003) derived in the working paper version of Onatski and Williams (2004):

$$\begin{aligned} L &= L_{Onatski-Williams(2004)} + Var(\Delta i) \\ &= Var(\pi_t) + 0.21Var(K_{t-1}) - 0.51Cov(\pi_t, \pi_{t-1}) + 0.24Var(w_t) \\ &\quad +0.24Cov(w_t, \pi_t) - 0.24Cov(w_t, w_{t-1}) - 0.24Cov(\pi_t, w_{t-1}) \\ &\quad +Var(\Delta i) \end{aligned}$$

We have added $Var(\Delta i)$ to the loss function to account for central bank's tendency to keep interest rate volatility in check and avoid regular violations of the non-negativity constraint on nominal interest rates. The optimized 2-, 3- and 4-parameter rules are reported in Table A-5. The robustness of rules derived under the simpler loss function is evaluated in Table A-6. As noted in the main text, 2-parameter rules with inflation and the output gap and 3-parameter rules with interest rate smoothing that replace the output gap with the deviation of output growth from trend are quite robust, but not the other rules.

Table A-5
Optimized Rules in SW Model
$L = L_{Onatski-Williams(2004)} + Var(\Delta i)$

Utility-Based Loss	ρ	α	β_0	β_1	β_Δ	Loss
4 Parameter Rule (Gaps)	0.97	0.52	0.60	-0.52		22.97
3 Parameter Rule (Gap)	0.76	0.89	0.11	-		23.00
2 Parameter Rule (Gap)	-	2.55	0.17	-		23.54
3 Parameters Rule (Growth)	1.14	0.16			0.53	22.17

Table A-6

Performance and Robustness of Rules with Utility-Based Loss

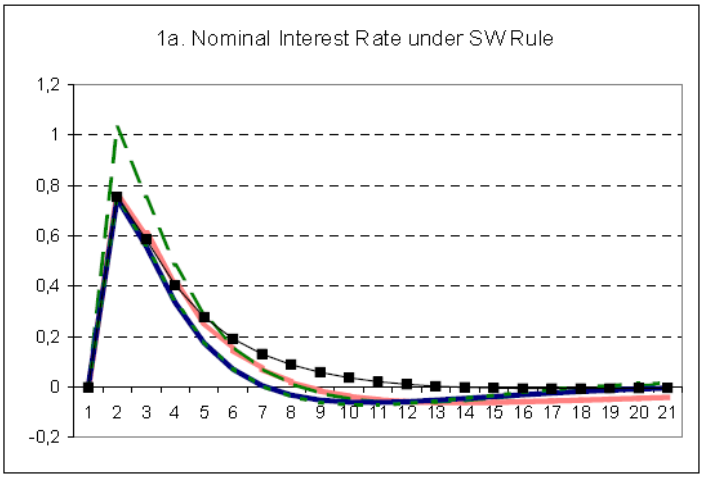
Compare the performance and robustness of rules optimized under the simple New-Keynesian (NK) loss function relative to the SW model utility-based loss function of Onatski and Williams (2004).

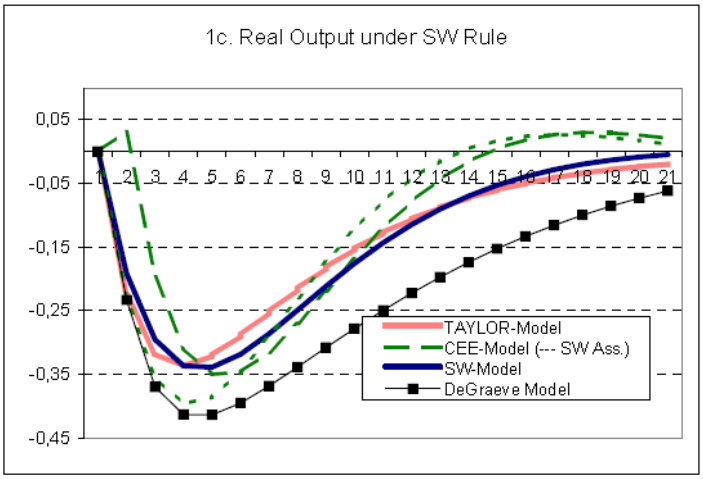
Rule	Loss
4 Parameter Rules (Gaps)	
Benchmark: **SW Model, SW-Loss**	**22.97**
SW, NK-Loss	26.02
TAYLOR, NK-Loss	110.8
ACEL, NK-Loss	24.91
3 Parameter Rules (Gap)	
Benchmark: **SW Model, SW-Loss**	**23.0**
SW, NK-Loss	25.08
TAYLOR, NK-Loss	79.88
ACEL, NK-Loss	4956
2 Parameter Rules (Gap)	
Benchmark: **SW Model, SW-Loss**	**23.54**
SW, NK-Loss	24.28
TAYLOR, NK-Loss	24.13
ACEL, NK-Loss	24.29
3 Parameter Rules (Growth)	
Benchmark: **SW Model, SW-Loss**	**22.17**
SW, NK-Loss	23.25
TAYLOR, NK-Loss	22.71
ACEL, NK-Loss	24.57

A-5. Sensitivity studies for "7. Conclusions and extensions / Robustness to other macroeconomic models"

We have extended our analysis of the performance and robustness of simple monetary policy rules by including the estimated DSGE model of the U.S. economy of De Graeve (2008) (DG Model in the following). This model includes a more detailed financial sector through a financial accelerator mechanism. The GDP response to a monetary policy shock remains very similar to the findings in Figure 1 in the main text (see Figure A4 below). As noted in the main text 2-parameter rules optimized in the DG model are not very robust to the other models. However, the model-averaging rules are still very robust when the DG model is included in the optimization, and remain very close to the model-averaging rules obtained with only three models. See Tables A7 to A11, which show the results with the DG model in comparison to the results with the other models reported previously in Tables 1 to 5 in the main text.

Figure A-4
Monetary policy shock under SW Rule in 4 Models





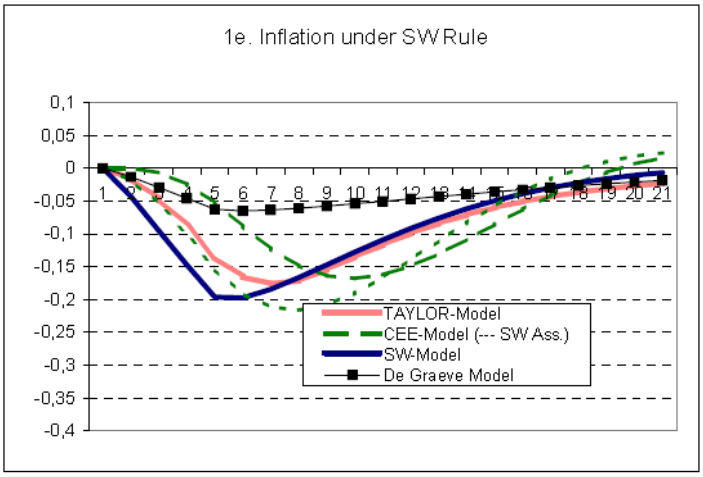


Table A-7
Optimal Simple Policy Rules[1)]

Rules: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1} + \beta_\Delta \Delta y_t$

Rule /Model	**Loss (λ_y = 0):** $Var(\pi) + Var(\Delta i)$					**Loss (λ_y=1):** $Var(\pi) + Var(y) + Var(\Delta i)$				
	ρ	α	β_0	β_1	β_Δ	ρ	α	β_0	β_1	β_Δ
	2 Parameters (Gap)[2)]					$\alpha\pi_t + \beta_0 y_t$				
TAYLOR		2.54	0.19				3.00	0.52		
SW		2.33	-0.10				2.04	0.26		
CEE/ACEL		4.45	0.28				2.57	0.45		
DG		1.34	0.31				1.51	0.60		
	3 Parameters (Gap)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$				
TAYLOR	0.98	0.37	0.09			0.98	0.21	0.53		
SW	1.06	0.49	0.01			1.13	0.012	0.015		
CEE/ACEL	0.97	0.99	0.02			2.84	7.85	-2.12		
DG	0.99	0.30	0.01			0.95	0.35	0.15		
	4 Parameters (Gaps)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$				
TAYLOR	0.98	0.37	0.07	0.02		0.96	0.18	0.41	0.19	
SW	1.06	0.46	-0.03	0.03		1.07	0.16	1.63	-1.62	
CEE/ACEL	1.01	1.11	0.18	-0.18		1.04	0.51	2.24	-2.30	
DG	0.99	0.30	0.04	-0.03		0.90	0.34	-0.24	0.40	
	3 Parameters (Growth)[3)]					$\rho i_{t-1} + \alpha\pi_t + \beta_\Delta \Delta y_t$				
TAYLOR	1.01	0.52			0.07	1.13	0.40			0.68
SW	1.03	0.48			-0.01	1.01	0.20			1.04
CEE/ACEL	1.02	1.07			-.002	0	3.71			.002
DG	1.00	0.30			-0.002	0.83	0.47			-0.02

Notes:

[1)] The loss function includes the variance of inflation and the variance of the first-difference of nominal interest rates with a weight of unity, $\lambda_{\Delta i}$=1. λ_y denotes the weight on the variance of the output gap.

[2)] In the Taylor model the output gap denotes the difference between actual and trend output. In the SW and ACEL models it is the difference to the level realized under flexible prices given current macroeconomic shocks.

[3)] The output growth measure Δy_t is defined relative to steady-state/trend output growth in all three models.

Table A-8
Increase in Loss when Reducing the Number of Parameters in the Rule
Percentage Increase (Increase in IIP)[1)]

Models	**Loss (λ_y = 0):** $Var(\pi) + Var(\Delta i)$		
	4 versus 3 Parameters (Gaps)	4 Parameters (Gaps) vs 3 Par. (Growth)	3 versus 2 Parameters (Gaps)
TAYLOR	0.12% (0.001)	13.5% (0.10)	278% (1.38)
SW	0.22% (0.001)	1.40% (0.01)	316% (0.78)
CEE/ACEL	5.10% (0.001)	10.0% (0.003)	229% (0.04)
DG	0.23%(0.001)	2.64% (0.01)	147% (0.35)
	Loss(λ_y = 1): $Var(\pi) + Var(y) + Var(\Delta i)$		
TAYLOR	1.81% (0.07)	67.1% (1.61)	98.8% (2.14)
SW	10.6% (0.47)	18.1% (0.76)	25.6% (1.17)
CEE/ACEL	14.4% (0.11)	36.7% (0.22)	9.67% (0.11)
DG	1.2% (0.04)	44.5% (0.91)	23.5%(0.57)

Notes:
[1)]The values in parentheses measure the increase in absolute loss in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

Table A-9
Robustness of Policy Rules
Increase in IIP[1)] when a rule optimized in model X is used in model Y relative to rule optimized in model Y.

Loss (λ_y=0): $Var(\pi)+Var(\Delta i)$				
IIP if evaluated in Model:	Rules optimized in TAYLOR Model			
	2 Parameters	3 Par. (Gap)	3 Par. (Growth)	4 Par. (Gaps)
SW	0.37	0.83	0.01	0.90
ACEL	0.03	0.12	0.01	0.14
DG	0.23	0.13	0.02	0.14
	Rules optimized in SW Model			
TAYLOR	0.27	0.13	0.03	0.15
ACEL	0.15	0.02	0.01	0.02
DG	M.E.[2)]	0.05	0.03	0.28
	Rules optimized in ACEL Model			
SW	0.54	0.11	0.10	0.09
TAYLOR	0.76	0.27	0.25	0.34
DG	0.80	0.16	0.19	0.14
	Rules optimized in DG Model			
TAYLOR	1.14 (2.30)[3)]	0.16	0.11	0.15
SW	3.20 (3.73)[3)]	0.11	0.03	0.11
ACEL	0.17	0.02	0.03	0.02
Loss(λ_y=1): $Var(\pi)+Var(y)+Var(\Delta i)$				
IIP if evaluated in:	Rules optimized in TAYLOR Model			
SW	0.17 (1.53)[3)]	5.41	0.66 (1.21)[3)]	7.18
ACEL	0.001	M.E.[2)]	0.31	M.E.[2)]
DG	0.40	2.07	0.87	2.31
	Rules optimized in SW Model			
TAYLOR	0.86 (2.64)[3)]	3.20	1.05 (2.28)[3)]	2.71
ACEL	0.03	0.21	0.44	0.13
DG	0.29	5.39	1.27	0.28
	Rules optimized in ACEL Model			
SW	0.07	108	1.69	0.53
TAYLOR	0.12	24.9	1.40	3.85
DG	0.28	15.1	0.88	0.72
	Rules optimized in DG Model			
TAYLOR	1.23 (2.89)[3)]	1.08	1.97	1.56
SW	2.09 (2.88)[3)]	0.96	1.54	1.74
ACEL	0.27	0.20	0.18	0.34

Notes:
[1)]The values in this table concern the increase in absolute loss under a particular rule relative to the comparable simple policy rule optimized in the respective model. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.
[2)]**M.E.** refers to indeterminacy and the existence of multiple self-fulfilling equilibria.
[3)]The values in parenthesis refer to the IIP that results from implementing the rule optimized in one model in any of the other model and comparing the generated loss to the loss that would be realized under 4-parameter rule optimized for that specific model.

Table A-10
Optimized Model-Averaging Rules

Objective: Min $\sum_{m \in M} \frac{1}{4}(Var(\pi_m) + Var(y_m) + Var(\Delta i_m))$

Rules: $i_t = \rho i_{t-1} + \alpha \pi_t + \beta_0 y_t + \beta_1 y_{t-1} + \beta_\Delta \Delta y_t$

Set of equally-weighted models: $M = \{SW, TAYLOR, ACEL, DG\}$	ρ	α	β_0	β_1	β_Δ
2 Parameters Rule (Gap)		2.64	0.53		
3 Parameters Rule (Gap)	1.03	0.42	0.22		
3 Parameters Rule (Growth)	1.09	0.25			0.72
4 Parameters rule (Gap)	1.06	0.21	0.68	-0.60	

Table A-11
Robustness of Model-Averaging Policy Rules (IIP[1])

Loss(λ_y=1): $Var(\pi) + Var(y) + Var(\Delta i)$				
IIP if evaluated in	2 Parameters	3 Par. (Gap)	3 Par. (Growth)	4 Par. (Gaps)
SW	0.10 (1.49) [2]	1.01	0.27 (0.94) [2]	0.48
TAYLOR	0.06 (2.19) [2]	0.59	0.13 (1.68) [2]	1.28
ACEL	0.01 (0.17) [2]	0.25	0.35 (0.41) [2]	0.16
DG	0.27 (0.75) [2]	0.11	0.99 (1.47) [2]	0.19

Notes:
[1] The values in this table concern the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.
[2] The values in parentheses refer to the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a 4-parameter rule optimized in model Y.

The New Keynesian Approach to Dynamic General Equilibrium Modeling: Models, Methods and Macroeconomic Policy Evaluation *

Sebastian Schmidt † Volker Wieland ‡

September 30, 2011

*Volker Wieland acknowledges funding support from European Community grant MONFISPOL under grant agreement SSH-CT- 2009-225149.

†Mailing address: Goethe University of Frankfurt, Grueneburgplatz 1, House of Finance, 60323 Frankfurt am Main, Germany, E-mail: s.schmidt@wiwi.uni-frankfurt.de.

‡Corresponding Author. Mailing address: Goethe University of Frankfurt, Grueneburgplatz 1, House of Finance, 60323 Frankfurt am Main, Germany, E-mail: wieland@wiwi.uni-frankfurt.de.

Abstract

This chapter aims to provide a hands-on approach to New Keynesian models and their uses for macroeconomic policy analysis. It starts by reviewing the origins of the New-Keynesian approach, the key model ingredients and representative models. Building blocks of current-generation dynamic stochastic general equilibrium (DSGE) models are discussed in detail. These models address the famous Lucas critique by deriving behavioral equations systematically from the optimizing and forward-looking decision-making of households and firms subject to well-defined constraints. State-of-the-art methods for solving and estimating such models are reviewed and presented in examples. The chapter goes beyond the mere presentation of the most popular benchmark model by providing a framework for model comparison along with a database that includes a wide variety of macroeconomic models. Thus, it offers a convenient approach for comparing new models to available benchmarks and for investigating whether particular policy recommendations are robust to model uncertainty. Such policy analysis is illustrated by evaluating the performance of simple policy rules across a range of recently-estimated models including some with financial market imperfections.

Keywords: Monetary macroeconomics, Keynesian models, New-Keynesian models, dynamic stochastic general equilibrium models, New Neoclassical synthesis, model comparison, rational expectations, policy evaluation, policy robustness, monetary and fiscal policy.

JEL-Codes: C51, C52, C61, C68, E12, E17, E52, E63

Contents

1 Introduction **1**

2 The New Keynesian Approach to Monetary Economics: A Brief History of Thought **3**

3 Building New Keynesian Models **6**

3.1 A Simple Model with Microeconomic Foundations 6

3.1.1 Households 7

3.1.2 Firms 9

3.1.3 The Government 11

3.1.4 Monetary Policy 12

3.1.5 Log-linearized System of Equations 12

3.1.6 Model Dynamics 14

3.2 Medium-Scale Models for Policy Analysis 16

3.2.1 Capital and Investment 19

3.2.2 Habit Formation in Consumption 21

3.2.3 Price indexation 23

3.2.4 Sticky wages 23

3.2.5 Financial market frictions 24

3.2.6 Model Dynamics 27

3.3 Using Structural Models for Policy Analysis: The Lucas Critique 28

4 Methods for Model Solution and Estimation **30**

4.1 Solving New-Keynesian Models 30

4.1.1 Linear Approximation 31

4.1.2 Solving a System of Linear Difference Equations 32

4.1.3 The Extended Path Solution Method for Nonlinear Models 35

4.1.4 Linear Quadratic Dynamic Programming: Value Function Iteration . . . 36
4.1.5 Perturbation Methods for Higher-Order Approximations 38
4.2 Estimating New-Keynesian Models 40
4.2.1 Bayesian Methods 40
4.2.2 Estimating a Small New Keynesian Model 44
4.3 Challenges for Model Estimation 47

5 A New Approach to Model Comparison and Policy Evaluation 50

6 Policy Evaluation and Policy Robustness under Model Uncertainty 57

7 Outlook: Unanswered Questions and Future Research 67

8 Appendices 69
8.1 Data sources and treatment 69
8.2 Augmented Models 71
8.3 A Database of Macroeconomic Models 74

Preliminary List of References 75

1 Introduction

What is New Keynesian Economics? In their 1991 introduction to a collection of seminal contributions Greg Mankiw and David Romer gave the following answer: (i) New Keynesian theories of business cycles posit that fluctuations in nominal variables like the money supply influence fluctuations in real variables, like output and employment; and (ii) real market imperfections such as imperfect competition or imperfect information also have an important influence on economic fluctuations. At the time, they contrasted New Keynesian thought with real business cycle theory that emphasized technological disturbances and perfect markets (cf. Kydland and Prescott (1982)). Constraints on price or wage adjustment constituted a central element of New Keynesian models of the economy. A first wave of New Keynesian models following the 1970s rational expectations revolution, such as Fischer (1977), Phelps and Taylor (1977) and Taylor (1979a,b) used long-term nominal contracts to explain how demand shifts cause real fluctuations even if expectations are rational and the shifts are anticipated.

The ensuing debate between real business cycle and New Keynesian theorists, and the successive extension and empirical application of both types of models, eventually triggered a second wave of New Keynesian models or monetary business cycle models that aimed to marry key ingredients of both approaches. The small-scale model of Goodfriend and King (1997) and Rotemberg and Woodford (1997) was quickly extended with additional decision aspects and constraints. These models, which are frequently referred to as New Keynesian dynamic stochastic general equilibrium (DSGE) models, are exemplified by the medium-scale model of the U.S. economy of Christiano et al. (2005). Nowadays, medium- to large-scale DSGE models are routinely used by economists at central banks and international institutions to evaluate monetary and fiscal stabilization policies.

The objective of this chapter is to explain how to build current-generation New Keynesian DSGE models, how to estimate them and how to use them for policy design. Given their in-

fluence on current macroeconomic thinking and policy analysis such a hands-on introduction should be useful for any reader interested in macroeconomics. However, several of the topics addressed in this chapter should also be of interest to a wider readership that uses computable general equilibrium modeling in many other areas of economic policy making. For example, the systematic handling of restrictions imposed by optimizing and forward-looking decision-making of economic agents subject to a variety of constraints that is practiced in the macro DSGE literature may be usefully applied in other areas. Furthermore, the methods used for approximating the solutions of nonlinear dynamic and stochastic models and for estimating them with economic data may easily be applied elsewhere. Finally, we review a new approach to model comparison that helps identifying robust policies under model uncertainty (see Wieland et al. (2011)). Such an undertaking appears particularly urgent to us, as some commentators have criticized macroeconomists in general and DSGE modelers in particular, for failing to foresee or to properly warn of the risk of a global financial crisis and recession. Clearly, this comparative approach could also be of benefit to practical model-based policy making in other arenas including international trade, economic development and climate change.

The remainder of this chapter proceeds as follows. Section 2 offers a brief history of thought and additional references regarding the development of New Keynesian macroeconomic models. Section 3 begins with a detailed presentation of the canonical New Keynesian model which provides the starting point for various extensions that help improve the model's empirical performance and usefulness for policy analysis. Emphasis is laid on the microeconomic foundations of the model and the implied cross-equation restrictions on the reduced-form system. We then discuss a selected number of extensions to the baseline model that have become popular in the literature. The section concludes with an illustration of the Lucas critique. Section 4 discusses methods for solving dynamic general equilibrium models and provides an introduction to Bayesian methods for model estimation. An example is given by the estimation of the small-scale New Keynesian model on data of the US economy. In the last part of the section

we address some remaining challenges for model estimation. Section 5 presents our approach to model comparison that allows for systematic and straightforward comparison and evaluation of macroeconomic models and alternative policies. Section 6 applies the comparative approach to evaluate the performance and robustness of monetary policy rules when the true model underlying the economy is unknown and the policymaker is instead confronted with a range of competing models. Section 7 concludes with an outlook on further research.

2 The New Keynesian Approach to Monetary Economics: A Brief History of Thought

The common characteristic of New Keynesian monetary models, compared with earlier models, is the combination of rational expectations, staggered price and wage setting, and policy rules. The term also is used to contrast these models with traditional Keynesian models without rational expectations. New Keynesian models rather than the traditional Keynesian models are the ones commonly taught in graduate schools because they capture how people's expectations and microeconomic behavior change over time in response to policy interventions and because they are empirically estimated and fit the data. They are therefore viewed as better for policy evaluation. In assessing the effect of government actions on the economy, it is important to take into account how households and firms adjust their spending decisions as their expectations of future government policy changes.

In the introduction, we have distinguished two waves of New Keynesian modeling in the last 35 years. Key driving factors of this scientific process included empirical failures of traditional approaches, intellectual challenges such as the Lucas critique, theoretical innovations such as the combination of nominal rigidities with forward-looking and optimizing behavior of economic agents and the invention of new modeling and estimation techniques. The first wave of New Keynesian models took off in the late 1970s. The apparent failure of traditional Keynesian

models to satisfactorily explain the 1970s stagflation raised many questions about the connection between inflation and economic activity and the role of monetary policy in stabilizing the economy. The famous Lucas critique underscored the need to account for the forward-looking and optimizing behavior of households and firms in macroeconomic models intended to be used for policy evaluation. Traditional Keynesian models were typically lacking these elements. Expectations where modeled as backward-looking, i.e. captured by past values of the respective variables, and model equations were not being directly related to individual optimization.

Innovations in the late 1970s and 1980s lead to the development of the first generation of New Keynesian models with rational expectations and nominal rigidities that allowed for interesting interactions between (systematic) monetary policy and real economic activity. These innovations included modeling of menu costs and overlapping wage and price contracts (Fischer (1977), Taylor (1979b), Calvo (1983)), new methods for solving linear and nonlinear dynamic models with rational expectations as well as successful estimation of such models using maximum likelihood techniques (Hansen and Sargent (1980), Fair and Taylor (1983)). First-generation New Keynesian models were extended, enlarged and eventually applied rather intensively in practical monetary policy analysis at central banks. We highlight the following three models from the 1990s that played an important role for U.S. monetary policy: Taylor's (1993) model of the G-7 economies, Fuhrer and Moore's (1995) model with relative-real-wage staggered contracts that helped explain U.S. inflation persistence, and the Federal Reserve's FRB-US model described, for example, in Reifschneider, Tetlow and Williams (1999). All three models are available for comparison and policy evaluation exercises from the model archive that is discussed in more detail in section 5.

Another challenge for Keynesian-style macroeconomic modeling arose from the real business cycle approach to macroeconomic fluctuations propounded by Kydland and Prescott (1982). Their extension of the neoclassical growth model to study the real (rather than monetary) sources of business cycles delivered a modeling approach that stringently enforced all the restrictions

following from the utility maximization of representative households and profit maximization of representative firms on the dynamics of macroeconomic variables. At the same time the real business-cycle approach put technological innovations forth as the main drivers of business innovations. Monetary policy has no real effects in the real-business-cycle world and therefore, stabilization policy is of minor concern. Goodfriend and King (1997) and Rotemberg and Woodford (1997) presented a first monetary business cycle model with such microeconomic foundations but also with nominal rigidities and imperfect competition. In this manner, New Keynesian research aims to incorporate Keynesian ideas into dynamic general equilibrium frameworks that are used in the real business cycle (RBC) literature. For this reason, the above-mentioned monetary business cycle model is alternatively referred to as the New Neoclassical Synthesis model or the New Keynesian DSGE model. The inclusion of nominal rigidities and imperfect competition had also been motivated by the failure of RBC models - as seen by part of the New Keynesian literature - to account for certain empirical regularities (Rotemberg and Woodford (1996), Galí (1999)).

Recent years have witnessed an explosion in New Keynesian modeling. Importantly, Christiano, Eichenbaum and Evans (2005) developed and estimated a medium-sized dynamic stochastic general equilibrium model with capital and investment, prices and wages, and a number of additional variety of adjustment costs and frictions. While they used impulse-response function matching techniques, Smets and Wouters (2003, 2007) showed how these models can be estimated more easily and effectively with Bayesian methods. This approach was quickly popularized and lead to wide-spread New Keynesian model building at central banks around the world. Levin, Wieland and Williams (2003) and Taylor and Wieland (2011) provide a comparison of these and earlier New-Keynesian macroeconomic models and their implications for monetary policy. New Keynesian models offer many uses for practical policy analysis. They can be utilized to assess the implications and desirability of alternative policies and institutional developments such as the creation of a common currency area in Europe. Medium-scale models

exhibiting a wide range of frictions have been developed as tools for forecasting exercises, for evaluations of the effects of policy changes and for historical decompositions, eg. Christiano et al. (2005), Smets and Wouters (2003, 2007), Adolfson et al. (2007b).

3 Building New Keynesian Models

New Keynesian business cycle models are characterized by a set of key assumptions and ingredients. Similar to real business cycle models, modern New Keynesian models are general equilibrium models. Equilibrium conditions are explicitly derived from the optimization problem of consumers and producers. A standard assumption is that agents have rational expectations, that is agents form model-consistent expectations conditional on the information available. Producers have market power over prices which facilitates the introduction of short-run nominal price rigidities. The presence of nominal rigidities is the key ingredient that distinguishes New Keynesian models from RBC models and that assigns an explicit role for monetary stabilization policy.

3.1 A Simple Model with Microeconomic Foundations

This section shortly reviews the simple stochastic monetary New Keynesian model that has become a much-used workhorse model and is now widely taught in the first-year macro sequence in graduate school (see Galí (2008), Galí and Gertler (2007), Goodfriend and King (1997), Walsh (2010), Woodford (2003)). The model economy is inhabited by households, monopolistic-competitive firms, the monetary authority and a government sector. Households decide how much to consume and how much labor to supply in order to maximize their lifetime utility. In turn, firms hire labor in order to produce differentiated goods. In contrast to the RBC literature, firms do not act under perfect competition but under monopolistic competition, which converts them from price-takers to price-setters. This assumption is necessary in order to

introduce price stickiness. Specifically, firms can reset prices only once in a while. Due to these nominal rigidities the monetary authority can affect real activity in the short run because the real interest rate will no longer be insensitive to movements in the monetary policy instrument, the short-term nominal interest rate. Finally, the government collects lump-sum taxes and consumes part of the final good. This setting is augmented with a set of stochastic shocks.

3.1.1 Households

We consider an economy that is made up of a large number of identical households. The representative household is characterized by her *preferences* regarding consumption, labor and real money balances as represented by her objective function

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[U\left(C_t, M_t/P_t\right) - V\left(H_t\right) \right]. \tag{1}$$

Equation (1) represents households' expected discounted *life-time utility*, where C_t denotes the household's consumption of a basket of differentiated goods, M_t measures her end-of-period money balances, P_t is the price of the consumption good basket in terms of money, and H_t denotes the number of hours worked. The inclusion of real money balances in the utility function is a common way to capture their transaction services, see e.g. Woodford (2003).[1] The consumption goods basket C_t consists of a continuum of differentiated goods

$$C_t \equiv \left[\int_0^1 C_t\left(i\right)^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2}$$

[1] Alternatively, one could model transactions frictions explicitly by introducing a cash-in-advance constraint on household consumption.

where $\varepsilon > 1$ and $C_t(i)$ denotes consumption of good i. The price index P_t is then defined as the minimum expenditure at which the household can buy one unit of C_t

$$P_t = \left[\int_0^1 P_t(i)^{1-\varepsilon} di \right]^{\frac{1}{1-\varepsilon}}, \tag{3}$$

where $P_t(i)$ denotes the price of good i. One can show that $P_t C_t = \int_0^1 P_t(i) C_t(i) di$ and

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} C_t. \tag{4}$$

Thus, household demand for good i depends on its relative price, $p_t(i) = \frac{P_t(i)}{P_t}$, with ε representing the elasticity of demand. A one percent increase in the relative price of good i leads to a reduction in the demand for this good of ε percent.

The period utility function $U(C, M/P)$ is assumed to be strictly increasing and concave in each of its arguments, and $V(H)$ is assumed to be increasing and convex. Finally, $0 < \beta < 1$ denotes the subjective discount factor. Under rational expectations, the representative household maximizes (1) subject to a sequence of budget constraints

$$P_t C_t + M_t + Q_{t,t+1} B_t \leq W_t H_t + M_{t-1} + B_{t-1} + T_t + \Gamma_t, \tag{5}$$

for all t. B_t represents the quantity of a one-period, riskless, nominal government bond paying one unit of money per bond in period $t+1$, where its price is denoted by $Q_{t,t+1}$. It holds $Q_{t,t+1} = \frac{1}{R_t}$, where R_t is the riskless one-period gross nominal interest rate. The nominal wage rate is denoted by W_t, T_t are (possibly negative) lump-sum transfers of the government and Γ_t denotes firms' profits distributed to the household sector. The optimality conditions of the optimization

problem read

$$\frac{1}{R_t} = \beta E_t \frac{U_C(C_{t+1}, m_{t+1})/P_{t+1}}{U_C(C_t, m_t)/P_t} \tag{6}$$

$$\frac{V_H(H_t)}{U_C(C_t, m_t)} = w_t \tag{7}$$

$$\frac{U_m(C_t, m_t)}{U_C(C_t, m_t)} = \frac{R_t - 1}{R_t}, \tag{8}$$

where $w_t = W_t/P_t$ is the real wage. U_C and U_m with $m = M/P$ denote the marginal utility of consumption and real money balances, respectively, and V_H measures the marginal disutility of labor. We will interpret these optimality conditions when summarizing the complete set of model equations.

3.1.2 Firms

The economy is inhabited by a continuum of firms of measure one. Each firm i possesses a production technology

$$Y_t(i) = A_t N_t(i), \tag{9}$$

where A_t denotes a common technology shock. In this simple model, labor is the only production input. Demand for good i is given by

$$Y_t(i) = C_t(i) + G_t(i), \tag{10}$$

where $G_t(i)$ denotes government purchases of good i, satisfying

$$G_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon} G_t. \tag{11}$$

The public consumption good basket G_t is defined equivalent to the private consumption good basket (2).

Firms are price-setters, however, following Calvo (1983), it is assumed that in a given period each firm can reset its price $P_t(i)$ only with probability $1-\theta$. Therefore, each period a fraction $1-\theta$ of firms reoptimizes its price while the remaining fraction θ of firms keep their price unchanged. Importantly, the probability of a change in the price by some firm i is independent of the time elapsed since its last price change. Sticky prices are an important feature of our model because it introduces *nominal rigidities*. As we will see, this allows monetary policy to affect real variables in the short run.

To produce output $Y_t(i)$, firms have to hire labor. Minimizing production costs for a given demand level subject to the production technology leads to

$$MC_t(i) = \frac{W_t}{A_t}, \tag{12}$$

where MC_t is the Lagrange multiplier representing marginal costs. In equilibrium, marginal costs of firm i equal the wage divided by the marginal product of labor. Note, that in our model marginal costs are identical across firms, $MC_t(i) = MC_t$. We can then formulate the optimization problem of firm i that resets its price in period t, taking into account that the price set today might be effective for some time and taking as given the demand for it's good, as follows

$$\max_{P_t(i)} \sum_{j=0}^{\infty} E_t Q_{t,t+j} \theta^j Y_{t+j}(i) \left[P_t(i) - MC_{t+j}\right], \tag{13}$$

subject to demand functions (4) and (11), where

$$Q_{t,t+j} = \beta^j E_t \frac{U_C\left(C_{t+j}, m_{t+j}\right)/P_{t+j}}{U_C\left(C_t, m_t\right)/P_t} \tag{14}$$

is the stochastic discount factor. The first-order condition then reads

$$\sum_{j=0}^{\infty} E_t Q_{t,t+j} \theta^j Y_{t+j} P_{t+j}^{\varepsilon} \left[P_t^*(i) - \frac{\varepsilon}{\varepsilon-1} MC_{t+j}\right] = 0, \tag{15}$$

where $P_t^*(i)$ is the optimal price set by firm i in period t. Equation (15) reveals that all firms reoptimizing their price in a given period will set the same price, $P_t^*(i) = P_t^*$. In the case of flexible prices, (15) reduces to $P_t^* = \frac{\varepsilon}{\varepsilon-1} MC_t$, that is the optimal price is a constant markup over contemporaneous marginal costs. Instead, in our sticky-price model, the optimal price is a markup on a weighted sum of current and expected future marginal costs. From (3) follows then that the aggregate price level is given by

$$P_t = \left[\theta P_{t-1}^{1-\varepsilon} + (1-\theta)(P_t^*)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}. \tag{16}$$

3.1.3 The Government

The government consumes part of the produced goods. Market clearing of all goods markets implies

$$Y_t = C_t + G_t, \tag{17}$$

where $Y_t = \left(\int_0^1 Y_t(i)^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}}$. When we simulate the model, we will assume that government expenditures as a share of steady state total output follow a simple AR(1) process. The government budget constraint reads

$$P_t G_t + B_{t-1} = \frac{B_t}{R_t} - T_t + M_t - M_{t-1}, \tag{18}$$

hence government spending is financed by a combination of one-period nominal government bonds, lump-sum taxes (negative transfers) and seigniorage revenues. Note, that fiscal policy is Ricardian, hence the exact choice of feasible tax and debt paths is irrelevant.

3.1.4 Monetary Policy

The monetary authority sets the short-term nominal interest rate, R_t. Here, we assume that the interest rate is set according to a simple rule which is known by all agents in the economy

$$\frac{R_t}{R} = \left(\frac{\pi_t}{\pi}\right)^{\tau_\pi} \left(\frac{Y_t^{gap}}{Y^{gap}}\right)^{\tau_Y} \upsilon_t, \tag{19}$$

where $\pi_t = P_t/P_{t-1}$ is the gross inflation rate between period $t-1$ and t, and Y_t^{gap} denotes the output gap, ie the deviation of actual output from some natural level, that will be defined explicitly below. Variables without a time subscript denote steady state values of the respective variable. Unsystematic components of interest rate policy are captured by the monetary policy shock υ_t.

3.1.5 Log-linearized System of Equations

After we have also imposed market clearing on the labor market and the market for money balances and government bonds, we can summarize the complete model by a set of log-linearized equations. Let $\hat{x}_t = log(x_t) - log(x)$, for some variable x_t, where x denotes the corresponding steady state level. That is, we express the variable in percentage deviations from its steady state. Section 4.1.1 contains an introduction to the procedure of (log)-linearisation. We log-linearize equations (6), (7), (8), (9), (15), (16), (17) and (19) around the non-stochastic steady state with

zero inflation, ie a gross steady state inflation rate of $\pi = 1$

$$\hat{C}_t = E_t\hat{C}_{t+1} - \frac{1}{\sigma}\left(\hat{R}_t - E_t\hat{\pi}_{t+1}\right) \tag{20}$$

$$\hat{H}_t = \frac{1}{\eta}\hat{w}_t - \frac{\sigma}{\eta}\hat{C}_t \tag{21}$$

$$\hat{m}_t = \frac{\sigma}{\sigma_m}\hat{C}_t - \frac{1}{(\beta^{-1}-1)\,\sigma_m}\hat{R}_t \tag{22}$$

$$\hat{Y}_t = \hat{A}_t + \hat{H}_t \tag{23}$$

$$\hat{Y}_t = \frac{C}{Y}\hat{C}_t + \hat{g}_t \tag{24}$$

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \kappa\hat{Y}_t^{gap} \tag{25}$$

$$\hat{R}_t = \tau_\pi\hat{\pi}_t + \tau_Y\hat{Y}_t^{gap} + \hat{\upsilon}_t, \tag{26}$$

where $\sigma \equiv -\frac{U_{C,C}(C,m)}{U_C(C,m)}C$, $\eta \equiv \frac{V_{H,H}(H)}{V_H(H)}H$, $\sigma_m \equiv -\frac{U_{m,m}(C,m)}{U_m(C,m)}m$ and $\kappa = \frac{(1-\beta\theta)(1-\theta)}{\theta}\left(\frac{\sigma}{C/Y} + \eta\right)$, and we assumed $U_{C,M}(C,m) = U_{M,C}(C,m) = 0$, ie utility is separable between consumption and real money balances. Equation (20) is the log-linearized consumption Euler equation (6) stating that consumption increases when expected future consumption increases or the ex-ante real interest rate, $\hat{R}_t - E_t\hat{\pi}_{t+1}$, decreases. Equation (21) can be interpreted as a labor supply equation, showing that the number of hours worked depends positively on the equilibrium real wage and negatively on the consumption level. Real money demand is given by equation (22). In this model, the money demand function only serves the purpose to determine the amount of money that the central bank has to supply given a certain nominal interest rate implied by the monetary policy rule. Equation (23) represents the production technology aggregated over all firms, and the resource constraint is shown in equation (24), where $\hat{g}_t \equiv \frac{G_t - G}{Y}$. Equation (25) results from combining equations (15) and (16), and represents the so-called New Keynesian Phillips curve, stating that the inflation rate depends on expected future inflation and the contemporaneous output gap. The latter is defined as $\hat{Y}_t^{gap} = \hat{Y}_t - \hat{Y}_t^{nat}$, i.e. as the percentage deviation of output from natural output,

that is the level of output in an economy without any frictions. The latter is defined as

$$\hat{Y}_t^{nat} = \frac{1}{\sigma/(C/Y)+\eta}\left[(1+\eta)\hat{A}_t + \frac{\sigma}{C/Y}\hat{g}_t\right]. \tag{27}$$

Finally, equation (26) is the log-linearized monetary policy rule. Government spending (as a share of total output), aggregate technology and the monetary policy shock are assumed to follow AR(1) processes

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \varepsilon_t^g \tag{28}$$

$$\hat{A}_t = \rho_A \hat{A}_{t-1} + \varepsilon_t^A \tag{29}$$

$$\hat{\upsilon}_t = \rho_\upsilon \hat{\upsilon}_{t-1} + \varepsilon_t^R, \tag{30}$$

where ε_t^j, $j \in \{g, A, R\}$, are zero mean, constant variance *iid* innovations. The set of equations can then be solved to obtain the solution functions describing the equilibrium dynamics of the endogenous variables. Importantly, unlike the structural equations these solution functions are not independent of monetary policy. Methods for model solution are discussed in 4.1, here we proceed instead with an analysis of the model dynamics.

3.1.6 Model Dynamics

A convenient way to depict the propagation mechanisms of the model is to calculate impulse response functions. Impulse responses isolate the propagation of a particular shock through the economy.

We calibrate the model parameters as follows. Assuming that the period length is one quarter, the subjective discount factor is set to $\beta = 0.99$ which implies a steady state annualized interest rate of around 4 percent. We set $\sigma = 1.5$ and $\eta = 1$. The preference parameter regarding real money balances is set to a rather high value of $\sigma_m = 110$, consistent with the empirical evidence

presented in Andres et al. (2006). The Calvo parameter is fixed at $\theta = 0.75$, implying that prices are reset on average every four quarters. Steady-state government spending as a share in GDP is set to 0.2. Finally, the parameters in the monetary policy function are $\tau_\pi = 1.5$ and $\tau_Y = 0.5/4$, similar to the rule discussed by Taylor (1993b) and the AR-coefficients of the three structural shocks in our model are set to $\rho_g = 0.85$, $\rho_A = 0.9$ and $\rho_\upsilon = 0.5$.

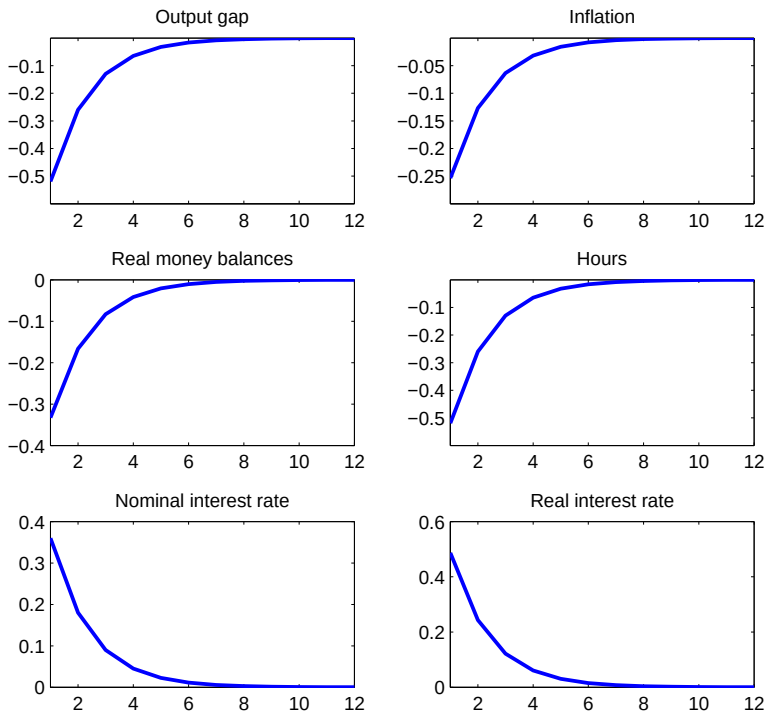
Figure 1 displays the dynamic responses of several variables to a monetary policy shock. Since prices are sticky, the increase in the nominal interest rate fosters the real rate to rise as well. This dampens aggregate demand by private households. In turn, firms require less labor which reduces equilibrium marginal costs and creates downward-pressure on inflation. Note, that monetary policy affects only the deviation of economic activity from its natural level, thus the responses of the output gap and output (not shown) are identical. Finally, the rise in the nominal interest rate and the reduction in output both serve to decrease households' demand for real money balances.

Figure 2 displays dynamic responses to a technology shock. The improvement in firms' production technology reduces production costs. However, due to sticky prices, only part of the firms can lower their prices immediately. Hence, the increase in aggregate demand is less than proportional to the improvement in technology. Therefore, equilibrium hours of work temporarily fall and the output gap becomes negative. The fall in inflation and the output gap induce the monetary authority to lower interest rates, however not by as much as would be needed in order to completely offset the decline in the two target variables. In the literature, the effects of exogenous changes in technology on various macroeconomic variables is controversial, see e.g. the literature overview by Galí and Rabanal (2004).

Figure 3 displays dynamic responses to a government spending shock. The increase in government demand for the composite consumption good stimulates aggregate demand. However, we observe that private consumption is partially crowded out and hence total output rises by less than the induced stimulus. The fall in private consumption results from anticipated higher taxes and from an intertemporal substitution effect due to the increase in interest rates following the

government spending shock.[2] The interest rate response is however not aggressive enough to prevent small increases in the output gap and inflation.

Figure 1: Monetary policy shock



3.2 Medium-Scale Models for Policy Analysis

While the baseline model described in the previous section is a useful starting point to understand some of the main ingredients of the NK modeling approach, more elaborated models are needed to perform forecast exercises and policy simulations. In particular, the baseline model has a hard time to capture the high degree of persistence observed for many macroeconomic variables. A popular route to address this issue is the incorporation of additional frictions.

[2]Part of the literature argues that empirical evidence points towards a positive response of private consumption to an increase in government spending. Galí et al. (2007) show that a New Keynesian model can match these dynamics when augmented with rule-of-thumb consumers.

Figure 2: Technology shock

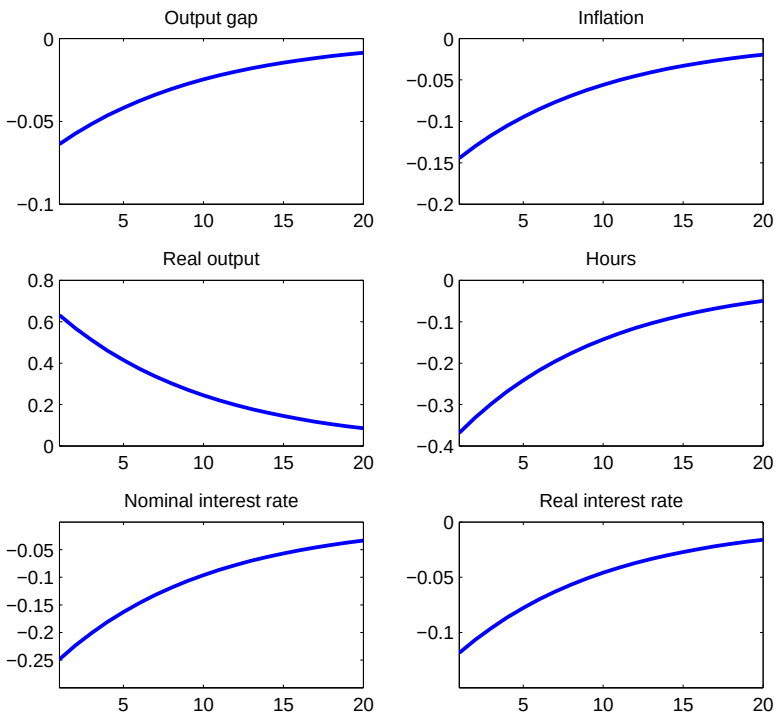
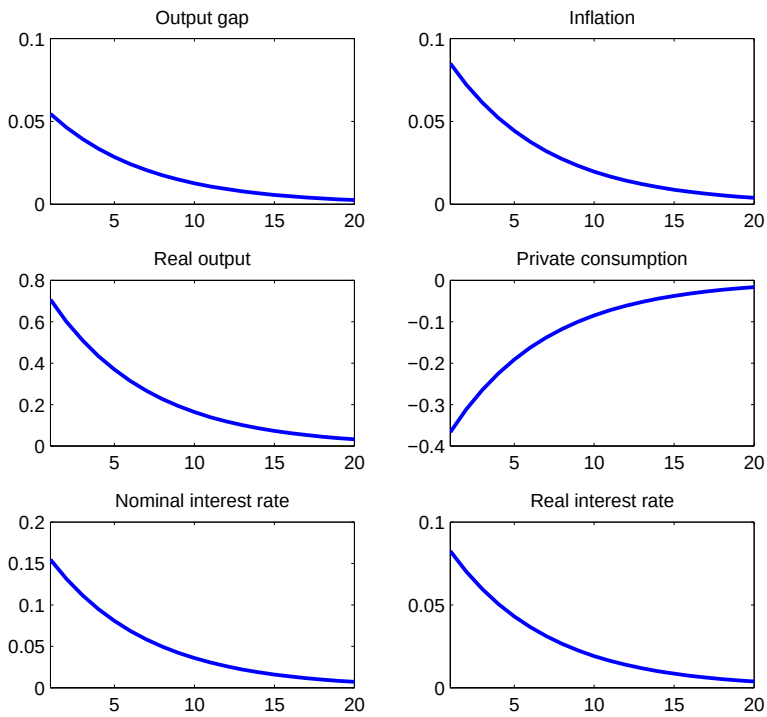


Figure 3: Government spending shock



The first medium-scale model that fully incorporates recent advances in terms of microeconomic foundations is presented in Christiano et al. (2005). Their model features additional nominal rigidities not present in the small-scale model such as staggered wage contracts and price and wage indexation, i.e. firms / households that cannot reoptimize their price / wage in a given period index it to lagged inflation. Furthermore, real frictions are employed. Christiano et al. (2005) introduce variable capital and allow for investment adjustment costs and variable capital utilization. Households exhibit habit formation in preferences for consumption. Christiano et al. (2005) estimate their model via minimization of the distance between the model and empirical impulse response functions to a monetary policy shock from a VAR for the US economy. Smets and Wouters (2003, 2007) employed Bayesian estimation techniques to estimate medium-scale models similar to the Christiano et al. (2005) model for the euro area and the US, respectively. The Bayesian likelihood approach has become popular among academics and central bankers working with DSGE models. In the following, we present a selection of frictions and rigidities that are often employed in medium-scale New Keynesian models and discuss their role for the model dynamics.

3.2.1 Capital and Investment

Typically, in medium-scale models production of the consumption good requires labor and capital services as inputs. Here, we follow the assumption of Christiano et al. (2005) that households own the economy's capital stock and rent capital services to firms in an economy-wide rental market for capital. The production function in the small-scale model, equation (9), is replaced by

$$Y_t(i) = A_t F\left(N_t(i), K_t^s(i)\right). \tag{31}$$

Here, K_t^S denotes capital services rent from households and F represents a Cobb-Douglas production function, $F\left(N_t(i), K_t^s(i)\right) = N_t(i)^{1-\alpha}\left(K_t^S(i)\right)^{\alpha}$, where $\alpha \in (0,1)$ denotes the capital

services share in production. The stock of physical capital, K_t, follows

$$K_t = (1-\delta) K_{t-1} + S(i_t, i_{t-1}), \tag{32}$$

where the parameter δ denotes the capital depreciation rate, i_t are purchases of the investment good and the function S represents the technology for the production of new capital goods as a function of current and past investment, and captures investment adjustment costs. Christiano et al. (2005) assume $S(i_t, i_{t-1}) = \left(1 - \tilde{S}\left(\frac{i_t}{i_{t-1}}\right)\right) i_t$, where $\tilde{S}(1) = \tilde{S}'(1) = 0$ and $\tilde{S}''(1) > 0$. New capital becomes productive with a one period lag, thus the amount of capital services in the current period can only be altered through the utilization rate of capital, u_t, which is set by the representative household

$$K_t^S = u_t K_{t-1}. \tag{33}$$

Variation in the utilization rate is subject to some costs, $a(u_t) K_{t-1}$, where a is an increasing, convex function with $a(1) = 0$ and the steady-state utilization rate satisfies $u = 1$. Both, the utilization costs and the household's earnings from renting capital services to the firms, $R_t^k u_t K_{t-1}$, enter the budget constraint, where R_t^k denotes the rental rate of capital. Cost minimization by the household then requires that the marginal benefit of raising the utilization rate equals marginal costs, that is $R_t^k = a'(u_t)$, or, after log-linearizing $\frac{1}{\sigma_a} \hat{R}_t^k = \hat{u}_t$, where $\sigma_a = \frac{a''(1)}{a'(1)} > 0$. Empirical estimates of σ_a typically point to a small value, see e.g. Christiano et al. (2005). In turn, this implies that the elasticity of the capital utilization rate with respect to the rental rate of capital is large. Christiano et al. (2005) find that a variable capital utilization rate is crucial for their model to generate the desired inertia in the inflation response to a monetary policy shock as well as a persistent output response. Without variable capital utilization, firms' costs for capital would be more sensitive with respect to an expansionary monetary policy shock, resulting in stronger inflationary pressures and weaker effects on real output. In the literature, model devices such as a variable utilization rate are referred to as real rigidities.

Alternatively to the assumption of competitive markets for production inputs it is sometimes assumed that these inputs are firm-specific. For instance, if capital is firm-specific, each individual firm accumulates capital only for its own use. This introduces another real rigidity, see e.g. Sveen and Weinke (2005) and Woodford (2005). With an economy-wide market for capital, an increase in demand in part of the firm sector will increase the rental price for capital for all firms. In contrast, with firm-specific capital, the individual firm's variable production costs are less affected by an increase in demand for some other firms' products. Importantly, firm-specific production inputs help to dampen the effect of an expansionary shock on inflation, see e.g. Eichenbaum and Fisher (2007).

Turning to the evolution of investment, the introduction of investment adjustment costs implies that the household's first order condition with respect to investment involves lagged as well as expected future investment

$$\hat{i}_t = \frac{1}{1+\beta}\hat{i}_{t-1} + \frac{\beta}{1+\beta}E_t\hat{i}_{t+1} + \frac{1}{(1+\beta)\tilde{S}''(1)}\hat{P}_t^k, \tag{34}$$

where $\hat{P}_t^k$ denotes the real value of the existing capital stock. The lagged investment variable helps to generate endogenous persistence not present in the purely forward-looking baseline New Keynesian model presented before. Specifically, investment and therefore output exhibit hump-shaped responses to an expansionary monetary policy shock.

3.2.2 Habit Formation in Consumption

In the presence of habit formation in consumption, household period utility from consumption does no longer depend solely on the level of current consumption but instead on the difference between current consumption and previous period's consumption level.[3] One distinguishes be-

[3] Alternatively, in some models the ratio between current consumption and previous period's consumption level enters the utility function. See Schmitt-Grohe and Uribe (2005) for an overview of modeling approaches to habit formation.

tween external and internal habits. The former relate the current level of consumption to aggregate past consumption, the latter to individual past consumption. Assuming internal habits, the representative household's objective function (1) would change to

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[U\left(C_t - bC_{t-1}, M_t/P_t\right) - V\left(H_t\right) \right], \tag{35}$$

where $b \in [0,1]$ is referred to as the habit parameter. In the case of $b = 0$, (35) reduces to the utility function of the standard model. With $b > 0$, aggregate current consumption will no longer depend solely on expected future consumption but also on past consumption and consumption Euler equation (20) is modified accordingly. Christiano et al. (2005) report a point estimate of $b = 0.65$. Similar to investment adjustment costs, consumption habits help to increase the degree of endogenous model persistence. Specifically, the response of consumption to exogenous shocks becomes more inertial and exhibits a hump-shaped pattern. Without habit formation, impulse responses of consumption to a monetary policy shock peak in the initial period and then return back to steady state monotonically as in our small-scale model.[4] To gain intuition, reconsider Euler equation (20) slightly rewritten as

$$E_t \Delta \hat{C}_{t+1} = \frac{1}{\sigma} \left(\hat{R}_t - E_t \hat{\pi}_{t+1} \right), \tag{36}$$

where Δ denotes the first difference operator. When interest rates are low, due to some expansionary shock, (36) implies that (expected) consumption growth is low, too. This requires that one starts from a high consumption level that decreases monotonically in the subsequent periods. Instead, with habit formation, it is the change in the growth rate of consumption that is related to the interest rate. In this case, low interest rates imply that the initially positive growth rate of consumption is declining, which translates into a hump-shaped impulse response pattern.

[4]The response of consumption to a monetary policy shock in the small-scale model is qualitatively equivalent to the response of the output gap shown in Figure 1.

3.2.3 Price indexation

In the small-scale model, the rate of inflation in the New Keynesian Phillips curve (25) is a purely forward-looking variable. Thus, the baseline model has a hard time to replicate the degree of inertia in the inflation rate found in the data. Empirical estimates often provide supportive evidence for a lagged inflation term on the right-hand side of (25), see e.g. Galí and Gertler (1999). One possibility to allow for lagged inflation in the NKPC of the structural model is the introduction of price indexation. Under (partial) price indexation, firms that do not receive a Calvo signal to reoptimize their price in a given period will instead mechanically increase previous period's price by an amount proportional to past inflation. This gives rise to the following NKPC

$$\hat{\pi}_t = \frac{\gamma}{1+\beta\gamma}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\gamma}E_t\hat{\pi}_{t+1} + \frac{\kappa}{1+\beta\gamma}\hat{Y}_t^{gap}, \tag{37}$$

where the parameter γ represents the degree of price indexation to past inflation. A similar expression can be derived under the assumption that part of the firms follow simple rules of thumb when deciding about their price, see Galí and Gertler (1999). While price indexation helps to produce additional inertia in the rate of inflation, there is little microeconomic evidence that firms change prices continuously, see e.g. Klenow and Malin (2010) and the references therein.

3.2.4 Sticky wages

In the baseline model presented before, households and firms interact in a perfectly competitive labor market. This assumption has rather unrealistic implications for wage dynamics. Medium-scale models therefore generally feature some form of staggered wage setting. Erceg et al. (2000) employ the Calvo setup to introduce sticky wages into the baseline model. In this model extension, households supply differentiated labor services over which they exhibit some monopolistic power. While each household supplies one type of labor service, firms employ all types of

labor to produce consumption goods. Similar to the price-setting framework, each period a randomly drawn fraction of households is allowed to reset their nominal wage while the remaining fraction demands the same wage as in the previous period. The optimization problem of a household that is allowed to reset its nominal wage in the current period then consists of choosing the nominal wage that maximizes her expected discounted lifetime utility, taking into account that she might not be able to reoptimize her nominal wage for some time in the future. Since labor income differs across households, individual private consumption does not have to be the same for all households. However, in the literature it is often assumed that households have access to complete asset markets that allow for full consumption risk sharing across households. The optimality condition for the wage setting decision results in a Phillips curve for wage inflation

$$\hat{\pi}_t^W = \beta E_t \hat{\pi}_{t+1}^W + \kappa_W \hat{Y}_t^{gap} - \xi_W \hat{w}_t^{gap}, \tag{38}$$

where $\hat{\pi}_t^W = \hat{W}_t - \hat{W}_{t-1}$ is the wage inflation rate, and $\hat{w}_t^{gap}$ denotes the real wage gap, i.e. the deviation of the actual real wage from its natural level that would obtain in the analogue model without price and wage rigidities. The wage inflation equation replaces labor supply equation (21) in the baseline model. The presence of sticky wages leads to a more muted response of real wages to monetary policy shocks.

3.2.5 Financial market frictions

The recent global financial crisis has drawn attention to the importance of an explicit modeling of financial market frictions in standard DSGE models. A prominent approach to integrate financial frictions into microfounded models of the macroeconomy is the so-called financial accelerator. We first present the main features of the financial accelerator model of Bernanke et al. (1999) and then discuss some more recent extensions.

Bernanke et al. (1999) introduce credit market imperfections into an otherwise standard New

Keynesian model with variable capital and demonstrate that these contribute to propagate and amplify the response of key macroeconomic variables to nominal and real shocks. Specifically, they consider an agency problem due to asymmetries of information in borrower-lender relationships. The economy is inhabited by three types of agents, risk-averse households, risk-neutral entrepreneurs and retailers. Entrepreneurs use capital and labor to produce wholesale goods that are sold to the retail sector which acts under monopolistic competition. Each period, entrepreneurs have to accumulate capital that becomes productive in the subsequent period. Bernanke et al. (1999) assume that entrepreneurs have finite live-time horizons which precludes the possibility that their aggregate wealth is increasing without bounds. Instead, entrepreneurs have to borrow from households via a financial intermediary to finance part of the new capital. The agency problem arises because the return to capital is prone to idiosyncratic risk and can only be observed by the financial intermediary if it pays some auditing cost.[5] Therefore, the entrepreneurs' net worth becomes a crucial determinant of their borrowing costs. If net worth is high, less of the capital acquisition has to be financed via external borrowing, thereby reducing the severity of the agency problem. The optimal contract in this environment turns out to be similar to a standard debt contract. The contract is characterized by a non-default loan rate, Z_t^j, and a threshold value for the idiosyncratic shock, ω^j, denoted by $\overline{\omega}^j$. The latter is defined as the minimum realization of the idiosyncratic shock required in order for the entrepreneur to be able to repay the loan

$$\overline{\omega}^j R_{t+1}^k Q_t K_t^j = Z_t^j B_t^j, \tag{39}$$

where R_t^k is the gross return to capital averaged across firms, Q_t is the price per unit of capital and K_t^j denotes the amount of capital acquired by entrepreneur j in period t for production in period $t+1$. The funds that have to be borrowed equal $B_t^j = Q_t K_t^j - N_t^j$, where N_t^j is net worth of entrepreneur j at the end of period t. Entrepreneurs accumulate net worth primarily from profits

[5] This framework refers to the so-called costly state verification problem of Townsend (1979).

from capital investment and to a minor part also from the supply of labor. If the realization of the idiosyncratic shock lies below the contractual threshold level, the entrepreneur defaults, the financial intermediary pays the auditing costs and occupies the entrepreneur's wealth left. Since the idiosyncratic loan risk is perfectly diversifiable, the opportunity costs of the financial intermediary equal the risk-free nominal interest rate. Any aggregate risk is absorbed by the risk-neutral entrepreneurs as specified in the contract. Each entrepreneur then has to choose the amount of capital it wants to buy. The optimality condition relates the ratio of external finance costs to the riskless rate and the ratio of capital expenditures to net worth. The aggregated log-linearized condition reads

$$E_t\left(\hat{R}^k_{t+1}\right) - \hat{R}_t = \chi\left(\hat{Q}_t + \hat{K}_t - \hat{N}_t\right), \tag{40}$$

where parameter $\chi > 0$ is a function of the structural model parameters.[6] Equation (40) shows that the so called external finance premium, i.e. the difference between the costs of external funds and the opportunity costs of internal funds, rises with the amount of external borrowing. As Bernanke et al. (1999) show, unexpected movements in the price of capital can have considerable effects on entrepreneurs' financial conditions. This in turn affects borrowing conditions, which influences investment decisions. For instance, an unexpected drop in the return to capital reduces net worth of a leveraged entrepreneur by more than one-for-one. The external finance premium rises, demand for capital decreases, investment decreases and the price of capital falls, which reduces entrepreneurial net worth even further. Importantly, the credit market feeds back into the real economy. The countercyclical movement in the external finance premium serves to amplify the response of macroeconomic aggregates such as output and investment to shocks.

Extensions of the baseline financial accelerator model include the consideration of nominal instead of real financial contracts, e.g. Christensen and Dib (2008), the incorporation of the

[6] See Bernanke et al. (1999) for the details of the derivation and conditions that permit aggregation.

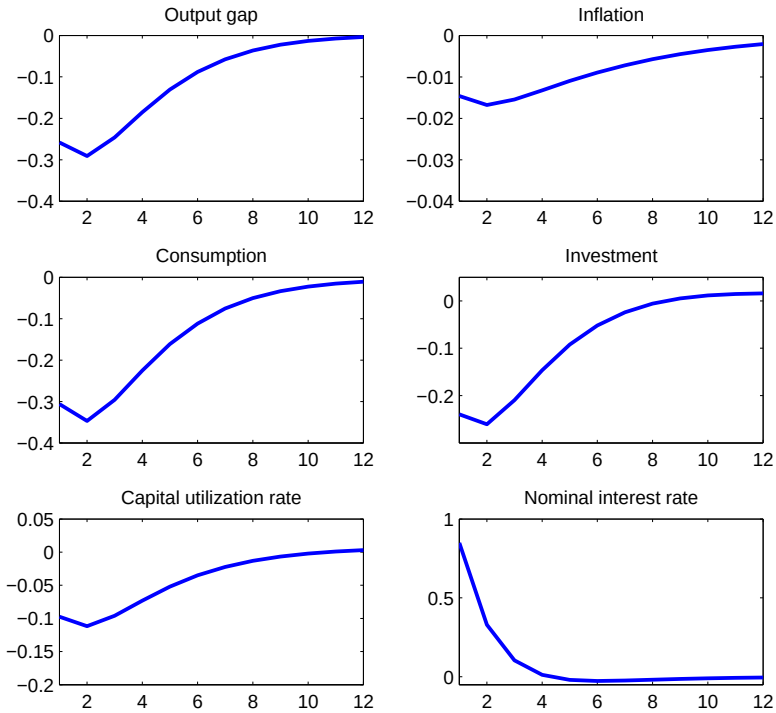
financial accelerator in a small open economy model, e.g. Gertler et al. (2007), or into a medium-scale New Keynesian model, e.g. De Graeve (2008). Meh and Moran (2010) consider the role of financial frictions in a DSGE model exhibiting an agency problem between banks and entrepreneurs, as in Bernanke et al. (1999), and between banks and their creditors, i.e. households. In this two-sided agency problem, not only entrepreneurs' wealth influences business cycle movements but also the capital position of banks. Iacoviello (2005) develops a financial accelerator model with financing constraints at the household level in form of collateral constraints tied to housing values, see also Kiyotaki and Moore (1997). Allowing for nominal debt contracts, he shows that the feedback of the financial market friction on the economy depends on the type of shock. Responses of output and consumer price inflation to a demand shock get amplified and propagated whereas the output response to supply shocks is mitigated. In the case of an unexpected increase in aggregate demand, goods prices and housing prices rise, which increases borrowers collateral value and reduces the real value of their debt. Since the borrowers in the model have a higher propensity to consume than the lenders, the net effect of this resource transfer from creditors to debtors is positive and serves to amplify the output response. On the contrary, a negative supply shock decreases inflation and therefore raises borrowers' real value of debt leading to a mitigated output response.

3.2.6 Model Dynamics

How are model dynamics affected by these additional frictions? In order to answer this question we can compare impulse responses in a prototypical medium-scale model to those in the small model. The DSGE model of the US economy estimated by Smets and Wouters (2007) incorporates capital with investment adjustment costs and variable capital utilization, habit formation in consumption, partial price indexation and sticky wages. Figure 4 displays the dynamic responses of several variables to a monetary policy shock in the Smets, Wouters model. To facilitate the comparison with the baseline model, monetary policy is assumed to follow rule (26). Indeed,

we observe hump-shaped impulse responses of consumption and investment. The responses of the output gap and inflation are more persistent than in the small-scale model, respectively. In fact, the effect of the policy shock on the considered variables persists beyond the effect on the nominal interest rate. Also, the inflation response is particularly subdued, being much smaller in magnitude than in our baseline model.

Figure 4: Monetary policy shock in the Smets, Wouters (2007) model



3.3 Using Structural Models for Policy Analysis: The Lucas Critique

Two key ingredients of the New Keynesian model that distinguish it from traditional Keynesian model paradigms are that (i) the decision rules of economic agents are based on optimization subject to constraints and (ii) agents' view of the future behavior of variables is formed under rational expectations. Importantly then, agents' decision rules inevitably vary with changes in

policy and this dependence becomes explicit in the system of reduced-form equations.

The Lucas critique, named after economist Robert E. Lucas and formulated in Lucas (1976), criticizes policy evaluation exercises based on estimated reduced-form relationships that - while potentially successful in short-term forecasting - fail to recognize this dependence. He argues that this kind of econometric model is unsuited for policy analysis because the estimated parameters are not policy-invariant.

To give an example, let us reconsider the hybrid New Keynesian Phillips curve introduced in section 3.3.3

$$\hat{\pi}_t = \frac{\gamma}{1+\beta\gamma}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\gamma}E_t\hat{\pi}_{t+1} + \frac{\kappa}{1+\beta\kappa}\hat{Y}_t^{gap} + \frac{1}{1+\beta\kappa}\varepsilon_t^{\pi}, \tag{41}$$

where we have added some zero mean, constant variance *iid* shock ε_t^{π}. For simplicity, let us assume that the policymaker can directly control the output gap $\hat{Y}_t^{gap}$ and monetary policy is described by the following rule

$$\hat{Y}_t^{gap} = \tau\hat{\pi}_t, \tag{42}$$

where $\tau < 0$. Substituting the policy rule into (41), we get

$$\hat{\pi}_t = \frac{\gamma}{1+\beta\gamma}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\gamma}E_t\hat{\pi}_{t+1} + \frac{\kappa\tau}{1+\beta\kappa}\hat{\pi}_t + \frac{1}{1+\beta\kappa}\varepsilon_t^{\pi}. \tag{43}$$

We guess the following solution for inflation

$$\hat{\pi}_t = a\hat{\pi}_{t-1} + b\varepsilon_t^{\pi}, \tag{44}$$

where the solution function parameters a and b are to be determined. This guess implies $E_t\hat{\pi}_{t+1} = a\hat{\pi}_t$. Substituting this expression into (43) and collecting terms, one can show that $a(\tau)$ is defined by the stable solution of the polynomial $a^2 - \frac{1}{\beta}(1+\beta\gamma-\kappa\tau)a + \frac{\gamma}{\beta} = 0$, and the second

parameter is defined by $b(\tau) = \frac{1}{1+\beta\gamma-\kappa\tau-\beta a(\tau)}$. As the notation underlines, the coefficients $a(\tau)$ and $b(\tau)$ depend on the policy rule parameter τ.

If we use the solution function to substitute out the expected inflation term in (41) and collect terms, we arrive at

$$\hat{\pi}_t = d_1\hat{\pi}_{t-1} + d_2\hat{Y}_t^{gap} + d_3\varepsilon_t^{\pi}, \tag{45}$$

where

$$d_1 = \frac{\gamma}{1+\beta[\gamma-a(\tau)]} \tag{46}$$

$$d_2 = \frac{\kappa}{1+\beta[\gamma-a(\tau)]} \tag{47}$$

$$d_3 = \frac{1}{1+\beta[\gamma-a(\tau)]}. \tag{48}$$

Equation (45) is reminiscent of a traditional Phillips curve. However, (45) is a reduced-form relationship, not a structural equation. Policy analysis based on empirical estimates of d_j, $j = 1,2,3$, that fails to recognize the parameter restrictions imposed by (46) - (48) is misleading for the parameters d_j, $j = 1,2,3$, are not invariant but will change in response to changes in the policy rule parameter τ. In contrast, when we use structural models to analyse alternative policies, we incorporate the parameter restrictions on the optimal decision rules of economic agents and therefore take into account that the behavioral reduced-form parameters will change systematically with changes in policy.

4 Methods for Model Solution and Estimation

4.1 Solving New-Keynesian Models

This subsection will inform the reader of commonly used approaches for solving New-Keynesian models: log-linearization, first- and second order numerical approximation and other methods for

approximating the solution of nonlinear models. References to available software will also be provided.

Consider a particular model m defined by the following system of nonlinear difference equations

$$E_t\left[\psi_m\left(x_{t+1}^m, x_t^m, v_t^m, \mu^m\right)\right] = 0, \tag{49}$$

where the x^ms are $n \times 1$ vectors of endogenous model variables. The model may include current values, lags and leads of endogenous variables. Such higher-order systems can be written as a first-order system by augmenting the x^m vectors accordingly. The model variables are functions of each other, of structural shocks, v_t^m, and of model parameters μ^m. In general, alternative approaches exist to solve the model in (49). We present three different solution procedures. The first consists of two steps, constructing a linear approximation of the system of nonlinear equations and getting the solution of the linear approximation. The second procedure presented is the extended path solution method which goes without prior linearization of the nonlinear system. Finally, the value function iteration procedure is presented in the context of a linear quadratic dynamic programming problem.

4.1.1 Linear Approximation

For the presentation of the linear approximation of system (49) we abstract from the stochastic model components

$$\psi_m\left(x_{t+1}^m, x_t^m, \mu^m\right) = 0. \tag{50}$$

A first-order Taylor series approximation around the non-stochastic steady state yields

$$0 \approx \psi\left(\overline{x}\right) + \frac{\partial \psi}{\partial x_t}\left(\overline{x}\right) \times \left(x_t - \overline{x}\right) + \frac{\partial \psi}{\partial x_{t+1}}\left(\overline{x}\right) \times \left(x_{t+1} - \overline{x}\right), \tag{51}$$

where we have slightly simplified notation in that the model index m and the dependence of $\psi(x_{t+1}, x_t)$ on the model parameters, μ, are not made explicit for the moment. The $n \times n$ matrix $\frac{\partial \psi}{\partial x_t}(\bar{x})$ represents the Jacobian of $\psi(x_{t+1}, x_t)$ with respect to x_t evaluated at the steady state $\bar{x}$. To obtain a log-linear approximation of the system as presented for the small-scale New Keynesian model in section 3.2, define

$$A \equiv \frac{\partial \psi}{\partial x_{t+1}}(\bar{x}) \times diag(\bar{x}), \qquad B \equiv -\frac{\partial \psi}{\partial x_t}(\bar{x}) \times diag(\bar{x}),$$

where $diag(\bar{x})$ is an $n \times n$ matrix with the elements of $\bar{x}$ on the main diagonal. The log-linear approximation of the nonlinear system then reads

$$A\hat{x}_{t+1} = B\hat{x}_t. \tag{52}$$

Following the notation in section 3.2, $\hat{x}_t$ denotes the percentage deviation of variable x_t from its steady state.

4.1.2 Solving a System of Linear Difference Equations

Various related ways exist to solve a system of linear (stochastic) difference equations such as (52). References include Blanchard and Kahn (1980), Uhlig (1999), Klein (2000), Sims (2001), and King and Watson (1998, 2002). Here we present the classical method of Blanchard and Kahn (1980). Their approach requires that the system of linear difference equations can be reformulated as

$$\begin{pmatrix} x_{1,t+1} \\ E_t x_{2,t+1} \end{pmatrix} = \Omega \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} + \Gamma f_t, \tag{53}$$

where the vector of possibly (log-)linearized endogenous variables x_t has been divided into an $n_1 \times 1$ vector of predetermined endogenous variables, $x_{1,t}$, and an $n_2 \times 1$ vector of free endogenous variables, $x_{2,t}$, with $n = n_1 + n_2$. The vector f_t contains exogenous forcing variables such as the exogenous technology variable in the small-scale model presented in section 3.2. The idea now is to transform the variables in a way that facilitates the solution of the system. First, we conduct a Jordan decomposition of matrix Ω

$$\Omega = \Lambda^{-1} J \Lambda, \tag{54}$$

Matrix J is called the Jordan canonical form of Ω and contains the eigenvalues of Ω on its diagonal. Partition J such that J_1 contains the eigenvalues that lie inside or on the unit circle and J_2 contains the eigenvalues that lie outside the unit circle

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix}. \tag{55}$$

The matrices Λ and Γ are then partitioned conformably

$$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}. \tag{56}$$

The existence of a unique solution of the model requires that the number of eigenvalues outside the unit circle equals the number of free endogenous variables, n_2.[7] If this condition is satisfied, we may now transform the vector $X_t = \begin{bmatrix} x'_{1,t} & x'_{2,t} \end{bmatrix}'$ according to

$$z_t \equiv \Lambda x_t. \tag{57}$$

[7] See Blanchard and Kahn (1980) for further details.

The linear system (53) can then be rewritten as

$$\begin{pmatrix} z_{1,t+1} \\ E_t z_{2,t+1} \end{pmatrix} = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} \begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} + \begin{pmatrix} \tilde{\Gamma}_1 \\ \tilde{\Gamma}_2 \end{pmatrix} f_t, \tag{58}$$

where $\tilde{\Gamma} \equiv \Lambda\Gamma$. The system of the transformed variables is diagonal.[8] Therefore, we can consider the subsystem related to the vector $z_{2,t}$ separately from the rest of the system. Solving for $z_{2,t}$, one obtains

$$z_{2,t} = J_2^{-1} E_t z_{2,t+1} - J_2^{-1} \tilde{\Gamma}_2 f_t. \tag{59}$$

If we iterate forward on this equation, we arrive at

$$z_{2,t} = -J_2^{-1} \sum_{j=0}^{\infty} \left(J_2^{-1}\right)^j \tilde{\Gamma}_2 E_t f_{t+j}, \tag{60}$$

where we have used the Law of Iterated Expectations. If we retransform the endogenous variables in (60), we obtain the solution functions for the free endogenous variables

$$x_{2,t} = -\Lambda_{22}^{-1}\Lambda_{21} x_{1,t} - \Lambda_{22}^{-1} J_2^{-1} \sum_{j=0}^{\infty} \left(J_2^{-1}\right)^j \tilde{\Gamma}_2 E_t f_{t+j}. \tag{61}$$

Finally, the solution functions for the predetermined endogenous variables can be obtained from substitution of (61) into (53)

$$x_{1,t+1} = \left(\Omega_{11} - \Omega_{12}\Lambda_{22}^{-1}\Lambda_{21}\right) x_{1,t} + \left(\Gamma_1 - \Omega_{12}\Lambda_{22}^{-1} J_2^{-1}\tilde{\Gamma}_2\right) f_t - \Omega_{12}\Lambda_{22}^{-1} J_2^{-1} \sum_{j=1}^{\infty} \left(J_2^{-1}\right)^j \tilde{\Gamma}_2 E_t f_{t+j}. \tag{62}$$

[8] If the number of distinct real eigenvalues is smaller than n so that Ω is not diagonalizable, then the superdiagonal elements corresponding to repeated eigenvalues are ones rather than zeros.

Here, matrix Ω has been partitioned conformably

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}. \tag{63}$$

4.1.3 The Extended Path Solution Method for Nonlinear Models

In the previous paragraphs we have discussed methods to solve linear or higher-order approximations of nonlinear models. The deterministic extended path (EP) method of Fair and Taylor (1983) can be used to obtain a numerical solution without prior linearization of the model equations and allows for the correct consideration of nonlinearities. The basic idea of the approach is to solve (49) under perfect foresight for $t = 1, ..., T$, each time setting future innovations to their expected value of zero. Of course, the perfect foresight assumption gives rise to an approximation error. The magnitude of this error depends on the degree of non-linearity in the model. Examinations of the accuracy of numerical solutions from the EP algorithm are documented in Gagnon (1990) and Taylor and Uhlig (1990) for the stochastic neoclassical growth model. The EP solution method proceeds as follows:

Step 1. We begin by choosing the initial length of a forecast horizon n and the initial conditions for the state variables. Then, set the innovations $v_{t+i} = 0$, for $i = 1, ..., n$.

Step 2. Next, guess values $\tilde{x}^j_{t+i}$ for x_{t+i}, $i = 0, ..., n$.

Step 3. Solve the perfect foresight model

$$\psi\left(\tilde{x}^j_{s+1}, x_s, v_s, \mu\right) = 0 \tag{64}$$

for x_s, $s = t, ..., t+n-1$. Standard nonlinear solution methods that can be used are described in e.g. Judd (1998) and Heer and Maussner (2005).

Step 4. Check whether the obtained values for x_{t+i} are within a selected tolerance criterion of the guesses $\tilde{x}^j_{t+i}$, for $i = 0, ..., n-1$. If not, return to step 2 and update your guesses $\tilde{x}^{j+1}_{t+i} = x_{t+i}$.

Otherwise, continue with step 5.

Step 5. Denote the values obtained for x_t by x_t^k, where k counts how many times step 5 has been reached. If $k = 1$, increase the forecast horizon n by one period and return to step 2. Otherwise, check whether the values x_t^k are within a selected tolerance criterion of the values x_t^{k-1}. If so, x_t^k is the numerical solution for x_t. If not, increase the forecast horizon n by one period and return to step 2.

Given a sequence of innovations $\{v_t\}_{t=1}^T$, a time series $\{x_t\}_{t=1}^T$ can then be generated using the described algorithm for $t = 1, 2, ..., T$.

4.1.4 Linear Quadratic Dynamic Programming: Value Function Iteration

Linear quadratic dynamic programming procedures are useful to solve problems that can be recasted as an optimization problem with quadratic objective function and linear constraints. Let us consider the following optimization problem

$$\max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t x_t' Q x_t \tag{65}$$

$$s.t. \quad s_{t+1} = As_t + Bu_t, \tag{66}$$

where (65) represents a quadratic objective function or a quadratic approximation to the nonlinear non-quadratic objective function. The $(n-m) \times 1$ vector u_t contains the control variables and the $m \times 1$ vector s_t the state variables. Let $x_t = [s_t' \quad u_t']'$ and Q a conformable $n \times n$ matrix. Maximization takes place subject to the linear constraints in (66). We guess that the value function is quadratic. The Bellman equation for the linear quadratic optimization problem is

$$s^T V s = \max_u \left\{ x^T Q x + \beta \left(s'\right)^T V s' \right\}, \tag{67}$$

subject to (66), where we have changed notation in line with the literature on dynamic programming. Notably, we ignore the time subscript. Leads are marked by a prime and a transpose is denoted by a T superscript. Matrix V of value function $s^T V s$ is negative semidefinite. Let us rewrite the Bellman equation as

$$\max_u \left\{ \begin{pmatrix} x \\ s' \end{pmatrix}^T \begin{pmatrix} Q & 0_{n\times m} \\ 0_{m\times n} & \beta V \end{pmatrix} \begin{pmatrix} x \\ s' \end{pmatrix} \right\}, \tag{68}$$

subject to (66). Define

$$C = \begin{pmatrix} A & B \end{pmatrix}, \qquad R = \begin{pmatrix} I_{n\times n} \\ C \end{pmatrix}. \tag{69}$$

Matrix $I_{n\times n}$ denotes the identity matrix of dimension n. We can then incorporate the linear constraints (66) as follows

$$\max_u \left\{ x^T R^T \begin{pmatrix} Q & 0_{n\times m} \\ 0_{m\times n} & \beta V \end{pmatrix} R x \right\}. \tag{70}$$

Define

$$W = R^T \begin{pmatrix} Q & 0_{n\times m} \\ 0_{m\times n} & \beta V \end{pmatrix} R, \tag{71}$$

and differentiate (70) with respect to the control vector u to get the first-order condition for the maximization problem. Solving the resulting condition for u, we obtain

$$u = -W^{-1}_{[m+1:n]\times[m+1:n]} W_{[m+1:n]\times[1:m]} s. \tag{72}$$

Let

$$F = -W^{-1}_{[m+1:n]\times[m+1:n]} W_{[m+1:n]\times[1:m]}. \tag{73}$$

If we substitute feedback rule (72) into (70), this leads to the following algebraic matrix Riccati equation

$$V = \begin{pmatrix} I_{m\times m} \\ F \end{pmatrix}^T R^T \begin{pmatrix} Q & 0_{n\times m} \\ 0_{m\times n} & \beta V \end{pmatrix} R \begin{pmatrix} I_{m\times m} \\ F \end{pmatrix}. \tag{74}$$

Value-function iteration involves iterations on

$$V^{(j+1)} = \begin{pmatrix} I_{m\times m} \\ F^{(j)} \end{pmatrix}^T R^T \begin{pmatrix} Q & 0_{n\times m} \\ 0_{m\times n} & \beta V^{(j)} \end{pmatrix} R \begin{pmatrix} I_{m\times m} \\ F^{(j)} \end{pmatrix}. \tag{75}$$

Under particular conditions, (75) will converge to a unique solution as $j \to \infty$, see e.g. Ljungqvist and Sargent (2004) and the references therein. In practice, we iterate until $|V^{(j+1)} - V^{(j)}| < \varepsilon$, for some small $\varepsilon > 0$. This provides us with an approximation of the solution for the policy function F.

4.1.5 Perturbation Methods for Higher-Order Approximations

Perturbation methods constitute a generalized approach to obtain linear or higher-order local approximations of the true model solution. As in the case of the linear approximation method presented above, the true solution is approximated in the neighborhood of some special point. The basic idea consists of finding a special case of the general problem for which the exact solution is known. The special case and its known solution is then used to compute approximate solutions of the general problem for points in the neighborhood of the special case with known solution. The following illustration of the approach is based on Judd (1998), chapter 13.

Consider the univariate problem

$$f(x,\sigma)=0, \tag{76}$$

where σ denotes some known parameter. It is assumed that for each value of σ (76) has a solution for x. Hence, (76) describes a system of equations in x for which the solution is unknown. Suppose however, that the equation can be solved for a specific value of σ, say $\sigma=0$. Let $x(\sigma)$ be an unknown function that satisfies $f(x(\sigma),\sigma)=0$. If f is differentiable, implicit differentiation of (76) leads to

$$f_x(x(\sigma),\sigma)x'(\sigma)+f_\sigma(x(\sigma),\sigma)=0. \tag{77}$$

For $\sigma=0$, we can find

$$x'(0)=-\frac{f_\sigma(x(0),0)}{f_x(x(0),0)}. \tag{78}$$

Furthermore, differentiating (77) leads to

$$\begin{aligned}&f_{xx}(x(\sigma),\sigma)\left(x'(\sigma)\right)^2+f_x(x(\sigma),\sigma)x''(\sigma)+2f_{x\sigma}(x(\sigma),\sigma)x'(\sigma)\\&+f_{\sigma\sigma}(x(\sigma),\sigma)=0,\end{aligned} \tag{79}$$

from which we obtain

$$x''(0)=-\frac{f_{xx}(x(0),0)(x'(0))^2+2f_{x,\sigma}(x(0),0)x'(0)+f_{\sigma\sigma}(x(0),0)}{f_x(x(0),0)}. \tag{80}$$

This allows us to compute a quadratic approximation to the solution using a second-order Taylor series expansion around $\sigma=0$

$$\hat{x}(\sigma)=x(0)+x'(0)\sigma+\frac{1}{2}x''(0)\sigma^2, \tag{81}$$

with $x'(0)$ and $x''(0)$ given by (78) and (80), respectively. Similarly, one can find higher-order approximations of $x(\sigma)$ using higher-order derivatives of $x(\sigma)$ obtained from further differentiation.

Judd (1998) contains a detailed treatment of perturbation methods. Algorithms for quadratic approximations and corresponding applications can be found in e.g. Collard and Juillard (2001), Kim et al. (2005) and Schmitt-Grohe and Uribe (2004). A ready-to-use computer implementation for Matlab programmed by Michel Juillard and his collaborators is available within DYNARE and may be downloaded from www.dynare.org.

4.2 Estimating New-Keynesian Models

Given the appealing characteristics of medium-size NK models for quantitative analysis, it is natural to take them to the data. This section provides an introduction to Bayesian estimation methods, which have become the dominant estimation technique in the DSGE literature. We illustrate the Bayesian estimation approach using the baseline New Keynesian model developed in section 3.2. The section concludes with an overview of prevailing challenges for model estimation.

4.2.1 Bayesian Methods

Bayesian methods allow us to estimate model parameters, to construct model forecasts and to conduct model comparisons. Here, we focus on the first application, model estimation. As Maximum Likelihood, Bayesian estimation takes a full information approach, that is, the econometrician's inference is based on the full range of empirical implications captured by the model. In the Bayesian context, a model is defined by a likelihood function and a prior. The likelihood function represents the data generating process, more specifically, it is the density of the data conditional on the structure of the model and conditional on the model parameters. Under the Maximum Likelihood approach the model parameters are interpreted as fixed and the observed

data represents a particular draw from the likelihood function. Parameter estimation then requires the maximization of the likelihood function. In contrast, the Bayesian approach interprets the parameters as random variables. Let μ represent model parameters as before, and let y be a sample of data observations to be explained by model M. Employing the rules of probability, the joint probability of (y, μ) conditional on model M is given by

$$p(y, \mu | M) = L(y | \mu, M)\, p(\mu | M), \tag{82}$$

or, alternatively

$$p(y, \mu | M) = p(\mu | y, M)\, p(y | M). \tag{83}$$

Here, $L(y|\mu, M)$ denotes the likelihood function. If we combine both equations in order to eliminate the joint probability terms, we receive Bayes' rule

$$p(\mu | y) = \frac{L(y|\mu)\, p(\mu)}{p(y)}, \tag{84}$$

where we keep in mind that we refer to a particular model M. The term $p(\mu|y)$ denotes the posterior distribution and $p(\mu)$ is the prior distribution. The posterior distribution allows us to make probabilistic statements with respect to the model parameters conditional on our model, the data and the prior. The posterior kernel reads

$$p(\mu | y) \propto L(y|\mu)\, p(\mu), \tag{85}$$

where the prior $p(\mu)$ contains any information about the parameters μ available to the econometrician not based on the sample of data observations. Thus, equation (85) may be interpreted as an updating rule that uses data observations to update the econometricians prior for the model parameters.

In practice, for most applications of interest the posterior distribution does not have a simple known form. Suppose, we are interested in a point estimate of the model parameters μ. One candidate would be the mean of the posterior distribution

$$E(\mu) = \int \mu p(\mu|y) d\mu. \tag{86}$$

Most of the time it is impossible to derive an analytical expression for this integral. Instead, one has to rely on computational methods. Popular in the context of DSGE models are Markov Chain Monte Carlo (MCMC) methods. The goal is to generate a Markov Chain $\{\mu_j\}$ that has the ergodic distribution $p(\mu|y)$, i.e. the posterior. Various algorithms exist to generate $\{\mu_j\}$. Here, we present the Metropolis-Hastings (MH) algorithm, a more detailed description can be found in e.g. Chib and Greenberg (1995). Since we cannot draw directly from the posterior distribution, we have to draw instead from some stand-in density, $q\left(\mu|\mu_{j-1}\right)$. Let the candidate draw from this stand-in density be denoted by μ^*. The candidate draw is accepted to be the next drawing μ_j with probability

$$\alpha\left(\mu^*|\mu_{j-1}\right) = \min\left\{1, \frac{p(\mu^*|y)\, q\left(\mu_{j-1}|\mu^*\right)}{p\left(\mu_{j-1}|y\right) q\left(\mu^*|\mu_{j-1}\right)}\right\}. \tag{87}$$

Importantly, it is sufficient to employ the posterior kernel (85). Let τ denote the draw from a uniform distribution over the interval $[0,1]$. The candidate draw μ^* is then accepted if $\alpha\left(\mu^*|\mu_{j-1}\right) > \tau$, otherwise we set $\mu_j = \mu_{j-1}$. This procedure is repeated J times. Note, that the acceptance probability will be relatively low if $q\left(\mu^*|\mu_{j-1}\right)$ is rather high and vice versa and it will be relatively high if $p(\mu^*|y)$ is rather high. The acceptance probability therefore adjusts for the fact that the stand-in density is different from the posterior density. If instead the stand-in density equals the posterior density, than the acceptance probability will always equal 1. To initialize the algorithm, one needs to specify a starting value μ_0. Often, numerical optimization is used to

determine the maximizer of the (log-)posterior kernel which is then used as a starting value.

Once we have a sequence of accepted draws $\{\mu_j\}$, we can approximate the mean of the posterior distribution as

$$\overline{\mu}_J = \frac{1}{J}\sum_{j=1}^{J}\mu_j. \tag{88}$$

More generally, let $f(\mu)$ be a function of the model parameters. The conditional expected value of this function can then be approximated by

$$\overline{f}_J = \frac{1}{J}\sum_{j=1}^{J} f(\mu_j). \tag{89}$$

Various instruments exist to assess the convergence of $\overline{f}_J$. A comparison of alternative convergence diagnostics can be found in Cowles and Carlin (1996). Here, instead we continue with the question how to choose the stand-in density $q(\mu|\mu_{j-1})$. A widely-used variant of the algorithm is the Random Walk Chain MH algorithm. The idea is to explore the neighborhood of an accepted draw. In this case, the candidate draw is generated from

$$\mu^* = \mu_{j-1} + \varepsilon, \tag{90}$$

where $\varepsilon \sim iid(0,\Sigma)$. This implies $q(\mu^*|\mu_{j-1}) = q(\mu_{j-1}|\mu^*)$, hence (87) simplifies to

$$\alpha(\mu^*|\mu_{j-1}) = \min\left\{1, \frac{p(\mu^*|y)}{p(\mu_{j-1}|y)}\right\}. \tag{91}$$

The choice of Σ is crucial for the efficiency of the sampler. A common approach is to use an estimate of the posterior covariance matrix scaled by some constant.

A second variant of the MH algorithm is called the Independence Chain MH algorithm. In this case, the stand-in density has the property $q(\mu|\mu_{j-1}) = q(\mu)$. In practice, it is important to select a stand-in density that has fatter tails than the posterior. For more detailed expositions of

the Bayesian approach in the context of DSGE models see e.g. An and Schorfheide (2007) and Del Negro and Schorfheide (2010).

4.2.2 Estimating a Small New Keynesian Model

To demonstrate the Bayesian methods just described, we estimate the small New Keynesian model developed in section 3.2. Prior to estimation, we consolidate the log-linearized model equations as follows. First, we combine the dynamic IS equation (20) with the aggregate resource constraint (24)

$$\hat{Y}_t = E_t\hat{Y}_{t+1} - E_t\Delta\hat{g}_{t+1} - \frac{1}{\tilde{\sigma}}\left(\hat{R}_t - E_t\hat{\pi}_{t+1}\right), \tag{92}$$

where $\frac{1}{\tilde{\sigma}} = \frac{1}{\sigma}C/Y$. One can also express this IS relation in terms of the output gap

$$\hat{Y}_t^{gap} = E_t\hat{Y}_{t+1}^{gap} - \frac{1}{\tilde{\sigma}}\left(\hat{R}_t - E_t\hat{\pi}_{t+1} - \hat{R}_t^{nat}\right). \tag{93}$$

Equation (93) contains a composite shock term

$$\hat{R}_t^{nat} = \tilde{\sigma}\left[E_t\left(\hat{Y}_{t+1}^{nat} - \hat{g}_{t+1}\right) - \left(\hat{Y}_t^{nat} - \hat{g}_t\right)\right], \tag{94}$$

which represents the natural rate of interest.[9] Second, we employ the NKPC

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \tilde{\kappa}\left(\tilde{\sigma} + \eta\right)\hat{Y}_t^{gap}, \tag{95}$$

where $\tilde{\kappa} = \frac{(1-\beta\theta)(1-\theta)}{\theta}$. Finally, we impose a slightly modified version of the interest rate rule (26)

$$\hat{R}_t = \tau_R\hat{R}_{t-1} + (1-\tau_R)\left(\tau_\pi\hat{\pi}_t + \tau_Y\hat{Y}_t^{gap}\right) + \varepsilon_t^R, \tag{96}$$

[9] Equation (93) reveals that if the policymaker keeps the real interest rate equal to the natural rate of interest, real output, Y_t, will always equal the natural level of output, Y_t^{nat}. From the NKPC (25), we observe that in this case inflation will also be fully stabilized.

where we have introduced an interest rate smoothing term. The vector of endogenous model variables then consists of the inflation rate, the nominal interest rate, output and the output gap. The exogenous variables g_t and a_t are again specified as AR(1) processes as in equation (28) and (29). The natural level of output and the natural interest rate are defined by (27) and (94), respectively.

We estimate the model using quarterly data on the US economy for 1966:I to 2007:II. The three data series employed comprise real per-capita quarter-to-quarter GDP growth (in percent), the average effective federal funds rate (in percent) and the quarter-to-quarter inflation rate based on the GDP Implicit Price Deflator (in percent). Real GDP per capita is constructed by dividing Real GDP by the level of the civilian non-institutional population over 16 from the FRED database of the St. Louis Fed. All three data series have been demeaned prior to estimation. A detailed description of the construction of the data series is provided in the appendix.

As discussed before, the Bayesian approach encompasses the specification of prior distributions. Depending on how informative or uninformative the priors are, they can add additional curvature to the likelihood function. Principally, economic theory can provide a valuable source to deduce one's priors. This is an important advantage of the use of structural models as opposed to reduced-form specifications as a priori information regarding the model's parameterization is typically much more readily available. In practice, specification of the prior distributions is also often based on pre-sample data or on evidence from micro data. We begin with a discussion of those parameters for which we impose fixed values in the estimation procedure. The subjective discount value is fixed at $\beta = 0.99$ which is in line with a steady state real interest rate of 4 percent in annualized terms. Furthermore, the data employed for estimation is unlikely to contain much information about the inverse of the elasticity of labor supply, η. We impose $\eta = 1$, a value often used in calibration exercises. We assume $\tilde{\kappa} \sim Gamma\,(0.08, 0.1)$ centered around a value in line with a Calvo parameter of $\theta = 0.75$ as in our baseline calibration. For the inverse of the intertemporal elasticity of substitution, we assume $\tilde{\sigma} \sim Gamma\,(1, 0.5)$, where the prior

mean is in line with a log-utility specification for consumption and zero steady state government spending. The priors for the policy rule parameters are loosely centered around values often used in the literature and are described by a Normal distribution in the case of the response coefficients to inflation and the output gap and by a Beta distribution for the response coefficient to the lagged interest rate. Relatively uninformative priors are used for the standard errors of the three exogenous innovations, each being described by an Inverse Gamma distribution. The complete summary of our priors is listed in Table 1. The estimation is conducted using the

Table 1: Prior distribution

Parameter	Density	Mean	Stand. deviation
$\tilde{\kappa}$	Gamma	0.08	0.1
$\tilde{\sigma}$	Gamma	1	0.5
τ_π	Gamma	1.5	0.25
τ_Y	Gamma	0.5	0.25
τ_R	Beta	0.5	0.2
ρ_g	Beta	0.8	0.2
ρ_A	Beta	0.8	0.2
σ_R	InvGamma	1	4
σ_g	InvGamma	1.5	4
σ_A	InvGamma	1.5	4

DYNARE software package.[10] The MH algorithm is used to generate 250,000 draws of which 33 percent are discarded as burn-in replications. We select the step size of the algorithm in line with an average acceptance ratio of around 35 percent. Table 2 shows the resulting mean and the 5 and 95 percentiles of the posterior distribution of the model parameters. The mean of the posterior estimates of $\tilde{\kappa}$ and $\tilde{\sigma}$ deviates quite substantially from the mean of the prior, which lies outside the reported confidence interval, respectively. The mean of the estimated monetary policy rule parameters reflect quite strong interest rate smoothing, a more than one-for-one long-run response to inflation and a positive response coefficient to the output gap, which is however

[10] We employ DYNARE version 4.1.3, see Juillard (1996) and Juillard (2001) for a general description of the software package.

Table 2: Posterior

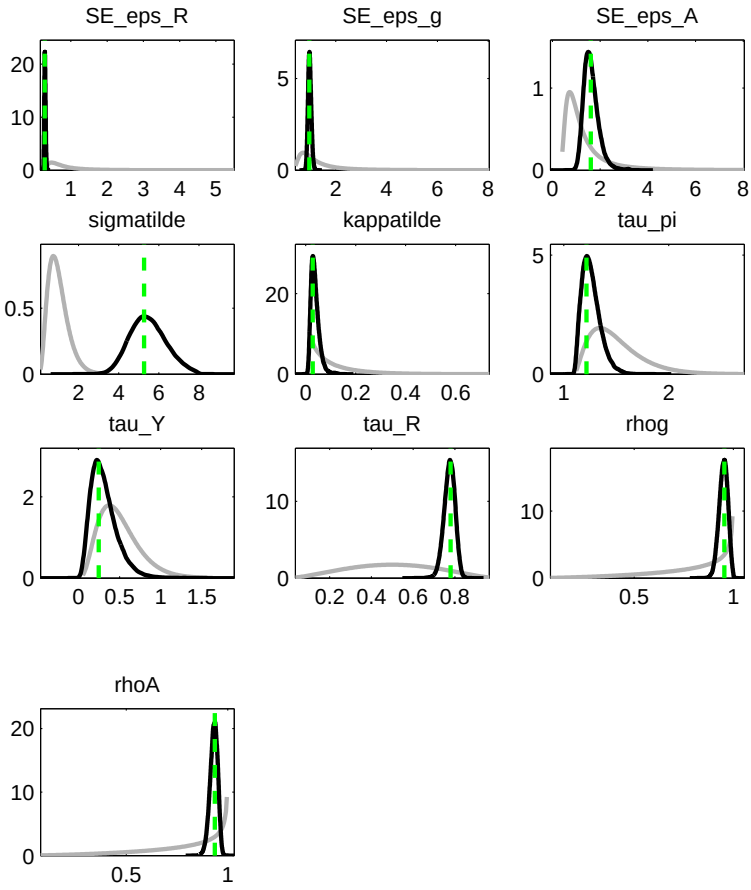
Parameter	Mean	5th percentile	95th percentile
$\tilde{\kappa}$	0.0375	0.0133	0.0615
$\tilde{\sigma}$	5.4574	3.9660	6.9018
τ_π	1.2607	1.1248	1.3881
τ_Y	0.3117	0.0666	0.5412
τ_R	0.7730	0.7311	0.8161
ρ_g	0.9492	0.9145	0.9861
ρ_A	0.9308	0.9003	0.9616
σ_R	0.2851	0.2541	0.3144
σ_g	0.9760	0.8749	1.0798
σ_A	1.6186	1.1403	2.0933

estimated less precisely than most of the other parameters. The two exogenous processes for the government spending shock and the technology shock are estimated to be very persistent with AR(1) coefficients of 0.95 and 0.93, respectively. Finally, the standard error of the government spending shock and the technology shock are estimated to be much larger than the standard error of the monetary policy shock. In summary, the data appears to be quite informative on the model parameters not fixed prior to the estimation procedure. The standard deviation of the posterior distribution is lower than the one of the prior distribution for all estimated parameters except of the inverse of the intertemporal elasticity of substitution, $\tilde{\sigma}$. Additional estimation output from DYNARE is shown in Figure 5, documenting the prior (solid gray line) and posterior (solid black line) distribution together with the posterior mode (dashed grey line) for all parameters. In each case, the posterior distribution has a single peak which is always close to the posterior mode.

4.3 Challenges for Model Estimation

Despite enormous progress in the development of computational methods that allow for increasing complexity of macroeconomic models, several issues remain. First and foremost, it is crucial to recognize that any DSGE model represents at best an approximation of the law of motion of

Figure 5: Prior and posterior distribution



the economy. A main source of misspecification are the cross-equation restrictions imposed by the DSGE models. Consider for instance the point estimate of the Phillips-curve parameter $\tilde{\kappa}$ in our small-scale model. From $\tilde{\kappa} = \frac{(1-\beta\theta)(1-\theta)}{\theta}$, a point estimate of $\tilde{\kappa} = 0.0375$ is consistent with $\beta = 0.99$ and $\theta = 0.8279$. The value for the Calvo parameter θ would imply that firms re-optimize their prices on average approximately every 6 quarters. Unfortunately, microeconomic evidence on price setting points towards much more frequent average changes in prices, see e.g. the survey by Klenow and Malin (2010). Many of the standard model devices in medium-scale New Keynesian models have been introduced to improve model fit. For instance, the incompatibility of our point estimate for $\tilde{\kappa}$ with micro evidence on price changes can be partly overcome by the introduction of firm-specific production inputs, see e.g. Eichenbaum and Fisher (2007). This does not mean that less abstraction should in general be preferred to more abstraction. As Kydland and Prescott (1996) put it

> To criticize or reject a model because it is an abstraction is foolish: all models are necessarily abstractions. A model environment must be selected based on the question being addressed.

However, when engaging in policy experiments, the parameter estimates obtained via Bayesian methods or Maximum Likelihood will only be invariant to changes in the policy regime if the model is not misspecified. Since computational power is limited, from a practical perspective the parameter estimates of models used for policy analysis should be approximately invariant to shifts in policy parameters. An example is provided by Cogley and Yagihashi (2010). They consider two New Keynesian models, one being the true data-generating model and the other representing an approximating model, and show that policy analysis based on the approximating model can still provide sensible recommendations for monetary policy.

A second problem that typically arises when estimating DSGE models is lack of identification of some of the structural parameters. Identification problems arise if different parameterizations of a model generate the same probability distribution. Since the elements of the coefficient ma-

trices of the model solution are usually highly nonlinear functions of the structural parameters, identification problems are of practical relevance. Canova and Sala (2009) categorize them as follows. First, observational equivalence occurs if the objective function, for instance the likelihood function, does not have a unique maximum given the mapping of structural parameters. Second, under-identification arises when structural parameters do not appear in the solution so that the objective function is independent of these parameters. Third, structural parameters may enter the objective function only proportionally, rendering them individually unidentifiable. This case is labeled partial identification by Canova and Sala (2009). Fourth, weak identification occurs when the objective function exhibits a unique maximum but is rather flat in some regions of the parameter space. In this case it is again difficult to identify the values of these parameters. A comparison of the prior and posterior distributions of model parameters can provide some first insights regarding identification problems. However, even if the posterior is shifted away from the prior, problems cannot be definitely ruled out as the shift might be due to some stability constraints. Canova and Sala (2009) recommend to consider a sequence of prior distributions with increasing variances to find evidence for identification problems. In the literature, often a subset of parameters is fixed in the estimation step, as we did with the stochastic discount factor and the inverse of the elasticity of labor supply when estimating the small New Keynesian model. However, in the case of partial identification of this parameter and some other parameter, estimates of the latter will depend on the calibration of the former. For a recent overview of techniques to determine conditions for identifiability, see Schorfheide (2011).

5 A New Approach to Model Comparison and Policy Evaluation

Today we have a large number of models at our disposal that aim to explain the behavior of the main aggregates of the world's economies. 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. Yet, systematic comparisons of the empirical implications of a large variety of available models are rare. One reason for the low number of model comparison projects surely has been that they required the input of many teams of researchers and multiple meetings to obtain a limited set of comparative findings. Examples include Bryant, Hooper and Mann (1993), Taylor (1999) and Hughes Hallett and Wallis (2004). This section presents a new coherent approach to macroeconomic model comparison developed by Wieland et al. (2011). This approach is based on a common computational platform that includes many well-known empirically estimated models and enables individual researchers to conduct model comparisons easily and on a large scale. The mechanics of the tool are illustrated with an exemplary comparison exercise.

The financial crisis and subsequent world recession has triggered much criticism of modern macroeconomics, including the New Keynesian approach and the stringent microeconomic foundations with representative agents and homogeneous expectations that are embodied in the recent DSGE models. Our comparative approach provides an avenue for setting different models against each other and for checking whether new modeling approaches perform equally well or better than existing approaches in fitting empirical benchmarks.

In section 4, we defined a model m by a system of nonlinear difference equations

$$E_t\left[\psi_m\left(x_{t+1}^m, x_t^m, v_t^m, \mu^m\right)\right] = 0. \tag{97}$$

The letter m is used to refer to a specific model that we would like to compare to some other model(s). The endogenous model variables, x_t^m, the structural shocks, v_t^m, and the model parameters, μ^m, need not be comparable across models. For any model m, we distinguish between two types of equations. Policy rules are denoted by g_m, and the other model equations and identities are captured by f_m. Similarly, we distinguish between policy shocks, η_t^m, and other shocks, ε_t^m,

with $v_t^m = (\eta_t^m, \varepsilon_t^m)$, and between policy-rule parameters, γ^m, and the rest of the model parameters, β^m, with $\mu^m = (\gamma^m, \beta^m)$. Thus, (97) may be rewritten as

$$E_t\left[g_m\left(x_{t+1}^m, x_t^m, \eta_t^m, \gamma^m\right)\right] = 0 \tag{98}$$

$$E_t\left[f_m\left(x_{t+1}^m, x_t^m, \varepsilon_t^m, \beta^m\right)\right] = 0. \tag{99}$$

As before, the model may include lags and further leads of endogenous variables. In this case, x_t^m has to be augmented accordingly. The innovations v_t^m have a zero mean and a constant covariance matrix Σ^m which can be partitioned as follows

$$\Sigma^m = \begin{pmatrix} \Sigma_\eta^m & \Sigma_{\eta,\varepsilon}^m \\ \Sigma_{\eta,\varepsilon}^m & \Sigma_\varepsilon^m \end{pmatrix}, \tag{100}$$

where we distinguish the covariance matrices of policy shocks, Σ_η^m, and other economic shocks, Σ_ε^m. Unless policy shocks are correlated with the other shocks, we set $\Sigma_{\eta,\varepsilon}^m = 0$.

Usually, a comparison of the empirical implications of two models, say $m = (1,2)$, cannot be based directly on either (97) or (98) and (99) as distinct models often feature different notation and definitions. Instead, first it is necessary to augment all models with a set of common, comparable variables, parameters, shocks and equations as summarized in Table 3. Note, that

Table 3: Common comparable variables, shocks, equations and parameters

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

these common objects are not indexed by m, this is because they are defined coherently for all models that are included in a comparison exercise. The common policy rules, $g(.)$, replace the

model-specific policy rules $g_m(.)$ so that model implications can be compared conditional on a particular policy specification. The model augmentation step also involves the definition of a set of additional model-specific equations. These equations define the common variables, z_t, in terms of the model-specific variables, x_t^m, and are denoted by $h_m(.)$. Importantly, the notation and definitions for all the other equations, variables, parameters and shocks is preserved. Consequently, an augmented model m consists of three components: First, the original set of model equations, $f_m(.)$, determining endogenous variables, excluding the model-specific policy rules, $g_m(.)$. Second, a set of new model-specific equations, $h_m(.)$, that define the common variables in terms of original model-specific endogenous variables with parameters θ^m. Third, the common policy rules $g(.)$ expressed in terms of common variables, z_t, common policy shocks η_t, and common policy rule parameters γ. These three components comprise the system of difference equations defining the augmented model m

$$E_t\left[f_m\left(x_{t+1}^m, x_t^m, \varepsilon_t^m, \beta^m\right)\right] = 0. \tag{101}$$

$$E_t\left[h_m\left(z_t, x_{t+1}^m, x_t^m, \theta^m\right)\right] = 0 \tag{102}$$

$$E_t\left[g\left(z_{t+1}, z_t, \eta_t, \gamma\right)\right] = 0. \tag{103}$$

Models augmented in this manner can be used for comparison exercises. For instance, one can compare the implications of a policy rule across models by constructing certain metrics based on the dynamics of the common endogenous variables in those alternative models. Before we consider such objects for comparison, we illustrate the model augmentation step with an example.

Let us suppose the vector of common comparable variables, z_t, consists of six variables, the annualized quarterly money market rate, i_t^z, discretionary government spending expressed as a share in GDP, g_t^z, the year-on-year rate of inflation, π_t^z, the annualized quarter-to-quarter rate of inflation, p_t^z, quarterly real GDP, y_t^z, and the output gap, q_t^z. The notation and the definitions for

the common variables is also summarized in Table 4.

Table 4: 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 common monetary and fiscal policy rules are assumed to be subject to random innovations $\eta_t = \begin{bmatrix} \eta_t^i & \eta_t^g \end{bmatrix}'$. We are now ready to introduce the small-scale New Keynesian model that we have estimated in the previous section into our comparison framework. Let us denote this model by $m = 1$. The original model and the augmented model are presented in Table 11 in the appendix. In the augmented version of the model the original equations, $f_1(.)$, are unchanged except for the original policy rule which is replaced by some common rule $g(.)$. The additional model-specific equations $h_1\left(., \theta^1\right)$ define the common comparable variables in terms of the model-specific variables. In the case of our simple model, this augmentation step might seem rather trivial, but it is nevertheless necessary to avoid comparing apples and oranges.

In order to illustrate how to conduct a model comparison, we need at least one more model. Here, we take the business cycle model from Ireland (2004) presented in Table 12 in the appendix and abbreviated henceforth by $m = 2$. It represents a stylized New Keynesian model with real money balance effects and quadratic adjustment costs in price setting. The model is estimated via Maximum Likelihood using quarterly US data from 1980:1 to 2001:3. It consists of a dynamic IS equation, a New Keynesian Phillips curve, a demand equation for real money balances, $\hat{m}_t$, an interest-rate rule, and AR(1) specifications for the three non-policy shocks. All variables are log-linearized around the non-stochastic steady state. The version of the model implemented in

the Macroeconomic Model Database and used here refers to the constrained estimate in Ireland (2004) with the parameter ω_2 fixed to a value of 0.25. Economically, this refers to the case of non-separable utility from consumption and real money balances. The augmented model is shown in the lower part of Table 12. The original model does not feature an output gap and a government spending shock. While the natural level of output and therefore the output gap can be derived in the common variables block based on the microfoundations of the model, one has to remain silent with regard to the common variable g_t^z.

The augmented models can then be solved conditional on some common policy rules using the methods outlined in section 4. The solution function for an augmented model m can be written as

$$\begin{pmatrix} z_t \\ x_t^m \end{pmatrix} = K_m(\gamma) \begin{pmatrix} z_{t-1} \\ x_{t-1}^m \end{pmatrix} + D_m(\gamma) \begin{pmatrix} \eta_t \\ \varepsilon_t^m \end{pmatrix}. \tag{104}$$

The reduced-form matrices $K_m(\gamma)$ and $D_m(\gamma)$ are functions of the common policy parameters, γ, and the model-specific non-policy parameters, β^m. Having obtained the solution functions for models $m = 1, 2$, one can construct objects for comparison based on the common comparable variables z_t. For instance, we might compare dynamic responses of the common variables to a common policy shock across models. The impulse response functions of an augmented model m to a common monetary policy shock, η_t^i, in period $t+j$, $j \geq 0$, are defined as

$$IR_{t+j}^m(\gamma, \eta_t^i) = \begin{pmatrix} E_t\left(z_{t+j}|z_{t-1}, x_{t-1}^m, \eta_t^i\right) - E_t\left(z_{t+j}|z_{t-1}, x_{t-1}^m\right) \\ E_t\left(x_{t+j}^m|z_{t-1}, x_{t-1}^m, \eta_t^i\right) - E_t\left(x_{t+j}^m|z_{t-1}, x_{t-1}^m\right) \end{pmatrix}. \tag{105}$$

For some $\eta_t^i > 0$, $IR_{t+j}^1(\gamma, \eta_t^i)$ represents the impulse responses of the estimated small-scale New Keynesian model to a positive monetary policy shock and similarly $IR_{t+j}^2(\gamma, \eta_t^i)$ for the Ireland (2004) model. Comparisons of impulse responses from different augmented models should be limited to common variables and common shocks. Such comparisons can provide interesting

insights into the monetary policy transmission channels of the included models. One may evaluate the distance between several models for a given characteristic of the model dynamics by defining some metric s. For instance, one might consider the difference in the cumulative sum of the response of some common variable to a monetary policy shock

$$s(\gamma, z) = \sum_{j=0}^{\infty} \left[IR^1_{t+j}\left(\gamma, \eta^i_t, z\right) - IR^2_{t+j}\left(\gamma, \eta^i_t, z\right) \right], \tag{106}$$

where the index z serves as a caveat that we can compare only the impulse responses of the common variables. To give an example, we impose the following common monetary policy rule from Smets and Wouters (2007) written in terms of common variables

$$i^z_t = 0.81 i^z_{t-1} + 0.39 p^z_t + 0.97 q^z_t - 0.90 q^z_{t-1} + \eta^i_t, \tag{107}$$

and compare impulse responses of the output gap, q^z, to a unitary monetary policy shock. For our two models $m = 1, 2$ we get $s(\gamma, q^z) = 1.14$. Further examples and a more detailed presentation of the formal approach to model comparison are provided in Wieland et al. (2011). We have also built a computational platform that includes many well-known empirically estimated models and allows individual researchers to conduct model comparisons and quantitative analysis of stabilization policies easily and on a large scale.[11] The Appendix contains a complete list of models currently available in the model archive.

[11] The Macroeconomic Model Data Base software can be downloaded from www.macromodelbase.com.

6 Policy Evaluation and Policy Robustness under Model Uncertainty

Lucas (1980) suggests the use of economic models as laboratories to evaluate alternative policies. Purely empirical approaches are unsuited for this purpose and real world experimentation is usually not an option. In a situation where no model's structure is considered completely satisfactory from a theoretical perspective and many competing models fit the historical data of key aggregates reasonably well, it is not advisable to base real-world policy recommendations on a single preferred model. Instead, researchers should help policy makers to develop robust policies. This strategy for policy advise is nicely formulated by McCallum (1999), who 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."

In this section, we demonstrate how one can use the above-mentioned model comparison platform for investigating the robustness of policies under model uncertainty. Objectives for policy makers may be formulated as Bayesian or Minimax decision criteria (see Kuester and Wieland (2010)).

The first question we want to raise is what recommendations can be drawn from economic models for the design of monetary policy. In particular, using a range of models rather than a single model will allow us to identify policy implications that are common to all or a large share of models as opposed to model-specific recommendations. The focus of our analysis is on simple interest rate rules, that is, the one-period nominal interest rate represents the monetary policy instrument and this instrument responds systematically to a small number of variables. For instance, the interest rate rules specified in (26) and (96) belong both to the group of simple rules. In principle, economic models can be used to evaluate the performance of a broad range of rules that may differ in their functional forms, in their instruments, and in the set of variables for

the instrument to respond to. However, there are several reasons to focus on simple interest rate rules. First, there is a broad consensus in the literature that an interest rate instrument is superior to for instance a money supply instrument, see e.g. the survey of Taylor and Williams (2010). Second, the gains from increasing either the number of leads and lags or the set of variables to which the instrument responds are very small, as has been shown by e.g. Levin et al. (1999). Furthermore, when model uncertainty is explicitly taken into account, it turns out that simple monetary policy rules are more robust than more complicated fully optimal rules.

We begin with the consideration of rules that specify the interest rate as a linear function of two variables, the year-on-year inflation rate and the output gap. Using the definition of common variables introduced above, we then have

$$i_t^z = \tau_\pi \pi_t^z + \tau_q q_t^z, \tag{108}$$

where i_t^z is the annualized quarterly nominal interest rate, π_t^z denotes the annual inflation rate and q_t^z is the output gap. The question which measure of inflation is most appropriate to be incorporated into the policy rule is of key interest. Evidence in Levin et al. (1999) and Levin et al. (2003) suggests that responding to a smoothed inflation rate like the year-on-year rate is more desirable in terms of stabilization performance than the one-period inflation rate even if the latter is the one that enters the central bank's policy objective. Interest rate rule (108) can be extended to include an interest rate smoothing term

$$i_t^z = \tau_i i_{t-1}^z + \tau_\pi \pi_t^z + \tau_q q_t^z. \tag{109}$$

Empirical estimates of policy rules often hint at quite substantive interest rate smoothing of monetary policy in practice. Economic models can be used to assess whether this is also a characteristic of optimal rules.

We assume that the central bank's objective is represented by the following quadratic loss func-

tion

$$L = Var\left(\pi^z\right) + \lambda_q Var\left(q^z\right) + \lambda_{\Delta i} Var\left(\Delta i^z\right), \tag{110}$$

where $Var\left(.\right)$ denotes the unconditional variance operator. The parameters $\lambda_q \geq 0$ and $\lambda_{\Delta i} \geq 0$ represent the central bank's preferences for reducing the variability of the output gap and of changes in the nominal interest rate relative to inflation variability. The form as well as the targets entering the loss function have been widely used in previous model-based analyses of monetary policy rules, especially in the context of policy experiments based on competing models. In particular, for $\lambda_{\Delta i} = 0$, (110) corresponds to a second-order approximation of household utility in the standard New Keynesian model evaluated under the unconditional expectations operator, see e.g. Woodford (2003). With real money balances entering the utility function, also the level of the nominal interest rate appears in such a linear-quadratic approximation.

We base our analysis on a set of state-of-the-art medium-scale New Keynesian models of the US economy with a focus on financial frictions. The model of De Graeve (2008) incorporates a financial accelerator a la Bernanke et al. (1999) as discussed in section 3.2.4 into a New Keynesian model with nominal price and wage frictions, habit formation in consumption, price and wage indexation, variable capital and investment adjustment costs. The model, henceforth referred to as the DG model, has been estimated on quarterly US data from 1954:1 to 2004:4 using Bayesian techniques. We also consider a second variant of the model, labeled DGnoff, in which we shut down the financial accelerator mechanism. Third, we consider the model of Iacoviello (2005) (IAC) which incorporates housing into a New Keynesian model. A financial accelerator arises in the IAC model due to the presence of borrowing constraints. The value of housing serves as collateral for firms and for part of the households. Unlike in Bernanke et al. (1999), debt contracts are denominated in nominal terms. The model also features variable capital and adjustment costs for housing and for capital. Model estimation has been conducted using calibration and impulse response function matching based on quarterly US data from 1974:1 to

2003:2. Finally, the model of Rabanal (2007) (RB) is similar to the model of De Graeve (2008) in that it exhibits nominal rigidities in price and wage setting, price and wage indexation, variable capital utilization, investment adjustment costs and habit formation in consumption. Unlike in the DG model, there is no financial accelerator present in the RB model. However, part of the firms have to pay their wage bill prior to their sales receipts, which forces them to borrow from a financial intermediary. A cost channel of monetary policy transmission arises where changes in the nominal interest rate have a direct effect on firms' marginal costs. The model has been estimated on quarterly US data from 1959:1 to 2004:4 using Bayesian techniques. Table 5 reports the optimized response coefficients of the two policy rules (108) and (109) for each of the four economic models. The left panel reports results for the case of equal weights on the variance

Table 5: Characteristics of optimized simple rules

Rule / Model	$\lambda_q = 0, \lambda_{\Delta i} = 1$			$\lambda_q = 1, \lambda_{\Delta i} = 1$		
	τ_i	τ_π	τ_q	τ_i	τ_π	τ_q
2 parameters						
DG		1.45	0.70		1.46	1.60
DGnoff		1.82	0.47		1.39	1.99
IAC		2.12	0.07		1.31	0.49
RB		2.43	0.27		2.44	1.20
3 parameters						
DG	1.00	0.28	0.01	0.90	0.46	0.68
DGnoff	1.01	0.22	0.01	0.98	0.16	0.87
IAC	1.14	0.75	-0.01	1.49	0.52	0.59
RB	1.05	0.66	0.12	1.07	0.54	0.56

Note: Optimized response coefficients for the two-parameter rule $i_t^z = \tau_\pi \pi_t^z + \tau_q q_t^z$ and the three-parameter rule $i_t^z = \tau_i i_{t-1}^z + \tau_\pi \pi_t^z + \tau_q q_t^z$ are reported. The parameters λ_q and $\lambda_{\Delta i}$ denote the weight on the variance of the output gap and on the variance of the change in the nominal interest rate in the central bank's loss function, respectively.

of inflation and the change in the nominal interest rate and no weight on the output gap variance in the central bank loss function (110) whereas the right panel shows results for the case of equal weights on the variances of all three variables. All two-parameter rules satisfy the so-called Taylor principle, named after economist John B. Taylor, which postulates that the nominal

interest rate should respond more than one-for-one to changes in inflation, $\tau_\pi > 1$. In many New Keynesian models the Taylor principle is a necessary and often also sufficient condition for the existence of a unique rational expectations equilibrium. A second characteristic common to all optimized two-parameter rules is a strictly positive response coefficient to the output gap. Thus, the interest rate is increased in order to dampen aggregate demand whenever it exceeds the natural level of output and vice verca. Despite these similarities of optimized simple rules across models, the response coefficients can differ quite substantially in terms of magnitude. The results for the 3-parameter rules reveal that in most of the cases a response coefficient to the lagged interest rate near unity is desirable. Rules with $\tau_i > 1$ are called superinertial rules. Rules that respond to the lagged interest rate introduce history dependence because future policy actions will depend in part on current economic conditions. It should be noted, however, that superinertial rules can lead to instability in backward-looking models. In order to compare the stabilization performance of the two-parameter and three-parameter rules, Table 6 reports the increase in absolute loss in terms of the implied inflation variability premium (IIP) when moving from the three-parameter to the two-parameter rule. The IIP, proposed by Kuester and Wieland (2010), translates the increase in the absolute loss into an equivalent increase in the standard deviation of inflation.[12] For instance, the entry in the third row and second column of Table 6 tells us that if the central bank weights inflation and output gap volatility equally and the IAC model represents the economy, then employing the optimized two-parameter rule instead of the three-parameter rule will result in an increase in the central bank's loss equivalent to an increase in the standard deviation of inflation of 1.43 percentage points. An increase in the standard deviation of inflation of this magnitude is of economic relevance. While the numbers are somewhat smaller for the other models, the results suggest that including the lagged interest rate in the policy rule leads to non-negligible improvements in the central bank's stabilization performance. A natural question

[12] In the literature, the analysis of policy rules is often based on the percentage increase in the central bank's loss instead of the absolute increase. This measure of relative policy performance can lead to misleading signals, as demonstrated by Kuester and Wieland (2010) and Taylor and Wieland (2011).

Table 6: Loss increase when reducing the number of parameters in the rule: IIP

Model	$\lambda_q = 0, \lambda_{\Delta i} = 1$	$\lambda_q = 1, \lambda_{\Delta i} = 1$
DG	0.41	0.49
DGnoff	0.51	0.75
IAC	0.58	1.43
RB	0.96	0.76

The increase in absolute loss when monetary policy follows the optimized two-parameter rule instead of the optimized three-parameter rule is reported in terms of the implied inflation variability premium. The IIP corresponds to the required increase in the standard deviation of the annual inflation rate that would imply an equivalent increase in absolute loss.

then is whether one should raise the number of parameters in the policy rule further. Taylor and Wieland (2011) consider 4-parameter rules that include the lagged output gap in addition to the three variables already included in our 3-parameter rules and find little gains from expanding the number of variables. Also, one might ask whether policy rules should respond to expectations of future inflation and the output gap instead of contemporaneous realizations. Levin et al. (2003) show that the benefits of such policy rules are in general limited. Furthermore, if the forecast-horizon is too long, the models become highly susceptible to equilibrium indeterminacy under this class of rules.

So far, we have implicitly assumed that the central bank knows the true model of the economy with certainty. What if the reference model of the central bank used for policy analysis is not a good representation of the economy and one of the other models constitutes a more valid representation? To address this question we evaluate each rule optimized for a particular model in the competing models. Table 7 reports the loss increase in terms of IIPs when a rule optimized for model Y is evaluated in the distinct model X relative to the performance of the model-consistent optimal rule in X of the same class. We find that for the two-parameter rules the IIPs are generally much lower when the central bank does not care about output gap stabilization, $\lambda_q = 0$. In this case only one experiment results in an IIP of more than 1 percentage points, namely when the rule optimized for the DG model is evaluated in the IAC model. In contrast, with equal weights

Table 7: Robustness of optimized simple rules: IIP

Model	2-parameter rules		3-parameter rules	
	$\lambda_q = 0, \lambda_{\Delta i} = 1$	$\lambda_q = 1, \lambda_{\Delta i} = 1$	$\lambda_q = 0, \lambda_{\Delta i} = 1$	$\lambda_q = 1, \lambda_{\Delta i} = 1$
		DG rule		
DGnoff	0.04	0.05	0.00	0.09
IAC	1.37	1.36	0.08	0.27
RB	0.82	0.85	0.19	0.23
		DGnoff rule		
DG	0.06	0.06	0.00	0.18
IAC	0.93	1.80	0.10	0.34
RB	0.24	1.20	0.28	1.38
		IAC rule		
DG	0.52	1.04	0.19	0.71
DGnoff	0.14	1.19	0.19	0.88
RB	0.05	0.92	0.31	0.24
		RB rule		
DG	0.21	0.28	0.11	0.12
DGnoff	0.09	0.31	0.10	0.21
IAC	0.56	0.93	0.22	0.13

The increase in absolute loss when the policy rule optimized for one model is evaluated in the other models is reported in terms of the implied inflation variability premium. The IIP corresponds to the required increase in the standard deviation of the annual inflation rate that would imply an equivalent increase in absolute loss.

on inflation and output gap variation, 5 of the 12 IIPs exceed one percentage point. However, for both loss functions we find that the rule optimized for the DG model with the financial friction shut down leads to almost no reduction in the performance of monetary policy in the full DG model relative to the model-consistent optimal rule. The same holds true when evaluating the rule optimized for the full DG model in the model variant without financial frictions. The IIP never exceeds 0.06 percentage points. Turning to the three-parameter rules, the corresponding IIPs turn out to be somewhat lower than under the two-parameter rules for the majority of experiments, though there are exceptions. Two three-parameter rules, the DG rule and the RB rule, turn

out to be relatively robust to model uncertainty, leading only to small increases in the absolute loss when implemented in competing models. Interestingly, even though the rule optimized for the DG model without financial frictions performs again almost as well as the model-consistent optimal rule in the full DG model, its performance differs from the latter rule when evaluated in other models. For instance, the DGnoff rule performs much worse in the RB model than the DG rule, when the central bank cares about output gap stabilization.

We can also use our set of models to evaluate rules optimized to other models considered in the literature. For instance, Taylor and Wieland (2011) offer a comparison of the Taylor (1993a) model (TAY), the Christiano et al. (2005) model (CEE) and the Smets and Wouters (2007) model (SW) of the US economy. Table 8 reports IIPs for optimized rules from Taylor and Wieland (2011). The documented IIPs confirm our previous finding that rules optimized for one model may exhibit a weak performance in other models. However, the results reported in Table 8 offer an additional insight. Optimized model-specific rules may even lead to desastrous outcomes when the true model turns out to be different from the reference model. In the case of the two-parameter rules the SW rule leads to equilibrium indeterminacy in the DGnoff model allowing for a multiplicity of possible equilibria. In the case of the 3-parameter rules and equal weights on all three target variables in the loss function, the rule optimized for the CEE model induces IIPs of tremendous size (between 12 and 18 percentage points) in the four models.

Clearly, these results suggest that we should take model uncertainty explicitly into account. So far, the design of policy rules has relied on a particular reference model. We now adopt a Bayesian perspective on the design of robust rules, following the approach proposed by Levin et al. (1999, 2003) and Brock et al. (2003). Under this approach the policy rule parameters are optimized by minimizing a weighted average of losses across models

$$\sum_{m\in M}\omega_m L_m = \sum_{m\in M}\omega_m\left[Var\left(\pi_m^z\right)+\lambda_q Var\left(q_m^z\right)+\lambda_{\Delta i}Var\left(\Delta i_m^z\right)\right], \tag{111}$$

Table 8: Robustness of rules from Taylor, Wieland (2009): IIP

Model	2-parameter rules		3-parameter rules	
	$\lambda_q = 0, \lambda_{\Delta i} = 1$	$\lambda_q = 1, \lambda_{\Delta i} = 1$	$\lambda_q = 0, \lambda_{\Delta i} = 1$	$\lambda_q = 1, \lambda_{\Delta i} = 1$
		TAY rule		
DG	0.26	0.90	0.07	0.06
DGnoff	0.12	0.95	0.07	0.06
IAC	0.42	0.36	0.26	0.18
RB	0.01	0.31	0.11	0.55
		CEE rule		
DG	0.79	0.92	0.16	14.19
DGnoff	0.71	1.01	0.17	18.18
IAC	0.87	0.24	0.05	12.00
RB	0.60	0.20	0.17	18.45
		SW rule		
DG	0.53	1.29	0.04	1.12
DGnoff	IND	1.45	0.04	1.65
IAC	0.01	0.26	0.02	0.40
RB	0.07	0.24	0.12	0.77

The increase in absolute loss under optimized rules from Taylor and Wieland (2011) relative to the model-specific optimized rule is reported in terms of the implied inflation variability premium. The IIP corresponds to the required increase in the standard deviation of the annual inflation rate that would imply an equivalent increase in absolute loss. IND refers to the case where a rule induces equilibrium indeterminacy in a particular model.

where $M = \{DG, DGnoff, IAC, RB\}$ and the parameters ω_m denote the weights on the models, representing the central bank's priors. Here, we consider flat priors, $\omega_m = \frac{1}{4}$ for all m, so that (111) is similar to model averaging. Table 9 reports the optimized response coefficients of the 2-parameter and 3-parameter model averaging rules. Results are shown for the case where annual inflation, the output gap and the change in the nominal interest rate enter the central bank's loss function with the same weight. For the two-parameter rule, the optimized response coefficient to inflation lies close to 2, being somewhat smaller than the optimal response coefficient in the RB model but larger than the optimal parameter value in the remaining three models.

Table 9: Characteristics of optimized model-averaging rules

Rule	τ_i	τ_π	τ_q
2 parameters		2.06	0.91
3 parameters	1.05	0.49	0.60

Note: Optimized response coefficients for the two-parameter rule $i_t^z = \tau_\pi \pi_t^z + \tau_q q_t^z$ and the three-parameter rule $i_t^z = \tau_i i_{t-1}^z + \tau_\pi \pi_t^z + \tau_q q_t^z$ in case of equal weights on the variance of the output gap and the variance on the change in the nominal interest rate, $\lambda_q = 1, \lambda_{\Delta i} = 1$, are reported.

The optimized three-parameter rule has a response coefficient to the lagged interest rate near unity and (short-run) response coefficients to inflation and the output gap which are smaller than under the two-parameter rule. Comparing the Bayesian three-parameter rule with the corresponding model-specific optimal rules in Table 5 it turns out that the parameter values of the model-averaging rule are relatively close to the rule that is optimal in the RB model. This is not too surprising given that our earlier experiments showed that the RB rule performs quite well in the three competing models. This brings us to the question how the model-averaging rules perform in the individual models in comparison to the model-specific optimal rules. Table 10 reports the corresponding IIPs for the two classes of rules. Beginning with the two-parameter rule we

Table 10: Optimized model-averaging rules: IIP

Model	2-parameter rule	3-parameter rule
DG	0.35	0.07
DGnoff	0.43	0.15
IAC	0.55	0.14
RB	0.05	0.02

Note: The increase in absolute loss in each model under a rule optimized by averaging over all models relative to the model-specific optimized rule of the same class is reported in terms of the implied inflation variability premium. The IIP corresponds to the required increase in the standard deviation of the annual inflation rate that would imply an equivalent increase in absolute loss.

observe that while the IIPs are strictly positive, i.e. the model averaging rule is not optimal based on any single model, the increase in loss relative to the optimal rule remains relatively modest.

The largest IIP of 0.55 percentage points is much smaller than the maximum IIP under any of the model-specific rules (see Table 7). That the costs of achieving robustness in terms of stabilization performance are small becomes even more eminent in case of the Bayesian 3-parameter rule. The maximum IIP amounts to 0.15 percentage points. The smallest IIP arises in case of the RB model due to the close similarity between the parameter values of the Bayesian rule and the RB rule.

Our results emphasize that rules fine-tuned to a particular model may lead to poor or even disastrous outcomes in alternative models. Policy experiments should therefore be based on a broad range of alternative models. Here, we have focused on models with financial frictions but in general it is desirable to consider much more diversity in terms of modeling approaches when we want to make sure that policy recommendations are robust.

7 Outlook: Unanswered Questions and Future Research

In the course of the recent financial crisis commentators have criticized the DSGE approach to macroeconomic modeling. DSGE models are blamed for their failing to predict the financial turmoil and its implications for the real economy. Indeed, the type of models that were used prior to the crisis did in general not include a realistic treatment of the banking sector and the involved macroeconomic risks. In response to this criticism, proponents of the DSGE approach have started to enhance the existing benchmark models to allow for a more detailed treatment of financial market frictions and the role of the banking sector in causing and propagating shocks to the macroeconomy. Wieland (2011) provides an overview of recent contributions along these lines. A common feature of these contributions is the attempt to analyse the link between the financial sector and the rest of the economy. Issues addressed include implications of financial market imperfections for business cycle dynamics, the role of financial frictions for optimal monetary policy and the effectiveness of unconventional policy measures. In the future, it will

be important to evaluate whether DSGE models extended along these lines will result in an improved forecasting performance in particular with regard to the evolution of major crises.

A second line of attack is concerned with the way expectations of economic agents are formed in standard DSGE models. The benchmark DSGE approach imposes rational expectations, as we do in our treatment throughout this chapter. However, critics question the reasonability of model inhabitants that calculate expectations under complete knowledge about the structural features of the economy. Given the important role of expectations in macroeconomics, it is indeed crucial to test the sensitivity of business cycle dynamics and policy recommendations with respect to alternative assumptions about agents' expectation formation. A range of different approaches for modeling less-than-fully rational expectations have been proposed in the economic literature that could be incorporated in a systematic comparison of alternative model paradigms. An alternative to rational expectations that has been known for a long time is adaptive learning. Under adaptive learning, inhabitants of the model economy re-estimate simple reduced-form specifications of model variables to form expectations. The parameter estimates of these reduced-form specifications are updated each period once new data becomes available. In this sense the economic agent acts like an econometrician. Examinations of the implications of adaptive learning for policy performance and business cycle dynamics include Orphanides and Williams (2006), Slobodyan and Wouters (2008) and Wieland (2009).

While adaptive learning relaxes the assumption of rational expectations, expectations usually remain homogeneous across agents. In contrast, empirical investigations of professional forecasts point to substantial diversity, see e.g. Wieland and Wolters (2011). Theoretical analyses show that such heterogeneity of expectations can amplify economic fluctuations and might therefore have important implications for macroeconomic policy. Recent contributions include Branch and McGough (2011), Branch and Evans (2011), Chiarella et al. (2007), De Grauwe (2011), Kurz et al. (2005) and Kurz (2009). In particular, an explicit treatment of belief diversity allows one to decompose the sources of economic volatility between heterogeneity in expectations and

other structural shocks. Again it would be interesting to compare the forecasting performance of models enhanced along this dimension to existing benchmark DSGE models with homogeneous expectations. The Macroeconomic Model Data Base provides a powerful platform to facilitate such comparisons of competing model paradigms. It would therefore be very interesting to incorporate models that feature some of the characteristics outlined above into the model archive.

8 Appendices

8.1 Data sources and treatment

The data series used for estimation of the small New Keynesian model in section 4.2 are defined as follows

$$\begin{aligned} YGR &= (1-L)\,ln\,(GDPC1/CNP16OV)*100 \\ INFL &= GDPDEF \\ INT &= FEDFUNDS/4, \end{aligned}$$

where L denotes the lag-operator. The original data sources are

- GDPC1: Real Gross Domestic Product (Billions of Chained 2005 Dollars, Seasonally Adjusted Annual Rate)
  Source: U.S. Department of Commerce - Bureau of Economic Analysis (via St. Louis Fed FRED database)

- CNP16OV: Civilian Non-institutional Population (Thousands, Not Seasonally Adjusted, Average of Monthly Data)
  Source: U.S. Department of Labor - Bureau of Labor Statistics (via St. Louis Fed FRED database)

- GDPDEF: Gross Domestic Product Implicit Price Deflator (Percent Change, Seasonally Adjusted)

  Source: U.S. Department of Commerce - Bureau of Economic Analysis (via St. Louis Fed FRED database)

- FEDFUNDS: Effective Federal Funds Rate (Percent, Averages of Daily Figures)

  Source: Board of Governors of the Federal Reserve System (via St. Louis Fed FRED database)

The data variables are related to the model variables via the following measurement equations:

$$
\begin{aligned}
YGR_t &= \hat{Y}_t - \hat{Y}_{t-1} + mean\,(YGR_t) \\
INFL_t &= \hat{\pi}_t + mean\,(INFL_t) \\
INT_t &= \hat{R}_t + mean\,(INT_t)
\end{aligned}
$$

8.2 Augmented Models

Table 11: The small-scale New Keynesian model

Description	Equations and definitions
Original Model	
variables	$x_t^1 = \left[\hat{R}_t \quad \hat{\pi}_t \quad \hat{Y}_t \quad \hat{Y}_t^{gap} \quad \hat{Y}_t^{nat} \quad \hat{R}_t^{nat} \quad \hat{g}_t \quad \hat{A}_t\right]'$
shocks	$\varepsilon_t^1 = \varepsilon_t^A, \quad \eta_t^1 = \left[\varepsilon_t^R \quad \varepsilon_t^g\right]'$
parameters	$\beta_1 = [\beta \quad \kappa \quad \tilde{\sigma} \quad \eta \quad \rho_g \quad \rho_A]', \quad \gamma_1 = [\tau_R \quad \tau_\pi \quad \tau_Y]'$
model equations	
$g_1(.)$	$\hat{R}_t = \tau_R \hat{R}_{t-1} + (1-\tau_R)\left(\tau_\pi \hat{\pi}_t + \tau_Y \hat{Y}_t^{gap}\right) + \varepsilon_t^R$
$f_1(.)$	$\hat{Y}_t^{gap} = E_t \hat{Y}_{t+1}^{gap} - \frac{1}{\tilde{\sigma}}\left(\hat{R}_t - E_t \hat{\pi}_{t+1} - \hat{R}_t^{nat}\right)$
	$\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \kappa \hat{Y}_t^{gap}$
	$\hat{Y}_t = \hat{Y}_t^{gap} + \hat{Y}_t^{nat}$
	$\hat{Y}_t^{nat} = \frac{1}{\tilde{\sigma}+\eta}\left[(1+\eta)\hat{A}_t + \tilde{\sigma}\hat{g}_t\right]$
	$\hat{R}_t^{nat} = \tilde{\sigma}\left[E_t\left(\hat{Y}_{t+1}^{nat} - \hat{g}_{t+1}\right) - \left(\hat{Y}_t^{nat} - \hat{g}_t\right)\right]$
	$\hat{g}_t = \rho_g \hat{g}_{t-1} + \varepsilon_t^g$
	$\hat{A}_t = \rho_A \hat{A}_{t-1} + \varepsilon_t^A$
Augmented Model	
z_t, η_t	$z_t = [i_t^z \quad g_t^z \quad \pi_t^z \quad p_t^z \quad y_t^z \quad q_t^z]', \quad \eta_t = \left[\eta_t^i \quad \eta_t^g\right]'$
$\gamma, g(.)$	$i_t^z = 0.81 i_{t-1}^z + 0.39 p_t^z + 0.97 q_t^z - 0.90 q_{t-1}^z + \eta_t^i$
$f_1(.)$	as defined above in original model
$h_1(z_t, E_t x_{t+1}^1, x_t^1, \theta^1)$	$i_t^z = 4\hat{R}_t$
	$g_t^z = \hat{g}_t$
	$\pi_t^z = \hat{\pi}_t + \hat{\pi}_{t-1} + \hat{\pi}_{t-2} + \hat{\pi}_{t-3}$
	$p_t^z = 4\hat{\pi}_t$
	$y_t^z = \hat{Y}_t$
	$q_t^z = \hat{Y}_t^{gap}$

Table 12: The New Keynesian model of Ireland (2004)

Description	Equations and definitions
Original Model	
variables	$x_t^2 = [\hat{r}_t \quad \hat{\pi}_t \quad \hat{y}_t \quad \hat{m}_t \quad \hat{a}_t \quad \hat{e}_t \quad \hat{z}_t]'$
shocks	$\varepsilon_t^2 = [\varepsilon_{at} \quad \varepsilon_{et} \quad \varepsilon_{zt}]', \quad \eta_t^2 = \varepsilon_{rt}$
parameters	$\beta_2 = [\omega_1 \quad \omega_2 \quad \gamma_1 \quad \gamma_2 \quad \gamma_3 \quad \psi \quad \pi \quad r \quad \rho_a \quad \rho_e \quad \rho_z]'$, $\gamma_2 = [\rho_r \quad \rho_\pi \quad \rho_y]'$
model equations	
$g_2(.)$	$\hat{r}_t = \rho_r \hat{r}_{t-1} + \rho_y \hat{y}_{t-1} + \rho_\pi \hat{\pi}_{t-1} + \varepsilon_{rt}$
$f_2(.)$	$\hat{y}_t = E_t \hat{y}_{t+1} - \omega_1 (\hat{r}_t - E_t \hat{\pi}_{t+1})$
	$+\omega_2 [(\hat{m}_t - \hat{e}_t) - (E_t \hat{m}_{t+1} - E_t \hat{e}_{t+1})]$
	$+\omega_1 (\hat{a}_t - E_t \hat{a}_{t+1})$
	$\hat{\pi}_t = \frac{\pi}{r} E_t \hat{\pi}_{t+1} + \psi \left[\frac{1}{\omega_1} \hat{y}_t - \frac{\omega_2}{\omega_1} (\hat{m}_t - \hat{e}_t) - \hat{z}_t \right]$
	$\hat{m}_t = \gamma_1 \hat{y}_t - \gamma_2 \hat{r}_t + \gamma_3 \hat{e}_t$
	$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{at}$
	$\hat{e}_t = \rho_e \hat{e}_{t-1} + \varepsilon_{et}$
	$\hat{z}_t = \rho_z \hat{z}_{t-1} + \varepsilon_{zt}$
Augmented Model	
z_t, η_t	$z_t = [i_t^z \quad \pi_t^z \quad p_t^z \quad y_t^z \quad q_t^z]', \quad \eta_t = \eta_t^i$
$\gamma, g(.)$	$i_t^z = 0.81 i_{t-1}^z + 0.39 p_t^z + 0.97 q_t^z - 0.90 q_{t-1}^z + \eta_t^i$
$f_2(.)$	as defined above in original model
$h_2(z_t, E_t x_{t+1}^2, x_t^2, \theta^2)$	$i_t^z = 4\hat{r}_t$
	$\pi_t^z = \hat{\pi}_t + \hat{\pi}_{t-1} + \hat{\pi}_{t-2} + \hat{\pi}_{t-3}$
	$4 p_t^z = \hat{\pi}_t$
	$y_t^z = \hat{y}_t$
	$q_t^z = \hat{y}_t - \frac{1}{1-\omega_2\gamma_1} [\omega_1 \hat{z}_t + \omega_2 (\gamma_3 - 1) \hat{e}_t]$

8.3 A Database of Macroeconomic Models

The following two tables summarize the models currently available in the data base.

MODELS AVAILABLE IN THE MACROECONOMIC MODEL DATABASE (VERSION 1.2)

References

Adolfson, Malin, Michael Andersson, Jesper Lindé, Matthias Villani and Anders Vredin. 2007a. Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks. *International Journal of Central Banking* 3(4).

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani. 2007b. Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics* 72.

Altig, David, Lawrence Christiano, Martin Eichenbaum and Jesper Lindé. 2005. Firm-specific capital, nominal rigidities and the business cycle. 2005. CEPR Discussion Papers 4858.

An, Sungbae and Frank Schorfheide. 2007. Bayesian Analysis of DSGE Models. *Econometric Reviews* 26(2-4).

Andres, Javier, David Lopez-Salido and Javier Valles. 2006. Money in an Estimated Business Cycle Model of the Euro Area. *The Economic Journal* 116.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist. 1999. The Financial Accelerator in a Quantitative Business Cycle Framework. In J. Taylor and M. Woodford (editors), *Handbook of Macroeconomics*, North-Holland, Amsterdam.

Blanchard, Olivier and Stanley Fischer. 1989. *Lectures on Macroeconomics*. The MIT Press.

Blanchard, Olivier and Charles M. Kahn. 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica* 48(5).

Branch, William A. and George W. Evans. 2011. Monetary policy with heterogeneous expectations. *Economic Theory* 47(2-3): 365-393.

Branch, William A. and Bruce McGough. 2011. Business cycle amplification with heterogeneous expectations. *Economic Theory* 47(2-3): 395-421.

Brayton, Flint, and Thomas Laubach. 2008. Documentation of Linearized FRB/US model. Manuscript.

Brock, William, Steven Durlauf and Kenneth West. 2003. Policy Evaluation in Uncertain Environments. Brookings Papers on Economic Activity 1:2003.

Bryant, Ralph, Dale Henderson, Gerald Holtham, Peter Hooper and Steven Symansky (editors). 1988. *Empirical Macroeconomics for Interdependent Economies*, Washington, D.C.: The Brookings Institution.

Bryant, Ralph, David Currie, Jacob Frenkel, Paul Masson and Richard Portes (editors). *Macroeconomic Policies in an Interdependent World*, Washington, D.C.: The Brookings Institution.

Bryant, Ralph, Peter Hooper and Catherine Mann (editors). 1993. *Evaluating Policy Regimes: New Research in Empirical Macroeconomics*, Washington, D.C.: The Brookings Institution.

Calvo, Guillermo. 1983. Staggered Prices in a Utility Maximizing Framework. *Journal of Monetary Economics* 12:383-398.

Canova, Fabio and Luca Sala. 2009. Back to square one: Identification issues in DSGE models. *Journal of Monetary Economics* 56.

Carabenciov, Ioan, Igor Ermolaev, Charles Freedman, Michel Juillard, Ondra Kamenik, Dmitry Korshunov and Douglas Laxton. 2008. A Small Quarterly Projection Model of the US Economy. IMF Working Paper 08/278.

Chiarella, Carl, Roberto Dieci and Xue-Zhong He. 2007. Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework. *Journal of Economic Behavior and Organization* 62: 408-427.

Chib, Siddhartha and Edward Greenberg. 1995. Understanding the Metropolis-Hastings Algorithm. *The American Statistician* 49(4).

Christensen, Ian and Ali Dib. 2008. The financial accelerator in an estimated New Keynesian model. *Review of Economic Dynamics* 11.

Christiano, Lawrence. 2002. Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics* 20(1-2).

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. 1999. Monetary Policy Shocks: What Have we Learned and to What End? in Taylor and Woodford (eds.), *Handbook of Macroeconomics* volume 1A, 65-148.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1).

Christoffel, Kai, and Keith Kuester. 2008. Resuscitating the wage channel in models with unemployment fluctuations. *Journal of Monetary Economics* 55.

Christoffel, Kai, Keith Kuester, and Tobias Linzert. 2009. The role of labor markets for euro area monetary policy. *European Economic Review* 53.

Clarida, Richard, Jordi Galí, and Mark Gertler. 1999. The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature* 37(4).

Clarida, Richard, Jordi Galí, and Mark Gertler. 2002. A Simple Framework for International Monetary Policy Analysis. *Journal of Monetary Economics* 49.

Coenen, Guenter, Peter McAdam, and Roland Straub. 2008. Tax reform and labour-market performance in the euro area: A simulation-based analysis using the New Area-Wide Model. *Journal of Economic Dynamics and Control* 32(8).

Coenen, Guenter, Athanasios Orphanides and Volker Wieland. Price Stability and Monetary Policy Effectiveness when Nominal Interest Rates are Bounded at Zero. *Berkeley Electronic Press: Advances in Macroeconomics* 4(1).

Coenen, Guenter and Volker Wieland. 2002. Inflation Dynamics and International Linkages: A Model of the United States, the Euro Area and Japan. ECB Working Paper Series 181.

Coenen, Guenter and Volker Wieland. 2005. A Small Estimated Euro Area Model with Rational Expectations and Nominal Rigidities. *European Economic Review* 49.

Cogan, John, Tobias Cwik, John Taylor, and Volker Wieland. 2010. New Keynesian versus Old Keynesian Government Spending Multipliers. *Journal of Economic Dynamics and Control* 34.

Cogley, Timothy and Takeshi Yagihashi. 2010. Are DSGE Approximating Models Invariant to Shifts in Policy? *The B.E. Journal of Macroeconomics (Contributions)* 10(1)

Collard, Fabrice and Michel Juillard. 2001. Accuracy of stochastic perturbation methods: the case of asset pricing models. *Journal of Economic Dynamics and Control* 25.

Cowles, Mary K. and Bradley P. Carlin. 1996. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association* 91.

De Graeve, Ferre. 2008. The external finance premium and the macroeconomy: US post-WWII evidence. *Journal of Economic Dynamics and Control* 32.

De Grauwe, Paul. 2011. Animal spirits and monetary policy. *Economic Theory* 47(2-3): 423-457.

Del Negro, Marco and Frank Schorfheide. 2010. Bayesian Macroeconometrics. Forthcoming in: J. Geweke, G. Koop, H. van Dijk (editors), *The Oxford Handbook of Bayesian Econometrics*, Oxford: Oxford University Press.

Dieppe, Alistair, Keith Kuester and Peter McAdam. 2005. Optimal Monetary Policy Rules for the Euro Area: An Analysis using the Area Wide Model. *Journal of Common Market Studies* 43(3).

Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte. 2007. Natural Rate Measures in an Estimated DSGE Model of the U.S. Economy. *Journal of Economic Dynamics and Control* 32.

Eichenbaum, Martin, and Jonas Fisher. 2007. Estimating the frequency of price re-optimization in Calvo-style models. *Journal of Monetary Economics* 54.

Erceg, Christopher, Luca Guerrieri, and Christopher Gust. 2008. Trade Adjustment and the Composition of Trade. *Journal of Economic Dynamics and Control* 32.

Erceg, Christopher, Dale Henderson, and Andrew Levin. 2000. Optimal Monetary Policy with Staggered Wage and Price Contracts. *Journal of Monetary Economics*46: 281-313.

Fagan, Gabriel, Jerome Henry and Ricardo Mestre. 2005. An Area-Wide Model for the Euro Area. *Economic Modelling* 22.

Fair, Ray C. and John B. Taylor. 1983. Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models. *Econometrica* 51(4).

Fischer, Stanley. 1977. Long-term contracts, rational expectations and the optimal money supply rule. *Journal of Political Economy* 85(1): 191-205.

Fuhrer Jeffrey and George Moore. 1995a. Inflation Persistence. *The Quarterly Journal of Economics* 110(1).

Fuhrer Jeffrey and George Moore. 1995b. Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output. *The American Economic Review* 85(1).

Funke, Michael, Michael Paetz, and Ernest Pytlarczyk. 2011. Stock market wealth effects in an estimated DSGE model for Hong Kong. *Economic Modelling* 28.

Gagnon, Joseph E. 1990. Solving the Stochastic Growth Model by Deterministic Extended Path. *Journal of Business and Economic Statistics* 8(1).

Galí, Jordi. 1999. Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *The American Economic Review* 89(1).

Galí, Jordi. 2008. *Monetary Policy, Inflation, and the Business Cycle*. Princeton University Press.

Galí, Jordi and Mark Gertler.1999. Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics* 44.

Galí, Jordi and Mark Gertler. 2007. Macroeconomic Modeling for Monetary Policy Evaluation. *Journal of Economic Perspectives* 21 (4).

Galí, Jordi, David López-Salido and Javier Vallés. 2007. Understanding the effects of government spending on consumption. *Journal of the European Economic Association* 5(1).

Galí Jordi and Tommaso Monacelli. 2005 Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies* 72.

Galí, Jordi and Pau Rabanal. 2004. Technology Shocks and Aggregate Fluctuations: How Well Does the Real Business Cycle Model Fit Postwar U.S. Data? *NBER Macroeconomics Annual*

Gelain, Paolo. 2010. The external finance premium in the euro area: A dynamic stochastic general equilibrium analysis. *North American Journal of Economics and Finance* 21.

Gertler, Mark, Simon Gilchrist, and Fabio Natalucci. 2007. External Constraints on Monetary Policy and the Financial Accelerator. *Journal of Money, Credit, and Banking* 39(2-3).

Gertler, Mark, and Peter Karadi. 2009. A model of unconventional monetary policy. Working Paper.

Goodfriend, Marvin and Robert King. 1997. The New Neoclassical Synthesis and the Role of Monetary Policy. *NBER Macroeconomics Annual* 231-282.

Gouvea, Solange, Andre Minella, Rafael Santos, and Nelson Souza-Sobrinho. 2008. Samba: Stochastic Analytical Model with a Bayesian Approach. Working Paper.

Hansen, Lars. P. and Thomas J. Sargent. 1980. Formulating and Estimating Dynamic Linear Rational Expectations Models. *Journal of Economic Dynamics and Control* 2: 7-46.

Heer, Burkhard and Alfred Maussner. 2005. *Dynamic General Equilibrium Modelling*, Berlin, Springer.

Hughes Hallett, Andrew and Kenneth Wallis (editors). 2004. *EMU Macroeconomic Model Comparison exercise for the Euroconference 7-8 June 2002*, Economic Modelling 21(5).

Iacoviello, Matteo. House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. 2005. *The American Economic Review* 95.

Ireland, Peter N. 2004. Money's Role in the Monetary Business Cycle. *Journal of Money, Credit and Banking* 36(6).

Ireland, Peter N. 2011. A New Keynesian Perspective on the Great Recession. *Journal of Money, Credit and Banking* 43(1).

Judd, Kenneth L. 1998. *Numerical Methods in Economics*, Cambridge, MA, London: MIT Press.

Juillard, Michel. 1996. 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, Michel. 2001. A Program for the Simulation of Rational Expectation Models. *Computing in Economics and Finance* 213.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg and Christopher A. Sims. 2005. Calculating and Using Second Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models. Manuscript.

Kimball, Miles. 1995. The Quantitative Analytics of the Basic Neomonetarist Model. *Journal of Money, Credit and Banking* 27(4).

King, Robert G. and Mark Watson. 1998. The Solution of Linear Difference Systems under Rational Expectations. *International Economic Review* 39(4).

King, Robert G. and Mark Watson. 2002. System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations. *Computational Economics* 20(1-2).

Kiyotaki, Nobuhiro and John Moore. 1997. Credit Cylces. *Journal of Political Economy* 105(2).

Klein, Lawrence (editor). 1991. *Comparative Performance of U.S. Econometric Models*, Oxford, Eng.: Oxford University Press.

Klein, Paul. 2000. Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control* 24(10).

Klenow, Peter J. and Benjamin A. Malin (2010). Microeconomic Evidence on Price-Setting. Prepared for the *Handbook of Monetary Economics*.

Kuester, Keith, and Volker Wieland. 2010. Insurance Policies for Monetary Policy in the Euro Area. *Journal of the European Economic Association* 8.

Kurz, Mordecai, Hehui Jin and Maurizio Motolese. 2005. The role of expectations in economic fluctuations and the efficacy of monetary policy. *Journal of Economic Dynamics and Control* 29: 2017-2065.

Kurz, Mordecai. 2009. Rational diverse beliefs and market volatility. In T. Hens and K. Schenk-Hoppe (editors): *Handbook of Financial Markets: Dynamics and Evolution*. North Holland.

Kydland, Finn E. and Edward C. Prescott. 1982. Time to Build and Aggregate Fluctuations. *E*conometrica 50 (6): 13451370.

Kydland, Finn E. and Edward C. Prescott. 1996. The Computational Experiment: An Econometric Tool. *Journal of Economic Perspectives* 10(1).

Laxton, Douglas and Paolo Pesenti. 2003. Monetary rule for small, open, emerging economies. *Journal of Monetary Economics* 50.

Levin, Andrew T., Volker Wieland, and John C. Williams. 1999. Robustness of Simple Policy Rules under Model Uncertainty. In J. Taylor (editor), *Monetary Policy Rules*, Chicago: University of Chicago Press.

Levin, Andrew T., Volker Wieland, and John C. Williams. 2003. The Performance of Forecast-Based Monetary Policy Rules under Model Uncertainty. *The American Economic Review* 93:622-645.

Levin, Andrew T., and John C. Williams. 2003. Robust Monetary Policy with Competing Reference Models. *Journal of Monetary Economics* 50:945-975.

Ljungqvist, Lars and Thomas J. Sargent. 2004. *Recursive Macroeconomic Theory*, Cambridge, MA, London: MIT Press.

Lubik, Thomas, and Frank Schorfheide. 2007. Do central banks respond to exchange rate movements? A structural investigation. *Journal of Monetary Economics* 54.

Lucas, Robert E. 1976. Econometric Policy Evaluation: A Critique. Carnegie-Rochester Conference Series on Public Policy 1.

Lucas, Robert E. 1980. Methods and Problems in Business Cycle Theory. *Journal of Money, Credit, and Banking* 12:696-715.

Mankiw, Gregory, and Ricardo Reis. 2007. Sticky information in general equilibrium. *Journal of the European Economic Association* 5(2-3).

Mankiw, Gregory, and David Romer (editors). 1991a. *New Keynesian Economics*, vol. 1, *Imperfect Competition and Sticky Prices*, MIT Press.

Mankiw, Gregory, and David Romer (editors). 1991b. *New Keynesian Economics*, vol. 2, *Coordination Failures and Real Rigidities*, MIT Press.

McCallum, Bennet. 1988. Robustness properties of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy* 29.

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

McCallum, Bennet and Edward Nelson. 1999. Performance of operational policy rules in an estimated semi-classical structural model. In J. Taylor (editor), *Monetary Policy Rules*, Chicago: University of Chicago Press.

Medina, Juan Pablo, and Claudio Soto. 2007. The Chilean Business Cycles through the Lens of a Stochastic General Equilibrium Model. Central Bank of Chile Working Papers 457.

Meh, Cesaire and Kevin Moran. 2010. The role of bank capital in the propagation of shocks. *Journal of Economic Dynamics and Control* 34.

Murchison, Stephen, and Andrew Rennison. 2006. ToTEM: The Bank of Canada's New Quarterly Projection Model. Bank of Canada Technical Report No. 97.

Orphanides, Athanasios. 2003. The quest for prosperity without inflation. *Journal of Monetary Economics* 50.

Orphanides, Athanasios and Volker Wieland. 1998. Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero. Finance and Economics Discussion Series 98-35, Board of Governors of the Federal Reserve System.

Orphanides, Athanasios and John C. Williams. 2006. Monetary Policy with Imperfect Knowledge. *Journal of the European Economic Association* 4(2-3): 366-375.

Phelps, Edmund and John B. Taylor. 1977. Stabilizing powers of monetary policy with rational expectations, *Journal of Political Economy* 85: 163-190.

Rabanal, Pau. 2007. Does inflation increase after a monetary policy tightening? Answers based on an estimated DSGE model. *Journal of Economic Dynamics and Control* 31.

Rabanal, Pau. 2009. Inflation Differentials between Spain and the EMU: A DSGE Perspective. *Journal of Money, Credit and Banking* 41(6).

Ratto, Marco, Werner Roeger, and Jan in 't Veld. QUEST III: An Estimated Open-economy DSGE Model of the Euro Area with Fiscal and Monetary Policy. *Economic Modelling* 26(1).

Ravenna, Federico, and Carl Walsh. 2006. Optimal monetary policy with the cost channel. *Journal of Monetary Economics* 53(2).

Reifschneider, David, Robert Tetlow and John Williams. 1999. Aggregate Disturbances, Monetary Policy and the Macroeconomy: The FRB/US Perspective. Federal Reserve Bulletin, 85(1).

Rotemberg, Julio. 1982. Sticky prices in the united states. *Journal of Political Economy* 90 (4).

Rotemberg, Julio, and Michael Woodford. 1996. Real-Business-Cycle Models and the Forecastable Movements in Output, Hours, and Consumption. *The American Economic Review* 86(1).

Rotemberg, Julio, and Michael Woodford. 1997. An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual* 12.

Rudebusch, Glenn and Lars Svensson. 1999. Policy rules for inflation targeting. In J. Taylor (editor), *Monetary Policy Rules*, Chicago: University of Chicago Press.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2004. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control* 28.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2005. Habit Persistence. In S. Durlauf and L. Blume (editors), *The New Palgrave Dictionary of Economics*, second edition.

Schorfheide, Frank. 2011. Estimation and Evaluation of DSGE Models: Progress and Challenges. NBER Working Paper 16781.

Sims, Christopher. 2001. Solving Linear Rational Expectations Models. *Journal of Computational Economics* 20(1-2).

Slobodyan, Sergey and Raf Wouters. 2007. Learning in an estimated medium-scale dsge model. Working Paper.

Slobodyan, Sergey and Raf Wouters. 2008. Estimating a medium-scale DSGE model with expectations based on small forecasting models. Mimeo.

Smets, Frank and Raf Wouters. 2003. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association* 1(5).

Smets, Frank and Raf Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *The American Economic Review* 97(3).

Sveen, Tommy and Lutz Weinke. 2005. New perspectives on capital, sticky prices, and the Taylor principle. *Journal of Economic Theory* 123.

Taylor, John B. 1979a. Staggered wage setting in a macro model. *The American Economic Review, Papers and Proceedings* 69 (2). Reprinted in N.G. Mankiw and D. Romer (editors). 1991. *New Keynesian Economics*, MIT Press.

Taylor, John B. 1979b. Estimation and control of a macroeconomic model with rational expectations. *Econometrica* 47(5).

Taylor, John B. 1980. Aggregate dynamics and staggered contracts. *Journal of Political Economy* 88(1).

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

Taylor, John B. 1993b. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39.

Taylor, John B. (editor). 1999. *Monetary Policy Rules*, The University of Chicago Press.

Taylor, John B. and Harald Uhlig. 1990. Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods. *Journal of Business and Economic Statistics* 8(1).

Taylor, John B. and Volker Wieland. 2011. Surprising comparative properties of monetary models: Results from a new data base. forthcoming in *Review of Economics and Statistics*.

Taylor, John B. and John Williams. 2010. Simple and Robust Rules for Monetary Policy. Prepared for the *Handbook of Monetary Economics*.

Tovar, Camilo. 2008. DSGE models and central banks. BIS Working Papers No 258.

Townsend, Robert M. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* 21.

Uhlig, Harald. 1999. A toolkit for analyzing nonlinear dynamic stochastic models easily. In Ramon Marimonm Andrew Scott (editors), *Computational Methods for the Study of Dynamic Economies*, Oxford: Oxford University Press.

Walsh, Carl E. 2010. *Monetary Theory and Policy*. Third edition. MIT Press.

Wieland, Volker (editor). 2009. *The Science and Practice of Monetary Policy Today*, Springer Science.

Wieland, Volker. 2009. Learning, endogenous indexation and disinflation in the New Keynesian model. In C. Walsh and K. Schmidt-Hebbel (editors), *Monetary Policy under Uncertainty and Learning*. Central Bank of Chile.

Wieland, Volker, Tobias Cwik, Gernot Mueller, Sebastian Schmidt, and Maik Wolters. 2011. A New Comparative Approach to Macroeconomic Modeling and Policy Analysis. Working Paper.

Wieland, Volker and Maik Wolters. 2011. The diversity of forecasts from macroeconomic models of the U.S. economy. *Economic Theory* 47(2-3): 247-292.

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

Woodford, Michael. 2005. Firm-Specific Capital and the New-Keynesian Phillips Curve. *International Journal of Central Banking* 1(2).

Yellen, Janet. 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", 2007.

Yun, Tack. 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics* 37.

1. Small Calibrated Models

1.1	NK_RW97	Rotemberg and Woodford (1997)
1.2	NK_LWW03	Levin et al. (2003)
1.3	NK_CGG99	Clarida et al. (1999)
1.4	NK_CGG02	Clarida et al. (2002)
1.5	NK_MCN99cr	McCallum and Nelson (1999), (Calvo-Rotemberg model)
1.6	NK_IR04	Ireland (2004)
1.7	NK_BGG99	Bernanke et al. (1999)
1.8	NK_GM05	Galí and Monacelli (2005)
1.9	NK_GK09	Gertler and Karadi (2009)
1.10	NK_CK08	Christoffel and Kuester (2008)
1.11	NK_CKL09	Christoffel et al. (2009)
1.12	NK_RW06	Ravenna and Walsh (2006)

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_FRB08	linearized by Brayton and Laubach (2008)
2.5	US_FRB08mx	linearized by Brayton and Laubach (2008), (mixed expectations)
2.6	US_SW07	Smets and Wouters (2007)
2.7	US_ACELm	Altig et al. (2005), (monetary policy shock)
	US_ACELt	Altig et al. (2005), (technology shocks)
	US_ACELswm	no cost channel as in Taylor and Wieland (2011) (mon. pol. shock)
	US_ACELswt	no cost channel as in Taylor and Wieland (2011) (tech. shocks)
2.8	US_NFED08*	based on Edge et al. (2007), version used for estimation in Wieland and Wolters (2011)
2.9	US_RS99	Rudebusch and Svensson (1999)
2.10	US_OR03	Orphanides (2003)
2.11	US_PM08	IMF projection model US, Carabenciov et al. (2008)
2.12	US_PM08fl	IMF projection model US (financial linkages), Carabenciov et al. (2008)
2.13	US_DG08	De Graeve (2008)
2.14	US_CD08	Christensen and Dib (2008)
2.15	US_IAC05	Iacoviello (2005)
2.16	US_MR07	Mankiw and Reis (2007)
2.17	US_RA07	Rabanal (2007)
2.18	US_CCTW10	Smets and Wouters (2007) model with rule-of-thumb consumers, estimated by Cogan et al. (2010)
2.19	US_IR11	Ireland (2011)

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. (2007b)
3.6	EA_QUEST3	QUEST III Euro Area Model of the DG-ECFIN EU, Ratto et al. (2009)
3.7	EA_CKL09	Christoffel et al. (2009)
3.8	EA_GE10**	Gelain (2010)

4. Estimated/Calibrated Multi-Country Models

4.1	G7_TAY93	Taylor (1993a) model of G7 economies
4.2	G3_CW03	Coenen and Wieland (2002) model of USA, Euro Area and Japan
4.3	EACZ_GEM03	Laxton and Pesenti (2003) model calibrated to Euro Area and Czech republic
4.4	G2_SIGMA08	The Federal Reserve's SIGMA model from Erceg et al. (2008) calibrated to the U.S. economy and a symmetric twin.
4.5	EAUS_NAWM08	Coenen et al. (2008), New Area Wide model of Euro Area and USA
4.6	EAES_RA09	Rabanal (2009)

5. Estimated Models of other Countries

5.1	CL_MS07	Medina and Soto (2007), model of the Chilean economy
5.2	CA_ToTEM10*	ToTEM model of Canada, based on Murchison and Rennison (2006), 2010 vintage
5.3	BRA_SAMBA08	Gouvea et al. (2008), model of the Brazilian economy
5.4	CA_LS07	Lubik and Schorfheide (2007), small-scale open-economy model of the Canadian economy
5.5	HK_FPP11	Funke et al. (2011), open-economy model of the Hong Kong economy