WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 11-5

EVALUATING DSGE MODEL FORECASTS OF COMOVEMENTS

Edward Herbst
University of Pennsylvania

Frank Schorfheide
University of Pennsylvania,
CEPR, NBER, and Visiting Scholar,
Federal Reserve Bank of Philadelphia

January 12, 2011

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Evaluating DSGE Model Forecasts of Comovements

Edward Herbst
University of Pennsylvania

Frank Schorfheide*
University of Pennsylvania
CEPR, NBER, and Visiting Scholar,
Federal Reserve Bank of Philadelphia

January 12, 2011

*Correspondence: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: herbstep@sas.upenn.edu (E. Herbst); schorf@ssc.upenn.edu (F. Schorfheide). We are very grateful to Rochelle Edge and Refet Gürkaynak for providing us with their real-time data set for the estimation of the Smets-Wouters model. We thank Frank Diebold as well as seminar participants at the 2010 FRB Philadelphia Real-Time Data Research Center Conference, the Riksbank, and the 2010 ESOBE in Rotterdam. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant SES 0617803. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/

Abstract

This paper develops and applies tools to assess multivariate aspects of Bayesian Dynamic Stochastic General Equilibrium (DSGE) model forecasts and their ability to predict comovements among key macroeconomic variables. We construct posterior predictive checks to evaluate the calibration of conditional and unconditional density forecasts, in addition to checks for root-mean-squared errors and event probabilities associated with these forecasts. The checks are implemented on a three-equation DSGE model as well as the Smets and Wouters (2007) model using real-time data. We find that the additional features incorporated into the Smets-Wouters model do not lead to a uniform improvement in the quality of density forecasts and prediction of comovements of output, inflation, and interest rates.

JEL CLASSIFICATION: C11, C32, C53, E27, E47

KEY WORDS: Bayesian Methods, DSGE Models, Forecast Evaluation, Macroeconomic Forecasting

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models use modern macroeconomic theory to explain and predict comovements of aggregate time series over the business cycle, which is one of the features that makes DSGE models attractive to central banks for forecasting and policy analysis.[1] The small existing literature on the assessment of DSGE model forecasts tends to focus on point forecasts, predominantly evaluated based on root-mean-squared error (RMSE) measures. The general finding is that DSGE model forecasts are comparable to standard autoregressive or vector autoregressive models in terms of RMSE but are dominated by sophisticated univariate or multivariate statistical time series models. However, RMSEs do not reflect the DSGE models' alleged strength, namely their ability to predict comovements between key macroeconomic variables. Moreover, the quadratic prediction error loss function underlying RMSE comparisons may not be the relevant loss for decision makers. In fact, central banks increasingly pay attention to density forecasts to assess the probability of particular events, such as inflation and output growth being above or below target, and to judge the uncertainty about future economic developments more generally. The contribution of this paper is to extend the evaluation of DSGE model forecasts in two dimensions. First, we develop and apply statistics that are designed to assess multivariate aspects of DSGE model forecasts. Second, rather than measuring the precision of point forecasts, our evaluation focuses on whether predictive densities are well calibrated in the sense that predicted probabilities of events are commensurable with actual frequencies in a sequential application of the forecasting procedure.

DSGE models are frequently estimated with Bayesian methods, in particular if the goal is to track and forecast macroeconomic time series. Bayesian inference delivers posterior predictive distributions that reflect uncertainty about latent state variables, parameters, and future realizations of shocks conditional on the available information. To assess a model's ability to predict comovements among macroeconomic variables, three statistics are considered in this paper. First, we derive conditional distributions from the model's joint predictive

[1] The ability to conduct policy experiments in which agents' decision rules are re-derived under the counterfactual policy is another important feature.

distribution. In particular, we condition on future realizations of one of the DSGE model variables. We assess whether the resulting conditional predictive densities are well calibrated in the sense that probability integral transforms (PITs) are uniformly distributed. Second, we compute point forecasts from the conditional predictive distributions and construct RMSE ratios of conditional versus unconditional forecasts, where the latter are computed from the marginal predictive distribution. We examine to what extent the actual RMSE ratios are commensurable with the model-implied RMSE ratios. Finally, we compare the predicted probability of events, such as inflation and output growth being below target, with the observed frequencies.

We evaluate the predictions of comovements for two DSGE models, a simple, three-equation New Keynesian model and the more elaborate model developed by Smets and Wouters (2007). The small-scale model reduces the New Keynesian monetary theory to an intertemporal Euler equation, a New Keynesian Phillips curve (NKPC), and a monetary policy rule and thus provides a natural benchmark. The Smets-Wouters (SW) model is a New Keynesian model that has been augmented by internal and external propagation mechanisms to improve its time series fit, and variants of the model have been considered in other studies of forecasting performance of DSGE models. Given the relevance of the exercise to the policy making process, we use a real-time data set constructed in Edge and Gürkaynak (2010). Using this data in our estimation ensures that the information set upon which the forecasts are based matches the one that was available to policymakers over the past 15 years.

The starting point of our analysis is the observation that output growth, inflation, and interest rate forecasts from the two models considered in this paper attain very similar RM-SEs. However, going beyond their respective means, the density forecasts of the two DSGE models are quite different in various dimensions. We find that the additional mechanisms incorporated into the SW model do not lead to a uniform improvement in the quality of the density forecasts and in the prediction of comovements compared with the small three-equation DSGE model. The PITs computed from the marginal density forecasts of the small-scale model are by and large uniform, at least over short horizons. The realized RM-

SEs are of similar magnitude as the model-implied RMSEs. For the SW model, on the other hand, the distribution of output growth PITs is skewed, and in general, the density forecast of output is too diffuse. The reason for this deficiency is most likely the counterfactual common trend restriction that the SW model imposes on output, consumption, and investment. In terms of predicting whether average output and inflation will fall above or below their long-run target values, the small-scale model also does a better job than the SW model as actual frequencies are closer to the predicted probabilities.

With respect to other dimensions of comovement predictions, the SW model comes out ahead. The estimated small-scale DSGE model exhibits a fairly strong correlation between output and inflation in the predictive distribution that would imply that knowing future output would lead to a substantial reduction in the forecast error of inflation. However, it turns out that the predictive distribution does not adequately capture the comovement among the two variables, as the actual RMSE of the conditional forecast is larger than the RMSE of the unconditional forecast. The SW model, on the other hand, implies that there is very little exploitable correlation between output and inflation as well as output and interest rates, which turns out to be consistent with an actual RMSE ratio that is close to one. The predictive distribution of the SW model does, however, contain a strong correlation between interest rates and inflation, which correctly implies that knowing future interest rates can substantially improve the precision of inflation forecasts, and vice versa. For the SW model we also find that the predictive distribution of inflation conditional on the federal funds rate delivers better calibrated probability statements than the unconditional predictive distribution, which is consistent with inflation being well modeled conditional on interest rates and a relatively poor fit of the monetary policy rule.

While, strictly speaking, the predictive distributions in a Bayesian framework are subjective, an important question for practitioners is whether they are well calibrated in view of the observed data. In repeated practical (as opposed to hypothetical) use of a statistical procedure such as sequential forecasting, it is desirable that the long-run average level of accuracy is no less than (and ideally equal to) the long-run average reported accuracy. Bayarri and Berger (2004) refer to this notion as *frequentist principle*. Geweke and Whiteman (2006) emphasize that Bayesian approaches to forecast evaluation are fundamentally different from non-Bayesian approaches, although they might appear superficially similar. In a Bayesian framework, there is no uncertainty about the predictive density given the specified collection of models. Predictive densities are simply constructed by the relevant conditioning. Non-Bayesian approaches, surveyed in Corradi and Swanson (2006), tend to adopt the notion of a "true" data-generating process (DGP) and try to approximate the predictive density inherent in the DGP with a member of a collection of probability distributions $\mathcal{P}_{t+h|t} = \{P^{\theta}_{t+h|t} | \theta \in \Theta\}$. To the extent that the forecaster faces uncertainty with respect to θ, there is uncertainty about the density forecast itself. In turn, non-Bayesian assessments of density forecasts try to account for this uncertainty.

Formally, we conduct so-called posterior predictive checks to assess the adequacy of DSGE model forecasts. A discussion of the role of predictive checks in Bayesian analysis can be found, for instance, in Lancaster (2004) and Geweke (2005). It is important to note that the objective of the forecast evaluation conducted in this paper is not to select the best forecasting model from a collection of DSGE models, to average predictive densities from competing DSGE models, or to "reject" or "accept" a particular DSGE model specification. Instead, the goal is to examine whether, despite some of the empirical deficiencies of DSGE models documented in the literature, e.g. Del Negro, Schorfheide, Smets, and Wouters (2007), the probability forecasts generated by the models are *adequate* in certain dimensions, as policymakers at central banks shift their attention from point forecasts to interval and density forecasts.[2] In general, the purpose of posterior predictive checks is to provide diagnostics that may spur creative thinking about new models.

One of the statistics used in our posterior predictive checks, namely PITs constructed from the DSGE models' predictive distributions, has a long tradition in the literature on density forecast evaluation. Dawid (1984) proposed a prequential approach to statistics, founded on the premiss that the purposes of statistical inference is to make adequate sequential probability forecasts for future observations. While there are many different ways of defining the adequacy of probability forecasts, Dawid (1984) suggested, among other things, to examine

[2]Bayesian methods to formally compare DSGE models in instances in which the model space spanned by a collection of DSGE models is deemed incomplete are discussed in Schorfheide (2000) and Geweke (2010).

whether PITs are *iid* uniformly $\mathcal{U}[0,1]$ distributed. Kling and Bessler (1989) construct probability density forecasts from a VAR for interest rates, money stock, consumer prices, and output. To examine whether the PITs are uniformly distributed, the authors compared their empirical distribution functions with the cumulative distribution function (cdf) of a $\mathcal{U}[0,1]$ random variable and reported the outcome of χ^2 tests. Diebold, Gunther, and Tay (1998) emphasize that PITs of a sequence of density forecasts should also be independent across time in addition to marginally uniformly distributed. Thus, the authors inspect autocorrelation functions of PITs (and nonlinear transformations of PITs) to assess density forecasts for the S&P 500 constructed from GARCH models. Diebold, Hahn, and Tay (1999) extend the PIT-based density forecast evaluation to multivariate models. More recently, Geweke and Amisano (2010) use PITs to evaluate Bayesian forecasts from competing univariate time series models of S&P 500 portfolio returns.

The remainder of the paper is organized as follows. Section 2 briefly reviews empirical findings regarding the forecasting performance of DSGE models. The predictive checks are described in Section 3, and the two DSGE models considered in this paper are summarized in Section 4. The empirical results are presented in Section 5 and Section 6 concludes. An Online Appendix[3] contains the log-linearized equilibrium conditions of the SW model as well as a description of how we use Kernel methods to compute moments and PITs from conditional densities based on a sequence of draws from a joint density.

2 Forecasting Record of DSGE Models

Most of the literature on evaluating the forecasting performance of DSGE models has focused on the evaluation of point forecasts based on RMSEs or the log determinant of the error covariance matrix of the forecasts. The latter criterion, henceforth "ln-det" statistic, had been proposed by Doan, Litterman, and Sims (1984).[4] Table 1 summarizes pseudo out-of-

[3]Available at http://www.econ.upenn.edu/s̃chorf/.

[4]The eigenvectors of the forecast error covariance matrix generate linear combinations of the model variables with uncorrelated forecast errors. The determinant equals the product of the eigenvalues and thereby measures the product of the forecast error variances associated with these linear combinations. The

sample RMSEs for U.S. data obtained with Smets and Wouters (2003, 2007) type DSGE models from the following five studies: (i) Del Negro, Schorfheide, Smets, and Wouters (2007), (DSSW); (ii) Smets and Wouters (2007) (SW); (iii) Edge, Kiley, and Laforte (2009) (EKL); (iv) Edge and Gürkaynak (2010) (EG); and (v) Schorfheide, Sill, and Kryshko (2010) (SSK). Since all studies differ with respect to the forecast period, sample standard deviations for the respective forecast periods are also reported. Unlike the other three studies, EKL and EG use real-time data. Overall, the RMSEs reported in DSSW are slightly worse than those in the other three studies. This might be due to the fact that DSSW use a rolling window of 120 observations to estimate their DSGE model and start forecasting in the mid-1980s, whereas the other papers let the estimation sample increase and start forecasting in the 1990s. Only EKL and EK are able to attain an RMSE for output growth that is lower than the sample standard deviation. The RMSEs for the inflation forecasts range from 0.21 to 0.29 and are very similar across studies. They are only slightly larger than the sample standard deviations. Finally, the interest rate RMSEs are substantially lower than the sample standard deviations because the forecasts are able to exploit the high persistence of the interest rate series.

EG compare univariate forecasts from the SW model estimated with real-time data against forecasts obtained from the staff of the Federal Reserve, the Survey of Professional Forecasters, and a Bayesian VAR. Based on RMSEs, they conclude that the DSGE model delivers forecasts that are competitive in terms of accuracy with those obtained from the alternative prediction methods. The evidence from Euro Area data is similar. Adolfson, Lindé, and Villani (2007) assess the forecasting performance of an Open Economy DSGE model during the period of 1994 to 2004 based on RMSEs, ln-det statistics, predictive scores, and the coverage frequency of interval forecasts. Overall, the authors conclude that the DSGE model compares well with more flexible time series models such as VARs. Christoffel, Coenen, and Warne (2010) examine the forecasting performance of the New Area Wide Model (NAWM), the DSGE model used by the European Central Bank. The authors evaluate the model's univariate forecast performance through RMSEs and its multivariate performance using the

more linear combinations exist that can be predicted with small forecast error variance, the smaller the ln-det statistic.

ln-det statistic. They find that the DSGE model is competitive with other forecasting models such as VARs of various sizes. The authors also find that the assessment of multivariate forecasts based on the ln-det statistic can sometimes be severely affected by the inability to forecast just one series, nominal wage growth.

The Bayesian VARs that serve as a benchmark in the above cited papers use a Minnesota prior but are typically not optimized with respect to their empirical performance. For instance, some of the dummy observations described in Sims and Zha (1998) and more recently discussed in Del Negro and Schorfheide (2010) that generate *a priori* correlations among VAR coefficients and have been found useful for prediction have been excluded from the construction of the prior distribution. Del Negro and Schorfheide (2004) and DSSW compare the forecasting performance of a three-equation New Keynesian DSGE model and a variant of the SW model to Bayesian VARs that use a prior distribution centered at the DSGE model restrictions. Both papers find that the resulting DSGE-VAR forecasts significantly better than the underlying DSGE model. Overall, the empirical evidence supports our claim that DSGE model forecasts are comparable to standard autoregressive or vector autoregressive models but can be dominated by more sophisticated univariate or multivariate time series models. Thus, in addition to considering the precision of DSGE model point forecasts, subsequently the emphasis will be on examining whether the probability density forecasts are well calibrated.

3 Econometric Approach: Predictive Checks

The forecast evaluation in this paper is based on posterior predictive checks. The actual sample is partitioned into $Y_{1:R}$ and $Y_{R+1:T}$ and we define $P = T - R$. In addition, let $Y^*_{R+1:T}$ be a hypothetical sample of length P. The predictive distribution for $Y^*_{R+1:T}$ is given by

$$p(Y^*_{R+1:T}|Y_{1:R}) = \int p(Y^*_{R+1:T}|\theta)p(\theta|Y_{1:R})d\theta. \qquad (1)$$

Draws from the predictive distribution can easily be obtained in two steps. First, generate a parameter draw $\tilde{\theta}$ from $p(\theta|Y_{1:R})$. Second, simulate a trajectory of observations $Y^*_{R+1:T}$ from the DSGE model conditional on $\tilde{\theta}$. The simulated trajectories can be converted into sample statistics of interest, $\mathcal{S}(Y^*_{R+1:T})$ to obtain an approximation for predictive distributions of these sample statistics. Finally, one can compute the value of the statistic $\mathcal{S}(Y_{R+1:T})$ based on the actual data and assess how far it lies in the tails of its predictive distribution. A $\mathcal{S}(Y_{R+1:T})$ far in the tails of its predictive distribution can be interpreted as a model deficiency. We now discuss the specific choices of checking functions $\mathcal{S}(\cdot)$.

PITs for One-Step Ahead Forecasts. Let $y_{i,t}$, $i = 1, \ldots, n$, denote the elements of the vector y_t and define the probability integral transform for the actual one-step ahead forecast of $y_{i,t+1}$ based on time t information by

$$z_{i,t,1} = \int_{-\infty}^{y_{i,t+1}} p(\tilde{y}_{i,t+1}|Y_{1:R}, Y_{R+1:t}) d\tilde{y}_{i,t+1}. \tag{2}$$

Likewise, the probability integral transform constructed for a forecast of $y^*_{i,t+1}$ based on the sample $(Y_{1:R}, Y^*_{R+1:t})$ can be defined as

$$z^*_{i,t,1} = \int_{-\infty}^{y^*_{i,t+1}} p(\tilde{y}^*_{i,t+1}|Y_{1:R}, Y^*_{R+1:t}) d\tilde{y}^*_{i,t+1}. \tag{3}$$

Thus, the sequence $\{z^*_{i,t,1}\}_{t=R}^{T-1}$ can be interpreted as a draw from the model-implied predictive distribution of the probability integral transforms. It is shown in Rosenblatt (1952) and Diebold, Gunther, and Tay (1998) that the $z^*_{i,t,1}$'s are *iid* $\mathcal{U}[0,1]$ distributed. This property can be exploited in a predictive check. For instance, suppose we divide the unit interval into J sub-intervals. According to the predictive distribution, the fraction of PITs in each sub-interval is equal to $1/J$. Paraphrasing Bayarri and Berger's (2004) *frequentist principle*, the fraction of actual PITs in each sub-interval should be close to the predicted fraction. The *closeness* can be assessed with a χ^2 goodness-of-fit statistic of the form

$$\mathcal{S}_\chi(z_{i,R,1}, \ldots, z_{i,T-1,1}) = \sum_{j=1}^{J} \frac{(n_j - P/J)^2}{P/J}, \tag{4}$$

where n_j is the number of PITs in the bin $[(j-1)/J, j/J]$.

PITs for Multi-Step Ahead and Conditional Forecasts. The multi-step generalization of (2) is given by

$$z_{i,t,h} = \int_{-\infty}^{y_{i,t+h}} p(\tilde{y}_{i,t+h}|Y_{1:R}, Y_{R+1:t}) d\tilde{y}_{i,t+1}. \tag{5}$$

Conditional on $(Y_{1:R}, Y_{R+1:t})$ the marginal distribution of $z_{i,t,h}$ remains $\mathcal{U}[0,1]$. However, the sequence $\{z_{i,t,h}\}_{t=R}^{T-h}$ is serially dependent, which is reflected in the posterior predictive distribution of statistics $S_\chi(z^*_{i,R,h}, \ldots, z^*_{i,T-h,h})$, such as the fraction of PITs falling into the J bins $[(j-1)/J, j/J]$ and the χ^2 statistic (4). The simulation-based approach taken in this paper resolves the problem of approximating the finite-sample distribution of statistics $S_\chi(\cdot)$ by directly generating draws from its distribution.

PITs for conditional predictive distributions $p(\tilde{y}_{i,t+h}|y_{j,t+h}, Y_{1:R}, Y_{R+1:t})$ are defined as

$$z_{i|j,t,h} = \int_{-\infty}^{y_{i,t+h}} p(\tilde{y}_{i,t+h}|y_{j,t+h}, Y_{1:R}, Y_{R+1:t}) d\tilde{y}_{i,t+1}. \tag{6}$$

Finally, the DSGE models are specified in terms of growth rates rather than log levels of output and prices. Thus, instead of forecasts of a growth rate between period $t+h-1$ and $t+h$, we consider forecasts of the average growth rate between period t and period $t+h$ defined as $\bar{y}_{i,t+1:t+h} = (1/h)\sum_{s=1}^{h} y_{i,t+s}$. In turn $h\bar{y}_{i,t+1:t+h}$ captures the total change between the forecast origin and period $t+h$. The corresponding PITs are defined by replacing $y_{i,t+h}$ and $\tilde{y}_{i,t+h}$ in (5) and (6) with $\bar{y}_{i,t+1:t+h}$ and $\tilde{\bar{y}}_{i,t+1:t+h}$, respectively.

RMSE Ratios. We also use RMSEs and ratios of conditional and unconditional RMSEs as in Schorfheide, Sill, and Kryshko (2010) to form a predictive check. Let $\hat{y}_{i|j,t+h}$ and $\hat{y}_{i,t+h}$ denote the means of the conditional and marginal predictive distribution of $y_{i,t+h}$. The RMSE ratio is defined as

$$\mathcal{R}(i|j,h) = \sqrt{\frac{\frac{1}{P-h}\sum_{s=0}^{P-h}(y_{i,R+s+h} - \hat{y}_{i|j,R+s+h})^2}{\frac{1}{P-h}\sum_{s=0}^{P-h}(y_{i,R+s+h} - \hat{y}_{i,R+s+h})^2}} \tag{7}$$

and has a straightforward extension to the prediction of multi-step averages $\bar{y}_{i,t+1:t+h}$. If the forecast errors are homoskedastic and normally distributed, then the RMSE ratio is one if the forecast errors are uncorrelated and less than one if the forecast errors are correlated.

Event Forecasts. Finally, we consider probability forecasts for events of the form $\{y_{i,t+h} \geq a,\ y_{j,t+h} \geq b\}$. Let $\mathbb{P}_t[\{y_{i,t+h} \geq a,\ y_{j,t+h} \geq b\}]$ denote the model predicted probability that the event occurs. The event forecasts are evaluated based on

$$\begin{aligned}\mathcal{P}(i,j,h) \;=\; & \frac{1}{P-h}\sum_{s=0}^{P-h}\Big(\{y_{i,R+s+h} \geq a,\ y_{j,R+s+h} \geq b\} \\ & -\mathbb{P}_{R+s}[\{y_{i,R+s+h} \geq a,\ y_{j,R+s+h} \geq b\}]\Big).\end{aligned} \tag{8}$$

The statistic $\mathcal{P}(i,j,h)$ measures the divergence of the actual frequency of the events $\{y_{i,t+h} \geq a,\ y_{j,t+h} \geq b\}$ from the probabilities of these events implied by the model.

Our econometric analysis is novel in the following dimensions. First, none of the previous studies on DSGE model forecast performance has examined the uniformity of PITs. Only Adolfson, Lindé, and Villani (2007) assessed the coverage frequency of interval forecasts for Euro Area forecasts generated with a small open economy DSGE model. Second, none of the existing papers have set up the forecast evaluation formally as a predictive check. While it is well known that the PITs for a sequence of one-step ahead forecasts are independently distributed, PITs of h-step ahead forecasts are serially correlated. Our simulation approach is able to capture this serial correlation. Third, while PITs based on predictive densities that are conditioned on future realizations of a subset of variables arise naturally in a multivariate density forecast evaluation setting as explained in Diebold, Hahn, and Tay (1999), they have not yet been applied to assess a DSGE model's ability to forecast comovements. Finally, while ratios of RMSEs of unconditional forecasts versus forecasts that are conditioned on the future realization of a subset of variables have been reported in Schorfheide, Sill, and Kryshko (2010), that paper did not provide a formal benchmark, such as percentiles of predictive distributions, against which the ratios could be evaluated. Moreover, Schorfheide, Sill, and Kryshko (2010) used a parametric approximation of the predictive density to form conditional forecasts, whereas we are using a nonparametric Kernel-based approximation of the predictive density.

4 The DSGE Models

The predictive checks are applied to two New Keynesian DSGE models. First, we consider a small-scale model that consists of three basic equations: a consumption Euler equation, a New Keynesian Phillips curve, and a monetary policy rule. The theoretical properties of this class of models are discussed extensively in Woodford (2003), and numerous versions that differ with respect to the specification of the exogenous shock processes and the formulation of the monetary policy rule have been estimated based on output, inflation, and interest rate data, see Schorfheide (2008) for a survey. Second, we generate forecasts from the SW model. This model has a richer structure that accounts for capital accumulation, variable capital utilization, wage rigidity in addition to price rigidity, and households' habit formation.

4.1 A Small-Scale Model

Empirical specifications of the canonical small-scale New Keynesian DSGE model differ with respect to the exogenous shock processes as well as the formulation of the monetary policy rule. Our version is identical to the one studied in the survey paper by An and Schorfheide (2007) and includes a technology growth, a government spending, and a monetary policy shock. The interest rate feedback rule implies a reaction to output growth deviations from steady state rather than to deviations of the level of output from a measure of potential output.

Log-Linearized Equilibrium Conditions. We briefly summarize the log-linearized equilibrium conditions associated with the small-scale DSGE model. The underlying decision problems of households and firms are described in detail in An and Schorfheide (2007). Let $\hat{x}_t = \ln(x_t/x)$ denote the percentage deviation of a variable x_t from its steady state x. The equilibrium can be approximated by an intertemporal Euler equation, a New Keynesian Phillips curve, and an interest rate feedback rule:

$$\begin{aligned} \hat{y}_t &= \mathbb{E}_t[\hat{y}_{t+1}] + \hat{g}_t - \mathbb{E}_t[\hat{g}_{t+1}] - \frac{1}{\tau}\left(\hat{R}_t - \hat{\mathbb{E}}_t[\pi_{t+1}] - \mathbb{E}_t[\hat{z}_{t+1}] \right) \qquad (9) \\ \hat{\pi}_t &= \beta \mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t) \\ \hat{R}_t &= \rho_R \hat{R}_{t-1} + (1-\rho_R)\psi_1 \hat{\pi}_t + (1-\rho_R)\psi_2 \left(\Delta \hat{y}_t + \hat{z}_t\right) + \epsilon_{R,t}. \end{aligned}$$

Here y_t denotes output, π_t inflation, and R_t nominal interest rates. The parameter β is the households' discount factor and τ is the inverse intertemporal elasticity of substitution. The parameter κ captures the slope of the Phillips curve. The monetary policy rule depends on the smoothing parameter ρ_R and the coefficients ψ_1 and ψ_2, which determine how strongly the central bank reacts to deviations of inflation and output growth from their target levels. The model economy is perturbed by three exogenous shocks. $\epsilon_{R,t}$ is a monetary policy shock, and $\hat{z}_t$ and $\hat{g}_t$ are AR(1) processes that capture total factor productivity growth and the evolution of government spending (as a fraction of output):

$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}, \quad \hat{g}_t = \rho_g \hat{g}_{t-1} + \epsilon_{g,t}. \tag{10}$$

Measurement Equations. The model is completed by a set of measurement equations that relate the model states to a set of observables. We assume that the time period t in the model corresponds to one quarter and that the following observations are available for estimation: quarter-over-quarter per capita GDP growth rates (YGR), annualized quarter-over-quarter inflation rates (PI), and annualized nominal interest rates (INT).[5] The three series are measured in percentages, and their relationship to the model variables is given by the following set of equations:

$$\begin{aligned} YGR_t &= \gamma^{(Q)} + \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \\ INF_t &= \pi^{(A)} + 4\hat{\pi}_t \\ FFR_t &= \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 4\hat{R}_t. \end{aligned} \tag{11}$$

The parameter $\gamma^{(Q)}$ captures the steady-state growth rate of output, which in this simple model is identical to the growth rate of the exogenous technology. $\pi^{(A)}$ is the annualized steady-state inflation rate, which is equal to the central bank's target rate. Finally, we use $r^{(A)}$ to denote the following transformation of the households' discount factor: $\beta^{-1} = 1 + r^{(A)}/400$.

Prior Distribution. Bayesian estimation of a DSGE model requires the specification of a prior distribution. This distribution is constructed as a product of marginals, which are summarized in Table 2. We use the same priors as in An and Schorfheide (2007) with one exception: The inflation coefficient in the monetary policy rule is fixed at $\psi_1 = 1.7$. It is well known in the literature that ψ_1 is difficult to identify. This lack of identification causes some numerical instabilities in the application of Markov-Chain Monte-Carlo (MCMC) methods. Since the predictive check requires us to estimate the DSGE model many times and the precise measurement of ψ_1 is not the objective of our analysis, we decided to fix the parameter.

[5]Here we follow An and Schorfheide (2007) and estimate the model based on annualized inflation and interest rates. The RMSE statistics reported in Section 5 are converted to quarter-over-quarter percentages.

Priors for the autocorrelations and standard deviations of the exogenous processes, the steady-state parameters $\gamma^{(Q)}$, $\pi^{(A)}$, and $r^{(A)}$, as well as the standard deviation of the monetary policy shock, are quantified based on regressions on pre-(estimation)-sample observations of output growth, inflation, and nominal interest rates. The priors for the policy rule coefficients ψ_2 and ρ_R are loosely centered around values typically associated with the "Taylor rule." The prior for the parameter that governs price stickiness is chosen based on micro-evidence on price setting-behavior provided. More formal methods for the elicitation of priors for DSGE model parameters are discussed in Del Negro and Schorfheide (2008).

Posterior Inference. To implement the posterior predictive checks, we need to generate draws from a sequence of posterior distributions $p(\theta|Y_{1:R+\tau})$ for $\tau = 0, \ldots, P-1$. For $\tau = 0$ we use the Random-Walk Metropolis (RWM) algorithm in An and Schorfheide (2007). Draws for $\tau > 0$ are also generated with the RWM algorithm. However, the covariance matrix of the proposal density is constructed by re-weighting the draws from $p(\theta|Y_{1:R+\tau-1})$ with the importance weight $p(y_{R+\tau}|\theta, Y_{1:R+\tau-1})$.

4.2 The Smets-Wouters Model

The SW model is second model considered in this paper. The model is a more elaborate version of the DSGE model presented in Section 4.1. Capital is a factor of intermediate goods production, and nominal wages, in addition to nominal prices, are rigid. The model is based on work by Christiano, Eichenbaum, and Evans (2005), who added various forms of frictions to a basic New Keynesian DSGE model in order to capture the dynamic response to a monetary policy shock as measured by a structural vector autoregression (VAR). In turn, Smets and Wouters (2003) augmented the Christiano-Eichenbaum-Evans model by additional shocks to be able to capture the joint dynamics of Euro Area output, consumption, investment, hours, wages, inflation, and interest rates. The 2007 version of the SW model contains a number of minor modifications of the 2003 model in order to optimize its fit on U.S. data. We use the 2007 model exactly as presented in SW and refer the reader to that article for details. The log-linearized equilibrium conditions are reproduced in the Online Appendix.

Measurement Equations. The SW model is estimated based on seven macroeconomic time series. The period t corresponds to one quarter and the measurement equations for output growth, inflation, interest rates, consumption growth, investment growth, wage growth, and hours worked are given by:

$$\begin{aligned} YGR_t &= \bar{\gamma} + \hat{y}_t - \hat{y}_{t-1} \\ INF_t &= \bar{\pi} + \hat{\pi}_t \\ FFR_t &= \bar{r} + \hat{R}_t \\ CGR_t &= \bar{\gamma} + \hat{c}_t - \hat{c}_{t-1} \\ IGR_t &= \bar{\gamma} + \hat{i}_t - \hat{i}_{t-1} \\ WGR_t &= \bar{\gamma} + \hat{w}_t - \hat{w}_{t-1} \\ HOURS_t &= \bar{l} + \hat{l}_t. \end{aligned} \tag{12}$$

Unlike in the small-scale model, for the estimation of the SW model both measured interest and inflation rates are not annualized. Moreover, since the neutral technology shock in the SW model is assumed to be stationary, the model variables are not transformed as in the small-scale model to induce stationarity, and the growth rate of the technology shock does not appear in the measurement equations.

Prior Distributions. Based on information that does not enter the likelihood function, SW fix the following five parameters in their estimation:

$$\delta = 0.025, \quad g_y = 0.18, \quad \lambda_w = 1.50, \quad \varepsilon_w = 10.0, \quad \varepsilon_p = 10.$$

We deviate from SW's analysis by some additional parameters:

$$\begin{aligned} \varphi &= 5.00, \quad \sigma_c = 1.5, \quad h = 0.7, \quad \xi_w = 0.7, \quad \sigma_l = 2, \\ \xi_p &= 0.7, \quad \iota_w = 0.5, \quad \iota_p = 0.5, \quad r_\pi = 2, \quad \alpha = 0.3. \end{aligned}$$

These parameter values are close to the posterior mean estimates reported in Smets and Wouters (2007). Our predictive check requires us to estimate the SW model several hundred times on recursive samples. Fixing the additional parameters ensures the numerical stability of our MCMC methods. The marginal prior distributions for the remaining parameters are identical to those used by SW and are summarized in Table 3. The joint prior density is constructed as the product of the marginal densities.

Posterior Inference. The posterior simulator is implemented in the same way as for the small-scale New Keynesian model.

5 Empirical Results

The empirical results are presented in three steps. First, we discuss the data set that is used to conduct the predictive check (Section 5.1). Second, we evaluate the marginal predictive distributions of output growth, inflation, and interest rates (Section 5.2). Finally we examine the prediction of comovements of the small-scale DSGE model and the SW model (Section 5.3). The results of the predictive checks reported below are based on 100 trajectories from predictive distribution $p(Y^*_{R+1:T}|Y_{1:R})$. Each simulated trajectory $Y^*_{R+1:T}$ is combined with $Y_{1:R}$, and the DSGE model is estimated recursively, starting with the sample $Y_{1:R}$ and ending with the sample $(Y_{1:R}, Y^*_{R+1:T-1})$. For each estimation, we generate 10,000 draws from the posterior using the MCMC approach described in Section 4.1. To construct recursive density forecasts, we use a subsample of 1,000 parameter draws $\tilde{\theta}$ and simulate 10 trajectories from $p(\tilde{Y}_{R+\tau+1:R+\tau+h}|\tilde{\theta}, Y_{1:R}, Y^*_{R+1:\tau})$ for each $\tilde{\theta}$. The same calculations are repeated along the observed history $(Y_{1:R}, Y_{R+1:T})$. The computation of the PITs based on the draws from the predictive distribution is described in the Online Appendix.

5.1 Data Set

For the evaluation of the density forecasts, we are using the real-time data set assembled by Edge and Gürkaynak (2010). These authors compared the accuracy of point forecasts from the SW model to those from the Fed's Greenbook. Thus, for each Greenbook publication date, EK compiled the time series that are used in the estimation of the SW model. Since the focus in our paper is not a comparison of the DSGE model and the Greenbook forecast we use only a subset of the data sets constructed by EK, namely those for the Greenbooks published in March, June, September, and December. We refer to the March forecast as the first-quarter forecasts, and the remaining forecasts are associated with Quarters 2 to 4.

The March forecasts are based on fourth-quarter releases from the previous year, meaning that the estimation period for the DSGE model effectively ends in Q4 of the preceding year. Thus, the first forecast in March is essentially a nowcast for Q1, and the subsequent forecasts are for Q2, Q3, and so forth. For each forecast origin, we refer to the "nowcast" as a one-step-ahead forecast and choose the maximum forecast horizon according to the maximum horizon of the corresponding Greenbook. Thus, we construct forecasts up to eight periods ahead for March origins, seven periods ahead for June origins, 10 periods ahead for September origins, and nine periods ahead for December origins. The first forecast origin in our analysis is March 1997 and the last forecast origin is December 2004, which provides us with 32 sets of forecasts.

Since there is strong empirical evidence that monetary policy as well as the volatility of macroeconomic shocks changed in the early 1980s, we estimate both DSGE models based on data sets that start in 1984:Q3.[6] We use both the small-scale DSGE model as well as the SW model to construct pseudo out-of-sample forecasts using the real-time data set. Forecast errors are calculated based on the first final release of each variable. The predictive distribution for the model checks – using the notation of Section 3 – is constructed conditional on $Y_{1:R}$. Period R corresponds to 1996:Q4, which is the last period in the estimation sample that is used to generate the March 1997 forecasts. When simulating from the predictive distribution $p(Y^*_{R+1+T}|Y_{1:R})$, we do not take into account the data revision process. For the analysis of the SW model, all real series are converted into per capita terms, whereas we do not convert GDP growth in per capita terms when estimating the small-scale DSGE model. Finally, all h-step forecasts refer to averages over the forecast horizon. Thus, using the notation of Section 3, we consider forecasts of $\bar{y}_{i,t+1:t+h}$ rather than $y_{i,t+h}$.

[6] We are using a conditional likelihood function that conditions on observations from 1983:Q3 to 1984:Q4.

5.2 Evaluation of Marginal Predictive Distributions

Figure 1 displays the RMSEs of pseudo out-of-sample forecasts of average output growth, inflation, and interest rates up to eight quarters ahead computed from the small-scale DSGE model and the SW model. The dashed lines indicate the 90% credible bands associated with the model-implied predictive distribution of the RMSEs. The RMSEs attained with the small-scale model and the SW model are of similar magnitude as those reported in Table 1. For the small DSGE model, the RMSEs fall within the 90% credible interval generated by the predictive distribution, with the exception of the inflation rate at longer horizons. For the SW model, the actual output and inflation RMSEs fall below the 5th percentile of their predictive distribution at most horizons. Conversely, the federal funds rate RMSEs are well within the bands associated with their predictive distribution.

The bands of the posterior predictive distribution supplement the information provided by the actual RMSEs in an important dimension. The actual output and inflation RMSEs in the SW model are in general much smaller than what one would expect from the estimated model. A potential explanation for this finding is that some of the exogenous shock processes in the SW model are highly persistent, in part because they have to capture deviations of output, consumption, and investment from the model-implied common trend. Highly persistent shocks in turn imply fairly large forecast error variances, which contribute to the RMSE. Thus, the inconsistency of the actual RMSEs with those from the predictive check is a reflection of a model deficiency.

Figure 2 displays the histograms of the unconditional PITs from both models at one-quarter and four-quarter horizons. To generate the histogram plots, we divide the unit interval into five equally sized subintervals and depict the fraction of PITs (measured in percent) computed from the actual data that fall in each bin. Since the PITs are uniformly distributed on the unit interval, we also plot the 20% line. Finally, the dashed lines indicate the 5th and 95th percentile of the predictive distribution for the fraction of PITs in each bin. The χ^2-goodness-of-fit statistics, which measure the squared distance of the bin heights from the 20% line, are summarized in Table 4. We also report the 90th percentiles (in parentheses) from the predictive distribution. In general, the χ^2 statistics conform with the graphical information provided in the figures.

At the short horizon, the density forecasts from the small DSGE model appear well calibrated for output and inflation, with the bars of the PIT histograms approximately equally sized and within the interval obtained from the predictive distribution. The distribution of the PITs associated with the federal funds rate for one-step-ahead forecasts appears somewhat skewed to the left. The behavior the federal funds rate over the evaluation period was persistently low, a feature the simple model was unable to capture. Looking at four-quarter averages, output forecasts still appear well calibrated, while the distribution of the PITs for inflation has deviated from uniformity. The density forecasts are too diffuse in view of the observed inflation rates, which leads to PITs that cluster in the center of the unit interval. For the federal funds rate, actual realizations are still falling too frequently in the left tail of the predictive distribution.

The PIT histograms for the SW model are similar for inflation and the federal funds rate, but different with respect to output. The one-quarter ahead forecast for output of the SW model is poorly calibrated in comparison to the small-scale DSGE model. The predictive distribution of output is too diffuse, as the PITs too frequently fall in the 0.2 to 0.6 interval. This deficiency is also apparent at longer horizons. In fact, it is consistent with the RMSE predictive check discussed above. The estimated shock processes generate a predictive distribution that is too diffuse. If the model is simulated forward, the counterfactual output growth RMSEs are much larger than the actual RMSEs because the simulated trajectories exhibit an unrealistically large output volatility.

Figure 3 displays the autocorrelation function of the one-step ahead PITs of the actual forecasts and the 90% credible interval associated with the predictive distribution. As emphasized by Diebold, Gunther, and Tay (1998), the PITs should be independent and thus serially uncorrelated. For both DSGE models, the PITs of the inflation forecasts indeed appear uncorrelated. The PITs for the observed interest rates, on the other hand, are highly correlated in particular at horizons one and two. This correlation is slightly higher in the small model. With respect to output, the two DSGE models differ quite substantially.

The output growth PITs based on the small-scale DSGE model exhibit a strong positive correlation, whereas they are essentially uncorrelated for the SW model. Deviations from independence are indicative of a dynamic misspecification in the sequence of predictive distributions, which overall is more pronounced for the small-scale model than the SW model. This finding is consistent with the SW model containing many additional features that were designed to capture the dynamics of U.S. data.

Figures 1, 2, and 3 taken together portray the collective performance of the marginal predictive distributions for output growth, inflation, and interest rates. Consider the one quarter-ahead forecast for output, for example. While the RMSEs of point forecasts from both models are similar, the predictive distributions exhibit important differences. The predictive distribution for the small model is well calibrated in the sense that the marginal distributions of the PITs appear to be uniform. However, the model forecasts are deficient in a dynamic sense. The positive correlation of the output growth PITs indicates that a too-high GDP forecast is likely to be followed by a too-high GDP forecast the next quarter. On the other hand, the output growth forecasts from the SW model appear dynamically well calibrated. Forecast errors in one quarter contain no information about errors in future quarters. However, in an unconditional sense the density forecasts are deficient because they predict output volatility much higher than what was observed in the data. While the density forecast performance of the two DSGE models is markedly different, no clear ranking emerges. Both models have deficiencies along different dimensions. The posterior predictive checks highlight deficiencies that are not apparent from a RMSE comparison.

5.3 Evaluating Predictions of Comovements

To assess the DSGE models' ability to predict comovements among macroeconomic aggregates, we now consider statistics computed from bivariate predictive distributions. Since a joint density of two random variables can be factorized into a marginal and a conditional density and since we have examined marginal densities in the previous subsection, we shall now focus on predictive densities that are conditioned on future realizations of output growth, inflation, or interest rates.

The starting point of the analysis is the relative precision of conditional and unconditional point forecasts. The unconditional point forecasts are the same that were used in Section 5.2 to compute the RMSEs displayed in Figure 1. To obtain forecasts conditional on future realizations, say of output growth, we compute the mean of the conditional predictive density. In a nutshell, if a joint predictive distribution is approximately normal and implies that two variables are highly correlated, then knowing the future realization of one of them should substantially reduce the RMSE when predicting the other. If the correlation structure in the predictive distribution is consistent with the observed comovements of the two variables, the predicted RMSE reduction should be commensurable with the actual reduction. Figure 4 displays the RMSE ratios of conditional and unconditional forecasts, $\mathcal{R}(i|j,h)$ in (7). For the small DSGE model, there are gains to forecasting output by conditioning on interest rates and vice versa at short horizons. Moreover, these gains are consistent with those implied by the predictive distribution. On the other hand, there is no reduction in RMSE associated with forecasts of GDP conditional on inflation and vice versa, which is largely inconsistent with the gain in efficiency predicted by the estimated model. Thus, the correlation structure in the predictive distribution does not appropriately capture the comovements of output and inflation.

The RMSE ratios for the SW model exhibit different behavior. There is hardly any gain using the conditional relationship between output and inflation and output and interest rates for forecasting. Moreover, this is, for the most part, consistent with the predictive distribution for the RMSE ratios. On the other hand, there are sizable gains from conditioning on interest rates to forecast inflation and vice versa. At the eight-quarter horizon, the reduction in RMSE associated with these conditional distributions is roughly 15 basis points (annualized) for the annualized inflation rate and 30 basis points for the annualized interest rate.

We proceed by examining the calibration of predictive distribution conditional on the average realization of the output, inflation, and interest rates at one- and four-quarter horizons, respectively. Figure 5 displays PITs based on density forecasts conditional on future output. For the small DSGE model, the PIT histograms for one-step ahead conditional forecasts look very similar to their unconditional counterparts in Figure 2. For average inflation over four-quarters, however, the predictive distribution conditional on output appears better calibrated than the unconditional counterpart, which overstates the volatility of inflation. This is reflected in the goodness-of-fit statistics reported in Table 4. The χ^2 statistic for the unconditional forecast lies much farther in the tail of its predictive distribution. For the SW model, the distribution of the PITs conditional on output looks similar to the distribution of the unconditional PITs.

The comovement of output growth and inflation with the federal funds rate is of particular importance to central banks, which often compare macroeconomic forecasts based on various hypothetical paths for short-term nominal interest rates. Thus, Figure 6 depicts PIT histograms for output and inflation forecasts that condition on average future interest rates. It is important to note that these histograms do not convey any information about the DSGE models' ability to generate accurate counterfactual policy predictions. We only examine whether the DSGE model is able to capture the comovement between interest rates and other macroeconomic variables under the actual policy.

For the small DSGE model, the histograms look very similar to their unconditional counterparts. One difference is that, at the four-quarter horizon, the conditional distribution for inflation appears better calibrated than the unconditional distribution. This is reflected in the χ^2 statistics reported in Table 4. Its value for the unconditional distribution exceeds the 90th percentile, whereas the value for the unconditional distribution does not. Recall from Figure 4 that at the four-quarter horizon, the RMSE ratio for inflation conditional on interest rates is greater than one (and far outside the predictive distribution). Since the RMSEs of the unconditional point forecasts were small relative to the predictive distribution, the RMSE ratio being greater than one is consistent with a correction of this deficiency in the conditional distribution. In short, a conditional distribution can be properly calibrated even when the marginal distribution is not.

For the SW model, at the short horizon the conditional distribution looks similar to the unconditional distribution. This is consistent with the observation that the RMSE ratios for one-quarter ahead forecasts are basically unity for output and inflation conditional on interest rates. At the longer horizon, there is a large gain in precision in forecasting inflation when conditioning on interest rates. While the PITS for inflation forecasts conditional on interest rates are still not distributed uniformly, the distribution is less skewed to the left. Using the model, conditional on knowing the average interest rate over one year, one is less likely to underpredict the inflation rate, relative to the case in which the interest rate is not known. This is consistent with the reduction in RMSE displayed in Figure 4.

We now consider a simple event study as a final dimension along which to evaluate the predictive densities. We partition the sample space for average output and inflation in four events: output growth and inflation are both above (below) their respective long-run targets, output growth is above and inflation is below target, and output is below and inflation is above target. These events summarize the models' ability to correctly forecast the directional movements of output and inflation, which are important in policy settings. If both output and inflation are above (below) their steady-state values, the policymaker has an incentive to raise (lower) interest rates. We assess how big the divergence is between the realized frequency of these events and the average probabilities of events implied by the model in Figure 7 for up to eight quarters ahead, see $\mathcal{P}(i, j, h)$ in (8).

The small DSGE model does a good job at predicting these events overall. It slightly underpredicts the event in which output is growing above trend and inflation is falling but these predictions still fall within the 90% band of the predictive distribution. The results for the SW model indicate that it performs slightly worse than the small DSGE model. The event in which output growth and inflation are above target is overpredicted by the model relative to the actual occurrences. Moreover, like the small DSGE model, the event in which output growth is above and inflation is below target is underpredicted.

The SW model also generates a predictive distribution for consumption, investment, hours, and wages, in addition to the three variables we have considered. Now we augment our earlier results with a selection of results on the predictive distribution for consumption, investment, and hours. Figure 8 displays the PITs based on forecasts of consumption, investment, and hours given future interest rates. We see that for one-quarter ahead forecasts, the predictive distributions appear well calibrated. The histogram for consumption is skewed slightly to the right, while the histograms for investment and hours have too much mass in the $[0.4, 0.6]$ quantile, but nearly all of the realizations lie within the 90% band of the predictive distribution. At four quarters, these undesirable features, which are also present in the unconditional density forecasts of investment and consumption, become much more pronounced. The SW model imposes a common growth rate on consumption, investment, and output. In the data, however, the growth rates are slightly different. While the estimated model generates a predictive distribution for investment that is correctly centered, it tends to underpredict consumption and overpredict output growth. Conditioning on the federal funds rate leads to a RMSE reduction of the hours worked forecast but has essentially no positive effect on the consumption and investment forecasts.

6 Conclusion

This paper develops and applies tools to assess multivariate aspects of Bayesian DSGE model density forecasts and their ability to predict comovements among key macroeconomic variables. The forecast evaluation is implemented through posterior predictive checks. These checks are based on three sets of statistics. First, we consider PITs computed from marginal and conditional density forecasts. Second, we compute ratios of RMSEs attained by point forecasts derived from marginal and conditional predictive densities. Finally, we study the difference between the frequency of occurrence and the model-implied probability of particular events.

The predictive checks are applied to a simple three-equation New Keynesian model as well as the more elaborate SW model. The starting point is the observation that the output, inflation, and interest rate forecasts from both models attain very similar RMSEs. However, these point forecasts capture only the mean of the density forecast. It turns out that the predictive densities of the two DSGE models are quite different in various dimensions, yet no clear ranking emerges. The additional features incorporated into the SW model do not lead to a uniform improvement in the quality of the density forecasts and in the prediction of comovements. The predictive density for output appears to be poorly calibrated. Surprisingly, the small scale model also seems to do better in terms of predicting whether output growth and inflation will be above or below target.

Moving forward, we hope that these predictive checks can be used as a diagnostic tool to assess model performance in a policy-relevant way and to spur new thinking about DSGE models, particularly in modeling the relationship between variables.

References

Adolfson, M., J. Lindé, and M. Villani (2007): "Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model," *Econometric Reviews*, 26(2-4), 289–328.

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

Bayarri, M., and J. Berger (2004): "The Interplay of Bayesian and Frequentist Analysis," *Statistical Science*, 19(1), 58–80.

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

Christoffel, K., G. Coenen, and A. Warne (2010): "Forecasting with DSGE Models," in *Oxford Handbook on Economic Forecasting*, ed. by M. Clements, and D. Hendry, p. forthcoming. Oxford University Press.

Corradi, V., and N. Swanson (2006): "Predictive Density Evaluation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann, vol. 1 of *Handbooks in Economics 24*, pp. 197–286. North Holland, Amsterdam.

Dawid, A. (1984): "Statistical Theory: The Prequential Approach," *Journal of the Royal Statistical Society, Series A*, 147(2), 278–292.

Del Negro, M., and F. Schorfheide (2004): "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45(2), 643 – 673.

——— (2008): "Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities)," *Journal of Monetary Economics*, 55(7), 1191–1208.

——— (2010): "Bayesian Macroeconometrics," in *Handbook of Bayesian Econometrics*, ed. by H. K. van Dijk, G. Koop, and J. Geweke. Ox.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007): "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 25(2), 123–162.

Diebold, F., T. Gunther, and A. Tay (1998): "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39(4), 863–883.

Diebold, F., J. Hahn, and A. Tay (1999): "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange," *Review of Economics and Statistics*, 81(4), 661–673.

Doan, T., R. Litterman, and C. A. Sims (1984): "Forecasting and Conditional Projections Using Realistic Prior Distributions," *Econometric Reviews*, 3(4), 1–100.

Edge, R., and R. Gürkaynak (2010): "How Useful Are Estimated DSGE Model Forecasts for Central Bankers," *Brookings Papers of Economic Activity*, p. forthcoming.

Edge, R., M. Kiley, and J.-P. Laforte (2009): "A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model," *Federal Reserve Board of Governors Finance and Economics Discussion Paper Series*, 2009-10.

Geweke, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons, Hoboken, New Jersey.

——— (2010): *Complete and Incomplete Econometric Models*. Princeton University Press, Princeton, New Jersey.

Geweke, J., and G. Amisano (2010): "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns," *International Journal of Forecasting*, 26(2), 216–230.

Geweke, J., and C. H. Whiteman (2006): "Bayesian Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. Granger, and A. Timmermann, vol. 1, pp. 3–80. North Holland, Amsterdam.

Kling, J., and D. Bessler (1989): "Calibration-Based Predictive Distributions: An Application of Prequential Analysis to Interest Rates, Money, Prices, and Output," *Journal of Business*, 62(4), 447–499.

Lancaster, T. (2004): *An Introduction to Modern Bayesian Econometrics*. Blackwell Publishing.

Rosenblatt, M. (1952): "Remarks on a Multivariate Transformation," *Annals of Mathematical Statistics*, 23(3), 470–472.

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

——— (2008): "DSGE Model-Based Estimation of the New Keynesian Phillips Curve," *FRB Richmond Economic Quarterly*, Fall Issue, 397–433.

Schorfheide, F., K. Sill, and M. Kryshko (2010): "DSGE Model-Based Forecasting of Non-Modelled Variables," *International Journal of Forecasting*, 26(2), 348–373.

Sims, C. A., and T. Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–968.

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

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Woodford, M. (2003): *Interest and Prices*. Princeton University Press.

Table 1: One-Step-Ahead Forecast Performance of DSGE Models

Study	Forecast Period	Output Growth (Q %)	Inflation (Q %)	Interest Rate (Q %)
Del Negro *et al.* (2007)	1985:IV to 2000:I	0.73	0.27	0.22
		(0.52)	(0.25)	(0.43)
Smets, Wouters (2007)	1990:I to 2004:IV	0.57	0.24	0.11
		(0.57)	(0.22)	(0.49)
Edge, Kiley, Laforte (2009)	1996:III to 2004:IV	0.45	0.29	0.21
		(0.57)	(0.20)	(0.49)
Edge, Gürkaynak (2010)	1997:I to 2006:III	0.49	0.21	0.12(*)
		(0.56)	(0.23)	(0.47)
Schorfheide, Sill, Kryshko (2010)	2001:I to 2007:IV	0.51	0.22	0.18
		(0.47)	(0.22)	(0.42)

Notes: Del Negro *et al.* (2007, Table 2): VAR approximation of DSGE model estimated based on rolling samples of 120 observations. Smets and Wouters (2007, Table 3): DSGE model is estimated recursively, starting with data from 1996:I. Edge, Kiley, and Laforte (2009, Table 5): DSGE model is estimated recursively with real-time data starting in 1984:II. Edge, Gürkaynak (2010, Figure 3): DSGE model is estimated recursively, with real-time data from 1997:I. (*) The interest rate entry refers to changes in the interest rate. Schorfheide, Sill, Kryshko (2010, Table 2): DSGE model is estimated recursively with data starting in 1984:I. Numbers in parentheses are sample standard deviations for the forecast period, computed from the Schorfheide, Sill, Kryshko data set. Q % is the quarter-to-quarter percentage change.

Table 2: Prior Distribution for Small-Scale Model

	Density	Para (1)	Para (2)
τ	Gamma	2.00	0.50
κ	Gamma	0.20	0.10
ψ_2	Gamma	0.50	0.25
ρ_R	Beta	0.50	0.20
ρ_G	Beta	0.80	0.10
ρ_Z	Beta	0.66	0.15
$r^{(A)}$	Gamma	0.50	0.50
$\pi^{(A)}$	Gamma	7.00	2.00
$\gamma^{(Q)}$	Normal	0.40	0.20
σ_R	InvGamma	0.40	4.00
σ_G	InvGamma	1.00	4.00
σ_Z	InvGamma	0.50	4.00

Notes: The following parameter is fixed: $\psi_1 = 1.70$. Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

Table 3: Prior Distribution for Smets-Wouters Model

	Density	Para (1)	Para (2)		Density	Para (1)	Para (2)
ψ	Beta	2.00	0.50	Φ	Normal	1.25	0.12
ρ	Beta	0.75	0.10	r_y	Normal	0.12	0.05
$r_{\Delta y}$	Normal	0.12	0.05	$\bar{\pi}$	Gamma	0.62	0.10
$100(\beta^{-1}-1)$	Gamma	0.25	0.10	$\bar{l}$	Normal	875	10.0
$\bar{\gamma}$	Normal	0.40	0.10	σ_a	Invgamma	0.10	2.00
σ_b	Invgamma	0.10	2.00	σ_g	Invgamma	0.10	2.00
σ_I	Invgamma	0.10	2.00	σ_r	Invgamma	0.10	2.00
σ_p	Invgamma	0.10	2.00	σ_w	Invgamma	0.10	2.00
ρ_a	Beta	0.50	0.20	ρ_b	Beta	0.50	0.20
ρ_g	Beta	0.50	0.20	ρ_I	Beta	0.50	0.20
ρ_r	Beta	0.50	0.20	ρ_p	Beta	0.50	0.20
ρ_w	Beta	0.50	0.20	μ_p	Beta	0.50	0.20
μ_w	Beta	0.50	0.20	ρ_{ga}	Beta	0.50	0.20

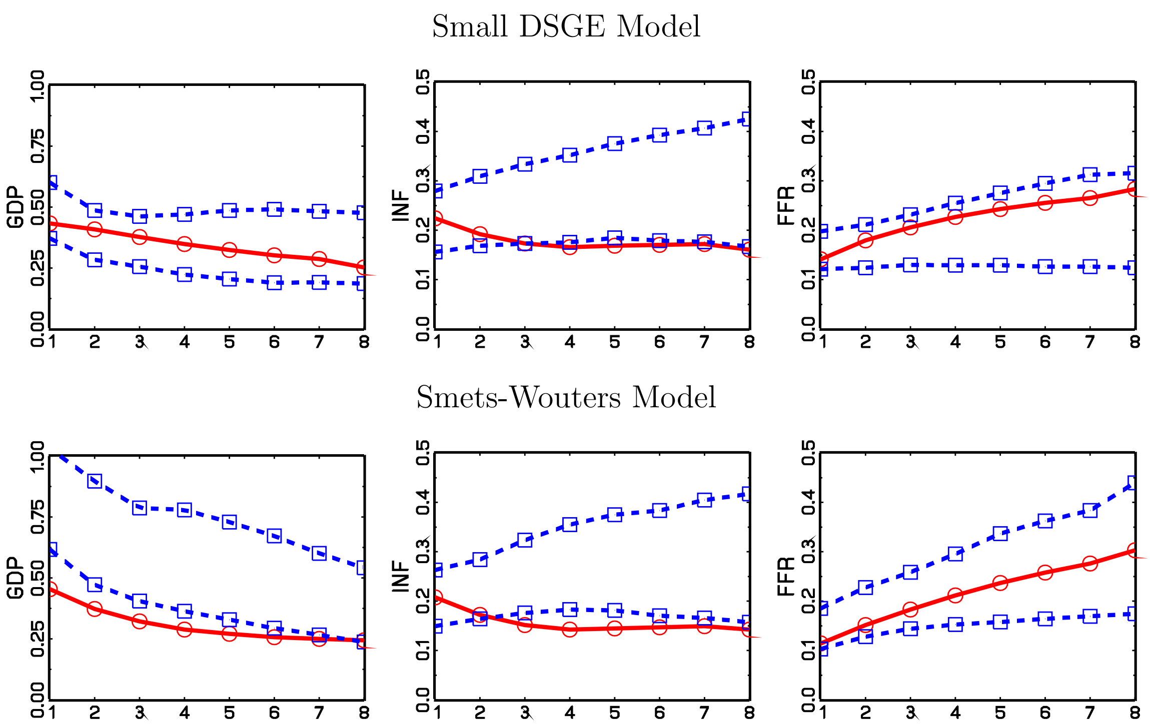
Notes: The following parameters are fixed in Smets and Wouters (2007): $\delta = 0.025$, $g_y = 0.18$, $\lambda_w = 1.50$, $\varepsilon_w = 10.0$, and $\varepsilon_p = 10$. In addition, we fix: $\varphi = 5.00$, $\sigma_c = 1.5$, $h = 0.7$, $\xi_w = 0.7$, $\sigma_l = 2$, $\xi_p = 0.7$, $\iota_w = 0.5$, $\iota_p = 0.5$, $r_\pi = 2$, $\alpha = 0.3$. Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

Table 4: χ^2-Goodness-Fit Statistics for PITs

	Model	Output Gr.		Inflation		Interest	
			1 Quarter Ahead				
Uncond.	Small-Scale	2.69	(8.00)	1.75	(8.31)	**8.00**	(7.38)
	Smets-Wouters	**23.3**	(7.38)	3.63	(7.69)	**8.94**	(8.31)
Cond on Y	Small-Scale			1.75	(8.31)	**12.7**	(6.75)
	Smets-Wouters			4.25	(8.63)	**8.94**	(8.00)
Cond on π	Small-Scale	3.31	(8.94)			6.44	(7.69)
	Smets-Wouters	**22.1**	(9.25)			**8.00**	(7.38)
Cond on R	Small-Scale	4.88	(8.00)	1.75	(7.69)		
	Smets-Wouters	**21.4**	(8.31)	3.00	(7.38)		
			4 Quarters Ahead				
Uncond.	Small-Scale	2.69	(13.9)	**14.3**	(13.9)	**14.6**	(14.3)
	Smets-Wouters	**16.4**	(15.8)	**17.7**	(13.6)	**13.9**	(13.0)
Cond on Y	Small-Scale			2.69	(14.9)	**23.9**	(13.9)
	Smets-Wouters			**17.7**	(13.0)	12.1	(13.0)
Cond on π	Small-Scale	3.31	(16.4)			**13.9**	(13.6)
	Smets-Wouters	**16.4**	(14.9)			9.88	(12.4)
Cond on R	Small-Scale	3.62	(12.1)	13.0	(14.6)		
	Smets-Wouters	12.7	(15.8)	**13.3**	(13.0)		

Notes: The values in parentheses correspond to the 90th percentile of the predictive distribution. A *boldface entry* indicates that the actual value exceeded the 90th percentile of the predictive distribution.

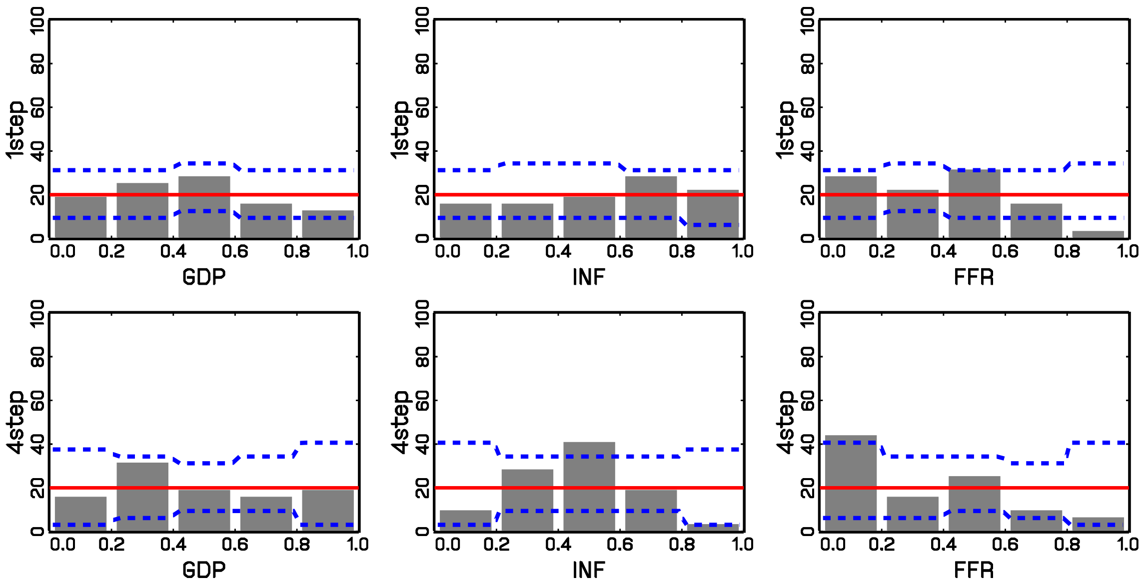
Figure 1: RMSEs of Unconditional Forecasts



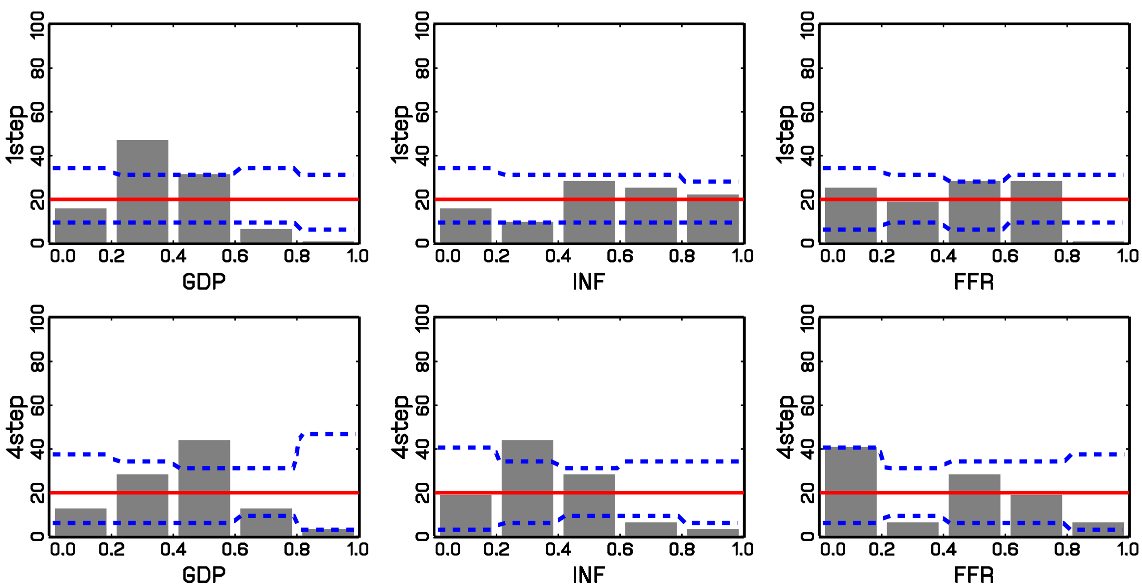
Notes: Root-mean-squared Errors (RMSEs) for forecasts of output growth (*GDP*), inflation (*INF*), and interest rates (*FFR*). The *solid line* corresponds to RMSEs associated with actual forecasts, and the *dashed line* signifies 90% credible intervals obtained from the predictive distribution.

Figure 2: PIT Histograms – Unconditional Forecasts

Small DSGE Model

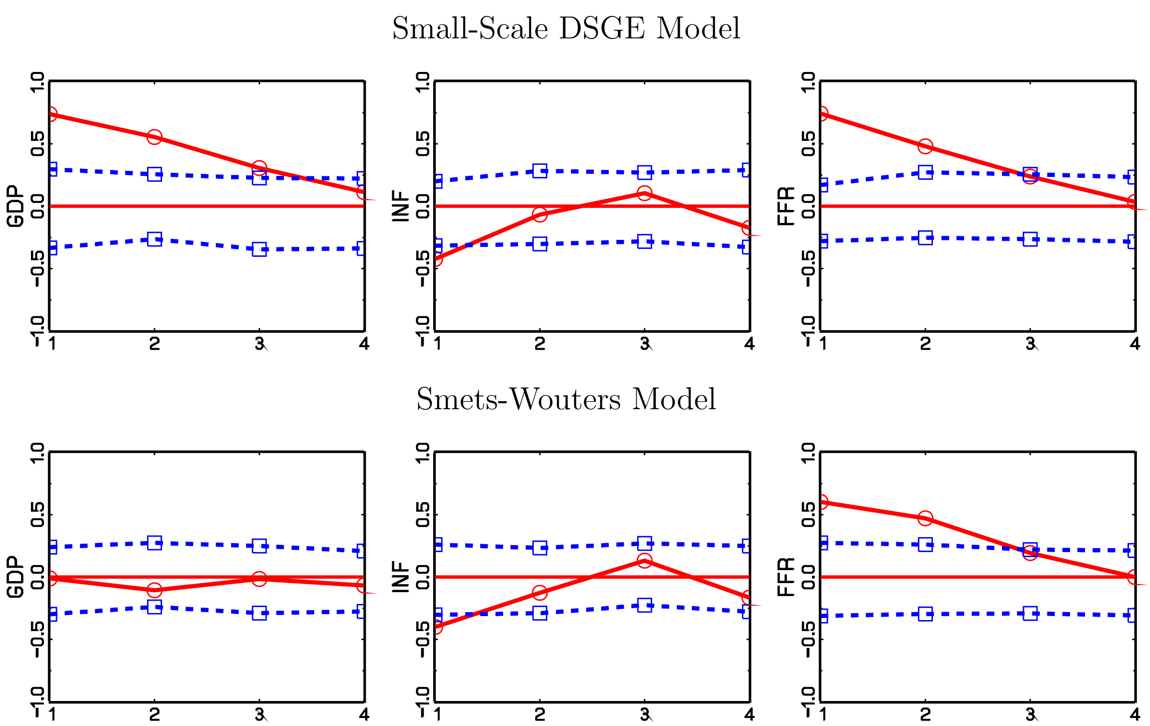


Smets-Wouters Model



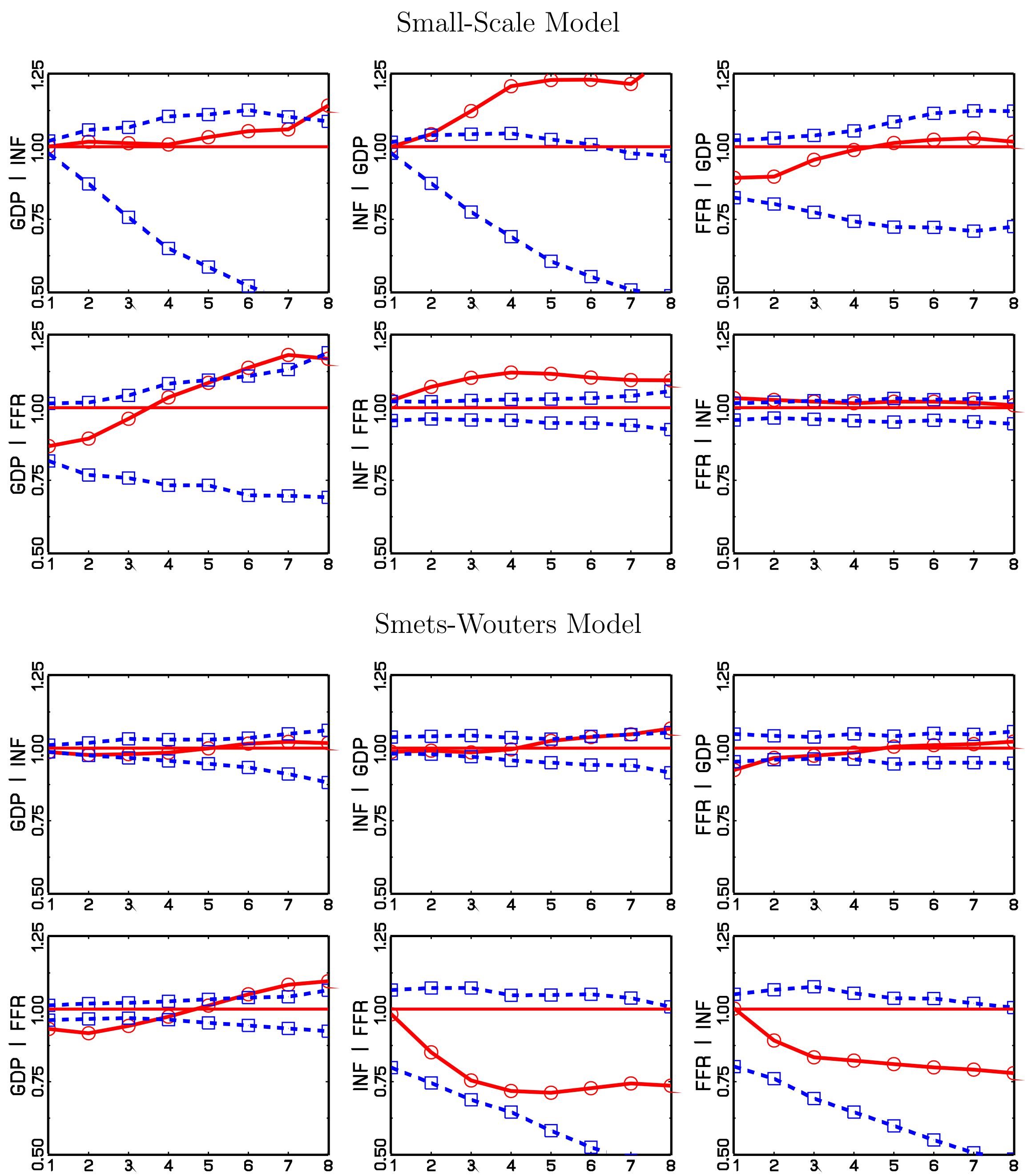
Notes: Probability integral transforms for forecasts of output growth (*GDP*), inflation (*INF*), and interest rates (*FFR*). *Bars* correspond to actuals, and *dashed bands* indicate 90% credible intervals obtained from the predictive distribution.

Figure 3: PIT AUTOCORRELATIONS OF ONE-STEP FORECASTS



Notes: Autocorrelation functions (ACFs) of PITs for forecasts of output growth (*GDP*), inflation (*INF*), and interest rates (*FFR*) of order 1 to 4 (x-axis). The *solid line* corresponds to ACFs associated with actual forecasts, and the *dashed line* signifies 90% credible intervals obtained from the predictive distribution.

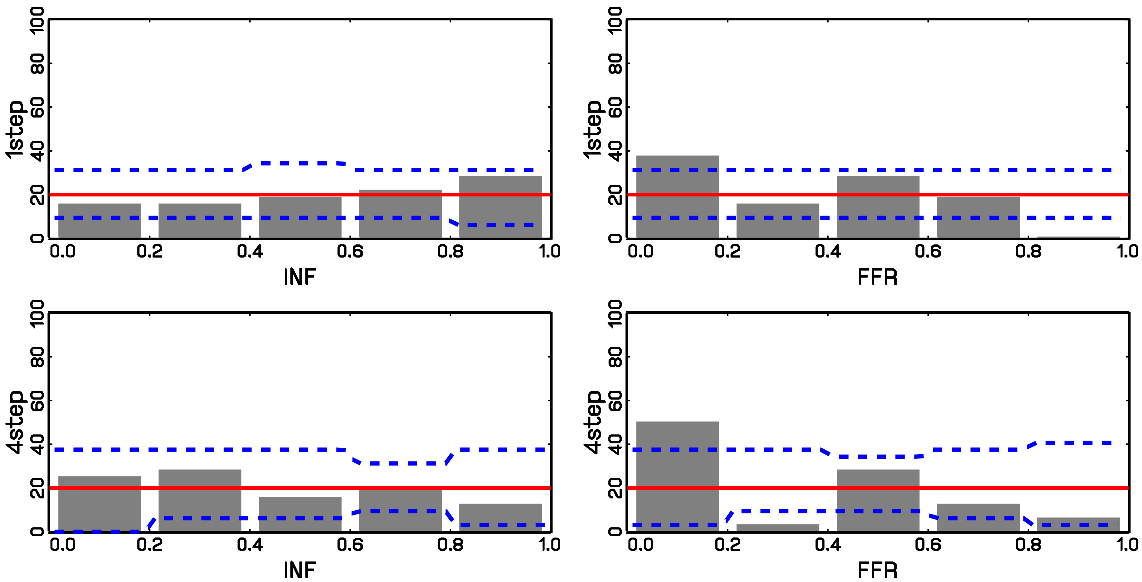
Figure 4: RMSE Ratios of Conditional and Unconditional Forecasts



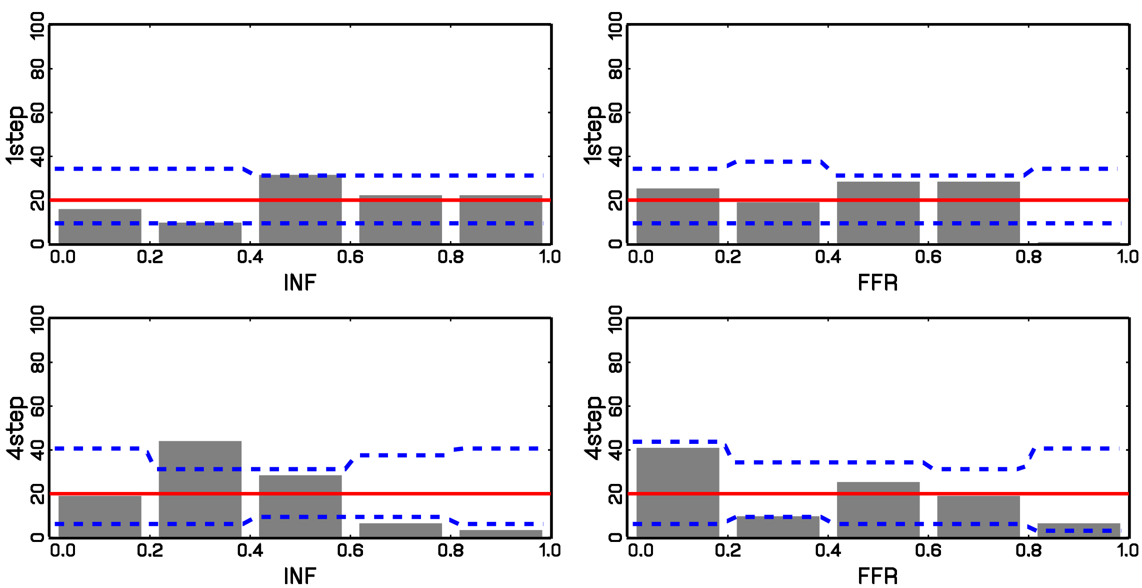
Notes: RMSE ratios for forecasts of output growth (*GDP*), inflation (*INF*), and interest rates (*FFR*): conditional on future realizations of other variables versus unconditional. The *solid line* corresponds to RMSE ratios associated with actual forecasts, and the *dashed line* signifies 90% credible intervals obtained from the predictive distribution.

Figure 5: PITs Histograms – Forecasts Given Future Output

Small-Scale Model



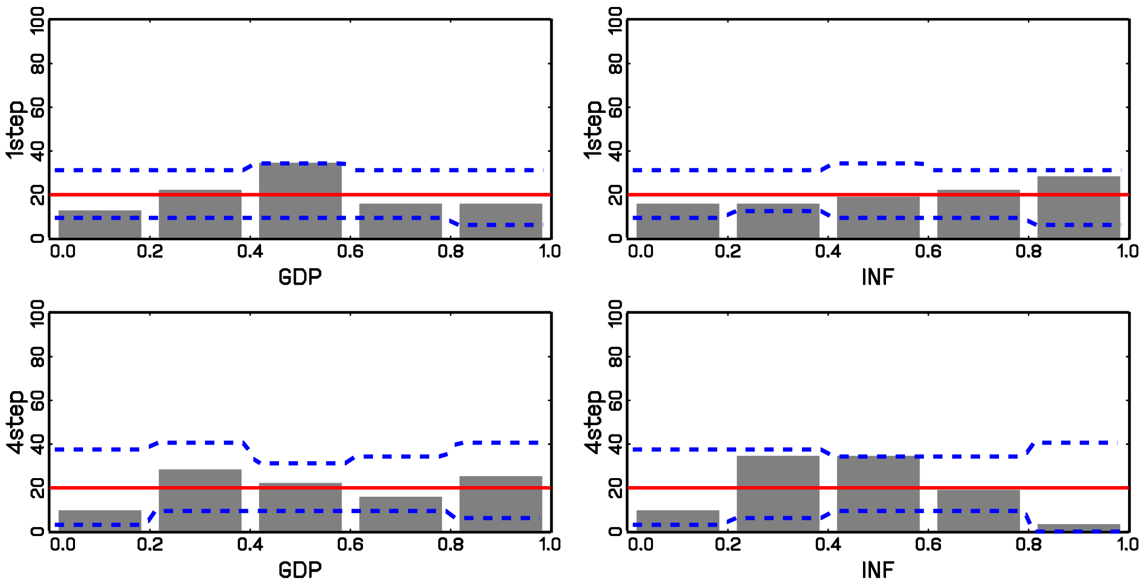
Smets-Wouters Model



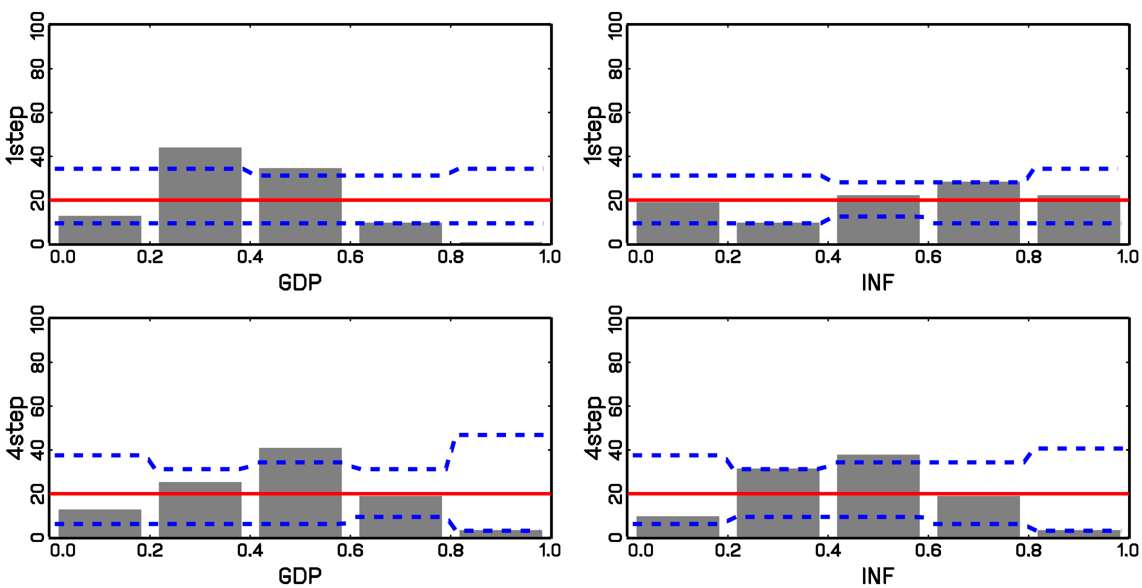
Notes: Probability integral transforms for forecasts of inflation (*INF*) and interest rates (*FFR*) conditional on actual future output growth. *Bars* correspond to actuals, and *dashed bands* indicate 90% credible intervals obtained from the predictive distribution.

Figure 6: PIT Histograms – Forecasts Given Future Interest Rates

Small-Scale Model



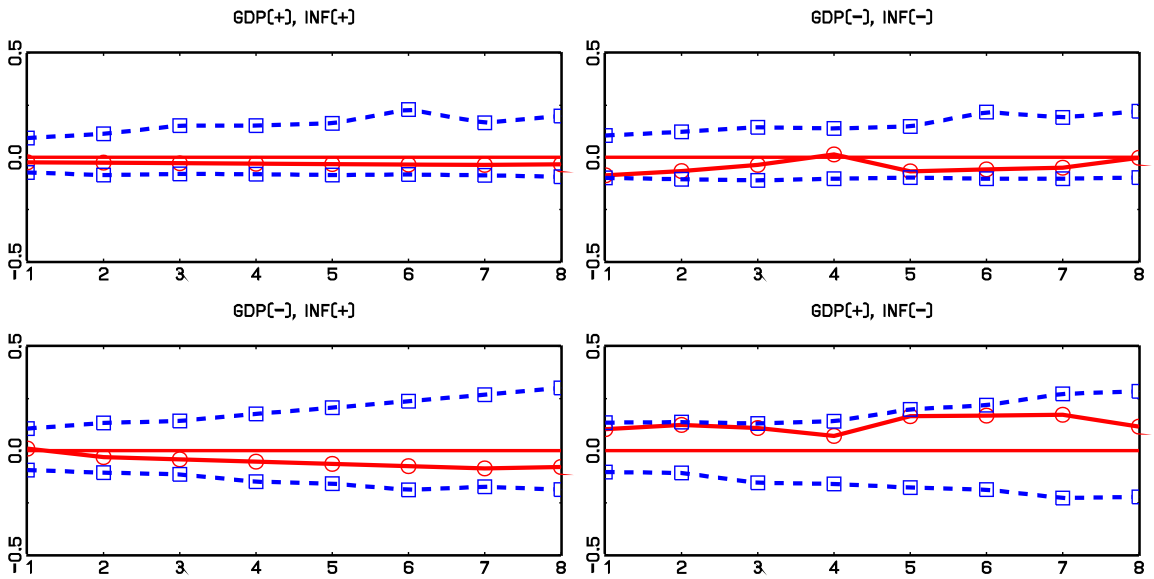
Smets-Wouters Model



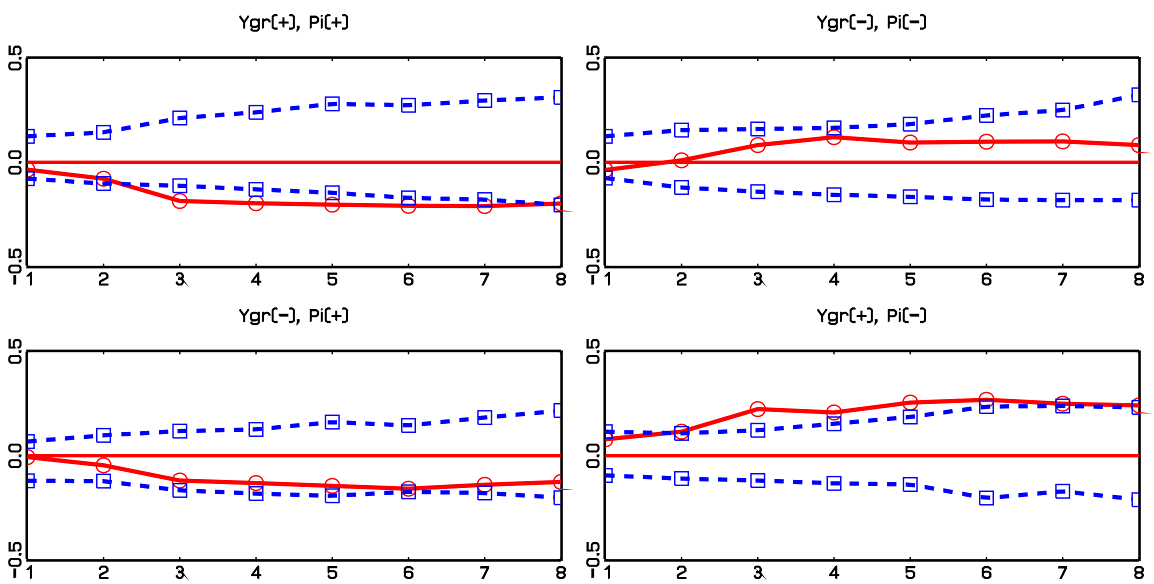
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and inflation (*INF*) conditional on actual future interest rates. *Bars* correspond to actuals, and *dashed bands* indicate 90% credible intervals obtained from the predictive distribution.

Figure 7: Event Occurrence minus Event Probabilities

Small-Scale Model



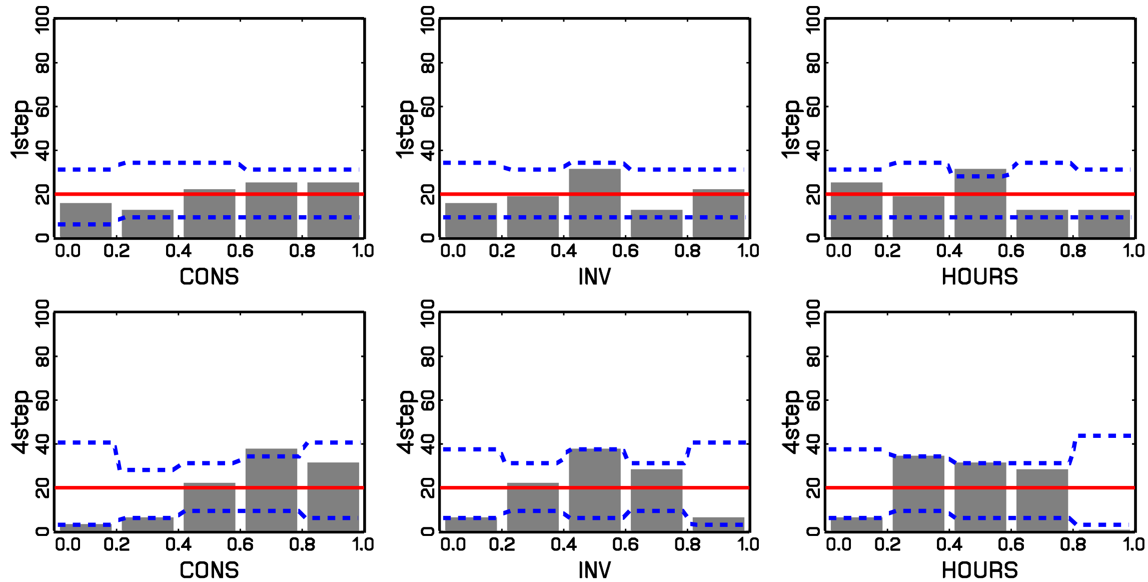
Smets-Wouters Model



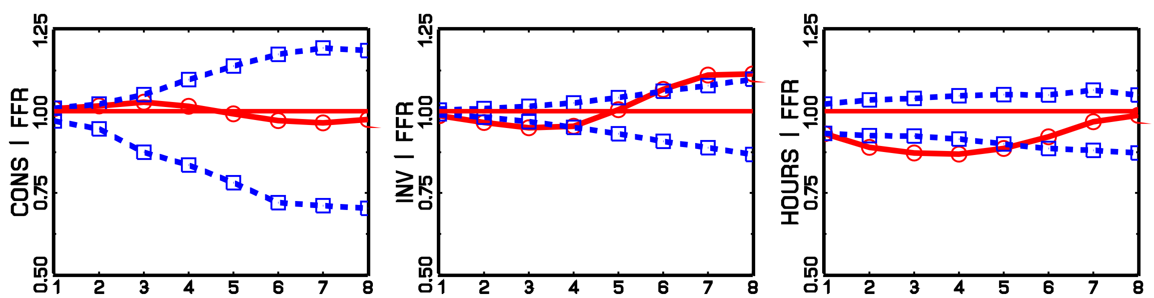
Notes: The *solid line* corresponds to differences associated with actual forecasts, and the *dashed line* signifies 90% credible intervals obtained from the predictive distribution.

Figure 8: Smets-Wouters Model: Consumption, Investment, Hours

PITs Histograms – Forecasts Given Future Interest Rates



RMSE Ratios of Conditional and Unconditional Forecasts



Notes: Top Panels: Probability integral transforms for forecasts of consumption growth (*CONS*), investment growth (*INV*), and hours worked (*HOURS*) conditional on actual future interest rates. *Bars* correspond to actuals, and *dashed bands* indicate 90% credible intervals obtained from the predictive distribution. Bottom panels: The *solid line* corresponds to RMSE ratios associated with actual forecasts, and the *dashed line* signifies 90% credible intervals obtained from the predictive distribution.

A Approximating PITs Based on the Output of a Posterior Simulator

We need to approximate the integral

$$\int_{-\infty}^{y_{t+1}} p(\tilde{y}_{t+1}|Y_{1:t})d\tilde{y}_{t+1}. \tag{A.1}$$

To do so, we use draws from the predictive distribution $p(\tilde{y}_{t+1}|Y_{1:t})$. Let $\{\tilde{y}_{t+h}^{(i)}\}_{i=1}^{n_{sim}}$ be a set of such draws. The approximation of the integral takes the form

$$z_{t,h} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \mathcal{I}\{\tilde{y}_{t+h}^{(i)} < y_{t+h}\}. \tag{A.2}$$

For step-size h, we can repeat this process from $t = T_0 + 1$ until $t = T_0 + S$, yielding a set of S probability transforms of h-step ahead forecasts.

We are also interested in probability integral transforms based on conditional predictive distributions, $p(y_{1,t+h}|y_{2,t+h}, Y_{1:t})$. We approximate this conditional distribution by kernel density estimation and compute the probability integral transform as follows:

$$z_{1,t,h} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \kappa \left(\frac{\tilde{y}_{2,t+h}^{(i)} - y_{2,t+h}}{b} \right) \mathcal{I}\{\tilde{y}_{1,t+h}^{(i)} < y_{1,t+h}\} \tag{A.3}$$

where b is the bandwidth and κ is the normalized kernel. For step-size h, we can repeat this process from $t = T_0 + 1$ until $t = T_0 + S$, yielding a set of S probability transforms of h-step ahead forecasts.

B The Smets-Wouters Model

The equilibrium conditions of the Smets and Wouters (2007) model take the following form:

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + r^{kss} k_y \hat{z}_t + \varepsilon_t^g \tag{A.4}$$

$$\begin{aligned} \hat{c}_t = {} & \frac{h/\gamma}{1+h/\gamma}\hat{c}_{t-1} + \frac{1}{1+h/\gamma}E_t\hat{c}_{t+1} + \frac{wl^{ss}(\sigma_c - 1)}{c^{ss}\sigma_c(1+h/\gamma)}(\hat{l}_t - E_t\hat{l}_{t+1}) \\ & - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}(\hat{r}_t - E_t\hat{\pi}_{t+1}) - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon_t^b \end{aligned} \tag{A.5}$$

$$
\begin{aligned}
\hat{i}_t &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}_{t+1} + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}_t + \varepsilon_t^i && \text{(A.6)}\\
\hat{q}_t &= \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}_{t+1} - \hat{r}_t + E_t\hat{\pi}_{t+1} + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t\hat{r}^k_{t+1} - \varepsilon_t^b && \text{(A.7)}\\
\hat{y}_t &= \Phi(\alpha\hat{k}^s_t + (1-\alpha)\hat{l}_t + \varepsilon^a_t) && \text{(A.8)}\\
\hat{k}^s_t &= \hat{k}_{t-1} + \hat{z}_t && \text{(A.9)}\\
\hat{z}_t &= \frac{1-\psi}{\psi}\hat{r}^k_t && \text{(A.10)}\\
\hat{k}_t &= \frac{(1-\delta)}{\gamma}\hat{k}_{t-1} + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon^i_t && \text{(A.11)}\\
\hat{\mu}^p_t &= \alpha(\hat{k}^s_t - \hat{l}_t) - \hat{w}_t + \varepsilon^a_t && \text{(A.12)}\\
\hat{\pi}_t &= \frac{\beta\gamma^{(1-\sigma_c)}}{1+\iota_p\beta\gamma^{(1-\sigma_c)}}E_t\hat{\pi}_{t+1} + \frac{\iota_p}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_{t-1} && \text{(A.13)}\\
&\quad - \frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_p)(1-\xi_p)}{(1+\iota_p\beta\gamma^{(1-\sigma_c)})(1+(\varphi-1)\varepsilon_p)\xi_p}\hat{\mu}^p_t + \varepsilon^p_t\\
\hat{r}^k_t &= \hat{l}_t + \hat{w}_t - \hat{k}_t && \text{(A.14)}\\
\hat{\mu}^w_t &= \hat{w}_t - \sigma_l\hat{l}_t - \frac{1}{1-h/\gamma}(\hat{c}_t - h/\gamma\hat{c}_{t-1}) && \text{(A.15)}\\
\hat{w}_t &= \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}(E_t\hat{w}_{t+1} + E_t\hat{\pi}_{t+1}) + \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}(\hat{w}_{t-1} - \iota_w\hat{\pi}_{t-1}) && \text{(A.16)}\\
&\quad - \frac{1+\beta\gamma^{(1-\sigma_c)}\iota_w}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_t + \frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_w)(1-\xi_w)}{(1+\beta\gamma^{(1-\sigma_c)})(1+(\lambda_w-1)\epsilon_w)\xi_w}\hat{\mu}^w_t + \varepsilon^w_t && \text{(A.17)}\\
\hat{r}_t &= \rho\hat{r}_{t-1} + (1-\rho)(r_\pi\hat{\pi}_t + r_y(\hat{y}_t - \hat{y}^*_t)) + r_{\Delta y}((\hat{y}_t - \hat{y}^*_t) - (\hat{y}_{t-1} - \hat{y}^*_{t-1})) + \varepsilon^r_t && \text{(A.18)}\\
\varepsilon^a_t &= \rho_a\varepsilon^a_{t-1} + \eta^a_t && \text{(A.19)}\\
\varepsilon^b_t &= \rho_b\varepsilon^b_{t-1} + \eta^b_t && \text{(A.20)}\\
\varepsilon^g_t &= \rho_g\varepsilon^a_{t-1} + \rho_{ga}\eta^a_t + \eta^g_t && \text{(A.21)}\\
\varepsilon^i_t &= \rho_i\varepsilon^i_{t-1} + \eta^i_t && \text{(A.22)}\\
\varepsilon^r_t &= \rho_r\varepsilon^r_{t-1} + \eta^r_t && \text{(A.23)}\\
\varepsilon^p_t &= \rho_r\varepsilon^p_{t-1} + \eta^p_t - \mu_p\eta^p_{t-1} && \text{(A.24)}\\
\varepsilon^w_t &= \rho_w\varepsilon^w_{t-1} + \eta^w_t - \mu_w\eta^w_{t-1} && \text{(A.25)}\\
\hat{y}^*_t &= c_y\hat{c}^*_t + i_y\hat{i}^*_t + r^{kss}k_y\hat{z}^*_t + \varepsilon^g_t && \text{(A.26)}\\
\hat{c}^*_t &= \frac{h/\gamma}{1+h/\gamma}\hat{c}^*_{t-1} + \frac{1}{1+h/\gamma}E_t\hat{c}^*_{t+1} + \frac{wl^{ss}(\sigma_c-1)}{c^{ss}\sigma_c(1+h/\gamma)}(\hat{l}^*_t - E_t\hat{l}^*_{t+1}) && \text{(A.27)}\\
&\quad - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}r^*_t - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon^b_t
\end{aligned}
$$

$$\hat{i}^*_t = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}^*_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}^*_{t+1} + \frac{1}{\phi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}^*_t + \varepsilon^i_t \quad \text{(A.28)}$$

$$\hat{q}^*_t = \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}^*_{t+1} - r^*_t + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t r^{k*}_{t+1} - \varepsilon^b_t \quad \text{(A.29)}$$

$$\hat{y}^*_t = \Phi(\alpha k^{s*}_t + (1-\alpha)\hat{l}^*_t + \varepsilon^a_t) \quad \text{(A.30)}$$

$$\hat{k}^{s*}_t = k^*_{t-1} + z^*_t \quad \text{(A.31)}$$

$$\hat{z}^*_t = \frac{1-\psi}{\psi}\hat{r}^{k*}_t \quad \text{(A.32)}$$

$$\hat{k}_t = \frac{(1-\delta)}{\gamma}\hat{k}^*_{t-1} + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon^i_t \quad \text{(A.33)}$$

$$\hat{\mu}^{p*}_t = \alpha(\hat{k}^{s*}_t - \hat{l}^*_t) - \hat{w}^*_t + \varepsilon^a_t \quad \text{(A.34)}$$

$$\hat{\mu}^{p*}_t = 1 \quad \text{(A.35)}$$

$$\hat{r}^{k*}_t = \hat{l}^*_t + \hat{w}^*_t - \hat{k}^*_t \quad \text{(A.36)}$$

$$\hat{\mu}^{w*}_t = -\sigma_l\hat{l}^*_t - \frac{1}{1-h/\gamma}(\hat{c}^*_t + h/\gamma\hat{c}^*_{t-1}) \quad \text{(A.37)}$$

$$\hat{w}^*_t = \mu^{w*}_t, \quad \text{(A.38)}$$

where

$$\gamma = \bar{\gamma}/100 + 1 \quad \text{(A.39)}$$

$$\pi^* = \bar{\pi}/100 + 1 \quad \text{(A.40)}$$

$$\bar{r} = 100(\beta^{-1}\gamma^{\sigma_c}\pi^* - 1) \quad \text{(A.41)}$$

$$r^k_{ss} = \gamma^{\sigma_c}/\beta - (1-\delta) \quad \text{(A.42)}$$

$$w_{ss} = \left(\frac{\alpha^\alpha(1-\alpha)^{(1-\alpha)}}{(\phi r^k_{ss})^\alpha}\right)^{\frac{1}{1-\alpha}} \quad \text{(A.43)}$$

$$i_k = (1-(1-\delta)/\gamma)\gamma \quad \text{(A.44)}$$

$$l_k = \frac{1-\alpha}{\alpha}\frac{r^k_{ss}}{w_{ss}} \quad \text{(A.45)}$$

$$k_y = \phi l_k^{(\alpha-1)} \quad \text{(A.46)}$$

$$i_y = (\gamma - 1 + \delta)k_y \quad \text{(A.47)}$$

$$c_y = 1 - g_y - i_y \quad \text{(A.48)}$$

$$z_y = r^k_{ss}k_y. \quad \text{(A.49)}$$

(A.50)