Evaluating DSGE Model Forecasts of Comovements

Frank Schorfheide*
University of Pennsylvania
CEPR and NBER

Edward Herbst
University of Pennsylvania

October 17, 2010

Preliminary Draft

Abstract

This paper develops and applies tools to assess multivariate aspects of DSGE model forecasts and their ability to predict comovements among key macroeconomic variables.

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

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

*Correspondence: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu (F. Schorfheide); herbstep@sas.upenn.edu (E. Herbst). 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.

1 Introduction

One key advantage of dynamic stochastic general equilibrium (DSGE) models over vector autoregressions (VARs) and the reason that central banks are considering them for projections and policy analysis is, that DSGE models use modern macroeconomic theory to explain and predict comovements of aggregate time series over the business cycle. Historical observations can be decomposed into contributions of the underlying exogenous disturbances, such as technology, preference, government spending, or monetary policy shocks. Future paths of the endogenous variables can be constructed conditional on particular realizations of monetary policy shocks that reflect potential future nominal interest rate paths. Unfortunately, these desirable attributes of DSGE model forecasts are not reflected in the root-mean-squared error (RMSE) measures that have been predominantly used to assess the forecast performance of DSGE models.

This paper develops and applies tools to assess multivariate aspects of DSGE model forecasts and their ability to predict comovements between key macroeconomic variables. DSGE models are typically analyzed with Bayesian methods, which can be easily used to generate draws from the model-implied h-step ahead predictive distribution for a vector of key macroeconomic variables. These predictive distributions reflect uncertainty about latent state variables, parameters, and future realizations of shocks. 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.[1]

More specifically, we examine whether predictive densities obtained from DSGE models are well calibrated in the sense that probability integral transforms (PITs) are uniformly distributed. The use of PITs has a long tradition in density forecast evaluation, which we will discuss in more detail below. The novel aspect of our evaluations is to construct statistics that measure whether the predicted change in forecast accuracy due to conditioning

[1] Bayarri and Berger (2004) refer to this notion as *frequentist principle*.

on future realizations of a subset of the model variables matches the actual change. We use these statistics to assess the models' ability to predict comovements among macroeconomic variables, which is one of the stated goals of DSGE modeling.

For a three equation New Keynesian DSGE model, we find that the marginal and conditional one-step, two-step, and four-step predictive densities for output and inflation are well calibrated, but do not observe a reduction in RMSE associated with conditioning on output or inflation. The predictive density for the interest rate is poorly calibrated, with the majority of realizations falling in the left tail of the distribution. When performing the same exercise with real time data, we find an that the predictive density for interest rates improves and the predictive density for inflation worsens. Now the predictive density for inflation is too wide, with too many realizations falling in the center of the distribution. The relationship between the interest rate and output is generally well calibrated and associated with reductions in RSME when forecasting using conditional predictive densities for short horizons, but performs poorly at longer horizons.

For the Smets-Wouters model, we find that the marginal predictive densities for inflation and output are not well calibrated in the sense the tails of the predictive densities are too large. That is, the realizations of output and inflation fall too frequently in the center of the predictive distribution. Conversely, the marginal predictive density for the Federal Funds funds appears well calibrated. In terms of comovements, there are large gains in conditioning on the federal funds rate for forecasting inflation but not for output growth as seem by the reduction of RMSEs. However, the predictive distributions of inflation and output growth given the Federal Funds Rate are too diffuse.

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.[2] Predictive densities are simply constructed by the relevant conditioning. Non-Bayesian approaches, surveyed in Corradi and Swanson

[2]Though there might be some uncertainty about the accuracy with which it is evaluated using computational methods.

(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 the goal 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. 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 policy makers at central banks shift their attention from point forecasts to interval and density forecasts.[3] 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 based on the predictive distributions associated with our DSGE models, 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 sequential probability forecasts for future observations. According to his *prequential principle*, any assessment of the overall adequacy of a sequence of forecasts should only depend on the forecasts actual made.[4] While there are many different ways of defining the adequacy of forecasts, Dawid (1984) suggested among other things to examine whether PITs are *iid* uniformly $\mathcal{U}[0,1]$ distributed.

[3]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).

[4]This principle asserts the irrelevance of hypothetical forecasts that might have been issued in circumstances that did not come about.

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 to 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 multi-step density forecasts should be both uniformly distributed as well as independent across time and provide a general proof. In addition to examining cumulative distribution functions, the authors also inspect autocorrelation functions of PITs (and nonlinear transformations of PITs) to assess density forecasts for 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 five univariate time series models of S&P500 portfolio returns.

Gneiting, Balabdaoui, and Raftery (2007) illustrate that uniformly distributed PITs do not imply that conditional mean forecasts are unbiased or that the predictive density is sharp. Thus, if a DSGE model passes out PIT-based predictive checks, it cannot be concluded that there do not exist other econometric models that provide superior forecasts. If PITs are found not to be uniformly distributed, they can in principle be re-calibrated. This idea was discussed in Dawid (1984) and applied to VAR forecasts by Kling and Bessler (1989). The latter use a sequential procedure to re-calibrate predictive distributions based on previously issued predictions and outcomes and found that the re-calibration improved the VAR forecasts. Diebold, Hahn, and Tay (1999) re-calibrated density forecasts of high-frequency returns on foreign exchange from JP Morgan's RiskMetrics. Due to the short sample that we have available for our empirical analysis we do not pursue a re-calibration of DSGE model forecasts in this paper.

The remainder of the paper is organized as follows. In Section 2 we briefly review empirical findings regarding the forecasting performance of DSGE models. Our econometric methodology is discussed 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.

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 root-mean-squared errors 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). In Table 1 we are summarizing pseudo-out-of-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, we report sample standard deviations over the respective forecast periods. 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.

Edge and Gürkaynak (2010) compare univariate forecasts from the Smets and Wouters (2007) 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. However, none of the procedures is particularly successful in forecasting large fluctuations. Overall, the predictability of inflation has been quite low over the past 10 to 15 years, because its persistence dropped substantially after the Great Moderation.

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 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 effected by the inability to forecast just one series, nominal wage growth.

3 Econometric Approach: Predictive Checks

In this paper we shall consider predictive checks based on posterior parameter distributions. We shall partition the actual sample into $Y_{1:R}$ and $Y_{R+1:T}$, defining $P = T - R$. 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

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

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 shall use predictive checks to assess the adequacy of model forecasts.

One-Step Ahead PITs: Denote the elements of the vector y_t by $y_{i,t}$, $i = 1, \ldots, n$ 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 (2004)'s *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 interval $[(j-1)/J, j/J]$.

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

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

Moreover, we are interested in assessing the conditional predictive distribution of $p(\tilde{y}_{i,t+h}|y_{1:(i-1),t+h}, Y_{1:R}, Y_{R+1:t})$, using

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

Conditional on $(Y_{1:R}, Y_{R+1:t})$ the marginal distribution of $z_{i,t,h}$ and $z_{i|1:(i-1),t,h}$ remain $\mathcal{U}[0,1]$. However, the sequence $\{z_{i,t,h}\}_{t=R}^{T-h}$ is serially dependent, which will be reflected in the posterior predictive distribution of this sequence and transformations such as (4).

RMSE ratios: Finally, we use RMSEs and ratios of conditional and unconditional RMSEs as in Schorfheide, Sill, and Kryshko (2010) to form a predictive check.

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 has 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. Moreover, our approach let us easily consider other statistics such as RMSEs or ratios of RMSEs. 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 DSGE model forecasts. 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 a 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

We will apply the predictive checks 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 is 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 model of Smets and Wouters (2007) (henceforth the SW model). The SW 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.

Agents' Decision Problems and Equilibrium Conditions. The perfectly competitive, representative, final good producing firm combines a continuum of intermediate goods indexed by $j \in [0, 1]$ using the technology

$$Y_t = \left(\int_0^1 Y_t(j)^{1-\nu} dj \right)^{\frac{1}{1-\nu}}. \tag{7}$$

Here $1/\nu > 1$ represents the elasticity of demand for each intermediate good. The firm takes input prices $P_t(j)$ and output prices P_t as given. Profit maximization implies that the demand for intermediate goods is

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t. \tag{8}$$

The relationship between intermediate goods prices and the price of the final good is

$$P_t = \left(\int_0^1 P_t(j)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}. \tag{9}$$

Intermediate good j is produced by a monopolist who has access to the following linear production technology:

$$Y_t(j) = A_t N_t(j), \tag{10}$$

where A_t is an exogenous productivity process that is common to all firms and $N_t(j)$ is the labor input of firm j. Labor is hired in a perfectly competitive factor market at the real wage W_t. Firms face nominal rigidities in terms of quadratic price adjustment costs

$$AC_t(j) = \frac{\phi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - \pi\right)^2 Y_t(j), \tag{11}$$

where ϕ governs the price stickiness in the economy and π is the steady state inflation rate associated with the final good. Firm j chooses its labor input $N_t(j)$ and the price $P_t(j)$ to maximize the present value of future profits

$$\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s Q_{t+s}\left(\frac{P_{t+s}(j)}{P_{t+s}}Y_{t+s}(j) - W_{t+s}N_{t+s}(j) - AC_{t+s}(j)\right)\right]. \tag{12}$$

Here, Q_{t+s} is the marginal value of a unit of the consumption good to the household, which is treated as exogenous by the firm.

The representative household derives utility from real money balances M_t/P_t and consumption C_t relative to a habit stock. We assume that the habit stock is given by the level of technology A_t. This assumption ensures that the economy evolves along a balanced growth path even if the utility function is additively separable in consumption, real money balances, and leisure. The household derives disutility from hours worked H_t and maximizes

$$\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s\left(\frac{(C_{t+s}/A_{t+s})^{1-\tau}-1}{1-\tau} + \chi_M \ln\left(\frac{M_{t+s}}{P_{t+s}}\right) - \chi_H H_{t+s}\right)\right], \tag{13}$$

where β is the discount factor, $1/\tau$ is the intertemporal elasticity of substitution, and χ_M and χ_H are scale factors that determine steady state real money balances and hours worked. We will set $\chi_H = 1$. The household supplies perfectly elastic labor services to the firms taking the real wage W_t as given. The household has access to a domestic bond market where nominal government bonds B_t are traded that pay (gross) interest R_t. Furthermore, it receives aggregate residual real profits D_t from the firms and has to pay lump-sum taxes T_t. Thus, the household's budget constraint is of the form

$$P_tC_t + B_t + M_t - M_{t-1} + T_t = P_tW_tH_t + R_{t-1}B_{t-1} + P_tD_t \tag{14}$$

The usual transversality condition on asset accumulation applies, which rules out Ponzi schemes.

Monetary policy is described by an interest rate feedback rule of the form

$$R_t = R_t^{*\,1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}}, \tag{15}$$

where $\epsilon_{R,t}$ is a monetary policy shock and R_t^* is the (nominal) target rate, which evolves according to

$$R_t^* = r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}}\right)^{\psi_2} \quad . \tag{16}$$

Here r is the steady state real interest rate, π_t is the gross inflation rate defined as $\pi_t = P_t/P_{t-1}$, and π^* is the target inflation rate, which in equilibrium coincides with the steady state inflation rate.

The fiscal authority consumes a fraction ζ_t of aggregate output Y_t, where $\zeta_t \in [0,1]$ follows an exogenous process. The government levies a lump-sum tax (subsidy) to finance any shortfalls in government revenues (or to rebate any surplus). The government's budget constraint is given by

$$P_t G_t + R_{t-1} B_{t-1} = T_t + B_t + M_t - M_{t-1}, \tag{17}$$

where $G_t = \zeta_t Y_t$. The market clearing conditions are given by

$$Y_t = C_t + G_t + AC_t \quad \text{and} \quad H_t = N_t. \tag{18}$$

Exogenous Processes. The model economy is perturbed by three exogenous processes. Aggregate productivity evolves according to

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t, \quad \text{where} \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}. \tag{19}$$

Thus, on average technology grows at the rate γ and z_t captures exogenous fluctuations of the technology growth rate. Define $g_t = 1/(1-\zeta_t)$. We assume that

$$\ln g_t = (1-\rho_g)\ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}. \tag{20}$$

Finally, the monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated. The three innovations are independent of each other at all leads and lags and are normally distributed with means zero and standard deviations σ_z, σ_g, and σ_R, respectively.

Log-Linearized Equilibrium Conditions and Measurement. Since the non-stationary technology process A_t induces a stochastic trend in output and consumption, it is convenient to express the model in terms of detrended variables $c_t = C_t/A_t$ and $y_t = Y_t/A_t$. The model economy has a unique steady state in terms of the detrended variables that is attained if the innovations $\epsilon_{R,t}$, $\epsilon_{g,t}$, and $\epsilon_{z,t}$ are zero at all times. The steady state inflation π equals the target rate π^*, $r = \frac{\gamma}{\beta}$, and $R = r\pi^*$. Let $\hat{x}_t = \ln(x_t/x)$ denote the percentage deviation of a variable x_t from its steady state x. Then the log-linearized equilibrium conditions can be summarized by a (consumption) Euler equation, a New Keynesian Phillips curve, and an interest rate feedback rule:

$$\begin{aligned}
\hat{y}_t &= \mathbb{E}_t[\hat{y}_{t+1}] + \hat{g}_t - \mathbb{E}_t[\hat{g}_{t+1}] - \frac{1}{\tau}\left(\hat{R}_t - \hat{\mathbb{E}}_t[\pi_{t+1}] - \mathbb{E}_t[\hat{z}_{t+1}]\right) \\
\hat{\pi}_t &= \beta\mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t) \\
\hat{R}_t &= \rho_R\hat{R}_{t-1} + (1-\rho_R)\psi_1\hat{\pi}_t + (1-\rho_R)\psi_2\left(\Delta\hat{y}_t + \hat{z}_t\right) + \epsilon_{R,t}
\end{aligned} \tag{21}$$

where $\kappa = \tau(1-\nu)/(\nu\pi^2\phi)$. The loglinearized exogenous shock processes take the form

$$\hat{g}_t = \rho_g\hat{g}_{t-1} + \epsilon_{g,t}, \quad \hat{z}_t = \rho_z\hat{z}_{t-1} + \epsilon_{z,t}. \tag{22}$$

The model is completed by defining 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). The three series are measured in percentages and their relationship to the model variables is given by the following set of equations:

$$\begin{aligned}
YGR_t &= \gamma^{(Q)} + \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \\
PI_t &= \pi^{(A)} + 4\hat{\pi}_t \\
INT_t &= \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 4\hat{R}_t.
\end{aligned} \tag{23}$$

The parameters $\gamma^{(Q)}$, $\pi^{(A)}$, and $r^{(A)}$ are related to the steady states of the model economy as follows

$$\gamma = 1 + \frac{\gamma^{(Q)}}{100}, \quad \beta = \frac{1}{1 + r^{(A)}/400}, \quad \pi = 1 + \frac{\pi^{(A)}}{400}.$$

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 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.

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 run 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 $\tau = 0, \ldots, P$. 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 second model considered in this paper is the one by Smets and Wouters (2007). 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 Appendix.

Measurement Equations: We estimate the SW model 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:

$$\begin{aligned} YGR_t &= \bar{\gamma} + \hat{y}_t - \hat{y}_{t-1} \\ PI_t &= \bar{\pi} + \hat{\pi}_t \\ INT_t &= \bar{r} + \hat{R}_t \\ CGR_t &= \bar{\gamma} + \hat{c}_t - \hat{c}_{t-1} \\ IGR_t &= \bar{\gamma} + \hat{i}_t - \hat{i}_{t-1} \\ WGR_t &= \bar{\gamma} + \hat{w}_t - \hat{w}_{t-1} \\ HOURS_t &= \bar{l} + \hat{l}_t. \end{aligned} \tag{24}$$

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 fixing the following 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.

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 present results for the small-scale DSGE model (Section 5.2) and finally we examine the predictive performance of the Smets-Wouters model (Section 5.3). The results of the predictive checks reported below are based on 50 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 $(Y_{1:R}, Y^*_{R+1:T})$. 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 $(Y_{1:R}, Y_{R+1:T})$.

5.1 Data Set

For the evaluation of the density forecasts we 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 DSGE model and Greenbook forecast we only use 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 8 periods ahead for March origins, 7 periods ahead for June origins, 10 periods ahead for September origins, and 9 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.[5] 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. In addition to a forecast evaluation based on real time data, we also construct a sequence of pseudo-out-of-sample forecasts for the small-scale model based on the data set associated with the December 2004 forecasts. We refer to these pseudo-out-of-sample forecasts as last-vintage as opposed to real-time forecasts. Finally, 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

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

terms when estimating the small-scale DSGE model.

5.2 Evidence from the Small-Scale Model

Empirical results are presented in terms of histograms for PITs, e.g. Figure 1, plots of RMSEs (or RMSE ratios) as a function of the forecast horizon, e.g. Figure 11, or tables with χ^2 goodness-of-fit statistics, e.g. Table 4. 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 of the subintervals. 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. In the RMSE plots the solid lines signify the actual RMSEs or RMSE ratios and the dashed lines indicate the 5th and 95th percentile of the predictive distribution. Finally, in the tables we report the values of the χ^2 statistic computed based on actual data as well as the 95th percentile of the predictive distribution. We first examine the results obtained from the last vintage of observations and then discuss the results from the real time data.

Last Vintage. The small-scale DSGE model generates a joint predictive distribution for output growth, inflation, and interest rates. We begin the density forecast evaluation by considering the marginal predictive distributions for these three variables. PIT histograms for one-step, two-step, and four-step ahead density forecasts are displayed in Figure 1. The inflation and output growth forecasts appear to be well calibrated as the percentages of PITs computed from actual data fall inside the intervals computed based on the predictive distribution. This is confirmed by the χ^2 statistics reported in Table 4, which measure the squared distance of the bin heights from the 20% line. All actuals except for the goodness-of-fit statistic computed for the four-step-ahead inflation forecasts are below the 95th percentile of their respective predictive distribution. The interest rate (Federal Funds rate) density forecasts, especially two and four steps ahead, are less well calibrated. Many interest rate realizations all into the left tail of the density forecasts leading to a fraction of PITs in the [0, 0.2] that is larger than the estimated model predicts.

We now consider bivariate predictive distributions generated by the small-scale DSGE model. Since a joint density of two random variables can be factored into a marginal and a conditional density and we have examined the marginal densities already, we shall now focus on density forecasts that are conditioned on future realizations of either output growth, inflation, and interest rates. If the DSGE model is able to capture the comovements between pairs of the three observables, then probability integrals computed from the conditional density of say output growth given actual future inflation are also uniformly distributed, which is shown formally in Diebold, Hahn, and Tay (1999). The results are summarized in Figures 2 to 4. Given the relatively poor performance of the marginal interest rate forecasts, it is not surprising that the forecasts conditional on future realizations of output growth or inflation are not well calibrated either. The forecasts of output given inflation and inflation given output, on the other hand, appear well calibrated. The corresponding χ^2 statistics in Table 4 computed from the actual forecasts are all below the 95th percentile of the predictive distribution.

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 the short-term nominal interest rate. Thus, Figure 4 depicts PIT histograms for output and inflation forecasts that condition on the actual future interest rate. It is important to note that these histograms do not convey any information about the DSGE model's ability to generate accurate counterfactual policy predictions. We only examine whether the DSGE model is able to capture the comovement between interest rate and other macroeconomic variables under the actual policy. The results are mixed at best. Judging from the χ^2 statistics in Table 4 the output growth forecasts conditional on interest rates pass the predictive check at all horizons, whereas the inflation forecasts only passes at the one-step horizon.

Finally, we turn to predictive checks based on RMSEs, which are summarized in Figures 5 and 6. The RMSEs are computed for point predictions, which are in turn based on the mean of the predictive distribution and thereby optimal under a quadratic loss function. To obtain forecasts conditional on future realizations of one of the other variables, we compute the mean

from the conditional predictive density. In a nutshell, if the predictive distribution 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. Figure 5 depicts actual RMSEs as well as the 5th and 95th percentile of the predictive distribution. The RMSEs for the output growth forecasts are around 0.5, which for one-step-ahead forecasts is in line with the values reported in Table 1. For the output growth forecasts the actual RMSEs lie within the bands computed from the predictive distribution. The small-scale DSGE model implies that conditioning on realized values of inflation should reduce the RMSE of output growth forecasts, which is more evident from the top left panel of Figure 6. This RMSE reduction however is not realized by the actual forecasts. The realized ratio of RMSEs is one or slightly larger, whereas the 5th percentile of the predictive distribution is about 0.75 for four-step-ahead forecasts. The evidence from the predictive check for RMSEs of inflation and interest rate forecasts is less favorable for the DSGE model. The estimated model assigns a lot of probability to inflation (interest rate) forecast RMSEs that are higher (lower) than the actual RMSE. Conditioning on one of the other two variables worsens the inflation forecasts and leaves the RMSE of the interest rate forecast essentially unchanged.[6]

Real Time Data. We now turn to results from the real-time data set. It is important to re-emphasize that we do not try to capture the data revision process with the measurement equation of the DSGE model and therefore data revisions are not reflected in predictive distribution that we are simulating. Figure 7 displays the PITs of unconditional forecasts for the small model using real time data and χ^2 statistics are reported in Table 5. The one-step-ahead forecasts (nowcasts) for all three variable appear to be well calibrated. At longer horizons, the distribution of the PITs for inflation and the Federal Funds rate deteriorates slightly. The actual interest rate falls too frequently in the left tail of the predictive distribution, while realizations of the inflation are dispersed much less than would be indicated by the predictive distribution. That is, the volatility of inflation implied by the model is too high, which is also evident from the predictive distribution for the inflation RMSE depicted in Figure 11. The realized RMSEs for inflation (and at longer horizons output) fall outside

[6]Unlike in Table 1 the inflation forecast RMSEs depicted in Figure 5 are based on *annualized* quarter-over-quarter inflation rates.

the 90% credible intervals of the predictive distribution. However, the realized RMSEs for the interest rate falls well inside the credible intervals, despite the poor left tail behavior of the predictive density.

Figure 8 shows the distribution of the PITs conditional on future output. The distributions for the PITs of output and interest rates look very similar to the distributions of the unconditional PITs. The one-step-ahead forecasts look well calibrated, while the four-step ahead conditional predictive distribution for inflation is too conservative. Evidently, conditioning on interest rates does not improve the models ability capture the volatility of the inflation rate. Figures 9 and 10 display the distribution of the PITs conditional on inflation and interest rates, respectively. Once again, the one-step-ahead forecasts look well-calibrated, while at longer horizons the distributions for output growth and inflation given interest rates and for interest rates given inflation to be mis-calibrated.

Turning to the Figure 11, the actual RMSEs errors for inflation and to lesser extent output fall in the lower tail of the predictive distributions. However, the actual RMSEs for interest rates are consistent with those implied by the model. Compared to the last-vintage data set (Figure 5), this behavior is reversed. Furthermore, the actual RMSEs are in general lower with real time data. Figure 12 displays the ratio of the conditional and unconditional RMSEs. Conditional on R, the ratio of the RMSEs falls well outside the predictive distribution for output and too a lesser extent inflation. At longer horizons, the actual interest rate often lies in the extreme tail of the distribution while output and inflation are moderate. With real-time data, there are actual gains in conditioning on realized output and inflation, whereas with the final vintage data set there were no gains.

5.3 Evidence from the Smets-Wouters Model

The Smets-Wouters model generates a joint predictive distribution for output, consumption, inflation, investment, hours, wages, the federal funds rates. We consider only model predictions for output, inflation, and the federal funds rate. The PITs histograms of the one-step, two-step, and four-step ahead marginal predictive density forecasts are displayed in Figure 13. The nowcasts (one-step forecast) for inflation and the federal funds seem to

be well calibrated, but the nowcast for GDP appears to less well calibrated. Evidently, for the two-step and four-step forecasts, the marginal predictive density for output and inflation is too disperse. Conversely, the predictive density for the Federal Funds rate appear well calibrated. The χ^2 statistics in Table 6 are consistent with these observations, with the realized statistic for output growth well above the 95th percentile of the predictive distribution at all horizons. The statistic for inflation exceeds the 95th percentile of the predictive distribution for two-step and four-step ahead forecasts. The χ^2 statistic associated with the Federal Funds rate does not exceed the 90th percentile of the predictive distribution at any horizon.

Figure 14 shows the histograms of PITs of the forecasts given future output. The nowcasts are well calibrated, but at all longer horizons there are too few realizations of inflation in the tails of the distribution. On the other hand, the conditional forecasts for the Federal Funds rate appear well calibrated. The χ^2 statistics in Table 6 confirm the conclusions draw from the histograms of the PITs. Figure 15 shows the histogram of the PITs of forecasts given future inflation. The conditional predictive density for output given inflation is poorly calibrated at all horizons. Essentially, the predictive distribution of output given inflation implied by the model is too disperse relative the to volatility of the data. This problem also occurs with the marginal predictive distribution for output, as seen in Figure 13. The conditional predictive density for interest rates given inflation is well calibrated. Turning to Figure 16, the histogram of conditional PITs given the Federal Funds Rate, the nowcast for inflation appears to the only well calibrated forecast. In general the features of these conditional predictive densities are similar to the marginal predictive distribution.

Next we consider the RMSE for each of the three series, along with the predictive distribution for these statistics constructed in the same way as the small-scale model RMSE predictive distribution. The first row of Figure 17 displays the actual RMSEs for up to eight-step forecasts based on the marginal predictive densities as well as the 5th and 95th percentile of the predictive distribution. As before, the actual RMSEs errors are consistent with those in Table 1. The one-step RMSEs for are 0.45 for output growth, 0.21 for inflation,

and 0.46 for the Federal Funds rate.[7] The RMSEs are essentially identical to those reported by Edge and Gürkaynak (2010), which suggests that our fixing of additional parameters did not adversely affect the fit of the model. The RMSEs for output and inflation are below 5th of their predictive distributions, consistent with the earlier observations that the predictive densities for these variables are too wide. Conversely, the RMSEs for the federal funds appear well within the bands given by the predictive distribution.

Finally, Figure 18 displays ratio of the conditional to unconditional RMSEs. We see that there is a substantial reduction in RMSE when forecasting inflation given the Federal Funds Rate or vice versa. This is consistent with the predictive distribution of the RMSEs. There is not any gain in conditioning on output for forecasting inflation or the interest rate; the predictive distribution indicates that one should not expect large reductions in RMSE from knowing GDP. Similar findings hold for forecasting GDP conditioning on inflation and the Federal Funds rate.

6 Conclusion

We have evaluated the forecasting performance of two DSGE models by examining whether the predictive densities are well calibrated in the sense that probability integral transforms are uniformly distributed. To accomplish this, we implemented predictive checks for the uniformity of the PITs for both the marginal and conditional predictive densities. We also examined the behavior of RMSEs of point forecasts for unconditional and conditional predictions against the distribution of RMSEs implied by the posterior.

For a three equation New Keynesian DSGE model, we find that the marginal and conditional one-step, two-step, and four-step predictive densities for output and inflation are well calibrated, but do not observe a reduction in RMSE associated with conditioning on output or inflation. The predictive density for the interest rate is poorly calibrated, with the majority of realizations falling in the left tail of the distribution. When performing the same exercise with real time data, we find an that the predictive density for interest rates

[7] Now the inflation rate is given in Q-o-Q percentages.

improves and the predictive density for inflation worsens. Now the predictive density for inflation is too wide, with too many realizations falling in the center of the distribution. The relationship between the interest rate and output is generally well calibrated and associated with reductions in RSME when forecasting using conditional predictive densities for short horizons, but performs poorly at longer horizons.

For the Smets-Wouters model, we find that the marginal predictive densities for inflation and output are not well calibrated in the sense the tails of the predictive densities are too large. That is, the realizations of output and inflation fall too frequently in the center of the predictive distribution. Conversely, the marginal predictive density for the Federal Funds funds appears well calibrated. In terms of comovements, there are large gains in conditioning on the federal funds rate for forecasting inflation but hardly any gains for forecasting output growth as seen by the reduction of RMSEs. However, the predictive distributions of inflation and output growth given the Federal Funds Rate are too diffuse.

This paper develops tools to assess the adequacy of the forecasting performance of DSGE models, with an emphasis on capturing the comovements of key macroeconomic variables. We hope that these predictive checks can be used to spur new thinking about DSGE models, particularly in the 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 (2008): "Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities)," *Journal of Monetary Economics*, 55(7), 1191–1208.

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.

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

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.

Gneiting, T., F. Balabdaoui, and A. Raftery (2007): "Probabilistic Forecasts, Calibration, and Sharpness," *Journal of the Royal Statistical Society, Series B*, 69(2), 243–268.

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.

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 (A %)
Del Negro *et al.* (2007)	1985:IV to 2000:I	0.73	0.27	0.87
		(0.52)	(0.25)	(1.72)
Smets, Wouters (2007)	1990:I to 2004:IV	0.57	0.24	0.43
		(0.57)	(0.22)	(1.97)
Edge, Kiley, Laforte (2009)	1996:III to 2004:IV	0.45	0.29	0.83
		(0.57)	(0.20)	(1.96)
Edge, Gürkaynak (2010)	1997:I to 2006:III	0.49	0.21	0.48$^{(*)}$
		(0.56)	(0.23)	(1.89)
Schorfheide, Sill, Kryshko (2010)	2001:I to 2007:IV	0.51	0.22	0.71
		(0.47)	(0.22)	(1.68)

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 forecast period, computed from the Schorfheide, Sill, Kryshko data set. Q % is the quarter-to-quarter percentage change, and A % is an annualized rate.

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; 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; 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 Tests for PITs (Small Model, Final Vintage)

	Output Growth		Inflation		Interest Rates	
			1 Quarter Ahead			
Uncond.	1.44	(8.31)	3.94	(10.81)	10.50	(10.19)
Cond on Y			3.63	(9.88)	12.38	(8.31)
Cond on π	1.75	(7.06)			10.50	(10.81)
Cond on R	5.50	(10.19)	3.94	(12.06)		
			2 Quarters Ahead			
Uncond.	3.31	(10.81)	10.19	(19.25)	20.81	(17.38)
Cond on Y			3.31	(15.81)	24.56	(13.63)
Cond on π	1.75	(8.31)			20.81	(18.94)
Cond on R	9.25	(11.13)	12.69	(21.44)		
			4 Quarters Ahead			
Uncond.	5.58	(16.83)	28.50	(21.42)	24.75	(21.83)
Cond on Y			5.58	(16.42)	28.50	(17.67)
Cond on π	4.75	(10.58)			23.50	(25.58)
Cond on R	12.67	(15.17)	23.08	(21.42)		

Notes: The values in parentheses correspond to the 90th percentile of the predictive distribution.

Table 5: χ^2 Tests for PITs (Small Model, Real Time)

	Output Growth		Inflation		Interest Rates	
			1 Quarter Ahead			
Uncond.	2.69	(8.63)	1.75	(10.19)	8.00	(9.25)
Cond on Y			1.75	(10.19)	12.69	(8.94)
Cond on π	3.31	(11.44)			6.44	(9.56)
Cond on R	4.88	(8.31)	1.75	(9.56)		
			2 Quarters Ahead			
Uncond.	13.31	(12.06)	31.12	(18.00)	17.69	(20.81)
Cond on Y			22.06	(17.69)	12.06	(21.44)
Cond on π	3.94	(11.13)			19.25	(23.31)
Cond on R	7.69	(10.19)	26.44	(16.44)		
			4 Quarters Ahead			
Uncond.	21.42	(14.75)	32.25	(23.92)	7.67	(28.92)
Cond on Y			25.58	(22.67)	11.83	(31.42)
Cond on π	14.75	(13.08)			6.42	(31.42)
Cond on R	13.50	(14.33)	32.25	(23.92)		

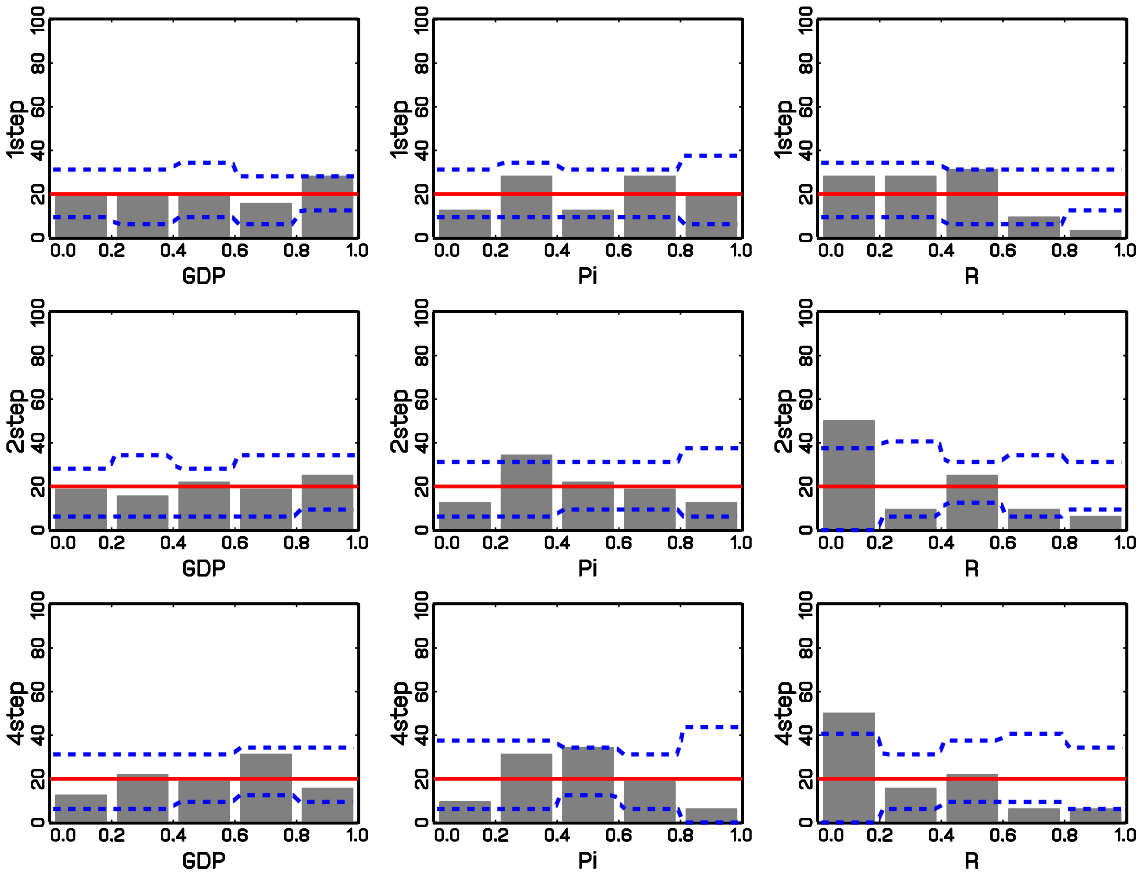
Notes: The values in parentheses correspond to the 90th percentile of the predictive distribution.

Table 6: χ^2 Tests for PITs (Smets Wouters, Real Time)

	Output Growth		Inflation		Interest Rates	
			1 Quarter Ahead			
Uncond.	23.31	(18.31)	3.62	(7.69)	8.94	(9.25)
Cond on Y			4.25	(9.88)	8.94	(10.19)
Cond on π	22.06	(16.44)			8.00	(10.19)
Cond on R	21.44	(14.56)	3.00	(7.06)		
			2 Quarters Ahead			
Uncond.	26.44	(9.25)	65.81	(18.62)	12.06	(29.88)
Cond on Y			54.25	(15.50)	10.19	(29.88)
Cond on π	24.56	(5.81)			10.19	(18.62)
Cond on R	27.69	(5.19)	27.69	(4.88)		
			4 Quarters Ahead			
Uncond.	28.92	(4.75)	77.25	(17.67)	9.75	(40.17)
Cond on Y			42.25	(17.25)	9.75	(35.58)
Cond on π	31.00	(3.08)			7.67	(23.08)
Cond on R	36.42	(7.25)	13.08	(7.67)		

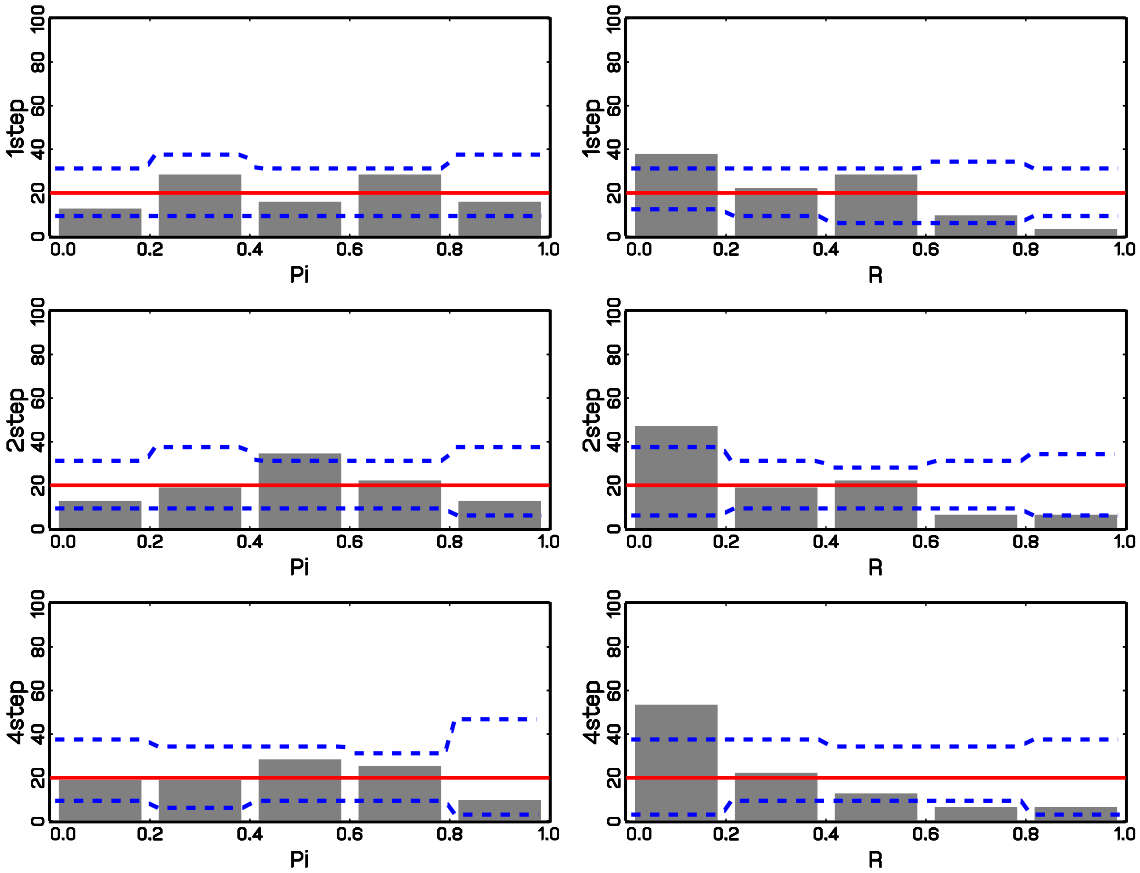
Notes: The values in parentheses correspond to the 90th percentile of the predictive distribution.

Figure 1: PITs of Unconditional Forecasts (Small Model, Last Vintage)



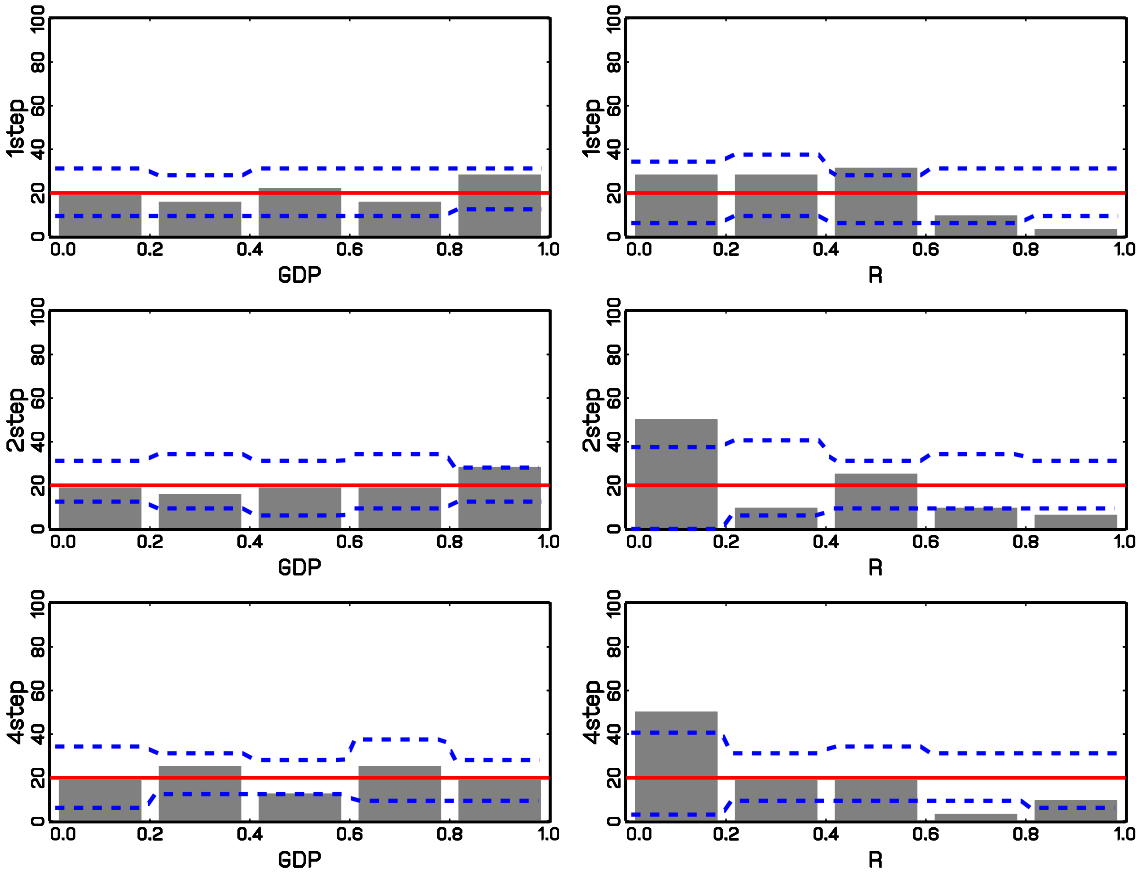
Notes: Probability integral transforms for forecasts of output growth (GDP), inflation (PI), and interest rates (R). *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 2: PITs of Forecasts Given Future Output (Small Model, Last Vintage)



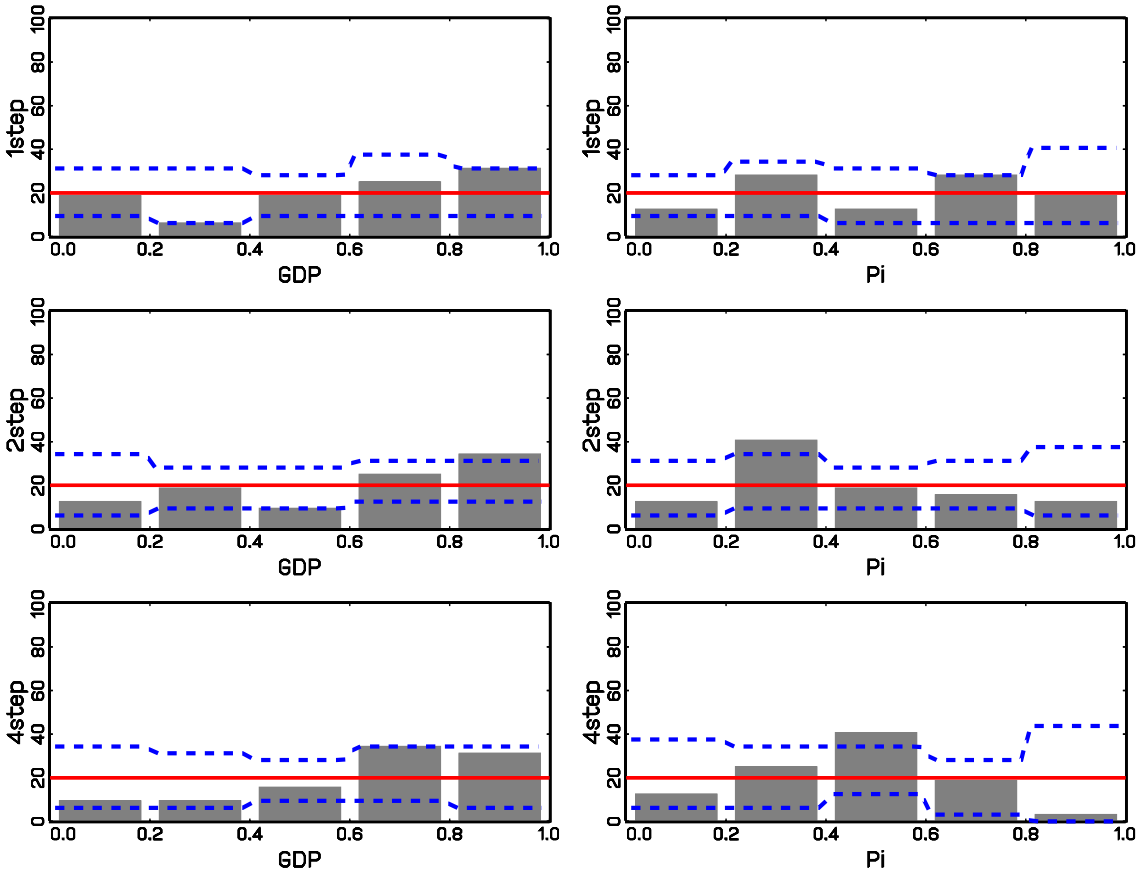
Notes: Probability integral transforms for forecasts of inflation (PI) and interest rates (R) conditional on actual future output growth. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 3: PITs of Forecasts Given Future Inflation (Small Model, Last Vintage)



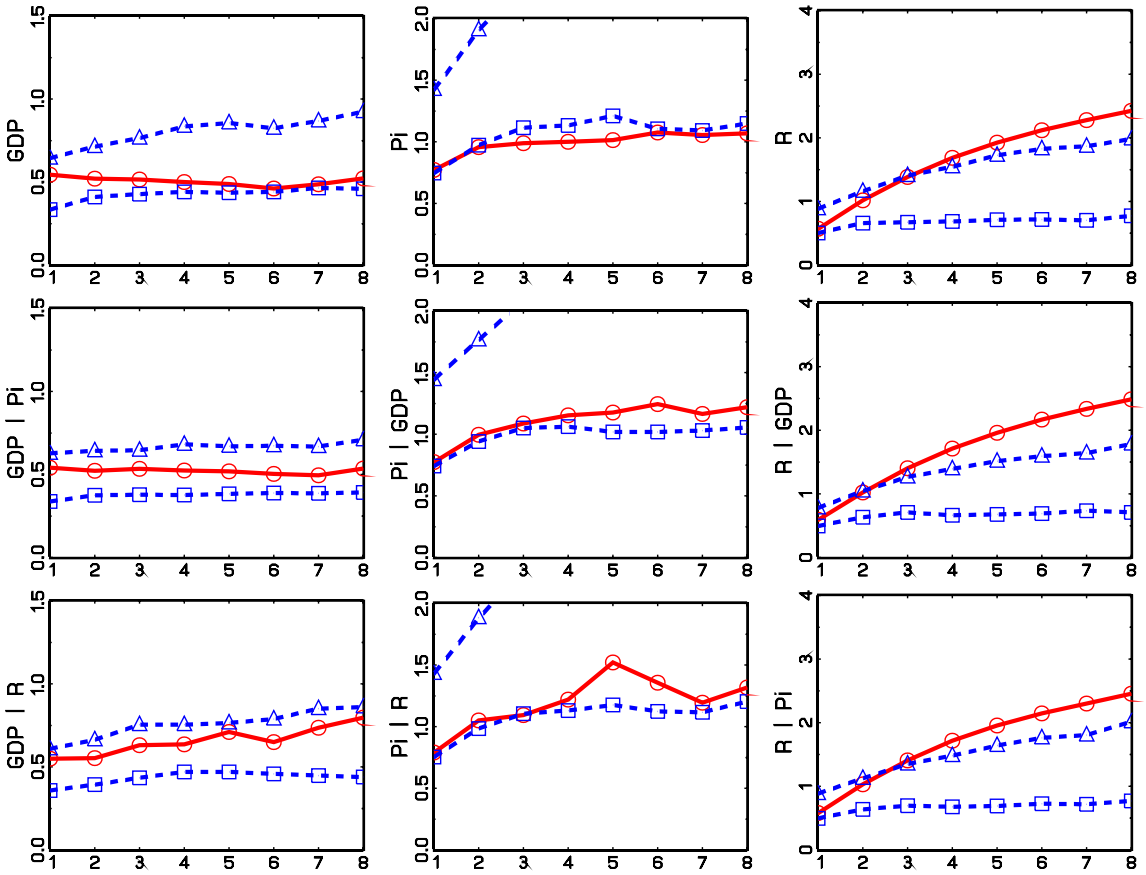
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and interest rates (*R*) conditional on actual future inflation. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 4: PITs of Forecasts Given Future Interest Rates (Small Model, Last Vintage)



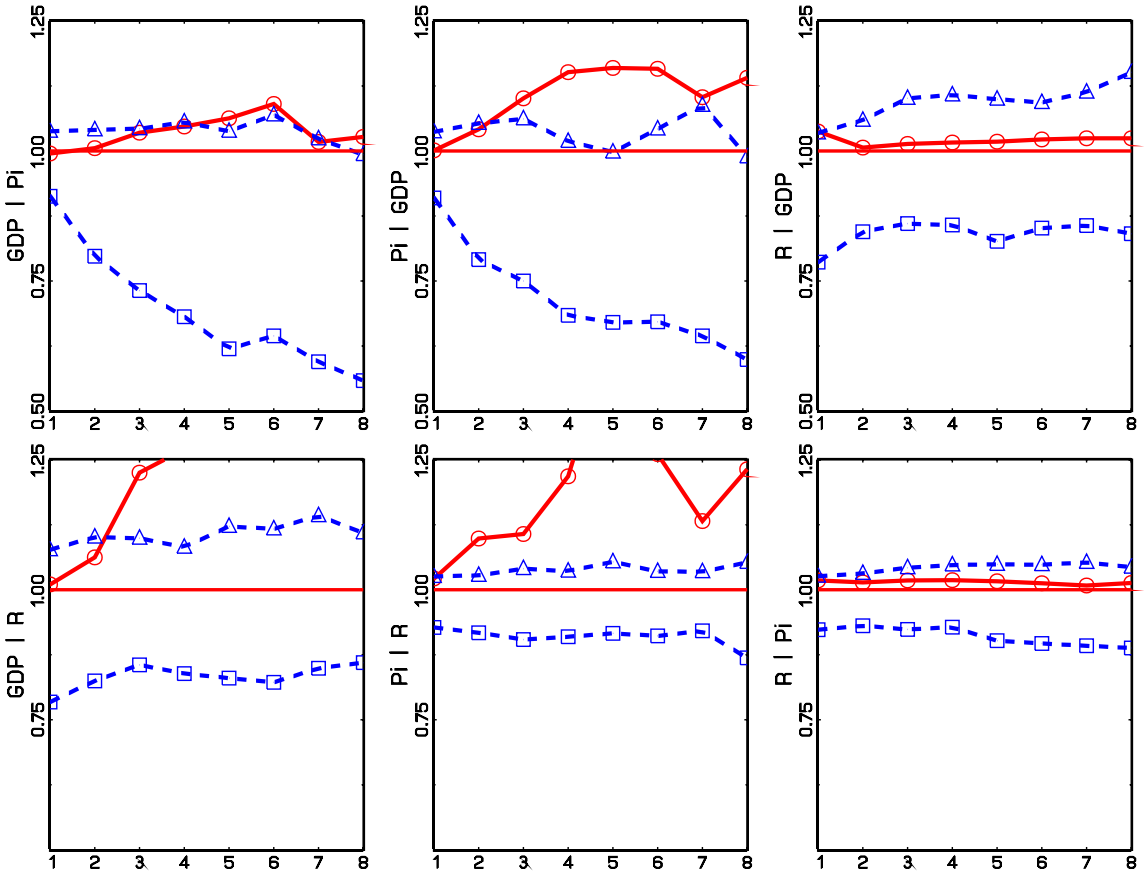
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and inflation (*PI*) conditional on actual future interest rates. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 5: RMSEs of Unconditional and Conditional Forecasts (Small Model, Last Vintage)



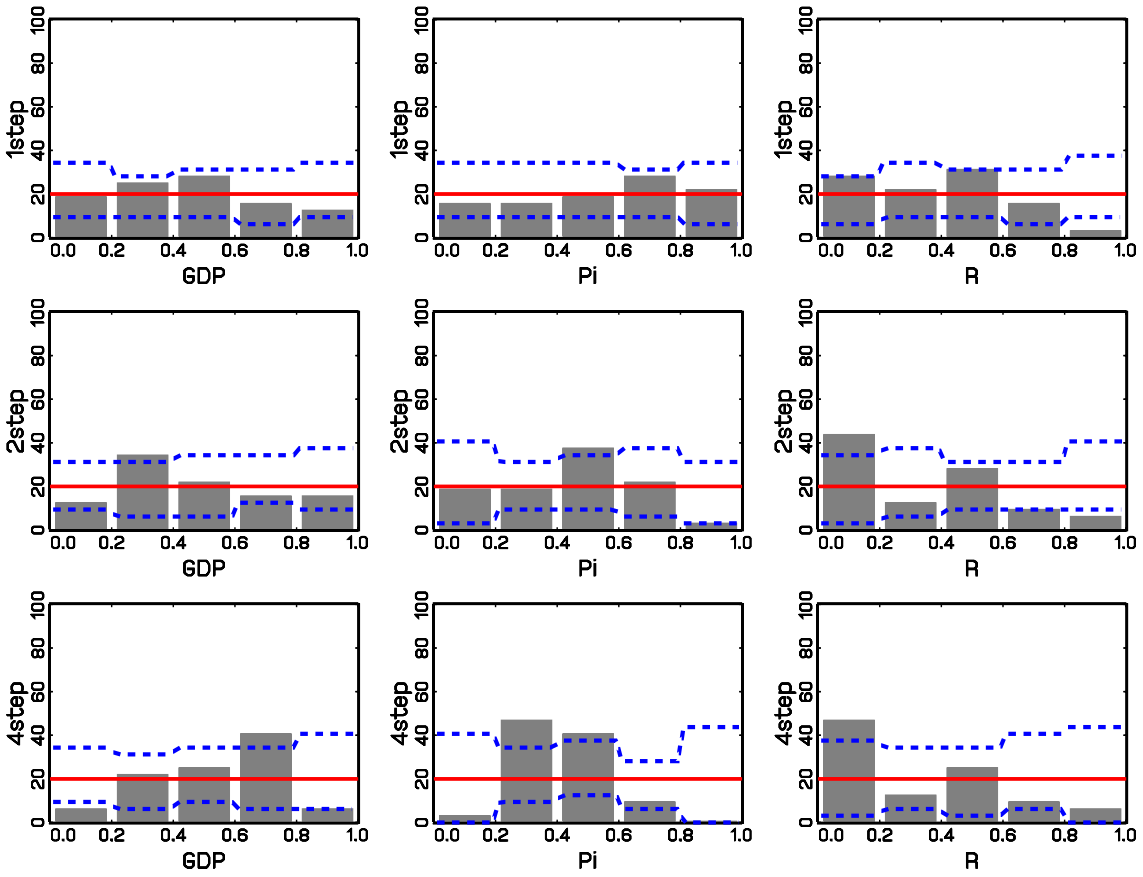
Notes: Root Mean Squared Errors (RMSEs) for forecasts of output growth (*GDP*, not annualized), inflation (*PI*, annualized), and interest rates (*R*, annualized) unconditional and conditional on future realizations of other variables. The solid line corresponds to RMSEs associated with actual forecasts and the dashed line signifies 90% credible intervals obtained from the predictive distribution.

Figure 6: RMSE Ratios of Unconditional and Conditional Forecasts (Small Model, Last Vintage)



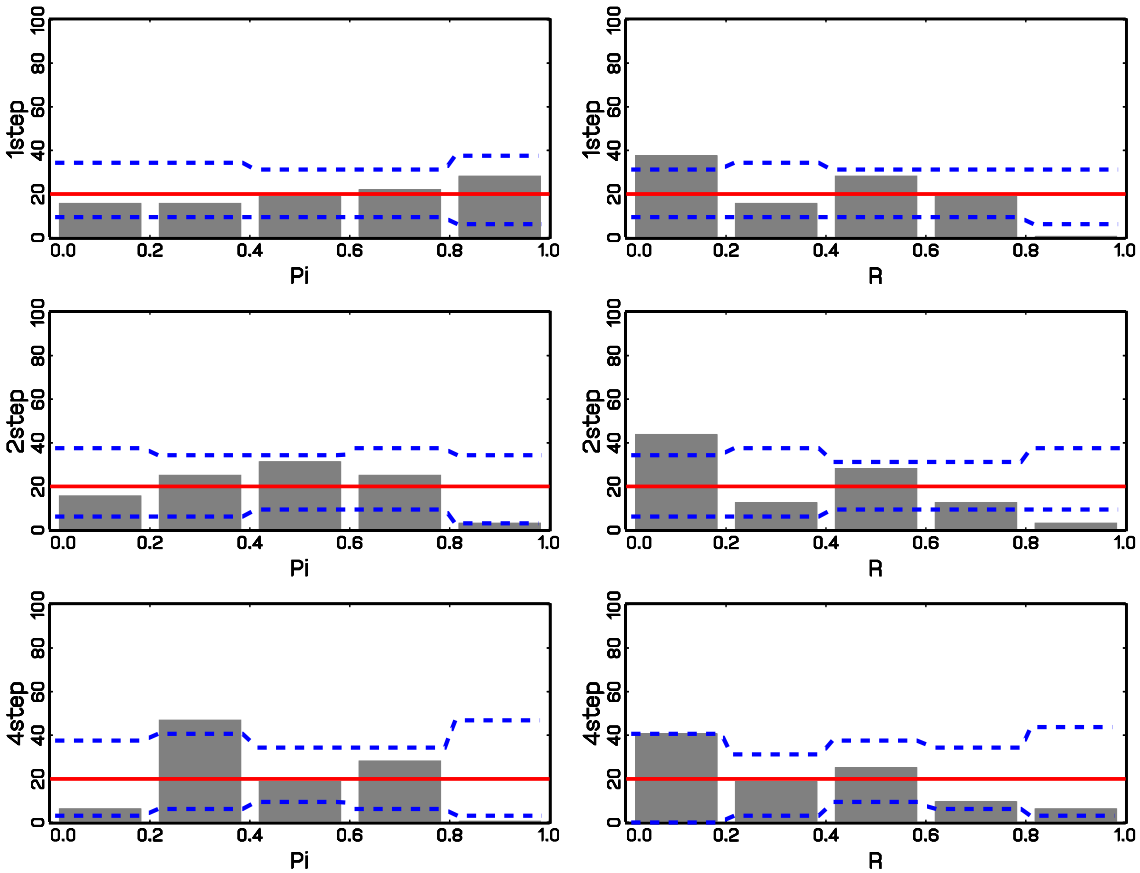
Notes: RMSE ratios for forecasts of output growth (GDP), inflation (PI), and interest rates (R): conditional on future realizations of other variables versus unconditional. The solid line corresponds to RMSE ratios associated with actual forecasts and the dashed line signifies 90% credible intervals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

Figure 7: PITs of Unconditional Forecasts (Small Model, Real Time)



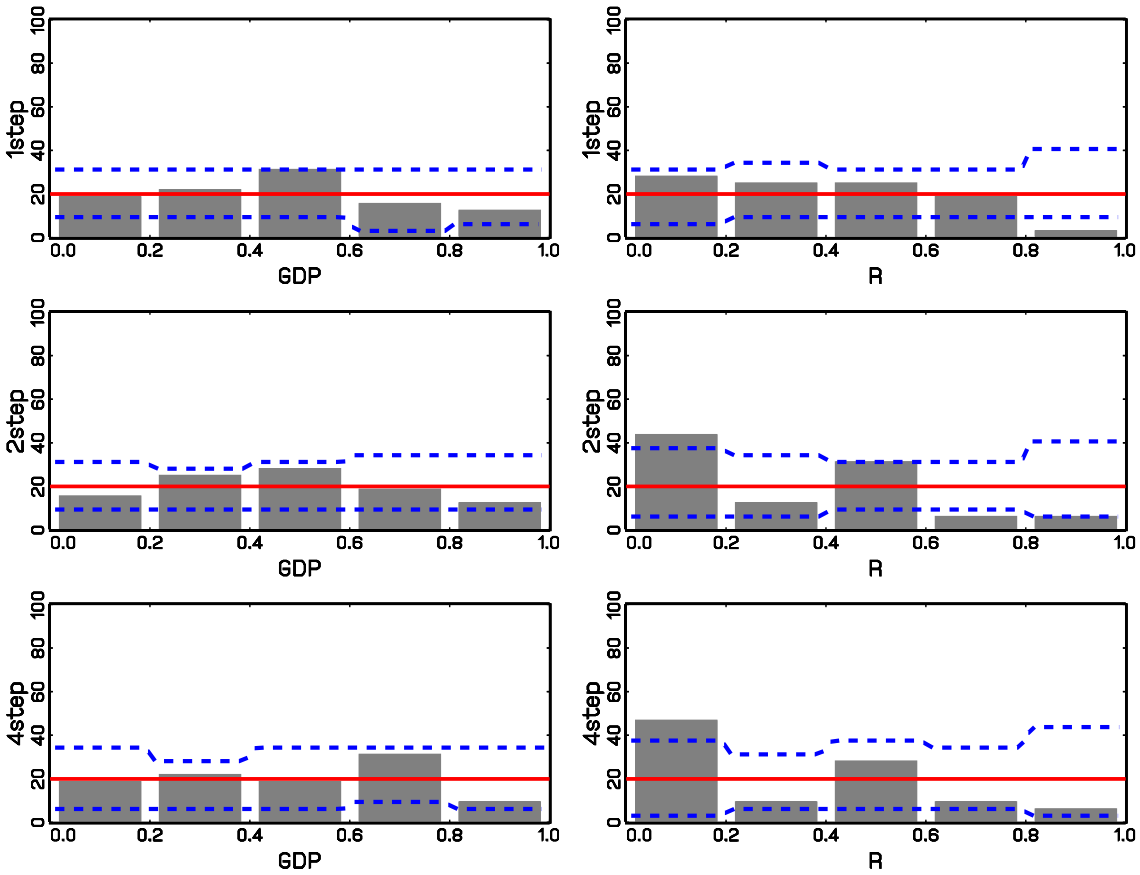
Notes: Probability integral transforms for forecasts of output growth (GDP), inflation (PI), and interest rates (R). *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 8: PITs of Forecasts Given Future Output (Small Model, Real Time)



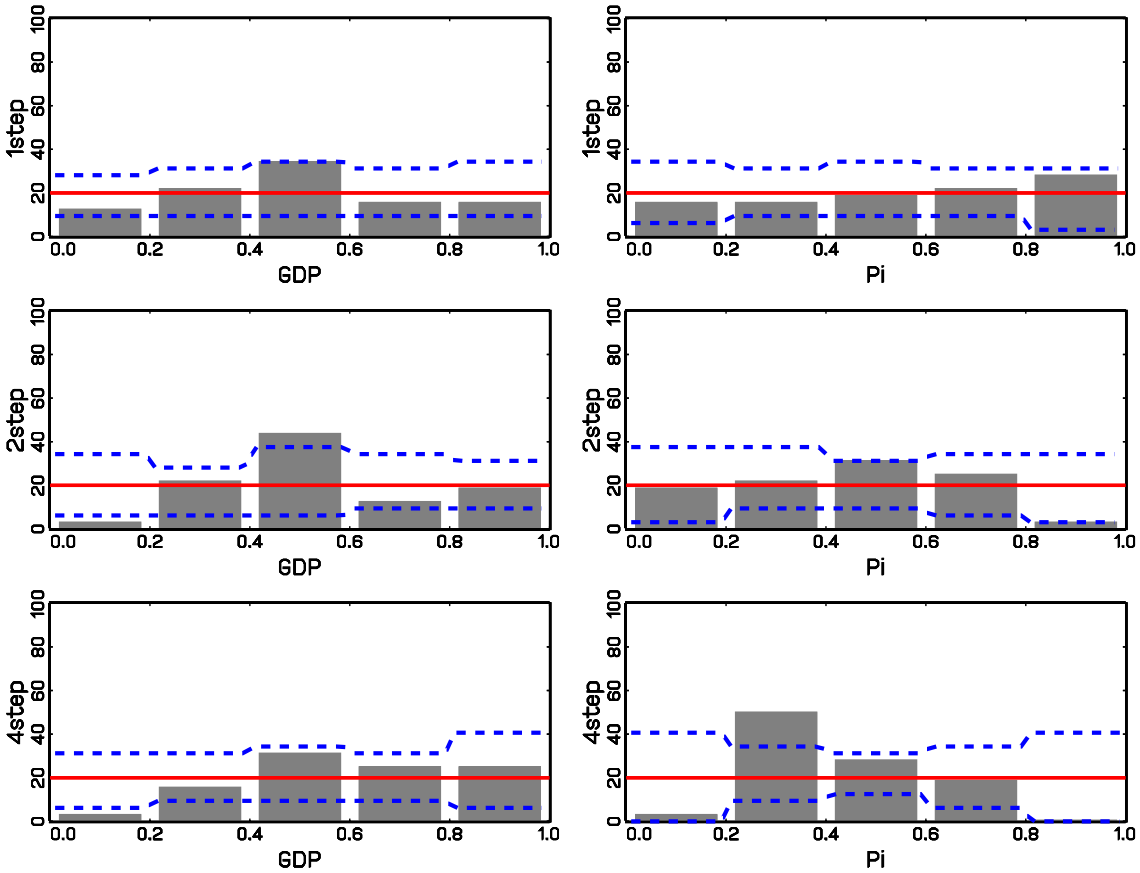
Notes: Probability integral transforms for forecasts of inflation (PI) and interest rates (R) conditional on actual future output growth. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 9: PITs of Forecasts Given Future Inflation (Small Model, Real Time)



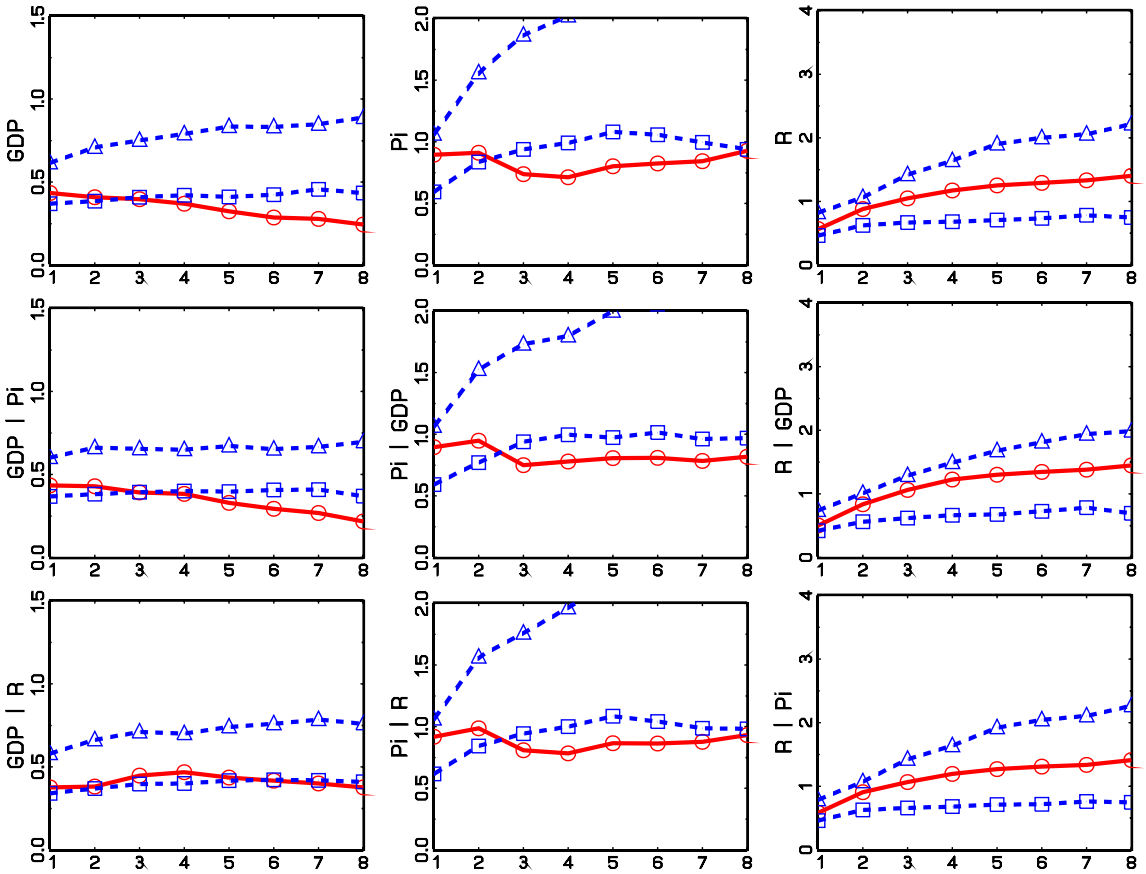
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and interest rates (*R*) conditional on actual future inflation. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 10: PITs of Forecasts Given Future Interest Rates (Small Model, Real Time)



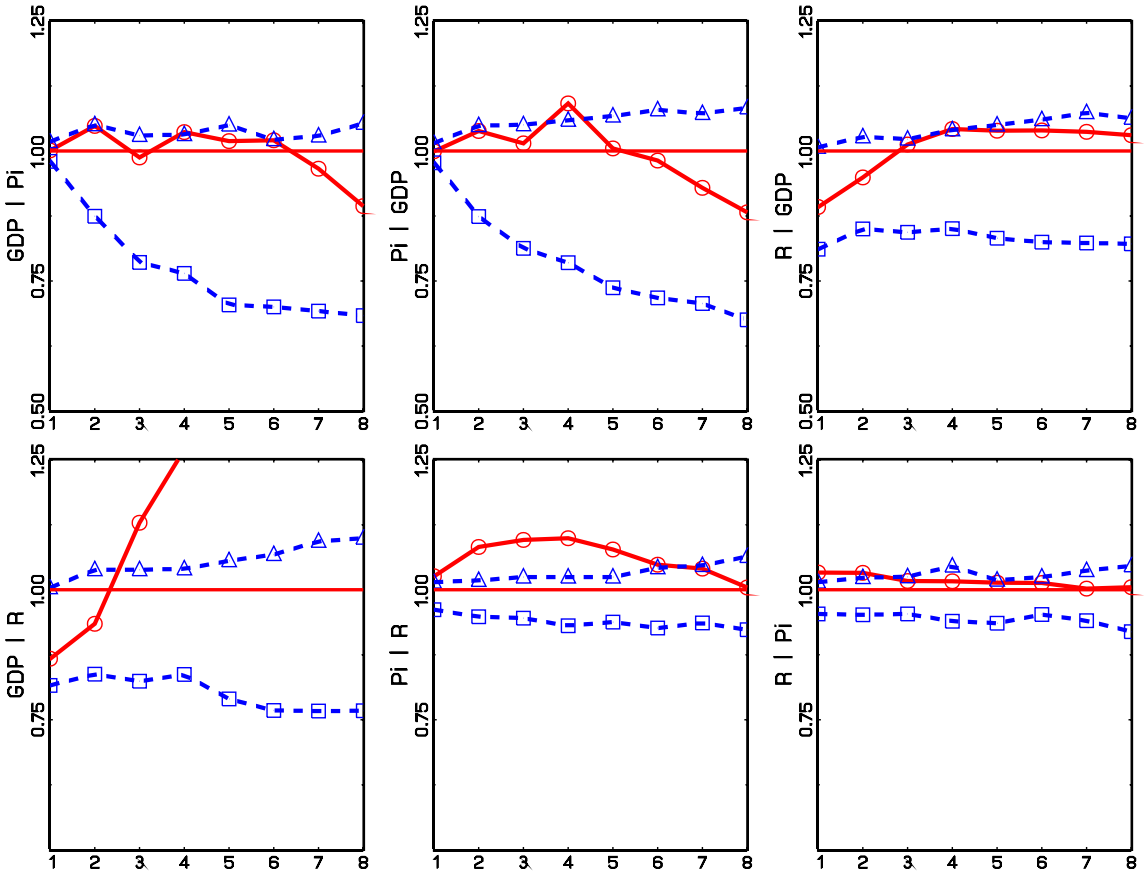
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and inflation (*PI*) conditional on actual future interest rates. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 11: RMSEs of Unconditional and Conditional Forecasts (Small Model, Real Time)



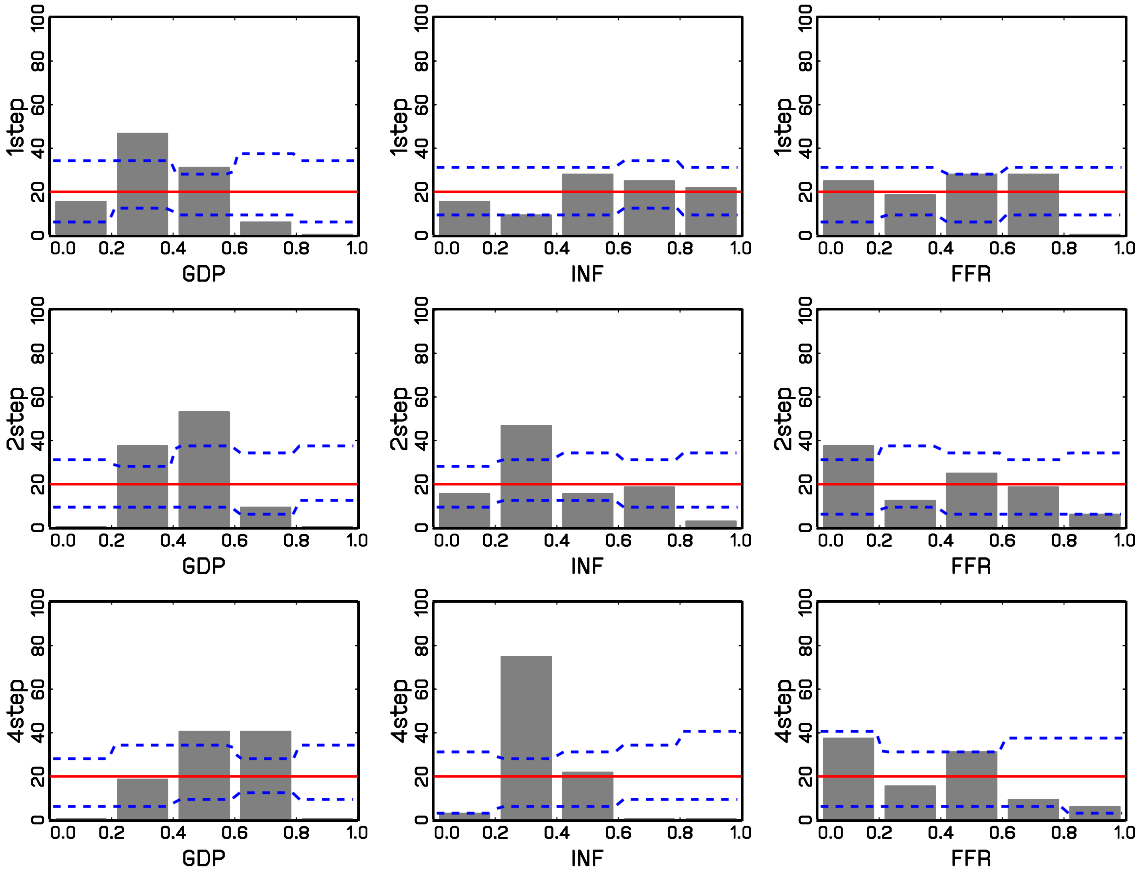
Notes: Root Mean Squared Errors (RMSEs) for forecasts of output growth (*GDP*, not annualized), inflation (*PI*, annualized), and interest rates (*R*, annualized) unconditional and conditional on future realizations of other variables. The solid line corresponds to RMSEs associated with actual forecasts and the dashed line signifies 90% credible intervals obtained from the predictive distribution.

Figure 12: RMSE Ratios of Unconditional and Conditional Forecasts (Small Model, Real Time)



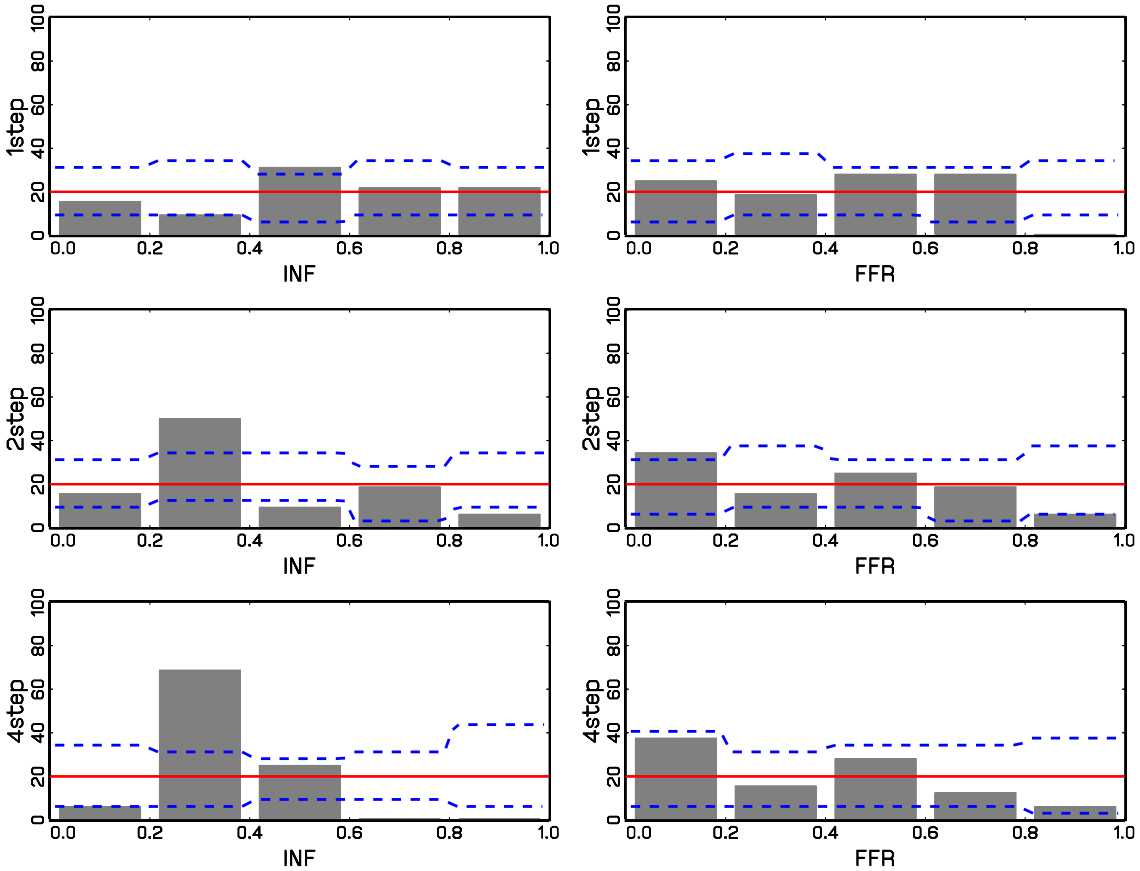
Notes: RMSE ratios for forecasts of output growth (GDP), inflation (PI), and interest rates (R): conditional on future realizations of other variables versus unconditional. The solid line corresponds to RMSE ratios associated with actual forecasts and the dashed line signifies 90% credible intervals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

Figure 13: PITs of Unconditional Forecasts for Output, Inflation and Interest Rates (Smets Wouters, Real Time)



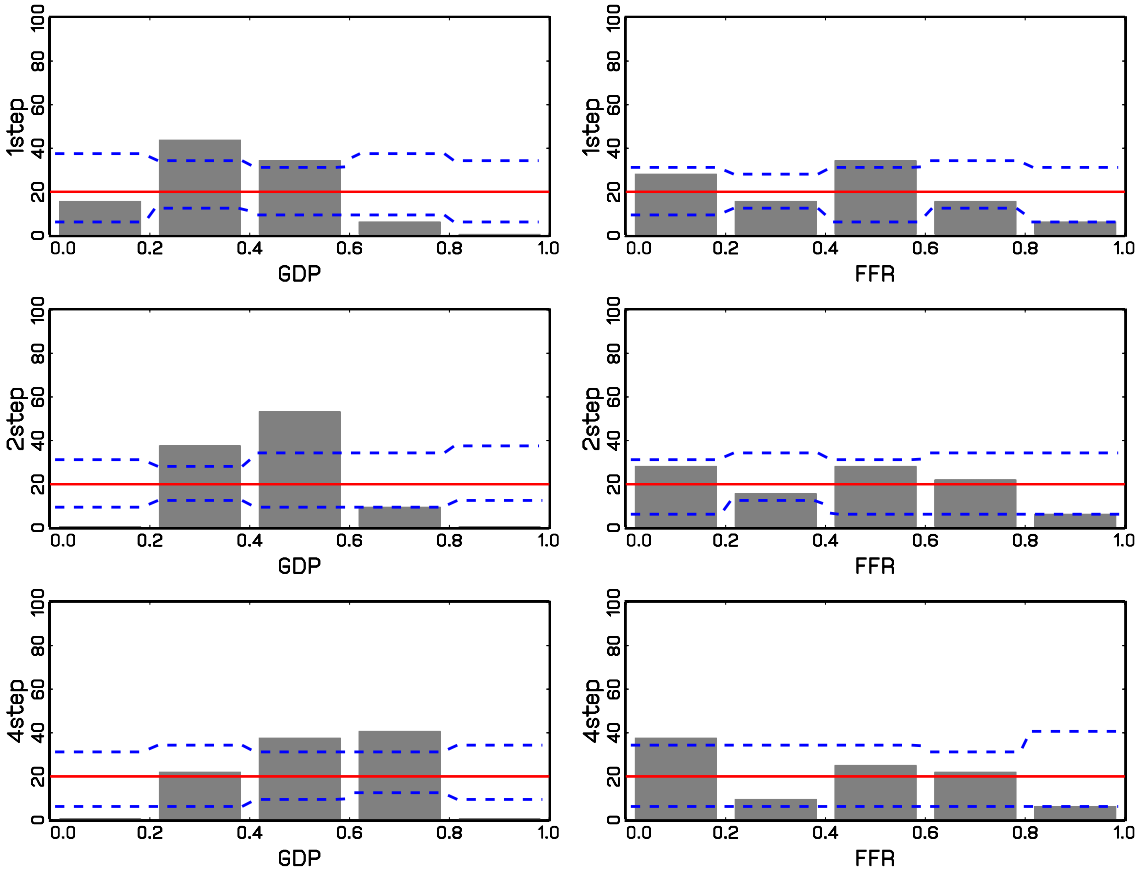
Notes: Probability integral transforms for forecasts of output growth (*GDP*), inflation (*PI*), and interest rates (*R*). *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 14: PITs of Forecasts for Inflation and Interest Rates Given Future Output (Smets Wouters, Real Time)



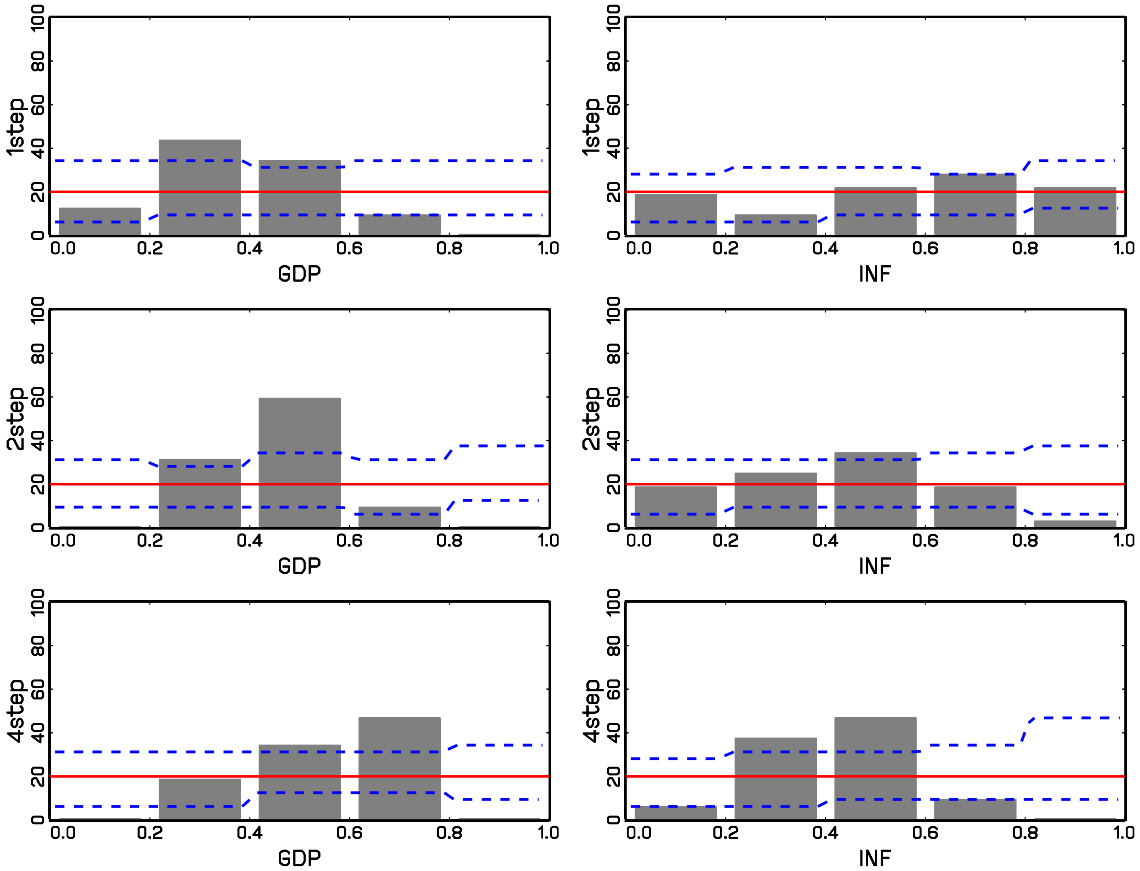
Notes: Probability integral transforms for forecasts of inflation (PI) and interest rates (R) conditional on actual future output growth. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 15: PITs of Forecasts for Output and Interest Rates Given Future Inflation (Smets Wouters, Real Time)



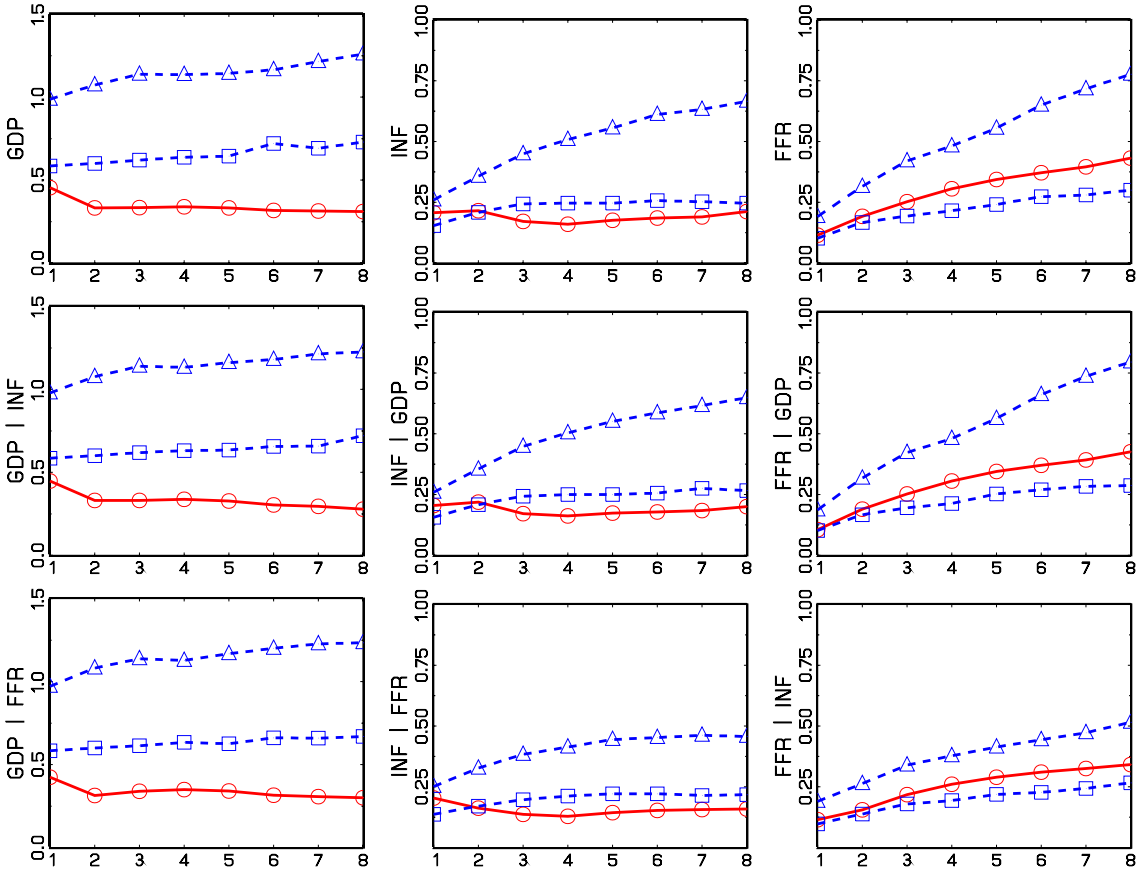
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and interest rates (*R*) conditional on actual future inflation. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 16: PITs of Forecasts for Output and Inflation Given Future Interest Rates (Smets Wouters, Real Time)



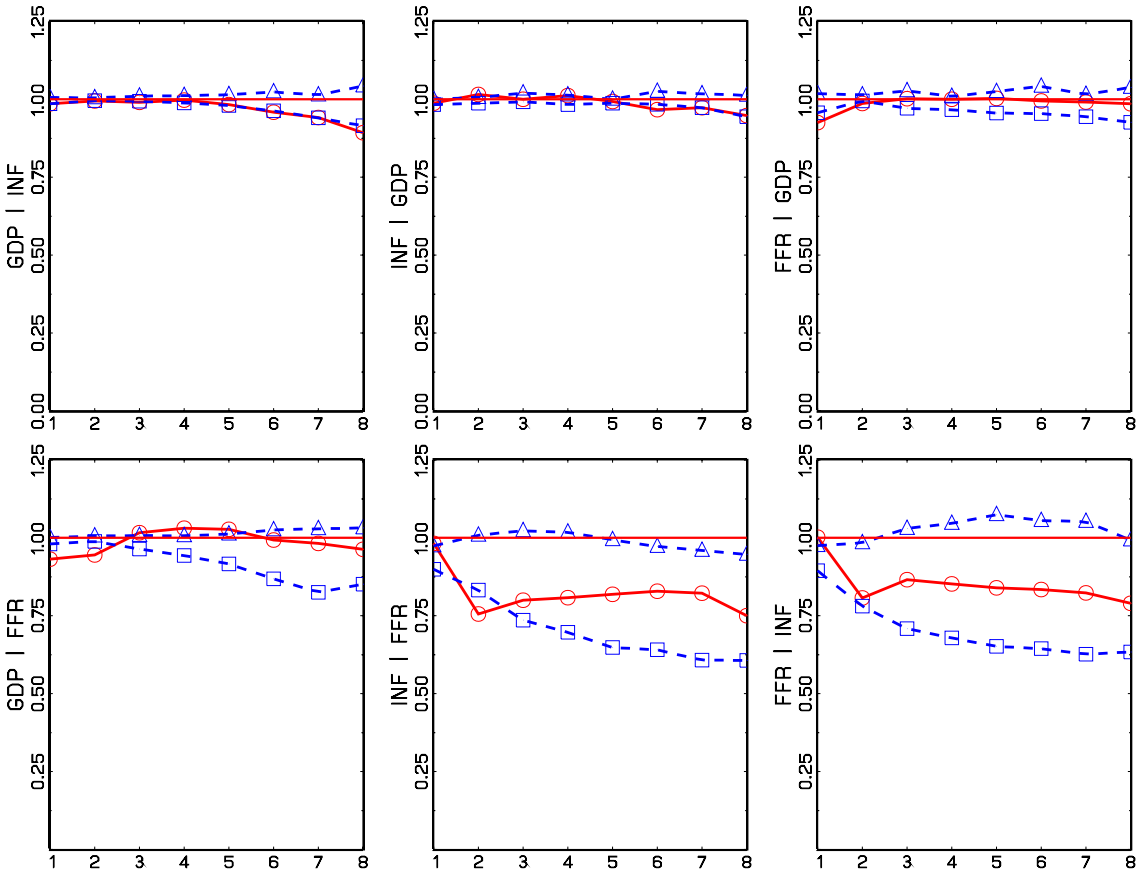
Notes: Probability integral transforms for forecasts of output growth (*GDP*) and inflation (*PI*) conditional on actual future interest rates. *Bars* correspond to actuals and dashed bands indicate 90% credible intervals obtained from the predictive distribution.

Figure 17: RMSEs of Unconditional and Conditional Forecasts (Smets Wouters, Real Time)



Notes: Root Mean Squared Errors (RMSEs) for forecasts of output growth (*GDP*, not annualized), inflation (*PI*, annualized), and interest rates (*R*, annualized) unconditional and conditional on future realizations of other variables. The solid line corresponds to RMSEs associated with actual forecasts and the dashed line signifies 90% credible intervals obtained from the predictive distribution.

Figure 18: RMSE Ratios of Unconditional and Conditional Forecasts (Smets Wouters, Real Time)



Notes: RMSE ratios for forecasts of output growth (GDP), inflation (PI), and interest rates (R): 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.

A A Short Review of Density Forecast Evaluation

The following properties of PITs are derived in Diebold, Gunther, and Tay (1998) and Diebold, Hahn, and Tay (1999).

Univariate density forecasts. Let $p(y_{t+h}|Y_{1:t})$ be a predictive distribution. Consider the probability integral transform of y_{t+h} with respect to $p(y_{t+h}|Y_{1:t})$,

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

Since $P_{t,h}$ is a cdf, it is nondecreasing over the y_{t+h} and we can apply a change of variables to find the density of $f(z_{t,h})$:

$$f(z_{t,h}) = p(P_{t,h}^{-1}(z_{t,h})|Y_{1:t}) \left| \frac{\partial P_{t,h}^{-1}(z_{t,h})}{\partial z_{t,h}} \right| = \frac{p(P_{t,h}^{-1}(z_{t,h})|Y_{1:t})}{p(P_{t,h}^{-1}(z_{t,h})|Y_{1:t})} = I\{0 \leq z_{t,h} \leq 1\}, \tag{A.2}$$

which means that the probability integral transforms are $z_{t,h} \sim \mathcal{U}[0,1]$.

Suppose y_{t+h} has probability density $p_{t,h}(y_{t,h}|Y_{1:t})$, $t = 1, \ldots, m$. For $h = 1$ the joint distribution of $\{y_{t+h}\}_{t=1}^m$ can be written as

$$p(y_{1+h}, \ldots, y_{m+h}|Y_{1:1}) = \prod_{t=1}^{m} p(y_{t+h}|Y_{1:t}). \tag{A.3}$$

One can write the density of the probability integral transforms using the same change of variables as before. Since the Jacobian of the transformation

$$\begin{vmatrix} \frac{\partial y_{1+1}}{\partial z_{1,1}} & \cdots & \frac{\partial y_{1+1}}{\partial z_{m,1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{m+1}}{\partial z_{1+1}} & \cdots & \frac{\partial y_{m+1}}{\partial z_{m,1}} \end{vmatrix}, \quad y_{t+1} = P_{t,1}^{-1}(z_{t,1})$$

is lower triangular, we can write

$$\begin{aligned} f(z_{1,1}, \ldots, z_{m,1}) &= \left(\prod_{t=1}^{m} \left| \frac{\partial P_{t,1}^{-1}(z_{t,1})}{\partial z_{t,1}} \right| \right) \left(\prod_{t=1}^{m} p(P_{t,1}^{-1}(z_{t,1})|Y_{1:t}) \right) \\ &= \prod_{t=1}^{m} \{0 \leq z_{t,1} \leq 1\} \end{aligned} \tag{A.4}$$

For $h > 1$ the probability integral transforms still have a $\mathcal{U}(0,1)$ distribution, but they are now dependent. It is straightforward to verify that the dependence dies out after h periods, that is, $z_{t,h}$ and $z_{t-s,h}$ are independent for $s \geq h$.

Multivariate density forecasts. In order to extend the above analysis to multivariate forecasts, let $y_t = [y_{1,t}, \ldots, y_{n,t}]'$ and define the conditional univariate probability integral transforms:

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

The same arguments that were used to examine the properties of univariate PITs can now be used to derive the properties of multivariate PITs.

Testing. The uniformity of the probability integrals transforms can be tested with a Kolmogorov-Smirnov tests, which measures the distance between the empirical distribution function of the PITs and the cdf of a $\mathcal{U}[0,1]$. Alternatively, Kling and Bessler (1989) used the following test. If one has a sequence of n PITs, then one expects (under the null hypothesis that the forecasts are well calibrated) any subinterval of the unit interval of length L to have Ln observed fractiles. If we have J nonoverlapping subintervals that exhaust the unit interval, then we can calculate the goodness of fit statistic

$$X^2 = \sum_{j=1}^{J} (a_j - L_j n)^2/(L_j n) \sim \chi^2(j-1),$$

where a_j is the actual number of observed fractiles in the interval j, and L_j is its length.

B Computing PITs

We need to approximate the integral

$$\int_{-\infty}^{y_{t+1}} p(\tilde{y}_{t+1}|Y_{1:t}) d\tilde{y}_{t+1}. \tag{A.5}$$

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

$$z_{t,h} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \mathcal{I}\{\tilde{y}_{t+h}^{(i)} < y_{t+h}\}. \tag{A.6}$$

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:

$$z_{1,t,h} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \kappa \left(\frac{\tilde{y}_{2,t+h}^{(i)} - y_{2,t+h}}{b} \right) \mathcal{I}\{\tilde{y}_{1,t+h}^{(i)} < y_{1,t+h}\} \tag{A.7}$$

where b is the bandwidth and κ is 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.

C The Smets-Wouters Model

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

$$\hat{y}_t = c_y\hat{c}_t + i_y\hat{i}_t + r^{kss}k_y\hat{z}_t + \varepsilon_t^g \tag{A.8}$$

$$\hat{c}_t = \frac{h/\gamma}{1+h/\gamma}\hat{c}_{t-1} + \frac{1}{1+h/\gamma}E_t\hat{c}_{t+1} + \frac{wl^{ss}(\sigma_c-1)}{c^{ss}\sigma_c(1+h/\gamma)}(\hat{l}_t - E_t\hat{l}_{t+1}) \tag{A.9}$$
$$-\frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}(\hat{r}_t - E_t\hat{\pi}_{t+1}) - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon_t^b$$

$$\hat{i}_t = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}_{t+1} + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}_t + \varepsilon_t^i \tag{A.10}$$

$$\hat{q}_t = \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}_{t+1} - \hat{r}_t + E_t\hat{\pi}_{t+1} + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t\hat{r}_{t+1}^k - \varepsilon_t^b \tag{A.11}$$

$$\hat{y}_t = \Phi(\alpha\hat{k}_t^s + (1-\alpha)\hat{l}_t + \varepsilon_t^a) \tag{A.12}$$

$$\hat{k}_t^s = \hat{k}_{t-1} + \hat{z}_t \tag{A.13}$$

$$\hat{z}_t = \frac{1-\psi}{\psi}\hat{r}_t^k \tag{A.14}$$

$$\hat{k}_t = \frac{(1-\delta)}{\gamma}\hat{k}_{t-1} + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^i \tag{A.15}$$

$$\hat{\mu}_t^p = \alpha(\hat{k}_t^s - \hat{l}_t) - \hat{w}_t + \varepsilon_t^a \tag{A.16}$$

$$\hat{\pi}_t = \frac{\beta\gamma^{(1-\sigma_c)}}{1+\iota_p\beta\gamma^{(1-\sigma_c)}}E_t\hat{\pi}_{t+1} + \frac{\iota_p}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_{t-1} \tag{A.17}$$
$$-\frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_p)(1-\xi_p)}{(1+\iota_p\beta\gamma^{(1-\sigma_c)})(1+(\varphi-1)\varepsilon_p)\xi_p}\hat{\mu}_t^p + \varepsilon_t^p$$

$$\hat{r}_t^k = \hat{l}_t + \hat{w}_t - \hat{k}_t \tag{A.18}$$

$$\hat{\mu}_t^w = \hat{w}_t - \sigma_l\hat{l}_t - \frac{1}{1-h/\gamma}(\hat{c}_t - h/\gamma\hat{c}_{t-1}) \tag{A.19}$$

$$\hat{w}_t = \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}(E_t\hat{w}_{t+1} + E_t\hat{\pi}_{t+1}) + \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}(\hat{w}_{t-1} - \iota_w\hat{\pi}_{t-1}) \tag{A.20}$$
$$-\frac{1+\beta\gamma^{(1-\sigma_c)}\iota_w}{1+\beta\gamma^{(1-\sigma_c)}}\hat{\pi}_t + \frac{(1-\beta\gamma^{(1-\sigma_c)}\xi_w)(1-\xi_w)}{(1+\beta\gamma^{(1-\sigma_c)})(1+(\lambda_w-1)\epsilon_w)\xi_w}\hat{\mu}_t^w + \varepsilon_t^w \tag{A.21}$$

$$\hat{r}_t = \rho\hat{r}_{t-1} + (1-\rho)(r_\pi\hat{\pi}_t + r_y(\hat{y}_t - \hat{y}_t^*)) + r_{\Delta y}((\hat{y}_t - \hat{y}_t^*) - (\hat{y}_{t-1} - \hat{y}_{t-1}^*)) + \varepsilon_t^r \tag{A.22}$$

$$\varepsilon_t^a = \rho_a\varepsilon_{t-1}^a + \eta_t^a \tag{A.23}$$

$$\varepsilon_t^b = \rho_b\varepsilon_{t-1}^b + \eta_t^b \tag{A.24}$$

$$\varepsilon_t^g = \rho_g\varepsilon_{t-1}^a + \rho_{ga}\eta_t^a + \eta_t^g \tag{A.25}$$

$$\varepsilon_t^i = \rho_i\varepsilon_{t-1}^i + \eta_t^i \tag{A.26}$$

$$\varepsilon_t^r = \rho_r\varepsilon_{t-1}^r + \eta_t^r \tag{A.27}$$

$$\varepsilon_t^p = \rho_r\varepsilon_{t-1}^p + \eta_t^p - \mu_p\eta_{t-1}^p \tag{A.28}$$

$$\varepsilon_t^w = \rho_w\varepsilon_{t-1}^w + \eta_t^w - \mu_w\eta_{t-1}^w \tag{A.29}$$

$$\tag{A.30}$$

$$\hat{y}_t^* = c_y \hat{c}_t^* + i_y \hat{i}_t^* + r^{kss} k_y \hat{z}_t^* + \varepsilon_t^g \tag{A.31}$$

$$\hat{c}_t^* = \frac{h/\gamma}{1+h/\gamma}\hat{c}_{t-1}^* + \frac{1}{1+h/\gamma}E_t\hat{c}_{t+1}^* + \frac{wl^{ss}(\sigma_c - 1)}{c^{ss}\sigma_c(1+h/\gamma)}(\hat{l}_t^* - E_t\hat{l}_{t+1}^*) - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}r_t^* - \frac{1-h/\gamma}{(1+h/\gamma)\sigma_c}\varepsilon_t^b \tag{A.32}$$

$$\hat{i}_t^* = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\hat{i}_{t-1}^* + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}E_t\hat{i}_{t+1}^* + \frac{1}{\phi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})}\hat{q}_t^* + \varepsilon_t^i \tag{A.33}$$

$$\hat{q}_t^* = \beta(1-\delta)\gamma^{-\sigma_c}E_t\hat{q}_{t+1}^* - r_t^* + (1-\beta(1-\delta)\gamma^{-\sigma_c})E_t r_{t+1}^{k*} - \varepsilon_t^b \tag{A.34}$$

$$\hat{y}_t^* = \Phi(\alpha k_t^{s*} + (1-\alpha)\hat{l}_t^* + \varepsilon_t^a) \tag{A.35}$$

$$\hat{k}_t^{s*} = k_{t-1}^* + z_t^* \tag{A.36}$$

$$\hat{z}_t^* = \frac{1-\psi}{\psi}\hat{r}_t^{k*} \tag{A.37}$$

$$\hat{k}_t = \frac{(1-\delta)}{\gamma}\hat{k}_{t-1}^* + (1-(1-\delta)/\gamma)\hat{i}_t + (1-(1-\delta)/\gamma)\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^i \tag{A.38}$$

$$\hat{\mu}_t^{p*} = \alpha(\hat{k}_t^{s*} - \hat{l}_t^*) - \hat{w}_t^* + \varepsilon_t^a \tag{A.39}$$

$$\hat{\mu}_t^{p*} = 1 \tag{A.40}$$

$$\hat{r}_t^{k*} = \hat{l}_t^* + \hat{w}_t^* - \hat{k}_t^* \tag{A.41}$$

$$\hat{\mu}_t^{w*} = -\sigma_l\hat{l}_t^* - \frac{1}{1-h/\gamma}(\hat{c}_t^* + h/\gamma\hat{c}_{t-1}^*) \tag{A.42}$$

$$\hat{w}_t^* = \mu_t^{w*} \tag{A.43}$$

With,

$$\gamma = \bar{\gamma}/100 + 1 \tag{A.44}$$

$$\pi^* = \bar{\pi}/100 + 1 \tag{A.45}$$

$$\bar{r} = 100(\beta^{-1}\gamma^{\sigma_c}\pi^* - 1) \tag{A.46}$$

$$r_{ss}^k = \gamma^{\sigma_c}/\beta - (1-\delta) \tag{A.47}$$

$$w_{ss} = \left(\frac{\alpha^\alpha(1-\alpha)^{(1-\alpha)}}{(\phi r_{ss}^k)^\alpha}\right)^{\frac{1}{1-\alpha}} \tag{A.48}$$

$$i_k = (1-(1-\delta)/\gamma)\gamma \tag{A.49}$$

$$l_k = \frac{1-\alpha}{\alpha}\frac{r_{ss}^k}{w_{ss}} \tag{A.50}$$

$$k_y = \phi l_k^{(\alpha-1)} \tag{A.51}$$

$$i_y = (\gamma - 1 + \delta)k_y \tag{A.52}$$

$$c_y = 1 - g_y - i_y \tag{A.53}$$

$$z_y = r_{ss}^k k_y \tag{A.54}$$

$$\tag{A.55}$$