Berg, Tim Oliver

**Working Paper**

# Forecast Accuracy of a BVAR under Alternative Specifications of the Zero Lower Bound

ifo Working Paper, No. 203

**Provided in Cooperation with:**
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

ifo Working Papers

Forecast Accuracy of a BVAR under
Alternative Specifications of the Zero Lower Bound

Tim Oliver Berg

August 2015

# Forecast Accuracy of a BVAR under Alternative Specifications of the Zero Lower Bound

## Abstract

This paper discusses how the forecast accuracy of a Bayesian vector autoregression (BVAR) is affected by introducing the zero lower bound on the federal funds rate. As a benchmark I adopt a common BVAR specification, including 18 variables, estimated shrinkage, and no nonlinearity. Then I entertain alternative specifications of the zero lower bound: replace the federal funds rate by its shadow rate, consider a logarithmic transformation, feed in monetary policy shocks, or utilize a rejection sampler. The latter two are also coupled with interest rate expectations from future contracts. The comparison is based on the accuracy of point and density forecasts of major U.S. macroeconomic series during the period 2009:1 to 2014:4. The results show that the performance of the specifications is greatly different, suggesting that this modeling choice is not innocuous. The introduction of the zero lower bound is not beneficial *per se*, but it depends on how it is done and which series is forecasted. With caution, I recommend the shadow rate specification and the rejection sampler combined with interest rate expectations to deal with the nonlinearity in the policy rate. Since the policy rate will remain low for some time, these findings could prove useful for practical forecasters.

JEL Code: C11, C52, C53,C55, E37.
Keywords: Forecasting, BVAR, zeroo lower bound, nonlinearity.

Tim Oliver Berg
Ifo Institute – Leibniz Institute for
Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
Phone: +49(0)89/9224-1689
berg@ifo.de

# 1   Introduction

Bayesian vector autoregressions (BVARs) are useful tools to forecast major macroeconomic time series. In contrast to classical VARs, they do not entail the risk of overparameterization and are hence able to exhaust the information contained in large datasets. A popular approach is that of Bańbura, Giannone, and Reichlin (2010), who use a variant of the Minnesota prior as proposed by Doan, Litterman, and Sims (1984) and Litterman (1980, 1986). Their prior shrinks forecasts towards a random walk for those variables showing substantial persistence and white noise for those that do not. The degree of shrinkage is controlled by a single hyperparameter, which can be easily computed or estimated in an optimal fashion.

In a forecast experiment the researcher has to make lots of choices with respect to model specification, estimation, and forecast construction. Because the specification choices affect the forecast accuracy of a BVAR, the literature has discussed issues such as: the appropriate number of variables included, whether variables should enter in (log-)levels or growth rates, whether a (log-)level specification should be complemented with priors favoring unit roots and cointegration, the treatment of the residual covariance matrix, the lag length, the selection of the shrinkage parameter, whether a recursive or rolling window estimation scheme should be used, and whether forecasts should be computed by iteration or the direct-step approach (see, e.g., Koop, 2013; Berg and Henzel, 2015; Carriero, Clark, and Marcellino, 2015, among others).

This paper discusses a further specification choice, which has not been addressed carefully in the literature. In particular, I investigate how the forecast accuracy of a BVAR is affected by introducing the zero lower bound on the federal funds rate. Researchers have dealt with the fact that the monetary policy instrument is constraint from below by 0% by either ignoring it or by adjusting their evaluation period such that the recent era of exceptionally low interest rates is avoided. For instance Clark and McCracken (2014) write: "We stop our forecast evaluation in 2007:Q4 to avoid possible complications of the zero lower bound constraints that became relevant in subsequent years." (see p. 27, ll. 25-26).

What Clark and McCracken (2014) mean by the relevance of the zero lower bound constraint becomes obvious when inspecting Figure 1. The figure shows the federal funds rate and forecasts thereof one quarter and four quarters ahead. While the federal funds rate is stuck between 0% and 0.25% since 2009, the BVAR attaches a lot probability mass to negative policy rates. And the BVAR also suggests a federal funds rate liftoff as early as mid-2011, despite the Federal Reserve's repeated promise to keep interest rates low for an "extended period." Such statements were part of the central bank's policy to anchor longer-term interest rate expectations during the zero lower bound period, known as forward guidance (see Bernanke, 2013).
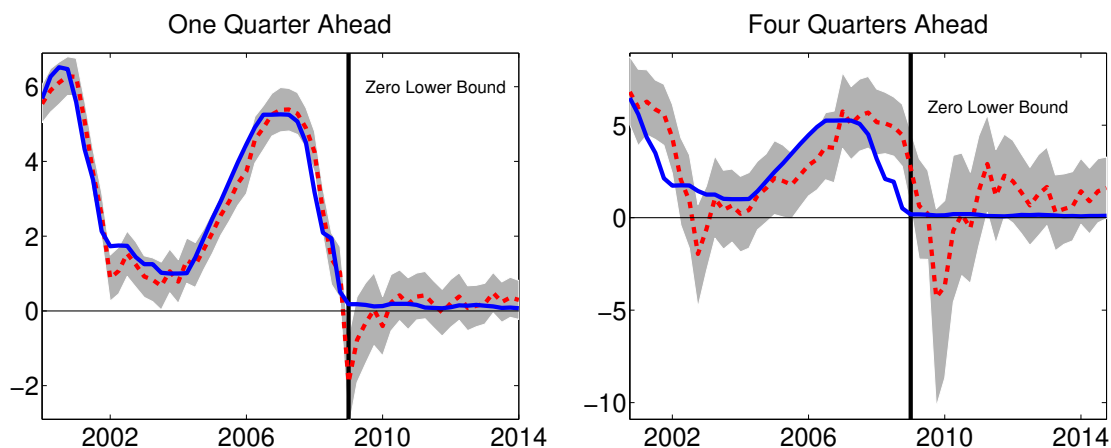
Figure 1: Federal Funds Rate – Forecast and Realization. Notes: this figure shows the point forecast (dashed line) together with the realization (solid line) and a 90% error band (shaded area) for the periods 2000:1 to 2014:1 (left panel) and 2000:4 to 2014:4 (right panel). x-axis: year; y-axis: percent.

The zero lower bound constraint and the Federal Reserve's forward guidance are challenges for any forecast model. However, ignoring them or discarding years of observations are no good options. And since the federal funds rate will most likely remain low for some time even after a liftoff, the results in this paper could prove useful for practical forecasters and researchers alike.

Hence I focus on the period from 2009 to 2014 and entertain alternative specifications of the zero lower bound. Also, I include interest rate expectations from future contracts to account for the Federal Reserve's forward guidance. Accordingly, most BVAR specifications considered in this paper are not linear and forecasts have nonstandard distributions.[1]

As a benchmark I adopt a common BVAR specification, including 18 variables that enter in quarterly (growth) rates. The prior is modeled as in Bańbura et al. (2010). Moreover, I use a long lag specification, while parsimony is achieved by optimal estimation of the shrinkage parameter as proposed by Giannone, Lenza, and Primiceri (2015). Forecasts are obtained by iteration using a recursive estimation scheme. My reading of the literature is that such a specification produces relatively accurate point and density forecasts for major macroeconomic series in normal times.

Then I entertain alternative specifications of the zero lower bound. First, I replace the federal funds rate by the shadow rate of Wu and Xia (2015). Second, I substitute the federal funds rate by its natural logarithm. Third, I feed shocks to monetary policy into the BVAR to prevent the policy rate from falling below zero. Fourth, I utilize the rejection sampler of Waggoner and Zha (1999) to truncate the predictive density of the federal funds rate at 0%. Finally, I couple the latter

[1]For linear models with Gaussian shocks, predictive densities are approximately multivariate normal (see Gelman, Hwang, and Vehtari, 2014).

two specifications with interest rate expectations from future contracts to rule out federal funds rate predictions that appear too high. As a first result I obtain that the alternative specifications produce very different predictive densities for the federal funds rate.

In the next step I discuss how this finding impacts on the forecast accuracy of the BVAR with respect to other key variables. Particularly, I rank the alternative specifications according to their point and density forecast accuracy for GDP growth, the unemployment rate, CPI inflation, and the treasury rate during the period from 2009:1 to 2014:1 up to four quarters ahead. For a robust assessment I consider several relative and absolute forecast metrics.

I obtain a large set of empirical results that can be summarized as follows. First, introducing the zero lower bound hardly affects the forecast accuracy of the BVAR for GDP growth, CPI inflation, and the unemployment rate one quarter ahead but does four quarters ahead. While this is by construction for the policy shock specifications, it is also observed for the other specifications. A practical forecaster should thus be more concerned about the zero lower bound when she is interested in long-term forecasts. Except for the treasury rate, for which this specification choice also matters in the short-term. Second, the performance of the alternative specifications is greatly different four quarters ahead, suggesting that this modeling choice is not innocuous. In particular, the introduction of the zero lower bound is not beneficial *per se*, but it depends on how it is done and which series is forecasted.

With caution, I recommend the shadow rate specification. Except for GDP growth, this specification shows a satisfactory forecast ability in absolute terms and relative to the benchmark. In contrast, both policy shock specifications are no good choices, while the logarithmic specifications is accurate for the treasury rate, but inflates forecast errors for any other series. Finally, the soft condition specification is useful only in combination with the financial market expectations, but then helps to achieve a better calibration of predictive densities.

The paper is organized as follows. Section 2 develops the BVAR and discusses major specification choices. Furthermore, I explain how the model is estimated and forecasts are obtained. Section 3 presents alternative specifications that introduce the zero lower bound into the BVAR. I also discuss how the predictive density of the federal funds rate is affected. Section 4 presents the data and explains the forecast experiment. Then I compare the alternative specifications in terms of forecast accuracy measures. Section 5 provides a summary of the results and concludes.

# 2 Forecasting with a BVAR

In this section I develop the BVAR and discuss major specification choices. Moreover, I explain how the model is estimated and forecasts are obtained.

## 2.1 BVAR

Consider the following VAR

$$y_t = c + B_1 y_{t-1} + ... + B_p y_{t-p} + u_t, \tag{1}$$

where $y_t$ is a $n \times 1$ vector of endogenous variables; $c$ is a $n \times 1$ vector of intercepts; $B_i$ are $n \times n$ matrices of coefficients; $i = 1, ..., p$ denotes the lags included; $u_t$ is a $n \times 1$ vector of normally distributed residual terms with zero mean and covariance matrix $\Sigma$; and data are available for $t = 1 - p, ..., T$. Let us denote $y = (y_1, ..., y_T)'$, $x_t = (y_{t-1}', ..., y_{t-p}', 1)'$, $x = (x_1, ..., x_T)'$, $B = (B_1, ..., B_p, c)'$, and $u = (u_1, ..., u_T)'$. The VAR in (1) can thus be written as $y = xB + u$. Moreover, let $\beta = \text{vec}(B)$ with $\text{vec}(\cdot)$ being the column stacking operator and $k = n(1 + np)$. Then $\beta$ is a $k \times 1$ vector containing all coefficients of the model.

In the forecast experiment the VAR is estimated on $n = 18$ variables including $p = 4$ lags of each (hence $k = 1,314$).[2] Such a large dimensional system of multivariate regressions is, however, not estimable without imposing additional prior beliefs on the parameters. Hence I follow common practice and use a variant of the Minnesota prior to deal with the dense parameterization of the model. The basic idea is that a white noise process is a reasonable description of the data generating process behind most macroeconomic series once transformed to stationarity. The VAR is hence centered around the prior mean $y_t = c + u_t$ and imposing the white noise prior amounts to shrinking all elements of $B_i$ towards zero.

In contrast to the original Minnesota prior (see Litterman, 1980, 1986), I do not assume $\Sigma$ to be known and diagonal, but use the generalized version of Kadiyala and Karlsson (1993, 1997) that allows for correlation among residuals. The evidence in Bańbura et al. (2010) suggests that a generalized Minnesota prior produces accurate forecasts for major macroeconomic series even though the $n(n+1)/2$ distinct elements of $\Sigma$ have to be estimated on top of the $k$ coefficients.

In particular, I consider a conjugate Normal-Inverse-Wishart prior of the following form:

$$\Sigma \sim \text{IW}(\Psi, d) \quad \text{and} \quad \beta | \Sigma \sim \text{N}(b, \Sigma \otimes \Omega), \tag{2}$$

---

[2]The choice for the lag length can be motivated by the frequency of the data, which is quarterly.

where $\otimes$ denotes the Kronecker product and the elements $\Psi$, $d$, $b$, and $\Omega$ are functions of hyperparameters. The conjugate prior implies a likelihood and posterior that come from the same family of distributions and hence makes Bayesian inference feasible even for large $n$.

Bańbura et al. (2010) implement the prior by constructing a set of artificial observations:

$$y^+ = \begin{bmatrix} 0_{np \times n} \\ \text{diag}\,(\sigma_1, ..., \sigma_n) \\ 0_{1 \times n} \end{bmatrix}, \qquad x^+ = \begin{bmatrix} \text{diag}\,(1, 2, .., p) \otimes \text{diag}\,(\sigma_1, ..., \sigma_n)\,/\lambda & 0_{np \times 1} \\ 0_{n \times np} & 0_{n \times 1} \\ 0_{1 \times np} & \epsilon \end{bmatrix}, \quad (3)$$

where $\text{diag}\,(\cdot)$ denotes a diagonal matrix. The $\sigma_j$'s account for the different scale and variability of the series and are set to the standard deviation of a residual from a univariate autoregression for the variable $y_{j,t}$ using the same lag length as in the VAR. The parameter $\epsilon$ is set to a small number ($10^{-4}$), reflecting a diffuse prior for the intercepts. Finally, the parameter $\lambda$ governs the degree of shrinkage and hence the tightness of the prior. As $\lambda \to \infty$ the prior becomes uninformative and posterior expectations coincide with the ordinary least squares (OLS) estimates. For $\lambda \to 0$ the posterior approaches the dogmatic prior. $\lambda$ is hence the key parameter in the BVAR and its selection is discussed in detail below.

The artificial observations are added on top of the actual data matrices, and the augmented regression model is used for inference:

$$y^* = x^* B + u^*, \tag{4}$$

where $y^* = (y', y^{+\prime})'$, $x^* = (x', x^{+\prime})'$, $u^* = (u', u^{+\prime})'$, and $u^{+\prime} = (u_1^+, ..., u_T^+)$. The latter contains the corresponding residual terms to $y^+$.

Let $\hat{\Sigma}$ and $\hat{\beta}$ be the covariance matrix and coefficients from an OLS regression of $y^*$ on $x^*$. The conditional posterior distributions of the covariance matrix and coefficients can be computed in closed form as a function of the shrinkage parameter:

$$\Sigma | \lambda, y \ \sim \ \text{IW}\left(\hat{\Sigma}, T + n + 2\right) \quad \text{and} \quad \beta | \Sigma, \lambda, y \sim \text{N}\left(\hat{\beta}, \Sigma \otimes \left(x^{*\prime} x^*\right)^{-1}\right). \tag{5}$$

In practice I get draws for the coefficients from the equivalent non-vectorized representation of their conditional posterior distribution:

$$B = \hat{B} + \text{chol}\left(\left(x^{*\prime} x^*\right)^{-1}\right) \times V \times \text{chol}\,(\Sigma)', \tag{6}$$

5

where chol $(\cdot)$ denotes the Cholesky decomposition and $V$ is a $k \times n$ standard Gaussian matrix. Drawing coefficients from (6) rather than (5) is much faster for large $n$ since the computation of the Kronecker product can be avoided (see also Carriero et al., 2015).

## 2.2 Choice for Shrinkage Parameter and MCMC Estimation

Following Giannone et al. (2015) I treat $\lambda$ as an additional unknown parameter that is estimated in an hierarchical fashion. The approach hence accounts for the uncertainty related to this specification choice and requires to add one more layer to the prior structure by placing a prior on the shrinkage parameter - a hyperprior $p(\lambda)$.

The marginal posterior for $\lambda$ after integrating out the posterior uncertainty about the model's parameters is

$$p(\lambda|y) \propto p(y|\lambda) \cdot p(\lambda), \tag{7}$$

where $\propto$ denotes proportionality. The marginal likelihood of the model (conditional on $\lambda$) has a well-known analytic form (see also Carriero et al., 2015):

$$p(y|\lambda) = \zeta^{-1} \times |\Psi + (y - xb)' (I + x\Omega x')^{-1} (y - xb)|^{-\frac{T+d}{2}}, \tag{8}$$

where

$$\zeta = \pi^{\frac{Tn}{2}} \times |(I + x\Omega x')^{-1}|^{-\frac{n}{2}} \times |\Psi|^{-\frac{d}{2}} \times \frac{\Gamma_n\left(\frac{d}{2}\right)}{\Gamma_n\left(\frac{T+d}{2}\right)}, \tag{9}$$

and $\Gamma_n(\cdot)$ denoting the n-variate gamma function. The prior moments are $b = \left(x^{+'}x^+\right)^{-1}x^{+'}y^+$, $\Omega = \left(x^{+'}x^+\right)^{-1}$, $\Psi = (y^+ - x^+b)'(y^+ - x^+b)$, and $d = n + 2$.

The hyperprior needs to be chosen and reflects how confident we are about the values for $\lambda$. I follow common practice and choose a diffuse hyperprior. In particular, I consider a Gamma density with mode equal to 0.2 and standard deviation of 0.4.

Since the joint posterior distribution for the parameters and $\lambda$ is not available in closed form, Giannone et al. (2015) recommend a Metropolis-Hastings algorithm to simulate the distribution. The sampler is a Markov chain Monte Carlo (MCMC) method and generates $\lambda$ from its marginal posterior with a Metropolis update. After convergence of the sampler, the covariance matrix and coefficients can be drawn from their posterior distribution conditional on $\lambda$ and forecasts computed in an iterative fashion. In detail, the algorithm works as follows.

Let $N_B$ be the number of burn-in-draws and $N_R$ the number of retained draws. The maximal forecast horizon is denoted by $H$. The steps are:

**Step 1:** Choose a starting point $\lambda^0$. I use the posterior mode, which is obtained by numerical optimization.[3] For $j = 1, ..., N_B + N_R$ run a loop over the following steps.

**Step 2:** Draw a proposal $\lambda^*$ from a jumping distribution $J\left(\lambda^*|\lambda^{j-1}\right) = N\left(\lambda^{j-1}, c \cdot \sigma_m^2\right)$, where $\sigma_m^2$ is the inverse of the Hessian computed at the posterior mode, and $c$ is a scaling constant chosen to obtain an acceptance ratio of about 20%.

**Step 3:** Compute the acceptance ratio:

$$r = \frac{p\left(\lambda^*|y\right)}{p\left(\lambda^{j-1}|y\right)}, \tag{10}$$

where $p\left(\lambda|y\right)$ is given by Equation (7).

**Step 4:** Randomly draw $\nu$ from $U\left(0, 1\right)$.

**Step 5:** Accept or discard the proposal $\lambda^*$ according to the following rule, and update, if necessary, the jumping distribution:

$$\lambda^j = \begin{cases} \lambda^* & : \quad \text{if} \quad \nu \leq r \\ \lambda^{j-1} & : \quad \text{otherwise} \end{cases} \tag{11}$$

If $j \leq N_B$ repeat the previous four steps, otherwise continue.

**Step 6:** Draw $\Sigma^j|\lambda^j, y$ and $\beta^j|\Sigma^j, \lambda^j, y$ from their conditional posterior in (5) and (6).

**Step 7:** Generate $u_{T+1}^j, ..., u_{T+H}^j$ from $u_t \sim N\left(0, \Sigma^j\right)$.

**Step 8:** Compute the 1-step-ahead forecast as

$$\hat{y}_{T+1}^j = c^j + B_1^j y_T + ... + B_p^j y_{T-p+1} + u_{T+1}^j, \tag{12}$$

while the h-step-ahead forecasts are obtained by iteration:

$$\hat{y}_{T+h}^j = c^j + B_1^j \hat{y}_{T+h-1}^j + ... + B_p^j \hat{y}_{T+h-p}^j + u_{T+h}^j, \tag{13}$$

where $h = 1, ..., H$ and $\hat{y}_{T+h}^j = y_{T+h-p}$ for $h \leq p$. $\quad\square$

This algorithm yields $\left(\hat{y}_{T+1}^j, ..., \hat{y}_{T+H}^j\right)_{j=N_B+1}^{N_B+N_R}$ as a sample of forecasts generated from the joint posterior distribution for the parameters and $\lambda$, where $N_B = 5,000$ and $N_R = 25,000$.

Figure A.1 suggests that the posterior distribution for $\lambda$ is well behaved and $N_B$ sufficiently large to let the Markov chain converge to its ergodic random distribution. The figure also shows

---

[3]I use the Matlab routine *fmincon*.

that the autocorrelation of the chain dies out quickly. Therefore I do not thin the chain to reduce dependence, but keep all draws for inference.

# 3 Alternative Specifications of the Zero Lower Bound

In this section I present alternative specifications that introduce the zero lower bound into the BVAR. Furthermore, I discuss how the predictive density of the federal funds rate is affected.

## 3.1 Specifications

In particular, I consider the following specifications:

**Benchmark** In the first specification I neglect the presence of the zero lower bound and construct forecasts using the algorithm shown in Section 2.2. This proceeding serves as the benchmark for all the other specifications. While the federal funds rate is constraint to be nonnegative, forecasts thereof may take negative values. And with the federal funds rate stuck between 0% and 0.25% since 2009, the policy instrument progressively loses importance for the U.S. economy over the evaluation period.

**Shadow Rate** In the second specification I replace the federal funds rate from 2009:1 onwards by the shadow short-term interest rate of Wu and Xia (2015)[4] and then construct forecasts as usual. Unlike the federal funds rate, the shadow rate can fall below zero in unconventional monetary policy environments (see also Figure A.2). Intuitively, it is the nominal interest rate with shortest maturity that would prevail in the absence of physical currency. Wu and Xia (2015) derive their shadow rate from a nonlinear term structure model estimated on forward rates and show that it interacts with the U.S. economy since 2009 in the same fashion as the federal funds rate did before. The shadow rate also reflects the unconventional monetary policy measures implemented by the Federal Reserve in recent times and is hence a reasonable proxy for the monetary policy stance during the zero lower bound period. Still, the hybrid federal funds/shadow rate series is hypothetical since economic agents cannot transact at shadow rates.

**Logarithmic** In the third specification I substitute the federal funds rate ($R_t$) by its natural logarithm. While the federal funds rate is bounded between 0 and $\infty$, the logarithmic transformation maps this into $(-\infty, \infty)$. The BVAR is hence estimated including $\log(R_t)$, and I obtain the predictive density for $R_t$ using the inverse log transformation, which makes sure that all posterior

---

[4]The Wu-Xia data are available here: https://www.frbatlanta.org/cqer/research/shadow_rate.aspx. I also experimented with the shadow rate estimates by Krippner (2013) with similar qualitative results.

draws for the federal funds rate are nonnegative. Cogley and Sargent (2002) follow a similar strategy to deal with the boundedness of the unemployment rate. Even though the logarithmic transformation truncates the predictive density of the federal funds rate at 0%, it also introduces a nonlinear relation between the monetary policy instrument and the economy, which may not be supported by the data.[5]

**Policy Shock** In the fourth specification I follow the standard procedure of the DSGE forecasting literature and feed shocks to monetary policy into the BVAR to prevent the federal funds rate from falling below zero (see, e.g., Del Negro and Schorfheide, 2013, among others). Under this specification the ordering of the variables in $y_t$ hence matters. In particular, I divide the sample into $n_1$ slow ($S_t$) and $n_2$ fast moving ($F_t$) variables and estimate the BVAR on $y_t = (S_t \ R_t \ F_t)'$. The former category includes production, employment, and price data, while the latter contains financial market variables. The monetary policy shock is then the innovation to $R_t$. To obtain a forecast sample, the algorithm above needs to be modified in the following way:

**Step 8 (*Policy Shock*):** Compute the 1-step-ahead forecast as

$$\hat{y}_{T+1}^j = c^j + B_1^j y_T + ... + B_p^j y_{T-p+1} + u_{T+1}^j. \tag{14}$$

If $\hat{R}_{T+1}^j \geq 0$ continue. Otherwise add a monetary policy shock ($\epsilon_{T+1}^j$) such that the federal funds rate is equal to zero, i.e. $\epsilon_{T+1}^j = \left(-\hat{R}_{T+1}^j / \sigma_{n_1+1}^j\right) S^j e$, where $\sigma_{n_1+1}^j$ is the standard deviation of the shock, $S^j = \text{chol}\left(\Sigma^j\right)$, and $e = (0_{n_1} \ 1 \ 0_{n_2})'$. Obtain the h-step-ahead forecasts by iteration and always add a monetary policy shock ($\epsilon_{T+h}^j$) if needed:

$$\hat{y}_{T+h}^j = c^j + B_1^j \hat{y}_{T+h-1}^j + ... + B_p^j \hat{y}_{T+h-p}^j + u_{T+h}^j, \tag{15}$$

where $h = 1, ..., H$ and $\hat{y}_{T+h}^j = y_{T+h-p}$ for $h \leq p$. $\square$

Hitting the zero lower bound in $T+h$ hence alters forecasts for slow moving variables in $T+h+1$ and later, while forecasts for fast moving variables are also affected in $T + h$. But none variable is affected in $T + h - 1$ and earlier since the monetary policy shocks are not anticipated.[6]

**Soft Condition** In the fifth specification I conform to Waggoner and Zha (1999), incorporating the zero lower bound using an acceptance-rejection algorithm. In particular, parameter constel-

---

[5]In the logarithmic BVAR a decline of the federal funds rate from 10% to 5% has the same effect on the economy as a decline from 1% to 0.5%, which is hardly consonant with standard economic theory.

[6]The DSGE forecasting literature also considers anticipated monetary policy shocks since reaching the zero lower bound may not come as a surprise to economic agents. However, this is not possible in a VAR.

lations that are not consistent with a nonnegative federal funds rate forecast receive zero prior weight and are discarded.[7] The truncated forecast distribution is sampled as follows.

**Step 8: (*Soft Condition*)** Compute the 1-step-ahead forecast as

$$\hat{y}_{T+1}^j = c^j + B_1^j y_T + ... + B_p^j y_{T-p+1} + u_{T+1}^j, \tag{16}$$

while the h-step-ahead forecasts are obtained by iteration:

$$\hat{y}_{T+h}^j = c^j + B_1^j \hat{y}_{T+h-1}^j + ... + B_p^j \hat{y}_{T+h-p}^j + u_{T+h}^j, \tag{17}$$

where $h = 1, ..., H$ and $\hat{y}_{T+h}^j = y_{T+h-p}$ for $h \leq p$. If $\hat{R}_{T+h}^j \geq 0$ for $h = 1, ..., H$ accept the draw, otherwise repeat Steps 6 and 7 and use the alternative draw $\tilde{\Sigma}^j, \tilde{\beta}^j, \tilde{u}^j$ to compute forecasts.  □

The Waggoner and Zha (1999) method can be computationally intense since the acceptance rate may be low in some periods. Moreover, this specification affects not only the predictive density, but also the posterior of the coefficients and the composition of the residual vector. For instance, shocks that imply a decrease in the policy rate, e.g. negative monetary policy shocks, receive a lower weight than those that lead to an increase, e.g. positive demand shocks.

**External Information (EI)** Finally, I combine the policy shock and soft condition specifications with external information, ruling out federal funds rate forecasts that appear too high. To obtain a reasonable upper bound, I use the expectations inherent in future contracts.[8] In particular, I restrict a federal funds rate forecast for the period $T + h$ to the interval $\hat{R}_{T+h}^j \in \left(0, 2 \times \hat{R}_{T+h}^f\right)$, where $\hat{R}_{T+h}^f$ denotes the financial market expectation. For example, if financial market participants expect the federal funds rate to be 0.5% next quarter, I ensure that forecasts from the BVAR are positive but not larger than 1%, which implies that the predictive mean roughly matches the market expectations. Conditioning forecasts on external information has become popular and proved to be beneficial (see, e.g., Clark and McCracken, 2014; Smets, Warne, and Wouters, 2014; Bańbura, Giannone, and Lenza, 2015, among others). In accordance with the algorithms above, I implement the upper bound either by adding negative monetary policy shocks such that the federal funds rate equals the upper bound (policy shock specification) or by discarding draws that violate the restriction (soft condition specification).

---

[7]This condition is soft in the sense that future values of the federal funds rate are restricted to a certain range, whereas a hard condition would fix forecasts at a single point, say at 1%.

[8]Federal funds rate future contracts are traded on the Chicago Board of Trade and historical prices are available from Datastream. For example, the mnemonic for the January 2012 contract is "CFF0112". See also Figure A.2 for a times series plot.

## 3.2 Discussion

Before the different specifications are ranked according to forecast accuracy, I discuss how they affect the predictive density of the federal funds rate, which is shown in Figure 2 for $h = 1$ and the period 2012:1 based on the 25,000 draws.[9] The figure suggests that the density is symmetric for the benchmark BVAR, while the dispersion is quite large. The benchmark also assigns a high probability to a negative federal funds rate. Replacing the federal funds rate by its shadow rate neither affects the shape nor the dispersion of the density, but shifts it completely into negative territory, which is consonant with the shadow rate concept. Under both specifications the BVAR is linear and densities approximately Gaussian. Simulating the BVAR with $\log(R_t)$ rather than $R_t$ introduces a nonlinearity. The density is approximately log-normal, right-skewed, and takes only nonnegative values. It is also more condensed than those of the other two specifications.

While the shadow rate and logarithmic specification build on replacing or transforming the policy rate before estimation, the policy shock and soft condition specifications include the federal funds rate, but manipulate its predictive density. For instance, the policy shock specification generates a density that is that of the benchmark, but truncated at zero. All negative draws are shifted by adding positive monetary policy shocks, producing a density with a lot probability mass at zero. When combined with external information, all draws above the upper bound are also shifted, and the density may be bimodal. The density under the soft condition specification is also truncated at zero, but less right-skewed. In combination with external information this specification leads to a density that appears uniform with mean close to the market expectations.

In sum, the alternative specifications produce very different predictive densities for the federal funds rate, and I provide a summary of their main properties in Table 1. Below I assess how this finding impacts on the forecast accuracy of the BVAR with respect to other key variables.

# 4 Forecast Comparison

In this section I first present the data and explain the forecast experiment. Thereafter I compare the alternative specifications in terms of point and density forecast accuracy measures.

## 4.1 Data and Forecast Experiment

The dataset includes 18 quarterly U.S. macroeconomic series for the period 1985:1 to 2014:4. The starting point coincides roughly with the beginning of the Great Moderation era and is also cho-

---

[9]I use this period as an example. The densities are qualitatively similar for other periods and horizons.
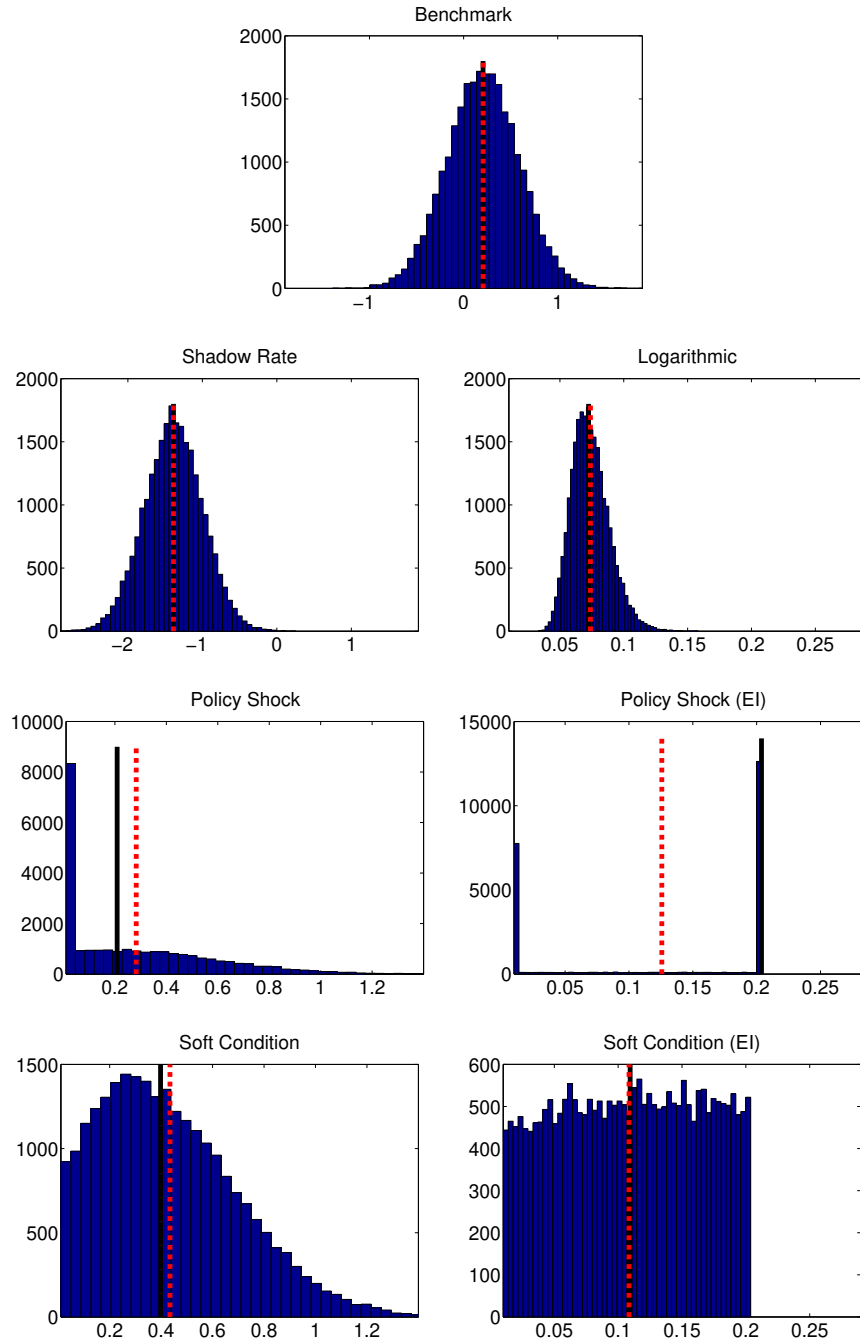
Figure 2: Predictive Density of Federal Funds Rate. Notes: this figure shows the predictive density of the federal funds rate (or its shadow rate) for $h = 1$ and the period 2012:1 under alternative specifications of the zero lower bound. The vertical lines denote the predictive mean (dashed) and median (solid), respectively. x-axis: percent; y-axis: number of posterior draws.

Table 1: Alternative Specifications of the Zero Lower Bound

| Specification | Policy Instrument | Predictive Density of $R_t$ (or $S_t$) | | | |
| --- | --- | --- | --- | --- | --- |
| | | Density | Support | Symmetry | Skewness |
| Benchmark | $R_t$ | Normal | $(-\infty, \infty)$ | Yes | None |
| Shadow Rate | $S_t$ | Normal | $(-\infty, \infty)$ | Yes | None |
| Logarithmic | $\log(R_t)$ | Log-Normal | $(0, \infty)$ | No | Right |
| Policy Shock | $R_t$ | n/a | $(0, \infty)$ | No | Right |
| Policy Shock (EI) | $R_t$ | n/a | $(0, b)$ | No | Right/Left |
| Soft Condition | $R_t$ | n/a | $(0, \infty)$ | No | Right |
| Soft Condition (EI) | $R_t$ | n/a | $(0, b)$ | No | Right/Left |

Notes: this table summarizes the main properties of the predictive density of the federal funds rate ($R_t$) or its shadow rate ($S_t$) under alternative specifications of the zero lower bound. n/a means that the density is nonstandard, while $b$ denotes an upper bound derived from external information (EI).

sen to avoid the monetarist experiment of the early 1980s. In most cases the series are obtained from the Federal Reserve Economic Database (FRED) and cover the following categories that are particularly relevant for all policymakers: national accounts, business activity, employment, prices, financial markets, and the federal funds rate as the monetary policy instrument. Except for those variables that are already expressed in rates (e.g. the unemployment rate) or assumed to be stationary in levels (e.g. the consumer sentiment), I transform all of them into quarterly growth rates, approximated by the first difference of their logarithm. This proceeding is consistent with the white noise prior imposed on the VAR coefficients. I provide a detailed description of the series and their exact transformation in Table 2. The table also shows the ordering of the variables, which is relevant for the policy shock specification.

For each specification I produce forecasts up to $H = 4$. Thanks to differencing I lose the first observation. Thus I begin with the estimation sample 1985:2 to 2008:4 and generate forecasts for 2009:1 to 2009:4. I then iterate forward, always updating the estimation sample one quarter at a time, until 2013:4, producing forecasts for 2014:1 to 2014:4.[10] This procedure yields a sequence of 21 forecast sets for each specification. The evaluation period hence runs from 2009:1 to 2014:1 for the shortest forecast horizon ($h = 1$) and from 2009:4 to 2014:4 for the longest ($h = 4$), and coincides with the recent zero lower bound period.

[10]Berg and Henzel (2015) show that BVARs benefit a lot in terms of forecast accuracy from recursive estimation schemes that use all available observations up to period $T$ compared to rolling window schemes that discard the most distant observations. Therefore I prefer the former to the latter.

Table 2: Dataset

| No. | Mnemonic | Label | Category | Source | S/R/F | Code |
|---|---|---|---|---|---|---|
| 01 | GDPC1 | Real Gross Domestic Product | National Accounts | FRED | S | 1 |
| 02 | PCECC96 | Real Personal Consumption Expenditures | National Accounts | FRED | S | 1 |
| 03 | GPDIC1 | Real Gross Private Domestic Investment | National Accounts | FRED | S | 1 |
| 04 | INDPRO | Industrial Production Index | Business Activity | FRED | S | 1 |
| 05 | TCU | Capacity Utilization: Total Industry | Business Activity | FRED | S | 3 |
| 06 | NAPM | ISM Manufacturing: PMI Composite Index | Business Activity | FRED | S | 3 |
| 07 | USUMCONSH | University of Michigan: Consumer Sentiment | Business Activity | TRDS | S | 3 |
| 08 | PAYEMS | All Employees: Total Nonfarm | Employment | FRED | S | 1 |
| 09 | UNRATE | Civilian Unemployment Rate | Employment | FRED | S | 3 |
| 10 | COMPRNFB | Real Compensation Per Hour: Nonfarm | Employment | FRED | S | 1 |
| 11 | PCECTPI | PCE: Chain-type Price Index | Prices | FRED | S | 1 |
| 12 | CPIAUCSL | CPI for All Urban Consumers: All Items | Prices | FRED | S | 1 |
| 13 | PPIACO | Producer Price Index: All Commodities | Prices | FRED | S | 1 |
| 14 | FEDFUNDS | Effective Federal Funds Rate | Policy Instrument | FRED | R | 3 |
| 15 | DGS10 | 10-Year Treasury Constant Maturity Rate | Financial Markets | FRED | F | 3 |
| 16 | BAA | Moody's Seasoned Baa Corporate Bond Yield | Financial Markets | FRED | F | 3 |
| 17 | S&PCOMP | S&P 500 Composite - Price Index | Financial Markets | TRDS | F | 2 |
| 18 | TWEXBPA | Real Trade Weighted U.S. Dollar Index: Broad | Financial Markets | FRED | F | 2 |

Notes: Mnemonic, label, and category as used in database. Source: FRED = Federal Reserve Economic Database; TRDS = Thomson Reuters Datastream. S = Slow moving; R = Policy instrument; F = Fast moving. Transformation code: $1 = 400 \cdot \log (x_t/x_{t-1})$; $2 = 100 \cdot \log (x_t/x_{t-1})$; $3 =$ None.

## 4.2 Forecast Metrics

The different specifications are compared in terms of several point and density forecast accuracy measures. Each metric represents an alternative loss function and computing all of them allows for a robust assessment of forecast performance. I summarize their main properties in Table 3.

For point forecasts I begin with the mean forecast error (MFE). Let $\bar{y}_{i,T|T-h}^{j}$ be the predictive mean for variable $i$ produced by approach $j$ and $y_{i,T}$ the realization. The MFE is computed as

Table 3: Forecast Metrics

| Metric | Short | Reference | Forecast | Evaluation | Measure |
|---|---|---|---|---|---|
| Mean Forecast Error | MFE | Point | Mean | Absolute | pp |
| Root Mean Squared Forecast Error | RMSFE | Point | Mean | Relative | pp |
| Mean Absolute Forecast Error | MAFE | Point | Median | Relative | pp |
| Log Predictive Density Score | LPDS | Density | PDF | Relative | n/a |
| Continuous Ranked Probability Score | CRPS | Density | CDF | Relative | pp |
| Probability Integral Transform | PIT | Density | CDF | Absolute | n/a |

Notes: this table summarizes the main properties of the forecast accuracy measures. The acronyms mean: PDF = Probability density function; CDF = Cumulative density function; pp = Percentage points.

$$\text{MFE}_{i,h}^{j} = \frac{1}{T_1 - T_0 + 1} \sum_{T=T_0+h}^{T_1+h} \left( \bar{y}_{i,T|T-h}^{j} - y_{i,T} \right), \tag{18}$$

where $T_0 = 95$ and $T_1 = 115$ denote the last observation of the first (2008:4) and final (2013:4) estimation sample, respectively. The MFE allows for an absolute assessment of the point forecast accuracy because it detects any systematic bias of the predictive mean. Whenever the MFE is positive (negative) the BVAR overpredicts (underpredicts) a particular variable.

For a forecaster with a quadratic loss function the predictive mean is also the optimal forecast and the root mean squared forecast error (RMSFE) is the appropriate measure to discriminate among forecast models (see Weiss, 1996):

$$\text{RMSFE}_{i,h}^{j} = \sqrt{\frac{1}{T_1 - T_0 + 1} \sum_{T=T_0+h}^{T_1+h} \left( \bar{y}_{i,T|T-h}^{j} - y_{i,T} \right)^2}. \tag{19}$$

The predictive mean is, however, not an informative measure of the location of a distribution if the forecast approach is nonlinear. For instance, the skewness of the predictive density of the federal funds rate may translate to the densities of other variables, and their predictive mean then differs from the mode and median. Hence I also compute the mean absolute forecast error (MAFE), for which the predictive median ($y_{i,T|T-h}^{m,j}$) is the optimal forecast:

$$\text{MAFE}_{i,h}^{j} = \frac{1}{T_1 - T_0 + 1} \sum_{T=T_0+h}^{T_1+h} |y_{i,T|T-h}^{m,j} - y_{i,T}|. \tag{20}$$

While the MFE is a useful measure to evaluate the absolute accuracy of a point forecast, the RMSFE and MAFE are informative only when compared across models. Thus I will report them for each specification $j$ relative to the benchmark $B$:

$$\left( \text{RMSFE}_{i,h}^{j} / \text{RMSFE}_{i,h}^{B} \right) \quad \text{and} \quad \left( \text{MAFE}_{i,h}^{j} / \text{MAFE}_{i,h}^{B} \right). \tag{21}$$

For density forecasts I start with the log predictive density score (LPDS), which is a standard metric to compare forecast models in a Bayesian setting (see, e.g., Geweke and Amisano, 2010, among others). The LPDS is obtained by computing the height of the predictive density at the realization, and I report its average across evaluation periods:

$$\overline{\text{LPDS}}_{i,h}^{j} = \frac{1}{T_1 - T_0 + 1} \sum_{T=T_0+h}^{T_1+h} \log p_{i,T|T-h}^{j}(y_{i,T}), \tag{22}$$

15

where $p^j_{i,T|T-h}(\cdot)$ is the predictive density, which is obtained by smoothing the forecast samples using a kernel density estimator.[11]

Some authors argue that the LPDS is sensitive to extreme observations and propose the continuous ranked probability score (CRPS) as a robust alternative (see, e.g., Ravazzolo and Vahey, 2014; Clark and Ravazzolo, 2015, among others). Hence I also compute the CRPS and document its average across evaluation periods:

$$\overline{\text{CRPS}}^j_{i,h} = \frac{1}{T_1 - T_0 + 1} \sum_{T=T_0+h}^{T_1+h} \text{CRPS}^j_{i,T|T-h}, \tag{23}$$

with

$$\text{CRPS}^j_{i,T|T-h} = \int_{-\infty}^{\infty} \left( P^j_{i,T|T-h}(y) - \mathbb{I}\,[y \geq y_{i,T}] \right)^2 dy, \tag{24}$$

where $P^j_{i,T|T-h}(\cdot)$ is the cumulative density function (CDF) to $p^j_{i,T|T-h}(\cdot)$. The CDF of the realization is given by the indicator function $\mathbb{I}\,[y \geq y_{i,T}]$ taking value 1 if $y \geq y_{i,T}$ and 0 otherwise.

In practice, I follow Gneiting and Raftery (2007) and calculate the CRPS using its alternative representation:

$$\text{CRPS}^j_{i,T|T-h} = \text{E}|Y - y_{i,T}| - \frac{1}{2}\text{E}|Y - Y'|, \tag{25}$$

where $Y$ and $Y'$ are independent draws from $p^j_{i,T|T-h}(\cdot)$. The expectation terms (E) are approximated using the forecast samples.[12]

Since the LPDS and CRPS are both silent about the absolute density forecast accuracy, I again report them for each specification relative to the benchmark:

$$\left( \overline{\text{LPDS}}^j_{i,h} - \overline{\text{LPDS}}^B_{i,h} \right) \quad \text{and} \quad \left( \overline{\text{CRPS}}^j_{i,h} / \overline{\text{CRPS}}^B_{i,h} \right). \tag{26}$$

Finally, I assess the absolute calibration of the predictive densities using the probability integral transform (PIT). For a given predictive density, the PIT is the corresponding CDF evaluated at the realization:

$$\text{PIT}^j_{i,T|T-h} = P^j_{i,T|T-h}(y_{i,T}). \tag{27}$$

---

[11]I use the Matlab routine *ksdensity* with default options. The frequently used Gaussian approximation of Adolfson, Lindé, and Villani (2007) is not applicable since densities may be asymmetric.

[12]Note that the CRPS can be viewed as the distributional analogue of the MAFE, since for $Y = Y'$, i.e. for point forecasts, the CRPS is the same as the MAFE (see also Arora, Little, and McSharry, 2013).

If the proposed predictive density is consistent with the (unobserved) true density, the sequence $\left\{\text{PIT}^{j}_{i,T|T-h}\right\}^{T_1+h}_{T=T_0+h}$ has an uniform distribution. I follow Diebold, Gunther, and Tay (1998) and provide a graphical assessment of the PITs. In particular, I divide the unit interval into $\hat{n} = 5$ equally sized bins and depict the number of PITs falling into any bin. If the predictive density would be calibrated correctly, each bin should contain $\hat{p} = 100\%/\hat{n} = 20\%$ of the PITs. To detect any significant departure from uniformity, I compute 90% confidence bands utilizing a normal approximation, i.e. $\hat{p} \pm 1.645\sqrt{\hat{p}(1-\hat{p})/(T_1-T_0+1)}$ (see also Rossi and Sehkposyan, 2014).

For all other forecast accuracy measures I consider the test of unconditional predictive ability of Giacomini and White (2006) to check for significance at the 1%, 5%, and 10% level. Asymptotic standard errors are calculated with the Newey and West (1987) method that accounts for heteroskedasticity and autocorrelation in forecast errors. While I could provide evidence for all variables, I restrict the forecast comparison to four major macroeconomic series: GDP growth, the unemployment rate, CPI inflation, and the treasury rate.[13]

## 4.3 Point Forecasts

I begin the forecast comparison by analyzing the MFE, which are reported for each specification and variable in Table 4. In each case, the MFE is in percentage points (pp). The table shows that the benchmark specification overpredicted GDP growth during the zero lower bound period, but not significantly. Replacing the federal funds rate by its shadow rate reinforces this bias, as I would expect additional stimulus to do. Simulating the BVAR with $\log(R_t)$ rather than $R_t$ also inflates the MFE. At $h = 4$ the logarithmic specification significantly overpredicts GDP growth by 1.6 pp. The small movements in the federal funds rate are large once the logarithm is taken, leading to substantial increases in growth forecasts that are not supported by the data. Furthermore, both policy shock specifications show MFE that are similar to those of the benchmark. Restricting the policy rate from below and/or above does not affect the location of the predictive mean in this case. In contrast, the soft condition specification produces point forecasts that are significantly above actual growth rates at all horizons. This bias could be the result of a large number of positive demand shocks added to the BVAR to prevent the policy rate from falling below 0%. Indeed, if the soft condition specification is combined with external information, and the size of the demand shocks limited, the MFE are close to zero.

Consistent with an overprediction of GDP growth, the specifications underpredict the unemployment rate. Yet, the MFE are small and in no case significantly different from zero. Similar

---

[13]In fact, I could not provide a complete forecast comparison for the federal fund rate since the shadow rate BVAR does not include the policy instrument.

## Table 4: Mean Forecast Errors

| Specification | Forecast Horizon $h$ (in Quarters) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| GDP Growth | | | | |
| Benchmark | 0.15 | 0.12 | 0.32 | 0.46 |
| Shadow Rate | 0.35 | 0.35 | 0.64 | 0.90* |
| Logarithmic | 0.87* | 0.81 | 1.13 | 1.60** |
| Policy Shock | 0.15 | 0.06 | 0.27 | 0.42 |
| Policy Shock (EI) | 0.15 | 0.07 | 0.30 | 0.44 |
| Soft Condition | 0.97** | 0.84* | 1.10** | 1.06** |
| Soft Condition (EI) | 0.10 | −0.14 | −0.08 | 0.10 |
| Unemployment Rate | | | | |
| Benchmark | −0.10 | −0.16 | −0.20 | −0.24 |
| Shadow Rate | −0.11 | −0.17 | −0.23 | −0.30 |
| Logarithmic | 0.02 | 0.04 | 0.02 | −0.05 |
| Policy Shock | −0.10 | −0.17 | −0.23 | −0.29 |
| Policy Shock (EI) | −0.10 | −0.17 | −0.22 | −0.28 |
| Soft Condition | −0.17 | −0.33 | −0.50 | −0.65 |
| Soft Condition (EI) | −0.11 | −0.16 | −0.18 | −0.18 |
| CPI Inflation | | | | |
| Benchmark | −0.68 | −0.43 | −0.10 | 0.14 |
| Shadow Rate | −0.83 | −0.69 | −0.39 | −0.11 |
| Logarithmic | −0.40 | −0.71 | −0.61 | −0.22 |
| Policy Shock | −0.69 | −0.42 | −0.01 | 0.40 |
| Policy Shock (EI) | −0.69 | −0.44 | −0.11 | 0.18 |
| Soft Condition | −0.18 | 0.10 | 0.62 | 1.16*** |
| Soft Condition (EI) | −0.42 | −0.26 | −0.07 | 0.19 |
| Treasury Rate | | | | |
| Benchmark | 0.25*** | 0.52*** | 0.84*** | 1.15*** |
| Shadow Rate | 0.22*** | 0.46*** | 0.75*** | 1.03*** |
| Logarithmic | 0.08 | 0.16 | 0.21 | 0.26 |
| Policy Shock | 0.32*** | 0.63*** | 0.98*** | 1.31*** |
| Policy Shock (EI) | 0.29*** | 0.55*** | 0.85*** | 1.13*** |
| Soft Condition | 0.54*** | 1.01*** | 1.41*** | 1.78*** |
| Soft Condition (EI) | 0.28** | 0.49** | 0.72** | 0.98** |

Notes: this table shows the mean forecast errors for the period 2009:$h$ to 2014:$h$. ***, **, and * indicate that the errors are significantly different from zero at the 1%, 5%, and 10% level, respectively.

results are obtained for CPI inflation. Except for the soft condition specification, which produces point forecasts that are well above actual rates at longer horizons. This finding confirms the idea that the soft condition specification puts a large weight on positive demand shocks to ensure a nonnegative policy rate. As for GDP growth, the bias can be removed when this specification is coupled with external information.

Finally, almost all specifications significantly overpredict the treasury rate at all horizons. For the benchmark the predictive mean is on average 1.15 pp higher than the realization at $h = 4$. Replacing the federal funds rate by its shadow rate or using the policy shock specifications does not help to achieve correct calibration. The soft condition specification even amplifies the bias, if not combined with external information. In fact, the logarithmic specification is the only way to produce point forecasts that are consonant with actual values, which could be due to a plausible predictive density for the federal funds rate influencing the treasury rate (see Figure 2 again).

I continue the forecast comparison by discussing the RMSFE, which are reported in Table 5 for the benchmark in levels (pp) and the other specifications as ratios relative to the benchmark. A ratio below unity means that a specification outperforms the benchmark. The table suggests that replacing the federal funds rate by its shadow rate delivers RMSFE for GDP growth that are slightly above those of the benchmark, but ratios are not significantly different from unity at any horizon. In contrast, the logarithmic approach shows RMSFE ratios that are significantly larger than one at long horizons. At $h = 4$ the RMSFE is 47% higher as compared to the benchmark. Moreover, both policy shock specifications produce similar RMSFE ratios of around unity. And despite its biased predictive mean, the soft condition specification does not significantly worse than the benchmark. In fact, if coupled with external information, the soft condition specification is the only model that delivers RMSFE ratios below unity at all horizons.

For the unemployment rate I detect a small but significant improvement of the shadow rate specification over the benchmark for $h = 2$ and later. In contrast, both soft condition specifications inflate the RMSFE a lot for the unemployment rate, while performing satisfactory for CPI inflation. The logarithmic specification does significantly worse for both at $h = 4$. Furthermore, the policy shock specification, with and without external information, is not significantly different from the benchmark in terms of RMSFE.

With respect to the treasury rate, the RMSFE vary a lot across specifications. Relative to the benchmark, the shadow rate specification displays small but significant gains at all horizons, while including $\log(R_t)$ rather than $R_t$ is even more profitable. The logarithmic specification cuts the RMSFE by 31% at $h = 4$. In contrast, the soft condition specification inflates the RMSFE a lot, with ratios between 1.34 and 1.50. Conditioning on financial market expectations is again

19

## Table 5: Root Mean Squared Forecast Errors

| Specification | Forecast Horizon $h$ (in Quarters) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| GDP Growth | | | | |
| Benchmark | 2.55 | 2.44 | 2.75 | 2.18 |
| Shadow Rate | 1.03 | 1.03 | 1.04 | 1.08 |
| Logarithmic | 1.06 | 1.10 | 1.17* | 1.47** |
| Policy Shock | 1.00 | 1.02 | 1.03 | 1.01 |
| Policy Shock (EI) | 1.00 | 1.03 | 1.03 | 1.02* |
| Soft Condition | 1.03 | 0.96 | 1.02 | 1.09 |
| Soft Condition (EI) | 0.95 | 0.89 | 0.91 | 0.96 |
| Unemployment Rate | | | | |
| Benchmark | 0.31 | 0.54 | 0.70 | 0.90 |
| Shadow Rate | 0.99 | 0.98* | 0.96** | 0.96*** |
| Logarithmic | 0.85 | 0.92 | 1.02 | 1.11* |
| Policy Shock | 1.00 | 1.03 | 1.05 | 1.06 |
| Policy Shock (EI) | 1.00 | 1.02 | 1.04 | 1.05 |
| Soft Condition | 1.25 | 1.34 | 1.42 | 1.42 |
| Soft Condition (EI) | 1.20 | 1.25 | 1.28* | 1.27** |
| CPI Inflation | | | | |
| Benchmark | 2.56 | 3.14 | 2.80 | 1.96 |
| Shadow Rate | 1.00 | 0.99 | 0.96 | 0.93** |
| Logarithmic | 0.99* | 1.07 | 1.11 | 1.19* |
| Policy Shock | 1.00 | 1.01 | 1.00 | 0.92 |
| Policy Shock (EI) | 1.00 | 1.01** | 1.01 | 0.89 |
| Soft Condition | 0.88 | 0.92 | 0.91 | 1.01 |
| Soft Condition (EI) | 0.88 | 0.90 | 0.88 | 0.86 |
| Treasury Rate | | | | |
| Benchmark | 0.48 | 0.88 | 1.20 | 1.52 |
| Shadow Rate | 0.96*** | 0.94*** | 0.94*** | 0.93*** |
| Logarithmic | 0.91* | 0.84* | 0.74* | 0.69* |
| Policy Shock | 1.09 | 1.04* | 1.04** | 1.05 |
| Policy Shock (EI) | 1.05 | 0.97 | 0.96 | 0.96 |
| Soft Condition | 1.50** | 1.41** | 1.38** | 1.34** |
| Soft Condition (EI) | 1.17 | 0.99 | 0.93 | 0.92 |

Notes: this table shows the root mean squared forecast errors for the period 2009:$h$ to 2014:$h$. The errors are provided in levels for the benchmark and as ratios relative to the benchmark for the other specifications. A ratio below unity means that the respective specification outperforms the benchmark. ***, **, and * indicate that errors are significantly different from each other at the 1%, 5%, and 10% level, respectively.

helpful in this case, leading to small but not significant gains at longer horizons compared to the benchmark. Moreover, the point forecast accuracy of the policy shock specification is similar to that of the benchmark, while introducing an upper bound for the federal funds rate improves its performance slightly. All in all, these RMSFE are in line with the MFE, which detected a large variation in the location of predictive means across specifications.

To conclude, I briefly comment on the MAFE, which are reported in Table 6. By and large, the findings are qualitatively similar to those based on RMSFE and differences between them concern significance only. For instance, the shadow rate specification shows significantly higher MAFE for GDP growth at longer horizons, while the benefits of this specifications regarding the unemployment rate are not significant anymore. Yet, the superior forecast ability of the shadow rate specification for CPI inflation and the treasury rate is confirmed. Moreover, the logarithmic specification still shows significantly lower errors for the latter at $h = 2$ and later.

## 4.4   Density Forecasts

Next I extend the forecast comparison to the predictive density by presenting the LPDS in Table 7. The scores are in levels for the benchmark and in differences to the benchmark for the other specifications. A positive difference thus means that a specification outperforms the benchmark. The table shows that all specifications produce similar LPDS for GDP growth, except the logarithmic specification that delivers low scores. At $h = 4$ its LPDS is 30% lower than that of the benchmark and the difference significant. Due to a biased predictive mean and a large variance, the logarithmic specification attaches a too low probability to actual growth rates.

For the unemployment rate I obtain very different LPDS that are also surrounded by a lot uncertainty. On the one side, both soft condition specifications deliver LPDS that are much smaller than those of the benchmark. At $h = 4$ their score is 59% lower with external information and 42% without, while only the former difference is significant. On the other side, the logarithmic specification outperforms the benchmark at $h = 1$ by 13%. However, its density forecast accuracy is deteriorating over time, showing an underperformance of 18% for the longest horizon. The logarithmic specification often shows this pattern of relatively accurate forecasts at $h = 1$ and poor performance at longer horizons, suggesting that the nonlinear relation between the federal funds rate and other macroeconomic variables is not supported by data, inflating forecast errors over time. Moreover, I find that both policy shock specifications produce scores that are a bit lower, while the shadow rate specification shows small but insignificant gains. For CPI inflation the scores are similar across specifications. Only the logarithmic specification shows a significantly lower LPDS at $h = 4$, while the shadow rate specification delivers significant gains.

Table 6: Mean Absolute Forecast Errors

| Specification | Forecast Horizon $h$ (in Quarters) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| GDP Growth | | | | |
| Benchmark | 1.88 | 1.96 | 1.93 | 1.62 |
| Shadow Rate | 1.01 | 1.02 | 1.02* | 1.07*** |
| Logarithmic | 1.14 | 1.12 | 1.27 | 1.50* |
| Policy Shock | 1.00 | 1.02 | 1.01 | 1.02 |
| Policy Shock (EI) | 1.00 | 1.03 | 1.01 | 1.03 |
| Soft Condition | 1.13 | 0.94 | 1.13 | 1.15* |
| Soft Condition (EI) | 1.04 | 0.89 | 1.06 | 1.04 |
| Unemployment Rate | | | | |
| Benchmark | 0.25 | 0.39 | 0.52 | 0.69 |
| Shadow Rate | 0.99 | 0.97 | 0.94 | 0.91 |
| Logarithmic | 0.87 | 0.99 | 1.11 | 1.17 |
| Policy Shock | 1.00 | 1.01 | 1.03 | 1.03 |
| Policy Shock (EI) | 1.00 | 1.01 | 1.03 | 1.04 |
| Soft Condition | 1.16 | 1.28 | 1.33 | 1.28 |
| Soft Condition (EI) | 1.13 | 1.24 | 1.33 | 1.36 |
| CPI Inflation | | | | |
| Benchmark | 2.01 | 2.07 | 1.84 | 1.35 |
| Shadow Rate | 1.02 | 1.02 | 0.94 | 0.89* |
| Logarithmic | 0.89 | 1.00 | 0.98 | 1.17 |
| Policy Shock | 1.00 | 1.01* | 1.01 | 0.98 |
| Policy Shock (EI) | 1.00 | 1.01 | 1.01 | 0.93 |
| Soft Condition | 0.89 | 0.92 | 0.96 | 1.07 |
| Soft Condition (EI) | 0.88 | 0.95 | 0.96 | 0.97 |
| Treasury Rate | | | | |
| Benchmark | 0.40 | 0.70 | 0.98 | 1.27 |
| Shadow Rate | 0.94** | 0.93*** | 0.92*** | 0.92*** |
| Logarithmic | 0.85 | 0.81* | 0.70** | 0.66** |
| Policy Shock | 1.07 | 1.07** | 1.05** | 1.08** |
| Policy Shock (EI) | 1.02 | 1.00 | 0.96 | 0.98 |
| Soft Condition | 1.45 | 1.47* | 1.43** | 1.39** |
| Soft Condition (EI) | 1.04 | 1.01 | 0.94 | 0.93 |

Notes: this table shows the mean absolute forecast errors for the period 2009:$h$ to 2014:$h$. The errors are provided in levels for the benchmark and as ratios relative to the benchmark for the other specifications. A ratio below unity means that the respective specification outperforms the benchmark. ***, **, and * indicate that errors are significantly different from each other at the 1%, 5%, and 10% level, respectively.

Table 7: Average Log Predictive Density Scores

| Specification | Forecast Horizon $h$ (in Quarters) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| GDP Growth | | | | |
| Benchmark | −2.39 | −2.34 | −2.43 | −2.33 |
| Shadow Rate | −0.01 | −0.02 | −0.01 | −0.05** |
| Logarithmic | −0.06 | −0.11 | −0.18* | −0.30** |
| Policy Shock | 0.01 | −0.00 | −0.01 | −0.01 |
| Policy Shock (EI) | 0.00 | −0.00 | −0.00 | −0.01 |
| Soft Condition | −0.04 | −0.02 | −0.02 | −0.01 |
| Soft Condition (EI) | 0.07 | 0.05 | 0.04 | 0.06 |
| Unemployment Rate | | | | |
| Benchmark | −0.24 | −0.75 | −0.96 | −1.21 |
| Shadow Rate | 0.03 | 0.02 | 0.03 | 0.04 |
| Logarithmic | 0.13 | 0.00 | −0.11 | −0.18** |
| Policy Shock | 0.00 | −0.03* | −0.04 | −0.04 |
| Policy Shock (EI) | 0.00 | −0.03* | −0.03 | −0.04* |
| Soft Condition | −0.28 | −0.40 | −0.46 | −0.42 |
| Soft Condition (EI) | −0.29 | −0.52* | −0.58** | −0.59** |
| CPI Inflation | | | | |
| Benchmark | −2.18 | −2.28 | −2.15 | −2.04 |
| Shadow Rate | −0.02 | 0.01 | 0.03 | 0.04** |
| Logarithmic | −0.05 | −0.10 | −0.06 | −0.08** |
| Policy Shock | −0.01 | 0.00 | 0.00 | 0.00 |
| Policy Shock (EI) | −0.01 | −0.01 | −0.00 | 0.03 |
| Soft Condition | 0.05 | 0.01 | −0.01 | −0.05 |
| Soft Condition (EI) | 0.06 | 0.04 | 0.02 | 0.05 |
| Treasury Rate | | | | |
| Benchmark | −0.71 | −1.33 | −1.69 | −1.96 |
| Shadow Rate | 0.04** | 0.07*** | 0.10*** | 0.10*** |
| Logarithmic | 0.09 | 0.21* | 0.37** | 0.47** |
| Policy Shock | −0.04 | −0.03 | −0.05* | −0.08** |
| Policy Shock (EI) | 0.00 | 0.04 | 0.04 | 0.05 |
| Soft Condition | −0.41*** | −0.53*** | −0.58*** | −0.60*** |
| Soft Condition (EI) | 0.00 | 0.08 | 0.13 | 0.13 |

Notes: this table shows the average log predictive density scores for the period 2009:$h$ to 2014:$h$. The scores are provided in levels for the benchmark and in differences to the benchmark for the other specifications. A positive difference means that the respective specification outperforms the benchmark. ***, **, and * indicate that scores are significantly different from each other at the 1%, 5%, and 10% level.

Replacing the federal funds rate by its shadow rate or logarithm is also profitable for the treasury rate in terms of LPDS. Both specifications produce significantly larger scores than the benchmark, and this benefit increases with the forecast horizon. In contrast, the soft condition specification displays significantly lower LPDS at all horizons. This incorrect calibration can be avoided by including external information. In this case, the soft condition specification delivers even higher scores than the benchmark, but not significantly. As before, the policy shock specification (EI) is as accurate as the benchmark, while excluding the external information leads to an inferior performance at longer horizons.

Before I assess the density forecast accuracy of the different specifications in absolute terms using PIT histograms, I briefly comment on the CRPS, which are shown in Table 8 for the benchmark in levels (pp) and the other specifications as ratios relative to the benchmark. A ratio below unity thus means that a specification outperforms the benchmark. Overall, the ratios are similar to those obtained under the LPDS, suggesting that the scores are not distorted by outliers, but are a reasonable description of the density forecast accuracy of the specifications.

I conclude the forecast comparison by presenting the PIT histograms in Figures 3 and 4 for $h = 1$ and $h = 4$, respectively. This proceeding is useful since the shape of the histograms allows me to explore the reasons behind the poor density forecast accuracy of some specifications. In particular, I obtain that most of them significantly overpredict GDP growth at $h = 4$. Except for the soft condition specification with external information, not a single observation falls into the fifth bin for any specification. However, this miscalibration is limited to the longest forecast horizon. For the shortest the densities are largely consistent with true densities, while the soft condition specification with external information again shows a particularly realistic calibration, in line with its low MFE and relatively strong performance in terms of LPDS and CRPS.

For the unemployment rate, I obtain u-shaped PIT histograms for all specifications at $h = 1$, meaning that a too low probability is assigned to average values. Noteworthy, the densities are often calibrated better for $h = 4$ than $h = 1$. For instance, the benchmark and both policy shock specifications produce histograms that seem to be uniform for $h = 4$. In contrast, the densities generated by both soft condition specifications are poorly calibrated, which is consistent with their underperformance in terms of LPDS and CRPS. The PIT histograms also suggest that all specifications underpredict CPI inflation at $h = 1$, but tend to overpredict the series at $h = 4$. In particular, the policy shock and soft condition specification show a strong overprediction.

Finally, the PIT histograms detect a poor calibration of nearly all specifications with respect to the treasury rate. At $h = 1$ the benchmark significantly overpredicts the series, while replacing the federal funds rate by its logarithm or using the soft condition specification (EI) helps to

## Table 8: Average Continuous Ranked Probability Scores

| Specification | Forecast Horizon $h$ (in Quarters) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| *GDP Growth* | | | | |
| Benchmark | 1.42 | 1.38 | 1.50 | 1.30 |
| Shadow Rate | 1.02 | 1.02 | 1.02 | 1.06* |
| Logarithmic | 1.07 | 1.11 | 1.19* | 1.40** |
| Policy Shock | 1.00 | 1.01 | 1.02 | 1.01 |
| Policy Shock (EI) | 1.00 | 1.02 | 1.02 | 1.02 |
| Soft Condition | 1.05 | 0.98 | 1.03 | 1.05 |
| Soft Condition (EI) | 0.97 | 0.93 | 0.96 | 0.97 |
| *Unemployment Rate* | | | | |
| Benchmark | 0.17 | 0.29 | 0.38 | 0.48 |
| Shadow Rate | 0.99 | 0.98* | 0.96* | 0.95* |
| Logarithmic | 0.87 | 0.96 | 1.06 | 1.15 |
| Policy Shock | 1.00 | 1.02 | 1.05 | 1.06 |
| Policy Shock (EI) | 1.00 | 1.02 | 1.05 | 1.06 |
| Soft Condition | 1.21 | 1.31 | 1.40 | 1.41 |
| Soft Condition (EI) | 1.18 | 1.27 | 1.37* | 1.41** |
| *CPI Inflation* | | | | |
| Benchmark | 1.35 | 1.52 | 1.34 | 1.02 |
| Shadow Rate | 1.01 | 0.99 | 0.96 | 0.94** |
| Logarithmic | 0.99 | 1.04 | 1.05 | 1.13* |
| Policy Shock | 1.00 | 1.01* | 1.01 | 0.97 |
| Policy Shock (EI) | 1.00 | 1.01** | 1.01 | 0.94 |
| Soft Condition | 0.92 | 0.96 | 0.97 | 1.07 |
| Soft Condition (EI) | 0.91 | 0.94 | 0.95 | 0.94 |
| *Treasury Rate* | | | | |
| Benchmark | 0.28 | 0.51 | 0.71 | 0.92 |
| Shadow Rate | 0.96** | 0.94*** | 0.93*** | 0.93*** |
| Logarithmic | 0.91 | 0.83* | 0.71** | 0.65** |
| Policy Shock | 1.07 | 1.04** | 1.05** | 1.06** |
| Policy Shock (EI) | 1.02 | 0.97 | 0.96 | 0.96 |
| Soft Condition | 1.49** | 1.47*** | 1.45** | 1.42** |
| Soft Condition (EI) | 1.07 | 0.98 | 0.92 | 0.90 |

Notes: this table shows the average continuous ranked probability scores for the period 2009:$h$ to 2014:$h$. The scores are provided in levels for the benchmark and as ratios relative to the benchmark for the other specifications. A ratio below unity means that the respective specification outperforms the benchmark. ***, **, and * indicate that scores are significantly different from each other at the 1%, 5%, and 10% level.
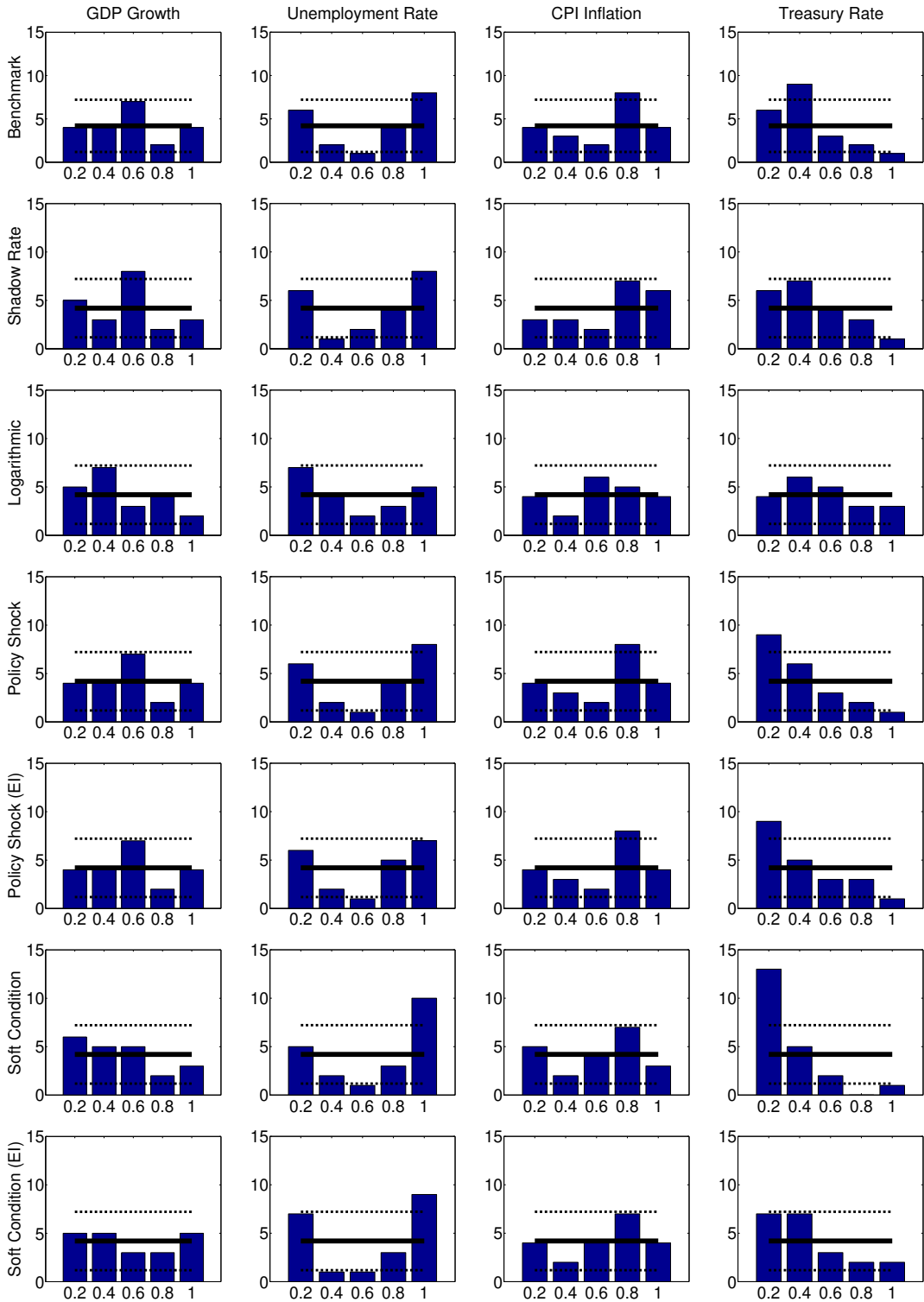
Figure 3: Probability Integral Transform Histograms ($h = 1$). Notes: this figure shows the probability integral transform histograms for $h = 1$ and the period 2009:1 to 2014:1. The solid line denotes uniformity. The dashed lines are a 90% confidence band. x-axis: quantile; y-axis: number of realizations.
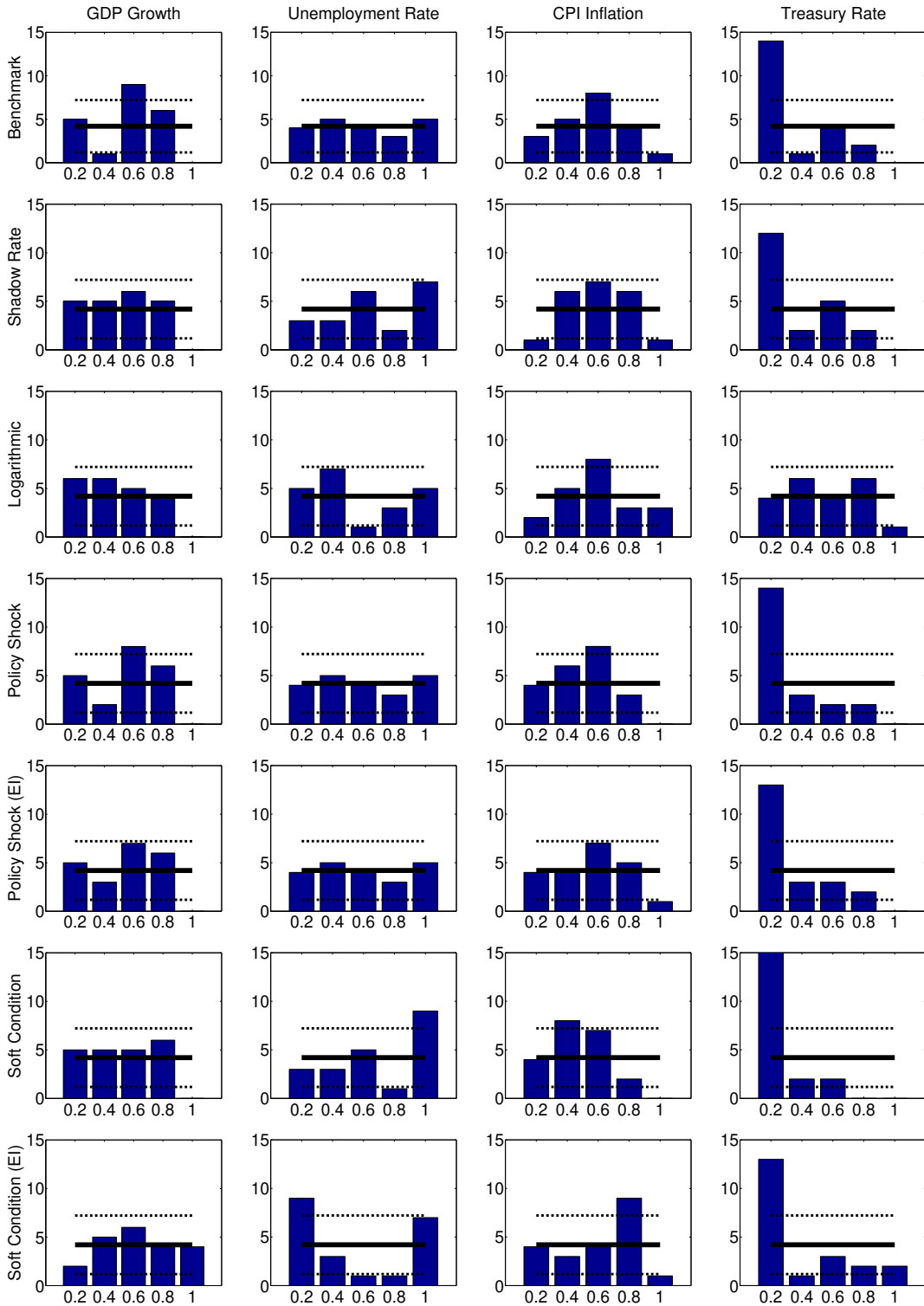
Figure 4: Probability Integral Transform Histograms ($h = 4$). Notes: this figure shows the probability integral transform histograms for $h = 4$ and the period 2009:4 to 2014:4. The solid line denotes uniformity. The dashed lines are a 90% confidence band. x-axis: quantile; y-axis: number of realizations.

achieve correct calibration. For $h = 4$, both specifications are, however, as inapt as the benchmark in producing a realistic predictive density for the treasury rate. Across both horizons, the logarithmic specification is the one that generates PIT histograms that come at least closest to uniformity, explaining why it performs relatively well in terms of LPDS and CRPS.

# 5   Summary and Conclusion

In this paper I discuss how the forecast accuracy of a BVAR is affected by introducing the zero lower bound on the federal funds rate. As a benchmark I adopt a common BVAR specification, including 18 variables, estimated shrinkage, and no nonlinearity. Then I entertain alternative specifications of the zero lower bound: replace the federal funds rate by its shadow rate, consider a logarithmic transformation, feed in monetary policy shocks, or use a rejection sampler. The latter two are also coupled with interest rate expectations from future contracts. The comparison is based on the accuracy of point and density forecasts of GDP growth, the unemployment rate, CPI inflation, and the treasury rate during the period 2009:1 to 2014:4. For a robust assessment I consider several relative and absolute forecast metrics.

I obtain a large set of empirical results in this paper. To condense the material, I provide a qualitative summary of them in Tables 9 and 10 for the shortest and longest forecast horizon, respectively. My interpretation of the results is as follows. First, the introduction of the zero lower bound hardly affects the forecast accuracy of the BVAR with respect to real activity and prices at $h = 1$ but does at $h = 4$. While this is by construction for the policy shock specifications, it is also observed for the other specifications. A practical forecaster should thus be more concerned about the zero lower bound when she is interested in long-term forecasts. Second, the performance of the alternative specifications is greatly different at $h = 4$, suggesting that this modeling choice is not innocuous. In particular, the introduction of the zero lower bound is not beneficial *per se*, but it depends on how it is done and which series is forecasted.

With caution, I recommend the shadow rate specification. Except for GDP growth, this specification shows a satisfactory forecast ability in absolute terms and relative to the benchmark. In contrast, both policy shock specifications are no good choices, while the logarithmic specifications is accurate for the treasury rate, but inflates forecast errors for any other series. Finally, the soft condition specification is useful only in combination with external information, but then helps to achieve a better calibration of predictive densities. Since the policy rate will remain low for some time, these findings could prove useful for practical forecasters.

28

Table 9: Summary of Results ($h = 1$)

| Specification | Forecast Metric | | | | | |
|---|---|---|---|---|---|---|
| | MFE | RMSFE | MAFE | LPDS | CRPS | PITS |
| **GDP Growth** | | | | | | |
| Benchmark | ✓ | | | | | ✓ |
| Shadow Rate | ✓ | | | | | ✗ |
| Logarithmic | ✗ | | | | | ✓ |
| Policy Shock | ✓ | | | | | ✓ |
| Policy Shock (EI) | ✓ | | | | | ✓ |
| Soft Condition | ✗ | | | | | ✓ |
| Soft Condition (EI) | ✓ | | | | | ✓ |
| **Unemployment Rate** | | | | | | |
| Benchmark | ✓ | | | | | ✗ |
| Shadow Rate | ✓ | | | | | ✗ |
| Logarithmic | ✓ | | | | | ✓ |
| Policy Shock | ✓ | | | | | ✗ |
| Policy Shock (EI) | ✓ | | | | | ✗ |
| Soft Condition | ✓ | | | | | ✗ |
| Soft Condition (EI) | ✓ | | | | | ✗ |
| **CPI Inflation** | | | | | | |
| Benchmark | ✓ | | | | | ✗ |
| Shadow Rate | ✓ | | | | | ✓ |
| Logarithmic | ✓ | ✓ | | | | ✓ |
| Policy Shock | ✓ | | | | | ✗ |
| Policy Shock (EI) | ✓ | | | | | ✗ |
| Soft Condition | ✓ | | | | | ✓ |
| Soft Condition (EI) | ✓ | | | | | ✓ |
| **Treasury Rate** | | | | | | |
| Benchmark | ✗ | | | | | ✗ |
| Shadow Rate | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ |
| Logarithmic | ✓ | ✓ | | | | ✓ |
| Policy Shock | ✗ | | | | | ✗ |
| Policy Shock (EI) | ✗ | | | | | ✗ |
| Soft Condition | ✗ | ✗ | | ✗ | ✗ | ✗ |
| Soft Condition (EI) | ✗ | | | | | ✓ |

Notes: this table shows a summary of the results for $h = 1$. Absolute metrics: ✓ = correct calibration, ✗ = incorrect calibration. Relative metrics: ✓ = gains over benchmark; ✗ = losses over benchmark; blank = no gains/losses over benchmark. Significance level is 10% in each case.

Table 10: Summary of Results ($h = 4$)

| Specification | Forecast Metric | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | MFE | RMSFE | MAFE | LPDS | CRPS | PITS |
| GDP Growth | | | | | | |
| Benchmark | ✓ | | | | | ✗ |
| Shadow Rate | ✗ | | ✗ | ✗ | ✗ | ✗ |
| Logarithmic | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Policy Shock | ✓ | | | | | ✗ |
| Policy Shock (EI) | ✓ | ✗ | | | | ✗ |
| Soft Condition | ✗ | | ✗ | | | ✗ |
| Soft Condition (EI) | ✓ | | | | | ✓ |
| Unemployment Rate | | | | | | |
| Benchmark | ✓ | | | | | ✓ |
| Shadow Rate | ✓ | ✓ | | | ✓ | ✓ |
| Logarithmic | ✓ | ✗ | | ✗ | | ✗ |
| Policy Shock | ✓ | | | | | ✓ |
| Policy Shock (EI) | ✓ | | | ✗ | | ✓ |
| Soft Condition | ✓ | | | | | ✗ |
| Soft Condition (EI) | ✓ | ✗ | | ✗ | ✗ | ✗ |
| CPI Inflation | | | | | | |
| Benchmark | ✓ | | | | | ✗ |
| Shadow Rate | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ |
| Logarithmic | ✓ | ✗ | | ✗ | ✗ | ✗ |
| Policy Shock | ✓ | | | | | ✗ |
| Policy Shock (EI) | ✓ | | | | | ✗ |
| Soft Condition | ✗ | | | | | ✗ |
| Soft Condition (EI) | ✓ | | | | | ✗ |
| Treasury Rate | | | | | | |
| Benchmark | ✗ | | | | | ✗ |
| Shadow Rate | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ |
| Logarithmic | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ |
| Policy Shock | ✗ | | ✗ | ✗ | ✗ | ✗ |
| Policy Shock (EI) | ✗ | | | | | ✗ |
| Soft Condition | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Soft Condition (EI) | ✗ | | | | | ✗ |

Notes: this table shows a summary of the results for $h = 4$. Absolute metrics: ✓ = correct calibration, ✗ = incorrect calibration. Relative metrics: ✓ = gains over benchmark; ✗ = losses over benchmark; blank = no gains/losses over benchmark. Significance level is 10% in each case.

# References

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

Arora, S., M. A. Little, and P. E. McSharry (2013). Nonlinear and Nonparametric Modeling Approaches for Probabilistic Forecasting of the US Gross National Product. *Studies in Nonlinear Dynamics and Econometrics 17*(4), 395–420.

Bańbura, M., D. Giannone, and M. Lenza (2015). Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections. *International Journal of Forecasting 31*(3), 739–756.

Bańbura, M., D. Giannone, and L. Reichlin (2010). Large Bayesian Vector Autoregressions. *Journal of Applied Econometrics 25*(1), 71–92.

Berg, T. O. and S. R. Henzel (2015). Point and Density Forecasts for the Euro Area Using Bayesian VARs. forthcoming in *International Journal of Forecasting*.

Bernanke, B. S. (2013). Communication and Monetary Policy. Speech at the National Economists Club Annual Dinner, Herbert Stein Memorial Lecture, Washington, D.C., November 19.

Carriero, A., T. E. Clark, and M. Marcellino (2015). Bayesian VARs: Specification Choices and Forecast Accuracy. *Journal of Applied Econometrics 30*(1), 46–73.

Clark, T. E. and M. W. McCracken (2014). Evaluating Conditional Forecasts from Vector Autoregressions. Working Paper 1413, Federal Reserve Bank of Cleveland.

Clark, T. E. and F. Ravazzolo (2015). Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility. *Journal of Applied Econometrics 30*(4), 551–575.

Cogley, T. and T. J. Sargent (2002). Evolving Post-World War II U.S. Inflation Dynamics. In B. S. Bernanke and K. Rogoff (Eds.), *NBER Macroeconomics Annual 2001*, Volume 16. Cambridge: MIT Press.

Del Negro, M. and F. Schorfheide (2013). DSGE Model-Based Forecasting. In G. Elliot and A. Timmerman (Eds.), *Handbook of Economic Forecasting*, Volume 24, pp. 57–140. North-Holland: Elsevier.

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

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

Gelman, A., J. Hwang, and A. Vehtari (2014). Understanding Predictive Information Criteria for Bayesian Models. *Statistics and Computing 24*(6), 997–1016.

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

Giacomini, R. and H. White (2006). Tests of Conditional Predicitve Ability. *Econometrica 74*(6), 1545–1578.

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior Selection for Vector Autoregressions. *Review of Economics and Statistics 97*(2), 412–435.

Gneiting, T. and A. E. Raftery (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association 102*(477), 359–378.

Kadiyala, K. R. and S. Karlsson (1993). Forecasting with Generalized Bayesian Vector Autoregressions. *Journal of Forecasting 12*(4), 365–378.

Kadiyala, K. R. and S. Karlsson (1997). Numerical Methods for Estimation and Inference in Bayesian VAR-Models. *Journal of Applied Econometrics 12*(2), 99–132.

Koop, G. (2013). Forecasting with Medium and Large Bayesian VARS. *Journal of Applied Econometrics 28*(2), 177–203.

Krippner, L. (2013). Measuring the Stance of Monetary Policy in Zero Lower Bound Environments. *Economics Letters 118*(1), 135–138.

Litterman, R. B. (1980). A Bayesian Procedure for Forecasting with Vector Autoregressions. Manuscript, Massachusetts Institute of Technology, Department of Economics.

Litterman, R. B. (1986). Forecasting With Bayesian Vector Autoregressions - Five Years of Experience. *Journal of Business and Economic Statistics 4*(1), 25–38.

Newey, W. K. and K. D. West (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica 55*(3), 703–708.

Ravazzolo, F. and S. P. Vahey (2014). Forecast Densities for Economic Aggregates from Disaggregate Ensembles. *Studies in Nonlinear Dynamics and Econometrics 18*(4), 367–381.

Rossi, B. and T. Sehkposyan (2014). Evaluating Predictive Densities of US Output Growth and Inflation in a Large Macroeconomic Data Set. *International Journal of Forecasting 30*(3), 662–682.

Smets, F., A. Warne, and R. Wouters (2014). Professional Forecasters and Real-Time Forecasting with a DSGE Model. *International Journal of Forecasting 30*(4), 981–995.

Waggoner, D. F. and T. Zha (1999). Conditional Forecasts in Dynamic Multivariate Models. *Review of Economics and Statistics 81*(4), 639–651.

Weiss, A. A. (1996). Estimating Time Series Models Using the Relevant Cost Function. *Journal of Applied Econometrics 11*(5), 539–560.

Wu, J. C. and F. D. Xia (2015). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. forthcoming in *Journal of Money, Credit, and Banking*.

# A    Additional Figures
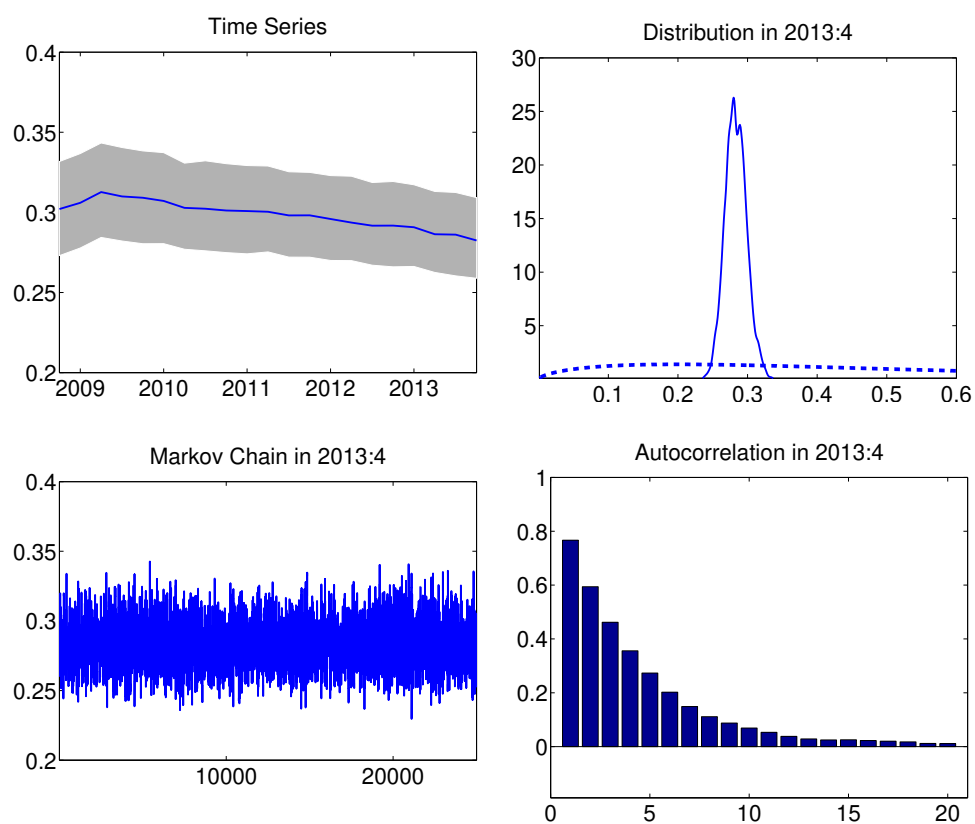


Figure A.1: Posterior Distribution of Shrinkage Parameter. Notes: this figure shows the posterior distribution of the shrinkage parameter. Top left: Median and 90% error band for all periods. Top right: Posterior distribution (solid) in final period and prior distribution (dashed). Bottom left: Markov chain in final period. Bottom right: Autocorrelation function in final period.
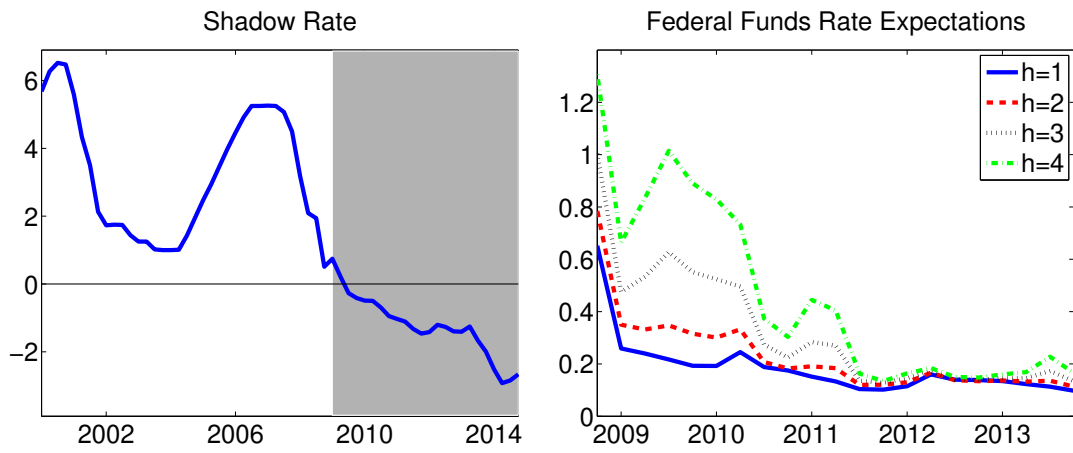
**Figure A.2: Data.** Notes: this figure shows the Wu-Xia shadow rate and the federal funds rate expectations from future contracts. The shaded area is the evaluation period. x-axis: year; y-axis: percent.

# Ifo Working Papers

No. 190    Auer, W. and N. Danzer, Fixed-Term Employment and Fertility: Evidence from German Micro Data, October 2014.

No. 189    Rösel, F., Co-Partisan Buddies or Partisan Bullies? Why State Supervision of Local Government Borrowing Fails, October 2014.

No. 188    Kauder, B., Incorporation of Municipalities and Population Growth – A Propensity Score Matching Approach, October 2014.

No. 187    Méango, R., Financing Student Migration: Evidence for a Commitment Problem, September 2014.

No. 186    Nagl, W. and M. Weber, Unemployment compensation and unemployment duration before and after the German Hartz IV reform, September 2014.

No. 185    Potrafke, N. and M. Reischmann, Explosive Target balances of the German Bundesbank, July 2014.

No. 184    Eppinger, P.S. and G.J. Felbermayr, Bilateral Trade and Similarity of Income Distributions: The Role of Second Moments, July 2014.

No. 183    Wohlrabe, K., Das FAZ-Ökonomenranking 2013: Eine kritische Betrachtung, Juli 2014.

No. 182    Lehmann, R. and A. Weyh, Forecasting employment in Europe: Are survey result helpful?, June 2014.

No. 181    Schinke, C., Government Ideology, Globalization, and Top Income Shares in OECD Countries, June 2014.

No. 180    Benz, S., M. Larch and M. Zimmer, The Structure of the German Economy, May 2014.

No. 179    Meier, V. and H. Rainer, Pigou Meets Ramsey: Gender-Based Taxation with Non-Cooperative Couples, May 2014.

No. 178    Kugler, F., G. Schwerdt und L. Wößmann, Ökonometrische Methoden zur Evaluierung kausaler Effeke der Wirtschaftspolitik, April 2014.

No. 177    Angerer, S., D. Glätzle-Rützler, P. Lergetporer and M. Sutter, Donations, risk attitudes and time preferences: A study on altruism in primary school children, March 2014.