

CAMP Working Paper Series  
No 8/2015

# Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts

Fabian Krüger, Todd E. Clark and Francesco Ravazzolo



© Authors 2015

This paper can be downloaded without charge from the CAMP website <http://www.bi.no/camp>

# Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts\*

Fabian Krüger<sup>†</sup>

Todd E. Clark<sup>‡</sup>

Francesco Ravazzolo<sup>§</sup>

August 15, 2015

## Abstract

This paper shows entropic tilting to be a flexible and powerful tool for combining medium-term forecasts from BVARs with short-term forecasts from other sources (nowcasts from either surveys or other models). Tilting systematically improves the accuracy of both point and density forecasts, and tilting the BVAR forecasts based on nowcast means and variances yields slightly greater gains in density accuracy than does just tilting based on the nowcast means. Hence entropic tilting can offer — more so for persistent variables than not-persistent variables — some benefits for accurately estimating the uncertainty of multi-step forecasts that incorporate nowcast information.

*Keywords:* Forecasting, Prediction, Bayesian Analysis

*JEL classification code:* E17, C11, C53

---

\*We thank Frank Schorfheide, Ellis Tallman, and the participants of the workshop on “Uncertainty and Economic Forecasting” (London, April 2014), the conference on “Uncertainty and Probabilistic Forecasting during the Financial and Economic Crisis” (Heidelberg, June 2014), the conference on “Advances in Applied Macro-Finance and Forecasting” (Istanbul, September 2014), the 8th CFE conference (Pisa, December 2014), the 25th EC2 conference on “Advances in Forecasting” (Barcelona, December 2014), the Bundesbank research seminar (Frankfurt, March 2015) and the International Symposium on Forecasting (Riverside, June 2015) for helpful comments. The first author gratefully acknowledges financial support from the European Union Seventh Framework Programme (grant agreement no. 290976). The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, Federal Reserve System, or Norges Bank.

<sup>†</sup>Heidelberg Institute for Theoretical Studies, [fabian.krueger@h-its.org](mailto:fabian.krueger@h-its.org)

<sup>‡</sup>Federal Reserve Bank of Cleveland, [todd.clark@clev.frb.org](mailto:todd.clark@clev.frb.org)

<sup>§</sup>Norges Bank and Centre for Applied Macro and Petroleum economics at BI Norwegian Business School, [francesco.ravazzolo@norges-bank.no](mailto:francesco.ravazzolo@norges-bank.no)

# 1 Introduction

It is commonly known that models such as vector autoregressions (VARs) or dynamic stochastic general equilibrium (DSGE) models that are effective in medium-term macroeconomic forecasting are not as effective at short-horizon forecasting. As a result, VARs and DSGE models are often combined with current-quarter forecasts, or nowcasts, from another source. One such source is a judgmental forecast from a central bank or a survey of professional forecasters, motivated by evidence that such forecasts often provide useful information beyond that contained in econometric models (e.g. Ang, Bekaert, and Wei, 2007; Faust and Wright, 2013). Alternatively, relatively accurate short-horizon forecasts can be obtained from bridging equations or factor models, surveyed in Banbura, Giannone, and Reichlin (2013) and Banbura, Giannone, Modugno, and Reichlin (2013). Compared to medium-term forecasting models, these nowcasting approaches improve near-term forecast accuracy by better adding up information in data releases for the current quarter and require dealing with differences in data release dates within the quarter (what is known as the “ragged edge” of data).

A number of methods for combining (VAR or DSGE) medium-term forecasts with nowcasts from another source have been used in the recent literature. Faust and Wright (2009) use short-horizon forecasts from the Federal Reserve Board’s Greenbook as jumping-off points (treating them as data, appended to the actual data) for forecasts obtained from autoregressive and factor-augmented autoregressive models of GDP growth and inflation. Similarly, Faust and Wright (2013) use current-quarter forecasts from the Survey of Professional Forecasters as jumping-off points for inflation forecasts from a range of autoregressive, Phillips curve, and DSGE models. Schorfheide and Song (2015) and Wolters (2015) treat nowcasts from the Greenbook as data in forming forecasts at subsequent horizons from, respectively, a Bayesian VAR (BVAR) and DSGE models. Del Negro and Schorfheide (2013) combine current quarter Blue Chip Consensus forecasts of GDP growth, inflation, and interest rates with DSGE model forecasts by treating the Blue Chip forecasts as noisy data for the quarter, using Kalman filter methods for signal extraction. Frey and Mokinski (2015) use survey nowcasts in estimating the parameters of a VAR. While we are not aware of published examples, in practice it also seems to be common to use conditional forecast methods (see e.g. Doan, Litterman, and Sims, 1984) to incorporate nowcast information into medium-term forecasts from BVARs. Finally, while this discussion and our analysis focuses on combining forecasts from different sources, an alternative approach is to specify a single model in mixed frequency data

(e.g., quarterly and monthly). For example, Schorfheide and Song (2015) and Giannone, Monti, and Reichlin (2014) develop mixed frequency BVAR and DSGE models, respectively.

As this review suggests, there is no single, standard approach for combining forecasts from medium-term projection models with short-term forecasts from other sources, either surveys or nowcasting models. In this paper, we examine the effects of using entropic tilting to combine such forecasts. Entropic tilting is a technique for modifying a baseline distribution such that it matches certain moment conditions of interest. Robertson, Tallman, and Whiteman (2005) introduced tilting into macroeconomic forecasting, using it to impose conditions on policy rates in a small BVAR forecasting model. Cogley, Morozov, and Sargent (2005) used tilting to produce BVAR forecasts that conditioned on information in the Bank of England’s forecast. More recently, Altavilla, Giacomini, and Ragusa (2013) use entropic tilting to combine survey-based forecasts of short-term interest rates with yield curve forecasts from econometric models, and Lewis and Whiteman (2015) use tilting to improve forecasts of tax revenues in Iowa. These studies primarily focus on point forecasts — not only tilting based on point forecasts but also measuring performance in terms of point forecast accuracy.

Compared to some other existing approaches for combining forecasts from multiple sources, tilting has the advantage of being highly flexible. This flexibility is needed here. In particular, merging a multi-step BVAR forecast density with an external nowcast is not a traditional density combination problem in the spirit of Stone (1961), Hall and Mitchell (2007), Geweke and Amisano (2011) and Gneiting and Ranjan (2013). All of these studies consider a set of densities  $f_1, \dots, f_n$  which refer to the same (univariate or multivariate) random variable. Our setting is different in two respects: first, the nowcast refers to a univariate random variable whereas the BVAR density is jointly for five forecast horizons. Second, the nowcast does not come as a full density but only as a set of moment conditions. Furthermore, compared to simpler approaches such as treating the nowcast as additional data, the flexibility of tilting permits the forecaster to properly capture uncertainty around the combined forecast.

Building on the aforementioned prior research, we use tilting to improve macroeconomic forecasts from BVARs by combining them with nowcasts from surveys and specialized models. Extending past research, we consider tilting the BVAR forecast distributions toward not just the means but also the variances of the nowcasts, and we consider the effects of tilting on the accuracy of not only point forecasts but also density forecasts. We also compare how proper combination

of forecasts via tilting affects estimates of forecast uncertainty compared to cruder approaches that do not account for nowcast uncertainty.

In our implementation, we focus on forecasts of (U.S.) GDP growth, the unemployment rate, inflation in GDP price index, and the 3-month Treasury bill (T-bill) rate. A range of studies have considered similar variable sets (e.g. Clark, 2011; D’Agostino, Gambetti, and Giannone, 2013). We use forecasts from a BVAR with stochastic volatility as in Clark and Ravazzolo (2015). The survey-based forecasts we consider are taken from the Survey of Professional Forecasters (SPF). We also consider model-based nowcasts (current-quarter forecasts); for GDP and inflation, the model uses the Bayesian mixed frequency formulation of Carriero, Clark, and Marcellino (2014), while for the unemployment and T-bill rates, we use small VARs in monthly data (to construct quarterly nowcasts), detailed below.

Broadly, our results show entropic tilting to be a flexible, powerful, and effective tool for combining forecasts from BVARs with external nowcasts. We show that tilting, like other approaches to combining BVAR forecasts with nowcasts, systematically improves the accuracy of point forecasts of standard macroeconomic variables. Extending previous work, we also find that tilting based on nowcast means systematically improves the accuracy of density forecasts from our BVAR. We go on to show that tilting the BVAR forecasts based on not only nowcast means but also nowcast variances yields slightly greater gains in density accuracy than does just tilting based on the nowcast means. For less persistent variables such as GDP growth, the accuracy gains tend to die out as the forecast horizon increases, but for unemployment and interest rates, the gains carry over to horizons as long as five quarters. Our results also show that tilting toward the nowcast mean and variance produces sharper forecast distributions than tilting toward the nowcast mean only. This is because the former approach incorporates the reduced variance of the nowcast – which uses intra-quarter information – whereas the latter approach implicitly conditions on the BVAR variance. Again, these effects are much more pronounced for the more persistent variables.

As to the merits of the survey-based (SPF) nowcasts compared to the model-based nowcasts, for GDP and inflation, survey forecasts from the SPF are hard to beat, so the BVAR is improved more by tilting toward the SPF nowcast than the model-based nowcasts. But for the unemployment and T-bill rates, our model-based nowcasts are more accurate than their SPF counterparts, with corresponding effects on the tilted BVAR forecasts. In a comparison of tilting on a variable-by-variable basis to tilting jointly toward the nowcasts for all four variables of the BVAR, we find

that the overall differences in forecast performance for the joint treatment of variables versus the individual treatment of variables are small.

The paper proceeds as follows. Sections 2 and 3 detail the data and models, respectively. Section 4 explains the implementation of tilting and provides examples. Section 5 provides our main results on entropic tilting. Section 6 presents comparisons to some related combination methods proposed in the literature, and Section 7 concludes. The appendix provides details of our priors and estimation algorithms and presents some analytical derivations mentioned in the paper. A supplementary appendix available online provides additional robustness results (for other combination methods, a shorter sample period, alternative prior settings, and some additional VAR specifications), as well as some additional analysis of the effects of tilting on the forecast distributions.

## 2 Data

### 2.1 Data for models

We use quarterly data to estimate BVAR models (detailed below) for growth of real GDP, inflation in the GDP price index or deflator (henceforth, GDP inflation), the unemployment rate, and the 3-month T-bill rate. We compute GDP growth as 400 times the log difference of real GDP and inflation as 400 times the log difference of the GDP price index, to put them in units of annualized percentage point changes. The unemployment rate and interest rate are also defined in units of percentage points (annualized in the case of the interest rate), with quarterly rates formed as within-quarter averages of monthly rates.

In constructing model-based nowcasts of growth, inflation, unemployment, and the T-bill rate using models detailed in the next section, we rely on a small set of other indicators. For nowcasting GDP growth, we use two monthly coincident indicators taken from Carriero, Clark, and Marcellino (2014): employment growth and the Institute of Supply Management's production index for manufacturing. For nowcasting GDP inflation, we use monthly inflation rates of the CPI ex food and energy, the CPI for food, the CPI for energy, the PPI for capital goods, and the price deflator for new one-family houses under construction. We form nowcasts of unemployment using monthly data on not only unemployment but also growth in payroll employment and new claims for unemployment insurance. Finally, we construct nowcasts of the T-bill rate using monthly data

on the average (for the month) T-bill rate and the 3-month and 6-month T-bill rates on the 15th of the month.

In forming all of our model-based forecasts and nowcasts, for those indicators subject to significant revisions and for which we can easily obtain the needed data, we use real-time data from the Federal Reserve Bank of Philadelphia’s Real Time Dataset for Macroeconomists (RTDSM). The variables for which we use real time data are the following: GDP, GDP price index, monthly unemployment, and monthly employment. Note that, for simplicity, we use “GDP” and “GDP price index” to refer to the output and price series to be forecast, even though the measures are based on GNP and a fixed weight deflator for some of the sample. As described in Croushore and Stark (2001), the quarterly vintages of the RTDSM are dated to reflect the information available around the middle of each quarter. In vintage  $t$ , the available GDP and GDP price index data run through period  $t - 1$ . For all remaining variables, we use currently available data obtained from either the FRED database of the Federal Reserve Bank of St. Louis or from the FAME database of the Federal Reserve Board of Governors: quarterly unemployment and T-bill rates, the Institute of Supply Management’s production index for manufacturing, new claims for unemployment insurance, the CPI ex food and energy, the CPI for food, the CPI for energy, the PPI for capital goods, and the price index for new home construction.

## 2.2 SPF forecast data

We obtain quarterly SPF forecasts of GDP growth, unemployment, GDP inflation, and the T-bill rate from the website of the Federal Reserve of Philadelphia. At each forecast origin, the available forecasts span five quarterly horizons, from the current quarter through the next four quarters. We take the point forecast to be the median of the SPF responses. In some entropic tilting results, we also use a measure of forecast uncertainty. In the presented results, we consider what Clements (2014) refers to as an *ex post* measure: the variance of recent forecast errors, which we compute over the previous 20 forecasts. Specifically, denote by  $\hat{Y}_{t,h}$  the (median) SPF forecast of  $Y_t$  at forecast horizon  $h$  (i.e., the forecast for  $t$  based on data up to  $t - h$ ). Then, our  $h$ -period error measure is computed as

$$\hat{\sigma}_{t,h}^2 = \frac{1}{20} \sum_{r=0}^{19} (Y_{t-D-r} - \hat{Y}_{t-D-r,h})^2,$$

where  $D$  reflects the delay (in quarters) with which the forecaster learns of the relevant realizations data. In line with the considerations in the next section, we set  $D = 2$  for all variables except T-bill

( $D = 1$ ). When considering model-based nowcasts, we instead compute the variance from the model's simulated predictive distribution.

### 2.3 Forecast evaluation sample

We evaluate forecasts from 1988:Q3 through 2013:Q2 (and over a pre-crisis sample of 1988:Q3-2007:Q4, in results presented primarily in the supplementary appendix). The start date of 1988:Q3 marks the earliest possible for a common sample size across variables; SPF forecasts of the T-bill rate do not begin until 1981:Q3, and we require additional observations for computing the forecast error variance at all horizons. For each forecast origin  $t$  starting with 1988:Q3, we estimate the forecast models and construct forecasts of quarterly values of all variables for periods  $t$  and beyond. Consistent with the availability of SPF forecasts, we report results for forecast horizons of 1-5 quarters ahead. In light of the time  $t - 1$  information actually incorporated in the quarterly BVAR models used for forecasting at  $t$ , the 1-quarter ahead forecast is a current quarter ( $t$ ) forecast, while the 2-quarter ahead forecast is a next quarter ( $t + 1$ ) forecast, etc. For the BVAR used to forecast the four variables of interest, the starting point of the model estimation sample is 1955:Q1; we use data for the 1948-54 period to set the priors on some parameters, as detailed in the appendix. For the GDP and inflation nowcasting models, the starting point of model estimation is always 1970:Q2 and 1965:Q1, respectively. For the unemployment rate and T-bill nowcasting models, the estimation samples begin with January 1955 and January 1965, respectively, reflecting data availability.

As discussed in such sources as Romer and Romer (2000), Sims (2002), and Croushore (2006), evaluating the accuracy of real-time forecasts requires a difficult decision on what to take as the actual data in calculating forecast errors. We follow studies such as Romer and Romer (2000) and Faust and Wright (2009) and use the second available estimates of GDP/GNP and the GDP/GNP deflator as actuals in evaluating forecast accuracy. In the case of  $h$ -quarter ahead forecasts made for period  $t + h$  with vintage  $t$  data ending in period  $t - 1$ , the second available estimate is taken from the vintage  $t + h + 2$  data set. In light of our abstraction from real-time revisions in quarterly unemployment and interest rates, we use final vintage data for evaluating forecasts of these series.

### 3 Models

This section provides the specifications of our models and an overview of the estimation methods. The priors and estimation algorithms are detailed in the appendix.

#### 3.1 BVAR specification

We focus on forecasts from a BVAR with random walk stochastic volatility, the specification that Clark and Ravazzolo (2015) found to perform relatively well in a comparison of the forecasting performance (both point and density — stochastic volatility is particularly important for density accuracy) of a range of autoregressive models with and without time-varying volatility.

Let  $y_t$  denote the  $k \times 1$  vector of model variables,  $B_0$  denote a  $k \times 1$  vector of intercepts, and  $B_i, i = 1, \dots, p$ , denote a  $k \times k$  matrix of coefficients on lag  $i$ . For our set of  $k = 4$  variables, we consider a VAR( $p$ ) model with stochastic volatility, with a lag length of  $p = 4$ :

$$\begin{aligned}
 y_t &= B_0 + \sum_{i=1}^p B_i y_{t-i} + v_t, \\
 v_t &= A^{-1} \Lambda_t^{0.5} \epsilon_t, \quad \epsilon_t \sim N(0, I_k), \quad \Lambda_t \equiv \text{diag}(\lambda_{1,t}, \dots, \lambda_{k,t}), \\
 \log(\lambda_{i,t}) &= \log(\lambda_{i,t-1}) + \nu_{i,t}, \quad i = 1, \dots, k, \\
 \nu_t &\equiv (\nu_{1,t}, \nu_{2,t}, \dots, \nu_{k,t})' \sim N(0, \Phi),
 \end{aligned} \tag{1}$$

where  $A$  is a lower triangular matrix with ones on the diagonal and non-zero coefficients below the diagonal, and the diagonal matrix  $\Lambda_t$  contains the time-varying variances of underlying structural shocks. This model implies that the reduced form variance-covariance matrix of innovations to the VAR is  $\text{var}(v_t) \equiv \Sigma_t = A^{-1} \Lambda_t A^{-1'}$ . Note that, as in Primiceri's (2005) implementation, innovations to log volatility are allowed to be correlated across variables; i.e.,  $\Phi$  is not restricted to be diagonal.

To estimate this BVAR, we use a Gibbs sampler, detailed in the appendix. Stochastic volatility is estimated with the algorithm of Kim, Shephard, and Chib (1998), as detailed in Primiceri (2005), and correcting the ordering of sampling steps as proposed by Del Negro and Primiceri (2015). The VAR coefficients are drawn from a conditional posterior distribution that is multivariate normal, with a GLS-based mean and variance given in Clark (2011). All of our reported results are based on samples of 5,000 posterior draws of the model parameters, obtained by retaining every 8th draw

of a total sample of 40,000 post-burn draws, with a burn period of 5,000 draws.<sup>1</sup>

The posterior distributions of forecasts reflect the uncertainty due to all parameters of the model and shocks occurring over the forecast horizon. To simulate the predictive density of the BVAR, from a forecast origin of period  $T$ , for each retained draw of the model parameters or latent states ( $B$ ,  $A$ ,  $\Lambda_t$  up through  $T$ , and  $\Phi$ ), we: (1) draw innovations to log volatilities for periods  $T + 1$  through  $T + H$  from a multivariate normal distribution with variance-covariance matrix  $\Phi$  and use the random walk model of  $\log \lambda_{t+h}$  to compute  $\lambda_{T+1}, \dots, \lambda_{T+H}$ ; (2) draw innovations to  $y_{T+h}$ ,  $h = 1, \dots, H$ , from a normal distribution with variance  $\Sigma_{T+h} = A^{-1}\Lambda_{T+h}A^{-1'}$ , and use the vector autoregressive structure of the model along with the coefficients  $B$  to obtain draws of  $y_{T+h}$ ,  $h = 1, \dots, H$ . We repeat steps (1) and (2) five times for each draw of the model parameters. This yields 25,000 draws of  $y_{T+h}$ , which we use to compute the forecast statistics of interest.

### 3.2 Nowcast model: GDP growth

To align with the typical timing of the Survey of Professional Forecasters, we use the Bayesian mixed frequency modeling approach of Carriero, Clark, and Marcellino (2014) to produce a current-quarter forecast of GDP growth with data available around the end of the first week of the second month of the quarter. More specifically, we forecast the quarterly growth rate of GDP in month two of the current quarter based on the regression:

$$\begin{aligned} y_t &= X_t' \beta + v_t, \\ v_t &= \lambda_t^{0.5} \epsilon_t, \epsilon_t \sim i.i.d.N(0, 1), \\ \log(\lambda_t) &= \log(\lambda_{t-1}) + \nu_t, \nu_t \sim i.i.d.N(0, \phi), \end{aligned} \tag{2}$$

where  $t$  is measured in quarters and the vector  $X_t$  contains predictors available at the time the forecast is formed.

The specification of the regressor vector  $X_t$  is a function of the way the monthly variables are sampled. For the timing we follow in this analysis, the vector  $X_t$  contains variables available at about the end of the first week of month 2 of quarter  $t$ . Specifically, in our implementation, it contains a constant, GDP growth in quarter  $t - 1$ , employment growth in month 1 of quarter  $t$ , and the ISM index in month 1 of quarter  $t$ . We use employment and the ISM because, for our

---

<sup>1</sup>All numerical computations in this paper were done using the R programming language R Core Team (2015), with some of the graphical illustrations based on the ggplot2 package Wickham (2009).

information timing, these are the two major coincident indicators that are available for forecasting GDP growth in the quarter. Our model with this small set of indicators performs comparably to models with the larger sets of indicators considered in Carriero, Clark, and Marcellino (2014).

### **3.3 Nowcast model: inflation in GDP price index**

Our nowcast model for inflation takes the same form as that described above for GDP growth, but with a different set of indicators included in  $X_t$ . While the information set of the typical SPF response has included just week 1 of month 2 of the quarter since the Philadelphia Fed took over the survey, prior to that time the information set (and survey response date) changed over time, and it was often later in the month. Accordingly, for simplicity, we construct nowcasts of GDP inflation using (inflation rates of) monthly price indexes released in the second half of month 2 of the quarter, for the CPI ex food and energy, the CPI for food, the CPI for energy, the PPI for capital goods, and the price index for new home construction. This set of indicators reflects major measures of consumption and investment prices, as typically available in the middle of the quarter.

### **3.4 Nowcast model: unemployment rate**

To align with current SPF timing, we obtain a nowcast of the quarterly average rate of unemployment by averaging the observed rate for month 1 of the quarter with forecasts for months 2 and 3. As noted above, the typical SPF response is based on an information set that includes labor market indicators for the first month of the quarter. We produce the forecasts of months 2 and 3 of the quarter using a BVAR(3) with stochastic volatility in monthly data, for the unemployment rate, growth in payroll employment, and new claims for unemployment insurance. We include unemployment claims in the model because they are commonly thought to be a leading indicator with some predictive content for the unemployment rate (e.g., Montgomery, Zarnowitz, Tsay, and Tiao, 1998), whereas employment is a major coincident indicator of the business cycle that might have predictive content for the unemployment rate, which has sometimes been considered to be a lagging indicator of the business cycle. This model takes the same basic form as the BVAR detailed above, except in monthly rather than quarterly data.

### 3.5 Nowcast model: T-bill rate

To align with SPF timing, we obtain a nowcast of the quarterly average 3-month T-bill rate by averaging the observed rate for month 1 of the quarter with forecasts for months 2 and 3. As SPF timing has shifted over time and respondents have access to a wide range of financial indicators, we incorporate information through the 15th of month 2 of the quarter (in the event the 15th is not a business day, we use the preceding business day). Specifically, to forecast the monthly T-bill rate for months 2 and 3 of the quarter, we use a BVAR(3) with stochastic volatility in which the variable vector  $y_t$  is monthly and contains the average 3-month T-bill rate in  $t$  and the 3-month and 6-month T-bill rates on the 15th of month  $t + 1$ . We include the daily rates in the model as a way of capturing current information that would be available to a forecaster under our timing assumption. We include the 6-month rate because, under the expectations hypothesis, it should contain information about the expected future path of the 3-month rate. This BVAR takes the same form as the one detailed above, except in monthly rather than quarterly data.

## 4 Entropic tilting: methodology and examples

This section first details the general implementation of entropic tilting and then provides examples of our use.

### 4.1 General methodology

In using tilting to incorporate information from survey forecasts or model-based nowcasts into medium-term forecasts from a BVAR with stochastic volatility, our starting point is a “raw” sample of  $I$  (possibly vector-valued) MCMC forecast draws,

$$f := \{\mathbf{y}_i\}_{i=1}^I,$$

where  $\mathbf{y}_i \in \mathbb{R}^p$ ,  $p \geq 1$ . In the following we interpret  $f$  as a discrete distribution with  $I$  possible outcomes, each of which has probability  $1/I$ . For simplicity, at this stage we suppress dependence on a certain variable, forecast origin date and forecast horizon. We consider modifying the distribution  $f$  by imposing the moment condition

$$\mathbb{E}g(\mathbf{y}) = \bar{g},$$

where  $g : \mathbb{R}^p \rightarrow \mathbb{R}^m$  and  $\bar{g} \in \mathbb{R}^m, m \geq 1$ . The following functional optimization problem is often called “entropic tilting”:

$$\min_{\tilde{f} \in \mathbb{F}} \text{KLIC}(\tilde{f}, f) \text{ subject to } \mathbb{E}_{\tilde{f}} g(\mathbf{y}) = \bar{g}. \quad (3)$$

Here  $\mathbb{F}$  denotes the class of all discrete distributions that can be constructed by re-weighting the draws from  $f$  in an admissible way (such that the weights are positive and sum to one). Furthermore,

$$\begin{aligned} \text{KLIC}(\tilde{f}, f) &= \sum_{i=1}^I \tilde{\pi}_i \log(I\tilde{\pi}_i) \\ &= \log(I) + \sum_{i=1}^I \tilde{\pi}_i \log(\tilde{\pi}_i) \end{aligned} \quad (4)$$

is the Kullback-Leibler divergence between the candidate distribution  $\tilde{f}$  (which places weight  $\tilde{\pi}_i$  on the  $i$ th MCMC draw) and  $f$  (which uses flat weights  $1/I$ ). Finally,

$$\mathbb{E}_{\tilde{f}} g(\mathbf{y}) = \sum_{i=1}^I \tilde{\pi}_i g(\mathbf{y}_i)$$

is the expectation of  $\mathbf{y}$  under  $\tilde{f}$ . As discussed by Robertson, Tallman, and Whiteman (2005) and others, the tilting solution is given by setting

$$\pi_i^* = \frac{\exp(\gamma^* g(\mathbf{y}_i))}{\sum_{i=1}^I \exp(\gamma^* g(\mathbf{y}_i))}, \quad (5)$$

$$\gamma^* = \arg \min_{\gamma} \sum_{i=1}^I \exp(\gamma (g(\mathbf{y}_i) - \bar{g})). \quad (6)$$

We should note the following broad implications. First, the solution of the tilting problem comes in the form of a set of weights for the existing sample  $f$ . The squared error and CRPS, which we use to evaluate forecasts (see Section 5), can be computed directly for this new set of weights, without the need for additional simulation. Second, in practice tilting requires solving the minimization problem in (6), whose dimension equals the number of moment conditions (below we consider dimensions of one, two, four, and eight). This is often easy to do, given that the objective function is usually globally convex, and computing the gradient with respect to  $\gamma$  (and passing it to a numerical optimizer) is straightforward. In our implementation, we use the `optim` function of the R programming language (R Core Team, 2015), together with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. In case the algorithm fails to converge, we impose a very

small penalty on the L2 norm of the candidate parameter  $\gamma$  in order to regularize the problem. The corresponding R code is available from the first author. Third, Equation (5) implies that the functional form of the tilting weights is determined by the choice of  $g(\cdot)$ ; we explore this point in our first example below. Finally, it is possible to ensure some smoothness on the tilted forecast distribution by targeting a higher dimensional vector  $\bar{g}$  of moment conditions. We explore this below by experimenting with different sets of moment conditions.

## 4.2 Tilting variants considered in this paper

In the results to be presented below, we will consider the following variants of entropic tilting. First, for a given variable – indicated by the index  $(k)$  – we tilt the BVAR forecast distribution of the vector  $\begin{bmatrix} y_{t+1}^{(k)} & \dots & y_{t+5}^{(k)} \end{bmatrix}$  to match a certain nowcast mean of variable  $k$  (dubbed “small m” below). Second, we tilt the same distribution to match a certain nowcast mean and variance for variable  $k$  (“small m/v”). Third, we consider the joint forecast distribution for the 20-dimensional vector  $\begin{bmatrix} y_{t+1}^{(1)} & \dots & y_{t+5}^{(4)} \end{bmatrix}$  comprising four variables and five forecast horizons. We tilt this distribution to simultaneously match the nowcast means of all four variables (“big m”). Finally, we again consider the full 20-dimensional distribution and tilt it to simultaneously match the nowcast means and variances for all four variables (“big m/v”). To avoid clutter, we henceforth suppress the superindex  $(k)$  whenever we refer to a representative variable.

## 4.3 Example: tilting the mean vs. mean and variance

In this section we illustrate how the forms of entropic tilting we will examine below are implemented and affect forecast distributions. In these examples, the forecast origin date is 2008:Q4, which is interesting because it coincides with the recent recession becoming much more severe, which the SPF nowcasts pick up in real time but the BVAR in quarterly data by itself is slower to detect. The  $p = 5$  variate vector of interest,  $\mathbf{y}_{t:t+4} = [y_t, y_{t+1}, y_{t+2}, y_{t+3}, y_{t+4}]'$ , contains the GDP growth rates from 2008:Q4 to 2009:Q4 (i.e., forecasts for GDP growth zero to four quarters ahead). The two panels of Figure 1 illustrate the following implementations of tilting:

- Targeting the SPF nowcast mean for GDP growth in 2008:Q4 (small m). This corresponds to

setting

$$\begin{aligned} g(\mathbf{y}_{t:t+4}) &= y_t, \\ \bar{g} &= -2.94. \end{aligned}$$

As expected, the figure shows that the tilted distribution is located left of the raw one; this is necessary to implement the SPF nowcast mean which is much smaller than the original one. In this case, the tilted density has a somewhat unconventional shape, featuring substantial probability mass at the lower end of its support.

- Targeting the SPF nowcast mean and variance (small  $m/v$ ) corresponds to

$$\begin{aligned} g(\mathbf{y}_{t:t+4}) &= [y_t, (y_t + 2.94)^2]' \\ \bar{g} &= [-2.94, 2.41]'. \end{aligned}$$

In this case, the tilted density again reaches the SPF nowcast mean, but the distribution is now bell-shaped and tighter than before. This is the result of targeting the nowcast variance in addition to the mean.

In these examples,  $\pi_{t,i}^*$  — the tilting weight on the  $i$ th MCMC draw — is a function of  $y_{t,i}$  (the first element of the vector  $\mathbf{y}_{t:t+4,i}$ ) alone; this follows from the specific choices of  $g(\cdot)$  made here. For each example, the solutions to the tilting weights are given by the following.

- Small  $m$ :

$$\pi_{t,i}^* = \frac{\exp(-0.40 y_{t,i})}{\sum_{i=1}^I \exp(-0.40 y_{t,i})}.$$

- Small  $m/v$ :

$$\pi_{t,i}^* = \frac{\exp(-0.58 y_{t,i} - 0.19 (y_{t,i} + 2.94)^2)}{\sum_{i=1}^I \exp(-0.58 y_{t,i} - 0.19 (y_{t,i} + 2.94)^2)}.$$

Figure 2 plots the relationship for the two tilting variants. As the solutions and charts make clear, the choice of  $g(\cdot)$  significantly affects the re-weighting of the draws in the tilted distribution. While the weight is a monotonic function of  $y_{t,i}$  in the first variant, the relationship is bell-shaped in the second variant.

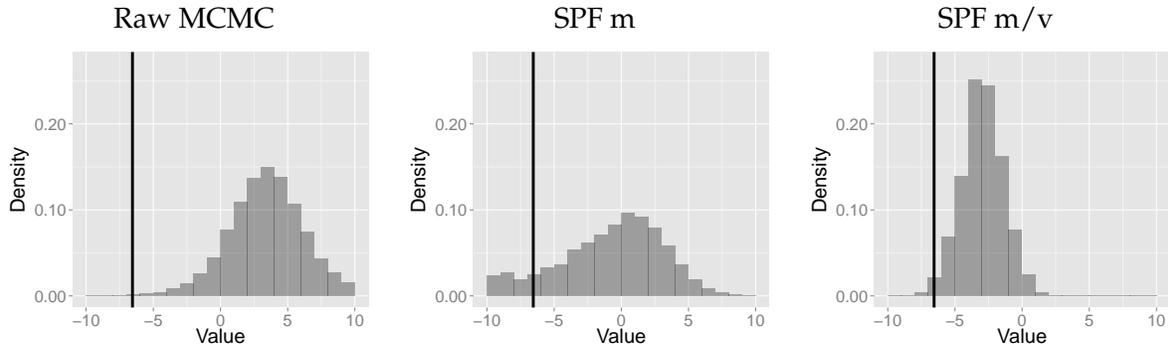


Figure 1: Histograms for raw and tilted samples. In each panel, the black vertical line shows the ex-post outcome of  $-6.55$ .

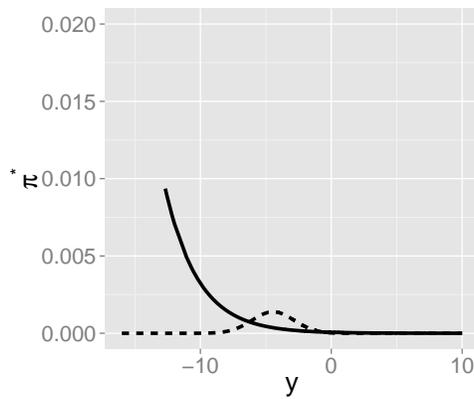


Figure 2: Tilting weight  $\pi_{t,i}^*$  as a function of the first element  $y_t$  of the vector  $y_{t:t+4}$ . The solid line corresponds to tilting toward the nowcast mean only; the dashed line corresponds to tilting toward the nowcast mean and variance.

#### 4.4 Example: spillover effects on longer horizon forecasts

The results just presented demonstrate how tilting toward an external nowcast mean and variance yields a combined nowcast density. However, imposing moment conditions on the nowcast  $y_t$  also affects other elements of the vector of forecasts,  $\mathbf{y}_{t:t+4} = [y_t, y_{t+1}, y_{t+2}, y_{t+3}, y_{t+4}]'$ . While effects on other forecast horizons are difficult to see in the non-parametric solutions provided above in equations (5) and (6), we can use a Gaussian benchmark case (extending the example in Robertson, Tallman, and Whiteman, 2005) to provide some intuition.

Consider a five-variate vector  $\mathbf{y}_{t:t+4}$ , and suppose a forecaster uses a multivariate normal distribution  $f = \mathcal{N}(\theta, \Sigma)$ , where  $\theta = [\theta_1 \dots \theta_5]'$  and  $\Sigma$  is a positive definite matrix with elements  $\Sigma_{i,j}$  (suppressing the dependence of the parameters on time and forecast horizon for simplicity). Consider the tilted density  $f^*$  which imposes that the first system variable have mean  $\mu_1$  and variance  $\Omega_{1,1}$ . Then,  $f^*$  is multivariate normal  $\mathcal{N}(\mu, \Omega)$ , with parameters

$$\mu_{2:5} = \theta_{2:5} + \Sigma_{1,1}^{-1} \Sigma_{1,2:5} (\mu_1 - \theta_1), \quad (7)$$

$$\Omega_{2:5,2:5} = \Sigma_{2:5,2:5} - \Sigma_{2:5,1} \Sigma_{1,1}^{-1} \Sigma_{1,2:5} \times \left(1 - \frac{\Omega_{1,1}}{\Sigma_{1,1}}\right), \quad (8)$$

$$\Omega_{2:5,1} = \Sigma_{2:5,1} \Sigma_{1,1}^{-1} \Omega_{1,1}, \quad (9)$$

where  $A_{i:j, k:l}$  denotes the matrix consisting of rows  $i : j$ , columns  $k : l$  of any matrix  $A$ . We write  $A_{i:j, k}$  if the “matrix” is a column vector, and  $A_{i, k:l}$  if it is a row vector. This Gaussian example yields the following implications.

- In the special case that  $y_t$  is fixed, such that  $\Omega_{1,1} = 0$ , we end up at the textbook formulas for conditioning in the multivariate normal distribution. That is, entropic tilting is exactly the same as conditional forecasting. It is also exactly the same as treating the nowcast as data or jumping-off points for forecasts at subsequent horizons (Faust and Wright, 2009, 2013). See Section 6.1 for further discussion of this equivalence.
- The special case that  $\Omega_{1,1} = \Sigma_{1,1}$  corresponds to a scenario in which the tilted variance for  $y_{t+1}$  is the same as the un-tilted variance. Interestingly, the same solution obtains when targeting a mean of  $\theta_1$  only, without making a tilting assumption about  $\Omega_{1,1}$  – see e.g. Altavilla, Giacomini, and Ragusa (2013, Section 3.1). This implies that, at least in the Gaussian case, targeting the mean only is equivalent to targeting the mean *and* the original variance.

- If  $\Omega_{1,1} < \Sigma_{1,1}$  and  $\Sigma_{2:5,1} \neq 0$  in (8), tilting “reduces” (in a matrix sense) the variance of the forecasts at other horizons.
- The magnitude of the impact on  $\theta_{2:5}$  and  $\Omega_{2:5,2:5}$  mainly depends on  $\Sigma_{2:5,1}$ , the correlation of the nowcast with the other horizons. This correlation matrix reflects the persistence of the time series, as predicted by the BVAR.

The results we obtain for our examples using the non-parametric entropic solution are broadly consistent with the implications of the Gaussian specification. In the interest of brevity, we suppress the details and provide a simple example here. As the forecast horizon increases from the current quarter through the following four quarters, tilting has more persistent effects on the forecasts of the unemployment and T-bill rates (the most persistent variables) than the forecasts of GDP growth and inflation. That is, tilting based on the nowcast for 2008:Q4 has larger effects on the 2009:Q4 forecasts for the unemployment and T-bill rates than on the 2009:Q4 forecasts for growth and inflation. Figure 3 illustrates these points for GDP and unemployment forecasts. The figure uses boxplots to visualize the raw and tilted distributions, for the nowcast (2008:Q4) and the one-year-ahead forecast (2009:Q4). For GDP (left panel), tilting leads to a strong downward revision of the nowcast. At the same time, the one-year-ahead forecast distribution is not strongly revised. For unemployment (right panel), the mean of the one-year-ahead forecast is revised almost as strongly as that of the nowcast itself, with the revision pointing to higher unemployment rates in both cases. The stronger spillover effects for unemployment (compared to GDP) are due to larger entries in  $\Sigma_{2:5,1}$ , which represents the persistence of the series (as predicted by the BVAR). For example, the predicted first-order autocorrelation is only 0.17 in the case of GDP, but 0.76 in the case of unemployment.

#### 4.5 Example: joint tilting across variables and forecast horizons

In the examples above, we performed tilting on a variable-by-variable basis. However, the methodology allows us to directly consider the full multivariate forecast distribution comprising all variables and forecast horizons. Forecasts tilted based on the joint set of nowcasts should be conceptually preferable. In practice one is likely to have available and to be interested in using nowcasts for all model variables. Tilting based on the set of nowcasts together yields a single set of forecasts of all variables that reflects the BVAR-captured historical relationships among the variables. To see the logic of it, consider the very simple (and parametric) approach of incorporating nowcasts

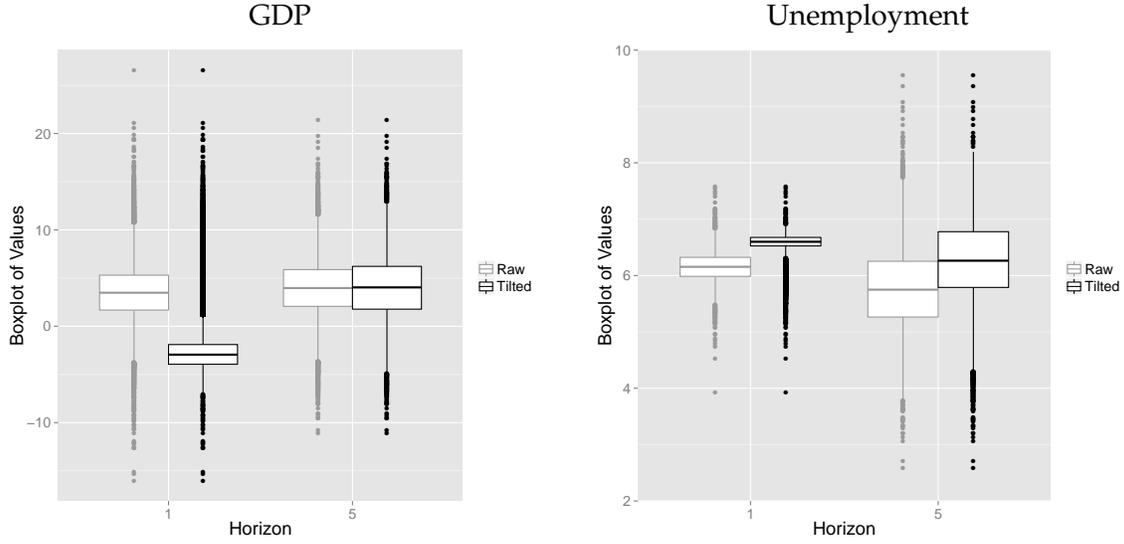


Figure 3: Vertical axis: boxplots of raw and tilted forecast distributions at origin date 2008:Q4. Horizontal axis: forecast horizon  $h$ . Boxes range from the 25 % to the 75 % quantile of a forecast distribution; the end of the upper vertical line is the 75% quantile plus 1.5 times the interquartile range. Forecast draws exceeding that value are plotted as points.

through Gaussian conditional forecasting. One could first condition on the nowcast for variable 1 and produce BVAR forecasts for all variables, then condition on the nowcast for variable 2 and produce BVAR forecasts for all variables, etc. This would of course produce an entire set of alternative forecasts for each variable, reflecting conditions imposed one at a time. In practice, it is more likely the case that the entire set of nowcast conditions would be imposed at once, to obtain a single set of forecasts that reflects the joint set of conditions. The reasoning is the same for a joint approach to entropic tilting.

Interestingly, big tilting turns out to be a more stringent version of the four small problems (one variable at a time). To see this, denote by  $f$  the full (20 dimensional) empirical MCMC distribution for all variables and horizons, by  $f^{(k)}$  the distribution for variable  $k$  (five dimensions = forecast horizons), and by  $\mathcal{C}^{(k)}$  the set of moment conditions imposed on variable  $k$ . Then, big tilting solves

$$\min_{\tilde{f}} \text{KLIC}(\tilde{f}, f) \quad \text{subject to} \quad \mathcal{C}^{(1)} \cup \dots \cup \mathcal{C}^{(4)}. \quad (10)$$

Small tilting for variable  $k$  solves

$$\min_{\tilde{f}^{(k)}} \text{KLIC}(\tilde{f}^{(k)}, f^{(k)}) \quad \text{subject to} \quad \mathcal{C}^{(k)}. \quad (11)$$

Notice that the candidate distributions  $\tilde{f}$  from (10) and  $\tilde{f}^{(k)}$  from (11) are both characterized by

a weight vector of dimension 25,000 (the number of MCMC draws), and the raw distributions  $f$  and  $f^{(k)}$  both feature flat weights. Hence, for a given weight vector, we have that  $\text{KLIC}(\tilde{f}, f) = \text{KLIC}(\tilde{f}^{(k)}, f^{(k)})$ , and thus (11) is equivalent to solving

$$\min_{\tilde{f}} \text{KLIC}(\tilde{f}, f) \quad \text{subject to} \quad \mathcal{C}^{(k)}. \quad (12)$$

Hence the minimization problem (12) for small tilting is a variant of the problem (10) for big tilting, featuring a less stringent set of constraints. This implies that big tilting will typically entail a more drastic move away from the baseline distribution compared to small tilting. In order to illustrate this point, we again consider the 2008:Q4 example, and tilting based on the SPF nowcast means and variances (big m/v, small m/v). Table 1 illustrates the logical necessity that the four small tilting approaches are KLIC-closer to raw MCMC than the big tilting approach (although, in this steep recession example, all approaches are fairly far away from the equal weights of raw MCMC, because the nowcast of growth is so different from the BVAR forecast). Similarly, Figure 4 presents Lorenz curves for the observation weights resulting from big versus small tilting. The figure shows that the weights for big tilting are highly unequal, with a fairly small number of influential MCMC draws (for example, the fifty largest weights sum up to 0.35). The weights for the small tilting problems are much more equal, i.e. the Lorenz curves in Figure 4 are left of the one for big tilting. Note that the small tilting method for inflation generates by far the most uniform weights (leftmost Lorenz curve), which is in line with the fact that in 2008:Q4 the current quarter MCMC forecast (2.72) is already close to the SPF mean nowcast (2.6). The supplementary appendix contains further analysis of the tilting weights, illustrating broader patterns of the weights over time. It also demonstrates that, given the sample sizes common in MCMC, numerical issues caused by unequal weights seem to have very little practical impact on forecasting performance.

Method	Variable	KLIC divergence from equal weights
big m/v	(all)	24993.831
	GDP	24992.512
m/v	UNE	24991.846
	INF	24990.030
	TBI	24991.320

Table 1: KLIC divergences for big versus small tilting in the 2008Q4 example. See Equation (4) for the underlying formula.

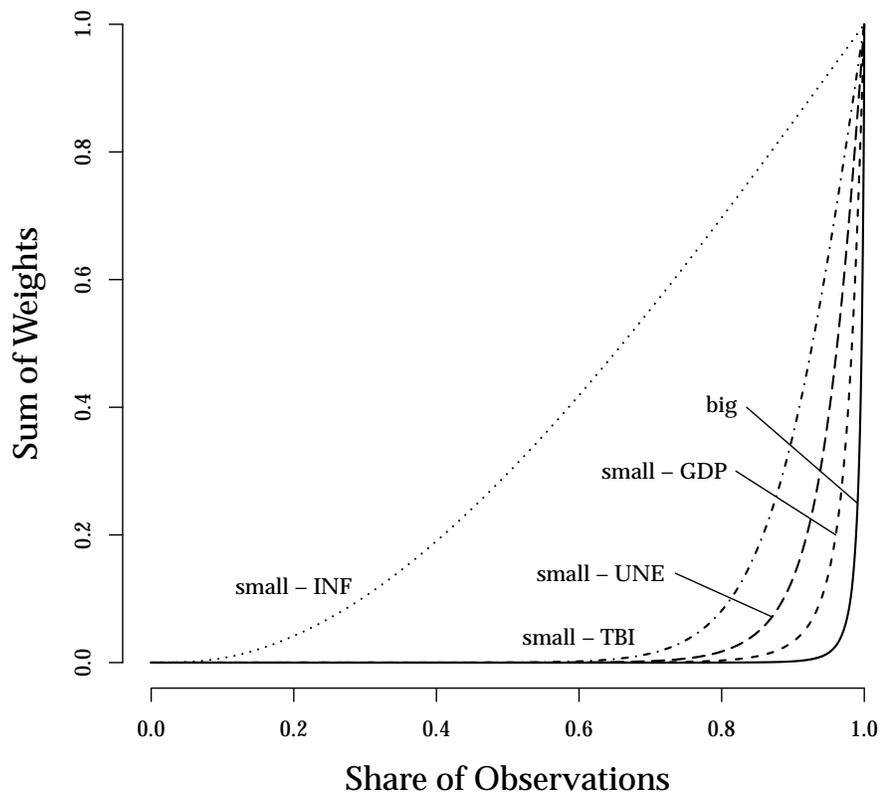


Figure 4: Lorenz curve of weights in the 2008Q4 example. *Reading example:* The topmost line indicates that for inflation, the 25 % smallest weights (horizontal axis) add up to roughly 0.07 (vertical axis).

## 5 Forecast results

We first consider the accuracy of point forecasts (defined as posterior means), using root mean square errors (RMSEs). We then consider density forecasts, using the average continuous ranked probability score (CRPS). Studies such as Gneiting and Raftery (2007) and Gneiting and Ranjan (2011) discuss the advantages of the CRPS over other measures. The CRPS, defined such that a lower number is a better score, is given by

$$\text{CRPS}_t(y_{t+h}^o) = \int_{-\infty}^{\infty} (F(z) - 1_{\{y_{t+h}^o \leq z\}})^2 dz, \quad (13)$$

where  $y_{t+h}^o$  denotes the observed outcome,  $F$  denotes the cumulative distribution function associated with the (posterior) predictive density  $f$ , and  $1_{\{y_{t+h}^o \leq z\}}$  denotes an indicator function taking value 1 if  $y_{t+h}^o \leq z$  and 0 otherwise. In our analysis,  $F$  takes the form of an empirical distribution function, whereby the observation weights are equal in the case of raw MCMC but not in the case of tilted distributions. We employ the algorithm by Hersbach (2000, Section 4) – which allows for non-equal weights – to compute the CRPS in both cases.

In order to test the statistical significance of differences in predictive performance, we consider pairwise tests of equal predictive accuracy (henceforth, EPA; Diebold and Mariano, 1995; West, 1996) in terms of either RMSE or CRPS. All EPA tests we conduct compare the raw BVAR forecasts against a given variant of entropic tilting, using two sided tests and standard normal critical values. Based on simulation evidence in Clark and McCracken (2013), in computing the variance estimator which enters the test statistic, we employ a rectangular kernel truncated at lag  $h - 1$  and incorporate the finite sample correction due to Harvey, Leybourne, and Newbold (1997). In the rare cases in which the rectangular kernel yields a negative variance estimate, we resort to Bartlett kernel weights (Newey and West, 1987) in order to ensure positivity. In these cases, we use the automatic bandwidth selection procedure of Newey and West (1994) as implemented in R’s `sandwich` package (Zeileis, 2004).

Our use of EPA tests based on normal critical values may be viewed as an approximation that simplifies an inference problem that, in our context, features many complexities — possible nesting of forecasts and tilting that bears similarities to conditional forecasting — not necessarily easily dealt with in the forecast evaluation literature (see e.g. Clark and McCracken, 2013, 2014).<sup>2</sup>

---

<sup>2</sup>At the one-step horizon, the tilted forecasts are, by construction, essentially the nowcasts, so the benchmark BVAR

Under the asymptotics of Giacomini and White (2006), a test of a null of equal forecast accuracy in the finite sample (at estimated model parameters) is generally normally distributed, subject to a requirement that the model parameters be estimated with a rolling sample of data. While we have not estimated the BVAR with a rolling sample of data, Monte Carlo evidence in Clark and McCracken (2013) indicates that, with nested models estimated with an expanding data sample (the approach we have used with our BVAR), EPA tests compared against normal critical values can be viewed as a somewhat conservative (modestly under-rejecting compared to nominal size) test for equal accuracy in the finite sample.

As noted in Section 4, we consider the following variants of entropic tilting: small  $m$ , in which we tilt the BVAR forecast distribution of the vector  $\begin{bmatrix} y_{t+1}^{(k)} & \dots & y_{t+5}^{(k)} \end{bmatrix}$  to match a certain nowcast mean of variable  $k$ ; small  $m/v$ , in which we tilt the forecast distribution of variable  $k$  to match the nowcast mean and variance; big  $m$ , in which we tilt the entire 20 element vector of variables and horizons to simultaneously match the nowcast means of all four variables; and big  $m/v$ , in which we tilt the entire 20 element vector of variables and horizons to simultaneously match the nowcast means and variances of all four variables. We separately apply each of the tilting variants to nowcasts from the SPF and from the nowcasting models described in Section 3. In the case of the SPF-based results, the variances used in tilting are computed as described in Section 2.2; for model-based nowcasts, the variances used in tilting are defined as the variance of the posterior distribution of BVAR forecasts for the period in question.

## 5.1 Comparison of current-quarter forecasts

Before examining the effects of entropic tilting of BVAR forecasts toward different nowcasts, it is useful to compare the accuracy of current quarter forecasts from the BVAR, the SPF, and the nowcasting models. Table 2 provides the RMSEs and CRPS scores of each current quarter forecast (except that we don't provide CRPS scores for the SPF forecasts because the SPF does not include the forecast density information needed to compute the CRPS over our sample). These results yield the following findings.

---

forecast and each tilted forecast are not nested, in which case the application of the EPA test is valid. At longer horizons, the picture is less clear; the tilted forecasts are functions of the nowcasts and the underlying BVAR forecasts. Under some conditions, at horizons of 2 or more periods, the tilted and BVAR forecasts could be seen as nested under a null of equal accuracy. Regardless, the multi-step tilted forecasts bear similarities to conditional forecasts; Clark and McCracken (2014) propose a modified test of EPA necessary for application to conditional forecasts.

- Consistent with previous studies, current-quarter forecasts from the SPF and the models designed for nowcasting are generally more accurate than the current quarter forecasts from the BVAR. For example, in the case of GDP growth over the pre-crisis sample, the SPF and mixed frequency nowcasting models have RMSEs of 1.580 and 1.682, respectively, compared to the BVAR’s RMSE of 1.975. The differences (for GDP growth and unemployment) are even larger in the full sample than in the pre-crisis sample.
- Compared to SPF, some of the nowcasting models yield better accuracy, while others yield less accuracy. For GDP growth, the mixed frequency nowcasting model is almost as accurate as SPF in the pre-crisis sample and modestly less accurate in the full sample, reflecting the better job the SPF did in picking up the sharp downturn of the Great Recession (see the discussion in Carriero, Clark, and Marcellino, 2014). For unemployment and the T-bill rate, the model-based nowcasts are at least somewhat more accurate than the SPF forecasts. These gains are likely due to the use of intra-quarter information about the predictand (c.f. Montgomery, Zarnowitz, Tsay, and Tiao, 1998, for similar results on unemployment). For instance, in the pre-crisis sample, the model-based nowcast of the T-bill rate has an RMSE of 0.066, compared to the SPF RMSE of 0.133.
- The CRPS scores move closely in line with the RMSEs, both qualitatively and in terms of the magnitude of improvements of nowcasts over current-quarter forecasts from the BVAR.

		<i>Pre Crisis (88:Q3 – 07:Q4)</i>				<i>Complete (88:Q3 – 13:Q2)</i>			
		GDP	UNE	INF	TBI	GDP	UNE	INF	TBI
<i>RMSE</i>	<i>SPF</i>	1.580	0.125	0.767	0.133	1.591	0.151	0.809	0.133
	<i>BMF</i>	1.682	0.095	0.861	0.066	1.899	0.095	0.985	0.072
	<i>BVAR</i>	1.975	0.157	0.876	0.406	2.390	0.235	0.938	0.406
<i>CRPS</i>	<i>BMF</i>	0.960	0.053	0.495	0.037	1.048	0.053	0.554	0.035
	<i>BVAR</i>	1.123	0.089	0.504	0.214	1.274	0.116	0.539	0.214

Table 2: Root Mean Squared Error and CRPS for different nowcasts (SPF = Survey of Professional Forecasters, BMF = Bayesian Mixed Frequency, BVAR = Bayesian VAR with stochastic volatility). SPF and BMF use data up to daily frequency; BVAR is based on quarterly data.

## 5.2 Main results

We now consider tilting longer-horizon forecasts based on just current-quarter forecasts. Table 3 presents the full-sample results (see the supplementary appendix for results from the pre-crisis

period). In light of the common central bank practice of reporting growth and inflation rates that are averages over four quarters, the table provides results for (annualized) quarterly forecasts four and five quarters ahead and for four-quarter averages four and five quarters ahead (in the columns “4\*” and “5\*”, respectively). These results yield the following key take-aways.

- In all cases, tilting forecasts based on just the nowcast (point or point and variance) from either the SPF or the nowcasting models improves the accuracy of point and density forecasts at horizons of one, two and three quarters. For example, in the results for GDP growth at the three quarters-ahead horizon, under the small  $m$  approach, tilting toward the nowcast from the mixed frequency model lowers the RMSE of the BVAR forecast from 2.656 to 2.572; the difference is significant at the 5% level (two-sided test). For the same sample and horizon, tilting the T-bill forecasts toward the model-based nowcasts (small  $m$  approach) lowers the RMSE of the BVAR from 1.056 to 0.821 (difference significant at 1% level). Tilting has quantitatively similar effects on density forecast accuracy as measured by the CRPS.
- At forecast horizons of four and five quarters, the performance of forecasts tilted toward nowcasts is more mixed. At these horizons, tilting has relatively little benefit for forecasts of GDP growth and inflation. But it has some benefit for forecasts of the more persistent variables, the unemployment and T-bill rates. As an example, at the five step horizon, tilting the T-bill forecasts toward the model-based nowcasts lowers the RMSE of the BVAR from 1.591 to 1.425 (difference significant at 1% level). Again, tilting has quantitatively similar effects on density forecast accuracy as measured by the CRPS. These patterns align with the observations drawn in the illustration of Section 4.4.
- Tilting the BVAR forecasts toward both the mean and variance of nowcasts ( $m/v$ ) — rather than just the mean or point nowcast ( $m$ )— yields small additional gains in density forecast accuracy. This pattern is very robust: In 23 out of 24 scenarios (variables and forecast horizons) covered by Table 3, the CRPS score of the best  $m/v$  specification is smaller than that of the best  $m$  specification. For example, in the case of the unemployment rate and  $h = 2$ , the best specification based on the mean only (BMF small  $m$ ) attains a CRPS of 0.146, whereas the best mean/variance specification (BMF big  $m/v$ ) attains a CRPS of 0.131. By comparison, the CRPS of the raw BVAR distribution is 0.217.
- Jointly considering the nowcasts of all four variables (“big  $m/v$ ”) versus considering all variables separately (“small  $m/v$ ”) tends to perform similarly well, with each approach outper-

forming the other in a number of scenarios. While one interpretation might be that joint treatment offers little overall advantage, an alternative interpretation might be that it is conceptually preferable for imposing tilting at a system level and does so at little (if any) cost, in terms of forecast accuracy.

### 5.3 Entropic tilting and nowcast uncertainty

Table 3 implies that tilting toward the nowcast mean *and variance* consistently yields better CRPS scores than tilting toward the mean only. For  $h = 1$ , this effect is simply a consequence of the nowcast distributions being more accurate than the BVAR ones, which is well known in the literature. Much more interestingly, the result also holds for  $h \geq 2$ , which suggests that the m/v approach produces more favorable spillover effects on the horizons that are not directly affected by tilting. Table 4 investigates this result in more detail, by reporting the length and coverage of central prediction intervals obtained from both approaches (nominal level of 70%). In particular, we define length as the spread between the 15th and 85th percentiles of the forecast distribution and report the average length over time, and we measure coverage as the percent of actual outcomes of each variable falling within the 70% confidence band.

For all variables and forecast horizons, we observe that the m/v specifications produce shorter prediction intervals than the m specifications, which implies sharper (i.e., more concentrated) forecast distributions. This result is natural: The SPF and model nowcasts generally have lower variance than the current quarter forecasts produced by the BVAR. While m/v imposes this information, m fails to do so. Instead, it penalizes the KLIC divergence from the BVAR distribution, and thus implicitly targets the BVAR variance (c.f. Section 4.4). These effects are clearest for the T-bill and unemployment rates, where the m/v approaches produce prediction intervals whose average lengths (over time) are often about 20-40 percent shorter than those of the m approaches. The differences are much smaller for GDP growth and inflation, where the average lengths of the prediction intervals typically differ by less than five percent.

Naturally, the reduced length of the m/v prediction intervals comes along with reduced coverage rates compared to m. For GDP, unemployment and inflation, the coverage rates of m/v are mostly still above 60 percent (recall that the nominal level is 70 percent). A similar statement holds for the T-bill rate and  $h \in \{2, 3\}$ . For the T-bill rate and  $h \in \{4, 5\}$ , the coverage rates of the m

Forecast Horizon		1	2	3	4	5	4*	5*
		GDP						
RMSE	raw	2.390	2.589	2.656	2.640	<b>2.617</b>	2.084	2.140
	SPF small m	<b>1.591**</b>	2.413*	2.540**	2.650	2.651	1.704*	2.076
	SPF small m/v	1.591**	<b>2.390*</b>	<b>2.511*</b>	2.637	2.631	<b>1.688*</b>	<b>2.052</b>
	SPF big m	1.591**	2.439	2.562	2.665	2.628	1.716*	2.093
	SPF big m/v	1.596**	2.435*	2.524*	<b>2.632</b>	2.631	1.700*	2.060
	BMF small m	1.899**	2.475**	2.572*	2.654	2.636	1.835*	2.096
	BMF small m/v	1.899**	2.464**	2.563*	2.643	2.627	1.829*	2.085
	BMF big m	1.900**	2.649	2.736	2.753	2.655	1.964	2.215
	BMF big m/v	1.917**	2.478*	2.611	2.699	2.670	1.873*	2.139
CRPS	raw	1.274	1.380	1.414	<b>1.408</b>	1.403	1.122	1.158
	SPF small m	0.985**	1.304*	1.358**	1.414	1.418	0.936**	1.122
	SPF small m/v	<b>0.906**</b>	<b>1.280**</b>	<b>1.335**</b>	1.411	1.411	<b>0.912**</b>	<b>1.113</b>
	SPF big m	0.993**	1.326	1.379	1.422	<b>1.398</b>	0.944*	1.127
	SPF big m/v	0.909**	1.300**	1.350*	1.411	1.403	0.920**	1.123
	BMF small m	1.074**	1.330**	1.376*	1.414	1.413	0.991**	1.133
	BMF small m/v	1.042**	1.317**	1.366**	1.412	1.410	0.977*	1.129
	BMF big m	1.077**	1.363	1.440	1.456*	1.432	1.043	1.190
	BMF big m/v	1.049**	1.320**	1.385	1.457	1.427	1.000*	1.159
		UNE						
RMSE	raw	0.235	0.464	0.706	0.940	1.147		
	SPF small m	0.151*	0.348	0.575	0.811	1.028		
	SPF small m/v	0.151*	0.350	0.573	0.808	1.024		
	SPF big m	0.151*	0.332	0.542	0.769	0.987		
	SPF big m/v	0.153*	0.335	0.551	0.778	0.995		
	BMF small m	0.095**	<b>0.239</b>	<b>0.470</b>	0.716	0.958		
	BMF small m/v	0.096**	0.251	0.470	<b>0.708</b>	<b>0.945</b>		
	BMF big m	0.095**	0.247	0.476	0.725	0.966		
	BMF big m/v	<b>0.092**</b>	0.256	0.483	0.727	0.964		
CRPS	raw	0.116	0.217	0.333	0.453	0.569		
	SPF small m	0.087**	0.179	0.280	0.396	0.510		
	SPF small m/v	0.083**	0.175*	0.276*	0.391	0.505		
	SPF big m	0.088**	0.173*	0.267*	0.377*	0.491		
	SPF big m/v	0.082**	0.167*	0.262*	0.372*	0.487		
	BMF small m	0.071**	0.146*	0.242	0.357	0.477		
	BMF small m/v	0.054**	0.135*	0.233*	<b>0.346</b>	<b>0.462</b>		
	BMF big m	0.072**	0.147*	0.241*	0.355*	0.474*		
	BMF big m/v	<b>0.053**</b>	<b>0.131**</b>	<b>0.231**</b>	0.346*	0.466*		

Table 3: **Empirical results for entropic tilting, complete sample (1988:Q3 – 2013:Q2)**. “RMSE” rows contain root mean squared errors. “CRPS” rows contain mean cumulative ranked probability scores. raw – MCMC output of BVAR-SV model. Alternative tilting targets: SPF small m – SPF mean nowcast for the same variable. SPF small m/v – SPF nowcast mean and variance for the same variable. SPF big m – SPF nowcast means for all four variables. SPF big m/v – SPF nowcast means and variances for all four variables. BMF small m, BMF small m/v, BMF big m and BMF big m/v are defined analogously. One and two stars indicate rejections of equal predictive ability at the five and one percent level (two sided tests; implementation details described in the beginning of Section 5).

<i>Forecast Horizon</i>		1	2	3	4	5	4*	5*
			INF					
RMSE	raw	0.938	0.996	1.012	1.051	1.170	0.717	0.795
	SPF small m	0.809**	0.917**	0.974	1.030	1.144	0.636*	0.746
	SPF small m/v	0.809**	0.915**	0.973	1.029	1.144	0.634*	0.746
	SPF big m	0.809**	0.908**	0.972	1.018	1.135	0.622**	0.735
	SPF big m/v	<b>0.807**</b>	<b>0.906**</b>	<b>0.966</b>	<b>1.003</b>	<b>1.128</b>	<b>0.617**</b>	<b>0.731</b>
	BMF small m	0.985	1.026	1.049	1.075	1.194	0.765	0.828
	BMF small m/v	0.985	1.027	1.050	1.075	1.196	0.766	0.829
	BMF big m	0.985	1.010	1.050	1.075	1.204	0.753	0.823
	BMF big m/v	0.986	1.002	1.051	1.054	1.197	0.751	0.817
CRPS	raw	0.539	0.578	0.575	0.610	0.687	0.413	0.463
	SPF small m	0.471**	0.540**	0.563	0.602	0.676	0.378*	0.444
	SPF small m/v	0.463**	0.534**	0.558	0.597	0.672	0.369*	0.438
	SPF big m	0.473**	0.544*	0.565	0.600	0.677	0.380*	0.445
	SPF big m/v	<b>0.462**</b>	<b>0.532**</b>	<b>0.555</b>	<b>0.582</b>	<b>0.663</b>	<b>0.361**</b>	<b>0.431</b>
	BMF small m	0.544	0.588	0.589	0.619	0.697	0.428	0.476
	BMF small m/v	0.553	0.591	0.590	0.618	0.697	0.430	0.476
	BMF big m	0.542	0.589	0.598*	0.625	0.701	0.431	0.479
	BMF big m/v	0.554	0.586	0.589	0.606	0.694	0.424	0.469
			TBI					
RMSE	raw	0.406	0.756	1.056	1.336	1.591		
	SPF small m	0.133**	0.529**	0.871**	1.183**	1.469**		
	SPF small m/v	0.132**	0.514**	0.854**	1.162**	1.450**		
	SPF big m	0.133**	0.522**	0.856**	1.152**	1.423**		
	SPF big m/v	0.134**	0.511**	0.847**	1.142**	1.418**		
	BMF small m	0.072**	0.473**	0.821**	1.137**	1.425**		
	BMF small m/v	0.073**	0.457**	0.800**	1.112**	1.402**		
	BMF big m	0.083**	0.465**	<b>0.791**</b>	<b>1.092**</b>	<b>1.372**</b>		
	BMF big m/v	<b>0.070**</b>	<b>0.447**</b>	0.815**	1.104**	1.387**		
CRPS	raw	0.214	0.419	0.611	0.794	0.970		
	SPF small m	0.159**	0.343**	0.530**	0.715**	0.898**		
	SPF small m/v	0.070**	0.272**	0.479**	0.680**	0.872**		
	SPF big m	0.162**	0.340**	0.522**	0.699**	0.872**		
	SPF big m/v	0.072**	0.271**	0.475**	0.670**	0.853**		
	BMF small m	0.152**	0.328**	0.511**	0.694**	0.876**		
	BMF small m/v	<b>0.041**</b>	0.239**	<b>0.445**</b>	0.645**	0.838**		
	BMF big m	0.165**	0.331**	0.499**	0.672**	0.845**		
	BMF big m/v	0.047**	<b>0.238**</b>	0.447**	<b>0.636**</b>	<b>0.822**</b>		

Table 3: continued.

approaches are already well below 70 percent, with the rates of  $m/v$  being even lower.

On balance, the increased sharpness of  $m/v$  appears to come at a small cost, in that the coverage rates are similarly close to (or far from) their nominal level as under the  $m$  variant. This assessment is consistent with the fact that the CRPS – which can be seen as a trade-off between sharpness and correct coverage, see e.g. Gneiting, Balabdaoui, and Raftery (2007) – consistently favors  $m/v$  over  $m$ .

Forecast Horizon		1	2	3	4	5	4*	5*
		GDP						
Coverage	SPF m	0.870	0.700	0.720	0.700	0.670	0.640	0.560
	SPF m/v	0.710	0.670	0.690	0.690	0.670	0.600	0.560
	BMF m	0.790	0.710	0.690	0.690	0.660	0.620	0.560
	BMF m/v	0.710	0.680	0.700	0.690	0.660	0.610	0.560
Length	SPF m	5.217	4.872	4.922	4.957	4.964	2.984	2.960
	SPF m/v	3.326	4.638	4.730	4.795	4.830	2.628	2.851
	BMF m	4.921	4.794	4.838	4.881	4.908	2.903	2.914
	BMF m/v	3.973	4.656	4.731	4.787	4.819	2.715	2.852
CRPS	SPF m	0.985	1.304	1.358	1.414	1.418	0.936	1.122
	SPF m/v	<b>0.906</b>	<b>1.280</b>	<b>1.335</b>	<b>1.411</b>	1.411	<b>0.912</b>	<b>1.113</b>
	BMF m	1.074	1.330	1.376	1.414	1.413	0.991	1.133
	BMF m/v	1.042	1.317	1.366	1.412	<b>1.410</b>	0.977	1.129
		UNE						
Coverage	SPF m	0.860	0.820	0.750	0.700	0.650		
	SPF m/v	0.760	0.730	0.730	0.650	0.590		
	BMF m	0.990	0.910	0.840	0.770	0.710		
	BMF m/v	0.830	0.770	0.730	0.680	0.620		
Length	SPF m	0.448	0.712	0.932	1.118	1.248		
	SPF m/v	0.292	0.562	0.799	0.999	1.152		
	BMF m	0.512	0.804	1.026	1.214	1.324		
	BMF m/v	0.245	0.544	0.794	1.005	1.176		
CRPS	SPF m	0.087	0.179	0.280	0.396	0.510		
	SPF m/v	0.083	0.175	0.276	0.391	0.505		
	BMF m	0.071	0.146	0.242	0.357	0.477		
	BMF m/v	<b>0.054</b>	<b>0.135</b>	<b>0.233</b>	<b>0.346</b>	<b>0.462</b>		
		INF						
Coverage	SPF m	0.790	0.740	0.760	0.720	0.700	0.830	0.790
	SPF m/v	0.690	0.720	0.750	0.710	0.700	0.830	0.780
	BMF m	0.660	0.700	0.710	0.750	0.710	0.770	0.770
	BMF m/v	0.620	0.700	0.710	0.730	0.690	0.740	0.760
Length	SPF m	2.036	2.238	2.418	2.595	2.777	1.822	2.052
	SPF m/v	1.665	2.121	2.318	2.499	2.680	1.660	1.944
	BMF m	2.008	2.222	2.411	2.581	2.767	1.810	2.045
	BMF m/v	1.823	2.160	2.357	2.536	2.719	1.728	1.990
CRPS	SPF m	0.471	0.540	0.563	0.602	0.676	0.378	0.444
	SPF m/v	<b>0.463</b>	<b>0.534</b>	<b>0.558</b>	<b>0.597</b>	<b>0.672</b>	<b>0.369</b>	<b>0.438</b>
	BMF m	0.544	0.588	0.589	0.619	0.697	0.428	0.476
	BMF m/v	0.553	0.591	0.590	0.618	0.697	0.430	0.476
		TBI						
Coverage	SPF m	0.970	0.810	0.680	0.600	0.470		
	SPF m/v	0.820	0.750	0.620	0.490	0.430		
	BMF m	0.990	0.820	0.700	0.620	0.530		
	BMF m/v	0.940	0.770	0.650	0.510	0.450		
Length	SPF m	0.987	1.403	1.716	1.996	2.241		
	SPF m/v	0.301	0.888	1.290	1.615	1.892		
	BMF m	1.024	1.434	1.741	2.022	2.260		
	BMF m/v	0.242	0.859	1.269	1.602	1.884		
CRPS	SPF m	0.159	0.343	0.530	0.715	0.898		
	SPF m/v	0.070	0.272	0.479	0.680	0.872		
	BMF m	0.152	0.328	0.511	0.694	0.876		
	BMF m/v	<b>0.041</b>	<b>0.239</b>	<b>0.445</b>	<b>0.645</b>	<b>0.838</b>		

Table 4: Impact of accounting for nowcast uncertainty, complete sample (1988:Q3 – 2013:Q2). Coverage and Length refer to central prediction intervals with a nominal level of 70 % (reported length is on average over time). CRPS reports the continuous ranked probability score (best = lowest number printed in bold). Note that the CRPS scores are identical to those in Table 3, and are reprinted here for ease of reference.

## 6 Comparisons to other combination methods

We next compare entropic tilting to two related methods that can be used for combining BVAR and external nowcasts. We first describe these methods and then present the results of the comparison.

### 6.1 Jumping-off approach (Faust and Wright, 2009, 2013)

The “jumping-off” method of Faust and Wright (2009, 2013) appends the nowcast to the actual data, thus treating it as known. Under Gaussianity, this approach is equivalent to conditional forecasting discussed in Section 4.4. To see this, suppose that  $y_t$  follows an AR(1) process, i.e.  $y_t = \phi y_{t-1} + \epsilon_t, \epsilon_t \sim N(0, \sigma^2)$ . The usual  $h$  step ahead forecast distribution for  $y_{t+h}$  is Gaussian with mean  $\phi^h y_t$  and variance  $\sigma^2 \sum_{j=0}^{h-1} \phi^{2j}$ . Under the jumping-off approach, the nowcast  $\mu_1$  is treated as data for period  $t + 1$ , and we form the forecast for period  $t + h$  as an  $(h - 1)$ -step ahead forecast using the pseudo-data for  $t + 1$ . Thus, the forecast distribution is Gaussian with mean  $\phi^{h-1} \mu_1$  and variance  $\sigma^2 \sum_{j=0}^{h-2} \phi^{2j}$ . Under the Gaussian conditional forecasting approach, the forecast for period  $t + h$  is formed under the condition that  $y_{t+1}$  take the nowcast value of  $\mu_1$ , without any uncertainty around it. Using the more general formulas of Section 4.4, it is easily checked (see the last section of the appendix) that this conditional forecast distribution coincides with the jumping-off approach.

In our empirical implementation, we approximate the BVAR forecast distribution for a given variable and dates  $t + 1, \dots, t + 5$  via a five-variate Gaussian, and then apply the method just illustrated, whereby the (SPF or model based) nowcast for date  $t + 1$  takes the role of  $\mu_1$ .

### 6.2 Soft conditioning (Waggoner and Zha, 1999)

Waggoner and Zha (1999) consider VAR forecasts that condition on a certain path for one or more of the system variables. A key example is to forecast the evolution of inflation and output growth, given a certain path of the federal funds rate. They also consider approximate (“soft”) conditions which formulate a corridor for some of the system variables. Their resulting algorithm (Algorithm 2 of their paper) simply prescribes to keep the simulated forecast draws which satisfy the specified conditions, and discard the other draws. We apply this method on a variable-by-variable basis, and require the forecast draws for the current quarter  $t$  to lie in the interval  $\left[ \hat{Y}_{t,1} \pm 1.96 \sqrt{\hat{\sigma}_{t,1}^2} \right]$ ,

where  $\hat{Y}_{t,1}$  and  $\hat{\sigma}_{t,1}^2$  denote the nowcast mean and variance for the variable of interest. We use 5,000 posterior draws of the BVAR parameters, as well as 20 forecast paths for each parameter draw (in the notation of Waggoner and Zha, we thus use an oversampling rate of  $n_2 = 20$ ). The resulting number of forecast paths that satisfy the nowcast condition is never below 207, and exceeds 5,000 in about 97% of all cases.

### 6.3 Empirical results

Table 5 summarizes the performance of the methods in terms of RMSE and CRPS.

- Tilting performs similarly to jumping-off and soft conditioning, in that differences in RMSE or CRPS across these methods are typically smaller than differences across nowcast types (model based versus survey).
- In some cases, tilting attains markedly better RMSE and CRPS results than soft conditioning at the current quarter horizon. This may be due to unrealistic behavior of the soft conditioning method in case the nowcast deviates substantially from the BVAR forecast. Figure 5 provides an example, based on the current quarter distribution for GDP growth in 2008:Q4. The soft conditioning distribution consists of 7879 draws between  $-5.49$  and  $-0.38$  (draws within the rectangle on the left panel). The draws are clearly skewed toward the right endpoint of the interval. This is because they are taken from the left tail of the (roughly bell shaped) BVAR distribution.

Overall, these results indicate that, in terms of RMSE and CRPS accuracy, the empirical performance of tilting is competitive with, but not necessarily better than, other state of the art methods for combining BVAR forecasts with external nowcasts. However, as described above, tilting has other advantages, in terms of properly accounting for nowcast uncertainty, as well as flexibility.

## 7 Conclusion

This paper is concerned with the problem of combining forecasts from a BVAR with nowcasts from other sources. This combination problem is non-standard, in that the BVAR implies a joint forecast distribution for several forecast horizons, whereas the nowcast information is restricted to

		SPF Nowcasts					BMF Nowcasts				
<i>Horizon</i>		1	2	3	4	5	1	2	3	4	5
GDP											
RMSE	Tilting	<b>1.591</b>	<b>2.390</b>	<b>2.511</b>	2.637	2.631	<b>1.899</b>	<b>2.464</b>	<b>2.563</b>	2.643	2.627
	Soft Conditioning	1.649	2.433	2.550	<b>2.633</b>	<b>2.622</b>	2.037	2.514	2.608	<b>2.637</b>	<b>2.620</b>
	Jumping-off		2.403	2.522	2.649	2.637		2.469	2.568	2.647	2.629
CRPS	Tilting	<b>0.906</b>	<b>1.280</b>	<b>1.335</b>	1.411	1.411	<b>1.042</b>	<b>1.317</b>	<b>1.366</b>	1.412	1.410
	Soft Conditioning	0.972	1.302	1.359	<b>1.408</b>	<b>1.407</b>	1.125	1.343	1.390	<b>1.407</b>	<b>1.406</b>
	Jumping-off		1.302	1.352	1.421	1.419		1.327	1.375	1.419	1.415
UNE											
RMSE	Tilting	<b>0.151</b>	<b>0.350</b>	<b>0.573</b>	<b>0.808</b>	<b>1.024</b>	0.096	0.251	0.470	0.708	0.945
	Soft Conditioning	0.152	0.355	0.585	0.823	1.039	<b>0.079</b>	0.257	0.487	0.728	0.962
	Jumping-off		0.351	0.574	0.810	1.027		<b>0.251</b>	<b>0.467</b>	<b>0.704</b>	<b>0.940</b>
CRPS	Tilting	<b>0.083</b>	<b>0.175</b>	<b>0.276</b>	<b>0.391</b>	<b>0.505</b>	0.054	0.135	<b>0.233</b>	<b>0.346</b>	<b>0.462</b>
	Soft Conditioning	0.086	0.176	0.281	0.399	0.515	<b>0.046</b>	<b>0.134</b>	0.238	0.354	0.474
	Jumping-off		0.177	0.277	0.392	0.506		0.136	0.234	0.347	0.464
INF											
RMSE	Tilting	<b>0.809</b>	<b>0.915</b>	<b>0.973</b>	1.029	<b>1.144</b>	0.985	1.027	1.050	1.075	1.196
	Soft Conditioning	0.853	0.954	0.984	<b>1.028</b>	1.153	<b>0.944</b>	<b>1.008</b>	<b>1.025</b>	<b>1.055</b>	<b>1.180</b>
	Jumping-off		0.915	0.974	1.030	1.144		1.025	1.050	1.074	1.195
CRPS	Tilting	<b>0.463</b>	<b>0.534</b>	<b>0.558</b>	0.597	<b>0.672</b>	0.553	0.591	0.590	0.618	0.697
	Soft Conditioning	0.501	0.555	0.561	<b>0.595</b>	0.676	<b>0.545</b>	<b>0.584</b>	<b>0.578</b>	<b>0.607</b>	<b>0.689</b>
	Jumping-off		0.534	0.560	0.600	0.675		0.593	0.592	0.620	0.699
TBI											
RMSE	Tilting	<b>0.132</b>	<b>0.514</b>	<b>0.854</b>	<b>1.162</b>	<b>1.450</b>	<b>0.073</b>	<b>0.457</b>	<b>0.800</b>	<b>1.112</b>	<b>1.402</b>
	Soft Conditioning	0.147	0.532	0.871	1.179	1.465	0.089	0.470	0.810	1.120	1.409
	Jumping-off		0.518	0.860	1.170	1.459		0.460	0.807	1.121	1.412
CRPS	Tilting	<b>0.070</b>	<b>0.272</b>	<b>0.479</b>	<b>0.680</b>	<b>0.872</b>	<b>0.041</b>	<b>0.239</b>	<b>0.445</b>	<b>0.645</b>	<b>0.838</b>
	Soft Conditioning	0.079	0.285	0.494	0.695	0.888	0.042	0.245	0.451	0.651	0.845
	Jumping-off		0.283	0.491	0.689	0.878		0.256	0.463	0.661	0.850

Table 5: Root Mean Squared Errors and CRPS scores (the smaller, the better) for the complete sample (1988:Q3 – 2013:Q2). Tilting is based on the small m/v variant as described below Table 3. The best performing method in each comparison is printed in bold.

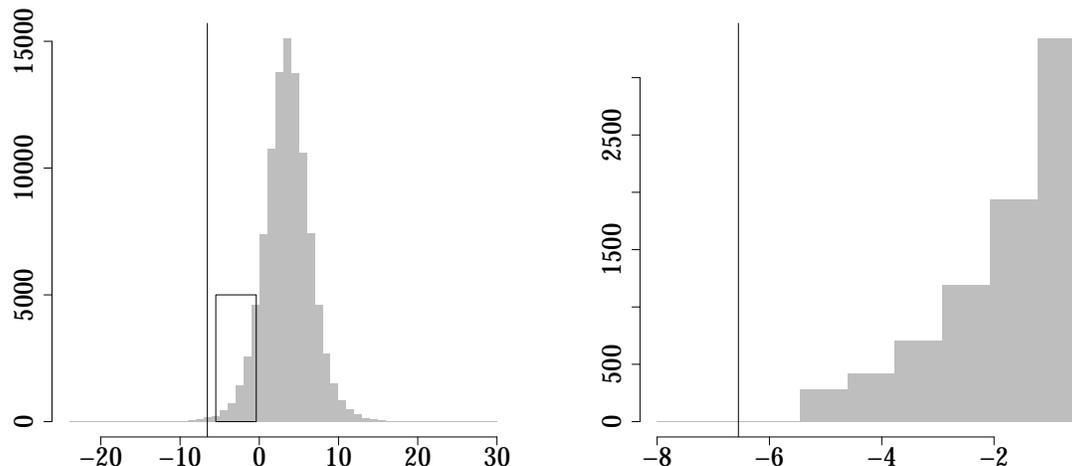


Figure 5: Left panel: histogram of all 100,000 BVAR draws for 2008:Q4 (current quarter forecasts). The 7,879 draws within the rectangle satisfy the soft condition imposed by SPF nowcasts. Right panel: zoomed histogram for the draws that satisfy the nowcast condition. In both panels, the vertical line marks the realizing value of  $-6.55$ .

mean and variance predictions for the current quarter. We argue that entropic tilting is a powerful tool to tackle these challenges; unlike other methods proposed in the literature, it does not require restrictive assumptions such as joint normality of the VAR system or zero variance of the nowcast.

In our empirical analysis, tilting systematically improves the accuracy of both point and density forecasts, and tilting the BVAR forecasts based on nowcast means and variances yields slightly greater gains in density accuracy than does just tilting based on the nowcast means. In a comparison of tilting on a variable-by-variable basis to tilting jointly toward the nowcasts for all four variables of the BVAR, we find that the overall differences in forecast performance for the joint treatment of variables versus the individual treatment of variables are small.

The analysis presented in this paper (in addition to results presented in the supplementary appendix) shows that the benefits of tilting are not limited to a specific empirical setup, but hold across a range of choices for both the external nowcast and the BVAR specification to which tilting is applied. Finally, our analysis in Section 6 documents that the empirical performance of tilting is competitive with, but not necessarily better than, other state of the art methods for combining information. Given their similar empirical performance, a user's choice of combination method may thus depend on additional factors such as theoretical appeal, flexibility and ease of use. We think that tilting is attractive along each of these dimensions.

## References

- ALTAVILLA, C., R. GIACOMINI, AND G. RAGUSA (2013): "Anchoring the Yield Curve Using Survey Expectations," Working Paper, UC London.
- ANG, A., G. BEKAERT, AND M. WEI (2007): "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?," *Journal of Monetary Economics*, 54, 1163 – 1212.
- BANBURA, M., D. GIANNONE, M. MODUGNO, AND L. REICHLIN (2013): "Now-Casting and the Real-Time Data Flow," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 195 – 237. Elsevier.
- BANBURA, M., D. GIANNONE, AND L. REICHLIN (2013): "Nowcasting," in *Oxford Handbook of Economic Forecasting*, ed. by M. P. Clements, and D. F. Hendry, pp. 193 – 224. Oxford University Press.
- CARRIERO, A., T. E. CLARK, AND M. MARCELLINO (2014): "Real-Time Nowcasting with a Bayesian Mixed Frequency Model with Stochastic Volatility," *Journal of the Royal Statistical Society: Series A*, forthcoming.
- CARTER, C. K., AND R. KOHN (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.
- CLARK, T. E. (2011): "Real-Time Density Forecasts From Bayesian Vector Autoregressions with Stochastic Volatility," *Journal of Business & Economic Statistics*, 29, 327–341.
- CLARK, T. E., AND M. W. MCCRACKEN (2013): "Advances in Forecast Evaluation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 1107–1201. Elsevier.
- (2014): "Evaluating Conditional Forecasts from Vector Autoregressions," Federal Reserve Bank of Cleveland Working Paper 14-13.
- CLARK, T. E., AND F. RAVAZZOLO (2015): "The Macroeconomic Forecasting Performance of Autoregressive Models with Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30, 551–575.
- CLEMENTS, M. P. (2014): "Forecast Uncertainty – Ex Ante and Ex Post: US Inflation and Output Growth," *Journal of Business & Economic Statistics*, 32, 206–216.

- COGLEY, T., S. MOROZOV, AND T. J. SARGENT (2005): "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System," *Journal of Economic Dynamics and Control*, 29, 1893–1925.
- COGLEY, T., AND T. J. SARGENT (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262–302.
- CROUSHORE, D. (2006): "Forecasting with Real-Time Macroeconomic Data," *Handbook of Economic Forecasting*, 1, 961–982.
- D'AGOSTINO, A., L. GAMBETTI, AND D. GIANNONE (2013): "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics*, 28, 82–101.
- DEL NEGRO, M., AND G. E. PRIMICERI (2015): "Time-Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *Review of Economic Studies*, forthcoming.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2013): "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 57–140. Elsevier.
- DIEBOLD, F. X., AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13, 253–263.
- DOAN, T., R. LITTERMAN, AND C. SIMS (1984): "Forecasting and Conditional Projection using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.
- DURBIN, J., AND S. J. KOOPMAN (2002): "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis," *Biometrika*, 89, 603–616.
- FAUST, J., AND J. H. WRIGHT (2009): "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset," *Journal of Business & Economic Statistics*, 27, 468–479.
- (2013): "Forecasting Inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 2–56. Elsevier.
- FREY, C., AND F. MOKINSKI (2015): "Forecasting with Bayesian Vector Autoregressions Estimated Using Professional Forecasts," *Journal of Applied Econometrics*, forthcoming.
- GEWEKE, J., AND G. AMISANO (2011): "Optimal Prediction Pools," *Journal of Econometrics*, 164, 130–141.

- GIACOMINI, R., AND H. WHITE (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578.
- GIANNONE, D., F. MONTI, AND L. REICHLIN (2014): "Exploiting the Monthly Data-flow in Structural Forecasting," Centre for Macroeconomics (CFM) Working Paper 1416.
- GNEITING, T., F. BALABDAOUI, AND A. E. RAFTERY (2007): "Probabilistic Forecasts, Calibration and Sharpness," *Journal of the Royal Statistical Society: Series B*, 69, 243–268.
- GNEITING, T., AND A. E. RAFTERY (2007): "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378.
- GNEITING, T., AND R. RANJAN (2011): "Comparing Density Forecasts using Threshold-and Quantile-weighted Scoring Rules," *Journal of Business & Economic Statistics*, 29, 411–422.
- (2013): "Combining Predictive Distributions," *Electronic Journal of Statistics*, 7, 1747–1782.
- HALL, S. G., AND J. MITCHELL (2007): "Combining Density Forecasts," *International Journal of Forecasting*, 23, 1–13.
- HARVEY, D., S. LEYBOURNE, AND P. NEWBOLD (1997): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281–291.
- HERSBACH, H. (2000): "Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems," *Weather and Forecasting*, 15, 559–570.
- KIM, S., N. SHEPHARD, AND S. CHIB (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.
- LEWIS, K. F., AND C. H. WHITEMAN (2015): "Empirical Bayesian Density Forecasting in Iowa and Shrinkage for the Monte Carlo Era," *Journal of Forecasting*, 34, 15–35.
- LITTERMAN, R. B. (1986): "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience," *Journal of Business & Economic Statistics*, 4, 25–38.
- MONTGOMERY, A. L., V. ZARNOWITZ, R. S. TSAY, AND G. C. TIAO (1998): "Forecasting the US Unemployment Rate," *Journal of the American Statistical Association*, 93, 478–493.
- NEWBY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

- NEWKEY, W. K., AND K. D. WEST (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.
- OMORI, Y., S. CHIB, N. SHEPHARD, AND J. NAKAJIMA (2007): "Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference," *Journal of Econometrics*, 140, 425–449.
- PRIMICERI, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.
- R CORE TEAM (2015): *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- ROBERTSON, J. C., E. W. TALLMAN, AND C. H. WHITEMAN (2005): "Forecasting Using Relative Entropy," *Journal of Money, Credit and Banking*, 37, 383–401.
- ROMER, C. D., AND D. H. ROMER (2000): "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90, 429–457.
- SCHORFHEIDE, F., AND D. SONG (2015): "Real-Time Forecasting with a Mixed-Frequency VAR," *Journal of Business & Economic Statistics*, 33, 366–380.
- SIMS, C. A. (2002): "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity*, 2002, 1–40.
- STONE, M. (1961): "The Opinion Pool," *The Annals of Mathematical Statistics*, 32, 1339–1342.
- WAGGONER, D. F., AND T. ZHA (1999): "Conditional Forecasts in Dynamic Multivariate Models," *Review of Economics and Statistics*, 81, 639–651.
- WEST, K. D. (1996): "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067–1084.
- WICKHAM, H. (2009): *ggplot2: Elegant Graphics for Data Analysis*. Springer.
- WOLTERS, M. H. (2015): "Evaluating Point and Density Forecasts of DSGE Models," *Journal of Applied Econometrics*, 30, 74–96.
- ZEILEIS, A. (2004): "Econometric Computing with HC and HAC Covariance Matrix Estimators," *Journal of Statistical Software*, 11, 1–17.

## Appendix

This appendix provides details of our priors and estimation algorithms and presents some analytical derivations establishing the equivalence of the jumping off approach to Gaussian conditional forecasting.

### A Prior for BVARs with stochastic volatility

We describe in this section the priors used with the BVAR-SV models to produce quarterly forecasts of the variables of interest and nowcasts of unemployment and the T-bill rate.

For the VAR coefficients, we use a conventional Minnesota prior, without cross-variable shrinkage (note that  $i$  and  $j$  refer to the row and column of  $B_l$ ):

$$\underline{\mu}_B \text{ such that } E[B_l^{(ij)}] = 0 \quad \forall i, j, l, \quad (14)$$

$$\underline{\Omega}_B \text{ such that } V[B_l^{(ij)}] = \begin{cases} \frac{\theta^2}{l^2} \sigma_j^2 & \text{for } l > 0, \\ \varepsilon^2 \sigma_i^2 & \text{for } l = 0. \end{cases} \quad (15)$$

Following common settings, we set  $\theta = 0.2$ ,  $\varepsilon = 1000$ , and the scale parameters  $\sigma_i^2$  at estimates of residual variances from AR( $p$ ) models from the estimation sample. With all of the variables of our VAR models defined so that they should be stationary, we set the prior mean of all the VAR coefficients to 0.

In the prior for the volatility-related components of the model, we use an approach to setting them similar to that of such studies as Cogley and Sargent (2005), Primiceri (2005) and Clark (2011). (The supplementary appendix provides evidence that our results are robust to alternative settings of the volatility-related priors.) The prior for  $A$  is uninformative, with a mean and variance for each row vector of  $\underline{\mu}_{a,i} = 0$ ,  $\underline{\Omega}_{a,i} = 1000^2 \cdot I_{i-1}$ ,  $i = 2, \dots, k$ . We make the priors on the volatility-related parameters loosely informative. The prior for  $\Phi$  is inverted Wishart, with mean of  $0.01 \times I_k$  and  $k + 1$  degrees of freedom. For the initial value of the log volatility of each equation  $i$ , we use a mean of  $\log \hat{\lambda}_{i,0,OLS}$  and variance of 4. To obtain  $\log \hat{\lambda}_{i,0,OLS}$ , we use the residuals from AR( $p$ ) models estimated over a training sample preceding the estimation sample. For each  $j = 2, \dots, k$ , we regress the residual from the AR model for  $j$  on the residuals associated with variables 1 through  $j - 1$  and compute the error variance  $\hat{\sigma}_{i,0}^2$ . We set the prior mean of log volatility in period 0 at  $\log \hat{\lambda}_{i,0,OLS} = \log \hat{\sigma}_{i,0}^2$ . For the quarterly  $\hat{\lambda}$  model and the unemployment nowcasting model, the

training sample is 1949-54; for the T-bill nowcasting model, the training sample is 1959-63. For the quarterly model and the unemployment nowcasting model, because a handful of the data vintages do not start until later than most others, we use the same prior mean on initial volatility for all vintages (forecast origins), computed using the last available vintage of data.

## B Prior for Bayesian mixed frequency models

We describe in this section the priors used for mixed frequency nowcasting models for GDP growth and inflation. (The supplementary appendix provides evidence that our results are robust to alternative settings of the priors detailed below.) Since the form of the prior is not dependent on the month timing  $m$ , in spelling out the prior we drop the index  $m$  from the model parameters for notational simplicity.

In all cases, for the coefficient vector  $\beta$ , we use a prior distribution that is normal, with mean 0 (for all coefficients) and variance that takes a diagonal, Minnesota-style form. The prior variance is Minnesota style in the sense that shrinkage increases with the lag (with the quarter, not with the month within the quarter), and in the sense that we impose more shrinkage on the monthly predictors than on lags of GDP growth or inflation. The shrinkage is controlled by two hyperparameters (in all cases, a smaller number means more shrinkage):  $\lambda_1$ , which controls the overall rate of shrinkage; and  $\lambda_2$ , which controls the rate of shrinkage on variables relative to GDP or GDP inflation.

At each forecast origin, the prior standard deviation associated with the coefficient on variable  $x_{i,j,t}$  of  $X_t$ , where  $i$  denotes the indicator (employment, etc.) and  $j$  denotes the month within the quarter  $t$  at which the indicator has been sampled, is specified as follows:

$$\text{sd}_{i,j,t} = \frac{\sigma_y}{\sigma_{i,j}} \lambda_1 \lambda_2. \quad (16)$$

For coefficients on lag 1 of  $y$ , the prior standard deviation is

$$\text{sd}_1 = \lambda_1. \quad (17)$$

Finally, for the intercept, the prior is uninformative:

$$\text{sd}_{int} = 1000\sigma_y. \quad (18)$$

In setting these components of the prior, for  $\sigma_y$  and  $\sigma_{i,j}$  we use standard deviations from AR(4) models for GDP growth or inflation and  $x_{i,j,t}$  estimated with the available sample of data.

In all of our results, the hyperparameters are set at values that may be considered very common in Minnesota-type priors (e.g. Litterman, 1986):  $\lambda_1 = 0.2$ , and  $\lambda_2 = 0.2$ .

Finally, in the prior for the volatility-related components of the model, our approach is similar to that used in such studies as Cogley and Sargent (2005), Primiceri (2005) and Clark (2011). For the prior on  $\phi$ , we use a mean of 0.035 and 5 degrees of freedom. For the period 0 value of volatility of each equation  $i$ , we use a prior of

$$\underline{\mu}_\lambda = \log \hat{\lambda}_{0,OLS}, \underline{\Omega}_\lambda = 4. \quad (19)$$

To obtain  $\log \hat{\lambda}_{0,OLS}$ , we use a training sample of 40 observations preceding the estimation sample to fit an AR(4) model to GDP growth or inflation.

## C Algorithm for BVAR with stochastic volatility

We estimate the BVAR-SV model with a five-step Gibbs sampling algorithm. Let  $X_t$  denote the collection of right-hand side variables of each equation of the VAR and  $B$  denote the vector of the system of VAR coefficients contained in  $B_i$ ,  $i = 0, \dots, p$ , as defined in the paper's equation (1).

Step 1: Draw the VAR coefficients  $B$  conditional on the history of  $\Lambda_t$ ,  $A$ , and  $\Phi$ .

The vector of coefficients is sampled from a conditional posterior distribution that is multivariate normal with mean  $\bar{\mu}_B$  and variance  $\bar{\Omega}_B$ , based on prior mean  $\underline{\mu}_B$  and variance  $\underline{\Omega}_B$ . Letting  $\Sigma_t = A^{-1}\Lambda_t A^{-1'}$ , the posterior mean and variance are:

$$\text{vec}(\bar{\mu}_B) = \bar{\Omega}_B \left\{ \text{vec} \left( \sum_{t=1}^T X_t y_t' \Sigma_t^{-1} \right) + \underline{\Omega}_B^{-1} \text{vec}(\underline{\mu}_B) \right\} \quad (20)$$

$$\bar{\Omega}_B^{-1} = \underline{\Omega}_B^{-1} + \sum_{t=1}^T (\Sigma_t^{-1} \otimes X_t X_t'). \quad (21)$$

Step 2: Draw the elements of  $A$  conditional on  $B$ , the history of  $\Lambda_t$ , and  $\Phi$ .

Following Cogley and Sargent (2005), rewrite the VAR as

$$A(y_t - X_t' B) = A \hat{y}_t \equiv \tilde{y}_t = \Lambda_t^{0.5} \epsilon_t, \quad (22)$$

where, conditional on  $B$ ,  $\hat{y}_t$  is observable. This system simplifies to a set of  $i = 2, \dots, k$  equations, with equation  $i$  having as dependent variable  $\hat{y}_{i,t}$  and as independent variables  $-1 \cdot \hat{y}_{j,t}$ ,  $j =$

$1, \dots, i-1$ , with coefficients  $a_{ij}$ . Multiplying equation  $i$  by  $\lambda_{i,t}^{-0.5}$  eliminates the heteroskedasticity associated with stochastic volatility. Then, proceeding separately for each transformed equation  $i$ , draw the  $i$ 'th equation's vector of coefficients  $a_i$  (a vector containing  $a_{ij}$  for  $j = 1, \dots, i-1$ ) from a normal posterior distribution with the mean and variance implied by the posterior mean and variance computed in the usual way. See Cogley and Sargent (2005) for details.

Step 3: Draw the elements of the states for the mixture distribution used to approximate the  $\chi^2$  distribution under the Kim, Shephard, and Chib (1998) algorithm, conditional on  $B, A$ , the history of  $\Lambda_t$ , and  $\Phi$ .

See Primiceri (2005) for details. However, we depart from Primiceri by using a 10 state approximation of the  $\chi^2$  distribution from Omori, Chib, Shephard, and Nakajima (2007) instead of the 7-state approximation from Kim, Shephard, and Chib (1998).

Step 4: Draw the elements of the variance matrix  $\Lambda_t$  conditional on  $B, A, \Phi$ , and the mixture states.

Following Primiceri (2005), the VAR can be rewritten as

$$A(y_t - X_t' B) \equiv \tilde{y}_t = \Lambda_t^{0.5} \epsilon_t,$$

where  $\epsilon_t \sim N(0, I_k)$ . Taking logs of the squares yields

$$\log \tilde{y}_{i,t}^2 = \log \lambda_{i,t}^2 + \log \epsilon_{i,t}^2, \quad i = 1, \dots, k.$$

The conditional volatility process is

$$\log(\lambda_{i,t}^2) = \log(\lambda_{i,t-1}^2) + \nu_{i,t}, \quad i = 1, \dots, k.$$

The estimation of the time series of  $\lambda_{i,t}^2$  uses the vector of the measured  $\log \tilde{y}_{i,t}^2$  and Primiceri's version of the Kim, Shephard, and Chib (1998) algorithm; see Primiceri for further detail (we depart from his implementation by using the Durbin and Koopman (2002) simulation smoother instead of the one proposed by Carter and Kohn (1994)).

Step 5: Draw the variance matrix  $\Phi$ , conditional on  $B$ , the history of  $\Lambda_t$ , and  $A$ .

Following Primiceri (2005), the sampling of  $\Phi$ , the variance of innovations to the log variances, is based on inverse Wishart priors and posteriors. The scale matrix of the posterior distribution is the sum of the prior mean  $\times$  the prior degrees of freedom and  $\sum_{t=1}^T \hat{\nu}_t \hat{\nu}_t'$ , where  $\hat{\nu}_t$  denotes the vector of innovations to the posterior draw of the volatilities for the set of variables.

## D Algorithm for Bayesian mixed frequency model

The mixed frequency nowcasting model with stochastic volatility is estimated with a Metropolis-within-Gibbs algorithm, used in such studies as Clark (2011) and Carriero, Clark, and Marcellino (2014). The posterior mean and variance of the coefficient vector are given by

$$\bar{\mu}_\beta = \bar{\Omega}_\beta \left\{ \sum_{t=1}^T \lambda_t^{-1} X_t y_t + \underline{\Omega}_\beta^{-1} \underline{\mu}_\beta \right\} \quad (23)$$

$$\bar{\Omega}_\beta^{-1} = \underline{\Omega}_\beta^{-1} + \sum_{t=1}^T \lambda_t^{-1} X_t X_t', \quad (24)$$

where we again omit the  $m$  index from the parameters for notational simplicity.

## E Analytics of jumping-off approach

Denote by  $\Sigma_{1h}$  the predicted covariance between 1-step and  $h$ -step ahead forecast errors (under the standard, unconditional approach without incorporating the nowcast). Furthermore, let  $\hat{y}_{t+h}^u$  denote the unconditional point forecast of  $y_{t+h}$ . For Gaussian conditioning, the results of Section 4.4 imply a point forecast of

$$\begin{aligned} \hat{y}_{t+h}^c &= \hat{y}_{t+h}^u + \frac{\Sigma_{1h}}{\Sigma_{11}} (\mu_1 - \hat{y}_{t+1}^u) \\ &= \phi^h y_t + \frac{\phi^{h-1} \sigma^2}{\sigma^2} (\mu_1 - \phi y_t) \\ &= \phi^{h-1} \mu_1. \end{aligned}$$

Given the nowcast condition on  $y_{t+1}$  (without uncertainty around it, so  $\Omega_{11} = 0$ ), the conditional forecast error variance is given by

$$\begin{aligned} \Omega_{hh} &= \Sigma_{hh} - \frac{\Sigma_{1h}^2}{\Sigma_{11}} \\ &= (1 + \phi^2 + \dots + \phi^{2h-2}) \sigma^2 - \frac{\phi^{2h-2} \sigma^4}{\sigma^2} \\ &= \sigma^2 \sum_{j=0}^{h-2} \phi^{2j}, \end{aligned}$$

thus coinciding with the jumping-off approach.

# Supplementary Appendix to “Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts”

This supplementary appendix provides additional robustness results — for other combination methods, a shorter sample period, alternative prior settings, and some additional VAR specifications — as well as some additional analysis of the effects of tilting on the forecast distributions.

## 1 Comparisons to other combination methods

We compare entropic tilting to two alternative methods that can be used for combining BVAR and external forecasts; these methods are more distantly related to entropic tilting, in that they employ external (survey) forecasts at longer horizons and not just at the current quarter horizon. We first describe these methods and then present the results of the comparison.

### 1.1 Linear pooling

A number of recent papers (e.g. Geweke and Amisano, 2011) have found that linear, equally weighted density forecast combinations are often superior to selecting a single method. Here we consider such an approach, combining the  $h$  step ahead BVAR forecast distribution with a Gaussian distribution constructed from the  $h$  step ahead SPF point forecast and its associated variance measure (see Section 2.2 in the paper). In practice, we simply augment 5,000 BVAR forecast draws with 5,000 draws from the SPF-based Gaussian distribution. We then compute the CRPS for this sample as described in Section 5 of the paper.

### 1.2 Long-run prior (Wright, 2013)

Wright (2013) found that using survey forecasts to set the prior on the steady state of a VAR in the steady state formulation of Villani (2009) and portions of the prior on the VAR coefficients

improved significantly on the accuracy of a conventional BVAR. We also compare the accuracy of our BVAR forecasts tilted toward nowcasts to the accuracy of Wright-type forecasts, under the prior denoted in Wright as D4.

Augmented to include stochastic volatility as in Clark (2011) — although the Wright (2013) specification did not include stochastic volatility, we include it to avoid giving our BVAR-SV specification an advantage — the model with the steady-state prior takes the form:

$$\begin{aligned}
 y_t - \mu &= \sum_{i=1}^p B_i(y_{t-i} - \mu) + v_t, \\
 v_t &= A^{-1} \Lambda_t^{0.5} \epsilon_t, \quad \epsilon_t \sim N(0, I_k), \quad \Lambda_t \equiv \text{diag}(\lambda_{1,t}, \dots, \lambda_{k,t}), \\
 \log(\lambda_{i,t}) &= \log(\lambda_{i,t-1}) + \nu_{i,t}, \quad i = 1, \dots, k, \\
 \nu_t &\equiv (\nu_{1,t}, \nu_{2,t}, \dots, \nu_{k,t})' \sim N(0, \Phi),
 \end{aligned}$$

where  $\mu$  is a vector of constants containing steady-state means.

In setting the prior on the VAR coefficients and  $\mu$ , we follow Wright’s usage of information from the Blue Chip consensus, along with his implementation. At each forecast origin, the prior mean on the steady states of  $\mu$  are set to the 6-10 year ahead forecast from Blue Chip. The prior variance for each steady state is set to be proportional to the difference between the top 10 and bottom 10 forecasts from Blue Chip. The prior mean on each AR(1) coefficient is set at the value implied by the quarterly forecasts available in each Blue Chip survey, obtained by solving the minimization problem described by Wright. The priors used at each forecast origin vary with the information in the associated Blue Chip survey available at that time. Finally, the overall tightness of the prior on the VAR coefficients is set at the same value (0.2) used in our baseline BVAR-SV specification.<sup>3</sup> The priors on the volatility-portion of the model are set just as they are for our baseline BVAR-SV model.

### 1.3 Empirical results

Table 1 summarizes the performance of the methods in terms of RMSE and CRPS.

- At the current quarter horizon, linear pooling performs worse than the nowcast based meth-

---

<sup>3</sup>Wright (2013) uses a tighter prior, but his model is larger than ours, and it is common in the literature to use tighter priors for larger models than smaller.

ods. At longer horizons, the performance of nowcast based methods versus linear pooling is more mixed. For GDP and inflation, linear pooling performs clearly better, which seems to be due to the increased accuracy of the SPF distribution (compared to the BVAR) at horizons beyond the current quarter. For T-bill and unemployment, the methods using SPF nowcasts perform similar to linear pooling.

- Targeting a long-run prior mean performs clearly worse than tilting, except for inflation at horizons 3 to 5.

## **2 Results for the pre-crisis period**

Table 2 is the analogue of the paper's Table 3, covering the pre-crisis sample (1988:Q3–2007:Q4) instead of the complete sample (1988:Q3–2013:Q2). Similarly, Table 3 is the pre-crisis analogue of Table 4. Broadly, results for the pre-crisis sample are very similar to those for the full sample.

		SPF Nowcasts				
<i>Horizon</i>		1	2	3	4	5
GDP						
RMSE	Tilting	<b>1.591</b>	2.390	2.511	2.637	2.631
	Linear Pooling	1.803	<b>2.114</b>	<b>2.279</b>	<b>2.368</b>	<b>2.378</b>
	Wright Prior	2.322	2.581	2.657	2.668	2.671
CRPS	Tilting	<b>0.906</b>	1.280	1.335	1.411	1.411
	Linear Pooling	0.982	<b>1.138</b>	<b>1.230</b>	<b>1.281</b>	<b>1.291</b>
	Wright Prior	1.236	1.373	1.410	1.419	1.427
UNE						
RMSE	Tilting	<b>0.151</b>	<b>0.350</b>	0.573	0.808	1.024
	Linear Pooling	0.175	0.364	<b>0.559</b>	<b>0.767</b>	<b>0.957</b>
	Wright Prior	0.229	0.455	0.697	0.937	1.151
CRPS	Tilting	<b>0.083</b>	<b>0.175</b>	0.276	0.391	0.505
	Linear Pooling	0.089	0.177	<b>0.271</b>	<b>0.378</b>	<b>0.482</b>
	Wright Prior	0.116	0.214	0.323	0.439	0.553
INF						
RMSE	Tilting	<b>0.809</b>	0.915	0.973	1.029	1.144
	Linear Pooling	0.841	<b>0.913</b>	<b>0.932</b>	0.985	1.076
	Wright Prior	0.935	0.969	0.956	<b>0.956</b>	<b>1.047</b>
CRPS	Tilting	<b>0.463</b>	0.534	0.558	0.597	0.672
	Linear Pooling	0.481	<b>0.527</b>	<b>0.535</b>	0.574	0.633
	Wright Prior	0.533	0.562	0.547	<b>0.560</b>	<b>0.620</b>
TBI						
RMSE	Tilting	<b>0.132</b>	<b>0.514</b>	<b>0.854</b>	<b>1.162</b>	<b>1.450</b>
	Linear Pooling	0.246	0.578	0.898	1.208	1.503
	Wright Prior	0.381	0.725	1.029	1.308	1.555
CRPS	Tilting	<b>0.070</b>	<b>0.272</b>	<b>0.479</b>	<b>0.680</b>	<b>0.872</b>
	Linear Pooling	0.110	0.309	0.510	0.708	0.896
	Wright Prior	0.203	0.405	0.597	0.776	0.944

Table 1: Root Mean Squared Errors and CRPS scores (the smaller, the better) for the complete sample (1988:Q3 – 2013:Q2). Tilting is based on the small m/v variant as described below the paper’s Table 3. The best performing method in each comparison is printed in bold.

<i>Forecast Horizon</i>		1	2	3	4	5	4*	5*
		GDP						
RMSE	raw	1.975	2.039	2.009	<b>1.991</b>	1.973	1.442	1.452
	SPF small m	<b>1.580*</b>	1.982	1.958	2.038	2.026	1.232*	1.459
	SPF small m/v	1.580*	1.961	<b>1.950*</b>	2.026*	2.010	<b>1.225*</b>	1.440
	SPF big m	1.580*	1.986	2.011	2.028	<b>1.950</b>	1.231	1.447
	SPF big m/v	1.587*	1.994	1.958	2.012	1.996	1.225	<b>1.427</b>
	BMF small m	1.682	1.973	1.967	2.033	2.011	1.252	1.450
	BMF small m/v	1.682	<b>1.960*</b>	1.963	2.019	2.000	1.249	1.440
	BMF big m	1.682	1.978	2.076	2.089	1.988	1.301	1.494
	BMF big m/v	1.679*	1.988	2.013	2.060	2.020	1.291	1.473
CRPS	raw	1.123	1.156	1.142	<b>1.134</b>	1.132	0.826	0.832
	SPF small m	0.934**	1.131	1.120	1.155*	1.152*	0.725	0.830
	SPF small m/v	<b>0.900**</b>	<b>1.116</b>	<b>1.110*</b>	1.154*	1.149	<b>0.710</b>	0.825
	SPF big m	0.942**	1.146	1.146	1.158	<b>1.116</b>	0.731	0.826
	SPF big m/v	0.904*	1.133	1.122	1.150	1.133	0.713	0.825
	BMF small m	0.980*	1.128	1.126	1.150	1.147	0.735	0.827
	BMF small m/v	0.956*	1.117*	1.119*	1.147	1.145	0.722	<b>0.824</b>
	BMF big m	0.993*	1.138	1.174	1.183*	1.136	0.767	0.852
	BMF big m/v	0.954*	1.127	1.142	1.180	1.153	0.744	0.846
		UNE						
RMSE	raw	0.157	0.274	0.398	0.511	0.605		
	SPF small m	0.125*	0.239	0.352	0.468	0.567		
	SPF small m/v	0.125*	0.240	0.353	0.468	0.565		
	SPF big m	0.125*	0.230*	0.333*	0.447	0.552		
	SPF big m/v	0.125*	0.230*	0.334*	0.446	0.545		
	BMF small m	<b>0.095**</b>	0.195*	0.305*	0.414*	0.517*		
	BMF small m/v	0.096**	0.191*	0.295*	0.407*	0.509*		
	BMF big m	0.095**	0.182*	0.282**	0.394*	0.502*		
	BMF big m/v	0.096**	<b>0.172**</b>	<b>0.272**</b>	<b>0.388*</b>	<b>0.498*</b>		
CRPS	raw	0.089	0.152	0.222	0.290	0.351		
	SPF small m	0.074**	0.137*	0.200	0.268	0.329		
	SPF small m/v	0.070**	0.133*	0.197	0.268	0.329		
	SPF big m	0.076*	0.135*	0.194*	0.260	0.323		
	SPF big m/v	0.069**	0.127**	0.186*	0.253	0.316		
	BMF small m	0.064**	0.122*	0.182*	0.244*	0.303*		
	BMF small m/v	<b>0.054**</b>	0.110*	0.169*	0.233*	0.293*		
	BMF big m	0.066**	0.121*	0.176*	0.239*	0.300*		
	BMF big m/v	0.054**	<b>0.103**</b>	<b>0.159**</b>	<b>0.224**</b>	<b>0.289*</b>		

Table 2: Empirical results for entropic tilting, pre-crisis sample (1988:Q3 – 2007:Q4). Design of the table corresponds to the paper’s Table 3.

<i>Forecast Horizon</i>		1	2	3	4	5	4*	5*
			INF					
RMSE	raw	0.876	0.996	0.969	1.007	1.134	0.696	0.782
	SPF small m	0.767*	0.921**	0.943	0.997	1.110	0.633	0.745
	SPF small m/v	0.767*	0.919**	<b>0.942</b>	0.996	1.110	0.632	0.744
	SPF big m	0.767*	<b>0.913*</b>	0.954	0.991	1.107	0.628	<b>0.740</b>
	SPF big m/v	<b>0.766*</b>	0.917*	0.947	<b>0.980</b>	<b>1.103</b>	<b>0.624</b>	0.741
	BMF small m	0.861	1.025	0.999	1.019	1.141	0.725	0.805
	BMF small m/v	0.861	1.026	1.000	1.020	1.143	0.726	0.806
	BMF big m	0.861	1.005	1.011	1.021	1.149	0.718	0.801
	BMF big m/v	0.859	1.008	0.994	1.004	1.127	0.713	0.792
CRPS	raw	0.504	0.571	0.548	0.585	0.661	0.400	0.453
	SPF small m	0.445*	0.533**	0.540	0.579	0.650	0.373	0.438
	SPF small m/v	0.436*	<b>0.529**</b>	0.537	0.577	0.648	0.366	0.434
	SPF big m	0.446*	0.534*	0.545	0.578	0.652	0.373	0.438
	SPF big m/v	<b>0.435*</b>	0.531*	<b>0.537</b>	<b>0.565</b>	<b>0.643</b>	<b>0.361</b>	<b>0.430</b>
	BMF small m	0.490	0.580	0.559	0.589	0.664	0.410	0.463
	BMF small m/v	0.494	0.583	0.560	0.588	0.665	0.412	0.463
	BMF big m	0.493	0.576	0.570*	0.595	0.667	0.412	0.465
	BMF big m/v	0.492	0.579	0.557	0.578	0.655	0.405	0.455
			TBI					
RMSE	raw	0.406	0.750	1.027	1.269	1.481		
	SPF small m	0.133**	0.523**	0.842**	1.120**	1.363**		
	SPF small m/v	0.132**	0.505**	0.824**	1.099**	1.343**		
	SPF big m	0.133**	0.518**	0.828**	1.098**	1.337*		
	SPF big m/v	0.133**	0.507**	0.823**	1.092**	1.333*		
	BMF small m	0.066**	0.478**	0.815**	1.101**	1.345**		
	BMF small m/v	0.067**	0.458**	0.793**	1.076**	1.323**		
	BMF big m	0.066**	0.486**	0.782**	<b>1.041**</b>	<b>1.283**</b>		
	BMF big m/v	<b>0.066**</b>	<b>0.449**</b>	<b>0.777**</b>	1.045**	1.291**		
CRPS	raw	0.214	0.413	0.586	0.741	0.885		
	SPF small m	0.157**	0.337**	0.508**	0.669**	0.819**		
	SPF small m/v	0.068**	0.268**	0.461**	0.637**	0.796**		
	SPF big m	0.158**	0.332**	0.497**	0.653**	0.800**		
	SPF big m/v	0.071**	0.270**	0.461**	0.634**	0.789**		
	BMF small m	0.150**	0.325**	0.498**	0.660**	0.811**		
	BMF small m/v	<b>0.040**</b>	0.241**	0.438**	0.617**	0.778**		
	BMF big m	0.155**	0.324**	0.478**	0.626**	0.773**		
	BMF big m/v	0.046**	<b>0.238**</b>	<b>0.427**</b>	<b>0.594**</b>	<b>0.752**</b>		

Table 2: continued.

Forecast Horizon		1	2	3	4	5	4*	5*
		GDP						
Coverage	SPF m	0.846	0.705	0.731	0.718	0.718	0.667	0.641
	SPF m/v	0.731	0.679	0.705	0.705	0.718	0.641	0.641
	BMF m	0.795	0.731	0.718	0.705	0.705	0.667	0.641
	BMF m/v	0.705	0.692	0.718	0.705	0.705	0.667	0.641
Length	SPF m	4.736	4.446	4.498	4.553	4.593	2.658	2.652
	SPF m/v	3.372	4.302	4.377	4.439	4.491	2.421	2.584
	BMF m	4.586	4.414	4.453	4.508	4.555	2.615	2.629
	BMF m/v	3.790	4.313	4.378	4.436	4.483	2.473	2.585
CRPS	SPF m	0.934	1.131	1.120	1.155	1.152	0.725	0.830
	SPF m/v	<b>0.900</b>	<b>1.116</b>	<b>1.110</b>	1.154	1.149	<b>0.710</b>	0.825
	BMF m	0.980	1.128	1.126	1.150	1.147	0.735	0.827
	BMF m/v	0.956	1.117	1.119	<b>1.147</b>	<b>1.145</b>	0.722	<b>0.824</b>
		UNE						
Coverage	SPF m	0.859	0.833	0.756	0.718	0.654		
	SPF m/v	0.833	0.756	0.744	0.654	0.615		
	BMF m	0.987	0.910	0.872	0.808	0.769		
	BMF m/v	0.833	0.821	0.744	0.705	0.679		
Length	SPF m	0.392	0.626	0.815	0.977	1.088		
	SPF m/v	0.275	0.513	0.721	0.893	1.025		
	BMF m	0.450	0.693	0.870	1.036	1.141		
	BMF m/v	0.239	0.489	0.705	0.886	1.027		
CRPS	SPF m	0.074	0.137	0.200	0.268	0.329		
	SPF m/v	0.070	0.133	0.197	0.268	0.329		
	BMF m	0.064	0.122	0.182	0.244	0.303		
	BMF m/v	<b>0.054</b>	<b>0.110</b>	<b>0.169</b>	<b>0.233</b>	<b>0.293</b>		
		INF						
Coverage	SPF m	0.795	0.718	0.744	0.705	0.705	0.795	0.756
	SPF m/v	0.718	0.692	0.731	0.692	0.705	0.795	0.744
	BMF m	0.679	0.705	0.705	0.744	0.705	0.744	0.756
	BMF m/v	0.628	0.705	0.705	0.718	0.679	0.731	0.744
Length	SPF m	1.893	2.074	2.243	2.408	2.574	1.685	1.898
	SPF m/v	1.602	1.988	2.171	2.339	2.505	1.564	1.820
	BMF m	1.869	2.062	2.240	2.398	2.566	1.680	1.897
	BMF m/v	1.732	2.021	2.205	2.372	2.540	1.626	1.864
CRPS	SPF m	0.445	0.533	0.540	0.579	0.650	0.373	0.438
	SPF m/v	<b>0.436</b>	<b>0.529</b>	<b>0.537</b>	<b>0.577</b>	<b>0.648</b>	<b>0.366</b>	<b>0.434</b>
	BMF m	0.490	0.580	0.559	0.589	0.664	0.410	0.463
	BMF m/v	0.494	0.583	0.560	0.588	0.665	0.412	0.463
		TBI						
Coverage	SPF m	0.962	0.782	0.641	0.564	0.462		
	SPF m/v	0.808	0.705	0.564	0.449	0.423		
	BMF m	0.987	0.782	0.667	0.590	0.513		
	BMF m/v	0.936	0.731	0.603	0.474	0.436		
Length	SPF m	0.935	1.284	1.542	1.775	1.979		
	SPF m/v	0.263	0.812	1.168	1.450	1.684		
	BMF m	0.973	1.302	1.553	1.786	1.988		
	BMF m/v	0.238	0.805	1.165	1.449	1.683		
CRPS	SPF m	0.157	0.337	0.508	0.669	0.819		
	SPF m/v	0.068	0.268	0.461	0.637	0.796		
	BMF m	0.150	0.325	0.498	0.660	0.811		
	BMF m/v	<b>0.040</b>	<b>0.241</b>	<b>0.438</b>	<b>0.617</b>	<b>0.778</b>		

Table 3: **Impact of accounting for nowcast uncertainty, pre-crisis sample (1988:Q3 – 2007:Q4).** Design of the table corresponds to the paper’s Table 4.

### 3 Prior robustness checks for nowcast models

In general, as documented in the paper, our broad results on tilting apply with nowcasts obtained from both the Survey of Professional Forecasters and the nowcasting models detailed in the paper. We have further verified the robustness of our results by comparing the accuracy of our model-based nowcasts obtained under prior specifications that differ from those of the baseline.

#### Mixed frequency models for GDP growth and inflation

In the mixed frequency models for GDP growth and inflation, we set the hyperparameters controlling shrinkage on the regression coefficients at values that may be considered very common in Minnesota-type priors:  $\lambda_1 = 0.2$  and  $\lambda_2 = 0.2$ . For the baseline prior on  $\phi$ , the mean is 0.035 and the degrees of freedom equal 5. In the present appendix we also consider three alternative specifications of the prior, as follows.

- alternative prior 1:  $\lambda_1 = 0.2$ ,  $\lambda_2 = 0.2$ ; prior mean on  $\phi = 0.01$
- alternative prior 2:  $\lambda_1 = 0.2$ ,  $\lambda_2 = 0.2$ ; prior mean on  $\phi = 0.0001$
- alternative prior 3:  $\lambda_1$  and  $\lambda_2$  optimized in real time to maximize the marginal likelihood of the corresponding constant volatility version of the mixed frequency model estimated with data available up to each forecast origin; prior mean on  $\phi = 0.035$ .

Whereas the baseline prior sets the mean for the variance of volatility innovations at the value used in Carriero, Clark, and Marcellino (2014), alternative prior 1 covers the baseline prior mean we use in the VARs (see below), and alternative prior 2 covers the much lower mean for variances of volatility innovations considered in such studies as Cogley and Sargent (2005), Primiceri (2005), and D'Agostino, Gambetti, and Giannone (2013). In contrast to the setting of  $\lambda_1$  and  $\lambda_2$  at typical Minnesota prior-like values, alternative prior 3 sets the hyperparameters to maximize model fit as measured by the marginal likelihood. Carriero, Clark, and Marcellino (2014) reported considering the choice of hyperparameters to instead maximize model fit as measured by past forecast accuracy (RMSE).

## Monthly VARs for unemployment and T-bill

In the baseline VAR models at the monthly frequency used to nowcast the unemployment rate and T-bill rate, for the volatility portion of the model, the prior for  $A$  is uninformative, with a mean and variance for each row vector of  $\underline{\mu}_{a,i} = 0$ ,  $\underline{\Omega}_{a,i} = 1000^2 \cdot I_{i-1}$ ,  $i = 2, \dots, k$ , and the prior for  $\Phi$  is inverted Wishart, with mean of  $0.01 \times I_k$  and  $k + 1$  degrees of freedom. In this appendix, we consider three alternative specifications of the prior for the volatility portion of the model, as follows.

- alternative prior 1: prior mean of  $\Phi = 0.0001 \times I_k$ ; prior on  $A$  same as in the baseline
- alternative prior 2: prior mean of  $\Phi = 0.035 \times I_k$ ; prior on  $A$  same as in the baseline
- alternative prior 3: prior mean of  $\Phi = 0.01 \times I_k$ ; for each row  $i$  of  $A$ ,  $i = 2, \dots, k$ , the prior mean is set to estimates obtained from OLS regressions over the training sample preceding the estimation sample, and the prior variance is set to  $10 \cdot I_{i-1}$ .

As noted in the paper, the baseline prior follows Cogley and Sargent (2005) and others in making the prior on  $A$  uninformative and Clark (2011), among others, in setting the prior on  $\Phi$  to assume for a decent amount of time variation in volatility. Alternative priors 1 and 2 keep the prior on  $A$  uninformative and cover the much lower mean for  $\Phi$  considered in such studies as Cogley and Sargent (2005) as well as the higher mean allowed in Clark (2011). Alternative prior 3 retains the baseline setting for the prior on  $\Phi$  and makes the prior on  $A$  modestly informative, with a mean obtained from OLS estimates over a training sample.

## Empirical results

Table 4 compares the point (RMSE) and density (CRPS) forecast accuracy of model-based nowcasts under the baseline and alternative priors, for the full sample of 1988:Q3–2013:Q2. For the BVAR-SV models at the monthly frequency used to nowcast unemployment and the T-bill rate, the results are very clear: the alternative priors yield results that are pretty much exactly the same as the baseline results, for both RMSE and CRPS accuracy. With these models, the estimation sample is relatively large due to the monthly frequency, helping reduce sensitivity to prior. For the mixed frequency models used to nowcast GDP growth and inflation, adjusting the prior can have some

effect on the results, but the effects are not large, and they can differ across variables. Reducing the prior mean of  $\Phi$  from 0.035 (baseline) to 0.01 (alternative prior 1) has very little effect on the RMSE and CRPS accuracy of growth and inflation forecasts. Further reducing the prior mean of  $\Phi$  to a very low level of 0.0001 (alternative prior 2) slightly reduces accuracy as measured by RMSE and CRPS, for both growth and inflation. Setting the shrinkage hyperparameters to maximize the marginal likelihood (alternative prior 3) has mixed effects, modestly reducing the accuracy of GDP growth forecasts and modestly increasing the accuracy of inflation forecasts.

RMSE				
prior specification	GDP	UNE	INF	TBI
baseline	1.899	0.095	0.985	0.072
alt. prior 1	1.898	0.095	0.988	0.072
alt. prior 2	1.909	0.095	1.026	0.072
alt. prior 3	2.025	0.096	0.897	0.072

CRPS				
prior specification	GDP	UNE	INF	TBI
baseline	1.048	0.053	0.554	0.035
alt. prior 1	1.046	0.053	0.556	0.035
alt. prior 2	1.082	0.053	0.580	0.035
alt. prior 3	1.128	0.054	0.506	0.035

Table 4: RMSEs and CRPS scores (the smaller, the better) for model-based nowcasts obtained under alternative prior specifications. The sample is the complete sample (1988:Q3 – 2013:Q2).

## 4 Applying tilting to alternative BVAR specifications

In our main analysis, we apply entropic tilting to predictive samples from a Bayesian VAR with stochastic volatility and prior distributions as presented in the paper’s appendix. A natural question is whether tilting is similarly beneficial when applied to MCMC output from other specifications. To answer this question, we consider the following alternatives:

- The Bayesian VAR used in the main text, but with three alternative prior specifications, as follows:
  - alternative prior 1: prior mean of  $\Phi = 0.0001 \times I_k$ ; prior on  $A$  same as in the baseline
  - alternative prior 2: prior mean of  $\Phi = 0.035 \times I_k$ ; prior on  $A$  same as in the baseline

– alternative prior 3: prior mean of  $\Phi = 0.01 \times I_k$ ; for each row  $i$  of  $A$ ,  $i = 2, \dots, k$ , the prior mean is set to estimates obtained from OLS regressions over the training sample preceding the estimation sample, and the prior variance is set to  $10 \cdot I_{i-1}$ .

- A time-varying parameter VAR(2), as detailed below
- The Wright (2013) model as described in Section 1.2 above

To ease computational burden, we use an MCMC sample size of 5,000 forecast draws (rather than  $5 \times 5,000$  draws, as in the main results). As documented in Section 5 below, differences across these two sample sizes tend to be very small.

Regarding the VAR with time-varying parameters and stochastic volatility, let  $X_t$  denote the collection of right-hand side variables of each equation of the VAR and  $B_t$  denote the period  $t$  value of the vector of all VAR coefficients (of dimension  $k(kp + 1) \times 1$ ). The VAR-TVP-SV model takes the form given in Cogley and Sargent (2005):

$$\begin{aligned}
 y_t &= X_t' B_t + v_t, \\
 B_t &= B_{t-1} + n_t, \text{ var}(n_t) = Q, \\
 v_t &= A^{-1} \Lambda_t^{0.5} \epsilon_t, \epsilon_t \sim N(0, I_k), \Lambda_t \equiv \text{diag}(\lambda_{1,t}, \dots, \lambda_{k,t}), \\
 \log(\lambda_{i,t}) &= \log(\lambda_{i,t-1}) + \nu_{i,t}, i = 1, \dots, k, \\
 \nu_t &\equiv (\nu_{1,t}, \nu_{2,t}, \dots, \nu_{k,t})' \sim N(0, \Phi),
 \end{aligned} \tag{1}$$

where  $A$  is a lower triangular matrix with ones on the diagonal and non-zero coefficients below the diagonal. The VAR coefficients follow random walk processes, with innovations that are allowed to be correlated across coefficients. The volatility portion of the VAR-TVP-SV model is the same as that of the baseline VAR specification used in the paper. Priors are set as detailed in Clark and Ravazzolo (2015).

For each of the specifications, Table 5 presents results on the impact of tilting. The table reports the (RMSE or CRPS) *difference* between the raw BVAR output and the small m/v variant of tilting. A positive difference means that tilting improves forecast performance. The table shows that applying tilting to all of the specifications is generally beneficial, except for GDP at the longest horizons, and for inflation using model nowcasts. Across all specifications, model nowcasts are more effective in the case of unemployment and T-bill, whereas SPF nowcasts are more effective for GDP and inflation. As regards the magnitude of the effects of tilting, there is no clear pattern

regarding which specification benefits most, with each specification attaining the largest benefits in a number of occasions (indicated by boldface printing in Table 5). In summary, the results from our main analysis are robust across a range of BVAR specifications to which tilting can be applied.

## 5 Analysis of tilting weights

As illustrated in Section 4.5 of the paper, the tilting weights are fairly unequal in some instances. This appendix provides some additional discussion of this issue and its potential consequences. For concreteness, the analysis in this section focuses on model based nowcasts and tilting towards the mean and variance (m/v).

Figure 1 visualizes the inequality of the tilting weights obtained at different forecast origin dates. Denote by  $\Pi_t^* = \{\pi_{t,(i)}^*\}_{i=1}^I$  the ordered set of tilting weights at date  $t$ , such that  $\pi_{t,(1)}$  is the smallest weight and  $\pi_{t,(I)}$  is the largest weight. The figure plots the Gini index,

$$G_t = \frac{2 \sum_{i=1}^I i \pi_{t,(i)}}{I \sum_{i=1}^I \pi_{t,(i)}} - \frac{I+1}{I},$$

against time  $t$ . Note that  $G_t = 0$  corresponds to perfect equality and  $G_t = 1$  corresponds to perfect inequality.

The figure shows that the big tilting variant (jointly targeting the nowcasts of all variables) produces the most unequal weight distributions (Gini coefficients exceed 0.5 for most dates). This is as expected from the example in Section 4.5. Among the small variants, we find that the tilting weights are most unequal for the T-bill rate and fairly equal for inflation. This suggests that the nowcast-implied moment conditions are “further away” from the BVAR for the T-bill rate than for inflation. Furthermore, the figure shows that the Gini coefficients are positively correlated over time.

A natural question is whether highly unequal tilting weights degrade the numerical performance of tilting, due to the reduced effective size of the tilted sample. To investigate this issue, we analyze forecast performance under two different sizes of the simulated sample to which tilting is applied: The setup from the main analysis (25,000 draws, composed of 5,000 parameter draws times five forecast draws per parameter draw), and a reduced sample (5,000 parameter draws, composed of a single forecast draw per parameter draw). Figure 2 provides a graphical illustration of the differences in RMSE and CRPS (averaged over the complete sample period). The figure

		SPF Nowcasts					BMF Nowcasts				
<i>Horizon</i>		1	2	3	4	5	1	2	3	4	5
GDP											
$\Delta$ RMSE	Basic BVAR-SV	0.799	0.199	0.145	0.002	-0.014	0.490	0.126	0.092	-0.003	-0.010
	Alt. Prior 1	<b>0.847</b>	0.216	0.148	-0.004	-0.029	<b>0.539</b>	0.143	0.100	-0.003	-0.015
	Alt. Prior 2	0.796	0.202	0.144	-0.022	-0.011	0.488	0.119	0.094	-0.015	-0.012
	Alt. Prior 3	0.794	0.190	0.139	-0.007	-0.014	0.485	0.122	0.094	-0.008	-0.013
	TVPSV	0.706	0.120	0.136	0.015	0.010	0.398	0.050	0.065	0.005	-0.001
Wright	0.732	<b>0.261</b>	<b>0.173</b>	<b>0.030</b>	<b>0.020</b>	0.423	<b>0.166</b>	<b>0.115</b>	<b>0.021</b>	<b>0.013</b>	
$\Delta$ CRPS	Basic BVAR-SV	0.368	0.100	0.078	-0.003	-0.008	0.232	0.063	0.048	-0.004	-0.007
	Alt. Prior 1	<b>0.391</b>	0.103	0.076	-0.007	-0.016	<b>0.251</b>	0.070	0.052	-0.003	-0.010
	Alt. Prior 2	0.368	0.097	0.079	-0.013	-0.011	0.229	0.061	0.049	-0.009	-0.009
	Alt. Prior 3	0.363	0.091	0.076	-0.007	-0.013	0.227	0.060	0.049	-0.006	-0.010
	TVPSV	0.300	0.060	0.074	0.007	-0.002	0.165	0.024	0.037	0.001	-0.007
Wright	0.329	<b>0.123</b>	<b>0.093</b>	<b>0.011</b>	<b>0.006</b>	0.192	<b>0.083</b>	<b>0.061</b>	<b>0.010</b>	<b>0.003</b>	
UNE											
$\Delta$ RMSE	Basic BVAR-SV	0.084	0.114	0.132	0.131	0.123	0.139	0.213	0.236	0.232	0.202
	Alt. Prior 1	0.083	0.114	<b>0.135</b>	0.135	0.126	<b>0.141</b>	0.197	0.240	0.231	0.182
	Alt. Prior 2	<b>0.085</b>	<b>0.115</b>	0.134	0.134	0.125	0.141	0.212	0.244	0.260	0.230
	Alt. Prior 3	0.083	0.113	0.131	0.130	0.119	0.140	0.181	0.166	0.130	0.033
	TVPSV	0.067	0.074	0.085	0.080	0.074	0.122	0.172	0.201	0.213	0.207
Wright	0.078	0.109	0.133	<b>0.136</b>	<b>0.129</b>	0.134	<b>0.224</b>	<b>0.270</b>	<b>0.277</b>	<b>0.254</b>	
$\Delta$ CRPS	Basic BVAR-SV	0.034	0.042	0.057	0.062	0.064	0.063	0.082	0.099	0.107	0.107
	Alt. Prior 1	0.033	0.042	0.057	0.062	0.065	<b>0.064</b>	0.077	0.097	0.103	0.095
	Alt. Prior 2	<b>0.034</b>	<b>0.043</b>	<b>0.059</b>	0.063	0.066	0.063	0.083	0.101	<b>0.117</b>	0.118
	Alt. Prior 3	0.033	0.041	0.056	0.061	0.062	0.062	0.077	0.086	0.087	0.074
	TVPSV	0.029	0.022	0.028	0.029	0.033	0.058	0.063	0.072	0.085	0.089
Wright	0.033	0.041	0.056	<b>0.063</b>	<b>0.067</b>	0.062	<b>0.089</b>	<b>0.106</b>	0.117	<b>0.120</b>	
INF											
$\Delta$ RMSE	Basic BVAR-SV	0.128	0.081	0.039	<b>0.022</b>	<b>0.027</b>	-0.047	-0.031	-0.038	-0.024	-0.025
	Alt. Prior 1	0.129	0.081	0.038	0.022	0.023	-0.046	-0.035	-0.044	-0.026	-0.027
	Alt. Prior 2	<b>0.130</b>	0.080	<b>0.042</b>	0.020	0.026	<b>-0.045</b>	-0.029	-0.037	-0.021	-0.024
	Alt. Prior 3	0.129	0.080	0.040	0.018	0.024	-0.047	-0.031	-0.039	-0.024	-0.026
	TVPSV	0.079	0.055	0.017	-0.003	0.002	-0.096	<b>-0.022</b>	<b>-0.015</b>	<b>-0.001</b>	<b>-0.002</b>
Wright	0.126	<b>0.085</b>	0.034	0.000	0.009	-0.049	-0.034	-0.045	-0.040	-0.042	
$\Delta$ CRPS	Basic BVAR-SV	0.076	0.045	0.017	0.013	<b>0.015</b>	-0.014	-0.012	-0.015	-0.008	-0.010
	Alt. Prior 1	<b>0.081</b>	0.044	0.016	<b>0.014</b>	0.014	<b>-0.011</b>	-0.016	-0.017	-0.008	-0.011
	Alt. Prior 2	0.078	0.044	<b>0.018</b>	0.012	0.015	-0.012	-0.012	-0.014	-0.008	-0.009
	Alt. Prior 3	0.077	0.044	0.017	0.011	0.014	-0.014	-0.013	-0.015	-0.009	-0.010
	TVPSV	0.050	0.030	0.008	-0.002	0.002	-0.039	<b>-0.007</b>	<b>-0.001</b>	<b>0.001</b>	<b>-0.001</b>
Wright	0.072	<b>0.047</b>	0.013	0.000	0.005	-0.018	-0.013	-0.018	-0.017	-0.017	
TBI											
$\Delta$ RMSE	Basic BVAR-SV	0.273	0.242	0.203	0.174	0.141	0.333	0.299	0.256	0.225	0.190
	Alt. Prior 1	0.271	<b>0.244</b>	<b>0.210</b>	<b>0.181</b>	0.141	0.330	<b>0.302</b>	<b>0.264</b>	0.236	0.193
	Alt. Prior 2	<b>0.274</b>	0.242	0.203	0.173	0.140	<b>0.334</b>	0.298	0.257	0.227	0.191
	Alt. Prior 3	0.274	0.243	0.206	0.178	0.142	0.334	0.299	0.259	0.227	0.190
	TVPSV	0.216	0.198	0.181	0.172	0.136	0.276	0.268	0.253	<b>0.240</b>	<b>0.201</b>
Wright	0.249	0.234	0.207	0.180	<b>0.145</b>	0.309	0.292	0.263	0.234	0.198	
$\Delta$ CRPS	Basic BVAR-SV	0.144	0.147	0.132	0.114	0.098	0.173	0.180	0.166	0.149	0.132
	Alt. Prior 1	<b>0.144</b>	<b>0.148</b>	<b>0.134</b>	0.117	0.100	<b>0.174</b>	0.180	0.166	0.153	0.135
	Alt. Prior 2	0.144	0.148	0.132	0.115	0.099	0.173	<b>0.180</b>	0.167	0.153	0.134
	Alt. Prior 3	0.144	0.148	0.132	0.115	0.099	0.173	0.179	0.166	0.149	0.132
	TVPSV	0.113	0.120	0.116	0.108	0.092	0.142	0.160	0.159	0.152	0.135
Wright	0.133	0.143	0.133	<b>0.118</b>	<b>0.104</b>	0.162	0.176	<b>0.167</b>	<b>0.153</b>	<b>0.139</b>	

Table 5: Performance improvements through tilting, for various baseline specifications. For each specification, the table reports the (RMSE or CRPS) difference between the raw and tilted BVAR. Positive numbers indicate that tilting improves forecast performance. Results refer to the “small m/v” variant of tilting and the complete sample period (1988:Q3 to 2013:Q2). The specification for which tilting yields the largest improvement is printed in bold.

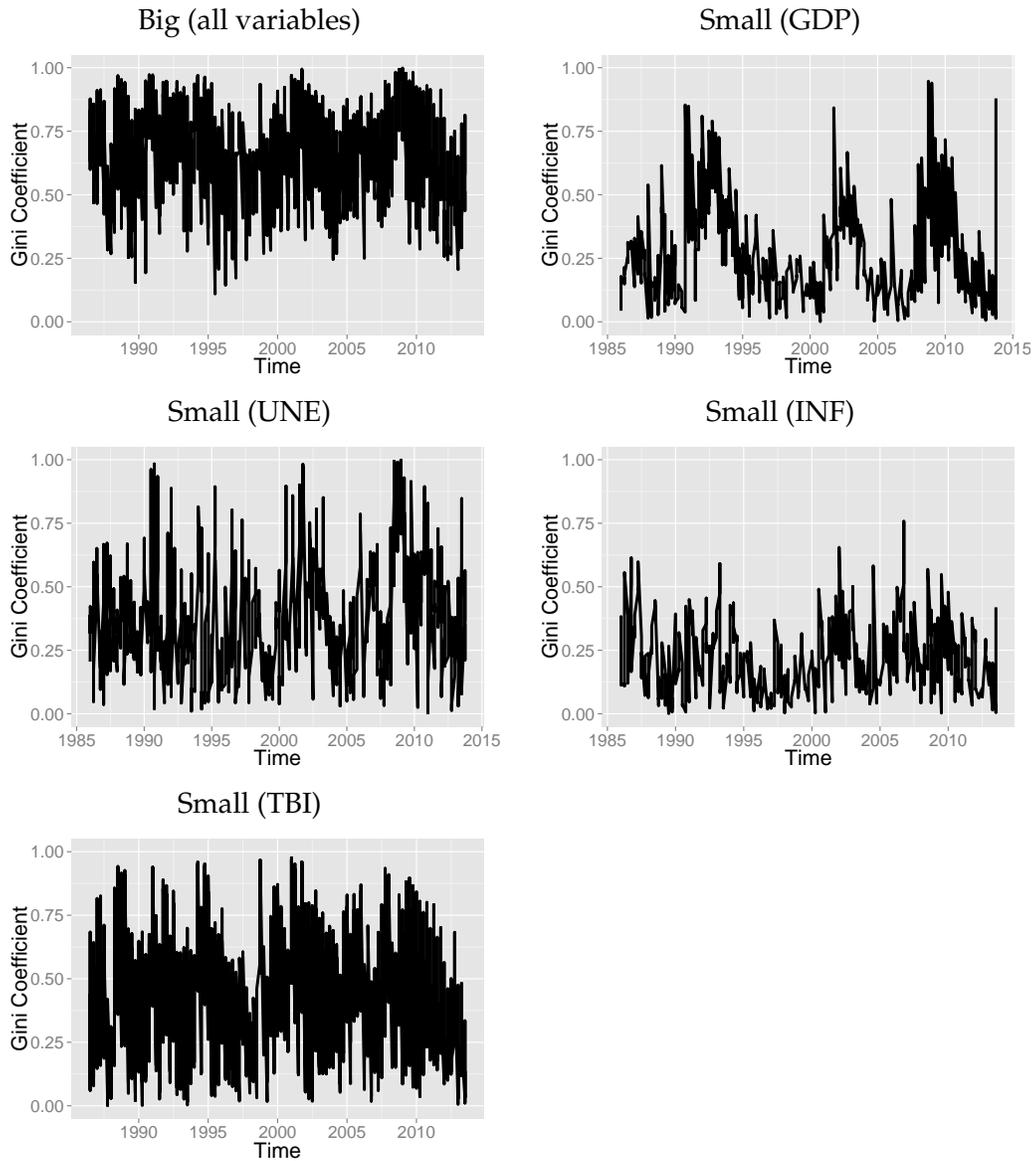


Figure 1: Gini coefficients plotted against time (forecast origin date), for tilting based on model nowcast means and variances (small  $m/v$ ).

shows that the differences are very small, and mostly (but not always) in favor of the larger sample size.

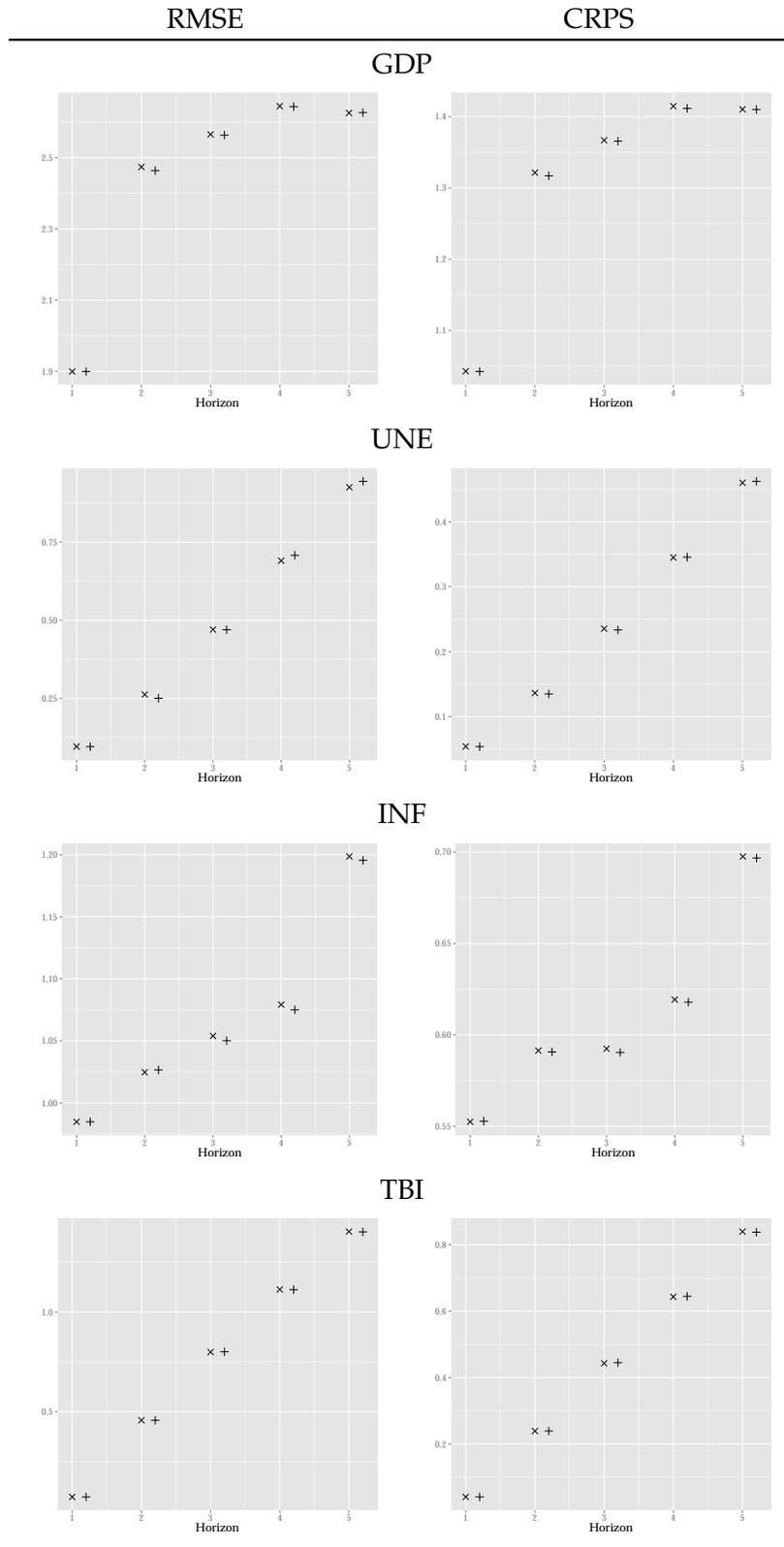


Figure 2: Average RMSE and CRPS for the small sample (5,000 draws, represented by crosses, "x") and the large sample ( $5 \times 5,000$  draws, represented by plus signs, "+"), for the complete evaluation period (1988:Q3 to 2013:Q2). Tilting is based on model nowcasts and the small m/v variant.

## References

- CARRIERO, A., T. E. CLARK, AND M. MARCELLINO (2014): "Real-Time Nowcasting with a Bayesian Mixed Frequency Model with Stochastic Volatility," *Journal of the Royal Statistical Society: Series A*, forthcoming.
- CLARK, T. E. (2011): "Real-Time Density Forecasts From Bayesian Vector Autoregressions with Stochastic Volatility," *Journal of Business & Economic Statistics*, 29, 327–341.
- CLARK, T. E., AND F. RAVAZZOLO (2015): "The Macroeconomic Forecasting Performance of Autoregressive Models with Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30, 551–575.
- COGLEY, T., AND T. J. SARGENT (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262–302.
- D'AGOSTINO, A., L. GAMBETTI, AND D. GIANNONE (2013): "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics*, 28, 82–101.
- GEWEKE, J., AND G. AMISANO (2011): "Optimal Prediction Pools," *Journal of Econometrics*, 164, 130–141.
- PRIMICERI, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.
- VILLANI, M. (2009): "Steady-State Priors for Vector Autoregressions," *Journal of Applied Econometrics*, 24, 630–650.
- WRIGHT, J. H. (2013): "Evaluating Real-Time Forecasts with an Informative Democratic Prior," *Journal of Applied Econometrics*, 28, 762–776.

**Centre for Applied Macro - and Petroleum economics (CAMP)**  
will bring together economists working on applied macroeconomic issues, with special emphasis on petroleum economics.

BI Norwegian Business School  
Centre for Applied Macro - Petroleum economics (CAMP)  
N-0442 Oslo

<http://www.bi.no/camp>