Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts

Fabian Krüger, Todd E. Clark & Francesco Ravazzolo

Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts

Fabian KRÜGER

Heidelberg Institute for Theoretical Studies, Computational Statistics Group, D-69118, Heidelberg, Germany
(*fabian.krueger@h-its.org*)

Todd E. CLARK

Economic Research Department, Federal Reserve Bank of Cleveland, P.O. Box 6387, Cleveland, OH 44101
(*todd.clark@clev.frb.org*)

Francesco RAVAZZOLO

Norges Bank and BI Norwegian Business School, Bankplassen 2, 0267, Oslo, Norway
(*francesco.ravazzolo@norges-bank.no*)

This article 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.

KEY WORDS: Bayesian analysis; Forecasting; Prediction.

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 (2013b) and Banbura et al. (2013a). 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) used 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) used 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) treated 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) combined 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) used 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) developed 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 article, 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) used entropic tilting to combine survey-based forecasts of short-term interest rates with yield curve forecasts from econometric models, and Lewis and Whiteman (2015) used 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 $f_1, \ldots, f_n$ 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 three-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 (2015), 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 article 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 online Appendix provides details of our priors and estimation algorithms and presents some analytical derivations mentioned in the article. Furthermore, it 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 three-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 (2015): 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 three-month and six-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 by 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) referred 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 consider-

ations 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 precrisis sample of 1988:Q3–2007:Q4, in results presented primarily in the 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 one-quarter ahead forecast is a current quarter (t) forecast, while the two-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 1948–1954 period to set the priors on some parameters, as detailed in the online 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 $t + h + 2$ dataset. 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 online 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, $i = 1, \ldots, p$, 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 $p = 4$:

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

where A is a lower triangular matrix with ones on the diagonal and nonzero coefficients below the diagonal, and the diagonal matrix Λ_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; that is, Φ is not restricted to be diagonal.

To estimate this BVAR, we use a Gibbs sampler, detailed in the online Appendix. Stochastic volatility is estimated with the algorithm of Kim, Shephard, and Chib (1998), as detailed by 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 5000 posterior draws of the model parameters, obtained by retaining every eighth draw of a total sample of 40,000 post-burn draws, with a burn period of 5000 draws.

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, Λ_t up through T, and Φ), we: (1) draw innovations to log volatilities for periods $T + 1$ through $T + H$ from a multivariate normal distribution with variance–covariance matrix Φ and compute $\lambda_{T+1}, \ldots, \lambda_{T+H}$; (2) draw innovations to y_{T+h}, $h = 1, \ldots, H$, 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}, $h = 1, \ldots, H$. 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 (2015) 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:

$$y_t = X_t' \beta + v_t,$$
$$v_t = \lambda_t^{0.5} \epsilon_t, \ \epsilon_t \sim \text{iid } N(0, 1), \quad (2)$$
$$\log(\lambda_t) = \log(\lambda_{t-1}) + \nu_t, \ \nu_t \sim \text{iid } N(0, \phi),$$

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 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 (2015).

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 et al. 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 three-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 three-month T-bill rate in t and the three-month and six-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 six-month rate because, under the expectations hypothesis, it should contain information about the expected future path of the three-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 $g : \mathbb{R}^p \to \mathbb{R}^m$ 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 reweighting the draws from f in an admissible way (such that the weights are positive and sum to one). Furthermore,

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

is the Kullback–Leibler divergence between the candidate distribution $\tilde{f}$ (which places weight $\tilde{\pi}_i$ on the ith 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

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

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

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 γ (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 γ 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 Article

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 $\left[y_{t+1}^{(k)} \dots y_{t+5}^{(k)} \right]$ 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 $\left[y_{t+1}^{(1)} \dots y_{t+5}^{(4)} \right]$ 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.
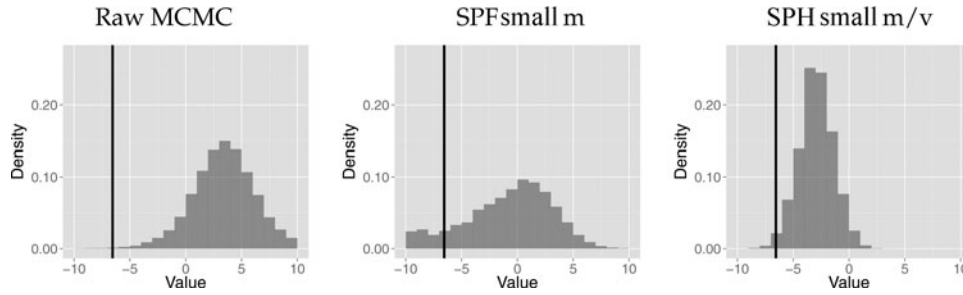
Figure 1. Histograms for raw and tilted samples. In each panel, the black vertical line shows the ex post outcome of -6.55.

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

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

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

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

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 ith 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 reweighting 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.

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 nonparametric 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 $\theta = [\, \theta_1 \ldots \theta_5 \,]'$ and Σ 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^*
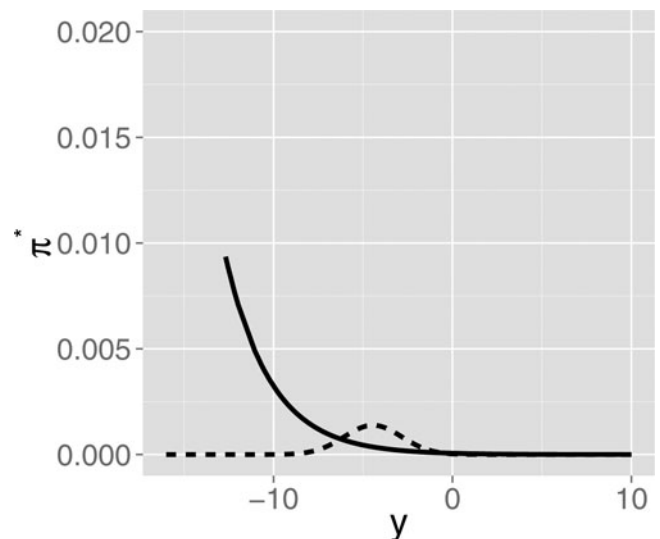

Figure 2. Tilting weight $\pi_{t,i}^*$, as a function of the first element y_t of the vector $\mathbf{y}_{t:t+4}$. The solid line corresponds to tilting toward the nowcast mean only; the dashed line corresponds to tilting toward the nowcast mean and variance.
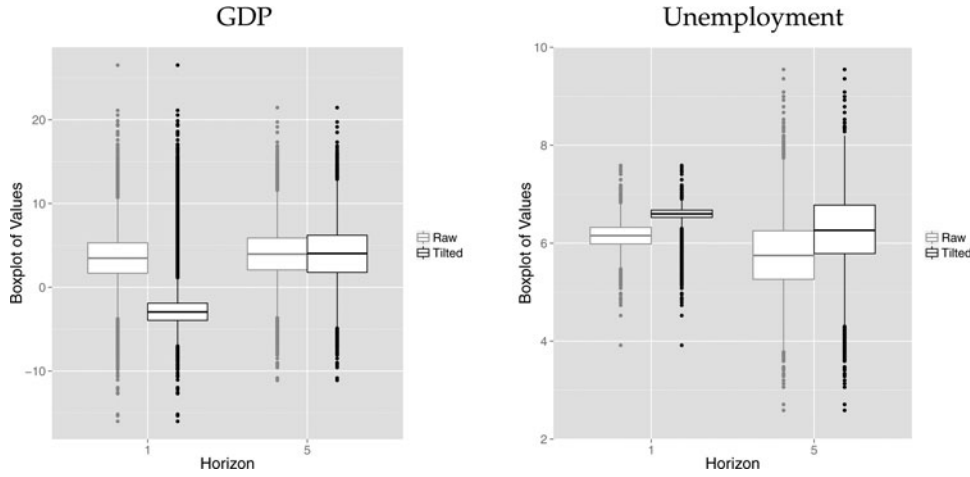
Figure 3. Vertical axis: boxplots of raw and tilted forecast distributions at origin date 2008:Q4. Horizontal axis: forecast horizon h. Boxes range from the 25% to the 75% quantile of a forecast distribution; the end of the upper vertical line is the 75% quantile plus 1.5 times the interquartile range. Forecast draws exceeding that value are plotted as points.

which imposes that the first system variable have mean μ_1 and variance $\Omega_{1,1}$. Then, f^* is multivariate normal $\mathcal{N}(\mu, \Omega)$, with parameters

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

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

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

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 untilted variance. Interestingly, the same solution obtains when targeting a mean of θ_1 only, without making a tilting assumption about $\Omega_{1,1}$—see, for example, Altavilla, Giacomini, and Ragusa (2013, sec. 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 nonparametric 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 through

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

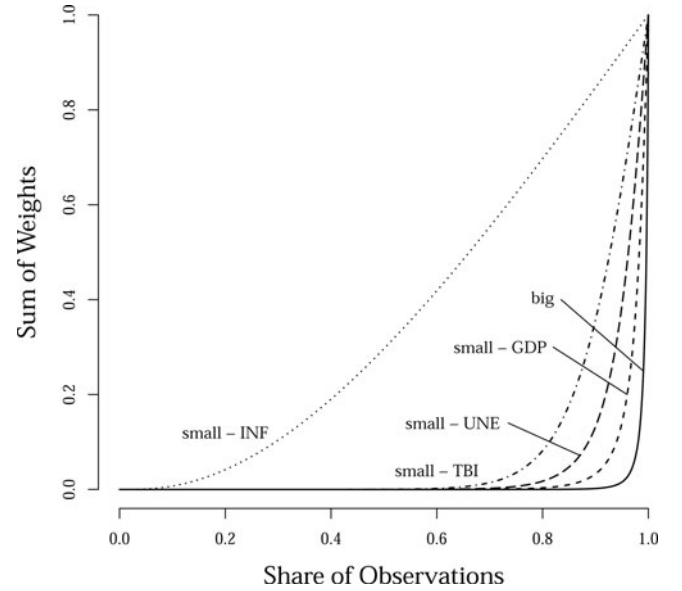| Method	Variable	KLIC divergence from equal weights
Big m/v	(all)	24993.831
Small m/v	GDP	24992.512
	UNE	24991.846
	INF	24990.030
	TBI	24991.320

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).

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

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

Small tilting for variable k solves

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

Note 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

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

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. 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 un-

equal, with a fairly small number of influential MCMC draws (e.g., the 50 largest weights sum to 0.35). The weights for the small tilting problems are much more equal, that is, 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 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.

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

$$\text{CRPS}_t(y^o_{t+h}) = \int_{-\infty}^{\infty} \left(F(z) - 1\{y^o_{t+h} \leq z\} \right)^2 dz, \quad (13)$$

where y^o_{t+h} denotes the observed outcome, F denotes the cumulative distribution function associated with the (posterior) predictive density f, and $1\{y^o_{t+h} \leq z\}$ denotes an indicator function taking value 1 if $y^o_{t+h} \leq z$ 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 nonequal weights—to compute the CRPS in both cases.

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)

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

NOTE: SPF and BMF use data up to daily frequency; BVAR is based on quarterly data.

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) 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).[1] 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 $\left[y_{t+1}^{(k)} \ldots y_{t+5}^{(k)} \right]$ 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 model-based predictive distribution 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 do not 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.

- 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 2015). 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

---

[1]At the one-step horizon, the tilted forecasts are, by construction, essentially the nowcasts, so the benchmark BVAR 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) proposed a modified test of EPA necessary for application to conditional forecasts.

to the use of intraquarter information about the predictand (see Montgomery et al. 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.

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 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 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 outperforming 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 (see 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% 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 5%.

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% (recall that the nominal level is 70%). 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 approaches are already well below 70%, 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

Table 3. Empirical results for entropic tilting, complete sample (1988:Q3 – 2013:Q2)

Forecast horizon		1	2	3	4	5	4*	5*
				GDP				
RMSE	Raw	2.390	2.589	2.656	2.640	**2.617**	2.084	2.140
	SPF small m	**1.591****	2.413*	2.540**	2.650	2.651	1.704*	2.076
	SPF small m/v	1.591**	**2.390***	**2.511***	2.637	2.631	**1.688***	**2.052**
	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*	**2.632**	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	**1.408**	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	**0.906****	**1.280****	**1.335****	1.411	1.411	**0.912****	**1.113**
	SPF big m	0.993**	1.326	1.379	1.422	**1.398**	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**	**0.239**	**0.470**	0.716	0.958		
	BMF small m/v	0.096**	0.251	0.470	**0.708**	**0.945**		
	BMF big m	0.095**	0.247	0.476	0.725	0.966		
	BMF big m/v	**0.092****	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*	**0.346**	**0.462**		
	BMF big m	0.072**	0.147*	0.241*	0.355*	0.474*		
	BMF big m/v	**0.053****	**0.131****	**0.231****	0.346*	0.466*		
				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	**0.807****	**0.906****	**0.966**	**1.003**	**1.128**	**0.617****	**0.731**
	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	**0.462****	**0.532****	**0.555**	**0.582**	**0.663**	**0.361****	**0.431**
	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

(Continued on next page)

Table 3. Empirical results for entropic tilting, complete sample (1988:Q3 – 2013:Q2) (*Continued*)

Forecast horizon		1	2	3	4	5	4*	5*
				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**	**0.791****	**1.092****	**1.372****		
	BMF big m/v	**0.070****	**0.447****	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	**0.041****	0.239**	**0.445****	0.645**	0.838**		
	BMF big m	0.165**	0.331**	0.499**	0.672**	0.845**		
	BMF big m/v	0.047**	**0.238****	0.447**	**0.636****	**0.822****		

NOTES: "RMSE" rows contain root mean squared errors. "CRPS" rows contain mean cumulative ranked probability scores. <u>raw</u>—MCMC output of BVAR-SV model. Alternative tilting targets: <u>SPF small m</u>—SPF mean nowcast for the same variable. <u>SPF small m/v</u>—SPF nowcast mean and variance for the same variable. <u>SPF big m</u>—SPF nowcast means for all four variables. <u>SPF big m/v</u>—SPF nowcast means and variances for all four variables. <u>BMF small m</u>, <u>BMF small m/v</u>, <u>BMF big m,</u> and <u>BMF big m/v</u> 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).

assessment is consistent with the fact that the CRPS—which can be seen as a trade-off between sharpness and correct coverage, see, for example, Gneiting, Balabdaoui, and Raftery (2007)—consistently favors m/v over m.

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, that is, $y_t = \phi y_{t-1} + \epsilon_t, \epsilon_t \sim N(0, \sigma^2)$. 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 μ_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 μ_1, without any uncertainty around it. Using the more general formulas of Section 4.4, it is eas-

ily checked (see Section 5 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 $t + 1, \ldots, t + 5$ 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 μ_1.

6.2 Soft Conditioning (Waggoner and Zha 1999)

Waggoner and Zha (1999) considered 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 article) 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 $[\hat{Y}_{t,1} \pm 1.96\sqrt{\hat{\sigma}_{t,1}^2}]$, 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 5000 in about 97% of all cases.

Table 4. Impact of accounting for nowcast uncertainty, complete sample (1988:Q3–2013:Q2)

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	**0.906**	**1.280**	**1.335**	**1.411**	1.411	**0.912**	**1.113**
	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	**1.410**	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	**0.054**	**0.135**	**0.233**	**0.346**	**0.462**		
				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	**0.463**	**0.534**	**0.558**	**0.597**	**0.672**	**0.369**	**0.438**
	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	**0.041**	**0.239**	**0.445**	**0.645**	**0.838**		

NOTES: Coverage and Length refer to central prediction intervals with a nominal level of 70 %. (reported length is on average over time). The CRPS scores are identical to those in Table 3, and are reprinted here for ease of reference.

Table 5. Root mean squared errors and CRPS scores (the smaller, the better) for the complete sample (1988:Q3–2013:Q2)

		SPF Nowcasts					BMF Nowcasts				
	Horizon	1	2	3	4	5	1	2	3	4	5
						GDP					
RMSE	Tilting	**1.591**	**2.390**	**2.511**	2.637	2.631	**1.899**	**2.464**	**2.563**	2.643	2.627
	Soft Conditioning	1.649	2.433	2.550	**2.633**	**2.622**	2.037	2.514	2.608	**2.637**	**2.620**
	Jumping-off		2.403	2.522	2.649	2.637		2.469	2.568	2.647	2.629
CRPS	Tilting	**0.906**	**1.280**	**1.335**	1.411	1.411	**1.042**	**1.317**	**1.366**	1.412	1.410
	Soft Conditioning	0.972	1.302	1.359	**1.408**	**1.407**	1.125	1.343	1.390	**1.407**	**1.406**
	Jumping-off		1.305	1.352	1.421	1.419		1.327	1.375	1.419	1.415
						UNE					
RMSE	Tilting	**0.151**	**0.350**	**0.573**	**0.808**	**1.024**	0.096	0.251	0.470	0.708	0.945
	Soft Conditioning	0.152	0.355	0.585	0.823	1.039	**0.079**	0.257	0.487	0.728	0.962
	Jumping-off		0.351	0.574	0.810	1.027		**0.251**	**0.467**	**0.704**	**0.940**
CRPS	Tilting	**0.083**	**0.175**	**0.276**	**0.391**	**0.505**	0.054	0.135	**0.233**	**0.346**	**0.462**
	Soft Conditioning	0.086	0.176	0.281	0.399	0.515	**0.046**	**0.134**	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	**0.809**	**0.915**	**0.973**	1.029	**1.144**	0.985	1.027	1.050	1.075	1.196
	Soft Conditioning	0.853	0.954	0.984	**1.028**	1.153	**0.944**	**1.008**	**1.025**	**1.055**	**1.180**
	Jumping-off		0.915	0.974	1.030	1.144		1.025	1.050	1.074	1.195
CRPS	Tilting	**0.463**	**0.534**	**0.558**	0.597	**0.672**	0.553	0.591	0.590	0.618	0.697
	Soft Conditioning	0.501	0.555	0.561	**0.595**	0.676	**0.545**	**0.584**	**0.578**	**0.607**	**0.689**
	Jumping-off		0.534	0.560	0.600	0.675		0.593	0.592	0.620	0.699
						TBI					
RMSE	Tilting	**0.132**	**0.514**	**0.854**	**1.162**	**1.450**	**0.073**	**0.457**	**0.800**	**1.112**	**1.402**
	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	**0.070**	**0.272**	**0.479**	**0.680**	**0.872**	**0.041**	**0.239**	**0.445**	**0.645**	**0.838**
	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

NOTES: 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.
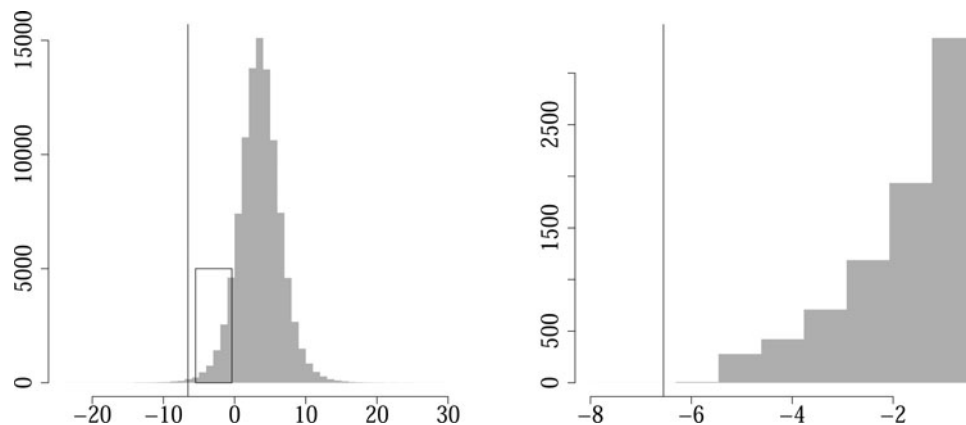

Figure 5. Left panel: histogram of all 100,000 BVAR draws for 2008:Q4 (current quarter forecasts). The 7879 draws within the rectangle satisfy the soft condition imposed by SPF nowcasts. Right panel: zoomed histogram for the draws that satisfy the nowcast condition. In both panels, the vertical line marks the realized value of −6.55.

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 7,879 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 article is concerned with the problem of combining forecasts from a BVAR with nowcasts from other sources. This combination problem is nonstandard, in that the BVAR implies a joint forecast distribution for several forecast horizons, whereas the nowcast information is restricted to 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 article (in addition to results presented in the 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 the 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.

SUPPLEMENTARY MATERIALS

Supplementary Appendix to "Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts": File providing additional results. (PDF file)

REFERENCES

Altavilla, C., Giacomini, R., and Ragusa, G. (2013), "Anchoring the Yield Curve Using Survey Expectations," Working Paper, University College London. [471,476]

Ang, A., Bekaert, G., and Wei, M. (2007), "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" *Journal of Monetary Economics*, 54, 1163–1212. [470]

Banbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013a), "Now-Casting and the Real-Time Data Flow," in *Handbook of Economic Forecasting* (vol. 2), eds. G. Elliott and A. Timmermann, Amsterdam, the Netherlands: Elsevier, pp. 195–237. [470]

Banbura, M., Giannone, D., and Reichlin, L. (2013b), "Nowcasting," in *Oxford Handbook of Economic Forecasting*, eds. M. P. Clements, and D. F. Hendry, Claredon: Oxford University Press, pp. 193–224. [470]

Carriero, A., Clark, T.E., and Marcellino, M. (2015), "Real-Time Nowcasting With a Bayesian Mixed Frequency Model With Stochastic Volatility," *Journal of the Royal Statistical Society,* Series A, 178, 837–862. [471,473,478]

Clark, T.E. (2011), "Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility," *Journal of Business & Economic Statistics*, 29, 327–341. [471,473]

Clark, T.E., and McCracken, M.W. (2013), "Advances in Forecast Evaluation," in *Handbook of Economic Forecasting* (vol. 2), eds. G. Elliott, and A. Timmermann, Amsterdam, the Netherlands: Elsevier, pp. 1107–1201. [478]

——— (2014), "Evaluating Conditional Forecasts From Vector Autoregressions," Working Paper 14-13, Federal Reserve Bank of Cleveland. [478]

Clark, T.E., and Ravazzolo, F. (2015), "The Macroeconomic Forecasting Performance of Autoregressive Models With Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30, 551–575. [471,472]

Clements, M.P. (2014), "Forecast Uncertainty–Ex Ante and Ex Post: U.S. Inflation and Output Growth," *Journal of Business & Economic Statistics*, 32, 206–216. [472]

Cogley, T., Morozov, S., and Sargent, T.J. (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. [471]

Croushore, D. (2006), "Forecasting With Real-Time Macroeconomic Data," *Handbook of Economic Forecasting*, 1, 961–982. [472]

Croushore, D., and Stark, T. (2001), "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, 105, 111–130. [472]

D'Agostino, A., Gambetti, L., and Giannone, D. (2013), "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics*, 28, 82–101. [471]

Del Negro, M., and Primiceri, G.E. (2015), "Time-Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *Review of Economic Studies*, 82, 1342–1345. [473]

Del Negro, M., and Schorfheide, F. (2013), "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting* (vol. 2), eds. G. Elliott and A. Timmermann, Amsterdam, the Netherlands: Elsevier, pp. 57–140. [470]

Diebold, F.X., and Mariano, R.S. (1995), "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13, 253–263. [478]

Doan, T., Litterman, R., and Sims, C. (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100. [470]

Faust, J., and Wright, J.H. (2009), "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset," *Journal of Business & Economic Statistics*, 27, 468–479. [470,472,476]

——— (2013), "Forecasting Inflation," in *Handbook of Economic Forecasting* (vol. 2), eds. G. Elliott, and A. Timmermann, Amsterdam, the Netherlands: Elsevier, pp. 2–56. [470,476]

Frey, C., and Mokinski, F. (2015), "Forecasting with Bayesian Vector Autoregressions Estimated using Professional Forecasts," *Journal of Applied Econometrics*, forthcoming. [470]

Geweke, J., and Amisano, G. (2011), "Optimal Prediction Pools," *Journal of Econometrics*, 164, 130–141. [471]

Giacomini, R., and White, H. (2006), "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578. [478]

Giannone, D., Monti, F., and Reichlin, L. (2014), "Exploiting the Monthly Data-Flow in Structural Forecasting," Working Paper 1416, Centre for Macroeconomics (CFM). [470]

Gneiting, T., Balabdaoui, F., and Raftery, A.E. (2007), "Probabilistic Forecasts, Calibration and Sharpness," *Journal of the Royal Statistical Society,* Series B, 69, 243–268. [479]

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

Gneiting, T., and Ranjan, R. (2011), "Comparing Density Forecasts Using Threshold-And Quantile-Weighted Scoring Rules," *Journal of Business & Economic Statistics*, 29, 411–422. [477]

——— (2013), "Combining Predictive Distributions," *Electronic Journal of Statistics*, 7, 1747–1782. [471]

Hall, S.G., and Mitchell, J. (2007), "Combining Density Forecasts," *International Journal of Forecasting*, 23, 1–13. [471]

Harvey, D., Leybourne, S., and Newbold, P. (1997), "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281–291. [478]

Hersbach, H. (2000), "Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems," *Weather and Forecasting*, 15, 559–570. [477]

Kim, S., Shephard, N., and Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison With ARCH Models," *Review of Economic Studies*, 65, 361–393. [473]

Lewis, K.F., and Whiteman, C.H. (2015), "Empirical Bayesian Density Forecasting in Iowa and Shrinkage for the Monte Carlo Era," *Journal of Forecasting*, 34, 15–35. [471]

Montgomery, A.L., Zarnowitz, V., Tsay, R.S., and Tiao, G.C. (1998), "Forecasting the U.S. Unemployment Rate," *Journal of the American Statistical Association*, 93, 478–493. [473,478]

Newey, W.K., and West, K.D. (1987), "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [478]

——— (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653. [478]

Primiceri, G.E. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852. [473]

R Core Team (2015), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. [474,484]

Robertson, J.C., Tallman, E.W., and Whiteman, C.H. (2005), "Forecasting Using Relative Entropy," *Journal of Money, Credit and Banking*, 37, 383–401. [471,474,475]

Romer, C.D., and Romer, D.H. (2000), "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90, 429–457. [472]

Schorfheide, F., and Song, D. (2015), "Real-Time Forecasting With a Mixed-Frequency VAR," *Journal of Business & Economic Statistics*, 33, 366–380. [470]

Sims, C.A. (2002), "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity*, 2002, 1–40. [472]

Stone, M. (1961), "The Opinion Pool," *The Annals of Mathematical Statistics*, 32, 1339–1342. [471]

Waggoner, D.F., and Zha, T. (1999), "Conditional Forecasts in Dynamic Multivariate Models," *Review of Economics and Statistics*, 81, 639–651. [481]

West, K.D. (1996), "Asymptotic Inference About Predictive Ability," *Econometrica*, 64, 1067–1084. [478]

Wickham, H. (2009), *ggplot2: Elegant Graphics for Data Analysis*, New York: Springer. [484]

Wolters, M.H. (2015), "Evaluating Point and Density Forecasts of Dsge Models," *Journal of Applied Econometrics*, 30, 74–96. [470]

Zeileis, A. (2004), "Econometric Computing With HC and HAC Covariance Matrix Estimators," *Journal of Statistical Software*, 11,1–17. [478]