working paper

18 09

Combining Survey Long-Run Forecasts and Nowcasts with BVAR Forecasts Using Relative Entropy

Ellis W. Tallman and Saeed Zaman

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 18-09 June 2018

Combining Survey Long-Run Forecasts and Nowcasts with BVAR Forecasts Using Relative Entropy

Ellis W. Tallman and Saeed Zaman

This paper constructs hybrid forecasts that combine both short- and long-term conditioning information from external surveys with forecasts from a standard fixed-coefficient vector autoregression (VAR) model. Specifically, we use relative entropy to tilt one-step ahead and long-horizon VAR forecasts to match the nowcast and long-horizon forecast from the Survey of Professional Forecasters. The results indicate meaningful gains in multi-horizon forecast accuracy relative to model forecasts that do not incorporate long-term survey conditions. The accuracy gains are achieved for a range of variables, including those that are not directly tilted but are affected through spillover effects from tilted variables. The forecast accuracy gains for inflation are substantial, statistically significant, and are competitive with the forecast accuracy from both time-varying VARs and univariate benchmarks. We view our proposal as an indirect approach to accommodating structural change and moving end points.

JEL Classification Codes: E17, C53, C11, C32.

Keywords: Bayesian analysis, relative entropy, survey forecasts, nowcasts, density forecasts, real-time data.

Suggested citation: Tallman, Ellis W., and Saeed Zaman. 2018. "Combining Survey Long-Run Forecasts and Nowcasts with BVAR Forecasts Using Relative Entropy." Federal Reserve Bank of Cleveland, Working Paper no. 18-09. https://doi.org/10.26509/frbc-wp-201809.

Ellis W. Tallman (ellis.w.tallman@clev.frb.org) is at the Federal Reserve Bank of Cleveland. Saeed Zaman (saeed.zaman@clev.frb.org) is at the Federal Reserve Bank of Cleveland and the University of Strathclyde. The authors gratefully acknowledge comments from Gary Koop, Michele Modugno, Aubrey Poon, their discussant Sharada Davidson, and participants at the 11th International Conference on Computational and Financial Econometrics, the Post-Graduate Research Away Day at the University of Strathclyde, and the International Institute of Forecasters 38th International Symposium on Forecasting. They thank Andrea Carriero and Todd Clark for sharing Matlab code estimating large BVAR with stochastic volatility.

1 Introduction

Macroeconomic forecasters often use atheoretical models for forecasting. Banbura, Giannone, and Reichlin (2010) show that large VARs containing more than 100 variables can work effectively, a finding that has contributed to a resurgence in the use of VARs in forecasting and policy analysis by both central banks and private forecasters.[1]

In this paper, we propose a technique to adjust the medium-term and long-horizon forecasts from a VAR toward plausible values proposed by judgmental forecasters. Specifically, we utilize the technique of relative entropy to alter the medium-term to long-horizon VAR forecast to match that of the real-time survey long-horizon forecast. It is well-established that long-run forecasts contained in published surveys of professional forecasters are reasonable because they adjust faster than the unrestricted long-run model forecasts in response to any exogenous and/or underlying shifts in the economy.

Thus far, the use of medium- and large-scale VARs for forecasting and policy analysis has generally been limited to time-invariant parameter VARs because of computational limitations and their ability to generate accurate forecasts.[2] More important, Aastveit, et al. (2016) show that the forecasting accuracy of constant parameter medium-scale VAR is competitive with both a small-scale time-varying parameter BVAR with stochastic volatility (similar to Primiceri 2005), with a medium-scale time-varying BVAR (built along the lines of Koop and Korobilis, 2013) and with a regime-switching VAR (Barnett, Mumtaz, and Theodoridis 2014). This finding lends credibility to the use of medium-scale constant parameter VAR models for forecasting, especially given their computational ease relative to the alternatives.

In constant parameter VAR models, the unrestricted long-run forecasts converge to or close to the ergodic mean of the estimation sample, which at times differs substantially with economists' view of the long-run values for particular variables. Furthermore, less persistent data series such as the growth of the real Gross Domestic Product (GDP) will revert to their sample mean quickly and the implied medium-term to long-term VAR forecasts essentially will be the ergodic mean.[3]

Long-run projections of the survey participants typically adapt more rapidly to secular trend changes, such as demographic trends usually not featured in VAR models, and therefore may be preferable proxies for underlying trend properties. Professional forecasters' expectations for the long run are plausible because survey participants have at their disposal indicators typically not included in the information set fed into the models,

[1]Koop (2013) complements the findings of Banbura, Giannone, and Reichlin (2010). Robertson and Tallman (1999) provide an accessible review of the vector autoregressions and the Bayesian estimation-widely used these days to estimate medium- and large-scale VARs.

[2]Time-varying VARs with stochastic volatility developed in Cogley and Sargent (2005) and Primiceri (2005) are limited to three and four variables. The work of Koop and Korobilis (2013) is an exception, as their approach permits time variation in the coefficients and, at the same time, allows for a larger set of variables without the heavy computational demands. We anticipate that, over time, their approach will gain wider popularity.

[3]Larger VARs (with stationary variables) estimated with a Minnesota prior and relatively fewer observations are much more prone to exhibit this behavior. This is because they require heavier prior shrinkage, leading to reduced influence on the variable from its own lags and from other variables in the system.

such as information about the value of the inflation target or demographic factors. By examining the recent trends in business investment and labor market dynamism, survey participants would be better able to extrapolate the trend in productivity going forward. Similarly, they can examine recent trends in birth rates and aging of the population to better assess the growth in the labor force going forward. The use of forward guidance by the central bank and, more generally, the era of more predictable monetary policy (through central bank communications) since the beginning of the financial crisis are other examples of important information at the disposal of the survey participants. As a result, survey participants (collectively) are likely to have more timely and more informed long-run projections.[4]

In combining the long-horizon survey forecast with the VAR forecast, an important question is: At which horizon should the external values for the long-run properties "bind" on the model? That horizon should be variable specific because some variables are more persistent than others and so their forecast will navigate back to the model-implied steady state at a slower speed than less persistent series. In determining the forecast horizon for tilting the long-horizon forecast for the variable of interest, we propose that the VAR estimate of the persistence of the series (e.g., real GDP) at the forecast origin be used to infer the forecast horizon over which the external long-run forecast takes over.

We generate conditional forecasts with moment conditions that match survey forecasts using the short-term forecast from the survey as the mean condition on a one-step-ahead VAR forecast[5], and the long-horizon forecast from the survey as the mean condition on VAR's future forecast horizon (i.e., variable specific). In a VAR, conditioning or tilting on some future horizon will influence the forecast starting from the jumping-off point all the way to the conditioned forecast horizon.

We use relative entropy for conditional forecasting owing to its ease of use, computational ease, and flexibility; it allows the forecaster to combine appropriately both the mean condition and the modeler's confidence in that mean condition (i.e., variance) as illustrated in Kruger, Clark, and Ravazzolo (2017). This is an important advantage if the interest is in density forecasts in addition to point forecasts.

Our main question of interest is whether we can achieve any meaningful gains in the forecast accuracy of the VAR variables over the forecast horizon of interest to policymakers by forcing the medium-term to long-horizon forecast of the select number of VAR variables to match the published survey forecasts (assumed to be reasonable proxies for the underlying trend). Essentially the forecast of interest is a hybrid forecast consisting of a *survey nowcast*, a *BVAR forecast*, and a *survey long-horizon forecast*.

Our empirical forecast evaluation results provide evidence of notable improvements in both the point and density forecast accuracy of VAR forecasts augmented with long-run survey forecasts (hybrid forecast) for most macroeconomic variables of interest. The accuracy gains are achieved for variables that are directly tilted and importantly also

[4]It is worth noting that even if these indicators were included as part of the model's information set, the time-invariant models such as VARs would extrapolate forward the trends prevailing over the entire estimation sample, which may not necessarily align with recent developments.

[5]There is a long list of papers documenting the usefulness of nowcasts in helping improve the forecasting accuracy for future horizons (e.g., Kruger, Clark, and Ravazzolo 2017; Knotek and Zaman 2017b).

for variables that are indirectly influenced through the spillover effects of the tilted variables. Over the forecast horizon of interest to monetary policymakers (i.e., 1 quarter to 12 quarters ahead), the gains in forecast accuracy are strongest for inflation and the federal funds rate. We infer that the apparent structural break in inflation and the extended experience of the federal funds rate near the zero lower bound (ZLB) are the crucial challenges in VAR models estimated with fixed coefficients. The gains are more notable for the model specifications estimated with a longer sample starting in 1959 compared to a shorter sample estimated from 1985 onward.[6] This is to be expected because a variable is likely to exhibit a greater number of mean shifts over a larger sample period compared to a smaller sample period. So the divergence between a variable's current trend estimate and the mean based on a larger sample period is likely to be larger compared to the divergence based on a recent sample period.

We summarize an additional set of findings as follows. First, we show that the point forecast accuracy of hybrid inflation forecasts from our modeling approach is competitive with univariate benchmark models that research has shown to be hard to beat: a univariate unobserved components model with stochastic volatility (Stock and Watson, 2007) and the random walk model (Atkeson and Ohanian, 2001). The result offers policymakers a practical contribution because among the many frustrations of monetary policymakers is the inability of multivariate models – which allow for feedback effects from policy to the real economy and inflation – to match the forecasting performance of the univariate forecasting models. The application in this paper provides one potential path.

Second, we show that allowing for stochastic volatility (SV) in the fixed-coefficient VAR marginally improves the forecast accuracy of the hybrid forecasts. The fact that we see only small gains from SV is in part because the density forecast from a hybrid approach is already centered around a more accurate mean informed by the survey. The gains of using SV in our empirical exercises appear small.

Third, we show that our hybrid forecasts generated from a fixed-coefficient VAR estimated with post-World War II data with or without stochastic volatility are competitive with the forecasts generated from computationally intensive time-varying VARs (also commonly known as adaptive VARs) that are built to explicitly account for structural changes. For inflation, our approach generates more accurate forecasts and the gains are statistically significant. This result is of practical importance because the use of adaptive VARs thus far is limited to three or four variables, yet there is general interest in the forecasts of many additional variables that our approach can easily handle.

Finally, we show (in the supplementary set of results reported in the Online Appendix) that the forecast accuracy of constant coefficient VARs tilted to match the survey conditions is competitive with alternative econometric approaches that have been developed to handle the issue of mean shifts (e.g., VARs in gaps that model variables as the deviation from their respective trends informed by survey expectations, or steady-state BVARs). The crucial difference between our approach and some of the existing approaches incorporating long-term survey expectations is that our approach influences forecasts' post-model estimation.

[6] In order to get a richer characterization of both the shock and parameter uncertainty, the preference is to use a longer estimation sample.

The paper is structured as follows. The next section discusses the related research. Section 3 describes the data and empirical models. Section 4 compares the survey long-run forecasts with the BVAR model long-run forecasts. Section 5 details our methodology of generating the hybrid forecasts. Section 6 reports the forecasting results. Section 7 concludes. Supplementary sets of results are reported in the companion Online Appendix.

2 Related Research

It is well-known that long-horizon forecasts from time-invariant VARs reflect time-series means and can be implausible proxies for long-run trends.[7] Villani (2009) proposed a fully Bayesian approach to adjust the long-run forecast of VAR models. His methodology for both stationary and co-integrated Bayesian VARs permits the forecaster to specify prior beliefs on the unconditional mean of the variable, including variance around that prior mean (commonly denoted in the literature as a steady-state BVAR). Using a seven-variable steady-state BVAR model estimated with data for the Swedish economy, he illustrated the improved forecast accuracy compared to an unrestricted BVAR.

Wright (2013) uses the steady-state BVAR technology of Villani (2009) to show that prior beliefs on the unconditional mean of the variable informed by a survey's (Blue Chip) long-run forecasts lead to systematic improvements in forecast accuracy for a range of U.S. macroeconomic variables, especially for inflation. The insights in Wright (2013) motivate our study, and the results in this paper echo many of those reported in Wright.[8]

Chan and Koop (2014) extend the steady-state BVAR technology by developing a methodology that allows for detection of changes in the steady states of included variables in an automatic fashion. However, their econometric approach relies on past data and therefore, it is likely to be slower to detect any shifts in variables' mean compared to professional forecasters (e.g., knowledge of forward guidance and extended zero lower bound). In their empirical exercise, they consider five variables (output growth, hours worked, labor share, inflation, and interest rates), and their methodology detects changes in the steady states for inflation and the interest rate when estimated over the sample 1954.Q3 through 2012.Q3. This finding coincides with our results on tilting VAR forecasts

[7]The empirical models that explicitly permit time variation in the dynamic coefficients and intercepts (e.g., Primiceri, 2005; Cogley and Sargent, 2005) or unobserved components models that allow for time-varying latent components (e.g., Stock and Watson, 2007; Tallman and Zaman, 2017) can address shifts in the mean. In this section we focus only on those studies that have applied innovative techniques to deal with the shifts in the mean in time-invariant VARs.

[8]In a related work, Frey and Mokinski (2016) extend the methodology of Wright (2013) by augmenting the vector of dependent variables of a steady-state VAR with their corresponding survey nowcasts and the relationship of the survey nowcasts with all the lagged dependent variables of the model is allowed to deviate from the coefficients capturing the relationship between dependent variables and the lagged dependent variables. The extent of the deviation is operationalized through the prior and is pinned down through the joint estimation of the data and the survey nowcasts. They show that doing so leads to more accurate forecasts compared to Wright (2013), as their approach acts as a form of Bayesian shrinkage technique that helps sharpen the parameter estimates of the unaugmented (i.e., original) steady-state VAR.

to long-term survey conditions. Specifically, our results indicate a notable improvement in forecast accuracy for inflation and the interest rate, largely because these two macroeconomic variables have exhibited sizable shifts in the mean since the 1950s.

Our approach has practical advantages over the steady-state BVAR: (1) our approach does not require the modeler to re-specify and re-parameterize the model, and (2) our approach does not require the VAR to be stationary or co-integrated. Therefore, modelers can continue to use their preferred BVAR/VAR specification and seamlessly integrate the relative entropy approach by simply calling a particular routine that takes as an input the output from the BVAR and the modeler's desired short-term and long-term forecasts informed by judgment, other models, or surveys. A related advantage is that the forecaster can get a sense of the differences between the forecast that imposes conditions compared to the one that does not without the need to re-specify the model. The forecaster can test the implications on the forecasts of a wide spectrum of restrictions without re-estimating the model each time. The technique of relative entropy generates several diagnostics that help the modeler assess the severity of tilting, that is, reflecting the presumed reliability of inferences drawn from the tilted distributions – and the role of the various moment conditions contributing to the severity.[9] In the Online Appendix A3 we report the results of a horserace comparing the forecast accuracy from an approach along the lines of Wright (2013) to our proposed approach.[10] Not surprisingly, both approaches perform comparably in a statistical sense; however, in rare instances, steady-state BVAR forecasts run into convergence issues, which may be an important factor for practical reasons.[11]

Another popular approach to anchor model forecasts to survey expectations is to model variables by first transforming them into a gap form (i.e., deviation from the long-run survey expectations) and then estimating them using a VAR (or a univariate regression for the single variable of interest).[12] The forecasts of the gap coming out of the VAR are then transformed back to the units of interest by adding the last estimate of the survey expectations available as of the forecast origin to construct the corresponding implied forecasts (see Faust and Wright, 2013 in the context of the univariate inflation case; see Clark and McCracken, 2010 and Zaman, 2013 in the case of the VAR).[13] By construction, the implied long-run forecasts from this approach would be close to the

[9]In an application of relative entropy, Cogely, Morozov, and Sargent (2005) use information from Bank of England's inflation fan chart to calibrate or "twist" the predictive densities generated from a time-varying BVAR. They use the diagnostic measures to assess the degree of twisting required in matching BVAR predictive density to the Bank of England's fan chart (which has a high judgmental component) and propose that these diagnostics be used to support the reasoning behind the judgment conditions.

[10]Both of these approaches use the same information set but differ in terms of model specification.

[11]In our specifications, the convergence issue arises because the posterior coefficients of the unemployment rate exhibit a nontrivial probability mass in the nonstationary region. Villani (2009) has highlighted the possibility of convergence issues due to this very reason. The issue can be resolved through the use of tighter priors that push the probability mass away from the nonstationary region and toward the stationary region.

[12]This approach parallels the approach commonly used in DSGE models, where variables are modeled as deviations from their steady states.

[13]The trend estimate (proxied by the survey expectations measure) is assumed to follow a random walk over the forecast horizon.

latest available estimate of the survey expectations plugged in as of the time the forecast is generated. The advantage of this approach is its simplicity and therefore it has gained traction in the literature over the past few years. A key drawback is that it requires a time series of survey expectations as long as the estimation sample (necessary for constructing the transformed gap variable). This issue may be more likely to bind for regions outside the United States and Europe for which publicly available survey forecasts have a shorter history. Another drawback is that forecasts may not necessarily converge precisely to the survey's long-run expectations (e.g., 2% for inflation), and thereby create a communication problem. The presence of an intercept term in the gap equation (say, for the inflation gap) captures the long-run historical deviation of the (inflation) gap from zero within the estimation sample. The estimate of the intercept term will be positive if inflation has exceeded the inflation trend on average during the sample, while it will be negative if inflation has been below trend on average. So an inflation forecast three years out may settle at a level that is lower than the trend estimate informed by the survey expectations (and the modeler's desired level). This issue could be resolved by using a steady-state BVAR (of Villani, 2009) that includes the variables transformed in to gap form and using a very tight prior for the steady-state value of zero (i.e., pushing the constant in the gap equation to zero, which translates into the model-implied long-run forecast to converge to the survey expectations);[14] see Clark (2011).[15] In the Online Appendix A4 we report the results of a horserace comparing the forecast accuracy from our proposed approach to a comparable VAR that models variables in gaps.[16] Not surprisingly, both approaches perform comparably in a statistical sense.

Giannone, Lenza, and Primiceri (2018) propose a prior to influence the joint long-run behavior of the VAR variables modeled in levels (denote the prior PLR). The motivation of PLR is similar to the application of the "sum of coefficients" and "co-integration" priors proposed by Sims and Zha (1998), the combination of which aims to reduce the explanatory power of deterministic components. The PLR prior requires defining the co-integrating relationships among the variables of the VAR model in the specification. This may be tractable for a small dimensional VAR but a daunting task for large VARs. Furthermore, long-run cointegration relationships may not be well established for many emerging countries, yet survey expectations data are now widely available for many countries. The advantage of their approach is that their prior is a conjugate and as such it can be conveniently implemented like other popular priors (Minnesota and sum of coefficients; Sims and Zha, 1998) by appending dummy observations to the data. However, based on our understanding their approach does not allow control over the specific values that the VAR variables converge toward in the long-run. That said, our proposed approach could even be used for a VAR estimated with PLR to push its forecasts toward survey expectations and hence allowing control over the specific equilibrium/steady-state values.

[14] Alternatively, it could be resolved by using the predictive densities of transformed variables from the steady-state BVAR and tilt them (through relative entropy) to the desired steady state (e.g., survey expectations).

[15] Clark (2011) extends a steady-state BVAR to allow for stochastic volatility of the sort entertained in Primiceri (2005) and Cogley and Sargent (2005).

[16] Both of these approaches use the same information set but differ in terms of model specification.

Our paper employs methods most closely associated with Kruger, Clark and Ravazzolo (2017) and Altavilla, Giacomini, and Ragusa (2017). Kruger, Clark and Ravazzolo (2017) apply relative entropy to tilt a one-quarter-ahead BVAR forecast to match the nowcast estimate from the Survey of Professional Forecasters (SPF).[17] They tilt the BVAR forecast to match both the nowcast mean and the nowcast variance (both informed by the SPF). By tilting both the mean and the variance, they find improvements in the density forecast accuracy. They compare the approach of relative entropy to the soft conditioning of Waggoner and Zha (1999), an alternate approach to adding mean and variance conditions to BVAR forecasts. The forecast accuracy for both approaches performs similarly on average, but at times, the soft conditioning approach results in unusual distributions corresponding to quarters for which conditions are imposed. Overall, Kruger, Clark and Ravazzolo (2017) find that relative entropy is the preferred approach because of its flexibility and its proper accounting of nowcast uncertainty. In our paper, we extend Kruger, Clark and Ravazzolo (2017): in addition to tilting the VAR forecast in the nowcast quarter, we tilt the medium- to long-horizon forecast coming out of the VARs to match the long-horizon forecast reported in the external survey of forecasters, the SPF. Our methodology parallels that of Altavilla, Giacomini, and Ragusa (2017), who also use relative entropy to tilt the segments of the yield curve forecasts from the term structure models to match survey expectations. They find that tilting ("anchoring") at the three-month horizon leads to the most improvement in the density forecasts for the entire yield curve.

3 Data and the Empirical Model

3.1 Data

Our empirical examination uses real-time data at a quarterly frequency. We estimate two VAR models (discussed below), a small-scale VAR (denoted Small VAR) consisting of five variables and a medium-sized VAR (denoted Medium VAR) consisting of 10 variables. The Small VAR consists of real GDP, the CPI, the unemployment rate, the federal funds rate, and the credit spread. Both real GDP and the CPI enter in annualized quarterly rates and the remaining three enter in levels. Our use of a Small BVAR is motivated by the fact that several papers on VAR forecasting employ it as a benchmark VAR containing the set of core variables thought to be of interest to central bankers. The Medium VAR adds to the small VAR variables shown to be useful in improving forecasts of the core variables. Forecasts of the additional variables may be of their own interest to central bankers, for example, productivity growth and wage inflation measures. Specifically, these additional five variables include real personal consumption expenditures, nonfarm business productivity, the employment cost index-wage and salary of private workers (ECI), nonfarm payroll employment, and the core CPI (i.e., the CPI excluding food and energy). All these variables are transformed to quarterly annualized growth rates. To compute the growth rates, we use 400 times the log difference formula. The unemploy-

[17]They also perform a separate exercise in which they tilt VAR forecasts to nowcasts from external nowcasting models. For inflation and GDP, they find that the SPF nowcasts are hard to outperform.

ment rate and two financial variables are defined in units of percentage points. As a result, both our VAR models are stationary VARs.

We use the Federal Reserve Bank of Philadelphia's real-time data set for macroeconomists and the Federal Reserve Bank of Saint Louis's ALFRED database to construct our real-time quarterly data set. The quarterly macroeconomic variables for which we collect data are: real GDP growth, real personal consumption growth, inflation in the consumer price index (CPI), inflation in the consumer price index excluding food and energy prices (core CPI), growth in productivity of the nonfarm business sector, growth in compensation (employment cost index: compensation of private-industry workers), growth in payroll employment, and the level of the unemployment rate. We also collect quarterly financial variables, which are realtime by construction and are downloadable from Haver Analytics. The financial variables we collect include the level of the federal funds rate and the credit spread, defined as the difference between the yield on ten-year Treasury bonds and the yield on BAA bonds.

All vintages of real-time quarterly data coincide with the survey data of the SPF, which is a quarterly survey released approximately in the middle of the middle month of the quarter. The real-time vintages start in 1994.Q1 and end in 2016.Q4. Each real-time vintage begins in 1959.Q4 The results in the main part of the paper focus on models estimated beginning in 1959.Q4. In the Online Appendix A2, we also report results for models estimated beginning in 1985.Q1.

For the purpose of forecast evaluation, we treat the truth as the third release when evaluating the forecasts from the small-scale VAR and we treat the truth as the latest available vintage (most revised; vintage available as of August 2017) when evaluating the forecasts from the medium-scale VAR. We use Haver Analytics to collect the most revised quarterly data for forecast evaluation. We also collect the survey nowcasts and survey long-horizon forecasts for real GDP growth, CPI inflation, the unemployment rate, and federal funds rate from the SPF, and the Blue Chip Economic Indicators (and Blue Chip Financial Forecasts). For illustrative purposes, we also collect the long-run projections from the Federal Open Market Committee's (FOMC) Summary of Economic Projections (SEP). For the formal forecast evaluation exercises reported in the main text, we use the nowcasts and long-horizon forecasts from the SPF.

3.2 Empirical Bayesian VAR Model

In this paper, all empirical examinations use vector autoregression (VAR) models.

3.2.1 BVAR model

A general representation of a VAR(p) model is as follows:

$$Y_t = A_c + A_1 Y_{t-1} + ... + A_p Y_{t-p} + \varepsilon_t, \qquad \varepsilon_t \sim N\left(0, \Sigma\right) \tag{1}$$

where $t = 1, ...T$, Y_t is an $n \times 1$ vector of n observed variables, $A_c = [c_1, c_2, ...c_n]$ is an $n \times 1$ vector of intercepts, $A_1, ...A_p$ are $n \times n$ matrices of coefficients, ε_t is an $n \times 1$ vector of error terms distributed normally with zero mean and variance matrix, $\Sigma = E\varepsilon_t\varepsilon_t^{'}$. A VAR is a system consisting of n equations and with each equation containing $k = n \times p + 1$ regressors, $n \times k$ parameters need to be estimated.

In our analysis reported in the main text of the paper, we consider two quarterly VAR specifications differing in the number of variables, with n=5, and n=10, respectively. The lag length, p is set to 4. This means that we have to estimate 41 coefficients per equation in our larger VAR (and 21 coefficients in the smaller VAR) with limited number of observations.

We estimate our VARs using Bayesian methods as discussed in Banbura, Giannone, and Reichlin (2010), Beauchemin and Zaman (2011,) Koop (2013), and Carriero, Clark, and Marcellino (2015) to reduce over-fitting. The estimation details and prior settings, including the approach of determining hyper parameters, are relegated to the Technical Appendix.

3.2.2 Generating point forecasts

Using the posterior mean of the coefficient matrix $\overline{A}$ in (7) and the VAR model in (1), the one-step-ahead forecast is computed as

$$\widehat{Y}_{t+1} = \overline{A}_c + \overline{A}_1 Y_t + \overline{A}_2 Y_{t-1} + \ldots + \overline{A}_p Y_{t-p+1} \tag{2}$$

Similarly $h = 2, \ldots, H-1$ step-ahead point forecasts are computed recursively.

$$\widehat{Y}_{t+h} = \overline{A}_c + \overline{A}_1 \widehat{Y}_{t+h-1} + \overline{A}_2 \widehat{Y}_{t+h-2} + \ldots + \overline{A}_p \widehat{Y}_{t+h-p}, \quad h = 1, \ldots, H \tag{3}$$

where $\widehat{Y}_{t+h-p} = Y_{t+h-p}$ for $h \leq p$.

We evaluate point forecasts using the widely used metric of mean squared forecast error (MSE),

$$MSE_{i,h} = \frac{\sum_{t=T_0}^{T_1-h} (Y_{i,t+h} - \widehat{Y}_{i,t+h})^2}{T_1 - h - T_0 + 1} \tag{4}$$

where i corresponds to the macroeconomic (e.g., real GDP growth) or financial (e.g., credit spread) variable of interest, T_0 denotes one quarter prior to the start of the evaluation period (e.g., 1993.Q4), T_1 the end of the evaluation period (2016.Q4 or 2006.Q4), h is the forecast horizon of interest (e.g., $h = 1, ..., H$), $Y_{i,t+h}$ is the actual realization, and $\widehat{Y}_{i,t+h}$ is the forecast.

3.2.3 Generating density forecasts

Our use of a Normal inverse-Wishart (N-IW) prior greatly simplifies the computation of posterior predictive densities from the VAR model in (1). If we let $Y^T = (y_1', \ldots, y_T')'$ be the entire history of the observables Y, then the posterior predictive density is the joint probability density of all possible future paths of the observables Y conditional on the history of the observables and the priors. This posterior predictive density is thus

$$p\left(Y^{T+1,T+H} \mid Y^T\right) = \int p\left(Y^{T+1,T+H}, \theta \mid Y^T\right) d\theta \tag{5}$$

where $p\left(Y^{T+1,T+H}, \theta \mid Y^T\right)$ is the joint posterior density of model parameters and future observables. This integrand can be expressed in a way such that it highlights the two sources of uncertainty embedded in the posterior predictive density:

$$p\left(Y^{T+1,T+H}, \theta \mid Y^T\right) = p\left(Y^{T+1,T+H} \mid Y^T, \theta\right) p\left(\theta \mid Y^T\right), \tag{6}$$

The first term, $p\left(Y^{T+1,T+H} \mid Y^T, \theta\right)$, on the right-hand side represents the uncertainty of future observables treating the observed data and model parameters as given. In other words, this density reflects the forecast uncertainty due to future shocks impacting the VAR model. The second term, $p\left(\theta \mid Y^T\right)$, is the parameter posterior distribution representing the parameter uncertainty within the estimation sample. The multi-step-ahead density forecasts are computed by using a Monte Carlo algorithm to sequentially draw Σ and VAR coefficients A from the respective posterior distributions defined in (6).[18]

We evaluate the performance of density forecasts using the metric continuous ranked probability score (CRPS) as proposed by Gneiting and Raftery (2007). CRPS measures the closeness between the actual realization and the predictive density: the closer the distribution, the smaller is the CRPS value and the more accurate is the predictive density. Accordingly, it is defined as

$$CRPS_t^{i,h}(Y_{i,t+h}) = \int_{-\propto}^{\propto} (F(z) - 1\left\{Y_{i,t+h} \leq z\right\})^2 dz = E_p|Y_{i,t+h}^* - Y_{i,t+h}| - 0.5 E_p|Y_{i,t+h}^* - Y_{i,t+h}^+| \tag{7}$$

where $Y_{i,t+h}$ is the actual realization, F is the cumulative distribution function corresponding to the predictive density f, $1\left\{Y_{i,t+h} \leq z\right\}$ is an indicator function that takes a value of 1 if $Y_{i,t+h} \leq z$ and a value of 0 otherwise, and $Y_{i,t+h}^*$ and $Y_{i,t+h}^+$ are independent random draws from $p\left(Y^{T+1,T+H}, \theta \mid Y^T\right)$.
The CRPS metric favors predictive densities that have higher probabilities near the actual realization. As defined above, a lower CRPS would be preferable to a higher score.

[18] We obtain 20,000 draws of VAR coefficients and error variance-covariance matrix to construct the empirical predictive distribution.

3.2.4 BVAR model with stochastic volatility

Past empirical research has provided overwhelming evidence of time variation in shifts in the volatility (i.e., standard deviation) of many macroeconomic variables. One direct way to capture the potential shifts in the volatility of the variable in a VAR model is to allow the standard deviation of the error terms to drift gradually over time (i.e., stochastic volatility).

Until recently, the use of equation-specific modeling of stochastic volatility in a VAR was restricted to a small VAR consisting of three or four variables. This was mainly due to the infeasibility of the estimation procedure for VARs of bigger dimensions. Specifically, the introduction of equation-specific volatility (i.e., multivariate stochastic volatility) breaks the symmetry between the VAR equations, which invalidates the equation-by-equation estimation of the VAR parameters (loss of Kronecker structure). The estimation can proceed only by first vectorizing the model and then sequentially drawing from a set of conditional posteriors. The difficulty arises when drawing the VAR parameters because to draw from the conditional posterior requires first constructing a variance matrix whose size equals $n^2 \times (lags + 1)$ and consequently the required CPU processing time increases astronomically for models with more than a few variables (see Carriero, Clark, and Marcellino, 2016). This very recent paper shows that this difficulty in estimation can be significantly mitigated by performing a simple triangularization of the VAR model which allows the VAR coefficients to be drawn equation by equation, substantially reducing the computational complexity. Their proposed simple new estimation procedure permits the estimation of VAR models with stochastic volatility of more than 100 variables with just a marginal increase in computational complexity. In addition, their proposed approach could be implemented in an existing algorithm of a constant parameter VAR with just a few additional steps.

Given the minimal additional computational requirement of introducing stochastic volatility in our VAR model(s), we investigate the usefulness of stochastic volatility in our forecasting exercise. Accordingly, we re-estimate our corresponding specifications of constant parameter VAR models augmented to allow for stochastic volatility as in Carriero, Clark, and Marcellino (2016). The appropriate modeling of time variation in the standard deviation of the error terms a priori would be expected to help better characterize the potential time variation in the forecast uncertainty around the point forecasts. In fact, there is a great deal of evidence that allowing for stochastic volatility does lead to improved density forecasts (e.g., Clark 2011; D'Agostino, Giannone, and Gambetti, 2013; Carriero, Clark, and Marcellino, 2016).

The stochastic volatility in the error term of equation 1 takes the following form (as in Carriero, Clark, and Marcellino, 2016; we have tried to keep the same notation for convenience),

$$\varepsilon_t = A^{-1}\Lambda_t^{0.5}\mu_t, \qquad \mu_t \sim i.i.d\ \ N(0, I_n) \tag{8}$$

$$\Sigma_t \equiv Var(\varepsilon_t) = A^{-1}\Lambda_t A^{\acute{-1}} \tag{9}$$

where matrix A is lower triangular with ones on the main diagonal and nonzero elements below it; $\Lambda_t \equiv \ diag(\lambda_{1,t}, ..., \lambda_{j,t})$; $\lambda_{j,t}$ represents the univariate stochastic volatilities that are assumed to evolve according to a geometric random walk

$$ln \ \lambda_{j,t} = ln \ \lambda_{j,t-1} + e_{j,t}, \qquad e_t \sim i.i.d \ N(0, \Phi) \tag{10}$$

The variance-covariance matrix Φ of innovations e_t is assumed to be full rank, i.e., correlation among innovations of different equations is permitted. Equation 10 constitutes the transition equation.
Next, the rescaled VAR error term and its j-th element can be written as,

$$\tilde{\varepsilon_t} = A\varepsilon_t \quad \tilde{\varepsilon_{j,t}} = \lambda_{j,t}^{0.5}\mu_{j,t}, \qquad j = 1, ...n \tag{11}$$

The observation equation is obtained by taking the log of the square of the rescaled VAR error term in equation 11,

$$ln \ \tilde{\varepsilon_{j,t}^2} = ln \ \lambda_{j,t} + ln \ \mu_{j,t}^2 \tag{12}$$

For complete estimation details, please refer to Carriero, Clark, and Marcellino (2016).

4 Real-time Long-horizon survey forecasts versus BVAR forecasts

In the United States, the SPF and the Blue Chip Economic Indicators (BC) are the two most widely known and easily available forecast surveys routinely published. The SPF is a quarterly survey released in the middle of the middle month of the quarter. Each SPF survey release contains the forecasts of the macroeconomic and financial variables for the current quarter (i.e., nowcasts) and forecasts up to four quarters ahead. For the survey carried out in the third quarter of the year, the SPF asks respondents for their estimates of the natural rate of unemployment. Similarly, for the surveys carried out in the first quarter, the SPF asks respondents their projections of long-run values (defined as 10 year annual average) of real GDP growth, the short-term interest rate (i.e., yield on the 3-month Treasury bill), and a few other variables. This is helpful because the SPF contains the projections for all the core set of variables of interest for this paper; for our purposes, we treat the SPF's projections of the natural rate of unemployment as the long-run forecast of the unemployment rate. Following the forecasting literature, for the SPF projections, we use the median projection, and for the BC we use the mean projection.

The BC is a monthly survey released at the end of the first week of the month. Since it is a monthly survey, the forecasts and nowcasts reflect the developments in the intra-quarterly flow of information. Two times a year, i.e., each March and October, the BC also reports the respondents' long run projections (defined as an average for the 7 to 11 years ahead). The BC reports long-run forecasts for all of our core variables of interest:

CPI inflation, real GDP growth, the unemployment rate, and the short-term interest rate. Our forecast evaluation exercises in this paper will use the nowcasts and long-run forecasts from the SPF,[19] but just to give a sense of how the evolution of the long-run forecasts compares across the two surveys, we report them side-by-side.[20]

Figures 1-2 plot the evolution of the macroeconomic forecasts from an unrestricted BVAR model (i.e., Medium VAR), the SPF, and the BC. To facilitate comparison, we list the mean of the estimation sample over which the BVAR model is estimated.

Figure 1 compares the real-time estimates of the long-run forecasts of real GDP growth (left panel) and the unemployment rate (right panel) from professional forecasters to the forecasts from a statistical BVAR model estimated recursively with real-time data. The real-time data we use to estimate our BVAR model would be a subset of the information set available to the professional forecasters. The professional forecasters would likely rely on a larger information set, including judgmental opinions of their own and of the subject matter experts along with their own econometric methodologies to come up with their forecasts.

The BVAR long-run projection for real GDP growth is roughly similar to the ergodic mean of the estimation sample. At the beginning of 1994, both the SPF and the BC were forecasting the underlying growth rate closer to 2.6 percent just four-tenths lower than that of the BVAR. As time rolled forward from 1994 through 1997, professional forecasters edged their estimate of long-run growth lower while the BVAR edged higher. By late 1999, professional forecasters had revised their forecasts back up by a couple of tenths to 2.5 percent compared to the BVAR's 3.4 percent. Moving into 2000, while the BVAR projection held steady at 3.4 percent, professional forecasters strongly revised up their projection by roughly six-tenths, to 3.1 percent. This upward revision was in response to stronger growth data the prior two years averaging more than 4 percent. The upward revision continued through 2005, at which point long-run forecasts by both survey forecasters and the BVAR were in agreement roughly at 3.4 percent. Beginning in 2006 and onward, the survey forecasters gradually lowered their growth forecasts, reaching 2.3 percent by the end of 2016. This rate of growth roughly matches the US economy's average growth rate since the start of the post-crisis recovery. At the same time, the BVAR forecast also edged lower but by a lot less magnitude (3.4 to 3.1 percent) compared to professional forecasters (from 3.4 to 2.3 percent).

In the case of the unemployment rate, while the BVAR-implied long-run forecast has fluctuated in a narrow range around 6 percent, the professional forecasters' estimate of the natural rate has evolved in line with the movements in the business cycle. For example, beginning in 1994 through the start of 2001, the estimate trended lower, but at the onset of the 2001 recession, it reversed and began to trend up until the beginning

[19]Our choice of the SPF is motivated in part because it is publicly available. Relatedly, Croushore (2010) documents the good inflation forecasting properties of long-run forecasts of CPI inflation from the SPF and suggests using them as a proxy for inflation expectations.

[20]Our results (not shown) are robust if we instead use nowcasts and long-run forecasts from the BC.

of the recovery. Thereupon it trended lower until the onset of the Great Recession. In response to a large upward spike in the unemployment rate reflecting the severity of the recession and the associated disruptions to the labor market, the professional forecasters rapidly adjusted upward their projections of the natural rate of unemployment. By the end of 2010, the professional forecasters' estimate of the natural rate of unemployment ranged between 5.8 and 6.0 percent, close to that implied by the BVAR. Thereafter, as the recovery picked up pace, professional forecasters adjusted their estimates downward reaching 4.8 percent by the end of 2016.

Figure 2 plots the projections corresponding to nominal variables, CPI inflation (left panel) and the short-term interest rate (right panel). The short-term interest rate projection from the SPF refers to the average yield on the three-month Treasury bill over the next 10 years. In the case of the BC, it refers to the average yield on the three-month Treasury bill between 7 and 11 years ahead. For the BVAR it is the long run projection of the federal funds rate. For CPI inflation, the SPF refers to the 10-year annual average CPI inflation rate, which is reported and available with every quarterly SPF release. The BC refers to 7-to-11 years ahead annual average CPI inflation rate.

Looking at the figure, a few things stand out. First, over our forecast evaluation sample, there is a noticeable downward trend in the projections. For CPI inflation, even though the BVAR projections gradually trend lower, both the SPF and BC projection are relatively stable from 1999 onwards. The BVAR projection trends lower in response to lower inflation readings relative to the sample mean, which includes high inflation readings from the 1970s. That said, by the end of 2016, the BVAR projection is still 1.5 percentage points higher than the SPF. Second, the important distinguishing feature of this chart compared to the previous one (for real variables) is the sizable gap between the forecasts from the professional forecasters and the BVAR forecasts. In the case of CPI inflation, the gap between the two on average is 2 percentage points, whereas in the case of the short-term interest rate, the gap averages 2.5 percentage points. Third, both the level and the evolution of both the SPF and the BC are notably similar.

Overall, these charts provide suggestive evidence that professional forecasters are quick to adjust their expectations. In reality, it is difficult to distinguish between transitory fluctuations and fluctuations associated with changes to the underlying trend. As a result, forecasters gradually learn about shifts in the underlying trend. Even then, their expectations adjust more rapidly than implied by the BVAR estimated with a longer sample of data.

5 Methodology for Tilting Forecasts

5.1 Relative Entropy

The technique of relative entropy, applied by Robertson, Tallman and Whiteman (2005) to economic forecasting, consists of modifying a given predictive distribution to a new predictive distribution such that it satisfies a given set of moment conditions while minimizing the relative entropy between two predictive distributions.

Let's begin with a predictive distribution, $p\left(Y^{T+1,T+H} \mid Y^T\right)$, corresponding to an n-dimensional random variable Y. We take a random sample of D draws $\{Y_i, i=1,...D\}$ from this predictive density. Since the sample is random, the corresponding weights are $\{w_i, i=1,...D\}$. Suppose now that the modeler wants to impose moment conditions contained in the matrix $\overline{g}$ on this predictive distribution $p\left(Y^{T+1,T+H} \mid Y^T\right)$ such that $\sum_{i=1}^{D} w_i p(Y_i^{T+1,T+H}) \neq \bar{g}$, i.e., the mean of the predictive distribution $p(.)$ is not equal to the mean condition required by the modeler (denote it as "new" information). For the predictive distribution to satisfy the new information requires modifying the original weights $\{w_i, i=1,...D\}$. The new weights $\{w_i^*, i=1,...D\}$ that satisfy this new information is equivalent to finding a new predictive distribution that is as close as possible to the original predictive density in the information-criterion sense.

Specifically, the relative entropy or the Kullback-Leibler Information Criterion (KLIC) of w^* to w is

$$K(w^* : w) = \sum_{i=1}^{D} w^* log(\frac{w_i^*}{w_i}) \tag{13}$$

We solve for new weights that minimize $K(w^* : w)$ and satisfy the following constraints

$$w_i^* \geq 0, \ \sum_{i=1}^{D} w_i^* = 1, \ \sum_{i=1}^{D} w_i^* p(Y_i^{T+1,T+H}) = \bar{g} \tag{14}$$

The first and second terms reflect the fact that weights are probabilities and so should be non-negative and sum to one. The third term represents the new moment conditions.

The solution to the above minimization problem using the method of Lagrange is

$$w_i^* = \frac{w_i \ \exp(\gamma' \ p(Y_i^{T+1,T+H}))}{\sum_{i=1}^{D} w_i \ \exp(\gamma' \ p(Y_i^{T+1,T+H}))} \tag{15}$$

where γ is the vector of Lagrange multipliers associated with the constraints. According to this, the original weights w have been "exponentially" tilted to produce the new weights w^*.[21]

[21] This is why the technique of relative entropy is also denoted by exponential tilting. See Kruger, Clark, and Ravazzolo (2017).

The vector of Lagrange multipliers (i.e. tilting parameters) can be obtained as a solution to the following minimization problem,

$$\gamma = arg \ \min_{\tilde{\gamma}} \sum_{i=1}^{D} w_i \ \ \exp(\acute{\tilde{\gamma}} \ [g(Y_i^{T+1,T+H}) - \bar{g}]) \tag{16}$$

Then using the newly computed weights, the updated expectation of other functions of interest can be computed simply as

$$\sum_{i=1}^{D} w_i^* h(Y^{T+1,T+H}) \tag{17}$$

If the interest is in the modified probabilistic density $g(Y^{T+1,T+H})$, which will be the case in our density forecast evaluation exercises, then as discussed in Cogley,Morogov, and Sargent (2005) importance sampling techniques could be used to redraw $Y_i^{T+1,T+H}$ from the original density $p\left(Y^{T+1,T+H} \mid Y^T\right)$ using the newly found weights, w^*, and can be achieved using the multinomial resampling algorithm of Gordon, Salmond, and Smith (1993). The steps of the algorithm (taken from Cogley,Morogov, and Sargent, 2005) are detailed in the Online Appendix A8.

In our forecast exercises, both the short-term and long-run survey forecasts will constitute the moment conditions on the BVAR posterior predictive densities. The median forecasts from the SPF will act as the mean conditions. In the Online Appendix A.5 we report the results of adding second order moment conditions (i.e. variance) around the nowcast mean conditions (similar in spirit to Kruger, Clark, and Ravazzolo, 2017).

In a VAR, conditioning or tilting on some future horizon will influence the forecast starting from the jumping-off point all the way to the conditioned forecast horizon. For example, if we tilt real GDP growth at forecast horizon h=6, then tilting it will likely impact the forecast trajectory from h=1 to h=5 for all the variables.[22] Simultaneously conditioning on multiple variables (in a system such as VAR) would result in forecast trajectories that reflect the cumulative effect of those conditions.

We use relative entropy as opposed to other approaches to conditional forecasting because of its ease of use, computational simplicity, and flexibility. Relative entropy is an effective and flexible conditional forecasting methodology, because it allows us to appropriately combine both the mean condition and the modeler's confidence in that mean condition (i.e., variance) as illustrated in Kruger, Clark, and Ravazzolo (2017). This is an important advantage if the interest is in density forecasts in addition to point forecasts. Furthermore, specifying the mean condition only would not result in the automatic shrinkage of the variance around the mean condition to zero; relative entropy will assume that the variance around that mean condition is the same as the unconditional. This

[22] In the Online Appendix A7 we provide an intuition behind this spillover feature using an analytical Gaussian example.

is an attractive feature since a commonly used VAR approach of conditional forecasting (e.g., Doan, Litterman, and Sims, 1984) would collapse the variance around the mean condition to zero.[23] This feature will be to our benefit for the forecasting exercises, because it does not require us to specify the variances around our survey forecasts used as conditions on the VAR forecasts (i.e., mean conditions).

5.2 Determining the Forecast Horizon for Tilting

In combining the survey long-run projections with the BVAR forecast, the initial inclination would be to combine the survey projections with the BVAR forecast at some very distant horizon. This assumption is valid, since the terminology "long-run" projection by definition suggests many years into the future. That said, several macroeconomic variables (transformed to growth rates) display little persistence and therefore tend to move back rapidly to their respective (unconditional) mean–the unrestricted long-run model forecast. The series characterizing this behavior are denoted as "mean" reverting. Real GDP growth fits into this category. The forecasts of real GDP growth from VAR models typically tend to move back toward the estimated mean within a year, i.e., less than four quarters. On the other extreme are series such as the unemployment rate, which are very persistent. Depending on the starting point, it may take them several years to move back toward the model-implied long run.

Therefore in combining the long-horizon survey forecast with the VAR forecast, an important and appropriate question is: At which horizon should that occur? We propose that the horizon should be variable specific, because some variables are more persistent compared to others and so their unconditional forecasts will evolve to their respective model-implied steady states at varying speeds. Failure to account for variable-specific horizon may result in forecasts that display gyrations from one steady state to another (i.e., from the model-implied long run to the one imposed by the modeler informed by outside information).

Accordingly, we suggest the following approach to help determine the forecast horizon for combination:

At each forecast origin t, retrieve the persistence estimates (i.e., slope parameters), corresponding to variable i from equation i of the VAR system in (1).

$$\rho_{i,t}^{+,BVAR} = \sum_{l=1}^{p} \bar{A}_{i,l}^{(i,i)} \tag{18}$$

[23] Alternative approaches to constructing conditional forecasts include Waggoner and Zha's (1999) soft conditioning which is an extension of Doan, Litterman, and Sims (1984) and the Kalman filter approach as in Banbura, Giannone, and Lenza (2015). Both of the latter two approaches can also allow for both mean and variance conditions. If we instead use the approach of Doan, Litterman and Sims (1984) to impose our conditions, we get very similar results for the point forecast evaluation.

where $\bar{A}_{i,l}^{(i,i)}$ represents the posterior mean estimate of the slope coefficient of variable i in equation i of the VAR system in (1). It reflects an estimate of variable i's persistence conditional on the BVAR system

The corresponding metric that roughly determines the number of quarters it takes to revert back to the BVAR's implied steady state is

$$h_{i,t}^{+,BVAR} = \frac{1}{1-\rho_{i,t}^{+,BVAR}} \tag{19}$$

The horizon, $h_{i,t}^{*}$ at which the survey long-run forecast is combined with the BVAR forecast for variable i is set as

$$h_{i,t}^{*} = \max\left\{P_t^Q, h_{i,t}^{+,BVAR}\right\} \tag{20}$$

where $P_t^Q - 1$ specifies the minimum number of quarters prior to which the long-horizon survey forecast takes over the BVAR forecast. The max operator insures that at least for the $P_t^Q - 2$ number of quarters following the nowcast quarter the hybrid forecast uses the BVAR forecast. In our exercises we set $P_t^Q = 5$ to reflect our preference to have a dynamic and informative forecast in the short to medium term. It is important to stress that the choice of P_t^Q does not influence our results; setting $P_t^Q = 0$ gives us very similar forecasting results because this choice binds only on real GDP but not other variables.[24]

Figure 3 provides a visual impression of the composition of the hybrid forecast derived from our modeling approach for our variables of interest. Specifically, the hybrid forecast is composed of three components: survey nowcast, BVAR forecast, and survey long-horizon forecast. The width of the component gives a sense of the number of quarters the particular component contributes to the construction of the hybrid forecast over our forecast sample. The orange shaded area indicates the survey nowcast plugged at forecast horizon h=1 for all four variables of interest. The pink shaded area corresponds to the BVAR forecast, and the blue shaded area corresponds to the survey long-run forecast. The changing width of the pink bars reflects the differences in the persistence of the variables: the more persistent the variable, the longer it takes for it to reach its long-run value, which in our exercise is the long-run forecast from the survey. Real GDP growth is the least persistent and the value of $h_{GDP,t}^{*}$ has equaled 5 quarters over our forecast evaluation. If we set $P^Q = 0$, the value of $h_{GDP,t}^{*}$ ranges between 3 and 4 quarters. The interest rate and the unemployment rate are on the other extreme; that is, they are very persistent and as such the corresponding values obtained for $h_{interestrate,t}^{*}$ ranged from 11 to 15 quarters (i.e., the survey forecast takes over the starting forecast horizon roughly three years out) and $h_{unemploymentrate,t}^{*}$ ranged from 22 to 26 quarters, respectively, over the forecast evaluation sample. CPI inflation is persistent but much less than the unemployment rate and the federal funds rate, as $h_{CPIinflation,t}^{*}$ ranged between 5 and 14 quarters.

[24]The use of h=5 quarters is consistent with the statement in Clark and McCracken (2008) that the forecast accuracy beyond the first year for the variable of interest is importantly influenced by the model's implied unconditional mean of the variable.

6 Results

6.1 Forecasting Exercise

We measure the forecast accuracy of the point forecasts (posterior means) based on the metric of the mean squared error (MSE) and relative MSE. A lower MSE is preferable to a higher one. And we measure the forecast accuracy of the density forecasts using the metric of continuous ranked probability score (CRPS). A lower CRPS score is preferable to a higher score. We evaluate the predictive accuracy using a recursively expanding window of data. Our main question of interest pertains to whether we achieve meaningful improvements in the forecast accuracy of the variables of interest (to monetary policymakers) by tilting the model-based forecasts to match the modeler's long-run value. In our examination, the modeler's long-run value equates to the (median of the) long-horizon SPF projections.

To answer this we perform a real-time out-of-sample forecasting evaluation over the period 1994.Q1 to 2016.Q4. We generate forecasts using a recursively expanding estimation window. Specifically, we begin by estimating our small BVAR model using real-time data beginning in 1959 through 1993.Q4 and iteratively generate unconditional forecasts one to forty quarters out. Then we re-estimate the model using an additional data point and again generate forecasts up to 40 quarters out. We repeat this recursive exercise until 2016.Q3. That is, the last estimation sample uses data from 1959 to 2016.Q3, and the forecasts span the period 2016.Q4 to 2026.Q3; with data for evaluation available through 2016.Q4, we would be able to evaluate the one-step-ahead forecast only. The forecasts generated through this recursive exercise are denoted raw BVAR forecasts. Next, using the technique of relative entropy we tilt the one-step-ahead model (mean) forecasts generated in the previous step to match the SPF nowcasts for real GDP growth, CPI inflation (and core CPI inflation[25] in the Medium BVAR), the unemployment rate, and the federal funds rate.[26] We denote the resulting tilted forecast (corresponding to all variables) as the baseline forecast. This **baseline forecast** tilts on the **nowcasts only**. Next, we generate another set of forecasts, but this time tilting the model forecasts **on both the nowcasts and the survey long-run projections** for the same variables.[27] We denote this forecast as the **hybrid forecast**. Finally, we evaluate and compare both the point forecast accuracy and density forecast accuracy corresponding to the raw BVAR,

[25]For core CPI inflation the SPF nowcasts aren't available until 2007.Q1. The inflation nowcasting models of Knotek and Zaman (2017a) or Modugno (2013) could be used to produce core CPI inflation nowcasts prior to 2007.Q1. For the sake of consistency and simplicity, we use core CPI nowcasts from the SPF starting in 2007.Q1.

[26]The SPF does not publish nowcasts for the federal funds rate. We follow the simple procedure of Knotek and Zaman (2017b) to construct an estimate. The procedure involves using the available daily reading as of the SPF survey date and using that to fill the missing trading days of the quarter (i.e., daily random walk). The average of the daily readings (which includes the daily data and the random walk forecast) within the quarter constitutes our nowcast estimate.

[27]In our examination, the SPF long-run projection for the 3-month Treasury bill is assumed to be the long-run projection for the federal funds rate. Historically there is a small gap between the two with the federal funds rate averaging roughly 30 basis points higher on quarterly basis compared to the 3-month Treasury bill.

baseline (also "Now Only"), and the hybrid forecasts (also "Now and LR") respectively. We repeat this procedure for the Medium BVAR. In evaluating the forecasts from the Small BVAR we use the third release as the truth, whereas in evaluating forecasts from the Medium BVAR, we use the latest available vintage as the truth. This allows us to check the robustness of our results to different definitions of the truth.

To check the robustness and sensitivity of our results to a forecast evaluation sample that excludes the Great Recession, we also perform forecast evaluation over the period 1994.Q1 to 2006.Q4. The results for this exercise are reported in the Online Appendix A1. The results on the smaller sample echo the themes of the results of the full sample.

To assess the statistical significance of the differences in forecast accuracy between the baseline and the hybrid forecasts, we follow Kruger, Clark, and Ravazzolo (2017) and Altavilla, Giacomini, and Ragusa (2017). We use the Diebold and Mariano (1995) and West (1996) test of equal predictive accuracy for pairwise comparisons of the RMSE using the two-sided tests of standard normal. In computing the test, we use the HAC variance estimator (an input into the test statistic) with the lag $h-1$ truncation parameter and adjust the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); see Clark and McCracken (2013). As emphasized in Kruger, Clark, and Ravazzolo (2017) the use of our test statistics based on standard normal critical values are likely to be on the conservative side and should be treated as an approximation that, in our case, deals with issues such as the nesting of forecasts and conditional forecasting (see Clark and McCracken, 2017).

6.1.1 Results with the Small BVAR model

Table 1 reports the real-time point forecast accuracy of real GDP growth, CPI inflation, the unemployment rate, the federal funds rate, and credit spread from the Small BVAR. The forecast evaluation is performed over the full sample spanning 1994.Q1 to 2016.Q4. The entries reported in the table correspond to the relative mean square for one quarter out (i.e., nowcast quarter) to 32 quarters out. Although our interest is in the forecast accuracy from 1 to 12 quarters out, we also report for horizons further out to give a sense of the persistence in accuracy gains or deteriorations. The top panel reports the ratio of the mean squared error of the baseline forecast relative to the raw unconditional BVAR forecast. The middle panel reports the ratio of the mean squared error of the hybrid forecast relative to the raw unconditional BVAR forecast. And the bottom panel reports the ratio of the mean squared error of the hybrid forecast relative to the mean squared error of the baseline forecast. The latter ratio would help provide a sense of the marginal gains or losses resulting from tilting toward the long-run survey forecasts only.

Beginning with the top panel, the ratios for the most part are smaller than one, suggesting that conditioning the BVAR with nowcasts (informed by the survey) helps improve forecasts for quarters beyond the nowcast quarter, consistent with Faust and Wright (2013; for inflation) and Kruger, Clark, and Ravazzolo (2017). The spillover effects of improved accuracy are longer lasting for persistent variables, CPI inflation, the unemployment rate, and the federal funds rate. For real GDP growth, the gains are relatively short-lived, since by one year out, the ratio is close to one, and by the end of the second year, the accuracy of the baseline forecast is marginally inferior compared to the raw BVAR forecast. These results echo the findings reported in Kruger, Clark, and Ravazzolo (2017). For the nowcast quarter ratios are significantly lower, suggesting that the SPF nowcasts are significantly more accurate than the BVAR model's one-step-ahead forecast. This result is to be expected, because the nowcast estimate incorporates developments in the intra-quarterly data from which the unconditional BVAR forecast abstracts.

In the middle panel, the ratios are less than one, and with the exception of the unemployment rate, the ratios reported are generally smaller than those reported in the top panel. This suggests that tilting the BVAR forecasts to match the survey long-horizon forecasts in addition to the survey nowcasts leads to further improvement in accuracy. A notable difference across the two panels is the significantly improved accuracy in forecast horizons further out. To get a sense of the extent of the gains in accuracy the bottom panel reports the ratio defined as the numbers in the middle panel relative to the numbers reported in the top panel. For example, six quarters out, the hybrid forecast for real GDP growth is on average 23 percent more accurate than the baseline forecast (i.e. forecast that only conditions on the survey nowcast) as indicated by a ratio of 0.77. The gains persist at least three years out. Digging deeper into error evaluation, the improvements in the accuracy of the real GDP growth forecast arise mainly from the Great Recession and the subsequent recovery. Indeed, looking at Figure 4, the forecast trajectories corresponding to the hybrid forecast align with the actual evolution more closely compared to those of the baseline forecast. The baseline forecast calls for stronger long-run projections, and as a result, the recursive baseline forecast trajectories continuously over-predict growth. However, the recursive trajectories from the hybrid forecast track the actual data relatively better because they rely on outside information for the underlying growth. The SPF forecast as a proxy for the long-run trend in the BVAR in turn helps improve the forecast accuracy of the entire forecast trajectory (one quarter ahead to 12 quarters ahead). For real GDP growth, for example the professional forecasters' assessment of the long-term trend indicates a lower growth potential of the economy, perhaps drawn from demographics or specific assessments of technological change.

For CPI inflation, the forecast accuracy gains are substantially higher, statistically significant, and persist throughout; the hybrid forecast for CPI inflation is roughly 40 percent more accurate three years out, and 50 percent more accurate eight years out. This is not surprising from the discussion in the previous section. The inflation process has exhibited pronounced changes in the underlying trend since the 1950s, and so accounting for those changes turns out to be very important in achieving improved accuracy. Figure 5 further illustrates this point as the recursive forecast trajectories corresponding to hybrid forecasts track the actual evolution more closely compared to the baseline forecast. The baseline forecast trajectories shown in Figure 5 navigate toward the much higher long-run value, informed by the mean of the estimation sample that includes the high inflation rates of the 1970s.

The improvements in accuracy for the federal funds rate are of similar magnitude to those for CPI inflation. The combination of improved inflation forecasts and conditional on judgmental survey-based long-run values of the federal funds rate results in highly accurate forecasts of the federal funds rate (policy rate). The improvements in the forecast accuracy by the end of the third year are 30 percent on average. Figure 6 provides a visual sense of the comparison in the dynamics of the federal funds rate forecast between the baseline and hybrid forecasts. Since the onset of the financial crisis, the FOMC has used forward guidance on the federal funds rate as an additional monetary policy tool. In addition, starting in 2009, each quarter, the FOMC provides forecasts (up to three years out) of the macroeconomic variables real GDP growth, core and headline inflation, and the unemployment rate, along with long-run forecasts of the federal funds rate (Summary of Economic Projections). This information is likely part of the information set of the professional forecasters, resulting in a more reasonable projection.

For the unemployment rate, the results are mixed. For the first six quarters, the hybrid forecast is marginally inferior, and for the subsequent six quarters, the hybrid forecast is marginally better but these results are not statistically significant. Digging deeper in the error evaluation, the marginal gains in accuracy reflect the improved accuracy of the hybrid forecasts over the Great Recession and post-recovery period. Figure 7 provides a visual assessment indicating that the hybrid forecast tracks relatively well the movements in the unemployment rate, especially beginning in 2008 through the first half of 2011. The full-sample results for the unemployment rate suggests that on net combining the BVAR forecasts with the survey long-horizon forecast does not have any material impact on accuracy; if anything, it worsens the point forecast accuracy in the very long term.

Table 2 reports the corresponding accuracy of the density forecasts. As discussed and defined in section 3.2.3, we use the continuous ranked probability score (CRPS) metric to evaluate our density forecasts. A lower CRPS score is preferable to a higher score. The numbers reported in Table 2 are the average relative CRPS. Specifically, for each recursive forecast run, we compute the relative CRPS by variable and forecast horizon, defined as the difference between two forecasts. We then take the average of the relative CRPS over our forecast evaluation sample. For example, in the top panel, the reported numbers correspond to the average of the CRPS of the baseline forecast minus the CRPS of the raw unconditional BVAR forecast. A negative number suggests that the baseline forecast is more accurate compared to the raw BVAR forecasts. The results of the density forecast evaluation echo the results of the point forecast evaluation reported in Table 1. That is, baseline forecasts are more accurate compared to the raw BVAR forecasts, and the hybrid forecasts are more accurate compared to both the baseline and the raw BVAR forecasts. The bottom panel illustrates that gains in the density forecast accuracy of the hybrid forecasts for CPI inflation, the federal funds rate, and credit spreads are substantial, statistically significant, and persist far into the future. For the unemployment rate, just as in the case of the point forecasts, the hybrid density forecasts are marginally inferior to the baseline forecasts but the differences are not statistically significant. In the results reported in the Online Appendix A5, incorporating the variance conditions of the nowcasts in addition to the nowcast mean conditions into the baseline forecasts and the hybrid forecasts lead to slight additional gains in the accuracy of density forecasts for all variables.

6.1.2 Results with the Medium BVAR model

Table 3 reports the forecast evaluation results for the forecasts from the Medium BVAR. This model has five additional variables compared to the small BVAR, and we report the forecasting accuracy of these additional variables. Just as in the case of the small BVAR, the forecasts of the four variables (real GDP, CPI inflation, the unemployment rate, and the federal funds rate) are tilted to match the respective SPF nowcasts and long-run forecasts. In addition, we now also tilt the forecasts of core CPI (beginning in 2007.Q1) to its respective nowcast and long-run forecast.[28] The estimation of the data and forecasts are generated using real-time data but the forecasts are evaluated using the latest available vintage. The use of the latest vintage for truth helps provide a check on the robustness of the results.

The top panel reports the results for the point forecast accuracy comparing the relative accuracy between baseline forecasts (i.e., tilting only on nowcasts) and hybrid forecasts (i.e., tilting on both nowcasts and survey long-horizon forecasts). The forecasting accuracy results for real GDP growth, CPI inflation, the unemployment rate, and federal funds rate are qualitatively similar to those reported in Table 1. More interestingly, this table gives a sense of the spillover effects on the accuracy of the model variables that have not been tilted (conditioned). Beginning with real consumption growth, on average over the evaluation sample the accuracy of the forecast of consumption growth based on the hybrid approach is marginally better compared to the baseline. Digging deeper into the forecast evaluation, we find that the forecast accuracy of the consumption growth forecasts based on the hybrid approach is substantially more accurate over the Great Recession and the subsequent recovery compared to the baseline. Recall that the hybrid forecast for real GDP is incorporating a more reasonable and rapidly evolving underlying trend growth compared to the baseline forecast, and because of a stronger historical correlation between real GDP and real consumption, the revised lower trajectory of real GDP in the post-crisis period helps lower the consumption trajectory, resulting in more accurate consumption forecasts.

The hybrid forecasts corresponding to core CPI inflation display the greatest gains in accuracy as the hybrid point forecasts are roughly 70 percent more accurate on average compared to the baseline for the 12-quarters ahead forecasts. And the gains are statistically significant. The improvement in core CPI inflation is due to the same reason as discussed earlier for CPI inflation in the case of the small BVAR. Not surprisingly, ECI compensation growth forecasts based on the hybrid approach exhibit accuracy gains at par with those of inflation. Payroll employment forecasts based on the hybrid approach also experience notable gains in accuracy and are statistically significant. The forecasting accuracy results for productivity growth and the unemployment rate suggest that adding survey long-horizon forecasts does not appear to help because the gains are not statistically significant. For the federal funds rate, the hybrid forecasts are more accurate in an absolute sense but are statistically significant only in forecast horizons beyond 12 quarters.

The density forecast accuracy exactly mirrors the accuracy results for the point fore-

[28] The long-run projection for core CPI is assumed to be the same as the SPF long-run projection of CPI inflation.

cast accuracy reported in the top panel. The greatest gains in density forecast accuracy for the hybrid forecasts are achieved for core CPI, CPI inflation, ECI compensation, payroll employment, and credit spreads. For the remaining variables, the hybrid forecasts generally indicate accuracy gains in an absolute sense but are not statistically significant.

Overall, our point and density forecasting results using real-time data provide strong evidence that tilting BVAR forecasts to match the long-run forecasts from the Survey of Professional Forecasters systematically leads to improved forecast accuracy for most variables over the forecast horizon of interest to monetary policymakers.

6.1.3 Results with the Medium BVAR model with stochastic volatility

Table 4 reports the forecast accuracy comparison between the hybrid forecast without stochastic volatility (denoted "BVAR (Now and LR)") and the hybrid forecast with stochastic volatility (denoted "BVAR with SV (Now and LR)"). Panel A reports the point forecast accuracy and Panel B reports the density forecast accuracy. Allowing for stochastic volatility leads to meaningful and persistent gains in point forecast accuracy for both core CPI inflation and nonfarm payroll employment. Shorter-term gains in the accuracy of real GDP forecasts are also achieved. For all other variables with the exception of credit spreads, stochastic volatility neither hurts nor helps as indicated by ratios that are close to one. In the case of credit spreads, the hybrid forecasts that employ stochastic volatility are on average 7 to 10 percent worse.

Allowing for stochastic volatility helps improve the density forecast accuracy (as evidenced by the negative values reported in Panel B) of the hybrid forecasts for all variables, with the exception of credit spreads. The gains for the most part are statistically significant. Just as in the case of point forecast accuracy, the density forecast accuracy corresponding to the hybrid forecast with stochastic volatility for credit spreads is meaningfully worse.

A result worth highlighting is the fact that the accuracy of both point and density forecasts of CPI inflation is not affected either positively or negatively by allowing for stochastic volatility. However, the accuracy for core CPI inflation is demonstrably improved through the use of stochastic volatility. Overall, our results indicate that allowing for stochastic volatility is useful because it helps improve both point and density forecast accuracy for most variables.

6.2 Inflation Forecast Accuracy of Tilted BVAR Compared to Univariate Benchmarks

Tilting the BVAR forecasts to match the long-horizon survey forecasts leads to meaningful gains in forecast accuracy for most variables, but the gains in point forecast accuracy are substantial for nominal variables such as price inflation and wage inflation. Given this, one empirical exercise worth investigating is how the accuracy of inflation forecasts from the tilting approach compares to hard-to-beat univariate benchmark models. Accordingly, we next compare the inflation forecast accuracy from our Medium BVAR (with and without stochastic volatility) using the two most well-known univariate benchmarks:

a random walk model (Atkeson and Ohanian, 2001) and univariate unobserved component with stochastic volatility (UCSV) model of Stock and Watson (2007).[29]

Random walk model of Atkeson and Ohanian (2001). For our forecasting exercise, the forecasts of CPI inflation into the future is computed by averaging the previous four available quarterly annualized readings of the CPI .
To construct a fair horserace, we set $\hat{\pi}_{t+1}$ equal to the *survey nowcast*

$$\hat{\pi}_{t+h} = 0.25(\hat{\pi}_{t+1} + \pi_t + \pi_{t-1} + \pi_{t-2}) \quad for\ h >= 2 \tag{21}$$

Univariate unobserved component with stochastic volatility (UCSV) model of Stock and Watson (2007). The superior accuracy of this model in forecasting inflation is well documented in numerous studies. The model decomposes inflation into two components – a stochastic trend component and a transitory component – and assumes time-varying variances of the respective shocks to these two components. The specification of this model is as follows (for ease of exposition we retain the notation used in Stock and Watson (2007)):

$$\pi_t = \tau_t + \eta_t, \qquad where\ \eta_t = \sigma_{\eta,t}\zeta_{n,t} \qquad \zeta_{n,t}\ \ is\ i.i.d.\ N(0, I_1) \tag{22}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t, \qquad where\ \varepsilon_t = \sigma_{\varepsilon,t}\zeta_{\varepsilon,t} \qquad \zeta_{\varepsilon,t}\ \ is\ i.i.d.\ N(0, I_1) \tag{23}$$

$$ln(\sigma^2_{\eta,t}) = ln(\sigma^2_{\eta,t-1}) + \nu_{\eta,t}, \qquad where\ \nu_{\eta,t}\ \ is\ i.i.d.\ N(0, \gamma I_1) \tag{24}$$

$$ln(\sigma^2_{\varepsilon,t}) = ln(\sigma^2_{\varepsilon,t-1}) + \nu_{\varepsilon,t}, \qquad where\ \nu_{\varepsilon,t}\ \ is\ i.i.d.\ N(0, \gamma I_1) \tag{25}$$

The model forecast for inflation infinite quarters into the future is simply the model's current estimated trend inflation rate.[30]

To construct a fair horserace, we set $\hat{\pi}_{t+1}$ equal to the *survey nowcast*. For forecasts $h >= 2$, we estimate the model through t+1, treating the survey nowcast as data and computing the updated trend estimate.

$$\hat{\pi}_{t+h} = \hat{\tau}_{t:t+1} \quad for\ h >= 2 \tag{26}$$

Table 5 reports the results comparing out-of-sample CPI inflation forecasting performance (both point and density) across four models: tilted Medium BVAR (hybrid), tilted Medium BVAR with SV (hybrid SV), random walk model, and UCSV model. The upper panel reports the results corresponding to the full sample, and the lower panel to the pre-crisis sample. In each panel, the first two rows report mean squared errors (MSE) from the hybrid and the hybrid with stochastic volatility, respectively, and the subsequent two rows report the ratio of the MSE from the hybrid (since the hybrid and the hybrid SV perform competitively) relative to the respective benchmarks (i.e., relative MSE). A ratio less than one suggests that CPI inflation forecasts from our BVAR

[29] See Tallman and Zaman (2017) for a broader examination of the forecasting properties of other models as well as these two models.

[30] The scalar parameter γ determines the smoothness of the stochastic volatility process. Following Stock and Watson (2007) we fix it at 0.2.

model tilted to match survey long-horizon and survey nowcasts (i.e., the hybrid forecast) are more accurate on average than those of the respective univariate benchmark model. The last three rows in each panel report the relative CRPS of the baseline forecast (i.e., the BVAR forecast tilted to nowcast only) to the UCSV, hybrid forecast relative to the UCSV, and hybrid forecast with SV relative to the UCSV. A positive number suggests that the density forecast from the UCSV model is more accurate on average.

As shown by the numbers reported in the table, the hybrid CPI point forecasts are more accurate on average compared to the random walk model and are competitive with those from the UCSV model. Even though in an absolute sense the hybrid forecasts are on average slightly more accurate, the gains are not statistically significant. This result holds for each of the forecast evaluation samples.

For density forecast accuracy, the UCSV forecast is substantially superior to the baseline forecast (as evidenced by positive values) and the gains are statistically significant throughout. But compared to the hybrid forecast, the UCSV forecast is only slightly better and accuracy gains are just marginally significant in a statistical sense in only a few horizons. This suggests that combining the long-horizon survey forecasts with the BVAR forecasts improves the density forecasts notably. Adding stochastic volatility to the hybrid forecasts (i.e., hybrid with SV) helps further improve the density forecast accuracy of the hybrid forecasts, because the accuracy gains of the UCSV in an absolute sense are not statistically significant.

The evidence that the tilted multivariate model (BVAR) estimated with constant parameters using a longer sample generates inflation forecasts that rival both the point and density forecast accuracy of forecasts from tough benchmarks is a worthwhile result because policymakers at central banks faced with inflation targeting may benefit from using models that allow for feedback between policy and the real economy and inflation. Such models have underperformed simple univariate approaches in terms of forecasting inflation and at times significantly so. The examination of our exercises suggests that the modeling technique we propose and illustrate may have some useful payoffs for policymakers.

6.3 Comparison to Time-varying coefficient VAR with SV

In light of the well-documented evidence of structural changes in the macroeconomic relationships, macroeconometricians have innovated the constant parameter VAR models along various dimensions, such as introducing time variation in the following components: dynamic coefficients, coefficient capturing the permanent component, standard deviation of the shocks, and the correlation structure of the shocks in the VAR models (e.g., Cogley and Sargent, 2005; Primiceri, 2005). The parameters are assumed to gradually evolve over time. These innovations to the VAR model (denoted TVP-VAR SV) allows it to explicitly model possible structural changes. Not surprisingly, researchers have documented the superior forecasting capability of the TVP-VAR SV models compared to the alternatives (see D'Agostino, Giannone, and Gambetti, 2013 for the US; Barnett, Mumtaz, and Theodoridis, 2014 for the UK). Interestingly, the forecast superiority of the TVP-VAR SV documented by previous research is most notable for inflation, the variable for which we document notable gains using our approach. As such, it will be worthwhile to run a horserace comparing the forecasting accuracy of TVP-VAR SV (as in Primiceri, 2005) to our approach of tilting a constant parameter VAR model to survey expectations (i.e., the hybrid approach).

A limitation of time-varying VAR models is that they can feasibly be estimated only with three or four variables mainly because of computational constraints. Therefore to run a fair horserace, we compare the forecast accuracy between the hybrid approach using the Small VAR and the TVP-VAR SV that uses the three macroeconomic variables of most interest to central bankers (real GDP, CPI inflation and unemployment rate). We also compare the forecast accuracy between the hybrid SV forecast (Small VAR with stochastic volatility) and the TVP-VAR SV. For a fair horserace, we also require the TVP-VAR SV one-step-ahead forecast to be tilted to match the more accurate survey nowcast (just like the hybrid forecast). This will ensure that competing models start from the same jumping-off point.

Table 6 reports the forecast accuracy comparison results between the hybrid forecasts and TVP-VAR SV. The top panel (i.e., Panel A) reports the results corresponding to point forecast accuracy and the bottom panel (i.e., Panel B) reports the results corresponding to density forecast accuracy. In both Panels A and B, the top portion reports the accuracy comparison between hybrid forecasts (without stochastic volatility) and TVP-VAR SV; the bottom portion reports the accuracy comparison between hybrid forecasts that allow for stochastic volatility and the TVP-VAR SV. In the case of the point forecast comparison, a ratio of less than one suggests that the hybrid (and the hybrid SV) point forecast accuracy is on average more accurate compared to the TVP-VAR SV. In the case of the density forecast comparison, a negative number suggests that the hybrid density forecast accuracy is on average more accurate compared to the TVP-VAR SV.

Beginning with the point forecast accuracy, the hybrid forecasts corresponding to real GDP growth and CPI inflation are generally more accurate in an absolute sense, and for inflation, the gains are statistically significant. The ability of the survey expectations to adapt more quickly compared to the TVP-VAR SV may partly explain the more accurate inflation forecasts from the hybrid approach. To give a sense of the evolution of the estimated trends from the TVP-VAR SV, the Online Appendix A6 includes figures plotting trend estimates from the TVP-VAR SV next to the survey expectations. In the case of the unemployment rate, the forecast accuracy is competitive with each other as the ratios are close to 1 and the differences are not statistically significant.

The hybrid point forecasts from the Small VAR model that incorporates stochastic volatility (hybrid SV) are on average more accurate compared to both the hybrid forecasts and the TVP-VAR SV as judged by ratios that are marginally smaller in magnitude compared to those reported in the top portion of Panel A. Moving next to the density forecast accuracy, the hybrid forecasts generally rival the density forecast accuracy of the TVP-VAR SV with the exception of the nowcast quarter, for which the TVP-VAR SV generates more accurate density forecasts for both real GDP growth and the unemployment rate. The density accuracy superiority of the TVP-VAR SV for real

GDP growth in the nowcast quarter diminishes when the hybrid forecast incorporates stochastic volatility. The density forecast accuracy of the hybrid SV and TVP-VAR SV is generally competitive with each other. It is worth noting that the hybrid forecast without stochastic volatility generally rivals the forecast accuracy of the TVP-VAR SV, but allowing for stochastic volatility in the constant coefficient VAR model (hybrid SV) leads to just slightly more accurate forecasts, resulting in the hybrid SV approach as the most accurate.

If we were to compare the forecast accuracy between the hybrid SV forecast from the Medium VAR to that of the TVP-VAR SV, the hybrid SV will clearly dominate both the point and density forecast accuracy. This is the case because the hybrid from the Medium VAR SV generates more accurate forecasts on average compared to the hybrid Small VAR SV as would be expected because of the larger data set in Medium VAR versus the Small VAR (Banbura, Giannone, and Reichlin, 2010; Carriero, Clark, and Marcellino, 2016).

The hybrid forecasts generated from our approach perform competitively with the time-varying VAR model for all four variables. We suggest that tilting VAR forecasts to survey long-term forecasts is an indirect method of accommodating structural change and moving end points. The important advantage of our approach is that we are not restricted to smaller data sets. For example, we could easily estimate a larger VAR with additional variables of interest to obtain forecasts that to some extent reflect adjustment in response to structural change. This effect would also be visible through spillover effects from the core variables that are directly tilted toward survey long-run expectations.

7 Conclusions

In this paper, we propose a technique to adjust the medium-term and long-horizon forecasts from a VAR toward plausible values proposed by judgmental forecasters. We construct a hybrid forecast of survey nowcast, VAR forecast, and long-run survey projection. Specifically, using the flexible and powerful technique of relative entropy, we tilt the VAR forecast both in the near term with the survey nowcast and in the long run with the survey long-run projection. The horizon at which the long-run survey projection is combined with the VAR forecast is variable specific and determined by the variable's estimated persistence at forecast origin.

We estimate two variants of VAR models (small and medium) using Bayesian methods with data going back to 1959 and recursively forecast from 1 to 40 quarters ahead beginning in 1994 through 2016 using real-time data vintages. We find that tilting BVAR forecasts to match the long-run forecasts from the Survey of Professional Forecasters systematically leads to improved forecast accuracy as indicated by lower RMSEs and lower CRPS for most variables over the forecast horizon of interest (i.e., 1 to 3 years out) to monetary policymakers. The greatest gains in accuracy are achieved for variables believed to have undergone marked shifts in their permanent components (e.g., inflation and the federal funds rate). We also show substantial forecast accuracy improvements for a host of variables (such as compensation, payroll employment, credit spreads) that are not directly tilted but are affected through the spillover effects of the tilted variables.

We also show that hybrid inflation forecasts rival those of relatively accurate univariate benchmark models. We view this as a useful practical contribution because among the many frustrations of monetary policymakers is the inability of multivariate models – which allow for feedback effects from policy to the real economy and inflation – to match the forecasting performance of the univariate forecasting models.

Finally, we show that the forecast accuracy from our approach rivals the accuracy of the time-varying coefficient VAR models that are built to account explicitly for structural changes. These results lead us to view our forecasting approach as an indirect method for accommodating structural change and shifting end points.

Although our examination focused on VAR models and surveys, in principle forecasts from any empirical model that can generate predictive densities could be tilted through relative entropy to match moment restrictions informed by a wide variety of sources including other models.

References

[1] Aastveit, Knut Are, Andrea Carriero, Todd E. Clark, and Massimiliano Marcellino (2016) Have Standard VARs Remained Stable Since the Crisis? *Journal of Applied Econometrics* doi:10.1002/jae.2555

[2] Altavilla, Carlo, Raffaella Giacomini, and Giuseppe Ragusa (2017) "Anchoring the Yield Curve Using Survey Expectations," *Journal of Applied Econometrics* 32(6): 1055-1068 doi:10.1002/jae.2588

[3] Atkeson, Andrew and Lee E. Ohanian (2001) Are Phillips Curves Useful for Forecasting Inflation? Federal Reserve Bank of Minneapolis *Quarterly Review* (Winter): 211. Retrieved from https://ideas.repec.org/a/fip/fedmqr/y2001iwinp2-11nv.25no.1.html

[4] Banbura, Marta, Domenico Giannone, and Lucrezia Reichlin (2010) Large Bayesian Vector Auto Regressions, *Journal of Applied Econometrics* 25: 71-92 doi:10.1002/jae.1137

[5] Banbura, Marta, Domenico Giannone, and Michele Lenza (2015) Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections, *International Journal of Forecasting*, 31 (3), pages 739-756 doi:10.1016/j.ijforecast.2014.08.013

[6] Barnett, Alina, Haroon Mumtaz, and Konstantinos Theodoridis (2014) Forecasting UK GDP Growth and Inflation under Structural Change: A Comparison of Models with Time-Varying Parameters, *International Journal of Forecasting*, 30: 129-143 doi:10.1016/j.ijforecast.2013.06.002

[7] Bauwens Luc, Michel Lubrano, and Jean Francois Richard (1999) Bayesian Inference in Dynamic Econometric Models, *Oxford University Press* doi:10.1093/acprof:oso/9780198773122.001.0001

[8] Beauchemin, Kenneth, and Saeed Zaman (2011) A Medium Scale Forecasting Model for Monetary Policy Federal Reserve Bank of Cleveland Working Paper no. 11-28. doi:/10.26509/frbc-wp-201128

[9] Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2015) Bayesian VARs: Specification Choices and Forecast Accuracy, *Journal of Applied Econometrics*, 30(1), pp 46-73 doi:10.1002/jae.2315

[10] Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2016) Large Vector Autoregressions with Stochastic Volatility and Flexible Priors, Federal Reserve Bank of Cleveland Working Paper, no. 16-17. doi:/10.26509/frbc-wp-201617

[11] Chan, Joshua and Gary Koop (2014) Modeling Breaks and Clusters in the Steady States of Macroeconomic Variable, *Computational Statistics and Data Analysis*, 76, 186-193 doi:10.1016/j.csda.2013.05.007

[12] Clark, Todd E. (2011) Real-Time Density Forecasts from Bayesian VARs with Stochastic Volatility, *Journal of Business Economics and Statistics*, 29: 327-341 doi:10.1198/jbes.2010.09248

[13] Clark, Todd E., and Michael W. McCracken (2008) Forecasting with Small Macroeconomic VARs in the Presence of Instabilities, *in Forecasting in the Presence of Structural Breaks and Model Uncertainty, D. Rapach and M. Wohar, eds., Emerald Publishing*, pp. 93-147 doi:10.1016/s1574-8715(07)00203-5

[14] Clark, Todd E., and Michael W. McCracken (2010) Averaging Forecasts from VARs with Uncertain Instabilities, *Journal of Applied Econometrics*, 5-29 doi:10.1002/jae.1127

[15] Clark, Todd E., and Michael W. McCracken (2013) Advances in Forecast Evaluation, *in Handbook of Economic Forecasting (vol. 2), eds. G. Elliott, and A.Timmermann, Amsterdam, the Netherlands: Elsevier*, pp. 11071201 doi:10.1016/B978-0-444-62731-5.00020-8

[16] Clark, Todd E., and Michael W. McCracken (2017) Tests of Predictive Ability for Vector Autoregressions Used For Conditional Forecasting, *Journal of Applied Econometrics*, 32: 533-553 doi:10.1002/jae.2529

[17] Clements, Michael P., (2014) Forecast Uncertainty-Ex Ante and Ex Post: U.S. Inflation and Output Growth, *Journal of Business and Economic Statistics* 32: 206-216 doi:10.1080/07350015.2013.859618

[18] Cogley, Timothy, and Thomas J. Sargent (2005) Drifts and Volatilities: Monetary Policies and Outcomes in the Post World War II U.S., *Review of Economic Dynamics* 8: 262-302 doi:10.1016/j.red.2004.10.009

[19] Cogley, Timothy, Sergei Morozov, and Thomas 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 doi:10.1016/j.jedc.2005.06.005

[20] Croushore, Dean (2010) An Evaluation of Inflation Forecasts from Surveys Using Real-time Data, *BE Journal of Macroeconomics* 10(1):1-32 doi:10.2202/1935-1690.1677

[21] D'Agostino, Antonello, Domenico Giannone, and Luca Gambetti (2013) Macroeconomic Forecasting and Structural Change, *Journal of Applied Econometrics*, 28 (1), pages 81-101 doi:10.1002/jae.1257

[22] Diebold, Francis X., and Roberto S. Mariano (1995) Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263 doi:10.2307/1392185

[23] Doan Tom, Robert Litterman, and Chris Sims (1984) Forecasting and Conditional Projection Using Realistic Prior Distributions, *Econometric Reviews*, 3: 1-100 doi:10.1080/07474938408800053

[24] Faust, Jon, and Jonathan Wright (2013) Inflation Forecasting, *In Handbook of Economic Forecasting, Elliot G. Timmermann A (eds)*, Vol. 2. North-Holland: Amsterdam; 2-56 doi:10.1016/b978-0-444-53683-9.00001-3

[25] Frey, Christoph., and Frieder Mokinski (2016) Forecasting with Bayesian Vector Autoregressions Estimated Using Professional Forecasts, *Journal of Applied Econometrics*, 31:1083-1099 doi:10.1002/jae.2483

[26] Geweke, John (1989) Bayesian Inference in Econometric Models Using Monte Carlo Integration, *Econometrica* 57, 1317-1339 doi:10.2307/1913710

[27] Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri (2018) Priors For The Long Run, *Journal of the American Statistical Association* 147 doi:10.1080/01621459.2018.1483826

[28] Gneiting, Tilmann, and Adrian E. Raftery (2007) Strictly Proper Scoring Rules, Predictions, and Estimations, *Journal of the American Statistical Association*, 102: 359-378. doi:10.1198/016214506000001437

[29] Gordon, Neil J., D.J. Salmond, and A.F.M. Smith (1993) A Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation, *IEEE Proceedings-F140*, pp.107-113 doi:10.1049/ip-f-2.1993.0015

[30] Harvey, David., Stephen Leybourne, and Paul Newbold (1997) Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting*, 13, 281-291 doi:10.1016/s0169-2070(96)00719-4

[31] Knotek II, Edward S., and Saeed Zaman (2017a) Nowcasting U.S. Headline and Core Inflation, *Journal of Money, Credit and Banking*, 49(5): 932-968 doi:10.1111/jmcb.12401

[32] Knotek II, Edward S., and Saeed Zaman (2017b) Financial Nowcasts and Their Usefulness in Macroeconomic Forecasting, Federal Reserve Bank of Cleveland Working Paper no. 17-02 doi:/10.26509/frbc-wp-201702

[33] Koop, Gary (2013) Forecasting with Medium and Large Bayesian VARs, *Journal of Applied Econometrics* 28, 177-203 doi:10.1002/jae.1270

[34] Koop, Gary and Dimitris Korobilis (2013) Large Time-Varying Parameter VARS, *Journal of Econometrics* 177: 185-198 doi:10.1016/j.jeconom.2013.04.007

[35] Kruger, Fabian, Todd E. Clark, and Francesco Ravazzolo (2017) Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts, *Journal of Business and Economic Statistics* doi:10.1080/07350015.2015.1087856

[36] Litterman, R.B., (1986) Forecasting With Bayesian Vector Autoregressions – Five Years of Experience, *Journal of Business and Economic Statistics* 4, 25-38. doi:10.1080/07350015.1986.10509491

[37] Modugno, Michele (2013) Now-casting Inflation Using High Frequency Data, *International Journal of Forecasting*, 29(4): 664-675 doi:10.1016/j.ijforecast.2012.12.003

[38] Primiceri, Giorgio E. (2005) Time Varying Structural Vector Autoregressions and Monetary Policy, *Review of Economic Studies*, 72, 821-852 doi:10.1111/j.1467-937x.2005.00353.x

[39] Robertson, John C., and Ellis W. Tallman (1999) Vector Autoregressions: Forecasting and Reality, Federal Reserve Bank of Atlanta *Economic Review*, Vol. 84:1 Retrieved from https://ideas.repec.org/a/fip/fedaer/y1999iq1p4-18nv.84no.1.html

[40] Robertson, John C., Ellis W. Tallman, and Charles H. Whiteman (2005) Forecasting Using Relative Entropy, *Journal of Money, Credit and Banking*, 37(3): 383-401 doi:10.1353/jmcb.2005.0034

[41] Sims, Christopher A., and Tao Zha (1998) Bayesian Methods for Dynamic Multivariate Models, *International Economic Review*, 39: 949-968 doi:10.2307/2527347

[42] Stock, James H. and Mark W. Watson (2007) Why Has U.S. Inflation Become Harder to Forecast?, *Journal of Money, Credit and Banking*, 39: 3-33 doi:10.1111/j.1538-4616.2007.00014.x

[43] Tallman, Ellis W., and Saeed Zaman (2017) Forecasting Inflation: Phillips Curve Effects on Services Price Measures, *International Journal of Forecasting*, 33(2): 442-457 doi:10.1016/j.ijforecast.2016.10.004

[44] Villani, Mattias (2009) Steady-State Priors for Vector Autoregressions, *Journal of Applied Econometrics*, 24: 630-650 doi:10.1002/jae.1065

[45] Waggoner, Daniel F., and Tao Zha (1999) Conditional Forecasts in Dynamic Multivariate Models, *Review of Economics and Statistics*, 81: 639-651 doi:10.1162/003465399558508

[46] West, Kenneth D. (1996) Asymptotic Inference About Predictive Ability, *Econometrica*, 64: 1067-1084 doi:10.2307/2171956

[47] Wright, Jonathan (2013) Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior, *Journal of Applied Econometrics*, 28: 762-776 doi:10.1002/jae.2268

[48] Zaman, Saeed (2013) Improving Inflation Forecasts in the Medium to Long Term, Federal Reserve Bank of Cleveland *Economic Commentary*, No. 2013-16 Retrieved from https://ideas.repec.org/a/fip/fedcec/y2013inov15n2013-16.html

[49] Zellner, Arnold (1971) An Introduction to Bayesian Inference in Econometrics, *Wiley: New York*

8 Technical Appendix

Specifically, we impose our prior beliefs about the coefficient estimates in $A_1, ...A_p$ and Σ using Normal inverse-Wishart (N-IW) conjugate priors.[31] The prior beliefs for the mean and variances of the coefficient matrices are as follows:

$$E[A_l^{(i,j)}] = \begin{cases} \delta_i & \text{if } i = j, l = 1 \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

$$Var[A_l^{(i,j)}] = \lambda^2 \frac{1}{l^2} \frac{\sigma_i^2}{\sigma_j^2}, \quad l = 1, ...p \tag{28}$$

The scale factor $\frac{1}{l^2}$ helps impose the prior belief that recent lags play a more influential role compared to the distant ones by proportionally shrinking the variances on the more distant lags (centered on the prior mean of zero). The prior parameter σ_i equals the standard deviation of the residuals obtained from regressing the variable y_i on its own p lags and a constant over the sample period up to time t. We set δ_i equal to the sum of the autoregressive coefficients obtained from regressing the variable y_i on its own p lags and a constant over a pre-forecast evaluation sample.[32] The hyperparameter λ governs the tightness of our prior beliefs. As $\lambda \longrightarrow 0$ the prior dominates and so posterior equals prior, i.e., the data have no say. On the other hand as $\lambda \longrightarrow \infty$, the prior's influence diminishes and so posterior estimates converge to OLS estimates.

The VAR model in (1) can be rewritten in compact matrix notation as

$$Y = XA + E \tag{29}$$

where Y=$[y_1, ..., y_T]'$ is a $T \times n$ matrix, X=$[x_1, ..., x_T]'$ is a $T \times k$ matrix co, $x_t = [1, y'_{t-1}, ..., y'_{t-p}]$ is a $1 \times k$ vector. $A = [A_c \ A_1 \ \cdots \ A_p]$ is a matrix of VAR coefficients of size $k \times n$, $E = [\varepsilon_1 \ \cdots \ \varepsilon_T]$ is a $T \times n$ matrix of innovation terms.

The conjugate Normal inverse-Wishart (N-IW) prior is:

$$vec\,(A)\,|\Sigma \sim N\left(vec\,(A_0)\,, \Sigma \otimes \Omega_0\right), \qquad \Sigma \sim iW\,(S_0, v_0) \tag{30}$$

where A_0, Ω_0, S_0, and v_0 are the prior parameters whose values are set based on the VAR model in (1) that satisfies the prior moment conditions specified in (2).
Since our prior N-IW is conjugate, the resulting conditional posterior distribution (i.e.,

[31]Natural conjugate priors such as N-IW have computational advantages and at the same time have competitive forecasting properties (see Koop, 2013)

[32]Since all the variables entering the VAR are stationary, if we instead set $\delta_i = 0$ for all i, we get qualitatively similar results. Studies such as Clark (2011) and Carriero et al. (2015) have used a value of 0.8 for variables that are known to exhibit persistence (e.g., the unemployment rate, inflation and the interest rate). Our results are very similar if we instead set $\delta_i = 0.8$

the product of the prior and likelihood function) is also N-IW.

$$vec\left(A\right)|\Sigma,Y\sim N\left(vec\left(\overline{A}\right),\Sigma\otimes\overline{\Omega}\right),\qquad\Sigma\sim iW\left(\overline{S},\overline{v}\right). \tag{31}$$

where

$$\overline{A}=\left(\Omega_0^{-1}+X'X\right)^{-1}\left(\Omega_0^{-1}A_0+X'Y\right) \tag{32}$$

$$\overline{\Omega}=\left(\Omega_0^{-1}+X'X\right)^{-1} \tag{33}$$

$$\overline{v}=v_0+T. \tag{34}$$

are the respective posterior mean estimates of the VAR model.
And,

$$\overline{S}=A_0+\widehat{\varepsilon}'\widehat{\varepsilon}+\widehat{A}'X'X\widehat{A}+A_0'\Omega_0^{-1}A_0-\widehat{A}'\overline{\Omega}^{-1}\widehat{A} \tag{35}$$

where $\widehat{A}=\left(X'X\right)^{-1}X'Y$ equals the OLS estimate of A and $\widehat{\varepsilon}=Y-X\widehat{A}$ are the OLS residuals (Zellner, 1971). We use the mixed estimation method of Litterman (1986) to implement the N-IW prior, which equates to appending the data matrices with dummy observations.

Previous research (e.g., Robertson and Tallman, 1999; Banbura et al., 2010) has documented further gains in forecast accuracy by imposing a 'sum of coefficients' (SOC) prior on the equations of the VAR. The hyper-parameter μ will govern the tightness of this prior.

8.0.1 Optimal values of Hyper-parameters using marginal likelihood

The values of λ and μ are set by maximizing the marginal likelihood of the model over a predefined two dimensional discrete grid of λ and μ. The optimization is performed over the pre-forecast evaluation sample, and optimal values obtained are kept fixed over the forecast evaluation sample.[33]

$$[\lambda^+,\mu^+]=arg\max_{[\lambda^g,\mu^g]}\ lnp(Y) \tag{36}$$

where

$$p(Y)=\int p(Y|\theta)p(\theta)d(\theta) \tag{37}$$

is the marginal likelihood, and θ is the set of all model coefficients. Given we are

[33] That said, if we instead fix the $\lambda=0.2$ and $\mu=1$ (widely used values) the results are qualitatively similar.

using an N-IW prior, the marginal density p(Y) can be solved in closed form as[34]

$$
\begin{aligned}
p(Y) \;&=\; (\frac{1}{\pi})^{\frac{nT}{2}} \times |(I + X\Omega_0 X\prime)^{-1}|^{\frac{n}{2}} \times |S_0|^{\frac{v_0}{2}} \\
&\times \frac{\Gamma_n(\frac{v_0+T}{2})}{\Gamma_n(\frac{v_0}{2})} \times |S_0 + (Y - XA_0)^{'}(I + X\Omega_0 X\prime)^{-1}(Y - XA_0)|^{-\frac{v_0+T}{2}}.
\end{aligned}
$$

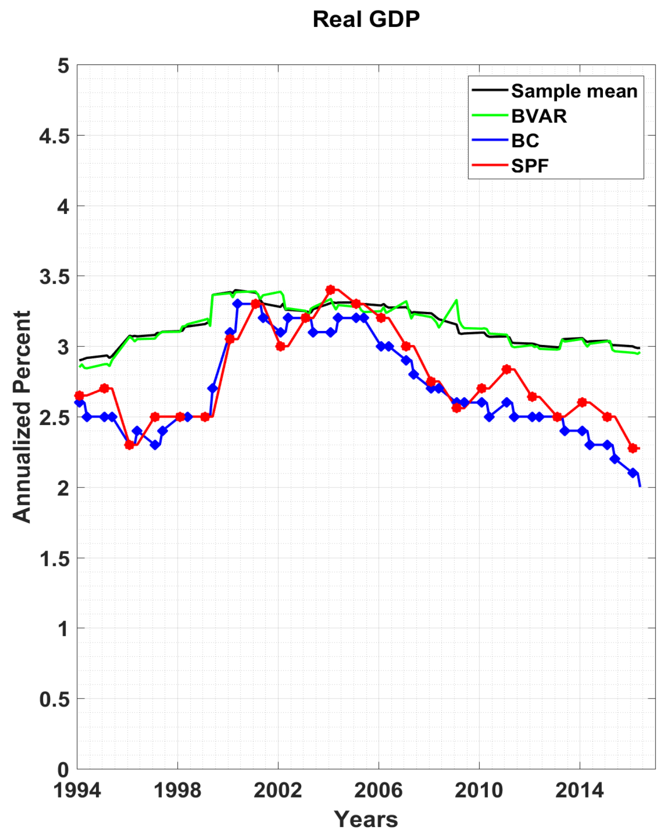
To evaluate the log marginal likelihood, we use the two-dimensional grid of discrete values defined as follows: λ^g=[0.050,0.10,0.15,0.20,0.30,0.40,0.50]; and μ^g=[0.1,0.15,0.2,0.25,0.50,1,1.5,2,2.5,3]; For the Small BVAR estimated with data starting in 1959, the optimal values obtained are $\lambda = 0.4$ and $\mu = 0.2$ and for the Medium VAR, optimal values are $\lambda = 0.3$ and $\mu = 0.25$. The values we obtain are close to values that other researchers have obtained through a grid search optimization.

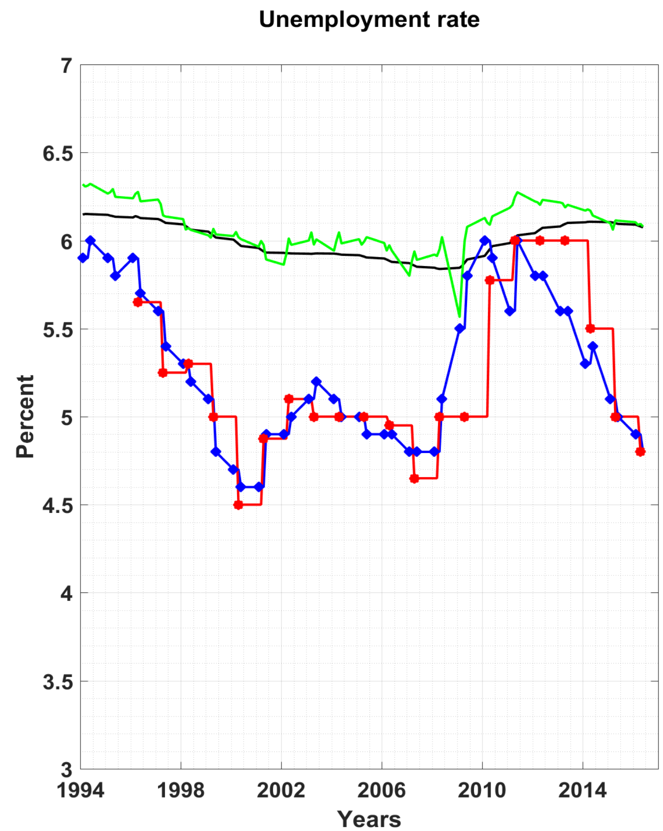
Note that prior specification for each equation is symmetric in its treatment of own lags of the dependent variable and lags of other variables. As such, we have a prior that is a natural conjugate (Normal inverse-Wishart prior), which proves to be convenient when solving for the model, because these priors can be implemented easily by augmenting the data matrices with dummy variables, permitting OLS to estimate the model equation by equation (for details see Banbura et al., 2010 and Carriero et al., 2015).

[34] Derivation details can be found in Bauwens et al. (1999), Carriero et al. (2015) and Giannone et al. (2015).

9 Tables and Figures

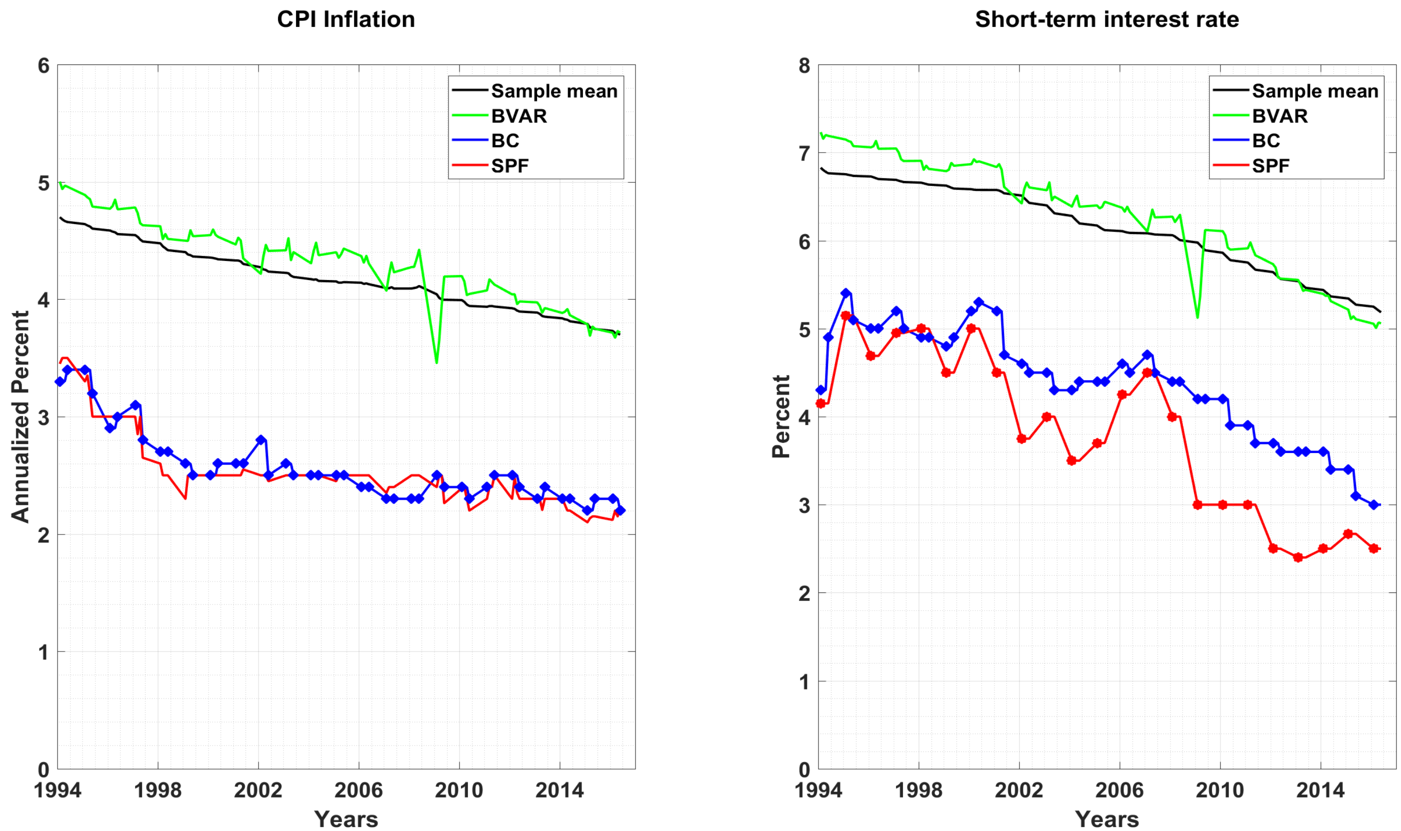
Figure 1: Real-Time Long Run Forecasts: GDP and Unemployment Rate





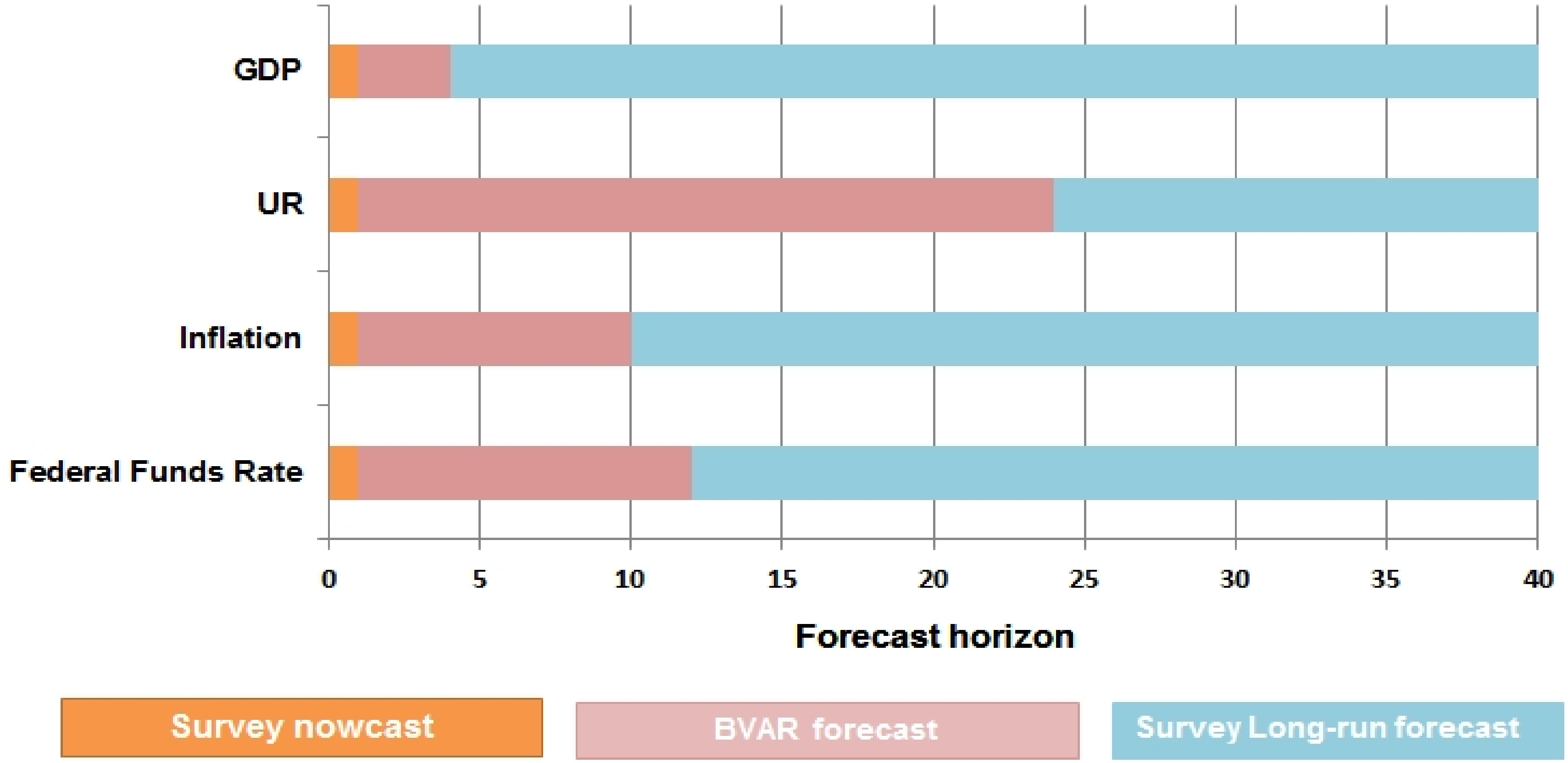
Notes for the figure: The figure plots the real-time long-run forecasts from the Survey of Professional Forecasters (SPF), the 10-year-ahead average forecast for real GDP growth and the estimate of the natural rate of unemployment; from the Blue Chip survey, the Figure plots 7 to 11 years ahead average. To facilitate comparison, also plotted are the real-time estimates of the long-run forecasts (40 quarters out) from a Small BVAR recursively estimated with data beginning in 1959 and a simple arithmetic sample mean computed recursively.

Figure 2: Real-Time Long Run Forecasts: CPI and Short-Term Interest Rate



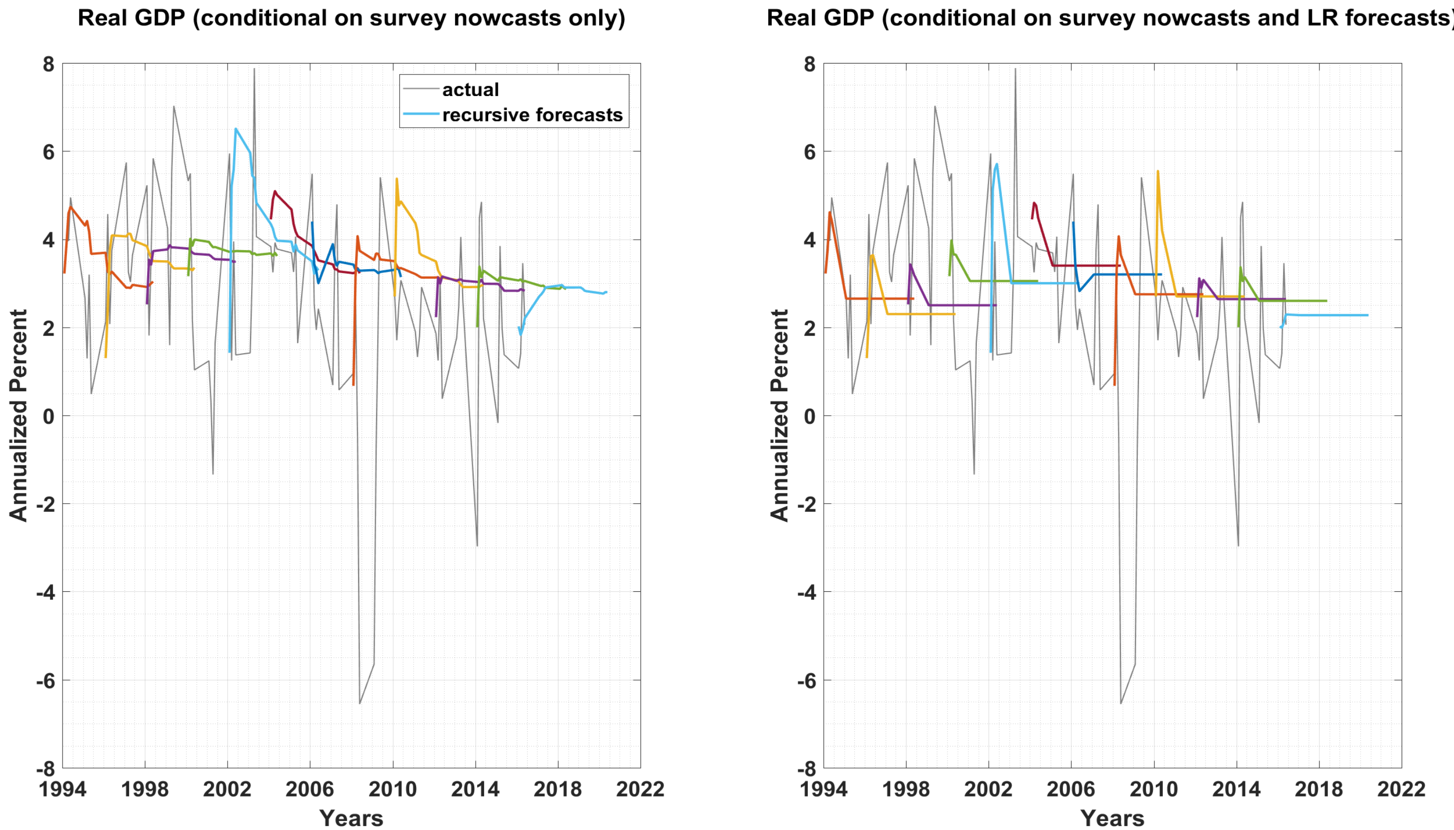
Notes for the figure: The figure plots the real-time long-run forecasts from the Survey of Professional Forecasters (SPF), the 10-year-ahead average forecast; from the Blue Chip survey, the Figure plots the 7 to 11 years ahead average. To facilitate comparison also plotted are the real-time estimates of the long-run forecasts (40 quarters out) from a Small BVAR recursively estimated with data beginning in 1959 and a simple arithmetic sample mean computed recursively. The short-term interest rate is the 3-month T-bill rate, but in the case of the BVAR, it refers to the federal funds rate.

Figure 3: Composition of Hybrid Forecast



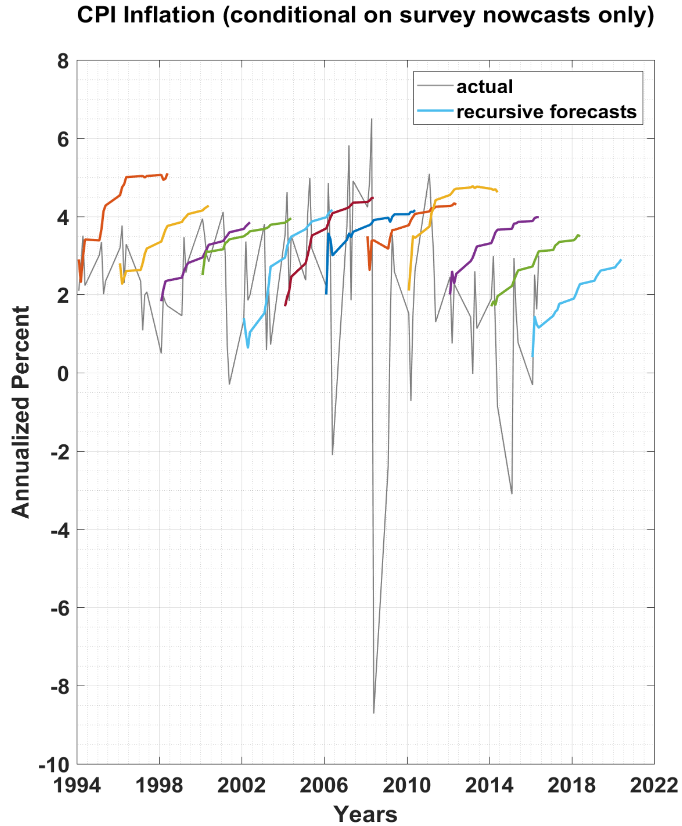
Notes: The figure illustrates the components of the hybrid forecast corresponding to our four variables of interest. The orange shaded area indicates the survey nowcast plugged at forecast horizon h=1 for all four variables. The pink shaded area corresponds to the BVAR forecast, and the blue shaded area corresponds to the survey long-run forecast. The changing width of the pink bars reflects the differences in the persistence of the variables: the more persistent the variable, the longer it takes for it to reach its long-run value, which in our exercise is the long-run forecast from the survey. The estimates are based on forecast evaluation exercises using the Small BVAR. On average in our forecast exercises, the hybrid forecast for real GDP consists of 3 quarters of the BVAR forecast, and 36 quarters of the survey long-horizon forecast. For the unemployment rate, it is 22 quarters of the BVAR forecast, and 17 quarters of the survey long-horizon. Similarly, for CPI inflation, 9 quarters of the BVAR forecast, and 30 quarters of the survey long-horizon forecast. Finally, for the federal funds rate we use 11 quarters of BVAR forecast, and 28 quarters of the survey long-horizon forecast. The estimates based on the Medium BVAR are very similar.

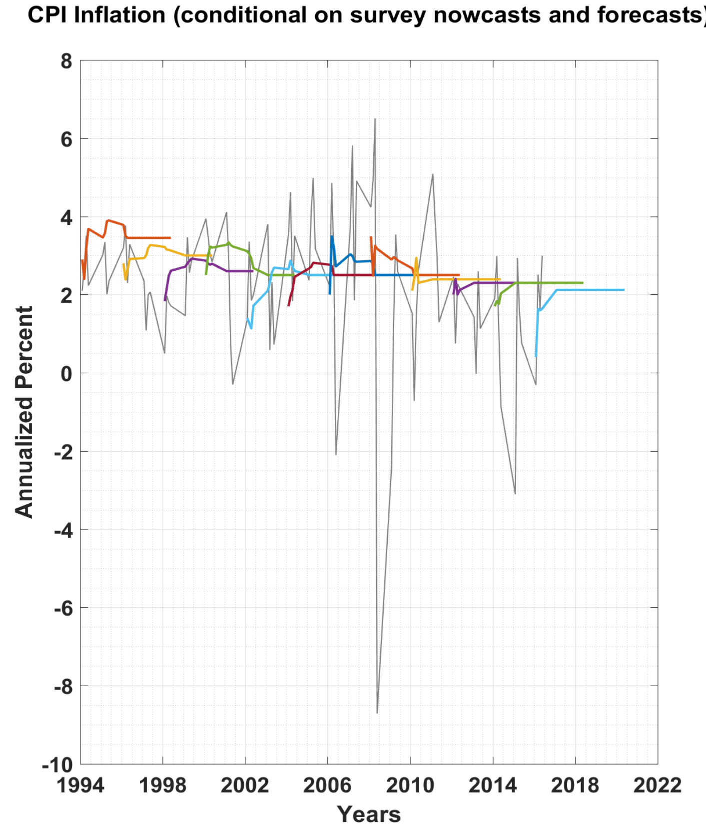
Figure 4: Recursive Point Forecasts for Real GDP Growth



Notes: The figure plots the recursive point forecasts of real GDP growth annualized from the Small BVAR at different points in time. Also plotted are the actual data (solid gray line). The forecasts are recursively generated beginning in 1994.Q1 through 2016.Q4, giving us a total of 92 recursive forecasts. Each forecast is up to 20 quarters ahead. For presentation purposes every 7th forecast is displayed, i.e., 12 out of 92. The left panel corresponds to the forecast that conditions only on the survey nowcasts. The right panel corresponds to the forecast that conditions on both the survey nowcasts and the survey long-horizon forecasts.

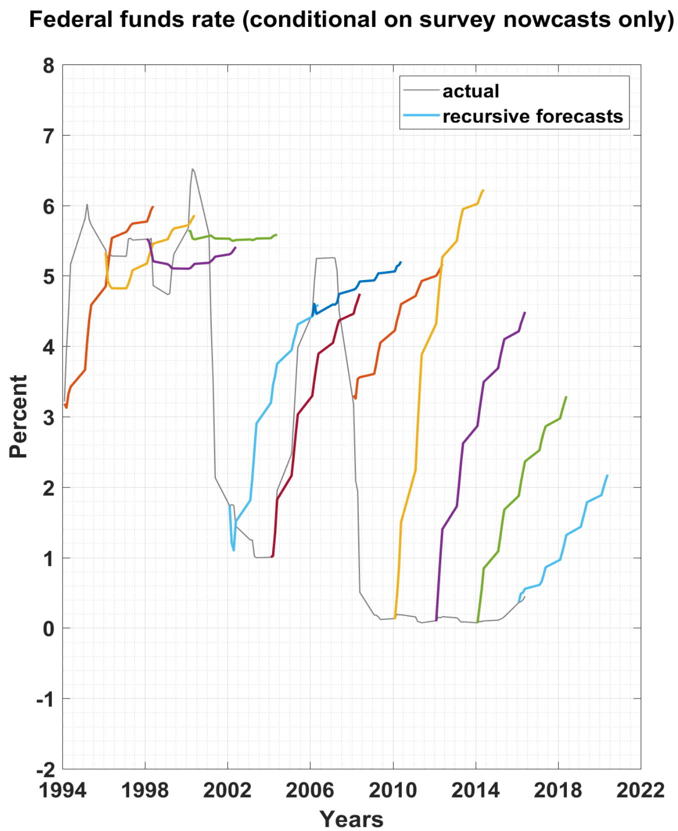
Figure 5: Recursive Point Forecasts for CPI Inflation

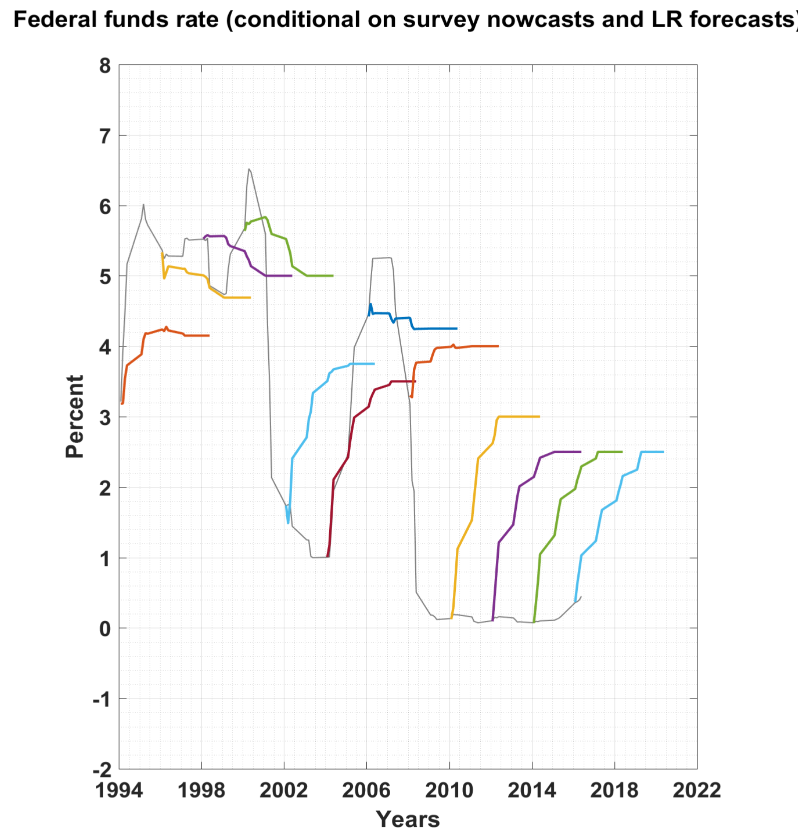




Notes: The figure plots the recursive point forecasts of headline CPI inflation at an annualized rate from the Small BVAR at different points in time. Also plotted are the actual data (solid gray line). The forecasts are recursively generated beginning in 1994.Q1 through 2016.Q4, giving us a total of 92 recursive forecasts. Each forecast is up to 20 quarters ahead. For presentation purposes every 7th forecast is displayed, i.e., 12 out of 92. The left panel corresponds to the forecast that conditions only on the survey nowcasts. The right panel corresponds to the forecast that conditions on both the survey nowcasts and the survey long-horizon forecasts.

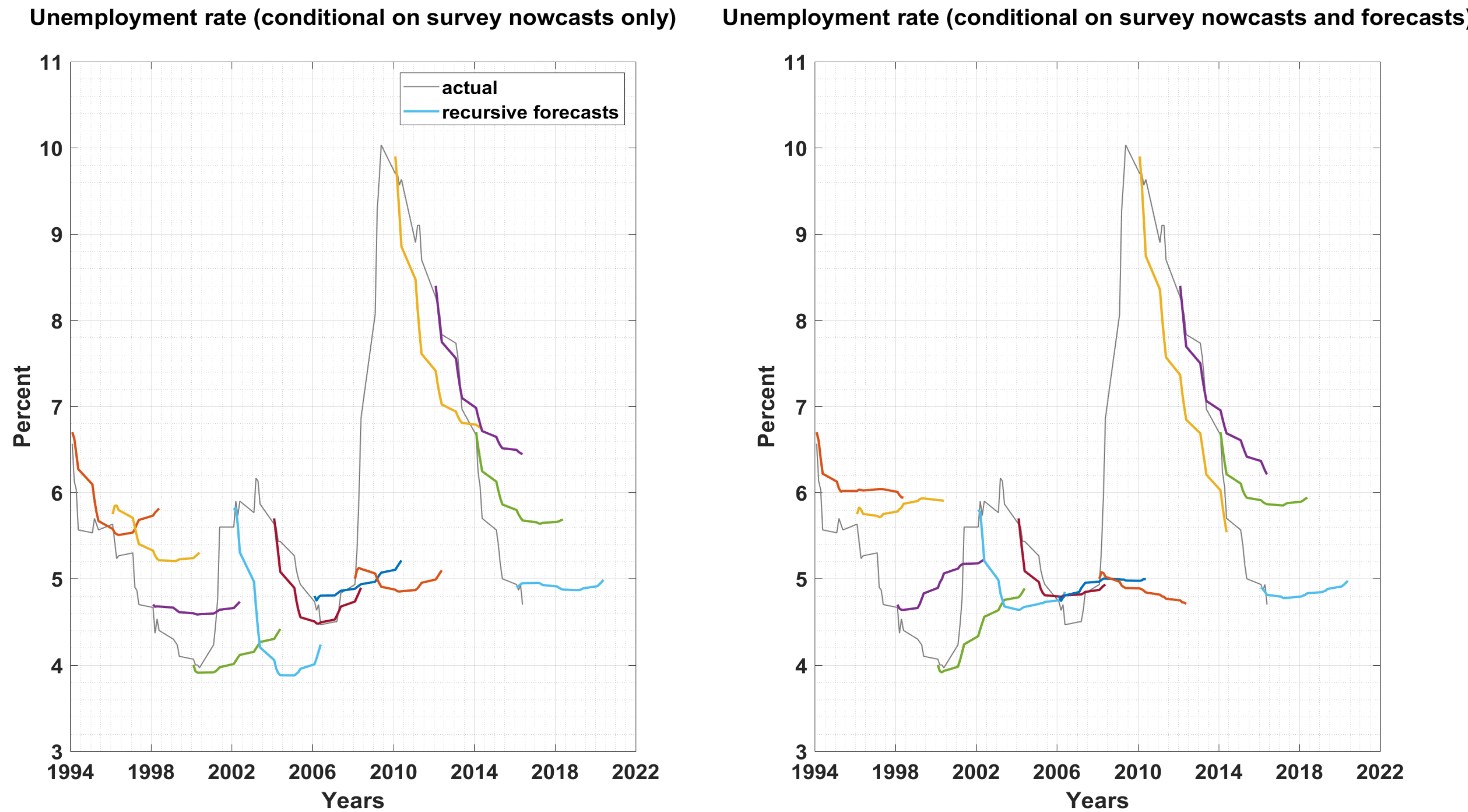
Figure 6: Recursive Point Forecasts for Federal Funds Rate





Notes: The figure plots the recursive point forecasts of the federal funds rate from the Small BVAR at different points in time. Also plotted are the actual data (solid gray line). The forecasts are recursively generated beginning in 1994.Q1 through 2016.Q4, giving us a total of 92 recursive forecasts. Each forecast is up to 20 quarters ahead. For presentation purposes every 7th forecast is displayed, i.e., 12 out of 92. The left panel corresponds to the forecast that conditions only on the survey nowcasts. The right panel corresponds to the forecast that conditions on both the survey nowcasts and the survey long-horizon forecasts.

Figure 7: Recursive Point Forecasts for Unemployment Rate



Notes: The figure plots the recursive point forecasts of the unemployment rate from the Small BVAR at different points in time. Also plotted are the actual data (solid gray line). The forecasts are recursively generated beginning in 1994.Q1 through 2016.Q4, giving us a total of 92 recursive forecasts. Each forecast is up to 20 quarters ahead. For presentation purposes, every 7th forecast is displayed, i.e., 12 out of 92. The left panel corresponds to the forecast that conditions only on the survey nowcasts. The right panel corresponds to the forecast that conditions on both the survey nowcasts and the survey long-horizon forecasts.

Table 1: Real-Time Out-of-Sample **Point** Forecasting Performance: **Small BVAR**

Full Sample (Recursive evaluation: 1994.Q1-2016.Q4)

Panel A: BVAR (Now Only) vs. Raw BVAR

Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Relative MSE: **BVAR (Now Only)** Forecast / BVAR Raw Forecast								
Real GDP	0.57***	1.01	1.06***	1.02***	1.01	1.03	1.02	1.00
CPI Inflation	0.28***	0.85***	0.88**	0.88*	0.93***	0.92**	0.99	1.00
Unemployment Rate	0.23***	0.69*	0.81*	0.88*	0.91	0.94	1.00	0.96
Federal Funds Rate	0.01***	0.66**	0.77**	0.83**	0.88**	0.91**	0.96**	0.96
Credit Spread	0.80**	0.97	0.98*	0.99*	1.01	1.02	1.04**	1.01

Panel B: BVAR (Now and LR) vs. Raw BVAR

Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Relative MSE: **BVAR (Now and LR)** Forecast / BVAR Raw Forecast								
Real GDP	0.57***	1.01	0.82	0.82*	0.89*	0.96	0.99	0.98
CPI Inflation	0.28***	0.71**	0.68***	0.61**	0.56***	0.57***	0.51***	0.51***
Unemployment Rate	0.23***	0.80**	0.86*	0.85*	0.85	0.87	1.01	1.20**
Federal Funds Rate	0.01***	0.61*	0.69*	0.69*	0.67**	0.62**	0.50**	0.45***
Credit Spread	0.80**	0.91**	0.88***	0.83***	0.81***	0.81***	0.95**	1.08***

Panel C: BVAR (Now and LR) vs. BVAR (Now Only)

Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Relative MSE: **BVAR (Now and LR)** Forecast / **BVAR (Now Only)** Forecast								
Real GDP	1.00	1.01	0.77*	0.80*	0.88	0.93	0.98	0.98
CPI Inflation	1.00	0.83*	0.78***	0.70**	0.60***	0.62***	0.52***	0.52***
Unemployment Rate	1.00	1.16	1.07	0.98	0.94	0.92	1.00	1.25***
Federal Funds Rate	1.00	0.92	0.90	0.84*	0.75**	0.69***	0.52**	0.47***
Credit Spread	1.00	0.94	0.90***	0.84***	0.81***	0.79***	0.91***	1.07***

Notes for Table: Panel A reports the relative mean squared error (MSE) defined as the MSE of the BVAR forecast conditional on survey nowcasts only divided by the MSE of the unconditional BVAR forecast; a ratio of less than 1 suggests that tilting the BVAR forecasts to survey nowcasts only is on average more accurate compared to unconditional BVAR forecasts. Panel B reports the relative MSE defined as MSE of the BVAR forecast conditional on both survey nowcasts and long-horizon forecasts divided by the MSE of the unconditional BVAR forecast; a ratio of less than 1 suggests that tilting the BVAR forecasts to survey nowcasts and long-horizon forecasts is on average more accurate compared to unconditional BVAR forecasts. Panel C reports the relative MSE defined as the MSE of the BVAR forecast conditional on survey nowcasts and long-horizon forecasts divided by the MSE of the BVAR forecast conditional on survey nowcasts only; a ratio of less than 1 suggests that tilting the BVAR forecasts to survey nowcasts and long-horizon forecasts is on average more accurate compared to tilting on just the survey nowcasts. The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistics use two-sided standard normal critical values.

Table 2: Real-Time Out-of-Sample **Density** Forecasting Performance: **Small BVAR**

Full Sample (Recursive evaluation: 1994.Q1-2016.Q4)								
Panel A: BVAR (Now Only) vs. Raw BVAR								
Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Mean (Relative CRPS Score: **BVAR Now Only** Forecast - BVAR Raw Forecast)								
Real GDP	-0.23***	0.00	0.05***	0.01***	0.01	0.02	0.00	-0.01
CPI Inflation	-0.45***	-0.08***	-0.07**	-0.08*	-0.04***	-0.05***	-0.01	0.01
Unemployment	-0.06***	-0.08**	-0.09*	-0.07*	-0.06*	-0.05	0.00	-0.02
Federal Funds	-0.12***	-0.13**	-0.12**	-0.12**	-0.10**	-0.08*	-0.05**	-0.08**
Credit Spread	-0.01**	-0.02*	-0.01*	-0.01*	0.00	0.01	0.02**	0.00

Panel B: BVAR (Now and LR) vs. Raw BVAR								
Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Mean (Relative CRPS Score: **BVAR Now and LR** Forecast - BVAR Raw Forecast)								
Real GDP	-0.22***	0.03	-0.12*	-0.10	-0.07	-0.02*	0.00	0.00
CPI Inflation	-0.44***	-0.18***	-0.19***	-0.27**	-0.31***	-0.28***	-0.38**	-0.44**
Unemployment	-0.06***	-0.06**	-0.07*	-0.08	-0.09	-0.09	-0.01	0.14**
Federal Funds	-0.12***	-0.14*	-0.15*	-0.20*	-0.26**	-0.35**	-0.66**	-0.94***
Credit Spread	-0.01**	-0.03**	-0.05***	-0.07***	-0.09***	-0.10***	-0.03***	0.03***

Panel C: BVAR (Now and LR) vs. BVAR (Now Only)								
Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Mean (Relative CRPS Score: **BVAR Now and LR** Forecast - **BVAR Now Only** Forecast)								
Real GDP	0.00	0.02	-0.17*	-0.11*	-0.08	-0.05	0.00	0.01
CPI Inflation	0.00	-0.10*	-0.12***	-0.19**	-0.28***	-0.24***	-0.37**	-0.45**
Unemployment	0.00	0.02	0.02	-0.01	-0.03	-0.04	-0.01	0.16***
Federal Funds	0.00	-0.01	-0.03	-0.07	-0.16**	-0.27***	-0.60**	-0.86***
Credit Spread	0.00	-0.01	-0.04***	-0.07***	-0.09***	-0.10***	-0.05***	0.03***

Notes for Table: Panel A reports the mean relative CRPS between the CRPS of BVAR forecast conditional on survey nowcasts and the of unconditional BVAR forecast; a negative value suggests that tilting the BVAR forecasts to survey nowcasts only is on average more accurate compared to unconditional BVAR forecasts. Panel B reports the mean relative CRPS between CRPS of BVAR forecast conditional on both survey nowcasts and long-horizon forecasts and CRPS of unconditional BVAR forecast; a negative value suggests that tilting the BVAR forecasts to survey nowcasts and long-horizon forecasts is on average more accurate compared to unconditional BVAR forecasts. Panel C reports the mean relative CRPS between CRPS of BVAR forecast conditional on survey nowcasts and long-horizon forecasts and CRPS of BVAR forecast conditional on survey nowcasts only; a negative value suggests that tilting the BVAR forecasts to survey nowcasts and long-horizon forecasts is on average more accurate compared to tilting on just the survey nowcasts. The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistic uses two-sided standard normal critical values.

Table 3: Real-Time Out-of-Sample Forecasting Performance: **Medium BVAR**

Full Sample (Recursive evaluation: 1994.Q1-2016.Q4)

Panel A: **Point** Forecast Accuracy of BVAR (Now and LR) vs. BVAR (Now Only)

Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Relative MSE: **BVAR (Now and LR)** Forecast / **BVAR (Now Only)** Forecast								
Real GDP	1.00	1.05	0.85	0.83	0.91	0.94	1.00	0.98
Real Consumption	1.06	0.97	0.97	1.00	0.98	1.04	1.05**	1.05
Core CPI Inflation	0.98	0.47*	0.59*	0.43**	0.32***	0.29***	0.18***	0.17***
CPI Inflation	1.00	0.79*	0.79***	0.72*	0.68***	0.73***	0.55***	0.46***
Productivity	0.99	1.01	1.07	0.93	1.02	1.00	0.98	1.01
Compensation	1.00	0.85**	0.70***	0.63***	0.47***	0.41***	0.27***	0.22***
Nonfarm Payroll	1.00	1.02	0.75**	0.69**	0.72**	0.74**	0.82***	0.90**
Unemployment	1.00	1.20	1.08	0.97	0.91	0.88	0.92	1.23
Federal Funds	1.00	0.88	0.97	0.94	0.88	0.81	0.57***	0.40***
Credit Spread	0.98	0.92	0.91***	0.88***	0.85***	0.83***	0.88***	1.05

Panel B: **Density** Forecast Accuracy of BVAR (Now and LR) vs. BVAR (Now Only)

	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Mean (Relative CRPS Score: **BVAR Now and LR** Forecast - **BVAR Now Only** Forecast)								
Real GDP	0.00	0.03	-0.10	-0.11	-0.06	-0.03	-0.01	-0.01
Real Consumption	0.02***	-0.01	-0.02	0.01	-0.01	0.02	0.04***	0.03*
Core CPI Inflation	0.00	-0.07*	-0.06**	-0.13**	-0.21***	-0.26***	-0.46***	-0.67***
CPI Inflation	0.00	-0.11*	-0.11**	-0.17**	-0.18***	-0.14**	-0.31***	-0.55**
Productivity	-0.02***	0.00	0.04	-0.03	0.03*	0.00	-0.02	0.01
Compensation	0.00	-0.03*	-0.07**	-0.11***	-0.18***	-0.21***	-0.52***	-0.83***
Nonfarm Payroll	0.00	0.01	-0.16**	-0.22**	-0.18**	-0.17**	-0.13**	-0.06
Unemployment	0.00	0.03	0.02	-0.02	-0.05	-0.08	-0.08	0.13*
Federal Funds	0.00	-0.01	0.02	-0.01	-0.06	-0.14	-0.57***	-1.23***
Credit Spread	0.00	-0.01	-0.03**	-0.05***	-0.06***	-0.08***	-0.07***	0.01

Notes for Table: The numbers reported in Panel A are relative mean squared errors: mean squared error conditional on nowcasts and long-horizon survey forecasts (Hybrid) / mean squared error conditional on nowcasts only (Baseline). So a ratio of less than 1 indicates that imposing the long-horizon survey forecast helps improve forecast accuracy. The numbers reported in Panel B are mean relative CRPS between CRPS of Hybrid forecast and CRPS of Baseline forecast; a negative value suggests Hybrid forecast is on average more accurate compared to the Baseline forecast The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistics use two-sided standard normal critical values. The model is a Bayesian VAR estimated with Minnesota and sum of coefficients priors.

Table 4: Real-Time Out-of-Sample Forecasting Performance: **Medium BVAR with SV**

Full Sample (Recursive evaluation: 1994.Q1-2016.Q4)

Panel A: **Point** Forecast Accuracy of BVAR with SV (Now and LR) vs. BVAR (Now and LR)

Series	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Relative MSE: **BVAR with SV (Now and LR)** Forecast / **BVAR (Now and LR)** Forecast								
Real GDP	1.00	0.95*	1.00	1.00	1.00	1.00	1.00	1.00
Real Consumption	1.07*	0.95**	0.98*	1.00	1.00	0.99	1.01	0.99
Core CPI Inflation	0.97	0.85	0.78***	0.86	0.93	1.00	0.94	0.89***
CPI Inflation	1.00	1.01	1.00	0.99	1.00	1.00	1.00	1.00
Productivity	1.00	0.97	0.98	0.98*	0.98	0.99	1.01	1.00
Compensation	1.07	1.05*	1.08*	1.00	1.06	1.03	0.95	0.85***
Nonfarm Payroll	0.95	0.87*	0.92*	0.93*	0.90*	0.92*	0.97*	0.94
Unemployment	1.00	0.93**	0.94**	0.95*	0.96	0.97	1.00	1.00
Federal Funds	1.00	0.88*	1.01	1.09	1.11	1.07	1.00	1.00
Credit Spread	1.03	1.05	1.09**	1.09**	1.09**	1.07**	1.09**	1.07**

Panel B: **Density** Forecast Accuracy of BVAR with SV (Now and LR) vs. BVAR (Now and LR)

	h=1Q	h=4Q	h=5Q	h=8Q	h=10Q	h=12Q	h=20Q	h=32Q
Mean (Relative CRPS Score: **BVAR with SV (Now and LR) - BVAR (Now and LR)**)								
Real GDP	-0.04**	-0.05**	-0.04*	-0.02	-0.02	-0.01	0.02	0.00
Real Consumption	-0.01	-0.05***	-0.02	0.00	0.01	0.00	0.02	0.00
Core CPI Inflation	-0.07***	-0.13***	-0.16***	-0.16***	-0.15***	-0.15***	-0.17***	-0.18***
CPI Inflation	0.03**	-0.02	-0.03	-0.04	-0.03	-0.04	-0.04	-0.07*
Productivity	0.00	-0.02	-0.02	0.00	0.00	-0.01	0.04*	0.03
Compensation	-0.02*	-0.04**	-0.04*	-0.07**	-0.07***	-0.09***	-0.12***	-0.16***
Nonfarm Payroll	-0.03***	-0.08***	-0.06***	-0.04**	-0.05**	-0.03	0.03***	-0.01
Unemployment	0.00	-0.02***	-0.03**	-0.03*	-0.03	-0.02	0.00	0.00
Federal Funds	-0.08**	-0.10**	-0.04	0.01	0.03	0.04	0.06*	-0.02
Credit Spread	0.00	0.01	0.03**	0.04**	0.04**	0.04*	0.08***	0.03

Notes for Table: The numbers reported in Panel A are relative mean squared errors: mean squared error conditional on nowcasts and long-horizon survey forecasts of BVAR with SV (Hybrid SV) / mean squared error conditional on nowcasts and long-horizon survey forecasts of BVAR (Hybrid). So a ratio of less than 1 indicates that Hybrid forecast with stochastic volatility is more accurate on average. Panel B reports the mean relative CRPS between CRPS of Hybrid forecast with SV and CRPS of Hybrid forecast without SV; a negative value suggests Hybrid SV forecast is on average more accurate compared to Hybrid forecast. The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistics use two-sided standard normal critical values. The model is a Bayesian VAR estimated with Minnesota and sum of coefficients priors.

Table 5: CPI Inflation Real-Time Out-of-Sample Point Forecasting Performance

Panel A: Full Sample (Recursive evaluation: 1994.Q1-2016.Q4)							
CPI Inflation	h=2Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q
MSE							
Medium BVAR (Hybrid)	4.33	4.27	4.31	4.36	4.46	4.58	4.89
Medium BVAR (Hybrid SV)	4.21	4.39	4.30	4.39	4.51	4.70	5.01
Relative MSE							
Hybrid / RW (AO)	0.80**	0.86*	0.85**	0.77	0.81*	0.89***	0.90
Hybrid / UCSV (SW)	0.97	0.99	0.98	0.94	0.93	0.99	1.00
Relative CRPS							
Baseline - UCSV	0.05	0.16**	0.16***	0.19***	0.25***	0.24**	0.43***
Hybrid - UCSV	0.00	0.04*	0.05	0.02	0.06*	0.10**	0.11
Hybrid with SV - UCSV	0.05	0.07	0.08	0.07	0.11	0.15	0.18

Panel B: Pre-Crisis Sample (Recursive evaluation: 1994.Q1-2006.Q4)							
CPI Inflation	h=2Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q	h=20Q
MSE							
Medium BVAR (Hybrid)	1.83	1.91	2.22	2.22	2.22	2.34	2.43
Medium BVAR (Hybrid SV)	1.84	1.90	2.16	2.20	2.25	2.40	2.46
Relative MSE							
Hybrid / RW (AO)	0.88*	0.82*	0.88	0.78*	0.78	0.97	1.23
Hybrid / UCSV (SW)	0.93	0.92	0.96	0.89	0.86	0.97	1.14
Relative CRPS							
Baseline - UCSV	0.01	0.05	0.09	0.13*	0.25***	0.35***	0.55***
Hybrid - UCSV	0.00	0.02	0.04	0.05	0.08	0.16**	0.27***
Hybrid with SV - UCSV	-0.01	-0.03	-0.02	-0.08	-0.01	0.04	0.12*

Notes for Table: The first two rows (in both panels) report the mean squared error (MSE) corresponding to the Medium BVAR tilted to match the SPF nowcast and the SPF long-horizon forecast (Hybrid) and the Hybrid with SV, respectively. Rows three and four in Panel A and Panel B report relative mean squared errors: mean squared error from Hybrid / mean squared error from the univariate model listed in the row. So a ratio of less than 1 indicates that tilted BVAR (Hybrid) is on average more accurate. Rows five, six, and seven in both panels report the relative CRPS of baseline, hybrid and hybrid SV relative to UCSV, respectively. A positive value suggests that the density forecast from the UCSV is more accurate on average. The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistics use two-sided standard normal critical values. All models use the SPF nowcast for the one-step-ahead forecast; as a result, the relative MSE is equal to one and is not reported.

Table 6: Real-Time Out-of-Sample Forecasting Accuracy Hybrid vs. TVP-VAR SV

Panel A: Point Forecast Accuracy (Recursive evaluation: 1994.Q1-2016.Q4)

	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q
Relative MSE: MSE **Hybrid** from Small BVAR / MSE **TVP-VAR SV**						
Real GDP	1.00	1.03	0.93	0.93*	0.95	1.03
CPI Inflation	1.00	0.99	0.94***	0.87***	0.83**	0.82**
Unemployment rate	1.00	1.06	1.05	1.02	1.02	1.02
Relative MSE: MSE **Hybrid SV** from Small BVAR / MSE **TVP-VAR SV**						
Real GDP	1.00	0.89	0.88	0.85	0.90	1.00
CPI Inflation	1.00	0.96	0.93**	0.86***	0.81***	0.80***
Unemployment rate	1.00	1.01	1.02	1.00	1.00	1.00

Panel B: Density Forecast Accuracy (Recursive evaluation: 1994.Q1-2016.Q4)

	h=1Q	h=4Q	h=6Q	h=8Q	h=10Q	h=12Q
Relative CRPS: CRPS **Hybrid** from Small BVAR - CRPS **TVP-VAR SV**						
Real GDP	0.11***	0.08*	0.02	0.02	0.04	0.07**
CPI Inflation	-0.05***	0.01	0.03	-0.01	-0.03	-0.03
Unemployment rate	0.01***	-0.01	-0.04	-0.06	-0.08	-0.08
Relative CRPS: CRPS **Hybrid SV** from Small BVAR - CRPS **TVP-VAR SV**						
Real GDP	0.02	-0.02	-0.02	0.01	0.04	0.11**
CPI Inflation	0.00	0.03	0.05	0.05	0.06	0.07
Unemployment rate	0.01***	-0.01	-0.03	-0.05	-0.06	-0.05

Notes for Table: The top panel compares the forecast accuracy of the hybrid (and hybrid SV) with that of the time varying parameter BVAR with SV (TVP-VAR SV). The first half of the top panel reports the Mean Square Error (MSE) corresponding to the Small BVAR tilted to match the SPF nowcast and SPF long-horizon forecast (Hybrid) relative to the TVP-SV VAR. The second half of the top panel report the relative mean squared errors: mean squared error from the Small BVAR with SV tilted to match the SPF nowcast and SPF long-horizon forecast (Hybrid SV) / mean squared error from TVP-VAR SV. So a ratio of less than 1 indicates that tilted BVAR (Hybrid or Hybrid SV) is on average more accurate. The bottom panel reports the corresponding density forecast accuracy performance. A negative value suggests that hybrid (or hybrid SV) density forecast is on average more accurate. The table reports statistical significance based on the Diebold-Mariano and West test with the lag $h-1$ truncation parameter of the HAC variance estimator and adjusts the test statistic for the finite sample correction proposed by Harvey, Leybourne, and Newbold (1997); *10 percent, **5 percent, and ***1 percent significance levels, respectively. The test statistics use two-sided standard normal critical values. All models use the SPF nowcast for the one-step-ahead forecast; as a result, the relative MSE is equal to one and is not reported.