Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR*

Michael W. McCracken† Michael T. Owyang‡ Tatevik Sekhposyan§

July 19, 2018

Abstract

We use a mixed-frequency vector autoregression (VAR) to obtain intraquarter point and density forecasts as new, high-frequency information becomes available. Our empirical model is specified at the lowest sampling frequency; high frequency observations are treated as different economic series occurring at the low frequency. As this type of data stacking results in a high-dimensional system, we rely on Bayesian shrinkage to mitigate parameter proliferation. We obtain high-frequency updates to forecasts by treating new data releases as conditioning information. The relative performance of the model is compared to, and shown to be competitive with, current and four-quarter-ahead forecasts from other time-series models and the Survey of Professional Forecasters' consensus forecasts.

Keywords: vector autoregression, stacked vector autoregression, mixed-frequency estimation, bayesian methods, conditional forecasts, nowcasting, forecasting

JEL Codes: C22, C52, C53

*We have benefited greatly from conversations with Bob Rasche. We are grateful to Domenico Giannone, Eric Ghysels and participants of seminars at the Bank of Canada, Federal Reserve Board, Ohio State University, University of Houston, 2013 CIRANO-CIREQ Workshop on Data Revisions in Macroeconomic Forecasting and Policy, 2013 $(EC)^2$ Conference, 2014 SNDE Conference, 2014 Midwest Macro Meeting, 2014 EABCN Conference, 2014 IAAE Meeting, 2015 Econometrics Workshop of the St. Louis Fed, 2015 Workshop on High-Dimensional Time Series in Macroeconomics and Finance, 2016 Texas Econometrics Camp, 2016 Econometric Society European Meetings, 2018 Kick off Workshop on Big Data and Macroeconomic Forecasting at JRC for comments and suggestions. Joe McGillicudy and Hannah G. Shell provided research assistance. The views expressed herein do not reflect the official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA; tel: (314) 444-8594; e-mail: Michael.W.McCracken@stls.frb.org

‡Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA; tel: (314) 444-8558; e-mail: Michael.T.Owyang@stls.frb.org

§Texas A&M University, 4228 TAMU, College Station, TX 77843, USA; tel.: (979) 862-8857; e-mail: tsekhposyan@tamu.edu

1 Introduction

Economic forecasting typically requires managing mixed-frequency data. Across any particular quarter, policymakers and professional forecasters analyze monthly, weekly, and sometimes even daily indicators of economic activity. For instance, the Federal Reserve Bank of Atlanta maintains the GDPNow series on its website, while the Federal Reserve Bank of New York regularly publishes a Nowcasting Report. Both of these forecast series use high frequency data releases to provide an updated view on the performance of the economy—often viewed through the lens of real GDP growth, which is released at a lower, quarterly frequency.

A variety of statistical methods enable the integration of high frequency variables into forecasting models that predict lower frequency variables such as real GDP growth. Corrado and Green (1988) and Parigi and Schlitzer (1995) use linear bridge equations to map monthly data into quarterly frequency when modeling GDP for the U.S. and Italy, respectively. Giannone, Reichlin, and Small (2008) and Blasques et al. (2016) use dynamic factor models to generate forecasts of U.S. GDP using monthly, weekly, and daily data releases. Zadrozny (1990) and Mittnik and Zadrozny (2005) employ a Kalman-filtering-based maximum-likelihood estimation method to modeling mixed-frequency data in a single vector autoregression (VAR) by treating the low frequency release as a missing value problem: GDP exists at a monthly frequency but we only observe it once every three months. Eraker et al. (2015), Schorfheide and Song (2015), and Brave, Butters, and Justiniano (2016) also treat the low frequency release as a missing value problem but use Bayesian methods. Foroni and Marcellino (2013) offer a review of these and other mixed-frequency models.

In this paper, we address the forecaster's mixed-frequency problem using a mixed-frequency vector autoregression estimated at the lowest common data frequency. As an example, suppose we want to forecast real GDP growth using the three monthly nonfarm payroll employment data releases that occur during the quarter. In this case, our VAR would be based on a four dimensional vector formed by the three monthly and one quarterly series. Ghysels (2016) and Bacchiocchi et al. (2016) have recently used such a model to investigate the role mixed-frequencies play in tests of Granger causality and in the construction of impulse response functions, respectively. In other fields—particularly engineering, theoretical aspects of this modeling approach (sometimes referred to as "blocking") have been analyzed [see Bittanti et al. (1988) and, more recently, in Chen et al. (2012)]. Our model can also be interpreted as a multivariate version of the univariate, unrestricted MIDAS model discussed in Foroni et al. (2015).

We evaluate our mixed-frequency VAR based on its ability to forecast in real time as we move within a quarter and higher frequency data becomes available. Our analysis focuses on the nowcasting problem (i.e., forecasting the current quarter), although we also provide results for a longer, four-quarter-ahead horizon. In our empirical implementation, we use real-time vintage data on twelve monthly frequency predictors and quarterly GDP, yielding a heavily parameterized, 37 dimensional VAR.[1] To handle the high dimensionality, we estimate the model using Bayesian techniques, where we allow data-driven shrinkage to resolve the bias-variance trade-off embedded in the forecasting problem. Specifically, we use an off-the-shelf estimation procedure for reduced-form VARs provided by Giannone, Lenza, and Primiceri (GLP; 2015) to obtain the posterior distribution of the parameters and the predictive densities.

To understand the nature of our model, it is instructive to compare it with the one posited by Carreiro, Clark, and Marcellino (CCM; 2015). In their model, CCM construct nowcasts of real GDP growth based on monthly indicators using the unrestricted MIDAS approach delineated in Foroni et al. (2015).[2] Among other results, they use four separate linear scalar models to nowcast current-quarter real GDP growth from information available as of the first week in each month—i.e., after the release of the Employment Situation Report. This direct multi-step (DMS) approach to GDP forecasting has the advantage of not having to form a complete model for all variables in the system as one would with a VAR. Bhansali (1997) and Schorfheide (2005) provide theoretical results showing that DMS approaches to forecasting can be more robust to model misspecification than iterated multi-step (IMS) forecasts generated by a fully-specified VAR. That said, Marcellino et al. (2006) show that empirically, IMS-based models typically lead to more accurate predictions relative to DMS-based models. Our model can be viewed as the VAR-based IMS analogue to the DMS system of equations that CCM use to forecast. Though working with a fully-specified VAR notably complicates the empirics, it comes with the benefit of providing a more general framework that can be used for multivariate forecasting and scenario-based conditional forecasting of the type often used by central banks.

Given the strong relationship between our model and the one in CCM, in our nowcasting exercises we compare our approach to theirs while also noting some advantages to our approach. To do so, we use the same monthly predictors, the same data release schedule, consider the same intraquarter forecast origins (as well as others) and set of models (small and large) that are comparable with their specification. With this structure in place, we directly compare our IMS approach to current-quarter point and density nowcasts of real GDP growth to CCM's DMS approach. We also compare these nowcasting results to SPF forecasts and predictions from a quarterly frequency AR(2). In addition, we use our mixed-frequency Bayesian VAR (MF-BVAR) to produce point and density forecasts of GDP growth and several of our high frequency variables at a longer, four-quarter-ahead horizon and compare the efficacy of these forecasts to those from the SPF.

We find that our modeling approach provides competitive nowcasts and four-quarter-ahead forecasts. The nowcasts, in general, are comparable to the ones obtained using our version of

[1] Prior to 1992, we forecast real GNP growth. Throughout the remainder we simply reference our target variable as GDP without also referencing GNP.

[2] The model used by CCM also introduces stochastic volatility, an issue we do not address.

the model proposed in CCM and can be as accurate as SPF at certain forecast origins. The four-quarter-ahead forecasts, on the other hand, are as good or better than those in the SPF for variables such as industrial production, housing starts and Treasury yields. In general, holding the dimension of the model fixed, more information improves the models in terms of both point and density forecasts. Even so, it is sometimes the case that a lower dimensional MF-BVAR, consisting of fewer stacked monthly series, provides more accurate point forecasts than those from a larger MF-BVAR with many stacked monthly series.

The remainder of the paper is organized as follows. Section 2 lays out the specification of our model, discusses the data, and delineates the estimation methodology. Section 3 describes the construction of point and density forecasts; Section 4 discusses the forecasting results and compares them to various alternatives. We conclude with Section 5.

2 The Setup

As noted in the introduction, our model is a reduced-form VAR, where the high-frequency (in our case, monthly) variables are stacked such that each monthly data series is represented by three quarterly variables within a common calendar quarter. We follow the estimation procedure of GLP, where we estimate the VAR with Minnesota priors, reconfigured for use with stationary data, and shrinkage is chosen optimally, such that it maximizes the marginal data density. In what follows, we describe the data and the real-time properties of the various series included in our VAR specification. We then provide more details on the stacked nature of the VAR and outline the GLP estimation procedure for completeness.

2.1 Data

Our choice of predictors aims to facilitate a close comparison of our forecasting results with other mixed-frequency models in the literature, particularly the ones in CCM. The latter, in turn, includes variables that have proven to be useful for forecasting U.S. real GDP growth (or GNP for earlier portions of our sample) and are followed by markets and policymakers. The complete set of our monthly predictors, their transformations and mnemonics are as follows: the S&P 500 composite index (log-change, "stprice"), the 3-month Treasury bill rate ("tbill"), the 10-year Treasury bond yield ("tbond"), the Institute for Supply Management manufacturing index ("ISM"), the ISM supplier deliveries index ("supdel"), the ISM new orders index ("orders"), total nonfarm payroll employment (log-change, "emp"), average weekly hours of production and supervisory workers (log-change, "hours"), real retail sales (the nominal series is deflated using the consumer price index; log-change, "RS"), industrial production (log-change, "IP"), housing starts (natural log, "starts"), and finally initial unemployment claims ("claims").[3] The transformations are chosen to induce stationarity in the series. If no transformation is listed, the variable is used in levels. In addition, growth rates have been annualized.

Because our model treats higher frequency series as multiple quarterly frequency series, the number of estimated parameters in our model grows faster than in a standard VAR of equal lag order. In order to mitigate parameter proliferation, we consider data only in monthly and quarterly frequencies. Financial variables are summarized at a monthly frequency, constructed as averages of daily observations. For similar reasons we do not use weekly releases of initial claims, instead, we choose the four-week moving average. The twelve monthly series, along with the quarterly GDP series (log-change), imply that our large mixed-frequency VAR has $12 \times 3 + 1 = 37$ dimensions. In order to parallel CCM results more closely, we also investigate the performance of a small mixed-frequency VAR that only uses five of the monthly variables yielding 16 dimensions.

While not necessary for estimating the model, we choose to organize our data based on the approximate release calendar. Hence, since the longest publication lag we have across our data series is one month, at the end of each month we have complete realizations for all the information associated with the previous month. To facilitate comparison with CCM, within each month we order the data as follows: (i) the S&P 500 composite index, 3-month Treasury bill rate, and 10-year Treasury bond yield are monthly averages released on the first day of the following month, (ii) the ISM manufacturing, supplier deliveries, and new orders indices are then released immediately prior to (iii) total nonfarm payroll employment and average weekly hours, followed by (iv) real retail sales, (v) industrial production, (vi) housing starts, and finally (vii) the four-week moving average of initial unemployment claims. GDP growth is observed following initial unemployment claims.

In all of our forecasting exercises, we use real-time monthly vintage data starting in January 1985 and ending in April 2017. The monthly sequence of 388 real-time vintages of our predictors and GDP are gathered from Haver Analytics, the ALFRED database hosted by the Federal Reserve Bank of St. Louis, and the Real-time data set for macroeconomists hosted by the Federal Reserve Bank of Philadelphia. Each vintage consists of observables dating back to the first quarter of 1970. Accordingly, the target quarters for nowcasting and forecasting span from 1985:Q1 to 2017:Q1.

2.2 Model

Our forecasting model is based on a standard reduced-form VAR, estimated at the lowest sampling frequency. To accomplish this using mixed frequency data, we treat multiple releases at the highest

[3] In 2001, the Census changed details in the construction of retail sales (RETAIL) and started releasing the new version (RSAFS). Therefore, we use RETAIL for all vintages up to 2001:06 and RSAFS for the vintages after. Morover, when RSAFS was released, the historical sample was extended back only to 1992. Therefore, when we use the RSAFS vintages, we splice the pre-1992:01 values from the last vintage of RETAIL in order for RSAFS to have data dating back to 1970.

frequency as separate observations modeled in a blocked linear form. For illustrative purposes, consider the case of one quarterly variable (e.g., real GDP) and one monthly variable (e.g., payroll employment) released in each of the three intraquarter months. Let $x_{t-\tau}$ represent the high frequency (monthly) variable and y_t denote the low frequency (quarterly) variable, for quarters $t = 1, ..., T$ and months within the quarter $\tau = \{0, 1/3, 2/3\}$. In this setup, each τ represents an intraquarter month, hence $x_{t-2/3}$, $x_{t-1/3}$, and x_t index data that are observed during the second and third calendar months of quarter $t-1$ and first calendar month of quarter t, respectively.

Define the vector of data releases as $Y_t = [x_{t-2/3}, x_{t-1/3}, x_t, y_t]'$. The reduced-form VAR is

$$Y_t = C + B_1 Y_{t-1} + ... + B_p Y_{t-p} + \varepsilon_t, \tag{1}$$

where B_l are $n \times n$ parameter matrices, p is the lag order of the VAR, C is a $n \times 1$ vector of intercepts, and $\varepsilon_t \sim N(\mathbf{0}, \Sigma)$. Let $X_t = [Y'_{t-1}, ..., Y'_{t-p}, 1]'$. We can then write (1) as

$$Y_t = DX_t + \varepsilon_t, \tag{2}$$

where $D = [B \; C]$ and $B = [B_1 \; ... \; B_p]$. It will further be useful to define $\beta = vec(D)$. In our empirical section, we use a lag order of $p = 1$ as selected recursively by BIC as well as a desire to align the lag structure with that in CCM.

In general, the model is flexible enough to be generalized to include Q quarterly variables and M monthly variables. In such an environment one would treat y_t as a $(Q \times 1)$ vector and $x_{t-\tau}$ as a $(M \times 1)$ vector, producing a VAR of dimension $n = Q + 3M$.

2.3 Priors and Estimation

The dimension of our VAR increases quickly: An additional monthly predictor adds three variables to the VAR. To handle the parameter proliferation problem, we estimate the model using Bayesian methods and utilize a hierarchical shrinkage prior. Bańbura et al. (2010) show that Bayesian VARs can forecast well, even with 100+ variables, when the shrinkage is chosen appropriately such that the prior tightness increases with the model size. Thus, given the size of our model, a careful choice of prior hyperparameters is important. We use the procedure outlined in GLP to choose hyperparameters that maximize the marginal data density for each estimation sample.

We employ a standard Normal-Inverse-Wishart prior distribution for the VAR parameters. More formally, let the priors be defined as

$$\Sigma \sim IW(\Psi, d),$$
$$\beta | \Sigma \sim N(0, \Sigma \otimes \Omega),$$

where the degrees of freedom parameter of the Inverse-Wishart distribution $d = n+2$, the minimum value that guarantees the existence of the prior mean for Σ. Ψ is a diagonal matrix where each element of the diagonal is set to the residual variance of an AR(1) process for the respective variable in the VAR.[4] Ω is a $k \times k$ matrix, $k = np + 1$, parameterized such that the prior covariance of the regression coefficients takes the following form:

$$cov\left((B_s)_{ij}, (B_r)_{hm} | \Sigma\right) = \begin{cases} \lambda^2 \frac{1}{s^2} \frac{\Sigma_{ih}}{\psi_j/(d-n-1)} & if \quad m = j \text{ and } r = s \\ 0 & otherwise \end{cases}.$$

Thus, while the coefficients $B_1, ..., B_p$ are assumed to be independent of each other, coefficients associated with the same variable are allowed to be contemporaneously correlated across different equations. In general, the prior imposes a tighter variance on the distant lags, though, given that in our specification $p = 1$, this feature of the prior is not relevant. The hyperparameter λ governs the overall tightness of the prior by controlling the scale of the variances and covariances of the VAR coefficients. The prior variance on the constant term C is diffuse.

This prior is standard, taken directly from GLP, but reconfigued for stationarity with VAR parameter means set to zero, and we rely on their procedure and accompanying codes to estimate the VAR such that the hyperparameters (in our case λ) impose an optimal amount of shrinkage consistent with the marginal data density criterion. GLP simulate the posterior distribution of λ based on a standard Metropolis algorithm, under the assumption that the prior for λ follows a Gamma distribution with a mode of 0.2 and standard deviation of 0.4, values consistent with those in Sims and Zha (1998). Given the posterior of the hyperparameter λ, we obtain the posterior distribution of the VAR parameters (β and Σ) by drawing from Normal-Inverse-Wishart posterior distributions implied by the conjugate prior. Our reported results are based on 5,000 draws.

The model is estimated every month given the new vintage of real-time data. When constructing the forecasts, we use the posterior distribution of the parameters obtained from the past vintage of monthly data. We hold the posterior constant within a month. For example, analysis over the first calendar quarter would unfold as follows. On January 1, we estimate the model using the vintage of data ending on December 31. The last row of this vintage will have a ragged edge because fourth quarter data continue to be released during January. We drop the ragged edge and estimate the model. Draws from this posterior are used throughout January. To ensure the real-time nature of the exercise, as we move across data releases within January, new vintages of monthly data supplant those available in the December 31st vintage and are used as predictors when forming the forecasts. A similar pattern of updating the posterior and constructing forecasts consistent with the real-time nature of the data continues as we move across February and March.

[4]Limited robustness analysis shows that treating diagonal elements of the scale matrix Ψ as hyperparameters does not generate meaningful gains; thus, we resort to GLP's default implementation of the Minnesota prior.

3 Forecasting

We formulate the intraquarter forecasting problem in a conditional forecasting framework. We treat quarter t information set as complete when the first (Advance) release of GDP is released. This happens at the end of the first calendar month in quarter $t+1$. At this point, we construct unconditional h-step-ahead point and density forecasts. As we move across the quarter, high frequency data becomes available. We thus update the h-step-ahead unconditional forecasts using existing results on conditional forecasting as presented in Waggoner and Zha (1999). The approximate release schedule of the high frequency data, which determines the sequence by which we update the forecasts, is delineated in Section 2.1 and depicted in Figure 1.

As Figure 1 shows, our results are based on forecast updates across four months. In the context of forecasting current, first quarter GDP, we begin with a two-quarter-ahead forecast formed conditional on all previous quarter information up to and including the first data release of January. We then update that forecast until obtaining end-of-January claims. When we observe the previous quarter GDP, we now form the one-quarter-ahead forecast using a complete set of previous quarter data. As we move into February, we then update this one-quarter-ahead forecast sequentially across each data release. This continues into March and finally ends with end-of-April claims. All together we obtain a sequence of 49 intraquarter forecasts for first quarter GDP based on the 12 monthly releases in each month and the Advance release of previous quarter GDP.

3.1 Point Forecasts and Predictive Densities

Our forecasting model is estimated using Bayesian techniques, providing a natural characterization of uncertainty. The time t predictive density of Y at horizon h, $p(Y_{t+h}|\mathbf{Y_t})$, is

$$p(Y_{t+h}|\mathbf{Y_t}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(Y_{t+h}|\mathbf{Y_t}, D, \Sigma)p(D|\Sigma, \mathbf{Y_t})p(\Sigma|\mathbf{Y_t})dDd\Sigma, \tag{3}$$

where $\mathbf{Y_t} = [Y_1, ..., Y_t]'$ represents the history of observables up to time t and $p(D|\Sigma, \mathbf{Y_t})$ and $p(\Sigma|\mathbf{Y_t})$ are the time t posterior distributions of the parameters in D and Σ, respectively.

For each saved draw, $D^{(i)}$ and $\Sigma^{(i)}$, from their respective posteriors, we obtain a forecast draw $\hat{Y}^{(i)}_{t+h|t}$ from the conditional predictive density $p(Y_{t+h}|\mathbf{Y_t}, D^{(i)}, \Sigma^{(i)})$. Collecting these draws across MCMC iterations yields the predictive density that accounts for the uncertainty in the estimated parameters (including the uncertainty associated with the set-up of the prior tightness) and the uncertainty from the unobservable future shocks. Then, based on arguments made in Gneiting (2011), we obtain the point forecast as the mean of the predictive density.

The posterior distributions $p(D|\Sigma, \mathbf{Y_t})$ and $p(\Sigma|\mathbf{Y_t})$ are readily available as a result of the reduced-form VAR estimation in GLP framework. Next, we consider the simulation from the conditional predictive density $p(Y_{t+h}|\mathbf{Y_t}, D^{(i)}, \Sigma^{(i)})$, keeping in mind that this distribution changes with each new high-frequency release.

3.2 Conditional Predictive Density Simulation

In this section, we describe how we obtain a forecast draw $\hat{Y}^{(i)}_{t+h|t} \sim p(Y_{t+h}|\mathbf{Y_t}, D^{(i)}, \Sigma^{(i)})$, conditional on the ith draw of the VAR parameters obtained from the MCMC. In what follows, we suppress the superscript i denoting the MCMC iteration for notational simplicity. As we intimated above, our forecasting procedure has two components, each of which is based on the composition of the information set at the time the forecast is constructed. Specifically, at the time the last-quarter GDP value is released and the information set is deemed complete, the forecast is constructed as an *unconditional* forecast. As the quarter progresses, the forecast is constructed as a *conditional* forecast, where the already-released intraquarter data are treated as restrictions in the forecasting model. In what follows, we delineate our unified approach for producing both h-period-ahead conditional and unconditional forecasts based on well-established results in Waggoner and Zha (1999), implemented with computational simplifications proposed in Jarociński (2010).

Since our forecasting approach is an iterative (rather than direct) one, we demonstrate how to draw $\hat{Y}_{t+h|t}$ for a general h-steps-ahead horizon, assuming we have already obtained draws for the forecasts $\hat{Y}_{t+h-1|t}, ..., \hat{Y}_{t+1|t}$. Let $\mu = (I - \Sigma_{l=1}^{p} B_l)^{-1}C$ represent the mean of Y_t implied by the VAR; then, $Z_t = Y_t - \mu$ is the demeaned vector of period$-t$ observables and $\hat{Z}_{t+h|t} = \hat{Y}_{t+h|t} - \mu$ are the demeaned forecast draws. Let $\hat{\mathbf{Z}}_{t+h-1|t} = [\hat{Z}'_{t+h-1|t}, \hat{Z}'_{t+h-2|t}, ..., \hat{Z}'_{t+h-p|t}]'$. When $h - j \leq 0$, $\hat{Z}_{t+h-j|t} = Z_{t+h-j}$, i.e. it represents observed data.

We are interested in the joint distribution of the one- to h-step-ahead forecasts of Y obtained at time t. Let $\mathbf{Y} = [Y'_{t+1},\ ...,\ Y'_{t+h}]'$, $\hat{\mathbf{Y}}^u = [\hat{Y}^{u\prime}_{t+1|t},\ ...,\ \hat{Y}^{u\prime}_{t+h|t}]'$, and $\hat{\mathbf{Z}}^u = [Z_t, \hat{Z}^{u\prime}_{t+1|t}\ ...,\ \hat{Z}^{u\prime}_{t+h-1|t}]'$, where superscript u denotes unconditional forecasts. Define $\Phi_j \Sigma^{1/2}$ as the matrix of orthogonalized impulse responses after $j = 1, ..., h$ periods and let

$$R_{nh\times nh} = \begin{pmatrix} \Sigma^{1/2} & 0 & & 0 & 0 \\ \Phi_1\Sigma^{1/2} & \Sigma^{1/2} & & 0 & 0 \\ & & ... & \Sigma^{1/2} & 0 \\ \Phi_{h-1}\Sigma^{1/2} & \Phi_{h-2}\Sigma^{1/2} & & \Phi_1\Sigma^{1/2} & \Sigma^{1/2} \end{pmatrix},$$

where $\Sigma^{1/2}$ can be obtained as a Choleski factor of Σ.

Suppose we know a subset m of the future values of $\mathbf{Y}$. Let ω be a $(m \times nh)$ selection matrix, such that $\omega\mathbf{Y}$ results in an $m \times 1$ vector, consisting of the known future elements of $\mathbf{Y}$. Let $r = \omega(\mathbf{Y} - \hat{\mathbf{Y}}^u)$. Then the known future values can be thought of as linear restrictions on future values of the error term $\varepsilon = [\varepsilon'_{t+1}, ..., \varepsilon'_{t+h}]' = (\mathbf{Y} - \hat{\mathbf{Y}}^u)$ of the form $\tilde{R}(I_h \otimes \Sigma^{-1/2})\varepsilon = r$.

Here, $\tilde{R}$ is a matrix of dimension $(m \times nh)$ and is formed by those rows in R that are associated with conditioning variables, i.e. elements of $\mathbf{Y}$ treated to be observable (in our case this would be the high frequency releases). For example, suppose we have one monthly and one quarterly variable and hence, for our model, $n = 4$. Now suppose we are forming a two-quarter-ahead forecast ($h = 2$) of GDP growth and we have observed the first two monthly releases in the current quarter. $\tilde{R}$ is the (2×8) matrix formed by stacking the first and second rows of R.

We can then draw the h-period-ahead conditional forecast $\hat{\mathbf{Y}}^c = [\hat{Y}^{c\prime}_{t+1|t}, \ \ldots, \ \hat{Y}^{c\prime}_{t+h|t}]'$ from the following conditional predictive distribution:

$$\hat{\mathbf{Y}}^c \sim N\left(\psi_{t+h|t}, \Psi_{t+h|t}\right), \tag{4}$$

where $\psi_{t+h|t}$ captures the conditional mean, while $\Psi_{t+h|t}$ is the conditional variance of the predictive density determined as

$$\begin{aligned} \psi_{t+h|t} &= \iota_h \otimes \mu + (I_h \otimes B)\hat{\mathbf{Z}}^u + \tilde{R}'(\tilde{R}\tilde{R}')^{-1} r, & (5)\\ \Psi_{t+h|t} &= R\left(I - \tilde{R}'(\tilde{R}\tilde{R}')^{-1}\tilde{R}\right)R', & (6)\end{aligned}$$

and ι_h is an $h-$dimensioned vector of ones.

The first two terms of (5) are invariant to whether we have observed any of the intraquarter data—i.e., whether the forecast is conditional or unconditional. The third term captures the adjustment to the mean as more information, either arising from intraquarter data releases or hypothetical scenarios about future path of the variables, becomes available. In the absence of any conditioning information, model residuals are unrestricted, $\tilde{R} \equiv 0$, and this third term does not affect the mean of the predictive density. In contrast, when conditioning information exists, this third term provides the adjustment needed to revise the unconditional forecast to a conditional one. As equation (6) indicates, conditioning information affects the variance of the predictive distribution as well, thus changing the characterization of forecast uncertainty arising due to unobservable future shocks in the model's reduced form errors. Conditioning information restricts the model residuals ($\tilde{R} \not\equiv 0$) and, in general, reduces the uncertainty of the predictive density.

We assume that the conditioning variables do not change the posterior distribution of the parameters and simulate the predictive density by relying on the computational simplifications provided in Jarociński (2010).[5] Let USV' denote the singular value decomposition of $\tilde{R}$. Define E as the $m \times m$ diagonal matrix of with m singular values of $\tilde{R}$ and let V_1 and V_2 denote matrices formed from the first m and remaining $nh - m$ columns of V respectively. Under our assumptions, Jarociński (2010) shows that $V_1 E^{-1} U' r + V_2 \eta$ for $\eta \sim N(0, I_{(nh-m \times nh-m)})$ has the same normal distribution as that in (4) and is computationally more efficient in many cases.

[5] Waggoner and Zha (1999) propose a Gibbs sampler that allows the relaxation of this assumption.

3.3 Competing Models

When forecasting GDP growth, we consider a few alternatives to our mixed-frequency Bayesian VAR. A simple and competitive alternative that has also been considered by CCM is an AR(2), which we reestimate each month with each new vintage of GDP. In doing so, we use the same prior and estimation procedure proposed by GLP and used for our MF-BVAR.

We then also report results associated with the DMS approach to forecasting developed in CCM. As noted previously, CCM models GDP growth directly and reports results on monthly nowcasts of GDP growth using the information available at the time of the Employment Situation Report in each month. As such, the model size changes over the quarter: in month one of the quarter the model is larger in size, while in month two it is the smallest. Table 1 in CCM describes the models explicitly. We extend the analysis in CCM to forecast origins other than the employment release. In particular, we consider origins associated with each data release using the schedule detailed in Section 2.1. This provides us with the ability to compare our IMS- and their DMS-based approaches to forecasting at each data release. The CCM-type models are again estimated using the GLP code described earlier, adjusted to work with autoregressive distributed lag type models. It is worth emphasizing that this will imply differences between our results and those obtained by CCM. In particular, our prior on quarterly and monthly series is symmetric, while CCM impose a distinct prior on series that are at the quarterly versus monthly frequencies. Moreover, our prior is optimized for each estimation sample consistent with the GLP procedure, while the CCM prior is fixed.

In addition, we consider both small and large versions of our model and that of CCM. These models are defined relative to those described in CCM. Our large model uses all twelve monthly predictors and one quarterly predictor to form a 37 dimensional VAR. These variables are described in Section 2.1 and Figure 1. Our small model only considers one quarterly and five monthly predictors including the ISM manufacturing index, payroll employment, industrial production, retail sales, and housing starts, and thus forms a 16-dimensional VAR.

Finally, we compare the accuracy of our small and large models with the mean of the responses from the Survey of Professional Forecasters. We do this for GDP growth forecasts, as well as for forecasts of some of the monthly variables in our system. For the GDP growth forecasts, we are able to do this directly because annualized quarterly GDP growth is a component of our model and is also part of the survey. For the monthly variables, especially those in log-levels or log-differences, the comparison is more complicated because the survey provides forecasts of quarterly aggregates. For example, the SPF forecasts of the 3-month and 10-year Treasury yields are quarterly averages of the daily series while SPF forecasts of retail sales, industrial production, and housing starts are quarterly averages of the monthly series. Our forecast of the 3-month and 10-year quarterly yield is formed by taking the average of the three months within the target quarter. Our forecast of the quarterly level of housing starts is formed by exponentiating the forecasts of log-starts and then averaging the three relevant months in the target quarter. Finally, our forecasts of the quarterly level of industrial production and retail sales are formed by extrapolating from the current level of the series based on the forecasted monthly growth rates at the one- through four-quarter-ahead horizons—and then averaging the three relevant months of the target quarter.[6] The remaining monthly series are either not in the SPF or, in the case of employment, have only been part of the survey for a brief time and hence are not included in the evaluation exercise.

3.4 Evaluation

We evaluate point forecasts using root-mean-squared error (RMSE). We calculate RMSEs after each intraquarter data release, obtaining a term structure of RMSEs as we move across the quarter. We consider two out-of-sample evaluation periods. The first evaluation period is the same as that in CCM, where we use real-time data to obtain nowcasts for the Advance release of GDP growth between 1985:Q1-2011:Q3. We also consider a comparably sized out-of-sample period ranging from 1992:Q2-2017:Q1, which has the advantage that it allows us to focus exclusively on forecasting real GDP growth rather than a mixture of GNP and GDP growth depending on the vintage.

We then evaluate the accuracy of our density forecasts based on the continuous ranked probability score (CRPS). Relative to other scoring functions, such as the log-scores, CRPS is less sensitive to outliers and puts higher weight on draws from the predictive distribution that are close to but not equal to the outcome [see Gneiting and Raftery (2007) and Gneiting and Ranjan (2011) for further discussion]. Similar to the RMSE, the CRPS is defined such that the lower the value, the better the score, and is given by

$$\begin{aligned} CRPS_t(y_{t+h}) &= \int_{-\infty}^{\infty} \left(P(z) - 1\{y_{t+h} \leq z\}\right)^2 dz \\ &= E_p \left|\hat{y}_{t+h|t} - y_{t+h}\right| - \frac{1}{2} E_p \left|\hat{y}_{t+h|t} - \hat{y}'_{t+h|t}\right|, \end{aligned} \tag{7}$$

where $P(.)$ denotes the cumulative distribution function associated with the predictive density $p(y_{t+h}|\mathbf{Y}_t)$, $1\{y_{t+h} \leq z\}$ denotes the indicator function, taking value 1 if the outcome $y_{t+h} \leq z$ and 0 otherwise, and $\hat{y}_{t+h|t}$ and $\hat{y}'_{t+h|t}$ are independent random draws from the conditional predictive density $p(y_{t+h}|\mathbf{Y}_t)$. We compute the CRPS using the empirical CDF-based approximation given in equation (9) of Krueger, et al. (2017). As for the case of RMSEs, we obtain average CRPS values for each intraquarter data release and report the term structure of CRPS values for each out-of-sample period.

[6] Industrial production re-bases itself seven times across all of our vintages. In order to ensure that the levels at the forecast origin align with those in the target quarter, we unwind the new base year back to the base year at the forecast origin.

Pairwise differences in RMSEs and CRPS values across models are evaluated using a standard normal approximation to t-type tests of predictive ability akin to that developed in Diebold and Mariano (1995). Newey and West (1987) standard error estimates are used with lag-orders equal to the forecast horizon plus one. All point and density forecasts of GDP are evaluated using the Advance release. Point and density forecasts of the monthly series are evaluated using the same vintage of data used to evaluate the GDP forecasts.

4 Forecasting Results

In this section, we delineate some of the advantages of our blocked, mixed-frequency VAR in the context of forecasting. We begin by evaluating the real-time accuracy of our point and density forecasts relative to a handful of competitors outlined in Section 3.3. We use figures to present measures of accuracy across all intraquarter horizons. For each figure, we also have a table that presents the value of the measure of accuracy, the ratio of these measures across models, and pairwise tests of equal accuracy – but only for those forecast origins associated with the Employment Situation Report. In our sequencing of data releases, this lines up with "hours." We then illustrate the usefulness of our model for producing scenario-based forecasts of low frequency variables when the conditioning variable is observed at a higher frequency.

4.1 Current Quarter Forecasts

In Figure 2, we plot the term structure of RMSEs for GDP growth nowcasts from each of the models as we move across the intraquarter forecast origins. The upper panel shows the results for the subsample used by CCM in which nowcasts are generated for 1985:Q1 through 2011:Q3, while the lower panel shows the results for the pure GDP subsample—i.e., nowcasts for 1992:Q2 through 2017:Q1. For comparison with the results in CCM, we begin forecasting on the first day of the first calendar month of quarter $t+1$ (e.g., January 1 when forecasting Q1 GDP). We then proceed across all data releases until the last high frequency release prior to the release of Advance GDP (e.g., late April when forecasting Q1 GDP). Each subfigure contains paths associated with six forecasts: the SPF, an AR(2), both small and large versions of the CCM models, and both small and large versions of our MF-BVARs. For both large models, we have 49 intraquarter updates associated with three monthly releases of the 12 monthly indicators and the quarter$-t$ GDP Advance release at the end of the first calendar month. The small models, on the other hand, are updated only at a subset of these forecast origins. More specifically, we obtain 16 updates to the forecasts, associated with three monthly releases of five monthly indicators and the quarter$-t$ GDP Advance release at the end of the first calendar month. Between updates we simply flatline the RMSE path. The AR(2) model updates it's forecasts at the end of each month—i.e., after a new release or a revision to previously released GDP numbers. The current quarter SPF is timed to arrive after the Employment Situation Report but prior to the release of retail sales. In the first calendar month of the quarter $t+1$, we report RMSEs for the one-quarter-ahead SPF forecasts released in the second calendar month of the previous quarter.

Perhaps the most obvious fact across both evaluation samples is that the SPF point forecasts of Advance GDP growth are extremely difficult to beat. Further, in the first subsample the best model-based forecasts tend to be those from the small-models—both CCM and the MF-BVAR, but even these are comparable to the SPF only early in the second month and at the end of the fourth month. In contrast, both of the large models generally improve in performance across the first two calendar months of the quarter, but then either stagnate or even deteriorate as we get closer to the GDP release date. For reasons that are not obvious to us, the RMSEs of the direct multi-step models both deteriorate sharply in month 2 before improving as we move across the remaining intraquarter releases. The AR(2) performs worse than the SPF, but is competitive to the rest of the forecasts. The larger models tend to have better performance relative to the AR(2) after the second month of the calendar quarter.

The models seem to perform much better relative to the SPF when we move forward into the pure GDP subsample (i.e., 1992:Q2 - 2017:Q1). As before, both small models are competitive with the SPF early in the second month and at the end of the fourth month but are now also competitive over a broader stretch of the quarter. While the large MF-BVAR is generally dominated by its smaller version, the large MF-BVAR has improved substantially and is competitive with the SPF in all periods except early in the first month. Relative to the first evaluation sample, the MF-BVAR models generally outperform the DMS models in terms of their point forecasts.

We assess the statistical significance of these results in Table 1. The diagonals in the table have two numbers that are obtained as of the release of the Employment Situation Report: (i) the RMSE (below the slash) and (ii) the CRPS value (above the slash). For example, the second diagonal entry in the top panel, 2.02\3.42, indicates that for this sample the AR(2) has an RMSE of 2.02, while the average CRPS is 3.42. The lower off-diagonal portion of each panel in the table reports the ratio of RMSEs in the row-model to the column-model, where numbers greater than one indicate that the column-model is nominally more accurate. Ratios in bold are statistically different from one at the 5% significance level. The majority of the statistically significant pairwise RMSE comparisons arise when comparing model-based forecasts to the SPF. This is particularly true in month 2. Across models, however, there are few differences that are statistically significant. The few that arise tend to do so later in the quarter and simply reinforce what we observed in the figure: the small models tend to be more accurate than the large models.

In Figure 3, we provide the same type of term structure but applied to mean CRPS values for each estimated model (and hence not the SPF) across both evaluation samples. Given the reduction in RMSEs in the second panel of Figure 2, it is not surprising to see that the CRPS values are lower for all models and all horizons. In broad terms, the CRPS paths of all models decline as we move across forecast origins regardless of evaluation sample. Again, the exception is that both of the DMS models experience a sharp deterioration in mean CRPS values as we move into month 2. Interestingly, while the large MF-BVAR did not generally perform the best among the models in terms of RMSEs, it typically has the lowest CRPS values across all intraquarter forecast origins.

In Table 1, the upper off-diagonal portion of each panel reports each row-model's mean CRPS as of the "hours" release relative to that of the column-model. Here, we find many more instances of statistically significant differences. Clearly, while the AR(2) has some value for the point forecasts, the CRPS values are typically much higher than that of the other models. In addition, it is often the case that the CRPS values from the MF-BVAR models are better than those from the DMS models used in CCM.

4.2 Four-quarter-ahead Forecasts

In Figures 4 and 5, we reproduce the same term structure of RMSE and CRPS values for the four-quarter-ahead forecasts. In contrast to the nowcasting results, we now forecast a wider range of series. Among the series in our dataset, the SPF forecasts GDP growth, industrial production, housing starts, payroll employment, and both the 3-month and 10-year Treasury yields. For the RMSE paths, we report results from both large and small MF-BVARs. The RMSE paths include forecasts from the SPF and the quarterly AR(2) forecasts of GDP; we omit the SPF when evaluating CRPS. Because CCM do not consider longer horizons forecasts, we do not include the DMS forecast in the comparison. In this section, we compare our model-based forecasts to the SPF forecasts for each of these series except for employment which has only been part of the survey since 2003:Q4.

The top panel of Figure 4 shows that, although the AR(2) outperforms it in month 1 and early in month 2, the SPF point forecasts of GDP growth are very difficult to beat at the four-quarter-ahead horizon. At no horizon do the models exhibit a lower RMSE than either the AR(2) or the SPF. In contrast, the lower four panels of Figure 4 show that the models outperform the SPF for variables other than GDP—in particular, industrial production and housing starts—at some intraquarter forecast origins. Four-quarter-ahead forecasts of industrial production and housing starts are as good as or better than the SPF at almost all horizons other than late in the second calendar month. Relative to the SPF, model-based forecasts of yields are a bit better for the 10-year than the 3-month but, in both cases, they track the RMSEs of the SPF pretty closely.

In Figure 5, we plot the mean CRPS values as in Figure 3, but for the four-quarter-ahead horizon. As was the case before, the values generally decline as we move across the intraquarter forecast origins. This is particularly true for industrial production, but also for housing starts. While not statistically significant, it is typically the case that the large MF-BVAR has a lower

CRPS value than that of the small model.

Table 2 reports the nominal measures of accuracy for four-quarter-ahead GDP forecasts and bolds those ratios for which a pairwise test of equal accuracy indicates statistical significance at the 5% level. In general, the results are statistically insignificant; however, we find that the large BVAR density forecasts dominate the small BVAR after every Employment Situation Report.

Table 3 reports the tests of equal predictive accuracy for industrial production, housing starts, 3-month and 10-year yields (when available). We do this because only the large MF-BVAR is used to form density forecasts of the 3-month and 10-year yields. We do not find much evidence that the models are significantly different from one another in terms of either point or density forecasts despite the fact that, in nominal terms, the large MF-BVAR tends to be more accurate for both.

4.3 Scenario Forecasting

The previous subsections indicate that point forecasts from the SPF are difficult to beat in terms of RMSEs. Even so, there is one thing our model can do that the SPF cannot – produce scenario forecasts designed to guide hypothetical policies. In addition, while perhaps not immediately obvious, the block structure of our VAR allows us to implement scenarios many other lower frequency models cannot. In this section, we provide an example of a high frequency policy-oriented scenario and compare it's implementation using our MF-BVAR to that of a standard BVAR in which the monthly frequency variables are aggregated to a quarterly frequency using monthly averages.

As an example, consider a central bank that uses a high frequency interest rate to conduct monetary policy. Suppose that the goal of the policy is to influence a low frequency series such as GDP. In a purely quarterly model, the high frequency policy rate would likely be averaged across all three months of the quarter, leaving the timing of the rate change obscured.[7] Instead, in our MF-BVAR, we can explicitly account for the timing of the policy rate change within the quarter.

We explore the effect of this timing in the following scenario adapted from the pattern of Federal Funds Rate changes made by the Federal Open Market Committee (FOMC) throughout 2017 but, to maintain consistency across sections of the paper, we use the 3-month Treasury yield as the policy rate. Suppose that on the last business day of January 2017, a scenario forecast is made that assumes the 3-month Treasury yield remains constant throughout February but rises 25 bps in March (e.g. at the March FOMC meeting). It then remains unchanged until June, at which time it rises another 25 bps. It is then assumed to remain constant throughout much of the year before rising another 25 bps in December and stays constant until the end of the year.

In our MF-BVAR, implementing this scenario is straightforward. We simply map the month-to-month changes in the policy rate directly into specific variables in our large MF-BVAR: the first, second, and third 3-month Treasury yields. For the quarterly BVAR, which we estimate using the

[7] See Knotek and Zaman (2017) for an exception.

same GLP code, we first form quarterly averages of the scenario and then form conditional forecasts using this low frequency approximation to the high frequency scenario.

We are interested in two issues related to forecasts based on this scenario. First, holding the forecast origin constant, are there substantive differences between the MF-BVAR and quarterly BVAR forecasts? Second, are there substantive differences among the MF-BVAR forecasts as we receive high frequency intraquarter information?

Based on the example above, Figure 6 plots multiple scenario forecasts of *total* annualized real GDP growth (cumulative sum of quarter on quarter growth rates) from 2016:Q4 for quarterly horizons one through four. Total, rather than quarter-to-quarter, growth is reported in order to align our forecasts with the fixed-event forecasts used by the FOMC. Each forecast is generated using the relevant real-time vintage of data exactly as we did for the one- and four-quarter-ahead forecasting exercises. Actuals are reported using data from the Advance release of 2017:Q4 GDP. It is useful to keep in mind that forecasts from the MF-BVAR will evolve as we obtain intraquarter information, while those from the quarterly BVAR will not.

In the first panel we report forecasts from both quarterly and MF-BVAR models made from the same, end-of-January forecast origin. At this origin, the quarterly BVAR forecasts are uniformly higher than those from the mixed frequency BVAR. These differences are as large as 75 bps at the one-quarter horizon but narrow substantially at the four-quarter horizon. We then update the MF-BVAR forecast twice as we move across February. The first update aligns with the Employment Report, while the second aligns with the claims report (and hence we observe all February releases). The first of these remains relatively close to the path predicted by the initial MF-BVAR forecast while the latter, in particular, reduces the near term forecast of GDP growth sharply.

The second panel reports comparably updated forecasts from the MF-BVAR but as we move across March. Not surprisingly, given the strong employment report in early March 2017, the first March revision shifts the forecasts upwards and is now much more in line with that of the quarterly BVAR – at least for horizons greater than one quarter. Even so, as further data arrives in March, this shift moderates.

The final panel again reports updated forecasts from the MF-BVAR but this time as we move across April. In contrast to March, reported employment growth fell sharply. As a consequence, the scenario paths both decline and are again closely in line with the realized values of total GDP growth – at least relative to those from the quarterly BVAR, which does not get updated as we move across the quarter.

While this is only one example of a scenario forecast, it is a realistic, policy-oriented one that suggests some advantages to our approach to modeling mixed frequencies. First, the mixed-frequency allows us to implement detailed high frequency scenarios that low frequency models cannot. In addition, since our model is readily revised as high frequency data is released, we are able to track how the scenario forecasts evolve within a quarter and by implication, evolve as new data is observed between policy meetings.

5 Conclusion

In this paper, we investigate the usefulness of a particular type of mixed frequency VAR in the context of real-time forecasting. In this model, multiple high frequency intraquarter observations are treated as distinct quarterly observations and a standard VAR is formed based on these series. Because this leads to a high dimensional VAR, we estimate the parameters using standard shrinkage-based Bayesian methods. In addition, since the model is just a Bayesian VAR, existing methods developed by Waggoner and Zha (1999) can be used to produce end of quarter forecasts as well as intraquarter forecasts that account for high frequency data releases.

In terms of both point and density forecasts, we find that our iterated multi-step approach to mixed frequency nowcasting of GDP growth performs as well as direct multi-step variants developed in Carriero, Clark and Marcellino (2015) and, depending on the specific version of our model, can be as accurate as the SPF at certain very short horizons. One advantage of our modeling approach is that the same model can be used for both near and longer horizon forecasting. As such, we also compare the forecasting performance of our model to the SPF at the four quarter horizon. While the SPF largely dominates when forecasting GDP growth, our model is as good or sometimes better when forecasting monthly variables such as industrial production, housing starts and, to a lesser extent 3-month and 10-year Treasury yields. Finally, we discuss the usefulness of our model for central bank-type scenario forecasting when the scenario is delineated using high frequency observables, but the object of interest is only observed at the lower frequency.

References

[1] Bacchiocchi, E., Bastianin, A., Missale, A., and E. Rossi (2016). "Structural Analysis with Mixed Frequencies: Monetary Policy, Uncertainty and Gross Capital Flows," *manuscript.*

[2] Bańbura, M., Giannone, D., and R. Lucrezia (2010). "Large Bayesian Vector Auto Regressions," *Journal of Applied Econometrics* 25(1), 71-92.

[3] Bhansali, R.J. (1997). "Direct Autoregressive Predictors for Multistep Prediction: Order Selection and Performance Relative to Plug-In Predictors," *Statistica Sinica* 7(2), 425-449.

[4] Bittanti, S., Colaneri, P., and G. De Nicolao (1988). "The Difference Periodic Riccati Equation for the Periodic Prediction Problem," *IEEE transactions on Automatic Control* 33(8): 706-712.

[5] Blasques, F., Koopman, S.J., Mallee, M., and Z. Zhang (2016). "Weighted Maximum Likelihood for Dynamic Factor Analysis and Forecasting with Mixed Frequency Data," *Journal of Econometrics* 193(2), 405-417.

[6] Brave, S., Butters, R.A., and A. Justiniano (2016). "Forecasting Economic Activity with Mixed Frequency Bayesian VARs," *Chicago Fed Working Paper* 2016-05.

[7] Carriero, A., Clark, T.E., and M. Marcellino (2015). "Realtime Nowcasting with a Bayesian Mixed Frequency Model with Stochastic Volatility," *Journal of Royal Statistical Society A* 178(4), 837-862.

[8] Chen, W., Anderson, B.D.O., Deistler, M., and A. Filler (2012). "Properties of Blocked Linear Systems," *Automatica* 48(10), 2520–2525.

[9] Corrado, C. and M. Green (1988). "Reducing Uncertainty in Short-term Projections: Linkage of Monthly and Quarterly Models," *Journal of Forecasting* 7(2), 77-102.

[10] Diebold, F.X. and R.S. Mariano (1995). "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13(3), 253-263.

[11] Eraker, B., Chiu, C.W., Foerster, A.T., Kim, T.B., and H.D. Seoane (2015). "Bayesian Mixed Frequency VARs," *Journal of Financial Econometrics* 13(3), 698-721.

[12] Foroni, C. and M. Marcellino (2013). "A Survey of Econometric Methods for Mixed-Frequency Data," *Norges Bank Working Papers* 2013/06.

[13] Foroni, C., Marcellino, M., and C. Schumacher (2015). "U-MIDAS: MIDAS Regressions with Unrestricted Lag Polynomials," *Journal of the Royal Statistical Society A* 178(1), 57-82.

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

[15] Giannone, D., Reichlin, L. and D. Small (2008). "Nowcasting: The Real-time Informational Content of Macroeconomic Data," *Journal of Monetary Economics* 55(4), 665-676.

[16] Ghysels, E. (2016). "Macroeconomics and the Reality of Mixed Frequency Data," *Journal of Econometrics* 193(2), 294-314.

[17] Gneiting, T. (2011). "Making and Evaluating Point Forecasts," *Journal of the American Statistical Association* 106(494), 746-762.

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

[19] Gneiting, T. and R. Ranjan (2011). "Comparing Density Forecasts Using Threshold and Quantile Weighted Proper Scoring Rules," *Journal of the Business and Economic Statistics* 29(3), 411-422.

[20] Jarociński, M. (2010). "Conditional Forecasts and Uncertainty about Forecast Revisions in Vector Autoregressions," *Economics Letters* 108(3), 257-259.

[21] Knotek, E. and S. Zaman (2017). "Financial Nowcasts and Their Usefulness in Macroeconomic Forecasting," *Federal Reserve Bank of Cleveland working paper 17–02*.

[22] Krueger, F., Lerch, S., Thorarinsdottir, T.L., and T. Gneiting (2017). "Probabilistic Forecasting and Comparative Model Assessment Based on Markov Chain Monte Carlo Output," *manuscript*.

[23] Marcellino, M., Stock, J.H., and M.W. Watson (2006). "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series," *Journal of Econometrics* 135(1-2), 499-526.

[24] Mittnik, S. and P.A. Zadrozny (2005). "Forecasting Quarterly German GDP at Monthly Intervals using Monthly Ifo Business Conditions Data," in *Ifo Survey Data in Business Cycle and Monetary Policy Analysis*, ed. by J.-E. Sturm, and T. Wollmershauser, pp. 19-48. Heidelberg: Physica-Verlag.

[25] Newey, W.K. and K.D. West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55(3), 703-708.

[26] Parigi, G. and G. Schlitzer (1995). "Quarterly Forecasts of the Italian Business Cycle by Means of Monthly Economic Indicators," *Journal of Forecasting* 14(2), 117-141.

[27] Schorfheide, F. (2005). "VAR Forecasting under Misspecification," *Journal of Econometrics* 128, 99-136.

[28] Schorfheide, F. and D. Song (2015). "Real-Time Forecasting with a Mixed-Frequency VAR," *Journal of Business and Economic Statistics* 33(3), 366-380.

[29] Sims, C.A. and T. Zha (1998). "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review* 39(4), 949-968.

[30] Waggoner, D.F. and T. Zha (1999). "Conditional Forecasts in Dynamic Multivariate Models," *The Review of Economics and Statistics* 81(4), 639-651.

[31] Zadrozny, P.A. (1990). "Forecasting U.S. GNP at Monthly Intervals with an Estimated Bivariate Time Series Model," *Economic Review*, Federal Reserve Bank of Atlanta, Nov., 2-15.

Table 1: RMSE's and Mean CRPS's of Nowcasts

		1985:Q1 to 2011:Q3					
		SPF	AR(2)	CCM (small)	CCM (large)	MF-BVAR (small)	MF-BVAR (large)
M1	SPF	1.75\NA					
	AR(2)	1.15	2.02\3.42	**1.14**	**1.21**	**1.18**	**1.27**
	CCM (small)	1.14	0.99	2.00\2.98	**1.05**	**1.03**	**1.11**
	CCM (large)	**1.21**	1.05	1.06	2.13\2.83	0.97	**1.05**
	MF-BVAR (small)	1.14	0.99	1.00	0.94	2.00\2.91	**1.08**
	MF-BVAR (large)	**1.25**	1.09	1.10	1.03	1.10	2.20\2.69
M2	SPF	1.41\NA					
	AR(2)	1.32	1.86\3.23	**1.22**	**1.21**	**1.31**	**1.44**
	CCM (small)	**1.42**	1.08	2.01\2.64	0.99	**1.07**	**1.18**
	CCM (large)	**1.41**	1.07	0.99	1.99\2.67	**1.09**	**1.19**
	MF-BVAR (small)	**1.38**	1.05	0.97	0.98	1.95\2.46	**1.10**
	MF-BVAR (large)	**1.33**	1.01	0.94	0.95	0.97	1.88\2.23
M3	SPF	1.41\NA					
	AR(2)	1.30	1.84\3.23	**1.43**	**1.31**	**1.63**	**1.65**
	CCM (small)	1.17	0.90	1.65\2.25	**0.91**	**1.14**	**1.15**
	CCM (large)	**1.24**	0.96	1.06	1.75\2.47	**1.25**	**1.26**
	MF-BVAR (small)	1.08	0.83	0.93	0.87	1.53\1.98	1.01
	MF-BVAR (large)	**1.18**	0.91	1.01	0.95	1.09	1.67\1.96
M4	SPF	1.41\NA					
	AR(2)	1.30	1.83\3.21	**1.50**	**1.28**	**1.71**	**1.63**
	CCM (small)	1.05	**0.81**	1.48\2.15	**0.85**	**1.14**	**1.09**
	CCM (large)	**1.25**	0.96	**1.19**	1.76\2.51	**1.34**	**1.28**
	MF-BVAR (small)	1.03	0.79	0.98	0.82	1.45\1.88	0.95
	MF-BVAR (large)	**1.27**	0.98	1.22	1.02	**1.24**	1.80\1.97

		1992:Q2 to 2017:Q1					
		SPF	AR(2)	CCM (small)	CCM (large)	MF-BVAR (small)	MF-BVAR (large)
M1	SPF	1.77\NA					
	AR(2)	1.13	2.00\3.21	**1.15**	**1.20**	**1.19**	**1.28**
	CCM (small)	1.08	0.96	1.92\2.80	**1.05**	**1.04**	**1.12**
	CCM (large)	1.11	0.98	1.02	1.96\2.67	0.99	**1.07**
	MF-BVAR (small)	1.06	0.94	0.98	0.96	1.88\2.70	**1.08**
	MF-BVAR (large)	1.12	1.00	1.04	1.01	1.06	1.99\2.51
M2	SPF	1.44\NA					
	AR(2)	1.28	1.84\3.03	**1.21**	**1.19**	**1.34**	**1.44**
	CCM (small)	**1.33**	1.04	1.91\2.50	0.98	**1.10**	**1.18**
	CCM (large)	**1.38**	1.08	1.04	1.99\2.55	**1.12**	**1.21**
	MF-BVAR (small)	**1.22**	0.96	0.92	0.88	1.76\2.27	**1.08**
	MF-BVAR (large)	**1.21**	0.95	0.91	**0.88**	0.99	1.75\2.11
M3	SPF	1.44\NA					
	AR(2)	1.26	1.81\3.03	**1.39**	**1.30**	**1.61**	**1.66**
	CCM (small)	1.17	0.93	1.68\2.18	0.93	**1.16**	**1.20**
	CCM (large)	**1.18**	0.94	1.01	1.70\2.34	**1.24**	**1.28**
	MF-BVAR (small)	1.01	0.80	**0.86**	0.85	1.44\1.88	1.03
	MF-BVAR (large)	1.05	0.83	0.90	0.88	1.04	1.51\1.82
M4	SPF	1.44\NA					
	AR(2)	1.27	1.82\3.03	**1.48**	**1.27**	**1.73**	**1.71**
	CCM (small)	1.01	**0.80**	1.45\2.04	**0.85**	**1.17**	**1.15**
	CCM (large)	**1.20**	0.95	1.19	1.73\2.39	**1.36**	**1.35**
	MF-BVAR (small)	0.91	**0.72**	0.90	**0.75**	1.30\1.75	0.99
	MF-BVAR (large)	1.05	0.83	1.03	0.87	**1.15**	1.50\1.78

Table 2: RMSE's and Mean CRPS's of 4-Quarter-Ahead Forecasts of GDP for the 1992:Q2 to 2017:Q1 Sample

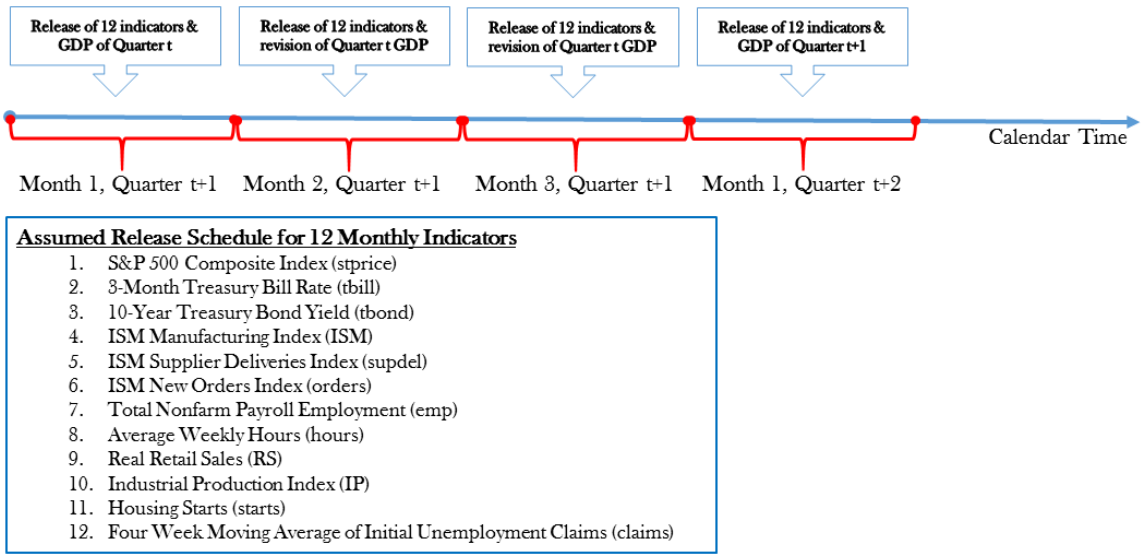
		SPF	AR(2)	MF-BVAR (small)	MF-BVAR (large)
M1	SPF	2.10\NA			
	AR(2)	0.98	2.06\3.30	0.98	1.03
	MF-BVAR (small)	1.07	1.09	2.24\3.35	**1.04**
	MF-BVAR (large)	1.06	1.08	0.99	2.22\3.21
M2	SPF	2.04\NA			
	AR(2)	1.00	2.04\3.27	0.98	1.04
	MF-BVAR (small)	1.09	1.09	2.23\3.33	**1.06**
	MF-BVAR (large)	1.09	1.09	1.00	2.23\3.14
M3	SPF	2.04\NA			
	AR(2)	1.00	2.05\3.28	0.98	1.05
	MF-BVAR (small)	1.12	1.12	2.29\3.34	**1.06**
	MF-BVAR (large)	1.12	1.12	1.00	2.28\3.14
M4	SPF	2.04\NA			
	AR(2)	1.00	2.05\3.28	0.99	1.05
	MF-BVAR (small)	1.11	1.11	2.28\3.33	**1.06**
	MF-BVAR (large)	1.12	1.12	1.00	2.29\3.14

Table 3: RMSE's and Mean CRPS's of 4-Quarter-Ahead Forecasts of Select Monthly Variables for the 1992:Q2 to 2017:Q1 Sample

		Industrial Production			Housing Starts		
		SPF	MF-BVAR (small)	MF-BVAR (large)	SPF	MF-BVAR (small)	MF-BVAR (large)
M1	SPF	5.63\NA			245.39\NA		
	MF-BVAR (small)	0.89	5.00\5.39	**1.09**	**0.89**	218.94\262.34	1.01
	MF-BVAR (large)	0.90	1.01	5.06\4.95	0.93	1.04	228.70\260.09
M2	SPF	4.58\NA			202.24\NA		
	MF-BVAR (small)	0.99	4.53\4.93	**1.08**	1.03	208.00\253.52	1.03
	MF-BVAR (large)	1.02	1.03	4.68\4.57	1.07	1.04	215.94\246.58
M3	SPF	4.58\NA			202.24\NA		
	MF-BVAR (small)	0.96	4.39\4.71	**1.07**	0.95	192.81\238.73	1.06
	MF-BVAR (large)	0.98	1.03	4.50\4.38	0.96	1.01	194.55\225.82
M4	SPF	4.58\NA			202.24\NA		
	MF-BVAR (small)	0.86	3.95\4.22	**1.10**	**0.92**	185.35\228.32	1.04
	MF-BVAR (large)	0.85	0.98	3.89\3.83	0.92	1.00	185.40\218.56
		3-Month Yield			10-Year Yield		
		SPF		MF-BVAR (large)	SPF		MF-BVAR (large)
M1	SPF	1.31\NA			1.04\NA		
	MF-BVAR (large)	0.99		1.30\1.80	0.98		1.02\1.30
M2	SPF	1.02\NA			0.90\NA		
	MF-BVAR (large)	1.09		1.11\1.61	1.03		0.93\1.18
M3	SPF	1.02\NA			0.90\NA		
	MF-BVAR (large)	1.07		1.08\1.56	0.97		0.87\1.12
M4	SPF	1.02\NA			0.90\NA		
	MF-BVAR (large)	1.00		1.02\1.52	0.96		0.87\1.11

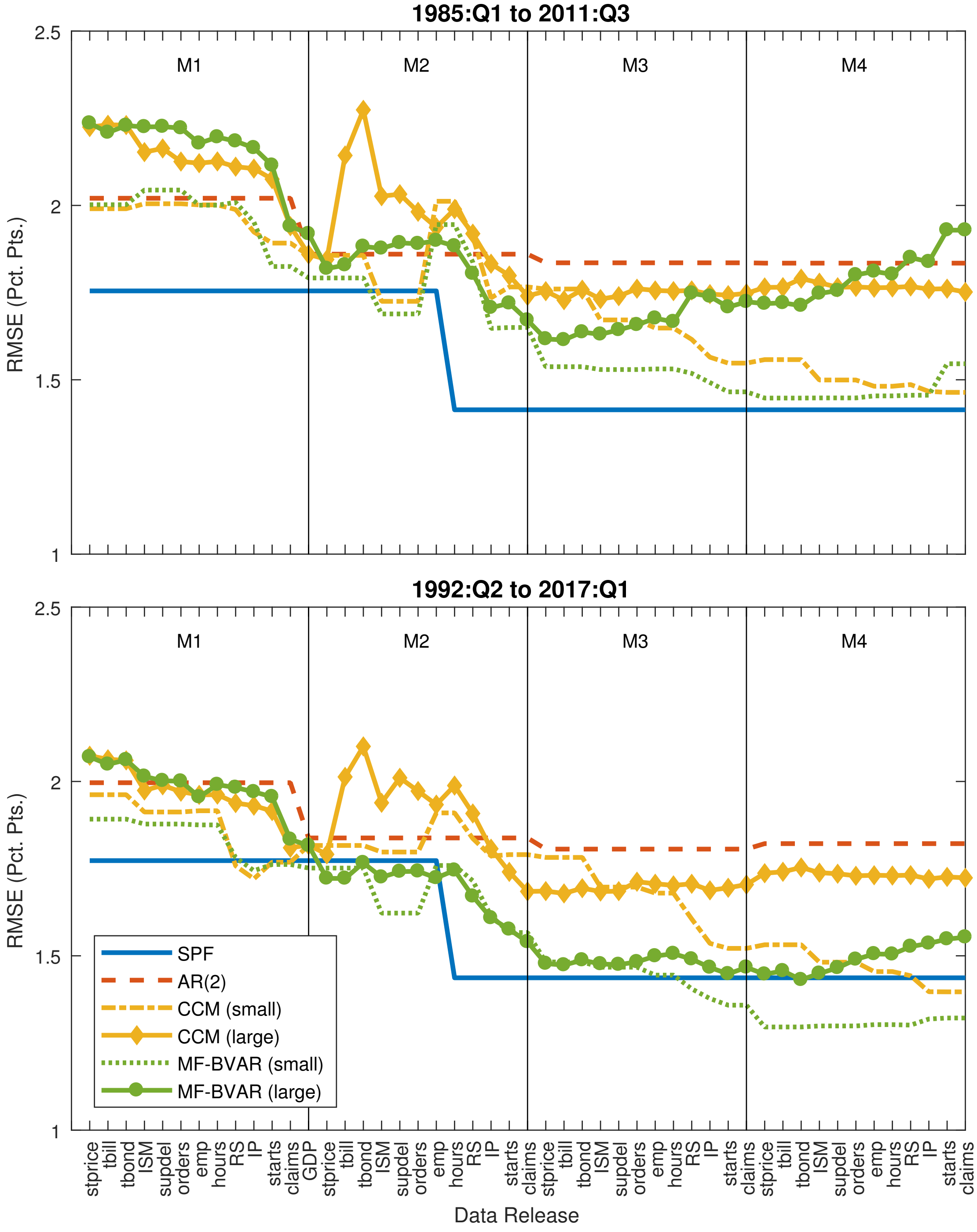
Notes to Tables 1-3: The tables show the RMSEs (below the slash) and mean CRPSs (above the slash) of each model on the diagonal. The lower off-diagonal portion of each panel in the table reports the ratio of RMSEs in the row-model to the column-model. The upper off-diagonal portion of each panel reports the ratio of average CRPS in the row-model to the column-model. The results are as of the "hours" release in each month. Off-diagonal numbers greater than one indicate that the column-model is nominally more accurate. Ratios in bold are statistically different from one at the 5% significance level.

Figure 1: Data Releases and Forecast Timing



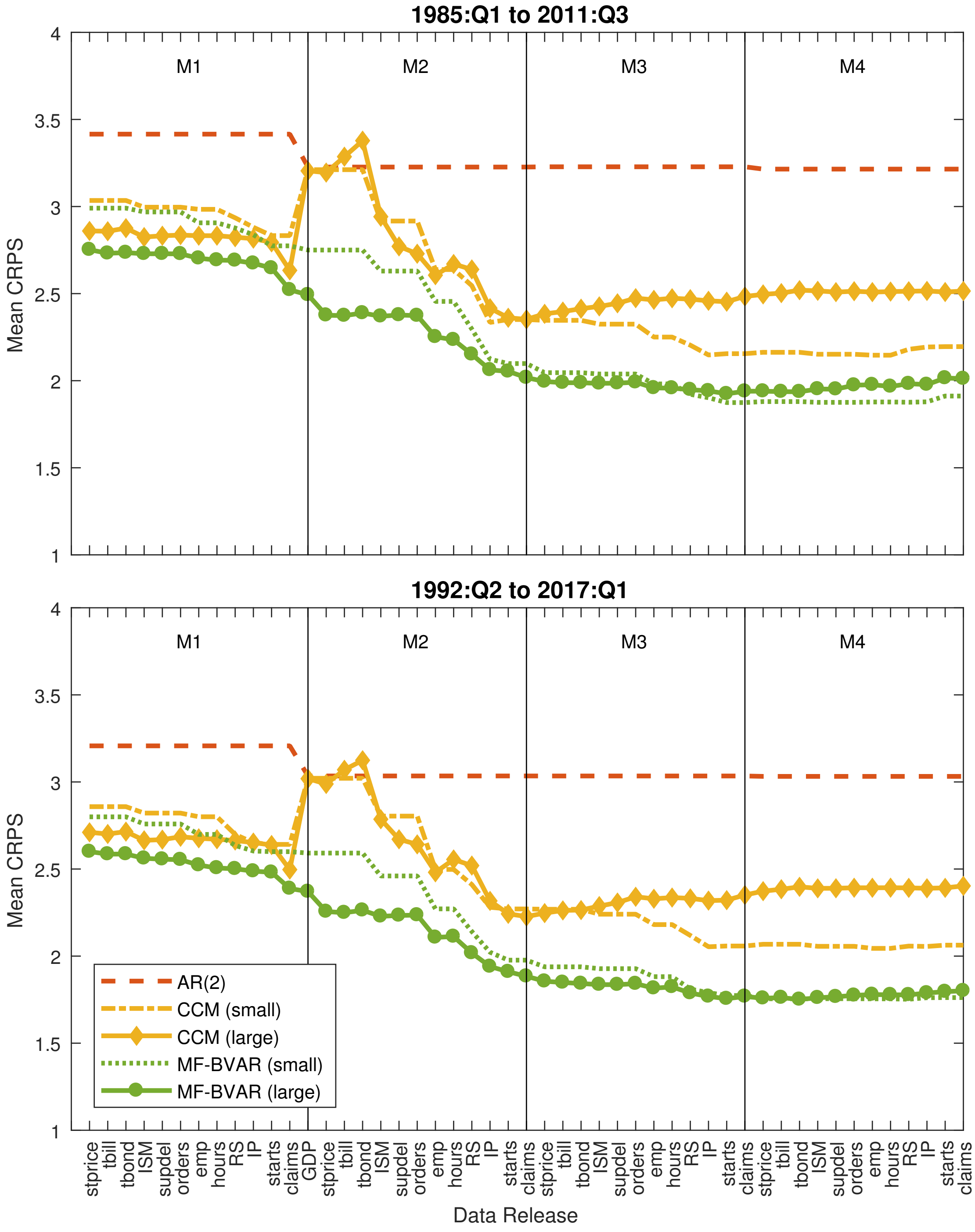
Note: The figure outlines the timing of the data releases and forecasts over the course of the quarter.

Figure 2: RMSE Paths for GDP Growth Nowcasts



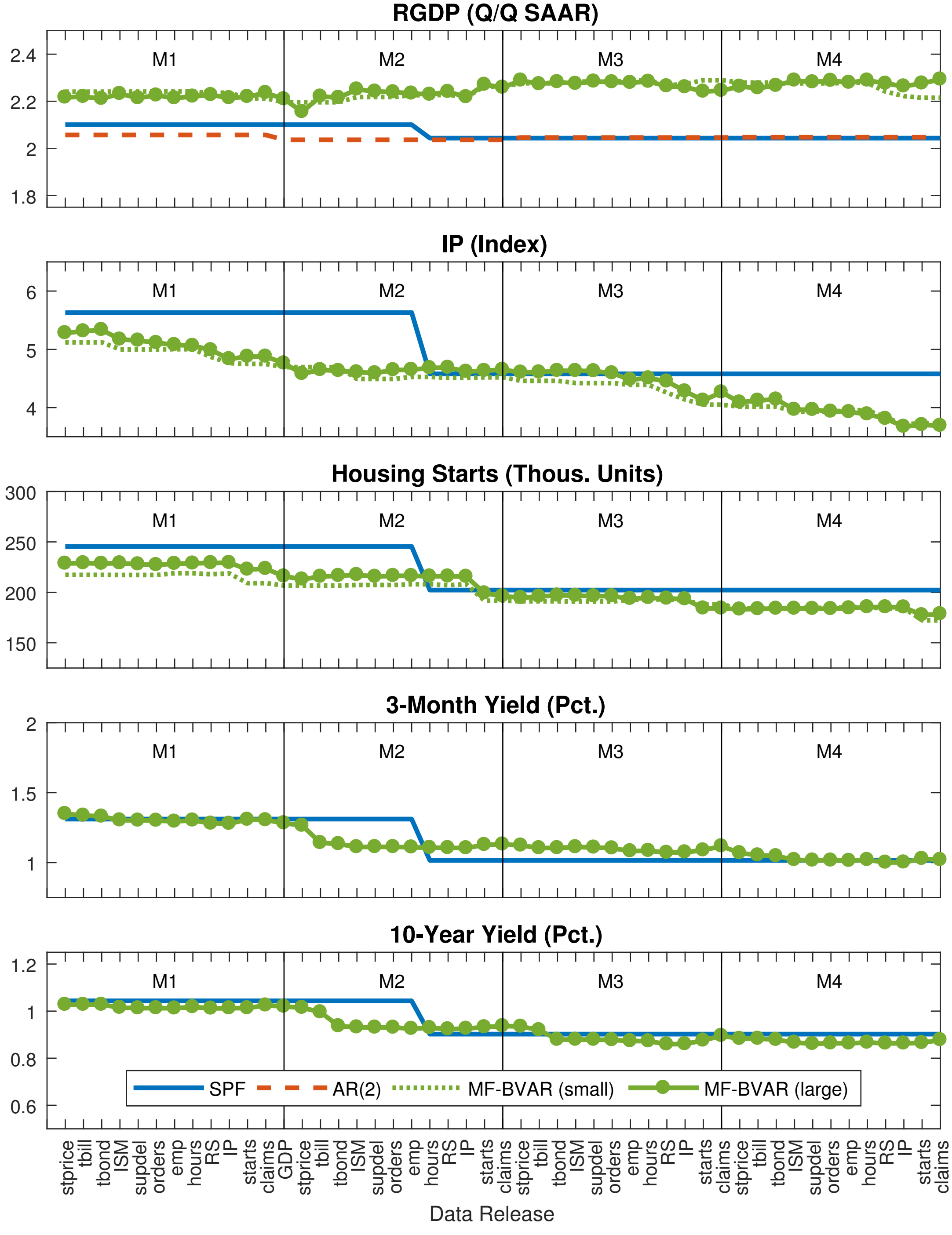
Note: The figure depicts the RMSE paths for nowcasts of the Advance release of GDP. The upper panel considers the evaluation period in CCM, while the lower panel shows the results associated with GDP (as opposed to GNP) forecasting. Each tick represents a data release in the respective month.

Figure 3: CRPS Paths for GDP Growth Nowcasts



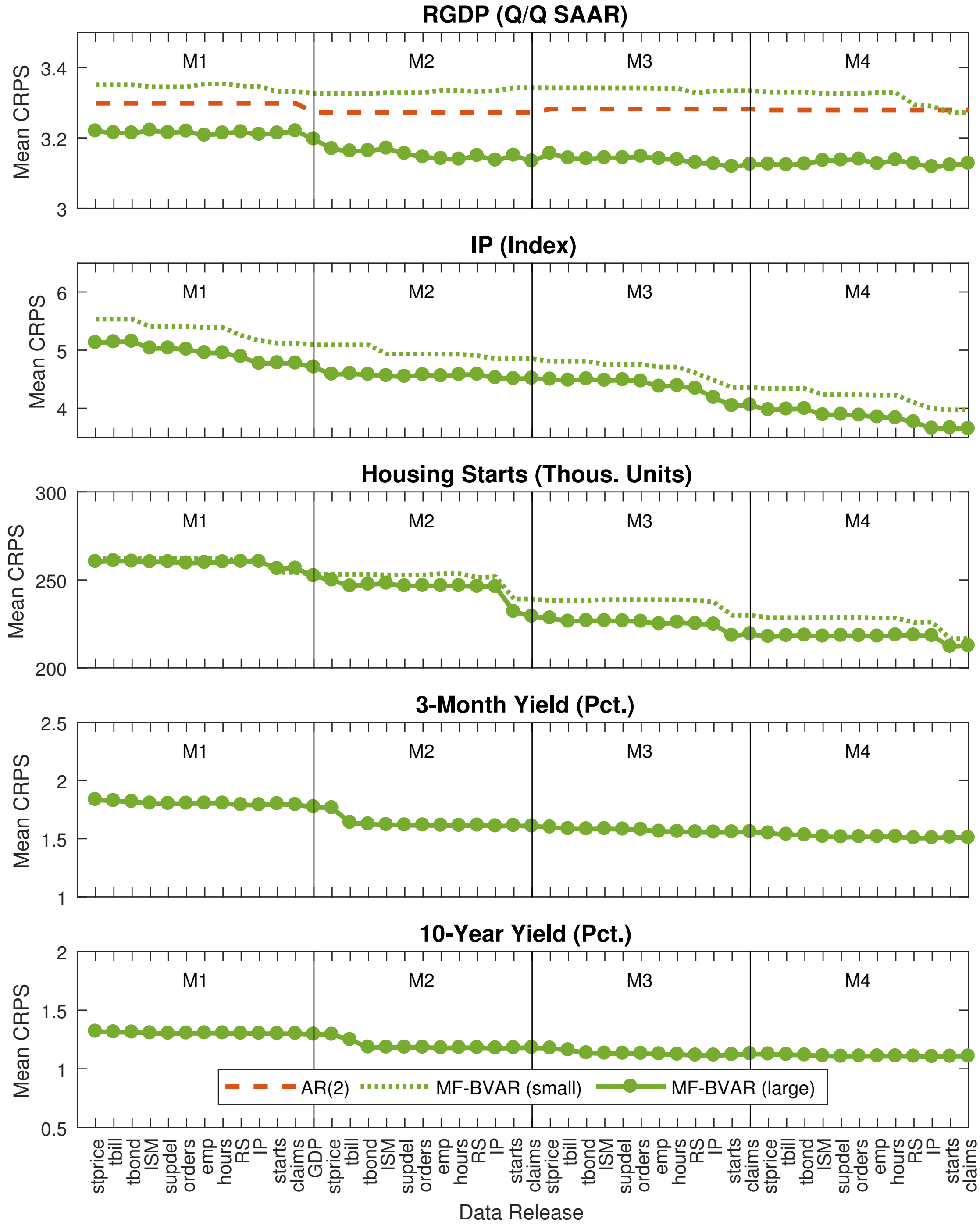
Note: The figure depicts the average CRPS paths for nowcasts of the Advance release of GDP. The upper panel considers the evaluation period in CCM, while the lower panel shows the results associated with GDP (as opposed to GNP) forecasting. Each tick represents a data release in the respective month.

Figure 4: RMSE Paths for Four-quarter-ahead Forecasts



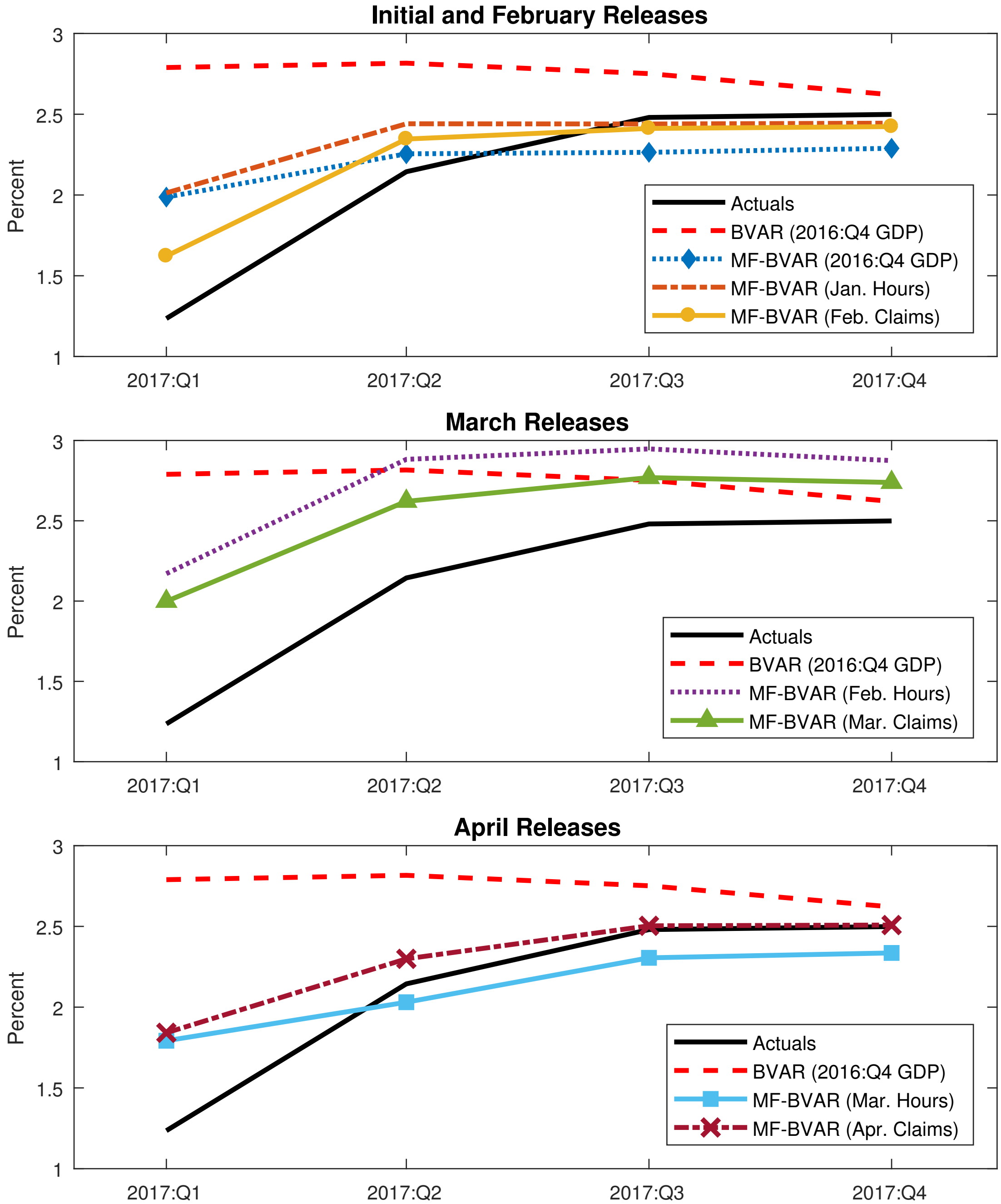
Note: The figure depicts the RMSE paths for four-quarter-ahead forecasts of the Advance release of GDP for the evaluation period of 1992:Q2-2017:Q1. The quarterly AR(2) model is included for GDP only. Each tick represents a data release in the respective month.

Figure 5: CRPS Paths for Four-quarter-ahead Forecasts



Note: The figure depicts the average CRPS paths for four-quarter-ahead forecasts of the Advance release of GDP for the evaluation period of 1992:Q2-2017:Q1. The quarterly AR(2) model is included for GDP only. Each tick represents a data release in the respective month.

Figure 6: Policy-rate-based Scenario Forecasts



Note: The figure depicts total GDP growth forecasts conditional on an assumed path of an interest rate. Each figure shows the forecasts from a quarterly BVAR as well as the actual realized value. Each panel further shows forecasts from the large MF-BVAR. (.) indicates the release of a variable the forecast is conditioned on.