CESifo WORKING PAPERS

8054
2020

January 2020

Real-Time Forecasting Using Mixed-Frequency VARs with Time-Varying Parameters

Markus Heinrich, Magnus Reif

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 8054*

Real-Time Forecasting Using Mixed-Frequency VARs with Time-Varying Parameters

Abstract

This paper provides a detailed assessment of the real-time forecast accuracy of a wide range of vector autoregressive models (VAR) that allow for both structural change and indicators sampled at different frequencies. We extend the literature by evaluating a mixed-frequency time-varying parameter VAR with stochastic volatility (MF-TVP-SV-VAR). Overall, the MF-TVP-SV-VAR delivers accurate now- and forecasts and, on average, outperforms its competitors. We assess the models' accuracy relative to expert forecasts and show that the MF-TVP-SV-VAR delivers better inflation nowcasts in this regard. Using an optimal prediction pool, we moreover demonstrate that the MF-TVP-SV-VAR has gained importance since the Great Recession.

JEL-Codes: C110, C530, C550, E320.

Keywords: time-varying parameters, forecasting, nowcasting, mixed-frequency models, Bayesian methods.

Markus Heinrich
University of Kiel
Institute for Statistics and Econometrics
Olshausenstr. 40 – 60
Germany - 24118 Kiel
heinrich@stat-econ.uni-kiel.de

Magnus Reif
ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstrasse 5
Germany – 81679 Munich
reif@ifo.de

2nd September 2018
We are thankful to Kai Carstensen, Todd Clark, Maik Wolters, and Jonathan Wright as well as the participants of the 6th Annual Conference of the International Association for Applied Econometrics and 10th Nordic Econometric Meeting for helpful comments and suggestions.

1 Introduction

Macroeconomists and, in particular, macroeconomic forecasters face two major challenges. First, there are structural changes within an economy. Second, in real time forecasters need to process unbalanced datasets due to indicators sampled at different frequencies and idiosyncratic publication lags. Concerning structural change, it is commonly found that particularly modeling time-varying volatilities enhances VAR-based inference and estimation, while fluctuations in the VAR coefficients are frequently considered to be less vital (for example, Sims and Zha, 2006; Chan and Eisenstat, 2017). This finding is confirmed with respect to forecasting by, for example, D'Agostino, Gambetti, and Giannone (2013). Since the onset of the Great Recession, which probably caused important structural shifts, modeling time-varying links between variables has attracted anew interest.[1] Concerning unbalanced datasets, a growing literature stresses the merits of mixed-frequency approaches in computing precise now- and forecasts, and tracking the current state of the economy in real time (for instance, Kuzin, Marcellino, and Schumacher, 2011; Foroni and Marcellino, 2014; Schorfheide and Song, 2015).

However, evidence regarding the forecast performance of models allowing for structural shifts in a mixed-frequency setting is rather sparse. This study aims at filling this gap by providing a detailed assessment of the real-time forecast accuracy of a bundle mixed-frequency models that allow for structural change. To this end, we estimate nonlinear and linear VARs with and without mixed-frequencies, including a fully-fledged model incorporating time-varying parameters, stochastic volatilities, and mixed-frequencies—a MF-TVP-SV-VAR. This analysis enables us to trace out the relative impact of the models' mixed-frequency part and the time-variation in the models' coefficients on the forecast accuracy. Our comparison relies on real-time out-of-sample now- and forecast accuracy of both point and density forecasts for three key US macroeconomic variables: GDP growth, CPI inflation, and the unemployment rate.

Overall, our forecast comparison provides two major findings. First, modelling structural change *and* intra-quarterly dynamics is beneficial for point and density forecasts, notably for now and short-term forecasts. In particular, the accuracy of inflation and unemployment rate forecasts can be substantially increased. Second, the MF-TVP-SV-VAR delivers very competitive point and density forecasts—on average over all variables it outperforms each competitor. Our results moreover suggest that the combination of mixed-frequencies, stochastic volatility, and time-varying parameters is particularly beneficial for inflation nowcasts computed with only little information about the respective quarters. In those cases, the MF-TVP-SV provides the largest gains in forecast accuracy. Taking a closer look at the mixed-frequency models' forecasts during the Great Recession reveals that allowing for time-variation in the VAR coefficients and stochastic volatility is superior relative to only one of these specifications for inflation and the unemployment rate.

We put our results to the test along two dimensions. First, since it is commonly found that a combination of forecasts from several models outperforms individual models, we augment our analysis with a forecast combination exercise, including equal weighting and an optimal

[1] See Ng and Wright (2013) for a survey of business cycle facts of the U.S. economy with a focus on the Great Recession.

prediction pool á la Geweke and Amisano (2011). We find that when estimating the models' weights, the mixed-frequency models consistently receive a large share of the entire probability mass. In particular, the MF-TVP-SV-VAR receives a high weight since the onset of the Great Recession, indicating the benefits of modelling structural change combined with modelling within-quarter dynamics in that period. Moreover, this optimal prediction pool provides strong gains in point and density forecast accuracy across all variables and horizons. It even outperforms the equal weighting combination scheme in almost each case. Second, since precise forecasts may stem from an accurate assessment of the current state of the economy (Sims, 2002), we assess whether the mixed-frequency models' satisfying forecast performance flows from more precise nowcasts. We assess this channel, by—in a first step—comparing the MF-VARs' predictions with those of the Survey of Professional Forecasters (SPF). SPF predictions are superior for GDP growth and the unemployment rate, while for inflation, the nonlinear MF-VARs provide slightly more accurate nowcasts. In a second step, we follow Schorfheide and Song (2015) and Wolters (2015) by augmenting the quarterly datasets with SPF nowcasts. We find that the gains in accuracy of the quarterly models are rather small with respect to GDP growth and inflation, while unemployment rate forecasts can be substantially improved.

Estimation of the models' TVP-SV part mainly follows Primiceri (2005). However, we treat those hyperparameters that relate to the amount of time-variation in the parameters as an additional layer and estimate them using Bayesian methods (Amir-Ahmadi, Matthes, and Wang, 2020).[2] Estimation of the models' MF part is based on the idea that lower-frequency variables can be expressed as higher-frequency variables with latent observations (Zadrozny, 1988).[3] Adopting this notion, Mariano and Murasawa (2010) derive a state-space representation for VARs with missing observations, called mixed-frequency VAR (MF-VAR). We follow Schorfheide and Song (2015) and apply the MF-VAR approach in a Bayesian framework.

On the one hand, this paper contributes to the ongoing discussion on how structural change affects VAR-based forecast performance. D'Agostino et al. (2013) forecast US inflation, unemployment, and short-term interest rates with TVP-SV-VARs and find that allowing for parameter instability significantly improves forecast accuracy. Barnett, Mumtaz, and Theodoridis (2014) and Clark and Ravazzolo (2015) underpin these findings and show that models with time-varying parameters improve forecast performance, especially regarding inflation forecasts. Focusing on the period since the Great Recession, Aastveit et al. (2017) provide strong evidence against constant parameter VARs and document that TVP-SV-VAR tend to perform best with small models. Banbura and van Vlodrop (2018) illustrate that accounting for time-varying means in a Bayesian VAR substantially increases long-term forecast accuracy.

On the other hand, this article extends the literature on forecasting with non-linear mixed-frequency VARs. Foroni, Guérin, and Marcellino (2015) introduce mixed-frequency Markov-switching VARs and provide evidence that modelling discrete regime shifts in a mixed-frequency setting is particularly beneficial with regard to nowcasting and short-term forecasting. Closely

[2] Amir-Ahmadi et al. (2020) show that the magnitude of the hyperparameters changes significantly when estimated on monthly data compared to quarterly data, which affects the time-variation in the model's coefficients.

[3] Alternative approaches are mixed data sampling (MIDAS) provided by Ghysels, Santa-Clara, and Valkanov (2004) and the mixed frequency VAR in a stacked system introduced by Ghysels (2016). For an assessment of the stacked approach with regard to forecasting, see McCracken, Owyang, and Sekhposyan (2015).

related to this analysis is the study by Götz and Hauzenberger (2018) that also uses a mixed-frequency VAR that allows for continuous parameter change. The latter, however, analyzes the forecast ability in a pseudo real-time setting of a more parsimonious model, restricting the parameter change to the intercept terms and employ common stochastic volatility, while we abstract form these restrictions.

The remainder of the paper is as follows. Section 2 provides a description of the dataset and outlines the forecast setup. Section 3 depicts the competing models and explains the estimation methodology. Section 4 describes the measures used for the forecast comparison. Section 5 presents the results. Section 6 concludes.

2 Data and forecast setup

2.1 Dataset

We use an updated version of the dataset used by Clark and Ravazzolo (2015) consisting of four macroeconomic time series, three of which are sampled at monthly frequency and one is observed quarterly. The quarterly series is US real GDP; the monthly series are CPI, the unemployment rate, and the 3-month Treasury bill rate. GDP and CPI enter the models in log first differences times 100 to obtain real GDP growth and CPI inflation in percentage point changes, respectively. The unemployment and interest rate remain untransformed. For the VARs estimated on quarterly frequency, the monthly indicators enter the models as quarterly averages. We obtain real-time data on inflation, unemployment, and GDP from the Archival FRED (ALFRED) database of the St. Louis Fed. Since the Treasury bill rate is not revised, we resort to the last available publication from the FRED database. The sample runs from January 1960 until September 2017. The first 8 years are used as a training sample to specify priors such that the estimation starts in January 1968.

Generally, macroeconomic variables are released with a publication lag, which implies that a certain vintage does not include the figures referring to the date of the vintage. The first release of quarterly GDP has a publication lag of roughly one month, thus, for example, the first figure for 2011Q4 is released at the end of 2012M1 and is then consecutively revised in the subsequent months. The value for the unemployment rate (CPI) is published in the first (second) week of the following month. Hence, following our previous example, at the end of 2012M1 the unemployment rate and CPI are available until 2011M12. Finally, the 3-month Treasury bill rate is available without any delay. Thus, we have so-called "ragged-edges" in our real-time dataset.

2.2 Forecast setup

To assess the predictions with regard to the intra-quarterly inflow of information, we follow Schorfheide and Song (2015) and establish three different information sets. We assume that the forecasts are generated around the middle of each month, when the current releases for GDP, CPI, and the unemployment rate are available.[4] The first information set, called I1,

[4]We follow Schorfheide and Song (2015) and replace the missing observations for the T-Bill rate in the last month of each recursion by the expected monthly average.

relates to the first month of each quarter such that the forecaster has information up to the middle of January, April, July, or October. In these months, the researcher has observations on inflation and unemployment until the end of the respective previous quarter and a first and preliminary estimate of GDP referring to the previous quarter. The second information set, called I2 (February, May, August, November), has one additional observation on inflation and unemployment referring to the current quarter and the first revision of GDP. The last set, I3 (March, June, September, December), includes one more observation on inflation and unemployment and the second GDP revision. Each information set is augmented with the observations of the T-Bill rate. Since the quarterly VARs cannot cope with "ragged-edges" in the data, we estimate them in each recursion based on the balanced information set I1, which accounts for new information only in terms of data revisions.

We use an expanding window to evaluate our forecasts for data vintages from January 1990 until September 2017. The predictions are evaluated based on quarterly averages, implying that for the mixed-frequency approaches we aggregate the predicted monthly time paths to quarterly frequency. To abstract from benchmark revisions, we evaluate GDP growth forecasts based on the second available estimate, that is the forecast for period $t+h$ is evaluated with the realization taken from the vintage published in $t+h+2$ (see, for example, Faust and Wright, 2009). Since the remaining variables are revised only rarely and slightly, we evaluate the forecast based on the latest vintage. The maximum forecast horizon h_{max} is set to 4 quarters. Thus, the mixed-frequency models generate forecasts for $h_m = 1, \ldots, 12$ months. Forecasts for horizons larger than one are obtained iteratively. We report results for 1, 2, 3, and 4 quarters ahead forecasts.

3 Models

Our baseline model is a standard VAR with all variables sampled at quarterly frequency. Based on this model, we evaluate the forecast performance of three extensions, namely, mixed-frequencies, stochastic volatilities, and time-varying parameters, as well as the forecast performance of combinations of these features. For the stochastic volatility models, we use random walk stochastic volatility, which is a parsimonious and competitive specification (Clark and Ravazzolo, 2015). Throughout the paper, $n = n_q + n_m$, where n, n_q, and n_m denote the number of total, quarterly, and monthly variables, respectively. Finally, p denotes the lag order.

3.1 Quarterly VAR

Our baseline quarterly VAR (Q-VAR) reads:

$$y_t = B_0 + \sum_{i=1}^{p} B_i y_{t-i} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Omega), \tag{1}$$

where y_t and B_0 denote $n \times 1$ vectors of variables and constants, respectively. B_i for $i, \ldots, p$ are $n \times n$ matrices of coefficients and Ω is the time-invariant $n \times n$ variance-covariance matrix.

3.2 Quarterly VAR with stochastic volatility

The quarterly VAR with stochastic volatility (Q-SV-VAR) does not assume constant residual variances and includes a law of motion for the (log) volatilities. Following Primiceri (2005), we decompose the time-varying covariance matrix of the reduced-form residuals into a lower-triangular matrix A_t and a diagonal matrix Σ_t according to:

$$A_t \Omega_t A_t' = \Sigma_t \Sigma_t', \tag{2}$$

where the diagonal elements of Σ_t are the stochastic volatilities. A_t has ones on the main diagonal and nonzero entries for the remaining lower triangular elements, describing the contemporaneous relationships between the volatilities. This allows to rewrite the VAR in (1) as:

$$y_t = B_0 + \sum_{i=1}^{p} B_i y_{t-i} + A_t^{-1} \Sigma_t u_t, \qquad u_t \sim N(0, I_n). \tag{3}$$

The laws of motion are modeled by defining σ_t as the vector of the diagonal elements of Σ_t and a_t as the vector of nonzero elements stacked by rows of A_t as follows:

$$\log \sigma_t = \log \sigma_{t-1} + e_t, \qquad e_t = (e_{1,t}, \ldots, e_{n,t})' \sim N(0, \Psi), \tag{4}$$

$$a_t = a_{t-1} + \upsilon_t, \qquad \upsilon_t = (\upsilon_{1,t}', \ldots, \upsilon_{n,t}')' \sim N(0, \Phi). \tag{5}$$

Ψ is diagonal and Φ is block diagonal where the blocks relate to the equations of the VAR in (3).

3.3 Quarterly VAR with time-varying parameter

The quarterly VAR with time-varying parameter is estimated in a homoscedastic specification (Q-TVP-VAR) and with stochastic volatility (Q-TVP-SV-VAR). The Q-TVP-VAR extends the baseline Q-VAR for a random walk process governing the evolution of the VAR coefficients:

$$y_t = Z_t' \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Omega), \tag{6}$$

$$\beta_t = \beta_{t-1} + \chi_t, \qquad \chi_t \sim N(0, Q), \tag{7}$$

where $Z_t = I_n \otimes [1, y_{t-1}', \ldots, y_{t-p}']$ contains all the right-hand side variables of the VAR, β_t is a $k_\beta \times 1$ vector of VAR coefficients, and $Q = diag(q_{\beta_1}^2, \ldots, q_{\beta_{k_\beta}}^2)$. For the Q-TVP-SV-VAR, the stochastic volatility part from (4) and (5) is added to the model.

3.4 Mixed-frequency VAR

Estimation of the mixed-frequency VAR (MF-VAR) follows the Bayesian state-space approach of Schorfheide and Song (2015), which can be combined with the former VAR specifications. To this end, we partition our vector of variables $y_t = [y_{q,t}', y_{m,t}']'$, where $y_{m,t}$ collects the monthly variables and $y_{q,t}$ denotes the quarterly variables at monthly frequency. Since the quarterly variables are observed only in the last month of each quarter, $y_{q,t}$ contains missing observations for the first and second month of each quarter. To construct the measurement equation, we

follow Mariano and Murasawa (2003) and assume that quarterly GDP in log levels ($\log Y_{q,t}$) can be expressed as the geometric mean of an unobserved monthly GDP ($\log \tilde{Y}_{q,t}$):

$$\log Y_{q,t} = \frac{1}{3}(\log \tilde{Y}_{q,t} + \log \tilde{Y}_{q,t-1} + \log \tilde{Y}_{q,t-2}). \tag{8}$$

This expression implies that the quarterly series is a first-order approximation to an arithmetic mean of the unobserved monthly series. To arrive at an expression for quarterly GDP growth ($y_{q,t}$) based on latent monthly GDP growth ($\tilde{y}_{q,t}$), we subtract $\log Y_{q,t-3}$ from (8):

$$\Delta_3 \log Y_{q,t} = y_{q,t} = \frac{1}{3}\tilde{y}_{q,t} + \frac{2}{3}\tilde{y}_{q,t-1} + \tilde{y}_{q,t-2} + \frac{2}{3}\tilde{y}_{q,t-3} + \frac{1}{3}\tilde{y}_{q,t-4}. \tag{9}$$

Combining the unobserved with the observed monthly variables in $\tilde{y}_t = [\tilde{y}'_{q,t}, y'_{m,t}]'$, we define the state vector by $z_t = [\tilde{y}'_t, \ldots, \tilde{y}'_{t-p+1}]$ and write the measurement equation as:

$$y_t = H_t z_t. \tag{10}$$

Assuming that GDP growth is ordered first in the model, H_t is given by:

$$H_t = \begin{bmatrix} H_{1,t} & H_{2,t} \end{bmatrix}', \tag{11}$$

$$H_{1,t} = \begin{bmatrix} 1/3 & 0_{1\times n-1} & 2/3 & 0_{1\times n-1} & 1 & 0_{1\times n-1} & 2/3 & 0_{1\times n-1} & 1/3 & 0_{1\times n-1} & 0_{1\times (p-4)n} \end{bmatrix}, \tag{12}$$

$$H_{2,t} = \begin{bmatrix} 0_{n-1\times 1} & I_{n-1} & 0_{n-1\times pn} \end{bmatrix}, \tag{13}$$

where $H_{1,t}$ translates the disaggregation constraint in (9) into the state-space framework. The missing observations in z_t are replaced by estimated states using the Carter and Kohn (1994) simulation smoother (hereafter CK) with a time-varying dimension of the state-space system (Durbin and Koopman, 2001).[5] If an indicator exhibits a missing observation in t, the corresponding entry in y_t and the corresponding row of H_t are deleted. The transition equation of the MF-VAR in state-space form is given by:

$$z_t = \mu + F z_{t-1} + \upsilon_t, \qquad \upsilon_t \sim N(0, S), \tag{14}$$

where μ and F contain the intercepts and AR-coefficients, respectively. S is a $pn \times pn$ variance-covariance matrix where the first $n \times n$ elements equal Ω and all remaining entries are zero.

We obtain the MF-SV-VAR by setting the first $n \times n$ elements of S to Ω_t using the decomposition in (2) and following the laws of motion in (4) and (5). The MF-TVP-VAR is obtained by allowing F to change over time according to (7). Including both specifications leads to the MF-TVP-SV-VAR. To summarize, we have a total of eight competing models:

1. MF-TVP-SV-VAR: Mixed-frequency VAR with time-varying parameters and stochastic

[5] To increase computational efficiency, we eliminate the monthly series, which are observed in each period of the balanced part of the sample, from the state vector for $t = 1, \ldots, T_B$, where T_B denotes the end of the balanced sample. For a detailed description of this "compact" system we refer to the appendix of Schorfheide and Song (2015).

volatility

2. MF-SV-VAR: Mixed-frequency VAR with stochastic volatility
3. MF-TVP-VAR: Mixed-frequency VAR with time-varying parameters
4. MF-VAR: Mixed-frequency VAR
5. Q-TVP-SV-VAR: Quarterly VAR with time-varying parameters and stochastic volatility
6. Q-SV-VAR: Quarterly VAR with stochastic volatility (benchmark)
7. Q-TVP-VAR: Quarterly VAR with time-varying parameters
8. Q-VAR: Quarterly linear VAR

3.5 Estimation procedure and prior specification

All models are estimated with Bayesian estimation techniques, since most models depend on a large number of parameters and thus make estimation based on frequentist approaches infeasible. The mixed-frequency models are estimated with 4 lags; the quarterly models are estimated with 2 lags.[6] In the following, we provide a brief description of the estimation procedure and the prior specifications. A detailed description is provided in Appendices A and B.

For the Q-VAR, we impose a Jeffrey's prior to abstract from shrinkage, since we use a small-scale VAR with only four variables. For the models' stochastic volatility part, we apply normal priors for the diagonal elements of Σ_t and the lower-triangular elements of A_t and obtain draws using the CK algorithm and the mixture sampler of Kim, Shephard, and Chib (1998) (hereafter KSC). Inverse-Wishart priors are applied for Ψ and Φ, respectively. For the SV-VAR and the MF-SV-VAR, we use normal priors for the VAR coefficients and obtain draws using the GLS-based posterior provided by Clark (2011). For the TVP models, we apply the Gibbs sampler of Del Negro and Primiceri (2015). Specifically, we apply the CK algorithm to draw the VAR coefficients, using a normal prior for β^T and an inverse-Wishart prior for Q.

The amount of time-variation in β_t, a_{it}, and $\log \sigma_{it}$ depends on the magnitude of the random walk variances Q, Ψ, and Φ and their corresponding prior distributions, which are—in part—determined by the hyperparameters k_Q, k_Ψ, and k_Φ:

$$p(Q) \sim IW(k_Q^2 \times T_0 \times V(\hat{\beta}_{OLS}), T_0), \tag{15}$$

$$p(\Psi) \sim IW(k_\Psi^2 \times (1+n) \times I_n, 4), \tag{16}$$

$$p(\Phi_i) \sim IW(k_\Phi^2 \times (i+1) \times V(\hat{A}_{i,OLS}), i+1), \qquad i = 1, \ldots, k-1, \tag{17}$$

where OLS denotes OLS estimates based on the training sample. The literature commonly adopts the hyperparameter values proposed by Primiceri (2005). However, these values are

[6] We set $p = 2$ for the quarterly model to be consistent with the literature on US data (see, e.g., Primiceri, 2005; D'Agostino et al., 2013; Clark and Ravazzolo, 2015). The monthly models have 4 lags to keep them computationally feasible. Furthermore, we require at least four lags to disaggregate quarterly GDP into monthly GDP (see (9)).

calibrated for a quarterly three-variable TVP-SV-VAR and it is not clear, whether they are useful in case of monthly data or other model specifications. Therefore, we follow Amir-Ahmadi et al. (2020) by implementing another layer of priors for those hyperparameters. Moreover, we split k_Q into k_{Q_C} and $k_{Q_{AR}}$, where k_{Q_C} relates to intercept coefficients and $k_{Q_{AR}}$ to AR-coefficients, respectively. By this means, we allow for different degrees of time-variation across these groups of coefficients. The latter is motivated by the observations that time-variation seems to be more pronounced in the intercepts than in AR-coefficients (see, for example, D'Agostino et al., 2013). To be agnostic about the coefficients' degree of time-variation, we use, for each hyperparameter, an inverse-Gamma prior with scale parameter and degrees of freedom equal to 0.1 and 2, respectively, as recommended by Amir-Ahmadi et al. (2020).

For the mixed-frequency models, we initialize the state vector with a normal prior. The prior mean is set to the observed values, implying that for GDP the within-quarter figures equal the quarterly observations. The prior variance is the identity matrix. After having drawn the latent states, the remaining coefficients are drawn conditional on the drawn states (instead of conditional on the observed data).

3.6 In-sample analysis

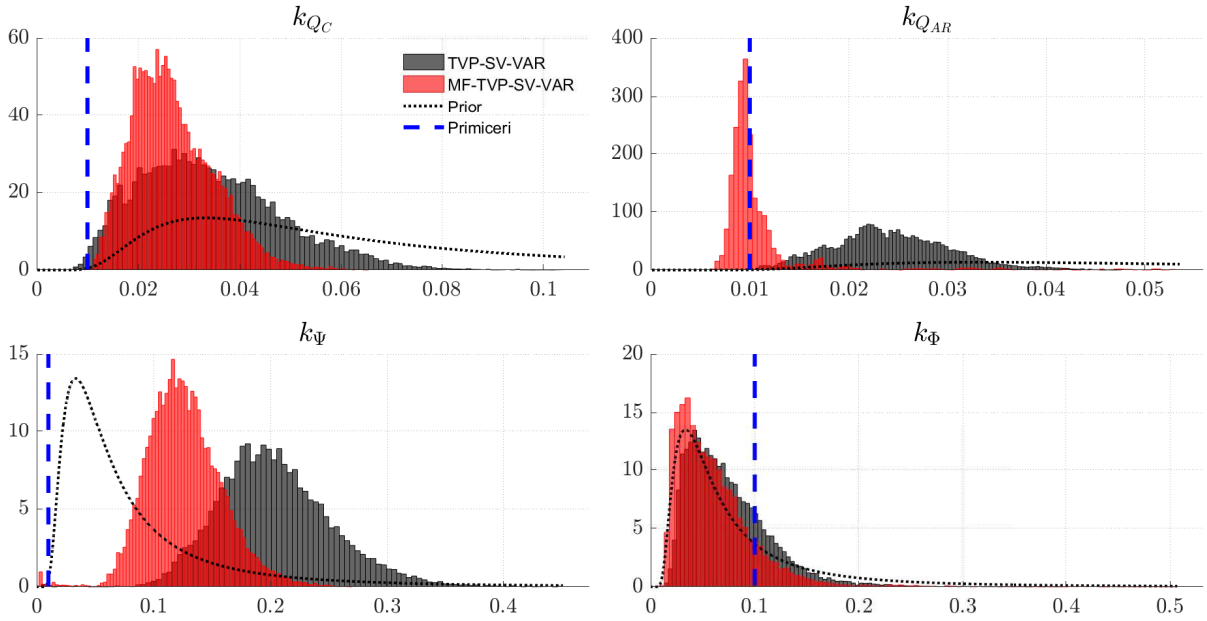
To illustrate the importance of modeling variability in volatility and the VAR-coefficients, as well as the hyperparameter estimation, this sections provides a brief in-sample analysis based on the final data vintage. Figure 1 depicts the posterior distributions of the estimated hyperparameters along with the values proposed by Primiceri (2005) (dashed lines) and the prior distributions (dotted lines). While a direct comparison between Primiceri's values and ours is not straightforward, Figure 1, nevertheless, provides interesting observations.[7] Although we impose the same prior for each hyperparameter, the posterior distributions differ considerably from each other and, in some cases, Primiceri's values. The distributions for k_{Q_C} assign only minor probability mass to Primiceri's values, but imply a stronger prior belief on time variation in the intercept terms. The posterior of $k_{Q_{AR}}$, is (almost) centered around Primiceri's value with a rather low variance for the MF-TVP-SV-VAR, while it induces a stronger prior belief on time variation in the AR coefficients according to the Q-TVP-SV-VAR. Finally, the posteriors of k_Ψ and k_Φ parameterize—for both models—a stronger prior belief about time-variation in the stochastic volatilities, but a weaker one for time-variation in the correlations among the residuals. These results suggest that both estimating the hyperparameters and allowing for heterogeneity among the hyperparameters might provide a better description of the data generating process and thus, might increase the forecast performance.

Figure 2 plots the posterior means of the standard deviations of the reduced-form residuals from the MF-TVP-SV-VAR and Q-TVP-SV-VAR.[8] We assume that the volatility estimates from the Q-TVP-SV-VAR are constant within a quarter to make them comparable across frequencies.

[7] On the one hand, we employ a different model specification. On the other hand, the remaining parts of the priors are based on a different training sample. Both can lead to different priors, in spite of identical hyperparameters, and hence, complicate comparison.

[8] We also examined the volatilites for different data vintages to investigate the impact of data revisions and different values for the hyperparameters. Analogous to Clark (2011), we obtain very similar estimates for the different vintages, and thus we only report results for the latest vintage.

Figure 1: Posterior distributions of hyperparameters



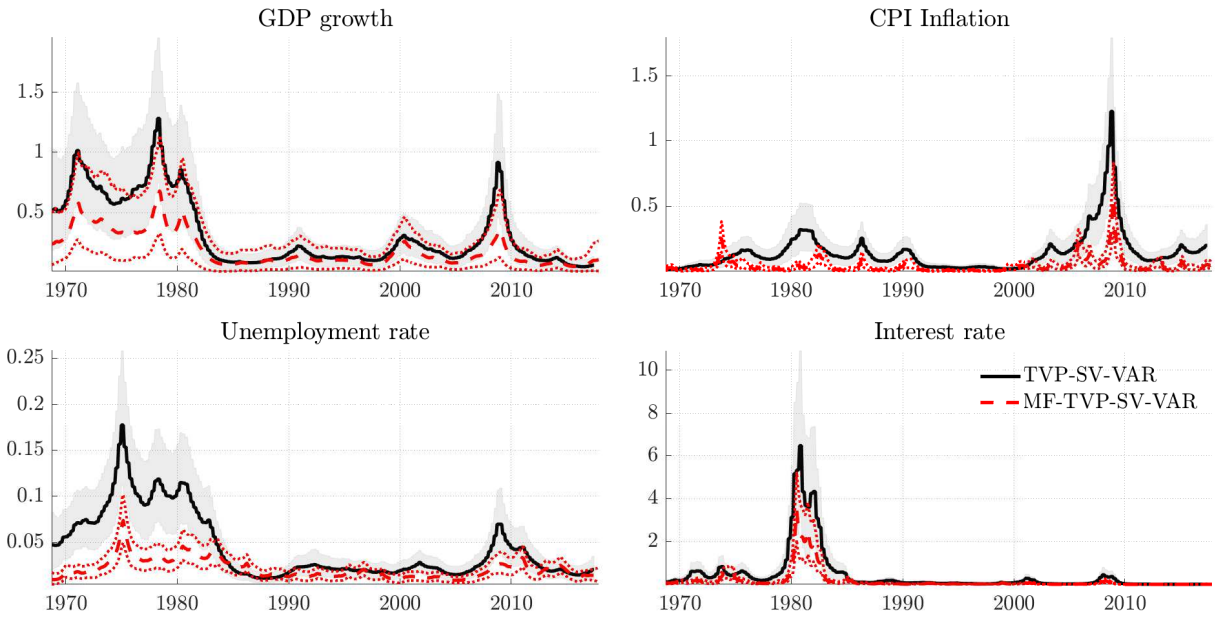
Notes: Figure shows the posterior distributions of the hyperparameters along with the respective prior distributions (dotted line) and the values proposed by Primiceri (2005) (dashed line).

The estimates of the Q-TVP-SV-VAR are smoother than those of Primiceri (2005), reflecting the weaker prior for time-variation in the residuals' correlations. Until the mid 1980s, the estimated volatilities are quite high and then fall sharply, indicating the beginning of the Great Moderation. Except for the increase during the burst of the dot-com bubble in 2000 and the rise during the Great Recession, they remain roughly at the levels of the mid 1980s. At the end of the sample, however, there is again a decline in volatility, indicating a time during which the US was remarkably less exposed to absolute shocks hitting the economy. Thus, as suggested by Clark (2009), the Great Recession seems to have simply interrupted, but not ended, the Great Moderation—the latest volatility estimates for GDP growth is the lowest of the entire sample.

The estimates from the MF-TVP-SV-VAR closely track the evolution of its quarterly counterpart. However, they are somewhat smaller, indicating that using monthly information absorbs part of the fluctuations in the volatility. This finding confirms the results of Carriero, Clark, and Marcellino (2015), who employ a Bayesian mixed-frequency model without time-variation in the AR-coefficients.

With regard to the VAR coefficient, we obtain—for both models—the largest variability for the intercept coefficients; the remaining parameters exhibit only minor time-variation (see Figures 7 and 8 in Appendix D). Overall, the results suggest that modeling variability in both volatility and the intercepts is more important for achieving precise forecasts than modeling time-varying autoregressive dynamics. Our results support the modeling strategy of Götz and Hauzenberger (2018), who specify time-variation only in the intercepts, but treat the hyperparameters as exogenous values.

Figure 2: Standard deviations of reduced-form residuals



Notes: Figure depicts the posterior means of the residual standard deviations from the last data vintage at monthly frequency. Quarterly estimates are assumed to be constant within a quarter. Shaded areas and dotted lines refer to 68% posterior probability bands.

3.7 Now- and forecasting

The quarterly models are estimated on balanced datasets containing all available information from the previous quarter. To generate the predictive distributions, we compute sequences of h_{max} normally distributed innovations with covariance Φ, Ψ, and Q to produce time paths for the elements of A_t, Σ_t, and β_t, respectively. Based on these trajectories, we simulate y_t h_{max} periods into the future. The first forecast is a nowcast, since it is generated in and refers to the respective current quarter.

Additional notation is helpful in describing how we obtain the predictive distributions of the mixed-frequency models. Let T_M denote the last month of the indicator that has the shortest publication lag and let $Z^{T_M} = [z_1, \ldots, z_{T_M}]$ denote the sequence of simulated state vectors. Note that the CK algorithm provides draws for the latent states until T_M. To obtain $Z^{T_m+1:T_m+h_{max}}$, we generate time paths for the elements of A_t, Σ_t, and β_t and simulate the state vector z_t forward using these time paths. Accordingly, if T_M belongs to I3, the CK algorithm provides draws for the entire last available quarter and by averaging over these draws we obtain the nowcasts. The forecasts are generated by averaging over the trajectories $Z^{T_m+1:T_m+h_{max}}$. However, if T_M belongs to I1 or I2, the CK algorithm does not provide draws of the latent states for the entire quarter since none of the indicators is available for the entire quarter. In this case, we average over the available CK draws and the simulated trajectories referring to this quarter to get the nowcast. The forecasts are calculated from the averages of the remaining trajectories.[9]

[9]For instance, in February, the T-Bill rate is available until February (T_M), while inflation and unemployment rate are available until January ($T_M - 1$). Hence, the CK algorithm provides draws for each indicator until T_M. The figures for March ($T_M + 1$) are generated using the time paths for A_t, Σ_t, and β_t. The forecast for the first quarter is the average over the figures referring to $T_M - 1$ to $T_M + 1$.

4 Forecast metrics

We evaluate the models' forecasts with respect to point and density forecasts. Subsequently, m, i, and h denote the model, variable, and forecast horizon, respectively, for the forecast sample $t = 1, \dots, N$. We measure point forecast accuracy using relative root mean squared errors:

$$\text{relative RMSE}_h^{i,m} = \frac{\sqrt{\frac{1}{N}\sum\left(\hat{y}_{t+h}^{i,m} - y_{t+h}^{i}\right)^2}}{\sqrt{\frac{1}{N}\sum\left(\hat{y}_{t+h}^{i,B} - y_{t+h}^{i}\right)^2}}, \tag{18}$$

where $\hat{y}_{t+h}^{i,B}$ refers to the forecast of the benchmark Q-SV-VAR.[10] We test for statistical differences in forecast accuracy by applying the Diebold and Mariano (1995) test.

Regarding density forecasts, we apply the continuous ranked probability score (CRPS). To compute the CRPS, we follow Gneiting and Ranjan (2011) and use the score function:

$$S(p_t^{i,m}, y_t^i, \nu(\alpha)) = \int_0^1 \text{QS}_\alpha(P_t(\alpha)^{-1}, y_t)\nu(\alpha)d\alpha, \tag{19}$$

where $\text{QS}_\alpha(P_t(\alpha)^{-1}, y_t^i) = 2(I\{y_t^i \leq P_t(\alpha)^{-1}\} - \alpha)(P_t(\alpha)^{-1} - y_t^i)$ is the quantile score for forecast quantile $P_t(\alpha)^{-1}$ at level $0 < \alpha < 1$. $I\{y_t^i \leq P_t(\alpha)^{-1}\}$ is an indicator function taking the value 1 when $y_t^i \leq P_t(\alpha)^{-1}$ and 0 otherwise. P_t^{-1} denotes the inverse of the cumulative predictive density function and $\nu(\alpha)$ is a weighting function. Using a uniform weighting scheme ($\nu(\alpha)=1$) and dividing by the number of generated densities yields the average CRPS:

$$\text{CRPS}_h^{i,m} = \frac{1}{N}\sum S(p_{t+h}^{i,m}, y_{t+h}^{i}, 1). \tag{20}$$

According to (20), a lower score indicates a better calibrated predictive density. We evaluate the CRPS as ratios relative to our benchmark:

$$\text{relative CRPS}_h^{i,m} = \frac{CRPS_h^{i,m}}{CRPS_h^{i,B}}. \tag{21}$$

We obtain approximate inference on whether the scores are significantly different from the benchmark by regressing the differences between the scores of each model and the benchmark on a constant. A t-test with Newey-West standard errors on the constant indicates whether these average differences are significantly different from zero (D'Agostino et al., 2013).

5 Results

In this section, we discuss the results from the forecast experiment. We evaluate both point and density forecasts. Regarding the point forecasts, we first assess the models' nowcast accuracy. Second, we evaluate the accuracy of the point forecasts and predictive densities with respect to

[10] Since several studies demonstrate that VARs with stochastic volatility outperform constant volatility VARs (see, for instance, Clark, 2011; Clark and Ravazzolo, 2015; Chiu, Mumtaz, and Pintér, 2017), we do not use the Q-VAR as our benchmark.

the subsequent quarters.[11] We provide results for the entire recursive sample (1995Q1–2017Q4) and for a shorter sample period of 2008Q1 until 2017Q4 to assess whether a possible structural break around the Great Recession affects the forecast performance.

5.1 Nowcast evaluation

Table 1 presents the results for the nowcast exercise taking into account the information sets I1 to I3. It provides three main takeaways. First, the mixed-frequency models outperform the quarterly models. On average, over all information sets and variables, the best nowcast performance is obtained by the MF-TVP-SV-VAR and the MF-SV-VAR, which improve on the benchmark (Q-SV-VAR) by roughly 35%. Second, most of the time, the nonlinear MF-models outperform the linear MF-VAR, indicating that—apart from using monthly information—parameter instability is beneficial also in a mixed-frequency setting. Third, the MF-models' relative performance improves with more information available, showing that the models are able to efficiently process the sequential data releases.

For GDP growth, only MF-models significantly outperform the benchmark. The best performance, for both samples, is obtained by the MF-SV-VAR. This result suggests that, from a nowcasting perspective, it is more important to account for the decline in output growth volatility than to account for changes in output growth dynamics. For inflation, the MF-TVP-SV-VAR delivers the best performance for the entire sample; it improves on the benchmark by, on average, 50%. Concerning the shorter sample, the MF-SV-VAR matches up with the MF-TVP-SV-VAR, suggesting that stochastic volatility has gained importance in the post-Great Recession period. Regarding both samples, the results indicate that notably with little information about the current quarter the MF-TVP-SV-VAR provides large gains in forecast accuracy relative to the competing models. The latter is particularly relevant because expert forecast, for example the SPF, are usually published in the second month of a quarter. With more information available, however, the differences towards the remaining MF-models vanish. For the unemployment rate, the MF-TVP-SV-VAR provides the most accurate nowcasts across all information sets with gains of about 40%. Though, it appears that non-linearity is not as important as for the remaining variables—the differences to the linear MF are minor. The latter is maybe not surprising given that the fluctuations in volatility of the unemployment rate are less pronounced compared to the remaining variables (see Figure 2).

In total, the nowcast exercise provides strong evidence in favor of nonlinear forecasting models. In particular, stochastic volatility seems to be a major determinant of precise nowcasts, which is consistent with, for instance, Carriero et al. (2015). Allowing for time-varying parameters without stochastic volatility improves accuracy relative to the benchmark but is—in most cases—inferior to models with stochastic volatility. Inflation nowcasts in turn benefit from combining both specifications.

[11] We abstract from evaluating the nowcasts with respect to predictive densities. Depending on the information sets, the nowcasts of the mixed-frequency models consist of quarterly averages over draws from the CK algorithm and realizations. Therefore, the nowcast densities of the mixed-frequency models are very narrow compared to the quarterly models and thus hardly comparable.

Table 1: Real-time nowcast RMSEs

Model	1990-2017 I1	1990-2017 I2	1990-2017 I3	2008-2017 I1	2008-2017 I2	2008-2017 I3
	GDP growth					
MF-TVP-SV-VAR	0.91	0.89	0.90	0.83	0.80	0.81
MF-SV-VAR	**0.85***	**0.80****	**0.78****	**0.80***	**0.74**	**0.72***
MF-TVP-VAR	0.88*	0.86*	0.82*	0.92	0.85	0.80
MF-VAR	0.99	0.96**	0.93	0.94	0.89	0.86
Q-TVP-SV-VAR	1.06	1.08	1.08	1.13	1.16	1.15
Q-TVP-VAR	0.98	0.99	0.99	0.99	1.00	1.01
Q-VAR	1.11***	1.12***	1.11***	1.12***	1.12***	1.12***
Q-SV-VAR	0.65	0.65	0.65	0.78	0.77	0.77
	Inflation					
MF-TVP-SV-VAR	**0.77***	**0.49****	**0.29****	**0.73***	0.46*	0.21*
MF-SV-VAR	0.86	0.52**	0.30***	0.77	**0.44**	0.21*
MF-TVP-VAR	0.85**	0.52*	**0.29*****	0.86*	0.50	0.22*
MF-VAR	0.89	0.53*	0.30***	0.80	0.45	**0.20***
Q-TVP-SV-VAR	0.87**	0.87**	0.88**	0.88	0.88	0.89
Q-TVP-VAR	0.93**	0.93**	0.93	0.94	0.94	0.95
Q-VAR	1.04***	1.04***	1.04***	1.03	1.03	1.03
Q-SV-VAR	0.56	0.56	0.56	0.76	0.76	0.76
	Unemployment rate					
MF-TVP-SV-VAR	**0.80***	**0.60****	**0.36*****	**0.74**	**0.54****	0.32**
MF-SV-VAR	0.83***	0.62***	0.38***	0.83**	0.59**	0.33**
MF-TVP-VAR	0.90	0.65**	**0.36*****	0.86	0.61**	**0.31****
MF-VAR	0.82***	0.61**	0.37***	0.80**	0.57**	0.32**
Q-TVP-SV-VAR	0.95	0.95	0.96	0.93	0.93	0.94
Q-TVP-VAR	1.01	1.01	1.01	1.01	1.01	1.02
Q-VAR	1.04*	1.04*	1.04	1.05	1.05	1.05
Q-SV-VAR	0.28	0.27	0.27	0.35	0.34	0.34

Notes: RMSEs are reported in absolute terms for the benchmark model (bottom row of each panel) and as ratios relative to the benchmark for the remaining models. A ratio below unity indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and information set. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, according to the Diebold-Mariano test with Newey-West standard errors.

5.2 Forecast evaluation

The results in Table 2 show that mixed-frequency VARs provide competitive forecasts even for higher horizons and for both samples.[12] In the case of the unemployment rate, modeling within-quarter dynamics is particularly beneficial—at each horizon even the worst performing mixed-frequency VAR outperforms the best performing quarterly VAR. Moreover, the results reveal that the models' forecast performance substantially differs across variables. The best relative performance, over all variables and horizons, is delivered by the MF-SV-VAR and the MF-TVP-SV-VAR; the corresponding RMSEs are roughly 10% lower than those of the benchmark.

[12]Since the marginal impact of an additional month of information becomes less important for forecasts at higher horizons, the RMSEs for higher horizons become similar across the information sets. For the forecast evaluation, we therefore compute total RMSEs by averaging over the entire forecast sample. Figure 9 in Appendix D plots the relative RMSE for each information set.

Table 2: Real-time forecast RMSEs

Model	1990-2017 h = 2	1990-2017 h = 3	1990-2017 h = 4	2008-2017 h = 2	2008-2017 h = 3	2008-2017 h = 4
	GDP growth					
MF-TVP-SV-VAR	1.02	1.00	1.02	1.00	1.03	1.04
MF-SV-VAR	**1.00**	1.01	1.00	**0.97**	1.01	1.01
MF-TVP-VAR	1.09***	1.13***	1.05	1.08	1.19*	1.06
MF-VAR	1.13***	1.18***	1.16***	1.11***	1.18***	1.21***
Q-TVP-SV-VAR	1.06	**0.98**	1.01	1.10	1.03	1.02
Q-TVP-VAR	1.03	0.99	**0.97**	1.08	1.02	**0.97**
Q-VAR	1.14***	1.14***	1.13***	1.17***	1.17***	1.18***
Q-SV-VAR	0.66	0.66	0.68	0.84	**0.86**	0.86
	Inflation					
MF-TVP-SV-VAR	**0.81****	0.84***	0.82***	**0.80****	0.87***	0.91***
MF-SV-VAR	0.90**	0.89***	0.86***	0.86**	0.87***	**0.87****
MF-TVP-VAR	0.93**	0.96**	0.93	0.97	1.05	1.11**
MF-VAR	0.98	1.09***	1.15***	0.92	1.02	1.06
Q-TVP-SV-VAR	0.84***	**0.83****	**0.81****	0.83***	**0.84****	**0.87****
Q-TVP-VAR	0.89***	0.89***	0.86***	0.92***	0.93***	0.94**
Q-VAR	1.08***	1.14***	1.22***	1.05**	1.10***	1.15***
Q-SV-VAR	0.63	0.64	0.64	0.89	0.87	0.78
	Unemployment rate					
MF-TVP-SV-VAR	**0.79****	0.88**	0.95	**0.76****	0.85	0.95
MF-SV-VAR	0.86***	0.92**	0.95*	0.85**	0.91*	0.94
MF-TVP-VAR	0.83***	**0.87****	**0.89****	0.80**	**0.83****	**0.85****
MF-VAR	0.84***	0.91**	0.94*	0.83**	0.90*	0.93
Q-TVP-SV-VAR	1.00	1.04	1.08	0.98	1.04	1.09
Q-TVP-VAR	1.03	1.05	1.07	1.03	1.05	1.08
Q-VAR	1.03***	1.03***	1.03***	1.04***	1.04***	1.04***
Q-SV-VAR	0.48	0.71	0.93	0.65	1.00	1.33

Notes: RMSEs are reported in absolute terms for the benchmark model (bottom row of each panel) and as ratios to the benchmark model for the remaining models. A ratio below unity indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, according to the Diebold-Mariano test with Newey-West standard errors.

For GDP growth, the benchmark is hard to beat. The MF-SV-VAR, the Q-TVP-VAR, and the Q-TVP-SV-VAR provide a better forecast performance for some horizons, albeit not statistically significant. For inflation, the Q-TVP-SV-VAR and the MF-TVP-SV-VAR deliver the best performance on average over all horizon, suggesting that time-variation in each coefficient is crucial for inflation forecasts. Thus, our results confirm the findings from previous studies based on quarterly models (see, among others, D'Agostino et al., 2013; Barnett et al., 2014; Faust and Wright, 2013) by use of mixed-frequency models. Moreover, while the TVP-models' performance tends to deteriorate in the shorter sample, the SV-models' performance enhances, again providing evidence that stochastic volatility has gained importance in the shorter sample. For the unemployment rate, the MF-models consistently outperform the benchmark, while the quarterly models fail to do. Hence, the results provide evidence that both intra-quarterly dynamics and

time-variation in the VAR-coefficients are particularly important.

In sum, the results are consistent with findings from previous studies, indicating that the gains in accuracy due to variations in the VAR-coefficients are smaller than the gains induced by stochastic volatility. However, using models with both features provides, on average over all variables, the most accurate forecasts. Finally, the results provide evidence that modeling within-quarter dynamics is beneficial also regarding short-term forecasts.

5.3 Comparison with survey-based forecasts

Since it is commonly found that survey-based forecasts are hard to beat (see, for example, Faust and Wright, 2013), we further assess the forecast performance of our MF-VARs relative to the forecasts provided by the SPF. To align the MF-VARs' information set with those of the SPF participants, we only resort to the forecasts from I2, providing 90 samples for the evaluation.[13]

Figure 3 depicts forecast errors of the MF-VARs relative to those of the SPF for the three variables. For GDP growth (left panel) and the unemployment rate (right panel), the SPF clearly outperforms each MF-VAR. The latter probably stems from the fact that survey participants consider a much broader information set than included in our small-scale VARs.[14] Regarding inflation, however, even small-scale MF-VARs provide very competitive nowcasts—each nonlinear specification slightly improves on the SPF, which itself is found to provide very accurate inflation nowcasts (Faust and Wright, 2013). For higher horizons, the SPF delivers more accurate inflation forecasts though. We moreover investigate whether the the quarterly VARs' forecast performance can be improved by conditioning the latter on the SPF nowcasts.[15] Figure 10 in the Appendix shows that this procedure indeed improves the Q-VARs' forecast accuracy. Overall, the gains are, however, small and die out quickly. In particular the MF-TVP-SV-VAR nevertheless provides very competitive predictions.

5.4 Predictive density evaluation

The results for the CRPS are displayed in Table 3. Overall, the results point to the usefulness of within-quarter information in delivering well calibrated predictive densities; the mixed-frequency models provide better results on average over all variables and horizons than their quarterly counterparts. The MF-TVP-SV-VAR provides the best performance with a reduction in CRPS of 19% followed by the MF-SV-VAR with 11% (on average over all variables and horizons). This emphasizes the importance of stochastic volatility for generating accurate predictive densities.

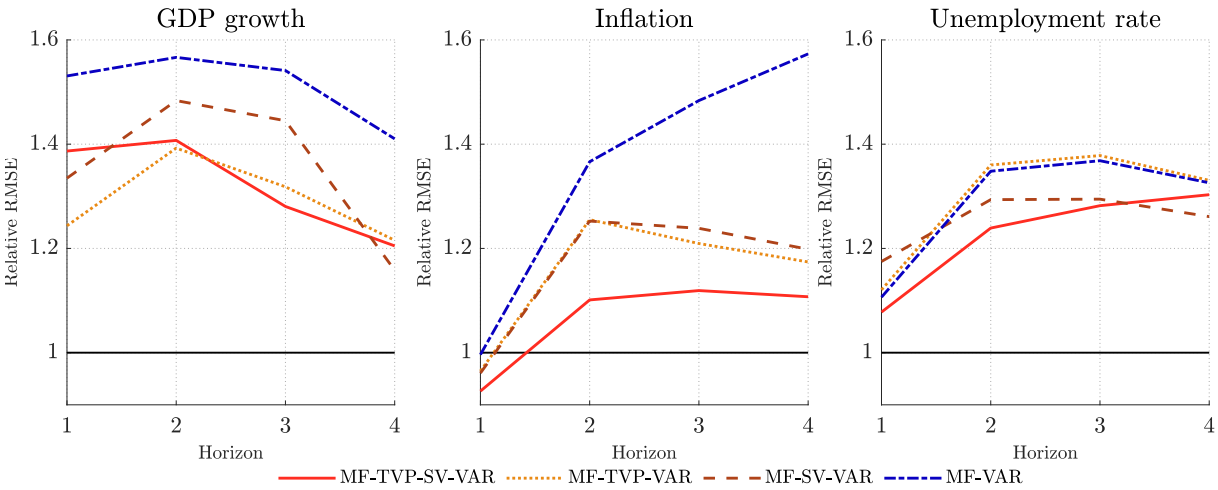
For GDP growth density forecasts, the benchmark is again difficult to beat—no model significantly improves on the benchmark. Only the MF-TVP-SV-VAR provides an (insignificant)

[13]The SPF participants' submission deadline for the first (second, third, fourth) quarter is the second to third week of February (May, August, November).

[14]For example, Brave, Butters, and Justiniano (2019) show that the forecast accuracy at medium-term horizons of a MF-VAR with regard to GDP growth tends to improve with more information included in the model.

[15]Specifically, we estimate the Q-VARs until T_B, add the SPF nowcasts for T_{B+1}, and compute forecasts for T_{B+2} until $T_{B+h_{max}/3}$. Note that we do not update the coefficients given the SPF nowcasts. Alternatively, one could also use more sophisticated methods for utilizing external forecasts, for example, entropic tilting. For a comparison of methods to combine external and model-based predictions see Krüger, Clark, and Ravazzolo (2017).

Figure 3: Comparison of MF-VARs with Survey of Professional Forecasters



Notes: Comparison of MF-VAR forecasts with mean forecasts from the Survey of Professional Forecasters (SPF). GDP growth and inflation forecasts from MF-VARs are transformed into annualized rates. Figures below (above) unity indicate that the models provides smaller (larger) forecast errors than the SPF. Sample: 1990–2017.

improvement, although its points forecast are worse than those of the benchmark. Regarding inflation, the results indicate two outcomes. First, the Q-TVP-SV-VAR and the MF-TVP-SV-VAR deliver the largest (and significant) improvements on the benchmark. Hence, as for point forecasts, it is important to model time-variation in both the parameters and the residual variances to obtain precise predictive densities. Second, including time-variation in the parameters does play a vital role since both the MF-TVP-VAR and the Q-TVP-VAR offer strong improvements of roughly 10% over the benchmark.

For the unemployment rate, the results are different from the point forecasts evaluation. In this case, the MF-TVP-SV-VAR delivers the best performance, improving on the benchmark by up to 14% followed by the MF-SV-VAR with 11%. The MF-TVP-VAR, which provides very accurate point forecasts, in turn performs slightly worse with gains of up to 9%. Moreover and in contrast to inflation, each mixed-frequency model improves both on the benchmark and on its quarterly counterpart. Thus, it is crucial to include intra-quarterly information and stochastic volatility to generate precise predictive densities for the unemployment rate.

In summary, the results of the predictive density evaluation support the findings from the point forecast evaluation. Using mixed-frequency models is beneficial over all variables and horizons. It significantly improves results for inflation and the unemployment rate. In addition, we confirm the importance of stochastic volatility in density forecasting by use of mixed-frequency VARs. We provide evidence that combining stochastic volatility, time-varying parameters, and mixed-frequencies significantly improves the accuracy of predictive densities. However, for the inflation rate adding mixed frequency does not pay-off.

Table 3: Real-time forecast CRPS

Model	1990-2017 h = 2	h = 3	h = 4	2008-2017 h = 2	h = 3	h = 4
	GDP growth					
MF-TVP-SV-VAR	**1.00**	**0.95**	**0.97**	**0.94**	**0.90**	**0.90**
MF-SV-VAR	1.02	1.02**	1.01	0.97	1.01	1.02
MF-TVP-VAR	1.18***	1.39***	1.59***	1.21***	1.51***	1.85***
MF-VAR	1.15***	1.20***	1.19***	1.14***	1.24***	1.29***
Q-TVP-SV-VAR	1.02	0.96	1.00	1.02	0.95	0.96
Q-TVP-VAR	1.02	1.00	1.00	1.02	0.96	0.93
Q-VAR	1.17***	1.17***	1.15***	1.20***	1.22***	1.25***
Q-SV-VAR	0.35	0.36	0.37	0.37	0.38	0.39
	Inflation					
MF-TVP-SV-VAR	**0.80*****	0.81***	0.77***	**0.79*****	0.87**	0.87**
MF-SV-VAR	0.92*	0.95***	0.93***	0.81***	0.88**	0.89**
MF-TVP-VAR	0.88***	0.90**	0.85***	0.95	1.09	1.10
MF-VAR	0.98	1.12	1.16**	0.86**	1.01	1.02
Q-TVP-SV-VAR	0.83***	**0.80*****	**0.76*****	0.84***	**0.83*****	**0.83*****
Q-TVP-VAR	0.88***	0.86***	0.81***	0.91**	0.92**	0.88***
Q-VAR	1.11***	1.17***	1.23***	1.08**	1.13***	1.16***
Q-SV-VAR	0.31	0.33	0.35	0.40	0.39	0.39
	Unemployment rate					
MF-TVP-SV-VAR	**0.80*****	**0.86***	**0.91**	**0.80****	**0.86**	0.92
MF-SV-VAR	0.84***	0.91***	0.94***	0.82***	0.87**	**0.90****
MF-TVP-VAR	0.89**	0.92	0.94	0.87	0.89	**0.90**
MF-VAR	0.84***	0.89***	0.92***	0.83**	0.88**	0.91*
Q-TVP-SV-VAR	1.00	1.02	1.02	0.99	1.02	1.04
Q-TVP-VAR	1.06**	1.06	1.06	1.06	1.06	1.07
Q-VAR	1.03**	1.02	1.02	1.05	1.05	1.06*
Q-SV-VAR	0.24	0.36	0.47	0.27	0.41	0.55

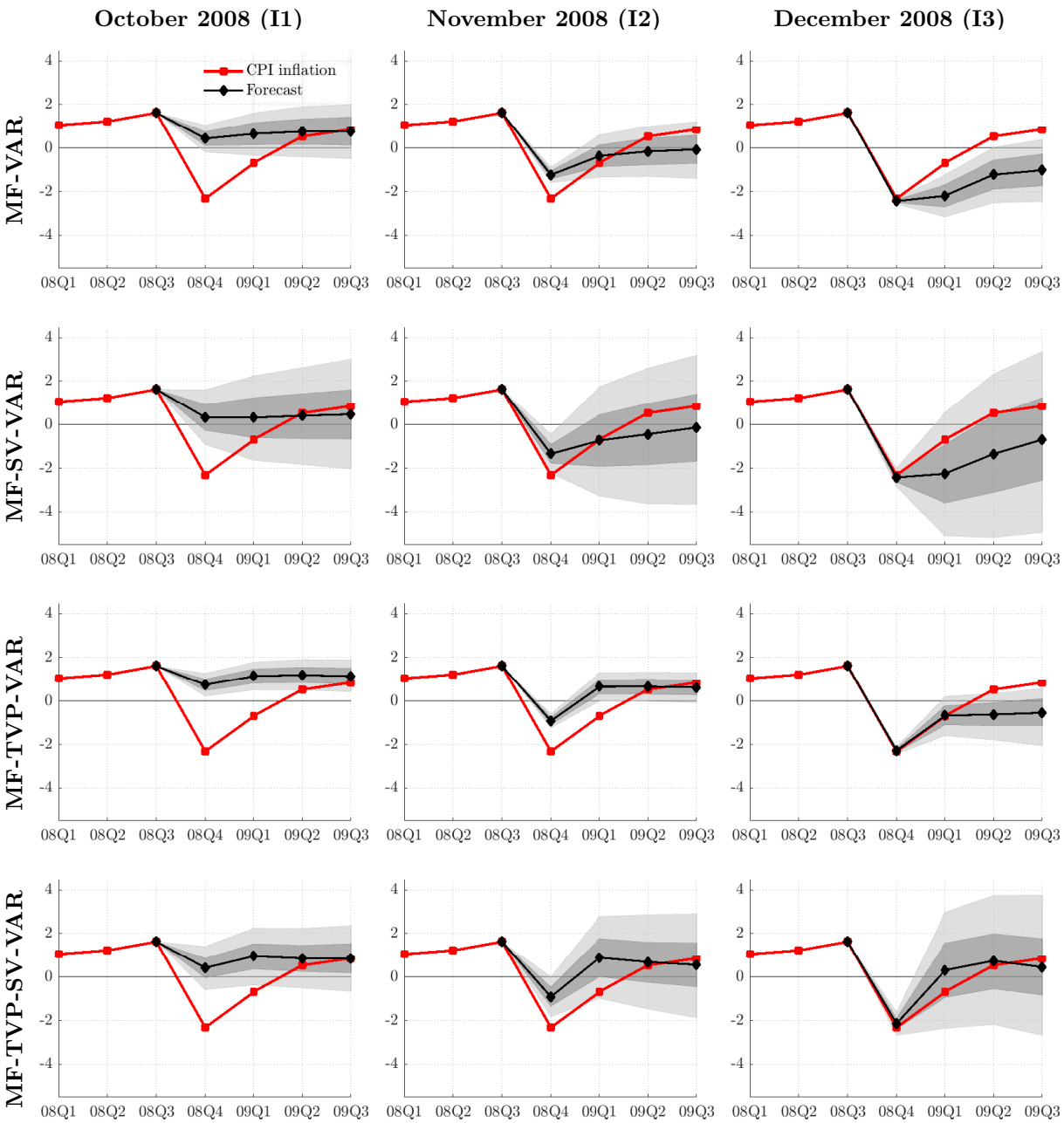
Notes: The scores are reported in absolute terms for the benchmark model (the bottom row of each panel) and as ratios to the benchmark for the remaining models. A ratio below unity indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, according to a t-test on the average difference in scores relative to the benchmark model with Newey-West standard errors.

5.5 Forecasting during the Great Recession

So far we have demonstrated that modeling intra-quarterly dynamics, on average, significantly improves forecast accuracy. Now we take a closer look at the MF-models' absolute performance during the Great Recession, which is of great interest, because many structural and nonstructural models failed to provide accurate forecasts for the steep contraction and the following upswing in 2008/2009. Figures 4 and 5 depict real-time quarter-on-quarter CPI inflation and the unemployment rate (red lines) along with both the means (black lines) and 60% as well as 90% error bands (shaded areas) from the predictive distributions, respectively. The figures' columns refer to the data vintages of October 2008 until December 2008 and demonstrate how the arrival of new data points affects the forecasts.

First, we consider the inflation forecasts computed with the vintage of October 2008 (first

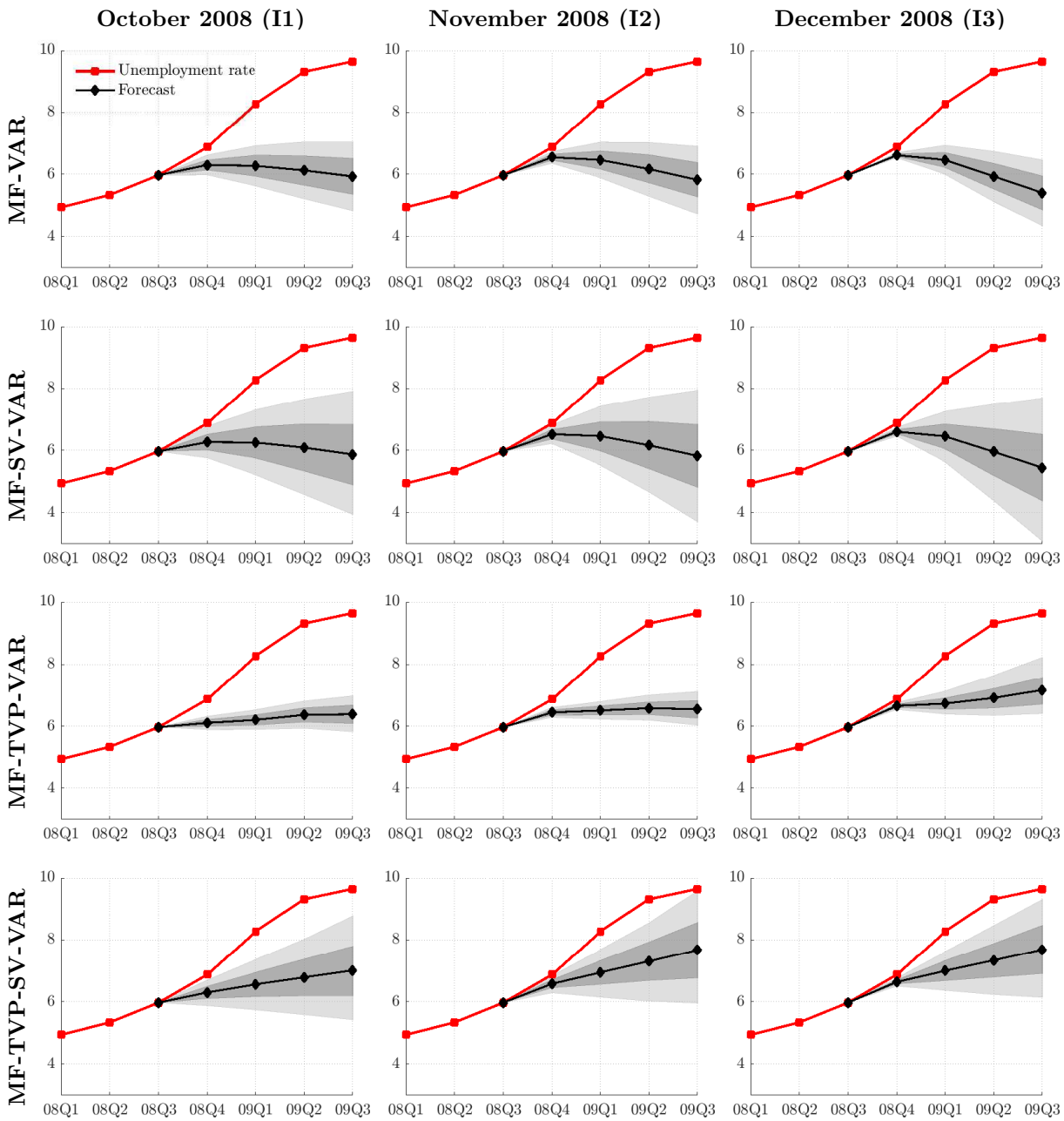
Figure 4: Inflation Forecasts During the Great Recession



Notes: Rows refer to mixed-frequency models; columns refer to the forecast origins, i.e., the information sets. Red line indicates quarter-on-quarter real-time CPI inflation; black line is the mean of the predictive distribution. Shaded areas are 60% and 90% error bands from the predictive distributions.

column). Note that in this month the models do not have any information on the current quarter except for the T-Bill rate of October. In October 2008, the models' posterior means are close to each other for each horizon—for the nowcast, all of them lie at roughly 0.5%, which is about three percentage points too high compared to the realization. The MF-VAR and the MF-TVP-VAR deliver narrow intervals, which assign only a small fraction of probability mass to negative inflation rates. The MF-SV-VAR and the MF-TVP-SV-VAR in turn generate much wider intervals, clearly including negative inflation rates. However, the realization is not included

Figure 5: Unemployment rate forecasts during the Great Recession



Notes: Rows refer to mixed-frequency models. columns refer to the forecast origins, i.e., the information sets. Red line indicates quarter-on-quarter real-time unemployment rate; black line is the mean of the predictive distribution. Shaded areas are 60% and 90% probability bands from the predictive distributions.

in any interval. In November 2008, the posterior means are still similar, but become much more pessimistic. The models correctly anticipate a negative inflation rate for 2008Q4 (approx. -1%). Thus, as indicated in Section 5.1, the forecast errors become remarkably smaller due to the additional monthly observations. Moreover, while the constant coefficient VARs predict a slow recovery with negative inflation rates until 2009Q3, the TVP-VARs correctly anticipates the recovery from 2009Q2 onward. In December 2008, the models produce a forecast error of almost zero for 2008Q4 with a narrow forecast interval. The subsequent recovery, however, is

best predicted by the MF-TVP-SV-model. Driven by the pessimistic nowcasts, the remaining models forecast negative inflation rates for the entire forecast horizon. For the unemployment rate (Figure 5), the MF-TVP-SV-VAR also provides the best performance. While the 90% intervals do not contain the realizations for 2009Q1 until 2009Q3, only the VARs with time-varying coefficients predicts a prolonged increase in the unemployment rate. This increase in turn is more pronounced according to the MF-TVP-SV-VAR.

In summary, these results illustrate that the mixed-frequency models can translate intra-quarterly information into more precise point and density forecasts. Furthermore, this example supports the findings from Sections 5.1 and 5.2; it demonstrates the importance of combining stochastic volatility with time-varying parameters for accurate now- and forecasts.

5.6 Forecast combination

Instead of estimating a model that directly captures parameter instability an obvious alternative is to estimate several models and generate forecasts using a combination of those. As shown by, for example, Clark and McCracken (2008), particularly forecasts from small-scale VARs can substantially benefit from this approach. Subsequently, we combine the forecasts from our models by applying both an equal-weighted combination scheme and the optimal prediction pool of Geweke and Amisano (2011). We derive the vector of optimal weights $\mathbf{w_t^{i,*}} = [w_{t,1}^i, \dots, w_{t,M}^i]'$ for variable i and M different models for each t by recursively minimizing the CRPS function:

$$\mathbf{w_t^{i,*}} = \arg\min_{w_t^i} \sum_{j=\tau+1}^{t} \left[\int_0^1 \text{QS}_\alpha \left(\sum_{m=1}^{M} w_{t,m}^i P_j(\alpha)^{-1}, y_j^i \right) \nu(\alpha) d\alpha \right], \tag{22}$$

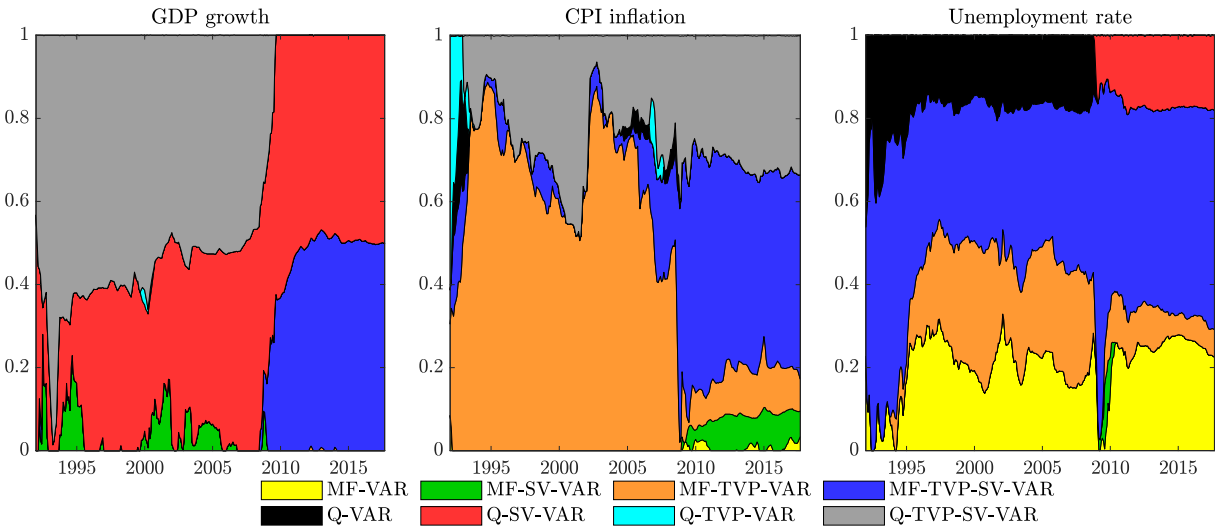
$$\text{s.t.: } w_{t,m}^i \geq 0, \text{ for } \quad m = 1, \dots, M \quad \text{and} \quad \sum_{m=1}^{M} w_{t,m}^i = 1 \ \forall t, \tag{23}$$

where τ denotes a two-year warm-up sample. On the one hand, time-varying weights are highly informative regarding shifts in the relative forecast performance among a set of competing models (Pettenuzzo and Timmermann, 2017). On the other hand, changing weights can reflect important changes in the underlying economic structure (Del Negro, Hasegawa, and Schorfheide, 2016).

Figure 6 displays the time-varying weights for $h = 2$.[16] For GDP growth, a great deal of the total probability mass is assigned to the Q-SV-VAR and the Q-TVP-SV-VAR until the Great Recession hits the US economy. Thereafter, the MF-TVP-SV-VAR quickly gains importance by receiving a weight of roughly 50%, outweighing the Q-SV-VAR. For inflation, the MF-TVP-VAR obtains the largest weight until the Great Recession. Afterwards, the MF-TVP-SV-VAR receives the highest weight, confirming the results that parameter instability in each coefficient is important for inflation forecasts. Regarding the unemployment rate, Figure 6 confirms the findings from the previous sections; almost the entire probability mass is consistently assigned to MF-models with the MF-TVP-SV-VAR receiving the largest share the entire sample (about 40%). Thus, for each variable, the model that receives the largest weight at the end of the sample includes time-variation in both the VAR coefficients and the residual variances. Moreover, the

[16] The figures for the remaining horizons look qualitatively similar and are available upon request.

Figure 6: Optimal prediction pools



Notes: Probability weights for different models according to optimal prediction pool for $h = 2$. The weights are derived by recursively solving the minimization problem in (22).

MF-models' share at the end of the sample, is larger than those QF-models, supporting the importance of modeling intra-quaterly dynamics.

Table 4 depicts the point forecast performance of both combinations schemes relative to the benchmark model.[17] Both combination schemes provide strong gains in forecast accuracy across all variables and horizons. Notably, the optimal prediction pool outperforms the equal weighted average in almost each case, providing evidence that the increasing relevance of MF-models depicted by Figure 6 actually results in more precise forecasts. In the case of GDP growth, combining the forecast from MF and QF-VARs does not only (significantly) improve on the benchmark, but also on each individual model (see Table 2). At the one-year ahead horizon, the RMSE of the combined forecast is, on average 12% lower. For inflation, we find that combining the individual forecast does provide better results than those of the best performing models. In fact, the Q-TVP-SV-VAR provides a slightly better performance. Regarding the unemployment rate, combining forecasts from several VARs reduces the relative RMSEs with respect to the best performing individual model at each horizon with gains ranging from 4% to 8%.

6 Conclusion

Several studies show that modeling structural change improves forecast accuracy. We contribute to this discussion by investigating whether allowing for structural change in a mixed-frequency VAR setup further improves performance.

We conduct a rigorous real-time out-of-sample forecast experiment and generate predictions for GDP growth, CPI inflation, and the unemployment rate. Our findings show that modeling monthly dynamics substantially improves forecast accuracy. Nowcasts and short-term forecasts

[17] The results regarding the density forecasts are qualitatively identical, which is why we do not report them. See Table 6 in the Appendix C.

Table 4: Real-time forecast combination RMSEs

Combination scheme	1990-2017 h = 2	1990-2017 h = 3	1990-2017 h = 4	2008-2017 h = 2	2008-2017 h = 3	2008-2017 h = 4
	GDP growth					
Equal-weighting	0.95**	0.91***	**0.87*****	0.96	0.92**	**0.84*****
Optimal prediction pool	**0.93*****	0.88***	**0.88*****	**0.94**	**0.89***	0.85***
	Inflation					
Equal-weighting	0.86***	0.90***	0.88***	0.87***	0.91***	0.91***
Optimal prediction pool	**0.82*****	**0.84*****	**0.83*****	**0.84*****	**0.88*****	**0.90*****
	Unemployment rate					
Equal-weighting	0.81***	0.83***	0.84***	0.82	0.84*	0.85**
Optimal prediction pool	**0.75*****	**0.79*****	**0.81*****	**0.74**	**0.78***	**0.81****

Notes: RMSEs are reported as ratios to the benchmark. A ratio below unity indicates that the combination scheme outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, according to the Diebold-Mariano test with Newey-West standard errors.

especially benefit from within-quarter information, while for longer horizons, the advantages vanish in most cases. The MF-TVP-SV-VAR provides, on average, the best point and density forecast performance. Both inflation and unemployment rate forecast benefit considerably from modeling both monthly dynamics and structural change. With regard to inflation, the MF-TVP-SV-VAR nowcasts are slightly more precise than those from the SPF. We obtain rather mixed results for the GDP growth; no model dominates over all horizons, though almost all nonlinear MF-models outperform their linear counterpart as well as the remaining quarterly models. Furthermore, we assess the forecast performance during the Great Recession and demonstrate how the inflow of monthly information alters inflation forecasts. We show that the combination of time-varying parameters and stochastic volatility yields overall the best performance for the downturn and subsequent recovery. Finally, using optimal prediction pools, we reveal the increased importance of the MF-VARs, notably the MF-TVP-SV-VAR, with the onset of the Great Recession, confirming the growing relevancy of modeling intra-quarterly dynamics and structural change.

Our models are small-scale VARs due to the large number of parameters that have to be estimated and our variables are rather standard in the literature. However, in the light of the recent developments regarding the usage of larger dataset for TVP-SV-VARs (Chan, 2019; Kapetanios, Marcellino, and Venditti, 2019; Petrova, 2019), our results suggest that introducing mixed-frequencies in these estimation procedures might lead to strong gains in nowcast accuracy.

References

Aastveit, K. A., A. Carriero, T. E. Clark, and M. Marcellino (2017). Have Standard VARS Remained Stable Since the Crisis? *Journal of Applied Econometrics 32*(5), 931–951.

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2020). Choosing Prior Hyperparameters: With Applications To Time-Varying Parameter Models. *Journal of Business & Economic Statistics 38*(1).

Banbura, M. and A. van Vlodrop (2018). Forecasting with Bayesian Vector Autoregressions with Time Variation in the Mean. Tinbergen Institute Discussion Paper 2018-025/IV, Tinbergen Institute.

Barnett, A., H. Mumtaz, and K. Theodoridis (2014). Forecasting UK GDP growth and inflation under structural change. A comparison of models with time-varying parameters. *International Journal of Forecasting 30*(1), 129–143.

Brave, S. A., R. A. Butters, and A. Justiniano (2019). Forecasting economic activity with mixed frequency BVARs. *International Journal of Forecasting 35*(4), 1692–1707.

Carriero, A., T. E. Clark, and M. Marcellino (2015). Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society: Series A 178*(4), 837–862.

Carter, C. K. and R. Kohn (1994). On Gibbs Sampling for State Space Models. *Biometrika 81*(3), 541–553.

Chan, J. C. C. (2019). Large hybrid time-varying parameter VARs. CAMA Working Papers 2019–77, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

Chan, J. C. C. and E. Eisenstat (2017). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics 33*(4), 509–532.

Chan, J. C. C. and I. Jeliazkov (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation 1*(1-2), 101–120.

Chiu, C.-W. J., H. Mumtaz, and G. Pintér (2017). Forecasting with VAR models: Fat tails and stochastic volatility. *International Journal of Forecasting 33*(4), 1124–1143.

Clark, T. E. (2009). Is the Great Moderation over? An empirical analysis. *Economic Review* (Q IV), 5–42.

Clark, T. E. (2011). Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility. *Journal of Business & Economic Statistics 29*(3), 327–341.

Clark, T. E. and M. W. McCracken (2008). Forecasting with Small Macroeconomic VARs in the Presence of Instabilities. In D. E. Rapach and M. E. Wohar (Eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, pp. 93–147. Emerald Publishing.

Clark, T. E. and F. Ravazzolo (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics 30*(4), 551–575.

D'Agostino, A., L. Gambetti, and D. Giannone (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics 28*(1), 82–101.

Del Negro, M., R. B. Hasegawa, and F. Schorfheide (2016). Dynamic prediction pools: An investigation of financial frictions and forecasting performance. *Journal of Econometrics 192*(2), 391–405.

Del Negro, M. and G. E. Primiceri (2015). Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum. *Review of Economic Studies 82*(4), 1342–1345.

Diebold, F. X. and R. S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics 13*(3), 253–263.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Number 9780198523543 in OUP Catalogue. Oxford University Press.

Faust, J. and J. H. Wright (2009). Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset. *Journal of Business & Economic Statistics 27*(4), 468–479.

Faust, J. and J. H. Wright (2013). Forecasting Inflation. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, pp. 2–56. Elsevier.

Foroni, C., P. Guérin, and M. Marcellino (2015). Markov-switching mixed-frequency VAR models. *International Journal of Forecasting 31*(3), 692–711.

Foroni, C. and M. Marcellino (2014). A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates. *International Journal of Forecasting 30*(3), 554–568.

Garthwaite, P. H., Y. Fan, and S. A. Sisson (2016). Adaptive optimal scaling of Metropolis-Hastings algorithms using the Robbins–Monro process. *Communications in Statistics-Theory and Methods 45*(17), 5098–5111.

Geweke, J. and G. Amisano (2011). Optimal prediction pools. *Journal of Econometrics 164*(1), 130–141.

Ghysels, E. (2016). Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics 193*(2), 294–314.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The midas touch: Mixed data sampling regression models. CIRANO Working Papers 2004s-20, CIRANO.

Gneiting, T. and R. Ranjan (2011). Comparing density forecasts using threshold and quantile weighted scoring rules. *Journal of Business & Economic Statistics 29*(3), 411–422.

Götz, T. B. and K. Hauzenberger (2018). Large mixed-frequency VARs with a parsimonious time-varying parameter structure. Discussion Paper 40/2018, Deutsche Bundesbank.

Kapetanios, G., M. Marcellino, and F. Venditti (2019). Large time-varying parameter VARs: A nonparametric approach. *Journal of Applied Econometrics 34*(7).

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies 65*(3), 361–393.

Krüger, F., T. E. Clark, and F. Ravazzolo (2017). Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts. *Journal of Business & Economic Statistics 35*(3), 470–485.

Kuzin, V., M. Marcellino, and C. Schumacher (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting 27*(2), 529–542.

Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics 18*(4), 427–443.

Mariano, R. S. and Y. Murasawa (2010). A Coincident Index, Common Factors, and Monthly Real GDP. *Oxford Bulletin of Economics and Statistics 72*(1), 27–46.

McCracken, M. W., M. T. Owyang, and T. Sekhposyan (2015). Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR. Working Paper 2015-30, Federal Reserve Bank of St. Louis.

Ng, S. and J. H. Wright (2013). Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling. *Journal of Economic Literature 51*(4), 1120–1154.

Petrova, K. (2019). A quasi-Bayesian local likelihood approach to time varying parameter VAR models. *Journal of Econometrics 212*(1), 286–306.

Pettenuzzo, D. and A. Timmermann (2017). Forecasting Macroeconomic Variables Under Model Instability. *Journal of Business & Economic Statistics 35*(2), 183–201.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies 72*(3), 821–852.

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

Sims, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity 2002*(2), 1–40.

Sims, C. A. and T. Zha (2006). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review 96*(1), 54–81.

Wolters, M. H. (2015). Evaluating Point and Density Forecasts of DSGE Models. *Journal of Applied Econometrics 30*(1), 74–96.

Zadrozny, P. A. (1988). Gaussian-likelihood of continuous-time armax models when data are stocks and flows at different frequencies. *Econometric Theory 4*(1), 108–124.

Appendix

A Priors

For models with time-varying VAR coefficients, priors are based on a training sample, which consists of the first 8 years of the entire sample. In the following, variables denoted with OLS refer to OLS quantities based on the training sample. The length of the training sample is denoted by T_0.

AR-coefficients

For the benchmark VAR, we implement a diffuse Jeffrey's prior:

$$p(\beta, \Sigma) \propto |\Sigma|^{-(n+1)/2}. \tag{A.1}$$

For the nonlinear models, we use normal priors for the VAR-coefficients. To keep the models comparable with respect to the VAR coefficients, we choose an uniformative prior. In case of the Q-SV-VAR and the MF-SV-VAR, we employ the following prior:

$$p(\beta) \sim N\left(0, 1000 \times I_{k_\beta}\right). \tag{A.2}$$

For the TVP-models, we draw the VAR coefficients using the CK algorithm and initialize it with the following prior:

$$p(\beta_0) \sim N(0, 4 \times V(\hat{\beta}_{OLS})). \tag{A.3}$$

The prior for the covariance of the AR-coefficients ($Q = diag(q^2_{\beta_1}, \ldots, q^2_{\beta_{k_\beta}})$) follows an inverse-Wishart distribution:

$$p(Q) \sim IW(k_Q^2 \times T_0 \times V(\hat{\beta}_{OLS}), T_0). \tag{A.4}$$

Since we assume that Q is diagonal, this is equivalent to an inverse gamma prior for each element where k_Q is split into k_{Q_C} and $k_{Q_{AR}}$ for the intercepts and AR-coefficients, respectively.

Stochastic volatilities

The stochastic volatilities are drawn via the CK algorithm. Thus, additional priors for the diagonal elements of Σ_0 ($\log \sigma_0$), and the lower-triangular elements of A_0 ($a_{i,0}$) are required. We follow Primiceri (2005) in defining these prior distributions as:

$$p(\log \sigma_0) \sim N(\log \hat{\sigma}_{OLS}, I_n), \tag{A.5}$$

$$p(A_0) \sim N(\hat{A}_{OLS}, 4 \times V(\hat{A}_{OLS})). \tag{A.6}$$

The priors for the covariance of $\log \sigma_0$ and A_0 are inverse-Wishart distributed:

$$p(\Psi) \sim IW(k_\Psi^2 \times (1+n) \times I_n, 4), \tag{A.7}$$

$$p(\Phi_i) \sim IW(k_\Phi^2 \times (i+1) \times V(\hat{A}_{i,OLS}), i+1), \qquad i = 1, \ldots, k-1, \tag{A.8}$$

where i denotes the respective VAR-equation that has non-zero and non-one elements in the lower-triangular matrix A_t, i.e., for $n=4$ it is equation 2, 3, and 4.

Latent observations

The missing values of the quarterly series expressed at monthly frequency are replaced with an estimated latent state by applying a time-dependent CK algorithm. We initialize the unobserved state variable z_t with z_0 as actual observations from the monthly variables and constant values for the quarterly variables in levels from the last observations of our training sample:

$$p(z_0) \sim N(z_L, I_{np}). \tag{A.9}$$

Hence, $z_L = [\tilde{y}_0', \ldots, \tilde{y}_{0-p+1}']$ where $\tilde{y}_i$ contains actual values, if observed, and constant values in levels, thus zero growth rates, for missing observations.

Hyperparameters

The variability of β_t, a_t, and $\log \sigma_t$ depends on Q, Ψ, and Φ, respectively, and thus on the hyperparameters k_{Q_C}, $k_{Q_{AR}}$, k_Ψ, and k_Φ. Therefore, we follow Amir-Ahmadi et al. (2020) and use priors for those hyperparameters. Specifically, we employ an inverse gamma distribution with scale parameter and degrees of freedom equal to 0.1 and 2, respectively:

$$p(k_i) \sim IG(2, 0.1), \qquad i = Q_C, Q_{AR}, \Phi, \Psi. \tag{A.10}$$

This parameterization implies a loose prior with a mode of 0.05 and an infinite variance.

B Specification of the Gibbs sampler

To estimate the models we employ a Gibbs sampler that consecutively draws from the conditional distribution. In the following, the general form of the MCMC algorithm according to Del Negro and Primiceri (2015) is outlined. To include the estimation of the hyperparameters, an additional Metropolis Hastings step is added to the Gibbs sampler. Denoting any vector of variables x over the sample T by $x^T = [x_1', \ldots, x_T']'$, the Gibbs sampler takes the following form:

1. Initialize β_t, Σ^T, A^T, s^T, Q, Ψ, Φ, k_Q, k_Φ, and k_Ψ.
2. Draw $\tilde{y}^T$ from $p(\tilde{y}^T | y^T, \beta^T, Q, \Sigma^T, A^T, \Psi, \Phi)$.
3. Draw β^T from $p(\beta^T | \tilde{y}^T, Q, \Sigma^T, A^T, \Psi, \Phi)$.

4. Draw Q from $p(Q|\tilde{y}^T, \beta^T, \Sigma^T, A^T, \Psi, \Phi)$.

5. Draw A^T from $p(A^T|\tilde{y}^T, \beta^T, Q, \Sigma^T, \Psi, \Phi)$.

6. Draw Φ from $p(\Phi|\tilde{y}^T, \beta^T, Q, \Sigma^T, A^T, \Psi)$.

7. Draw Ψ from $p(\Psi|\tilde{y}^T, \beta^T, Q, \Sigma^T, A^T, \Phi)$.

8. Draw s^T from $\tilde{p}(s^T|\tilde{y}^T, \beta^T, Q, \Sigma^T, A^T, \Psi, \Phi)$.

9. Draw Σ^T from $\tilde{p}(\Sigma^T|\tilde{y}^T, \beta^T, Q, A^T, s^T, \Psi, \Phi)$.

10. Draw k_{Q_C} from $p(k_{Q_C}|Q_C) = p(Q_C|k_{Q_C})p(k_{Q_C})$.
    Draw $k_{Q_{AR}}$ from $p(k_{Q_{AR}}|Q_{AR}) = p(Q_{AR}|k_{Q_{AR}})p(k_{Q_{AR}})$.
    Draw k_Ψ from $p(k_\Psi|\Psi) = p(\Psi|k_\Psi)p(k_\Psi)$.
    Draw k_ϕ from $\prod_{i=1}^{k-1} p(k_\Phi|\Phi_i) = p(\Phi_i|k_\Phi)p(k_\Phi)$.

The second step of this Gibbs sampler refers to drawing the latent observations. Since there are no latent observations in the quarterly models, the Gibbs sampler omits Step 2 for these models. Steps 3 to 8 belong to the block of drawing the joint posterior of $\tilde{p}(\theta, s^T|\tilde{y}^T, \Sigma^T)$ by drawing θ from $p(\theta|\tilde{y}^T, \Sigma^T)$ where $\theta = [\beta^T, A^T, Q, \Phi, \Psi]$. Subsequently, we draw s^T from $\tilde{p}(s^T|\tilde{Y^T}, \Sigma^T, \theta)$, and then Σ_t from $\tilde{p}(\Sigma_t|s^T, \theta)$. $\tilde{p}$ denotes the draws based on the approximate likelihood due to the KSC step, while p refers to draws based on the true likelihood (for further detail, see Del Negro and Primiceri, 2015). In Step 10, we include the Metropolis-Hastings within the Gibbs sampler to draw the hyperparameters.

For ease of exposition, in the following we use $\tilde{y}^T$ to indicate the data used in each step of the algorithm. If one considers quarterly models, however, $\tilde{y}^T$ has to be replaced by y^T. We employ 50000 burn-in iterations of the Gibbs sampler for each model and use every 4th draw of 20000 after burn-in draws for posterior inference.

Step 2: Drawing latent states z_t

Let $z_T = [z_1, \ldots, z_T]$ denote the sequence of state vectors consisting of the unobserved monthly states. Draws for z_t are obtained by using the CK algorithm, i.e., we run the Kalman filter until T to obtain $z_{T|T}$ as well as $P_{T|T}$ and draw z_T from $N(z_{T|T}, P_{T|T})$. Subsequently, for $t = T-1, \ldots, 1$ we draw z_t from $N(z_{t|t}, P_{t|t})$ by recursively updating $z_{t|t}$ and $P_{t|t}$.

Step 3: Drawing the AR-coefficient β^T

Conditional on the drawn states or the actual data, sampling the AR-coefficients proceeds as in Step 2 using the CK algorithm. In order to decrease computation time and efficiently simulate the draws, we apply the precision sampling approach by Chan and Jeliazkov (2009).

Step 4: Drawing the covariance of the VAR-coefficients Q

The posterior of the covariance of VAR-coefficients is inverse-Wishart distributed with scale matrix $\overline{Q} = Q_0 + e_t' e$, $e_t = \Delta\beta_t'$, and degrees of freedom $df_Q = T + T_0$, where Q_0 and T_0 denote the prior scale for Q and prior degrees of freedom, respectively.

Step 5: Drawing the elements of A^T

To draw the elements of A_T, we follow Primiceri (2005) and rewrite the VAR in (6) as follows:

$$A_t(\tilde{y}_t - Z_t'\beta_t) = \tilde{y}_t^* = \Sigma_t u_t, \tag{A.11}$$

where, taking into account that β_T and $\tilde{y}_t$ are known, y_t^* is observable. Due to the lower-triangular structure of A_t^{-1}, this system can be written as a system of k equations:

$$\hat{\tilde{y}}_{1,t} = \sigma_{1,t} u_{1,t}, \tag{A.12}$$

$$\hat{\tilde{y}}_{i,t} = -\hat{y}_{[1,i-1]} a_{i,t} + \sigma_{i,t} u_{i,t}, \qquad i = 2, \ldots, k, \tag{A.13}$$

where $\hat{\tilde{y}}_{[1,i-1]} = [\hat{\tilde{y}}_{1,t}, \ldots, \hat{\tilde{y}}_{i-1,t}]$. $\sigma_{i,t}$ and $u_{i,t}$ refer to the i-th elements of σ_t and u_t. Thus, under the block diagonal assumption of Φ, the RHS of equation i does not include $\hat{y}_{i,t}$, implying that one can recursively obtain draws for $a_{i,t}$ by applying an otherwise ordinary CK algorithm equation-wise.

Step 6: Drawing the covariance Φ_i of the elements of A^T

Φ_i has an inverse-ishart posterior with scale matrix $\overline{\Phi}_i = \Phi_{0,i} + \epsilon_{i,t}'\epsilon_{i,t}$, $\epsilon_{i,t} = \Delta a_{i,t}'$, and degrees of freedom $df_{\Phi_i} = T + df_{\Phi_{i,0}}$ for $i = 1, \ldots, k$. $\Phi_{0,i}$, and $df_{\Phi_{i,0}}$ denote prior scale and prior degrees of freedom, respectively.

Step 7: Drawing the covariance Ψ of log-volatilities

As in Step 6, Ψ has an inverse-Wishart distributed posterior with scale matrix $\overline{\Psi} = \Psi_0 + \varepsilon_t'\varepsilon_t$, $\varepsilon_t = \Delta \log \sigma_t'^2$, and degrees of freedom $df_\Psi = T + df_{\Psi_0}$, where Ψ_0 and df_{Ψ_0} denote the prior scale and the prior degrees of freedom, respectively.

Step 8: Drawing the states of the mixture distribution s^T

Conditional on the volatilities, we independently draw a new value for the indicator matrix s^T from (see Kim et al., 1998):

$$PR(s_{i,t} = j | \tilde{y}^{**}, h_{i,t}) \propto q_j f_N(\tilde{y}^{**} | 2h_{i,t} + m_j - 1.2704, \nu_j^2). \tag{A.14}$$

Step 9: Drawing the volatilities

The elements of Σ_t are drawn using the KSC algorithm. To this end, we employ the VAR rewritten as in (A.11). Taking squares and logarithms, we get

$$\tilde{y}_t^{**} = 2h_t + \nu_t, \tag{A.15}$$

and for the volatility process:

$$h_t = h_{t-1} + \varepsilon_t, \tag{A.16}$$

where $\tilde{y}_{i,t}^{**} = \log((\tilde{y}_{i,t}^*)^2 + c)$, $\nu_{i,t} = \log u_{i,t}^2$, $h_{i,t} = \log \sigma_{i,t}$, and c is set to a small but positive number to increase the robustness of the estimation process. To transform this

non-Gaussian system (ν_t is distributed according to a χ^2-distribution with one degree of freedom) into a Gaussian system, we resort to Kim et al. (1998) and consider a mixture of seven normal densities with component probabilities q_j, means $m_j - 1.2704$, and variances ν_j^2. The values for $\{q_j, m_j, \nu_j^2\}$ are chosen to match the moments of the log $\chi^2(1)$ distribution are given by Table 5.

Table 5: Gaussian mixtures for approximating the log-$\chi^2(1)$

ω	q_j	m_j	ν_j^2
1	0.0073	-10.1300	5.7960
2	0.1056	-3.9728	2.6137
3	0.0000	-8.5669	5.1795
4	0.0440	2.7779	0.1674
5	0.3400	0.6194	0.6401
6	0.2457	1.7952	0.3402
7	0.2575	-1.0882	1.2626

Kim et al. (1998).

Step 10: Drawing the hyperparameters $k_{Q_C}, k_{Q_{AR}}, k_\Psi$, and k_Φ

The prior hyperparameters of the scale matrix of the variance covariance matrix Q, Ψ, and Φ are drawn with a Metropolis within Gibbs step. Amir-Ahmadi et al. (2020) show that the acceptance probability for each draw i can be simplified to:

$$\alpha_{k_X}^i = min\left(\frac{p(X|k_X^*)p(k_X^*)q(k_X^*|k_X^{i-1})}{p(X|k_X^{i-1})p(k_X^{i-1})q(k_X^{i-1}|k_X^*)}, 1\right), \tag{A.17}$$

where $X = \{Q_C, Q_{AR}, \Psi, \Phi\}$. Q_C and Q_{AR} refer to the diagonal elements of Q with respect to the intercepts and AR-coefficients, respectively. $p(X|k_X^*)$ denotes the prior distribution of X, while $p(k_X^*)$ indicates the prior for the hyperparameter. $q(k_X^*|k_X^{i-1})$ labels the proposal distribution. We apply a random walk chain algorithm:

$$k_X^* = k_X^{i-1} + \xi_t, \qquad \xi_t \sim N(0, \sigma_{k_X}^2). \tag{A.18}$$

The standard deviation σ_{k_X} is adjusted according to the method proposed by Garthwaite, Fan, and Sisson (2016):

$$\sigma_{k_X}^i = \sigma_{k_X}^{i-1} + c(\alpha^{i-1} - \alpha^*)/(i-1), \tag{A.19}$$

where $\alpha^* = 0.4$ is the target acceptance rate and $c = 1/[\alpha^*(1-\alpha^*)]$ is the optimal step size. We initialize k_X with the values used by Primiceri (2005), $k_Q = 0.01$, $k_\Psi = 0.1$, and $k_\Phi = 0.01$, and the standard deviation by $\sigma_{k_X} = 0.01$.

C Additional Results

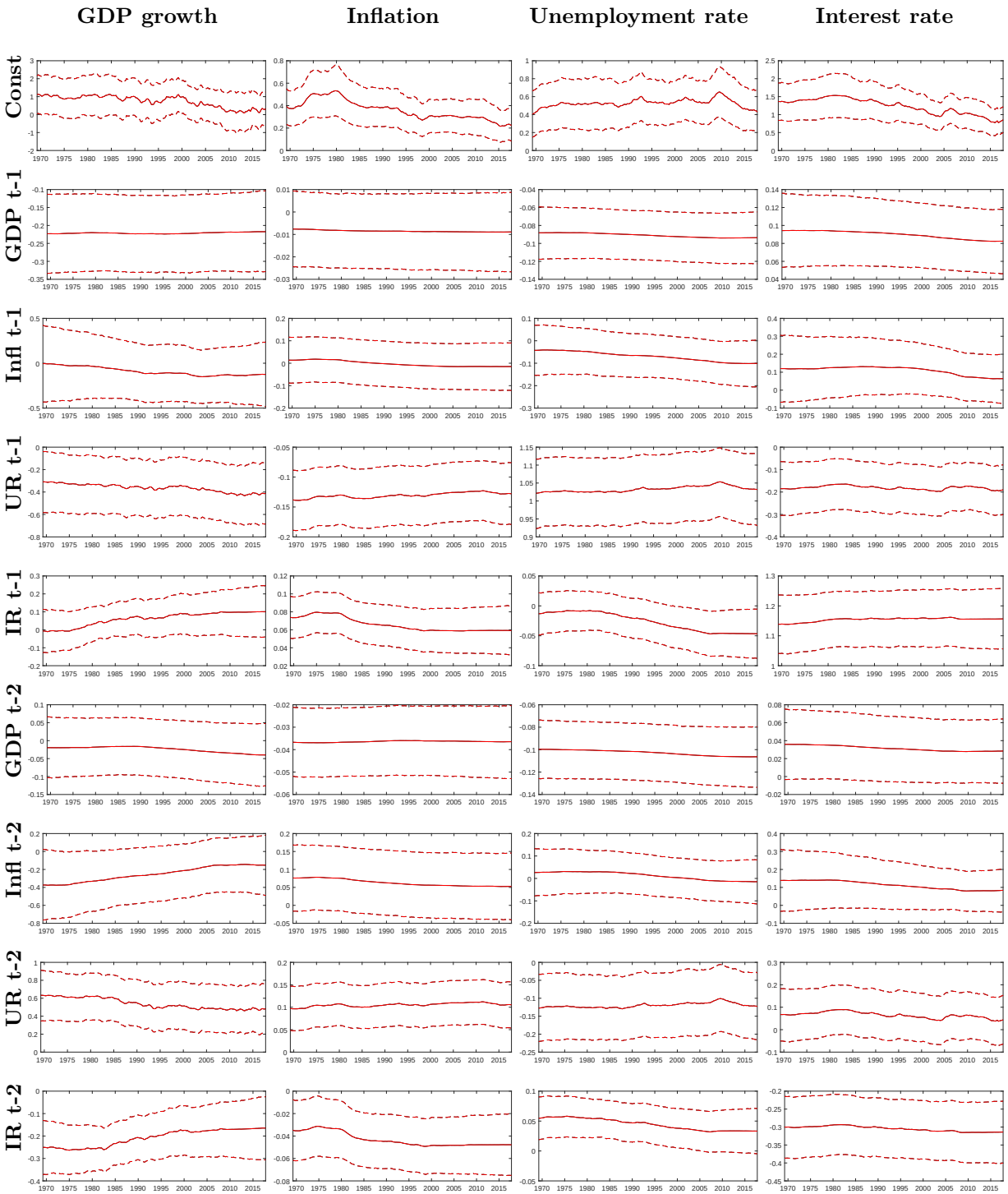
Table 6: Real-time forecast combination CRPS

Combination scheme	1990-2017 h = 2	1990-2017 h = 3	1990-2017 h = 4	2008-2017 h = 2	2008-2017 h = 3	2008-2017 h = 4
	GDP growth					
Equal-weighting	1.01	0.93*	0.87***	0.95	0.90*	**0.81*****
Optimal prediction pool	**0.93***	**0.85*****	**0.86*****	**0.93*****	**0.86*****	0.83***
	Inflation					
Equal-weighting	0.89***	0.89***	0.83***	0.92**	0.87**	**0.77*****
Optimal prediction pool	**0.81*****	**0.79*****	**0.74*****	**0.83*****	**0.85*****	0.84***
	Unemployment rate					
Equal-weighting	0.82***	0.81***	0.80***	0.86***	0.84***	0.83**
Optimal prediction pool	**0.76*****	**0.77*****	**0.76*****	**0.77*****	**0.77*****	**0.77*****

Notes: The scores are reported as ratios to the benchmark. A ratio below unity indicates that the combination scheme outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, according to a t-test on the average difference in scores relative to the benchmark model with Newey-West standard errors.

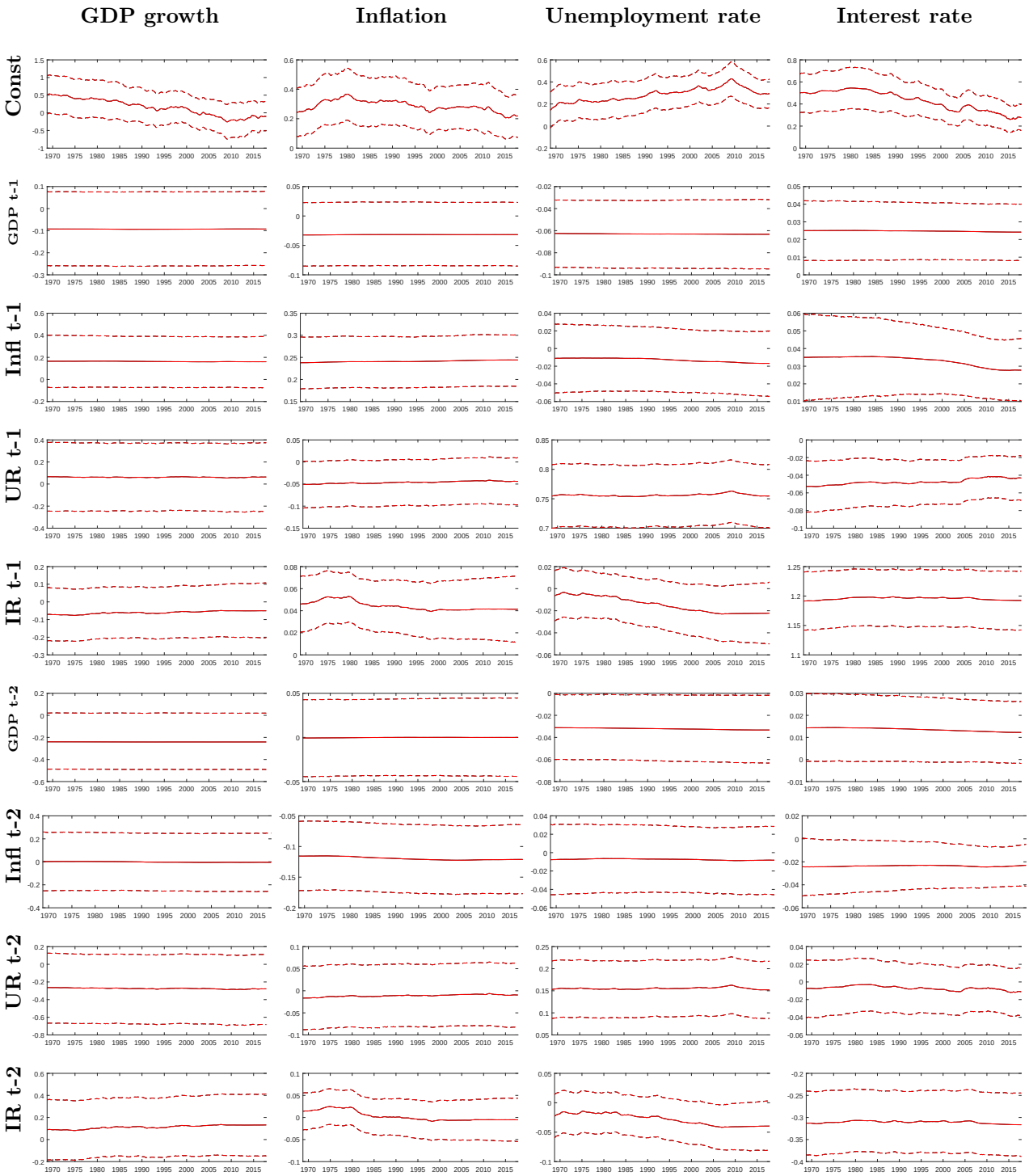
D Additional Figures

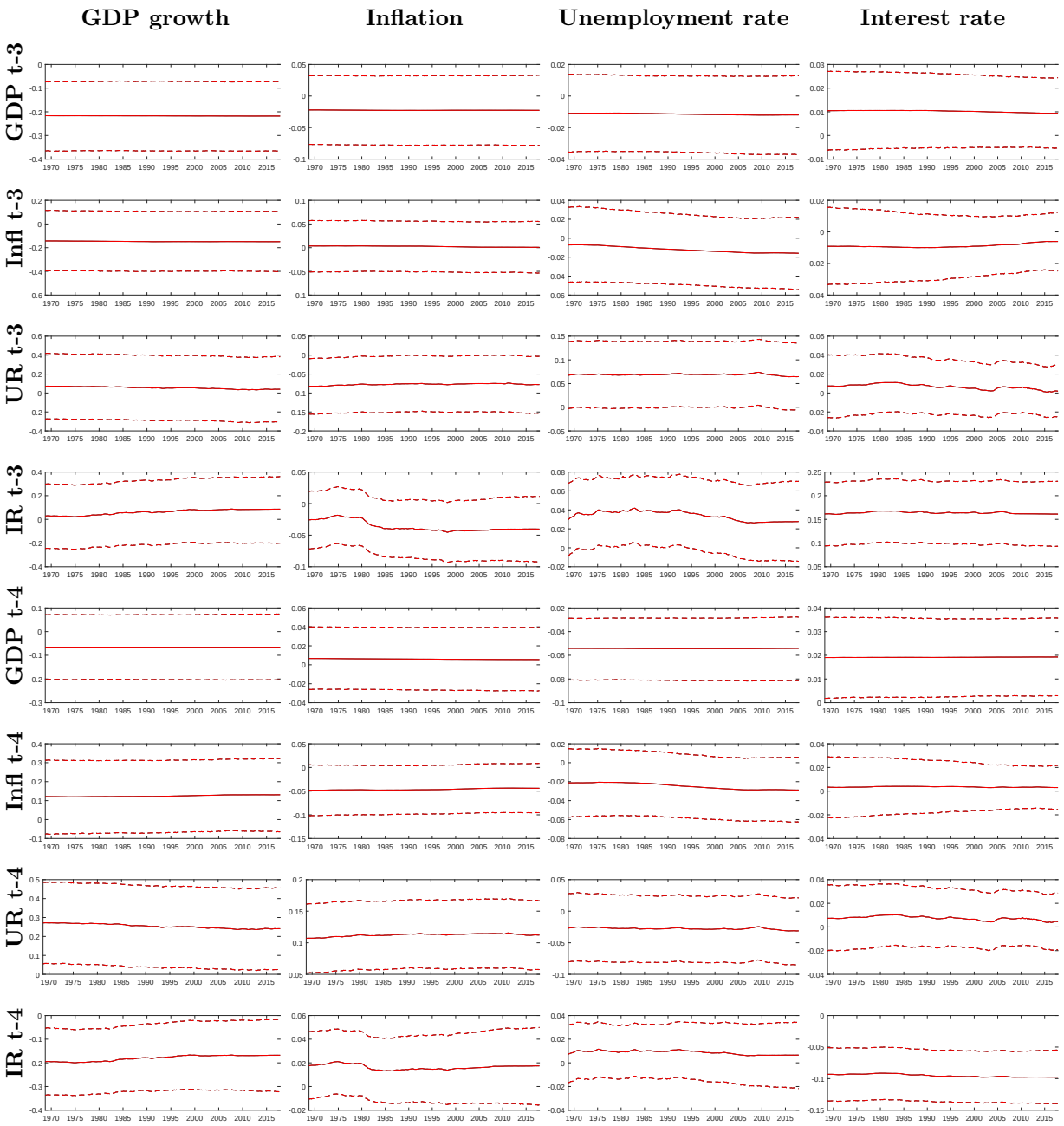
Figure 7: Time-Varying Parameters of the Q-TVP-SV-VAR



Notes: Figure depicts the time-varying parameters from the Q-TVP-SV-VAR. Columns refer to the variable and rows to the constant/lagged variable on which the variable is regressed. The dashed lines indicate 68% error bands. Results are based on the last data vintage

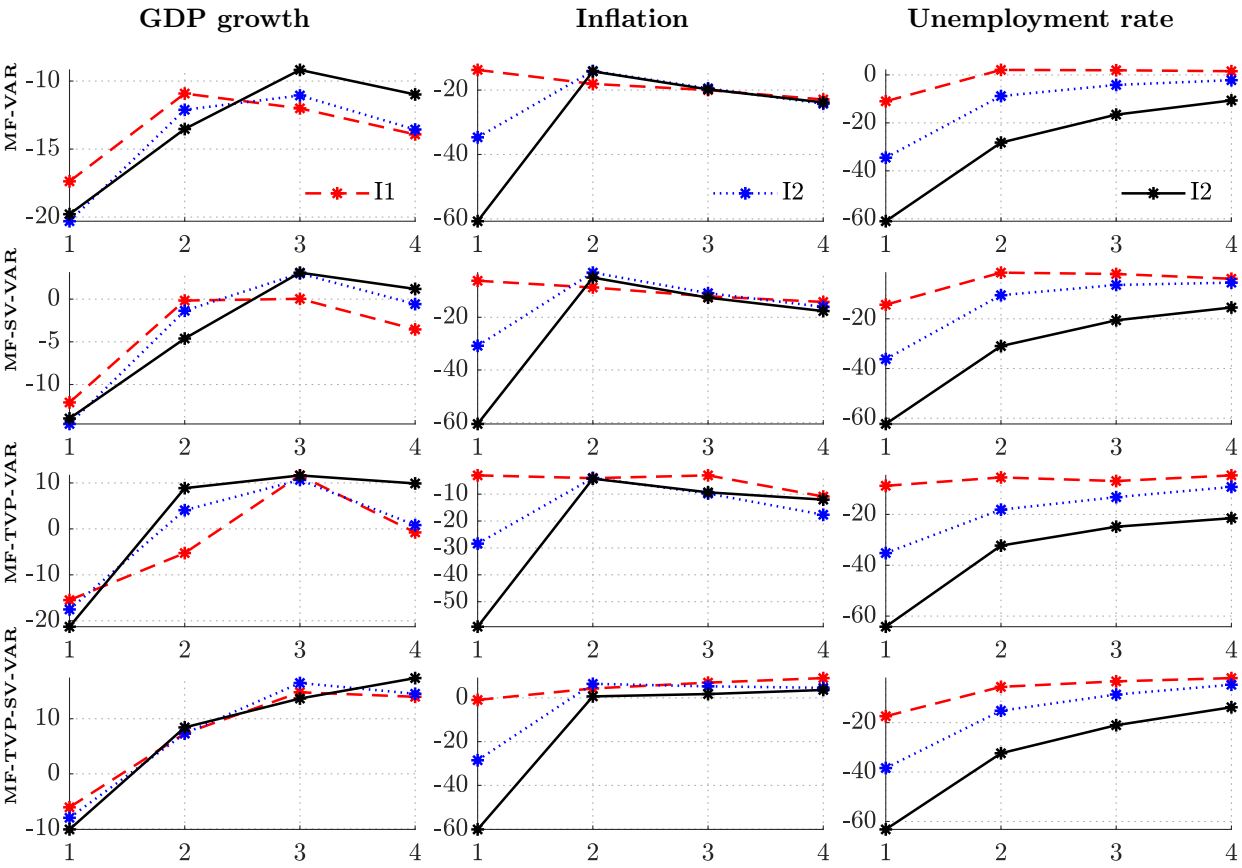
Figure 8: Time-Varying Parameters of the MF-TVP-SV-VAR





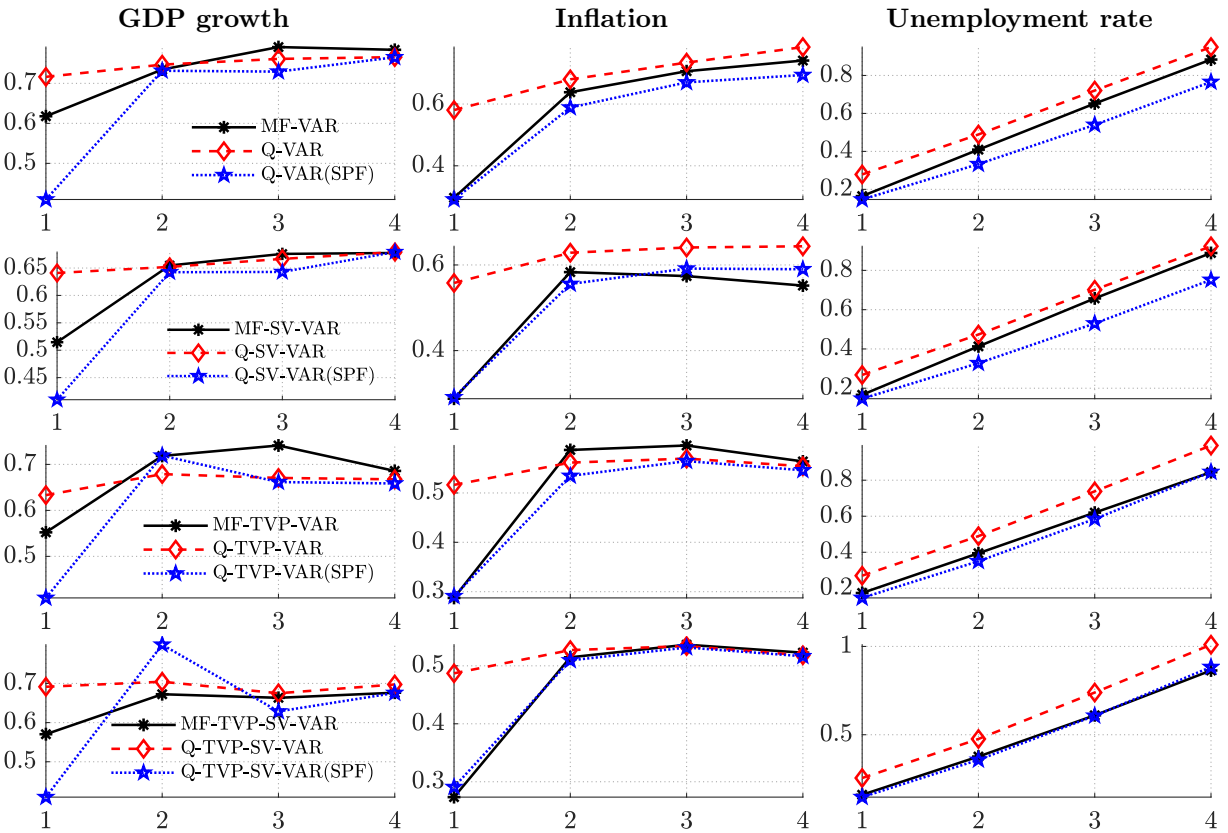
Notes: Figure depicts the time-varying parameters from the MF-TVP-SV-VAR. Columns refer to the variable and rows to the constant/lagged variable on which the variable is regressed. The dashed lines indicate 68% error bands. Results are based on the last data vintage.

Figure 9: Relative RMSEs



Notes: Figure depicts the relative RMSEs in terms of percentage gains compared to the benchmark model. Red, blue, and black lines refer to the information sets I1, I2, and I3 as outlined in Section 2.2, respectively. Sample: 1990–2017.

Figure 10: Comparison of model-based forecast with Survey of Professional Forecasters



Notes: Figure depicts the RMSEs of the mixed-frequency VARs (solid lines), the quarterly VARs (dashed lines), and the quarterly VARs conditional on the SPF nowcasts (dotted lines) for the the four horizons. To match the information set of SPF participants and models only forecasts from I2 are considered. Rows refer to models, columns refer to variables. Sample: 1990–2017.