Heinrich, Markus; Reif, Magnus

Working Paper

Forecasting using mixed-frequency VARs with time-varying parameters

ifo Working Paper, No. 273

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Heinrich, Markus; Reif, Magnus (2018) : Forecasting using mixed-frequency VARs with time-varying parameters, ifo Working Paper, No. 273, ifo Institute - Leibniz Institute for Economic Research at the University of Munich, Munich

This Version is available at:
https://hdl.handle.net/10419/191278

ifo WORKING PAPERS

273
2018

October 2018

Forecasting using mixed-frequency VARs with time-varying parameters

Markus Heinrich, Magnus Reif

ifo INSTITUTE

Leibniz Institute for Economic Research
at the University of Munich

Impressum:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.cesifo-group.de

An electronic version of the paper may be downloaded from the ifo website:
www.cesifo-group.de

ifo Working Paper No. 273

Forecasting using mixed-frequency VARs with time-varying parameters

Abstract

We extend the literature on economic forecasting by constructing a mixed-frequency time-varying parameter vector autoregression with stochastic volatility (MF-TVP-SV-VAR). The latter is able to cope with structural changes and can handle indicators sampled at different frequencies. We conduct a real-time forecast exercise to predict US key macroeconomic variables and compare the predictions of the MF-TVP-SV-VAR with several linear, nonlinear, mixed-frequency, and quarterly-frequency VARs. Our key finding is that the MF-TVPSV-VAR delivers very accurate forecasts and, on average, outperforms its competitors. In particular, inflation forecasts benefit from this new forecasting approach. Finally, we assess the models' performance during the Great Recession and find that the combination of stochastic volatility, time-varying parameters, and mixed-frequencies generates very precise inflation forecasts.

JEL code: C11, C53, C55, E32

Keywords: Time-varying parameters, forecasting, mixed-frequency models, Bayesian methods

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

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

This version: September 2, 2018

1 Introduction

Macroeconomists and in particular macroeconomic forecasters commonly face two major challenges. First, there are structural changes within an economy. Second, many economic time series are sampled at different frequencies and released with different publication lags. While several studies have shown that allowing for either structural changes or mixed-frequencies considerably improves forecast performance, Carriero, Clark and Marcellino (2013) assess a combination of both specifications in a univariate model. However, a multivariate assessment is still missing. Accordingly, the main contribution of this paper is to fill this gap and examine the real-time forecast performance of a model incorporating drifting coefficients and indicators observed at different frequencies. Our main finding is that this forecasting approach delivers accurate forecasts for the variables considered and in most cases significantly improves upon existing approaches, especially for inflation forecasts.

Our work relates to two strands of the literature. The first strand concerns the importance of modeling structural change in forecasting. To account for both changes in the comovements of variables demonstrated by Cogley and Sargent (2001, 2005) and the decline of business cycle volatility highlighted by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000), time-varying parameter VARs with stochastic volatility (TVP-SV-VAR) have become a frequently used tool.[1]

The second strand deals with the fact that many key macroeconomic variables, for instance GDP, are unavailable at frequencies higher than quarterly, while most key indicators for these variables are published at a higher frequency. As an alternative to models that require all variables to be sampled at the same frequency, in the recent past, mixed-frequency models have gathered much research interest (see Foroni and Marcellino, 2013, for s survey). This class of models has two advantages: First, the researcher can refrain from any kind of time (dis)aggregation to use for example quarterly and monthly variables in one model. Second, by jointly modelling high and low frequency variables, the researcher is better able to track the economic development in real time and assess the usefulness and impact of higher frequency information on the predictions.

This paper combines these two strands of literature by using a mixed-frequency TVP-SV-VAR (MF-TVP-SV-VAR) based on Cimadomo and D'Agostino (2016) to forecast in real time four US macroeconomic variables, namely GDP growth, CPI inflation, the unemployment rate and a short-term interest rate. As a combination of a MF-VAR and a TVP-SV-VAR, it can cope with indicators sampled at different frequencies and unbalanced datasets. To disentangle the relative impact of the model's mixed-frequency part and the time variation in the model's coefficients on the forecast accuracy, we compare the forecast performance of the MF-TVP-SV-VAR with several other specifications. The latter include constant parameter VARs with and without mixed-frequencies and time-varying VARs without a mixed-frequency part. Furthermore, we evaluate the intra-quarterly inflow of information with regard to the current-quarter estimates (nowcasts).

[1] See Galí and Gambetti (2009), Baumeister and Benati (2013) or Koop and Korobilis (2014) for examples of structural analysis using TVP-SV-VARs.

The estimation of the mixed-frequency part is based on the idea that variables observed at a lower frequency can be expressed as higher frequency variables with missing observations (Zadrozny, 1988).[2] 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.

The estimation of the TVP part basically follows Primiceri (2005). However, we treat those hyperparameters, which determine the amount of time variation in the parameters, as an additional layer and estimate them using Bayesian methods (Amir-Ahmadi, Matthes and Wang, 2018).[3] We generate forecasts up to one year ahead and evaluate the forecasts with respect to both point and density forecasts. For the point forecast evaluation we resort to root mean squared forecast errors, while the predictive densities are evaluated in terms of scoring rules.

Overall, our results provide evidence that the combination of mixed-frequencies, stochastic volatility and time-varying parameters provides very competitive point and density forecasts for each variable considered. We show that both nowcasts and forecasts significantly benefit from modeling intra-quarterly dynamics. In particular, the novel MF-TVP-SV-VAR generates, on average, more precise inflation forecasts than those generated by any other model considered. Using probability integral transforms, we compare the predictive densities of inflation forecasts generated by both the MF-TVP-SV-VAR and a quarterly TVP-SV-VAR and demonstrate that the former delivers an improved description of the data, especially in the short-run. In fact, the MF-TVP-SV-VAR provides a better estimate of the actual uncertainty surrounding the point estimate and is able to produce more forecasts corresponding to the upper percentiles of the empirical distribution. Finally, we examine the mixed-frequency models' inflation forecasts during the Great Recession. We show that allowing for time variation in the VAR coefficients and stochastic volatility leads to more precise predictions for the steep downturn and the subsequent recovery than considering only on of these specifications. Regarding the remaining variables, the results are mixed; for unemployment rate forecasts drifting coefficients are sufficient, for interest rate and GDP growth forecasts stochastic volatility yields precise forecasts.

On the one hand, this paper contributes to the ongoing discussion on how structural change affects forecast performance. D'Agostino, Gambetti and Giannone (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. A detailed assessment of the forecast performance of models with time-varying coefficients relative to a variety of other nonlinear and linear time series approaches is provided by both Barnett, Mumtaz and Theodoridis (2014) and Clark and Ravazzolo (2015). They underpin the findings of D'Agostino *et al.* (2013) and show that models with time-varying parameters increase the forecast performance in particular for inflation forecasts. Banbura and van Vlodrop (2018) illustrate that accounting for time-varying means in a Bayesian VAR substantially increases long-term forecast accuracy. Antolin-Diaz, Drechsel and Petrella (2017) provide evidence in favor of decline in long-run US output growth

[2] An alternative approach, called mixed data sampling (MIDAS), is provided by Ghysels, Santa-Clara and Valkanov (2004). For an assesment of this approach with regard to forecasting see Clements and Galvão (2008).

[3] Amir-Ahmadi *et al.* (2018) show that the magnitude of the hyperparameters significantly changes when estimated on monthly data compared to quarterly data, which affects the time-variation in the model's coefficients.

and demonstrate that modeling this decline in a DFM increases nowcast accuracy.

On the other hand, this article extends the literature on forecasting with mixed-frequency models. Since the work of Giannone, Reichlin and Small (2008), investigating the marginal impact of new information on nowcast accuracy, several studies have underpinned the benefits of modelling different frequencies with regard to forecasting.[4] The studies by Foroni, Guérin and Marcellino (2015), Barsoum and Stankiewicz (2015), and Bessec and Bouabdallah (2015) extend this literature by considering mixed-frequency models with discrete regime switches.

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, of which three are sampled at monthly frequency and one is observed quarterly. The quarterly series is US real GDP, whereas 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 to obtain real GDP growth rates and CPI inflation, respectively. The unemployment and interest rate remain untransformed. For the VARs estimated on quarterly frequency, the monthly indicators enter the models as quarterly averages, while for the mixed-frequency, we do not apply any further transformation. 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 2016. The first ten years are used as a training sample to specify priors. Thus, the actual model estimation starts in January 1970.

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 exhibits a publication lag of roughly one month, thus – for example – the first figure for 2011Q4 is released at the end of 2012M1 and then gets consecutively revised in the subsequent months 2012M2 and 2012M3. 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.

[4]For example, Clements and Galvão (2008), Wohlrabe (2009), Kuzin, Marcellino and Schumacher (2011), Foroni and Marcellino (2014), and Mikosch and Neuwirth (2015).

2.2 Forecast setup

In assessing the predictions we follow Schorfheide and Song (2015) and establish three different information sets. We assume that the forecasts are generated at the end of each month, when all current releases for the indicators are available. The first information set, called I1, relates to the first month of each quarter such that the forecaster has information up to the end 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 further observation on inflation and unemployment and the second GDP revision. Each information set is further augmented with the observations of the T-Bill rate.

To assess the intra-quarterly inflow of information, we evaluate the nowcasts separately per information set. Since the quarterly VARs, though, cannot cope with "ragged-edges" in the data, we estimate them in each recursion based on the balanced information set I1. The latter only accounts for new information in terms of data revisions.

We employ an expanding window to evaluate our forecasts for data vintages from January 1995 until September 2016, providing 261 estimation samples. The last one-year-ahead forecast refers to the third quarter 2017. The predictions are evaluated based on quarterly averages, implying that for the mixed-frequency approaches we time-aggregate the predicted monthly time paths to quarterly frequency. To abstract from benchmark revisions, definitional changes and other unforeseeable changes, we evaluate the GDP growth forecasts based on the second available estimate, i.e. the forecast for period $t+h$ is evaluated with the realization taken from the vintage published in $t+h+2$. 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, which implies that 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 upon 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 models exhibiting stochastic volatility, we use random walk stochastic volatility, which is a parsimonious and competitive specification (Clark and Ravazzolo, 2015). Throughout the paper, we use n as the number of variables which can be further split into n_q for quarterly and n_m for monthly variables, respectively, such that $n = n_q + n_m$. 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 intercepts, 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) refrains from the assumption of 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 and A_t has ones on the main diagonal and nonzero numbers for the remaining lower triangular elements, describing the contemporaneous relationships between the volatilities. This enables us to rewrite the VAR in (1) as

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

The law of motions 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:

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

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

with Ψ being diagonal and Φ being block diagonal where each block is related to each equation of the VAR in equation (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 augments the baseline Q-VAR with random walk processes 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 and β_t is the vectorized matrix of the VAR coefficients. For the Q-TVP-SV-VAR the stochastic volatility part from equations (4) and (5) is added to the model. Thus, the heteroscedastic model specification allows for changes in the magnitude of the shocks and for changes in the propagation of these shocks, whereas the homoscedastic version only accounts for the latter changes.

3.4 Mixed-frequency VAR

The 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 in a straightforward way. 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, which potentially contain missing observations due to "ragged-edges" in the dataset. $y_{q,t}$ denotes the quarterly variables at monthly frequency. Since the quarterly variables are only observed 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 adopt the notion of 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 (see Mitchell *et al.*, 2005). Note that $\log Y_{q,t}$ is only observed every third month, whereas $\log \tilde{Y}_{q,t}$ is never observed. To get an expression for quarterly GDP growth based on latent monthly GDP growth denoted by $\tilde{y}_{q,t}$, we subtract $\log Y_{q,t-3}$ from (8):

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

where lower case variables refer to logs. Combining the unobserved with the observed monthly variables in $\tilde{y}_t = [\tilde{y}'_{q,t}, y'_{m,t}]'$, we define the state vector as $z_t = [\tilde{y}'_t, \ldots, \tilde{y}'_{t-p+1}]$ 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. To

replace the missing observations in z_t with estimated states, we follow Durbin and Koopman (2001) and employ a time-depending vector of observables y_t and a time-varying matrix H_t. If an indicator exhibits a missing observation in period t, the corresponding entry in y_t and the corresponding row of H_t are deleted. Finally, 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 the singular variance-covariance matrix with all entries except for the upper left $n \times n$ submatrix being zero.

To introduce stochastic volatility into the mixed-frequency framework, we assume that S is a time dependent $pn \times pn$ matrix where the first $n \times n$ elements are equal to Ω_t with the same decomposition as in (2) and following the same law of motions as in (4) and (5). This yields the MF-SV-VAR. 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 for our forecast experiment:

1. MF-TVP-SV-VAR: Mixed-frequency VAR with time-varying parameters and stochastic 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, while the quarterly models are estimated with 2 lags.[5] 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 baseline Q-VAR we impose a Jeffrey's prior in order to abstract from shrinkage, since we use a small-scale VAR with only four variables. For the Q-SV-VAR we apply the algorithm of

[5]We fix the lag order for the quarterly model at 2 to be consistent with the literature on US data (see, e.g., Primiceri, 2005; D'Agostino *et al.*, 2013; Chan and Eisenstat, 2017). The monthly models have 4 lags to keep it computational feasible since each additional lag increases the number of parameters by $n \times n \times T$. Furthermore, at least four lags are required to disaggregate quarterly GDP into monthly GDP (see equation (9)).

Carter and Kohn (1994) (hereafter CK) to draw the VAR coefficients and the mixture sampler of Kim, Shephard and Chib (1998) (hereafter KSC) to draw the (log) volatilities.[6] We use normal priors for the diagonal elements of Σ_t and the lower-triangular elements of A_t. Inverse Wishart priors are applied for the variance covariance matrix Ψ and Φ, respectively. We adopt the CK algorithm for the Q-TVP-VAR with a normal prior for β_t and inverse Wishart prior for the variance-covariance matrix Q. The Q-TVP-SV-VAR combines both prior specifications of the Q-SV-VAR and Q-TVP-VAR and is estimated using the Gibbs sampler as in Del Negro and Primiceri (2015).

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, which are – in part – determined by the corresponding prior distributions. Hence, assigning sensible priors is crucial for the model estimation. The literature on TVP-SV-VAR commonly follows Primiceri (2005). However, these priors are calibrated for quarterly TVP-SV-VARs and it is not clear, whether they are useful considering monthly data or other model specifications. Thus, we follow Amir-Ahmadi *et al.* (2018) and abstract from using partly exogenous values for the scale matrix of the inverse Wishart prior by implementing another layer of priors for those hyperparameters.

The latent states in the mixed-frequency part of model 1 to 4 are estimated using a CK algorithm with the modification of Durbin and Koopman (2001), which enables us to cope with "ragged-edges" in the dataset and yields draws for each missing indicator until the end of the sample. We initialize the latent states of the CK algorithm with a normal prior based on monthly constant GDP values throughout the quarter for the mean and an identity matrix for the variance-covariance matrix. After having drawn the latent states, the implementation of the remaining specification is straightforward: Instead of conditional on the observed data, the remaining coefficients are drawn conditional on the drawn states. Each model is estimated using 60000 draws. For posterior inference we use each 5th draw from the last 10000 draws.

To provide an impression of the importance of modeling changes in volatility over time, Figure 1 plots the posterior means of standard deviations of reduced form residuals from the MF-TVP-SV-VAR and Q-TVP-SV-VAR using the latest vintage of data.[7] We assume that the volatility estimates from the Q-TVP-SV-VAR are constant within a quarter to make them comparable across frequencies. The estimates of the Q-TVP-SV-VAR closely match the patterns of previous analysis and show significant time-variation (see, for instance, Primiceri (2005), Clark and Ravazzolo (2015)). 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 roughly remain on the levels of the mid 1980s. At the end of the sample, however, there is again a decline in volatility indicating a time in which the US is remarkably less exposed to absolute shocks hitting

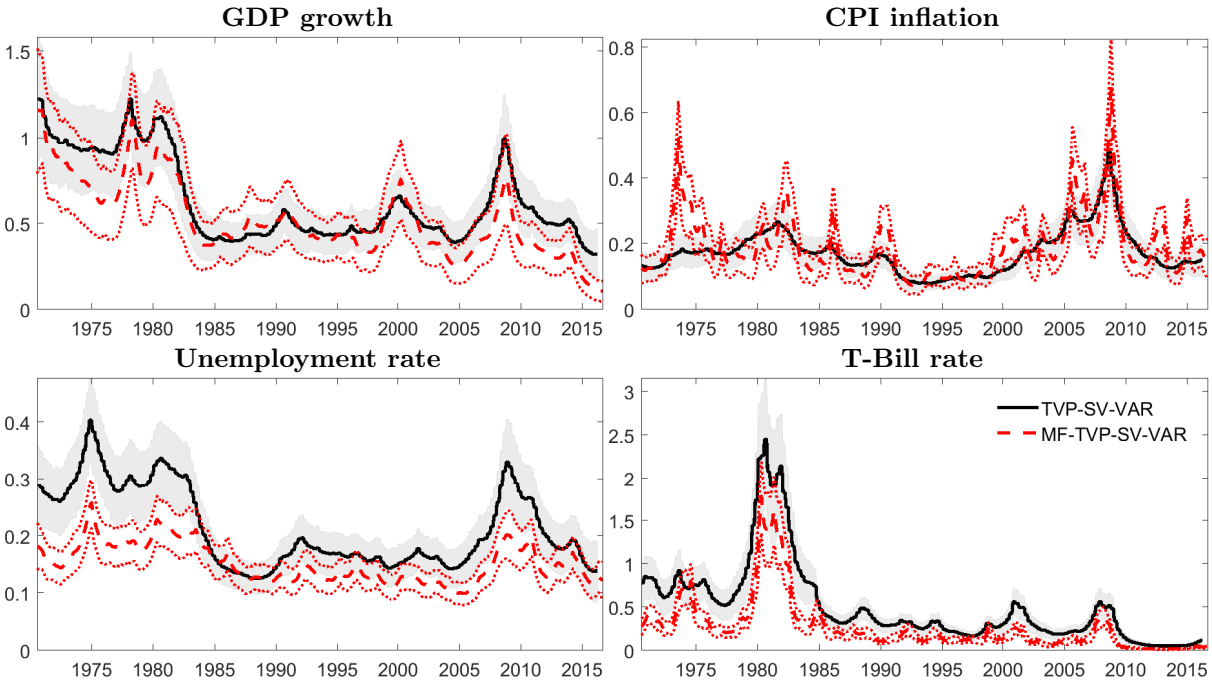
[6] Drawing the VAR coefficients using CK is equivalent to a GLS transformation of the model. Another possibility to draw the volatilities is the algorithm of Jacquier, Polson and Rossi (1995) which draws the volatilities one at a time. This single-move procedure is, however, computationally much less efficient and yields draws, which are more autocorrelated (see Kim *et al.*, 1998).

[7] We also examined the volatilites for different data vintages to investigate the impact of data revisions and different values for the hyperparameters. Analog to Clark (2012), though, we obtain very similar estimates for the different vintages, which is why we only report the results for the latest vintage.

the economy. Thus, as already suggested by Clark (2009) and Gadea Rivas, Gómez-Loscos and Pérez-Quirós (2014), the Great Recession seems to have only interrupted but not ended the Great Moderation. In fact, the latest volatility estimates for GDP growth and the unemployment rate are the lowest of the entire sample.

The estimates from the MF-TVP-SV-VAR closely track the evolution of the quarterly estimates. Except for CPI inflation they are, however, somewhat smaller indicating that using monthly information absorbs part of the fluctuations in the volatility. The latter is a confirmation of the results from Carriero, Clark and Marcellino (2015), who employ a Bayesian mixed-frequency model without time variation in the AR-coefficients. The change in the VAR parameters over time is in turn far less pronounced (see Figure 4 in Appendix D). For both models, the largest variability is obtained for the coefficients of the interest rate equation. For instance, the impact of past interest rates on the current interest rate has increased, while the impact of inflation on interest rates has dropped to almost zero, which reflects the binding zero lower bound. Overall, the results for both models suggest that modeling variability in volatility seems to be important for achieving precise forecasts.

Figure 1: Posterior Means of Standard Deviations of Reduced Form Residuals



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

3.6 Now- and forecasting

The quarterly models are estimated on balanced datasets containing all the available information from the previous quarter. To generate the now- and forecasts, we follow Cogley, Morozov and Sargent (2005) and 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 the vector of endogenous variable, y_t, h_{max} periods into the future. The first forecast is a nowcasts, since it is generated in and refers to the respective current quarter.

To describe how we obtain the predictive distributions for now- and forecasts from the mixed-frequency approaches, additional notation is useful. Let T_M denote the last month of the indicator, which has the shortest publication lag. Further, let $Z^{T_M} = [z_1, \ldots, z_{T_M}]$ denote the trajectory of simulated state vectors. Note that the CK algorithm provides draws for the latent states until T_M, which is why Z^{T_M} consists only of CK draws. To obtain $Z^{T_m+1:T_m+h_{max}}$, we again 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.[8]

4 Forecast metrics

The forecasts from our models are evaluated with respect to point and density forecasts. In the following, M, i, and h denote the model, variable, and forecast horizon, respectively, for the forecast sample $t = 1, \ldots, N$.[9] The point forecast accuracy is measured in terms of the root mean squared errors (RMSE):

$$\text{RMSE}_i^h = \sqrt{\frac{1}{N}\sum \left(\hat{y}_{t+h}^i - y_{t+h}^i\right)^2}. \tag{15}$$

The RMSE is though only a useful tool to assess the accuracy of a model when compared across different models, hence we report the RMSEs as ratios relative to a benchmark model

$$\text{relative RMSE}_{M,i}^h = \frac{\text{RMSE}_{M,i}^h}{\text{RMSE}_{B,i}^h}, \tag{16}$$

where $\text{RMSE}_{B,i}^h$ refers to the RMSE of the benchmark Q-SV-VAR estimated with quarterly data.[10] To provide a formal test of whether the difference in forecast accuracy are significant, we apply the test provided by Diebold and Mariano (1995).

While the evaluation of economic forecasts focused for a long time solely on evaluating point

[8] 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 indicators 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 of T_{M-1} to T_{M+1}.

[9] To facilitate readability, we drop subscript M indicating the respective model in the following in most cases.

[10] Since several studies have demonstrated that VARs with stochastic volatility outperform constant volatility VARs (see, for instance, Clark, 2012; Clark and Ravazzolo, 2015; Chiu, Mumtaz and Pintér, 2017), we abstract from using the Q-VAR as our benchmark.

forecasts, in the recent past the uncertainty around the forecasts has become an important issue. To take into account this uncertainty, i.e. the remaining part of the predictive density, which is neglected by the measures outlined above, we further evaluate the predictive densities. Since the true density is, however, not observed, evaluating the predictive density is less straightforward than evaluating point forecasts. The idea behind the evaluation of density forecasts is to compare the distribution of observed data with the predictive density and assess whether the latter provides a realistic picture of the reality. To this end, we apply both the log predictive scores and the continuous ranked probability score (CRPS). The log predictive score is computed by evaluating the predictive density at the realization.[11] In the following, we report average log scores:

$$\text{LS}_i^h = \frac{1}{N} \sum \log p_t(y_{t+h}^i), \tag{17}$$

where $p_t(\cdot)$ indicates the predictive density. According to (17), a higher average log score implies a more exact predictive density.[12] Again, we report results relative to the benchmark:

$$\text{relative LS}_{M,i}^h = LS_{M,i}^h - LS_{B,i}^h, \tag{18}$$

where $LS_{B,i}^h$ refers to the log score of the benchmark model. Furthermore, we evaluate the predictive densities in terms of the CRPS introduced by Matheson and Winkler (1976). As highlighted by for example Gneiting and Raftery (2007), CRPSs are both better able to evaluate forecasts close but not equal to the realization and less sensitive towards extreme outcomes. To compute the CRPS, we follow Gneiting and Ranjan (2011) and use the score function:

$$S(p_t, y_{t+h}^i, \nu(\alpha)) = \int_{-0}^{1} \text{QS}_\alpha(P_t(\alpha)^{-1}, y_{t+h}^i)\nu(\alpha)d\alpha, \tag{19}$$

where $\text{QS}_\alpha(P_t(\alpha)^{-1}, y_{t+h}^i) = 2(I\{y_{t+h}^i < P_t(\alpha)^{-1}\} - \alpha)(P_t(\alpha)^{-1} - y_{t+h}^i)$ is the quantile score for forecast quantile $P_t(\alpha)^{-1}$ at level $0 < \alpha < 1$. $I\{y_{t+h}^i < P_t(\alpha)^{-1}\}$ is an indicator function taking the value one if $y_{t+h}^i < P_t(\alpha)^{-1}$ and zero otherwise. P_t^{-1} denotes the inverse of the cumulative predictive density function and $\nu(\alpha)$ is a weighting function. Applying a uniform weighting scheme, yields the average CRPS:

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

Following (20), a lower value indicates a better score, which is evaluated as a ratio relative to

[11] The log predictive score goes back to Good (1952) and has become a commonly accepted tool for comparing the forecast performance of different models (see Geweke and Amisano, 2010; Clark, 2012; Jore, Mitchell and Vahey, 2010, among other).

[12] Since the predictive densities are not necessarily normal, the commonly used quadratic approximation of Adolfson, Lindé and Villani (2007) may not be appropriate, which is why we smooth the empirical forecast distribution using a kernel estimator to obtain the predictive distribution.

our benchmark model:

$$\text{relative CRPS}^h_{M,i} = \frac{CRPS^h_{M,i}}{CRPS^h_{B,i}} \tag{21}$$

To obtain an approximate inference on whether the scores are significantly different from the benchmark, we follow D'Agostino *et al.* (2013) and regress 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.

Finally, we compute probability integral transforms (PITs) as in Diebold, Gunther and Tay (1998). The PIT is defined as the CDF corresponding to the predictive density evaluated at the respective realizations

$$z^i_{t+h} = \int_{-\infty}^{y^i_{t+h}} p_t(u)du \quad \text{for } t = 1, \dots, N. \tag{22}$$

Thus, with regard to the respective predictive density, the PIT denotes the probability that a forecast is less or equal to the realization. If the predictive densities equal the true densities, z^i_{t+h} is uniformly distributed over the 0-1 interval. To assess the accuracy of the predictive density according to the PIT, it is convenient to divide the unit interval into k equally sized bins and count the number of PITs in each bin. If the predictive density equals the actual density, each bin contains N/k observations.

5 Results

In this Section, we discuss the results from our forecast experiment. Regarding the point forecasts, first we assess the nowcast accuracy of the models and resort to the information sets outlined in Section 2.2. Second, we evaluate the accuracy of the point forecasts and predictive densities with respect to horizons larger than 1, i.e. the respective subsequent quarters.[13] We provide results for the entire recursive sample (1995Q1–2016Q4) and for a shorter sample period of 2007Q1 until 2016Q4. The latter enables us to assess whether the anew rise in volatility during the Great Recession and the subsequent slow down affects the forecast performance.

5.1 Nowcast evaluation

Table 1 presents the results for the nowcasts taking into account the information sets I1 to I3. The results can be compared along five dimensions: Quarterly- vs. mixed-frequencies, fixed-coefficients vs. time-varying coefficients, across information sets as well as variables, and across samples. First, we discuss the results with respect to the full sample depicted in the left panel of Table 1.

[13] We abstract from evaluating the nowcasts with respect to predictive densities. The nowcasts of the mixed-frequency models consist, depending on the information set, 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, are hardly comparable.

Comparing the mixed-frequency models with the quarterly models reveals that the MF-models generate more accurate nowcasts for each variable and each information set. On average, over all information sets and variables the best nowcast performance is obtained by the MF-TVP-SV-VAR, which improves the benchmark (Q-SV-VAR) by roughly 42%. The remaining mixed-frequency models provide on average gains ranging from 38% (MF-VAR) to 40% (MF-SV-VAR) indicating that apart from using monthly information parameter variability is beneficial. Except for the Q-TVP-SV-VAR which provides roughly the same performance as the benchmark, all quarterly models deliver an inferior nowcast performance and thus provide – in line with the literature – strong support for mixed-frequency approaches.

Turning to the variables, we first look at GDP growth. In this case, the best performing model (MF-SV-VAR) provides up to 14% more accurate nowcasts compared to the benchmark. In contrast, the best performing quarterly model cannot outperform the benchmark. A similar pattern emerges for inflation; the best performance is again provided by a mixed-frequency model (MF-TVP-SV-VAR), which improves the benchmark by up to 60%, while the quarterly models enhance the benchmark by at best by 7%. Moreover, the MF-TVP-SV-VAR provides, on average, much more precise forecasts then the Q-TVP-SV-VAR, which is already found to deliver precise inflation forecasts (Faust and Wright, 2013). The largest difference between quarterly- and mixed-frequency models, is obtained for the unemployment rate. In this case the MF-TVP-VAR improves the benchmark by roughly 65%, while the quarterly models do not improve at all.

A comparison of the fixed-coefficients models with the time-varying coefficients models reveals that stochastic volatility seems to be a major determinant of precise nowcasts, which is consistent with for instance Carriero *et al.* (2015). In all but one case (unemployment rate at I1), the best performing model includes stochastic volatility. 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. In particular, inflation nowcasts benefit, however, from combining both specifications. For instance, the relative RMSE of the MF-TVP-SV-VAR is about 5 percentage points lower than the one of the MF-TVP.

Finally, we compare the RMSEs across information sets. In most cases using more information – as expected – significantly reduces the RMSEs. The improvements for the best performing models in case of inflation forecasts range from 13% at I1 to 61% at I3. With regard to the unemployment rate, the increases in forecast accuracy are even higher with 18% at I1 and 66% at I3. Concerning to GDP growth, though, more information seems not to provide a strong increase in forecast accuracy; the relative RMSEs for the best performing model (MF-TVP-SV-VAR) go from 0.89 at I1 to 0.86 at I3. The latter provides some evidence that the variables used are potentially not the best predictors for GDP growth and selecting the variables more carefully would further improve GDP growth forecasts. Since the goal of this paper is not to find the best GDP growth forecast, we leave this question for further research.

Overall, the results for the shorter sample are very close to the ones of the full sample. The right panel of Table 1 shows that the relative nowcast performance of each model remains almost unchanged for unemployment and interest rate. However, with regard to GDP growth and inflation, the MF-models' relative performance improves in the shorter sample, which is

Table 1: Real-time Nowcast RMSEs

Model	1995-2016			2008-2016		
	I1	**I2**	**I3**	**I1**	**I2**	**I3**
	GDP growth					
MF-TVP-SV-VAR	0.91	0.90	0.91	0.86	0.84	0.87
MF-SV-VAR	**0.89**	**0.86**	**0.86**	**0.81**	**0.77**	**0.77**
MF-TVP-VAR	0.98	0.96	0.88	0.95	0.94	0.82
MF-VAR	0.99	0.93	0.91	0.98	0.88	0.90
Q-TVP-SV-VAR	1.06	1.09	1.07	1.04	1.08	1.08
Q-TVP-VAR	1.07	1.11	1.03	1.08	1.10	0.99
Q-VAR	1.12***	1.13***	1.13***	1.13	1.14***	1.14***
Q-SV-VAR	0.63	0.62	0.63	0.80	0.78	0.78
	Inflation					
MF-TVP-SV-VAR	**0.87***	**0.64****	**0.39****	**0.85**	**0.59**	0.31
MF-SV-VAR	0.99	0.68*	**0.39****	0.93	0.61	**0.30**
MF-TVP-VAR	0.92*	0.69**	0.40**	0.92	0.67	0.36
MF-VAR	0.99	0.69*	**0.39****	0.95	0.62	**0.30**
Q-TVP-SV-VAR	0.94	0.93	0.94	0.94	0.92	0.93
Q-TVP-VAR	0.93	0.95	0.98	0.98	0.96	0.99
Q-VAR	1.03	1.02	1.13	1.02	1.01	1.00
Q-SV-VAR	0.26	0.26	0.26	0.34	0.33	0.34
	Unemployment rate					
MF-TVP-SV-VAR	0.82***	0.61***	**0.34*****	**0.78***	**0.60****	0.34**
MF-SV-VAR	0.83***	0.61***	**0.34*****	0.85***	**0.60****	**0.32****
MF-TVP-VAR	**0.79*****	**0.60*****	**0.34*****	**0.78****	0.61*	0.34*
MF-VAR	0.86***	0.61***	**0.34****	0.84**	**0.60****	**0.32****
Q-TVP-SV-VAR	1.02	1.08	1.04	1.05	1.07	1.07
Q-TVP-VAR	1.07	1.04	1.03	1.09	1.10	1.04
Q-VAR	1.05***	1.07***	1.07***	1.06	1.07**	1.07**
Q-SV-VAR	0.27	0.26	0.26	0.36	0.35	0.35
	Interest rate					
MF-TVP-SV-VAR	**0.43*****	0.16***	–	**0.36***	**0.15***	–
MF-SV-VAR	0.44***	**0.15*****	–	**0.36***	**0.15***	–
MF-TVP-VAR	0.45***	0.17***	–	0.41*	0.16*	–
MF-VAR	0.59***	0.19***	–	0.57	0.18*	–
Q-TVP-SV-VAR	0.96	0.94	0.94	0.87**	0.86*	0.86*
Q-TVP-VAR	1.13**	1.15*	1.14*	1.09	1.07	1.10
Q-VAR	1.48***	1.46***	1.47*	1.56	1.52*	1.51*
Q-SV-VAR	0.36	0.36	0.36	0.37	0.38	0.39

Notes: The models are detailed in Section 3. RMSEs are reported in absolute terms for the benchmark model (the bottom row of each panel). For the remaining the RMSEs are expressed as ratios relative to the benchmark model. A figure below unity indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and information set. $*$, $**$, and $***$ denote significance on the 15%, 10%, and 5% level, respectively, according to the Diebold-Mariano test with Newey-West standard errors. At I3 no interest rate forecast is computed by the mixed-frequency models, since the entire quarter is available.

characterized by a larger volatility of the series. The strongest gains are obtained for the MF-SV-VAR. Its GDP growth and inflation forecasts are roughly 8% more precise. Therefore and in contrast to the entire sample, the best performance in the shorter sample on average is provided

by the MF-SV-VAR, suggesting that stochastic volatility has become more important for precise nowcasts.

5.2 Forecast evaluation

Now we turn to investigate the forecast performance. Since the marginal impact of an additional month of information becomes less important for forecasts at higher horizons, the RMSEs for higher horizons become more similar across the information sets.[14] Therefore, in the following we compute total RMSEs by averaging over the entire forecast sample.

The results from Table 2 indicate that mixed-frequency VARs provide competitive forecasts even for higher horizons and for both samples. Indeed, in case of unemployment and interest rate forecasts, modeling within-quarter dynamics is particularly beneficial, since even the worst performing mixed-frequency VAR outperforms the best performing quarterly VAR on each horizon. In the following, we focus on the results for the full sample.

Overall, the most accurate forecasts for all indicators and on average is again provided by the MF-TVP-SV-VAR. It outperforms the baseline Q-VAR over all horizons and variables on average by roughly 12%. The best performance is obtained for the interest rate with an average gain of roughly 30% relative to the benchmark. Apart from the latter model, the MF-SV-VAR and the MF-TVP-VAR also provide very accurate forecasts with average gains of around 10%. The MF-VAR yields roughly the same performance as the best performing quarterly model, namely the Q-TVP-SV-VAR. Both improve the forecast on average over all horizons and variables by about 2%.

Concerning the variables individually shows that the gains in forecast accuracy differ substantially across the models. However, quarterly models are only able to outperform the mixed-frequency models for GDP growth. Nevertheless, all of the RMSEs – except for the MF-VAR – are in this case quite close to each other. For inflation, the best performance over all horizons is provided by the MF-TVP-SV-VAR. In particular, for one-quarter-ahead forecasts it generates by far the most accurate predictions. On larger horizons the Q-TVP-SV-VAR delivers virtually identical RMSEs, which indicates both that using monthly information become less important for higher horizons and that using time variation in all coefficients is crucial for inflation forecasts. The latter confirms the results of previous studies based on quarterly models (see D'Agostino *et al.*, 2013; Barnett *et al.*, 2014; Clark and Ravazzolo, 2015) by use of mixed-frequency models. As for the nowcasts, the most accurate unemployment rate forecasts are obtained by the MF-TVP-VAR. The differences to the MF-TVP-SV-VAR and MF-SV-VAR are, though, small. Using only quarterly data in turn provides significantly inferior RMSEs. The largest differences between the quarterly and the mixed-frequency models are obtained for the interest rate, where the RMSE of the best performing mixed-frequency model is roughly one third of the size of the RMSE of the best performing quarterly model (0.69 vs. 0.96).

Comparing the results across samples reveals that the relative RMSEs are very similar for each variable and model suggesting that the sample has only minor influence on the results. However, while the models which incorporating stochastic volatility improve their relative forecast accuracy

[14] Figure 6 in Appendix E plots the relative RMSE for each information set.

in the shorter sample, the performance of the models without this feature tends to deteriorate. The latter indicates that in this more volatile phase, stochastic volatility is more important for achieving precise forecasts. Moreover and in contrast to the nowcast evaluation, the best performance on average is provided by the MF-TVP-SV-VAR, which improves the benchmark by 17% and faintly outperforms the MF-SV-VAR.

Overall, the results are consistent with findings from previous studies, suggesting 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 more accurate forecasts for all variables in most cases.

5.3 Predictive density evaluation

The results for continuous rank probability scores (CRPS) are displayed in Table 3. The benchmark is reported in levels, while for the remaining models the scores are reported as ratios relative to the benchmark. We focus on CRPS in more detail since it is more sensitive to distance and less sensitive to outliers than the log scores.[15]

We draw three main conclusions from the results. First, the sample period has only minor impact on the relative accuracy of the predictive densities. In fact, the CRPS are overall very similar, which is why we do not discuss the results for both samples separately but cover them jointly.

Second, using within-quarter information significantly improves predictive densities; the mixed-frequency models provide better results on average over all variables and horizons. The best performance is again provided by the MF-TVP-SV-VAR with an average improvement of roughly 13%. In contrast, the best performing quarterly model (Q-TVP-SV-VAR) improves the benchmark on average only by 2%. Moreover, with regard to the unemployment rate, the worst performing mixed-frequency model (MF-VAR) improves the benchmark on average by 7%, and thus is better than each quarterly model, indicating that mixed-frequency is an important feature for unemployment forecasts. For the interest rate, we obtain a similar picture; apart from the Q-TVP-SV-VAR none of the quarterly models outperform the mixed-frequency models. For GDP growth and inflation, the results are less obvious. The most accurate GDP growth forecasts over all horizons are provided by the Q-TVP-SV-VAR, though the differences to its mixed-frequency counterpart are very small. Regarding inflation, only the mixed-frequency models with time-varying parameters outperform their quarterly counterparts. Investigating the forecast performance across the different horizons shows that the differences between the quarterly and mixed-frequency become smaller with increasing horizons. Thus, within-quarter information is more valuable with respect to short-term forecasting.

Third, concerning the variables again individually reveals that models using stochastic volatility and/or time-varying VAR-coefficients generate in most cases more accurate predictive densities than models without these features. As for the point forecast performance, the best performance for the interest rate forecasts is provided by the MF-SV-VAR indicating that variation in

[15] In general, the predictive distributions of the MF-models have a lower variance than the ones of the quarterly benchmark. Therefore, outlier realizations get a very low log score in case of MF-models, which distorts the overall results. However, as depicted by Table 4 in Appendix C both measures provide qualitatively similar results.

Table 2: Real-time Forecast RMSEs

Model	1995-2016 h = 2	1995-2016 h = 3	1995-2016 h = 4	2008-2016 h = 2	2008-2016 h = 3	2008-2016 h = 4
	GDP growth					
MF-TVP-SV-VAR	1.07	1.03	1.00	1.08	1.00	0.99
MF-SV-VAR	1.00	1.00	0.98	0.98	1.00	0.99
MF-TVP-VAR	**0.99**	0.99	1.01	**0.97**	**0.95**	1.00
MF-VAR	1.12***	1.14***	1.12***	1.11***	1.16***	1.15***
Q-TVP-SV-VAR	1.03	**0.98**	**0.96**	1.00	0.96	**0.94**
Q-TVP-VAR	1.17***	1.08**	1.07	1.17***	1.04	1.09
Q-VAR	1.17***	1.16***	1.15***	1.23***	1.19***	1.20***
Q-SV-VAR	0.63	0.64	0.66	0.82	0.85	0.85
	Inflation					
MF-TVP-SV-VAR	**0.87****	**0.91*****	**0.93*****	**0.85****	**0.92*****	**0.93*****
MF-SV-VAR	0.98	0.96	0.94	0.97	0.96	0.93
MF-TVP-VAR	0.93	0.95	0.94	0.93	0.97	0.96
MF-VAR	1.01	1.03	1.05	0.98	0.99	0.97
Q-TVP-SV-VAR	0.93**	0.93***	0.94***	0.92**	0.92***	0.94***
Q-TVP-VAR	0.93	0.95*	0.97	0.91	0.94	0.98
Q-VAR	1.00	1.07***	1.09***	0.99	1.06**	1.06
Q-SV-VAR	0.26	0.24	0.23	0.34	0.30	0.29
	Unemployment rate					
MF-TVP-SV-VAR	0.86*	0.93	0.98	0.86***	0.94***	0.98***
MF-SV-VAR	0.86**	0.91***	0.95	0.85***	0.91**	0.94
MF-TVP-VAR	**0.83****	**0.87****	**0.91**	**0.82****	**0.88****	**0.90***
MF-VAR	0.88**	0.94	0.94	0.85**	0.92	0.96
Q-TVP-SV-VAR	1.04	1.04	1.04	1.05	1.04	1.04
Q-TVP-VAR	1.02	1.01	0.98	1.12**	1.10***	1.05**
Q-VAR	1.07***	1.08***	1.08***	1.07**	1.07**	1.08**
Q-SV-VAR	0.48	0.73	0.97	0.68	1.04	1.39
	Interest rate					
MF-TVP-SV-VAR	0.59***	0.76***	0.84***	0.47***	0.59***	0.62***
MF-SV-VAR	**0.57*****	**0.71*****	**0.78*****	**0.44*****	**0.53*****	**0.57*****
MF-TVP-VAR	0.62***	0.81***	0.91***	0.56***	0.79***	0.89**
MF-VAR	0.72***	0.87***	0.94**	0.71***	0.90***	0.97**
Q-TVP-SV-VAR	0.95	0.96	0.97	0.81**	0.77***	0.81***
Q-TVP-VAR	1.14**	1.19***	1.17***	1.20**	1.33***	1.29***
Q-VAR	1.15***	1.16***	1.17***	1.34***	1.45***	1.48***
Q-SV-VAR	0.67	0.96	1.24	0.64	0.85	1.10

Notes: The models are detailed in Section 3. RMSEs are reported in absolute terms for the benchmark model (the bottom row of each panel). For the remaining the RMSEs are expressed as ratios relative to the benchmark model. A figure below unity indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. *, **, and *** denote significance on the 15%, 10%, and 5% level, respectively, according to the Diebold-Mariano test with Newey-West standard errors.

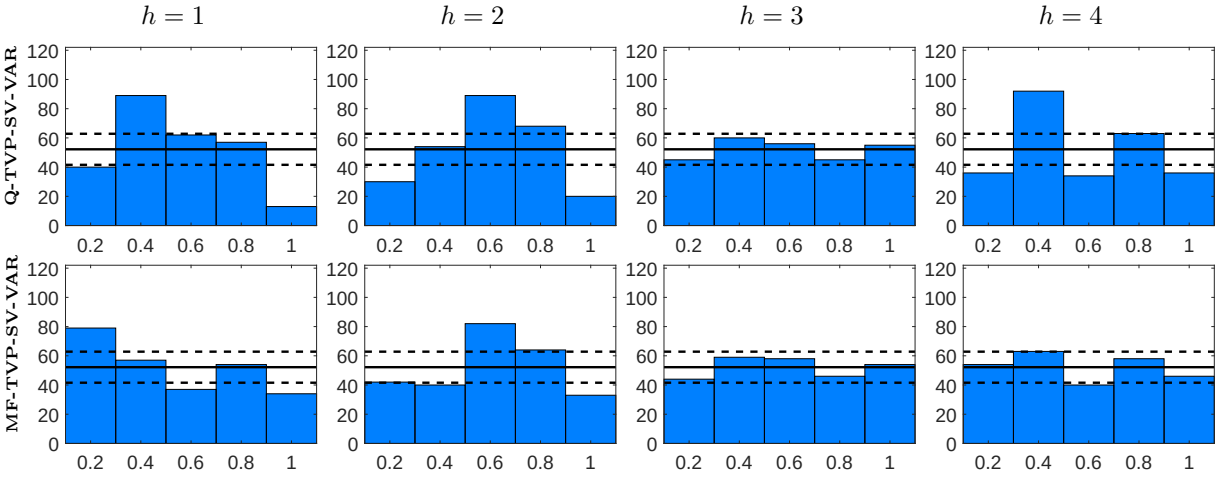
the VAR-coefficients is only a minor issue in this case. The results for the unemployment rate and inflation are similar to the point forecast performance. In case of the unemployment rate, the MF-SV-VAR and the MF-TVP-VAR provide virtually identical performances indicating that one can use either SV-models or TVP-models or both features. For inflation, however, it is cru-

cial to combine stochastic volatility and time-varying VAR-coefficients to get precise predictive densities.

In summary, the results of the predictive density evaluation support the findings from the point forecast evaluation. Using mixed-frequency is beneficial irrespective of time-variation in parameters, stochastic volatility, variables, and forecast horizons. As shown by several studies using quarterly models, stochastic volatility significantly improves predictive densities (Carriero *et al.*, 2015; Carriero, Clark and Marcellino, 2016; Chiu *et al.*, 2017; Huber, 2016, among other). We are able to confirm this finding using mixed-frequency models. Moreover, we underpin the results of D'Agostino *et al.* (2013) and Clark and Ravazzolo (2015) by demonstrating that combining stochastic volatility and time-varying parameters in general improves the accuracy of predictive densities using both quarterly and mixed-frequency models.

To get a better picture of the predictive densities' calibration, we compute probability integral transforms (PITs). For the sake of brevity, Figure 2 only presents the inflation predictions of the Q-TVP-SV-VAR (upper panel) and MF-TVP-SV-VAR (bottom panel).[16] To ensure uniformity, each bin in Figure 12 should contain 20% of the forecasts. The most apparent difference between both models' histograms is that the MF-TVP-SV-VAR is much better in capturing the right tail of the distribution than the Q-TVP-SV-VAR, in particular at short horizons. Moreover, the histograms of the Q-TVP-SV-VAR are hump-shaped for $h = 1$ and $h = 2$, indicating that the predictive densities are too wide and the uncertainty around the point estimate is overestimated. This pattern is less pronounced for the MF-TVP-SV-VAR with histograms closer to uniformity. In summary, our results indicate that omitting within-quarter dynamics and computing quarterly averages from monthly variables neglects valuable information , which in most cases significantly improves forecast accuracy.

Figure 2: Probability Integral Transforms for Inflation Forecasts



Note: The rows refer to the PITs of the Q-TVP-SV-VAR and MF-TVP-SV-VAR, respectively. The columns refer to the forecast horizons. The solid line indicates uniformity and the dashed lines 90% confidence bands as in Rossi and Sekhposyan (2014). Sample: 1995-2016.

[16]The PITs for the remaining variables, models, and horizons are presented in Appendix G, showing a similiar picture and further supports the good performance of the MF-TVP-SV-VAR in terms of predictive density calibration.

Table 3: Real-time Forecast CRPS

Model	1995-2016 h = 2	1995-2016 h = 3	1995-2016 h = 4	2008-2016 h = 2	2008-2016 h = 3	2008-2016 h = 4
	GDP growth					
MF-TVP-SV-VAR	1.05	1.00	**0.97**	1.03	**0.93**	**0.93***
MF-SV-VAR	1.04	1.01	0.98	1.00	1.00	0.98
MF-TVP-VAR	**0.99**	1.05*	1.06***	**0.94***	0.96	1.03
MF-VAR	1.13***	1.17***	1.13***	1.15***	1.21***	1.19***
Q-TVP-SV-VAR	1.02	**0.98**	0.98	0.99	0.95	0.95
Q-TVP-VAR	1.15***	1.15***	1.14***	1.12***	1.08	1.11**
Q-VAR	1.19***	1.21***	1.18***	1.27***	1.24***	1.25***
Q-SV-VAR	0.35	0.34	0.36	0.44	0.46	0.46
	Inflation					
MF-TVP-SV-VAR	**0.90*****	**0.88*****	**0.88*****	**0.88****	**0.88*****	**0.89****
MF-SV-VAR	1.00	0.95*	0.95**	0.97	0.92*	0.9**
MF-TVP-VAR	0.97	0.92***	0.93***	0.98	0.93***	0.94
MF-VAR	1.06	1.02	1.05	1.00	0.93	0.91*
Q-TVP-SV-VAR	0.94***	0.91***	0.90***	0.93***	0.9***	0.91***
Q-TVP-VAR	0.96	0.94***	0.96	0.93	0.90	0.95
Q-VAR	1.03	1.06***	1.08***	1.01	1.02	1.00
Q-SV-VAR	0.12	0.12	0.12	0.15	0.15	0.14
	Unemployment rate					
MF-TVP-SV-VAR	0.84***	0.90***	0.94	0.89	0.95	0.97
MF-SV-VAR	0.83***	**0.87*****	**0.90*****	**0.82*****	**0.86*****	**0.88*****
MF-TVP-VAR	**0.83*****	0.87***	0.92**	0.85**	0.89*	0.91
MF-VAR	0.89***	0.93**	0.96	0.86***	0.90**	0.93*
Q-TVP-SV-VAR	1.02	1.03	1.02	1.05**	1.05**	1.04
Q-TVP-VAR	1.03	1.05	1.04	1.02	1.02	1.01
Q-VAR	1.10***	1.10***	1.11***	1.08***	1.10***	1.11***
Q-SV-VAR	0.23	0.35	0.47	0.34	0.53	0.72
	Interest rate					
MF-TVP-SV-VAR	0.53***	0.70***	0.78***	0.39***	0.52***	0.58***
MF-SV-VAR	**0.52*****	**0.68*****	**0.75*****	**0.37*****	**0.49*****	**0.55*****
MF-TVP-VAR	0.7***	0.91	1.00	0.72***	0.99	1.09
MF-VAR	0.90*	1.03	1.04	0.97	1.14	1.16*
Q-TVP-SV-VAR	0.97*	0.97	0.98	0.85***	0.84***	0.86***
Q-TVP-VAR	1.30***	1.31***	1.29***	1.46***	1.54***	1.53***
Q-VAR	1.43***	1.35***	1.30***	1.73***	1.75***	1.71***
Q-SV-VAR	0.34	0.50	0.66	0.29	0.41	0.55

Notes: The models are detailed in Section 3. 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 on the 15%, 10%, and 5% level, respectively, according to a t-test on the average difference in scores relative to the benchmark model with Newey-West standard errors.

5.4 Forecasting during the Great Recession

In the previous Sections we demonstrated that modeling intra-quarterly dynamics significantly improves forecast accuracy on average, in particular with regard to the novel MF-TVP-SV-VAR.

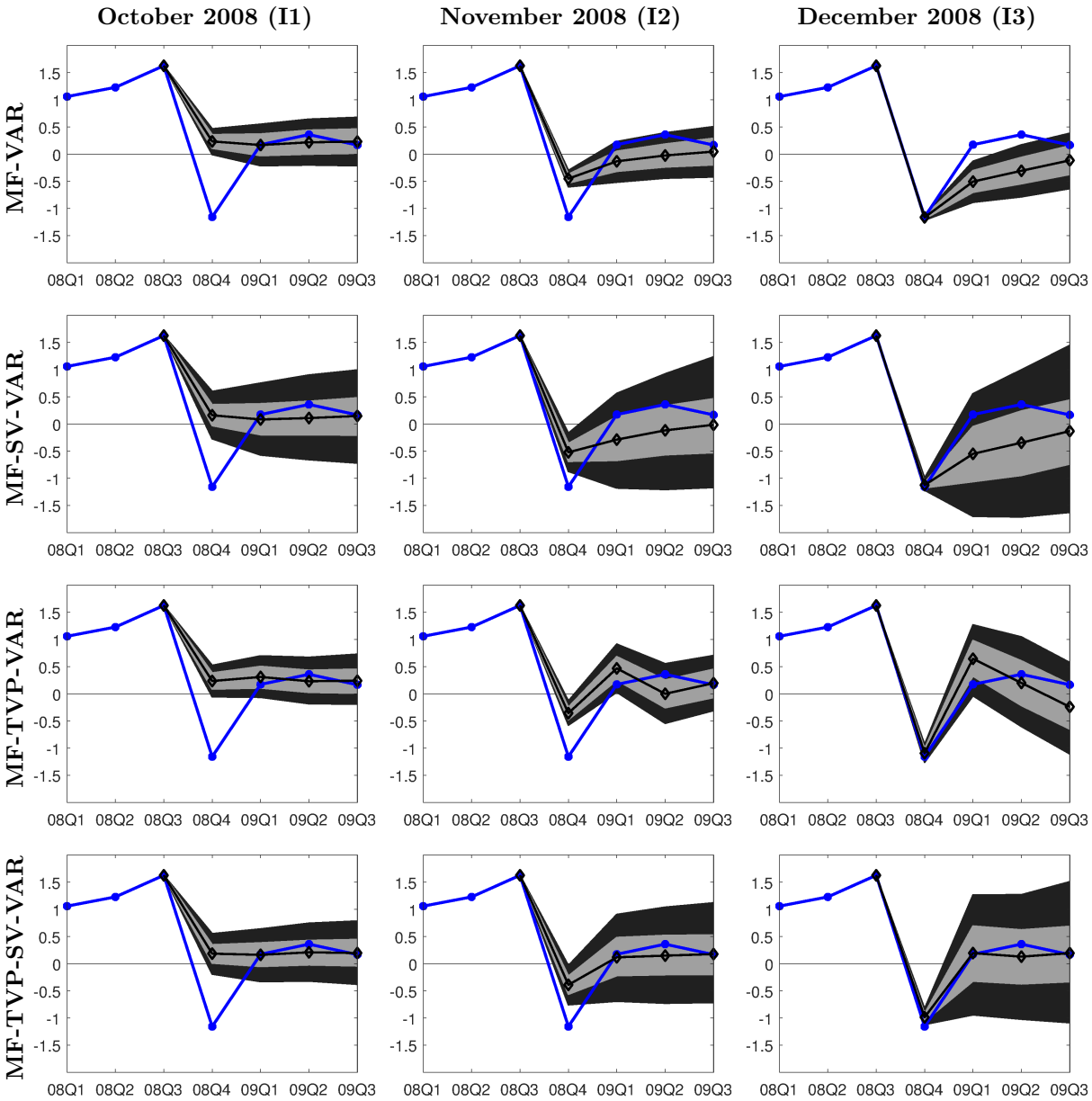
Subsequently, we take a closer look on the MF-models' absolute performance during the Great Recession. The latter is of great interest, because many structural and nonstructural models failed in providing accurate forecasts for the steep contraction and the following upswing in 2008/09. Since the MF-models perform especially well for forecasting inflation, Figure 3 depicts real-time quarter-on-quarter CPI inflation growth (blue line) along with both the means (black line) and 60% as well as 90% error bands (shaded areas) from the predictive distributions, respectively.[17] The columns of the figure refer to the data vintages of October 2008 until December 2008 and demonstrate how the arrival of new data points affects the forecasts. Let us consider the forecasts computed with the vintage of October 2008 (the first column). Note that in this month the models do not have any information on the current quarter expect for the T-Bill rate observation of October. At this data vintage, the models' posterior means are rather close to each other for each horizon – for the current quarter all of them lie at roughly 0.25%, which is approximately one percentage point too high compared to the realization. In contrast, the forecast intervals show noteworthy differences. The MF-VAR and the MF-TVP-VAR deliver narrow intervals, which assign only a small fraction of probability mass to negative inflation rates – for the nowcast almost no probability mass. In contrast, the MF-SV-VAR and the MF-TVP-SV-VAR generate much wider intervals, clearly including negative growth rates. However, the realization is not included in any interval. In November 2008 the posterior means are still close to each other, but become a lot of more pessimistic. Now each model correctly anticipates a negative growth rate for the nowcast – the nowcasts drop to about -0.5%. Thus, as already indicated in Section 5.1, the forecast errors become remarkably smaller due to the additional monthly observations. Moreover, there are considerable differences in the posterior means for higher horizons. The models with fixed VAR-coefficients predict a very slow recovery with negative inflation rates until 2009Q3. The TVP-models in turn predict – in line with the realizations – positive rates from 2009Q1 onwards; the MF-TVP-SV-VAR almost exactly predicts the growth rate for 2009Q1. The same pattern holds for the forecasts from December 2008. Now each model produces a forecast error of almost zero for the nowcast with a very narrow forecast interval. The subsequent recovery is, however, much better predicted by the TVP-models. 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 underpins the findings from Section 5.1 and 5.2; it demonstrates the importance of stochastic volatility for accurate nowcasts and the relevancy of time-varying parameters for precise forecasts. To improve the forecast accuracy on average it is recommended to combine both specifications.

6 Conclusion

Several studies have shown that modeling structural change improves forecast accuracy. We contribute to this discussion by investigating whether allowing for structural change in a mixed-frequency setup can further increase this performance. To this end, we use a Bayesian VAR that incorporates both time-varying parameters and stochastic volatility and is able to cope with

[17] Figures for the remaining variables are provided in Appendix F.

Figure 3: Inflation Forecasts during the Great Recession



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

indicators sampled at different frequencies.

We conduct a rigorous real-time out-of-sample forecast experiment and generate predictions for GDP growth, CPI inflation, the unemployment rate, and the 3-month Treasury bill rate. Our findings show that modeling monthly dynamics leads to substantial gains in forecast accuracy. Especially, nowcasts and short-term forecasts benefit from within-quarter information, while on longer horizons the advantages vanish. The novel MF-TVP-SV-VAR provides on average the best point and density forecast performance. In particular, inflation forecasts considerably benefit from modeling both monthly dynamics and structural change. For the remaining variables

the picture is less clear-cut. The MF-SV-VAR delivers the best forecasts for the interest rate, while the MF-TVP-VAR provides superior forecasts for the unemployment rate. We obtain rather mixed results for the GDP growth rate forecasts with no model dominating over all horizons, though almost all nonlinear MF-models outperform their linear counterparts as well as the remaining quarterly models. Finally, we assess the forecast performance during the Great Recession and demonstrate how the inflow of monthly information alters the inflation forecasts. We show that SV-models achieve the best performance for the downturn, while TVP-models are more precise in the subsequent recovery. Using the combined specification (MF-TVP-SV-VAR) is superior on average.

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. Thus, future research should focus on how to process larger dataset in this model framework and on how to select the most informative indicators.

References

Adolfson, M., Lindé, J. and Villani, M. (2007). Forecasting Performance of an Open Economy DSGE Model. *Econometrics Review*, **26** (2-4), 289–328.

Amir-Ahmadi, P., Matthes, C. and Wang, M.-C. (2018). Choosing Prior Hyperparameters: With Applications To Time-Varying Parameter Models. *Journal of Business & Economic Statistics*, **forthcoming**.

Antolin-Diaz, J., Drechsel, T. and Petrella, I. (2017). Tracking the Slowdown in Long-Run GDP Growth. *The Review of Economics and Statistics*, **99** (2), 343–356.

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

Barnett, A., Mumtaz, H. and Theodoridis, K. (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.

Barsoum, F. and Stankiewicz, S. (2015). Forecasting GDP growth using mixed-frequency models with switching regimes. *International Journal of Forecasting*, **31** (1), 33 – 50.

Baumeister, C. and Benati, L. (2013). Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound. *International Journal of Central Banking*, **9** (2), 165–212.

Bessec, M. and Bouabdallah, O. (2015). Forecasting GDP over the Business Cycle in a Multi-Frequency and Data-Rich Environment. *Oxford Bulletin of Economics and Statistics*, **77** (3), 360–384.

Carriero, A., Clark, T. and Marcellino, M. (2013). *Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility*. Working Paper 9312, Centre for Economic Policy Research, London.

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

—, — and — (2016). Common Drifting Volatility in Large Bayesian VARs. *Journal of Business & Economic Statistics*, **34** (3), 375–390.

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

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

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

Cimadomo, J. and D'Agostino, A. (2016). Combining Time Variation and Mixed Frequencies: An Analysis of Government Spending Multipliers in Italy. *Journal of Applied Econometrics*, **31** (7), 1276–1290.

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

— (2012). Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility. *Journal of Business & Economic Statistics*, **29** (3), 327–341.

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

Clements, M. P. and Galvão, A. B. (2008). Macroeconomic Forecasting With Mixed-Frequency Data: Forecasting Output Growth in the United States. *Journal of Business & Economic Statistics*, **26** (4), 546–554.

Cogley, T., Morozov, S. and Sargent, T. J. (2005). Bayesian fan charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system. *Journal of Economic Dynamics & Control*, **29** (11), 1893–1925.

— and Sargent, T. J. (2001). Evolving Post-World War II US inflation Dynamics. In *NBER Macroeconomics Annual 2001*, *NBER Chapters*, vol. 16, National Bureau of Economic Research, Inc, pp. 331–373.

— and — (2005). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics*, **8** (3), 262–302.

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

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

Diebold, F. X., Gunther, T. A. and Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, **39** (4), 863–883.

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

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. No. 9780198523543 in OUP Catalogue, Oxford University Press.

Faust, J. and Wright, J. H. (2013). Chapter 1 - Forecasting Inflation. In G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 2, Elsevier, pp. 2–56.

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

— and Marcellino, M. (2013). *A Survey of Econometric Methods for Mixed-Frequency Data*. Working Paper 2013/06, Norges Bank.

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

Gadea Rivas, M., Gómez-Loscos, A. and Pérez-Quirós, G. (2014). *The two greatest. Great recession vs. great moderation*. Working Paper 1423, Banco de Espana.

Galí, J. and Gambetti, L. (2009). On the Sources of the Great Moderation. *Amercian Economic Journal: Macroeconomics*, **1** (1), 26–57.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, **26** (2), 216–230.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004). *The MIDAS Touch: Mixed Data Sampling Regression Models*. CIRANO Working Papers 2004s-20, CIRANO.

Giannone, D., Reichlin, L. and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55** (4), 665–676.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102** (477), 359–378.

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

Good, I. J. (1952). Rational decisions. *Journal of the Royal Statistical Society. Series B (Methodological)*, **14** (1), 107–114.

Huber, F. (2016). Density forecasting using Bayesian global vector autoregressions with stochastic volatility. *International Journal of Forecasting*, **32** (3), 818–837.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1995). *Models and Priors for Multivariate Stochastic Volatility*. CIRANO Working Papers 95s-18, CIRANO.

Jore, A. S., Mitchell, J. and Vahey, S. P. (2010). Combining forecast densities from VARs with uncertain instabilities. *Journal of Applied Econometrics*, **25** (4), 621–634.

Kim, C.-J. and Nelson, C. R. (1999). Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a MarkovSwitching Model of the Business Cycle. *The Review of Economics and Statistics*, **81** (4), 608–616.

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

Koop, G. and Korobilis, D. (2014). A new index of financial conditions. *European Economic Review*, **71**, 101–116.

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

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

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

Matheson, J. E. and Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management Science*, **22** (10), 1087–1096.

McConnell, M. M. and Perez-Quiros, G. (2000). Output Fluctuations in the United States: What Has Changed since the Early 1980's? *American Economic Review*, **90** (5), 1464–1476.

Mikosch, H. and Neuwirth, S. (2015). *Real-time forecasting with a MIDAS VAR*. Working Paper 15-377, KOF Swiss Economic Institute, ETH Zurich.

Mitchell, J., Smith, R. J., Weale, M. R., Wright, S. and Salaza, E. L. (2005). An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth. *The Economic Journal*, **115** (501), 108–129.

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

Rossi, B. and Sekhposyan, T. (2014). Evaluating predictive densities of US output growth and inflation in a large macroeconomic data set. *International Journal of Forecasting*, **30** (3), 662–682.

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

Wohlrabe, K. (2009). *Forecasting with mixed-frequency time series models*. Munich Dissertations in Economics 9681, University of Munich, Department of Economics.

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

Apart from the VAR with constant volatilities, which uses a Jeffrey's prior, the priors for the remaining model specifications are based on a training sample. The latter consists of the first ten years of the entire sample. In the following, variables denoted with OLS refer to OLS quantities based on this training sample. The length of the trainings sample is denoted by T_0.

AR-coefficients

To keep the models comparable, we draw the AR-coefficients for each nonlinear specification using the CK algorithm with the following prior:

$$p(\beta_0) \sim N(\hat{\beta}_{OLS}, 4 * V(\hat{\beta}_{OLS})) \tag{A.1}$$

In case of the VAR-SV, we use the first draw of the backward recursion of the CK algorithm, i.e. $\beta_{T|T}$, for each period. For the benchmark VAR we implement a diffuse Jeffrey's prior:

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

The prior for the covariance of the AR-coefficients Q follows an inverse Wishart distribution:

$$p(Q) \sim IW(k_Q^2 \times T_0 \times V(\hat{\beta}_{OLS}), T_0). \tag{A.3}$$

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 priors distribution as:

$$p(\log \sigma_0) \sim N(\log \hat{\sigma}_{OLS}, I_n) \tag{A.4}$$

$$p(A_0) \sim N(\hat{A}_{OLS}, 4 * V(\hat{A}_{OLS})). \tag{A.5}$$

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

$$p(\Psi) \sim IW(k_\Psi^2 \times (1+n) \times I_n, 4) \tag{A.6}$$

$$p(\Phi_i) \sim IW(k_\Phi^2 \times (i+1) \times V(\hat{A}_{i,OLS}), i+1) \qquad i = 1, \ldots, k-1, \tag{A.7}$$

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 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:

$$p(z_0) \sim N(z_L, I_{np}). \tag{A.8}$$

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, k_Ψ, and k_Φ. Therefore, we follow Amir-Ahmadi *et al.* (2018) and use priors for those hyperparameters. Since we do not have any a priori knowledge about the true values of any of our models, we use uniform priors:

$$p(k_i) \sim U(1e^{-10}, 1) \qquad i = Q, \Phi, \Psi. \tag{A.9}$$

The lower and upper bound of the distribution are chosen to cover a broad range of possible values, including the default values used by Primiceri (2005).[18]

B Specification of the Gibbs sampler

To estimate the models we employ a Gibbs sampler, which consecutively draws from the conditional distribution. In the following, the general form of the MCMC algorithm according to Primiceri (2005) with the correction from Del Negro and Primiceri (2015) is outlined. In order 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 $p(\Phi | \tilde{y}^T, \beta^T, Q, \Sigma^T, A^T, \Psi)$.

[18] To ensure convergence of the MH-algorithm in case of the MF-TVP-SV-VAR, the lower bound for k_Q is choosen higher as the value in Primiceri (2005).

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 from $p(k_Q|Q) = p(Q|k_Q)p(k_Q)$
    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 quaterly models, the Gibbs sampler 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 details, see Del Negro and Primiceri (2015)). In step 10, we include the Metropolis-Hastings within the Gibbs sampler to draw our Hyperparameters.

For ease of exposition, we use in the following $\tilde{y}^T$ 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 5th draw of 10000 after burn-in draws for posterior inference.

2. Step: 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 $P_{t|t}$.

3. Step: 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.

4. Step: 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, i.e. the size of the training sample, respectively.

5. Step: 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:

$$A_t(\tilde{y}_t - Z_t'\beta_t) = \tilde{y}_t^* = \Sigma_t u_t \tag{A.10}$$

where, taking into account that β_T and $\tilde{y}_t$ is 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:

$$\hat{\tilde{y}}_{1,t} = \sigma_{1,t} u_{1,t} \tag{A.11}$$

$$\hat{\tilde{y}}_{i,t} = -\hat{y}_{[1,i-1]} a_{i,t} + \sigma_{i,t} u_{i,t} \qquad i = 2, \ldots, k, \tag{A.12}$$

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.

6. Drawing the covariance Φ_i of the elements of A^T
   Φ_i has an inverse Wishart 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 degree of freedoms, respectively.

7. Step: 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 degree of freedoms, respectively.

8. Step: 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):

$$PR(s_{i,t} = j | \tilde{y}^{**}, h_{i,t}) \propto q_j f_N(\tilde{y}^{**} | 2h_{i,t} + m_j - 1.2704, \nu_j^2). \tag{A.13}$$

9. Step: 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.10). Taking squares and logarithms, we get

$$\tilde{y}_t^{**} = 2r_t + \nu_t \tag{A.14}$$

and for the volatility process:

$$h_t = h_{t-1} + \varepsilon_t \tag{A.15}$$

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 Chi-squared 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 and read as follows:

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).

Conditional on s^T – the indicator matrix, governing composition of the mixture distribution for every ν_t, $t = 1, \ldots, T$ – the CK algorithm enables us to recursively get draws for

$$h_{t|t+1} = E(h_t|h_{t+1}, \tilde{y}^t, A^T, B^T, Q, s^T, \Psi, \Phi) \tag{A.16}$$

$$H_{t|t+1} = VAR(h_t|h_{t+1}, \tilde{y}^t, A^T, B^T, Q, s^T, \Psi, \Phi). \tag{A.17}$$

10. Step: Drawing the hyperparameters k_Q, k_Ψ, 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.* (2018) show that the acceptance probability for each draw i can be simplified to:

$$\alpha_{k_X}^i = min\left(\frac{p(X|k_X^*)p(k_X^*)q(k_X^*|k_X^{i-1})}{p(X|k_X^{i-1})p(k_X^{i-1})q(k_X^{i-1}|k_X^*)}, 1\right), \tag{A.18}$$

where $X = \{Q, \Psi, \Phi\}$. $p(X|k_X^*)$ denotes 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 the random walk chain algorithm:

$$k_X^* = k_X^{i-1} + \xi_t \qquad \xi_t \sim N(0, \sigma_{k_X}^2), \tag{A.19}$$

and the standard deviation σ_{k_X} is adjusted in every 500th step of the burn-in period by:

$$\sigma_{k_X}^{adj} = \sigma_{k_X} \frac{\overline{\alpha}_{k_X}}{\alpha^*}, \tag{A.20}$$

where $\overline{\alpha}_{k_X}$ is the average acceptance rate over the 500 draws and $\alpha^* = 0.4$ is the target acceptance rate. 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 Log scores

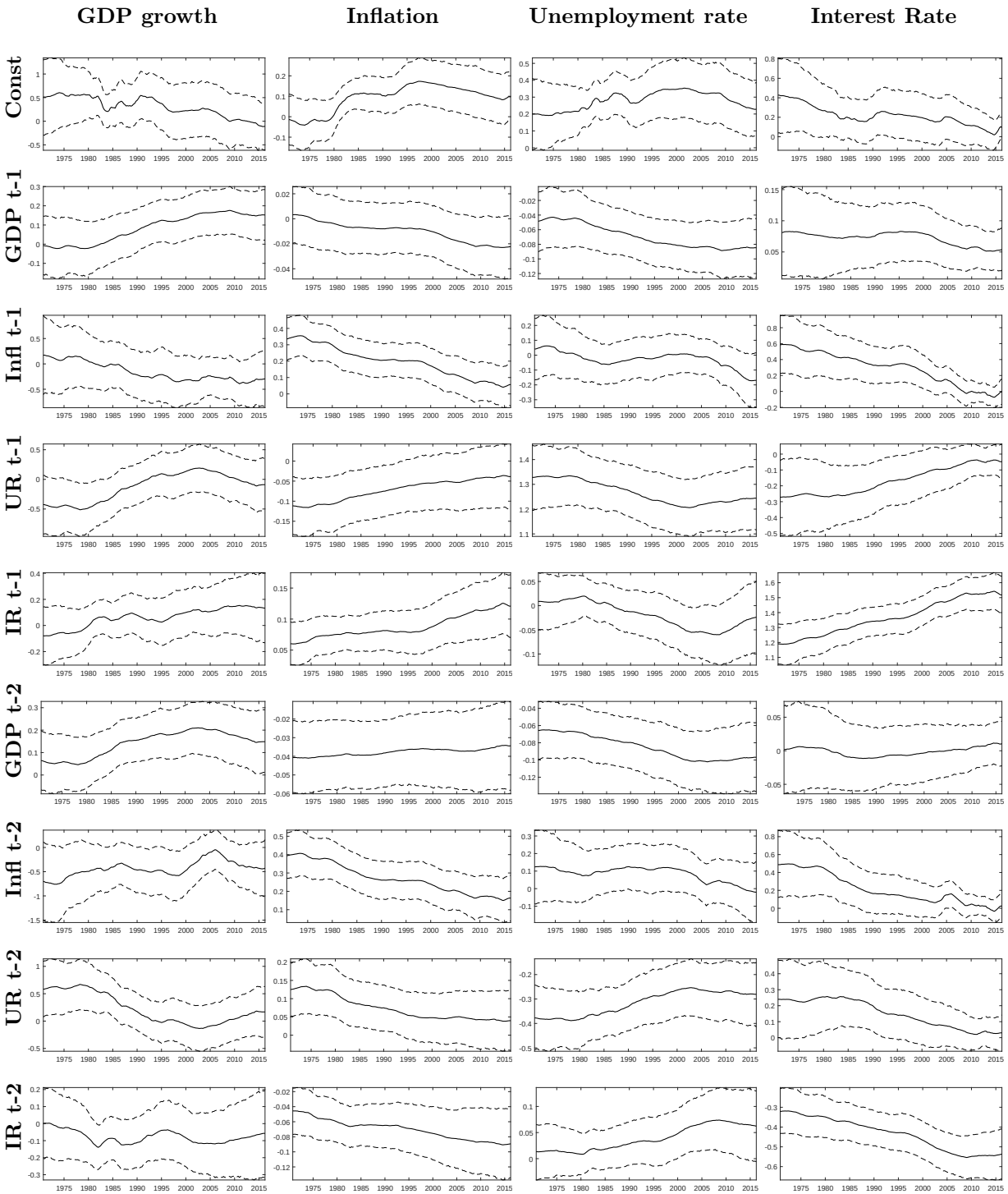
Table 4: Real-time Forecast LSs

Model	1995-2016 h = 2	1995-2016 h = 3	1995-2016 h = 4	2008-2016 h = 2	2008-2016 h = 3	2008-2016 h = 4
	GDP growth					
MF-TVP-SV-VAR	-0.33***	**0.04**	**0.08***	-0.47	**0.18*****	**0.17*****
MF-SV-VAR	-1.61***	-0.16	0.01	-3.03**	-0.37	0
MF-TVP-VAR	-0.04	-0.15***	-0.15***	0.02	-0.03	-0.11***
MF-VAR	-0.2***	-0.21***	-0.17***	-0.35***	-0.28***	-0.26***
Q-TVP-SV-VAR	**0**	0.01	0	**0.04**	0.08	0.05
Q-TVP-VAR	-0.16***	-0.24***	-0.24***	-0.1*	-0.12	-0.17*
Q-VAR	-0.21***	-0.26***	-0.24***	-0.28***	-0.27***	-0.3***
Q-SV-VAR	-0.94	-0.91	-0.97	-1.19	-1.24	-1.25
	Inflation					
MF-TVP-SV-VAR	**0.07**	**0.12*****	**0.16*****	**0.11**	**0.09**	**0.15*****
MF-SV-VAR	-0.03	0.03	0.05	0.04	0.09	0.12**
MF-TVP-VAR	-0.81	-0.65	-0.17	-2.01	-1.71	-0.5
MF-VAR	-2.23*	-0.62	-0.52	-5.35	-1.38	-1.05
Q-TVP-SV-VAR	0.07***	0.07***	0.1***	0.11***	0.03	0.04
Q-TVP-VAR	-0.53	-0.3	-0.33	-1.24	-0.71	-0.76
Q-VAR	-1.56	-0.51	-0.63	-3.75	-1.1	-1.34
Q-SV-VAR	0.18	0.09	0.11	-0.06	-0.1	-0.05
	Unemployment rate					
MF-TVP-SV-VAR	-0.02	0.38**	**0.34***	0.06	0.66*	0.63
MF-SV-VAR	**0.21*****	0.34**	0.14	**0.3*****	0.66	0.27
MF-TVP-VAR	0.21***	**0.35****	0.3	0.2**	0.75***	0.71
MF-VAR	-1.07	-0.67	-0.82	-2.79	-1.68	-2.02
Q-TVP-SV-VAR	0.04	0.24	0.26	0.04	0.56	0.61
Q-TVP-VAR	0.05	0.26	0.31	0.23	**0.81**	**0.93**
Q-VAR	-0.63***	-0.23	-0.31**	-1.37**	-0.4	-0.63*
Q-SV-VAR	-0.55	-1.18	-1.52	-1.02	-2.09	-2.55
	Interest rate					
MF-TVP-SV-VAR	0.77***	0.43***	0.3***	0.9***	0.59***	0.46***
MF-SV-VAR	**0.8*****	**0.47*****	**0.35*****	**1.03*****	**0.7*****	**0.58*****
MF-TVP-VAR	0.15	-0.12	-0.18**	-0.18*	-0.44***	-0.49***
MF-VAR	-0.14	-0.27***	-0.23***	-0.5***	-0.59***	-0.53***
Q-TVP-SV-VAR	0.01	0	-0.04	0.02	0.02	0.02
Q-TVP-VAR	-0.46***	-0.48***	-0.45***	-0.83***	-0.83***	-0.82***
Q-VAR	-0.6***	-0.52***	-0.42***	-1.01***	-0.93***	-0.82***
Q-SV-VAR	-0.8	-1.18	-1.49	-0.37	-0.78	-1.12

Notes: The models are detailed in Section 3. The scores are reported in absolute terms for the benchmark (the bottom row of each panel) and in differences to the benchmark for the remaining models. A positive difference indicates that the model outperforms the benchmark. Bold figures indicate the best performance for the variable and horizon. $*$, $**$, and $***$ denote significance on the 15%, 10%, and 5% level, respectively, according to a t-test on the average difference in scores relative to the benchmark model with Newey-West standard errors.

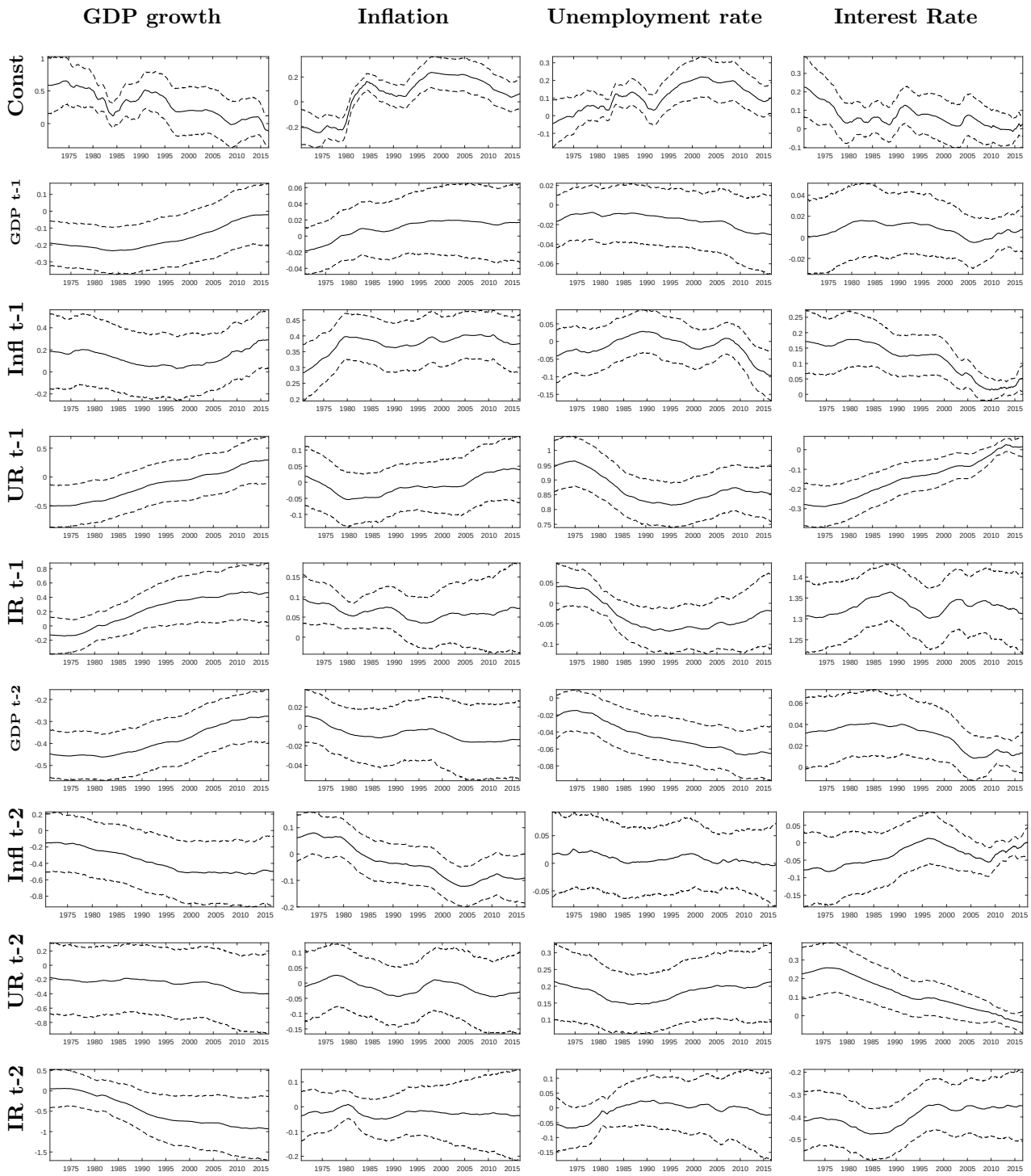
D Time-varying parameters

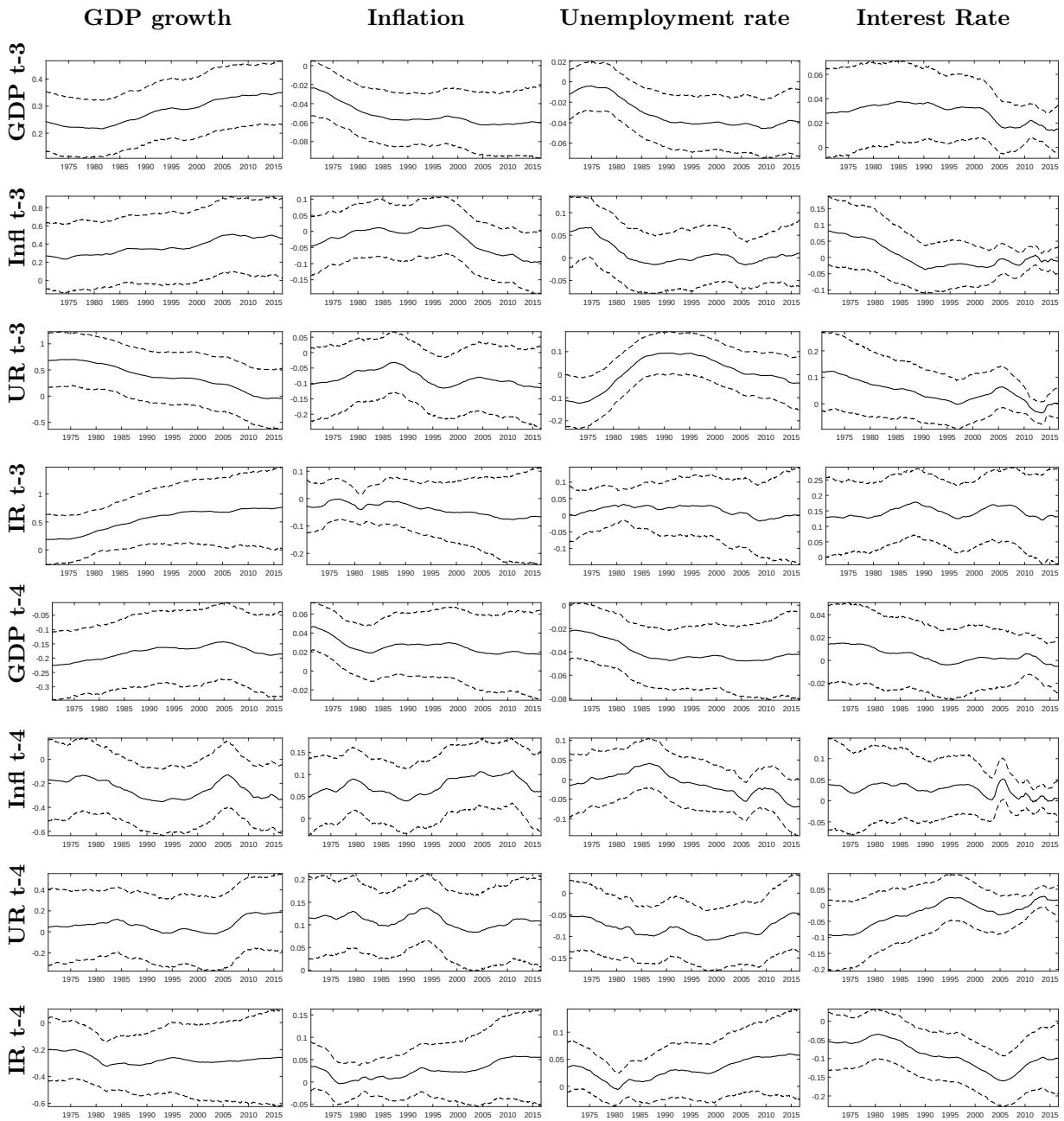
Figure 4: Time-Varying Parameters of the Q-TVP-SV-VAR



Note: 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 5: Time-Varying Parameters of the MF-TVP-SV-VAR

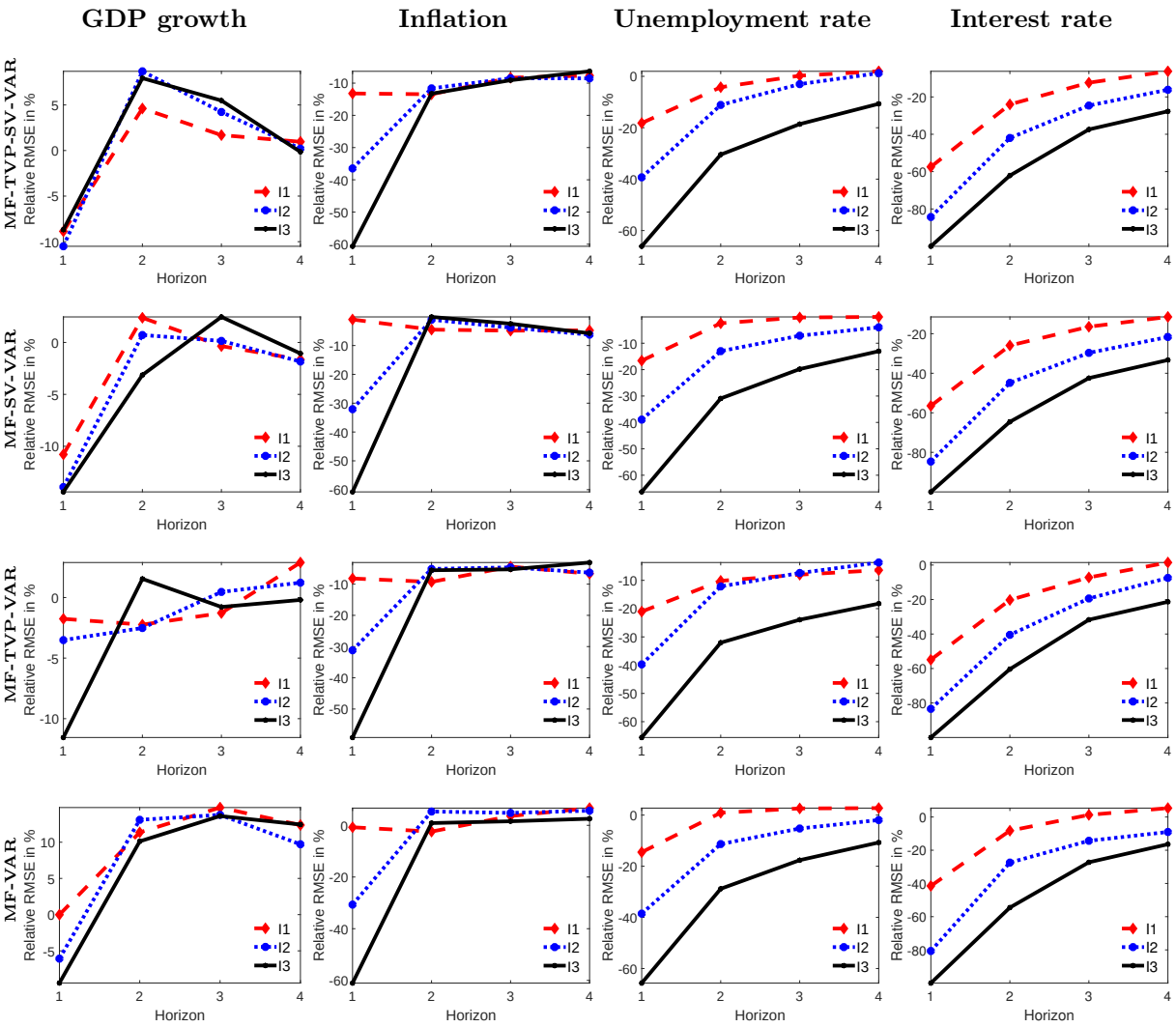




Note: Figure depicts the time-varying parameters from the MF-TVP-SV-VAR. Columns refer to the variable and rows to the constant/lagged variable on which the variable is regressed. The dashed lines indicate 68% error bands. Results are based on the last data vintage.

E Relative RMSEs

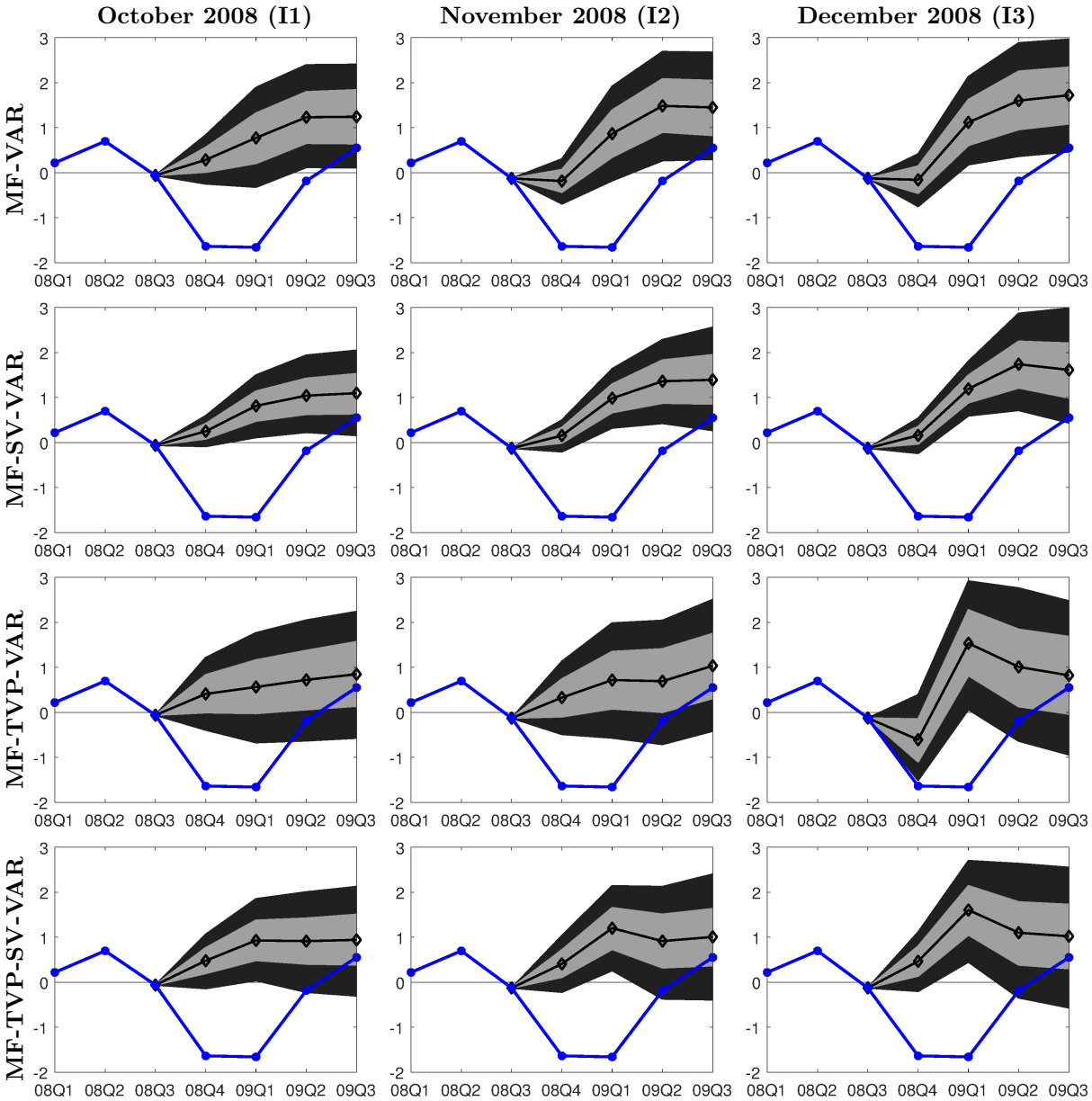
Figure 6: Relative RMSEs



Note: 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 outline in Section 2.2, respectively.

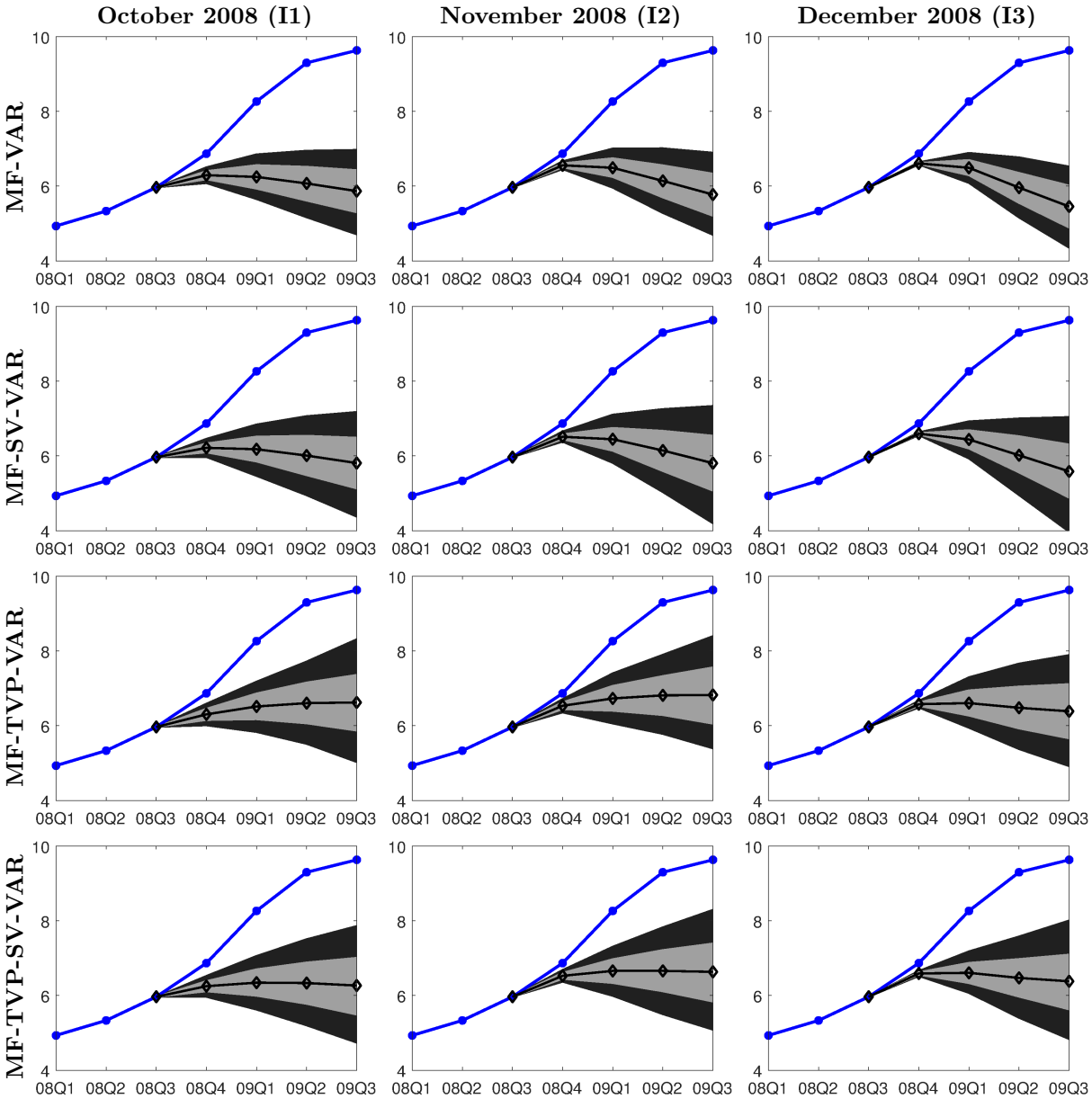
F Forecasting during the Great Recession - GDP, Unemployment Rate, and Interest Rate

Figure 7: GDP Forecasts during the Great Recession



Note: The rows refer to mixed-frequency models. The columns refer to the forecast origins, i.e. the information sets. The blue line indicates quarter-on-quarter real-time GDP growth, the black line is the mean of the predictive distribution. Shaded areas are 60% and 90% error bands from the predictive distributions.

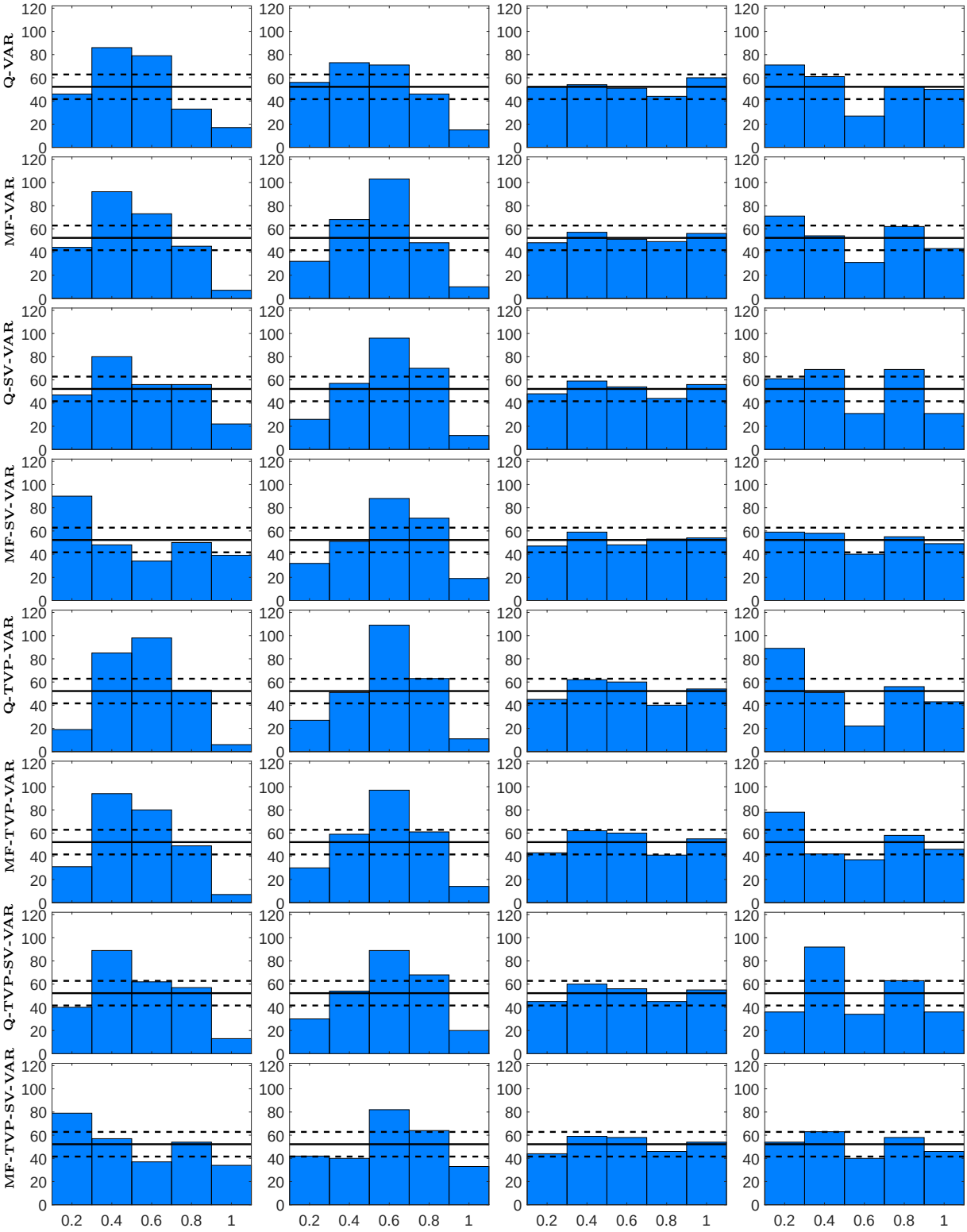
Figure 8: Unemployment Forecasts during the Great Recession



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

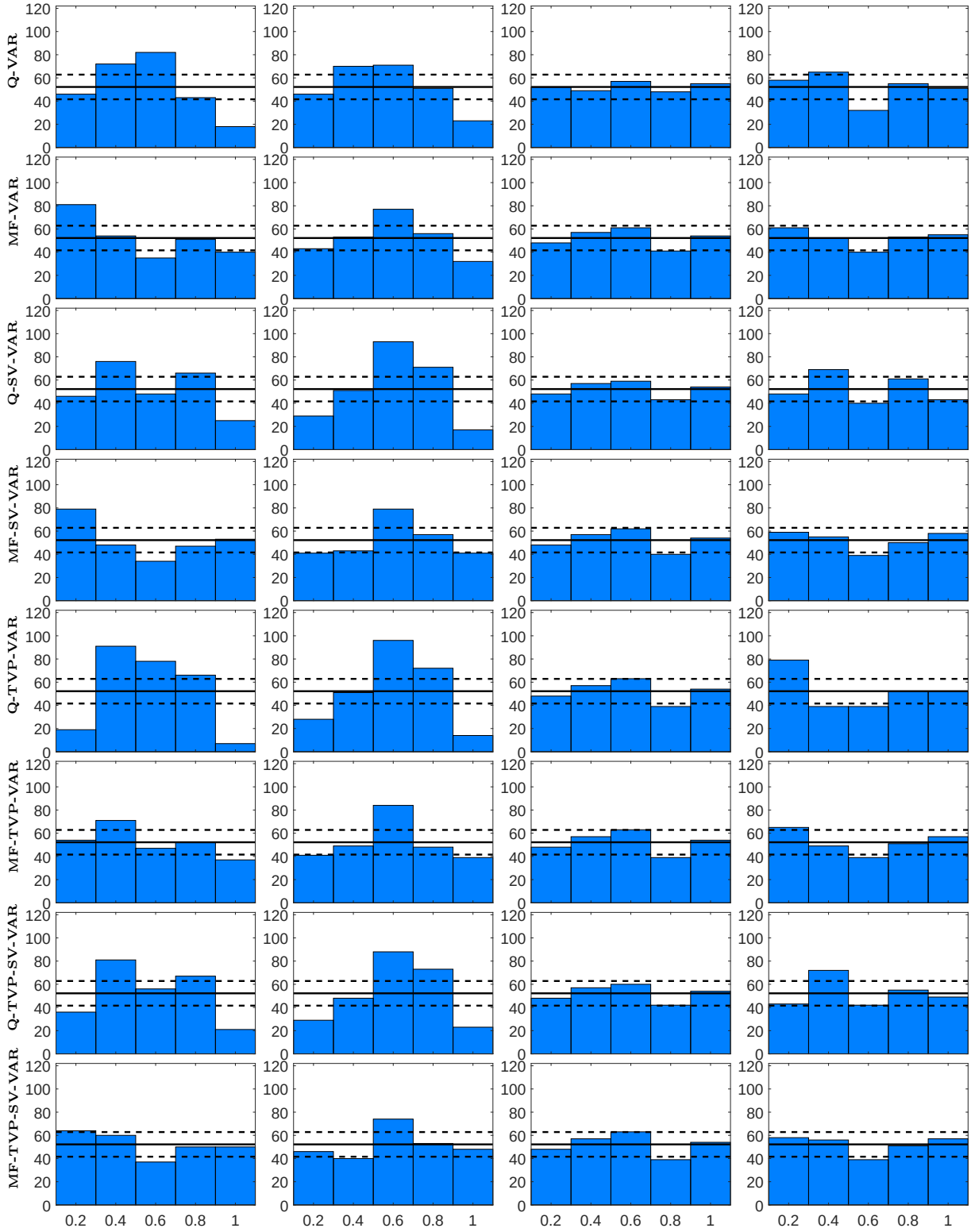
G Probability Integral Transforms

Figure 9: Probability Integral Transforms for Inflation Forecasts



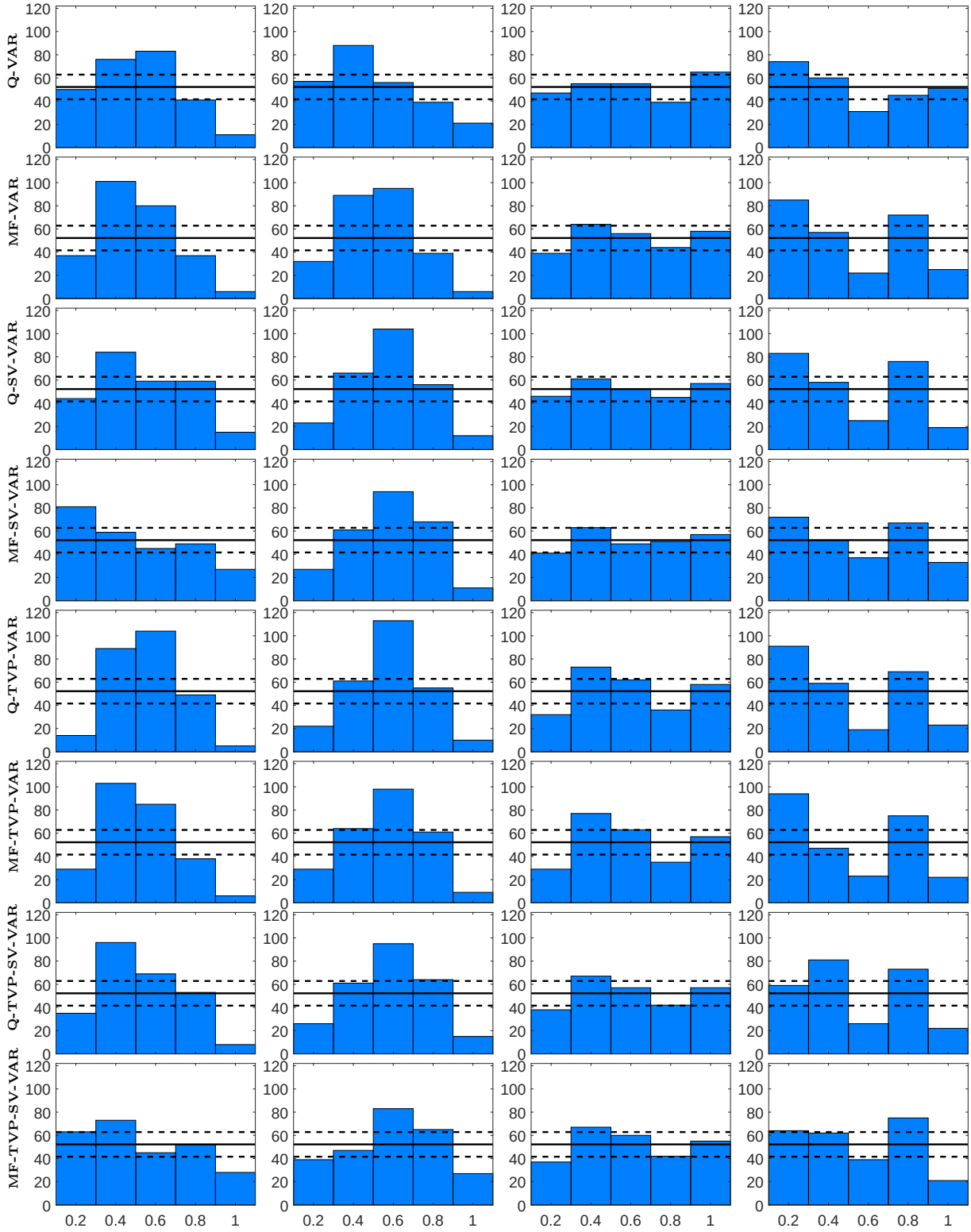
Note: The rows refer to the PITs of each model. The columns refer to the forecast horizons. The solid line indicates uniformity and the dashed lines 90% confidence bands as in Rossi and Sekhposyan (2014).

Figure 10: Probability Integral Transforms for GDP Forecasts



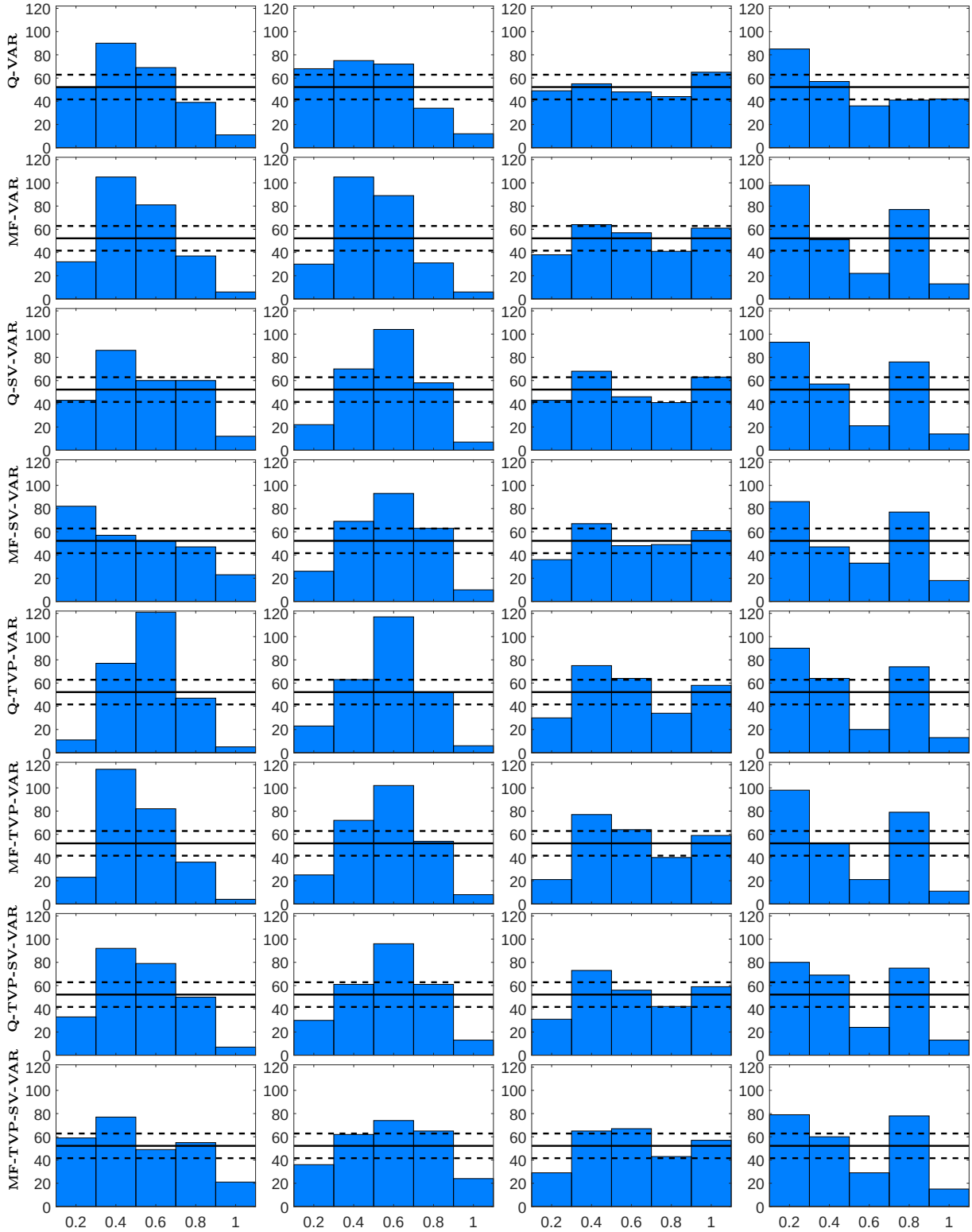
Note: The rows refer to the PITs of each model. The columns refer to the forecast horizons. The solid line indicates uniformity and the dashed lines 90% confidence bands as in Rossi and Sekhposyan (2014).

Figure 11: Probability Integral Transforms for Unemployment Rate Forecasts



Note: The rows refer to the PITs of each model. The columns refer to the forecast horizons. The solid line indicates uniformity and the dashed lines 90% confidence bands as in Rossi and Sekhposyan (2014).

Figure 12: Probability Integral Transforms for Interest Rate Forecasts



Note: The rows refer to the PITs of each model. The columns refer to the forecast horizons. The solid line indicates uniformity and the dashed lines 90% confidence bands as in Rossi and Sekhposyan (2014).

ifo Working Papers

No. 272 Potrafke, N., The globalisation-welfare state nexus: Evidence from Asia, October 2018.

No. 271 Sandkamp, A. and S. Yang, Where Has the Rum Gone? Firms' Choice of Transport Mode under the Threat of Maritime Piracy, October 2018.

No. 270 Weissbart, C., Decarbonization of Power Markets under Stability and Fairness: Do They Influence Efficiency?, October 2018.

No. 269 Hausfeld, J. and S. Resnjanskij, Risky Decisions and the Opportunity of Time, October 2018.

No. 268 Bornmann, L., K. Wohlrabe and S. Gralka, The Graduation Shift of German Universities of Applied Sciences, October 2018.

No. 267 Potrafke, N., Does public sector outsourcing decrease public employment? Empirical evidence from OECD countries, October 2018.

No. 266 Hayo, B. and F. Neumeier, Central Bank Independence in New Zealand: Public Knowledge About and Attitude Towards the Policy Target Agreement, October 2018.

No. 265 Reif, M., Macroeconomic Uncertainty and Forecasting Macroeconomic Aggregates, October 2018.

No. 264 Wohlrabe, K., F. de Moya Anegon and L. Bornmann, How efficiently produce elite US universities highly cited papers? A case study based on input and output data, October 2018.

No. 263 Schwefer, M., Sitting on a Volcano: Domestic Violence in Indonesia Following Two Volcano Eruptions, September 2018.

No. 262 Vandrei, L., Does Regulation Discourage Investors? Sales Price Effects of Rent Controls in Germany, June 2018.

No. 261 Sandkamp, A.-N., The Trade Effects of Antidumping Duties: Evidence from the 2004 EU Enlargement, June 2018.

No. 260 Corrado, L. and T. Schuler, Financial Bubbles in Interbank Lending, April 2018.

No. 259 Löffler, M., A. Peichl and S. Siegloch The Sensitivity of Structural Labor Supply Estimations to Modeling Assumptions, March 2018.

No. 258 Fritzsche, C. and L. Vandrei, Causes of Vacancies in the Housing Market – A Literature Review, March 2018.

No. 257 Potrafke, N. and F. Rösel, Opening Hours of Polling Stations and Voter Turnout: Evidence from a Natural Experiment, February 2018.

No. 256 Hener, T. and T. Wilson, Marital Age Gaps and Educational Homogamy – Evidence from a Compulsory Schooling Reform in the UK, February 2018.

No. 255 Hayo, B. and F. Neumeier, Households' Inflation Perceptions and Expectations: Survey Evidence from New Zealand, February 2018.

No. 254 Kauder, B., N. Potrafke and H. Ursprung, Behavioral determinants of proclaimed support for environment protection policies, February 2018.

No. 253 Wohlrabe, K., L. Bornmann, S. Gralka und F. de Moya Anegon, Wie effizient forschen Universitäten in Deutschland, deren Zukunftskonzepte im Rahmen der Exzellenzinitiative ausgezeichnet wurden? Ein empirischer Vergleich von Input- und Output-Daten, Februar 2018.

No. 252 Brunori, P., P. Hufe and D.G. Mahler, The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees, January 2018.

No. 251 Barrios, S., M. Dolls, A. Maftei, A. Peichl, S. Riscado, J. Varga and C. Wittneben, Dynamic scoring of tax reforms in the European Union, January 2018.

No. 250 Felbermayr, G., J. Gröschl and I. Heiland, Undoing Europe in a New Quantitative Trade Model, January 2018.