TIME-VARYING MIXED-FREQUENCY VECTOR AUTOREGRESSIVE MODELS*

Thomas B. Götz† and Klemens Hauzenberger‡

Deutsche Bundesbank

September 9, 2015

Abstract

To simultaneously consider mixed-frequency time series, their joint dynamics, and possible structural changes, we introduce a time-varying parameter mixed-frequency vector autoregressive model (TVP-MF-VAR). To keep our approach from becoming too complex in comparison with time-invariant MF-VARs, we limit time variation to the intercepts and error variances. We estimate our model using two techniques: an approximate one using forgetting factors in the prediction step of the Kalman filter and exponentially weighted moving averages (EWMA) for the error variances; and one based on exact Bayesian inference. Both approaches allow us to evaluate moderately large VARs (up to around 10 variables) in a recursive forecasting exercise in a reasonable amount of time. For eight variables in Germany, we examine the performance of our TVP-MF-VAR model variants and compare them to the MF-VAR of Schorfheide and Song (2015). Our empirical results demonstrate the feasibility and usefulness of our methods, even in the presence of only mild structural changes.

JEL Codes: C32, C51, C53
JEL Keywords: Forecasting, Mixed Frequencies, Time-Varying Parameters, Bayesian VAR, Forgetting Factors

*We thank Josef Hollmayr for valuable comments on an earlier version of the paper. The views expressed in this paper are solely ours and should not be interpreted as reflecting the views of the Deutsche Bundesbank.

†Thomas B. Götz, Deutsche Bundesbank, Macroeconomic Analysis and Projection Division, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main. Email: thomas.goetz@bundesbank.de

‡Klemens Hauzenberger, Deutsche Bundesbank, Macroeconomic Analysis and Projection Division, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main. Email: klemens.hauzenberger@bundesbank.de
1 Introduction

Many macroeconomic models for short-term forecasting include two groups of observations with different release patterns: a few quarterly and monthly targets, focusing on the gross domestic product (GDP), industrial production and the nominal side; and several leading indicators, focusing on monthly surveys and financial market data. But many of the existing forecasting models ignore at least one of three related features and challenges: the mismatch in sampling frequencies, the possibility of (smooth) structural changes, and the joint dynamics between the target variables and the indicators. The ends of the current spectrum of existing models are, perhaps, the mixed-frequency vector autoregressive model of Schorfheide and Song (2015), the common-frequency time-varying parameter VAR of Koop and Korobilis (2013) and the univariate Markov-switching MI(xed) DA(ta) S(ampling) model of Guérin and Marcellino (2013).

In this paper we attempt to address, at least to some extent, all three features and challenges simultaneously by constructing a time-varying mixed-frequency VAR (TVP-MF-VAR hereafter). Accounting for the mismatch in sampling frequencies, first and foremost between monthly and quarterly observations, has established itself as the standard in professional short-term forecasting. The benefits of modeling different frequencies for forecasting have been shown throughout the respective field of the literature (see inter alia Andreou et al., 2013, Foroni et al., 2015 or Götz et al., 2015). But besides the permanent hunting for lower forecast errors, a mixed-frequency model gives the analyst an invaluable tool for the current assessment of the economy; with several forecast updates during a month. Modeling of time-varying parameters, however, and its potential use for forecasting is so not clear a-priori because of increased computational complexity and it may thus be case dependent (see, e.g., D’Agostino et al., 2013 and Eickmeier et al., 2011).1

1 Barnett et al. (2014) compare the forecasting performance of numerous models allowing for time-variation in the parameters. Using data for the UK they find that simple VARs estimated over rolling samples deliver the most accurate predictions of GDP growth, while more sophisticated time-varying parameter models yield substantially better inflation forecasts.
We take the mixed-frequency VAR of Schorfheide and Song (2015) as a starting point (and benchmark model) and implement time variation in the parameters and error variances in slightly simplified variations of the models presented in Cogley and Sargent (2005) and Primiceri (2005). We essentially strip down the number of time-varying parameters to a bare minimum—the intercepts and error variances—to keep our model from becoming too complex.

The mixed-frequency VAR of Schorfheide and Song (2015) is sampled at the high frequency, implying that latent high-frequency observations of the low-frequency series need to be estimated. An alternative to this parameter-driven (Cox et al., 1981) approach is the observation-driven mixed-frequency VAR of Ghysels (2015) that depends exclusively on observable data.\(^2\) Ghysels aligns his VAR with the low frequency and stacks the corresponding high- (and low-) frequency observations into a single vector.\(^3\) Hence, in contrast to the MF-VAR of Schorfheide and Song (2015), the model of Ghysels (2015) does not deliver, for instance, a monthly GDP series, that might be a desirable byproduct for policymakers.

To estimate our TVP-MF-VAR we will first provide a full Bayesian treatment. In a second approach we take up ideas from online predictions in engineering to control processes on the fly (see, e.g., Kulhavý and Zarrop, 1993, and Raftery et al., 2010) and from finance (see, e.g., Zumbach, 2006). In this way we further reduce complexity in the time-varying part of our model. Koop and Korobilis (2013) exploited these techniques for macroeconomic forecasting in large TVP-VARs. The key here are so called “forgetting factors”. Rather than specifying the state space fully and estimating it by Markov Chain Monte Carlo (MCMC) methods, which can be computationally demanding in larger

\(^2\)This model represents the multivariate extension of MIDAS models, first introduced in Ghysels et al. (2004). Because of their simplicity and wide range of applicability, MIDAS regressions became popular for forecasting macroeconomic or financial time series (see Ghysels et al., 2005, Clements and Galvão, 2008, Götze et al., 2015 among many others). Foroni and Marcellino (2013) provide an excellent survey on the subject matter. The development of time-varying MIDAS models has been initiated in Schumacher (2014) with related approaches such as the model of Guérin and Marcellino (2013).

\(^3\)The augmentation of the observation-driven mixed-frequency VAR to allow for time variation in the parameters and error variances, and a comparison with the model presented here, is beyond the scope of this paper, but already on our future research agenda.
systems, we specify the degree of parameter evolution by simple predetermined factors. We dub our two approaches “exact” and “approximate” and examine to what extent the simplicity-complexity trade-off affects the forecast performance in moderately large TVP-MF-VARs.

Closely related to our work is the paper of Cimadomo and D’Agostino (2015), who, simultaneously and independently, developed a time-varying VAR model allowing for variables sampled at various frequencies. The way in which they overcome the mixed-frequency data feature is slightly different from our approach, though. In particular, they fix the available low-frequency values and interpolate the remaining data points, while we fill in all high-frequency values, but require their average to equal the available low-frequency figure. Furthermore, allowing for time variation in all of the parameters restricts their model to a few variables and lags, 4 whereas our approach can handle systems of moderate size, even using a full Bayesian approach. Of course, they could easily handle larger systems with the help of forgetting factors. Finally, looking at the macroeconomic effects of government spending shocks, the main focus of their work is different from ours. However, their model could be readily used for forecasting as well.

Focusing on eight indicators of the German economy we inspect the performance of our methods, with a special attention to the regular forecasting exercise carried out at the Deutsche Bundesbank. Our results are promising and show that, even if there is even slight evidence for time variation in the data, forecast accuracy gains may be obtained in comparison to time-invariant MF-VAR models and a simple benchmark model.

We use some modeling and notational conventions throughout the paper. Our description will mostly rely on a VAR of order one to keep the model as clearly laid out as possible. As always, any VAR of higher lag-order can be transformed in its VAR(1) companion form. In cases we need to distinguish between the VAR(1) companion form and the original VAR(p) we will switch from capital to small letters. Furthermore, as implied by the approach of Schorfheide and Song (2015), we write the entire model in

---

4Their system consists of three indicators, whereby the amount of lags is, to best of our knowledge, not provided.
monthly frequency. An expression \( z^\tau \) denotes the complete history of a generic variable (or parameter) \( z_t \) up to time \( \tau \), i.e., \( z^\tau = [z'_1, \ldots, z'_\tau]' \). We reserve the letter \( y^T \) (and \( Y^T \)) for the incomplete data matrix with missing values, and \( x^T \) (and \( X^T \)) for the same data matrix but with the lower frequency variable replaced by its latent monthly equivalent and all other missing values filled.

2 The Mixed-Frequency Model with Time Variation

In this section we present our model and outline, first, how we deal with the inherent mixed-frequency data feature and, second, how we introduce time-varying parameters and stochastic volatility. Our approach can be roughly described as a combination of the mixed-frequency VAR of Schorfheide and Song (2015, henceforth MF-VAR) and variants of the time-varying VARs of Cogley and Sargent (2005) and Primiceri (2005). The deviations to them arise through two simplifications: we limit time variation to the intercepts and volatilities; and all underlying innovation matrices driving the parameters in the laws of motion are diagonal. Simplifying in this direction leads to robust inference, whichever data we observe, and to a management-friendly runtime, whenever we observe new data during a month and re-estimate the model. Inevitably, some trade-off is necessary here, but in the end the amount of time variation we allow for should be sufficient to potentially improve the forecast performance compared to the benchmark MF-VAR. That is our main goal.

2.1 Block I: The Mixed-Frequency Part

We first start by describing the mixed-frequency framework and the associated state-space representation of our extended version of the MF-VAR. Let \( x_{m,t} \) and \( x_{q,t} \) be the \( n_m \times 1 \) and \( n_q \times 1 \) vectors of the original monthly variables and the latent monthly values of the quarterly variables. Further, let \( X_{m,t} = [x'_{m,t}, \ldots, x'_{m,t-p-1}]' \) and \( X_{q,t} = [x'_{q,t}, \ldots, x'_{q,t-p-1}]' \) be the \( pm_m \times 1 \) and \( pm_q \times 1 \) vectors of current and lagged observations.
in the companion form (if \( p > 1 \)) of our TVP-MF-VAR:

\[
\begin{bmatrix}
X_{m,t} \\
X_{q,t}
\end{bmatrix} =
\begin{bmatrix}
\Phi_{mm} & \Phi_{mq} & \Phi_{mc} \\
\Phi_{qm} & \Phi_{qq} & \Phi_{qc}
\end{bmatrix}
\begin{bmatrix}
X_{m,t-1} \\
X_{q,t-1}
\end{bmatrix} +
\begin{bmatrix}
\Phi_{mc,t} \\
\Phi_{qc,t}
\end{bmatrix}
+ 
\begin{bmatrix}
U_{m,t} \\
U_{q,t}
\end{bmatrix},
\]  

(1)

or more compactly as

\[
X_t = \Phi Z_t + \Phi_c,t + U_t \quad \text{and} \quad u_t \sim N \left( 0, \Sigma_t \right).
\]

\( U_t = [U_{m,t}, U_{q,t}] = [u'_{m,t}, 0, u'_{q,t}, 0]' \) has the obvious vector-dimension for the zeros. To identify the two parts of the intercept, a sufficient condition is to set \( \Phi_{c,0} = 0 \).

To address the ragged-edge feature of macroeconomic data sets (see, e.g., Marcellino and Schumacher, 2010) the time index \( t = 1, \ldots, T_b \) denotes the “balanced” sample running until the last month for which we have all observations available, while \( T_b + 1, \ldots, T \) denotes the rest of the sample until the month for which we observe at least one variable. We also use the convention that \( t = 1, 4, 7, \ldots \) refers to the first month in a given quarter. Compared to the MF-VAR we now have time-varying intercepts and a time-varying residual variance-covariance matrix.

The reason why we separate in (1) the original monthly variables from the latent monthly series of the quarterly variables is that, for \( t = 1, \ldots, T_b \), it is sufficient to consider a “reduced” state-transition equation for the latent monthly series, \( x_{q,t} \), only. Our preferred state-vector is therefore \( S_{q,t} = [x'_{q,t}, \ldots, x'_{q,t-p}]' \) and we write the state-transition equation as

\[
S_{q,t} = \begin{bmatrix}
\phi_{qq} & 0 \\
I_{p_{mq}} & 0
\end{bmatrix} S_{q,t-1} + 
\begin{bmatrix}
\phi_{qm} \\
0
\end{bmatrix} Y_{m,t-1} + 
\begin{bmatrix}
\phi_{qc,t} \\
0
\end{bmatrix} + 
\begin{bmatrix}
u_{q,t} \\
0
\end{bmatrix},
\]  

(2)
or more compactly

\[ S_{q,t} = \Gamma_s S_{q,t-1} + \Gamma_y Y_{m,t-1} + \Gamma_{c,t} + \Gamma_u u_t, \]

in which the \( p n_m \times 1 \) vector \( Y_{m,t} \) denotes the actual observations of the original monthly observations, i.e., \( Y_{m,t} = X_{m,t} \). The zero matrices and the \( \phi \)'s have the obvious dimensions. By acknowledging that \( X_{m,t} \) is fully observed for \( t \leq T_b \) we have reduced the size of the state-vector from \( (p + 1)(n_m + n_q) \) to \( (p + 1)n_q \).⁵

The corresponding measurement equation for the monthly and quarterly series, linking observations with the model, takes the form

\[
\begin{bmatrix}
y_{m,t} \\
y_{q,t}
\end{bmatrix} =
\begin{bmatrix}
0 & \phi_{mq,1} & \phi_{mq,2} & \phi_{mq,3p} \\
1/3 I_{n_q} & 1/3 I_{n_q} & 1/3 I_{n_q} & 0
\end{bmatrix}
S_{q,t} +
\begin{bmatrix}
\phi_{mm} \\
0
\end{bmatrix}
Y_{m,t-1} +
\begin{bmatrix}
\phi_{mc,t} \\
0
\end{bmatrix} +
\begin{bmatrix}
u_{m,t}
\end{bmatrix}
\]

(3)

or more compactly, including a switching mechanism,

\[ M_t y_t = M_t [\Lambda_s S_{q,t} + \Lambda_y Y_{m,t-1} + \Lambda_{c,t} + \Lambda_u u_t], \]

where the selector matrix \( M_t \) selects both equations in (3), if we are in the last month of a quarter, or merely the first equation, if we are not. A few special features of the state-space representation (2) and (3) are worth commenting. From the first matrix on the right-hand side of (3) it is clear that the number of lags must be larger or equal than three, i.e., \( p \geq 3 \). Furthermore, the inclusion of contemporaneous states requires the state vector \( S_{q,t} \) to be of dimension \((p + 1)n_q\), for otherwise we would loose the \( p \)-th lag in the measurement equation. Moreover, the aggregation of the latent monthly series to match the quarterly observations in (3) requires to specify the quarterly variables in (log) levels. This is the natural choice from a Bayesian perspective (see, e.g., Uhlig, 1994). A unit root, for instance, is just one possibility which may receive high posterior weight depending on the evidence in the data.⁶

⁵Why we need \( p + 1 \) will become clear momentarily.

⁶Of course, what works better in terms of forecasting—levels or growth rates—is ultimately an empirical question. In a large-scale assessment of specification choices and forecast accuracy, Carriero et al.
When we move to the ragged edge of our data set, i.e., $t = T_b + 1, \ldots, T$, we switch to the larger state vector, as we also have to take care of missing observations of the original monthly variables, $x_{m,t}$. Then, for $x_t = [x'_{m,t}, x'_{q,t}]'$, the state vector becomes $S_t = [x'_t, \ldots, x'_{t-p}]'$, and the proper selector matrix $\tilde{M}_t$ in the measurement equation selects the variables and states that are actually observed over the ragged edge $t = T_b + 1, \ldots, T$.\footnote{In our specific forecast setting with GDP, or other main national accounts aggregates as the only source for quarterly variables, no quarterly observation will become available over the ragged edge, $t = T_b + 1, \ldots, T$, and $y_{q,t}$ will be an empty set. If one, however, includes quarterly survey data this may not be the case.}

Modifying the “reduced” state-space model (2) and (3) yields the “full” state-transition and measurement equations,

$$
S_t = \begin{bmatrix} \Phi & 0 \\ I_{pn} & 0 \end{bmatrix} S_{t-1} + \begin{bmatrix} \Phi_{c,t} \\ 0 \end{bmatrix} + \begin{bmatrix} U_t \\ 0 \end{bmatrix},
$$

and

$$
\tilde{M}_t \begin{bmatrix} y_{m,t} \\ y_{q,t} \end{bmatrix} = \tilde{M}_t \begin{bmatrix} I_{nm} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 I_{n_q} & 0 & 1/3 I_{n_q} & 0 & 1/3 I_{n_q} \end{bmatrix} S_t,
$$

or more compactly

$$
S_t = \Phi^* S_{t-1} + \Phi_{c,t}^* + U_t^* \quad \text{and} \quad \tilde{M}_t y_t = \tilde{M}_t^* S_t.
$$

Conditional on the parameters and variance-covariances, the measurement equations (3) and (5) are linear and have Gaussian innovations with known variance. Standard estimation algorithms, such as Carter and Kohn (1994), based on Kalman filter and smoother techniques are therefore available to elicit the latent states $S_t$. For more details we refer to the Appendix and to Schorfheide and Song (2015)."
2.2 Block II: The Time-Varying Parameter and Volatility Part

To close our model we first need to define laws of motion for the time-varying parameters in (1). In particular, let the \( n = n_m + n_q \) time-varying intercepts evolve as random walks,

\[
\phi_{c,t} = \phi_{c,t-1} + \nu_t, \quad \nu_t \sim N(0, Q). \tag{6}
\]

We have sneaked in one detail here and in (1): we separate a constant part \( \phi_c \) from the time-varying intercept \( \phi_{c,t} \). This distinction will be convenient when we form our priors and estimate the TVP-MF-VAR. We further assume that the \( n \times n \) matrix \( Q \) of innovations is diagonal, which has some merits regarding computing time and numerical stability in an otherwise automated re-estimation during each update of the forecasts—two criteria we are particularly interested in. The diagonal elements of the innovation matrix \( Q \) are inverse-Gamma variates in this case, compared to the multivariate inverse-Wishart form we obtain, if we assume \( Q \) to be a full matrix (as in Primiceri, 2005). As a consequence, the required computation for the inverse reduces from order \( O(n^2) \) to \( nO(1) \).

For the time-varying variance-covariance matrix in (1) we first take the triangular reduction of \( \Sigma_t \), defined by

\[
A^{-1}\Sigma_tA^{-1}' = H_tH_t', \tag{7}
\]

in which \( A^{-1} \) is a lower triangular matrix with ones on the diagonal and \( H_t \) is a diagonal matrix with entries \( h_{i,t} \) for \( i = 1, \ldots, n \). As is standard in TVP-VARs with stochastic volatility, we assume that \( h_t \) follows a geometric random walk, i.e.,

\[
\log h_t = \log h_{t-1} + \eta_t, \quad \eta_t \sim N(0, W). \tag{8}
\]

Following our principle of simplifying matters (without loosing too much generality) we assume the innovation matrix \( W \) to be diagonal with inverse-Gamma distributed elements.
3 Estimation

The two blocks in the preceding section can be seen as constituting the two main conditional distributions we deal with: the latent states given the parameters and the observed mixed-frequency data, and the parameters given the transformed common-frequency data based on the latent states. Formally, $S_t | \phi, \phi_c^T, \Sigma^T, Q, W, y^T$ and $\phi, \phi_{c,t}, \Sigma_t | S^T, Q, W, y^T$.

We consider two distinct numerical solution methods for the joint distribution of $S_t$, $\phi$, $\phi_{c,t}$, and $\Sigma_t$: one relying on standard Gibbs sampling, and one relying on forgetting factors (see, e.g., Koop and Korobilis, 2013) and exponentially weighted moving averages for the error variances (see Zumbach, 2006). We dub the two ways to numerically solve the model as “exact” and “approximate” method. We use the term “exact” to refer to a full Bayesian inference that specifies and derives prior and posterior distributions for all parameters and variance-covariances in the above TVP-MF-VAR (see Section 3.1).\(^8\)

The term “approximate” should indicate that we replace some distributions by forgetting factors or other shortcuts (see Section 3.2). The motivation to consider an approximate version is to decrease the computational burden without losing significant forecast performance.\(^9\) As a corollary, the approximate method could also handle large TVP-MF-VARs where full Bayesian inference becomes computationally infeasible.

The priors for the time-invariant parts of our model follow closely the ones of Schorfheide and Song (2015). As such, we are using a Minnesota prior for $\phi$ and $\Sigma$, implemented by merging dummy observations with the estimation sample. We will provide more details on prior issues, especially regarding the time-varying parts of our model, and their implications, as we lay out the workings of our “exact” and “approximate” solution methods. It suffices to say here, that the dummy variable approach to the Minnesota prior provides us with a set of initial constant-parameter estimates $\phi_0$ and $\Sigma_0$. Specifically, we have a

---

\(^8\)Of course, any numerical solution is an approximation to some degree.

\(^9\)This property would be especially desirable from an applied perspective, e.g., central banks or other institutions, where forecasts for the current analysis need to be updated within a few hours, once new data points have been released.
conjugate inverted Wishart prior of the form

\[ \text{vec} (\phi_0) | \Sigma_0 \sim N \left( \text{vec} (\phi), \Omega \otimes \Sigma_0 \right) \quad \text{and} \quad \Sigma_0 \sim W \left( \Sigma_0^{-1}, 2 \Sigma_0 \right). \]  

The priors \( \phi, \Omega, \) and \( V_\Sigma \) follow directly from a least square regression based on dummy observations (see the Appendix of Schorfheide and Song, 2015, for details). To fully specify the dummy variable approach, we take priors for the mean and variance of our data from a pre-sample, and we set the degrees of freedom in the inverse-Wishart distribution \( v_\Sigma = 2n. \) Conditional on \( \phi_0 \) and \( \Sigma_0 \), we get the prior \( S_0 \) for the reduced state vector as

\[ S_0 = N \left( S_{0|0}, \text{Var} (S_{0|0}) \right), \]

where we derive the mean and variance from running the Kalman filter based on the state-space model (2) and (3) over the pre-sample.

### 3.1 Exact Method

For the full Bayesian inference we use the following sequential MCMC sampling scheme over our two main conditional distributions. Appendix B gives a step-by-step overview of the sampler.

**Block I:** We estimate the latent states \( S_t \), that include the latent monthly values of quarterly variables and missing observations over the ragged edge, recursively using the Kalman filter and smoother as laid out in Section 2.1. Specifically, given the prior distribution (10), the conditional posterior distribution becomes

\[ S_t | \phi, \phi_c^T, \Sigma^T, Q, W, y^T \sim N \left( S_{t|T}, V_{t|T} \right), \quad t = 1, \ldots, T. \]

While, in principle, standard textbook formulas apply for \( S_{t|T} \) and \( V_{t|T} \), casting the mixed-
frequency data feature into the required form of the Kalman filter forward and backward recursions (essentially the algorithm of Carter and Kohn, 1994) takes some work. We relegate a summary of the equations to Appendix A.

**Block II:** Up to now we were always referring to two conditional posterior distributions, but drawing the time-varying intercepts and volatilities involves several conditional distributions. Besides limiting time-variation in the parameters to the intercepts, the required steps are more or less the same as in Cogley and Sargent (2005). The limitation, however, leads to an additional step in which we have to estimate all parameters other than the time-varying intercepts. So the first conditional posterior distribution is given by

\[
\phi | S^T, \phi^T, \Sigma^T, Q, W, y^T \sim N(\bar{\phi}, \nabla_\phi).
\]

(12)

Noting that \( \hat{x}_t = x_t - \phi_{c,t} \) is the regressand net of the time-varying part of the intercepts, the posterior mean and variance take the usual form, i.e.,

\[
\nabla_\phi = \left( \sum_{t=1}^{T^*} (Z_t^* \otimes I_n) \Sigma_t^{-1} (Z_t^* \otimes I_n)' \right)^{-1} \quad \text{and} \quad \bar{\phi} = \nabla_\phi \left( \sum_{t=1}^{T^*} (Z_t^* \otimes I_n) \Sigma_t^{-1} \hat{x}_t^* \right),
\]

augmented by dummy observations to impose our Minnesota prior beliefs. That is why we have added asterisks ("*") to the variables in the expressions above.

We use the same specification of the Minnesota prior as in Schorfheide and Song (2015). Hence, besides the standard hyperparameter \( \gamma \), controlling the tightness of the prior through scaling all the variances and covariances, we have two extensions aboard: the “sum-of-coefficients” and the “dummy-initial-observation” prior whose variances are controlled by the hyperparameters \( \mu \) and \( \delta \) (see also the discussion in Giannone et al., 2015). We deviate, however, from Schorfheide and Song (2015) in one minor detail: we use a fixed quadratic decay rate at which the prior variance decreases with the number of lags (see, e.g., Bañaìbura et al., 2010).\(^{11}\)

\(^{11}\)In the terminology of Schorfheide and Song (2015) our \( \gamma \), \( \mu \), and \( \delta \) coincide with their \( \gamma_1 \), \( \gamma_4 \), and \( \gamma_5 \). We fix the values corresponding to \( \gamma_2 \) and \( \gamma_3 \) to one.
The second conditional posterior distribution draws the time-varying intercepts from

\[ \phi_{c,t} | \mathbf{x}^T, \phi, \Sigma^T, Q, W, y^T \sim N(\phi_{c,t|T}, V_{\phi,t|T}) , \]  

with \( \tilde{x}_t = x_t - (Z_t \otimes I_n)' \phi \), and priors \( \phi_{c,0|0} = 0 \) and \( V_{\phi,0|0} = M_c (\Omega \otimes \Sigma_0) M_c' \). The selector matrix \( M_c \) picks the rows and columns associated with the intercepts. The posterior mean and variance then follow from the Kalman filter forward and backward recursions. For details see the Appendix of Primiceri (2005). The Kalman filter recursions involve the innovation matrix \( Q \) in the law of motion (6) for \( \phi_{c,t} \). This requires an additional conditional posterior distribution for draws of \( Q \). Since we assume \( Q \) to be diagonal we can work with inverse Gamma distributions element-by-element. Starting with the conjugate prior distribution \( Q_{i,i} = IG(v_Q, V_Q) \), we have the following posterior:

\[ Q_{i,i} = IG(v_Q, V_{i,Q}) \]  

with \( V_{i,Q} = V_Q + 0.5 \sum_{t=2}^T ((\phi_{i,t} - \phi_{i,t-1})^2) \) and \( v_Q = v_Q + 0.5 (T - 1) \) for \( i = 1, \ldots, n \).

We choose the hyperpriors in the prior distribution to be fairly diffuse and uninformative: \( V_Q = 0.00001 \) and \( v_Q = np + 2 \). The value for the degrees of freedom, however, needs to be sufficiently large to avoid the time-varying intercepts from occasionally running out of bounds (see, e.g., Primiceri, 2005). This property is especially important for a stable forecasting tool, when the model is re-estimated on an ever-expanding data set.

Finally, we have the conditional posterior distribution for the log-volatilities \( h_{i,t} \) in (8) with the associated conjugate prior distribution \( \log h_{i,t} \sim N(\log \Sigma_{0,i,i}, 10) \). The prior variance of 10 is quite large on a log-scale and gives therefore more weight on sample information. Drawing the log-volatilities follows the approach of Kim et al. (1998) in which the distribution is approximated by a mixture of normal distributions; see also Primiceri (2005) and especially the corrigendum of Del Negro and Primiceri (2013)) for more details.\(^{12}\) It suffices to say here that, given our priors at time 0, the “structural”

\(^{12}\)We implement Kim et al. (1998) using the modification of Omori et al. (2007) in which the approxi-
residuals $\tilde{e}_t = (\tilde{x}_t - \phi_{c,t}) A^{-1}$, and the state $\pi_t$ in the mixture of normals, we get

$$\log h_{i,t}|S^T, \phi_c, \phi_t^T, Q, W, \pi^T, y^T \sim N (\log h_{i,t}|T, V_{h,t}|T)$$ (15)

for $i = 1, \ldots, n$ from the usual Kalman filter forward and backward recursions. To ultimately close the sampling sequence we need to specify prior and posterior distributions for the innovations $W_{i,i}$ driving the log-volatilities in (8). Specifically, starting with $W_{i,i} = IG (v_W, V_W)$ we get

$$W_{i,i} = IG (\bar{v}_W, \bar{V}_{i,W})$$ (16)

with $\bar{V}_{i,W} = V_W + 0.5 \sum_{t=2}^{T} \left( (\log h_{i,t} - \log h_{i,t-1})^2 \right)$ and $\bar{v}_W = v_W + 0.5 (T - 1)$ for $i = 1, \ldots, n$. In order to avoid ill-behavior in the latent monthly series $x_{q,t}$ we choose a fairly informative prior: $V_W = 1$ and $v_W = 4$. Having obtained the log-volatilities, the residual variance-covariance matrix $\Sigma_t$ follows from reversing the triangular reduction in (7).

### 3.2 Approximate Method

The main goal of our approximate method is to decrease the computational burden in the time-varying parameter block, while we leave the mixed-frequency and ragged-edge block untouched. What we do is to strip down Block II such that a one-shot estimation is feasible. The one-shot conception is the essential feature of online predictions in engineering applications (see, e.g., Kulhavý and Zarrop, 1993, and Raftery et al., 2010), brought to macroeconomic forecasting with large TVP-VARs by Koop and Korobilis (2013). The required modification toward an algorithm, in which the Kalman filter laid out in Block II of Section 3.1 needs to be run only once, is to replace the steps involving draws of $Q$ by a forgetting factor, and the ones involving $h_t$ and $W$ by an exponentially weighted moving average estimator (EWMA) that is controlled by a decay factor (see, e.g., Zumbach, 2006 and Koop and Korobilis, 2013).

Before going into details, leaving Block I unaltered and implementing a one-shot so-
olution for Block II, sequential iteration over the two blocks is akin to a stochastic implementation of an expectation-maximization (EM) algorithm (see, e.g., Deloyn et al., 1999). The mixed-frequency and ragged-edge part of the TVP-MF-VAR constitutes the simulation-based E-step, and the time-varying parameter part is the M-step. The EM-algorithm has gained popularity in factor models to deal with the ragged-edge problematic (see, e.g., Stock and Watson, 2002). The great advantage of our approximate method with EM features is that the sequential iterations take only a few minutes until convergence. The disadvantage is that optimization is only locally ensured. A safeguard against getting into a local optimum would require running several chains of different starting values or evolutionary algorithms (see Jank, 2006). In every case, this would require to increase the computational demand, thereby erasing some of the advantages compared to our “exact” full Bayesian method. Another disadvantage comes from having predetermined forgetting and decay factors: any uncertainty estimate surrounding our forecasts will thus be understated.

The details that distinguish the approximate method from the exact one is best seen in the prediction step of the Kalman filter where we update the variance by

\[ V_{\phi,t|t-1} = V_{\phi,t-1|t-1} + Q. \]

If we replace this equation by

\[ V_{\phi,t|t-1} = \frac{1}{\lambda} V_{\phi,t-1|t-1}, \tag{17} \]

in which \(0 < \lambda \leq 1\) is the forgetting factor, we get rid of estimating or simulating \(Q\). Details and an interpretation of \(\lambda\) can be found in, for instance, Raftery et al. (2010). The prior specifications for \(Q\) in Cogley and Sargent (2005) and Primiceri (2005) are roughly similar to a choice of \(\lambda = 0.99\); reflecting a fairly modest amount of time variation.
In turn, our EWMA estimator for the residual variance-covariance matrix is given by

\[ \Sigma_t = \kappa \Sigma_{t-1} + (1 - \kappa) \tilde{u}_t \tilde{u}_t' \]  

(18)

with the residual \( \tilde{u}_t \) following directly from the Kalman filter prediction step: \( x_t - (Z_t \otimes I_n)' \phi - \phi_{c,t|t-1} \). For the decay factor \( \kappa \), Zumbach (2006) suggests values in the region of [0.94, 0.98]. The required initial condition \( \Sigma_0 \) comes from a pre-sample; see equation (9). To decrease the influence \( \Sigma_0 \) may have on parameter estimates and ultimately on the mixed-frequency part of our model, we also run a Kalman smoothing step for \( \Sigma_t \) and \( \phi_{c,t} \). Estimating \( \phi \) and \( \phi_{c,t} \) follows the steps of our exact method, only now we take the means in (12) and (13) instead of drawing the parameters.

In the rest of the paper we label the approximate model version with forgetting and decay factors \( \lambda \) and \( \kappa \) as TVP-MF-VAR(\( \lambda, \kappa \)).

4 Empirical Analysis

4.1 Data

We collect eight macroeconomic variables for Germany from the database of the Deutsche Bundesbank: the consumer price index (CPI), the size of the working population (WORK), the Euro InterBank Offered Rate (EURIBOR), the Euro-Dollar exchange rate (EXCH), the industrial production index (IPI), the ifo business climate index (IFO), the oil price (OIL) and real gross domestic product (GDP). All series are sampled at the monthly frequency except for GDP, which is, of course, available on a quarterly basis. CPI, the size of the working population, the industrial production index, the ifo index and GDP are seasonally adjusted as well as corrected for so-called calendar effects.\(^{13}\) Following Schorfheide and Song (2015), all indicators enter the model in log levels except for the

\(^{13}\)Due to various public holidays, e.g., the re-unification day or New Year’s Day, so-called "bridge days" often have a notable effect. A similar reasoning applies to school holidays.
EURIBOR and the exchange rate, which are left unchanged.\footnote{Following the same paper, we do not standardize the series prior to estimation.} Table 1 provides detailed information about the series under consideration.

All indicators are available as of 1991:Q1 (1991:M1). Since we downloaded the data on 7 July, 2015, the dataset is characterized by the familiar ragged-edge structure (see, e.g., Marcellino and Schumacher, 2010) due to publication delays of several variables. To be more precise, the industrial production index and the size of the working population are available until 2015:M5; for the remaining monthly indicators the June 2015 figures are already available. GDP, however, appears with a delay of about 1.5 months implying that the latest available quarterly data point is 2015:Q1.

### 4.2 Layout of the Forecast Exercise

We base the evaluation of our models on the forecast accuracy with respect to three of our seven indicators: CPI, the industrial production index and GDP. To this end we conduct a forecast exercise in pseudo-real time. More precisely, we consider an increasing sequence of estimation samples starting with the period 1995:M1-2007:M1 until 1995:M1-2015:M6.\footnote{The first four years are taken as pre-sample to generate the priors using dummy observations as in Schorfheide and Song (2015). See Section 3 for details.} Note that we always adjust the ragged-edge structure of the data set corresponding to the final period of our estimation sample. Subsequently, we compute 12 monthly forecasts for each indicator,\footnote{Here, we assume that predictions are made on a monthly basis, always on the 20\textsuperscript{th} working day. We could easily align the timing of the forecasts to the publication dates of most hard (around the 4\textsuperscript{th} working day) and soft (around the 20\textsuperscript{th} working day) indicators. For the time being, though, we stick to one forecast per month.} whereby the respective publication delay of GDP determines the first period predicted. For example, in 2007:M1 the latest available GDP observation corresponds to 2006:Q3. Hence, we would look at forecasts corresponding to the period 2006:M10-2007:M9.\footnote{Obviously, due to the smaller publication delay, some values of the monthly indicators corresponding to the respective forecast period are already available. However, as this is a byproduct of the ragged-edge data structure, it does not influence the relative forecast performance of the models under consideration.}
Using the corresponding actual values, we compute root mean squared forecast errors (RMSFEs hereafter), whereby we average monthly GDP-forecasts corresponding to a certain quarter. However, when evaluating the forecast accuracy of our models we need to take the different publication delays of our variables of interest into account (see also Schorfheide and Song, 2015). For GDP it is actually decisive which month of the quarter constitutes the end of the estimation sample: Being in the first month, the average of the first three forecasts is, in fact, a backcast as it corresponds to the previous quarter (Bańbura et al., 2011). Consequently, the next forecast represents a nowcast, whereas the 3rd and 4th values are “truly” forecasts, with quarterly horizons 1 and 2. Now, at the end of the second month the GDP-value corresponding to the previous quarter got published, causing the first figure to be a nowcast and the remaining ones to be 1-, 2- and 3-quarter-ahead forecasts. Analogously for the third month.

Considering the increasing information content as we move from one month to the next, we obtain RMSFEs for, in total, 12 different forecast horizons. Counting the amount of months between the month we are in and the beginning of the reference quarter, we can label these horizons $h_{GDP} = -4, -3, \ldots, 7$; $h_{GDP} = -4$ corresponds to the backcast made at the end of the first month, $h_{GDP} = -3$ refers to the nowcast made in the third month, and so forth, until $h_{GDP} = 7$ applies to the 3-quarter-ahead forecast made in the second month.

The situation is slightly different for CPI and the industrial production index due to their higher sampling frequency. Let us start with the industrial production index. With forecasts computed on the 20th working day of a month, we always obtain one backcast, one nowcast and up to 8-, 9- or 10-month-ahead-forecasts (depending on whether we are in the first, third or second month of a quarter, respectively). For CPI, which becomes available just when forecasts are computed, we deal with forecasts only. The relationship between forecast horizons and the end dates of the estimation period is identical to the one for the industrial production index.\(^{18}\) Obviously, we group the different back-, now-
\[^{18}\]The reason for these ”shifts” in forecast horizons is, of course, that the forecast period is determined by the publication delay of GDP. Using the same example as above, in 2007:M1 the forecast period starts
and forecasts accordingly when computing RMSFEs in our forecast exercise.

Note that the RMSFE measure of forecast accuracy is a function of the forecast error variance and the squared bias, i.e.,

\[
RMSFE = \sqrt{E(e_{t+h,j}^2)} = \sqrt{\text{Var}(e_{t+h,j}) + [E(e_{t+h,j})]^2}
\]

for each model \(j\). Given that the sample counterparts of \(E(e_{t+h,j}^2)\) and \(E(e_{t+h,j})\) correspond to the mean squared forecast error (MSFE hereafter) and the mean forecast error (MFE hereafter), we can compute a measure of the forecast error variance:

\[
\hat{\text{Var}}(e_{t+h,j}) = MSFE - MFE^2. \tag{19}
\]

By comparing this quantity to the RMSFE for each model and horizon, we can investigate in how far our main outcomes and conclusions are driven by the forecast error variance on the one hand, and the bias on the other hand (see, e.g., Carriero et al., 2015b).

Aside from the forecast exercise explained above, we present various results corresponding to the most recent data set, i.e., the one downloaded on 7 July, 2015, in detail. To be more precise, we have a close look at the evolvement of our time-varying parameters and variances, the forecasts of our variables of interest, as well as the estimates of monthly GDP.

Finally, for the TVP-MF-VAR models we have to obey the laws of motion for the time-varying parameters, given in (6) and (8), over the forecast period. For now, we follow common practice (D’Agostino et al., 2013) and assume the intercepts and volatilities to remain at their last available values. In this sense, the forecasts we compute are "no-change" iterated forecasts. But an extension of the forecast exercise with respect to alternative ways to extrapolate the time-varying parameters is left for future research.

\[\text{Marcellino et al. (2006) compare iterated and direct forecasts without finding a clear winner among the two options. However, iterated forecasts tend to do slightly better and remove the need to re-estimate the model for each horizon.}\]
4.3 Additional Input Choices

Throughout the entire analysis we use a lag length of four (months). Although smaller than the lag length of six chosen in Schorfheide and Song (2015), we consider it reasonable given the increased complexity of our approach. The grid of hyperparameters for the Minnesota prior is chosen as follows:

\[ \{0.1, 0.3, 0.5\} \otimes \{0.5, 1, 2\} \otimes \{0.5, 1, 2\} \].

For the remaining hyperparameters we choose the combination of values, which leads to a maximization of the marginal likelihood (ML or marginal data density in the terminology of Schorfheide and Song, 2015). To keep the computational burden feasible, we select the hyperparameters based on the most recent data set and the model of Schorfheide and Song (2015) only. Subsequently, we fix the ML-maximizing values for \(\gamma\), \(\mu\) and \(\delta\), and conduct the forecast exercise laid out in Section 4.2.

As explained in Section 3.1, the degree of time variation in the intercepts and the error variances is determined by the data itself when following the exact approach. When using the approximate method, though, it is captured by the forgetting factor \(\lambda\) and the decay factor \(\kappa\). In principle, we could, for each estimation sample in our forecast exercise, run over a grid of values and choose the combination, which maximizes the ML. This is what Koop and Korobilis (2013) do in their TVP-VAR model. However, our approach is computationally much more demanding due to the additional task to resolve the mixed-frequency data feature. Hence, for now, we consider four variants of the approximate model obtained by the intersection of the following values:

\[ \lambda = \{0.97, 0.99\}, \ k = \{0.94, 0.98\} \].

Note that these values cover one rapid- and one slow-change-scenario for both, the intercepts and the error variances.
In the forecast exercise we treat each of the resulting four approximate TVP-MF-VAR variants as a separate model and compare them to the exact TVP-MF-VAR model (simply labeled TVP-MF-VAR). To investigate whether allowing for time variation leads to gains in forecast accuracy, we take the model of Schorfheide and Song (2015) as a benchmark. In addition, we also consider forecasts from a Random Walk model with drift for the variables of interest such that we end up with the following seven models:

- TVP-MF-VAR(0.97,0.94)
- TVP-MF-VAR(0.99,0.94)
- TVP-MF-VAR(0.97,0.98)
- TVP-MF-VAR(0.99,0.98)
- TVP-MF-VAR (á la Schorfheide & Song)
- Random Walk with drift

For the MF-VAR and TVP-MF-VAR we generate 15,000 and 12,500 draws from the posterior distributions of the parameters (and error variances) as well as the monthly GDP observations. With a burn-in period of 10,000 draws, we use the remaining 5,000 and 2,500 to compute approximations of the respective posterior moments. As far as the approximate TVP-MF-VAR models are concerned, we found that 1,000 draws are enough to initialize the processes. Furthermore, another 1,000 draws seem to yield accurate approximations of the posterior moments, emphasizing the computational gains we are able to achieve using the approximate approach.

Following D’Agostino et al. (2013), we alter the number of draws during the forecast exercise: In the first round, i.e., for the estimation period 1995:M1-2007:M1, we set the number of draws as outlined in the preceding paragraph. From the second round on, though, we initialize the relevant quantities with the means obtained in the previous rounds.

---

20 We could also evaluate the outcomes of common-frequency VAR models estimated by OLS. Here, we would solve the mixed-frequency data feature by temporally aggregating the monthly indicators, e.g., taking the average for flow variables and the last monthly observation for stock variables (see Silvestrini and Veredas, 2008 or Marcellino, 1999). We, however, forgo the comparison with such models for two reasons: first, as argued before, we view the mixed-frequency data feature as a naturally given one, implying that a "voluntary" loss of information would be incurred when temporally aggregating the monthly series; This is not what we have in mind. Second, such a comparison is essentially what Schorfheide and Song (2015) already do when analyzing the forecasting performance of their MF-VAR. Having shown that their model generally outperforms the common-frequency one, a comparison with the MF-VAR should axiomatically suffice.
estimation and adjust the number of draws as follows: 6,000 (of which 5,000 are retained and 1,000 discarded) for the MF-VAR, 3,500 (2,500 retained, 1,000 discarded) for the TVP-MF-VAR model, and 1,500 (1,000 retained, 500 discarded) for the approximate model variants.

4.4 Results

4.4.1 Latest Available Data Set

Before investigating the forecast accuracy of our TVP-MF-VAR models and the various benchmark approaches, let us focus on a few empirical results obtained using the latest available data set. In particular, Figure 1 plots the time-varying intercepts corresponding to the eight equations in our TVP-MF-VAR and obtained using the exact approach (solid black and dotted lines), the TVP-MF-VAR(0.97,0.94) model as a representative of the approximate model group (dashed line), and the MF-VAR (solid grey line).

Focusing on the results of the exact TVP-MF-VAR, we hardly detect time variation in the intercepts, except in the equations for CPI, the size of the working population, and GDP to a lesser extent. The medians of the time-varying intercepts are sometimes above (e.g., for the ifo index) and sometimes below zero (e.g., for the oil price), whereby zero-values are usually included in the 68% confidence interval for the entire time span (except for the ifo index). Finally, none of the intercepts suggests a systematic upward or downward trend (except GDP to a slight degree).

The TVP-MF-VAR(0.97,0.94) model, on the other hand, often yields some time variation in the intercepts, which develops rather smoothly. For CPI, the size of the working population and GDP the movements are smaller than in the exact model variant, but for the remaining equations the approximate TVP-MF-VAR version implies more time

21 The values are scaled to ease readability, whereby the respective factors are: CPI - 100, WORK - 100, EURIBOR - 10, EXCH - 10, IPI - 10, IFO - 10, OIL - 10, GDP - 100.
variation. All in all, the outcomes hardly differ from the time-invariant MF-VAR, though. An exception is the size of the working population, for which the median implied by the MF-VAR lies below the one of the TVP-MF-VAR(0.97,0.94) model, which, in turn, is far below the one obtained using the exact model variant. The results, thus, point toward relatively little time variation in the intercepts of the data at hand.

Let us investigate the degree to which the data show evidence of time variation in the error variances, i.e., stochastic volatility. Figure 2 displays the diagonal elements of the time-varying covariance matrix corresponding to the equations of our eight variables. Again, the results were obtained using the exact TVP-MF-VAR, whereby the corresponding median values of the MF-VAR and TVP-MF-VAR(0.97,0.94) models are included in the same way as before.22

Here, structural instabilities are apparent for all indicators in both, the exact and the approximate model versions. Hence, it seems that, for indicators with nearly constant intercepts, time variation predominantly enters their error variances. This effect is especially marked for the oil price. Noteworthily, the medians of the exact and approximate TVP-MF-VARs move together quite closely, whereby the ones underlying the exact approach often show more pronounced rises in volatility (see, e.g., the graphs of the EURIBOR and GDP). With the exception of CPI, the medians implied by the MF-VAR roughly constitute the average of the respective time-varying quantities.

The slightly declining error variance of the size of the working population is consistent with the calming of the job market in the last decade. The fluctuations of a couple of indicators feature a large peak due to the financial crisis: the EURIBOR, the exchange rate to some extent, the industrial production, the ifo index, and GDP show sharply rising error variances around 2009. The stochastic volatility of the oil price captures familiar features of the underlying series, i.e., rising fluctuations as of 1999, a period of high, but

22The respective scaling factors for the volatilities are: CPI - 100000, WORK - 1000000, EURIBOR - 10, EXCH - 1000, IPI - 1000, IFO - 1000, OIL - 100, GDP - 10000.
relatively stable prices between 2010 and 2014, and an increase in volatility as of 2014 due to recent drops in the oil price.

As discussed briefly in the Introduction, a byproduct of the parameter-driven (Cox et al., 1981) MF-VAR model, which serves as a basis to our TVP-MF-VAR model, is the production of high-frequency, in this case monthly, GDP figures. Figure 3 plots the corresponding monthly GDP series obtained using our approach (black lines) and the MF-VAR (grey dashed line). To be more precise, the solid black lines constitute a 95% Bayesian confidence interval, whereby the dotted line inside represents the median. It turns out that the outcomes are very similar (noticing that the red dotted line is also well within the solid black lines). Naturally, the picture can be seen as a high-frequency mirror image of the corresponding quarterly GDP plot.23

[INSERT FIGURE 3 HERE]

Finally, Figure 4 shows the forecasts for our three target variables, i.e., CPI, the industrial production index and GDP. Since the latest available GDP observation corresponds to 2015:Q1, we focus, for all indicators, on the forecast period 2015:M4-2016:M3. However, due to the series’ publication delays (see Table ?? or Section 4.2) some CPI- and IPI-values already got published. This is why the graphs corresponding to both indicators start with actual values on the left hand side.

[INSERT FIGURE 4 HERE]

Note that the distribution of the CPI forecasts (nowcasts for 2015:M7) starts off quite narrow, but becomes ever-wider as we increase the forecast horizon. For the industrial production index and GDP, this feature appears slightly more moderate. Interestingly, in the lower two graphs the confidence bands feature a small “kink” when moving from backcasting to nowcasting (from 2015:M6 to 2015:M7). This indicates that estimates of the current month’s values are slightly less volatile than those of last month’s figures.

23Not displayed here for representational ease.
Let us summarize the empirical results thus far. While there seems to be relatively little time variation in the intercepts of most indicators in question, the corresponding error variances clearly indicate the presence of structural instabilities in the model. As already mentioned before, it may be that most of the time variation is simply transferred into the error variances instead of the intercept coefficients. In any case, the sometimes large differences to the outcomes of the MF-VAR, at least for the error variances, motivate a more thorough investigation of the forecast accuracy of our TVP-MF-VAR models.

4.5 Forecast Comparison

As outlined in Section 4.2, we evaluate our different models based on their forecast accuracy with respect to our three target variables, CPI, the industrial production index and GDP. The respective forecast horizons differ, depending on whether we consider one of the monthly series \((h = -1, \ldots, 10)\) or quarterly GDP \((h = -4, \ldots, 7)\). Let us also remind the reader that the time-varying coefficients remain at their latest-available values when computing forecasts.

Now, Table 2 contains the outcomes for CPI and the industrial production index, whereas Table 3 displays the ones for GDP. The figures represent RMSFES relative to the ones obtained using the Random Walk model with drift, i.e., values larger than one indicate that the benchmark model is superior to the one in question. Focusing, first, on the top half of Table 2, i.e., on forecasting CPI, it seems that neither one of the TVP-MF-VAR variants, approximate or exact, nor the MF-VAR outperform the benchmark model. The Random Walk model with drift even appears to dominate many of the more evolved models for larger horizons (say for \(h > 5\)). Furthermore, note that while the exact TVP-MF-VAR performs equally well as the best approximate model variant for \(h < 6\), it clearly outperforms them for larger horizons. A similar observation can be made when considering the outcomes of the time-invariant MF-VAR, whereby the degree to which the exact TVP-MF-VAR dominates for \(h > 5\) is much smaller.

The bottom half of Table 2, i.e., the outcomes for the industrial production index,
reveals that both, the MF-VAR and some of the TVP-MF-VAR versions, beat the Random Walk model with drift. The gains in forecast accuracy diminish, however, for larger $h$. The exact TVP-MF-VAR outperforms its approximate counterparts with $\kappa = 0.94$, but not the ones with $\kappa = 0.98$, which, as far as TVP-MF-VAR models are concerned, lead to the best results for the industrial production index. Hence, the relative RMSFEs of the approximate TVP-MF-VAR models point toward a moderate degree of time variation in the volatilities leading to better forecasts. Especially for back- and nowcasting ($h = -1, 0$), it seems as if the performance of the exact model version lies in-between those of the two approximate model groups, TVP-MF-VAR($\cdot, 0.94$) and TVP-MF-VAR($\cdot, 0.98$). For forecasting ($h > 0$), its performance is generally comparable to the best approximate model variant, though. Nevertheless, none of the TVP-MF-VAR models is able to beat the MF-VAR, which clearly yields the overall most accurate forecasts of the industrial production index.

Turning to Table 3 and the results for predicting quarterly GDP, we conclude that all models (almost) always outperform the Random Walk model with drift. Most importantly, the exact and two of the approximate TVP-MF-VARs (the ones with $\kappa = 0.98$) very often lead to clearly superior forecasts than the MF-VAR. Consequently, adding the possibility of time variation in the intercepts and the volatilities seems to be very beneficial for forecasting in this case. Given the gains achievable and the importance of the target variable in question, this is a remarkable result.

Once again the outcomes of the exact TVP-MF-VAR version are similar to the ones underlying the TVP-MF-VAR($\cdot, 0.98$) model; sometimes it yields the overall lowest RMSFE (e.g., for $h = -4$), sometimes it lies in-between the two approximate versions (e.g., for $h = 0$), and sometimes it performs slightly worse (e.g., for $h = 1$). Overall, the best outcomes are obtained using TVP-MF-VAR(0.99,0.98), which implies that setting $\lambda$ and $\kappa$ in this way covers the degree of time variation rather well in this case. Hence, the simplicity-complexity trade-off seems to tip in favor of the former here. Note, though,

24The exceptions occurring for the TVP-MF-VAR($\cdot, 0.94$) and $h > 4$, as well as for TVP-MF-VAR and $h = 7$.  

26
that the exact model variant is most likely dominating its competitors when the “optimal” $\lambda$- and $\kappa$-values fall outside the grid of values we consider. $\lambda = 0.99$ and $\kappa = 0.98$ yielding the most accurate GDP forecasts shows, once again, that a mild level of time variation leads to the best forecast results. In light of the discussion in Section 4.4.1 and Figures 1 and 2, i.e., the rather small amount of time variation visible for the latest available data set, these results have a confirmatory touch. Consequently, a situation, in which smaller $\lambda$- and $\kappa$-values prove beneficial for forecasting, i.e., dealing with data that contain more structural instabilities, should lead to the TVP-MF-VAR models dominating the time-invariant MF-VAR even more clearly.

Comparing Tables 2 and 3 with Tables 4 and 5, i.e., investigating the contribution of variance (and, by formula (19), bias) to RMSFEs, it emerges that the forecast accuracy results are mainly driven by the variance of forecast errors for smaller horizons (roughly $h < 4$ for CPI and the industrial production index, and $h < 0$ for GDP), whereas the bias has a greater effect for larger ones. For the two monthly variables this effect goes in opposite directions: relative forecast error standard deviations increase with $h$ in comparison to relative RMSFEs for CPI, while they decrease with $h$ for the industrial production index. In other words, bias decreases with $h$ for CPI and increases with $h$ for the industrial production index. The outcomes for GDP are similar to the ones for the industrial production index. Interestingly, for all target variables, the influence of the bias appears weaker for the exact TVP-MF-VAR and the MF-VAR than for the approximate TVP-MF-VAR model variants.

5 Conclusion

In this paper, we have introduced time-varying mixed-frequency VAR models for macroeconomic forecasting purposes, whereby we restrict time variation to the intercept coefficients and error variances to keep the approach computationally feasible for systems of

\footnote{This is because under the exact model variant, we let the data decide upon the degree of time variation.}
moderate size. We have shown how estimation of the model requires the separation into two separate blocks; one, in which the mixed-frequency and ragged-edge features of the data are handled, and another, in which the parameters get estimated. With respect to the first block we rely on the approach of Schorfheide and Song (2015), which served as main competitor for our method. In the second block we follow two different routes; one based on exact Bayesian inference, and an approximate approach based on forgetting and decay factors (see Koop and Korobilis, 2013). While the exact version lets the data determine the degree to which time variation applies in the case at hand, it is predetermined by the respective factors in the approximate counterpart. We also discussed the inherent simplicity-complexity trade-off the user faces in this situation.

We demonstrated the feasibility and usefulness of our approach within an empirical analysis involving eight indicators of the German macroeconomy. Not only did we show how to employ our models for the last available data, i.e., in real time, we also investigated their forecast accuracies, with respect to three target variables, in a pseudo-real time forecast exercise. Our results indicate that, even in the presence of only mild structural instabilities, forecast gains are possible by allowing time variation to enter a mixed-frequency system. It is to be expected that the merits of including time-varying parameters and stochastic volatility are even more pronounced for data sets, that contain more evidence for structural instabilities. Hence, applying our methodology to such an alternative data set should be the first step in extending the results presented here. Additionally, this analysis could be the starting point toward TVP-MF-VAR models suitable for systems of larger size (along the lines of Koop and Korobilis, 2013 or Banbura et al., 2010).
References


A Details of Block I

Given $S_{0|0} = S_0$ and $V_{0|0} = \text{Var}(S_{0|0})$ from (10) the forward pass of the Kalman filter is given by the following system of equations:

The mixed-frequency part $t = 1, \ldots, T_b$

\[
S_{t|t-1} = \Gamma_s S_{t-1|t-1} + \Gamma_y y_{m,t-1} + \Gamma_c, \\
V_{t|t-1} = \Gamma_s V_{t-1|t-1} + \Gamma_u \Sigma_{qq,t} \Gamma_u', \\
K_{f,t} = (V_{t|t-1} (M_t \Lambda_s)' + \Gamma_u \Sigma_{mq,t} (M_t \Lambda_u)') \left( (M_t \Lambda_s) V_{t|t-1} \Lambda_s' + \Lambda_u \Sigma_{mm,t} \Lambda_u' \right)^{-1}, \\
\hat{y}_t = (M_t \Lambda_s) S_{t|t-1} + (M_t \Lambda_y) y_{m,t-1} + (M_t \Lambda_c), \\
S_{t|t} = S_{t|t-1} + K_{f,t} (M_t y_t - \hat{y}_t), \\
V_{t|t} = V_{t|t-1} - K_{f,t} (\tilde{M}_t^* V_{t|t-1}).
\]

The ragged-edge part $t = T_b + 1, \ldots, T$

\[
S_{t|t-1} = \Phi^* S_{t-1|t-1} + \Phi^* e, \\
V_{t|t-1} = \Phi^* V_{t-1|t-1} \Phi^* + [I_n, 0]' \Sigma_e, \\
K_{f,t} = (V_{t|t-1} \tilde{M}_t^* ) \left( \tilde{M}_t^* V_{t|t-1} \tilde{M}_t^* \right)^{-1}, \\
\hat{y}_t = \tilde{M}_t^* S_{t|t-1}, \\
S_{t|t} = S_{t|t-1} + K_{f,t} (\tilde{M}_t y_t - \hat{y}_t), \\
V_{t|t} = V_{t|t-1} - K_{f,t} (\tilde{M}_t^* V_{t|t-1}).
\]

Once we arrive at the end of the forward recursions we draw $S_T \sim N(S_{T|T}, V_{T|T})$ and run the backward pass from $T - 1$ until time zero. At each recursion we then draw a new set of states, i.e. $S_t \sim N(S_{t|t+1}, V_{t|t+1})$. Note that during the backward pass we no longer need to check whether we are in the last month of a quarter or where there are missing values over the ragged edge.
The ragged-edge part $t = T - 1, T - 2, \ldots, T_b$

\begin{align*}
S_{t+1|t} &= \Phi^* S_{t|t} + \Phi^* c, \\
V_{t+1|t} &= \Phi^* V_{t|t} \Phi^*, \\
S_{t|t+1} &= S_{t|t} + V_{t|t} \Phi^* V_{t+1|t}^{-1} (S_{t+1} - S_{t+1|t}), \\
V_{t|t+1} &= V_{t|t} - V_{t|t} \Phi^* V_{t+1|t}^{-1} \Phi^* V_{t|t}.
\end{align*}

The mixed-frequency part $t = T_b - 1, T_b - 2, \ldots, 0$

\begin{align*}
S_{t+1|t} &= \Gamma_s S_{t|t} + \Gamma_y y_{m,t} \Gamma_{c,t}, \\
V_{t+1|t} &= \Gamma_s V_{t|t} \Gamma_s + \Gamma_u \Sigma_q \Gamma_{u'}, \\
S_{t|t+1} &= S_{t|t} + V_{t|t} \Gamma_s V_{t+1|t}^{-1} (S_{t+1} - S_{t+1|t}), \\
V_{t|t+1} &= V_{t|t} - V_{t|t} \Gamma_s V_{t+1|t}^{-1} \Gamma_s V_{t|t}.
\end{align*}

B Summary of the Steps in the MCMC sampler

The MCMC sampler of our exact method in Section 3.1 takes the following form:

0. Initialize $\phi, \phi^T_c, \Sigma^T, S_{0|0}, \text{Var}(S_{0|0}), Q,$ and $W$. Set $j = 1$.

1. Sample $S^T_{(j)}$ from $p \left( S^T_{(j)} | \phi_{(j-1)}, \phi^T_{c,(j-1)}, \Sigma^T_{(j-1)}, Q_{(j-1)}, W_{(j-1)}, y^T \right)$.

2. Sample $\phi_{(j)}$ from $p \left( \phi_{(j)} | S^T_{(j)}, \phi^T_{c,(j-1)}, \Sigma^T_{(j-1)}, Q_{(j-1)}, W_{(j-1)}, y^T \right)$.

3. Sample $\phi^T_{c,(j)}$ from $p \left( \phi^T_{c,(j)} | S^T_{(j)}, \phi_{(j)}, \Sigma^T_{(j-1)}, Q_{(j-1)}, W_{(j-1)}, y^T \right)$.

4. Sample $Q_{(j)}$ from $p \left( Q | S^T_{(j)}, \phi_{(j)}, \phi^T_{c,(j)}, \Sigma^T_{(j-1)}, W_{(j-1)}, y^T \right)$.

5. Sample $\pi^T_{(j)}$ from $p \left( \pi^T_{(j)} | S^T_{(j)}, \phi_{(j)}, \phi^T_{c,(j)}, \Sigma^T_{(j-1)}, Q_{(j)}, W_{(j-1)}, y^T \right)$.

6. Sample $\Sigma^T_{(j)}$ from $p \left( \Sigma^T_{(j)} | S^T_{(j)}, \phi_{(j)}, \phi^T_{c,(j)}, Q_{(j)}, W_{(j-1)}, \pi^T_{(j)}, y^T \right)$.

7. Sample $W_{(j)}$ from $p \left( Q | S^T_{(j)}, \phi_{(j)}, \phi^T_{c,(j)}, \Sigma^T_{(j)}, Q_{(j)}, y^T \right)$.

8. Increment $j$ by one and repeat steps 1 to 8 until desired number of draws are saved.
C Figures

Figure 1: Time-Varying-Parameters

Note: For the time period 1995:M1-2015:M6, the medians of the time-varying intercepts underlying the exact TVP-MF-VAR are plotted as solid black lines (16th and 84th percentiles, implying a 68% confidence interval, as dotted lines). The corresponding medians of the TVP-MF-VAR(0.97,0.94) and the MF-VAR are shown as dashed black and solid grey lines.
Figure 2: Volatilities

Note: This figure contains the same information as figure 1, but for the medians of the time-varying diagonal elements of the covariance matrix.
Figure 3: Monthly GDP

Note: This figure displays, for the time period 1995:M1-2015:M3, the monthly GDP values obtained using the exact TVP-MF-VAR (black lines) and the MF-VAR (grey dotted line). The black dotted line represents the median values, whereas the solid black lines cover a 95% confidence interval.
Figure 4: Forecasts CPI, IPI and GDP

Note: This figure plots, for CPI, the industrial production index and GDP and the period 2015:M4-2016:M3, the medians of the forecasts computed using the exact TVP-MF-VAR as solid black lines (16th and 84th percentiles, implying a 68% confidence interval, as dotted lines).
## D Tables

Table 1: Data & Stylized Release Calendar Overview

<table>
<thead>
<tr>
<th>Series ID</th>
<th>Transf.</th>
<th>Publ. Day</th>
<th>Publ. Lag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPI</td>
<td>Logarithm</td>
<td>20</td>
<td>./.</td>
<td>Consumer price index, 2010=100, seasonally and calendar adjusted</td>
</tr>
<tr>
<td>WORK</td>
<td>Logarithm</td>
<td>22</td>
<td>1 Month</td>
<td>Absolute number of people employed according to ESVG 2010, in thousands, seasonally and calendar adjusted</td>
</tr>
<tr>
<td>EURIBOR</td>
<td>./.</td>
<td>22</td>
<td>./.</td>
<td>Euro InterBank Offered Rate, three months, monthly average, in percentages (before 1999: German three-month interest rate)</td>
</tr>
<tr>
<td>EXCH</td>
<td>./.</td>
<td>22</td>
<td>./.</td>
<td>Euro-USD exchange rate (before 1999: Deutsche Mark-USD exchange rate)</td>
</tr>
<tr>
<td>IPI</td>
<td>Logarithm</td>
<td>4</td>
<td>2 Month</td>
<td>Industrial production index, 2010=100, seasonally and calendar adjusted</td>
</tr>
<tr>
<td>IFO</td>
<td>Logarithm</td>
<td>17</td>
<td>./.</td>
<td>ifo business climate index, manufacturing industry, seasonally and calendar adjusted</td>
</tr>
<tr>
<td>OIL</td>
<td>Logarithm</td>
<td>22</td>
<td>./.</td>
<td>Price (in EUR) per barrel of crude oil, Brent</td>
</tr>
<tr>
<td>GDP</td>
<td>Logarithm</td>
<td>30*</td>
<td>1 Quarter</td>
<td>Gross domestic product, in chained prices of previous year, in billion Euros, seasonally and calendar adjusted</td>
</tr>
</tbody>
</table>

Note: All indicators are sampled at monthly frequency except for GDP, which is available on a quarterly basis. The table contains information about transformations applied to the variables (Transf.), the working day within each month (*-quarter) on which the respective indicator gets published (Publ. Day), the publication delay or lag (Publ. Lag), and a description of each series (Description). Within our stylized release calendar framework we assume each month to contain exactly 22 working days.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Horizon (in months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>MF-VAR</td>
<td>0.99</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>MF-VAR</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Note: This table displays, for forecasts of CPI and the industrial production index, the RMSFEs of our four approximate TVP-MF-VAR model variants, the exact TVP-MF-VAR model and the MF-VAR of Schorfheide and Song (2015) relative to the ones of the Random Walk model with drift. The estimation samples underlying the analysis go from 1995:M1-2007:M1 to 1995:M1-2015:M6.
Table 3: RMSFEs relative to the Random Walk model with drift; GDP

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Horizon (months between current one and start of reference quarter)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-4</td>
</tr>
<tr>
<td>GDP</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>MF-VAR</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Note: This table contains the same information as Table 2, but for forecasts of GDP. The horizons represent the amount of months between the current one and the start of the reference quarter (see Section 4.2).
Table 4: Forecast error standard deviation relative to the Random Walk model with drift; CPI and IPI

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Horizon (in months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>1.02</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>1.02</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>1.02</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>1.02</td>
</tr>
<tr>
<td>CPI</td>
<td>TVP-MF-VAR</td>
<td>1.04</td>
</tr>
<tr>
<td>CPI</td>
<td>MF-VAR</td>
<td>1.01</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>1.14</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>1.13</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>0.97</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>0.96</td>
</tr>
<tr>
<td>IPI</td>
<td>TVP-MF-VAR</td>
<td>1.07</td>
</tr>
<tr>
<td>IPI</td>
<td>MF-VAR</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Note: This table displays, for forecasts of CPI and the industrial production index, the square roots of the forecast error variances, i.e., the square root of the difference between the MSFE and the squared MFE (average bias), of our (TVP-)MF-VAR models as ratios of the benchmark. For the rest see the notes to Table 2.
Table 5: Forecast error standard deviation relative to the Random Walk model with drift; GDP

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Horizon (months between current one and start of reference quarter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>TVP-MF-VAR(0.97,0.94)</td>
<td>0.66 0.77 0.79 0.68 0.71 0.75 0.71 0.71 0.73 0.77 0.75 0.78</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.94)</td>
<td>0.70 0.75 0.79 0.66 0.74 0.77 0.68 0.74 0.76 0.76 0.78 0.80</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.97,0.98)</td>
<td>0.56 0.65 0.83 0.71 0.75 0.82 0.76 0.78 0.79 0.74 0.77 0.76</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR(0.99,0.98)</td>
<td>0.54 0.63 0.81 0.70 0.75 0.82 0.76 0.78 0.78 0.75 0.78 0.77</td>
</tr>
<tr>
<td></td>
<td>TVP-MF-VAR</td>
<td>0.52 0.65 0.88 0.71 0.76 0.86 0.79 0.84 0.87 0.84 0.89 0.91</td>
</tr>
<tr>
<td></td>
<td>MF-VAR</td>
<td>0.67 0.74 0.86 0.76 0.80 0.86 0.83 0.87 0.88 0.87 0.92 0.92</td>
</tr>
</tbody>
</table>

Note: This table contains the same information as Table 4, but for forecasts of GDP. For the interpretation of the horizons see the notes to Table 3 or Section 4.2.