

Discussion Paper

Deutsche Bundesbank
No 40/2018

**Large mixed-frequency VARs with
a parsimonious time-varying parameter structure**

Thomas B. Götz
Klemens Hauzenberger

Editorial Board:

Daniel Foos
Thomas Kick
Malte Knüppel
Jochen Mankart
Christoph Memmel
Panagiota Tzamourani

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet <http://www.bundesbank.de>

Reproduction permitted only if source is stated.

ISBN 978-3-95729-508-8 (Printversion)

ISBN 978-3-95729-509-5 (Internetversion)

Non-technical summary

Research Question

In macroeconomic forecasting one is often confronted with the task of predicting a few target variables using a range of indicators. Existing modeling approaches usually disregard at least one of the following three features: the possible presence of (smooth) structural changes, mixing publication frequencies and delays, or the joint dynamics between the target and indicator variables. This leads to the following questions: Can we develop a model that accounts for all of these features simultaneously? If so, does it deliver competitive short-term forecasts of, say, GDP, or does computational complexity erode any potential gains?

Contribution

We introduce a time-varying mixed-frequency vector autoregressive model that can handle various desirable features at the same time. In order to be able to include a comparably large number of variables within such a computationally intensive approach, we introduce time variation parsimoniously: only the intercepts are allowed to vary over time and time variation in the error variances (stochastic volatility) is restricted to be common across all variables. This novel econometric modeling approach enables researchers to investigate a system of comparably large size, that accounts for mixed frequencies, structural changes and the joint dynamics of its variables simultaneously.

Results

Compared to a benchmark model, that disregards time variation altogether, the restrictions with respect to time variation enable us to estimate our model in a fairly small amount of time. In an empirical application with eleven U.S. macroeconomic variables the results support the presence of time variation in the data, both in terms of time-varying intercepts and common stochastic volatility. An out-of-sample forecast exercise with three selected target variables (including GDP) reveals that – compared to the benchmark model – our approach yields considerable gains with respect to both point and density forecasts.

Nichttechnische Zusammenfassung

Fragestellung

Bei der Erstellung makroökonomischer Prognosen gilt es oftmals, einige wenige Zielvariablen mit Hilfe einer Reihe von Indikatoren fortzuschreiben. Dabei vernachlässigen existierende Prognosemodelle in der Regel mindestens eines der folgenden drei Charakteristika: (graduelle) Strukturbrüche, gemischte Publikationsfrequenzen und -verzögerungen sowie potenzielle Wechselwirkungen zwischen den Zielgrößen und den Indikatoren. Dies wirft die folgenden Fragen auf: Ist es möglich, ein Prognosemodell zu entwickeln, das all jene Charakteristika gleichzeitig berücksichtigt? Falls ja, kann das Modell wettbewerbsfähige Kurzfristprognosen, z.B. des BIP, liefern oder überlagert die Komplexität des Modells potenzielle Vorteile?

Beitrag

Wir stellen ein zeitvariierendes vektor-autoregressives Modell mit gemischten Frequenzen vor, das alle zuvor genannten Charakteristika simultan berücksichtigen kann. Um eine vergleichsweise große Anzahl von Variablen innerhalb eines solch rechenaufwendigen Ansatzes betrachten zu können, modellieren wir Zeitvariationen in sparsamer Weise. So erlauben wir lediglich den konditionalen Mittelwerten (d.h. den Konstanten in jeder Gleichung des Systems) sich über die Zeit zu verändern; ferner wird die Zeitvariation in den Varianzen der Fehler (stochastische Volatilität) auf einen Wert für alle Variablen restringiert. Dieser neuartige Modellierungsansatz erlaubt es dem Anwender, Systeme vergleichsweise großer Dimension, die gemischte Frequenzen, Strukturbrüche und Wechselwirkungen zwischen den Variablen zulassen, zu untersuchen.

Ergebnisse

Verglichen mit einem Modell, welches Zeitvariation gänzlich ausschließt, lässt sich unser Modell aufgrund der Restriktionen in Bezug auf Zeitvariation recht schnell schätzen. In einer empirischen Untersuchung mit elf makroökonomischen Variablen für die USA weisen die Resultate auf Zeitvariation in den Daten hin, sowohl bei den konditionalen Mittelwerten als auch bei der gemeinsamen stochastischen Volatilität. Eine empirische Prognosestudie mit drei ausgewählten Zielgrößen (darunter das BIP) ergibt, dass unsere Methode gegenüber dem Vergleichsmodell zu nennenswerten Verbesserungen in Bezug auf Punkt- und Dichteprognosen führt.

LARGE MIXED-FREQUENCY VARs WITH A PARSIMONIOUS TIME-VARYING PARAMETER STRUCTURE*

Thomas B. Götz[†] and Klemens Hauzenberger[‡]

Deutsche Bundesbank

Abstract

To simultaneously consider mixed-frequency time series, their joint dynamics, and possible structural changes, we introduce a time-varying parameter mixed-frequency VAR. To keep our approach from becoming too complex, we implement time variation parsimoniously: only the intercepts and a common factor in the error variances vary over time. We can therefore estimate moderately large systems in a reasonable amount of time, which makes our modifications appealing for practical use. For eleven U.S. variables, we examine the performance of our model and compare the results to the time-constant MF-VAR of [Schorfheide and Song \(2015\)](#). Our results demonstrate the feasibility and usefulness of our method.

JEL Codes: C32, C51, C53

JEL Keywords: Mixed Frequencies, Time-Varying Intercepts, Common Stochastic Volatility, Bayesian VAR, Forecasting

*We thank Todd Clark, Michael Funke, Josef Hollmayr, Dimitris Korobilis, Boriss Siliverstovs, Mike West, and seminar participants at the 2015 NBP Workshop on Forecasting, the 2015 DIW Macroeconomic Workshop, the 2015 CFE, the KOF Swiss Economic Institute, the 2016 SNDE, and the 2018 ECB Workshop on how to treat trends in macro-econometrics for valuable comments and discussions. The views expressed in this paper are solely ours and should not be interpreted as reflecting the views of the Deutsche Bundesbank or the Eurosystem.

[†]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, Monetary Policy Division, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main. Email: klemens.hauzenberger@bundesbank.de

1 Introduction

The biblical story of Joseph and the Pharaoh about seven years of plenty followed by seven lean years was, perhaps, the first accurate and documented forecast. Of course, modern economic forecasting is no longer about interpreting dreams, but similar in vein with today's practice is the attempt to translate available information from all kinds of sources into forecasts that are as accurate as possible.

Often, a few quarterly and monthly target variables, such as the gross domestic product (GDP) or the unemployment rate, are at the center of interest. For all these target variables we can then select between a large number of potentially useful indicators, such as surveys and financial market data, released earlier than the actual target variable. Given these two groups of observations and the different release patterns we develop a mixed-frequency model that accounts for possible (smooth) structural changes, can handle a moderately large number of variables, and, as we write the model as a VAR, includes the joint dynamics between the target variables and the indicators. Many of the existing forecasting models ignore at least one of these features. For instance, the mixed-frequency VAR of [Schorfheide and Song \(2015\)](#) excludes time variation and, as such, cannot deal with structural changes; the large time-varying parameter VAR of [Koop and Korobilis \(2013\)](#) requires all variables to be transformed into one common frequency; and the univariate Markov-switching MI(xed) DA(ta) S(ampling) model of [Guèrin and Marcellino \(2013\)](#) ignores the joint dynamics between the targets and indicators.

The main contribution of our paper, then, is to address all these features simultaneously by constructing a (moderately) large time-varying mixed-frequency VAR. 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 macroeconomic forecasting. The benefits of modeling different frequencies have been shown throughout the respective field of the literature (see inter alia [Andreou et al., 2013](#), [Forni et al., 2015](#) or [Götz et al., 2016](#)). But besides the permanent hunting for lower

forecast errors, a mixed-frequency model gives the economist an invaluable tool for the current analysis, with updated forecasts of the lower-frequency variable whenever new releases of indicators flow in during a month. Modeling of time variation, however, and its potential usefulness for forecasting is not so clear a-priori. Increased computational complexity, and with that numerical instabilities and pitfalls, can quickly erode any gains. The usefulness for forecasting, especially for large models, may thus depend on the way one introduces time variation (see, e.g., [D’Agostino et al., 2013](#), [Eickmeier et al., 2015](#) and [Barnett et al., 2014](#)).

We take the Bayesian mixed-frequency VAR (MF-VAR hereafter) of [Schorfheide and Song \(2015\)](#) as the starting point. Their model is sampled at the higher frequency, implying that latent high-frequency observations of the low-frequency series, for instance GDP, need to be estimated on top of the actual VAR parameters, time varying or not. Given already our concerns that computational complexity may limit the usefulness of models for forecasting, we focus on implementing time variation in a parsimonious way. Essentially, we let only the intercepts vary over time and restrict time-variation in the error-variances (stochastic volatility) to be common across all variables as in [Carriero et al. \(2015b\)](#). Taken together, modeling time-varying intercepts and common stochastic volatility may overcome possible losses in accuracy that could result from ignoring one or the other effect. Time-varying intercepts could capture variable-specific effects that would otherwise distort the estimation of the common volatility factor. The obvious benefit of simplifying our model along these dimensions lies in the reduced complexity, which, in turn, allows us to estimate moderately large systems in a reasonable amount of time. We further improve upon computational efficiency by using sparse and block-banded matrix algebra as, for instance, in [Chan and Jeliazkov \(2009\)](#) and [Chan \(2018\)](#). Even when computing power is available in abundance, running time is still an issue from a practical point of view. Especially so when many models need to be updated several times during a month, as is the case in the regular forecasting process at central banks and other institutions.

In our application we consider a VAR with the same eleven U.S. macroeconomic vari-

ables as in [Schorfheide and Song \(2015\)](#): three at quarterly frequency, including GDP, investment and government spending, and eight monthly variables, including the consumer price index, the federal funds rate and others. Besides our main model with time-varying intercepts and common stochastic volatility we also run intermediate versions, turning off either one of the two features. We find that, compared to [Schorfheide and Song \(2015\)](#), considerable gains are possible for point and density forecasts, especially so for our main model. For the intermediate versions, the time-varying intercept seems to be somewhat more beneficial for the forecast performance than the common stochastic volatility factor.

1.1 Motivating Experiment

Before formally introducing our model, we provide an example to motivate that, compared to a fully fledged time-varying model (as in [Primiceri, 2005](#)) with mixed frequencies, limiting time variation to the intercepts and a common stochastic volatility factor has indeed some merits. First, the amount of time-varying parameters and as such the computational burden decreases tremendously: only n instead of $n + n^2p$ VAR and one instead of $n(n + 1)/2$ variance-covariance parameters need to be estimated (n and p are size and lag order of the VAR). [Table 1](#) illustrates the computational gains as functions of various combinations of n and p (combinations of large values of n and p are computationally not feasible). While we already save more than 50% running time in a three-variable VAR with three lags, the gains go up to almost 90% in a VAR with six variables. Note that increasing the lag length seems comparably less harmful than enlarging the number of variables.

Second, if the forecast accuracy of, say, GDP as one of the main target variables deteriorates compared to the fully-fledged model, any gains in terms of computing time and tractability would not warrant our parsimonious modeling approach. We therefore conduct a quick forecast exercise using a three-variable subversion (GDP, CPI, Fed funds rate) of the larger model we will present later. [Table 2](#) shows root mean squared forecast errors (RMSFE) and average log predictive likelihoods (logPL) for GDP of our parsi-

monious model relative to the fully-fledged model. RMSFE ratios larger than one and logPLs smaller than zero would imply that our parsimonious modeling approach may not be warranted empirically.

The only case, however, in which the fully-fledged model seems to have an edge over ours, is the first density forecast at horizon -1 (i.e., a backcast actually). In all other instances forecast accuracy improves.

1.2 Related Literature

An alternative to the so-called parameter-driven (Cox et al., 1981) approach we are using to model the mixed frequencies is Ghysels (2016). Their approach depends exclusively on observable data.¹ The VAR is stacked with the low frequency and stacks the corresponding high- (and low-) frequency observations into a single vector.² Hence, in contrast to the MF-VAR, the model of Ghysels (2016) does not deliver, for instance, a monthly GDP series that might be a desirable byproduct.

Closer related to our work are three other simultaneously and independently written papers. One is by Cimadomo and D’Agostino (2016) who develop a small structural time-varying mixed-frequency VAR model for estimating the effects of government spending shocks. Their paper has to be praised for extending the estimation of a fully fledged time-varying parameter model, such as Primiceri (2005), to the mixed-frequency case. This generality, however, does not come without its costs and restricts their model to include only a few variables and lags. Antolin-Diaz et al. (2017), the second paper, use a dynamic factor model (DFM) with time-varying intercepts and stochastic volatility to capture long-run shifts in the variables of interest with a focus on the drivers of the slowdown in

¹The observation-driven 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 Clements and Galvão, 2008 or Götz et al., 2016, 2014 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).

²The augmentation of the observation-driven mixed-frequency VAR to allow for time variation, and a comparison with the model presented here, is beyond the scope of this paper.

real U.S. GDP growth. The third one is [Banbura and van Vlodrop \(2018\)](#). Although their time-varying VAR model includes only a single frequency it shares a common feature with our model: time variation is restricted to certain parameters. Unlike us, however, they model time variation in the unconditional mean rather than the intercept (i.e., conditional mean). While this complicates the estimation (see, e.g., [Villani, 2009](#)), it is typically easier to elicit informative priors for the unconditional mean. [Banbura and van Vlodrop \(2018\)](#) take even one promising step further and link the unconditional mean to long-run Consensus forecasts.

2 The Modeling Framework

Our framework is a combination of different methods and modeling approaches. We take the mixed-frequency VAR of [Schorfheide and Song \(2015\)](#) as the core and combine it with elements from the time-varying parameter (e.g., [Primiceri, 2005](#)) and the common stochastic volatility (e.g., [Carriero et al., 2015b](#)) literature.

We use some modeling and notational conventions throughout the paper. We refer to Block I as the part of the model that deals with the mixed-frequency and, related to that, the ragged-edge problem. Block II, then, covers modeling and estimating the (time-varying) VAR parameters. As implied by the approach of [Schorfheide and Song \(2015\)](#), we write the entire model in the higher, monthly frequency.³ An expression z^τ denotes the complete history of a generic variable (or parameter) z_t up to time τ , i.e., $z^\tau = [z'_1, \dots, z'_\tau]'$. We reserve the letter y^T (and Y^T) for the incomplete data matrix which contains missing values either through the mismatch in sampling frequency or because of variable-specific publication lags (i.e., ragged edge); x^T (and X^T), then, denotes the corresponding complete data matrix that includes estimates of the latent monthly equivalents of the quarterly variables and the missing values over the ragged edge.

³An application of our approach to different frequency pairs, e.g., month/week or year/quarter, follows straightforwardly.

2.1 Block I: The Mixed-Frequency Problem

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 $X_{m,t} = [x'_{m,t}, \dots, x'_{m,t-p+1}]'$ and $X_{q,t} = [x'_{q,t}, \dots, x'_{q,t-p+1}]'$ be the $pn_m \times 1$ and $pn_q \times 1$ vectors of current and p lagged observations. We can then write our model in a way where the distinction between the two frequencies becomes apparent, i.e.,

$$\begin{bmatrix} x_{m,t} \\ x_{q,t} \end{bmatrix} = \begin{bmatrix} \Phi_{mm} & \Phi_{mq} & \Phi_{mc} \\ \Phi_{qm} & \Phi_{qq} & \Phi_{qc} \end{bmatrix} \begin{bmatrix} X_{m,t-1} \\ X_{q,t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} \Phi_{mc,t} \\ \Phi_{qc,t} \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ u_{q,t} \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} u_{m,t} \\ u_{q,t} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{mm,t} & \Sigma_{mq,t} \\ \Sigma_{qm,t} & \Sigma_{qq,t} \end{bmatrix} \right).$$

We deliberately breakup the intercepts into a constant and time-varying part, imposing the initial condition $\Phi_{mc,0} = \Phi_{qc,0} = 0$. As it will become clear at the estimation stage, breaking up the intercepts in this way proves convenient.

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

The reason why we separate in (1) the original monthly variables from the latent monthly series of the quarterly variables is that, for $t = 1, \dots, T_b$, 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 $S_t = [x'_{q,t}, \dots, x'_{q,t-p}]'$ and we write the state-

transition equation as

$$S_t = \begin{bmatrix} \Phi_{qq} & 0 \\ I_{pn_q} & 0 \end{bmatrix} S_{t-1} + \begin{bmatrix} \Phi_{qm} \\ 0 \end{bmatrix} Y_{m,t-1} + \begin{bmatrix} \Phi_{qc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{q,t} \\ 0 \end{bmatrix}, \quad (2)$$

or more compactly as

$$S_t = \Gamma_s S_{t-1} + \Gamma_y Y_{m,t-1} + \Gamma_{c,t} + \Gamma_u u_{q,t},$$

in which the $pn_m \times 1$ vector $Y_{m,t}$ denotes the actual observations of the original monthly observations, i.e., $Y_{m,t} = X_{m,t}$. By acknowledging that $X_{m,t}$ is fully observed for $t \leq T_b$ we can work with a reduced state-vector of dimension $(p+1)n_q$, instead of $(p+1)(n_m + n_q)$. Why we use $p+1$ here will become clear momentarily.

The corresponding measurement equation takes the form

$$y_t = \begin{bmatrix} y_{m,t} \\ y_{q,t} \end{bmatrix} = \begin{bmatrix} 0 & \Phi_{mq,1} & \Phi_{mq,2} & \Phi_{mq,3:p} \\ 1/3 I_{n_q} & 1/3 I_{n_q} & 1/3 I_{n_q} & 0 \end{bmatrix} S_t + \begin{bmatrix} \Phi_{mm} \\ 0 \end{bmatrix} Y_{m,t-1} + \begin{bmatrix} \Phi_{mc} + \Phi_{mc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ 0 \end{bmatrix} \quad (3)$$

if we are in the last month of a quarter and contains only the first equation in (3) otherwise. Introducing an appropriate switching mechanism, denoted by the selector matrix W_t , we can write (3) for all t more compactly as

$$W_t y_t = W_t (\Lambda_s S_t + \Lambda_y Y_{m,t-1} + \Lambda_{c,t} + \Lambda_u u_{m,t}).$$

A few special features of the state-space representation (2) and (3) are worth commenting. From the aggregation scheme it is clear that the number of lags must be larger or equal than two, i.e., $p \geq 2$.⁴ The inclusion of contemporaneous states requires the

⁴In the case of $p = 2$ the last column in the first matrix on the right-hand side of (3) vanishes.

state vector S_t to be of dimension $(p + 1)n_q$, for otherwise we would lose the p -th lag in (3). 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).⁵

When we move to the ragged edge of our data set, i.e., $t = T_b + 1, \dots, T$, we switch to the larger state vector, as we also have to take care of missing observations of the original monthly variables. Then, for $x_t = [x'_{m,t}, x'_{q,t}]'$, the state vector becomes $S_t = [x'_t, \dots, x'_{t-p}]'$, and the proper selector matrix W_t^* in the measurement equation selects the variables and states that are actually observed over the ragged edge.⁶ Modifying the “reduced” state-space model (2) and (3) yields the “full” state-transition and measurement equations,

$$S_t = \begin{bmatrix} \Phi_{mm,1} & \Phi_{mq,1} & \cdots & \Phi_{mm,p} & \Phi_{mq,p} & 0 \\ \Phi_{qm,1} & \Phi_{qq,1} & \cdots & \Phi_{qm,p} & \Phi_{qq,p} & 0 \\ & & I_{p(n_m+n_q)} & & & 0 \end{bmatrix} S_{t-1} + \begin{bmatrix} \Phi_{mc} + \Phi_{mc,t} \\ \Phi_{qc} + \Phi_{qc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ u_{q,t} \\ 0 \end{bmatrix}, \quad (4)$$

with the obvious dimensions for the zero matrices, and

$$W_t^* \begin{bmatrix} y_{m,t} \\ y_{q,t} \end{bmatrix} = W_t^* \begin{bmatrix} I_{n_m} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 I_{n_q} & 0 & 1/3 I_{n_q} & 0 & 1/3 I_{n_q} & 0 \end{bmatrix} S_t, \quad (5)$$

or more compactly

$$S_t = \Phi^* S_{t-1} + \Phi_{c,t}^* + [I_n, 0]' u_t, \quad u_t \sim N(0, \Sigma_t), \quad \text{and} \quad W_t^* y_t = W_t^* M S_t.$$

⁵Of 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. \(2015a\)](#) find that specifications in levels and growth rates deliver by and large comparable forecasts. In this paper we focus on the specification in levels, the route taken in [Schorfheide and Song \(2015\)](#). If one, however, wants to specify the model in growth rates, the aggregation in the second equation of (3) must be replaced, specifically by the one in [Mariano and Murasawa \(2003\)](#), and the minimum lag order would increase to four.

⁶In 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, $t = T_b + 1, \dots, T$, and $y_{q,t}$ will be an empty set. If one, however, includes quarterly survey data this may not be the case.

Conditional on the parameters and residual 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 [Appendix A](#).

2.2 Block II: Parameters and Time Variation

To close our model we first need to define a law of motion for the time-varying intercepts in (1). In particular, we let them evolve in the usual way as random walks, i.e.,

$$\Phi_{c,t} = \Phi_{c,t-1} + \nu_t, \quad \nu_t \sim N(0, Q), \quad \text{and} \quad Q = \begin{bmatrix} Q_m & 0 \\ 0 & Q_q \end{bmatrix}. \quad (6)$$

We have sneaked in one detail here: we separate the laws of motion for the monthly and quarterly variables. The distinction may prove convenient if one has different prior beliefs about the amount of time-variation in the two groups of variables. We further assume that the innovation matrices Q_m and Q_q are diagonal. This assumption is not essential, and Q_m and Q_q could be replaced by full matrices with only small modifications of the estimation procedure.

Second, we need to specify how the time-varying variance-covariance matrix in (1) can be driven by a single common factor. The commonality can be treated either as additive (see, e.g., [Mumtaz and Surico, 2012](#)) or, as we do, multiplicative, adopting the common stochastic volatility model variant of [Carriero et al. \(2015b\)](#). While the representation (1) of our model was convenient in Block I to get a better picture of the different groups of parameters in a mixed-frequency environment, it will be more convenient now to work with the VAR in standard matrix form. Then, from stacking the $n = n_m + n_q$ vector of data $[x'_{m,t}, x'_{q,t}]'$, the lagged data including an intercept $[x'_{m,t-1}, x'_{q,t-1}, \dots, x'_{m,t-p}, x'_{q,t-p}, 1]'$, the

time-varying intercepts $[\Phi'_{m,t}, \Phi'_{q,t}]'$, and the residuals $[u'_{m,t}, u'_{q,t}]'$ over $t = 1, \dots, T$ we get

$$X = \Phi Z + \Phi_c^T + U \quad (7)$$

with respective dimensions $n \times T$, $(np + 1) \times T$, $n \times T$, and $n \times T$. The link between the coefficient matrix Φ here and the one in (1) is

$$\Phi = \begin{bmatrix} \Phi_{mm,1} & \Phi_{mq,1} & \cdots & \Phi_{mm,p} & \Phi_{mq,p} & \Phi_{mc} \\ \Phi_{qm,1} & \Phi_{qq,1} & \cdots & \Phi_{qm,p} & \Phi_{qq,p} & \Phi_{qc} \end{bmatrix}.$$

Now let us consider a general Kronecker structure for the variance-covariance matrix of U (see, e.g., [Chan, 2018](#)):

$$u = \text{vec}(U) \sim N(0, \Omega \otimes \Psi), \quad (8)$$

in which $\text{vec}(\cdot)$ denotes a column stacking operator and \otimes is the Kronecker product. The $T \times T$ and $n \times n$ matrices Ω and Ψ can be interpreted as governing the serial and cross-sectional variance-covariance structures in the data. The common stochastic volatility model of [Carriero et al. \(2015b\)](#) can then be nested in (8) by suitably defining $\Omega = \text{diag}(\exp(h_1), \dots, \exp(h_T))$ as diagonal matrix and

$$h_t = \rho h_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_h^2) \quad (9)$$

with $|\rho| < 1$. The $n \times n$ blocks along the diagonal of $\Omega \otimes \Psi$ correspond to the time-varying variance-covariance matrices in (1).

3 Bayesian Estimation

The two blocks in the preceding section constitute the two main conditional distributions over which we now build our Markov chain Monte Carlo (MCMC) sampling scheme.

Besides working out a proper MCMC sampler, we also focus on applying computationally efficient sparse and block-banded matrix algebra whenever possible. Exploiting fast and efficient matrix algebra in software packages such as MATLAB, GAUSS, or R complements our general strategy of simplifying matters in order to build a model that can handle a moderately large number of variables (i.e., reducing the state-vector in Block I and using a parsimonious specification of time variation).

3.1 Priors and Hyperparameters

To begin with, we form our baseline prior for Φ and Ψ in (7) and (8) following the Minnesota tradition (see, e.g., [Litterman, 1986](#)), assuming that (most) macroeconomic variables are unit root processes, assuming that we do not know much about how they drift, and assuming that more distant lags have less influence in a VAR. So in case we assume a unit root process, we set the prior parameter of the own first lag to one and to zero everywhere else. We use three hyperparameters to control how strongly we believe in our prior: the first two, λ_1 and λ_2 , separately control the prior on the coefficients of the endogenous variables and the intercept. If $\lambda_1, \lambda_2 \rightarrow 0$ the posterior shrinks towards the prior and otherwise, if $\lambda_1, \lambda_2 \rightarrow \infty$, the posterior mimics ordinary least squares (OLS) estimates. The third hyperparameter, λ_3 , sets the decay rate with which the prior variance decreases with increasing lag length. To account for different scales of the variables in the Minnesota prior, we use the residual variances from simple univariate AR(1) regressions over a pre-sample. We implement our prior beliefs through dummy observations, X_d and Z_d , and refer to [Bańbura et al. \(2010\)](#) and [Schorfheide and Song \(2015\)](#) for details on how to construct these “artificial” observations.

We further refine the baseline Minnesota prior by the so-called sum-of-coefficients and dummy-initial-observation priors. The hyperparameter κ_1 and κ_2 control the variances of these priors. In the case of the sum-of-coefficients prior, $\kappa_1 \rightarrow 0$ implies the presence of a unit root in each equation without any cointegrating relation, while for the dummy-initial-observation prior, $\kappa_2 \rightarrow 0$ implies the presence of an unspecified number of unit

roots in the system, leaving the possibility of cointegration. For $\kappa_1, \kappa_2 \rightarrow \infty$ the priors become uninformative. See [Sims and Zha \(1998\)](#) on the advantages of these refinements, and [Bańbura et al. \(2010\)](#) and [Schorfheide and Song \(2015\)](#) on how they can be imposed through additional dummy variables.

Our prior beliefs for Φ and Ψ are natural conjugate and belong to the family of normal inverse Wishart distributions, i.e.,

$$\text{vec}(\Phi)|\Psi \sim N\left(\text{vec}(\Phi_0), V_{\Phi_0} \otimes \Psi\right) \quad \text{and} \quad \Psi \sim IW\left(U_0 U_0', T_d - np - 1\right), \quad (10)$$

in which T_d denotes the length of the constructed dummy variables. The prior parameters Φ_0 , V_{Φ_0} , and U_0 follow from simply regressing the dummy variable X_d on Z_d , i.e.,

$$\Phi_0 = \left(Z_d Z_d'\right)^{-1} \left(Z_d X_d'\right), \quad V_{\Phi_0} = \left(Z_d Z_d'\right)^{-1}, \quad \text{and} \quad U_0 = X_d - \Phi_0 Z_d. \quad (11)$$

Now, given these prior beliefs we can elicit an initial value S_0 (and its covariance V_{S_0}) in the state-transition equation (2) from running the Kalman filter forward recursion based on the state-space model (2) and (3) over a pre-sample $t = -\tau + 1, \dots, 0$ (see [Appendix A](#) for the equations of the Kalman filter recursions). Note that we implicitly assume the model to be time-invariant over the pre-sample, i.e., $\Phi_c^\tau = 0$ and $\Omega = I_\tau$. This implicit assumption is also one of the reasons why it is convenient to break up the intercept into a constant and time-varying part.

For the time-varying intercepts in (6), $\Phi_{c,t} = [\Phi'_{mc,t}, \Phi'_{qc,t}]'$, we assume independent inverse Gamma priors of the form:

$$Q_m \sim IG\left(Q_0, k_{Q_m}^2 \cdot Q_0 \cdot \Psi_{1:n_m}\right) \quad \text{and} \quad Q_q \sim IG\left(Q_0, k_{Q_q}^2 \cdot Q_0 \cdot \Psi_{n_m+1:n}\right), \quad (12)$$

in which k_{Q_m} and k_{Q_q} are the hyperparameters controlling the amount of time variation we a-priori allow for, and $\Psi_{1:n_m}$ and $\Psi_{n_m+1:n}$ denote the corresponding diagonal elements of the (cross-sectional) prior variance-covariance Ψ in (10). In the expression above we

can explicitly set different priors for the amount of time variation in the intercepts of the monthly and quarterly variables. In the empirical application, however, we will not make use of this option and set $k_{Q_q} = k_{Q_m} = k_Q$. Our point here is that the specification of the Q s (12) is fairly general and could be further generalized to allow for different priors for each intercept. As an example, choosing $k_{Q_i} \rightarrow 0$ would more or less shut down time variation of the intercept in equation i .

Finally, we assume independent priors for the AR(1) coefficient ρ in the law of motion for the common stochastic volatilities and the innovation variance σ_h^2 ; see equation (9). Specifically, we have the following two prior distributions,

$$\rho \sim N\left(\rho_0, V_{\rho_0}\right) \mathbb{I}(|\rho| < 1) \quad \text{and} \quad \sigma_h^2 \sim IG\left(\nu_{h_0}, V_{h_0}\right), \quad (13)$$

with a truncated normal distribution ensuring that the law of motion is not explosive.

While, so far, we have been rather general in the description of our priors, Table 3 summarizes the values of priors and hyperparameters for our empirical application. The specific values should be interpreted as a good starting point for an arbitrary data set; they represent a mix of the related literature, such as Bańbura et al. (2010), Chan and Jeliazkov (2009), and Chan (2018). Maximizing the (approximate) marginal data density, as in Schorfheide and Song (2015), by searching through a grid of hyperparameters goes beyond the scope of our paper. The specific values in Table 3 are therefore not directly comparable with Schorfheide and Song (2015). The optimal selection of hyperparameters will be part of a follow-up paper.

3.2 Posterior Analysis

Having specified priors and hyperparameters, we can now obtain posterior draws by sequentially sampling from Block I, I.a) $p(S^T | \Phi, \Psi, \Phi_c^T, h^T, y^T)$, and from Block II, II.a) $p(\Phi, \Psi | \Phi_c^T, h^T, S^T, y^T)$, II.b) $p(\Phi_c^T | \Phi, \Psi, h^T, Q, S^T, y^T)$, II.c) $p(Q | \Phi_c^T)$, II.d) $p(h^T | \Phi, \Psi, \Phi_c^T, \rho, \sigma_h^2, S_q^T, S^T, y^T)$, II.e) $p(\sigma_h^2 | h^T, \rho)$, and II.f) $p(\rho | h^T, \sigma_h^2)$. For brevity, all the conditioning

sets above contain only quantities actually required in each step to compute the posterior.

In Step I.a we draw the latent states S^T by recursively applying the Kalman filter to the state space model laid out in Section 2.1. Given the initialization $S_{0|0} = S_0$ and $V_{0|0} = V_{S_0}$ the conditional posterior distribution is Gaussian,

$$S_t | \Phi, \Psi, \Phi_c^T, h^T, y^T \sim N(S_{t|T}, V_{t|T}), \quad t = 1, \dots, T. \quad (14)$$

While, in principle, standard textbook formulas apply for the posterior quantities $S_{t|T}$ 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.

As shown by Chan (2018), given the Kronecker structure in the variance-covariance matrix (8) and the natural conjugate prior for (Φ, Ψ) , the usual result of a normal inverse Wishart posterior distribution with analytical expressions holds. This result also goes through with time-varying intercepts by simply rewriting (7) as $\tilde{X} = X - \Phi_c^T = \Phi Z + U$. Specifically, in Step II.a we first sample Ψ marginally and then conditional on Ψ we draw Φ , i.e.,

$$\begin{aligned} \Psi | \Phi_c^T, h^T, S_q^T, S^T, y^T &\sim IW(\hat{\Psi}, \nu_0 + T) \\ \text{vec}(\Phi) | \Psi, \Phi_c^T, h^T, S_q^T, S^T, y^T &\sim N(\text{vec}(\hat{\Phi}), K_{\Phi}^{-1} \otimes \Psi). \end{aligned} \quad (15)$$

with the following set of posterior parameters:

$$\begin{aligned} K_{\Phi} &= V_{\Phi_0}^{-1} + Z\Omega^{-1}Z' \quad \text{with} \quad \Omega = \text{diag}(\exp(h_1), \dots, \exp(h_T)), \\ \hat{\Phi} &= K_{\Phi}^{-1} (V_{\Phi_0}^{-1}\Phi_0 + Z\Omega^{-1}\tilde{X}'), \\ \hat{\Psi} &= \Psi_0 + \Phi_0 V_{\Phi_0}^{-1} \Phi_0' + \tilde{X}\Omega^{-1}\tilde{X}' - \hat{\Phi} K_{\Phi} \hat{\Phi}'. \end{aligned} \quad (16)$$

Now, drawing from the normal distribution in (15) would require the inversion of the $(np + 1) \times (np + 1)$ matrix K_{Φ} and the Cholesky decomposition of the large $n(np + 1) \times n(np + 1)$ matrix $K_{\Phi}^{-1} \otimes \Psi$. By acknowledging the Kronecker structure we can limit the computational cost to two separate Cholesky decompositions of K_{Φ} and Ψ . Specifically,

Chan (2018) shows that instead of conventionally drawing Φ in (15), we can equivalently compute Φ as (this follows from standard results on the matrix normal distribution)

$$\Phi = \hat{\Phi} + (C'_K \backslash D) C_\Psi, \quad (17)$$

in which C_K and C_Ψ are the respective Cholesky decompositions, D denotes a $(np+1) \times n$ matrix of independent $N(0, 1)$ random variables, and the backslash operator “ \backslash ” refers to the unique solution to a triangular system of the generic form $Cu = e$ obtained by forward substitution, i.e., $C \backslash e = C^{-1}e$.⁷ The advantage of this solution method in triangular systems is that the “inversion” problem requires only the same number of operations as the multiplication Cu . We can further exploit the computational advantages from forward substitution by rewriting the expression for $\hat{\Phi}$ in (16) in terms of the Cholesky decomposition $C_K C'_K = K_\Phi$, i.e.,

$$\hat{\Phi} = \left(C_K C'_K \right)^{-1} \left(V_{\Phi_0}^{-1} \Phi_0 + Z \Omega^{-1} \tilde{X}' \right) = C'_K \backslash \left(C_K \backslash \left(V_{\Phi_0}^{-1} \Phi_0 + Z \Omega^{-1} \tilde{X}' \right) \right). \quad (18)$$

For drawing the time-varying intercepts Φ_c^T in Step II.b we use the same efficient forward substitution method to the matrix inversion problem. In addition, we will introduce sparse and block-banded matrix algebra which further speeds up computation. Note that we can write the law of motion for $\Phi_{c,t}$ in (6) by stacking over $t = 1, \dots, T$, i.e.,

$$H \text{vec}(\Phi_c^T) = \text{vec}(\nu^T), \quad \text{vec}(\nu^T) \sim N(0, I_T \otimes Q), \quad (19)$$

⁷The software packages MATLAB uses this efficient solution method by default when using the backslash (or slash) operator. Forward (and backward) substitution is also available in other software packages such as GAUSS and R.

in which the $Tn \times Tn$ matrix H contains the law of motion as follows:

$$H = \begin{bmatrix} I_n & & & & & \\ -I_n & I_n & & & & \\ & -I_n & I_n & & & \\ & & & \ddots & \ddots & \\ & & & & -I_n & I_n \end{bmatrix}. \quad (20)$$

Premultiplying (19) by H^{-1} implies that the joint density $\text{vec}(\Phi_c^T)|Q$ is given by

$$\text{vec}(\Phi_c^T)|Q \sim N\left(0, (H'(I_T \otimes Q)^{-1}H)^{-1}\right). \quad (21)$$

The conditional Gaussian posterior distribution follows directly from Bayes' rule by combining the joint density above with the data, i.e.,⁸

$$\Phi_c^T|\Phi, \Psi, h^T, Q, S_q^T, S^T, y^T \sim N(\hat{\Phi}_c^T, P_{\Phi_c}^{-1}), \quad (22)$$

in which the mean $\hat{\Phi}_c^T$ and the precision matrix P_{Φ_c} are given by

$$\begin{aligned} \hat{\Phi}_c^T &= P_{\Phi_c}^{-1} \left((\Omega^+ \otimes \Psi^{-1}) \cdot \text{vec}(\tilde{X} - \Phi Z) \right), \\ P_{\Phi_c} &= H'(I_T \otimes Q)^{-1}H + \Omega^+ \otimes \Psi^{-1}. \end{aligned} \quad (23)$$

Ω^+ is defined as the inverse of a diagonal matrix, i.e., $\Omega^+ = \text{diag}(\exp(-h_1), \dots, \exp(-h_T))$. Since P_{Φ_c} is a sparse and block-banded matrix, its Cholesky decomposition $C_{P_{\Phi_c}} C_{P_{\Phi_c}}' = P_{\Phi_c}$ can be computed efficiently and can be used in the same way as in (17) and (18) to draw the time-varying intercepts Φ_c^T , i.e.,

$$\Phi_c^T = C_{P_{\Phi_c}}' \setminus \left(C_{P_{\Phi_c}} \setminus \left((\Omega^+ \otimes \Psi^{-1}) \cdot \text{vec}(\tilde{X} - \Phi Z) \right) + D \right), \quad (24)$$

⁸The derivation is standard and can be found in references such as [Gelman et al. \(2003\)](#); see also [Chan and Jeliazkov \(2009\)](#).

in which D is now a $Tn \times 1$ vector of independent $N(0, 1)$ random variables. Note that to better disentangle the constant and time-varying parts of the intercept we demean Φ_c^T after each draw. Compared to standard samplers used in time-varying parameter models based on the Kalman filter and smoother recursions, such as [Carter and Kohn \(1994\)](#), the above precision sampler generates faster run times; see [Chan and Jeliazkov \(2009\)](#).

In Step II.c, drawing the diagonal elements of the innovation matrix Q in (6) is straightforward and standard results for the conditional inverse Gamma posterior distribution apply. Drawing the parameters h^T , ρ , and σ_h^2 in the common stochastic volatility part—Steps II.d, II.e, and II.f—follows exactly the algorithm in [Chan \(2018\)](#) and [Chan and Hisiao \(2014\)](#). Specifically, for h^T it first involves computing the mode and the corresponding negative Hessian of the underlying density by a Newton-Raphson algorithm. We then use the resulting proposal distribution to directly sample h^T using an acceptance-rejection Metropolis-Hastings step.

4 Application

To evaluate how the forecasting performance and other properties of our mixed-frequency VAR with time-varying intercepts and common stochastic volatility performs, we consider an empirical analysis using a dataset of eleven U.S. macroeconomic variables.

Aside from our “full” model, i.e., the one including common stochastic volatility and time-varying intercepts, we consider two “intermediate” model versions: a mixed-frequency VAR with time-varying intercepts but time-invariant volatilities and a mixed-frequency VAR with common stochastic volatility but constant intercepts. The outcomes of these model variants allow us to identify the individual contributions of both modifications and to unveil interactions between the two. We label the corresponding three models TVi-CSV-MF-VAR, TVi-MF-VAR and CSV-MF-VAR. Since the model of [Schorfheide and Song \(2015\)](#) served as a starting point to our analysis, we assess the usefulness of our proposed modifications against their MF-VAR. Table 5 provides an overview of the

models under consideration. All models have $p = 6$ monthly lags as in [Schorfheide and Song \(2015\)](#).

We run 12,000 iterations of our MCMC sampler, but discard the first 2,000 as burn-in phase.⁹ As outlined in Section 3.1, we reserve the initial four years of data (1986 to 1989) as pre-sample to initialize our prior beliefs.

4.1 Data

We collect eleven macroeconomic variables for the U.S. from the FRED database of the Federal Reserve Bank of St. Louis: gross domestic product (GDP), private domestic investment (INV), government expenditures (GOV), the unemployment rate (UNR), hours worked (HRS), the consumer price index (CPI), the industrial production index (IPI), the personal consumption expenditures index (PCE), the federal funds rate (FF), the treasury bond yield (TB), and the S&P 500 index (SP500). Note that the first three series are sampled at quarterly frequency, whereas the remaining ones are occurring on a monthly basis.¹⁰ Variables sampled at even higher frequencies, such as the federal funds rate, are time-aggregated to monthly frequency. All series are seasonally adjusted except for the financial variables. Following [Schorfheide and Song \(2015\)](#), all indicators enter the model in log-levels; exceptions are the unemployment rate, the federal funds rate and the treasury bond yield, which are scaled by 100 to make their scale comparable to the variables in log-levels. Table 4 provides detailed information about the series under consideration.

We have downloaded the data on July 31st, 2017, close of business. All variables start from 1986:Q1 (1986:M1) and are characterized by the familiar ragged-edge structure (see, e.g., [Marcellino and Schumacher, 2010](#)) because of different publication delays (see Table 4). As an example, the industrial production index is available until 2017:M6, whereas the

⁹We use MATLAB to carry out all computations.

¹⁰In contrast to [Schorfheide and Song \(2015\)](#) we use domestic instead of fixed investment and an index for personal consumption expenditures. Because of changes in the base years, fixed investment and actual PCE are available in chain-linked dollars only for a comparably short period of time. [Schorfheide and Song \(2015\)](#) circumvent this problem implicitly by conducting their analysis in real time, i.e., by using the respective data vintage at each point a forecast is made.

federal funds rate for 2017:M7 just got published. As we assume forecasts to be computed at the end of each month, the ragged-edge structure obtained when downloading the data is representative of our analysis. Note that we consider second releases for the quarterly variables, implying that they get released halfway through the respective quarter. In other words, on 31st of July the latest available GDP-figure corresponds to the first quarter.

4.2 Running Time

Before turning to the heart of our application we briefly comment on some computational aspects. As argued throughout the methodological part of the paper, we face a trade-off between considerably fast estimation and model complexity when extending the mixed-frequency VAR of [Schorfheide and Song \(2015\)](#). The mixed-frequency time-varying VAR of [Cimadomo and D’Agostino \(2016\)](#), for instance, adopts the full-blown time-varying parameter model (TVP-VAR) of [Primiceri \(2005\)](#), which marks more or less the upper bound in terms of complexity. While their paper needs to be praised for merging a mixed-frequency VAR with the prototypical TVP-VAR, such a complex model structure forces them to make sacrifices in terms of model size. They restrict themselves to three variables and two lags. One could perhaps add a few more variables or lags before reaching the limit of their sampler. Our application with eleven variables and six lags, in contrast, could be easily enlarged to, say, 20 variables.¹¹

Here, we address the trade-off between computational demand and model complexity by parsimoniously specifying time variation and by using efficient matrix algebra as outlined in Section 3.2. In Table 6 we show the running times of our models compared to the benchmark MF-VAR of [Schorfheide and Song \(2015\)](#).¹² It turns out that the addition of time-varying intercepts alone bears little extra computational costs. In contrast, the computation of the common stochastic volatility is a bit more demanding, as we need

¹¹Computationally even more demanding than increasing the cross-section is the time dimension. Starting the sample in 1970 rather than 1990 makes a considerable difference in terms of running time.

¹²Note that Block I is identical for all models implying that any differences must stem from the estimation procedure in Block II.

to run through an acceptance-rejection Metropolis-Hastings step. Our results of running time slightly less than two times that of the MF-VAR are roughly in line with [Chan \(2018\)](#), reporting a relative running time of 2.19 for a common-frequency Bayesian VAR with common stochastic volatility (compared to one without).

4.3 Full Sample Estimation

We now turn to the first main set of empirical results. Using the full sample from 1990:M1 to 2017:M7 (recalling that we discard a pre-sample of four years prior to 1990 and that GDP is only available until 2017Q1), we take a closer look at the estimated intercepts, common stochastic volatility, and the latent monthly GDP series.

Figure 1 reports the time-varying intercepts of three of our eleven variables – quite coincidentally those that we consider in our recursive out-of-sample forecast exercise: GDP, the unemployment rate and the federal funds rate.¹³ In particular, the solid (red) line represent the median value of the intercept in each time period for the TVi-CSV-MF-VAR, whereas the dotted (blue) line marks the corresponding outcome for the time-invariant MF-VAR model. Note that the corresponding median values have been demeaned to make them comparable in terms of scale.

The data clearly support the presence of time variation in the intercepts. Note that our choice $k_Q = \sqrt{0.001}$ does not particularly penalize time variation in the intercepts; larger values for k_Q would favor a higher degree of time variation at the cost of incurring more outliers rather than smoothly changing trends. Empirically, capturing slowly changing trends is, perhaps, the main reason why our model should have an edge over the time-invariant MF-VAR. As VAR-intercepts typically reflect the conditional mean of a series we cannot simply interpret their movements as fluctuations in the (long-run) growth rates.¹⁴ What we may deduct, however, is the change in the contribution of the intercept to the

¹³Results for the remaining variables are available upon request.

¹⁴In the dynamic factor model of [Antolin-Diaz et al. \(2017\)](#) without lagged endogenous variables or the time-varying VAR of [Banbura and van Vlodrop \(2018\)](#) the time-varying intercepts directly measure the unconditional mean or (long-run) growth rate.

variable in the respective equation between two points: In 2017 the intercept contributes about 0.35 percentage points less to monthly GDP growth than it did in 1990. For the unemployment rate, though, the intercept’s contribution increased over time by about 0.28 percentage points. For the fed funds rate its contribution started off negatively, then increased by about 0.1% percentage points by mid 2010, yet ended with almost nil impact at the end of the sample. This may reflect the fact that the rate remained very close to the zero lower bound at the end of the sample, thereby not resembling much of an individual time-varying effect.

Figure 2 provides a closer look into the common stochastic volatility factor. In particular, it plots the standardized values of h_t , i.e., $\exp(h_t/2)$. First and as expected, there is considerable time variation in the volatilities. Apart from smaller rises and falls, one spike clearly stands out: the one for the Great Recession in 2007. At the end of the sample, the results provide rather low evidence for time variation; the solid (red) line lies even beneath the dotted (blue) one. This may also hint towards the aftermath of the crisis being characterized by less erratic movements of the time series involved, similar to the fed funds rate outlined before.

Figure 3 plots the growth rates of the median values of the monthly GDP series (scaled by 3 to make them comparable to quarter-on-quarter rates). It turns out that the outcomes are fairly similar. Monthly GDP of the TVi-CSV-MF-VAR, however, often extends beyond the estimates under the MF-VAR, especially when larger changes in growth rates occur. This is most likely due to much more pronounced (common) stochastic volatility, also suggested by the – once again – less volatile behavior at the end of the sample period.

4.4 Forecasting Exercise

We base the evaluation of our models on the forecast accuracy with respect to the three target variables considered already above: the unemployment rate, the federal funds rate and GDP. To this end we conduct a forecast exercise in pseudo-real time and consider

an increasing sequence of estimation samples, always imposing the ragged-edge structure, starting with the period 1990:M1-2006:M2 until 1990:M1-2017:M7. We also impose the laws of motion in (6) and (9) for the time-varying intercepts and the common stochastic volatility over the forecast horizon. Subsequently, we compute up to twelve monthly forecasts for each indicator, assuming that we are always at the end of the month.¹⁵

The specific forecast period for the quarterly series is determined by their respective publication delays. For example, in 2007:M1 the latest available GDP observation corresponds to 2006:Q3. Hence, we would look at forecasts over the period 2006:M10-2007:M9. When evaluating these forecasts it is 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 (see, e.g., [Bańbura et al., 2011](#)). Consequently, the next value represents a nowcast, whereas the third and fourth figures are “truly” forecasts (with quarterly horizons of one and two). At the end of the second and third month, however, the GDP-value corresponding to the previous quarter is already available, causing the first figure to be a nowcast and the remaining ones to be one-, two and three-quarter-ahead forecasts. Considering the increasing information content as we move from one month to the next, we obtain performance measures for, in total, twelve different forecast horizons. Counting the amount of months between the moment we compute the forecast and the end of the reference quarter, we label these horizons $h^{GDP} = -1, 0, \dots, 10$: $h^{GDP} = -1$ corresponds to the backcast made at the end of the first month, $h^{GDP} = 0, 1$ refer to the nowcasts made at the end of the second and third month, and so forth, until $h^{GDP} = 10$ applies to the three-quarter-ahead forecast made at the end of the second month.

The situation is slightly different for our two monthly target variables, which are characterized by much shorter publication delays. For the unemployment rate, at the end of month t , we always need to compute one nowcast, as we observe it with a delay of $t - 1$, the rest are forecasts. For the federal funds rate, which becomes available just when we

¹⁵One could align the timing of the forecasts to the publication date of each indicator and update the forecasts several times a month.

start a new forecasting round, we deal with forecasts only.

To evaluate our forecasts we use two metrics: root mean squared forecast errors (RMSFE) for point forecasts, and average log predictive likelihoods (logPL) for density forecasts. We refer the reader to any standard textbook for the precise formulae. Tables 7 to 9 contain the outcomes for our three target variables in turn. Each table shows the results for point forecasts in the top half and the ones for density forecasts in the bottom half. For our full model and two intermediate versions we report the forecast accuracy measures relative to the MF-VAR of [Schorfheide and Song \(2015\)](#). Hence, for point forecasts values smaller than one and for density forecasts values larger than zero indicate a better forecast performance. Note that absolute RMSFEs (multiplied by 100) are provided for the MF-VAR.

For the unemployment rate the point forecasts of the TVi-MF-VAR clearly beat the MF-VAR. The models containing common stochastic volatility, however, lead to sacrifices in terms of point forecast accuracy. Like the model just including a time-varying intercept, TVi-CSV-MF-VAR outperforms the benchmark, yet to a lesser extent for small horizons due to the effect of stochastic volatility. The situation is similar for density forecasts, except that also the CSV-MF-VAR model leads to forecast improvements in almost all cases. The full model, including both channels of time variation, is thus the preferred choice most of the time.

The results for the federal funds rate feature a couple of similarities. Once again, the time-varying intercept seems to be more beneficial for the forecast performance than common stochastic volatility. Interestingly, point forecasts still benefit from the latter, whereas density forecasts only do so for small horizons. Like before, the TVi-CSV-MF-VAR model, which combines both features, delivers large improvements over the time-invariant MF-VAR of [Schorfheide and Song \(2015\)](#).

Finally, and most importantly due to its role as the most important summarizing indicator for the economic activity, we consider the outcomes for GDP. For both, point and density forecasts, we mostly record small improvements of the intermediate models

over the benchmark. When being combined in the full model, though, the results are oftentimes even better. Hence, there appear to be some diversifying effects when adding both channels of time variation at the same time such that we achieve comparably large forecast gains with respect to point and density forecasts of GDP.

All in all, it emerges that time-varying intercepts are somewhat more beneficial for forecasting than common stochastic volatility, whereby the interactions between the two features in the full model often lead to the best outcomes overall. Our full model thus presents itself as a useful alternative to the standard MF-VAR model of [Schorfheide and Song \(2015\)](#).

5 Conclusion

We introduce time-varying mixed-frequency VAR models for macroeconomic forecasting purposes. Since combining time-varying parameters with mixed frequencies requires considerable computational effort, we restrict time variation to the intercepts and error variances. For the error variances we further reduce model complexity by forcing the stochastic volatilities to be common across all variables. This frees us from computational restrictions usually implied by increasing the number of variables and lags ([Carriero et al., 2015b](#)). We show 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 we estimate the VAR parameters. With respect to the first block we rely on the approach of [Schorfheide and Song \(2015\)](#), which also serves as main competitor for our method. In the second block the Kronecker structure of the error terms, for which the common stochastic volatility restriction is a special case, allows us further to exploit gains from using sparse and block-banded matrix algebra (see, e.g., [Chan, 2018](#)).

We demonstrate the feasibility and usefulness of our approach within an empirical analysis involving eleven U.S. macroeconomic time series. We discuss several computa-

tional aspects of our “full” model and two “intermediate” ones, where either one of the two channels, through which time variation can enter, is switched off. It turns out that costs in terms of computing time are acceptable given the gains in flexibility. We show that the data support time variation in both the intercepts and volatilities. Finally, a pseudo-real time forecasting exercise reveals that, first and foremost, our models yield overall better point and density forecasts for the three target variables in question, among which is GDP. Furthermore, in our specific case, time-varying intercepts seem to be somewhat more beneficial for forecasting than common stochastic volatility.

References

- Andreou, E., Ghysels, E., and Kourtellis, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251.
- Antolin-Diaz, J., Drechsel, T., and Petralia, I. (2017). Tracking the slowdown in long-run GDP growth. *Review of Economics and Statistics*, 99(2):343–356.
- Bañbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector autoregressions. *Journal of Applied Econometrics*, 25(1):71–92.
- Bañbura, M., Giannone, D., and Reichlin, L. (2011). Nowcasting. In Clements, M. P. and Hendry, D. F., editors, *The Oxford Handbook on Economic Forecasting*, pages 193–224. Oxford University Press, New York.
- Banbura, M. and van Vlodrop, A. (2018). Forecasting with Bayesian vector autoregressions with time variation in the mean. Tinbergen Institute Discussion Papers 18-025/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.
- Carriero, A., Clark, T. E., and Marcellino, M. (2015a). Bayesian VARs: Specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30(1):46–47.
- Carriero, A., Clark, T. E., and Marcellino, M. (2015b). Common drifting volatility in large Bayesian VARs. *Journal of Business & Economic Statistics*, DOI: 10.1080/07350015.2015.1040116.
- Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

- Chan, J. C. (2018). Large Bayesian VARs: A flexible Kronecker error covariance structure. *Journal of Business & Economic Statistics*, 0:1–12.
- Chan, J. C. and Hisiao, C. Y. (2014). Estimation of stochastic volatility models with heavy tails and serial dependence. In Jeliazkov, I. and Yang, X.-S., editors, *Bayesian Inference in the Social Sciences*. John Wiley & Sons, Hoboken, NJ.
- Chan, J. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1-2):101–120.
- 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.
- Clements, M. and Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics*, 26(4):546–554.
- Cox, D. R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., and Lauritzen, S. L. (1981). Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics*, 8(2):93–115.
- D’Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.
- Eickmeier, S., Lemke, W., and Marcellino, M. (2015). Classical time-varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):493–533.
- Froni, C. and Marcellino, M. (2013). A survey of econometric methods for mixed-frequency data. Working Paper 2013/06, Norges Bank.

- Faroni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1):57–82.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman and Hall, New York, 2nd edition.
- Ghysels, E. (2016). Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics*, forthcoming.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models. CIRANO Working Papers 2004s-20.
- Götz, T. B., Hecq, A., and Urbain, J.-P. (2014). Forecasting mixed frequency time series with ECM-MIDAS models. *Journal of Forecasting*, 33:198–213.
- Götz, T. B., Hecq, A., and Urbain, J.-P. (2016). Combining forecasts from successive data vintages: An application to U.S. growth. *International Journal of Forecasting*, 32(1):61–74.
- Guèrin, P. and Marcellino, M. (2013). Markov-switching MIDAS models. *Journal of Business & Economic Statistics*, 31(1):45–56.
- Koop, G. and Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics*, 177(2):185–198.
- Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.
- Marcellino, M. and Schumacher, C. (2010). Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP. *Oxford Bulletin of Economics and Statistics*, 72(4):518–550.
- 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.

- Mumtaz, H. and Surico, P. (2012). Evolving international inflation dynamics: World and country-specific factors. *Journal of the European Economic Association*, 10(4):716–734.
- Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852.
- Schorfheide, F. and Song, D. (2015). Real-time forecasting with a mixed-frequency VAR. *Journal of Business & Economic Statistics*, 33(3):366–380.
- Schumacher, C. (2014). MIDAS regressions with time-varying parameters: An application to corporate bond spreads and GDP in the euro area. Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Time Series Analysis.
- Sims, C. A. and Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39(4):949–968.
- Uhlig, H. (1994). What macroeconomists should know about unit roots. *Econometric Theory*, 10(3-4):645–671.
- Villani, M. (2009). Steady-state priors for vector autoregressions. *Journal of Applied Econometrics*, 24(4):630–650.

A Details of Block I

Given $S_{0|0} = S_0$ and $V_{0|0} = V_{S_0}$ the forward pass of the Kalman filter, over the balanced and ragged-edge part of the data, is given by the following system of equations:

The balance part $t = 1, \dots, T_b$

$$\begin{aligned}
S_{t|t-1} &= \Gamma_s S_{t-1|t-1} + \Gamma_y Y_{m,t-1} + \Gamma_{c,t}, \\
V_{t|t-1} &= \Gamma_s V_{t-1|t-1} \Gamma'_s + \Gamma_u \Sigma_{qq,t} \Gamma'_u, \\
K_{f,t} &= (V_{t|t-1} \Lambda'_s + \Gamma_u \Sigma_{mq,t} \Lambda'_u) (\Lambda_s V_{t|t-1} \Lambda'_s + \Lambda_u \Sigma_{mm,t} \Lambda'_u \\
&\quad + \Lambda_s \Gamma_u \Sigma_{qm,t} \Lambda'_u + \Lambda_u \Sigma_{mq,t} \Gamma'_u \Lambda'_s)^{-1}, \\
\hat{y}_t &= \Lambda_s S_{t|t-1} + \Lambda_y Y_{m,t-1} + \Lambda_{c,t}, \\
S_{t|t} &= S_{t|t-1} + K_{f,t} (y_t - \hat{y}_t), \\
V_{t|t} &= V_{t|t-1} - K_{f,t} (\Lambda_s V_{t|t-1} + \Lambda_u \Sigma_{mq,t} \Gamma'_u).
\end{aligned}$$

In case t corresponds to the last month of a quarter we need to premultiply all Λ 's by the selector matrix W_t , defined in (3).

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

$$\begin{aligned}
S_{t|t-1} &= \Phi^* S_{t-1|t-1} + \Phi_{c,t}^*, \\
V_{t|t-1} &= \Phi^* V_{t-1|t-1} \Phi^{*'} + [I_n, 0]' \Sigma_t, \\
K_{f,t} &= (V_{t|t-1} W_t^{*'}) (W_t^* V_{t|t-1} W_t^{*'})^{-1}, \\
\hat{y}_t &= W_t^{*'} S_{t|t-1}, \\
S_{t|t} &= S_{t|t-1} + K_{f,t} (W_t^* y_t - \hat{y}_t), \\
V_{t|t} &= V_{t|t-1} - K_{f,t} (W_t^* V_{t|t-1}).
\end{aligned}$$

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 the beginning. 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 whether there are

missing values.

The ragged-edge part $t = T - 1, T - 2, \dots, T_b$

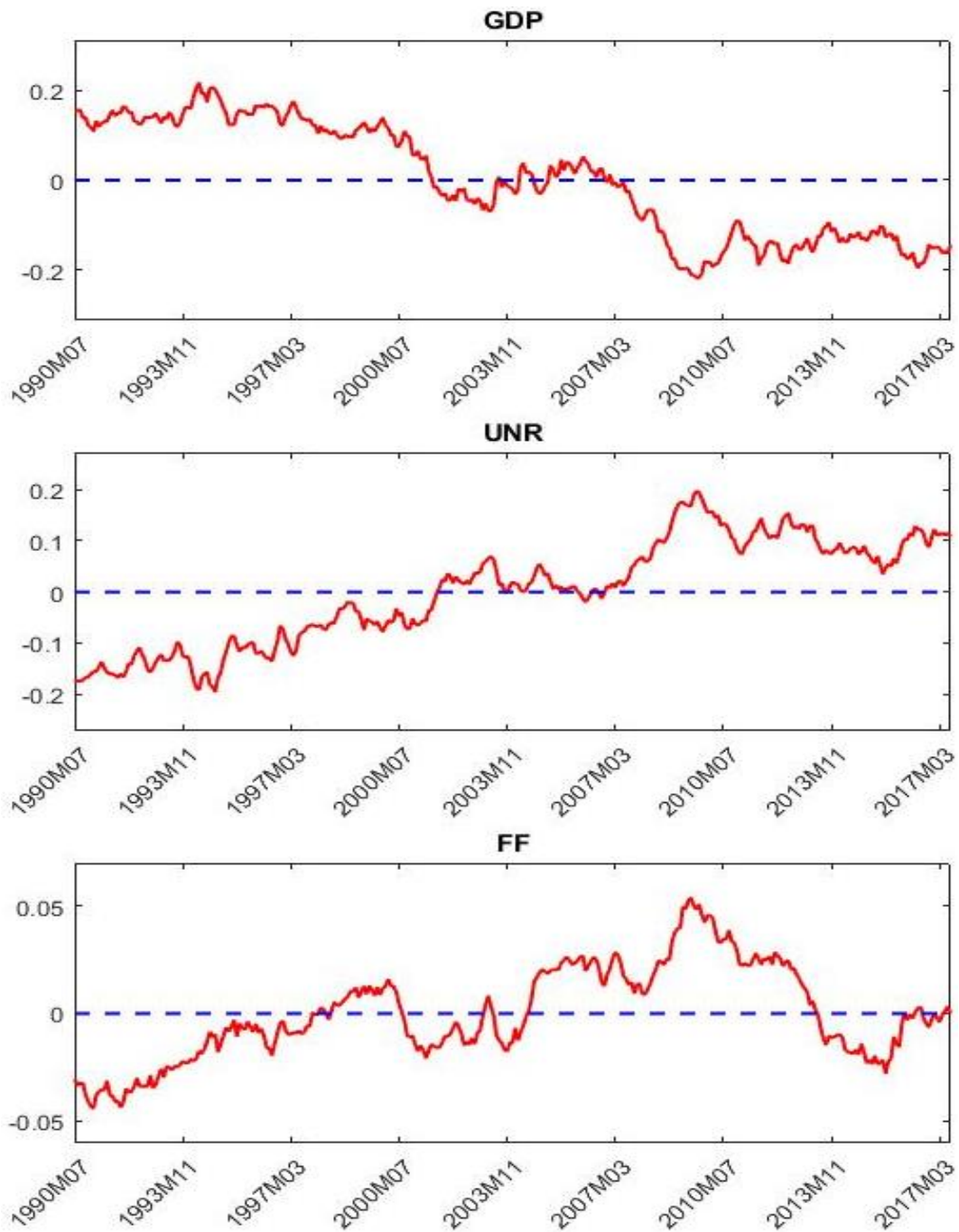
$$\begin{aligned}
S_{t+1|t} &= \Phi^* S_{t|t} + \Phi_{c,t}^*, \\
V_{t+1|t} &= \Phi^* V_{t|t} \Phi^{*'}, \\
S_{t|t+1} &= S_{t|t} + V_{t|t} \Phi^{*'} V_{t+1|t}^{-1} (S_{t+1} - S_{t+1|t}), \\
V_{t|t+1} &= V_{t|t} - V_{t|t} \Phi^{*'} V_{t+1|t}^{-1} \Phi^* V_{t|t}.
\end{aligned}$$

The balanced part $t = T_b - 1, T_b - 2, \dots, 1$

$$\begin{aligned}
S_{t+1|t} &= \Gamma_s S_{t|t} + \Gamma_y y_{m,t} + \Gamma_{c,t}, \\
V_{t+1|t} &= \Gamma_s V_{t|t} \Gamma'_s + \Gamma_u \Sigma_{qq,t} \Gamma'_u, \\
S_{t|t+1} &= S_{t|t} + V_{t|t} \Gamma'_s V_{t+1|t}^{-1} (S_{t+1} - S_{t+1|t}), \\
V_{t|t+1} &= V_{t|t} - V_{t|t} \Gamma'_s V_{t+1|t}^{-1} \Gamma_s V_{t|t}.
\end{aligned}$$

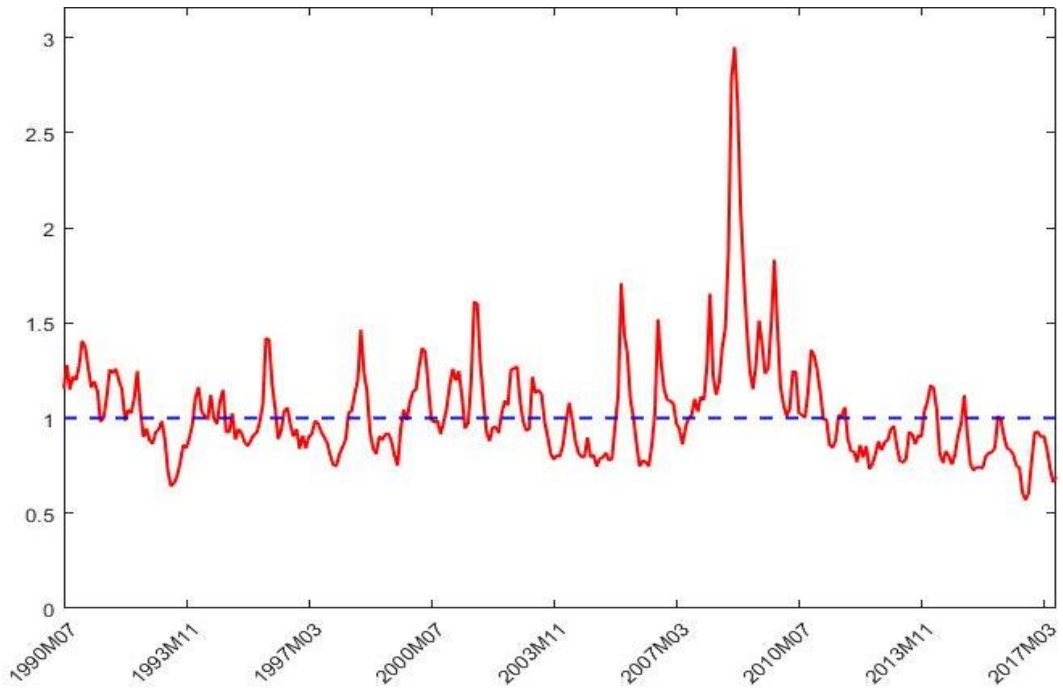
B Figures

Figure 1: Selected Time-Varying Intercepts



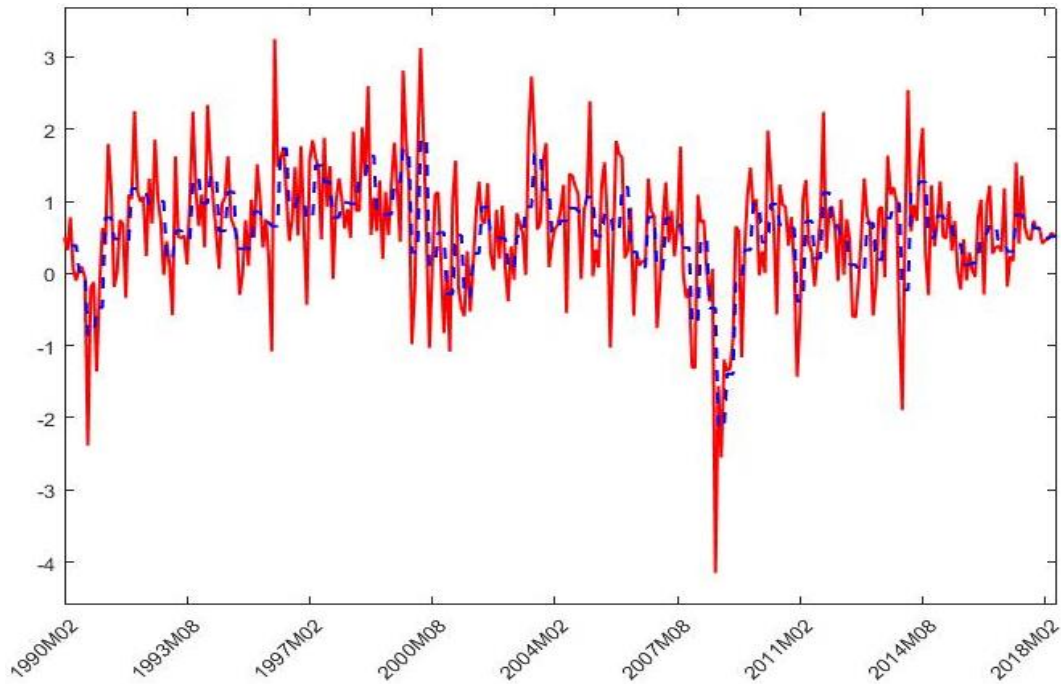
Notes: Demeaned medians of the time-varying intercepts ($\Phi_c + \Phi_c^T$); multiplied by 100, changes between two points are thus in units of percentage points.

Figure 2: Common Stochastic Volatility



Notes: Medians of the common stochastic volatility in standard deviations, i.e., $\exp(h_t/2)$.

Figure 3: Monthly GDP



Notes: Growth rates of the medians of monthly GDP; scaled by 3 to make them comparable to quarter-on-quarter rates.

C Tables

Table 1: Running Time Advantages Relative to a Fully-Fledged Model

Number of lags, p	Number of variables, n		
	3	4	6
3	42.9	34.4	10.6
6	33.9	16.4	
12	25.0		

Note: Numbers are percent of running time of our parsimonious mixed-frequency model with time-varying intercepts and common stochastic volatility relative to a fully-fledged model (Primiceri, 2005, extended to the mixed-frequency case). The $n = 3, 4, 6$ variables represent a subset of the ones described in Table 4: GDP, CPI, Fed funds rate, unemployment rate, industrial production, and personal consumption expenditures.

Table 2: Forecast Performance Relative to a Fully-Fledged Model

Horizon	-1	1	3	5	7
Rel. RMSFE	1.02	0.84	0.95	0.95	0.86
Rel. logPL	-0.64	0.04	0.18	0.18	0.14

Note: Comparison of our parsimonious mixed-frequency model with time-varying intercepts and common stochastic volatility and a fully-fledged model (Primiceri, 2005, extended to the mixed-frequency case). The specification is $n = 3$ and $p = 3$; the $n = 3$ variables represent a subset of the ones described in Table 4: GDP, CPI, Fed funds rate. The forecast horizon is defined as the amount of months between the moment a forecast is made (assumed to be at the end of each month) and the end of the reference quarter; -1 is a backcast, 0 and 1 are nowcasts, the remaining horizons refer to forecasts. RMSFE ratios smaller than one and logPLs larger than zero imply that our parsimonious model performs better.

Table 3: Selected Values for Priors and Hyperparameters

Parameter	Value	Description
λ_1	0.2	controls prior variance of endogenous VAR coefficients
λ_2	10^5	controls prior variance of intercepts
λ_3	1	quadratic decay rate of prior variance with increasing lags
κ_1	2	tightness of sum-of-coefficients prior
κ_2	2	tightness of dummy-initial-observation prior
ν_0	$n + 3$	prior degrees of freedom in inverse Wishart for Ψ
k_{Q_m}	$\sqrt{0.001}$	prior amount of time variation of intercepts (monthly)
k_{Q_q}	$\sqrt{0.001}$	prior amount of time variation of intercepts (quarterly)
Q_0	48	prior shape parameter for Q s (equal to length of pre-sample)
ρ_0	0.9	prior for AR(1) coefficient in law of motion for CSV
V_{ρ_0}	0.2^2	prior variance of ρ_0
ν_{h_0}	5	prior shape parameter for the CSV innovation variance σ_h^2
V_{h_0}	$0.01(\nu_{h_0} - 1)$	prior scale parameter for the CSV innovation variance σ_h^2

Table 4: Data and Stylized Release Calendar

Series ID	Transf.	Pub-Lag	Description	FRED Mnemonic
GDP	Log	1 Qrt*	Real gross domestic product, sa	GDPC1
INV	Log	1 Qrt*	Real gross private domestic investment, sa	GPDIC1
GOV	Log	1 Qrt*	Real government consumption expenditures and gross investment, sa	GCEC1
UNR	1/100	1 Mth	Civilian unemployment rate, sa	UNRATE
HRS	Log	1 Mth	Index of aggregate weekly hours, sa	AWHI
CPI	Log	1 Mth	Consumer price index for all urban consumers, sa	CPIAUCSL
IPI	Log	1 Mth	Industrial production index, sa	INDPRO
PCE	Log	1 Mth	Personal consumption expenditures index, sa	PCEPI
FF	1/100	./.	Effective federal funds rate, nsa	FEDFUNDS
TB	1/100	./.	10-year treasury constant maturity rate, nsa	GS10
SP500	Log	./.	S&P 500 stock index, adjusted close price, nsa	SP500**

Note: GDP, INV and GOV are available on a quarterly basis, the remaining series are sampled at monthly frequency. The table contains information about transformations applied to the variables (“Transf.”); the publication delay or lag (“Pub-Lag”), whereby second releases are considered for *quarterly series implying that they get published halfway through the respective quarter; a description of each series (“Description”); and the corresponding FRED mnemonic of the St. Louis Fed (**index not available anymore in FRED, downloaded from the database of the Deutsche Bundesbank). Within our stylized release calendar framework we assume each month to contain exactly 22 working days. Data downloaded on the July 31st, 2017, close of business.

Table 5: Overview of Competing Models

Model	Description
MF-VAR	Mixed-frequency VAR à la Schorfheide and Song (2015)
TVi-MF-VAR	Mixed-frequency VAR with <i>only</i> time-varying intercepts
CSV-MF-VAR	Mixed-frequency VAR with <i>only</i> common stochastic volatility
TVi-CSV-MF-VAR	Mixed-frequency VAR with time-varying intercepts <i>and</i> common stochastic volatility

Table 6: Running Time Relative to the MF-VAR

Model	Relative Time
MF-VAR	1.00
TVi-MF-VAR	1.32
CSV-MF-VAR	1.88
TVi-CSV-MF-VAR	1.92

Table 7: Forecast Performance Relative to the MF-VAR for the Unemployment Rate

relative RMSFE												
Model/Horizon	-1	0	1	2	3	4	5	6	7	8	9	10
MF-VAR (abs.)		0.17	0.24	0.31	0.39	0.48	0.59	0.70	0.81	0.92	1.03	1.15
TVi-MF-VAR		0.92	0.89	0.87	0.85	0.83	0.85	0.86	0.87	0.87	0.87	0.87
CSV-MF-VAR		1.05	1.09	1.12	1.12	1.11	1.08	1.06	1.05	1.04	1.21	1.01
TVi-CSV-MF-VAR		0.98	0.97	0.98	0.96	0.93	0.91	0.90	0.88	0.87	0.86	0.85
relative logPL												
Model/Horizon	-1	0	1	2	3	4	5	6	7	8	9	10
TVi-MF-VAR		0.07	0.10	0.15	0.24	0.35	0.50	0.80	1.17	1.24	1.29	1.21
CSV-MF-VAR		-0.02	-0.01	0.03	0.13	0.25	0.47	0.65	0.94	0.71	0.82	0.75
TVi-CSV-MF-VAR		0.02	0.06	0.15	0.29	0.42	0.67	0.94	1.42	1.66	1.86	1.85

Note: For the unemployment rate the table displays root mean squared forecast errors (RMSFE) and log predictive likelihoods (logPL) of the TVi-MF-VAR, the CSV-MF-VAR, and the TVi-CSV-MF-VAR relative to the MF-VAR of [Schorfheide and Song \(2015\)](#). The absolute RMSFEs corresponding to the benchmark model are also displayed in the top row. The estimation samples underlying the analysis go from 1990:M1-2006:M2 to 1990:M1-2017:M7. As we compute forecasts at the end of each month, the values corresponding to the previous month are already available and thus the blank column at $h^{UNR} = -1$.

Table 8: Forecast Performance Relative to the MF-VAR for the Federal Funds Rate

Model/Horizon	relative RMSFE											
	-1	0	1	2	3	4	5	6	7	8	9	10
MF-VAR (abs.)			0.14	0.26	0.37	0.47	0.57	0.66	0.76	0.85	0.95	1.04
TVi-MF-VAR			0.94	0.95	0.95	0.96	0.96	0.95	0.95	0.95	0.96	0.97
CSV-MF-VAR			1.01	0.99	0.99	0.98	0.97	0.96	0.96	0.96	0.96	0.97
TVi-CSV-MF-VAR			0.95	0.93	0.93	0.93	0.92	0.91	0.91	0.91	0.92	0.93

Model/Horizon	relative logPL											
	-1	0	1	2	3	4	5	6	7	8	9	10
TVi-MF-VAR			0.09	0.11	0.08	0.09	0.14	0.08	0.13	0.06	0.14	0.14
CSV-MF-VAR			0.34	0.40	0.36	0.07	-0.24	-0.26	-0.39	-0.45	-0.53	-0.50
TVi-CSV-MF-VAR			0.37	0.41	0.55	0.48	0.52	0.50	0.50	0.42	0.46	0.53

Note: As we compute forecasts at the end of each month, the values corresponding to the current month are already available and thus the blank columns at $h^{FF} = -1, 0$. For the rest see Table 7.

Table 9: Forecast performance relative to the MF-VAR for GDP

Model/Horizon	relative RMSFE											
	-1	0	1	2	3	4	5	6	7	8	9	10
MF-VAR (abs.)	0.58	0.59	0.58	0.99	1.08	1.04	1.43	1.53	1.47	1.92	2.03	1.97
TVi-MF-VAR	0.98	0.99	1.07	0.99	0.98	1.07	0.95	0.97	1.04	0.92	0.93	0.99
CSV-MF-VAR	0.96	0.93	0.96	0.94	0.94	0.97	0.97	0.98	1.00	0.99	0.99	1.01
TVi-CSV-MF-VAR	0.92	0.89	0.95	0.95	0.91	1.00	0.95	0.94	0.99	0.96	0.94	0.99

Model/Horizon	relative logPL											
	-1	0	1	2	3	4	5	6	7	8	9	10
TVi-MF-VAR	0.06	0.15	-0.10	0.03	0.22	0.14	0.06	0.34	0.23	0.08	0.16	0.08
CSV-MF-VAR	0.13	0.19	0.09	0.03	0.21	0.22	0.02	0.30	0.19	0.01	0.07	0.05
TVi-CSV-MF-VAR	0.16	0.22	0.10	0.06	0.25	0.23	0.06	0.37	0.26	0.10	0.18	0.13

Note: Quarterly forecasts are obtained by averaging the respective monthly figures. For the rest see Table 7.