

Mikosch, Heiner; Neuwirth, Stefan

Working Paper

Real-time forecasting with a MIDAS VAR

KOF Working Papers, No. 377

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Mikosch, Heiner; Neuwirth, Stefan (2015) : Real-time forecasting with a MIDAS VAR, KOF Working Papers, No. 377, ETH Zurich, KOF Swiss Economic Institute, Zurich, <https://doi.org/10.3929/ethz-a-010414894>

This Version is available at:

<https://hdl.handle.net/10419/114451>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

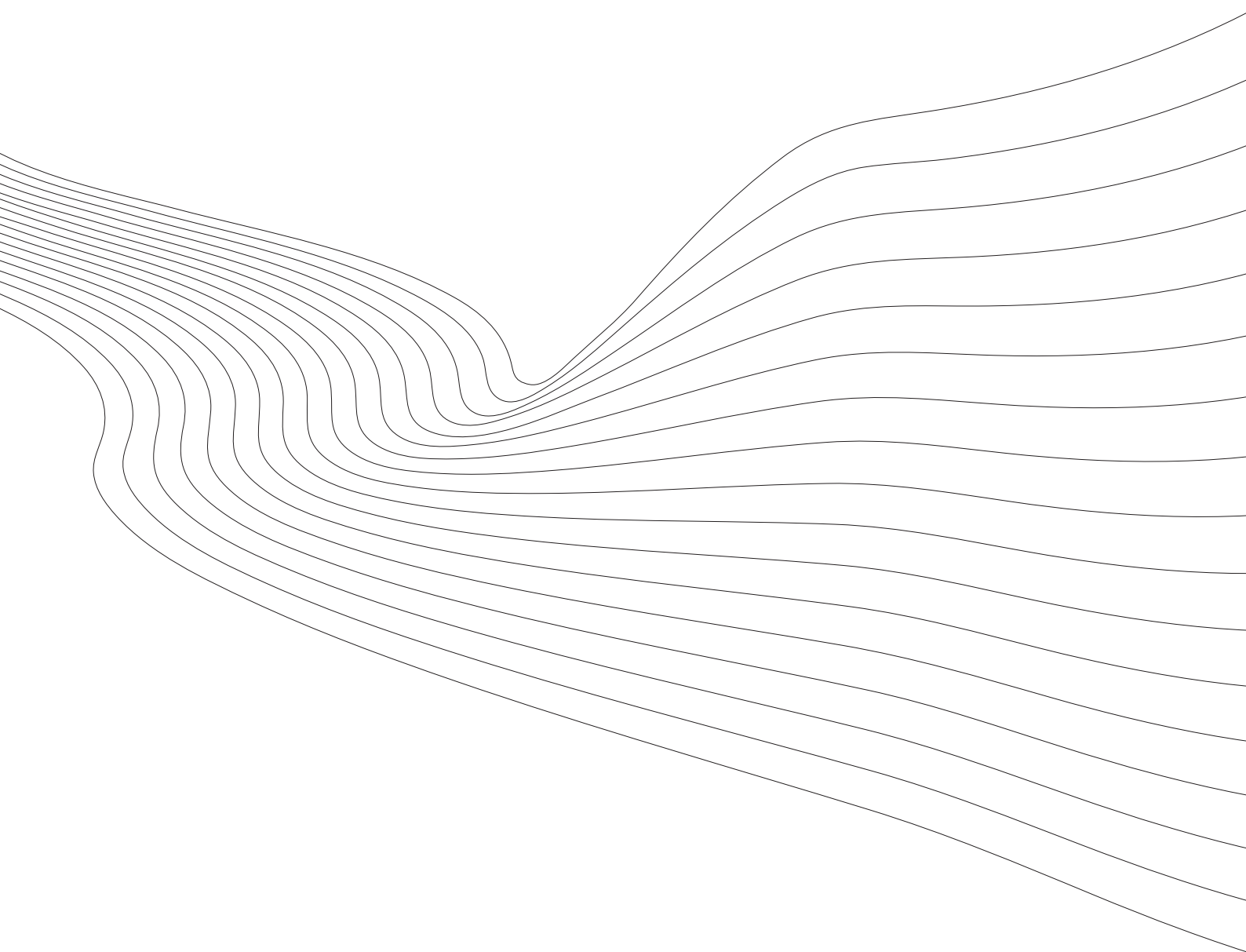
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

KOF Working Papers

Real-Time Forecasting with a MIDAS VAR

Heiner Mikosch and Stefan Neuwirth



KOF

KOF Swiss Economic Institute
ETH Zurich
LEE G 116
Leonhardstrasse 21
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Real-Time Forecasting with a MIDAS VAR

Heiner Mikosch and Stefan Neuwirth*

This version: April 5, 2015

Abstract

This paper presents a MIDAS type mixed frequency VAR forecasting model. First, we propose a general and compact mixed frequency VAR framework using a stacked vector approach. Second, we integrate the mixed frequency VAR with a MIDAS type Almon lag polynomial scheme which is designed to reduce the parameter space while keeping models flexible. We show how to recast the resulting non-linear MIDAS type mixed frequency VAR into a linear equation system that can be easily estimated. A pseudo out-of-sample forecasting exercise with US real-time data yields that the mixed frequency VAR substantially improves predictive accuracy upon a standard VAR for different VAR specifications. Forecast errors for, e.g., GDP growth decrease by 30 to 60 percent for forecast horizons up to six months and by around 20 percent for a forecast horizon of one year.

JEL classifications: C53, E27

Keywords: Forecasting, mixed frequency data, MIDAS, VAR, real time

*Both authors: KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21 LEE, CH-8092 Zurich, Switzerland. E-mails: mikosch@kof.ethz.ch and neuwirth@kof.ethz.ch. The authors thank Dirk Drechsel, Alexander Rathke, Samad Sarferaz, Frank Schorfheide, Christian Schumacher and Peter Zadrozny for helpful comments and suggestions. Part of the paper was written while Mikosch was a visiting researcher at Bank of Finland.

1 Introduction

Vector autoregression (VAR) models are a standard tool for macroeconomic forecasting due to their ease of use, their flexibility and the ability to produce coherent forecasts for multiple variables (Stock and Watson, 2001 and Karlsson, 2013, e.g.). A challenge to joining multiple macroeconomic variables in a VAR is that macroeconomic variables are usually sampled at different frequencies. For instance, GDP comes at a quarterly frequency whereas inflation is published monthly and short term interest rates are quoted at a daily or even higher frequency. The traditional solution is to simply time-aggregate all higher frequency series to the frequency of the lowest frequency series in the sample. A VAR including GDP, inflation and a interest rate will then be a quarterly frequency VAR (QFVAR).

However, such time aggregation comes at costs. First, any new data release, which occurs *within* the lowest frequency, can only be taken into account after the end of each lowest frequency period. To stick to the above example, a QFVAR does not allow to consider inter-quarterly inflation and interest rate releases before the end of a quarter. This delayed processing of information potentially impairs forecasts and nowcasts. Second, the time aggregation implies a peculiar constraint on the parameters attached to higher frequency variables which induces a possible loss of information. When time aggregating for instance a monthly flow variable to quarterly frequency all observations are weighted equally with one third. Different weights for the monthly observations could be more appropriate (i.e. the last month of a quarter could be more important than the first) but this cannot be incorporated in a single frequency setup.

To overcome the aforementioned drawbacks, we integrate a mixed frequency VAR (MFVAR) framework with MIXed DATA Sampling (MIDAS). The MIDAS approach has originally been developed by Ghysels and co-authors for forecasting with single equations.¹ Our paper makes the following contributions. First, we propose a stacked vector MFVAR framework which is general (multiple variables and multiple frequencies) but designed to still be compact and tractable. Previous expositions are limited to two frequencies for reasons of tractability (Ghysels, 2012 and the applications in Francis *et al.*, 2011 and Foroni *et al.*, 2014). Second, we augment the MFVAR with a MIDAS type non-linear Almon lag polynomial scheme (Almon, 1965). The Almon lag polynomial prevents overparametrization even with long lag lengths while, at the same time, keeping the regression model flexible. Comparisons between alternative MIDAS weight functions in a single equation context yielded that the Almon lag polynomial is a preferable choice (Mikosch and Zhang, 2014). Subsequently, we show how to transform the resulting stacked vector non-linear MIDAS type MFVAR into a linear equation system. The MIDAS parameters can then be easily estimated equation by equation using standard ordinary least squares (OLS). Due to this linear transformation the stacked vector MIDAS type MFVAR becomes feasible for estimation with multiple variables. Previous research was limited to stacked vector MIDAS type MFVARs with only few (mostly two) variables. The reason being that, in

¹See Ghysels *et al.* (2004), Ghysels *et al.* (2007), Andreou *et al.* (2010), e.g. Cf. the applications by Clements and Galvão (2008, 2009), Armesto *et al.* (2010), Drechsel and Scheufele (2012), Andreou *et al.* (2013), Ferrara *et al.* (2014), Mikosch and Zhang (2014), among many others. The MIDAS approach is related to earlier work on distributed lag models (Solow, 1960, Almon, 1965, Jorgenson, 1966, Tsurumi, 1971, Schmidt, 1974, Lütkepohl, 1981, Judge *et al.*, 1985, Mitchell and Speaker, 1986, among many others).

absence of a linear transformation, the non-linear MIDAS type MFVARs had to be estimated directly which is cumbersome, if not infeasible, in a VAR context involving more than two variables (provided one does not resort to auxiliary restrictions).

Ultimately, we conduct a pseudo out-of-sample forecast evaluation exercise using quarterly and higher frequency US real-time data. Our MFVAR substantially improves forecast accuracy upon a standard QFVAR for various different VAR specifications (three, six or twelve variables, three, six or twelve months of lagged information). Root mean squared forecast errors for GDP growth get reduced by 30 to 50 percent for forecast horizons up to six months and by about 20 percent for a forecast horizon of one year. For inflation the gains are even bigger with improvements of up to 90 percent in the very short run. For forecasts of two years ahead our model improves forecasts upon a QFVAR by around 20 percent. Even bigger are the improvements when forecasting the short term interest rate where the MFVAR is constantly better than a QFVAR for all forecast horizons. Further, we find that augmentation of a stacked vector MFVAR with an Almon lag polynomial scheme has a distinct advantage for specifications with longer lag structures. While a MFVAR with fully unrestricted parameters still yields considerable forecast improvements over a standard QFVAR when using few lags, these improvements vanish almost completely for longer lags. An Almon augmented MFVAR instead still yields high improvements over a QFVAR for longer lags.

This paper relates to a small but growing literature on mixed frequency VAR models which fall in two basic classes.² The first class uses a state space approach. The general idea is to conceptually assume the mixed frequencies away by reformulating each lower frequency series as a partially latent high frequency series. The Kalman filter or, in a Bayesian context, the Gibbs sampler then provide the possibility to estimate the partially latent VAR process. See [Zadrozny \(1988, 1990\)](#), [Kuzin *et al.* \(2011\)](#), [Bai *et al.* \(2013\)](#) and [Forni and Marcellino \(2014\)](#) for state space type MFVAR models using a non-Bayesian version of the Kalman filter, [Mariano and Murasawa \(2010\)](#) for a state space type MFVAR using the EM algorithm, and [Chiu *et al.* \(2012\)](#) and [Schorfheide and Song \(2013\)](#) for state space type MFVARs using the Gibbs sampler.³

The second class – to which our paper belongs – employs a stacking approach: all variable observations, that pertain to the same lowest frequency period in the sample, are stacked into one vector. For instance, when the sample includes quarterly GDP, monthly inflation and a daily interest rate, the quarterly frequency is the lowest frequency in the sample and, thus, the stacked vector for each quarter t includes one quarterly GDP observation, three monthly inflation observations and 60 daily interest rate observations (assuming a month has 20 trading days). Based on the stacked vectors a VAR process of any desired order can then be modeled. Further, in order to achieve parsimony, distributed lag or MIDAS style polynomial structures can then be imposed on the parameter space of the stacked vector type MFVAR, resulting in a stacked vector MIDAS type MFVAR. The MIDAS approach aims to reduce the parameter space while keeping models flexible. [Ghysels \(2012\)](#) provides a rich exposition of a stacked vector MIDAS type MFVAR model. Our

²Following [Cox \(1981\)](#) the first class may be referred to as parameter-driven, while the second approach may be qualified as observation-driven.

³Important contributions on non-VAR state space models are [Mariano and Murasawa \(2003\)](#), [Aruoba *et al.* \(2009\)](#), [Giannone *et al.* \(2008\)](#), e.g.

model is similar to Ghysels’ model in some respects and different in other respects (see Appendix 1 for a comparison). Building on Ghysels (2012), Francis *et al.* (2011) employ a stacked vector MIDAS type MFVAR to study the effects of monetary policy shocks on macroeconomic aggregates using daily interest rate data and monthly macroeconomic indicators. Foroni *et al.* (2014) introduce Markov switching to both a state space type and a stacked vector MIDAS type bivariate MFVAR model and compare these (and several other bivariate) models in terms of predictive accuracy.

The remainder of the paper is structured as follows: Section 2 presents our stacked vector MIDAS type mixed frequency VAR framework. In particular, we show how to recast the stacked vector MIDAS type non-linear MFVAR into a linear equation system that can be easily estimated equation by equation using standard OLS. Subsequently, we analyze whether our stacked vector MIDAS type mixed frequency VAR is helpful for forecasting in real time. Section 3 describes the pseudo out-of-sample forecast evaluation set up and the real-time data used, and Section 4 presents the results of the empirical assessment. Finally, Section 5 provides conclusions and possible directions of further research.

2 A general MIDAS VAR framework

Let there be a set of time series variables of different frequencies. Denote each variable by $y_{i,t-1+\tau_i/\mathcal{T}_i}$ where i is the variable subscript with $i = 1, \dots, I$ and $t - 1 + \tau_i/\mathcal{T}_i$ is the time period subscript.⁴ $t = 1, \dots, T$ denotes the lowest frequency periods such that \mathcal{T}_i is the frequency of the i -th variable in each period t , and $\tau_i = 1, \dots, \mathcal{T}_i$ denotes the subperiods of the i -th variable in each period t . For instance, when the set of variables comprises quarterly, monthly and daily variables, $t = 1, \dots, T$ denotes the quarters, $\mathcal{T}_i = 1/3/90$ is the frequency of any quarterly/monthly/daily variable in each quarter t and $\tau_i = 1/1, 2, 3/1, \dots, 90$ denotes the quarter/months/days in each quarter t (assuming that a quarter has 90 days).

For ease of exposition we assume in this section that each variable $y_{i,t-1+\tau_i/\mathcal{T}_i}$ with $i \in \{1, \dots, I\}$ is released simultaneously with each variable $y_{j,t-1+\tau_j/\mathcal{T}_j}$ with $j \in \{1, \dots, I\}$ and with $t - 1 + \tau_j/\mathcal{T}_j = t - 1 + \tau_i/\mathcal{T}_i$. Further, we assume for ease of exposition that each variable $y_{i,t-1+\tau_i/\mathcal{T}_i}$ is released directly after the end of period $t - 1 + \tau_i/\mathcal{T}_i$ (no ragged edge). For instance, any quarterly variable is released directly after the end of a quarter simultaneously with any other quarterly variable and with any monthly variable observation on the third month of that quarter. Our empirical application presented in Section 4 explicitly models ragged edges and delayed releases.

Appendix 2 illustrates the MIDAS VAR framework with the simple case of just one monthly and one quarterly variable.

⁴We model the time period subscript as $t - 1 + \tau_i/\mathcal{T}_i$ in order to let τ_i denote the τ_i -th subperiod in period t . Clements and Galvão (2008, 2009), e.g., model the time period subscript as $t - \tau_i/\mathcal{T}_i$ so that τ_i denotes the $(\mathcal{T}_i - \tau_i)$ -th subperiod. Both variants are equally possible.

The stacking approach

We stack all observations of the i -th variable in period $t - p$ with $p \in \mathbb{N}_0$ starting with the latest observation and ending with the earliest observation to get

$$\mathbf{y}_{i,t-p}^{\mathcal{T}_i \times 1} \equiv \begin{bmatrix} y_{i,t-p-1+\frac{\mathcal{T}_i}{\tau_i}} \\ \vdots \\ y_{i,t-p-1+\frac{1}{\tau_i}} \end{bmatrix}. \quad 5$$

Doing this for all I variables yields I variable vectors for period $t - p$. Next, we stack these I variable vectors to get

$$\mathbf{y}_{t-p}^{\sum_{i=1}^I \mathcal{T}_i \times 1} \equiv \begin{bmatrix} \mathbf{y}_{1,t-p} \\ \vdots \\ \mathbf{y}_{I,t-p} \end{bmatrix}.$$

Further, we lag each element in the variable vector $y_{i,t-p}$ by the period $\frac{1}{\tau_i}$ to get

$$\mathbf{y}_{i,t-p-\frac{1}{\tau_i}}^{\mathcal{T}_i \times 1} \equiv \begin{bmatrix} y_{i,t-p-1+\frac{\mathcal{T}_i}{\tau_i}-\frac{1}{\tau_i}} \\ \vdots \\ y_{i,t-p-1+\frac{1}{\tau_i}-\frac{1}{\tau_i}} \end{bmatrix}.$$

Doing this for all periods $t - p$ with $p = 0, 1, \dots, P$ results in $P + 1$ vectors for the i -th variable. We stack these $P + 1$ vectors to get

$$\mathbf{x}_{i,t}^{(P+1) \cdot \mathcal{T}_i \times 1} \equiv \begin{bmatrix} \mathbf{y}_{i,t-0-\frac{1}{\tau_i}} \\ \mathbf{y}_{i,t-1-\frac{1}{\tau_i}} \\ \vdots \\ \mathbf{y}_{i,t-P-\frac{1}{\tau_i}} \end{bmatrix}.$$

$\mathbf{x}_{i,t}$ contains all past observations of the i -th variable until (and including) period $t - P$. For ease of exposition we set $P_i = P$ for all variables. Our empirical application presented in Section 4 allows P_i to be different for each variable. Iterating the last two stacking steps for all variables leaves us with I variable vectors $\mathbf{x}_{1,t}, \dots, \mathbf{x}_{I,t}$. In a last step, we stack these I variable vectors to get the data vector

$$\mathbf{x}_t^{(P+1) \cdot \sum_{i=1}^I \mathcal{T}_i \times 1} \equiv \begin{bmatrix} \mathbf{x}_{1,t} \\ \vdots \\ \mathbf{x}_{I,t} \end{bmatrix}. \quad 6$$

Stacked vector mixed frequency VAR

The stacked vectors from the previous subsection build the basis for our stacked vector mixed frequency VAR framework. In particular, we model each element of vector \mathbf{y}_t ,

⁵In case of the lowest frequency variable, \mathbf{y}_{t-p} consists of just one observation.

⁶Stacking of lagged variables into a vector \mathbf{x}_t follows Hamilton (1994, p. 292) and Hayashi (2000, p. 397).

namely each $y_{i,t-1+\frac{\tau_i}{T_i}}$, as a regression function of the data vector \mathbf{x}_t :

$$\begin{aligned} y_{i,t-1+\frac{\tau_i}{T_i}} &= \mathbf{a}_{i,\tau_i} \mathbf{x}_t + \epsilon_{i,t-1+\frac{\tau_i}{T_i}} \\ &= \mathbf{a}_{i,\tau_i,1} \dots \mathbf{a}_{i,\tau_i,I} \begin{bmatrix} \mathbf{x}_{1,t} \\ \vdots \\ \mathbf{x}_{I,t} \end{bmatrix} + \epsilon_{i,t-1+\frac{\tau_i}{T_i}}. \end{aligned} \quad (1)$$

$\epsilon_{i,t-1+\frac{\tau_i}{T_i}}$ is an error term, and $\mathbf{a}_{i,\tau_i,1}, \dots, \mathbf{a}_{i,\tau_i,I}$ are selection vectors defined as

$$\mathbf{a}_{i,\tau_i,j} \equiv [\mathbf{0}_{i,\tau_i,j,1} \quad \boldsymbol{\alpha}_{i,\tau_i,j} \quad \mathbf{0}_{i,\tau_i,j,2}] \quad (2)$$

$1 \times (P+1) \cdot \mathcal{T}_j$

which relates period τ_i -observation of the i -th variable to all observations of the j -th variable with $j = 1, \dots, I$.⁷ The three parts of $\mathbf{a}_{i,\tau_i,j}$ are explained in turn.

First, $\mathbf{0}_{i,\tau_i,j,1}$ denotes a row vector of zero parameters with length $\mathcal{T}_j - \lceil \frac{\tau_i}{T_i} \mathcal{T}_j \rceil$.⁸ This vector of zero parameters relates $y_{i,t-1+\tau_i/T_i}$ to those observations of the j -th variable which are published simultaneously with or later than $y_{i,t-1+\tau_i/T_i}$, i.e. in periods $t-1+\tau_i/T_i, \dots, t-1+(\mathcal{T}_i-1)/T_i$. The reason for setting these relations to zero is as follows: Any observation which is published simultaneously or later than observation $y_{i,t-1+\tau_i/T_i}$ cannot be used for forecasting observation $y_{i,t-1+\tau_i/T_i}$. Hence, these simultaneous or future observations must be disregarded during the parameter estimation stage. This can be done by discarding these observations, or by setting their parameters to zero.

Second, $\boldsymbol{\alpha}_{i,\tau_i,j}$ denotes a row vector of parameters with length $K_{i,j} \equiv P \cdot \mathcal{T}_j + \lceil \frac{\mathcal{T}_j}{T_i} \rceil$:

$$\boldsymbol{\alpha}_{i,\tau_i,j} \equiv [\alpha_{i,\tau_i,j,1} \quad \dots \quad \alpha_{i,\tau_i,j,K_{i,j}}]. \quad (3)$$

$1 \times K_{i,j}$

This parameter vector relates $y_{i,t-1+\tau_i/T_i}$ to past observations of the j -th variable from period $\mathcal{T}_j - \lceil \frac{\tau_i}{T_i} \mathcal{T}_j \rceil + 1$ until period $\mathcal{T}_j - \lceil \frac{\tau_i}{T_i} \mathcal{T}_j \rceil + K_{i,j}$. $\boldsymbol{\alpha}_{i,\tau_i,j}$ will be further discussed in the next subsection.

Third, $\mathbf{0}_{i,\tau_i,j,2}$ denotes a row vector of zero parameters with length $\lceil \frac{\tau_i}{T_i} \mathcal{T}_j \rceil - \lceil \frac{\mathcal{T}_j}{T_i} \rceil$. This vector of zero parameters relates $y_{i,t-1+\tau_i/T_i}$ to those past observations of the j -th variable which have been published earlier than period $\mathcal{T}_j - \lceil \frac{\tau_i}{T_i} \mathcal{T}_j \rceil + K_{i,j}$. Given the choice of lag length P these observations are too “old” to be considered for forecasting observation $y_{i,t-1+\tau_i/T_i}$. Hence, these observations must be disregarded during the parameter estimation stage. Again, this can be done by discarding these observations, or by setting their parameters to zero.

In a next step, we stack the selection vectors $\mathbf{a}_{i,1,j}, \dots, \mathbf{a}_{i,\mathcal{T}_i,j}$ to get

$$A_{i,j} \equiv \begin{bmatrix} \mathbf{a}_{i,\mathcal{T}_i,j} \\ \vdots \\ \mathbf{a}_{i,1,j} \end{bmatrix}.$$

$\mathcal{T}_i \times \mathcal{T}_j \cdot (P+1)$

⁷ $\tau_i = j$ or $i \neq j$.

⁸The ceiling function $\lceil z \rceil \equiv \min\{n \in \mathbb{Z} | n \geq z\}$, where z is a real number and n is an integer from the set of all integers \mathbb{Z} .

$A_{i,j}$ contains the (zero and alpha) parameters which relate the I observations of the i -th variable in period t to $x_{j,t}$, i.e. to all observations of the j -th variable from period $t - P - 1$ to period $t - \frac{1}{\mathcal{T}_i}$. Then, we collect all I times I matrices $A_{i,j}$ in the big matrix

$$\underset{\sum_{i=1}^I \mathcal{T}_i \times (P+1) \cdot \sum_{j=1}^I \mathcal{T}_j}{A} \equiv \begin{bmatrix} A_{1,1} & \dots & A_{1,I} \\ \vdots & \ddots & \vdots \\ A_{I,1} & \dots & A_{I,I} \end{bmatrix}.$$

Iterating Equation (2) for all \mathcal{T}_i observations of the i -th variable in period t , namely for $y_{i,t-1+1/\mathcal{T}_i}, \dots, y_{i,t-1+\mathcal{T}_i/\mathcal{T}_i}$, yields \mathcal{T}_i error terms $\epsilon_{i,t-1+1/\mathcal{T}_i}, \dots, \epsilon_{i,t-1+\mathcal{T}_i/\mathcal{T}_i}$. We stack these error terms to get

$$\underset{\mathcal{T}_i \times 1}{\boldsymbol{\epsilon}_{i,t}} \equiv \begin{bmatrix} \epsilon_{i,t-1+\frac{\mathcal{T}_i}{\mathcal{T}_i}} \\ \vdots \\ \epsilon_{i,t-1+\frac{1}{\mathcal{T}_i}} \end{bmatrix}.$$

Further iterating the above error term stacking for all variables gives I error term vectors $\boldsymbol{\epsilon}_{1,t} \dots \boldsymbol{\epsilon}_{I,t}$. We again stack these error term vectors to get

$$\underset{\sum_{i=1}^I \mathcal{T}_i \times 1}{\boldsymbol{\epsilon}_t} \equiv \begin{bmatrix} \boldsymbol{\epsilon}_{1,t} \\ \vdots \\ \boldsymbol{\epsilon}_{I,t} \end{bmatrix}.$$

The mixed frequency vector autoregressive (VAR) process then generally writes

$$\begin{bmatrix} \mathbf{y}_{1,t} \\ \vdots \\ \mathbf{y}_{I,t} \end{bmatrix} = \begin{bmatrix} A_{1,1} & \dots & A_{1,I} \\ \vdots & \ddots & \vdots \\ A_{I,1} & \dots & A_{I,I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1,t} \\ \vdots \\ \mathbf{x}_{I,t} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_{1,t} \\ \vdots \\ \boldsymbol{\epsilon}_{I,t} \end{bmatrix},$$

or more compactly

$$\mathbf{y}_t = A\mathbf{x}_t + \boldsymbol{\epsilon}_t. \quad (4)$$

Unrestricted and Almon MIDAS type mixed frequency VAR

The core building blocks of the big matrix A are the parameter vectors $\boldsymbol{\alpha}_{i,\tau_i,j}$. $\boldsymbol{\alpha}_{i,\tau_i,j}$ has been generally defined in Equation (3) and relates $y_{i,t-1+\tau_i/\mathcal{T}_i}$ to all past observations of the j -th variable from period $\mathcal{T}_j - \lceil \frac{\tau_i}{\mathcal{T}_i} \mathcal{T}_j \rceil + 1$ until period $\mathcal{T}_j - \lceil \frac{\tau_i}{\mathcal{T}_i} \mathcal{T}_j \rceil + K_{i,j}$.⁹ We model the elements of $\boldsymbol{\alpha}_{i,\tau_i,j}$ in two alternative ways.

First, an intuitive strategy is to regard each element of $\boldsymbol{\alpha}_{i,\tau_i,j}$ as an unrestricted parameter. Equation (2) is then a linear regression equation with $\sum_{j=1}^I (\mathcal{T}_j - \lceil \frac{\tau_i}{\mathcal{T}_i} \rceil)$ zero restrictions and with $\sum_{j=1}^I K_{i,j}$ unknown unrestricted parameters. The unrestricted parameters in

⁹In case $j = i$ the relation is between observations of the same variable, in case $j \neq i$ the relation is between observations of two different variables.

matrix A of Equation (4) can thus be easily estimated row by row via ordinary least squares (OLS). The resulting unrestricted MIDAS (U-MIDAS) VAR is very flexible in terms of parametrization, but gets overparameterized/looses parsimony as the number of employed lags grows. [Foroni *et al.* \(2012\)](#) originally proposed the U-MIDAS approach for single equation forecasting models.

Second, following Ghysels and co-authors ([Ghysels *et al.*, 2007](#), [Andreou *et al.*, 2010](#), e.g.) one can regard $\alpha_{i,\tau_i,j}$ as a vector of unknown weights $\alpha_{i,\tau_i,j,1}, \dots, \alpha_{i,\tau_i,j,K_{i,j}}$, where each weight is a function of the unknown parameter vector $\theta_{i,\tau_i,j}$ and of the lag index k with $k = 1, \dots, K_{i,j}$. Accordingly, Equation (4) can be specified as

$$\mathbf{y}_t = A(\Theta, \mathcal{K})\mathbf{x}_t + \boldsymbol{\epsilon}_t, \quad (5)$$

where Θ denotes the space of Almon parameter vectors and \mathcal{K} is the lag index space. We model each weight function as a Almon lag polynomial,

$$\alpha_{i,\tau_i,j,k} = \alpha_{i,\tau_i,j,k}(\boldsymbol{\theta}_{i,\tau_i,j}, k) = \alpha_{i,\tau_i,j,k}(\theta_{i,\tau_i,j,0}, \dots, \theta_{i,\tau_i,j,Q}, k) = \sum_{q=0}^Q \theta_q k^q, \quad (6)$$

where $Q \in \mathbb{N}$ denotes the polynomial order.¹⁰

The Almon lag polynomial combines three features which makes it particularly attractive for a stacked vector mixed frequency VAR. The first feature concerns parameter parsimony: The polynomial order Q determines the number of θ -parameters to be estimated. By setting $Q \ll K_{i,j}$ the number of parameters to be estimated gets reduced. This prevents parameter proliferation or overfitting even for a long lag length $K_{i,j}$. The second feature concerns model flexibility: Even when Q is a small number the form of the function $\alpha_{i,\tau_i,j,k}$ with respect to k – i.e. the relative importance of any lagged observation as compared to any other lagged observation – is flexible and depends on the Almon parameter values $\boldsymbol{\theta}_{i,\tau_i,j}$. Figure 1 illustrates this graphically. In practice, $\boldsymbol{\theta}_{i,\tau_i,j}$ is estimated and, thus, the relative weight of each lagged observation is optimally chosen by the data.

– Figure 1 here –

The third feature concerns estimation: Since each Almon lag polynomial weight function $\alpha_{i,\tau_i,j,k}$ is highly non-linear in its parameters $\boldsymbol{\theta}_{i,\tau_i,j}$, OLS is not feasible for a direct estimation of the θ -parameters in matrix A in Equation (5). Further, since each row of A contains a multitude of Almon lag polynomial weight functions, $\alpha_{i,\tau_i,1,1}, \dots, \alpha_{i,\tau_i,1,K}, \alpha_{i,\tau_i,2,1}, \dots, \alpha_{i,\tau_i,2,K_{i,j}}, \dots, \dots, \alpha_{i,\tau_i,I,1}, \dots, \alpha_{i,\tau_i,I,K}$, even non-linear optimization methods reach their limits in practice. Specifically, each row of A contains $\sum_{j=1}^I K_{i,j}$ Almon lag polynomial weight functions. Fortunately, as we show in the next subsection, when the weight functions are modeled as Almon lag polynomials Equation (5) can be recast such that OLS estimation of the θ -parameters becomes feasible.

In sum, Almon MIDAS brings together three goals which are often in a trade-off position to each other: parsimony in terms of parametrization, model flexibility and feasibility concerning estimation.

¹⁰The literature knows yet other weight functions (see [Ghysels *et al.*, 2007](#), e.g.)

Recasting the Almon MIDAS type MFVAR in linear form

The Almon MIDAS type MFVAR model is both parsimonious and flexible. However, the model contains a multitude of highly non-linear terms which makes a direct estimation infeasible. In the following we show how to recast the Almon MIDAS type MFVAR model in a linear equation system which can be easily estimated via OLS.

Equation (6) has defined the Almon parameter vector as

$$\underset{1 \times (Q+1)}{\boldsymbol{\theta}_{i,\tau_i,j}} \equiv \left[\theta_{i,\tau_i,j,0} \ \dots \ \theta_{i,\tau_i,j,Q} \right]. \quad (7)$$

We stack the series of Almon parameter vectors $\boldsymbol{\theta}_{i,\tau_i,1}, \dots, \boldsymbol{\theta}_{i,\tau_i,I}$ into

$$\underset{1 \times I \cdot (Q+1)}{\boldsymbol{\theta}_{i,\tau_i}} \equiv \left[\boldsymbol{\theta}_{i,\tau_i,1} \ \dots \ \boldsymbol{\theta}_{i,\tau_i,I} \right]. \quad (8)$$

Further, we define a transformation matrix

$$M_{i,\tau_i,j} \equiv \underset{(Q+1) \times (P+1) \cdot \mathcal{T}_j}{\begin{bmatrix} \mathbf{0}_{i,\tau_i,j,1} & 1 & 1 & 1 & \dots & 1 & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & 1 & 2 & 3 & \dots & K_{i,j} & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & 1 & 4 & 9 & \dots & K_{i,j}^2 & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & \cdot & \cdot & \cdot & \dots & \cdot & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & \cdot & \cdot & \cdot & \dots & \cdot & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & \cdot & \cdot & \cdot & \dots & \cdot & \mathbf{0}_{i,\tau_i,j,2} \\ \mathbf{0}_{i,\tau_i,j,1} & 1 & 2^Q & 3^Q & \dots & K_{i,j}^Q & \mathbf{0}_{i,\tau_i,j,2} \end{bmatrix}}$$

and condense the series of transformation matrices $M_{i,\tau_i,1}, \dots, M_{i,\tau_i,J}$ into the block matrix

$$M_{i,\tau_i} \equiv \underset{I \cdot (Q+1) \times (P+1) \cdot \sum_{j=1}^I \mathcal{T}_j}{diag}(M_{i,\tau_i,1}, \dots, M_{i,\tau_i,I}). \quad (9)$$

Next, we premultiply the data vector $\mathbf{x}_{i,t}$ with the block matrix to get a row vector of transformed data

$$\underset{I \cdot (Q+1) \times 1}{\mathbf{x}_{i,\tau_i,t}^*} \equiv M_{i,\tau_i} \mathbf{x}_t. \quad (10)$$

Notably, M_{i,τ_i} and, hence, also $\mathbf{x}_{i,\tau_i,t}^*$ is specific to the subperiod τ_i -observation of the i -th variable. Using Equation (10) and the fact that

$$\mathbf{a}_{i,\tau_i} = \boldsymbol{\theta}_{i,\tau_i} M_{i,\tau_i}$$

we can rewrite the basic regression equation (2) as

$$y_{i,t-1+\frac{\tau_i}{T_i}} = \boldsymbol{\theta}_{i,\tau_i} \mathbf{x}_{i,\tau_i,t}^* + \epsilon_{i,t-1+\frac{\tau_i}{T_i}}. \quad (11)$$

We stack the Almon parameter vectors $\boldsymbol{\theta}_{i,1}, \dots, \boldsymbol{\theta}_{i,\tau_i}$ defined in Equation (8) to get

$$\underset{\mathcal{T}_i \times I \cdot (Q+1)}{\Theta_i} \equiv \begin{bmatrix} \boldsymbol{\theta}_{i,\tau_i} \\ \vdots \\ \boldsymbol{\theta}_{i,1} \end{bmatrix},$$

Doing this for each variable index $i = 1, \dots, I$ yields the series $\Theta_1, \dots, \Theta_I$ which can be stacked to the big Almon parameter matrix

$$\Theta_{\sum_{i=1}^I \mathcal{T}_i \times I \cdot (Q+1)} \equiv \begin{bmatrix} \Theta_1 \\ \vdots \\ \Theta_I \end{bmatrix}. \quad (12)$$

Further, upon transposition we stack all transformed data vectors specific to the i -th variable in period t starting with the latest subperiod and ending with the earliest subperiod to get

$$X_{i,t}^*_{\mathcal{T}_i \times I \cdot (Q+1)} \equiv \begin{bmatrix} \mathbf{x}'_{i,\mathcal{T}_i,t} \\ \vdots \\ \mathbf{x}'_{i,1,t} \end{bmatrix}.$$

Iterating the last stacking step for all variables leaves us with I matrices $X_{1,t}^*, \dots, X_{I,t}^*$. We stack these matrices again to get the transformed data matrix

$$X_t^*_{\sum_{i=1}^I \mathcal{T}_i \times I \cdot (Q+1)} \equiv \begin{bmatrix} X_{1,t}^* \\ \vdots \\ X_{I,t}^* \end{bmatrix}. \quad (13)$$

Next, we define a selection matrix

$$S_{L \times L \cdot L} \equiv (\mathbf{I}_{(L)} * \mathbf{I}_{(L)})'$$

where $\mathbf{I}_{(L)}$ denotes the identity matrix of size $L \equiv \sum_{i=1}^I \mathcal{T}_i$ and where “*” is the (column-wise) Khatri-Rao product. I. e.

$$\mathbf{I} * \mathbf{I} \equiv [\mathbf{I}_l \otimes \mathbf{I}_l]_l, \quad (14)$$

where l denotes the l -th column of \mathbf{I} with $l = 1, \dots, L$.¹¹

As is easily seen, the stacked vector MIDAS type non-linear MFVAR from Equation (5) can then be reformulated as

$$\mathbf{y}_t = S \text{vec}(\Theta X_t^{*'}) + \boldsymbol{\epsilon}_t. \quad (15)$$

Using the fact that for any two matrices B with size $R \times T$ and C with size $U \times V$

$$\text{vec}(BC) = (C' \otimes I_{(R)}) \text{vec}(B),$$

we further rewrite Equation (15) as a linear equations systems

$$\mathbf{y}_t = S(X_t^* \otimes I_{(L)}) \text{vec}(\Theta) + \boldsymbol{\epsilon}_t. \quad (16)$$

$S(X_t^* \otimes I_{(L)})$ with size $L \times I \cdot (Q + 1) \cdot L$ builds the right-hand side data matrix and $\text{vec}(\Theta)$ with size $I \cdot (Q + 1) \cdot L \times 1$ is the parameter vector. Equation (16) conforms to the standard VAR matrix notation in the sense that each element, $y_{i,t-1+\tau_i/\mathcal{T}_i}$, in the

¹¹See [Khatri and Rao \(1968\)](#).

left-hand side variable vector, \mathbf{y}_t , is a function of always the same right-hand side data matrix, namely of $S(X_t^* \otimes I_{(L)})$. On the other hand, Equation (16) deviates from the standard VAR matrix notation in that the data come in matrix form (but not in vector form) while the parameters come in vector form (but not in matrix form).

Importantly, the θ -parameters in Equation (16) can be easily estimated row by row via OLS. As each row of $S(X_t^* \otimes I_{(L)})$ contains only $I \cdot (Q + 1)$ non-zero elements, a single row can only be used to estimate $I \cdot (Q + 1)$ θ -parameters in $\text{vec}(\Theta)$. Hence, all rows in Equation (16) are needed to fully estimate the $I \cdot (Q + 1) \cdot L$ parameters in $\text{vec}(\Theta)$. Upon estimation of $\text{vec}(\Theta)$, the α -weights in matrix A of the general stacked vector MFVAR equation (4) can be calculated using Equation (6).

3 Data and forecast evaluation set up

Can our stacked vector MIDAS type mixed frequency VAR help improve forecasts upon a standard single frequency VAR? In order to answer this question we conduct a pseudo-out-of sample forecast evaluation with US real time data.

The real-time data used in our analysis comprise quarterly real seasonal adjusted GDP as well as the monthly consumer price index, industrial production and housing starts. Data sources are the Archival Federal Reserve Economic Data (ALFRED) published by the Federal Reserve Bank of St. Louis and the Federal Reserve Bank of Philadelphia Real-Time Data Set for Macroeconomists (RTDSM). In addition, we employ a set of time series that are not subject to data revisions: the ISM indices for manufacturing and supplier delivery times, S&P 500 stock market index, 3-month treasury bill yield, 10-year treasury bond yield, and average weekly hours worked by production and supervisory workers. Following common practice we time-aggregate all variables with a higher than monthly frequency by using their end-of-month values (Carriero *et al.*, 2012 and Schorfheide and Song, 2013, e.g.) Table 1 provides precise data definitions. We use different data transformations for our forecast evaluation, namely year-on-year growth rates, 3-month growth rates as well as month-on-month growth rates.

– Table 1 here –

The real-time dataset comprises 344 vintages covering the time frame January 1970 to August 2013. The first vintage starts in January 1970 and includes all data available until the end of January 1985. In order to use a rolling window setup each following vintage is shifted by one month, i.e. the second vintage starts in February 1970 and includes all data available until the end of February 1985, and so on.

We forecast at the end of each month m . In a first step, we estimate the model using the first vintage and forecast all monthly and quarterly variables 1 to 24 months ahead of release. We then re-estimate the model using the second vintage and generate again 1 to 24-months-ahead forecasts. This procedure gets iterated until the 315th vintage, which starts in August 1995 and includes all data available until end of August 2010, resulting for each variable in a series of h -months ahead forecasts with $h = 1, \dots, 24$ months ahead of

the actual realization of the variable.¹² Forecast errors are then calculated by subtracting each forecast from its actual realization. Finally, the forecast errors are used to calculate a root mean square forecast error (RMSFE) for each variable and each forecast horizon h , $RMSFE_{i,h}$.

The choice of actual realizations is a delicate issue in a real time data context (cf. the discussions in [Croushore, 2006](#), [Romer and Romer, 2000](#), and [Sims, 2002](#)). There have been several benchmark revisions in the time series we use; the latest revision for GDP occurred in mid-2014 and included a substantial redefinition of gross fixed capital formation which accounts for 20 percent of GDP. A forecaster in, e.g., 1985 could not have predicted such a definition change. Thus we follow [Romer and Romer \(2000\)](#), [Faust and Wright \(2009\)](#) and [Carriero *et al.* \(2012\)](#) and use the second estimate of quarterly GDP as the actual realization with which we compare our forecasts. We use this principle also for the monthly variables, i.e. we use the second estimate of CPI when calculating the forecast error. As a robustness test we also check the forecast performance with the first and third estimate of the variables.

In order to reproduce the available information a forecaster would have had at each forecast date, i.e. at the end of each month m , we need to take differing publication lags – so called ragged edges – into account. For most monthly variables the (first) release for each month is not available directly at the end of that month, but is published with a time lag of up to one month. We call these variables lag variables. Hence, when forecasting at the end of a month m the aforementioned releases cannot be used for forecasting but must be backcasted themselves using all information available until the end of m . Only the S&P 500 index, the 10-year treasury bond yield and the 3-month treasury bill yield are available directly at the end of each month and, hence, can be used for back-, now- and forecasting. The latter three variables are released daily and are never revised. As stated above, we time-aggregate the variables to monthly frequency. Equally, for the quarterly variable, namely GDP, the (first) release for each quarter q is published with a time lag of up to one month. Consequently, when forecasting at the end of a quarter q the releases for quarter q cannot be used for forecasting but must be backcasted using all information available until the end of q .

Our benchmark model is a standard quarterly frequency VAR (QFVAR) where all higher than quarterly frequency series are time-aggregated to quarterly frequency. Our choice of benchmark model is motivated by the great popularity of VARs as a workhorse for macroeconomic forecasting ([Karlsson, 2013](#)). VAR models are a standard tool for forecasting in many policy-oriented institutions and organisations due to their ease of use, their flexibility and the ability to produce coherent forecasts for multiple variables. Thus, our aim is to produce a forecasting model which improves upon the long proven method of VARs.¹³

¹² h -months ahead, hence, inter-quarterly forecasts are also generated for quarterly GDP.

¹³It might be interesting to compare the MIDAS VAR with a state space type mixed frequency VAR or with a variety of alternative forecasting models. However, such a comparative analysis goes beyond the scope of this paper.

Forecast performance for the i -th variable at forecast horizon h is measured by the difference between the RMSFE of the MIDAS VAR and the QFVAR in percent of the RMSFE of the QFVAR,

$$\Delta RMSFE_{i,h} = 100 * \left(\frac{RMSFE_{i,h}^{MIDAS\ VAR} - RMSFE_{i,h}^{QFVAR}}{RMSFE_{i,h}^{QFVAR}} \right).$$

We refer to $\Delta RMSFE_{i,h}$ as the relative change in the RMSFE. The *more negative* $\Delta RMSFE_{i,h}$ is, the *better* performs the MIDAS VAR relative to the QFVAR in terms of predictive power. MIDAS VAR and benchmark QFVAR are always specified such that they have the same amount of lagged information available. Thus, differences in forecast performance between competing models can solely accrue from two sources: how flexible – or parsimonious – the models are in terms of parametrization, and whether they can incorporate higher frequency data updates (new releases or revisions).

On a last note, it is not the purpose of this paper to find the best MIDAS VAR specification. Rather, we want to know whether MIDAS VARs *robustly* improves upon standard QFVARs. For this we compare MIDAS VAR and QFVAR for alternative reasonable specifications, i.e. for alternative variables included and alternative lag structures. None of these specifications dominates the others in terms of forecasting accuracy. Hence, the MIDAS VAR must outperform the QFVAR for all specifications in order for the improvement to be robust.

4 Empirical results

The following sections present the results of the forecast evaluation outlined in Section 3. We compare MIDAS VARs and QF-VARs for four alternative specification setups: few variables and few lags, few variables and many lags, higher number of variables and few lags, and higher number of variables and many lags. Competing MIDAS VARs and QFVARs always have the same amount of lagged information available. Thus, differences in forecast performance between competing models can solely accrue from two sources: how flexible – or parsimonious – the models are in terms of parametrization, and whether they can incorporate higher frequency data updates (new releases or revisions).

3-variable MIDAS VAR with 3-month memory

In a first step it might be interesting to see whether MIDAS VAR models improve forecasts upon a standard VAR when model specifications are kept minimal. Following common practice in macroeconomics our minimum VAR specification includes three variables only: GDP growth, consumer price inflation and a (3-month) short-term interest rate (treasury bill rate).

Figure 2 presents results for an Almon MIDAS VAR of order 1 and a U-MIDAS VAR both with one quarterly GDP growth lag, three monthly inflation lags and three monthly interest rate lags.¹⁴ The benchmark model is a quarterly frequency VAR (QFVAR) with one

¹⁴When three lags are employed an Almon MIDAS VAR of order 2 delivers identical results as a U-MIDAS VAR.

quarterly lag of GDP growth, inflation and interest rate, respectively. As regards forecasting GDP growth, the two MIDAS VAR specifications reduce the RMSFE by more than 30 percent compared to the QFVAR for forecast horizons of one to six months. The RMSFE improvement gradually declines as the forecast horizon increases, but is still substantial for longer horizons: One year before publication of GDP the MIDAS VAR forecasts are still more than 20 percent better than the QFVAR forecasts. For forecast horizons of 22 months or longer the MIDAS VARs yield no improvement over the QFVAR anymore. The zig-zag pattern of the relative RMSFE is due to the mixed frequency structure of our setup: the MIDAS VAR gets new information every month whereas the data of the QFVAR are only updated every third month when new quarterly information is available.

For inflation, the Almon MIDAS VAR and the U-MIDAS VAR yield very high RMSFE reductions for short forecast horizons (more than 65 percent for months 1 to 3). On the other hand, RMSFE improvements quickly vanish as the forecast horizon grows. And for forecast horizons of one year onwards the relative change in the RMSFE decreases again. The reason for this pattern is that the QFVAR forecast errors initially *decrease* with an increase in the forecast horizon, but then increase again from a forecast horizon of one year onwards. In contrast, the forecast errors of the MIDAS VAR slowly and steadily increase as the forecast horizon becomes bigger (which is what one would expect). When it comes to forecasting the interest rate, the Almon MIDAS VAR and the U-MIDAS VAR improve forecasts substantially throughout all considered forecast horizons as compared to the QFVAR. The RMSFE reductions reach more than 75 percent for horizons of one to three months. Forecast improvements gradually decline with an increase in the forecast horizon, but even two years before GDP publication the MIDAS VARs reduce the RMSFE by 5 percent or more compared to the QFVAR. All previous findings remain robust when we iterate the forecast evaluation for the first or third GDP estimate instead of the second estimate (see Section 3).

– Figure 2 here –

3-variable MIDAS VAR with 6- or 12-month memory

VAR models of higher order than the ones presented in the previous section arguably achieve better forecasting performance for longer forecast horizons. Thus, it is important to know whether MIDAS VARs still improve forecasts upon a standard VAR for specifications with longer time series memory. Figure 3 shows the results for an Almon MIDAS VAR of order 1 and a U-MIDAS VAR both with two quarterly GDP growth lags, six monthly inflation lags and six monthly interest rate lags. The benchmark model here is a QFVAR with two quarterly lags of GDP growth, inflation and interest rate, respectively, where two lags are chosen in order to provide each competing model with the same amount of lagged information (see Section 3).

To allow for even longer memory, Figure 4 depicts results for an Almon MIDAS VAR of order 1 and a U-MIDAS VAR both with four quarterly GDP growth lags, twelve monthly inflation lags and twelve monthly interest rate lags. In accordance with the above reasoning about the appropriate information set, the benchmark model now is a QFVAR with four quarterly lags of GDP growth, inflation and interest rate, respectively. Generally, the RMSFE patterns in Figures 3 and 4 strongly resemble the patterns in Figure 2: the two MIDAS VAR specifications yield substantial improvements in forecast accuracy upon

the QFVAR for GDP growth, inflation as well as the interest rate. That said, for the 12-month memory specification the U-MIDAS VAR performs clearly worse than the Almon VAR when it comes to forecasting GDP growth. When the forecast horizon exceeds 13 months the U-MIDAS VAR forecasts are even less accurate than the QFVAR forecasts. This finding is not surprising: U-MIDAS specifications easily become overparametrized when the lag order grows. As a consequence, U-MIDAS is less suitable for – and is actually not meant for – forecasting with longer lags.

We iterated the analysis with an Almon MIDAS VAR of order 2 (not shown in the figures). The RMSFE patterns very closely resemble the patterns of the order 1 Almon MIDAS VAR.

– Figures 3 and 4 here –

6-variable MIDAS VAR with 3-month memory

The previous minimum scale VAR specifications including only GDP growth, inflation and an short-term interest rate provide us with first evidence on the potential of MIDAS VAR models for forecasting. However, in order to improve predictive accuracy forecasters usually employ VAR models with more than just the aforementioned variables. So, do MIDAS VARs still improve forecasts upon standard VARs when more variables are included? To answer this question we include three additional variables which are commonly considered to be helpful for now- or forecasting GDP growth: industrial production in month-on-month growth rates, housing starts in year-on-year growth rates and the Standard & Poor’s 500 stock market index in month-on-month growth rates.¹⁵

Figure 5 depicts results for an Almon MIDAS VAR of order 1 and a U-MIDAS VAR both with one quarterly GDP growth lag and three monthly lags of the five monthly variables. The benchmark model here is a QFVAR with one quarterly lag of each of the aforementioned six variables. Thus, as before each competing model has the same amount of lagged information such that differences in forecast performance cannot accrue from differences in the (lagged) information set. The RMSFE patterns in Figure 5 strongly resemble the patterns presented in the previous subsections: the two MIDAS VAR specifications substantially improve predictive accuracy upon the QFVAR for GDP growth, inflation and the interest rate.

– Figure 5 here –

6-variable MIDAS VAR with 6- or 12-month memory

The previous section has shown that small scale MIDAS VARs substantially improve predictive accuracy upon a standard small scale VAR when only very short time series

¹⁵Indeed, we find that a small scale MIDAS VAR (with industrial production, housing starts, S&P 500, GDP growth, inflation and the interest rate) improves forecast accuracy upon a minimum MIDAS VAR (with only the latter three variables). Equally, the corresponding small scale benchmark QFVAR performs better than the corresponding minimum scale benchmark QFVAR. A rigorous variable selection procedure might deliver small scale specifications with still greater forecasting ability. That said, it is not the goal of this paper to find the best model specification. Rather, we compare MIDAS VARs and standard VARs for several sensible, yet alternative model specifications in order to see whether MIDAS VARs robustly outperform standard VARs.

memory is taken into account (one quarterly lag or three monthly lags). Is this gain is robust to allowing for longer time series memory (which increases the risk of over-parametrization)?

Figure 6 (or Figure 7) shows the results for an Almon MIDAS VAR of order 1 with two (or four) quarterly GDP growth lags and six (or twelve) monthly lags of inflation, the short-term interest rate, industrial production, housing starts and the S&P 500 stock market index. The benchmark model is a QFVAR with two (or four) quarterly lags for each of the aforementioned six variables such that each competing model has again the same amount of lagged information (see Section 3). The Almon MIDAS VAR still largely outperforms the benchmark QFVAR in terms of forecast accuracy at least for shorter horizons. A comparison between the RMSFE improvements resulting from the longer memory Almon MIDAS VARs in Figures 6 and 7 with the RMSFE improvements from the 3-month memory Almon MIDAS VAR in Figure 5 yields the following differences: Regarding GDP growth, the forecast improvement vanishes earlier (at a forecast horizon of 16 instead of 19 months) and turns into a deterioration for higher forecast horizons. In contrast, for inflation and the interest rate, the longer horizon forecast improvement is substantially larger.

The predictive accuracy of the U-MIDAS VAR substantially deteriorates for longer memory small scale specifications. Figure 6 shows that for a specification with two quarterly and six monthly lags the U-MIDAS VAR is generally outperformed by the Almon MIDAS VAR when it comes to forecasting GDP growth, inflation or the short-term interest rate (an exception being the very short-term forecasts for inflation and the interest rate). For the specification with four quarterly lags and twelve monthly lags the forecast performance of the U-MIDAS VAR is so poor that we do not show it in Figure 7. U-MIDAS does not help for – and is actually not meant for – forecasting with longer memory because of over-parametrization. In contrast, Almon MIDAS keeps models parsimonious despite long lags.

Again, iterating the analysis with an Almon MIDAS VAR of order 2 yields RMSFE patterns that closely resemble the patterns of the order 1 Almon MIDAS VAR (not shown in the figures).

– Figures 6 and 7 here –

12-variable MIDAS VAR with 3-month memory

Using a MIDAS setup does not only give an advantage when using a model with longer memory but also when using a bigger amount of variables, both when short or long memories are considered. To show this we increase our variable set by six additional variables, namely the (10-year) long-term interest rate (treasury bond yield), hours worked by production and supervisory workers in 3-month growth rates, the ISM index for manufacturing (= ISM total) in levels, the ISM index for supplier delivery times both in levels and year-on-year growth rates and the month-on-month growth in housing starts. In order to improve precision of the parameter estimates we increase the estimation sample by 5 years. The forecast comparison now starts in January 1990 only instead of January 1985.

Figure 8 depicts results for an Almon MIDAS VAR of order 1 and a U-MIDAS VAR both with one quarterly GDP growth lag and three monthly lags of the eleven monthly

variables. The benchmark model here is a QFVAR with one quarterly lag of each of the aforementioned twelve variables giving again each competing model the same amount of lagged information. The two MIDAS VAR specifications substantially improve predictive accuracy upon the QFVAR for GDP growth, inflation and the interest rate over all forecast horizon.

– Figure 8 here –

12-variable MIDAS VAR with 6- or 12-month memory

As shown in the previous section the 12-variable U-MIDAS and Almon MIDAS VARs deliver a forecast improvement over a 12-variable QFVAR when considering only a short memory of 1 quarter (3 months). Does this result hold when taking into account longer time series memories? Figure 9 (10) shows the relative RMSFE changes of the Almon MIDAS VAR of order 1 with two (four) quarterly lags of GDP growth and six (twelve) monthly lags of the aforementioned eleven variables as compared to a QFVAR with two (four) quarterly lags for each of the twelve variables. The improvement in forecast accuracy of the Almon MIDAS VAR relative to the QFVAR benchmark is even higher than for the short memory specification shown in Figure 8. Apparently, the Almon MIDAS VAR is rather robust against overparametrization, while the QFVAR tends to get overparametrized with both a large number of variables and a large number of lags. Not surprisingly, the U-MIDAS VAR suffers even more from overparametrization. We do not show the U-MIDAS VAR in Figures 9 and 10 because of its very poor relative forecast performance.

– Figures 9 and 10 here –

5 Conclusion

VAR models are widely used for macroeconomic forecasting and policy analysis (Stock and Watson, 2001 and Karlsson, 2013, e.g.). An initial challenge for VARs is that the variables they include must have the same frequency, whereas macroeconomic time series usually come at different frequencies. For instance, GDP is published every quarter, inflation is a monthly variable and interest rates come at a daily or even higher frequency. The traditional solution is to time-aggregate all variables to a common frequency. A VAR including GDP, inflation and an interest rates series will then be a quarterly frequency VAR. However, the time-aggregation comes with inconveniences: First, higher frequency data releases can only be taken into account with a delay. For instance, a quarterly frequency VAR can only consider inter-quarterly inflation or other indicator releases after the end of a quarter. Second, the time aggregation implies a peculiar constraint on the parameters attached to higher frequency variables which is potentially suboptimal. Whether or to what degree these inconveniences affect the forecasting ability of VAR models is ultimately an empirical question. We find that time-aggregation substantially impairs the forecasting power of VAR models even for higher forecast horizons.

Against this backdrop, a number of authors have developed VAR models which can deal with variables of differing frequencies. Zdrozny (1988, 1990), Mariano and Murasawa (2010), Kuzin *et al.* (2011), Chiu *et al.* (2012), Bai *et al.* (2013) and Schorfheide and Song

(2013) propose mixed frequency VAR (MFVAR) models using a state space approach. Ghysels (2012) develops a MFVAR using a stacked vector approach (cf. applications in Francis *et al.*, 2011 and Forni *et al.*, 2014). Both approaches have their merits and promises (Kuzin *et al.*, 2011, Bai *et al.*, 2013 and Forni and Marcellino, 2014). We make the following contributions to this emerging research on MFVARs. First, we propose a general and yet tractable stacked vector MFVAR framework. Previous expositions are limited to VARs with only two different frequencies for reasons of tractability. Second, we augment the stacked vector MFVAR with a MIDAS type non-linear Almon lag polynomial scheme (Almon, 1965) which is designed to prevent overparametrization even with long lag lengths while, at the same time, keeping the VAR model flexible.¹⁶ In turn, we show how to transform the resulting stacked vector non-linear MFVAR into a linear equation system which can be easily estimated via OLS. Due to the linear transformation the stacked vector MIDAS type MFVAR becomes feasible for estimation with multiple variables. Previous stacked vector MIDAS type MFVARs were limited to only few variables. The reason is that, in absence of a linear transformation, the non-linear MFVARs had to be estimated directly which is cumbersome, if not infeasible, in a VAR context involving multiple variables (provided no auxiliary restrictions).

Using quarterly and higher frequency US real-time data we test the forecast performance of our MIDAS type MFVAR against a quarterly frequency VAR for various different specifications (three, six or twelve variables, three, six or twelve months of lagged information). The MIDAS type MFVAR yields root mean squared forecast error reductions of 30 to 50 percent for forecast horizons up to six months and of about 20 percent for a forecast horizon of one year. For inflation and interest rates forecast error reductions are even bigger. According to our results, augmentation of a stacked vector MFVAR with an Almon lag polynomial scheme has a distinct advantage for specifications with longer lag structures. While a MFVAR with fully unrestricted parameters still yields considerable forecast improvements over a standard QFVAR when using few lags, these improvements vanish almost completely for longer lags. An Almon augmented MFVAR instead still yields high improvements over a QFVAR for longer lags.

The VAR specifications in our empirical application are intentionally kept simple. For instance, each variable's own lags are kept as flexible concerning polynomial order as its other variables lags. A model specification procedure with respect to polynomial flexibility and lag length will boost the predictive power of the MIDAS type MFVAR. Further, the linearly transformed stacked vector MIDAS type MFVAR could be estimated using Bayesian methods instead of OLS. This would allow to employ Bayesian VAR specifications with prior types that are known to deliver a better forecast performance than non-Bayesian or flat prior Bayesian VAR specifications (see Giannone *et al.*, 2014, e.g.). We leave these steps for future research.

¹⁶See a discussion in Judge *et al.* (1985).

References

- ALMON, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica*, pp. 178–196. [back to page [2](#), [18](#), [29](#)]
- ANDREOU, E., GHYSELS, E. and KOURTELLOS, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, **158**, 246–261. [back to page [2](#), [8](#)]
- , — and — (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, **31(2)**, 240–251. [back to page [2](#)]
- ARRESTO, M. T., ENGEMANN, K. M. and OWYANG, M. T. (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review*, **92(6)**, 521–36. [back to page [2](#)]
- ARUOBA, S. B., DIEBOLD, F. X. and SCOTTI, C. (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, **27** (4). [back to page [3](#)]
- BAI, J., GHYSELS, E. and WRIGHT, J. H. (2013). State space models and MIDAS regressions. *Econometric Reviews*, **32(7)**, 779–813. [back to page [3](#), [17](#), [18](#)]
- CARRIERO, A., CLARK, T. E. and MARCELLINO, M. (2012). Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility. *Federal Reserve Bank of Cleveland Working Papers*, **12-27**, Federal Reserve Bank of Cleveland. [back to page [11](#), [12](#)]
- CHIU, C. W. J., ERAKER, B., FOERSTER, A. T., KIM, T. B. and SEOANE, H. D. (2012). Estimating VARs sampled at mixed or irregular spaced frequencies: A Bayesian approach. *Federal Reserve Bank of Kansas City Research Working Papers*, **RWP 11-11**, Federal Reserve Bank of Kansas City. [back to page [3](#), [17](#)]
- CLEMENTS, M. P. and GALVÃO, A. B. (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics*, **26(4)**, 546–554. [back to page [2](#), [4](#)]
- and GALVÃO, A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*, **24**, 1187–1206. [back to page [2](#), [4](#)]
- COX, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics*, **8**, 93–115. [back to page [3](#)]
- CROUSHORE, D. (2006). *Forecasting with Real-Time Macroeconomic Data*, Elsevier, *Handbook of Economic Forecasting*, vol. 1, chap. 17, pp. 961–982. [back to page [12](#)]
- DRECHSEL, K. and SCHEUFELE, R. (2012). Bottom-up or direct? Forecasting German GDP in a data-rich environment. *Swiss National Bank Working Papers*, **2012-16**, swiss National Bank. [back to page [2](#), [29](#)]

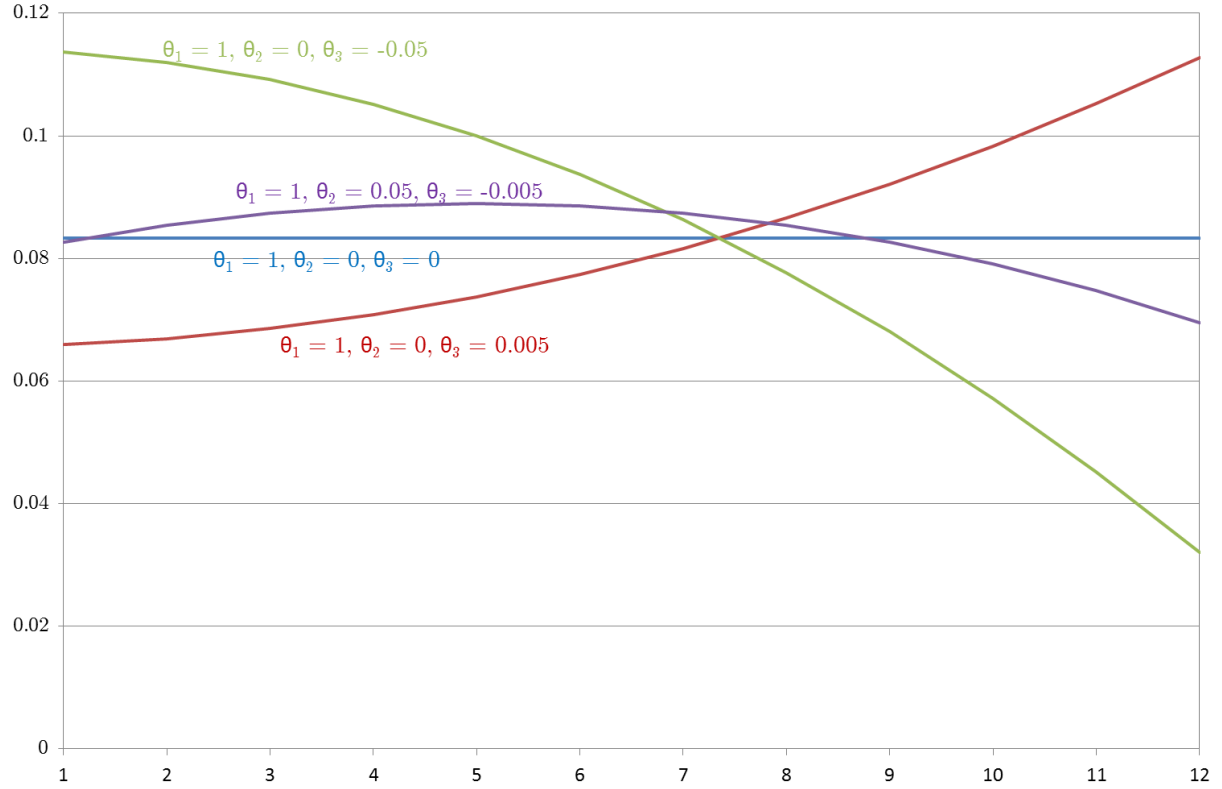
- FAUST, J. and WRIGHT, J. H. (2009). Comparing greenbook and reduced form forecasts using a large realtime dataset. *Journal of Business & Economic Statistics*, **27** (4), 468–479. [back to page [12](#)]
- FERRARA, L., MARSILLI, C. and ORTEGA, J.-P. (2014). Forecasting growth during the Great Recession: Is financial volatility the missing ingredient? *Economic Modelling*, **36**, 44–50. [back to page [2](#)]
- FORONI, C., GUÉRIN, P. and MARCELLINO, M. (2014). Markov-switching mixed-frequency VAR models. *International Journal of Forecasting*, **forthcoming**. [back to page [2](#), [4](#), [18](#)]
- and MARCELLINO, M. (2014). A comparison of mixed frequency approaches for nowcasting euro area macroeconomic aggregates. *International Journal of Forecasting*, **30**, 554–568. [back to page [3](#), [18](#)]
- , — and SCHUMACHER, C. (2012). U-MIDAS: MIDAS regressions with unrestricted lag polynomials. *CEPR Discussion Papers*, **8828**, C.E.P.R. Discussion Papers. [back to page [8](#), [29](#)]
- FRANCIS, N. R., GHYSELS, E. and OWYANG, M. T. (2011). The low-frequency impact of daily monetary policy shocks. *Federal Reserve Bank of St. Louis Working Papers*, **2011-009C**, Federal Reserve Bank of St. Louis. [back to page [2](#), [4](#), [18](#)]
- GHYSELS, E. (2012). Macroeconomics and the reality of mixed frequency data. *mimeo, University of North Carolina*. [back to page [2](#), [3](#), [4](#), [18](#), [28](#), [29](#)]
- , SANTA-CLARA, P. and VALKANOV, R. (2004). The MIDAS touch: Mixed data sampling regression models. *Working Paper*, University of California, Los Angeles. [back to page [2](#)]
- , SINKO, A. and VALKANOV, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, **26**(1), 53–90. [back to page [2](#), [8](#), [29](#)]
- GIANNONE, D., LENZA, M. and PRIMICERI, G. E. (2014). Prior selection for vector autoregressions. *Review of Economics and Statistics*, **forthcoming**. [back to page [18](#)]
- , REICHLIN, L. and SMALL, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55**(4), 665–676. [back to page [3](#)]
- HAMILTON, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press. [back to page [5](#)]
- HAYASHI, F. (2000). *Econometrics*. Princeton, NJ: Princeton University Press. [back to page [5](#)]
- JORGENSON, D. W. (1966). Rational distributed lag functions. *Econometrica*, pp. 135–149. [back to page [2](#)]
- JUDGE, G. G., GRIFFITHS, W. E., HILL, R. C., LÜTKEPOHL, H. and LEE, T.-C. (1985). *The Theory and Practice of Econometrics*. New York: John Wiley and Sons. [back to page [2](#), [18](#)]

- KARLSSON, S. (2013). *Forecasting with Bayesian Vector Autoregression*, Elsevier, *Handbook of Economic Forecasting*, vol. 2, chap. 15, p. 791897. [back to page [2](#), [12](#), [17](#)]
- KHATRI, C. G. and RAO, C. R. (1968). Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhya*, **30**, 167–180. [back to page [10](#)]
- KUZIN, V., MARCELLINO, M. and SCHUMACHER, C. (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting*, **27**, 529–542. [back to page [3](#), [17](#), [18](#)]
- LÜTKEPOHL, H. (1981). A model for non-negative and non-positive distributed lag functions. *Journal of Econometrics*, **16** (2), 211–219. [back to page [2](#)]
- 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. [back to page [3](#)]
- and — (2010). A coincident index, common factors, and monthly real GDP. *Oxford Bulletin of Economics and Statistics*, **72**(1), 27–46. [back to page [3](#), [17](#)]
- MIKOSCH, H. and ZHANG, Y. (2014). Forecasting Chinese GDP growth with mixed frequency data: Which indicators to look at? *KOF Working Papers*, **14-359**, KOF Swiss Economic Institute, ETH Zurich. [back to page [2](#), [29](#)]
- MITCHELL, D. W. and SPEAKER, P. J. (1986). A simple, flexible distributed lag technique: The polynomial inverse lag. *Journal of Econometrics*, **31** (3), 329–340. [back to page [2](#)]
- ROMER, D. H. and ROMER, C. D. (2000). Federal reserve information and the behavior of interest rates. *American Economic Review*, **90** (3), 429–457. [back to page [12](#)]
- SCHMIDT, P. (1974). A modification of the Almon distributed lag. *Journal of the American Statistical Association*, **69** (347), 679–681. [back to page [2](#)]
- SCHORFHEIDE, F. and SONG, D. (2013). Real-time forecasting with a mixed-frequency VAR. *NBER Working Papers*, **19712**, National Bureau of Economic Research. [back to page [3](#), [11](#), [17](#)]
- SIMS, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity*, **33** (2), 1–62. [back to page [12](#)]
- SOLOW, R. M. (1960). On a family of lag distributions. *Econometrica*, pp. 393–406. [back to page [2](#)]
- STOCK, J. H. and WATSON, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives*, **15**(4), 101–115. [back to page [2](#), [17](#)]
- TSURUMI, H. (1971). A note on Gamma distributed lags. *International Economic Review*, pp. 317–324. [back to page [2](#)]
- ZADROZNY, P. A. (1988). Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies. *Econometric Theory*, **4**(1), 108–124. [back to page [3](#), [17](#)]

— (1990). Forecasting U.S. GNP at monthly intervals with an estimated bivariate time series model. *Economic Review*, **November**, 2–15. [back to page [3](#), [17](#)]

6 Figures and Tables

Figure 1: Almon lag polynomial



y -axis: Almon lag polynomial weight of order $Q = 2$, $\alpha_{i,\tau_i,j,k}(\theta_{i,\tau_i,j,0}, \theta_{i,\tau_i,j,1}, \theta_{i,\tau_i,j,2}, k)$, attached to observation $y_{j,t-1+\tau_i/\mathcal{T}_i-k/\mathcal{T}_j}$ conditional on different values of $\theta_{i,\tau_i,j,0}$, $\theta_{i,\tau_i,j,1}$ and $\theta_{i,\tau_i,j,2}$. See Equation (6). x -axis: $k = 1, \dots, 12$ (= first to twelfth lag). Sum of weights normalized to one for ease of exposition.

Table 1: Data overview

VARIABLE NAME	FREQUENCY	REAL TIME	SOURCE	NOTES
<i>Real seasonal adjusted GDP</i>	quarterly	yes	RTDSM	
<i>Industrial production index: manufacturing</i>	monthly	yes	RTDSM	From 1997M12 to 1998M10 historical data from 1969 to 1970 were not available in the data set. As there do not seem to be any revisions in between, we make the assumption that there were no revisions.
<i>Consumer price index</i>	monthly	yes	ALFRED	For some months two vintages were available. In those cases we used the later publications. This was the case for 2000M9, 2005M2, 2006M2, 2007M2, 2008M2, 2009M2, 2010M2, 2011M2 and 2012M2.
<i>Housing starts</i>	monthly	yes	RTDSM	There were no observations available for the first publication of 1995M11 and 1995M12. We make the assumption that in those cases no revision took place and use the same values as in the second publication.
<i>ISM index for manufacturing (= ISM total)</i>	monthly	no	Institute for Supply Management	
<i>ISM index for supplier delivery times</i>	monthly	no	Institute for Supply Management	
<i>Average weekly hours of production and supervisory workers</i>	monthly	no	US Bureau of Labor Statistics	
<i>S&P 500 stock market index</i>	monthly	no	Thomson Reuters Datastream	
<i>3-month treasury bill yield</i>	monthly	no	Thomson Reuters	
<i>10-year treasury bond yield</i>	monthly	no	Thomson Reuters	

Figure 2: MIDAS VAR with 3 variables and 3 months of lagged information

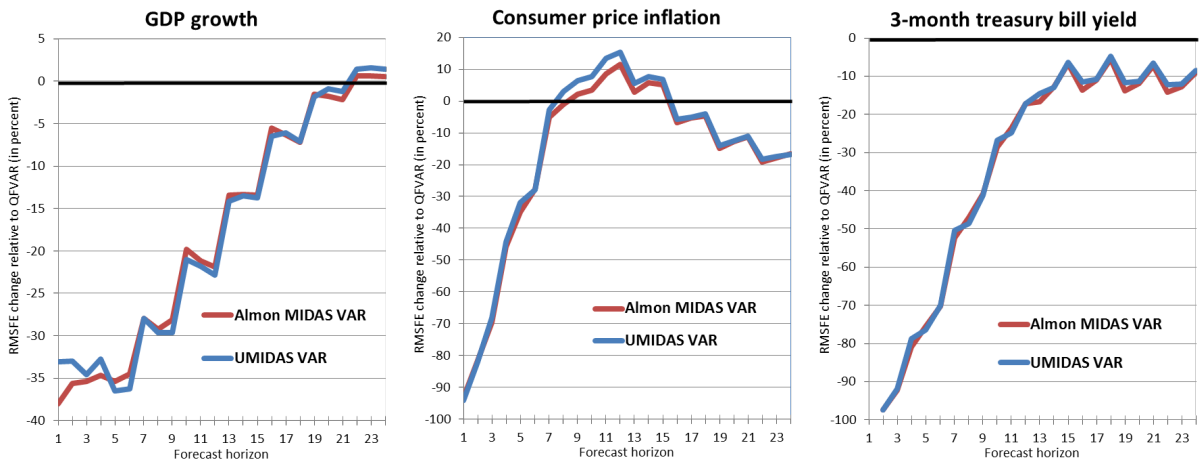


Figure 3: MIDAS VAR with 3 variables and 6 months of lagged information

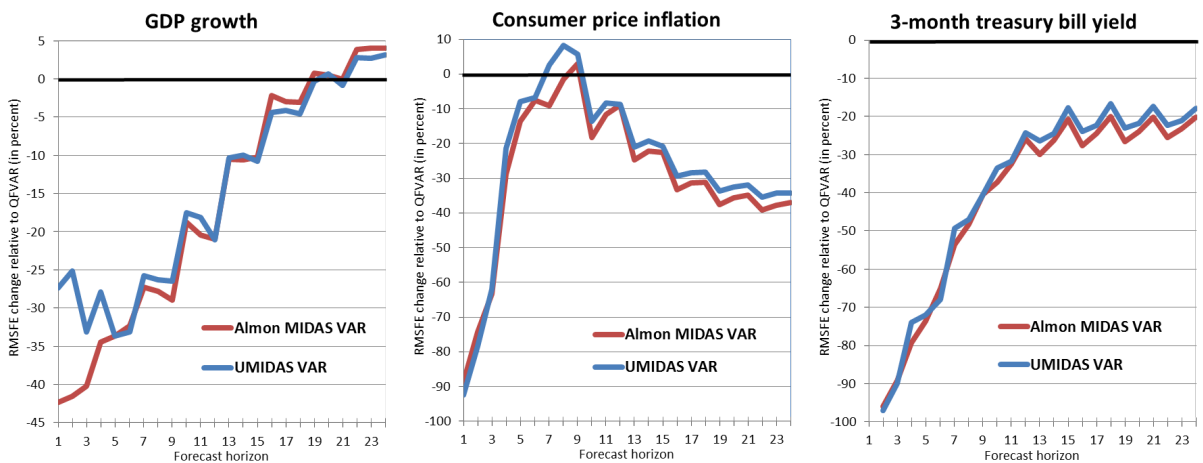
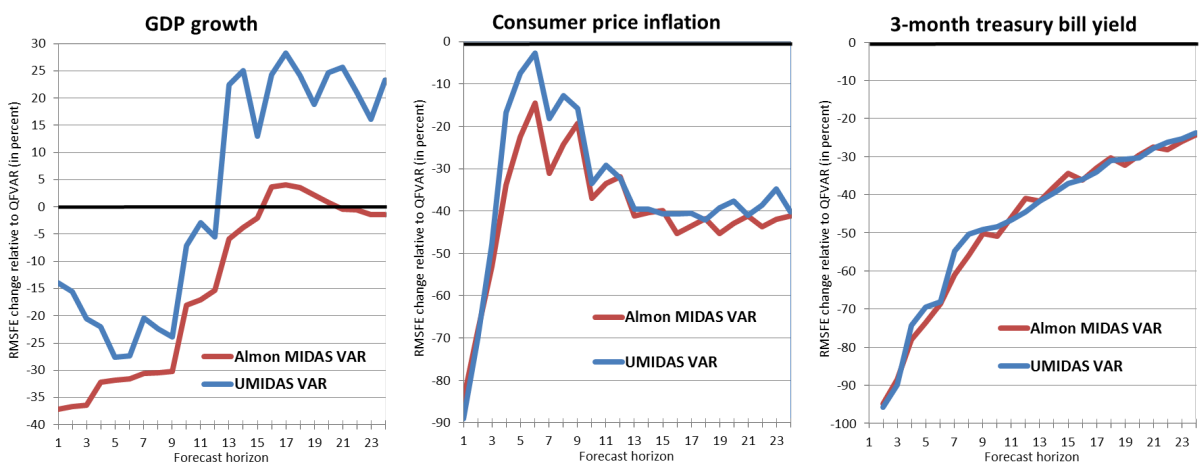


Figure 4: MIDAS VAR with 3 variables and 12 months of lagged information



x -axis: Forecast horizon, i.e. months prior to GDP growth/inflation/interest rate release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between RMSFE of MIDAS VAR and RMSFE of QFVAR in percent of RMSFE of QFVAR.

Figure 5: MIDAS VAR with 6 variables and 3 months of lagged information

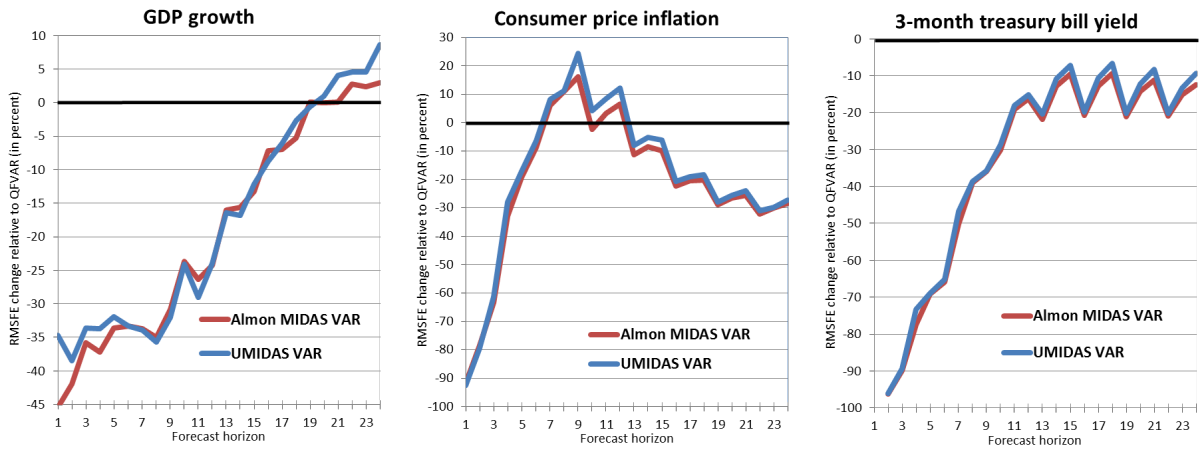


Figure 6: MIDAS VAR with 6 variables and 6 months of lagged information

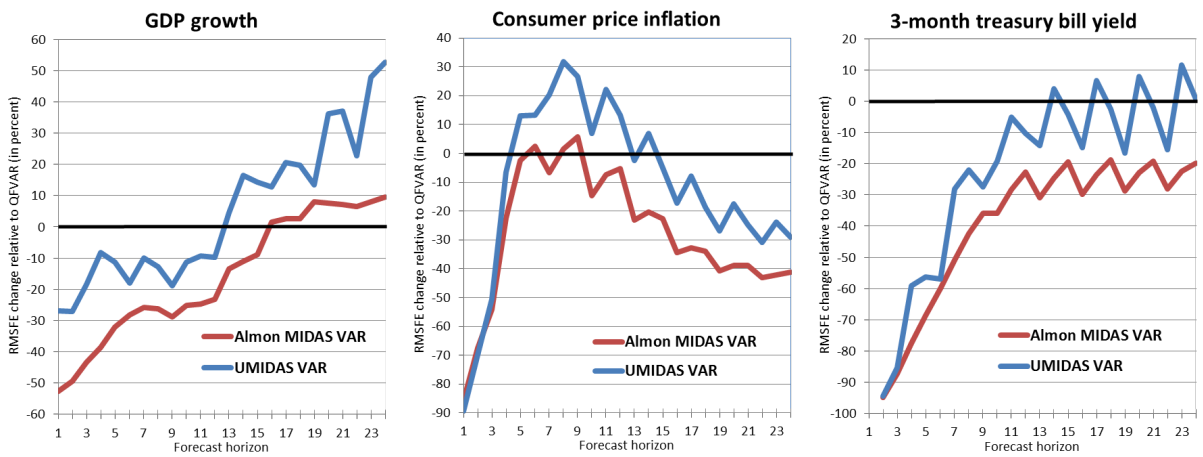
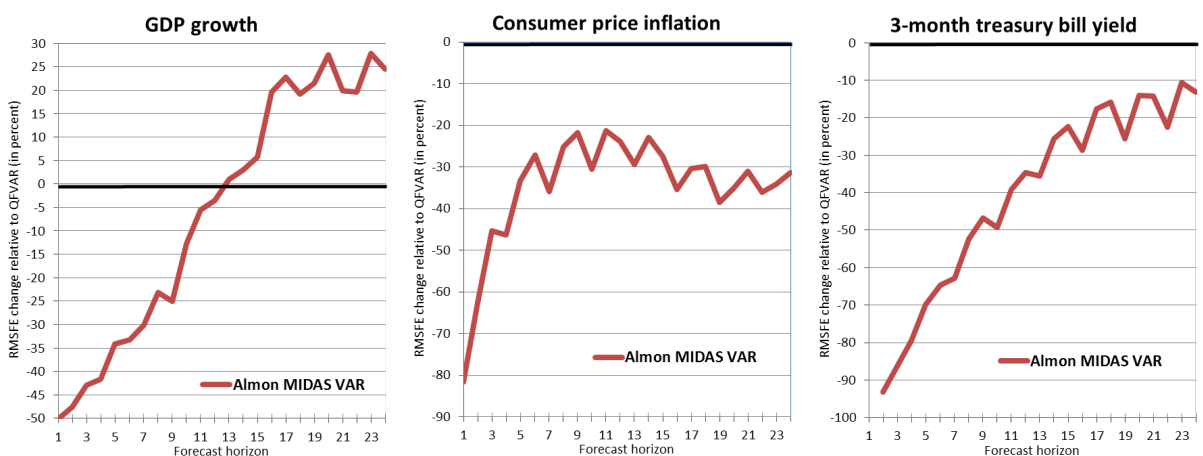


Figure 7: MIDAS VAR with 6 variables and 12 months of lagged information



x -axis: Forecast horizon, i.e. months prior to GDP growth/inflation/interest rate release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between RMSFE of MIDAS VAR and RMSFE of QFVAR in percent of RMSFE of QFVAR.

Figure 8: MIDAS VAR with 12 variables and 3 months of lagged information

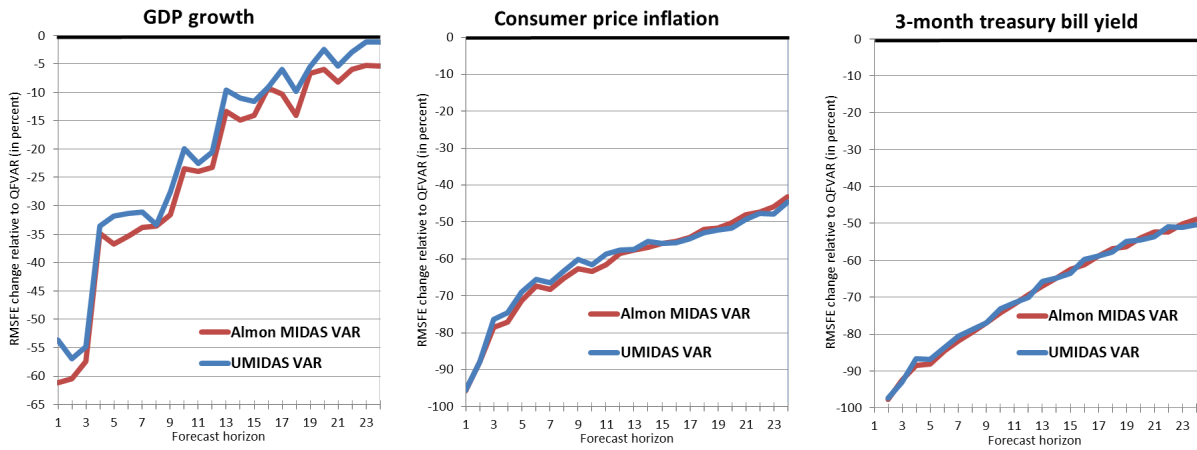


Figure 9: MIDAS VAR with 12 variables and 6 months of lagged information

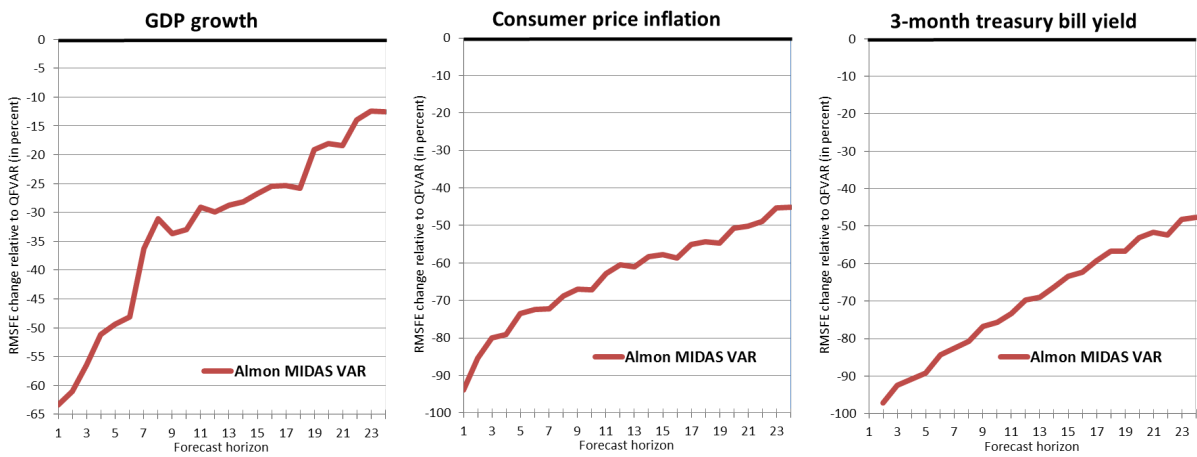
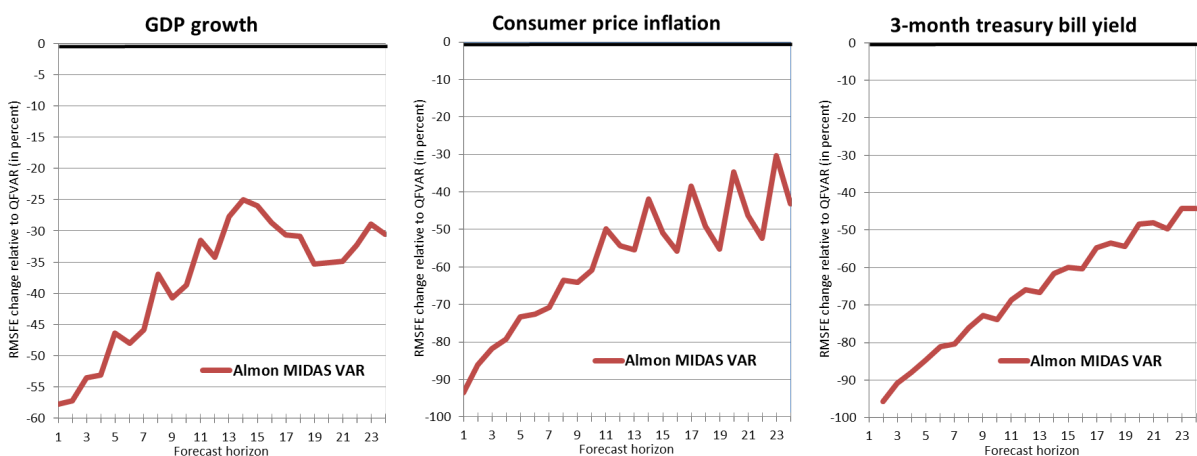


Figure 10: MIDAS VAR with 12 variables and 12 months of lagged information



x -axis: Forecast horizon, i.e. months prior to GDP growth/inflation/interest rate release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between RMSFE of MIDAS VAR and RMSFE of QFVAR in percent of RMSFE of QFVAR.

Appendix 1: Comparison with Ghysels (2012)

Ghysels (2012) is the first to develop a stacked vector MIDAS type mixed frequency VAR. Our model is similar to Ghysels' model in some respects and different in other respects. First, Ghysels (2012) mainly concentrates on impulse response analysis, whereas we focus on forecasting.

Second, while Ghysels' framework in principle allows for more multiple frequencies it becomes conceptually untractable as the number of frequencies grows. Ghysels limits his exposition to two frequencies (Ghysels, 2012, p. 3). In contrast, our MFVAR framework is designed for multiple variables and multiple frequencies while still being relatively compact and straightforward.

Third, Ghysels' and our framework differ in the way they handle the relationships between variable observations that pertain to the same lowest frequency period t with $t = 1, \dots, T$ and, thus, are stacked in the same vector (for instance, the relationship between the GDP observation of a quarter and the first monthly inflation observation of the same quarter, or the relationship between the second and the first monthly inflation observation of the same quarter). These within- t relationships are crucial for now- and forecasting since they allow to incorporate data releases or revisions which occur *within* period t (*within* a quarter, for instance). In order to capture the aforementioned relationships Ghysels (2012, p. 12ff) proposes a scheme that augments a reduced form MFVAR with a matrix structure derived from the Choleski lower triangular decomposition. In contrast, we propose to capture the within- t relationships via a recursive form type MFVAR scheme. While the two schemes may ultimately be transformed into each other they fundamentally differ in terms of exposition.

Fourth, while both frameworks adopt a stacking approach, the stacking is done in different ways resulting in quite different model setups and solutions. Ghysels starts with stacking all variable observations that pertain to the same lowest frequency period t into one vector. Then he stacks all resulting t -vectors with $t = 1, \dots, T$.¹⁷ In contrast, we start with stacking all lagged observations that belong to the same variable into one vector possibly extending over several periods t (nucleus vector). Thereupon we stack all resulting i -vectors with $i = 1, \dots, I$. The reason for pursuing this way is as follows: A single MIDAS type polynomial possibly spans a larger number of lagged observations of a variable and, hence, possibly extends over several lowest frequency periods t . For instance, a single MIDAS type polynomial with two parameters might span twelve lags of a monthly variable; provided that the lowest frequency period in the data sample is a quarter, the polynomial thus extends over four periods t .¹⁸ By making the i -vectors the nuclei of our stacked vector system and imposing a (separate) MIDAS type polynomial on each nucleus, we can build a system of MIDAS blocks. Not only can this system of MIDAS blocks be written in very compact notation, but it gives also a good basis for further transformation.

Fifth, the MIDAS literature knows various weighting functions which are quite different

¹⁷Actually, Ghysels does not show this second step, but presents the VAR directly in sum or lag operator notation (Ghysels, 2012, Equation (2.1) or (2.5)). Still, the second step would result from simply rewriting the sum notation into matrix notation.

¹⁸In contrast, one would not want a single polynomial to span several variables, at least unless there does exist a specific justification of the resulting parameter restriction.

from each other in several respects. Actually, the only uniting feature is that they can all be used in a MIDAS context. Ghysels stresses that various different MIDAS type specifications might be employed in the MFVAR context sketching some of them in turn (Ghysels, 2012, ch. 2.3 with reference to Appendix A). He concentrates on the beta probability density function (pdf) which is highly non-linear and conceptually very convenient for reducing the parameter space while keeping models flexible. A problem is that, due the non-linearity of the function, estimation becomes cumbersome (if not infeasible) in a context involving more than two variables. We propose to employ the Almon lag polynomial (Almon, 1965). The Almon polynomial is also highly non-linear but can be rewritten in linear form. We devote a substantial part of our theory section to showing how a stacked vector MIDAS type MFVAR involving Almon lag polynomials can be recast into a linear equation system.¹⁹ Comparisons between alternative MIDAS type polynomials in a single equation context yield that the Almon lag polynomial is a preferable choice (Mikosch and Zhang, 2014). This is because it can be rewritten in linear form and hence linear estimation becomes feasible (see also Drechsel and Scheufele, 2012).

Appendix 2: MIDAS VAR example for one quarterly and one monthly variable

In this section we give a simple illustrative example of the MIDAS VAR framework presented in Section 2. Let there be one monthly variable, $y_{1,t-1+\tau_1/3}$ and one quarterly variable, $y_{2,t-1+\tau_2/1}$, with $t = 1, \dots, T$, $\tau_1 = 1, 2, 3$ and $\tau_2 = 1$. Further let $P = 1$. In this case,

$$\mathbf{y}_{t-p} = \begin{bmatrix} y_{1,t-p} \\ y_{1,t-p-\frac{1}{3}} \\ y_{1,t-p-\frac{2}{3}} \\ y_{2,t-p} \end{bmatrix}, \quad 20$$

and

$$\mathbf{x}_t = \begin{bmatrix} y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{1,t-1} \\ y_{1,t-1\frac{1}{3}} \\ y_{1,t-1\frac{2}{3}} \\ y_{1,t-2} \\ y_{2,t-1} \\ y_{2,t-2} \end{bmatrix}.$$

¹⁹Further, both Ghysels' and our paper employs a MFVAR specification that refrains from any reduction or restriction of the parameter space. This is the unrestricted MIDAS (U-MIDAS) approach proposed by Forni *et al.* (2012). U-MIDAS works well for short lag structures where the parameter space is relatively small a priori. Ghysels also uses step functions following Ghysels *et al.* (2007).

The mixed-frequency VAR process from Equation (4) then writes

$$\begin{bmatrix} y_{1,t} \\ y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_{1,3,1,1} & \alpha_{1,3,1,2} & \alpha_{1,3,1,3} & \alpha_{1,3,1,4} & 0 & 0 & \alpha_{1,3,2,1} & \alpha_{1,3,2,2} \\ 0 & \alpha_{1,2,1,1} & \alpha_{1,2,1,2} & \alpha_{1,2,1,3} & \alpha_{1,2,1,4} & 0 & \alpha_{1,2,2,1} & \alpha_{1,2,2,2} \\ 0 & 0 & \alpha_{1,1,1,1} & \alpha_{1,1,1,2} & \alpha_{1,1,1,3} & \alpha_{1,1,1,4} & \alpha_{1,1,2,1} & \alpha_{1,1,2,2} \\ \alpha_{2,1,1,1} & \alpha_{2,1,1,2} & \alpha_{2,1,1,3} & \alpha_{2,1,1,4} & \alpha_{2,1,1,5} & \alpha_{2,1,1,6} & \alpha_{2,1,2,1} & \alpha_{2,1,2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{1,t-1} \\ y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{1,t-2} \\ y_{2,t-1} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{1,t-\frac{1}{3}} \\ \epsilon_{1,t-\frac{2}{3}} \\ \epsilon_{2,t} \end{bmatrix}.$$

To keep the example simple we assume that all lags are modeled as Almon lag polynomials of order $Q = 1$. To exemplify the data transformation, consider the block transformation matrix of Equation (9) for the third observation of the first variable ($i = 1, \tau_1 = 3$)

$$M_{1,3} = \text{diag}(M_{1,3,1}, M_{1,3,2}) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

As shown in Equation (10) this matrix can be used to build the transformed data vector

$$\begin{aligned} \mathbf{x}_{1,3,t}^* &= \begin{bmatrix} \mathbf{x}_{1,3,1,0,t}^* \\ \mathbf{x}_{1,3,1,1,t}^* \\ \mathbf{x}_{1,3,2,0,t}^* \\ \mathbf{x}_{1,3,2,1,t}^* \end{bmatrix}_{4 \times 1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{1,t-1} \\ y_{1,t-\frac{4}{3}} \\ y_{1,t-\frac{5}{3}} \\ y_{1,t-2} \\ y_{2,t-1} \\ y_{2,t-2} \end{bmatrix} \\ &= \begin{bmatrix} y_{1,1,t-\frac{1}{3}} + 1 \cdot y_{1,t-\frac{2}{3}} + 1 \cdot y_{1,t-1} + 1 \cdot y_{1,t-\frac{4}{3}} \\ y_{1,1,t-\frac{1}{3}} + 2 \cdot y_{1,t-\frac{2}{3}} + 3 \cdot y_{1,t-1} + 4 \cdot y_{1,t-\frac{4}{3}} \\ 1 \cdot y_{2,t-1} + 1 \cdot y_{2,t-2} \\ 1 \cdot y_{2,t-1} + 2 \cdot y_{2,t-2} \end{bmatrix} \end{aligned}$$

Further, we construct the big parameter matrix of Equation (12) which contains all Almon parameters of the MFVAR:

$$\Theta = \begin{bmatrix} \Theta_1 \\ \Theta_2 \end{bmatrix} = \begin{bmatrix} \theta_{1,3} \\ \theta_{1,2} \\ \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} = \begin{bmatrix} \theta_{1,3,1,0} & \theta_{1,3,1,1} & \theta_{1,3,2,0} & \theta_{1,3,2,1} \\ \theta_{1,2,1,0} & \theta_{1,2,1,1} & \theta_{1,2,2,0} & \theta_{1,2,2,1} \\ \theta_{1,1,1,0} & \theta_{1,1,1,1} & \theta_{1,1,2,0} & \theta_{1,1,2,1} \\ \theta_{2,1,1,0} & \theta_{2,1,1,1} & \theta_{2,1,2,0} & \theta_{2,1,2,1} \end{bmatrix}$$

Accordingly, the big data matrix of Equation (13), which contains all transformed data, writes

$$X_t^* = \begin{bmatrix} X_{1,t}^* \\ X_{2,t}^* \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1,3,t}^{*'} \\ \mathbf{x}_{1,2,t}^{*'} \\ \mathbf{x}_{1,1,t}^{*'} \\ \mathbf{x}_{2,1,t}^{*'} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1,3,1,0,t}^* & \mathbf{x}_{1,3,1,1,t}^* & \mathbf{x}_{1,3,2,0,t}^* & \mathbf{x}_{1,3,2,1,t}^* \\ \mathbf{x}_{1,2,1,0,t}^* & \mathbf{x}_{1,2,1,1,t}^* & \mathbf{x}_{1,2,2,0,t}^* & \mathbf{x}_{1,2,2,1,t}^* \\ \mathbf{x}_{1,1,1,0,t}^* & \mathbf{x}_{1,1,1,1,t}^* & \mathbf{x}_{1,1,2,0,t}^* & \mathbf{x}_{1,1,2,1,t}^* \\ \mathbf{x}_{2,1,1,0,t}^* & \mathbf{x}_{2,1,1,1,t}^* & \mathbf{x}_{2,1,2,0,t}^* & \mathbf{x}_{2,1,2,1,t}^* \end{bmatrix}.$$

With $L = \sum_{i=1}^I \mathcal{T}_i = 4$ the selection matrix of Equation (14) writes

$$S_{4 \times 16} = (\mathbf{I}_4 * \mathbf{I}_4)' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The above matrices can then be used to rewrite the Almon MIDAS type MFVAR from Equation (5) into the representation given in Equation (15). As shown in Section 2 we can further transform the MFVAR into the linear system of equations displayed in Equation (16), namely $\mathbf{y}_t = S(X_t^* \otimes I_L) \text{vec}(\Theta) + \boldsymbol{\epsilon}_t$. In case of this example with one monthly and one quarterly variable,

$$S(X_t^* \otimes I_4) = \begin{bmatrix} \mathbf{x}_{1,3,1,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,3,1,1,t}^* & 0 & 0 & 0 & \dots \\ 0 & \mathbf{x}_{1,2,1,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,2,1,1,t}^* & 0 & 0 & \dots \\ 0 & 0 & \mathbf{x}_{1,1,1,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,1,1,1,t}^* & 0 & \dots \\ 0 & 0 & 0 & \mathbf{x}_{2,1,1,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{2,1,1,1,t}^* & \dots \\ \dots & \mathbf{x}_{1,3,2,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,3,2,1,t}^* & 0 & 0 & 0 \\ \dots & 0 & \mathbf{x}_{1,2,2,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,2,2,1,t}^* & 0 & 0 \\ \dots & 0 & 0 & \mathbf{x}_{1,1,2,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{1,1,2,1,t}^* & 0 \\ \dots & 0 & 0 & 0 & \mathbf{x}_{2,1,2,0,t}^* & 0 & 0 & 0 & \mathbf{x}_{2,1,2,1,t}^* \end{bmatrix}$$

and

$$\text{vec}(\Theta) = \begin{bmatrix} \theta_{1,3,1,0} \\ \theta_{1,2,1,0} \\ \theta_{1,1,1,0} \\ \theta_{2,1,1,0} \\ \theta_{1,3,1,1} \\ \theta_{1,2,1,1} \\ \theta_{1,1,1,1} \\ \theta_{2,1,1,1} \\ \theta_{1,3,2,0} \\ \theta_{1,2,2,0} \\ \theta_{1,1,2,0} \\ \theta_{2,1,2,0} \\ \theta_{1,3,2,1} \\ \theta_{1,2,2,1} \\ \theta_{1,1,2,1} \\ \theta_{2,1,2,1} \end{bmatrix},$$

so that the linear system of equations of Equation (16) actually writes

$$\begin{bmatrix} y_{1,t} \\ y_{1,t-\frac{1}{3}} \\ y_{1,t-\frac{2}{3}} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \theta_{1,3,1,0} \cdot \mathbf{x}_{1,3,t,1}^* + \theta_{1,3,1,1} \cdot \mathbf{x}_{1,3,t,2}^* + \theta_{1,3,2,0} \cdot \mathbf{x}_{1,3,t,3}^* + \theta_{1,3,2,1} \cdot \mathbf{x}_{1,3,t,4}^* \\ \theta_{1,2,1,0} \cdot \mathbf{x}_{1,2,t,1}^* + \theta_{1,2,1,1} \cdot \mathbf{x}_{1,2,t,2}^* + \theta_{1,2,2,0} \cdot \mathbf{x}_{1,2,t,3}^* + \theta_{1,2,2,1} \cdot \mathbf{x}_{1,2,t,4}^* \\ \theta_{1,1,1,0} \cdot \mathbf{x}_{1,1,t,1}^* + \theta_{1,1,1,1} \cdot \mathbf{x}_{1,1,t,2}^* + \theta_{1,1,2,0} \cdot \mathbf{x}_{1,1,t,3}^* + \theta_{1,1,2,1} \cdot \mathbf{x}_{1,1,t,4}^* \\ \theta_{2,1,1,0} \cdot \mathbf{x}_{2,1,t,1}^* + \theta_{2,1,1,1} \cdot \mathbf{x}_{2,1,t,2}^* + \theta_{2,1,2,0} \cdot \mathbf{x}_{2,1,t,3}^* + \theta_{2,1,2,1} \cdot \mathbf{x}_{2,1,t,4}^* \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{1,t-\frac{1}{3}} \\ \epsilon_{1,t-\frac{2}{3}} \\ \epsilon_{2,t} \end{bmatrix}.$$

The θ -parameters in this system can be easily estimated row-by-row using OLS. They are then used to calculate the above α -weights.