

REGULAR ARTICLE

# A new approach for estimating VAR systems in the mixed-frequency case

Lukas Koelbl<sup>1</sup> · Manfred Deistler<sup>2</sup>

Received: 31 August 2016 / Revised: 29 October 2017 / Published online: 3 February 2018 © The Author(s) 2018. This article is an open access publication

**Abstract** In this paper we present a new estimation procedure named MF-IVL for VAR systems in the case of mixed-frequency data, where the data maybe, e.g., stock or flow data. The main idea of this new procedure is to project the slow components on the present and past fast ones in order to create instrumental variables. This procedure is shown to be generically consistent. Our claim is that the procedure is fast and more accurate when compared to the extended Yule-Walker procedure. A comparison of these two procedures is given by simulation.

Keywords High-frequency VAR · Mixed-frequency data · Estimation procedure

# **1** Introduction

We propose a simple and fast algorithm for estimating the parameters in a multivariate high-frequency VAR system from mixed-frequency data. The VAR system is of the form

$$y_{t} = \begin{pmatrix} y_{t}^{f} \\ y_{t}^{s} \end{pmatrix} = A_{1}y_{t-1} + \dots + A_{p}y_{t-p} + \nu_{t}, \ t \in \mathbb{Z},$$
(1.1)

 Manfred Deistler deistler@tuwien.ac.at
 Lukas Koelbl lukas.koelbl@accenture.com

- <sup>1</sup> Accenture Digital, Accenture Austria, Vienna, Austria
- <sup>2</sup> Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

where  $A_i \in \mathbb{R}^{n \times n}$  and the AR order *p* is given. Throughout we assume the stability condition

$$\det (a(z)) \neq 0 \ |z| \le 1, \tag{1.2}$$

where  $a(z) = I_n - A_1 z - \cdots - A_p z^p$ . Here z is used for the complex variable as well as for the backward shift on the integers  $\mathbb{Z}$ . We assume that  $(v_t)$  is white noise and we only consider the stable steady state solution  $y_t = a(z)^{-1}v_t$ . The innovation covariance matrix

$$\Sigma_{\nu} = \mathbb{E}\left(\nu_{t}\nu_{t}^{T}\right) > 0 \tag{1.3}$$

is assumed to be non-singular. The parameter space for the high-frequency models considered is:

$$\Theta = \left\{ \left( A_1, \dots, A_p \right) \mid \det\left( a(z) \right) \neq 0, \ |z| \le 1 \right\} \times \left\{ \operatorname{vech}\left( \Sigma_{\nu} \right) \mid \Sigma_{\nu} = \Sigma_{\nu}^T, \ \Sigma_{\nu} > 0 \right\},$$
(1.4)

where vech is the half-vectorization. Note that the conditions (1.2) and (1.3) define an open subset in the euclidean space  $\mathbb{R}^{pn^2+n(n+1)/2}$ . A subset of  $\Theta$  is called generic if it contains an open and dense subset of  $\Theta$ . The VAR system (1.1) can be written in state space form as

$$\underbrace{\begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{x_{t+1}} = \underbrace{\begin{pmatrix} A_1 \cdots A_{p-1} & A_p \\ I_n & & \\ & \ddots & \\ & & I_n & 0 \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{x_t} + \underbrace{\begin{pmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\mathcal{B}} v_t, \quad (1.5)$$

$$y_t = (A_1 \cdots A_p) x_t + v_t. \quad (1.6)$$

In this paper we consider the problem of estimating the parameters of the n-dimensional high-frequency VAR model (1.1) using mixed-frequency data. We actually observe mixed-frequency data of the form

$$\begin{pmatrix} y_t^f \\ w_t \end{pmatrix}, \tag{1.7}$$

where

$$w_t = \sum_{i=1}^{N} c_i y_{t-i+1}^s, \qquad (1.8)$$

where  $c_i \in \mathbb{R}$ ,  $1 < N \in \mathbb{N}$  and at least one  $c_i \neq 0$ . Here the  $n_f$ -dimensional, say, fast component  $y_t^f$  is observed at the highest (sampling) frequency  $t \in \mathbb{Z}$  and the

 $n_s$ -dimensional slow component  $w_t$  is observed only for  $t \in N\mathbb{Z}$ , i.e. for every *N*-th time point. In this paper we assume that  $n_f \geq 1$ . The population second moments, which can be directly observed, are of the form

$$\gamma^{ff}(h) = \mathbb{E}\left(y_{t+h}^{f}\left(y_{t}^{f}\right)^{T}\right), \quad h \in \mathbb{Z},$$
  
$$\gamma^{wf}(h) = \mathbb{E}\left(w_{t+h}\left(y_{t}^{f}\right)^{T}\right), \quad h \in \mathbb{Z},$$
  
$$\gamma^{ww}(h) = \mathbb{E}\left(w_{t+h}w_{t}^{T}\right), \quad h \in N\mathbb{Z}.$$
 (1.9)

Generic identifiability of the high-frequency parameters  $A_i$ , i = 1, ..., p and  $\Sigma_{\nu}$  has been shown in Anderson et al. (2016) (Theorems 2 and 3). Estimation procedures, in particular, a procedure based on the extended Yule-Walker (XYW) equations [see Chen and Zadrozny (1998)] and a procedure based on the Gaussian Likelihood as well as an EM algorithm are discussed in Koelbl et al. (2016) and Koelbl (2015). There it is shown that the MLE as well as the EM estimator heavily depend on the initial estimator used. The purpose of this paper is to describe an estimation procedure which can be used as an initial estimator, e.g. for the EM algorithm, but also as an estimator on its own, because it is easy to calculate, consistent and outperforms the estimator based on the XYW equations.

#### 2 The mixed-frequency IVL estimator

#### 2.1 The stock case

For the case of stock variables (i.e.  $c_1 = 1, c_i = 0, i = 2, ..., N$ ) the second moments, which can be directly observed, are:

$$\gamma^{ff}(h) = \mathbb{E}\left(y_{t+h}^{f}\left(y_{t}^{f}\right)^{T}\right), \quad h \in \mathbb{Z},$$
  
$$\gamma^{sf}(h) = \mathbb{E}\left(y_{t+h}^{s}\left(y_{t}^{f}\right)^{T}\right), \quad h \in \mathbb{Z},$$
  
$$\gamma^{ss}(h) = \mathbb{E}\left(y_{t+h}^{s}\left(y_{t}^{s}\right)^{T}\right), \quad h \in N\mathbb{Z}.$$
 (2.1)

In Anderson et al. (2016) it is shown that the system parameters can be generically reconstructed with the help of the extended Yule-Walker (XYW) equations, which can be constructed by postmultiplying Eq. (1.1) by  $\left(y_{t-j}^f\right)^T$ , j = 1, ..., np and forming expectations:

$$\mathbb{E}\underbrace{\left[y_t\left(\left(y_{t-1}^f\right)^T,\ldots,\left(y_{t-np}^f\right)^T\right)\right]}_{=:Z_1}$$

$$= (A_1, \dots, A_p) \mathbb{E} \left[ \begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \left( \begin{pmatrix} y_{t-1}^f \end{pmatrix}^T, \dots, \begin{pmatrix} y_{t-np}^f \end{pmatrix}^T \right) \right]_{=:Z_0}$$

Note, that the second moments on the left as well as those on the right hand side of the above equation can be directly observed in the mixed-frequency stock case. In Anderson et al. (2016) Theorem 2, it is shown that  $Z_0$  has generically full row rank and therefore we generically obtain  $(A_1, \ldots, A_p) = Z_1 Z_0^T (Z_0 Z_0^T)^{-1}$ . The XYW estimators are obtained by replacing the population second moments by their sample counterparts:

$$\hat{\gamma}^{ff}(h) = \frac{1}{T} \sum_{t=1}^{T-h} y_{t+h}^f \left( y_t^f \right)^T, \quad h \ge 0,$$
(2.2)

$$\hat{\gamma}^{ff}(h) = \hat{\gamma}^{ff}(-h)^T, \qquad (2.3)$$

$$\hat{\gamma}^{sf}(h) = \frac{1}{T/N} \sum_{t} y_{Nt}^{s} \left( y_{Nt-h}^{f} \right)^{T}, \qquad (2.4)$$

where the estimator of  $\gamma^{sf}(h)$  has only (approximately) <sup>1</sup>/*N*-th of the summands compared to the estimator of  $\gamma^{ff}(h)$  due to the missing observations [see Koelbl et al. (2016)].

The new estimation procedure proposed is as follows: The basic idea is to generate instrumental variables by projecting the slow components  $y_t^s$  on the space generated by present and a sufficient number of lagged fast components  $y_j^f$ . To be more precise, let, for a suitable chosen  $1 \le k \le t$ ,  $\mathcal{H}_k^f(t) = \text{span}\left\{y_j^f: t - k \le j \le t\right\}$  be the Hilbert space spanned by the one-dimensional components of the  $y_j^f$  in the underlying space of square integrable random variables  $\mathcal{L}^2$  over  $(\Omega, \mathcal{A}, P)$  and let  $x_{t|t-1}^k$  denote the (componentwise) projection of the state  $x_t$  onto  $\mathcal{H}_k^f(t-1)$  written as  $x_{t|t-1}^k = \mathbb{P}_{\mathcal{H}_k^f(t-1)}(x_t)$ . Projecting the state Eq. (1.5) onto  $\mathcal{H}_k^f(t)$ , we obtain, using an obvious notation,

$$x_{t+1|t}^{k} = \mathcal{A}x_{t|t-1}^{k} + \left\{ \mathcal{A}\left(x_{t|t}^{k} - x_{t|t-1}^{k}\right) + \mathcal{B}v_{t|t}^{k} \right\}.$$
 (2.5)

In a first step we show that the matrix  $\mathbb{E}\left(x_{t|t-1}^{k}\left(x_{t|t-1}^{k}\right)^{T}\right)$  is generically non-singular for  $k \ge np-1$ : For  $k_0 = np-1$  and  $Y_{t,k}^{-} = \left(\left(y_t^{f}\right)^{T}, \left(y_{t-1}^{f}\right)^{T}, \dots, \left(y_{t-k}^{f}\right)^{T}\right)^{T}$ let  $\Gamma^{ff}(k) = \mathbb{E}\left(Y_{t,k}^{-}\left(Y_{t,k}^{-}\right)^{T}\right)$ . It follows that  $\Gamma^{ff}(k_0) > 0$  which is a direct conse-

1206

quence of  $\Sigma_{\nu} > 0$ . The projection  $x_{t|t-1}^{k_0}$  is obtained by using the OLS formula

$$x_{t|t-1}^{k_0} = \underbrace{\mathbb{E}\left(x_t \left(Y_{t-1,k_0}^{-}\right)^T\right)}_{Z_0} \Gamma^{ff} (k_0)^{-1} Y_{t-1,k_0}^{-}$$

and therefore generically

$$\mathbb{E}\left(x_{t|t-1}^{k_0}\left(x_{t|t-1}^{k_0}\right)^T\right) = Z_0\Gamma^{ff}(k_0)^{-1}Z_0^T > 0$$
(2.6)

since  $Z_0$  has generically full row rank. This implies that generically

$$\mathcal{A} = \left(\mathbb{E}x_{t+1|t}^{k} \left(x_{t|t-1}^{k}\right)^{T}\right) \left(\mathbb{E}x_{t|t-1}^{k} \left(x_{t|t-1}^{k}\right)^{T}\right)^{-1}$$
(2.7)

holds, since  $x_{t|t-1}^k$  is uncorrelated with  $\left(x_{t|t}^k - x_{t|t-1}^k\right)$  and  $v_{t|t}^k$ . Note that, for  $k \ge p-1$ ,

$$x_{t|t-1}^{k} = \begin{pmatrix} y_{t-1}^{f} \\ \mathbb{P}_{\mathcal{H}_{k}^{f}(t-1)}(y_{t-1}^{s}) \\ \vdots \\ y_{t-p}^{f} \\ \mathbb{P}_{\mathcal{H}_{k}^{f}(t-1)}(y_{t-p}^{s}) \end{pmatrix}.$$

An estimator of the state  $x_{t+1|t}^k$ , denoted by  $\hat{x}_{t+1|t}^k$ , can be constructed as follows: W.l.o.g. let p = 2 and N = 2. The first *n* components of  $\hat{x}_{t+1|t}^k$  can be estimated by projecting  $y_t$  onto  $y_t^f, \ldots, y_{t-k}^f$ . This can be done by estimating  $\beta_1$  in

$$y_t = \beta_1 Y_{t\,k}^- + \varepsilon_t, \quad t \in 2\mathbb{Z},\tag{2.8}$$

Let  $\hat{\beta}_{1,T}$  denote the OLS estimator of  $\beta_1$ . Then we obtain  $(I_n, 0)$   $\hat{x}_{t+1|t}^k = \hat{\beta}_{1,T}Y_{t,k}^-$ ,  $t \in \mathbb{Z}$ . The second *n* components of  $x_{t+1|t}^k$  must be, due to the mixed-frequency structure and N = 2, estimated in a different way: Analogously to (2.8) we can construct

$$y_{t-1} = \beta_2 Y_{t,k}^- + \zeta_t, \tag{2.9}$$

but now we cannot directly observe the left hand side of (2.9). Therefore, we must shift (2.9) to

$$y_t = \beta_2 Y_{t+1,k}^- + \zeta_{t+1}, \quad t \in 2\mathbb{Z}$$
(2.10)

🖉 Springer

which directly leads us to the OLS estimator  $\hat{\beta}_{2,T}$  of  $\beta_2$ , since the left as well as the right hand side of (2.10) can be directly observed. In a last step we can construct the remaining part of the state with the help of  $(0, I_n) \hat{x}_{t+1|t}^k = \hat{\beta}_{2,T} Y_{t,k}^-$ ,  $t \in \mathbb{Z}$ , which leads us to

$$\hat{x}_{t+1|t}^{k} = \begin{pmatrix} \hat{\beta}_{1,T} \\ \hat{\beta}_{2,T} \end{pmatrix} Y_{t,k}^{-}, \quad t \in \mathbb{Z}.$$
(2.11)

Using these instrumental variables, we can estimate A according to (2.7):

$$\hat{\mathcal{A}}_{T} = \left(\sum_{t=p+1}^{T-1} \hat{x}_{t+1|t}^{k} \left(\hat{x}_{t|t-1}^{k}\right)^{T}\right) \left(\sum_{t=p+1}^{T} \hat{x}_{t|t-1}^{k} \left(\hat{x}_{t|t-1}^{k}\right)^{T}\right)^{-1}.$$
 (2.12)

**Theorem 1** Under the additional assumptions, that  $\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} v_t v_t^T = \Sigma_v$  a.s., we have  $\lim_{T\to\infty} \hat{\mathcal{A}}_T = \mathcal{A}$  a.s.

*Proof* Again we assume that p = 2 and N = 2. The above condition  $\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} v_t v_t^T = \Sigma_v$  implies, see Hannan and Deistler (2012) Theorem 4.1.1, that

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T-h} y_{t+h} y_t^T = \gamma (h) , \qquad (2.13)$$

where  $\gamma(j) = \mathbb{E}\left(y_t y_{t-j}^T\right)$ . In a next step, let us write (2.6) as

$$\mathbb{E}\left(x_{t|t-1}^{k_{0}}\left(x_{t|t-1}^{k_{0}}\right)^{T}\right) = Z_{0}\Gamma^{ff}(k_{0})^{-1}\Gamma^{ff}(k_{0})\Gamma^{ff}(k_{0})^{-1}Z_{0}^{T}$$

$$= \binom{\beta_{1}}{\beta_{2}}\Gamma^{ff}(k_{0})\binom{\beta_{1}}{\beta_{2}}^{T}$$
(2.14)

Equation (2.13) implies that  $\hat{\beta}_{1,T}$  and  $\hat{\beta}_{2,T}$  are consistent estimators for  $\beta_1$  and  $\beta_2$ , respectively. Thus, we obtain that

$$\begin{split} \lim_{T \to \infty} \frac{1}{T} \sum_{t=k+2}^{T+1} \hat{x}_{t|t-1}^{k_0} \left( \hat{x}_{t|t-1}^{k_0} \right)^T &= \lim_{T \to \infty} \frac{1}{T} \sum_{t=k+2}^{T+1} \left( \hat{\beta}_{1,T} \atop \hat{\beta}_{2,T} \right) Y_{t-1,k}^- \left( Y_{t-1,k}^- \right)^T \left( \hat{\beta}_{1,T} \atop \hat{\beta}_{2,T} \right)^T \\ &= \lim_{T \to \infty} \left( \frac{\hat{\beta}_{1,T}}{\hat{\beta}_{2,T}} \right) \hat{\Gamma}^{ff} \left( k_0 \right) \left( \frac{\hat{\beta}_{1,T}}{\hat{\beta}_{2,T}} \right)^T \\ &= \mathbb{E} \left( x_{t|t-1}^{k_0} \left( x_{t|t-1}^{k_0} \right)^T \right). \end{split}$$

Deringer

An analogous result can be shown for  $\mathbb{E}\left(x_{t+1|t}^{k}\left(x_{t|t-1}^{k}\right)^{T}\right)$ . This concludes the proof.

Of course the choice of k is important for estimating the system parameters. Our approach is to regress  $y_t^s$  on  $y_t^f, \ldots, y_{t-k}^f$  and to determine the maximum lag k by using AIC. Note that the structure of the matrix A, as far as the a priori zeros and ones are concerned, is not preserved by the estimation procedure (2.12). For this reason, we define a new estimator for the system parameters as

$$\hat{\hat{\mathcal{A}}}_T = \begin{pmatrix} (I_n, 0, \dots, 0) \, \hat{\mathcal{A}} \\ (I_{n(p-1)}, 0) \end{pmatrix}.$$
(2.15)

Clearly,  $\hat{\mathcal{A}}_T$  is also consistent. As shown in Anderson et al. (2016), the innovation covariance matrix  $\Sigma_{\nu}$  can be generically consistently estimated according to the following formula

$$\operatorname{vec}\left(\Sigma_{\nu}\right) = \left(\left(\mathcal{G}\otimes\mathcal{G}\right)\left(I_{(np)^{2}}-\left(\mathcal{A}\otimes\mathcal{A}\right)\right)^{-1}\left(\mathcal{G}^{T}\otimes\mathcal{G}^{T}\right)\right)^{-1}\operatorname{vec}\left(\gamma(0)\right) \quad (2.16)$$

where  $\mathcal{G} = (I_n, 0, ..., 0)$  and where  $\otimes$  denotes the Kronecker symbol. Let  $\hat{\Sigma}_{\nu}$  denote the corresponding estimator.

Note that the estimator  $\hat{A}_T$  (denoted by MF-IVL estimator) neither necessarily gives a stable AR system, nor is  $\hat{\Sigma}_{\nu}$  necessarily positive definite. Projecting a symmetric matrix on the space of positive definite symmetric matrices is in a certain sense a standard procedure (see Higham 1989; Koelbl 2015). Projecting unstable system parameters on the space of stable ones is described in Koelbl (2015) and, for the univariate case, in Orbandexivry et al. (2013). Projecting slow variables on fast lagged variables is also mentioned in Ghysels et al. (2007).

#### 2.2 The flow case

For the case of the more general observation scheme (1.8), we proceed as follows: Let

$$z_t = \sum_{i=1}^{N} c_i y_{t-i+1} = \begin{pmatrix} \sum_{i=1}^{N} c_i y_{t-i+1}^f \\ w_t \end{pmatrix}.$$
 (2.17)

From (1.5) we obtain

$$\underbrace{\begin{pmatrix} z_t \\ z_{t-1} \\ \vdots \\ z_{t-p+1} \end{pmatrix}}_{f_{t+1}} = \mathcal{A} \underbrace{\begin{pmatrix} z_{t-1} \\ z_{t-2} \\ \vdots \\ z_{t-p} \end{pmatrix}}_{f_t} + \mathcal{B} \left( \sum_{i=1}^N c_i v_{t-i+1} \right).$$
(2.18)

Springer

Let  $f_{t+1|t}^k$  denote the projection of  $f_{t+1}$  on the space  $\mathcal{H}_k^f(t)$ . Projecting both sides of (2.18) on the space  $\mathcal{H}_k^f(t)$  we get in an obvious notation

$$f_{t+1|t}^{k} = \mathcal{A}f_{t|t-N}^{k} + \left\{ \mathcal{A}\left(f_{t|t}^{k} - f_{t|t-N}^{k}\right) + \mathcal{B}\left(\sum_{i=1}^{N} c_{i}v_{t-i+1|t}\right) \right\}.$$
 (2.19)

Post-multiplying (2.19) by  $\left(f_{t|t-N}^k\right)^T$  and taking the expectations we obtain

$$\mathbb{E}\left(f_{t+1|t}^{k}\left(f_{t|t-N}^{k}\right)^{T}\right) = \mathcal{A}\mathbb{E}\left(f_{t|t-N}^{k}\left(f_{t|t-N}^{k}\right)^{T}\right).$$
(2.20)

Again, identifiability of the system parameters can been shown if we show that the matrix  $\mathbb{E}\left(f_{t|t-N}^{k}\left(f_{t|t-N}^{k}\right)^{T}\right)$  is non-singular. This is proved as follows: For  $k_{0} = np - 1$  it follows that

$$f_{t|t-N}^{k_0} = \underbrace{\mathbb{E}\left(f_t\left(Y_{t-N,k_0}^{-}\right)^T\right)}_{Z_0^g} \Gamma^{ff}(k_0)^{-1} Y_{t-N,k_0}^{-}$$

since  $\Gamma^{ff}(k_0)$  is again positive definite. Therefore it follows that  $\mathbb{E}\left(f_{t|t-N}^k\left(f_{t|t-N}^k\right)^T\right)$ =  $Z_0^{g}\Gamma^{ff}(k_0)^{-1}\left(Z_0^{g}\right)^T$  is generically non-singular since  $Z_0^{g}$  has generically full row rank (see Koelbl 2015). Using (2.20), a consistent estimation procedure is obtained analogously to the stock case described above. The innovation covariance matrix  $\Sigma_{\nu}$  can be estimated as in Koelbl et al. (2016).

### **3** Simulations

In this section we present a simulation study comparing the accuracy of IVL with the accuracy of the XYW estimator and comparing these procedures as initial estimators for the EM algorithm. We consider the following data generating processes corresponding to the following two models:

*Example 1* Model 1 (which was also presented in Koelbl et al. (2016)) is of the form:

$$y_t = \begin{pmatrix} -1.2141 & 1.1514 \\ -0.9419 & 0.8101 \end{pmatrix} y_{t-1} + v_t,$$
(3.1)

and Model 2 is of the form:

$$y_{t} = \begin{pmatrix} 1.5284 & 0.2727 & 1.0181 \\ 1.6881 & -1.5235 & -1.1424 \\ -0.6785 & 1.0936 & 1.2108 \end{pmatrix} y_{t-1} + \begin{pmatrix} -0.8089 & 0.4224 & 0.1477 \\ -0.4461 & -0.9209 & -0.3154 \\ -0.0496 & 0.6999 & -0.0982 \end{pmatrix} y_{t-2} + \nu_{t}.$$
(3.2)

	Estimators	Model 1		Model 2	
		Absolute	Relative	Absolute	Relative
HF	YW	0.002	1	0.017	1
MF	XYW	0.315	133.67	0.721	43.52
	IVL	0.056	23.53	0.075	4.55
	EM-XYW	0.006	2.38	0.066	3.99
	EM-IVL	0.004	1.74	0.027	1.64

Table 1 Absolute and relative mean squared errors of the system parameters

In both cases the innovations are standard normally distributed, i.e.  $v_t \sim \mathcal{N}(0, I_i)$ , i = 2, 3.

The simulation study reports the mean squared errors

$$MSE\left(\hat{\theta}\right) = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n^2 p} \left(\theta_i - \hat{\theta}_i^j\right)^2$$

for the parameters  $\theta = \text{vec}(A_1)$  and  $\theta = \text{vec}(A_1, A_2)$ , respectively. The sample size is T = 500 and we performed  $m = 10^3$  simulation runs. Only the case of stock variables has been considered. We put N = 2 and  $n_s = 1$ . The following estimation procedures are compared in this study: The Yule-Walker estimator obtained from highfrequency data, denoted by HF-YW. This estimator serves as an overall benchmark and therefore also the mean squared errors relative to the mean squared errors of the HF-YW estimators are presented. By MF-XYW we denote the mixed-frequency XYW estimator, by MF-IVL the mixed-frequency estimator introduced in the paper. By MF-EM-XYW we denote the mixed-frequency EM algorithm initialized with the XYW estimator and MF-EM-IVL the mixed-frequency EM algorithm initialized with the MF-IVL estimator, respectively. Table 1 summarizes the results.

Note that for the two models MF-IVL outperforms MF-XYW as far as the overall mean squared errors are concerned. This also holds for the estimators for the individual system- as well as for the corresponding estimates of the noise parameters. When used as initial estimators, again, MF-IVL outperforms MF-XYW. In addition, the number of iterations for the EM algorithm decreases for both models when initialized with the MF-IVL instead of the MF-XYW.

## **4** Conclusions

This paper proposes a new estimation procedure in the framework of VAR models and mixed-frequency data. The procedure is obtained by creating instrumental variables by projecting the slow variables on present and past fast ones. We show generic consistency of the system parameters for stock and flow variables. Simulations are presented to compare the properties of our procedure compared to the XYW estimator.

Both procedures are less accurate when compared to the MLE, our procedure however outperforms the XYW estimator.

Acknowledgements Open access funding provided by Austrian Science Fund (FWF). Support by the Oesterreichische Nationalbank (Oesterreichische Nationalbank, Anniversary Fund, Project Number: 16546) and the FWF (Austrian Science Fund under contract P24198/N18) is gratefully acknowledged. We thank Peter Zadrozny, U.S. Bureau of Labor Statistics, Washington D.C., and an unknown reviewer for helpful comments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Anderson BDO, Deistler M, Felsenstein E, Funovits B, Koelbl L, Zamani M (2016) Multivariate AR systems and mixed frequency data: g-identifiability and estimation. Econom. Theory 32(4):793–826
- Chen B, Zadrozny PA (1998) An extended yule-walker method for estimating a vector autoregressive model with mixed-frequency data. Adv. Econom. 13:47–73
- Ghysels E, Sinko A, Valkanov R (2007) MIDAS regressions: further results and new directions. Econom. Rev. 26(1):53–90
- Hannan EJ, Deistler M (2012) The statistical theory of linear systems. SIAM Classics in Applied Mathematics, Philadelphia
- Higham JN (1989) Matrix nearness problems and applications. In: Gover M, Barnett S (eds) Applications of matrix theory. Oxford University Press, Oxford, pp 1–27
- Koelbl L (2015) VAR systems: g-identifiability and asymptotic properties of parameter estimates for the mixed-frequency case. PhD thesis, Vienna University of Technology
- Koelbl L, Braumann A, Felsenstein E, Deistler M (2016) Estimation of VAR systems from mixed-frequency data: the stock and the flow case. Adv. Econom. 35:43–73
- Orbandexivry F-X, Nesterov Y, Van Dooren P (2013) Nearest stable system using successive convex approximations. Automatica 49:1195–1203