A New Approach for Estimating VAR Systems in the Mixed-Frequency Case

Lukas Koelbl[*], Manfred Deistler[†]

August 31, 2016

Key Words: high-frequency VAR, mixed-frequency data, estimation

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.

[*]Lukas Koelbl is with the Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, e-mail: lukas.koelbl@tuwien.ac.at

[†]Corresponding author. Manfred Deistler is with the Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, e-mail: deistler@tuwien.ac.at.

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

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \nu_t,\ t \in \mathbb{Z}, \tag{1.1}$$

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

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

where $a(z) = I - A_1 z - \cdots - A_p z^p$. Here z is used for the complex variable as well as for the backward shift on the integers $\mathbb{Z}$. We assume that (ν_t) is white noise and we only consider the stable steady state solution $y_t = a(z)^{-1}\nu_t$. The innovation covariance matrix $\Sigma_\nu = \mathbb{E}\left(\nu_t \nu_t^T\right)$ is assumed to be non-singular. The VAR system (1.1) can be written in state space form as

$$\underbrace{\begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{x_{t+1}} = \underbrace{\begin{pmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{x_t} + \underbrace{\begin{pmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\mathcal{B}} \nu_t, \tag{1.3}$$

$$y_t = \begin{pmatrix} A_1 & \cdots & A_p \end{pmatrix} x_t + \nu_t. \tag{1.4}$$

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

$$\begin{pmatrix} y_t^f \\ w_t \end{pmatrix}, \tag{1.5}$$

where

$$w_t = \sum_{i=1}^{N} c_i y_{t-i+1}^s, \tag{1.6}$$

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 $n_f \geq 1$.

Generic identifiability of the high-frequency parameters A_i, $i = 1, \ldots, p$ and Σ_ν has been shown in Anderson et al. (2016). 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); 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 both as an estimator on its own as well as an initial estimator, because it is easy to calculate and outperforms the estimator based on the XYW equations.

2 The Mixed-Frequency IVL Estimator

For the case of stock variables (i.e. $c_1 = 1$, $c_i = 0$, $i = 2, \ldots, N$) the 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 $1 \leq k \leq t$, $\mathcal{H}_k^f(t) = \text{span}\left\{y_j^f : t-k \leq j \leq t\right\}$ 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 $(\Omega, \mathcal{A}, \mathbb{P})$ and let $x_{t|t-1}^k$ denote the (componentwise) projection of the state x_t onto $\mathcal{H}_k^f(t-1)$. Projecting the state equation (1.3) onto $\mathcal{H}_k^f(t)$, we obtain

$$x_{t+1|t}^k = \mathcal{A}x_{t|t-1}^k + \left\{\mathcal{A}\left(x_{t|t}^k - x_{t|t-1}^k\right) + \mathcal{B}\nu_{t|t}^k\right\}. \tag{2.1}$$

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 nonsingular for $k \geq np-1$: For $k_0 = np-1$ and $Y_{t,k}^- = \left(\left(y_t^f\right)^T, \left(y_{t-1}^f\right)^T, \ldots, \left(y_{t-k}^f\right)^T\right)^T$ it follows that $\Gamma^{ff}(k_0) = \mathbb{E}\left(Y_{t,k_0}^-\left(Y_{t,k_0}^-\right)^T\right) > 0$,

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

and therefore generically

$$\mathbb{E}\left(x_{t|t-1}^{k_0}\left(x_{t|t-1}^{k_0}\right)^T\right) = Z_0\Gamma^{ff}(k_0)^{-1}Z_0^T > 0 \tag{2.2}$$

since Z_0 has generically full row rank (see Anderson et al. (2016)).

This implies that

$$\mathcal{A} = \left(\mathbb{E}x_{t+1|t}^k\left(x_{t|t-1}^k\right)^T\right)\left(\mathbb{E}x_{t|t-1}^k\left(x_{t|t-1}^k\right)^T\right)^{-1} \tag{2.3}$$

since $x^k_{t|t-1}$ is uncorrelated with $\left(x^k_{t|t} - x^k_{t|t-1}\right)$ and $\nu^k_{t|t}$. Note that, for $k \geq p-1$,

$$x^k_{t|t-1} = \begin{pmatrix} y^f_{t-1} \\ \mathbb{P}_{\mathcal{H}^f_k(t-1)}\left(y^s_{t-1}\right) \\ \vdots \\ y^f_{t-p} \\ \mathbb{P}_{\mathcal{H}^f_k(t-1)}\left(y^s_{t-p}\right) \end{pmatrix}.$$

In the next step the instruments $\hat{x}^k_{t+1|t}$ are obtained by using the coefficient of the regression, e.g. y^s_t on $y^f_t, \ldots, y^f_{t-k}$ in order to approximate y^s_t, and the coefficient of the regression y^s_t on $y^f_{t+1}, \ldots, y^f_{t-k+1}$ in order to approximate y^s_{t-1} for $t \in N\mathbb{Z}$. Using this instrumental variables, we estimate $\mathcal{A}$ according to (2.3). From what was said above, under suitable conditions guaranteeing consistent estimation of the respective population second moments by their sample counterparts, the estimator of the system parameters $A_1, \ldots, A_p$ can easily be shown to be consistent (see e.g. Hannan (1970); Koelbl (2015)). As shown in Anderson et al. (2016), the innovation covariance matrix Σ_ν can be generically consistently estimated according to the following formula

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

where $\mathcal{G} = (I_n, 0, \ldots, 0)$, $\otimes$ denotes the Kronecker symbol and $\gamma(j) = \mathbb{E}\left(y_t y^T_{t-j}\right)$. Of course the choice of k is important for estimating the system parameters. Our approach is to regress y^s_t on $y^f_t, \ldots, y^f_{t-k}$ and to determine the maximum lag k by using AIC. Note that the structure of the matrix $\mathcal{A}$, as far as the a priori zeros and ones are concerned, is not preserved by the estimation procedure described.

Note that the estimator (denoted by MF-IVL estimator) described above does not necessarily give a stable AR system, nor a positive definite innovation covariance matrix. Projecting a symmetric matrix on the space of positive definite 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).

For the case of the more general observation scheme (1.6), we proceed as follows: Let

$$z_t = \sum_{i=1}^{N} c_i y_{t-i+1} = \begin{pmatrix} \sum_{i=1}^{N} c_i y^f_{t-i+1} \\ w_t \end{pmatrix}. \tag{2.4}$$

From (1.3) we obtain

$$\underbrace{\begin{pmatrix} z_t \\ z_{t-1} \\ \vdots \\ z_{t-p+1} \end{pmatrix}}_{f_{t+1}} = \mathcal{A} \underbrace{\begin{pmatrix} z_{t-1} \\ z_{t-2} \\ \vdots \\ z_{t-p} \end{pmatrix}}_{f_t} + \mathcal{B} \left(\sum_{i=1}^{N} c_i \nu_{t-i+1} \right). \tag{2.5}$$

Let $f^k_{t+1|t}$ denote the projection of f_{t+1} on the space $\mathcal{H}^f_k(t)$. Projecting both sides of (2.5) on the space $\mathcal{H}^f_k(t)$ we get in an obvious notation

$$f^k_{t+1|t} = \mathcal{A} f^k_{t|t-N} + \left\{ \mathcal{A} \left(f^k_{t|t} - f^k_{t|t-N} \right) + \mathcal{B} \left(\sum_{i=1}^{N} c_i \nu_{t-i+1|t} \right) \right\}. \tag{2.6}$$

Post-multiplying (2.6) by $\left(f^k_{t|t-N} \right)^T$ and taking the expectations we obtain

$$\mathbb{E} \left(f^k_{t+1|t} \left(f^k_{t|t-N} \right)^T \right) = \mathcal{A} \mathbb{E} \left(f^k_{t|t-N} \left(f^k_{t|t-N} \right)^T \right). \tag{2.7}$$

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

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

since $\Gamma^{ff}(k_0)$ is again positive definite. Therefore it follows that $\mathbb{E}\left(f_{t|t-N}^{k}\left(f_{t|t-N}^{k}\right)^{T}\right) = Z_0^{\mathrm{g}}\Gamma^{ff}(k_0)^{-1}(Z_0^{\mathrm{g}})^{T}$ is generically non-singular since Z_0^{g} has generically full row rank (see Koelbl (2015)). Using (2.7), a consistent estimation procedure is obtained analogously to the stock case described above. The innovation covariance matrix Σ_ν 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 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:

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

and Model 2 is of the form:

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

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

		Model 1		Model 2	
	Estimators	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

The simulation study reports the mean squared errors

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

for the parameters $\theta = \mathrm{vec}\,(A_1)$ and $\theta = \mathrm{vec}\,(A_1, A_2)$, 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 has been compared in this study: The Yule-Walker estimator obtained from high-frequency 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.

5 Acknowledgments

Support by the Oesterreichische Nationalbank (Oesterreichische Nationalbank, Anniversary Fund, project number: 16546) and the FWF (Austrian Science Fund under contract P24198/N18).

References

References

Anderson, B. D. O., Deistler, M., Felsenstein, E., Funovits, B., Koelbl, L., and Zamani, M. (2016). Multivariate AR Systems and Mixed Frequency Data: g-Identifiability and Estimation. *Econometric Theory*, 32(4):793–826.

Chen, B. and Zadrozny, P. A. (1998). An Extended Yule-Walker Method for

Estimating a Vector Autoregressive Model with Mixed-Frequency Data. *Advances in Econometrics*, 13:47–73.

Hannan, E. J. (1970). *Multiple Time Series*. John Wiley & Sons, New York.

Higham, J. N. (1989). Matrix Nearness Problems and Applications. In Gover, M. and Barnett, S., editors, *Applications of Matrix Theory*, pages 1–27. Oxford University Press.

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., and Deistler, M. (2016). Estimation of VAR Systems from Mixed-Frequency Data: The Stock and the Flow Case. *Advances in Econometrics*, 35:43–73.

Orbandexivry, F.-X., Nesterov, Y., and Van Dooren, P. (2013). Nearest Stable System Using Successive Convex Approximations. *Automatica*, 49:1195–1203.