Efficient Estimation of State-Space Mixed-Frequency VARs: A Precision-Based Approach

Joshua Chan* Aubrey Poon** Dan Zhu***

*Purdue University

**Örebro University, Sweden

***Monash University

ESCoE Conference 2022

Motivation

- Mixed-Frequency (MF) VARs have become increasingly popular tool within the empirical macroeconomics for forecasting and nowcasting.
- There are two common approaches to handle the mixed-frequency variables:
  - Stacked approach - Ghysels (2016) JoE
  - State-space approach - Schorfheide and Song (2015) JBES
- Main goal of the paper: **We develop an efficient precision-based algorithm for estimating MF-State-Space VARs.**

State-Space Vs Stacked Approach

- The key advantage the MF-State-Space VARs have over the MF-Stacked VARs is that models all the variables at the highest observed frequency.
  - Allows the researcher to obtain the interpolated historical estimates of the low-frequency at higher frequency.
  - More timely nowcast.
- The main drawback of the MF-State-Space VARs is its estimation tends be computationally intensive.
  - This is mainly due to the simulation of large number of latent states (missing observations of the low-frequency variables) using standard filtering and smoothing techniques.

MF-State-Space VARs: A Precision-Based Approach

- Since the seminal work by Chan and Jeliazkov (2009) and McCausland, Miller, and Pelletier (2011), precision-based samplers have been used in a wide range of empirical applications:
  - Trend inflation: Chan, Koop, and Potter (2013, 2016).
  - Macroeconomic Forecasting: Cross, Hou, and Poon (2020).
  - DFM: Kaufmann and Schumacher (2019) and Beyeler and Kaufmann (2021).
  - Uncertainty: Cross et al. (2021).
- Our main contribution: **We develop a precision-based sampling approach to state-space models with missing observations.**

MF-State-Space VARs: A Precision-Based Approach (cont.)

- Drawing the missing observations of the low-frequency variable from the precision-based sampler involves two main steps:
  1. We derive the conditional distribution of the missing observations of the VAR.
  2. We sample from this conditional distribution subject to a linear (inter-temporal) constraint.

The Conditional Distribution of the Missing Observations

- Let us define a MF-VAR with p lags for $\mathbf{y}_t = (\mathbf{y}_t^{o'}, \mathbf{y}_t^{u'})'$ of dimension $n = n_o + n_u$

$$\mathbf{y}_t = \mathbf{b}_0 + \mathbf{B}_1\mathbf{y}_{t-1} + \ldots + \mathbf{B}_p\mathbf{y}_{t-p} + \epsilon_t, \epsilon_t \sim N(0, \Sigma), \quad (1)$$

- For example, a standard monthly-quarterly MF-VAR:
  - $\mathbf{y}_t^o$ consists of n_o monthly variables that are observed at every month t.
  - $\mathbf{y}_t^u$ consists of n_u quarterly variables at monthly frequency that are only observed every 3 months.

The Conditional Distribution of the Missing Observations (cont.)

- We can $\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_T')'$, and rewrite (1) into a standard linear regression matrix form

$$\mathbf{HY} = \mathbf{c} + \epsilon, \epsilon \sim N(0, \Xi), \tag{2}$$

- where $\mathbf{c} = \mathbf{1}_{T-p} \otimes \mathbf{b}_0$, $\epsilon = (\epsilon_1', \ldots, \epsilon_T')'$, $\Xi = \mathbf{I}_{T-p} \otimes \Sigma$, and

$$\mathbf{H} = \begin{bmatrix} -\mathbf{B}_p & \cdots & -\mathbf{B}_1 & \mathbf{I}_n & \mathbf{O}_n & \cdots & \cdots & \mathbf{O}_n \\ \mathbf{O}_n & -\mathbf{B}_p & \cdots & -\mathbf{B}_1 & \mathbf{I}_n & \mathbf{O}_n & \cdots & \mathbf{O}_n \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \mathbf{O}_n & \cdots & \mathbf{O}_n & -\mathbf{B}_p & \cdots & -\mathbf{B}_1 & \mathbf{I}_n & \mathbf{O}_n \\ \mathbf{O}_n & \cdots & \cdots & \mathbf{O}_n & -\mathbf{B}_p & \cdots & -\mathbf{B}_1 & \mathbf{I}_n \end{bmatrix}.$$

- Note here $\mathbf{H}$ is a banded and sparse matrix.

The Conditional Distribution of the Missing Observations (cont.)

- Next, we can write $\mathbf{Y}$ as a linear combination of the observed (high-frequency) and unobserved (low-frequency) variables as:

$$\mathbf{Y} = \mathbf{M}_o\mathbf{Y}^o + \mathbf{M}_u\mathbf{Y}^u, \quad (3)$$

- where $\mathbf{Y}^o = (\mathbf{y}_1^{o'}, \ldots, \mathbf{y}_T^{o'})'$ and $\mathbf{Y}^u = (\mathbf{y}_1^{u'}, \ldots, \mathbf{y}_T^{u'})'$. Both $\mathbf{M}_o$ and $\mathbf{M}_u$ are selection matrices.
- We can substitute (3) into (2) which give us

$$\mathbf{H}(\mathbf{M}_o\mathbf{Y}^o + \mathbf{M}_u\mathbf{Y}^u) = \mathbf{c} + \epsilon, \epsilon \sim N(0, \Xi), \quad (4)$$

The Conditional Distribution of the Missing Observations (cont.)

- From equation (4), we can derive the joint density of $\mathbf{Y}^u$ conditional on $\mathbf{Y}^o$ and model parameters, $p(\mathbf{Y}^u|\mathbf{Y}^o, \mathbf{B}, \Sigma)$

$$\propto \exp\left\{-\frac{1}{2}\left[\mathbf{Y}^{u'}\mathbf{M}_u'\mathbf{H}'\Xi^{-1}\mathbf{H}\mathbf{M}_u\mathbf{Y}^u - 2\mathbf{Y}^{u'}\mathbf{M}_u'\mathbf{H}'\Xi^{-1}(\mathbf{c} - \mathbf{H}\mathbf{M}\mathbf{Y}^o)\right]\right\},$$

- If we assume $\mathbf{K}_{\mathbf{Y}^u} = \mathbf{M}_u'\mathbf{H}'\Xi^{-1}\mathbf{H}\mathbf{M}_u$ and $\mu_{\mathbf{Y}^u} = \mathbf{K}_{\mathbf{Y}^u}^{-1}(\mathbf{M}_u'\mathbf{H}'\Xi^{-1}(\mathbf{c} - \mathbf{H}\mathbf{M}\mathbf{Y}^o)$, then by the completing the square of $\mathbf{Y}^u$, we can write the conditional density of $\mathbf{Y}^u$

$$p(\mathbf{Y}^u|\mathbf{Y}^o, \mathbf{B}, \Sigma) \propto \exp\left\{-\frac{1}{2}\left[(\mathbf{Y}^u - \mu_{\mathbf{Y}^u})'\mathbf{K}_{\mathbf{Y}^u}^{-1}(\mathbf{Y}^u - \mu_{\mathbf{Y}^u})\right]\right\},$$

The Conditional Distribution of the Missing Observations (cont.)

- Thus, the conditional distribution of the missing observations given the observed data is Gaussian
$$(\mathbf{Y}^u|\mathbf{Y}^o, \mathbf{B}, \Sigma) \sim N(\mu_{\mathbf{Y}^u}, \mathbf{K}_{\mathbf{Y}^u}),$$
- Since $\mathbf{H}$, Ξ and $\mathbf{M}_u$ are all band matrices, so is the precision matrix $\mathbf{K}_{\mathbf{Y}^u}$. Therefore,we can use the precision-based sampler of Chan and Jeliazkov (2009) to draw $\mathbf{Y}^u$ efficiently.

Inter-temporal restrictions

- So far the vector of missing observation $\mathbf{Y}^u$ is unrestricted. However, in practice, inter-temporal constraints on $\mathbf{Y}^u$ are often imposed to map the missing values of the observed values of the low-frequency variables.
- For example, common inter-temporal constraint employed for log-differenced variables is the Mariano and Murasawa (2003, 2010)

$$\tilde{y}^u_{i,t} = \frac{1}{3}y^u_{i,t} + \frac{2}{3}y^u_{i,t-1} + y^u_{i,t-2} + \frac{2}{3}y^u_{i,t-3} + \frac{1}{3}y^u_{i,t-4}, \qquad (5)$$

- where $y^u_{i,t}$ is the missing monthly value of the i-th variable at month t, and $\tilde{y}^u_{i,t}$ is the corresponding observed quarterly value. We can stack (5) over time

$$\tilde{\mathbf{Y}}^u = \mathbf{M}_a\mathbf{Y}^u,$$

- where $\mathbf{M}_a$ is a matrix containing all the inter-temporal constraints over time.

Inter-temporal restrictions (cont.)

- How do we sample from $\mathbf{Y}^u$ subject to the inter-temporal (linear) constraint?
  1. We first draw $\mathbf{Z} \sim N(\mu_{\mathbf{Y}^u}, \mathbf{K}_{\mathbf{Y}^u}^{-1})$ from the unconstrained distribution.
  2. We then correct for the constraint by computing $\mathbf{Y}^u = \mathbf{Z} + \mathbf{K}_{\mu}^{-1}\mathbf{M}_a^{'}(\mathbf{M}_a\mathbf{K}_{\mu}^{-1}\mathbf{M}_a^{'})^{-1}(\tilde{\mathbf{Y}}^u - \mathbf{M}_a\mathbf{Z})$. (See Rue and Held (2005) and Cong, Chen, and Zhou (2017))
- It can be shown that $\mathbf{Y}^u$ has the correct distribution, i.e., it follows the $N(\mu_{\mathbf{Y}^u}, \mathbf{K}_{\mathbf{Y}^u}^{-1})$ distribution satisfying the constraint $\tilde{\mathbf{Y}}^u = \mathbf{M}_a\mathbf{Y}^u$.

Simulation study

- We conduct simulation to assess the accuracy of our proposed precision-based sampler.
- All DGPs follows a VAR structure with $p = 5$ lags

$$\mathbf{y}_t = \mathbf{B}_0 + \mathbf{B}_1\mathbf{y}_{t-1} + \ldots + \mathbf{B}_4\mathbf{y}_{t-4} + \mathbf{B}_5\mathbf{y}_{t-5} + \epsilon_t, \epsilon_t \sim N(0, \Omega),$$

- $\mathbf{y}_t = (y_t^o, y_t^u)$ is a $n \times 1$ vector that contains the mixed-frequency data generated
- y_t^o is as an $n_o \times 1$ vector of the high-frequency variables (monthly).
- y_t^u is as an $n_u \times 1$ vector of the low-frequency variables (quarterly).
- We assume $\mathbf{B}_0 = 0.01 \times 1_n$, $\mathbf{B}_1 = 0.5 \times \mathbf{I}_n$, $\mathbf{B}_2 = 0.05 \times \mathbf{I}_n$, $\mathbf{B}_3 = 0.001 \times \mathbf{I}_n$, $\mathbf{B}_4 = 0.0001 \times \mathbf{I}_n$, $\mathbf{B}_5 = 0.00005 \times \mathbf{I}_n$, $\Omega = 0.01 \times \mathbf{I}_n$ and $T = 300$.
- We also assume the standard inter-temporal constraint of (5).

Simulation study (cont.)

- We consider 6 different types of DGPs that vary across the no. of monthly and quarterly variables in the model.
- We estimate the MF-State-Space VAR using four different approaches
  1. Precision-based sampler (Hard)
  2. Precision-based sampler (Soft) - A small error is assumed in the inter-temporal constraint of (5).
  3. Carter and Kohn simulation smoother - Schorfheide and Song (2015) JBES code
  4. Durbin and Koopman simulation smoother
- For each DGP, we estimate the model with $R = 10$ parallel chains.

Simulation study (cont.)

Table: Mean squared errors of the estimated missing observations and computation times using four methods: the proposed precision-based method with hard inter-temporal constraints (Precision-hard), the precision-based method with soft inter-temporal constraints (Precision-soft), the simulation smoother of Carter and Kohn (1994) implemented in Schorfheide and Song (2015) (KF), and the Durbin and Koopman (DK) simulation smoother.

		MSE				Computation time (minutes)			
n_u	n_o	Precision-hard	Precision-soft	DK	CK	Precision-hard	Precision-soft	DK	CK
1	5	0.004	0.004	0.005	0.005	0.7	0.6	7	5
1	10	0.005	0.005	0.005	0.006	3	3	31	9
1	15	0.005	0.005	0.005	0.005	13	13	61	16
5	5	0.005	0.005	0.005	0.006	8	4	23	24
5	10	0.005	0.005	0.005	0.006	16	12	51	35
5	15	0.005	0.005	0.005	0.006	38	35	106	51

Simulation study (cont.)

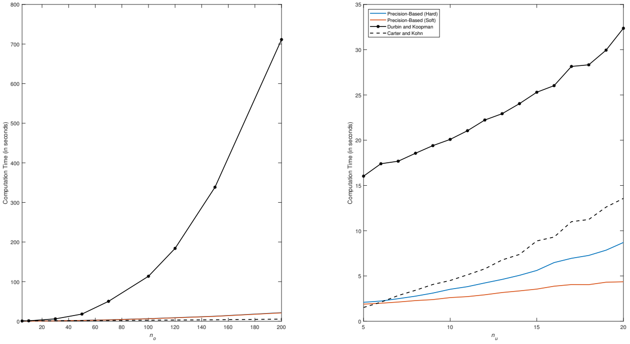


Figure: Computation times of obtaining 10 draws againist n_o and n_u, the numbers of observed and partially unobserved variables, respectively, with $T = 300$ and $p = 5$. The four methods are: precision-based sampler with hard inter-temporal constraints (Precision-hard), precision-based sampler with soft constraints (Precision-soft), the simulation smoother of Carter and Kohn (1994) implemented in Schorfheide and Song (2015) (KF), and the Durbin and Koopman (DK) simulation smoother.

Simulation study (cont.)

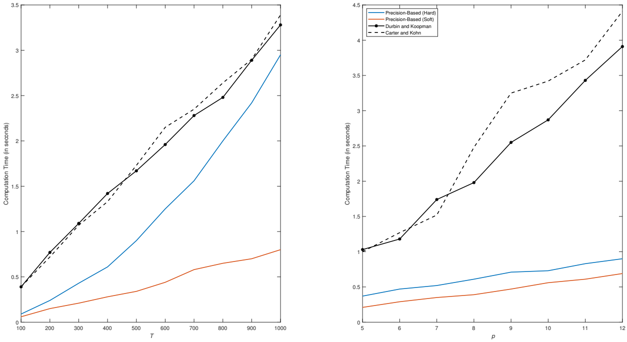
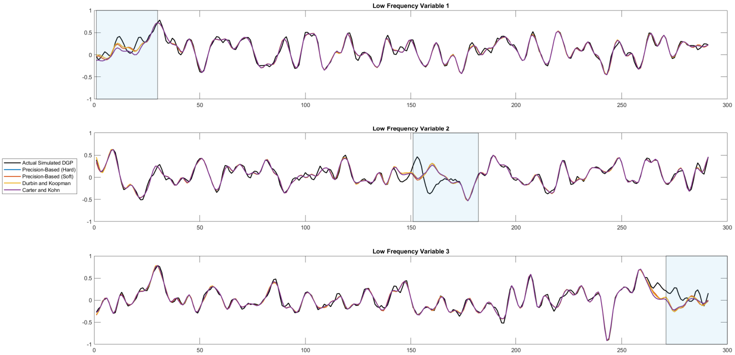


Figure: Computation times of obtaining 10 draws against T and p, the numbers of time periods and lags, respectively, with $n_u = 5$ and $n_o = 10$. The four methods are: precision-based sampler with hard inter-temporal constraints (Precision-hard), precision-based sampler with soft constraints (Precision-soft) and the simulation smoother of Carter and Kohn (1994) implemented in Schorfheide and Song (2015) (KF) and the Durbin and Koopman (DK) simulation smoother.

Simulation study - Unbalanced Missing Data Panel (cont.)

Figure: Posterior Mean of the Low-Frequency Variables against the Actual Simulated DGP



Notes: All estimates displayed in the graphs are transformed via the inter-temporal constraint of (5). The black line is the actual simulated data. The blue and orange line are the posterior estimates from the precision-based method of the hard and soft conditions, respectively. The yellow and purple line are the posterior estimates from the Durbin and Koopman, and the Carter and Kohn simulation smoother, respectively. The light shaded area denotes the time periods that contains the large missing data pattern for each low-frequency variables.

Simulation study - Unbalanced Missing Data Panel (cont.)

Table: Mean squared errors of the estimated missing observations and computation times using four methods: the proposed precision-based method with hard inter-temporal constraints (Precision-hard), the precision-based method with soft inter-temporal constraints (Precision-soft) and the simulation smoother of Durbin and Koopman (DK), and Carter and Kohn (CK). (For $R = 10$ parallel chains).

		MSE				Computation time (minutes)			
n_u	n_o	Precision-hard	Precision-soft	DK	CK	Precision-hard	Precision-soft	DK	CK
3	10	0.006	0.006	0.006	0.006	8	7	40	14

Empirical Application

- We applied our proposed mixed-frequency precision-based sampler on two popular empirical macroeconomic applications.
  1. We show that our proposed precision-based sampler can be incorporated efficiently within a state-of-the-art large BVAR-SV with global-local priors.
  2. We extend the methodology in Caldara and Herbst (2019) by proposing a novel mixed-frequency Bayesian Proxy VAR (MF-BP-VAR).

Empirical Application - MF-BP-VAR

- Caldara and Herbst (2019) considers a monthly frequency five-equation BP-VAR which consists of:
  - Fed funds rate, log of industrial production, unemployment rate, log of producer price index and measure of corporate spread.
  - Sample: 1994M1-2007M6
- We extend their model to include log of real GDP observed at the quarterly frequency, 1994Q1-2007Q2. Thus, we impose an inter-temporal constraint of

$$y_t^{GDP,Q} = \frac{1}{3}(y_t^{GDP,m} + y_{t-1}^{GDP,m} + y_{t-2}^{GDP,m}),$$

- We preserved all the model assumptions as specified in Caldara and Herbst (2019).

Empirical Application - MF-BP-VAR

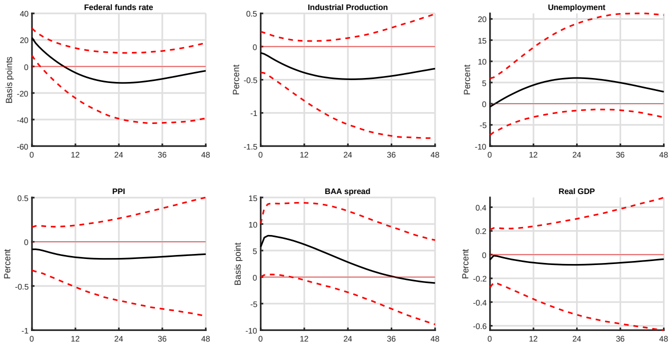


Figure: Impulse Responses to a Monetary Policy Shock

Conclusion

- We have developed a new precision-based sampler for drawing the missing observations of the low-frequency variables within a MF-State-Space VAR.
- In the simulation study, we show that our proposed precision-based sampler is slightly more accurate and computationally efficient on average compared to standard filtering and smoothing techniques.
- We show how our proposed precision-based sampler can be applied to two popular empirical macroeconomic applications.
- Extension: We are currently extending our precision-based sampler to conditional forecasting for large VARs.