

B State space representation of the model

Presented here are the details for the state-space representation of equation (11) as specified by the equations (5)-(10), for $p = 1$, $r = 3$, and a single quarterly variable (y_t^Q is of dimension 1 as in the benchmark model):

$$\underbrace{\begin{pmatrix} x_t \\ y_t^Q \end{pmatrix}}_{\bar{x}_t} = \underbrace{\begin{pmatrix} \mu \\ \tilde{\mu}_Q \end{pmatrix}}_{\bar{\mu}} + \underbrace{\begin{pmatrix} \Lambda & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 \\ \Lambda_Q & 2\Lambda_Q & 3\Lambda_Q & 2\Lambda_Q & \Lambda_Q & 0 & 1 & 2 & 3 & 2 & 1 \end{pmatrix}}_{Z(\theta)} \underbrace{\begin{pmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \varepsilon_t \\ \varepsilon_t^Q \\ \varepsilon_{t-1}^Q \\ \varepsilon_{t-2}^Q \\ \varepsilon_{t-3}^Q \\ \varepsilon_{t-4}^Q \end{pmatrix}}_{\xi_t} \quad (12)$$

$$\begin{pmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \varepsilon_t \\ \varepsilon_t^Q \\ \varepsilon_{t-1}^Q \\ \varepsilon_{t-2}^Q \\ \varepsilon_{t-3}^Q \\ \varepsilon_{t-4}^Q \end{pmatrix} = \underbrace{\begin{pmatrix} A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{diag}(\alpha_1, \dots, \alpha_n) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha_Q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}}_{T(\theta)} \begin{pmatrix} f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ f_{t-5} \\ \varepsilon_{t-1} \\ \varepsilon_{t-1}^Q \\ \varepsilon_{t-2}^Q \\ \varepsilon_{t-3}^Q \\ \varepsilon_{t-4}^Q \\ \varepsilon_{t-5}^Q \end{pmatrix} + \underbrace{\begin{pmatrix} u_t \\ 0 \\ 0 \\ 0 \\ 0 \\ e_t \\ e_t^Q \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\eta_t}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{n,t})'$ and $e_t = (e_{1,t}, e_{2,t}, \dots, e_{n,t})'$.

The block-specific factor structure further implies that

$$\Lambda = \begin{pmatrix} \Lambda_{N,G} & \Lambda_{N,N} & 0 \\ \Lambda_{R,G} & 0 & \Lambda_{R,R} \end{pmatrix}, \quad \Lambda_Q = (\Lambda_{Q,G} \quad 0 \quad \Lambda_{Q,R}),$$

$$f_t = \begin{pmatrix} f_t^G \\ f_t^N \\ f_t^R \end{pmatrix}, \quad A_1 = \begin{pmatrix} A_{1,G} & 0 & 0 \\ 0 & A_{1,N} & 0 \\ 0 & 0 & A_{1,R} \end{pmatrix}, \quad Q = \begin{pmatrix} Q_G & 0 & 0 \\ 0 & Q_N & 0 \\ 0 & 0 & Q_R \end{pmatrix}.$$

Hence, the parameters of the model are:

$$\theta = (\mu, \tilde{\mu}_Q, \text{vec}(\Lambda_{N,G})', \text{vec}(\Lambda_{N,N})', \text{vec}(\Lambda_{R,G})', \text{vec}(\Lambda_{R,R})', \Lambda_{Q,G}, \Lambda_{Q,R}, A_{1,G}, A_{1,N}, A_{1,R}, \\ Q_G, Q_N, Q_R, \alpha_1, \dots, \alpha_n, \alpha_Q, \sigma_1, \dots, \sigma_n, \sigma_Q)'$$

The state space representation can be easily modified to include an arbitrary number of quarterly variables n_Q (e.g., the model with disaggregated data contains six quarterly variables). In that case, y_t^Q , $\tilde{\mu}_Q$, ε_t^Q and e_t^Q will be vectors of length n_Q . Λ_Q will be a

matrix of size $n_Q \times r$ and α_Q will be a $n_Q \times n_Q$ diagonal matrix. Finally, the scalars in the lines of $Z(\theta)$ and $T(\theta)$ corresponding to y_t^Q and ε_t^Q need to be replaced by $n_Q \times n_Q$ identity or zero matrices.

C EM algorithm

The parameters θ of the state-space form of equation (12) are estimated by the expectation maximization (EM) algorithm. The algorithm is a popular solution to problems for which latent or missing data yield a direct maximization of the likelihood function intractable or computationally difficult.¹³ The basic principle behind the EM is to write the likelihood in terms of observable as well as latent variables (in our case, in terms of \bar{x}_t and $\bar{\xi}_t$, $t = 1, \dots, T_v = \max_i T_{i,v}$) and, given the available data Ω_v ,¹⁴ obtain the maximum likelihood estimates in a sequence of two alternating steps. Precisely, iteration $j+1$ would consist of the following steps:

- E-step - the expectation of the log-likelihood conditional on the data is calculated using the estimates from the previous iteration, $\theta(j)$,
- M-step - the new parameters, $\theta(j+1)$, are estimated through the maximization of the expected log-likelihood (from the previous iteration) with respect to θ .

Below we provide the details of the implementation of the EM algorithm for the state-space representation of equation (12) (based on the results in Bańbura and Modugno, 2010).

We first estimate μ and $\tilde{\mu}_Q$ by sample means and use the de-measured data throughout the EM steps.

To deal with missing observations in \bar{x}_t we follow Bańbura and Modugno (2010) and introduce selection matrices W_t and W_t^Q . They are diagonal matrices of size n and 1, respectively, with 1s corresponding to the nonmissing values in x_t and y_t^Q , respectively.

For the sake of simplicity, we first consider the case without restrictions on Λ , Λ_Q , A_1 and Q implied by block-specific factors.

The maximization of the expected likelihood (M-step) with respect to θ in the $(r+1)$ -iteration would yield the following expressions:

- The matrix of loadings for the monthly variables:

$$\text{vec}(\Lambda(j+1)) = \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_t f_t' | \Omega_v] \otimes W_t \right)^{-1} \text{vec} \left(\sum_{t=1}^T W_t x_t \mathbb{E}_{\theta(j)} [f_t' | \Omega_v] - W_t \mathbb{E}_{\theta(j)} [\varepsilon_t f_t' | \Omega_v] \right). \quad (13)$$

¹³See Dempster, Laird, and Rubin (1977) for a general EM algorithm and Shumway and Stoffer (1982) or Watson and Engle (1983) for application to state-space representations.

¹⁴ $\Omega_v \subseteq \{y_1, \dots, y_{T_v}\}$ because some observations in y_t are missing.

- The matrix of loadings for the quarterly variables:

Let $f_t^{(p)} = [f_t', \dots, f_{t-p+1}']'$ and $D = \sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_t^{(5)} f_t^{(5)' | \Omega_v}] W_t^Q$ and $\tilde{\epsilon}_t^Q = \epsilon_t^Q + 2\epsilon_{t-1}^Q + 3\epsilon_{t-2}^Q + 2\epsilon_{t-3}^Q + \epsilon_{t-4}^Q$. The unrestricted row vector of factor loadings for y_t^Q is given by

$$(\bar{\Lambda}_Q^{ur}(j+1))' = D^{-1} \left(\sum_{t=1}^T W_t^Q y_t^Q \mathbb{E}_{\theta(j)} [f_t^{(5)' | \Omega_v}] - W_t^Q \mathbb{E}_{\theta(j)} [\tilde{\epsilon}_t^Q f_t^{(5)' | \Omega_v}] \right)'.$$

For the restricted $\bar{\Lambda}_Q = (\Lambda_Q \ 2\Lambda_Q \ 3\Lambda_Q \ 2\Lambda_Q \ \Lambda_Q)$ it holds that $C\bar{\Lambda}_Q' = 0$ with

$$C = \begin{bmatrix} I_r & -\frac{1}{2}I_r & 0 & 0 & 0 \\ I_r & 0 & -\frac{1}{3}I_r & 0 & 0 \\ I_r & 0 & 0 & -\frac{1}{2}I_r & 0 \\ I_r & 0 & 0 & 0 & -I_r \end{bmatrix}.$$

Consequently the restricted $\bar{\Lambda}_Q$ is given by:

$$(\bar{\Lambda}_Q(j+1))' = (\bar{\Lambda}_Q^{ur}(j+1))' - D^{-1}C'(CD^{-1}C')^{-1}C(\bar{\Lambda}_Q^{ur}(j+1))'.$$

- The autoregressive coefficients in the factor VAR:

$$A_1(j+1) = \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_t f_{t-1}' | \Omega_v] \right) \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_{t-1} f_{t-1}' | \Omega_v] \right)^{-1}. \quad (14)$$

- The covariance matrix in the factor VAR:

$$Q(j+1) = \frac{1}{T} \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_t f_t' | \Omega_v] - A_1(j+1) \sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_{t-1} f_t' | \Omega_v] \right). \quad (15)$$

- The autoregressive coefficients in the AR representation for the idiosyncratic component of the monthly variables:

$$\alpha_i(j+1) = \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [\varepsilon_{i,t} \varepsilon_{i,t-1} | \Omega_v] \right) \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [(\varepsilon_{i,t-1})^2 | \Omega_v] \right)^{-1} \quad i = 1, \dots, n, Q.$$

- The variance in the AR representation for the idiosyncratic component of the monthly variables:

$$\sigma_i^2(j+1) = \frac{1}{T} \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [(\varepsilon_{i,t})^2 | \Omega_v] - \alpha_i(j+1) \sum_{t=1}^T \mathbb{E}_{\theta(j)} [\varepsilon_{i,t-1} \varepsilon_{i,t} | \Omega_v] \right) \quad i = 1, \dots, n, Q.$$

The conditional expectations (the E-step) in the expressions above are computed using the Kalman smoother on the state-space representation of equation (12) with the previous iteration parameters $\theta(j)$. The initial parameters $\theta(0)$ are obtained on the basis of principal components analysis (in the spirit of the two-step method of Doz, Giannone, and Reichlin, 2006b).

To account for the restrictions imposed by group-specific factors, we would split the parameters in Λ into $\Lambda_N = (\Lambda_{N,G} \ \Lambda_{N,N})$ and $\Lambda_R = (\Lambda_{R,G} \ \Lambda_{R,R})$ and obtain the $j+1$ -iteration

of Λ_N by modifying the formula in equation (13) as

$$\text{vec}(\Lambda_N(j+1)) = \left(\sum_{t=1}^T \mathbb{E}_{\theta(j)} [f_t^{G,N} f_t^{G,N'} | \Omega_v] \otimes W_t^N \right)^{-1} \text{vec} \left(\sum_{t=1}^T W_t^N x_t^N \mathbb{E}_{\theta(j)} [f_t^{G,N'} | \Omega_v] - W_t^N \mathbb{E}_{\theta(j)} [\varepsilon_t^N f_t^{G,N'} | \Omega_v] \right),$$

where $f_t^{G,N} = (f_t^{G'} f_t^{N'})'$, x_t^N and ε_t^N are the subvectors of x_t and ε_t containing only nominal variables and idiosyncratic components, respectively. W_t^N can be obtained from W_t by discarding all the rows and columns corresponding to the real data. The updating formulas for Λ_R can be obtained in an analogous fashion. To obtain restricted versions of A_1 and Q we can use equations (14) and (15) for each of the factors f_t^G , f_t^N , f_t^R , separately.