Online Appendix to "Estimating Shadow Policy Rates in a Small Open Economy and the Role of Foreign Factors"

Jorge Fornero Markus Kirchner Carlos Molina*

Central Bank of Chile†

May 2021

This appendix provides a detailed description of the estimation of the dynamic factor model through the Expectations Maximization (EM) algorithm (Appendix A), the data sources and variable definitions used (Appendix B), and the selection of the number of factors based on the ABC criterion (Appendix C). Furthermore, it presents the details of the recursive estimation during the global financial crisis episode (Appendix D) and the estimation of the Taylor rules for Chile (Appendix E).

*Authors' e-mail addresses: jfornero@bcentral.cl; mkirchner@bcentral.cl; cmolina@bcentral.cl.

†The views and conclusions presented in this paper are exclusively those of the authors and do not necessarily reflect the position of the Central Bank of Chile or its Board members.

A EM Algorithm for Dynamic Factor Model with Missing Observations and Restrictions

This appendix provides the details of the expectation-maximization (EM) algorithm used to estimate the model. The description follows Shumway and Stoffer (1982) and Bańbura and Modugno (2014).

Let $y_t = [y_{1,t}, y_{2,t}, \dots, y_{n,t}]'$, $t = 1, \dots, T$, denote a stationary vector process of n observed variables standardized to mean 0 and unit variance. We will derive the EM algorithm for the case when only a subset of $Y = \{y_1, \dots, y_T\}$, denoted by $\Omega_T \subseteq Y$ is observed. We assume that y_t admits a dynamic factor model (DFM) representation in terms of r normally distributed unobserved common factors collected in the vector f_t and normally distributed idiosyncratic components collected in the vector $e_t = [e_{1,t}, e_{2,t}, \dots, e_{n,t}]'$. The measurement equation of the DFM is given by

$$y_t = \Lambda f_t + e_t, \qquad e_{i,t} \sim N(0, \sigma_i^2), \tag{1}$$

where Λ is an $n \times r$ matrix of factor loadings. The dynamics of the common factors are assumed to follow a stationary vector auto-regressive (VAR) process with 1 lag:

$$f_t = A f_{t-1} + u_t, \qquad u_t \sim N(0, Q), \tag{2}$$

where A is an $r \times r$ matrix of auto-regressive coefficients and Q denotes the $r \times r$ variance-covariance matrix of the shocks to the common factors. The common and idiosyncratic shocks are assumed to be uncorrelated at all leads and lags. The common shocks may be cross-sectionally correlated, but the idiosyncratic shocks are assumed to be cross-sectionally uncorrelated. In addition, the idiosyncratic shocks are assumed to be serially uncorrelated. We further assume that $f_0 \sim N(\mu, \Sigma)$.

The EM algorithm (Dempster et al., 1977) is a method for maximizing a likelihood function for problems with incomplete or latent data. This algorithm was adapted to estimate general dynamic linear models with unobserved components by Shumway and Stoffer (1982) and small DFMs by Watson and Engle (1983). It was extended to large DFMs by Doz et al. (2011) based on earlier work by Giannone et al. (2005) and Giannone et al. (2008). Bańbura and Modugno (2014) modified the algorithm to estimate the parameters of a DFM with an arbitrary pattern of missing data, e.g., due to mixed data frequencies.[1] They also show how to adapt the algorithm for the case of a serially correlated idiosyncratic component, and for cases with restrictions on the parameters of the model, as in Bork (2009) and Bork et al. (2009).

In our application, we will be interested in restrictions on Λ and A of the form

$$\Lambda = \begin{bmatrix} \Lambda_{11} & 0 \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \tag{3}$$

where Λ_{11} (A_{11}) is an $n_1 \times n_1$ ($r_1 \times r_1$) matrix, Λ_{21} (A_{21}) is an $n_2 \times n_1$ ($r_2 \times r_1$) matrix and Λ_{22} (A_{22}) is an $n_2 \times n_2$ ($r_2 \times r_2$) matrix, with $n_1 + n_2 = n$ ($r_1 + r_2 = r$). The vectors of observed variables and unobserved factors are partitioned accordingly:

$$y_t = \begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix}, \qquad f_t = \begin{bmatrix} f_t^1 \\ f_t^2 \end{bmatrix},$$

where y_t^1 and y_t^2 (f_t^1 and f_t^2) are $n_1 \times 1$ and $n_2 \times 1$ ($r_1 \times 1$ and $r_2 \times 1$) vectors, respectively.

The algorithm consists of two steps that are iterated to convergence: an expectation and a maximization step. In the expectation step (E-step), the missing observations are estimated conditional on the observed data and guess values of the parameters. In the maximization step (M-step), the maximum likelihood estimates of the parameters are calculated by maximizing the conditional expectation of the likelihood derived in the first step.

A.1 Expectation Step

Let $l_{Y,F}(\theta)$ denote the joint log-likelihood of the observed data and the latent factors, where $F = \{f_0, f_1, \dots, f_T\}$, and where θ collects the parameters of the model. In the E-step, the expectation of the log-likelihood conditional on

[1] An alternative method to estimate high-dimensional DFMs, based on a time-varying state-space representation, has been proposed by Jungbacker et al. (2011).

the observed data is calculated using the estimates from the previous iteration, $\theta(j)$:

$$L(\theta, \theta(j)) = E_{\theta(j)} \left[l_{Y,F}(\theta) | Y \right].$$

To describe this step, we cast the model in state-space form:

$$\begin{aligned} y_t &= \Lambda f_t + e_t, \qquad e_t \sim N(0, R), &(4)\\ f_t &= A f_{t-1} + u_t, \qquad u_t \sim N(0, Q), &(5) \end{aligned}$$

with

$$R = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix}$$

and $f_0 \sim N(\mu, \Sigma)$. Then, the parameters of the model are $\theta = \{\mu, \Sigma, \Lambda, R, A, Q\}$. Based on (4) and (5), Kalman filtering and smoothing techniques can be used to calculate $l_{Y,F}(\theta)$ and to obtain the required moments of the latent factors. In particular, we will need $f_{t|T} = E_{\theta(j)} \left(f_t | \Omega_T \right)$, $P_{t|T} = E_{\theta(j)} \left[\left(f_t - f_{t|T} \right) \left(f_t - f_{t|T} \right)' | \Omega_T \right]$ and $P_{t,t-1|T} = E_{\theta(j)} \left[\left(f_t - f_{t|T} \right) \left(f_{t-1} - f_{t-1|T} \right)' | \Omega_T \right]$, where $E_{\theta(j)} \left(\cdot | \Omega_T \right)$ denotes the expectation operator conditional on the available data Ω_T and the parameters $\theta(j)$.

Under the Gaussian assumption, the prediction error decomposition of the joint density function yields the following sequence of conditional density functions representing the joint log-likelihood:

$$\begin{aligned} l_{Y,F}(\theta) &= \text{constant} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \left(f_0 - \mu \right)' \Sigma^{-1} \left(f_0 - \mu \right) - \frac{T}{2} \log |Q| \\ &\quad - \frac{1}{2} \sum_{t=1}^{T} \left(f_t - A f_{t-1} \right)' Q^{-1} \left(f_t - A f_{t-1} \right) - \frac{T}{2} \log |R| - \frac{1}{2} \sum_{t=1}^{T} \left(y_t - \Lambda f_t \right)' R^{-1} \left(y_t - \Lambda f_t \right) \\ &= \text{constant} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr} \left[\Sigma^{-1} \left(f_0 - \mu \right) \left(f_0 - \mu \right)' \right] - \frac{T}{2} \log |Q| \\ &\quad - \frac{1}{2} \text{tr} \left[Q^{-1} \sum_{t=1}^{T} \left(f_t - A f_{t-1} \right) \left(f_t - A f_{t-1} \right)' \right] - \frac{T}{2} \log |R| - \frac{1}{2} \text{tr} \left[R^{-1} \sum_{t=1}^{T} \left(y_t - \Lambda f_t \right) \left(y_t - \Lambda f_t \right)' \right], \end{aligned}$$

where the second equality uses the cyclic property of the trace operator applied to the respective quadratic forms (i.e., tr$(x'Zx)$ =tr(Zxx')). Taking expectations conditional on the observed data Ω_T and the parameters $\theta(j)$, the previous expression can be re-written as follows:

$$\begin{aligned} L(\theta, \theta(j)) &= \text{constant} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr} \left[\Sigma^{-1} \left(E_{\theta(j)} \left(f_0 f_0' | \Omega_T \right) - \mu f_{0|T}' - f_{0|T} \mu' + \mu \mu' \right) \right] \\ &\quad - \frac{T}{2} \log |Q| - \frac{1}{2} \text{tr} \left[Q^{-1} \sum_{t=1}^{T} \left(\begin{array}{c} E_{\theta(j)} \left(f_t f_t' | \Omega_T \right) - E_{\theta(j)} \left(f_t f_{t-1}' | \Omega_T \right) A' \\ - A E_{\theta(j)} \left(f_{t-1} f_t' | \Omega_T \right) + A E_{\theta(j)} \left(f_{t-1} f_{t-1}' | \Omega_T \right) A' \end{array} \right) \right] \\ &\quad - \frac{T}{2} \log |R| - \frac{1}{2} \text{tr} \left[R^{-1} \sum_{t=1}^{T} E_{\theta(j)} \left[\left(y_t - \Lambda f_t \right) \left(y_t - \Lambda f_t \right)' | \Omega_T \right] \right]. \end{aligned}$$

Using the decomposition $f_t = f_{t|T} + \left(f_t - f_{t|T} \right)$ for $t = 0, \ldots, T$, as well as the fact that $E_{\theta(j)} \left(f_t - f_{t|T} | \Omega_T \right) = 0$ yields

$$\begin{aligned} L(\theta, \theta(j)) &= \text{constant} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr} \left\{ \Sigma^{-1} \left[\left(f_{0|T} - \mu \right) \left(f_{0|T} - \mu \right)' + P_{0|T} \right] \right\} \\ &\quad - \frac{T}{2} \log |Q| - \frac{1}{2} \text{tr} \left[Q^{-1} \left(M_1 - M_2 A' - A M_2' + A M_3 A' \right) \right] \\ &\quad - \frac{T}{2} \log |R| - \frac{1}{2} \text{tr} \left\{ R^{-1} \sum_{t=1}^{T} E_{\theta(j)} \left[\left(y_t - \Lambda f_t \right) \left(y_t - \Lambda f_t \right)' | \Omega_T \right] \right\}, \qquad (6) \end{aligned}$$

where

$$M_1 = \sum_{t=1}^{T} \left(f_{t|T} f'_{t|T} + P_{t|T} \right), \tag{7}$$

$$M_2 = \sum_{t=1}^{T} \left(f_{t|T} f'_{t-1|T} + P_{t,t-1|T} \right), \tag{8}$$

$$M_3 = \sum_{t=1}^{T} \left(f_{t-1|T} f'_{t-1|T} + P_{t-1|T} \right). \tag{9}$$

Now consider the expectation $E_{\theta(j)}\left[(y_t - \Lambda f_t)(y_t - \Lambda f_t)' | \Omega_T \right]$. To characterize this expectation, following Bańbura and Modugno (2014), let

$$y_t = W_t y_t + (I_n - W_t) y_t = y_t^{(1)} + y_t^{(2)},$$

where W_t is an $n \times n$ diagonal matrix with ones corresponding to the non-missing entries in y_t and 0 otherwise, and where the $n \times 1$ vector $y_t^{(1)}$ ($y_t^{(2)}$) contains the non-missing (missing) observations at time t with 0 in place of missing (non-missing) observations. We have

$$\begin{aligned} (y_t - \Lambda f_t)(y_t - \Lambda f_t)' &= \left[W_t (y_t - \Lambda f_t) + (I_n - W_t)(y_t - \Lambda f_t) \right] \left[W_t (y_t - \Lambda f_t) + (I_n - W_t)(y_t - \Lambda f_t) \right]' \\ &= W_t (y_t - \Lambda f_t)(y_t - \Lambda f_t)' W_t + (I_n - W_t)(y_t - \Lambda f_t)(y_t - \Lambda f_t)' (I_n - W_t) \\ &\quad + W_t (y_t - \Lambda f_t)(y_t - \Lambda f_t)' (I_n - W_t) + (I_n - W_t)(y_t - \Lambda f_t)(y_t - \Lambda f_t)' W_t. \end{aligned} \tag{10}$$

By the law of iterated expectations (i.e., $E(X|Y) = E\left[E(X|Z, Y)|Y \right]$)),

$$E_{\theta(j)}\left[(y_t - \Lambda f_t)(y_t - \Lambda f_t)' | \Omega_T \right] = E_{\theta(j)}\left[E_{\theta(j)}\left[(y_t - \Lambda f_t)(y_t - \Lambda f_t)' | F, \Omega_T \right] | \Omega_T \right].$$

Then, defining $e_t^{(2)} = (I_n - W_t)(y_t - \Lambda f_t)$, we have

$$E_{\theta(j)}\left[W_t (y_t - \Lambda f_t)(y_t - \Lambda f_t)' (I_n - W_t) | F, \Omega_T \right] = \left(y_t^{(1)} - W_t \Lambda f_t \right) E_{\theta(j)}\left(e_t^{(2)\prime} | F, \Omega_T \right) = 0, \tag{11}$$

$$E_{\theta(j)}\left[(I_n - W_t)(y_t - \Lambda f_t)(y_t - \Lambda f_t)' W_t | F, \Omega_T \right] = E_{\theta(j)}\left(e_t^{(2)} | F, \Omega_T \right) \left(y_t^{(1)} - f'_t \Lambda' W_t \right) = 0, \tag{12}$$

and

$$E_{\theta(j)}\left[(I_n - W_t)(y_t - \Lambda f_t)(y_t - \Lambda f_t)' (I_n - W_t) | F, \Omega_T \right] = (I_n - W_t) R(j) (I_n - W_t). \tag{13}$$

In addition, we have

$$\begin{aligned} E_{\theta(j)}\left[W_t (y_t - \Lambda f_t)(y_t - \Lambda f_t)' W_t | \Omega_T \right] &= y_t^{(1)} y_t^{(1)\prime} - W_t \Lambda E_{\theta(j)}(f_t | \Omega_T) y_t^{(1)\prime} - y_t^{(1)} E_{\theta(j)}(f'_t | \Omega_T) \Lambda' W_t \\ &\quad + W_t \Lambda E_{\theta(j)}(f_t f'_t | \Omega_T) \Lambda' W_t. \end{aligned} \tag{14}$$

Taking the conditional expectation of (10) and using (11) through (14) yields

$$\begin{aligned} E_{\theta(j)}\left[(y_t - \Lambda f_t)(y_t - \Lambda f_t)' | \Omega_T \right] &= y_t^{(1)} y_t^{(1)\prime} - W_t \Lambda E_{\theta(j)}(f_t | \Omega_T) y_t^{(1)\prime} - y_t^{(1)} E_{\theta(j)}(f'_t | \Omega_T) \Lambda' W_t \\ &\quad + W_t \Lambda E_{\theta(j)}(f_t f'_t | \Omega_T) \Lambda' W_t + (I_n - W_t) R(j) (I_n - W_t) \\ &= y_t^{(1)} y_t^{(1)\prime} - W_t \Lambda f_{t|T} y_t^{(1)\prime} - y_t^{(1)} f'_{t|T} \Lambda' W_t + W_t \Lambda \left(f_{t|T} f'_{t|T} + P_{t|T} \right) \Lambda' W_t \\ &\quad + (I_n - W_t) R(j) (I_n - W_t). \end{aligned} \tag{15}$$

Inserting (15) into (6) yields

$$\begin{aligned} L(\theta, \theta(j)) &= \text{constant} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr} \left\{ \Sigma^{-1} \left[\left(f_{0|T} - \mu \right) \left(f_{0|T} - \mu \right)' + P_{0|T} \right] \right\} \\ &\quad - \frac{T}{2} \log |Q| - \frac{1}{2} \text{tr} \left[Q^{-1} \left(M_1 - M_2 A' - A M'_2 + A M_3 A' \right) \right] \\ &\quad - \frac{T}{2} \log |R| - \frac{1}{2} \text{tr} \left\{ R^{-1} \sum_{t=1}^{T} \left[\begin{array}{c} y_t^{(1)} y_t^{(1)\prime} - W_t \Lambda f_{t|T} y_t^{(1)\prime} - y_t^{(1)} f'_{t|T} \Lambda' W_t \\ + W_t \Lambda \left(f_{t|T} f'_{t|T} + P_{t|T} \right) \Lambda' W_t \\ + (I_n - W_t) R(j) (I_n - W_t) \end{array} \right] \right\}. \end{aligned}$$

We further have (which will be useful for maximization with respect to the parameters Λ and A):

$$\begin{aligned}
L(\theta,\theta(j)) &= \text{constant} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\text{tr}\left\{\Sigma^{-1}\left[\left(f_{0|T}-\mu\right)\left(f_{0|T}-\mu\right)' + P_{0|T}\right]\right\} \\
&\quad -\frac{T}{2}\log|Q| - \frac{1}{2}\left[\text{tr}\left(Q^{-1}M_1\right) - \text{tr}\left(A'Q^{-1}M_2\right) - \text{tr}\left(M_2'Q^{-1}A\right) + \text{tr}\left(Q^{-1}AM_3A'\right)\right] \\
&\quad -\frac{T}{2}\log|R| - \frac{1}{2}\left\{\begin{array}{c} \text{tr}\left(R^{-1}y_t^{(1)}y_t^{(1)\prime}\right) - \text{tr}\left(f_{t|T}y_t^{(1)\prime}R^{-1}\Lambda\right) - \text{tr}\left(\Lambda'R^{-1}y_t^{(1)}f_{t|T}'\right) \\ +\text{tr}\left[R^{-1}W_t\Lambda\left(f_{t|T}f_{t|T}' + P_{t|T}\right)\Lambda'\right] \\ +\text{tr}\left[R^{-1}\left(I_n - W_t\right)R(j)\left(I_n - W_t\right)\right] \end{array}\right\} \\
&= \text{constant} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\text{tr}\left\{\Sigma^{-1}\left[\left(f_{0|T}-\mu\right)\left(f_{0|T}-\mu\right)' + P_{0|T}\right]\right\} \\
&\quad -\frac{T}{2}\log|Q| - \frac{1}{2}\left[\begin{array}{c} \text{tr}\left(Q^{-1}M_1\right) - \text{vec}\left(A\right)'\left(M_2'\otimes I_r\right)\text{vec}\left(Q^{-1}\right) \\ -\text{vec}\left(M_2\right)'\left(I\otimes Q^{-1}\right)\text{vec}\left(A\right) + \text{vec}\left(A\right)'\left(M_3\otimes Q^{-1}\right)\text{vec}\left(A\right) \end{array}\right] \\
&\quad -\frac{T}{2}\log|R| - \frac{1}{2}\sum_{t=1}^{T}\left\{\begin{array}{c} \text{tr}\left(R^{-1}y_t^{(1)}y_t^{(1)\prime}\right) - \text{vec}\left(y_t^{(1)}f_{t|T}'\right)'\left(I_r\otimes R^{-1}\right)\text{vec}\left(\Lambda\right) \\ -\text{vec}\left(\Lambda\right)'\left(I_r\otimes R^{-1}\right)\text{vec}\left(y_t^{(1)}f_{t|T}'\right) \\ +\text{vec}\left(\Lambda\right)'\left[\left(f_{t|T}f_{t|T}' + P_{t|T}\right)\otimes R^{-1}W_t\right]\text{vec}\left(\Lambda\right) \\ +\text{tr}\left[R^{-1}\left(I_n - W_t\right)R(j)\left(I_n - W_t\right)\right] \end{array}\right\},
\end{aligned}$$

where the first equality uses a cyclic property of the trace operator, i.e., tr$(ABCD)$ =tr$(BCDA)$ =tr$(CDAB)$ =tr$(CABC)$ and the fact that $W_tR^{-1}W_t = R^{-1}W_t = W_tR^{-1}$, while the second equality uses the properties tr(AB) =vec$(A')'$vec(B), tr(ABC) =vec$(A')'\,(C'\otimes I)$vec(B) =vec$(A')'\,(I\otimes B)$vec(C) and tr$(ABCD)$ =vec$(D')'\,(C'\otimes A)$vec(B), as well as the fact that M_1, M_3 and $P_{t|T}$ are symmetric matrices.

The required conditional moments of the latent factors are obtained from the Kalman filtering and smoothing recursions. These recursions are easily adapted for missing observations (see Durbin and Koopman, 2012, Section 4.10), by re-writing the observation equation as follows:

$$y_t^* = \Lambda^* f_t + e_t^*, \qquad e_t^* \sim N(0, R^*),$$

where $y_t^* = H_t y_t$, $\Lambda^* = H_t\Lambda$, $e_t^* = H_t e_t$ and $R^* = H_t R H_t'$, where H_t is a matrix whose rows correspond to the rows of W_t that are also rows of I_n (i.e., the rows of H_t are a subset of the rows of I_n). Starting at $f_{0|0}$ and $P_{0|0}$, the Kalman filter gives

$$\begin{aligned}
f_{t|t-1} &= Af_{t-1|t-1}, & (16)\\
P_{t|t-1} &= AP_{t-1|t-1}A' + Q, & (17)\\
\Sigma_t &= \Lambda^* P_{t|t-1}\Lambda^{*\prime} + R^*, & (18)\\
K_t &= P_{t|t-1}\Lambda^{*\prime}\Sigma_t^{-1}, & (19)\\
\varepsilon_t &= y_t^* - \Lambda^* f_{t|t-1}, & (20)\\
f_{t|t} &= f_{t|t-1} + K_t\varepsilon_t, & (21)\\
P_{t|t} &= \left(I - K_t\Lambda^*\right)P_{t|t-1}, & (22)
\end{aligned}$$

for $t = 1, \ldots, T$. Starting at $f_{T|T}$ and $P_{T|T}$, the Kalman smoother gives

$$\begin{aligned}
J_{t-1} &= P_{t-1|t-1}A'P_{t|t-1}^{-1}, & (23)\\
f_{t-1|T} &= f_{t-t|t-1} + J_{t-1}\left(f_{t|T} - f_{t|t-1}\right), & (24)\\
P_{t-1|T} &= P_{t-1|t-1} + J_{t-1}\left(P_{t|T} - P_{t|t-1}\right)J_{t-1}', & (25)
\end{aligned}$$

for $t = T, T-1\ldots, 1$. The following backward recursions give the covariance $P_{t,t-1|T}$ for $t = T, T-1\ldots, 2$ (see Shumway and Stoffer, 1982):

$$P_{t-1,t-2|T} = P_{t-1|t-1}J_{t-2}' + J_{t-1}\left(P_{t,t-1|T} - AP_{t-1|t-1}\right)J_{t-2}', \qquad (26)$$

where

$$P_{T,T-1|T} = (I - K_T \Lambda^*) A P_{T-1|T-1}. \tag{27}$$

The Kalman filter also gives the value of the log-likelihood of the observed data:

$$l_{\Omega^T}(\theta) = \text{constant} - \frac{1}{2}\sum_{t=1}^{T} \log|\Sigma_t(\theta)| - \frac{1}{2}\sum_{t=1}^{T} \varepsilon_t(\theta)' \Sigma_t(\theta)^{-1} \varepsilon_t(\theta), \tag{28}$$

where we have emphasized the dependence of the innovations and their conditional variance on the parameters.

A.2 Maximization Step

In the M-step, the parameters are re-estimated by maximizing the expected log-likelihood over θ:

$$\theta(j) = \arg\max_{\theta} L(\theta, \theta(j)),$$

possibly subject to linear restrictions on a subset of $\eta_\theta \geq 0$ elements of θ, in the form

$$H_\theta \text{vec}(\theta) = \kappa_\theta,$$

where κ_θ is an $\eta_\theta \times 1$ vector that collects the restrictions and H_θ is an $\eta_\theta \times \text{length}(\theta)$ matrix. This maximization problem can be written in Lagrangian form:

$$\mathcal{L}(\theta, \theta(j), \lambda_\theta) = L(\theta, \theta(j)) - \lambda_\theta' (\kappa_\theta - H_\theta \text{vec}(\theta)),$$

where λ_θ is an $\eta_\theta \times 1$ vector of Lagrangian multipliers. In our application, the restrictions on Λ and A given in (3) may be written in the form $H_\Lambda \text{vec}(\Lambda) = \kappa_\Lambda$ and $H_A \text{vec}(A) = \kappa_A$, where κ_Λ is an $\eta_\Lambda \times 1$ vector, κ_A is an $\kappa_A \times 1$ vector, H_Λ is an $\eta_\Lambda \times nr$ matrix, and H_A is an $\eta_A \times r^2$ matrix. The associated Lagrangian multipliers λ_Λ and λ_A are of size $\eta_\Lambda \times 1$ and $\kappa_A \times 1$, respectively.

Consider first the derivative of $\mathcal{L}(\theta, \theta(j), \lambda_\theta)$ with respect to vec(Λ):

$$\begin{aligned}
\left.\frac{\partial L(\theta, \theta(j))}{\partial \text{vec}(\Lambda)}\right|_{\Lambda=\Lambda(j+1)} &= \frac{1}{2}\sum_{t=1}^{T} \left.\frac{\partial \left\{ \begin{array}{c} \text{vec}\left(y_t^{(1)} f_{t|T}'\right)' \left(I_r \otimes R(j)^{-1}\right) \text{vec}(\Lambda) \\ +\text{vec}(\Lambda)' \left(I_r \otimes R(j)^{-1}\right) \text{vec}\left(y_t^{(1)} f_{t|T}'\right) \\ -\text{vec}(\Lambda)' \left[\left(f_{t|T} f_{t|T}' + P_{t|T}\right) \otimes R(j)^{-1} W_t\right] \text{vec}(\Lambda) \end{array} \right\}}{\partial \text{vec}(\Lambda)}\right|_{\Lambda=\Lambda(j+1)} + \lambda_\Lambda' H_\Lambda \\
&= \sum_{t=1}^{T} \left\{ \begin{array}{c} \text{vec}\left(y_t^{(1)} f_{t|T}'\right)' \left(I_r \otimes R(j)^{-1}\right) \\ -\text{vec}(\Lambda(j+1))' \left[\left(f_{t|T} f_{t|T}' + P_{t|T}\right) \otimes W_t R(j)^{-1}\right] \end{array} \right\} + \lambda_\Lambda' H_\Lambda,
\end{aligned}$$

where we have used the identities $\partial(Ax)/\partial x = A$, $\partial(x'A)/\partial x = A'$, $\partial(x'Ax)/\partial x = 2x'A$, and $(A \otimes B)' = A' \otimes B'$. Setting the last expression to zero and using the property $(A \otimes B)(C \otimes D) = AC \otimes BD$ (if the products AC and BD exist) yields

$$\begin{aligned}
\lambda_\Lambda' H_\Lambda &= \sum_{t=1}^{T} \left\{ \text{vec}(\Lambda(j+1))' \left[\left(f_{t|T} f_{t|T}' + P_{t|T}\right) \otimes W_t R(j)^{-1}\right] - \text{vec}\left(y_t^{(1)} f_{t|T}'\right)' \left(I_r \otimes R(j)^{-1}\right) \right\} \\
&= \sum_{t=1}^{T} \left\{ \text{vec}(\Lambda(j+1))' \left[\left(f_{t|T} f_{t|T}' + P_{t|T}\right) \otimes W_t\right] - \text{vec}\left(y_t^{(1)} f_{t|T}'\right)' \right\} \left(I_r \otimes R(j)^{-1}\right). \qquad (29)
\end{aligned}$$

Solving (29) for $H_\Lambda' \lambda_\Lambda$ yields

$$\begin{aligned} H_\Lambda' \lambda_\Lambda &= \left(I_r \otimes R(j)^{-1}\right)\left[M_5 \text{vec}\left(\Lambda(j+1)\right) - \text{vec}\left(M_4\right)\right] \\ &= \left(I_r \otimes R(j)^{-1}\right) M_5 \left[\text{vec}\left(\Lambda(j+1)\right) - M_5^{-1}\text{vec}\left(M_4\right)\right] \\ &= \left(I_r \otimes R(j)^{-1}\right) M_5 \left[\text{vec}\left(\Lambda(j+1)\right) - \text{vec}\left(\Lambda_u(j+1)\right)\right], \end{aligned} \tag{30}$$

where

$$M_4 = \sum_{t=1}^{T} y_t^{(1)} f_{t|T}', \tag{31}$$

$$M_5 = \sum_{t=1}^{T} \left[\left(f_{t|T} f_{t|T}' + P_{t|T}\right) \otimes W_t\right], \tag{32}$$

and where $\Lambda_u(j+1)$ denotes the unrestricted estimate satisfying

$$\text{vec}\left(\Lambda_u(j+1)\right) = M_5^{-1}\text{vec}\left(M_4\right). \tag{33}$$

Pre-multiplying (30) by $\left[\left(I_r \otimes R(j)^{-1}\right) M_5\right]^{-1}$ yields

$$M_5^{-1}\left(I_r \otimes R(j)\right) H_\Lambda' \lambda_\Lambda = \text{vec}\left(\Lambda(j+1)\right) - \text{vec}\left(\Lambda_u(j+1)\right). \tag{34}$$

where we have used the properties $(AB)^{-1} = B^{-1}A^{-1}$ (if A and B are square matrices) and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$. Pre-multiplying (34) by H_Λ, plugging in the constraint $H_\Lambda \text{vec}(\Lambda(j+1)) = \kappa_\Lambda$ and (33), and solving for λ_Λ yields:

$$\lambda_\Lambda = \left[H_\Lambda M_5^{-1}\left(I_r \otimes R(j)\right) H_\Lambda'\right]^{-1}\left[\kappa_\Lambda - H_\Lambda \text{vec}\left(\Lambda_u(j+1)\right)\right].$$

Replacing the solution for λ_Λ in (29) and solving for vec$(\Lambda(j+1))$ yields

$$\begin{aligned} \text{vec}\left(\Lambda(j+1)\right) &= \text{vec}\left(\Lambda_u(j+1)\right) \\ &\quad + M_5^{-1}\left(I_r \otimes R(j)\right) H_\Lambda' \left[H_\Lambda M_5^{-1}\left(I_r \otimes R(j)\right) H_\Lambda'\right]^{-1}\left[\kappa_\Lambda - H_\Lambda \text{vec}\left(\Lambda_u(j+1)\right)\right]. \end{aligned} \tag{35}$$

Now consider the derivative of $\mathcal{L}(\theta, \theta(j), \lambda_\theta)$ with respect to vec(A):

$$\begin{aligned} \left.\frac{\partial L(\theta,\theta(j))}{\partial \text{vec}(A)}\right|_{A=A(j+1)} &= \frac{1}{2}\left.\frac{\partial \text{tr}\begin{bmatrix} \text{vec}(A)'\left(M_2' \otimes I_r\right)\text{vec}\left(Q(j)^{-1}\right) \\ +\text{vec}\left(M_2\right)'\left(I_r \otimes Q(j)^{-1}\right)\text{vec}(A) \\ -\text{vec}(A)'\left(M_3 \otimes Q(j)^{-1}\right)\text{vec}(A) \end{bmatrix}}{\partial \text{vec}(A)}\right|_{A=A(j+1)} + \lambda_A' H_A \\ &= \text{vec}\left(Q(j)^{-1}\right)'\left(M_2 \otimes I_r\right) - \text{vec}\left(A(j+1)\right)'\left(M_3 \otimes Q(j)^{-1}\right) + \lambda_A' H_A. \end{aligned}$$

Setting the last expression to zero and solving for $H_A' \lambda_A$ yields

$$H_A' \lambda_A = \left(M_3 \otimes Q(j)^{-1}\right)\text{vec}\left(A(j+1)\right) - \left(M_2' \otimes I_r\right)\text{vec}\left(Q(j)^{-1}\right), \tag{36}$$

Pre-multiplying both sides of (36) by $\left(M_3 \otimes Q\left(j\right)^{-1}\right)^{-1}$ yields

$$\begin{aligned}
\left(M_3^{-1} \otimes Q\left(j\right)\right) H_A' \lambda_A &= \left(M_3^{-1} \otimes Q\left(j\right)\right)\left(M_3 \otimes Q\left(j\right)^{-1}\right) \text{vec}\left(A(j+1)\right) - \left(M_3^{-1} \otimes Q\left(j\right)\right)\left(M_2' \otimes I_r\right) \text{vec}\left(Q\left(j\right)^{-1}\right) \\
&= \left(I_r \otimes I_r\right) \text{vec}\left(A(j+1)\right) - \left(M_3^{-1} \otimes Q\left(j\right)\right)\left(I_r \otimes Q\left(j\right)^{-1}\right) \text{vec}(M_2) \\
&= I_{r^2} \text{vec}\left(A(j+1)\right) - \left(M_3^{-1} \otimes I_r\right) \text{vec}(M_2) \\
&= \text{vec}\left(A(j+1)\right) - \text{vec}\left(M_2 M_3^{-1}\right) \\
&= \text{vec}\left(A(j+1)\right) - \text{vec}\left(A_u(j+1)\right), \qquad (37)
\end{aligned}$$

where

$$\text{vec}\left(A_u(j+1)\right)) = \text{vec}\left(M_2 M_3^{-1}\right). \qquad (38)$$

Pre-multiplying (37) by H_A, plugging in the constraint $H_A \text{vec}(A\left(j+1\right)) = \kappa_A$ and (38), and solving for λ_A yields

$$\lambda_A = \left[H_A \left(M_3^{-1} \otimes Q\left(j\right)\right) H_A'\right]^{-1} \left[\kappa_A - H_A \text{vec}\left(A_u(j+1)\right)\right].$$

Replacing the solution for λ_A in (36) and solving for vec$(A\left(j+1\right))$ yields

$$\begin{aligned}
\text{vec}\left(A(j+1)\right) &= \text{vec}\left(A_u(j+1)\right) \\
&\quad + \left(M_3^{-1} \otimes Q\left(j\right)\right) H_A' \left[H_A \left(M_3^{-1} \otimes Q\left(j\right)\right) H_A'\right]^{-1} \left[\kappa_A - H_A \text{vec}\left(A_u(j+1)\right)\right]. \qquad (39)
\end{aligned}$$

As $L(\theta, \theta(j))$ does not have to be maximized simultaneously with respect to all parameters (see, e.g., McLachlan and Krishnan, 2008), the following expressions are obtained by differentiating $L(\breve{\theta}, \theta(j))$, where $\breve{\theta} = \{\mu, \Sigma, \Lambda(j+1), R, A(j+1), Q\}$. The derivative of $L(\breve{\theta}, \theta(j))$ with respect to Q is given by

$$\begin{aligned}
\left.\frac{\partial L(\breve{\theta}, \theta(j))}{\partial Q}\right|_{Q=Q(j+1)} &= -\frac{T}{2} \left.\frac{\partial \log|Q|}{\partial Q}\right|_{Q=Q(j+1)} - \frac{1}{2} \left.\frac{\partial \text{tr}\left[Q^{-1}\left(\begin{array}{c} M_1 - M_2 A\left(j+1\right)' - A\left(j+1\right) M_2' \\ +A\left(j+1\right) M_3 A\left(j+1\right)' \end{array}\right)\right]}{\partial Q}\right|_{Q=Q(j+1)} \\
&= -\frac{T}{2} Q(j+1)^{-1} + \frac{1}{2}\left[Q(j+1)^{-1}\left(\begin{array}{c} M_1 - M_2 A\left(j+1\right)' - A\left(j+1\right) M_2' \\ +A\left(j+1\right) M_3 A\left(j+1\right)' \end{array}\right) Q(j+1)^{-1}\right],
\end{aligned}$$

where we have used the identities $\partial \log|X| / \partial X = X^{-1}$ and $\partial \text{tr}\left(X^{-1}A\right)/\partial X = -X^{-1}AX^{-1}$, and the fact that M_1 and M_3 are symmetric matrices. Setting the last expression to zero and solving for $Q(j+1)$ yields

$$Q(j+1) = \frac{1}{T}\left(M_1 - M_2 A\left(j+1\right)' - A\left(j+1\right) M_2' + A\left(j+1\right) M_3 A\left(j+1\right)'\right). \qquad (40)$$

The derivative of $L(\breve{\theta}, \theta(j))$ with respect to the i-th element on the diagonal of R is given by

$$\begin{aligned}
\left.\frac{\partial L(\theta, \theta(j))}{\partial \sigma_i^2}\right|_{\sigma_i^2=\sigma_i^2(j+1)} &= -\frac{T}{2} \left.\frac{\partial \log|R|}{\partial \sigma_i^2}\right|_{\sigma_i^2=\sigma_i^2(j+1)} \\
&\quad -\frac{1}{2}\text{tr}\left\{\left.\frac{\partial R^{-1}}{\partial \sigma_i^2}\right|_{\sigma_i^2=\sigma_i^2(j+1)} \sum_{t=1}^{T}\left[\begin{array}{c} \left(y_t^{(1)} - W_t \Lambda(j+1) f_{t|T}\right) \\ \left(y_t^{(1)} - W_t \Lambda(j+1) f_{t|T}\right)' \\ +W_t \Lambda(j+1) P_{t|T} \Lambda(j+1)' W_t \\ +\left(I_n - W_t\right) R(j) \left(I_n - W_t\right) \end{array}\right]\right\} \\
&= -\frac{T}{2} \sigma_i^2(j+1)^{-1} \\
&\quad +\frac{1}{2} \sigma_i^2(j+1)^{-2} \sum_{t=1}^{T}\left[\begin{array}{c} \left(y_{i,t}^{(1)} - \left(W_t \Lambda(j+1) f_{t|T}\right)_{ii}\right)^2 \\ +\left(W_t \Lambda(j+1) P_{t|T} \Lambda(j+1)' W_t\right)_{ii} \\ +\left(\left(I_n - W_t\right) R(j) \left(I_n - W_t\right)\right)_{ii} \end{array}\right],
\end{aligned}$$

where X_{ii} denotes the i-the element on the diagonal of matrix X. Setting the last expression to zero and solving for $\sigma_i^2(j+1)$ yields

$$\sigma_i^2(j+1) = \frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{c}\left(y_{i,t}^{(1)} - \left(W_t\Lambda(j+1)f_{t|T}\right)_{ii}\right)^2 + \left(W_t\Lambda(j+1)P_{t|T}\Lambda(j+1)'W_t\right)_{ii} \\ + \left(\left(I_n - W_t\right)R(j)\left(I_n - W_t\right)\right)_{ii}\end{array}\right].$$

Then, the updated estimate of R is

$$\begin{aligned} R(j+1) &= \operatorname{diag}\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{c}\left(y_t^{(1)} - W_t\Lambda(j+1)f_{t|T}\right)\left(y_t^{(1)} - W_t\Lambda(j+1)f_{t|T}\right)' \\ +W_t\Lambda(j+1)P_{t|T}\Lambda(j+1)'W_t + \left(I_n - W_t\right)R(j)\left(I_n - W_t\right)\end{array}\right]\right\} \\ &= \operatorname{diag}\left\{\frac{1}{T}\left[M_6 - M_7 - M_7' + M_8 + M_9\right]\right\}, \end{aligned} \tag{41}$$

where

$$M_6 = \sum_{t=1}^{T} y_t^{(1)} y_t^{(1)\prime}, \tag{42}$$

$$M_7 = \sum_{t=1}^{T} y_t^{(1)} f_{t|T}'\Lambda(j+1)'W_t, \tag{43}$$

$$M_8 = \sum_{t=1}^{T} W_t\Lambda(j+1)\left(f_{t|T}f_{t|T}' + P_{t|T}\right)\Lambda(j+1)'W_t, \tag{44}$$

$$M_9 = \sum_{t=1}^{T}\left(I_n - W_t\right)R(j)\left(I_n - W_t\right). \tag{45}$$

The derivative of $L(\theta,\theta(j))$ with respect to μ is

$$\begin{aligned} \left.\frac{\partial L(\theta,\theta(j))}{\partial\mu}\right|_{\mu=\mu(j+1)} &= -\frac{1}{2}\left.\frac{\partial \operatorname{tr}\left\{\Sigma^{-1}\left(f_{0|T} - \mu\right)\left(f_{0|T} - \mu\right)'\right\}}{\partial\mu}\right|_{\mu=\mu(j+1)} \\ &= \left(f_{0|T}' - \mu(j+1)'\right)\Sigma^{-1}, \end{aligned}$$

such that the update for the initial mean is

$$\mu(j+1) = f_{0|T}. \tag{46}$$

Defining $\check{\theta} = \{\mu(j+1), \Sigma, \Lambda, R, A, Q\}$, the derivative of $L(\check{\theta},\theta(j))$ with respect to Σ is

$$\begin{aligned} \left.\frac{\partial L(\check{\theta},\theta(j))}{\partial\Sigma}\right|_{\Sigma=\Sigma(j+1)} &= -\frac{1}{2}\left.\frac{\log|\Sigma|}{\partial\Sigma}\right|_{\Sigma=\Sigma(j+1)} \\ &\quad -\frac{1}{2}\left.\frac{\partial \operatorname{tr}\left\{\Sigma^{-1}\left[\left(f_{0|T} - \mu(j+1)\right)\left(f_{0|T} - \mu(j+1)\right)' + P_{0|T}\right]\right\}}{\partial\Sigma}\right|_{\Sigma=\Sigma(j+1)} \\ &= -\frac{1}{2}\Sigma(j+1)^{-1} + \frac{1}{2}\Sigma(j+1)^{-1}P_{0|T}\Sigma(j+1)^{-1}, \end{aligned}$$

such that the update for the initial variance-covariance matrix is

$$\Sigma(j+1) = P_{0|T}. \tag{47}$$

A.3 Initial Parameters

To initialize the algorithm, the set of parameters $\theta(0)$ is estimated by principal components adapted to consider the partition of the data and factors in an external and domestic block and the restrictions imposed in the equations (1) and (2) in the main text.

A.3.1 External block

The estimation of the external factors and their associated loadings relies on the variance-covariance matrix of the observables. More specifically, this matrix of standardized variables, y^E, is estimated as $\widehat{Cov}(y^E) = y^{E'}y^E/(T-1)$. Then, $\hat{\Lambda}^{EE}$ is calculated as the eigenvectors of $\widehat{Cov}(y^E)$, normalized and ordered decreasingly. Once $\hat{\Lambda}^{EE}$ is obtained, the factors are estimated as linear combinations of y^E, where the weights correspond to the factor loadings: $\hat{f}^E = y^E\hat{\Lambda}^{EE}$.

A.3.2 Domestic block

Once the external factors and loadings are estimated, it is possible to estimate the domestic block imposing the block exogeneity restrictions. This is done following the iterative approach proposed by Boivin and Giannoni (2007); see also Charnavoki and Dolado (2014) and Kamber et al. (2016). The idea behind this approach is that the estimated external factors can be used to orthogonalize (or "externally correct") the estimated domestic factors to control for the effect of foreign factors on domestic block. This guarantees that the estimated domestic factors capture only the dynamics of domestic variables not captured by the foreign factors. Formally, this is achieved using the following iterative approach:

1. Estimate $\{\hat{\Lambda}^{EE}, \hat{f}_t^E\}_{t=1}^T$ and $\{\hat{\Lambda}^{DD(0)}, \hat{f}_t^{D(0)}\}_{t=1}^T$ with principal components independently to obtain K^E external and K^D domestic factors.
2. For $i = 0$, regress y_t^D on $\hat{f}_t^E$ and $\hat{f}_t^{D(i)}$ using Ordinary Least Squares (OLS) to obtain $\hat{\Lambda}^{ED(i)}$ (which corresponds to the OLS coefficients associated with $\hat{f}_t^E$).
3. Compute $\hat{y}_t^{D(i)} = y_t^D - \hat{\Lambda}^{ED(i)}\hat{f}_t^E$.
4. Estimate $\{\hat{\Lambda}^{DD(i+1)}, \hat{f}_t^{D(i+1)}\}_{t=1}^T$ using the first K^D principal components of $\hat{y}_t^{D(i)}$.
5. Repeat steps 2-4 for $i = 1, 2, 3, \ldots$, until convergence of $\{\hat{f}_t^{D(i+1)}\}_{t=1}^T$.

A.3.3 Factor VAR

The block exogeneity restrictions present in the factor VAR are imposed by a reparametrization of the unrestricted version of the model that contains the desired linear constraints (i.e., exogeneity of the external factors with respect to the domestic factors). As a consequence of this reparametrization, not all the regressors are shared throughout all the equations present in the system. Thus, standard equation-by-equation OLS does not yield efficient parameter estimates Zellner (1962). This issue can be tackled by applying Feasible Generalized Least Squares (FGLS), which yields asymptotically efficient and consistent parameter estimates by estimating the covariance matrix of the residuals via an iterative approach. Also, since the VAR under consideration is stationary, standard asymptotic inference continues to apply. For further details see, e.g., Lütkepohl (2007).

A.4 EM algorithm pseudo-code

The procedure for the EM computations is as follows:

1. Initialize the parameters using the initial estimators $\theta(0) = \{\mu(0), \Sigma(0), \Lambda(0), A(0), R(0), Q(0)\}$.

On iteration j ($j = 1, 2, \ldots$):

2. Obtain the smoothed moments of the common factors for $t = 1, \ldots, T$ using the Kalman filter conditional on the estimated parameters, $\theta(j)$.

3. Compute the log-likelihood of the observed data.
4. Obtain the updated estimates, $\theta(j+1)$.
5. Repeat steps 2 through 4 until convergence.
6. Obtain estimates of the shadow rate and other missing data using the measurement equation.

For the convergence criterion, it should be smaller than 1×10^{-6}. We consider the relative change in the log-likelihood of the observed data, $[l_Y(\theta(j+1)) - l_Y(\theta(j))] / |l_Y(\theta(j))|$.

B Data

Table B1: Data sources and variable definitions

Variable	**Transformation**	**Source**	**Id**
DOMESTIC BLOCK			
Interest Rates			
Monetary Policy Rate		CBC	MPR
Interest Rate Swap 90d ("Promedio Cámara")		CBC	3M
Interest Rate Swap 180d ("Promedio Cámara")		CBC	6M
Interest Rate Swap 1y ("Promedio Cámara")		CBC	1Y
Interest Rate Swap 2y ("Promedio Cámara")		CBC	2Y
Interest Rate Swap 3y ("Promedio Cámara")		CBC	3Y
Interest Rate Swap 5y ("Promedio Cámara")		CBC	5Y
Interest Rate Swap 7y ("Promedio Cámara")		CBC	7Y
Interest Rate Swap 10y ("Promedio Cámara")		CBC	10Y
Bond rate 1y		CBC	bcp 1Y
Bond rate 2y		CBC	bcp 2Y
Bond rate 5y		CBC	bcp 5Y
Bond rate 10y		CBC	bcp 10Y
Monetary Aggregates			
Monetary Base or M0	YoY	CBC	M0
M1	YoY	CBC	M1
M2	YoY	CBC	M2
Balance Sheet: Assets			
Net External Position	YoY	CBC	NEP
Internal Credit	YoY	CBC	NEP
Bonds and notes	YoY	CBC	B&N
EXTERNAL BLOCK			
Interest Rates			
Effective federal fund rate		FRED	FYFF
Rate of U.S Treasury Bills 3 months		FRED	FYGM3
Rate of U.S Treasury Bills 6 months		FRED	FYGM6
Yield of U.S Treasury Bonds 1 year		FRED	FYGT1
Yield of U.S Treasury Bonds 2 year		FRED	FYGT2
Yield of U.S Treasury Bonds 5 year		FRED	FYGT5
Yield of U.S Treasury Bonds 10 year		FRED	FYGT10
Yield of U.S Treasury Bonds 20 year		FRED	FYGT20
Overnight indexed swap (OIS)		Bloomberg	OIS3M
Monetary Aggregates			
Monetary Base or M0	YoY	CBC	FMFBA_rb
M1	YoY	FRED	FM1
M2	YoY	FRED	F M2
Balance Sheet: Assets			
Total Assets	YoY	CBC	TotAss
Total Fed. Res. Securities held outright	YoY	CBC	TotSec
Percentage of long-term U.S Treasury securities <5 y	YoY	CBC	per 5y
Percentage of long-term U.S Treasury securities <10 y	YoY	CBC	per 10y
Percentage of long-term U.S Treasury securities >10 y	YoY	CBC	per GT10y
Balance Sheet: Liabilities			
Total Reserves	YoY	CBC	totRes
Excess of reserves	YoY	CBC	excessRes
Required reserves	YoY	CBC	reqRes

Notes: CBC = Central Bank of Chile and FRED = Federal Reserve Bank of St Louis.

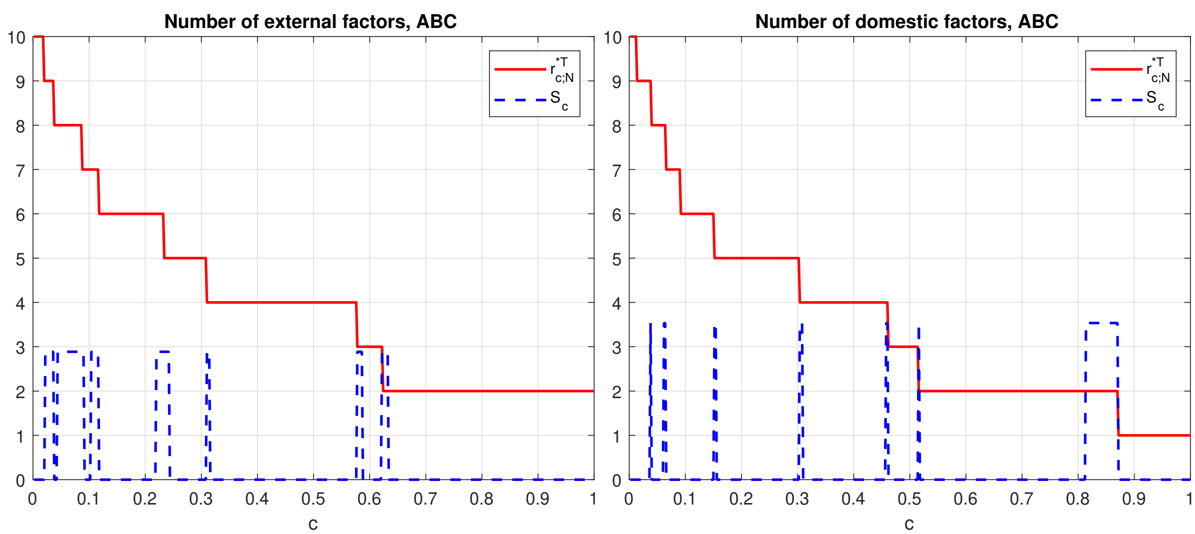
C Selection of the Number of Factors

This appendix describes the details of the methodology and analysis carried out to determine the number of factors.

The data of this application consist of 39 time series for 218 months (September 2002 to October 2020). For this dataset, the criteria proposed by Bai and Ng (2002) have tended to overestimate the number of factors. Therefore, the ABC criterion was implemented following Alessi et al. (2010). This criterion follows the idea of the procedure and arguments by Hallin and Liska (2007). Briefly, the ABC criterion multiplies the penalty functions of Bai and Ng (2002) by a constant (that is, $c = 1$ in Bai & Ng), then through an iterative process the constants are sensitized in an admissible range. Next, the number of factors $\hat{r_c}$, more robust for each value of the constant is estimated. Finally, the researcher decides the number of factors provided by the most robust model ($S_c = 0$). The codes were taken from the authors' website.

The ABC criterion is applied to the external and domestic data block and results are presented in Figure B1 on the left and right panels, respectively. On the vertical axis, the number of factors is presented and on the horizontal axis the values of the constant (c=1 by default in Bai & Ng). For the external block, it can be seen that stability zones are found by observing the functions $\hat{r_c}$ flat and S_c = 0 for: 2, 4, 5 and 6 factors. Following the selection criteria suggested by Alessi et al. (2010), the decision is based on the longest flat section $\hat{r_c}$. Note that higher (lower) values of the constant will tend to select a more (less) parsimonious model. In our case, following the above criteria, we select 4 external factors because they exhibit the most extensive zone of stability in the intermediate constant values of the range. Using a similar criterion for the domestic block, 5 domestic factors are considered (Figure B1, right panel).

Figure B1: ABC criterion



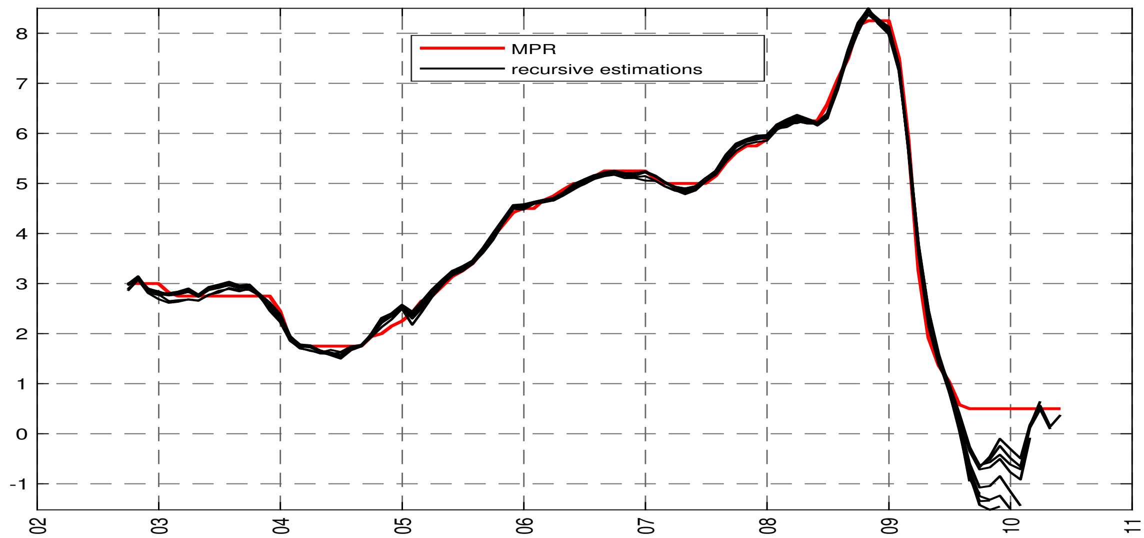
Notes: The solid red line indicates the optimal number of factors for different values of c. The blue dotted line indicates if this number of factors is stable across different subsamples. The idea is to choose the number of factors so that they are stable for different values of the constant and for different sample sizes, that is, where the blue dotted line is zero and the red line remains flat. These figures are standard and are generated using codes available from the authors' website, for more details see Alessi et al. (2010).

In summary, following the ABC criterior, the baseline models considers 4 external factors and 5 domestic factors. As shown in the main text, the main results are robust to small changes in the specification of the baseline model.

D Recursive Estimation Around the Global Financial Crisis

This section presents a recursive estimate of the SMPR during the global financial crisis episode. The model specification includes four domestic and foreign factors given that, as we illustrate in section 4.2, the fith domestic factor (F9) is strongly linked to the internal credit stimulus, which is related to the implementation of the FCIC and other special credit support programs in the context of the Covid-19 pandemic, which was not implemented during the financial crisis.

Figure D1: Recursive estimation



Notes: The black lines depict the recursive estimates of the SMPRs from August 2009 to May 2010. The thin red line depicts the effective MPR.

This exercise shows that as new data comes in during the crisis, the methodology tends to revise more the SMPR than in normal times, lending support to the view that around turning points confidence bands could be wider.

E Taylor Rule Estimation for Chile

The Taylor rule is adapted to capture the conduct of monetary policy in Chile. Chile has had an inflation targeting scheme in place since 1999, where the objective is expected inflation, which means that there is no MPR reaction to transitory shocks and idiosyncratic inflation surprises. Furthermore, the tradition of a mining country means that the most appropiate measure of inflationary pressure is the gap measured from non-mining output. The frequency of publication of GDP is quarterly.

The following rule is estimated:

$$i_t - in_t = c_1(i_{t-1} - in_{t-1}) + (1 - c_1)(c_2(E_t\pi_{t+4} - 3\%) + c_3y_t) + \epsilon_t, \tag{48}$$

where the variable to be explained is the interest rate gap, measured as the difference between the MPR and the nominal neutral rate (see Aldunate et al., 2019). The arguments are one lag in the interest rate gap (smoothing), the 11-month inflation expectations from the Economic Expectations Survey carried out by the CBC, and the non-mining output gap. Parameters c_1 to c_3 are estimated with standard econometric methods.

Table E1: Taylor rule estimates using different interest rates

Dep. Var	MPR	MPR	DFM SMPR	DFM SMPR
C_1	0.65***	0.65***	0.62***	0.62***
	(0.04)	(0.09)	(0.04)	(0.06)
C_2	1.93***	1.97***	2.1***	2.55***
	(0.38)	(0.61)	(0.4)	(0.52)
C_3	0.36***	0.36***	0.38***	0.34***
	(0.06)	(0.07)	(0.06)	(0.09)
R^2	0.95	0.94	0.95	0.95
N. obs.	57	57	57	57
Method	OLS	GMM	OLS	GMM

Notes: Quarterly frequency estimates from 2006Q2 to 2020Q3. When using the MPR, we include dummies to account for the FLAP (coefficients not reported). Standard errors in parenthesis. The estimation of the Taylor rules assumes a neutral MPR of the order of 4% Aldunate et al. (2019)), the constant is not reported. *** denotes that estimates are statistically different from zero at 1% levels.

References

Aldunate, R., F. Bullano, M. Canales, G. Contreras, A. Fernández, J. Fornero, M. García, B. García, J. Pena, M. Tapia, and R. Zuniga (2019). Estimation of structural parameters of the Chilean economy. Central Bank of Chile (in Spanish).

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved Penalization for Determining the Number of Factors in Approximate Factor Models. *Statistics and Probability Letters 80*, 1806–1813.

Bai, J. and S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica 70*(1), 191–221.

Bańbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics 29*(1), 133–160.

Boivin, J. and M. P. Giannoni (2007). Global Forces and Monetary Policy Effectiveness. In *International Dimensions of Monetary Policy*. National Bureau of Economic Research, Inc.

Bork, L. (2009). Estimating US Monetary Policy Shocks Using a Factor-Augmented Vector Autoregression: An EM Algorithm Approach. CREATES Research Papers 2009-11, Department of Economics and Business Economics, Aarhus University.

Bork, L., H. Dewachter, and R. Houssa (2009). Identification of Macroeconomic Factors in Large Panels. CREATES Research Papers 2009-43, Department of Economics and Business Economics, Aarhus University.

Charnavoki, V. and J. Dolado (2014). The effects of global shocks on small commodity-exporting economies: Lessons from canada. *American Economic Journal: Macroeconomics 6*(2), 207–37.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B 39*(1), 1–38.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics 164*(1), 188–205.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford: Oxford University Press.

Giannone, D., L. Reichlin, and L. Sala (2005). Monetary Policy in Real Time. In *NBER Macroeconomics Annual 2004, Volume 19*. National Bureau of Economic Research, Inc.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Hallin, M. and R. Liska (2007). Determining the Number of Factors in the General Dynamic Factor Model. *Journal of the American Statistical Association 102*(478), 603–617.

Jungbacker, B., S. Koopman, and M. van der Wel (2011). Maximum likelihood estimation for dynamic factor models with missing data. *Journal of Economic Dynamics and Control 35*(8), 1358–1368.

Kamber, G., O. Karagedikli, M. Ryan, and T. Vehbi (2016). International spill-overs of uncertainty shocks: Evidence from a FAVAR. CAMA Working Paper 61/2016, Crawford School of Public Policy, Australian National University.

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer Berlin Heidelberg.

McLachlan, G. J. and T. Krishnan (2008). *The EM Algorithm and Extensions* (2nd ed.). Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley.

Shumway, R. H. and D. S. Stoffer (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis 3*(4), 253–264.

Watson, M. W. and R. F. Engle (1983). Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression models. *Journal of Econometrics 23*(3), 385–400.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association 57*(298), 348–368.