
DMFM and Tensor 3PRF: A Nowcasting Framework for Mixed-Frequency Data in the Euro Area

Matteo Barigozzi[†] Massimiliano Marcellino[‡] Davide Delfino[‡]

[†] Alma Mater Studiorum – Università di Bologna

[‡] Bocconi University

Contents

1 Vector Models	5
1.1 Data Preparation	5
1.1.1 Data	5
1.1.2 Factor-Number Selection under Missing Data	5
1.1.3 Predictor Selection Based on GDP Correlation	6
1.2 Mixed-Frequency TPRF (MF–TPRF)	8
1.2.1 Standard Three-Pass Regression Filter	9
1.2.2 MF Three-Pass Regression Filter: Preparation	11
1.2.3 Pass 1: Quarterly Regressions	11
1.2.4 Pass 2: Monthly Factor Extraction	12
1.2.5 Pass 3: Quarterly U–MIDAS Regression	12
1.2.6 Handling Mixed Frequency and Ragged Edges	13
1.2.7 Nowcasting Structure	14
1.2.8 Lag-Length Selection in the MF–TPRF	15
1.3 MF-TPRF - Results	16
1.3.1 FR - France	19
1.3.2 IT - Italy	21
1.3.3 DE - Germany	23
1.3.4 ES - Spain	25
2 Matrix Models	27
3 Tensor-Based TPRF	28
3.1 T-MF-TPRF	28
3.1.1 Step 1 of the Tensor 3PRF: Single Proxy and Multiple Proxies	29
3.1.2 Step 2: Estimation of Monthly Tensor Factors	32
3.1.3 Step 3: Quarterly U-MIDAS Regression and Real-Time Nowcasting	33
3.1.4 Nowcasting Structure	33
3.1.5 EM Algorithm under a Tensor-Valued Predictor Structure	34
A Additional Results Accounting for Alternative Specifications	37
A.1 ES – Spain	37

A.2 FR – France	38
---------------------------	----

Table 1: Model structure and section references

Reference by section	Vector models			Matrix models	
	MF-TPRF	DFM	Tensor MF-TPRF	DMFM	
Data preparation	Bai and Ng (2008) and De Mol, Giannone, and Reichlin (2008) ^a	^a	^a	^a	^a
Factor selection	Xiong and Pelger (2023) and Ahn and Horenstein (2013) ^b	^b	Cen and Lam (2025) and Yu et al. (2022)		^c
Model reference	Hepenstrick and Marcellino (2016)	Bañbara and Modugno (2014)	...		Matteo Barigozzi and Trapin (2025)

Notes: ^a The data-preparation reference Bai and Ng (2008) applies to all models. ^b The factor-selection references Xiong and Pelger (2023) and Ahn and Horenstein (2013) apply to both vector models.

1 Vector Models

1.1 Data Preparation

1.1.1 Data

- **Source:** mixed-frequency *EA–MD–QD* dataset from M. Barigozzi and Lissona (2024).
- **Transformations:** all series are pre-transformed to ensure stationarity.¹
- **Time span:** April 2000 – October 2025.²
- **Crises covered:**
 - Global Financial Crisis (2008–2009): no special treatment for any class of variables.
 - COVID pandemic (March 2020 – July 2021): masking of real variables (those most affected by the crisis).
- **Types of missingness:**
 - Regular missingness from quarterly variables.
 - COVID-induced blocks of missing values for real variables during the masked period.
 - Ragged edge from heterogeneous publication calendars, handled in the pseudo real-time forecasting exercise.

1.1.2 Factor-Number Selection under Missing Data

Given the arbitrary missingness pattern in our large dataset, selecting the number of factors on a different subset (e.g. only monthly or only quarterly data, or excluding the COVID period) would be inconsistent with the final specification. We therefore follow an alternative approach.

All-purpose Covariance Estimator and Eigenvalue–Ratio

- Let $X \in \mathbb{R}^{T \times N}$ be the standardized panel with missing entries.
- Define the missingness matrix and the masked data as

$$w_{t,i} = \begin{cases} 1, & X_{t,i} \text{ observed}, \\ 0, & X_{t,i} \text{ missing}, \end{cases} \quad \tilde{X}_{t,i} = w_{t,i} X_{t,i},$$

i.e. missing values are imputed by zero in \tilde{X} .

¹Each quarterly variable is placed in the third month of the corresponding quarter, regardless of its original transformation.

²The sample starts in April 2000 to avoid missing values in the annual growth rate at the beginning of the time span, when the EA aggregate did not yet exist in the previous year.

- Following Xiong and Pelger (2023), estimate the covariance as

$$\tilde{\Sigma}_{ij} = \frac{1}{\Theta_{ij}} \sum_{t:w_{t,i}w_{t,j}=1} X_{t,i}X_{t,j}, \quad \Theta_{ij} = \sum_{t=1}^T w_{t,i}w_{t,j}.$$

- Let $\lambda_1 \geq \dots \geq \lambda_N$ be the eigenvalues of $\tilde{\Sigma}$. For a given K_{\max} , select

$$\hat{r} = \arg \max_{1 \leq j \leq K_{\max}} \frac{\lambda_j}{\lambda_{j+1}},$$

as in Ahn and Horenstein (2013).³

Otherwise, as an alternative, we could reconstruct the dataset by first estimating factors using the method of Xiong and Pelger (2023), substituting the estimated common component for the missing entries, and then apply the classical Bai–Ng information criteria (Bai and Ng 2002), which assume a clean, balanced, and exogenous panel.

1.1.3 Predictor Selection Based on GDP Correlation

Preprocessing Besides targeting the factors (especially during the TPRF) Bai and Ng (2008) suggest that preselecting variables most correlated with the target can provide a stronger factor structure. Here we introduce three method to preselect variables during the data preparation ⁴.

- Let y_τ denote quarterly GDP.
- Let $X_{m,t}$ and $X_{q,\tau}$ be monthly and quarterly predictors.
- Monthly series are aggregated to quarterly frequency using an operator $\omega(L)$ (sum or average depending on the variable type):

$$X_{\tau,j}^{(m \rightarrow q)} = \omega(L)X_{m,t,j}.$$

Selection criteria (country-by-country) Selection is performed separately for each country to reflect heterogeneity. Thus, for each country c and each predictor x_j :

1. Fixed-number criterion

- Compute absolute correlation

$$\rho_j = |\text{corr}(X_{\cdot,j}^{(m \rightarrow q)}, y)|.$$

³This factor–number selection is parallel to the one used in the matrix formulation, via the tensor extensions of Cen and Lam (2025) and Yu et al. (2022) applied to Xiong and Pelger (2023) and Ahn and Horenstein (2013), respectively.

⁴LASSO by De Mol, Giannone, and Reichlin (2008) can be an alternative additional method to extract the most targeted predictors

- Keep the n_m monthly and n_q quarterly predictors with largest $|\rho_j|$ according to the user parameter's specification.

2. Correlation-threshold criterion

- Select all predictors with

$$|\text{corr}(X_{\cdot,j}^{(m \rightarrow q)}, y)| \geq \tau_m, \quad |\text{corr}(X_{\cdot,j}^{(q)}, y)| \geq \tau_q.$$

3. F-test significance criterion

- For each predictor, run a simple regression of $X_{\tau,j}^{(m \rightarrow q)}$ on y_τ .
- Compute the F -statistic for $H_0 : \beta_j = 0$ and retain j if

$$p\text{-value}(F_j) \leq \tau_F.$$

1.2 Mixed-Frequency TPRF (MF–TPRF)

Overview and motivation

- Kelly and Pruitt (2015) propose the Three-Pass Regression Filter (3PRF) as a simple, OLS-based method to extract *targeted factors* for forecasting a specific variable using a large set of predictors.
- The procedure is:
 - computationally inexpensive,
 - easy to implement (only linear regressions),
 - designed to focus on the predictive content of high-dimensional predictors.
- However, the standard 3PRF relies on assumptions usually violated in macroeconomic applications:
 - a single frequency,
 - a balanced panel without ragged edges,
 - no temporary data unavailability.
- To address these issues, Hepenstrick and Marcellino (2016) extend 3PRF to the mixed-frequency framework (MF–3PRF), allowing:
 - a quarterly target y_τ to coexist with higher-frequency predictors x_t ,
 - explicit handling of ragged edges and unbalanced panels in x_t .

1.2.1 Standard Three-Pass Regression Filter

Model setup

- Let F_t denote latent factors, and consider the model:

$$y_{t+1} = \beta' F_t + \eta_{t+1}, \quad z_t = \Lambda F_t + \omega_t, \quad x_t = \Phi F_t + \varepsilon_t,$$

where:

- y_{t+1} is the target variable to be forecast ⁵,
- z_t is a low-dimensional proxy (e.g. asset returns, macro indicator, or typically the target itself),
- x_t is a large-dimensional set of predictors exhibiting balancedness.

Pass 1: Targeting step

- For each predictor $x_{i,t}$, run the regression

$$x_{i,t} = z_t' \phi_i + u_{i,t},$$

and stack the OLS estimates into

$$\hat{\Phi}_z' = (Z' Z)^{-1} Z' X,$$

where Z collects the proxies z_t and X the predictors x_t .

- Interpretation: $\hat{\Phi}_z$ measures how strongly each predictor responds to the proxy z_t and “targets” the part of X that is most relevant for forecasting y .

Pass 2: Factor extraction

- Using $\hat{\Phi}_z$, extract targeted factors from the high-dimensional predictors via

$$\hat{F}_t = (\hat{\Phi}_z' \hat{\Phi}_z)^{-1} \hat{\Phi}_z' x_t,$$

or, in matrix form,

$$\hat{F} = X \hat{\Phi}_z (\hat{\Phi}_z' \hat{\Phi}_z)^{-1}.$$

- Interpretation: \hat{F}_t summarizes the portion of X_t that is most correlated with the proxy z_t and therefore most relevant for forecasting y .

⁵the target is always excluded from the matrix of predictors X .

Pass 3: Forecast regression

- Regress the target on the extracted factors:

$$y_{t+1} = \beta' \hat{F}_t + \eta_{t+1}, \quad \hat{\beta} = (\hat{F}' \hat{F})^{-1} \hat{F}' Y.$$

- The forecast is then given by

$$\hat{y}_{t+1} = \hat{\beta}' \hat{F}_t.$$

- Interpretation: the final regression uses a small number of targeted factors, rather than the full set of predictors, delivering a low-dimensional but information-rich forecasting model.

1.2.2 MF Three-Pass Regression Filter: Preparation

- Let y_τ and z_τ be observed at quarterly frequency, and let x_t be observed monthly.
- Following Hedenstrom and Marcellino (2016), quarters are indexed by τ and months by t .
- Temporal aggregation from monthly to quarterly is defined by ⁶:

$$\omega(L) = \omega_0 + \omega_1 L + \cdots + \omega_{k-1} L^{k-1}, \quad k = 3,$$

1.2.3 Pass 1: Quarterly Regressions

Aggregation and regressions Consider for simplicity to run this steps on a centered set of predictors.

- Monthly predictors and proxies are first aggregated to quarterly frequency ⁷:

$$x_{i,\tau} = \omega(L)x_{i,t}, \quad z_\tau = \omega(L)z_t.$$

- Here,

$$x_{i,\tau} \in \mathbb{R}, \quad z_\tau \in \mathbb{R}^L$$

collects the quarterly aggregated values.

- For each predictor $i = 1, \dots, N$, estimate the quarterly regression

$$x_{i,\tau} = z'_\tau \alpha_i + u_{i,\tau}, \quad \alpha_i \in \mathbb{R}^L.$$

Stacked representation and OLS

- Stacking across predictors:

$$x_\tau = \mathbf{A}_z z_\tau + u_\tau, \quad \mathbf{A}_z = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_N \end{pmatrix} \in \mathbb{R}^{N \times L}.$$

- In matrix form, with $Z \in \mathbb{R}^{T_Q \times L}$ and $X \in \mathbb{R}^{T_Q \times N}$ collecting quarterly observations,

$$\widehat{\mathbf{A}}'_z = (Z'Z)^{-1} Z'X$$

is the OLS estimator of the loading matrix \mathbf{A}_z .

⁶e.g. quarterly growth is given by the simple sum $\omega(L) = 1 + L + L^2$.

⁷In the empirical exercise we treat stocks as averages over the quarter and flows as the sum

Robust F-test refinement

- After computing the OLS estimator $\widehat{\mathbf{A}}_z$, we refine the loading matrix by testing, for each predictor i , the joint null hypothesis

$$H_{0,i} : \alpha_i = 0$$

using a robust Wald/F-test based on heteroskedasticity- and autocorrelation-robust covariance matrices (e.g. White or Newey–West).

- If the null $H_{0,i}$ cannot be rejected at a given significance level, we set $\widehat{\alpha}_i = 0$, effectively excluding predictor i from the construction of the factors in the next step.
- This screening additionally contributes in removing series that do not co-move significantly with the proxy vector z_τ , thereby reducing noise in the factor extraction while retaining predictors that are genuinely informative for the target.

1.2.4 Pass 2: Monthly Factor Extraction

Projection on targeted loadings

- At the monthly frequency, the cross-section is projected onto the estimated loadings:

$$x_t = \widehat{\mathbf{A}}_z F_t + \varepsilon_t, \quad x_t \in \mathbb{R}^N, \quad F_t \in \mathbb{R}^L.$$

- Cross-sectional OLS yields the targeted factors

$$\widehat{F}_t = (\widehat{\mathbf{A}}_z' \widehat{\mathbf{A}}_z)^{-1} \widehat{\mathbf{A}}_z' x_t.$$

Stacked factors

- Stacking over months t gives

$$\widehat{F} = X_M \widehat{\mathbf{A}}_z (\widehat{\mathbf{A}}_z' \widehat{\mathbf{A}}_z)^{-1}, \quad \widehat{F} \in \mathbb{R}^{T_M \times L},$$

where X_M collects the monthly predictors.

1.2.5 Pass 3: Quarterly U–MIDAS Regression

Within-quarter factor positions

- Monthly factors are grouped into their positions within each quarter:

$$F_t \longrightarrow F_\tau^Q = \begin{pmatrix} F_\tau^{(1)} \\ F_\tau^{(2)} \\ F_\tau^{(3)} \end{pmatrix} \in \mathbb{R}^{3L},$$

where $F_\tau^{(1)}, F_\tau^{(2)}, F_\tau^{(3)}$ denote the factors in the first, second, and third month of quarter τ .

Quarterly regression and estimation

- Without loss of generality, we extend the U–MIDAS specification by adding an AR(1) term in the quarterly target:

$$y_\tau = \beta_0 + \rho y_{\tau-1} + B' F_{\tau-1}^Q + \eta_\tau, \quad B \in \mathbb{R}^{3L}.$$

- In matrix form, let \hat{F}^Q stack the lagged factor blocks $F_{\tau-1}^Q$ and let Y collect y_τ . Define the regressor matrix

$$X^Q = [\mathbf{1}, y_{-1}, \hat{F}^Q],$$

where $\mathbf{1}$ is a column of ones and y_{-1} stacks $y_{\tau-1}$. The parameter vector

$$\theta = \begin{pmatrix} \beta_0 \\ \rho \\ B \end{pmatrix}$$

is estimated by OLS as

$$\hat{\theta} = [(X^Q)' X^Q]^{-1} (X^Q)' Y, \quad Y \in \mathbb{R}^{T_Q}.$$

- The AR(1) component $\rho y_{\tau-1}$ is then included in all subsequent nowcasts, so that each real-time forecast combines information from lagged GDP and from the high-frequency factors.

1.2.6 Handling Mixed Frequency and Ragged Edges

As already mentioned, y_τ and z_τ are observed at quarterly frequency, but nothing prevents some $x_{i,t}$ be observed at the same lower frequency.

Selection matrices

- Let A_i be a *selection matrix* for series i , mapping the full time series into the vector of observed entries:

$$x_{i,\text{obs}} = A_i x_i, \quad A_i \in \{0, 1\}^{m_i \times T_M},$$

where m_i is the number of observed monthly values for series i .

- Missing or lower-frequency data correspond to rows of A_i that skip certain months.

EM-style imputation

- At iteration j , missing values of series i are updated using the current factor estimates:

$$\hat{x}_{i,t}^{(j)} = \hat{F}^{(j-1)} \hat{\Phi}_i^{(j-1)} + A_i (A_i' A_i)^{-1} (x_{i,\text{obs}} - A_i \hat{F}^{(j-1)} \hat{\Phi}_i^{(j-1)}).$$

- Intuitively, the first term is the model-implied common component, and the second term corrects it to exactly match the observed entries.
- Factors are then re-estimated (e.g. via PCA) on the completed dataset, and the procedure is iterated until convergence.
- To implement successfully this step the following procedures needs to be followed:

- standardize the set of predictors,
- extract the number of factors,
- Run the EM,
- aggregate at quarterly level the reconstructed dataset,
- center the data again to run the MF-TPRF or add a constant to the regressions in Step 1

1.2.7 Nowcasting Structure

In the empirical application, the real-time forecasting exercise is computed taking into account the publication delays of each indicator. For every month in the evaluation sample (from January 2017 to October 2025), the information set reflects only data released up to that date. The estimation sample, spanning from April 2000 to December 2016, is used to determine the number of factors and the lag length in the U-MIDAS regression.⁸

Within-quarter nowcasts with AR(1)

- Given monthly factors $(F_\tau^{(1)}, F_\tau^{(2)}, F_\tau^{(3)})$ and an AR(1) term in the quarterly regression, the nowcast of y_τ is updated as new monthly information arrives:

$$\begin{aligned}\hat{y}_\tau^{(1)} &= \hat{\beta}_0 + \hat{\rho} y_{\tau-1} + \hat{\beta}_1 F_\tau^{(1)}, \\ \hat{y}_\tau^{(2)} &= \hat{\beta}_0 + \hat{\rho} y_{\tau-1} + \hat{\beta}_1 F_\tau^{(1)} + \hat{\beta}_2 F_\tau^{(2)}, \\ \hat{y}_\tau^{(3)} &= \hat{\beta}_0 + \hat{\rho} y_{\tau-1} + \hat{\beta}_1 F_\tau^{(1)} + \hat{\beta}_2 F_\tau^{(2)} + \hat{\beta}_3 F_\tau^{(3)}.\end{aligned}$$

- These correspond to nowcasts after the first, second, and third month of the quarter, respectively. The AR(1) component $\hat{\rho} y_{\tau-1}$ enters all three nowcasts with the same coefficient.

⁸The baseline results are obtained with an expanding window. Given the large shocks associated with COVID, a rolling-window exercise is a natural robustness check.

1.2.8 Lag-Length Selection in the MF–TPRF

MF–TPRF requires choosing the number of quarterly lags L for the factor blocks in the U–MIDAS target equation. We first run Pass 1–2 of the MF–TPRF as described in the previous subsection to obtain the quarterly factors $F_\tau^{(1)}, F_\tau^{(2)}, F_\tau^{(3)}$. In Pass 3, we fix an AR(1) term for the target and then select L as follows.

- Motivated by standard macroeconomic dynamics, we impose an AR(1) term in y_τ and, for each candidate $L \leq L_{\max}$, estimate the quarterly regression

$$y_\tau = \beta_0 + \rho y_{\tau-1} + \sum_{\ell=1}^L \left(F_{\tau-\ell}^{(1)} \gamma_\ell^{(1)} + F_{\tau-\ell}^{(2)} \gamma_\ell^{(2)} + F_{\tau-\ell}^{(3)} \gamma_\ell^{(3)} \right) + \varepsilon_\tau.$$

- Conditional on the AR(1) specification, we compute information criteria for each L :

$$\hat{L}_{\text{AIC}} = \arg \min_L \text{AIC}(L), \quad \hat{L}_{\text{BIC}} = \arg \min_L \text{BIC}(L).$$

- The selected \hat{L} thus provides a data-driven lag choice for the factor blocks, given the fixed AR(1) dynamics in the target equation.

1.3 MF-TPRF - Results

Parameters for the Empirics The following table reports the parameters used for the results reported in this section for all the considered countries ⁹.

Table 2: Common parameter settings for the pseudo real-time MF-TPRF exercise

Category	Parameter	Value	Description
Sample windows	$t_{\text{start}}^{\text{est}}$	2000Q2	Start of the estimation sample (<code>start_est</code> = 2000-04-01).
	$t_{\text{start}}^{\text{eval}}$	2017M1	Start of the evaluation window (<code>start_eval</code> = 2017-01-01).
	$t_{\text{end}}^{\text{eval}}$	2025M10	End of the evaluation window (<code>end_eval</code> = 2025-10-01).
COVID window	$t_{\text{start}}^{\text{covid}}$	2020M3	COVID shock start (<code>covid_start</code> = 2020-03-01).
	$t_{\text{end}}^{\text{covid}}$	2021M7	COVID shock end (<code>covid_end</code> = 2021-07-01).
	covid_mask_m	TRUE	Monthly indicators are masked during the COVID window.
Target & selection	covid_mask_q	FALSE	Quarterly indicators are not masked during the COVID window.
	Target variable	GDP	Quarterly real GDP growth is the target series (<code>target</code> = "GDP").
	Selection method	corr_threshold	Variable pre-selection based on correlation thresholds.
Variable selection	thr_m	0.10	Correlation threshold for monthly variables.
	thr_q	0.85	Correlation threshold for quarterly variables.
MF-TPRF	AR(p)	1	AR order in the U-MIDAS regression for y_t (AR(1)).
	L_{\max}	3	Maximum number of quarterly lags of the factors in U-MIDAS (<code>Lmax</code> = 3).
	Robust_F	TRUE	Use robust F-tests in Step 1 to zero out insignificant proxy loadings.
	α	0.10	Significance level for robust t/F tests on proxy coefficients.
	robust_type	NW	Robust covariance type: Newey–West HAC (<code>robust_type</code> = "NW").
	nw_lag	1	Lag order in the Newey–West covariance estimator.

⁹This set is just a convenient one but nothing prevents to playing with it.

Nowcast procedure

- **Define samples.**
 - Choose an *estimation sample* (from start of the dataset up to `start_eval`).
 - Choose an *evaluation sample* (from `start_eval` to `end_eval`).
- **Hyperparameter selection on the estimation sample.**
 - Apply the unbalancedness scheme to reflect publication lags at the end of the estimation sample.
 - Run the EM algorithm to complete the mixed-frequency dataset and extract common factors.
 - Aggregate the completed data to quarterly frequency and align it with the quarterly target y_t .
 - Select the number of proxies L_{proxy} via eigenvalue ratio (ER) on the monthly covariance matrix, reconstructed using the "All purpose estimator".
 - Select the U-MIDAS lag length L_{midas} (having p_{AR} fixed) by minimizing an information criterion (AIC/BIC) in the quarterly regression.
- **Set fixed hyperparameters.**
 - Fix $(L_{\text{proxy}}, L_{\text{midas}})$ for the entire evaluation period, based only on the estimation sample ¹⁰.
- **Expanding-window real-time loop (evaluation sample).**
 - For each monthly date t in the evaluation window:
 - * Apply the unbalancedness scheme to X using only observations up to t .
 - * Run the EM algorithm again to complete the mixed-frequency dataset at date t .
 - * Split the completed data into monthly and quarterly blocks and aggregate monthly series to quarter.
 - * Determine how many quarters of the target y_t are available at date t .
 - * Construct low-frequency (X^{LF}) and high-frequency (X^{HF}) predictors up to the last observed quarter.
- **MF-TPRF estimation at each date t .**
 - *Step 1 (proxy regression):* Regress each low-frequency variable on the fixed set of proxies using OLS with a robust F-test to zero out non-significant loadings.
 - *Step 2 (factor extraction):* Use the selected loadings to construct monthly factors from X^{HF} .
 - *Step 3 (U-MIDAS + AR):* Build a quarterly regression of y_t on current and lagged factors (up to L_{midas}) and lagged y_{t-j} (AR part).

¹⁰Due to the pandemic crises upheavals we could also consider to re-estimate such parameters.

- *Step 4 (nowcasting):* Translate the quarterly regression into monthly nowcasts for the current target quarter, at M1/M2/M3, given the information set at date t .
- **Store pseudo real-time nowcasts.**
 - For each monthly date t , save the nowcast produced at M1, M2, or M3 of the target quarter.
 - Collect all nowcasts into three series (M1, M2, M3) and a combined panel for forecast evaluation (e.g., RMSFE pre-/post-COVID).

1.3.1 FR - France

In-sample fit and Q4 2025 GDP nowcast

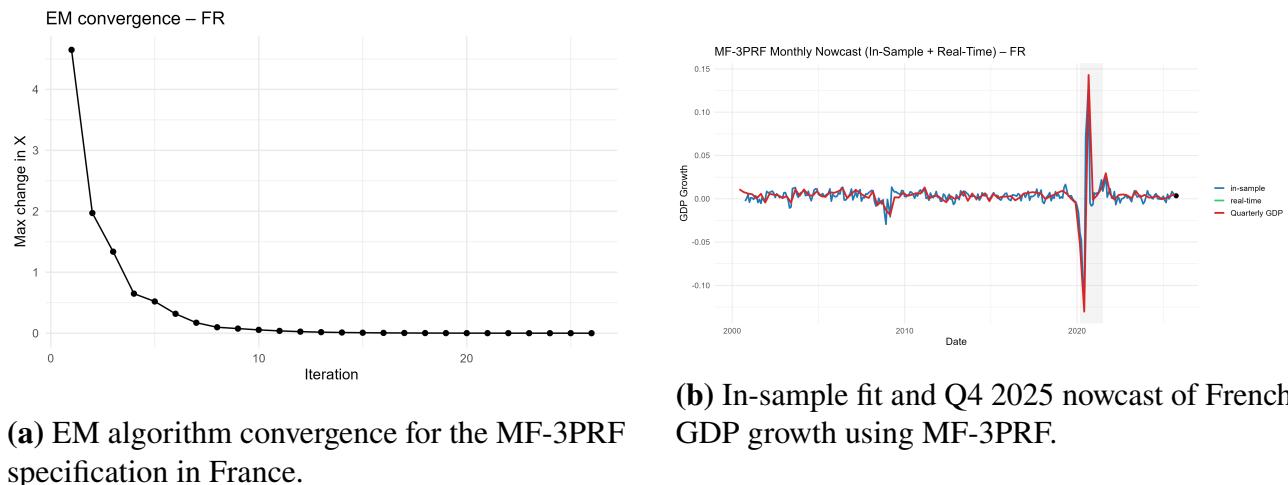


Figure 1: France: EM convergence (panel a) and in-sample fit with Q4 2025 GDP nowcast (panel b).

Table 3: MF-3PRF RMSFE by period – France

Period	M1	M2	M3
Full sample	0.0207	0.0122	0.0050
Pre-COVID	0.0055	0.0047	0.0041
COVID period	0.0481	0.0274	0.0077
Post-COVID	0.0066	0.0049	0.0044

Expanding-window nowcasts

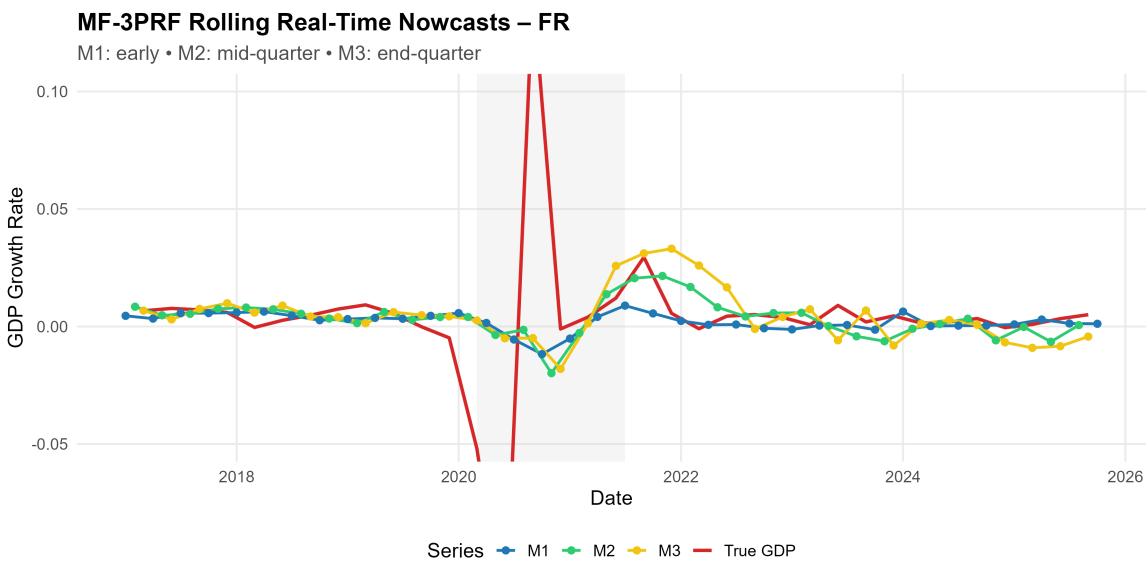


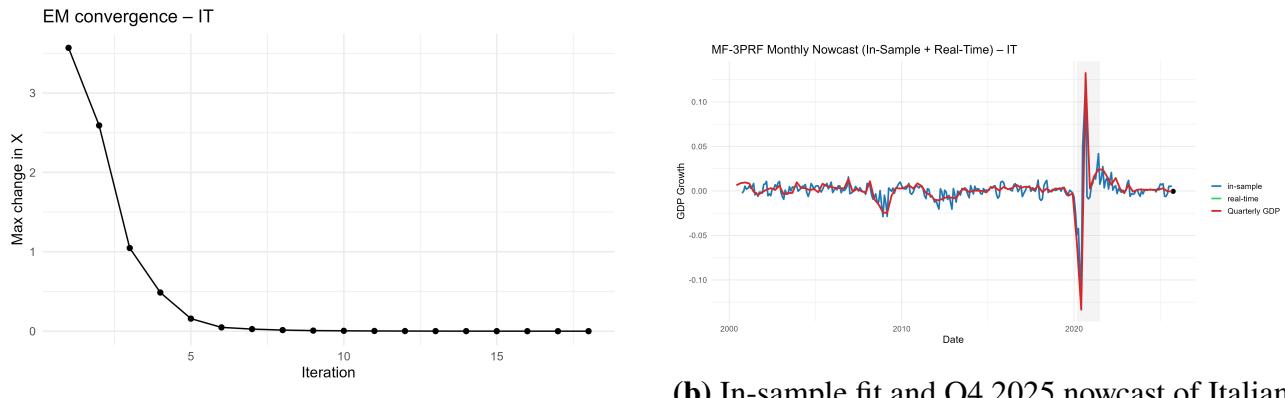
Figure 2: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for French GDP.

Table 4: MF-3PRF expanding-window RMSFE by period – France

Period	M1	M2	M3
Full sample	0.0410	0.0410	0.0422
Pre-COVID	0.0047	0.0049	0.0052
COVID period	0.0984	0.0975	0.0986
Post-COVID	0.0060	0.0089	0.0146

1.3.2 IT - Italy

In-sample fit and Q4 2025 GDP nowcast



(a) EM algorithm convergence for the MF-3PRF specification in Italy.

(b) In-sample fit and Q4 2025 nowcast of Italian GDP growth using MF-3PRF.

Figure 3: Italy: EM convergence (panel a) and in-sample fit with Q4 2025 GDP nowcast (panel b).

Table 5: MF-3PRF RMSFE by period – Italy

Period	M1	M2	M3
Full sample	0.0234	0.0142	0.0063
Pre-COVID	0.0047	0.0041	0.0043
COVID period	0.0549	0.0327	0.0112
Post-COVID	0.0073	0.0050	0.0048

Expanding-window nowcasts

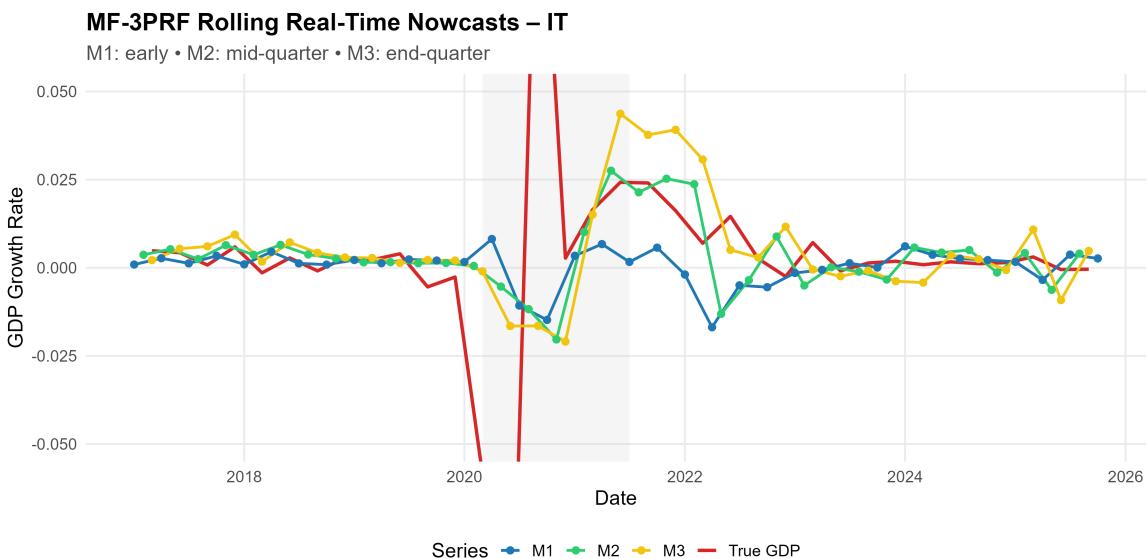


Figure 4: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for Italian GDP.

Table 6: MF-3PRF expanding-window RMSFE by period – Italy

Period	M1	M2	M3
Full sample	0.0366	0.0351	0.0347
Pre-COVID	0.0033	0.0033	0.0040
COVID period	0.0866	0.0830	0.0818
Post-COVID	0.0105	0.0097	0.0105

1.3.3 DE - Germany

In-sample fit and Q4 2025 GDP nowcast

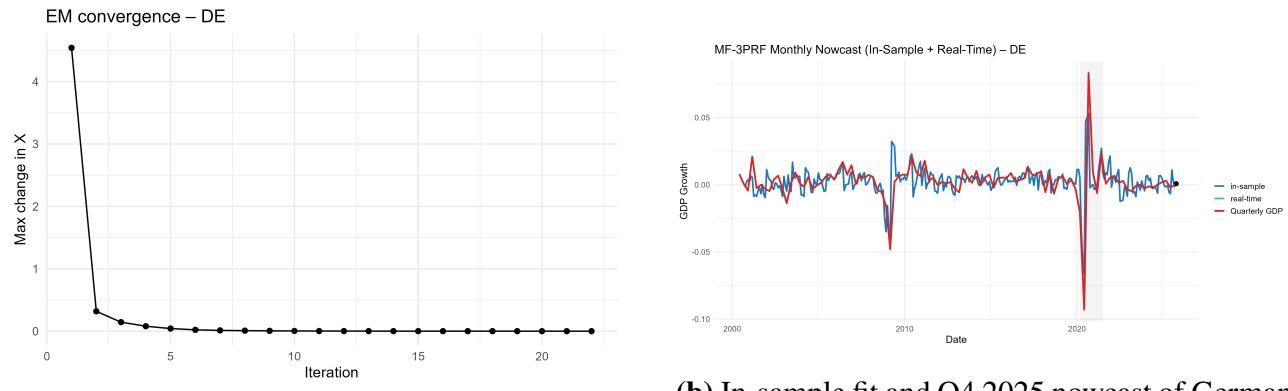


Figure 5: Germany: EM convergence (panel a) and in-sample fit with Q4 2025 GDP nowcast (panel b).

Table 7: MF-3PRF RMSFE by period – Germany

Period	M1	M2	M3
Full sample	0.0145	0.0114	0.0102
Pre-COVID	0.0063	0.0032	0.0038
COVID period	0.0316	0.0242	0.0200
Post-COVID	0.0072	0.0072	0.0078

Expanding-window nowcasts

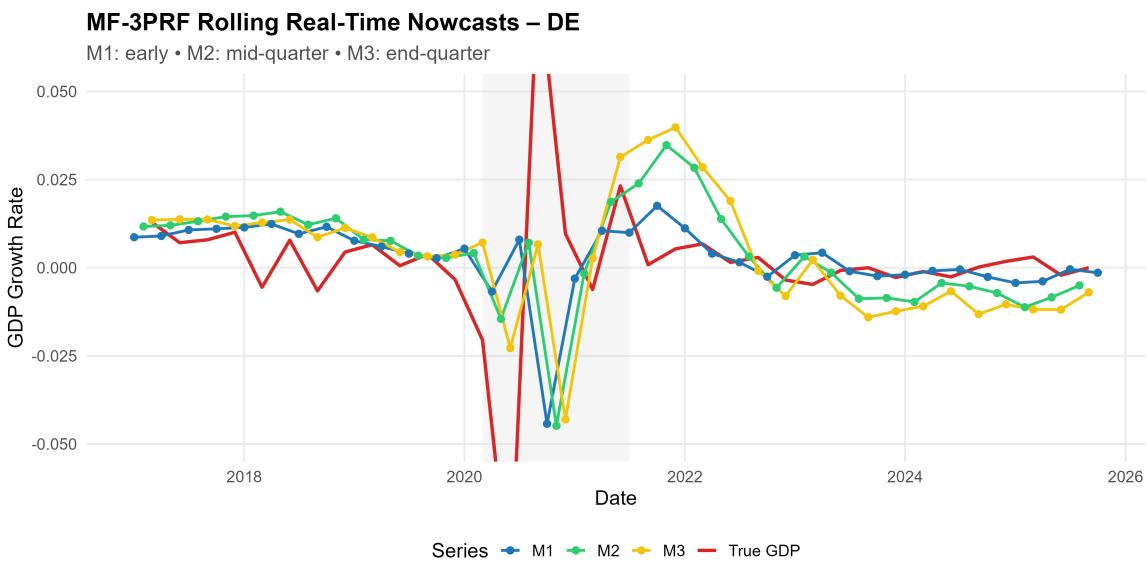


Figure 6: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for German GDP.

Table 8: MF-3PRF expanding-window RMSFE by period – Germany

Period	M1	M2	M3
Full sample	0.0227	0.0235	0.0237
Pre-COVID	0.0078	0.0095	0.0082
COVID period	0.0530	0.0509	0.0491
Post-COVID	0.0051	0.0125	0.0162

1.3.4 ES - Spain

In-sample fit and Q4 2025 GDP nowcast

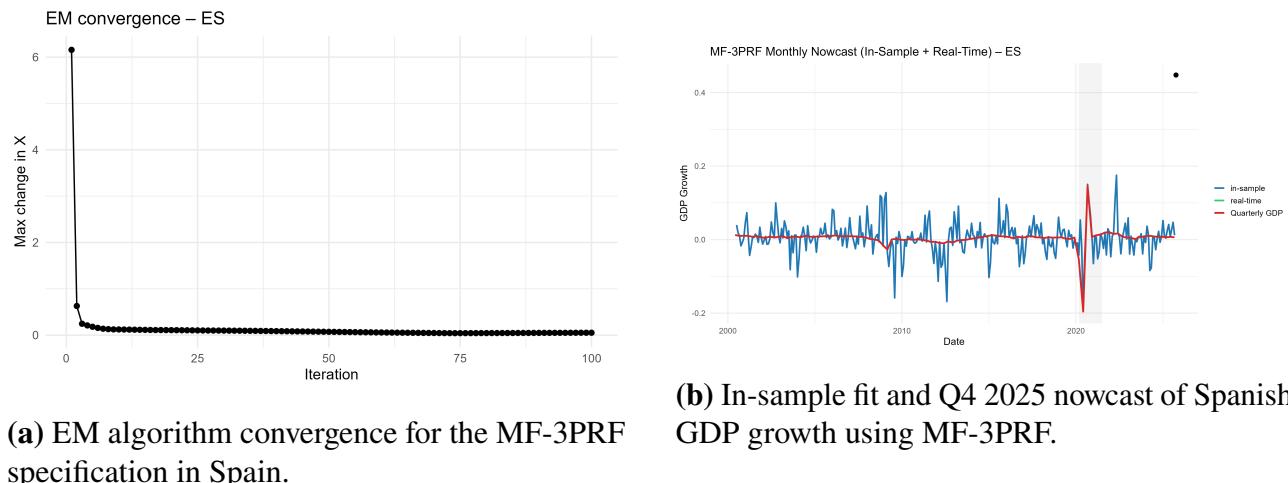


Figure 7: Spain: EM convergence (panel a) and in-sample fit with Q4 2025 GDP nowcast (panel b).

Table 9: MF-3PRF RMSFE by period – Spain

Period	M1	M2	M3
Full sample	0.0690	0.0559	0.0043
Pre-COVID	0.0369	0.0362	0.0035
COVID period	0.1409	0.0881	0.0077
Post-COVID	0.0428	0.0526	0.0027

Expanding-window nowcasts

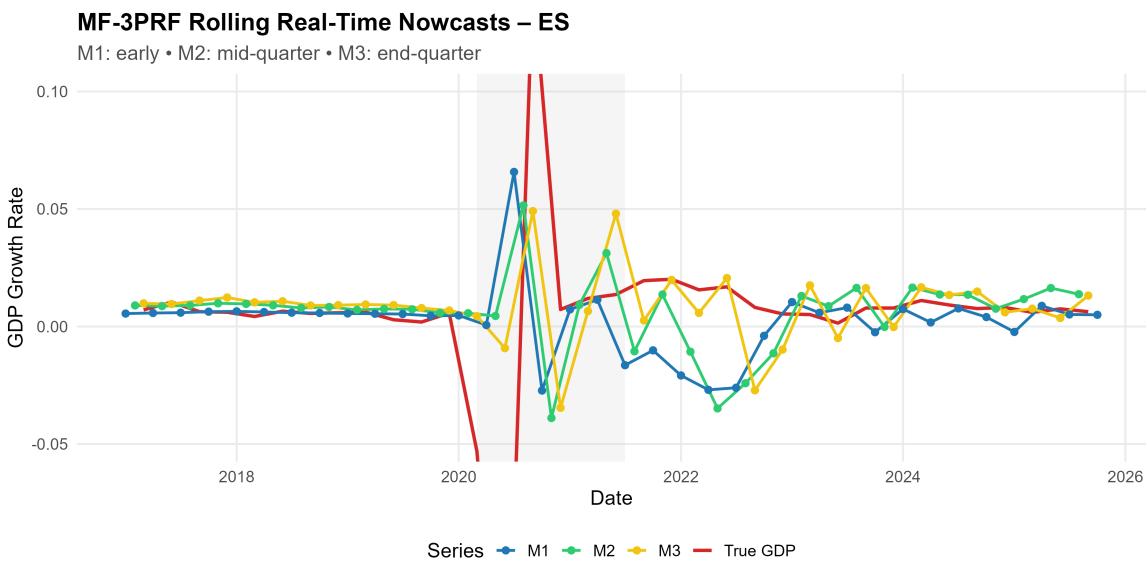


Figure 8: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for Spanish GDP.

Overall, this specification delivers a weaker fit for Spain than for the other Euro area countries. However, when additional predictors are included, the EM algorithm converges quickly and the nowcasts improve substantially. An alternative specification for Spain is reported in the Appendix.¹¹

Table 10: MF-3PRF expanding-window RMSFE by period – Spain

Period	M1	M2	M3
Full sample	0.0552	0.0655	0.0500
Pre-COVID	0.0023	0.0037	0.0050
COVID period	0.0971	0.1188	0.1008
Post-COVID	0.0542	0.0620	0.0393

¹¹This pattern suggests a potential advantage of tensor factor models. In vector models, each country must be tuned separately over a large grid of specifications, whereas matrix/tensor models exploit cross-country information and can achieve a more coherent fit across economies.

2 Matrix Models

3 Tensor-Based TPRF

Here we extend the classical Three-Pass Regression Filter to the case where predictors are *matrices* (e.g. country \times variable) observed at mixed frequencies over time.¹² This approach preserves the proxy-based targeting idea while exploiting the tensor structure of the data.

3.1 T-MF-TPRF

- **Step 0: Completion of matrix-valued predictors**

- As discussed above, monthly predictors are matrices (e.g. countries \times indicators) affected by:
 - * ragged edges,
 - * publication delays,
 - * blocks of missing values due to COVID masking.
- We vectorize each matrix and apply a static-factor EM algorithm in order to:
 - * estimate common factors and loadings,
 - * impute missing entries in a model-consistent way,
 - * obtain a completed monthly panel of matrix-valued predictors.
- After reconstruction, we reshape the completed data back into its original tensor form ($T \times P_1 \times P_2$).

- **Step 1: Proxy-based targeting in the matrix domain**

- Given a scalar proxy (e.g. GDP averaged across countries), we construct proxy-weighted moment matrices between this proxy and the completed predictors.
- For each proxy, we compute a truncated SVD of these moment matrices to extract a low-rank *bilinear* pattern:
 - * left singular vectors \Rightarrow loadings over the P_1 dimension (e.g. countries),
 - * right singular vectors \Rightarrow loadings over the P_2 dimension (e.g. indicators).
- This is the matrix analogue of the targeting step in the standard TPRF.

- **Step 2: Tensor factors preserving matrix structure**

- Instead of stacking everything into a single long vector, we keep the data in matrix form and extract factors directly as *matrix factors*. This preserves:
 - * the cross-country structure,
 - * the cross-indicator structure.
- This tensor formulation avoids losing information about the two-dimensional geometry of the predictors.

¹²This structure can be represented as a tensor of dimension $T \times P_1 \times P_2$ during empirics.

- **Step 3: Mixed-frequency forecasting via U–MIDAS**

- Tensor factors in Step 2 are naturally estimated at monthly frequency directly from the matrices. At this stage, the tensor can be vectorized for convenience in the regression step.
- For each quarter, we organise the monthly factors into three components

$$(F_\tau^{(1)}, F_\tau^{(2)}, F_\tau^{(3)}),$$

corresponding to the three months of quarter τ .

- These monthly tensor factors enter a U–MIDAS regression for the quarterly target (e.g. GDP), preserving the mixed-frequency nature:
 - * the regressors retain their monthly information content,
 - * the target is quarterly and linked through MIDAS lags.

Through this approach, we respect the tensor (matrix) structure of the data, handle missingness from ragged edges and COVID masking in a unified way, and, most importantly, retain the proxy-driven targeting philosophy of the original TPRF.

3.1.1 Step 1 of the Tensor 3PRF: Single Proxy and Multiple Proxies

Step 1 of the Tensor Mixed-Frequency Three-Pass Regression Filter (T-3PRF) is formalised in this section, beginning with the simplest case of a single proxy ($L = 1$) and then extend the formulation to an arbitrary number of proxies ($L > 1$). Throughout, quarterly observations are indexed by $\tau = 1, \dots, T_q$, and each quarterly observation of the mixed-frequency dataset is a matrix

$$X_\tau \in \mathbb{R}^{P_1 \times P_2},$$

where P_1 denotes the number of cross-sectional units (e.g. countries) and P_2 the number of variables per unit.

Case $L = 1$: Regression on a single proxy

Let the proxy be a scalar time series $Z_\tau \in \mathbb{R}$. Step 1 of the T-3PRF specifies the matrix regression model

$$X_\tau = Z_\tau AB^\top + U_\tau, \quad A \in \mathbb{R}^{P_1 \times r}, \quad B \in \mathbb{R}^{P_2 \times r}, \quad (1)$$

where AB^\top is a rank- r coefficient matrix. The parameters (A, B) are obtained by solving the least-squares problem

$$(A, B) = \arg \min_{\text{rank}(AB^\top) \leq r} \sum_{\tau=1}^{T_q} \|X_\tau - Z_\tau AB^\top\|_F^2. \quad (2)$$

The objective (2) expands as

$$\sum_{\tau} \|X_\tau\|_F^2 - 2 \sum_{\tau} Z_\tau \langle X_\tau, AB^\top \rangle + \sum_{\tau} Z_\tau^2 \|AB^\top\|_F^2.$$

Differentiating with respect to AB^\top yields the normal equation

$$\sum_{\tau=1}^{T_q} Z_\tau X_\tau = \left(\sum_{\tau=1}^{T_q} Z_\tau^2 \right) AB^\top. \quad (3)$$

Define the weighted moments

$$M \equiv \sum_{\tau=1}^{T_q} Z_\tau X_\tau, \quad S \equiv \sum_{\tau=1}^{T_q} Z_\tau^2.$$

Then the least-squares solution satisfies

$$AB^\top = S^{-1}M. \quad (4)$$

Since multiplying a matrix by a scalar does not affect its rank- r structure, the entire information about the targeted component AB^\top is contained in

$$M = \sum_{\tau=1}^{T_q} Z_\tau X_\tau \in \mathbb{R}^{P_1 \times P_2}. \quad (5)$$

Thus, M is a sufficient statistic for estimating the rank- r targeted pattern.

We now have a connection with low-rank approximation since the constrained problem (2) is equivalent to

$$\min_{\text{rank}(C) \leq r} \|M - C\|_F^2, \quad C = AB^\top S, \quad (6)$$

because (1) implies that the least-squares fit depends on X_τ only through their aggregated moment M . Therefore, Step 1 reduces to finding the *best rank- r approximation* of M .

By the Eckart–Young–Mirsky theorem, the solution is given by the truncated SVD:

$$M = U\Sigma V^\top, \quad A = U[:, 1:r], \quad B = V[:, 1:r]. \quad (7)$$

The SVD provides the unique (up to rotation) minimiser of (6) and hence of the original regression problem (2). In other words, performing least squares under a rank constraint is *mathematically equivalent to computing the best rank- r SVD approximation of M* , from where we can finally estimate AB^\top

Case $L > 1$: Regression on multiple proxies

Let the proxy vector at quarter τ be

$$z_\tau = \begin{pmatrix} Z_\tau^{(0)} \\ \vdots \\ Z_\tau^{(L-1)} \end{pmatrix} \in \mathbb{R}^L,$$

and collect all proxies across time into

$$Z = \begin{pmatrix} z_1^\top \\ \vdots \\ z_{T_q}^\top \end{pmatrix} \in \mathbb{R}^{T_q \times L}.$$

The Step 1 regression generalises to

$$X_\tau = \sum_{\ell=0}^{L-1} Z_\tau^{(\ell)} A^{(\ell)} B^{(\ell)\top} + U_\tau, \quad A^{(\ell)} \in \mathbb{R}^{P_1 \times r}, B^{(\ell)} \in \mathbb{R}^{P_2 \times r}. \quad (1)$$

Each proxy $Z^{(\ell)}$ targets its own rank- r bilinear pattern $A^{(\ell)} B^{(\ell)\top}$, in direct analogy with the multiple-proxy structure of the classical 3PRF.

The least-squares criterion is

$$\min_{\{A^{(\ell)}, B^{(\ell)}\}} \sum_{\tau=1}^{T_q} \left\| X_\tau - \sum_{\ell=0}^{L-1} Z_\tau^{(\ell)} A^{(\ell)} B^{(\ell)\top} \right\|_F^2. \quad (2)$$

First we need to specify the partial residual equation. When the parameters $\{A^{(j)}, B^{(j)}\}_{j \neq \ell}$ are held fixed, the subproblem for proxy ℓ reduces to a single-regressor matrix regression. Define the partial residual

$$R_\tau^{(\ell)} = X_\tau - \sum_{j \neq \ell} Z_\tau^{(j)} A^{(j)} B^{(j)\top}.$$

Then the subproblem satisfies

$$R_\tau^{(\ell)} = Z_\tau^{(\ell)} A^{(\ell)} B^{(\ell)\top} + U_\tau^{(\ell)}. \quad (*)$$

The normal equation is identical to the previous case. By solving $(*)$ using least squares we gain the normal equation

$$\sum_{\tau=1}^{T_q} Z_\tau^{(\ell)} R_\tau^{(\ell)} = \left(\sum_{\tau=1}^{T_q} (Z_\tau^{(\ell)})^2 \right) A^{(\ell)} B^{(\ell)\top}.$$

Define the proxy-specific weighted moment

$$M^{(\ell)} \equiv \sum_{\tau=1}^{T_q} Z_\tau^{(\ell)} R_\tau^{(\ell)} \in \mathbb{R}^{P_1 \times P_2}, \quad (3)$$

and the scalar weight

$$S^{(\ell)} \equiv \sum_{\tau=1}^{T_q} (Z_\tau^{(\ell)})^2.$$

Then the least-squares solution must satisfy

$$A^{(\ell)} B^{(\ell)\top} = (S^{(\ell)})^{-1} M^{(\ell)}.$$

Since multiplication by the scalar $(S^{(\ell)})^{-1}$ does not affect the rank- r structure, the matrix $M^{(\ell)}$ is a

sufficient statistic for the bilinear pattern targeted by proxy ℓ .

Again, the connection with low-rank approximation is straightforward. The constrained least-squares problem for proxy ℓ is equivalent to

$$\min_{\text{rank}(C) \leq r} \|M^{(\ell)} - C\|_F^2, \quad C = S^{(\ell)} A^{(\ell)} B^{(\ell)\top}.$$

Thus, for each proxy, Step 1 reduces to computing the best rank- r approximation of the weighted moment matrix $M^{(\ell)}$. By the Eckart–Young–Mirsky theorem, the optimal solution is its truncated SVD:

$$M^{(\ell)} = U^{(\ell)} \Sigma^{(\ell)} V^{(\ell)\top}, \quad A^{(\ell)} = U_{[:,1:r]}^{(\ell)}, \quad B^{(\ell)} = V_{[:,1:r]}^{(\ell)}.$$

Interpretation. Each proxy $Z^{(\ell)}$ extracts a different “targeted” direction in the space of matrix-valued observations. The matrices $A^{(\ell)}$ contains cross-sectional patterns across the P_1 units, while $B^{(\ell)}$ encode cross-variable patterns across the P_2 features. Despite the matrix structure of X_τ , Step 1 remains a *time-series regression*, where the truncated SVD provides the best rank- r bilinear approximation to the proxy-weighted moment matrix $M^{(\ell)}$.

3.1.2 Step 2: Estimation of Monthly Tensor Factors

Let $X_t \in \mathbb{R}^{P_1 \times P_2}$ denote the high-frequency (monthly) observation of the matrix-valued dataset at time t , and let $\{A^{(\ell)}, B^{(\ell)}\}_{\ell=1}^L$ be the loading matrices obtained in Step 1 from the L targeted proxies. For each proxy ℓ , the tensor regression model takes the form

$$X_t = A^{(\ell)} F_t^{(\ell)} B^{(\ell)\top} + U_t^{(\ell)}, \quad A^{(\ell)} \in \mathbb{R}^{P_1 \times r_1}, \quad B^{(\ell)} \in \mathbb{R}^{P_2 \times r_2}, \quad (4)$$

where $F_t^{(\ell)} \in \mathbb{R}^{r_1 \times r_2}$ is the tensor factor associated with proxy ℓ at time t , and $U_t^{(\ell)}$ is an idiosyncratic error matrix.

Pass 2 is a cross-sectional tensor regression. In analogy with the classical TPRF of Hegenstrick and Marcellino (2016), the monthly factors are obtained by estimating, for each time period t and for each proxy ℓ , the coefficient matrix $F_t^{(\ell)}$ that minimizes the Frobenius-norm loss:

$$\hat{F}_t^{(\ell)} = \arg \min_{F \in \mathbb{R}^{r_1 \times r_2}} \|X_t - A^{(\ell)} F B^{(\ell)\top}\|_F^2. \quad (5)$$

Problem (5) admits a closed-form solution, given by a doubly-projected bilinear least-squares estimator:

$$\hat{F}_t^{(\ell)} = (A^{(\ell)\top} A^{(\ell)})^{-1} A^{(\ell)\top} X_t B^{(\ell)} (B^{(\ell)\top} B^{(\ell)})^{-1}. \quad (6)$$

This expression is the tensor analogue of the cross-sectional regression used in the vector-valued TPRF. The estimator performs a left-projection of X_t onto the row-space of $A^{(\ell)}$ and a right-projection onto the column-space of $B^{(\ell)}$, thus producing an $r_1 \times r_2$ matrix of monthly latent factors.

For each time t , the full set of tensor factors is composed by Stacking across proxies and it is represented by:

$$\widehat{F}_t = \{\widehat{F}_t^{(1)}, \widehat{F}_t^{(2)}, \dots, \widehat{F}_t^{(L)}\}.$$

In subsequent steps of the Mixed-Frequency TPRF, these matrices can be vectorised and concatenated as $\text{vec}(\widehat{F}_t^{(\ell)}) \in \mathbb{R}^{r_1 r_2}$ for use in the MIDAS regression of quarterly GDP. Indeed, after this step we have achieved the goal of extracting factors and loadings from a matrix, that was our purpose from the beginning.

3.1.3 Step 3: Quarterly U-MIDAS Regression and Real-Time Nowcasting

Let y_τ denote the quarterly target variable (e.g., GDP growth), observed at quarterly dates $\tau = 1, \dots, T_q$. From Step 2, for each proxy $\ell = 1, \dots, L$ and for each month $m = 1, 2, 3$ within quarter τ , we obtain a matrix-valued monthly factor $F_{m,\tau}^{(\ell)} \in \mathbb{R}^{k_1 \times k_2}$. In order to use these components in a linear quarterly regression, we apply the vectorisation operator:

$$f_{m,\tau}^{(\ell)} \equiv \text{vec}(F_{m,\tau}^{(\ell)}) \in \mathbb{R}^{k_1 k_2}.$$

To construct the MIDAS regressors, let K denote the maximum number of quarterly MIDAS lags. For each quarter τ , the full regressor vector is formed by stacking all vectorised monthly tensor factors and their lags:

$$X_\tau = [f_{1,\tau}^{(1)}, f_{2,\tau}^{(1)}, f_{3,\tau}^{(1)}, \dots, f_{1,\tau}^{(L)}, f_{2,\tau}^{(L)}, f_{3,\tau}^{(L)}, f_{1,\tau-1}^{(1)}, \dots, f_{3,\tau-K+1}^{(L)}].$$

Hence, $X_\tau \in \mathbb{R}^{3LK \cdot (k_1 k_2)}$. ¹³

3.1.4 Nowcasting Structure

From this point onwards it is possible to run two practical exercises: forecasting and nowcasting.

U-MIDAS forecasting regression. Following the classical TPRF logic of Hepenstrick and Marcellino (2016), the quarterly forecasting step is performed using the *one-step-ahead* regression

$$y_{\tau+1} = \beta_0 + X_\tau^\top \beta + \varepsilon_{\tau+1}, \quad \tau = K, \dots, T_q - 1, \tag{7}$$

where $\beta_0 \in \mathbb{R}$ and $\beta \in \mathbb{R}^{3LK(k_1 k_2)}$ are estimated by ordinary least squares. Equation (7) is the tensor-valued analogue of the U-MIDAS regression used in the vector-based Mixed-Frequency TPRF.

Real-time nowcasting In real time, the quarter $\tau + 1$ is observed only partially. Let $M_\tau \in \{1, 2, 3\}$ denote the number of monthly observations currently available for quarter $\tau + 1$. For each proxy ℓ , the

¹³Although the dimension increases with the number of proxies, lags and tensor modes, macroeconomic datasets typically exhibit very low intrinsic rank. In practice, L is small, quarterly lags rarely exceed $K = 1$ or $K = 2$, and tensor ranks ($k_1 k_2$) are extremely modest, preventing any explosion in the size of the regressor.

corresponding available factors are $F_{m,\tau+1}^{(\ell)}$ for $m = 1, \dots, M_\tau$. We construct the truncated regressor vector

$$X_{\tau+1}^{(\text{rt})} = \left[f_{1,\tau+1}^{(1)}, \dots, f_{M_\tau,\tau+1}^{(1)}, \dots, f_{1,\tau+1}^{(L)}, \dots, f_{M_\tau,\tau+1}^{(L)} \right],$$

which includes only the information available in real time.

The real-time nowcast of $y_{\tau+1}$ is then given by

$$\hat{y}_{\tau+1}^{(\text{rt})} = \beta_0 + (X_{\tau+1}^{(\text{rt})})^\top \beta, \quad (8)$$

where the unavailable factor blocks (months $m > M_\tau$) are omitted from the regressor. Equation (8) yields three nested nowcasts:

- **first-month nowcast** ($M_\tau = 1$): uses only $f_{1,\tau+1}^{(\ell)}$;
- **second-month nowcast** ($M_\tau = 2$): uses $f_{1,\tau+1}^{(\ell)}$ and $f_{2,\tau+1}^{(\ell)}$;
- **full-quarter nowcast** ($M_\tau = 3$): uses all three monthly factor blocks.

To summarise, Step 3 extends the classical U-MIDAS forecasting regression to the tensor-valued factor representation obtained in Steps 1 and 2, while maintaining the real-time structure of the Mixed-Frequency TPRF.

3.1.5 EM Algorithm under a Tensor-Valued Predictor Structure

In the mixed-frequency setting, each monthly predictor is a matrix

$$X_t \in \mathbb{R}^{P_1 \times P_2}, \quad t = 1, \dots, T,$$

typically featuring arbitrary missing entries induced by ragged edges, different publication calendars, or variables masked during the COVID period. Although the Tensor-TPRF model preserves a bilinear structure for the *targeted regression* step, the EM algorithm used to impute the missing predictor data is applied in the *vectorized* domain, following the standard procedure for static-factor models.

Vectorization. We define the vector representation

$$x_t = \text{vec}(X_t) \in \mathbb{R}^N, \quad N = P_1 P_2,$$

obtained by stacking the columns of X_t . Collecting all observations yields the $T \times N$ panel

$$X_{\text{vec}} = \begin{pmatrix} x_1^\top \\ \vdots \\ x_T^\top \end{pmatrix}.$$

Vectorization does not impose any structural constraint nor does it interfere with the bilinear representation used later in the T-TPRF. Its purpose is simply to map matrix-valued predictors into the N -dimensional setting required by the classical EM algorithm for static-factor models.

Selection matrices. For each series $i = 1, \dots, N$, let A_i denote the binary selection matrix that extracts the observed entries of $x_{i,t}$:

$$x_{i,\text{obs}} = A_i x_i, \quad A_i \in \{0, 1\}^{m_i \times T}.$$

This representation accommodates arbitrary missing patterns, including mixed-frequency observations and ragged edges.

E-step. Given factor estimates $\widehat{F}^{(j-1)}$ and loadings $\widehat{\Phi}^{(j-1)}$, the completed value of series i at iteration j is

$$\widehat{x}_i^{(j)} = \widehat{F}^{(j-1)} \widehat{\Phi}_i^{(j-1)} + A_i (A_i' A_i)^{-1} \left(x_{i,\text{obs}} - A_i \widehat{F}^{(j-1)} \widehat{\Phi}_i^{(j-1)} \right).$$

This formula replaces missing entries with their projection onto the factor space, adjusted to match exactly the observed data.

M-step (PCA update). Once the completed dataset

$$\widehat{X}^{(j)} = (\widehat{x}_1^{(j)}, \dots, \widehat{x}_N^{(j)})$$

is available, the factors are updated via principal components:

$$\widehat{F}^{(j)} = \text{PC}_r(\widehat{X}^{(j)}), \quad \widehat{\Phi}^{(j)} = (\widehat{F}^{(j)}) \widehat{X}^{(j)},$$

where $\text{PC}_r(\cdot)$ denotes the rank- r principal components.

Reconstruction of tensor-valued predictors. Finally, the completed predictors are reshaped back to their tensor form:

$$\widehat{X}_t = \text{mat}(\widehat{x}_t, P_1, P_2), \quad t = 1, \dots, T.$$

The Tensor-TPRF estimation then operates on the reconstructed matrices \widehat{X}_t in Step 1, where the bilinear low-rank constraints are imposed through the truncated SVD of the proxy-weighted moment matrices $M^{(\ell)}$.

References

- Ahn, Seung C and Alex R Horenstein (2013). “Eigenvalue ratio test for the number of factors”. In: *Econometrica* 81.3, pp. 1203–1227.
- Bai, Jushan and Serena Ng (2002). “Determining the number of factors in approximate factor models”. In: *Econometrica* 70.1, pp. 191–221.
- (2008). “Forecasting economic time series using targeted predictors”. In: *Journal of Econometrics* 146.2, pp. 304–317.
- Bańbura, Marta and Michele Modugno (2014). “Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data”. In: *Journal of applied econometrics* 29.1, pp. 133–160.
- Barigozzi, M. and C. Lissone (2024). *EA-MD-QD: Large Euro Area and Euro Member Countries Datasets for Macroeconomic Research (Version 12.2023)*. Data set. URL: <https://doi.org/10.5281/ZENODO.10514668>.
- Barigozzi, Matteo and Luca Trapin (2025). “Quasi maximum likelihood estimation of high-dimensional approximate dynamic matrix factor models via the EM algorithm”. In: *arXiv preprint arXiv:2502.04112*.
- Cen, Zetai and Clifford Lam (2025). “Tensor time series imputation through tensor factor modelling”. In: *Journal of Econometrics* 249, p. 105974.
- De Mol, Christine, Domenico Giannone, and Lucrezia Reichlin (2008). “Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?” In: *Journal of Econometrics* 146.2, pp. 318–328.
- Hepenstrick, Christian and Massimiliano Marcellino (2016). *Forecasting with Large Unbalanced Datasets: The Mixed Frequency Three-Pass Regression Filter*. Tech. rep. Swiss National Bank.
- Kelly, Bryan and Seth Pruitt (2015). “The three-pass regression filter: A new approach to forecasting using many predictors”. In: *Journal of Econometrics* 186.2, pp. 294–316.
- Xiong, Ruoxuan and Markus Pelger (2023). “Large dimensional latent factor modeling with missing observations and applications to causal inference”. In: *Journal of Econometrics* 233.1, pp. 271–301.
- Yu, Long et al. (2022). “Projected estimation for large-dimensional matrix factor models”. In: *Journal of Econometrics* 229.1, pp. 201–217.

A Additional Results Accounting for Alternative Specifications

Just to provide a different example for Spain (the most problematic Country), and France (the most proficient), here we consider exactly the same specification described during the parameter's selection discussion, but imposing manually a set of 38 Monthly Variables and 10 Quarterly Variables.

A.1 ES – Spain

In-sample fit and Q4 2025 GDP nowcast (alternative specification)

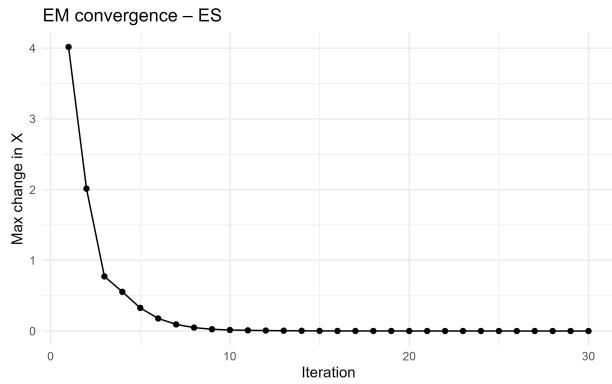


Figure 9: EM algorithm convergence for the alternative MF-3PRF specification in Spain.

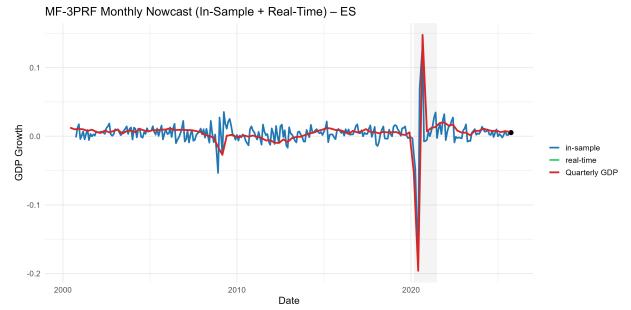


Figure 10: In-sample fit and Q4 2025 nowcast of Spanish GDP growth (alternative specification).

Table 11: MF-3PRF RMSFE by period – Spain (alternative specification)

Period	M1	M2	M3
Full sample	0.0312	0.0187	0.0107
Pre-COVID	0.0080	0.0082	0.0083
COVID period	0.0722	0.0425	0.0201
Post-COVID	0.0108	0.0058	0.0065

Expanding-window nowcasts (alternative specification)

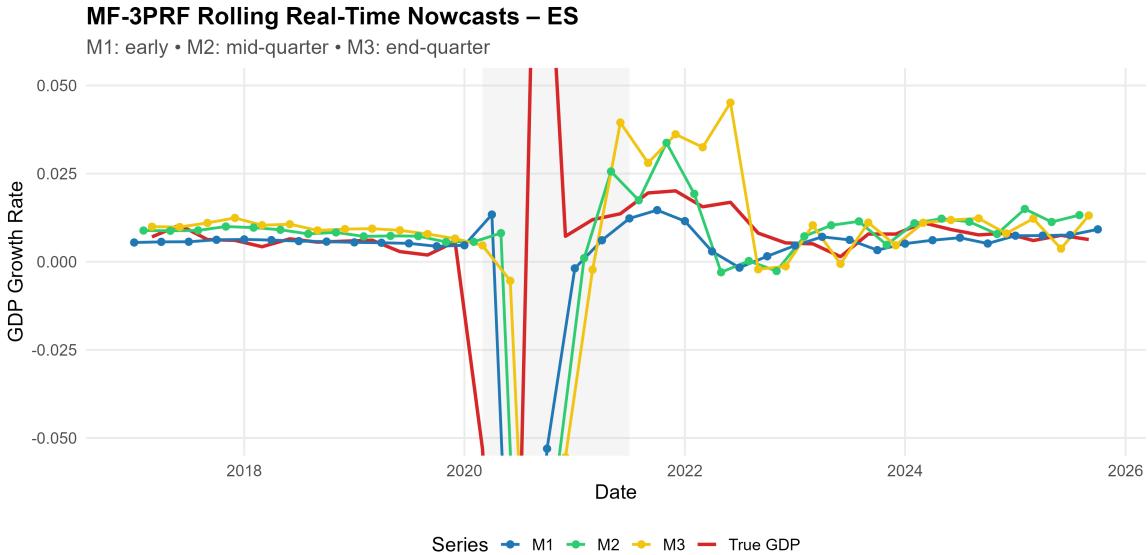


Figure 11: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for Spanish GDP (alternative specification).

Table 12: MF-3PRF expanding-window RMSFE by period – Spain (alternative specification)

Period	M1	M2	M3
Full sample	0.0660	0.0666	0.0655
Pre-COVID	0.0020	0.0032	0.0044
COVID period	0.1592	0.1602	0.1571
Post-COVID	0.0055	0.0077	0.0101

A.2 FR – France

In-sample fit and Q4 2025 GDP nowcast (alternative specification)

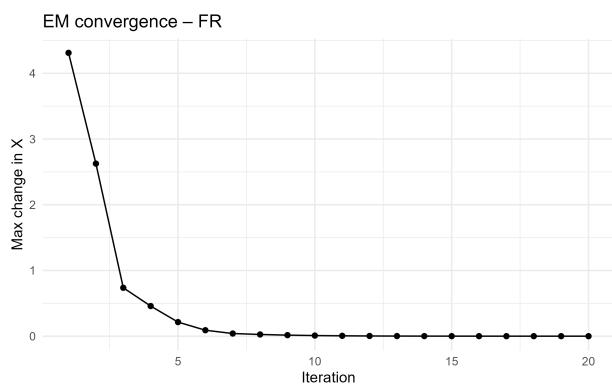


Figure 12: EM algorithm convergence for the alternative MF-3PRF specification in France.

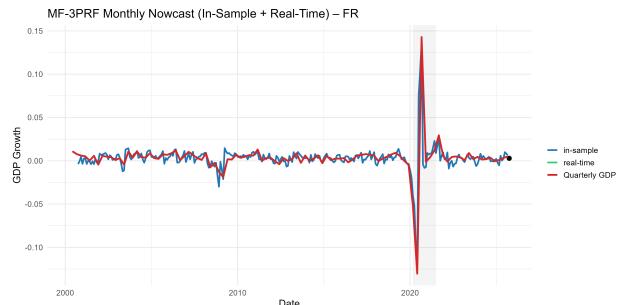


Figure 13: In-sample fit and Q4 2025 nowcast of French GDP growth (alternative specification).

Expanding-window nowcasts (alternative specification)

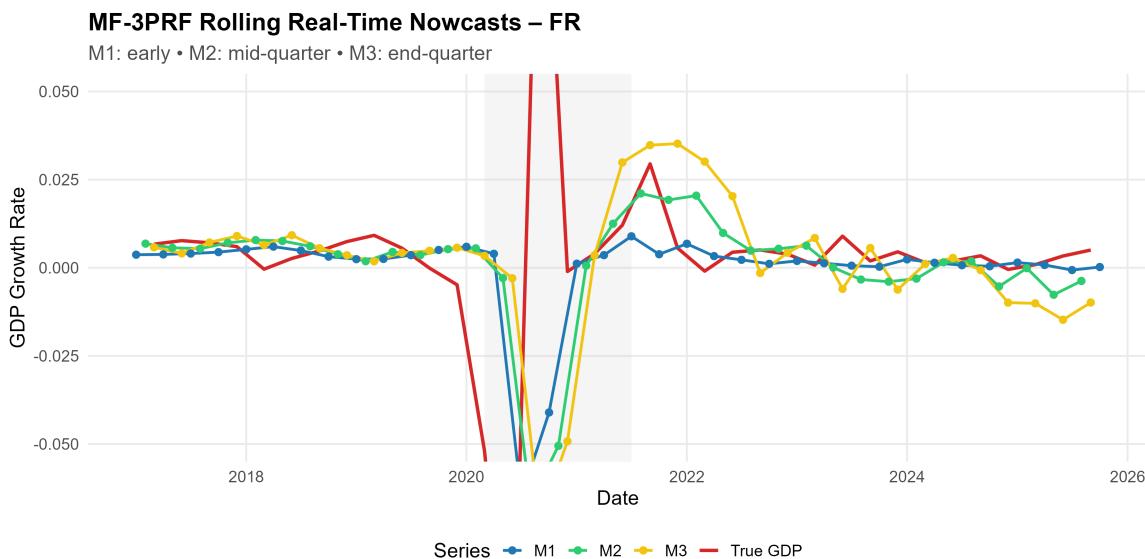


Figure 14: Expanding-window pseudo real-time nowcasts (M1, M2, M3) for French GDP (alternative specification).