Nowcasting Macroeconomic Variables with a Sparse Mixed Frequency Dynamic Factor Model
–Supplementary Material–

Domenic Franjic[*] Karsten Schweikert[†]

University of Hohenheim University of Hohenheim

[Latest update: October 30, 2024]

A Additional Simulation Results

	$s = 0.00$				$s = 0.80$			
	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$
$R = 1$	-0.424	-0.349	-0.059	-0.188	-0.005	-0.004	-0.033	-0.046
$R = 2$	-0.472	-0.406	-0.160	-0.057	-0.009	-0.009	-0.048	-0.002
$R = 3$	-0.483	-0.285	-0.064	0.022	-0.321	-0.212	-0.375	0.039
$R = 4$	-0.211	-0.080	-0.251	-0.199	-0.207	-0.078	-0.004	-0.065

Note: ρ refers to the average absolute correlation-coefficients of the measurement errors. MSNE reductions are computed as $1-S$, where S is given by $S = \sum_{i=1}^{500}(\hat{x}_{T,0,i,\text{SDFM}} - x_{T,0,i})^2 / \sum_{i=1}^{500}(\hat{x}_{T,0,i,\text{DFM}} - x_{T,0,i})^2$, where $\hat{x}_{T,0,i,\text{SDFM}}$ represents the ith nowcast of the SDFM model, $\hat{x}_{T,0,i,\text{DFM}}$ represents the ith nowcast of the benchmark DFM model, and $x_{T,0,i}$ represents the realisation of the variable of interest at time point T.

Table 1: MSNE reduction over 1000 simulated nowcasting exercises restricted to a single penalty for all factors using $T = 100$ observations for $N = 50$ variables

[*]Address: University of Hohenheim, Core Facility Hohenheim & Institute of Economics, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany, e-mail: *franjic@uni-hohenheim.de*

[†]Address: University of Hohenheim, Core Facility Hohenheim & Institute of Economics, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany, e-mail: *karsten.schweikert@uni-hohenheim.de*

	$s = 0.00$				$s = 0.80$			
	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$
$R = 1$	0.411	0.415	0.481	0.486	0.513	0.491	0.473	0.450
$R = 2$	0.403	0.404	0.516	0.509	0.490	0.497	0.527	0.505
$R = 3$	0.375	0.436	0.480	0.501	0.412	0.420	0.484	0.523
$R = 4$	0.433	0.452	0.479	0.515	0.447	0.470	0.491	0.522

Note: The SDFM model hyper-parameters are validated optimising the BIC. ρ refers to the average absolute correlation-coefficients of the measurement errors. The ratios of relative MSNE reductions are computed as $\frac{1}{500}\sum_{i=1}^{500} \mathbb{1}\left((\hat{x}_{T,0,i,\text{SDFM}} - x_{T,0,i})^2 < (\hat{x}_{T,0,i,\text{DFM}} - x_{T,0,i})^2\right)$, where $\hat{x}_{T,0,i,\text{SDFM}}$ represents the ith nowcast of the SDFM model, $\hat{x}_{T,0,i,\text{DFM}}$ represents the ith nowcast of the benchmark DFM model, and $x_{T,0,i}$ represents the realisation of the variable of interest at time point T.

Table 2: Ratio of relative MSNE reductions over 1000 simulated nowcasting exercises restricted to a single penalty for all factors using $T = 200$ observations for $N = 100$ variables

	$s = 0.00$				$s = 0.80$			
	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$
$R = 1$	-0.483	-0.225	-0.212	-0.034	0.017	-0.194	-0.008	-0.022
$R = 2$	-0.424	-0.049	-0.005	-0.001	-0.024	-0.081	-0.050	-0.109
$R = 3$	-0.413	0.010	-0.147	0.012	-0.493	-0.089	0.034	0.024
$R = 4$	-0.358	0.022	0.014	0.013	-0.313	0.038	0.037	-0.060

Note: The SDFM model hyper-parameters are validated optimising the BIC. The SDFM model hyper-parameters are validated optimising the BIC. ρ refers to the average absolute correlation-coefficients of the measurement errors. MSNE reductions are computed as $1 - S$, where S is given by $S = \sum_{i=1}^{500}(\hat{x}_{T,0,i,\text{SDFM}} - x_{T,0,i})^2 / \sum_{i=1}^{500}(\hat{x}_{T,0,i,\text{DFM}} - x_{T,0,i})^2$, where $\hat{x}_{T,0,i,\text{SDFM}}$ represents the ith nowcast of the SDFM model, $\hat{x}_{T,0,i,\text{DFM}}$ represents the ith nowcast of the benchmark DFM model, and $x_{T,0,i}$ represents the realisation of the variable of interest at time point T.

Table 3: MSNE reduction over 1000 simulated nowcasting exercises restricted to a single penalty for all factors using $T = 200$ observations for $N = 100$ variables

	$s = 0.00$				$s = 0.80$			
	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$	$\bar{\rho} \approx 0.2$	$\bar{\rho} \approx 0.3$	$\bar{\rho} \approx 0.4$	$\bar{\rho} \approx 0.5$
$R = 1$	0.432	0.486	0.479	0.496	0.518	0.493	0.490	0.456
$R = 2$	0.399	0.510	0.517	0.478	0.486	0.523	0.520	0.523
$R = 3$	0.435	0.528	0.551	0.510	0.373	0.542	0.550	0.510
$R = 4$	0.432	0.532	0.538	0.489	0.432	0.544	0.545	0.523

Note: The SDFM model hyper-parameters are validated optimising the BIC. ρ refers to the average absolute correlation-coefficients of the measurement errors. The ratios of relative MSNE reductions are computed as $\frac{1}{500}\sum_{i=1}^{500} \mathbb{1}\left((\hat{x}_{T,0,i,\text{SDFM}} - x_{T,0,i})^2 < (\hat{x}_{T,0,i,\text{DFM}} - x_{T,0,i})^2\right)$, where $\hat{x}_{T,0,i,\text{SDFM}}$ represents the ith nowcast of the SDFM model, $\hat{x}_{T,0,i,\text{DFM}}$ represents the ith nowcast of the benchmark DFM model, and $x_{T,0,i}$ represents the realisation of the variable of interest at time point T.

Table 4: Ratio of relative MSNE reductions over 500 simulated nowcasting exercises restricted to a single penalty for all factors using $T = 200$ observations for $N = 100$ Variables

B Algorithms

In the upcoming section, the following conventions are used. The symbol $\sqcup$ denotes the operation of adding an element to the end of a sequence. Let $\mathcal{S}$ be a sequence of integer values and $\mathbf{A} = (a_{n,m})_{n=1,m=1}^{N,M}$ be a matrix. Assuming $\min \mathcal{S} \geq 1$ and $\max \mathcal{S} \leq M$, $\mathbf{A}[\mathcal{S}, \cdot]$ refers to the submatrix of $\mathbf{A}$ that is constructed by concatenating the rows with index corresponding to the integers in $\mathcal{S}$ in order. For a vector $\mathbf{v} = (v_n)_{n=1}^{N}$, and $\min \mathcal{S} \geq 1$ and $\max \mathcal{S} \leq N$, $\mathbf{v}[\mathcal{S}]$ denotes the subvector consisting of the elements of $\mathbf{v}$ with corresponding index $n \in \mathcal{S}$. Similarly, $\mathbf{v}[n_1 : n_m]$ refers to the subvector consisting of the elements with index n_1 to index n_m. Adding an element v at the end of a vector is denoted as $[\mathbf{v}, v]$. Analogously, $[\mathbf{v}_1, \mathbf{v}_2]$ concatenates two vectors $\mathbf{v}_1, \mathbf{v}_2$. Initialising an empty vector is denoted as $\mathbf{v} = [\cdot]$. The operation of removing an element v_n from a vector $\mathbf{v}$ is defined as $]\mathbf{v}, v_n[:= (v_1, \ldots, v_{n-1}, v_{n+1}, \ldots, v_N)$. For two vectors $\mathbf{v}_1, \mathbf{v}_2$ of equal length, the element-wise division is defined as $\mathbf{v}_1 \oslash \mathbf{v}_2 = (v_{1,1}/v_{1,2}, \ldots, v_{n,1}/v_{n,2})$.

Algorithm 1 Sparse Principal Components Analysis (Zou and Hastie, 2020)

$\mathbf{X}' = \mathbf{UDV}'$
$\mathbf{A} \leftarrow (\mathbf{v}_1, \ldots, \mathbf{v}_R)$, where $\mathbf{v}_1, \ldots, \mathbf{v}_R$ correspond to the first R columns of $\mathbf{V}$.
Set $\tilde{\mathbf{A}}$ to a Matrix with entries equal to the double precision floating point maximum
Set conversion threshold $\epsilon > 0$
while $\|\mathbf{A} - \tilde{\mathbf{A}}\|_F > \epsilon$ **do**
    $\tilde{\mathbf{A}} \leftarrow \mathbf{A}$
    for $r \in \{1, \ldots, R\}$ **do**
        Use Algorithm 2 to solve

$$\widehat{\boldsymbol{\lambda}}_r \leftarrow \underset{\boldsymbol{\lambda}_r}{\operatorname{argmin}} \left\{ (\boldsymbol{\alpha}_r - \boldsymbol{\lambda}_r)' \mathbf{X}_\tau \mathbf{X}_\tau' (\boldsymbol{\alpha} - \boldsymbol{\lambda}_r) + \kappa_2 \|\boldsymbol{\lambda}_r\|_2 + \kappa_{1,r} \|\boldsymbol{\lambda}_r\|_1 \right\} \tag{B.1}$$

    end for
    $\widehat{\boldsymbol{\lambda}} \leftarrow (\widehat{\boldsymbol{\lambda}}_1, \ldots, \widehat{\boldsymbol{\lambda}}_R)$
    Compute the singular value decomposition of $\mathbf{XX}'\widehat{\boldsymbol{\lambda}}$, i.e., $\mathbf{XX}'\widehat{\boldsymbol{\lambda}} = \mathbf{UDV}'$.
    $\mathbf{A} \leftarrow \mathbf{UV}'$
end while
return $\widehat{\boldsymbol{\Lambda}} = \widehat{\boldsymbol{\lambda}}$

Algorithm 2 Least Angle Regression (Efron et al., 2004; Zou and Hastie, 2005, 2020)

Set N to the number of rows of $\mathbf{X}$
Set T to the number of columns of $\mathbf{X}$
$\mathbf{y} \leftarrow \mathbf{X}\boldsymbol{\alpha}_r$
$\mathbf{c} \leftarrow \left| \frac{1}{\sqrt{1+\kappa_2}} \mathbf{X}\mathbf{y} \right|$

$p \leftarrow \max \mathbf{c}$

$\mathcal{I} \leftarrow (1, \ldots, N)$

$\mathcal{A} \leftarrow \emptyset$

$d \leftarrow 0$

$\mathbf{s} \leftarrow [\cdot]$

$\widehat{\boldsymbol{\lambda}}_r \leftarrow \mathbf{0}_N$

$\widehat{\mathbf{q}} \leftarrow \mathbf{0}_N$

Set E to double precision floating point maximum

Set threshold values M for $0 < M \leq N$ and/or $\kappa_{1,r}$

while $2p\sqrt{1+\kappa_2} > \kappa_{1,r} \wedge |\mathcal{A}| < M$ **do**

  $\widehat{c} \leftarrow \max |\mathbf{x}_n \mathbf{y}|$ for $n \in \mathcal{I}$ and $\mathbf{x}_n \in \mathbf{X}$

  $\widehat{n} \leftarrow \text{argmax}|\mathbf{x}_n \mathbf{y}|$ for $n \in \mathcal{I}$ and $\mathbf{x}_n \in \mathbf{X}$

  if $d = 0 \wedge |\mathcal{A}| < N$ **then**

    $\mathcal{A} \leftarrow \mathcal{A} \sqcup \{\widehat{n}\}$

    $\mathcal{I} \leftarrow \mathcal{I} \backslash \{\widehat{n}\}$

    $\mathbf{s} \leftarrow [\mathbf{s}, \text{sign}(\widehat{c})]$

    if $|\mathcal{A}| = 1$ **then**

      $\mathbf{L} \leftarrow \sqrt{\mathbf{x}_n \mathbf{x}_n + \kappa_2}/(1+\kappa_2)$

    else

      Update the lower Cholesky matrix $\mathbf{L}$ by $\mathbf{x}_{\widehat{n}}$

    end if

  end if

  if $d = 1 \vee |\mathcal{A}| = N$ **then**

    $d \leftarrow 0$

  end if

  $\mathbf{g} \leftarrow \mathbf{s}$

  $\mathbf{g} \leftarrow \mathbf{v}_1$, where $\mathbf{v}_1$ is the solution to $\mathbf{L}\mathbf{v}_1 = \mathbf{g}$

  $\mathbf{g} \leftarrow \mathbf{v}_2$, where $\mathbf{v}_2$ is the solution to $\mathbf{L}'\mathbf{v}_2 = \mathbf{g}$

  $a \leftarrow 1/\sqrt{\mathbf{g}'\mathbf{s}}$

  $\mathbf{w} \leftarrow a\mathbf{g}$

  $\mathbf{u} \leftarrow \left[\frac{1}{\sqrt{1+\kappa_2}}\mathbf{X}[\mathcal{A}, \cdot]'\mathbf{w}, \frac{\sqrt{\kappa_2}}{\sqrt{1+\kappa_2}}\mathbf{w}\right]$

  $\boldsymbol{\gamma} \leftarrow -1 \cdot (\widehat{\boldsymbol{\lambda}}_r[\mathcal{A}] \oslash \mathbf{w})$

  if $0 < \max \boldsymbol{\gamma}$ **then**

    $\tilde{\gamma} \leftarrow \max \boldsymbol{\gamma}$

    $n \leftarrow \text{argmin}\gamma_i$ for $\gamma_i \in \boldsymbol{\gamma}$

  else

    $\tilde{\gamma} \leftarrow E$

  end if

$\widehat{\gamma} \leftarrow \widehat{c}/a$
if $|\mathcal{A}| < N$ **then**
$\boldsymbol{\alpha} \leftarrow \mathbf{X}[\mathcal{A}, \cdot]\mathbf{u}[1:T] + \sqrt{\kappa_2}\mathbf{u}[(T+1):|\mathcal{A}|]$
for $m \in (1, \ldots, |\mathcal{I}|)$ **do**
$h_1 \leftarrow (\widehat{c} - c_{i_m})/(a - \alpha_m)$, where $i_m \in \mathcal{I}$, $c_{i_m} \in \mathbf{c}$, and $\alpha_m \in \boldsymbol{\alpha}$
$h_2 \leftarrow (\widehat{c} + c_{i_m})/(a + \alpha_m)$, where $i_m \in \mathcal{I}$, $c_{i_m} \in \mathbf{c}$, and $\alpha_m \in \boldsymbol{\alpha}$
if $0 < h_1 < \widehat{\gamma}$ **then**
$\widehat{\gamma} \leftarrow h_1$
end if
if $0 < h_2 < \widehat{\gamma}$ **then**
$\widehat{\gamma} \leftarrow h_2$
end if
end for
end if
$\gamma^{\dagger} \leftarrow \min\{\tilde{\gamma}, \widehat{\gamma}\}$
$\boldsymbol{\lambda}^{\dagger} \leftarrow \widehat{\boldsymbol{\lambda}}_r$
$\widehat{\boldsymbol{\lambda}}_r \leftarrow \widehat{\boldsymbol{\lambda}}_r + \gamma^{\dagger}\mathbf{w}$
$\widehat{\mathbf{q}} \leftarrow \widehat{\mathbf{q}} - \gamma^{\dagger}\mathbf{u}$
$p^{\dagger} \leftarrow p - |\gamma a|$
if $2p^{\dagger}\sqrt{1+\kappa_2} > \kappa_{1,r}$ **then**
$q_1 \leftarrow 2p^{\dagger}\sqrt{1+\kappa_2}$
$q_2 \leftarrow 2p\sqrt{1+\kappa_2}$
$\widehat{\boldsymbol{\lambda}}_r \leftarrow \left(\frac{q_2-\kappa_{1,r}}{q_2-q_1}\widehat{\boldsymbol{\lambda}}_r + \frac{\kappa_{1,r}-q_1}{q_2-q_1}\boldsymbol{\lambda}^{\dagger}\right)\left(\sqrt{1+\kappa_2}\right)^{-1}$
$p \leftarrow p^{\dagger}$
else
if $\tilde{\gamma} < \widehat{\gamma}$ **then**
$d \leftarrow 1$
$\widehat{\boldsymbol{\lambda}}_r[n] \leftarrow 0$
$\mathcal{I} \leftarrow \mathcal{I} \sqcup \{n\}$
$\mathbf{s} \leftarrow \left]\mathbf{s}, s_{|\mathcal{A}|}\right[$
$\mathcal{A} \leftarrow \mathcal{A} \backslash \{n\}$
Downdate the lower Cholesky matrix $\mathbf{L}$ by $\mathbf{x}_n$
end if
end if
end while
return $\widehat{\boldsymbol{\lambda}}_r$

References

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, 32(2):407 – 499.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

Zou, H. and Hastie, T. (2020). *elasticnet: Elastic-Net for Sparse Estimation and Sparse PCA*. R package version 1.3.