Wavelet Estimation for Dynamic Factor Models with Time-Varying Loadings

Duván Humberto Cataño, Carlos Vladimir Rodríguez-Caballero and Daniel Peña

CREATES Research Paper 2019-23

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Wavelet Estimation for Dynamic Factor Models with Time-Varying Loadings

Duván Humberto Cataño*[†], Carlos Vladimir Rodríguez-Caballero[‡], and Daniel Peña [§]

Abstract

We introduce a non-stationary high-dimensional factor model with time-varying loadings. We propose an estimation procedure based on two stages. First, we estimate common factors by principal components. Afterwards, in the second step, considering the factors estimates as observed, the time-varying loadings are estimated by an iterative procedure of generalized least squares using wavelet functions. We investigate the finite sample features of the proposed methodology by some Monte Carlo simulations. Finally, we use this methodology to study the electricity prices and loads of the Nord Pool power market.

Keywords: Factor models, wavelet functions, generalized least squares, electricity prices and loads.

1 Introduction

Factor models have been widely used in the last decade due to their ability to explain the structure of a common variability among time series by a small number of unobservable common factors. In this sense, these models are used to reduce dimensionality on complex systems. The studies of factor models have encompassed the stationary and nonstationary frameworks by different estimation methods, see among many others, Pena and Box (1987), Forni et al. (2000), Stock and Watson (2002), Bai and Ng (2002), Bai (2003), Forni et al. (2004), and Forni et al. (2005) for the stationary cases, and Bai and Ng (2004), Peña and Poncela (2006),

*The author would like to thank financial supports from CAPES, CNPq, and University of São Paulo, Brasil.

†University of Antioquia, Colombia.

‡Corresponding author. Department of Statistics, ITAM, Mexico & CREATES, Aarhus University, Denmark. Address: Río Hondo No.1, Col. Progreso Tizapán, Álvaro Obregón, CDMX. 01080. Mexico. E-mail: vladimir.rodriguez@itam.mx

§Department of Statistics and Institute UC3M-BS of Financial Big Data. Universidad Carlos III de Madrid, Getafe, Spain.

Barigozzi et al. (2016), and Rodríguez-Caballero and Ergemen (2017) for the non-stationary cases.

One non stationary framework of special interest is the factor model with loadings that vary over time. Motta et al. (2011) introduced deterministic smooth variations in factor loadings and proposed estimation procedures based on locally weighted generalized least squares using kernel functions under a time-domain approach. Similarly, Eichler et al. (2011) allowed for a non stationary dynamic structure in a factor model and considered deterministic time-dependent functions in the factor loadings. Their estimation procedure can be seen as a time-varying spectral density matrix of the underlying process. Furthermore, Mikkelsen et al. (2018) proposed a factor model with time-varying loadings that evolve as stationary VAR processes. They employed Kalman filter procedures to obtain the maximum likelihood estimators of the parameters of the factor loadings.

In this paper, we use wavelet functions to define smooth variations of loadings in high-dimensional factor models. Our model can be useful in some economic applications when the dynamics is driven by smooth progressive variations, whose cumulative effects cannot be simply ignored, see Su and Wang (2017), and Bai and Han (2016), for details. In our framework, to capture the common smooth variations in the vector of time series, the parameters in the loading matrix are assumed to be well aproximated by deterministic functions over time.

The estimation procedure consists of two steps. First, the common factors are estimated by standard principal component analysis (PCA). Then, in the second step, considering the factors as known, factor loadings are estimated by an iterative procedure that combines generalized least squares (GLS) using wavelet functions. We show that factors estimated by principal components are consistent after controlling the magnitude of the loadings instabilities. We highlight that a necessary requirement for such consistency is that common factors need to be independent of the functions of factor loadings. This model is extended for the case when variables are non-stationary where the factor are first estimated by the approach of Peña and Poncela (2006). We use Monte Carlo simulations to show that the method correctly identify the factors and loadings. Finally, we use the methodology proposed with non-stationary variables to analyze an electricity market: the Nord Pool power market. We use the electricity system prices and loads throughout two years and a half to show that factor loadings are not invariant along time, illustrating the useful of our model in energy markets. We find that some features of electricity prices and loads, see e.g. Weron (2007), and Weron (2014), are well extracted by common factors estimates, and by time-varying loadings estimates.

The remainder of the paper is organized as follows. Section 2 introduces the model, shows the consistency of principal components with stationary and non-stationary variables, and introduces the wavelet functions. Section 3 discusses the proposed estimation procedure. Monte Carlo simulation studies are presented in Section 4, and an empirical illustration is provided in Section 5. Section 6 concludes. Throughout the paper we use bold-unslanted letters for matrices, bold-slanted letters for vectors and unbold (normal) letters for scalars. We denote by

$tr(\cdot)$ the trace operator, by $rank(\mathbf{A})$ the rank of a matrix $\mathbf{A}$, by $\mathbb{I}_n$ the identity matrix of dimension n, by $\otimes$ the Kronecker product and by $\|\cdot\|$ the Frobenius (Euclidean) norm, i.e., $\|\mathbf{A}\| = \sqrt{tr(\mathbf{A}'\mathbf{A})}$.

2 The model

The model we propose is as follows

$$\mathbf{Y}_t = \mathbf{X}_t + \mathbf{e}_t, \tag{1}$$
$$\mathbf{X}_t = [\mathbf{\Lambda}_0 + \mathbf{\Lambda}(t/T)]\mathbf{F}_t, \tag{2}$$

where the common component, $\mathbf{X}_t$, is a $N-$dimensional locally stationary process, in sense of Dahlhaus et al. (1997b) and the loadings are defined in the re-scaled time, $u = t/T \in [0,1]$. $\mathbf{F}_t$ are the unobservable common factors and $\mathbf{e}_t$, the idiosyncratic component, is a sequence of weakly dependent variables. $\mathbf{\Lambda}(u) = \{\lambda_{ij}(u),\ i = 1,\dots,N;\ j = 1,\dots,r$, is the time-dependent factor loading matrix. In this respect, the influence of $\mathbf{F}_t$ on the observed process varies over time and $\mathbf{\Lambda}(u)$ capture the smooth variations of the loadings with respect $\mathbf{\Lambda}_0$, a constant loading matrix.

When $\mathbf{\Lambda}(u) = 0, \quad \forall u \in [0,1]$, the standard factor model is considered, therefore the model proposed in (1) can be seen as a generalization of the standard factor model of Bai (2003).

Definition 1. *The sequence $\mathbf{Y}_t$ in (1) follows a factor model with time-varying loadings if:*

a. For $N \in \mathbb{N}$, there is a function with

$$\begin{aligned} \mathbf{\Lambda}(\cdot) \ : \ [0,1] &\rightarrow \mathbb{R}^{N\times r} \\ u &\mapsto \mathbf{\Lambda}(u) \end{aligned}$$

such that $\forall\ T \in \mathbb{N}$,

$$\Gamma_Y(u) = [\mathbf{\Lambda}_0 + \mathbf{\Lambda}(u)]\Gamma_F[\mathbf{\Lambda}_0 + \mathbf{\Lambda}(u)]' + \Gamma_e,$$

where $rank[\mathbf{\Lambda}_0 + \mathbf{\Lambda}(u)] = r$ and $\Gamma_F = \mathbb{V}ar(\mathbf{F}_t)$, is a positive definite diagonal matrix.

b. $\Gamma_e = var(\mathbf{e}_t)$ is a positive definite matrix.

2.1 Assumptions

With an arbitrary constant $Q \in \mathbb{R}^+$, the assumptions of the model in (1) following Bai and Ng (2002) are as follows,

Assumption A. ***Factors:***

A_1 $\mathbb{E}\|\boldsymbol{F}_t\|^4 < Q$,

A_2 $T^{-1}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{F}_t' \xrightarrow{p} \Gamma_F$ *when* $T \to \infty$, *where* Γ_F *is a positive definite diagonal matrix.*

Assumption B. ***Factor loadings:***

B_1 $\|\boldsymbol{\lambda}_{i0}\| \leq \bar{\lambda}$ *and* $\|\boldsymbol{\Lambda}_0'\boldsymbol{\Lambda}_0/N - D\| \to 0$, *when* $N \to \infty$, *with* $r \times r$ *positive definite matrix* D, *where* $\boldsymbol{\lambda}_{i0}$ *is the* $i-th$ *row of* $\boldsymbol{\Lambda}_0$,

B_2 $\sup_{u\in(0,1)}\|\boldsymbol{\lambda}_i(u)\| \leq \bar{\lambda} < \infty$, *where* $\boldsymbol{\lambda}_i(u)$ *is the* $i-th$ *row of* $\boldsymbol{\Lambda}(u)$,

B_3 $\lambda_{ij}(u) \in L^2[0,1]$, *for* $i = 1,\dots,N$ *and* $j = 1,\dots,r$.

Assumption C. ***Idiosyncratic terms:***

C_1 $\mathbb{E}(e_{it}) = 0, \mathbb{E}|e_{it}|^4 < Q$,

C_2 $\mathbb{E}[e_{it}e_{jt}] = \tau_{ij,t}$, *with* $|\tau_{ij,t}| < |\tau_{ij}|$, *for a constant* $|\tau_{ij}|$ *and* $N^{-1}\sum_{i,j=1}^{N}|\tau_{ij}| < Q$, *for all* t,

C_3 $\mathbb{E}[N^{-1}\sum_{i=1}^{N}e_{is}e_{it}] = \gamma_N(s,t)$, $|\gamma_N(s,s)| < Q$, *for all* (s,t); *and* $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}|\gamma_N(s,t)| < Q$.

Assumption D. ***Time-varying factor loadings and factors:***
K_{1NT}, K_{2NT} and K_{3NT} are functions such that the following conditions are fulfilled for any $n,m,k,q = 1,\dots,r$ and $\lambda_{ij}(s/T) \equiv \lambda_{ij}(s)$.

D_1 $\sup_{s,t}\sum_{i,j}^{N}|\lambda_{in}(s)\lambda_{jm}(t)||\mathbb{E}[F_{ns}F_{mt}]| < K_{1NT}$,

D_2 $\sum_{s,t=1}^{T}\sum_{i,j=1}^{N}|\lambda_{in}(s)\lambda_{jm}(s)||\mathbb{E}[F_{ns}F_{ms}F_{kt}F_{qt}]| < K_{2NT}$,

D_3 $\sum_{s=1}^{T}\sum_{i,j=1}^{N}|\lambda_{in}(s)\lambda_{jm}(s)\lambda_{ik}(t)\lambda_{jq}(t)||\mathbb{E}[F_{ns}F_{ms}F_{kt}F_{qt}]| < K_{3NT}$.

Assumption E. ***Independence:***
The process e_{it} and F_{js} are independent of each other for any (i,j,s,t).

Assumption A imposes standard moment conditions, that is, the unobservable factors have finite fourth moments and their covariance converges in probability to a positive definite matrix. Assumptions B_1 and B_2 ensure that each factor has a nontrivial contribution to the variance of $\mathbf{Y}_t$. Assumption B_3 ensures existence of the expansion in the wavelet function for the factor loadings. Assumptions C allows dependence in the idiosyncratic process. Assumption D is required to guarantee the consistency of the principal components. Finally, independence between factors and the idiosyncratic term is provided in the Assumption E.

2.2 Principal component estimators

We use PCA to estimate the factors $\mathbf{F}_t$. It is well-known that principal components are obtained solving the optimization problem

$$\min_{\mathbf{F},\mathbf{\Lambda}}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it}-\boldsymbol{\lambda}_i'\mathbf{F}_t)^2, \tag{3}$$

where $\mathbf{F}=(\mathbf{F}_1,\mathbf{F}_2,\dots,\mathbf{F}_T)'$ is a $T\times r$ matrix and $\mathbf{\Lambda}$ is a $N\times r$ matrix. We need to impose some restrictions to guarantee the identification of the parameters as usual. Solving for $\mathbf{\Lambda}$, the normalization $\mathbf{F}'\mathbf{F}=\mathbb{I}_r$ provides the necessary number of restrictions. With this, minimize (3) is equivalent to maximize $tr[\mathbf{F}'(\mathbf{YY}')\mathbf{F}]$, where $\mathbf{Y}=(\mathbf{Y}_1,\dots,\mathbf{Y}_T)'$. Then, the estimated factor matrix, $\tilde{\mathbf{F}}$, is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T\times T$ matrix $\mathbf{YY}'$. It is well-known that this solution is not unique, that is, any orthogonal rotation of $\tilde{\mathbf{F}}$ is also a solution. See Bai et al. (2008) for more details.

The following theorem is a modified version of Theorem 1 in Bates et al. (2013) and shows that, under the assumptions previously stated, it is possible to consistently estimate any rotation of the factors by principal components even if the loadings are time-varying.

Theorem 1. *Under Assumptions A-E, there exists an $r\times r$ matrix H such that*

$$T^{-1}\sum_{t=1}^{T}\|\tilde{\mathbf{F}}_t-H'\mathbf{F}_t\|^2=O_p(R_{NT}), \tag{4}$$

when $N,T\to\infty$, where $R_{NT}=\max\left\{\frac{1}{N},\frac{1}{NT},\frac{K_{1NT}}{N^2},\frac{K_{2NT}}{N^2T^2},\frac{K_{3NT}}{N^2T^2}\right\}$ with K_{1NT}, K_{2NT}, and K_{3NT} defined in the Assumption D. Furthermore, $H=(\mathbf{\Lambda}_0'\mathbf{\Lambda}_0/N)(\mathbf{F}'\tilde{\mathbf{F}}/T)V_{NT}^{-1}$, where V_{NT} is a diagonal matrix of the r largest eigenvalues of the matrix $(NT)^{-1}\mathbf{YY}'$.

Proof. See the appendix A.1.

Theorem 1 points out that the average squared deviation between the estimated factors and the space spanned by a rotation of the true factors will vanish at rate R_{NT}, which is similar to that in Bai and Ng (2002). Note that from (4), the estimated common factors, $\tilde{\mathbf{F}}_t$, are identified through a rotation, then, principal components converge to a rotation of the true common factors $H'\mathbf{F}_t$.

2.3 Nonstationary factors

In many areas, as in economics and finance, there is strong evidence in favor of the presence of non-stationarity processes, which have repeatedly found in many empirical studies with economic or financial time series. Consequently, it is natural to think that many panel data may include non-stationary economic or financial variables. There has been some debate concerning the use of differenced variables, the main argument being discussed is whether differencing the series causes a severe loss of information. In this paper, to treat with non-stationary variables, we use the methodology proposed by Peña and Poncela (2006) who, assuming that $\mathbf{Y}_t \sim I(d)$ with a positive integer d, use generalized covariance matrices, which is defined as

$$\mathbf{C}_y(k) = \frac{1}{T^{2d+d'}} \sum_{t=k+1}^{T} (\mathbf{Y}_{t-k} - \bar{\mathbf{Y}}_t)(\mathbf{Y}_t - \bar{\mathbf{Y}}_t), \tag{5}$$

where $\bar{\mathbf{Y}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{Y}_t$ and d' can be either 0 or 1. In this framework, a consistent estimate for the factors are:

$$\hat{\mathbf{F}} = \mathbf{Y}\hat{\mathbf{\Lambda}}, \tag{6}$$

where $\hat{\mathbf{F}} = (\hat{\mathbf{F}}_1, \hat{\mathbf{F}}_2, \ldots, \hat{\mathbf{F}}_T)'$ is a $T \times r$ matrix, $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_T)'$ is a $T \times N$ matrix and $\hat{\mathbf{\Lambda}}$ is a $N \times r$ matrix composed by the first r eigenvectors of $\mathbf{C}_y(k)$. Following the same reasoning as in Theorem 1, it can be shown that the average squared deviation between the estimated factors (6) and the space spanned by a rotation of the true factors will vanish when $(N, T) \to \infty$.

2.4 Wavelets

The basic idea of a wavelet is to construct infinite collections of translated and scaled versions of the scale function $\phi(t)$ and the wavelet $\psi(t)$ such as $\phi_{jk}(t) = 2^{j/2}\phi(2^j t - k)$, and $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$ for $j, k \in \mathbb{Z}$. Suppose that $\{\phi_{lk}(\cdot)\}_{k\in\mathbb{Z}} \cup \{\psi_{jk}(\cdot)\}_{j\geq l; k\in\mathbb{Z}}$ forms an orthonormal basis of $L^2(\mathbb{R})$, for any coarse scale l. A key point is to construct ϕ and ψ with a compact support that generates an orthonormal system, which has location in time-frequency. From this, we can get parsimonious representations for a wide class of wavelet functions, see Chiann and Morettin (2005) and Porto et al. (2008) for details. In some applications, these functions are defined in a compact set such as $[0, 1]$. We consider this compact set for functions $\lambda_{ij}(u)$, for $i = 1, \ldots, N$ and $j = 1, \ldots, r$ defined in (2). Then, it is necessary to consider an orthonormal system that generates $L^2[0, 1]$. For the construction of these orthonormal systems we follow the procedure by Cohen and Ryan (1995), that generates multiresolution levels $\tilde{V}_0 \subset \tilde{V}_1 \subset \cdots$, where the spaces $\tilde{V}_j$ are generated by $\tilde{\psi}_{jk}$. Negative values of j are not necessary since $\tilde{\phi} = \tilde{\phi}_{00} = 1$, and if $j \leq 0$, $\tilde{\psi}_{jk}(u) = 2^{-j/2}$, see Vidakovic (2009) for more details and different approaches. Therefore, for any function $\lambda(u) \in L^2[0, 1]$, can be expand in series of orthogonal functions

$$\lambda(u) = \alpha_{00}\phi(u) + \sum_{j\geq 0}\sum_{k\in I_j} \beta_{jk}\psi_{jk}(u), \tag{7}$$

where we take $l = 0$ and $I_j = \{k : k = 0, \ldots, 2^j - 1\}$. For each j, the set I_j generates values of k such that β_{jk} belongs to the scale 2^j. For example, for $j = 3$, there are eight wavelet coefficients in the scale 2^3, whereas for $j = 2$, only four coefficients in the scale 2^2.

Some applications consider the equation in (7) for a maximum resolution level J, through

$$\lambda(u) \approx \alpha_{00}\phi(u) + \sum_{j=0}^{J-1}\sum_{k\in I_j}\beta_{jk}\psi_{jk}(u). \tag{8}$$

In this way, the function $\lambda(u)$ approximates to the space $\tilde{V}_J$. In this paper, we use ordinary wavelets as in Dahlhaus et al. (1997a) due to their performance is suitable in the case of smooth functions. Particularly, Daubechies $D8$ and Haar wavelets of compact supports are employed.

3 Estimation of time-varying loadings by wavelets

We consider the process in (1) with r common factors $(r < N)$ to discuss the estimation procedure of the time-varying loadings. From now on, we consider the loading matrix, $[\mathbf{\Lambda}(u) + \mathbf{\Lambda}_0]$, as a unique function over time $\mathbf{\Lambda}(u)$, given by

$$\mathbf{Y}_t = \mathbf{\Lambda}(u)\mathbf{F}_t + \mathbf{e}_t, \tag{9}$$

with $t = 1, 2, \ldots, T$, and $u = t/T \in [0, 1]$. In matrix form, we have

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{rt} \\ \vdots \\ Y_{Nt} \end{bmatrix} = \begin{bmatrix} \lambda_{11}(u) & \lambda_{12}(u) & \ldots & \lambda_{1r}(u) \\ \lambda_{21}(u) & \lambda_{22}(u) & \ldots & \lambda_{2r}(u) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{r1}(u) & \lambda_{r2}(u) & \ldots & \lambda_{rr}(u) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N1}(u) & \lambda_{N2}(u) & \ldots & \lambda_{Nr}(u) \end{bmatrix} \begin{bmatrix} F_{1t} \\ F_{2t} \\ \vdots \\ F_{rt} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{rt} \\ \vdots \\ e_{Nt} \end{bmatrix}, \tag{10}$$

where $\mathbf{Y}_t$ is an $N-$dimensional vector of time series, $\mathbf{\Lambda}(u)$ is the time-varying loading matrix with $\lambda_{ij}(u) \in L^2[0, 1]$, for $i = 1, 2, \ldots, N$, and $j = 1, 2, \ldots, r$. $\mathbf{F}_t$ is the common factor, and $\mathbf{e}_t$ is the idisyncratic process. From (9), and the Assumption E, the structure of the covariance matrix of the process $\mathbf{Y}_t$ is written as

$$\Gamma_Y(u) = \mathbf{\Lambda}(u)\Gamma_F\mathbf{\Lambda}'(u) + \Gamma_e, \quad \forall u \in [0, 1],$$

that means, the variance of the common component is $\Gamma_X(u) = \mathbf{\Lambda}(u)\Gamma_F\mathbf{\Lambda}'(u)$.

For the construction of the time-varying loadings estimates, we first assume that the estimator of the r common factors are obtained by PCA of the $N-$dimensional time series $\mathbf{Y}_t$. Then, functions of time-varying loadings are approximated in series of orthogonal wavelets as in (8), for a fixed resolution level

$J < T$,

$$\lambda_{mn}(u) = \alpha_{00}^{(mn)}\phi(u) + \sum_{j=0}^{J-1}\sum_{k\in I_j}\beta_{jk}^{(mn)}\psi_{jk}(u). \tag{11}$$

The values of j, k vary depending on the resolution level in the wavelet decomposition. We choose the maximum resolution J, such that $2^{J-1} \leq \sqrt{T} \leq 2^J$, see Dahlhaus et al. (1997a) for details of this selection. In practice, the coefficients $\alpha_{00}^{(mn)}, \beta_{00}^{(mn)}, \beta_{10}^{(mn)}, \ldots, \beta_{J-1,2^J-1}^{(mn)}$ are obtained for a particular estimation method. In this paper, we use GLS to estimate these coefficients and to reconstruct the loadings functions.

Let $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{Nt})$ be N time series with $t = 1, 2, \ldots, T$ which are generated by

$$\mathbf{Y}_t = \mathbf{\Lambda}(u)\tilde{\mathbf{F}}_t + \mathbf{e}_t, \tag{12}$$

where the r common factors, $\tilde{\mathbf{F}}_t$, are estimated by principal components. Each loading function $\lambda_{mn}(u)$ is written as in (11), then when plugging each $\lambda_{mn}(u)$ into (10), we have

$$\underbrace{\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1T} \\ \vdots \\ Y_{r1} \\ \vdots \\ Y_{rT} \\ \vdots \\ Y_{N1} \\ \vdots \\ Y_{NT} \end{bmatrix}}_{vec(\mathbf{Y})} = \underbrace{\begin{bmatrix} \Psi_{\tilde{\mathbf{F}}}^{(1)} & \Psi_{\tilde{\mathbf{F}}}^{(2)} & \ldots & \Psi_{\tilde{\mathbf{F}}}^{(r)} & \ldots & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \Psi_{\tilde{\mathbf{F}}}^{(1)} & \Psi_{\tilde{\mathbf{F}}}^{(2)} & \ldots & \Psi_{\tilde{\mathbf{F}}}^{(r)} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \Psi_{\tilde{\mathbf{F}}}^{(1)} & \Psi_{\tilde{\mathbf{F}}}^{(2)} & \ldots & \Psi_{\tilde{\mathbf{F}}}^{(r)} \end{bmatrix}}_{\mathbf{\Theta}} \underbrace{\begin{bmatrix} \boldsymbol{\beta}^{(1)} \\ \boldsymbol{\beta}^{(2)} \\ \boldsymbol{\beta}^{(3)} \\ \vdots \\ \boldsymbol{\beta}^{(r)} \\ \vdots \\ \boldsymbol{\beta}^{(N)} \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} e_{11} \\ \vdots \\ e_{1T} \\ \vdots \\ e_{r1} \\ \vdots \\ e_{rT} \\ \vdots \\ e_{N1} \\ \vdots \\ e_{NT} \end{bmatrix}}_{vec(\mathbf{e})}, \tag{13}$$

where

$$\Psi_{\tilde{\mathbf{F}}}^{(i)} = \begin{bmatrix} \phi(1/T)\tilde{F}_{i1} & \psi_{00}(1/T)\tilde{F}_{i1} & \ldots & \psi_{J-1,2^J-1}(1/T)\tilde{F}_{i1} \\ \phi(2/T)\tilde{F}_{i2} & \psi_{00}(2/T)\tilde{F}_{i2} & \ldots & \psi_{J-1,2^J-1}(2/T)\tilde{F}_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi(T/T)\tilde{F}_{iT} & \psi_{00}(T/T)\tilde{F}_{iT} & \ldots & \psi_{J-1,2^J-1}(T/T)\tilde{F}_{iT} \end{bmatrix},$$

are $T \times 2^J$ matrices for $i = 1, 2, \ldots, r$ and $\mathbf{O}$ is $T \times 2^J$ null matrix.

Let $\Psi_{r\tilde{F}} = [\Psi_{\tilde{F}}^{(1)}, \ldots, \Psi_{\tilde{F}}^{(r)}]$ be a $T \times r2^J$ matrix, then

$$\mathbf{\Theta} = \mathbb{I}_N \otimes \Psi_{r\tilde{F}}$$

is $NT \times Nr2^J$ matrix that depends on the estimated factors, $\tilde{\mathbf{F}}_t$, the wavelets $\psi(u)$, and the resolution level J, with vector of parameters $\boldsymbol{\beta}^{(m)} = \left(\boldsymbol{\beta}^{(m1)}, \boldsymbol{\beta}^{(m2)}, \ldots, \boldsymbol{\beta}^{(mr)}\right)'$ of dimension $r2^J \times 1$ for $m = 1, 2, \ldots, N$, where $\boldsymbol{\beta}^{(mn)} = \left(\alpha_{00}^{(mn)}, \beta_{00}^{(mn)}, \beta_{10}^{(mn)}, \ldots, \beta_{J-1,2^J-1}^{(mn)}\right)'$.

Each $\boldsymbol{\beta}^{(m)}$ is composed by the wavelets coefficients of the $m-$th row of the matrix $\boldsymbol{\Lambda}(u)$. Therefore, the total number of wavelets parameters to be estimated is $2^J Nr$.

Hence, the model in (13) can be represented in a linear model form as

$$vec(\mathbf{Y}) = \boldsymbol{\Theta}\boldsymbol{\beta} + vec(\mathbf{e}),$$

where $vec(\mathbf{Y})$ is the response vector and $\boldsymbol{\Theta}$ is the usual design matrix in regression analysis. Assuming that the covariance matrix of the idiosyncratic errors, Γ_e, is known, then the GLS estimator of the coefficients $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{\Theta}'\Sigma_e^{-1}\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}'\Sigma_e^{-1}vec(\mathbf{Y}),$$

where Σ_e is a $NT \times NT$ matrix defined as

$$\Sigma_e = \Gamma_e \otimes \mathbb{I}_T = \begin{bmatrix} \mathbb{I}_T\gamma_{e,11} & \mathbb{I}_T\gamma_{e,12} & \cdots & \mathbb{I}_T\gamma_{e,1N} \\ \mathbb{I}_T\gamma_{e,21} & \mathbb{I}_T\gamma_{e,22} & \cdots & \mathbb{I}_T\gamma_{e,2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{I}_T\gamma_{e,N1} & \mathbb{I}_T\gamma_{e,N2} & \cdots & \mathbb{I}_T\gamma_{e,NN} \end{bmatrix}$$

These results provide linear estimators of wavelets coefficients for the time-varying loadings assuming that the covariance matrix of the idiosyncratic error is known. Some procedures, such as maximum likelihood methods, are not computationally efficient for estimating such model, since the number of parameters tends to be very large. In this light, we use GLS to simplify the implementation.

Note that we can use different basis of wavelet functions $\phi(u)$, and $\psi_{jk}(u)$ for each $\lambda_{mn}(u)$. In the simulation section, we use similar basis to simplify the exposition.

3.1 Estimation

The estimation procedure is based on a general two-step procedure, which can be executed by the following algorithm:

Step 1. Use principal components to estimate $\mathbf{F}_t$ as

$$\tilde{\mathbf{F}} = \sqrt{T}[v_1, v_2, \ldots, v_r],$$

where v_i is the eigenvector corresponding to the $i-$th largest eigenvalue, λ_i for $i = 1, \ldots, r$, of the matrix $(NT)^{-1}\mathbf{Y}\mathbf{Y}'$. Here, $\mathbf{Y}$ and $\tilde{\mathbf{F}}$ denote $T \times N$

and $T \times r$ matrices, respectively. Next, the loading function matrix, $\mathbf{\Lambda}(t)$, is approximated by the wavelets

$$\lambda_{mn}(t) = \alpha_{00}^{(mn)}\phi(t) + \sum_{j=0}^{J-1}\sum_{k \in I_j} \beta_{jk}^{(mn)}\psi_{jk}(t),$$

with $I_j = \{k : k = 0, 1, \ldots, 2^j - 1\}$, $m = 1, \ldots, N$, and $n = 1, \ldots, r$. Then, we write the equation (12) as

$$vec(\mathbf{Y}) = \mathbf{\Theta}(\tilde{\mathbf{F}}, \psi, \phi)\boldsymbol{\beta} + vec(\mathbf{e}),$$

where $vec(\mathbf{Y})$ and $vec(\mathbf{e})$ are $NT \times 1$ vectors, and the dimensions of $\mathbf{\Theta}(\tilde{\mathbf{F}}, \psi, \phi)$ and $\boldsymbol{\beta}$ are $(NT \times 2^J Nr)$ and $(2^J Nr \times 1)$, respectively, where J indicates the resolution level chosen in the wavelet expansions.

Step 2. Estimate by GLS the wavelet coefficients as

$$\hat{\boldsymbol{\beta}} = (\mathbf{\Theta}'\Sigma_e^{-1}\mathbf{\Theta})^{-1}\mathbf{\Theta}'\Sigma_e^{-1}\mathbf{Z},$$

where $\mathbf{\Theta}(\tilde{\mathbf{F}}, \psi, \phi) \equiv \mathbf{\Theta}$, $\mathbf{Z} = vec(\mathbf{Y})$, and $\Sigma_e = \mathbb{I}_{NT}$ is used as initial value.

Step 3. Using the estimated coefficient in the Step 2, the loadings are obtained as

$$\hat{\mathbf{\Lambda}}(t)^{(0)} = \{\hat{\lambda}_{mn}^{(0)}(t)\}_{m=1,\ldots,N}^{n=1,\ldots,r},$$

where

$$\hat{\lambda}_{mn}^{(0)}(t) = \hat{\alpha}_{00}^{(mn)}\phi(t) + \sum_{j=0}^{J-1}\sum_{k \in I_j} \hat{\beta}_{jk}^{(mn)}\psi_{jk}(t).$$

Step 4. With $\hat{\mathbf{\Lambda}}(t)^{(0)}$, obtain residuals, $\mathbf{Y}_t - \hat{\mathbf{\Lambda}}(t)^{(0)}\tilde{\mathbf{F}}_t = \hat{\mathbf{e}}_t^{(0)}$. Then, compute

$$\hat{\Gamma}_e^{(0)} = \sum_{t=1}^{T} \hat{\mathbf{e}}_t^{(0)}\hat{\mathbf{e}}_t^{(0)\prime}/T.$$

Step 5. Back to step 2 with $\Sigma_e = \hat{\Gamma}_e^{(0)}$. Iterate n-times the procedure to obtain the sequences $\{\hat{\mathbf{\Lambda}}(t)^{(i)}, \hat{\Gamma}_e^{(i)}\}_{i=1,\ldots,n}$. Stop the iteration when

$$\|\hat{\mathbf{\Lambda}}(t)^{(i-1)} - \hat{\mathbf{\Lambda}}(t)^{(i)}\| < \delta,$$

for any small $\delta > 0$, where $\|\cdot\|$ denotes the Frobenius norm for $t = 1, \ldots, T$.

4 Monte Carlo simulation

We examine the finite-sample properties of the estimation procedure proposed above using a Monte Carlo study. The model in (1) is generated as

$$\begin{aligned} Y_{it} &= \boldsymbol{\lambda}_i'(t)\mathbf{F}_t + e_{it}, \quad i = 1, \ldots, N \quad \text{and} \quad t = 1, \ldots, T, \\ F_{kt}(1 - \theta_k B) &= \eta_{kt}, \quad k = 1, \ldots, r. \ \boldsymbol{\eta}_t \sim \mathcal{N}_r(0, diag\{1 - \theta_1^2, \ldots, 1 - \theta_r^2\}), \\ \mathbf{e}_t &\sim \mathcal{N}_N(0, \Gamma_e), \end{aligned}$$

where the matrix Γ_e is generated by a couple of different structures: i) $\Gamma_e = \{\gamma^{|i-j|}\}_{i,j=1,\ldots,N}$, that is a Toeplitz matrix, and ii) a diagonal matrix. Furthermore, at instant t, Y_{it} denotes the $i-$th time series, $\boldsymbol{\lambda}_i'(t) = (\lambda_{i1}(t), \ldots, \lambda_{ir}(t))$ is the vector of loadings, which are generated alternately by some smooth functions. We discuss a couple of functions used below. $\mathbf{F}_t = (F_{1t}, \ldots, F_{rt})'$ is the vector of factors, and $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{rt})'$ and $\mathbf{e}_t = (e_{1t}, \ldots, e_{Nt})'$ are vectors of idiosyncratic terms which are independent to each other.

In our Monte Carlo study, the model is generated with $N \in \{20, 30\}$ cross-sectional units and $T \in \{512, 1024, 2048\}$ sample sizes. We consider for simplicity only two common factors, that is $r = 2$. Furthermore, three values for θ are considered; $\theta \in \{0, 0.5, 1\}$. For the case with $\theta = 1$, $\boldsymbol{\eta}_t$ is simulated as $\mathcal{N}_r(0, diag\{\theta_1^2, \ldots, \theta_r^2\})$. Two values for Γ_e; a Toeplitz matrix $\Gamma_e = Toep$ with $\gamma = 0,7$ for correlated noise and $\Gamma_e = Diag$ for uncorrelated, where the entries of the diagonal matrix are generated by an uniform distribution $U(0.5, 1.5)$. Note that simulated data are standardized before extracting the principal components. Common factors are estimated by principal components in cases with $\theta < 1$, and by the procedure of Peña and Poncela (2006) in the case with $\theta = 1$. All simulations are based on 1000 replications of the model.

We rotate the obtained factors in order to compare proposed estimations with the actual factors. The optimal rotation A^* is obtained by maximizing $tr[corr(\mathbf{F}, \tilde{\mathbf{F}}A)]$. The solution is given by $A^* = VU$ where V and U are orthogonal matrices of the decomposition $corr(F, \tilde{F}) = USV'$. When the number of k principal components is not equal to the number of factors r, we rotate the first $l = \min\{k, r\}$ principal components, see Eickmeier et al. (2015). Both estimated and simulated factors are re-scaled to keep the same standard deviation, then

$$\tilde{\mathbf{F}}_k^* = \frac{\sigma(F_k)}{\sigma(\tilde{F}_k)} \tilde{\mathbf{F}}_k, \quad k = 1, \ldots, r, \tag{14}$$

where $\tilde{\mathbf{F}}_k$ is the $k-$th column of the matrix of the rotated principal components $\tilde{\mathbf{F}}A^*$.

As explained before, these rotated factors are now treated as observed variables in the regression model

$$vec(\mathbf{Y}) = \boldsymbol{\Theta}(\tilde{\mathbf{F}}^*, \psi, \phi)\boldsymbol{\beta} + vec(\mathbf{e}),$$

to estimate the wavelet coefficients, $\boldsymbol{\beta}$, where $\tilde{\mathbf{F}}^* = (\tilde{\mathbf{F}}_1^*, \ldots, \tilde{\mathbf{F}}_r^*)$.

To investigate the performance of the estimation procedure, estimated and simulated factors are compared as following:

i) The precision of the estimation factors is measured by the $R^2_{\tilde{F},F}$ statistics as in Bates et al. (2013), given by

$$R^2_{\tilde{F},F} = \frac{tr[\mathbf{F}'\tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}'\mathbf{F}]}{tr[\mathbf{F}'\mathbf{F}]}, \tag{15}$$

where $\tilde{\mathbf{F}}$ is the $T \times r$ matrix of estimated factors as in (14) and $\mathbf{F}$ is the $T \times r$ matrix of the actual factors, that is the simulated ones. This statistics is a multivariate R^2 in a regression of the actual factors on the principal components. When the canonical correlation of the actual and estimated factors tends to one, then $R^2_{\tilde{F},F} \to 1$ as well.

ii) The precision of the estimation loadings is measured by the mean square errors (MSE) between estimated and actual loadings, as in Motta et al. (2011). The MSE is computed as follows

$$MSE(v) = (NT)^{-1} \sum_{t=1}^{T} \|\hat{\boldsymbol{\Lambda}}^{(v)}(t) - \boldsymbol{\Lambda}(t)\|,$$

for $v = 1, \ldots, 1000$. The estimator of the factor loadings matrix, $\boldsymbol{\Lambda}(t)$, is chosen by a path such that

$$\hat{\boldsymbol{\Lambda}}(t) = \{\hat{\boldsymbol{\Lambda}}^{(m)}(t) : MSE_m = median\{MSE(1), \ldots, MSE(1000)\}\}. \tag{16}$$

Table 1 shows the results of estimations of (15) and (16). As can be seen, the methodology proposed in this paper performs very well in relatively small samples regardless of size distortion between N and T. As seen in Table 1, $R^2_{\tilde{F},F}$ is relatively high indicating a good performance of the estimator, although the precision is a bit reduced when increasing the value of θ. These findings are maintained for the both type of wavelets used and even when we allow for cross-correlation between idiosyncratic errors. Furthermore, inspecting the MSE in Table 1, we find that the MSEs decrease as T increases in all cases, even if the common factors are serially correlated and even if idiosyncratic errors are cross-correlated. Furthermore, another finding indicates that in general, factor loadings using the wavelet D8 perform better than the wavelet Haar. We think that such findings are reasonable due to the smoothness of the wavelet D8 in contrast with the other one. The wavelet Haar should be implemented when dynamics of the factor loadings have breaks or perhaps some aggressive jumps. These structures are not considered in the current paper, but these possible features of the wavelet Haar is part of an another research and is out of the present scope.

Finally, in Figures 1 and 2, we display the good performance of the methodology to estimate the factor loadings. We choose a couple of different smooth functions for each value of θ and in both types of wavelets for comparison purposes. The following functions are considered: i) $\lambda_{1,12}(t) = 0.4\cos -3\pi t$, and ii) $\lambda_{2,8}(t) = 0.6(0.7\sqrt{t} - 0.5\sin 1.2\pi t)$, where $\lambda_{1,12}$ indicates the loading of the first factor of the cross-sectional unit $i = 12$, and $\lambda_{2,8}$, the loadings of the second factor of the unit $i = 8$. Figures display the actual and estimated time-varying loadings as well as their bootstrap confidence interval at 95% with $B = 100$ replications, following de A. Moura et al. (2012). In such figures, we can see that the methodology works well independently of the value taken in θ.

5 Application

In this section, we provide an application of the model proposed to study comovements in loads and prices of the Nord Pool power market.

Nord Pool runs the leading power market in Europe and operates in the day-ahead and intraday markets. Elspot is the day-ahead auction market, where participants act in a double auction and submit their supply and demand orders (spot prices and loads) for each individual hour of the next day. The market is in equilibrium when demand and supply curves intersect to each other at the system prices and loads for each hour. The hourly system prices and loads series are announced as 24 dimensional vectors which are determined simultaneously. See Bredesen and Nilsen (2013) for more information about the operation and history of the market.

Electricity markets have particular features that do not exist in another type of commodity market. Particularly, the non-storability of electricity provokes that the price behavior shows an excessive volatility, possible negative prices and many spikes along time. These characteristics are intrinsically more linked to prices than to loads which make them very different, however both series include strong intraday, weekly, and yearly seasonality, see Weron (2007).

Hourly electricity prices and loads have been mostly studied by univariate time series methods, see Weron (2014) for a rich review. Until the last years, some authors have explored these series by multivariate techniques as high dimensional factor models in different power markets. Using data from the Iberian Electricity Market (MIBEL), seasonal factor have been extracted in the works of Alonso et al. (2011) and Garcıa-Martos et al. (2012), whereas Alonso et al. (2016) propose to employ model averaging of factor models to improve the performance of forecasting. The Pennsylvania - New Jersey - Maryland (PJM) interconnection market is studied by Maciejowska and Weron (2015), who estimate factor models for forecasting evaluation using hourly and zonal prices. Furthermore, the Nord Pool power market is studied in the works of Ergemen et al. (2016), who study the long-term relationship between system prices and loads, and Rodríguez-Caballero and Ergemen (2017) who use regional prices in a multi-level setting. However, all these empirical studies consider that factor loadings do not vary along time. In this sense, to the best of our knowledge, we are the first to consider a factor model with time-varying loadings for the power market.

We consider a balanced panel data set consisting of $N = 24$ hourly prices and loads for each day from 13th March 2016 to 31th December 2018, yielding a total of $T = 1024$ daily observations in each hour. The series are downloaded from the Nord Pool ftp server and prices are denominated in Euros per Mwh of load. Figures 3 and 4 display six time series in logs from which we can observe some characteristics in both time series. First, electricity system prices and loads vary differently over the months with a common pattern in the evolution of hourly series. Second, the price series show many spikes which are related to the own features of the commodity as explained above. Third, the seasonal variation is stronger in loads than in prices series even if this component is present in both as discussed in Ergemen et al. (2016). Fourth, electricity prices and loads have nonstationary performances, see e.g. Haldrup et al. (2010), Haldrup and Nielsen (2006), and Koopman et al. (2007). These papers focus on the feature of electricity prices and loads have autocorrelation functions that decay at a hyperbolic rate, suggesting the use of fractionally integrated processes. In this paper, we keep in the $I(1)$ case to focus on the main ideas. Then, we set our estimation on the non-stationary approach discussed in the section 2.3. The possibility of long-memory dynamics is beyond the scope of the present paper and is not further explored.

We estimate the model in (12) with $r = 2$ for modeling the commonality of hourly system prices and loads. The literature has consistently used two common factors in both prices and loads. Since we work with heteroscedastic time series, we use (in first differences) the procedure proposed by Alessi et al. (2010) which introduce a tuning multiplicative constant in the penalty function to improve the criteria of Bai and Ng (2002). We also find two common factors in prices and loads which are in line with the literature. Using these factors, we explain 95% of the variation in the panel of electricity prices and 97% of electricity loads. These percentages are slightly higher than that found in Ergemen et al. (2016) but our period of time is much shorter than the one covered in their study. In this respect, what may be provoking such differences is that we do not cover the years in which the Nord Pool power market experienced a number of events in the delineation and market infrastructure.

Besides the estimated model, our study differs of Ergemen et al. (2016) and Rodríguez-Caballero and Ergemen (2017) in two aspects: first, we do not seasonally adjust and detrend each panel element of prices and loads as they do in their study. This is because we find that the seasonality is strongly extracted not only in common factors but also in factor loadings which evolve along time in our approach. Second, we do not follow their estimation procedure based on a fractionally differencing and integrating back strategy. Instead, we follow the procedure proposed by Peña and Poncela (2006) using the common eigenstructure of the generalized covariance matrices as discussed before. Using the methodlogy proposed in that paper, we set that the first common factor in loads and prices are nonstationary, $I(1)$, whereas that the second common factors in both series are stationary, $I(0)$. These findings are in line with the non-stationary literature on electricity modeling, see e.g. Ergemen et al. (2016), Rodríguez-Caballero and Ergemen (2017), for instance.

Figures 5 and 8 display estimates of common factors of hourly system loads and prices, respectively. In addition, Figures 6 and 7 show time-varying loadings of the first and second common factors of hourly system loads, respectively, while Figures 9 and 10 show the time-varying loadings estimates of hourly system prices. We only display the results for six hours for exposition purposes: 04:00, 08:00, 12:00, 16:00, 20:00, and 24:00 hrs. Note that with these hours, we can observe different performances of loadings between working and non-working hours, which exemplify the intrinsic nature of the market. We also estimate the standard factor model by PCA for comparison purposes. Static factor loadings are represented in Figures 6, 7, 9 and 10 by horizontal red lines. Common factors by static and time-varying approaches overlap, consequently static factors are not displayed.

As seen in Figure 5, the first factor captures the strong seasonal component showing possible weekly, and monthly periodicity in the hourly system loads. Empirical studies have also documented this regular behavior in such frequencies both in the Nord pool power market and in other energy markets. The second factor seems to capture mainly a kind of weekly variability, which occurs mostly during working hours as seen in Figure 3.

A further inspection on Figure 6 indicates that factor loadings corresponding to the first common factor of system loads are all positive showing smooth variations along time. Loadings have positive peaks more frequent during the summer. In this regard, time-varying loadings capture a kind of readjustment of system loads when the demand reaches a maximum level (in winter) and a minimum level (in summer). This makes sense with the nature of this commodity. Moreover, time-varying loadings in the remaining months remain stable, oscillating smoothly around the static factor loading. In turn, factor loadings corresponding to the second common factor of system loads do not have a regular behavior in contrast to those of the first factor, see Figure 7. However, these loadings are positive during night hours and negative during working hours as also discussed in Ergemen et al. (2016). This indicates that the second common factor plays a positive role during non-working hours and have negative contributions along the working hours. In this sense, such changes in the demand of electricity around the day may provoke the strong variability of the second common factor as discussed before.

With respect to hourly system prices, at a first glance, Figure 8 shows that estimated common factors exhibit some stylized facts of electricity prices. Volatility clustering is being captured by the first factor, while, excessive price spikes occurring in 2018 by March, July, and August are extracted by both common factors, but mostly by the second one. Similarly to the case of system loads, the factor loadings of the first common factor of system prices are all positive (Figure 9), whereas those of the second factor are positive during the night and negative during working hours (Figure 10). However, unlike system loads, we do not find that factor loadings in the case of prices have regular periodic movements, although in general, loadings in some consecutive hours are very similar (5-7, 15-18 hrs, for instance). The latter indicates that the impact of common factors to explain price variability will be very similar as the hours approach.

In conclusion, this empirical study leaves open many possibilities for future research. First, testing whether loadings vary or if they are constant over time would improve the specification of factor models and could help to understand market behavior. Second, in the context of the liberalization of energy markets, short-term forecasting of prices and loads are essential, then, in a subsequent paper, we are investigating if loads/prices forecasts obtained by time-varying factor models are better than those provided by standard or dynamic factor models. Third, as discussed in the Monte Carlo study, the factor loadings using the wavelet D8 perform better than Haar in situation where time series evolve smoothly as in the case of system loads. In this sense, a deeper study on electricity prices could help us understand if the wavelet Haar represents a better choice for modeling the performance of factor loadings, since electricity prices are more volatile than loads.

6 Concluding remarks

In the last years, factor models has been widely used in many branches due to the flexibility of treating high dimensional panels. However, most of the standard approaches have some limitations when assuming that factor loadings are invariant along time. In this respect, we relaxed such an assumption and focused on the study of factor models with time-varying loadings.

In this paper, we proposed a two-step procedure based on GLS with wavelet functions to estimate factor models, whose loadings are defined as smooth and continuous functions of time. We consider stationary as well as non-stationary variables to cover more possibilities of application. The finite-sample properties of our estimator were supported by some Monte Carlo simulations. Our findings indicate that the methodology proposed performed very well regardless of size distortion between N and T and even in relatively small samples. We also found that Wavelet D8 estimates were more attractive than wavelet Haar estimates due to its smoothness. This indicates that Wavelet D8 should provide better estimations when factor loadings do not have sudden changes.

Finally, we motivated the empirical relevance of allowing for time-varying loadings in factor models by studying the electricity loads and prices of the Nord Pool power market. At a first glance, we found that factor loadings seem to vary over time, which has not been discussed in the empirical literature on electricity markets. Furthermore, we also found that the strong seasonality, which have been documented in many other studies, is extracted not only by common factors, but also by factor loadings, which may improve the forecasting of electricity prices and loads.

References

Alessi, L., Barigozzi, M., and Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters*, 80(23):1806–1813.

Alonso, A., Bastos, G., and García-Martos, C. (2016). Electricity price forecasting by averaging dynamic factor models. *Energies*, 9(8):600.

Alonso, A. M., García-Martos, C., Rodríguez, J., and Jesús Sánchez, M. (2011). Seasonal dynamic factor analysis and bootstrap inference: application to electricity market forecasting. *Technometrics*, 53(2):137–151.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, pages 135–171.

Bai, J. and Han, X. (2016). Structural changes in high dimensional factor models. *Frontiers of Economics in China; Beijing*, 11(1):9–39.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bai, J. and Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1177.

Bai, J., Ng, S., et al. (2008). Large dimensional factor analysis. *Foundations and Trends® in Econometrics*, 3(2):89–163.

Barigozzi, M., Lippi, M., and Luciani, M. (2016). Non-stationary dynamic factor models for large datasets. *arXiv preprint arXiv:1602.02398*.

Bates, B. J., Plagborg-Møller, M., Stock, J. H., and Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177(2):289–304.

Bredesen, H. and Nilsen, T. (2013). *Power to the people. The first 20 years of Nordic power-market integration.* Nord Pool Spot and Nasdaq OMX Commodities.

Chiann, C. and Morettin, P. A. (2005). Time-domain estimation of time-varying linear systems. *Journal of Nonparametric Statistics*, 17(3):365–383.

Cohen, A. and Ryan, R. D. (1995). *Wavelets and multiscale signal processing*. Springer.

Dahlhaus, R., Eichler, M., and Sandkühler, J. (1997a). Identification of synaptic connections in neural ensembles by graphical models. *Journal of neuroscience methods*, 77(1):93–107.

Dahlhaus, R. et al. (1997b). Fitting time series models to nonstationary processes. *The annals of Statistics*, 25(1):1–37.

de A. Moura, M. S., Morettin, P. A., Toloi, C. M. C., and Chiann, C. (2012). Transfer function models with time-varying coefficients. *Journal of Probability and Statistics*, 2012:31.

Eichler, M., Motta, G., and Von Sachs, R. (2011). Fitting dynamic factor models to non-stationary time series. *Journal of Econometrics*, 163(1):51–70.

Eickmeier, S., Lemke, W., and Marcellino, M. (2015). Classical time varying factor-augmented vector autoregressive models estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):493–533.

Ergemen, Y. E., Haldrup, N., and Rodríguez-Caballero, C. V. (2016). Common long-range dependence in a panel of hourly nord pool electricity prices and loads. *Energy Economics*, 60:79–96.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82(4):540–554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2004). The generalized dynamic factor model consistency and rates. *Journal of Econometrics*, 119(2):231–255.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, 100(471):830–840.

Garcıa-Martos, C., Rodrıguez, J., and Sanchez, M. (2012). Forecasting electricity prices by extracting dynamic common factors: application to the iberian market. *IET Generation, Transmission & Distribution*, 6(1):11–20.

Haldrup, N., Nielsen, F., and Nielsen, M. (2010). A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics*, 32(5):1044–1058.

Haldrup, N. and Nielsen, M. Ø. (2006). A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1):349–376.

Koopman, S. J., Ooms, M., and Carnero, M. A. (2007). Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association*, 102(477):16–27.

Maciejowska, K. and Weron, R. (2015). Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships. *Computational Statistics*, 30(3):805–819.

Mikkelsen, J. G., Hillebrand, E., and Urga, G. (2018). Consistent estimation of time-varying loadings in high-dimensional factor models. *Journal of Econometrics*.

Motta, G., Hafner, C. M., and von Sachs, R. (2011). Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory*, 27(6):1279–1319.

Pena, D. and Box, G. E. (1987). Identifying a simplifying structure in time series. *Journal of the American statistical Association*, 82(399):836–843.

Peña, D. and Poncela, P. (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, 136(4):1237–1257.

Porto, R., Morettin, P., and Aubin, E. C. Q. (2008). Wavelet regression with correlated errors on a piecewise holder class. *Statistics & Probability Letters*, 78(16):2739–2743.

Rodríguez-Caballero, C. V. and Ergemen, Y. E. (2017). Estimation of a dynamic multilevel factor model with possible long-range dependence. Technical report, Universidad Carlos III de Madrid. Departamento de Estadística.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179.

Su, L. and Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, 198(1):84–101.

Vidakovic, B. (2009). *Statistical modeling by wavelets*, volume 503. John Wiley & Sons.

Weron, R. (2007). *Modeling and forecasting electricity loads and prices: A statistical approach*, volume 403. John Wiley & Sons.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030 – 1081.

Appendix

A Technical appendix

A.1 Proof of Theorem 1

Proof. Let $\mathbf{Y} = (Y_1, \ldots, Y_T)'$ be a $T \times N$ matrix, and let V_{NT} be a $r \times r$ diagonal matrix composed by the r largest eigenvalues of the matrix $(NT)^{-1}\mathbf{Y}\mathbf{Y}'$. By definition of eigenvectors and eigenvalues, we have

$$\frac{1}{NT}\mathbf{Y}\mathbf{Y}'\tilde{\mathbf{F}} = \tilde{\mathbf{F}}V_{NT} \iff \frac{1}{NT}\mathbf{Y}\mathbf{Y}'\tilde{\mathbf{F}}V_{NT}^{-1} = \tilde{\mathbf{F}}, \tag{17}$$

where $\tilde{\mathbf{F}}'\tilde{\mathbf{F}} = \mathbb{I}_r$. The model defined in (1) and (2) can be re-written as

$$\begin{aligned} \mathbf{Y}_t &= [\mathbf{\Lambda}_0 + \mathbf{\Lambda}(t)]\mathbf{F}_t + \mathbf{e}_t \\ &= \mathbf{\Lambda}_0\mathbf{F}_t + \mathbf{\Lambda}(t)\mathbf{F}_t + \mathbf{e}_t \\ &= \mathbf{\Lambda}_0\mathbf{F}_t + \mathbf{w}_t + \mathbf{e}_t, \end{aligned}$$

where $\mathbf{w}_t = \mathbf{\Lambda}(t)\mathbf{F}_t$. Now, we define the following $T \times N$ matrices, $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_T)'$ and $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_T)'$. Note that the model in (1), can also be re-written in matrix form as

$$\mathbf{Y} = \mathbf{F}\mathbf{\Lambda}_0' + \mathbf{w} + \mathbf{e}.$$

Consequently, after taking products we get

$$\mathbf{Y}\mathbf{Y}' = \mathbf{F}\mathbf{\Lambda}_0'\mathbf{\Lambda}_0\mathbf{F}' + \mathbf{F}\mathbf{\Lambda}_0'(\mathbf{e} + \mathbf{w})' + (\mathbf{e} + \mathbf{w})\mathbf{\Lambda}_0\mathbf{F}' + (\mathbf{e} + \mathbf{w})(\mathbf{e} + \mathbf{w})'.$$

Then, from definitions of $\tilde{\mathbf{F}}_t$ in (17) and the matrix rotation H, we can write for a fixed t

$$\begin{aligned} \tilde{\mathbf{F}}_t - H'\mathbf{F}_t &= (NT)^{-1}V_{NT}^{-1}\tilde{\mathbf{F}}'\mathbf{Y}\mathbf{Y}_t' - V_{NT}^{-1}(\tilde{\mathbf{F}}'\mathbf{F}/T)(\mathbf{\Lambda}_0'\mathbf{\Lambda}_0/N)\mathbf{F}_t \\ &= \tfrac{V_{NT}^{-1}}{NT}[\tilde{\mathbf{F}}'(\mathbf{F}\mathbf{\Lambda}_0'\mathbf{\Lambda}_0\mathbf{F}_t + \mathbf{F}\mathbf{\Lambda}_0'(w_t + e_t) + (w + e)\mathbf{\Lambda}_0\mathbf{F}_t + (e + w)(e_t + w_t)) \\ &\quad -(\tilde{\mathbf{F}}'\mathbf{F})(\mathbf{\Lambda}_0'\mathbf{\Lambda}_0)\mathbf{F}_t] \\ &= \tfrac{V_{NT}^{-1}}{NT}[\underbrace{\tilde{\mathbf{F}}'\mathbf{F}\mathbf{\Lambda}_0'e_t}_{D_{1t}} + \underbrace{\tilde{\mathbf{F}}'e\mathbf{\Lambda}_0\mathbf{F}_t}_{D_{2t}} + \underbrace{\tilde{\mathbf{F}}'ee_t}_{D_{3t}} + \underbrace{\tilde{\mathbf{F}}'\mathbf{F}\mathbf{\Lambda}_0'w_t}_{D_{4t}} + \underbrace{\tilde{\mathbf{F}}'w\mathbf{\Lambda}_0\mathbf{F}_t}_{D_{5t}} + \underbrace{\tilde{\mathbf{F}}'ww_t}_{D_{6t}} + \\ &\quad \underbrace{\tilde{\mathbf{F}}'ew_t}_{D_{7t}} + \underbrace{\tilde{\mathbf{F}}'we_t}_{D_{8t}}] \\ &= V_{NT}^{-1}\textstyle\sum_{i=1}^{8} D_{it}. \end{aligned}$$

Then, after applying squared norms in both sides, adding on t, and dividing by T in $V_{NT}^{-1}\sum_{i=1}^{8} D_{it}$, we get by the Loèv inequality

$$\frac{1}{T}\sum_{t=1}^{T}\|\tilde{\mathbf{F}}_t - H'\mathbf{F}_t\|^2 \leq \|V_{NT}^{-1}\|^2 8\sum_{i=1}^{8}\left(\frac{1}{T}\sum_{t=1}^{T}\|D_{it}\|^2\right) \tag{18}$$

where

$$\begin{array}{ll} D_{1t} = \tilde{\mathbf{F}}'\mathbf{F}\mathbf{\Lambda}_0' e_t/NT & D_{2t} = \tilde{\mathbf{F}}' e\mathbf{\Lambda}_0\mathbf{F}_t/NT \\ D_{3t} = \tilde{\mathbf{F}}' ee_t/NT & D_{4t} = \tilde{\mathbf{F}}'\mathbf{F}\mathbf{\Lambda}_0' w_t/NT \\ D_{5t} = \tilde{\mathbf{F}}' w\mathbf{\Lambda}_0\mathbf{F}_t/NT & D_{6t} = \tilde{\mathbf{F}}' ww_t/NT \\ D_{7t} = \tilde{\mathbf{F}}' ew_t/NT & D_{8t} = \tilde{\mathbf{F}}' we_t/NT \end{array}$$

Then, from Mikkelsen et al. (2018) (Lemma A.1), V_{NT} converges to a definite positive matrix, therefore $\|V_{NT}^{-1}\| = O_p(1)$. Considering Assumptions A-E and properties of principal components, for each term $D_{it}, i = 1, \ldots, 8$, we have

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}\|D_{1t}\|^2 &= O_p(N^{-1}), \\ T^{-1}\sum_{t=1}^{T}\|D_{2t}\|^2 &= O_p(N^{-1}T^{-1}), \\ T^{-1}\sum_{t=1}^{T}\|D_{3t}\|^2 &= O_p(N^{-1}T^{-1}), \\ T^{-1}\sum_{t=1}^{T}\|D_{4t}\|^2 &= O_p(N^{-2}K_{1NT}), \\ T^{-1}\sum_{t=1}^{T}\|D_{5t}\|^2 &= O_p(N^{-2}T^{-2}K_{2NT}), \\ T^{-1}\sum_{t=1}^{T}\|D_{6t}\|^2 &= O_p(N^{-2}T^{-2}K_{3NT}), \\ T^{-1}\sum_{t=1}^{T}\|D_{7t}\|^2 &= O_p(N^{-2}K_{1NT}), \\ T^{-1}\sum_{t=1}^{T}\|D_{8t}\|^2 &= O_p(N^{-2}K_{1NT}). \end{aligned}$$

Finally, the right-hand side of equation (18) is a sum of variables with orders $\left\{\frac{1}{N}, \frac{1}{NT}, \frac{K_{1NT}}{N^2}, \frac{K_{2NT}}{N^2T^2}, \frac{K_{3NT}}{N^2T^2}\right\}$, respectively, and the proof is now completed. □

B Tables

Table 1: $T \in \{512, 1024, 2048\}$, $N \in \{20, 30\}$, and $r = 2$. The measure of consistency of the estimated unobservable factors as well as the estimated factor loadings are presented in the report.

			Wavelet functions											
			Haar						D8					
			$\theta = 0$		$\theta = 0.5$		$\theta = 1$		$\theta = 0$		$\theta = 0.5$		$\theta = 1$	
N	T	Γ_e	$R^2_{\tilde{F},F}$	MSE_m	$R^2_{\tilde{F},F}$	MSE_m	$R^2_{\tilde{F},F}$	MSE_m	$R^2_{\tilde{F},F}$	MSE_m	$R^2_{\tilde{F},F}$	MSE_m	$R^2_{\tilde{F},F}$	MSE_m
	512		0.9341	0.0103	0.8542	0.1970	0.7450	0.1562	0.9447	0.0011	0.8655	0.0023	0.7714	0.1191
20	1024	$Diag$	0.9421	0.0076	0.8135	0.0900	0.7501	0.1202	0.9347	0.0082	0.8764	0.0011	0.7439	0.0876
	2048		0.9127	0.0020	0.8236	0.0053	0.7312	0.0931	0.9470	0.0005	0.8455	0.0009	0.7546	0.0045
	512		0.9236	0.0098	0.8456	0.0128	0.7611	0.1091	0.9560	0.0071	0.8521	0.0127	0.8131	0.1093
30	1024	$Diag$	0.9547	0.0051	0.8125	0.0090	0.7112	0.0859	0.9487	0.0047	0.8671	0.0079	0.7963	0.0759
	2048		0.9341	0.0011	0.8268	0.0027	0.7324	0.0738	0.9498	0.0020	0.8601	0.0060	0.8213	0.0028
	512		0.8731	0.0672	0.6932	0.0891	0.6711	0.2201	0.8911	0.0128	0.6871	0.0593	0.7242	0.1223
20	1024	$Toep$	0.8634	0.0501	0.6853	0.0702	0.7012	0.1901	0.9026	0.0100	0.7001	0.0337	0.6832	0.0693
	2048		0.8911	0.0137	0.7723	0.0433	0.7321	0.1642	0.8971	0.0091	0.7307	0.0108	0.6841	0.0031
	512		0.8531	0.0472	0.7984	0.0621	0.6812	0.0912	0.9021	0.0117	0.7815	0.0311	0.7483	0.1114
30	1024	$Toep$	0.8671	0.0231	0.7730	0.0539	0.7122	0.0734	0.9126	0.0090	0.7270	0.0276	0.7126	0.0554
	2048		0.8451	0.0111	0.8200	0.0311	0.7212	0.0819	0.8732	0.0068	0.7305	0.0109	0.7531	0.0037

Notes: The DGP is $Y_{it} = \boldsymbol{\lambda}_i'(t)\mathbf{F}_t + e_{it}$, where $i \in \{20, 30\}$ and $T \in \{512, 1024, 2048\}$, $F_{kt}(1 - \theta_k B) = \eta_{kt}$, with $k \in \{1, 2\}$. Idiosyncratic terms are independently generated as $\boldsymbol{\eta}_t \sim \mathcal{N}_1(0, diag\{1-\theta_1^2, 1-\theta_2^2\})$ with $\theta \in \{0, 0.5, 1\}$ defining the degree of serial correlation among factors, and $\mathbf{e}_t \sim \mathcal{N}_N(0, \Gamma_e)$ with Γ_e defined as diagonal and Toeplitz matrices. Common factors are estimated by principal components in cases with $\theta < 1$, while by the procedure of Peña and Poncela (2006) in the case with $\theta = 1$. $R^2_{\tilde{F},F}$ is the R^2 of a regression of actual on estimates factors. MSE_m is the median of the MSE between the actual and estimated factor loadings. Haar and D8 wavelet functions are used in the study. All experiments are based on 1000 replications.

C Figures

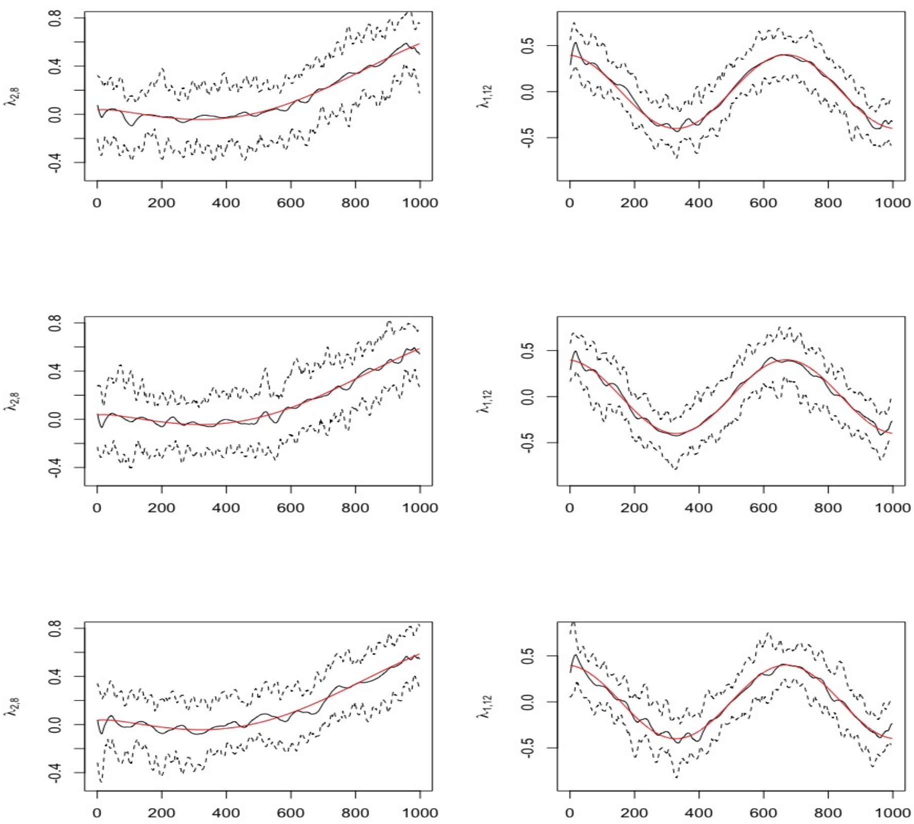

Figure 1: Comparison between the actual factor loadings in solid red line, the estimated factor loadings in solid black line, and Bootstrap confidence interval at 95% in dashed black line. By column, from left to right: λ_{12}, and λ_{28}. By row, from top to bottom: $\theta = 0$, $\theta = 0.5$, and $\theta = 1$. $\Gamma_e = Toep$, $N = 20$, $T = 1024$ and Wavelet Haar.

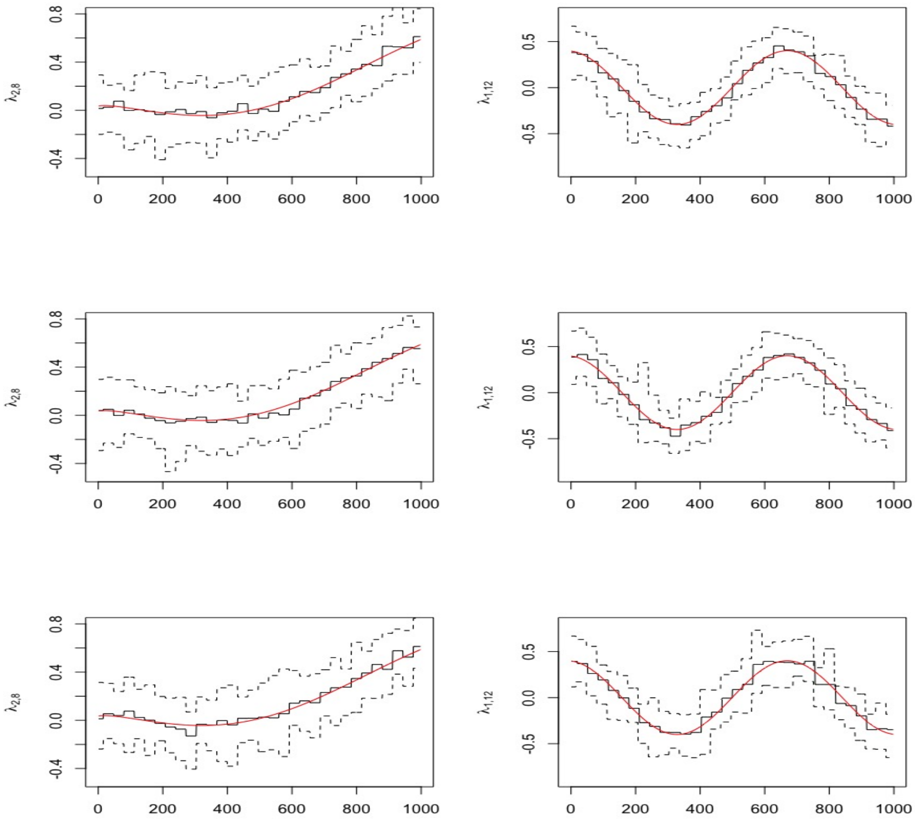

Figure 2: Comparison between the actual factor loadings in solid red line, the estimated factor loadings in solid black line, and Bootstrap confidence interval at 95% in dashed black line. By column, from left to right: λ_{12}, and λ_{28}. By row, from top to bottom: $\theta = 0$, $\theta = 0.5$, and $\theta = 1$. $\Gamma_e = Toep$, $N = 20$, $T = 1024$ and Wavelet D8.

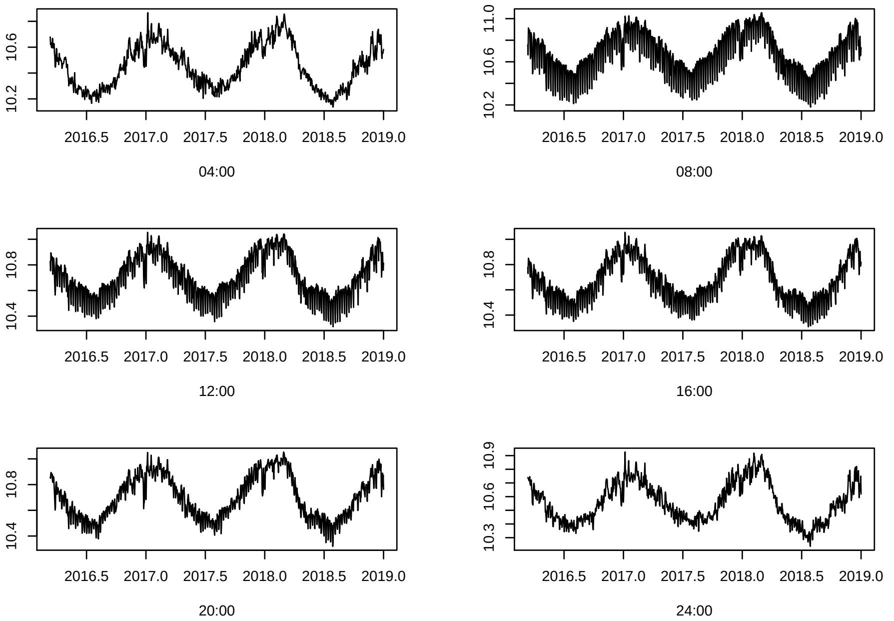

Figure 3: Hourly system loads in logs for six different hours showing working and non-working hours performances, 12 March 2016 to 31 December 2018.

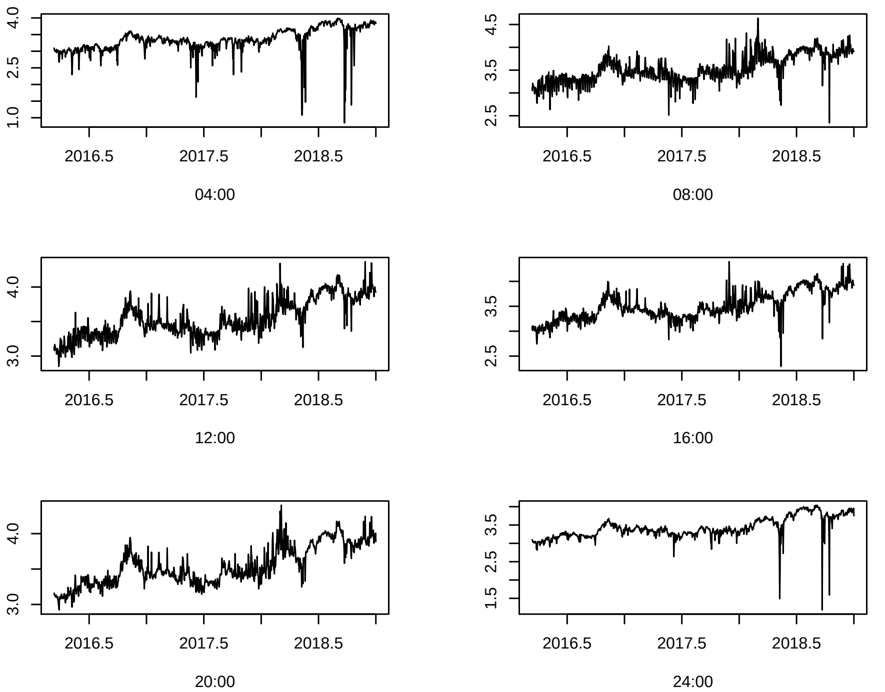

Figure 4: Hourly system prices in logsfor six different hours showing working and non-working hours performances, 12 March 2016 to 31 December 2018.

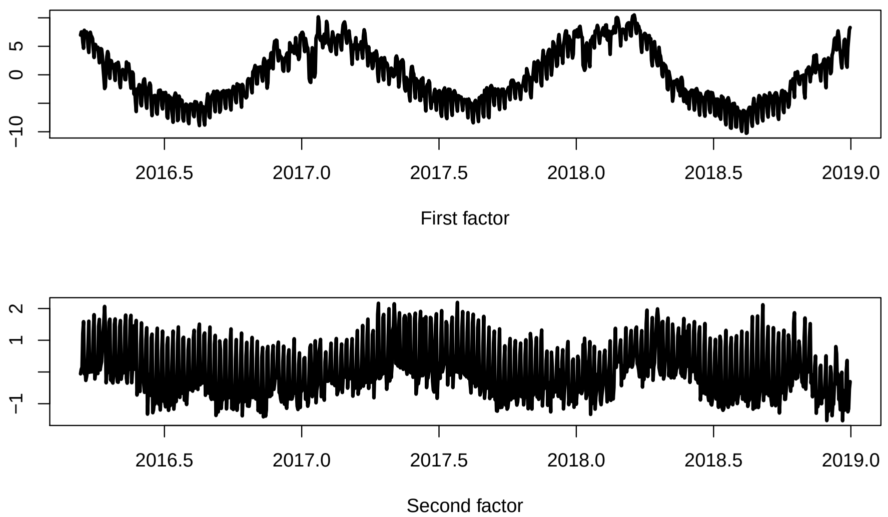

Figure 5: Common factors of hourly system loads.

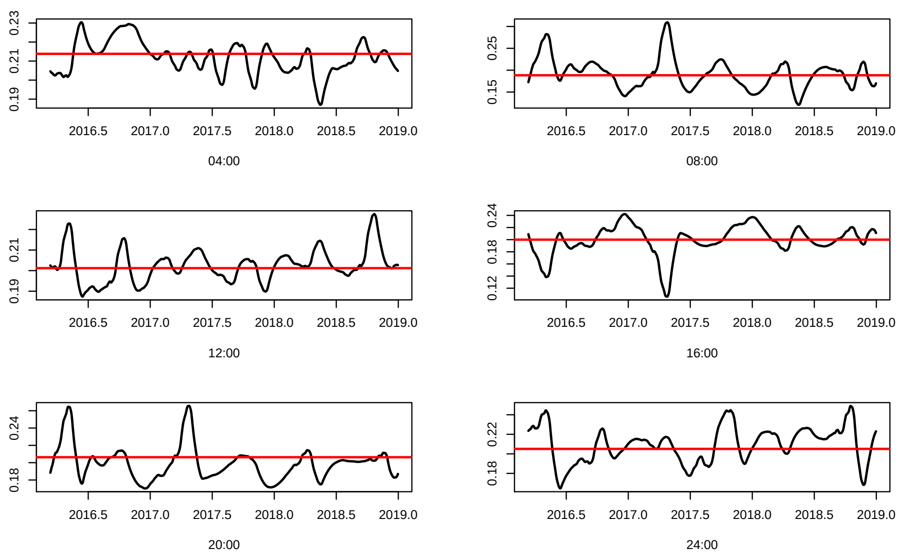

Figure 6: Time-varying loadings of the first factor of hourly system loads. The horizontal red line represents the level of loadings by the standard method.

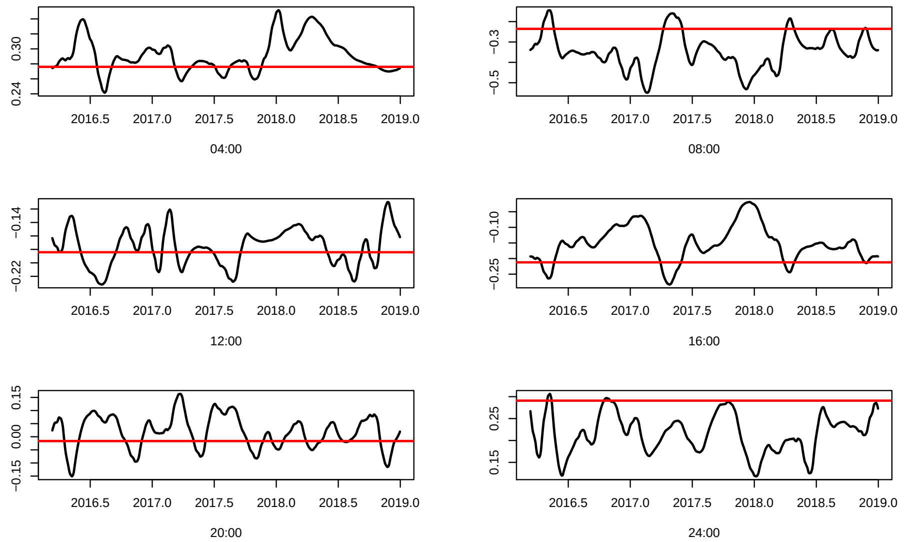

Figure 7: Time-varying loadings of the second factor of hourly system loads. The horizontal red line represents the level of loadings by the standard method.

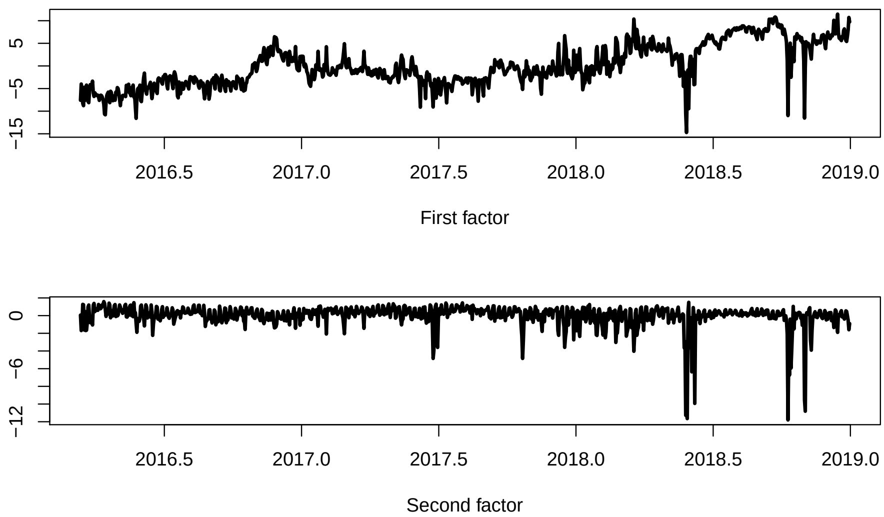

Figure 8: Common factors of hourly system prices.

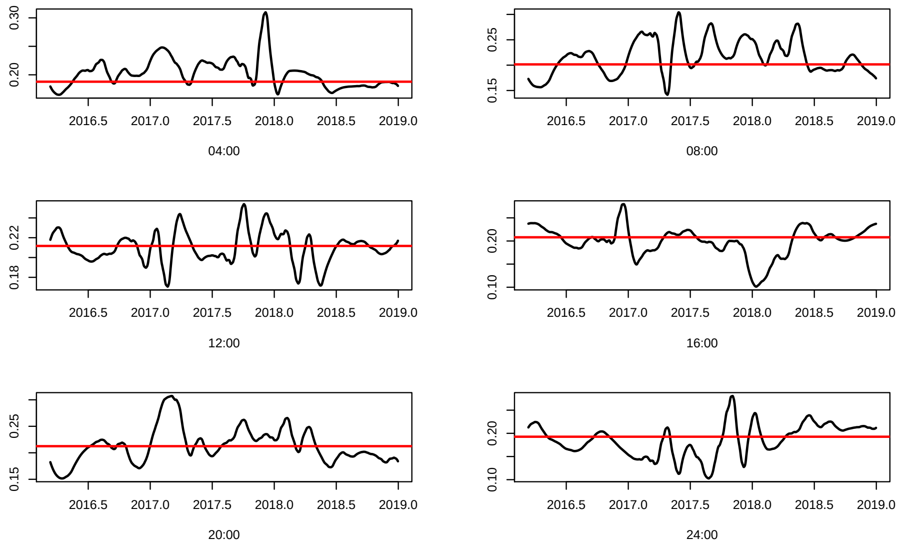

Figure 9: Time-varying loadings of the first factor of hourly system prices. The horizontal red line represents the level of loadings by the standard method.

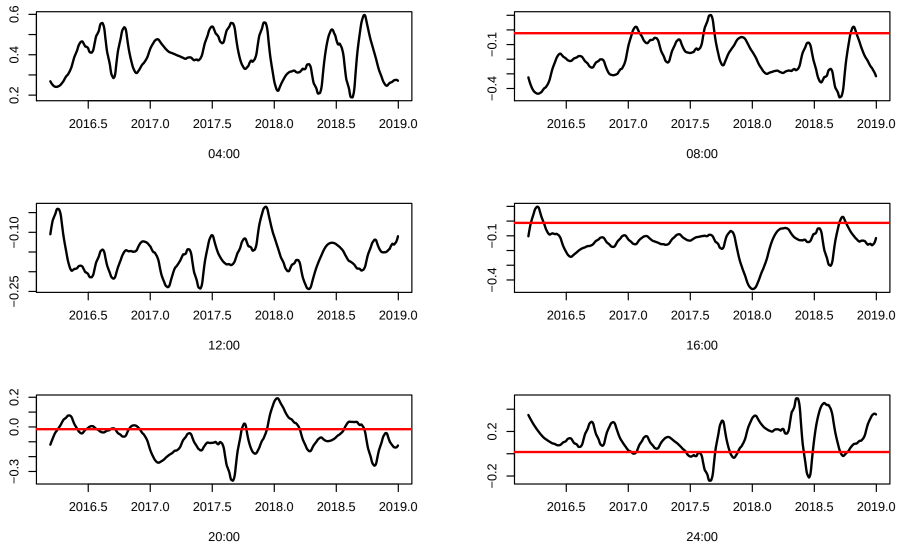

Figure 10: Time-varying loadings of the second factor of hourly system prices. The horizontal red line represents the level of loadings by the standard method.

Research Papers 2019

Center for Research in Econometric Analysis of Time Series