Estimation of the Semiparametric Factor Model: Application to Modelling Time Series of Electricity Spot Prices.

Liebl, Dominik

University of Cologne

2010

Estimation of the Semiparametric Factor Model: Application to Modelling Time Series of Electricity Spot Prices.

Dominik Liebl *

* Address:
Seminar für Wirtschafts– und Sozialstatistik, Lehrstuhl Prof. Mosler, Universität zu Köln, Albertus–Magnus–Platz, 50923 Köln Germany
liebl@statistik.uni-koeln.de

1 Introduction

Classical univariate and multivariate time series models have problems to deal with the high variability of hourly electricity spot prices. We propose to model alternatively the daily mean electricity supply functions using a dynamic factor model. And to derive, subsequently, the hourly electricity spot prices by the evaluation of the estimated supply functions at the corresponding hourly values of demand for electricity. Supply functions are price (EUR/MWh) functions, that increase monotonically with demand for electricity (MW). Apart from this new conceptual approach, that allows us to represent the auction design of energy exchanges in a most natural way, our main contribution is an extraordinary simple algorithm to estimate the factor structure of the dynamic factor model. We decompose the time series into a functional "spherical component" and an univariate "scaling component". The elements of the spherical component are all standardized having unit size such that we can robustly estimate the factor structure. This algorithm is much simpler than procedures suggested in the literature. In order to use a parsimonious labeling we will refer to the daily mean supply curves simply as "price curves".

The Dynamic Semiparametric Factor Model (DSFM) of [4] and the follow up application to electricity spot prices in [3] are close to our approach, but there are two important differences. Firstly, the authors model the hourly spot prices directly as a multivariate time series and therein fail to mirror the auction design (i.e. the data generating process) at electricity exchanges. As a result, they are able to explain only about 80% of the variation in hourly spot prices at the European Electricity Exchange while we are able to explain over 98% of the variation using the same number of factors. Secondly, they use an iterating optimization algorithm to estimate the factor structure, whereas we use principal component analysis for sparse functional data [6] to estimate the factor structure of the spherical component. And we show that the estimated factor structure of the spherical component is also the factor structure of the original series.

2 Functional Dynamic Factor Model

We model the prices, Y_{ti}, as observations of an underlying smooth price curve, X_t, such that

$$Y_{ti} = X_t(u_{ti}) + \varepsilon_{ti} \quad \text{with } t = 1, 2, \ldots, T. \tag{1}$$

Where $X_t(.)$ is a smooth monotone random function of adjusted demand[1] $u \in \mathcal{U}$ with $\mathcal{U}$ being a closed and bounded subspace of $\mathbb{R}$. We will set, without loss of generality, $\mathcal{U} = [0, 1]$. The index $i = 1, \ldots, N_t$ in u_{ti} refers to the i-th order statistic of the observed hourly adjusted demand values, u_{th}. The noise term, ε_{ti}, is assumed to be independently distributed for each t and i, with $\mathrm{E}(\varepsilon_{ti}) = 0$ and $\mathrm{Var}(\varepsilon_{ti}) = \sigma_\varepsilon^2$. An example of some raw data vectors $\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{tN_t})'$ can be seen in figure 1. Note that, some prices Y_{ti} have to be treated as outliers, and we use N_t to refer to the amount of prices per day t, that is used in the estimation procedure. An example of outlier prices can be seen in the left panel of figure 3.

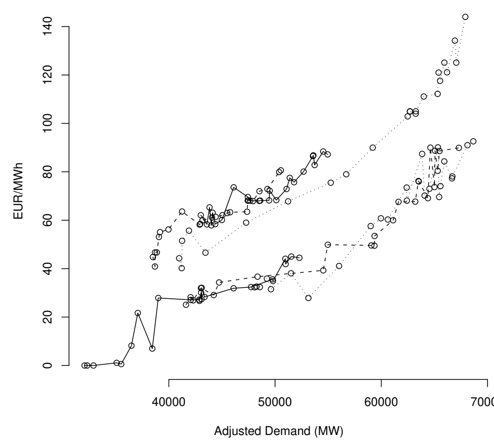

Figure 1: Three consecutive days from two different arbitrary weeks.

Dynamic factor models are a very successful approach to analyze high dimensional time series data. Our case is a special case of the generalized dynamic factor models considered in [2] and corresponds to the dynamic factor model in [4]. The factor structure, F, consists of unknown non parametric functions, $f_1, \ldots, f_K$, that have to be estimated from

[1] Adjusted demand means: Original demand values minus electricity from wind-power. Because of its privileged status of renewable energy sources, the market price of electricity is not valid for wind-power.

the data. The $K < \infty$ functionals of the estimated factor structure, $\hat{F} = [\hat{f}_1, \ldots, \hat{f}_K]$, are required to be mutually orthonormal to each other and to be an optimal empirical basis such that

$$X_t \approx \sum_{k=1}^{K} \hat{\beta}_{tk} \hat{f}_k = \hat{\beta}_t' \hat{F}. \tag{2}$$

More precisely, the factor structure, $\hat{F} = [\hat{f}_1, \ldots, \hat{f}_K]$, shall define the best possible projection from the space $\mathcal{H}_T \subset L^2(\mathcal{U})$ spanned by the sampled functions, $X_1, \ldots, X_T$, into a K dimensional subspace of $\mathcal{H}_T$, where "best possible" is understood with respect to the mean squared error sense,

$$\sum_{t=1}^{T} ||X_t - \sum_{k=1}^{K} \hat{\beta}_{tk} \hat{f}_k||_2^2 = \min_{v_1,\ldots,v_K} \sum_{t=1}^{T} \min_{\vartheta_1,\ldots,\vartheta_K} ||X_t - \sum_{k=1}^{K} v_{tk} \vartheta_k||_2^2, \tag{3}$$

with respect to all possible $\vartheta_1, \ldots, \vartheta_t \in L^2(\mathcal{U})$ and $v_{t1}, \ldots, v_{tK} \in \mathbb{R}$. We use $||.||_2$ to denote the L2-norm, in its functional version $||f||_2 = \sqrt{\int_0^1 f(u)^2 du}$ for functions $f \in L^2(\mathcal{U})$, and its euclidean version $||y||_2 = \sqrt{\sum_{i=1}^N y_i^2}$ for vectors $y \in \mathbb{R}^N$. Note that this definition of a factor structure, $\hat{F}$, is also fulfilled by any rigid rotations, $\hat{F}^* = \mathbf{T}\hat{F}$, where $\mathbf{T}$ is any orthonormal $K \times K$-matrix such that $\mathbf{T}\mathbf{T}' = \mathbf{T}'\mathbf{T} = I_K$.

It is well known that the first $K < \infty$ empirical eigenfunctions, let's say $f_{1T}, \ldots, f_{KT}$, of the sample covariance operator,

$$\begin{aligned} \rho_T\, g &= \int_0^1 \sigma_T(u,v)\, g(v) dv, \quad \text{for all } \; g \in L^2(\mathcal{U}), \\ \text{where} \quad \sigma_T(u,v) &= T^{-1} \sum_{t=1}^{T} X_t(u) X_t(v), \quad \text{with } \; u, v \in \mathcal{U}, \end{aligned}$$

can define such a best possible projection from the space $\mathcal{H}_T = \operatorname{span}(X_1, \ldots, X_T) \subset L^2(\mathcal{U})$ into a K-dimensional subspace of $\mathcal{H}_T$. In our general setting, where $(X_1, \ldots, X_T)$ is allowed to be any collection of functional random variables the sample covariance operator, $\rho_T\, g$, generally does not converge to a population counterpart and the empirical eigenfunctions and eigenvalues cannot be interpreted as variance components in the classical sense. This sample dependence of $F_T = [f_{1T}, \ldots, f_{KT}]$ is not different from other dynamic factor models as in [4].

Unfortunately, given the unrestrictive assumptions on the series (X_t), the spectral decomposition of the empirical covariance operator, $\rho_T\, g$, generally cannot be used to estimate a factor structure, F_T. As long as the process (X_t) is not stationary, its elements are likely to be of very different orders of magnitude, which will have a dramatic distortion effect on the sample covariance function, σ_T. But, contrary to the claim of the authors in [4], we do not need stationarity in order to use spectral decomposition of the sample covariance operator to estimate a factor structure for the functions $X_1, \ldots, X_T$.

Proposition 2.1 *Given the model in (2), if a factor structure $\hat{F}$ defines the best projection from the space $\mathcal{H}_T = span(X_1, \dots, X_T)$ into a K dimensional subspace $\mathcal{H}_T^K \subset \mathcal{H}_T$, then it also defines the best projection from the space $\mathcal{H}_T^* = span(\frac{X_1}{||X_1||_2}, \dots, \frac{X_T}{||X_T||_2})$ into the same K dimensional subspace $\mathcal{H}_T^K$.*

This proposition is trivially true, because $\mathcal{H}_T = \text{span}(X_1, \dots, X_T)$ is a vector space and therefore is closed under scalar multiplication, such that $\mathcal{H}_T = \mathcal{H}_T^*$. Different scales $X_t\, c_t$, with $c_t \neq 0$, will simply cause reciprocal scales of $\hat{\beta}/c_t$ in the minimization (3). As a consequence from proposition 2.1 we can also estimate a factor structure for the original series, (X_t), from the standardized series $(\frac{X_t}{||X_t||_2})$.

3 The Algorithm

The idea is to decompose the time series, $(\mathbf{Y}_t)$, into its "spherical" component that can be used to estimate the K-dimensional factor structure F and its "scaling" component that can be used to rescale the approximated spherical process to its original size.

Definition The spherical component of the factor model in equation (2) is given by the multivariate series,

$$\left(\frac{\mathbf{Y}_t - \mu_T(u_t)}{||\mathbf{Y}_t - \mu_T(u_t)||_2}\right)_{t=1,\dots,T}. \tag{4}$$

With $u_t = (u_{t1}, \dots, u_{N_t1})$ and $\mu_T = T^{-1}\sum_{t=1}^T X_t$ being the sample mean function.

Definition The scaling component is given by the univariate series,

$$(||\mathbf{Y}_t - \mu_T(u_t)||_2)_{t=1,\dots,T}\,. \tag{5}$$

From a mathematical perspective, it is not necessary to subtract the sample mean, $\mu_T \in \mathcal{H}_T = \text{span}(X_1, \dots, X_T)$, from the discretization vectors, $\mathbf{Y}_T$. This simply subtracts the constant vector $\bar{\hat{\beta}} = (T^{-1}\sum_t^T \hat{\beta}_{t1}, \dots, T^{-1}\sum_t^T \hat{\beta}_{tK})'$ from the process $(\hat{\beta}_t) = (\hat{\beta}_{t1}, \dots \hat{\beta}_{tK})'$. But, from a practical perspective, the subtraction of the sample mean, μ_T, helps to avoid rounding errors caused by floating point computation. Particularly, when the sizes of different vectors $\mathbf{Y}_t$ are of very different orders of magnitude, as in our application.

By construction, the elements of the spherical component, $\left(\frac{\mathbf{Y}_t - \mu_T(u_t)}{||\mathbf{Y}_t - \mu_T(u_t)||_2}\right)$, are all of the same order of magnitude, such that the factor structure, F, can be estimated by the spectral decomposition of the spherical sample covariance operator,

$$\begin{aligned}\tilde{\rho}_T\, g &= \int_0^1 \tilde{\sigma}_T(u,v)\, g(v)dv, \quad \text{for all } \; g \in L^2(\mathcal{U}),\\ \text{where} \quad \tilde{\sigma}_T(u,v) &= T^{-1}\sum_{t=1}^T \frac{\mathbf{Y}_t(u) - \mu_T(u)}{||\mathbf{Y}_t(u) - \mu_T(u)||_2}\,\frac{\mathbf{Y}_t(v) - \mu_T(v)}{||\mathbf{Y}_t(v) - \mu_T(v)||_2},\end{aligned}$$

without distortion effects. This estimation algorithm is by far less costly with respect to computation time and much simpler to implement than the iterative procedure in [4].

4 Application

The estimation of a factor structure, F, for the daily mean electricity supply functions, X_t, is made a bit more difficult by the sparseness of the data. The observed discretization points, $\mathbf{Y}_t$, of the price functions, X_t, are not uniformly distributed over the whole domain $\mathcal{U} = [0, 1]$, but over sub parts of $\mathcal{U}$. This is a slightly different form of sparseness as it is discussed in [5] and [6], where sparseness is referred to the situation with only a few discretization points per function. Nevertheless the smoothing approaches suggested by [5], to estimate the mean function and the covariance operator, as well as the PACE estimation procedure of [6], to estimate the loadings parameters, are directly applicable to our situation of sparse data. The empirical covariance function, $\tilde{\sigma}_T$, and the first four factors, $f_{1T}, \ldots, f_{4T}$, can be seen in figure 2. The estimated factor structure explains about 98.5% of the total variance of the price curves, such that we can reduce the high dimensional problem to a $K = 3$-dimensional problem without much loss of generality.

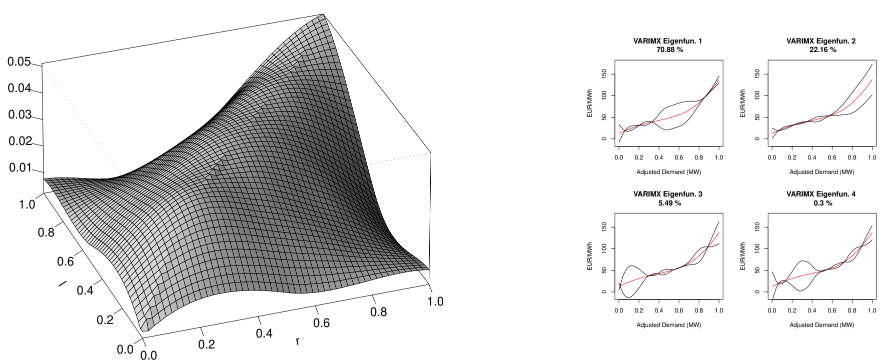

Figure 2: LEFT PANEL Empirical covariance function, $\tilde{\sigma}_T$, of the spherical component. RIGHT PANEL First four functionals of the estimated factor structure.

In the left panel of figure 3 we plot one estimated price function, $\hat{X}_t$, of an arbitrary day, t, with its corresponding raw data vector, $\mathbf{Y}_t$, as well as two outlier prices, that are excluded from the estimation procedure. In the right panel of figure 3 we show hourly electricity spot prices of one arbitrary week. The hourly fitted prices are determined by the evaluation of the estimated price functions, $\hat{X}_t$, at the corresponding hourly values of adjusted demand, u_{th}, for electricity. Note, that the proposed dynamic factor model may be easily combined with already developed approaches to model and forecast demand for electricity such as in [1].

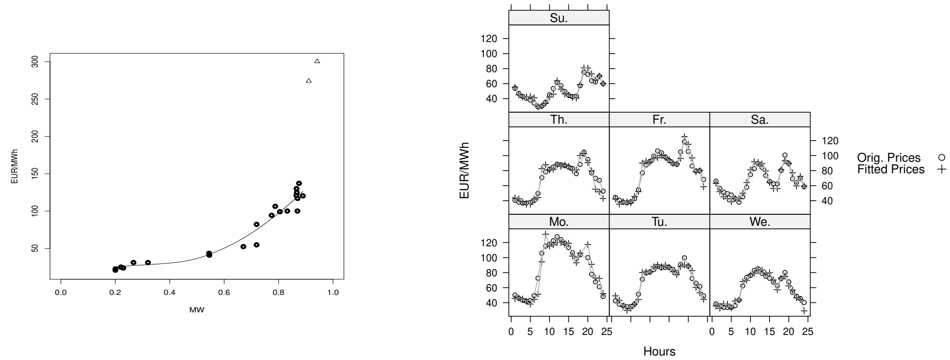

Figure 3: Left Panel Single fitted price curve with observed raw prices (circle points) and outlier prices (triangle points). Right Panel Hourly fitted prices and original prices.

References

[1] J Antoch, L Prchal, MR Rosa, and P Sarda. Functional linear regression with functional response: Application to prediction of electricity consumption. *Functional and Operatorial Statistics*, pages 5219–5219, 2008.

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

[3] W K Härdle and S Trück. The dynamics of hourly electricity prices. 2010.

[4] B U Park, E Mammen, W Härdle, and S Borak. Time Series Modelling With Semiparametric Factor Dynamics. *Journal of the American Statistical Association*, 104(485): 284–298, 2009.

[5] J G Staniswalis and J J Lee. Nonparametric Regression Analysis of Longitudinal Data. *Journal of the American Statistical Association*, 93(444): 1403–1418, 1998.

[6] F Yao, H G Müller, and J L Wang. Functional Data Analysis for Sparse Longitudinal Data. *Journal of the American Statistical Association*, 100(470): 577–590, 2005.