The Annals of Applied Statistics
2013, Vol. 7, No. 3, 1562–1592
DOI: 10.1214/13-AOAS652

MODELING AND FORECASTING ELECTRICITY SPOT PRICES: A FUNCTIONAL DATA PERSPECTIVE

BY DOMINIK LIEBL

Universities of Cologne and Bonn

Classical time series models have serious difficulties in modeling and forecasting the enormous fluctuations of electricity spot prices. Markov regime switch models belong to the most often used models in the electricity literature. These models try to capture the fluctuations of electricity spot prices by using different regimes, each with its own mean and covariance structure. Usually one regime is dedicated to moderate prices and another is dedicated to high prices. However, these models show poor performance and there is no theoretical justification for this kind of classification. The merit order model, the most important micro-economic pricing model for electricity spot prices, however, suggests a continuum of mean levels with a functional dependence on electricity demand.

We propose a new statistical perspective on modeling and forecasting electricity spot prices that accounts for the merit order model. In a first step, the functional relation between electricity spot prices and electricity demand is modeled by daily price-demand functions. In a second step, we parameterize the series of daily price-demand functions using a functional factor model. The power of this new perspective is demonstrated by a forecast study that compares our functional factor model with two established classical time series models as well as two alternative functional data models.

1. Introduction. Time series of hourly electricity spot prices have peculiar properties. They differ substantially from time series of equities and other commodities because electricity still cannot be stored efficiently and, therefore, electricity demand has an untempered effect on the electricity spot price [Knittel and Roberts (2005)].

The development of models for electricity spot prices was triggered by the liberalization of electricity markets in the early 1990s. Hourly electricity spot prices are usually considered to be multivariate (24-dimensional) time series since for each day t the 24 intra-day spot prices are settled simultaneously the day before [Huisman, Huurman and Mahieu (2007)].

However, classical time series models adopted for electricity spot prices such as autoregressive, jump diffusion or Markov regime switch models reduce the multivariate time series to univariate time series either by taking daily averages of the 24 hourly spot prices [Weron, Bierbrauer and Trück (2004), Kosater and Mosler

Received February 2012; revised March 2013.

Key words and phrases. Functional factor model, functional data analysis, time series analysis, fundamental market model, merit order curve, European Energy Exchange, EEX.

(2006) and Koopman, Ooms and Carnero (2007)] or by considering each hour h separately [Karakatsani and Bunn (2008)]. These unnatural aggregations and separations of the data necessarily come with great losses in information.

Our model, a functional factor model (FFM), is not a mere adaption of a classical time series model but is motivated by the data-generating process of electricity spot prices itself. Pricing in power markets is explained by the merit order model. This model assumes that the spot prices at electricity exchanges are based on the marginal generation costs of the last power plant that is required to cover the demand. The resulting so-called merit order curve reflects the increasing generation costs of the installed power plants. Often, nuclear and lignite plants cover the minimal demand for electricity. Higher demand is mostly served by hard coal and gas fired power plants.

Due to its importance, the merit order model is referred to as a fundamental market model [Burger, Graeber and Schindlmayr (2008), Chapter 4]. Essentially, the consideration of this fundamental model yields to the superior forecast performance of our FFM in comparison to state of the art time series models and alternative functional data models.

It is important to emphasize that the merit order model is not a static model. The merit order curve rather depends on the variations of the daily prices for raw materials, the prices of CO_2 certificates, the weather, plant outages and maintenance schedules of power plants.

The merit order curve is most important for the explanation of electricity spot prices in the literature on energy economics and justifies our view on the set of hourly electricity spot prices $\{y_{t1}, \dots, y_{t24}\}$ of day t. We do not interpret them as 24-dimensional vectors but rather as noisy discretization points of a smooth price-demand function X_t, which can be formalized as follows:

$$y_{th} = X_t(u_{th}) + \varepsilon_{th},$$

where u_{th} denotes electricity demand at hour h of day t and ε_{th} is assumed to be a white noise process.

The price-demand function $X_t(u)$ can be seen as the empirical counterpart of the merit order curve estimated nonparametrically from the $N = 24$ hourly price-demand data pairs $(y_{t1}, u_{t1}), \dots, (y_{tN}, u_{tN})$. Five exemplary estimated price-demand functions $\hat{X}_t(u)$ are shown in the lower panel of Figure 1. Figure 2 visualizes the temporal evolution of the time series of price-demand functions by showing the univariate time series $\hat{X}_1(u), \dots, \hat{X}_T(u)$ for a fixed value of electricity-demand $u = 58{,}000$ MW for the whole observed time span of $T = 717$ work days (Mo.–Fr.) from January 1, 2006 to September 30, 2008.

In order to capture the dynamic component of the price-demand functions, we assume them to be generated by a functional factor model defined as

$$X_t(u) = \sum_{k=1}^{K} \beta_{tk} f_k(u),$$

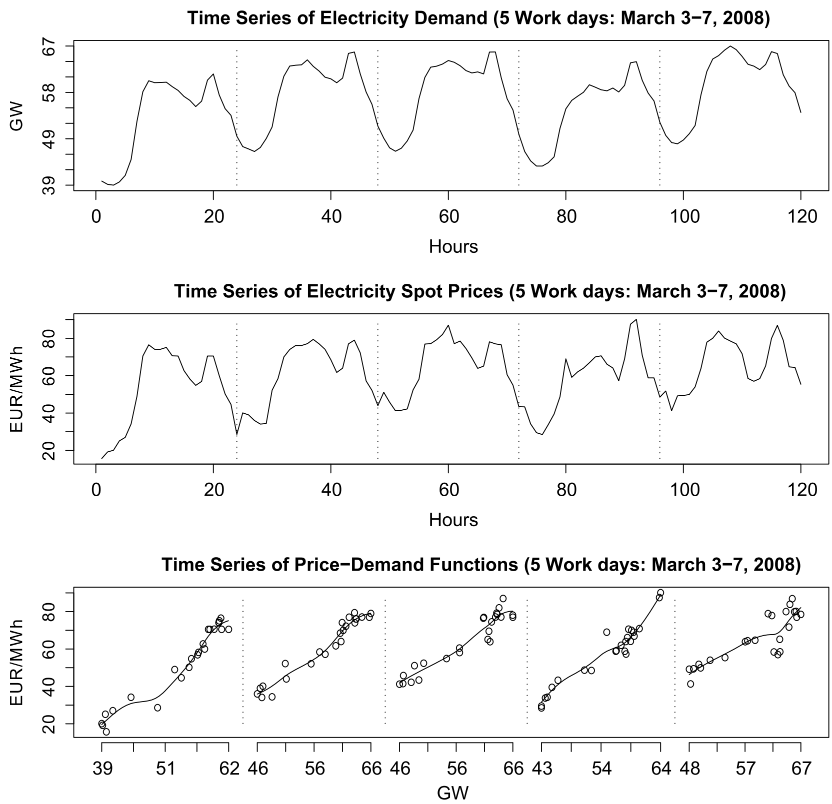

FIG. 1. UPPER PANEL: *time series of electricity demand* (u_{th}), *measured in GW* (1 *GW* $= 1000$ *MW*). MIDDLE PANEL: *electricity spot prices* (y_{th}). LOWER PANEL: *price-demand functions* ($\hat{X}_t$) *with noisy discretization points* $(y_{t1}, u_{t1}), \ldots, (y_{tN}, u_{tN})$.

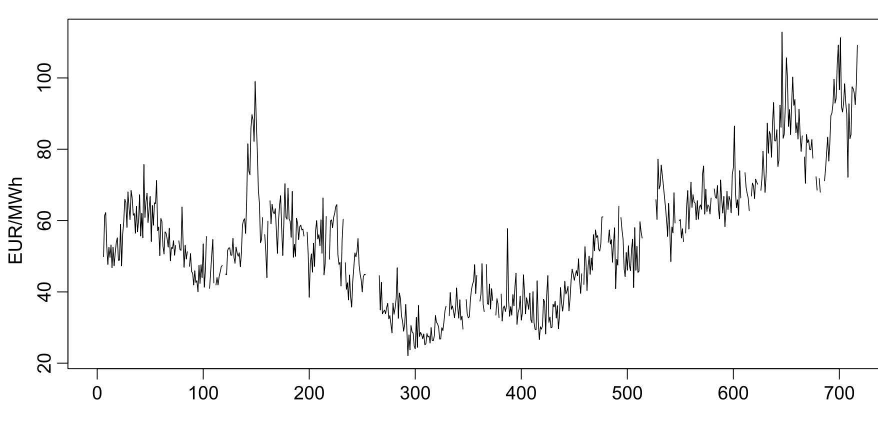

FIG. 2. *Univariate time series of fitted price-demand functions* $\hat{X}_1(u), \ldots, \hat{X}_T(u)$ *evaluated at* $u = 58{,}000$ *MW. Gaps correspond to holidays.*

where the factors or basis functions f_k are time constant and the corresponding scores β_{tk} are allowed to be nonstationary time series.

We do not specify a constant mean function in our FFM, since we allow the time series of price-demand functions $(X_t(u))$ to be nonstationary. Consequently, the classical interpretation of the factors f_k as perturbations of the mean does not apply—as common in the literature on dynamic (functional) factor models; see, for example, Hays, Shen and Huang (2012).

Note that the five price-demand functions in the lower panel of Figure 1 are observed on different domains. This distinguishes our functional data set from classical functional data sets, where all functions are observed on a common domain. We refer to this feature as *random domains* and its consideration in Sections 4.3 and 4.5 is a central part of our estimation procedure.

We use a two-step estimation procedure. The first step is to estimate the daily price-demand functions $\hat{X}_t$ by cubic spline smoothing for all days $t \in \{1, \dots, T\}$. The second step is to determine a $K < \infty$ dimensional common functional basis system $\{f_1, \dots, f_K\}$ for the estimated price-demand functions $\hat{X}_1, \dots, \hat{X}_T$. Given this system of basis functions, we model the estimated daily price-demand functions by a functional factor model—using the basis functions as common factors. The fitted discrete hourly electricity spot prices $\hat{y}_{th}$ are then obtained through the evaluation of the modeled price-demand functions at the corresponding hourly values of demand for electricity, formally written as $\hat{y}_{th} = \hat{X}_t(u_{th})$.

Functional data analysis (FDA) can share our perspective on electricity spot prices. A broad overview of many different FDA methods can be found in the monographs of Ramsay and Silverman (2005) and Ferraty and Vieu (2006). Particularly, Chapter 8 in Ramsay and Silverman (2005) and the nonparametric methods for computing the empirical covariance function as proposed in Staniswalis and Lee (1998), Yao, Müller and Wang (2005), Hall, Müller and Wang (2006) and Li and Hsing (2010) are important methodological references for this paper.

The application of models from the functional data literature to the electricity market data is not new. For example, there is a vast literature on modeling and forecasting electricity *demand*; see, for example, Ferraty and Vieu (2006) and Antoch et al. (2010). However, modeling and forecasting electricity spot prices is much more difficult than modeling and forecasting electricity demand. The semi-functional partial linear model (SFPL) of Vilar, Cao and Aneiros (2012) is one of the very rare cases in which FDA methods are used to forecast electricity spot prices.

Two very recent examples of other functional factor models are given by the functional factor analysis in Liu et al. (2012) and the functional dynamic factor model (FDFM) in Hays, Shen and Huang (2012). Liu et al. (2012) propose a new rotation scheme for the functional basis components. Hays, Shen and Huang (2012) model a time series of yield curves and estimate their model by the EM algorithm. In contrast to the FDFM of Hays, Shen and Huang (2012), we do not have to make a priori assumptions on the stochastic properties of the time series of scores in order to estimate our model components. Furthermore, we are able to model and forecast functional time series observed on random domains.

Very close to the FDFM of Hays, Shen and Huang (2012) is the Dynamic Semiparametric Factor Model (DSFM) of Park et al. (2009). As our functional factor model the DSFM does not need a priori assumptions on the time series of scores. This and the fact that the DSFM was already successfully applied to electricity prices [Borak and Weron (2008) and Härdle and Trück (2010)] makes the DSFM a perfect competitor for our FFM.

The main difference between the FDFM of Hays, Shen and Huang (2012) and the DSFM of Park et al. (2009) in comparison to our FFM is that the FFM can deal with functional times series observed on random domains. Furthermore, Park et al. (2009) use an iterating optimization algorithm to estimate the basis functions of the DSFM, whereas we standardize the elements of the time series (X_t) so that we can robustly estimate the basis functions by functional principal component analysis. Our estimation procedure is much simpler to implement and faster with respect to computational time than the Newton–Raphson algorithm suggested in Park et al. (2009).

The next section is devoted to the introduction of our data set and to a critical consideration of the stylized facts of electricity spot prices usually claimed in the electricity literature. In Section 3 we present our functional factor model and in Section 4 its estimation. An application of the model to real data is presented in Section 5. Finally, the performance of the functional factor model is demonstrated by an extensive forecast study in Section 6.

2. Electricity data. We demonstrate our functional factor model by modeling and forecasting electricity spot prices of the German power market traded at the European Energy Exchange (EEX) in Leipzig. The German power market is the biggest power market in Europe in terms of consumption. The wholesale market is fragmented into an Over The Counter (OTC) market and the EEX. While the OTC market has a continuous trade, the EEX has a single uniform price auction with a gate closure for the day ahead market at 12 p.m. the day before physical delivery. Although three-fourths of the trading volume is settled via bilateral OTC contracts, the EEX spot price is of fundamental importance as benchmark and reference point for other markets, such as OTC or forward markets [Grimm, Ockenfels and Zoettl (2008), Chapter 1].

The data for this analysis stem from three different publicly available sources. The hourly spot prices of the German electricity market are provided by the European Energy Exchange (www.eex.com), hourly values of Germany's gross electricity demand are provided by the European Network of Transmission System Operators for Electricity (www.entsoe.eu), and German wind power infeed data are provided by the EEX Transparency Platform (www.transparency.eex.com). The data set used in our application is provided as part of the supplementary material; see Liebl (2013).

In the German electricity market, as in most of the electricity markets in the world, renewable energy sources are usually provided with purchase guarantees. Therefore, the hourly values of gross electricity demand are not relevant for the pricing at the EEX but rather the hourly values of gross demand minus the hourly electricity infeeds from renewable energy sources. We consider only wind power infeed data since the influences of other renewable energy sources such as photovoltaic and biomass on electricity spot prices are still negligible for the German electricity market (and their explicit consideration essentially would lead to the same results).

The data consists of pairs (y_{th}, u_{th}) with y_{th} denoting the electricity spot price and u_{th} the electricity demand of hour $h \in \{1, \dots, 24\}$ at day t. We define electricity demand u_{th} as the gross electricity demand of hour h and day t minus the wind power infeed of electricity at the corresponding hour h and day t.

The data set analyzed in this article covers $T = 717$ work days (Mo.–Fr.) within the time horizon from January 1, 2006 to September 30, 2008. For the sake of clarity, only working days are considered in our analysis since for weekends there are different compositions of the power plant portfolio. The same reasoning applies to holidays and so-called Brückentage, which are extra days off that bridge single working days between a bank holiday and the weekend. Therefore, we set all holidays and Brückentage to NA-values.

As a referee noted, the time span of our data set is peculiar. Starting around January 2007, a price bubble for raw commodities such as coal and gas was formed, which induced a strong increase in the electricity spot prices. Interestingly, the increase in the electricity spot prices is hardly visible in the original time series as shown in Figure 6. But it catches the eye in the plot of Figure 2, which shows the time series of price-demand functions $(\hat{X}_t(u)|u)$ evaluated for a certain value of electricity demand $u = 58{,}000$ MW. The reason is that at this relatively high value of electricity demand usually coal and gas fired power plants cover the demand.

Very few (only 0.5%) of the data pairs (y_{th}, u_{th}) with prices $y_{th} > 200$ EUR/MWh have to be treated as outliers since they cannot be explained by the merit order model. Even in exceptional situations the marginal costs of electricity production do not exceed the value of 200 EUR/MWh. Prices above this threshold are referred to as price spikes and have to be explained using an additional scarcity premium [Burger, Graeber and Schindlmayr (2008), Chapter 4]. The analysis of price spikes is a research topic on its own [Christensen, Hurn and Lindsay (2009)] and is not within the scope of this paper.

We exclude the outliers for the estimation of our model and denote the amount of data pairs of day t used for estimation by $N_t \leq N = 24$. Nevertheless, we use the whole data set, including the outliers, in order to assess the forecast performance of our model in Section 6.

Review: *Stylized facts of electricity data.* Our functional perspective on electricity spot prices allows us to review critically the so-called "stylized facts" of hourly electricity spot prices (y_{th}). Usually, time series of electricity spot prices are assumed (i) to have deterministic daily, weekly and yearly seasonal patterns, (ii) to show price dependent volatilities, and (iii) to be stationary (after controlling for the seasonal patterns); see Huisman and De Jong (2003), Knittel and Roberts (2005), Kosater and Mosler (2006), Huisman, Huurman and Mahieu (2007) and many others.

At first glance these stylized facts seem to be reasonable; see the middle panel in Figure 1. However, the first two stylized facts, (i) and (ii), are misleading since both have their origins in the time series of electricity demand: the characteristics of electricity demand are rather carried over to the time series of electricity spot prices.

This can be explained by a micro-economic point of view, again using the merit order model. The merit order curve induces a monotone increasing supply function for electricity, which implies higher electricity spot prices for higher values of electricity demand, where electricity demand can be considered as inelastic. Given this micro-economic point of view, we can regard the daily supply functions for electricity as diffusers in the transmission from electricity demand u_{th} to the electricity spot price y_{th}.

Additional diffusion comes from the variations of the daily supply functions caused by the varying input-costs of, for example, coal and gas. Compare to this the time series of electricity demand with the time series of electricity spot prices shown in the upper and middle panels of Figure 1, respectively. The seasonal patterns of electricity spot prices are just a diffused version of the smoother seasonal patterns of electricity demand.

Price dependent volatility (ii) can be explained by the slope of the merit order curve, which is increasing with electricity demand. Changes in electricity demand have greater price effects for greater values of electricity demand and therefore cause greater volatilities than is the case for lower values of electricity demand.

Stationarity (iii) has to be considered critically, too. Recently, Bosco et al. (2010) were able to show empirically that electricity spot prices at the EEX have a unit root. The authors point out that the stationarity assumption might be wrong in markets that are influenced by price-enhancing sources such as prices for coal and gas since time series of coal and gas prices are commonly found to be nonstationary. Our functional factor model allows for nonstationarity in the time series of price-demand functions (X_t) and, in fact, tests indicate that the estimated series of price-demand functions is nonstationary; see Section 5.2.

This short review of electricity spot prices demonstrates that electricity data are complex with dynamics induced by the variations of the merit order curve (mainly caused by varying input-costs) and separate additional dynamics induced by electricity demand. To the best of our knowledge, our functional factor model is the first model that allows for a separate consideration of these two stochastic sources. The variations dedicated to the dynamics of the merit order curve are captured by the price-demand functions and modeled by our functional factor model.

The problem of modeling and forecasting electricity demand is "out-sourced" and the statistician can choose powerful specialized models for time series of gross electricity demand [Antoch et al. (2010)] and time series of wind power [Lau and McSharry (2010)]. This separation corresponds to the real data generating process.

3. Functional factor model. As mentioned above, electricity spot prices $y_{t1}, \dots, y_{t24}$ are actually one-day-ahead future prices since they are settled simultaneously at day $t-1$. This implies that there is some degree of uncertainty about the next day world in the electricity spot price y_{th}, which we model nonparametrically as

$$y_{th} = X_t(u_{th}) + \varepsilon_{th}. \tag{1}$$

The error terms ε_{th} are assumed to be i.i.d. white noise errors with finite variance $\mathrm{V}(\varepsilon_{th}) = \sigma_\varepsilon^2$ and each function X_t is assumed to be continuous and square integrable.

For each function X_t the values of electricity demand u_{th} are only observed within random sub-domains $\mathcal{D}(X_t) = [a_t, b_t]$, where $[a_t, b_t] \subseteq [A, B] \subset \mathbb{R}$. The unobserved univariate time series (a_t) and (b_t) are assumed to be time series processes with $A \le a_t < b_t \le B$ and marginal p.d.f.s of a_t and b_t given by $f_a(z_a) > 0$ and $f_b(z_b) > 0$ for all $z_a, z_b \in [A, B]$ and $t \in \{1, \dots, T\}$.

The price-demand functions are relatively homogeneous. All of them look very similar to the five randomly chosen price-demand functions shown in the lower panel of Figure 1. The underlying reason for this homogeneity is that, on the one hand, the merit order curve induces rather simple monotone increasing price-demand functions. On the other hand, the general portfolio of power plants, which is reflected by the merit order curve, is changing very slowly and can be considered as constant over the period of our analysis. We formalize this homogeneity of the price-demand functions by the assumption that the time series of price-demand functions (X_t) is generated by a functional factor model with time constant basis functions.

Given this assumption, every price-demand function X_t can be modeled by the same set of $K < \infty$ (unobserved) basis functions $f_1, \dots, f_k, \dots, f_K$ with $f_k \in L^2[A, B]$, which span the K-dimensional functional space $\mathcal{H}_K \subset L^2[A, B]$ such that we can write

$$X_t(u) = \sum_{k=1}^{K} \beta_{tk} f_k(u) \qquad \text{for all } u \in [a_t, b_t], \tag{2}$$

where the common basis functions f_k as well as the scores β_{tk} are unobserved and have to be determined from the data. We use the usual orthonormal identification restrictions for the basis functions, which require that $\int_A^B f_k^2(u)\,du = 1$ and $\int_A^B f_k(u) f_l(u)\,du = 0$ for all $k < l \in \{1, \dots, K\}$.

The K real time series $(\beta_{t1}), \ldots, (\beta_{tK})$ are defined as

$$\begin{pmatrix} \beta_{t1} \\ \vdots \\ \beta_{tK} \end{pmatrix} = \begin{pmatrix} \int_{a_t}^{b_t} f_1^2 & \cdots & \int_{a_t}^{b_t} f_1 f_K \\ \vdots & \ddots & \vdots \\ \int_{a_t}^{b_t} f_1 f_2 & \cdots & \int_{a_t}^{b_t} f_K^2 \end{pmatrix}^{-1} \begin{pmatrix} \int_{a_t}^{b_t} f_1 X_t \\ \vdots \\ \int_{a_t}^{b_t} f_K X_t \end{pmatrix} \tag{3}$$

and are allowed to be arbitrary nonstationary processes. Note that for $a_t = A$ and $b_t = B$ the definition of the scores β_{tk} corresponds to the classical definition, given by $\beta_{tk} = \int_A^B X_t(u) f_k(u)\,du$.

In the following section we propose an estimation algorithm for the functional factor model.

4. Estimation procedure. As outlined in Sections 1 and 2, we do not observe the series (X_t) directly but have to estimate each price-demand function X_t from the corresponding data pairs $(y_{t1}, u_{t1}), \ldots, (y_{tN_t}, u_{tN_t})$. After this initial estimation step, which is discussed in Section 4.1, we show in Section 4.2 how to determine an orthonormal K-dimensional basis system $\{f_1, \ldots, f_K\}$ for the classical functional data case when all price-demand functions $X_1, \ldots, X_T$ are observed on the deterministic domain $\mathcal{D}(X_t) = [A, B]$. In Section 4.3 we generalize the determination of the orthonormal K-dimensional basis system $\{f_1, \ldots, f_K\}$ to our case, where the price-demand functions X_t are observed only on random domains $\mathcal{D}(X_t) = [a_t, b_t]$. Finally, we define our estimator $\{\hat{f}_1, \ldots, \hat{f}_K\}$ in Section 4.4.

As usual for (functional) factor models, the set of factors $\{f_1, \ldots, f_K\}$ in (2) is only determined up to orthonormal rotations. Furthermore, the determination of an orthonormal K-dimensional basis system $\{\hat{f}_1, \ldots, \hat{f}_K\}$ for a given series $(\hat{X}_t)$ is, in the first instance, a mere algebraic problem. But it is also a statistical estimation problem in the sense that $\hat{\mathcal{H}}_K$, with $\hat{\mathcal{H}}_K = \operatorname{span}(\hat{f}_1, \ldots, \hat{f}_K)$, is a consistent estimator of the theoretical counterpart $\mathcal{H}_K$. The crucial assumption is that X_t comes from the FFM (2). Consistency of the estimation follows from the consistency of the single nonparametric estimators $\hat{X}_t(u)$, which converge in probability against $X_t(u)$ as $N_t \to \infty$ for all $u \in [a_t, b_t]$ and all $t \in \{1, \ldots, T\}$ [Benedetti (1977)]. Below in Section 4.5 we consider this issue in more detail.

4.1. *Estimation of the price-demand functions X_t.* The estimation of the functions X_t from the data pairs $(y_{t1}, u_{t1}), \ldots, (y_{tN_t}, u_{tN_t})$ is done by minimizing

$$d(\mathcal{X}|t) = \sum_{h=1}^{N_t} \big(y_{th} - \mathcal{X}(u_{th})\big)^2 + b \int_{a_t}^{b_t} \big(D^2 \mathcal{X}(u)\big)^2\,du \tag{4}$$

over all twice continuously differentiable functions $\mathcal{X}$, where $D^2\mathcal{X}$ denotes the 2nd derivative of $\mathcal{X}$ and $b > 0$ is a preselected smoothing parameter. Spline theory assures that any solution $\hat{X}_t$ of the minimization problem (4) can be expanded by a natural spline basis [de Boor (2001)]. Therefore, we can use the expansion $\mathcal{X}(u) = \mathbf{c}'\boldsymbol{\phi}(t)$, where $\boldsymbol{\phi}$ is the $(N_t + 2)$-vector of natural spline basis functions of degree 3 and $\mathbf{c}$ is the $(N_t + 2)$-vector of coefficients over which equation (4) is minimized. This procedure is usually denoted as cubic spline smoothing and the interested reader is referred to the monographs of de Boor (2001) and Ramsay and Silverman (2005).

An important issue that remains to be discussed is the selection of the smoothing parameter b. Usually, the optimal smoothing parameter b^{opt} is chosen by (generalized) cross-validation such that the trade-off between bias and variance of the estimate $\hat{X}_t$ is optimized asymptotically with respect to the mean integrated squared error (MISE) criterion. However, our aim is not an optimal single estimate $\hat{X}_t$ but rather an optimal estimation of the basis system $\{f_1, \dots, f_K\}$ for which we can use the information of all price-demand functions $X_1, \dots, X_T$.

Consequently, we do not have to optimize the MISEs of the single estimators $\hat{X}_t$ but those of their weighted averages $\hat{f}_1, \dots, \hat{f}_K$. In this case an undersmoothing parameter $\underline{b}_K < b^{\text{opt}}$ has to be chosen. This was discussed for the first time in Benko, Härdle and Kneip (2009). The underlying reason is that the estimators $\hat{f}_1, \dots, \hat{f}_K$ essentially are weighted averages over all $\hat{X}_1, \dots, \hat{X}_T$. Averaging reduces the overall variance and therefore opens the possibility for a further reduction in the MISEs of the estimators $\hat{f}_1, \dots, \hat{f}_K$ by a further reduction of the bias-component through choosing $\underline{b}_K < b^{\text{opt}}$ in the minimization of (4). Benko, Härdle and Kneip (2009) propose to approximate an optimal undersmoothing parameter $\underline{b}_K$ by minimizing the following cross-validation criterion:

$$\text{CV}(b_K) = \sum_{t=1}^{T}\sum_{i=1}^{N_t}\left\{y_{th} - \sum_{k=1}^{K}\hat{\gamma}_{tk}\hat{f}_{k,-t}(u_{th})\right\}, \tag{5}$$

over $0 \le b_K \le \infty$, where $\hat{\gamma}_{tk}$ are the OLS estimators of $\hat{\beta}_{tk}$ and $\hat{f}_{k,-t}$ denote the estimators of f_{tk} based on the data pairs (y_{sh}, u_{sh}) with $s \in \{1, \dots, t-1, t+1, \dots, T\}$. We denote the estimators of X_t based on an undersmoothing parameter $\underline{b}_K$ by $\tilde{X}_1, \dots, \tilde{X}_T$ and those based on b^{opt} by $\hat{X}_1, \dots, \hat{X}_T$.

As can be seen in (5), an optimal undersmoothing parameter $\underline{b}_K$ depends on the dimension K. The problem of choosing K can be seen as a model selection problem, which generally can be solved using information criteria. For our application in Section 5 we use the simple cumulative variance criterion as well as the AIC type criterion proposed in Yao, Müller and Wang (2005).

4.2. *Estimation of the basis system* $\{f_1, \dots, f_T\}$. Our estimation procedure uses the property that any orthonormal basis system $\{f_1, \dots, f_K\}$ of the series (X_t) has to fulfill the minimization problem

$$\sum_{t=1}^{T}\left\| X_t - \sum_{k=1}^{K}\beta_{tk}f_k\right\|_2^2 = \min_{B_K}\sum_{t=1}^{T}\min_{\gamma_{t1},\dots,\gamma_{tK}\in\mathbb{R}}\left\| X_t - \sum_{k=1}^{K}\gamma_{tk}g_k\right\|_2^2 \tag{6}$$

over all possible K-dimensional orthonormal basis systems $B_K = \{g_1, \dots, g_K\}$, where $g_1, \dots, g_K \in L^2[A, B]$ and $\|\cdot\|_2$ denotes the functional L2 norm $\|x\|_2 = \sqrt{\int_A^B x^2(u)\,du}$ for any $x \in L^2[A, B]$. This property is a direct consequence of the FFM (2).

The minimization problem (6) can be used to define an estimator for a basis system $\{f_1, \dots, f_K\}$. We would only have to replace the unobserved functions X_t with their undersmoothed estimators $\tilde{X}_t$ and try to find a basis system that minimizes the right-hand side (rhs) of (6) using functional principal component analysis (FPCA).

Before we present the analytic solution we adjust the minimization problem (6). This adjustment yields a robustification, which is needed since we allow the K time series $(\beta_{t1}), \dots, (\beta_{tK})$ to be nonstationary processes. The nonstationarity of the scores $(\beta_{t1}), \dots, (\beta_{tK})$ implies that different functions X_t and X_s can be of very different orders of magnitude, that is, $\|X_t\|_2 \ll \|X_s\|_2$. In such cases, the squared L2-norm on the rhs of (6) sets an overproportional weight on functions with great orders of magnitude, and a functional principal component estimator based on (6) would be distorted toward those functions X_s that have great orders of magnitude $\|X_s\|_2$.

If we were only interested in the determination of some set of basis functions $\{f_1, \dots, f_K\}$ that spans the same space $\mathcal{H}_K$ as the set of functions $\{X_1, \dots, X_T\}$, we would not have to care about functions X_s with great orders of magnitude $\|X_s\|$. However, if we are interested in the interpretation of the basis functions f_k, we want them to be representative for all functions $X_1, \dots, X_T$.

A general solution to this problem is to replace the price-demand functions X_t in (6) with their standardized counterparts $X_t^* = X_t/\|X_t\|$, which have equal orders of magnitude $\|X_t^*\| = 1$ for all $t \in \{1, \dots, T\}$. Using this replacement yields the following new minimization problem:

$$(7) \qquad \sum_{t=1}^{T} \left\| X_t^* - \sum_{k=1}^{K} \beta_{tk}^* f_k^* \right\|_2^2 = \min_{B_K} \sum_{t=1}^{T} \min_{\gamma_{t1}, \dots, \gamma_{tK} \in \mathbb{R}} \left\| X_t^* - \sum_{k=1}^{K} \gamma_{tk} g_k \right\|_2^2 .$$

Solving equation (7) by FPCA generally will yield different basis functions f_k^* than solving (6). However, both minimization problems (6) and (7) are equivalent in the sense that both sets of basis functions $\{f_1^*, \dots, f_K^*\}$ and $\{f_1, \dots, f_K\}$ are equivalent up to orthonormal rotations and therefore span the same space $\mathcal{H}_K$. The standardization yields to a simple base change, which can be seen by the fact that the original price-demand functions X_t can be written in terms of the basis functions f_k^* as $X_t = \sum_{k=1}^{K} (\|X_t\| \cdot \beta_{tk}^*) f_k^*$.

The standardization of all price-demand functions X_t in the minimization problem (7) allows us to establish a nondistorted estimator $\{\hat{f}_1, \dots, \hat{f}_K\}$ that represents all price-demand functions equally well. This approach is similar to robust estimation procedures proposed by Locantore et al. (1999) and Gervini (2008) but differs conceptually insofar as we do not consider any functional observation X_t as an outlier.

We construct our estimator $\{\hat{f}_1, \dots, \hat{f}_K\}$ from the analytic solution of the minimization problem (7). The solutions of the inner minimization problem with respect to the scores γ_{tk} are given by least squares theory, and we can write

$$\sum_{t=1}^{T} \left\| X_t^* - \sum_{k=1}^{K} \beta_{tk}^* f_k^* \right\|_2^2 = \min_{B_K} \sum_{t=1}^{T} \| X_t^* - P_K^g X_t^* \|_2^2, \tag{8}$$

where P_K^g, defined as $P_K^g X_t^* = \sum_{k=1}^{K} (\int_A^B X_t^*(v) g_k(v)\, dv) g_k$, is a linear projection operator that projects the standardized price-demand functions X_t^* into the subspace of $L^2[A, B]$ spanned by the orthonormal basis system $B_K = \{g_1, \dots, g_K\}$.

It is well known that a solution of the minimization problem (8) with respect to all K-dimensional orthonormal basis systems B_K can be determined by FPCA. This so-called "best basis property" of the empirical eigenfunctions $e_{T,1}, \dots, e_{T,K}$ is of central importance for this paper; see Section 8.2.3 in Ramsay and Silverman (2005), among others. Note that the eigenvalues $\lambda_{T,k}$ with $k \in \{1, \dots, K\}$ may be of multiplicity $L > 1$; in this case $e_{T,k} \in E_k$ with $E_k = \operatorname{span}(e_{Tk,1}, \dots, e_{Tk,L})$.

A solution of (8) is given by the set of eigenfunctions $\{e_{T,1}, \dots, e_{T,K}\}$ that belong to the first K greatest eigenvalues $\lambda_{T,1} > \lambda_{T,2} > \cdots > \lambda_{T,K} > 0$ of the empirical covariance operator Γ_T defined as

$$(\Gamma_T x)(u) = \int_A^B \gamma_T(u, v) x(v)\, dv \qquad \text{for all } x \in L^2[A, B], \tag{9}$$

where the empirical covariance function $\gamma_T(u, v)$ is defined as a local linear surface smoother in (10). We use this nonparametric version of $\gamma_T(u, v)$, since it can be applied to the classical case of deterministic domains $\mathcal{D}_t(X_t) = [A, B]$ as well as to the case of random domains $\mathcal{D}_t(X_t) = [a_t, b_t]$ discussed in the following Section 4.3. Contrary to this, the classical textbook definition of $\gamma_T(u, v)$ cannot be applied to the case of random domains.[1]

4.3. *Random domains $\mathcal{D}(X_t) = [a_t, b_t]$.* From a computational perspective, functional data observed on random domains cause problems similar to sparsely observed functional data. For the latter case there is already a broad stream of literature based on the papers of Staniswalis and Lee (1998), Yao, Müller and Wang (2005), Hall, Müller and Wang (2006) and Li and Hsing (2010).

We follow Yao, Müller and Wang (2005) and compute the covariance function γ_T by local linear surface smoothing. Here, $\gamma_T(u, v) = \beta_{T,0}$ and $\beta_{T,0}$ is determined

[1]The classical definition is given by $\gamma_T(u, v) = T^{-1} \sum_{t=1}^{T} X_t(u) X_t(v)$.

by minimizing

$$\sum_{t=1}^{T}\sum_{i,j=1}^{N_t}\kappa_2\left(\frac{(u_{ti}-u)}{b_\gamma},\frac{(u_{tj}-v)}{b_\gamma}\right) \times \bigl\{X_t^*(u_{ti})X_t^*(u_{tj})-f\bigl(\beta_T,(u,v),(u_{ti},u_{tj})\bigr)\bigr\}^2 \tag{10}$$

over $\beta_T=(\beta_{T,0},\beta_{T,11},\beta_{T,12})'\in\mathbb{R}^3$, where $f(\beta_T,(u,v),(u_{ti},u_{tj}))=\beta_{T,0}+\beta_{T,11}(u-u_{ti})+\beta_{T,12}(v-u_{tj})$, u_{ti} are the observed values of electricity demand, b_γ is the smoothing parameter that can be determined, for instance, by (generalized) cross-validation, and $\kappa_2:\mathbb{R}^2\to\mathbb{R}$ is a two-dimensional kernel function such as the multiplicative kernel $\kappa_2(x_1,x_2)=\kappa(x_1)\kappa(x_2)$ with κ being a standard univariate kernel such as the Epanechnikov kernel. See Yao, Müller and Wang (2005) and Fan and Gijbels (1996) for further details.

4.4. *The estimator* $\{\hat{f}_1,\dots,\hat{f}_K\}$. Given the analytic solution of the minimization problem (8), we can now define the estimator $\{\hat{f}_1,\dots,\hat{f}_K\}$. The only thing that we have to do is to replace the standardized price-demand functions X_t^* in (10) by their undersmoothed and standardized estimators $\tilde{X}_t^*$, where $\tilde{X}_t^*=\tilde{X}_t/\|\hat{X}_t\|_2$.

Note that we scale the undersmoothed estimator $\tilde{X}_t$ with the L2 norm of the estimator $\hat{X}_t$, which is optimally smoothed with respect to the *single* observation X_t. Undersmoothing of the price-demand functions is always important if the target quantity, such as the covariance function $\gamma_T(u,v)$, consists of an average over all functions. The approximation of the norm $\|X_t\|$ does not involve averages over all functions, such that we are better off to use the norm of the classically smoothed curves $\|\hat{X}_t\|$.

Let us denote the estimator of the empirical covariance operator Γ_T by $\hat{\Gamma}_T$, defined as

$$(\hat{\Gamma}_T x)(u)=\int_A^B \hat{\gamma}_T(u,v)x(v)\,dv \qquad \text{for all } x\in L^2[A,B], \tag{11}$$

where $\hat{\gamma}_T(u,v)$ is determined by minimizing equation (10) after replacing $X_t^*=X_t/\|X_t\|$ by $\tilde{X}_t^*=\tilde{X}_t/\|\hat{X}_t\|_2$. Accordingly, we denote the first K ordered eigenvalues and the corresponding eigenfunctions of $\hat{\Gamma}_T$ by $\hat{\lambda}_{T,1}>\cdots>\hat{\lambda}_{T,K}$ and $\hat{e}_{T,1},\dots,\hat{e}_{T,K}$.

The estimator $\{\hat{f}_1,\dots,\hat{f}_K\}$ is then defined as any orthonormal rotation of the orthonormal basis system $\{\hat{e}_{T,1},\dots,\hat{e}_{T,K}\}$ determined by (8). The trivial case would be to use the empirical eigenfunctions $\hat{e}_{T,1},\dots,\hat{e}_{T,K}$ directly as basis functions such that $\hat{f}_k=\hat{e}_{T,k}$ for all $k\in\{1,\dots,K\}$. It is generally left to the statistician to choose an appropriate orthonormal rotation scheme, which facilitates the interpretation. In our application we use the well-known VARIMAX-rotation.

Following our assumptions on the data generating process in (2), we use the basis system $\{\hat{f}_1, \dots, \hat{f}_K\}$ in order to re-estimate the functions $X_1, \dots, X_T$ by

$$\hat{X}_t^f = \sum_{k=1}^{K} \hat{\beta}_{tk} \hat{f}_k, \tag{12}$$

where the parameters $\hat{\beta}_{tk}$ are defined according to (3). This is a crucial step of the estimation procedure. Given that our model assumption in (2) is true, the original single cubic smoothing splines estimates $\hat{X}_t$ will be much less efficient estimators of the price-demand functions X_t than the estimators $\hat{X}_t^f$ since the latter use the information of the whole data set.

4.5. *A note on convergence.* Assume that we are able to observe the (unobservable) set of functions $\{X_1, \dots, X_T\}$ as defined in (2) but with deterministic domains $\mathcal{D}(X_t) = [A, B]$. In this case the K empirical eigenfunctions $e_{T,1}, \dots, e_{T,K}$ can be determined from the empirical covariance operator Γ_T as defined in (9) based on the classical definition of the empirical covariance function $\gamma_T(u, v) = K^{-1} \sum_{k=1}^{K} X_{t_k}(u) X_{t_k}(v)$. Actually, only a subset of at least K linear independent functions, say, $X_{t_1}, \dots, X_{t_K}$, would suffice to determine the K empirical eigenfunctions $e_{T,1}, \dots, e_{T,K}$.

In this case, the determination of the basis system $\{e_{T,1}, \dots, e_{T,K}\}$ is a mere algebraic problem. Furthermore, the space $\mathcal{H}_K$ spanned by the basis system $\{e_{T,1}, \dots, e_{T,K}\}$ does not depend on the data. By the definition in (2), two sets of functions $\{X_1, \dots, X_T\}$ and $\{X_1, \dots, X_{T'}\}$ span the same space $\mathcal{H}_K$ for all $T' \geq T \geq K$, such that also the corresponding basis systems $\{e_{T,1}, \dots, e_{T,K}\}$ and $\{e_{T',1}, \dots, e_{T',K}\}$ span the same space $\mathcal{H}_K$.

Note that we do not observe the functions X_t but the noisy discretization points $y_{th} = X_t(u_{th}) + \varepsilon_{th}$. Starting with the first scenario of deterministic domains $\mathcal{D}(X_t) = [A, B]$, the determination of the estimated basis system $\{\hat{e}_{T,1}, \dots, \hat{e}_{T,K}\}$ can be done from the estimated empirical covariance operator $\hat{\Gamma}_T$ defined in (11). Again, this on its own is a mere algebraic problem but it yields to our consistency argument.

The estimated eigenfunction $\hat{e}_{Tk}$ can be written as a continuous function of $\hat{X}_1, \dots, \hat{X}_T$, say, $\hat{e}_{Tk} = g_k(\hat{X}_1, \dots, \hat{X}_T)$. By the continuous mapping theorem, $\hat{e}_{Tk} = g_k(\hat{X}_1, \dots, \hat{X}_T)$ converges to $e_{Tk} = g_k(X_1, \dots, X_T)$ as $\hat{X}_t(u) \to X_t(u)$, for example, in probability as $N_t \to \infty$ for all $u \in [A, B]$ and $t \in \{1, \dots, T\}$ with $T \geq K$. We additionally have to assume that all involved smoothing parameters go against zero appropriately fast such that $N\underline{b}_K \to \infty$, $Nb^{\text{opt}} \to \infty$, and $NTb_\gamma \to \infty$ [Benedetti (1977)]. Eigenfunctions e_k are determined up to sign changes and it is assumed that the correct signs are chosen.

In this sense we can state that $\hat{\mathcal{H}}_K = \text{span}(\hat{e}_{T,1}, \dots, \hat{e}_{T,K})$ converges to $\mathcal{H}_K = \text{span}(e_{T,1}, \dots, e_{T,K})$, which is all we can achieve for (functional) factor models, since the single factors f_k remain unidentifiable.

Finally, it only remains to consider the scenario of random domains $\mathcal{D}(X_t) = [a_t, b_t]$. Also, in this case any two points $(u, v) \in [A, B]$ have to be covered by at least K price-demand functions, which are fulfilled asymptotically. By our assumptions the time series (a_t) and (b_t) are processes with $A \le a_t < b_t \le B$ and the marginal p.d.f.s of a_t and b_t are given by $f_a(z_a) > 0$ and $f_b(z_b) > 0$ for all $z_a, z_b \in [A, B]$ and $t \in \{1, \dots, T\}$. This yields that

$$\Pr\bigl(a_t \in [A, A+\varepsilon]\bigr) > 0 \quad \text{and} \quad \Pr\bigl(b_t \in [B-\varepsilon, B]\bigr) > 0 \qquad \text{for any } \varepsilon > 0,$$

such that for $T \to \infty$ with probability one, there are sub-series (a_s) and (b_s) of (a_t) and (b_t) for which the boundary points A and B are accumulation points. From this it follows that as $T \to \infty$ we can find always more than K functions X_t that cover the points $u, v \in [A, B]$.

To conclude, consistency of our estimation procedure relies on our model assumptions in (1) and (2) and is driven by the consistency of the first step estimators of the price-demand functions $X_t(u)$.

5. Application. In this section we apply our estimation procedure of the FFM described in Section 4 to the data set described in Section 2. In Section 5.1 we show how to interpret the factors and demonstrate an exemplary analysis of the scores and in Section 5.2 we validate the crucial model assumptions.

A drawback of the cross-validation criterion in (5) is that it depends on the unknown dimension K. Therefore, first, we determine optimal undersmoothing parameters $\underline{b}_K$ for several values of K and, second, choose the dimension K, which minimizes the AIC of Yao, Müller and Wang (2005).

The AIC type criterion indicates an optimal dimension of $K = 2$ (AIC values in Table 1 are shown as differences from the lowest AIC value). These first two basis functions are able to explain 99.95% of the variance. The minimization of the cross-validation criterion (5) for $K = 2$ yields an undersmoothing parameter of $\underline{b}_2$ that is only two-tenths of the usual cross-validation smoothing parameter b^{opt}; see Table 1.

Based on the undersmoothed and scaled estimators $\tilde{X}_t^* = \tilde{X}_t / \|\hat{X}_t\|$, we compute the estimator $\hat{\gamma}_T$ of the empirical covariance function γ_T by local linear surface

TABLE 1

Undersmoothing parameters $\underline{b}_K$ (shown as fractions of the usual cross-validation smoothing parameter b^{opt}), AIC values (shown as differences from the lowest AIC value) and cumulative variances for the dimensions $K \in \{1, 2, 3\}$

K	$\underline{b}_K / b^{\mathrm{opt}}$	AIC	Cum. var.
1	0.1	596.4	92.62%
2	0.2	0	99.95%
3	0.3	52.9	99.97%

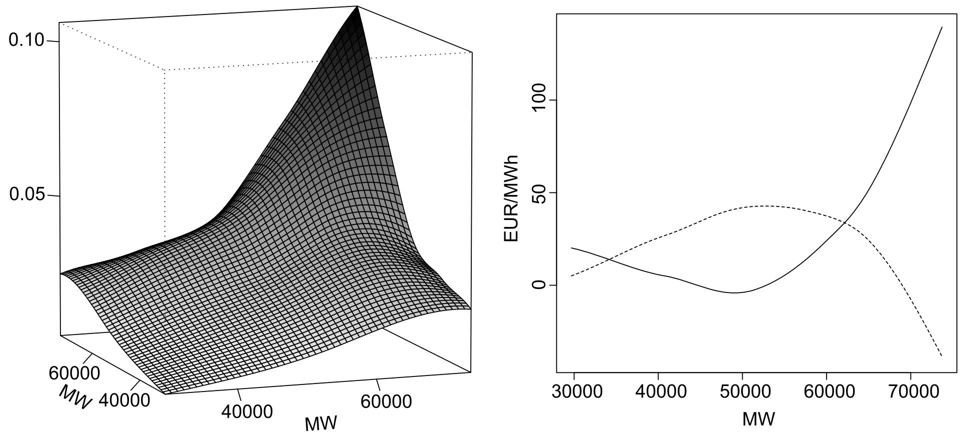

FIG. 3. LEFT PANEL: *Empirical covariance function* $\hat{\gamma}_T$. RIGHT PANEL: *VARIMAX rotated basis functions* $\hat{f}_1$ (*solid line*) *and* $\hat{f}_2$ (*dashed line*), *scaled by the average scores* $\bar{\hat{\beta}}_{\cdot 1}$ *and* $\bar{\hat{\beta}}_{\cdot 2}$, *respectively.*

smoothing, as explained in Section 4.4. The result is shown in the left panel of Figure 3. The plot of the estimator $\hat{\gamma}_T$ shows clearly that the sample variance of the standardized price-demand functions $\tilde{X}_t^*$ increases monotonically with electricity demand.

The estimators of the first two empirical eigenfunctions $\hat{e}_{T,1}$ and $\hat{e}_{T,2}$ are determined from the eigendecomposition of a discretized version of the estimated empirical covariance function $\hat{\gamma}_T$ using an equidistant grid of $n \times n$ discretization points of the plane $[A, B]^2$. The estimation of the smooth eigenfunctions by discretizing the smooth covariance function is common in the FDA literature; see, for example, Rice and Silverman (1991).

In order to find an appropriate number n of discretization points, there is the following trade-off, which has to be considered: on the one hand, n must be small enough that the algorithm to compute the eigendecomposition runs stable. On the other hand, n must be great enough that the $n \times n$-matrix of discretization points forms a good approximation to the covariance function. The choice of $n = 50$ appears to be appropriate for our application. As a robustness check we also tried values of n ranging from 20 to 70, which yield nearly identical results.

We rotate the basis system of the estimated eigenfunctions $\{\hat{e}_{T,1}, \hat{e}_{T,2}\}$ by the VARIMAX-criterion in order to get interpretable basis functions $\hat{f}_1$ and $\hat{f}_2$. The two rotated basis functions $\hat{f}_1$ and $\hat{f}_2$ explain 58.63% and 41.32% of the total sample variance of the price-demand functions $\hat{X}_t$.

It is convenient to choose an appropriate scaling of the graphs of the basis functions $\hat{f}_1$ and $\hat{f}_2$ in order to plot them with a reasonable order of magnitude. We scale the graphs by their corresponding average scores $\bar{\hat{\beta}}_{\cdot i} = T^{-1} \sum_{t=1}^{T} \hat{\beta}_{ti}$ for

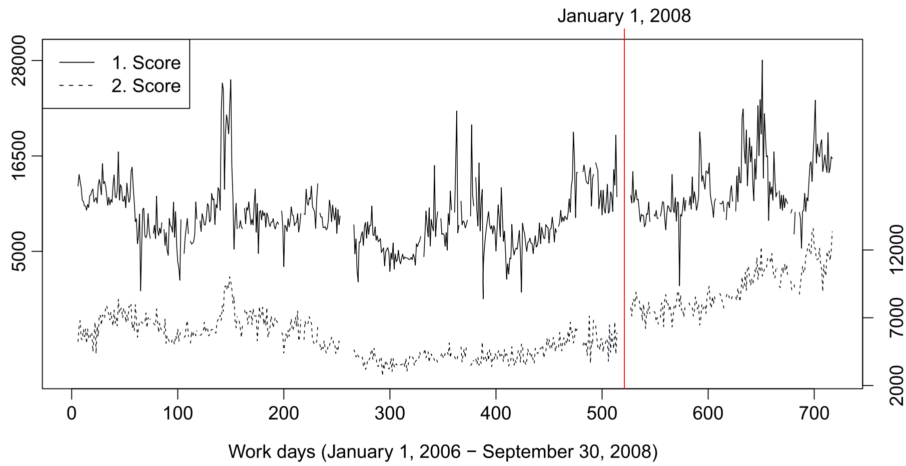

FIG. 4. *Time series of the first scores* ($\hat{\beta}_{t1}$) (*solid line*) *and second scores* ($\hat{\beta}_{t2}$) (*dashed line*). *The vertical red line separates the initial learning sample from the initial forecasting sample. Gaps in the time series correspond to holidays.*

$i \in \{1, 2\}$. In the right panel of Figure 3 the graph of $\hat{f}_1\bar{\hat{\beta}}_{\cdot 1}$ is plotted as a solid line, whereas the graph of $\hat{f}_2\bar{\hat{\beta}}_{\cdot 2}$ is plotted as a dashed line.

Given the basis system $\{\hat{f}_1, \hat{f}_2\}$, we re-estimate the functions $X_1, \ldots, X_T$ by (12) such that

$$\hat{X}_t^f = \hat{\beta}_{t1}\hat{f}_1 + \hat{\beta}_{t2}\hat{f}_2.$$

To simplify the notation, we write $\hat{X}_t = \hat{X}_t^f$ from now on. The coefficients $\hat{\beta}_{t1}$ and $\hat{\beta}_{t2}$ are determined by OLS regressions of $\hat{X}_t$ simultaneously on $\hat{f}_1$ and $\hat{f}_2$ after discretizing the functions at the N_t observed values of electricity demand $u_{t1}, \ldots, u_{tN_t}$. The time series of the scores are shown in Figure 4.

5.1. *Interpretation of the factors and exemplary analysis of the scores.* Remember that we do not use a mean function in our FFM. Consequently, the classical interpretation of the factors $\hat{f}_k$ as perturbations of the mean does not apply. A reasonable interpretation of the estimated factors $\hat{f}_1$ and $\hat{f}_2$ can be derived from the classical micro-economic point of view on electricity spot prices; see also the discussion in Section 2.

This point of view allows us to interpret the price-demand functions $\hat{X}_t(u) = \hat{\beta}_{t1}\hat{f}_1(u) + \hat{\beta}_{t2}\hat{f}_2(u)$ as daily *empirical merit order curves* or *empirical supply functions*, where the shape of the curves $\hat{X}_t$ is determined by the factors $\hat{f}_1$ and $\hat{f}_2$ and the scores $\hat{\beta}_{t1}$ and $\hat{\beta}_{t2}$. For example, steep empirical supply functions have high score ratios $\hat{\beta}_{t1}/\hat{\beta}_{t2}$ and vice versa. Since steep supply functions are associated with high prices, we could interpret the first factor $\hat{f}_1$ as the high-price component and the second factor $\hat{f}_2$ as the moderate-price component. In general,

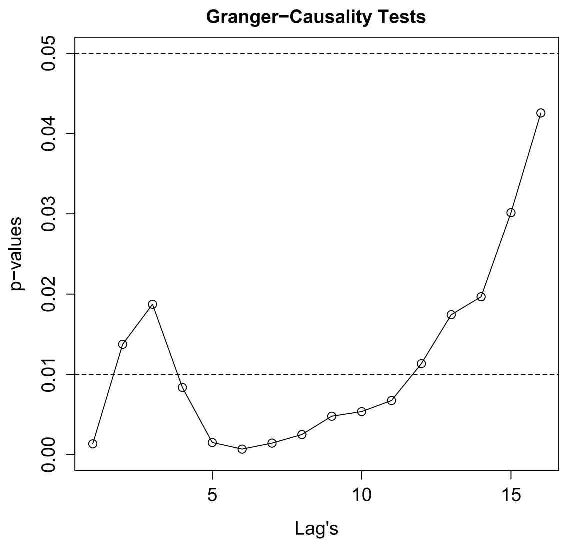

FIG. 5. *p-values of Granger-causality tests of whether the time-varying steepness of the price-demand functions* [*quantified as time series of score ratios* $(\hat{\beta}_{t1}/\hat{\beta}_{t2})$] *is Granger-caused by past values of the time series of extreme temperatures.*

any interpretation of the factors has to be done with caution since they are only identified up to orthonormal rotations.

Particularly, the scores $\hat{\beta}_{t1}$ and $\hat{\beta}_{t2}$ are useful for a further analysis of the dynamics of the empirical supply functions. For example, researchers or risk analysts, who wish to predict days with high electricity prices, could try to predict days with steep empirical supply functions $\hat{X}_t$.

Days with steep supply functions represent market situations with capacity constraints, that is, situations in which power plants with high generation costs are needed to supply the demanded amount of electricity. There are several causes for capacity constraints, such as extreme temperatures or power plant outages.

In fact, the time-varying steepness of the empirical supply functions [quantified as time series of score ratios $(\hat{\beta}_{t1}/\hat{\beta}_{t2})$] is Granger-caused by the time series of extreme temperatures (defined as absolute temperature deviations from the mean temperature), where the temperature data is available from the German Weather Service (www.dwd.de). Figure 5 shows the p-values of the corresponding Granger-causality tests [Granger (1969)].

5.2. *Validation of the model assumptions.* The overall in-sample data fit of the estimated spot prices $\hat{y}_{th} = \hat{X}_t(u_{th})$, measured by the R^2-parameter, is given by $R^2 = 0.92$ and indicates a good model fit. Nevertheless, our implicit stability assumption in (2) that $X_t \in \mathcal{H}_K$ for all $t \in \{1, \dots, T\}$, that is, that all functions X_t are elements of the same space $\mathcal{H}_K$, may be seen as critical.

In our context it is impossible to validate the stability assumption by statistical tests such as in Benko, Härdle and Kneip (2009) since we do not assume that the time series of the scores (β_{t1}) and (β_{t2}) are stationary. However, we can compare different basis systems estimated from subsets of the data with each other. The stability assumption can be seen as supported if all of these subset-basis functions span the same space $\hat{\mathcal{H}}_K$.

We define half-yearly data-subsets 6-1, 6-2, 7-1, 7-2 and one nine-month data-subset 8-1 by choosing according to subsets of the index set $\{1,\dots,T\}$ and investigate the R^2-parameters from subset regressions—such as, for example, $\hat{e}_1^{(6\text{-}1)}(u)$ simultaneously on $\hat{e}_1^{(6\text{-}2)}(u)$ and $\hat{e}_2^{(6\text{-}2)}(u)$. This assesses whether the eigenfunction $\hat{e}_1^{(6\text{-}1)}$ can be seen as an element of the space spanned by the basis system $\{\hat{e}_1^{(6\text{-}2)}, \hat{e}_2^{(6\text{-}2)}\}$.

The results are given in Table 2 and clearly support our assumption that $X_t \in \mathcal{H}_K$ for all $t \in \{1,\dots,T\}$. The R^2-values with respect to the first eigenfunctions $\hat{e}_1^{(6\text{-}1)}, \hat{e}_1^{(6\text{-}2)}, \dots, \hat{e}_1^{(8\text{-}1)}$ and $\hat{e}_{T,1}$ are all greater than or equal to 0.99. Also, the R^2-values with respect to the second eigenfunctions $\hat{e}_2^{(6\text{-}1)}, \hat{e}_2^{(6\text{-}2)}, \dots, \hat{e}_2^{(8\text{-}1)}$ and $\hat{e}_{T,2}$ indicate no clear violation of our model assumption.

The R^2-values with respect to the second eigenfunctions are systematically smaller than those with respect to the first eigenfunctions, since the first order bias term of an estimated eigenfunction is inversely related to the pairwise distances of its eigenvalue to all other eigenvalues; see Benko, Härdle and Kneip (2009), Theorem 2(iii). By construction, these distances are greatest for the first eigenvalue.

Finally, we test for (non-)stationarity of the time series of the scores $(\hat{\beta}_{t1})$ and $(\hat{\beta}_{t2})$ using the usual testing procedures such as the KPSS-tests for stationarity and ADF-tests for nonstationarity (with a 5%-significance level for all tests). The results allow us to assume that the time series of the scores $(\hat{\beta}_{t1})$ and $(\hat{\beta}_{t2})$ are

TABLE 2

Descriptive validation of the assumption that $X_t \in \mathcal{H}_K$ for all $t \in \{1,\dots,T\}$. The list elements are R^2-values, which stem from subset regressions of, for example, the eigenfunction $\hat{e}_1^{(6\text{-}1)}$ on the eigenfunctions $\{\hat{e}_1^{(6\text{-}2)}, \hat{e}_2^{(6\text{-}2)}\}$ in the upper left case with $R^2 = 0.99$

	$\hat{e}_1^{(6\text{-}1)}$ $\hat{e}_2^{(6\text{-}1)}$	$\hat{e}_1^{(6\text{-}2)}$ $\hat{e}_2^{(6\text{-}2)}$	$\hat{e}_1^{(7\text{-}1)}$ $\hat{e}_2^{(7\text{-}1)}$	$\hat{e}_1^{(7\text{-}2)}$ $\hat{e}_2^{(7\text{-}2)}$	$\hat{e}_1^{(8\text{-}1)}$ $\hat{e}_2^{(8\text{-}1)}$
$\{\hat{e}_1^{(6\text{-}1)}, \hat{e}_2^{(6\text{-}1)}\}$	— —	0.99 0.95	1.00 0.96	1.00 0.89	1.00 0.99
$\{\hat{e}_1^{(6\text{-}2)}, \hat{e}_2^{(6\text{-}2)}\}$	0.99 0.95	— —	0.99 0.83	1.00 0.98	1.00 0.95
$\{\hat{e}_1^{(7\text{-}1)}, \hat{e}_2^{(7\text{-}1)}\}$	1.00 0.96	0.99 0.83	— —	1.00 0.78	1.00 0.95
$\{\hat{e}_1^{(7\text{-}2)}, \hat{e}_2^{(7\text{-}2)}\}$	1.00 0.89	1.00 0.98	1.00 0.78	— —	1.00 0.89
$\{\hat{e}_1^{(8\text{-}1)}, \hat{e}_2^{(8\text{-}1)}\}$	1.00 0.99	1.00 0.95	1.00 0.95	1.00 0.89	— —
$\{\hat{e}_{T1}, \hat{e}_{T2}\}$	1.00 0.99	1.00 0.96	1.00 0.95	1.00 0.90	1.00 0.99

nonstationary. Detailed reports are not shown for reasons of space but can be reproduced using the R-Code provided as part of the supplementary material; see Liebl (2013).

This section demonstrates a very good and stable in-sample fit of our FFM. Of course, this cannot guarantee a good out-of-sample performance.

6. Forecasting. For our forecasting study we divide the data set into a learning sample of days $t \in \{1, \dots, T_L\}$ and a forecasting sample of days $t \in \{T_L + 1, \dots, T\}$, where the initial $T_L + 1$ corresponds to January 1, 2008 and T to September 30, 2008. The learning sample is used to estimate the parameters and the forecasting sample is used to assess the forecast performance. We enlarge the learning sample after each ℓ days ahead forecast by one day. For $\ell \in \{1, \dots, 20\}$ days ahead forecasts this leads to $T - T_L - (\ell - 1) = 717 - 521 - (\ell - 1) = 197 - \ell$ work days that can be used to assess the forecast performance of our model.

Figure 6 shows the whole data set of $717 \cdot 24 = 17{,}208$ hourly electricity spot prices along with the indicated time spans of the initial learning and forecasting sample. Gaps in the time series correspond to holidays. At least from a visual perspective, the learning sample and the forecasting sample are of comparable complexity.

In the following Section 6.1 we discuss forecasting of electricity spot prices using the FFM. In Section 6.2 we formally introduce four competing forecasting models (two classical and two FDA models), and in Section 6.3 we compare their predictive performance.

As noted above for the FFM, the two competing FDA models also use learning data with electricity spot prices below 200 EUR/MWh only. In contrast, the

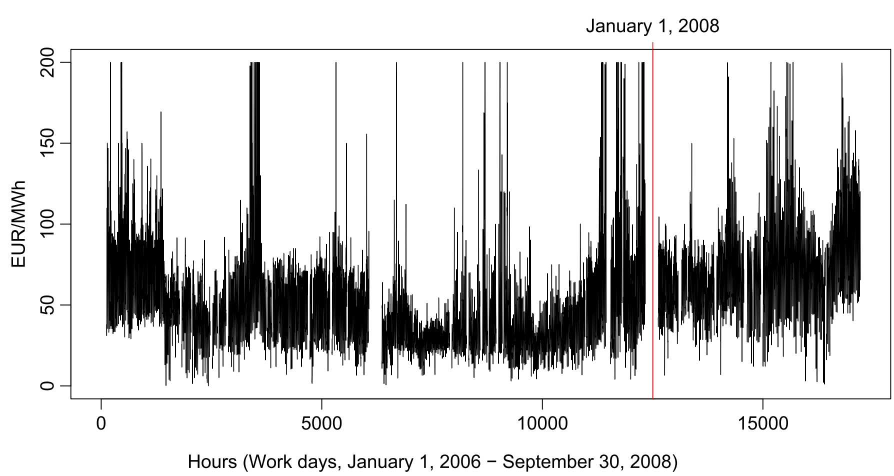

FIG. 6. *The whole data set of* 17,208 *hourly electricity spot prices. The vertical red line separates the initial learning sample from the initial forecasting sample. Gaps in the time series correspond to holidays.*

forecasting sample retains all data, including those with spot prices above 200 EUR/MWh. Advanced outlier forecast procedures, which might yield better predictive performances, are beyond the scope of this paper.

6.1. *Forecasting with the FFM.* The computation of the ℓ days ahead forecast $\hat{y}_{T_L,h}(\ell) \in \mathbb{R}$ of the electricity spot price $y_{T_L+\ell,h}$ given the information set of the learning sample, say, $\mathcal{I}_{T_L}$, involves the computation of the conditional expectation of a nonlinearly transformed random variable, namely, $E[f_k(u_{T_L+\ell,h})|\mathcal{I}_{T_L}]$. We approximate the latter using the naive plug-in predictor $\hat{f}_k(\hat{u}_{T_L,h}(\ell))$, where $\hat{u}_{T_L,h}(\ell) = E[u_{T_L+\ell,h}|\mathcal{I}_{T_L}]$. This yields to

$$\hat{y}_{T_L,h}(\ell) = \hat{\beta}_{T_L,1}(\ell)\hat{f}_1(\hat{u}_{T_L,h}(\ell)) + \hat{\beta}_{T_L,2}(\ell)\hat{f}_2(\hat{u}_{T_L,h}(\ell)), \tag{13}$$

where $\hat{\beta}_{T_L,1}(\ell)$, $\hat{\beta}_{T_L,2}(\ell)$ and $\hat{u}_{T_L,h}(\ell)$ are the ℓ days ahead forecasts of the scores $\hat{\beta}_{T_L+\ell,1}$, $\hat{\beta}_{T_L+\ell,2}$ and of the electricity demand value $u_{T_L+\ell,h}$.

The naive plug-in predictor $\hat{f}_k(\hat{u}_{T_L,h}(\ell))$ is a rather simple approximation of the conditional expectation $E[f_k(u_{T_L+\ell,h})|\mathcal{I}_{T_L}]$. Here, it performs very well because the basis functions are relatively smooth. In the case of more complex basis functions, it might be necessary to improve the approximation using higher order Taylor expansions of $\hat{f}_k$ around $\hat{u}_{T_L,h}(\ell)$.

We use the following univariate SARIMA$(0, 1, 6) \times (0, 1, 1)_5$-models to forecast the time series of the scores $(\hat{\beta}_{ti})$ with $i \in \{1, 2\}$:

$$(1 - B)(1 - B^5)\hat{\beta}_{ti} = \left(1 + \sum_{l=1}^{6} \delta_{il} B^l\right)(1 + \delta_i^S B^5)\omega_{ti}, \tag{14}$$

where B is the back shift operator. In order to ensure that the SARIMA models (14) are not sample dependent, we select them from a set of reasonable alternative SARIMA models, where all of them are confirmed by the usual diagnostics on the residuals. Each of the confirmed models is applied to different subsets of the learning sample and the final model selection is done by the AIC.[2] As usual, the ℓ days ahead forecasts $\hat{\beta}_{T_L,1}(\ell)$ and $\hat{\beta}_{T_L,2}(\ell)$ are given by the conditional expectations of $\hat{\beta}_{T_L+\ell,1}$ and $\hat{\beta}_{T_L+\ell,2}$ given the data from the learning sample; see, for example, Brockwell and Davis (1991).

A first visual impression of the forecast performance is given in Figure 7, which compares the 24 hourly spot prices y_{th} with the 1 day ahead forecast of the price-demand function X_t. The 1 day ahead forecast of the price-demand function X_t is defined as

$$\hat{X}_{T_L}(\ell) = \hat{\beta}_{T_L,1}(\ell)\hat{f}_1 + \hat{\beta}_{T_L,2}(\ell)\hat{f}_2 \in L^2[A, B]. \tag{15}$$

[2]The interested reader is referred to the R-Code provided as part of the supplementary material; see Liebl (2013).

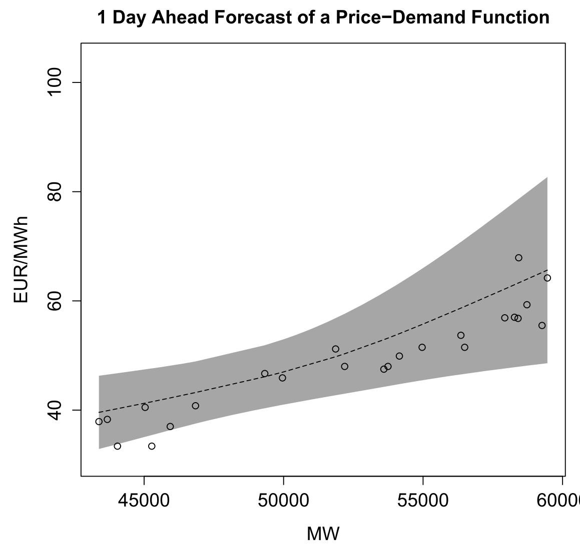

FIG. 7. *A comparison of the* 24 *hourly electricity spot prices* y_{th} (*circles*) *and the* 1 *day ahead forecast of the price-demand function* X_t (*dashed line*) *of January* 25, 2008. *The* 95% *forecast interval is plotted as a gray shaded band.*

Additionally, a 95% forecast interval is plotted as a gray shaded band. The forecast interval is computed on the basis of the 95% forecast intervals of the SARIMA forecasts $\hat{\beta}_{T_L,1}(\ell)$ and $\hat{\beta}_{T_L,2}(\ell)$ and has to be interpreted as a conditional forecast interval given the realizations $\hat{f}_1$ and $\hat{f}_2$.

In order to be able to forecast the hourly electricity spot prices y_{th}, we also have to forecast the hourly values of electricity demand u_{th}; see equation (13). Given our definition of electricity demand in Section 2, a ℓ days ahead forecast of electricity demand $u_{T_L,h}(\ell)$ involves forecasting gross demand for electricity as well as wind power infeed data. The statistician has to choose appropriate models—one for gross electricity demand such as that proposed in Antoch et al. (2010) and another for wind power such as that proposed in Lau and McSharry (2010). For the sake of simplicity, we use the two reference cases of a "persistence" and an "ideal" forecast of electricity demand:

persistence The persistence (or "no-change") forecast $\hat{u}^{\text{persi}}_{T_L,h}(\ell)$ is given by the last value of electricity demand that is still within the learning sample, that is, $\hat{u}^{\text{persi}}_{T_L,h}(\ell) = u_{T_L,h}$.

ideal The ideal forecast is given by $u_{T_L+\ell,h}$ itself, that is, $\hat{u}^{\text{ideal}}_{T_L,h}(\ell) = u_{T_L+\ell,h}$.

This yields a range for possible electricity demand forecasts with bounds that can be easily interpreted.

A first visual comparison of the observed hourly electricity spot prices $y_{T_L+1,h}$ with their 1 day ahead forecasts $\hat{y}_{T_L,h}(1)$ is given in Figure 8. The left panel

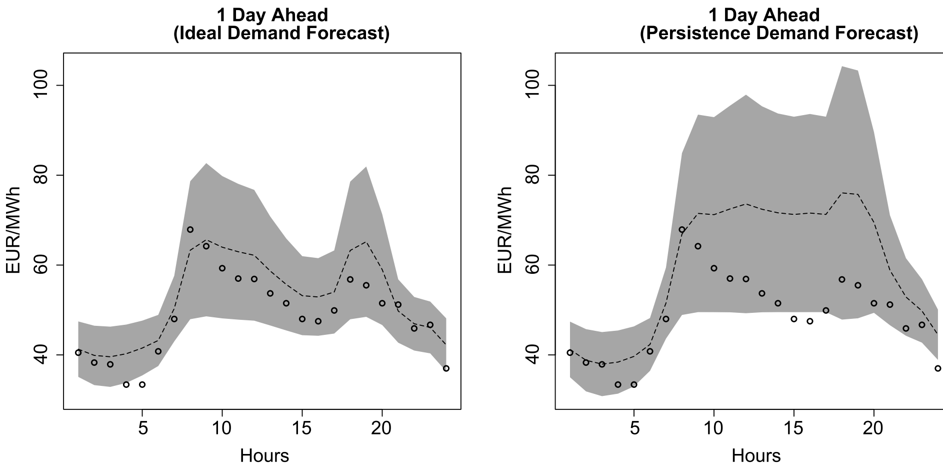

FIG. 8. LEFT PANEL: *comparison of the spot prices* y_{th} (*circles*) *and the* 1 *day ahead forecasts* $\hat{y}_{T_L,h}(\ell)$ (*dashed line*) *of January* 25, 2008 *based on ideal demand forecasts*. RIGHT PANEL: *comparison of the spot prices* y_{th} (*circles*) *and the* 1 *day ahead forecasts* $\hat{y}_{T_L,h}(\ell)$ (*dashed line*) *of January* 25, 2008 *based on persistence demand forecasts*. BOTH PANELS: *the* 95% *forecast intervals are plotted as gray shaded bands*.

demonstrates the ideal forecast case and shows the spot prices y_{th} (circles) and their 1 day ahead forecasts $\hat{y}_{T_L,h}(1)$ (dotted line) based on the electricity demand forecasts $\hat{u}^{\text{ideal}}_{T_L,h}(1)$. The right panel demonstrates the persistence case and shows the spot prices y_{th} (circles) and their 1 day ahead forecasts $\hat{y}_{T_L,h}(1)$ (dotted line) based on the electricity demand forecasts $\hat{u}^{\text{persi}}_{T_L,h}(1)$. The 95% forecast intervals are plotted as gray shaded bands. The forecast interval shown in the right panel is much broader than that shown in the left panel. This is because the forecasted electricity spot prices based on the persistence electricity demand forecasts are too high, and higher electricity spot prices have broader 95% forecast intervals; see Figure 7.

6.2. *Competing forecast models*. In this section we introduce the four competing forecast models (two classical and two FDA models). The two classical models, referred to as AR and MR models, are archetypal representatives of the classical approaches in the literature on forecasting electricity spot prices; see, for example, Kosater and Mosler (2006). The AR model is an autoregressive model and the MR model is the Markov regime switch model for electricity spot prices proposed by Huisman and De Jong (2003).

The two FDA models are the above-discussed DSFM model of Park et al. (2009) and the semi-functional partial linear (SFPL) model of Vilar, Cao and Aneiros (2012). Both of these FDA models have been successfully applied to forecast electricity spot prices [Härdle and Trück (2010) and Vilar, Cao and Aneiros (2012)] and are expected to be more challenging competitors for our FFM than the two classical models.

Before the formal introduction of the four alternative forecast models, we need some unifying notation. The problem is that the two classical models, AR and MR, are designed to forecast only daily *aggregated* peakload and baseload spot prices defined as

$$y_t^P = \log\left(\frac{1}{12}\sum_{h=9}^{20} y_{th}\right) \quad \text{and} \quad y_t^B = \log\left(\frac{1}{24}\sum_{h=1}^{24} y_{th}\right).$$

In contrast to this, the three FDA models are designed to forecast the *hourly* electricity spot prices y_{th}. Therefore, we define the forecasts of the peakload-aggregates $\hat{y}_{T_L}^P(\ell|\texttt{Model})$ and baseload-aggregates $y_{T_L}^B(\ell|\texttt{Model})$ of the FDA models as

$$\hat{y}_{T_L}^P(\ell|\texttt{Model}) = \log\left(\frac{1}{12}\sum_{h=9}^{20} \hat{y}_{T_L,h}(\ell|\texttt{Model})\right)$$

and

$$\hat{y}_{T_L}^B(\ell|\texttt{Model}) = \log\left(\frac{1}{24}\sum_{h=1}^{24} \hat{y}_{T_L,h}(\ell|\texttt{Model})\right),$$

where $\hat{y}_{T_L,h}(\ell|\texttt{Model})$ is the ℓ days ahead hourly electricity spot price forecast of the $\texttt{Model} \in \{\text{FFM}, \text{DSFM}, \text{SFPL}\}$. By Jensen's inequality, these definitions yield aggregated forecasts of the FDA models, which tend to be too high, that is, $E[y_{T_L+\ell}^A|\mathcal{I}_{T_L}] \le \hat{y}_{T_L}^A(\ell|\texttt{Model})$ with $A \in \{P, B\}$. Therefore, the RMSEs of the FDA models shown in Figure 9 tend to be inflated and can be interpreted as being conservative.

In the following we formally introduce the four competing forecast models. Further details can be found in Kosater and Mosler (2006), Park et al. (2009) and Vilar, Cao and Aneiros (2012).

AR. The first benchmark model is the classical AR(1) model with an additive constant drift component and a time-varying deterministic component. The AR model can be defined as

$$y_t^A = d^A + g_t^A + \alpha y_{t-1}^A + \omega_t^A, \qquad \omega_t^A \sim \mathcal{N}(0, \sigma_{\omega^A}^2), \tag{16}$$

where $A \in \{P, B\}$ refers to the type of aggregation (peakload or baseload), d^A is the constant drift parameter, and g_t^A captures daily, weekly and yearly deterministic effects of the peakload and baseload prices, respectively.

MR. The second benchmark model is the Markov regime switch model proposed by Huisman and De Jong (2003). The MR model extends the AR model (16) and distinguishes between two different regimes $R_t^A \in \{M, S\}$, where M denotes the regime of moderate prices and S denotes the regime of price spikes. The MR model can be defined as

$$y^A_{M,t} = d^A + \alpha^A y^A_{M,t-1} + \omega^A_{M,t}, \qquad (17)$$
$$y^A_{S,t} = \mu^A_S + \omega^A_{S,t},$$

where $A \in \{P, B\}$ refers to the type of aggregation (peakload or baseload), $\omega^A_{M,t} \sim \mathcal{N}(0, \sigma^2_{M^A})$ and $\omega^A_{S,t} \sim \mathcal{N}(0, \sigma^2_{S^A})$. The conditional probabilities of the transitions from one regime to another given the regime at $t-1$ are captured by the transition matrix

$$\begin{pmatrix} \mathrm{P}(R^A_t = M | R^A_{t-1} = M) & \mathrm{P}(R^A_t = M | R^A_{t-1} = S) \\ \mathrm{P}(R^A_t = S | R^A_{t-1} = M) & \mathrm{P}(R^A_t = S | R^A_{t-1} = S) \end{pmatrix} = \begin{pmatrix} q & 1-p \\ 1-p & p \end{pmatrix}$$

and have to be estimated, too.

DSFM. The third model, the DSFM of Park et al. (2009), is a functional factor model, which is very similar to our FFM. Its application to electricity spot prices, as suggested by Härdle and Trück (2010), differs from our application, since it models the hourly spot prices y_{th} based on the classical time series point of view on electricity spot prices. That is, Härdle and Trück model and forecast nonparametric price-hour functions, say, $\chi_t(h)$, and thereby fail to consider the merit order model. The DSFM can be written as

$$y_{th} = \chi_t(h) + \omega_{th}, \qquad h \in \{1, \dots, 24\}, \qquad (18)$$

with $\chi_t \in L^2[1, 24]$ defined as

$$\chi_t(h) = f^{\mathrm{DSFM}}_0(h) + \sum_{l=1}^{L} \beta^{\mathrm{DSFM}}_{tl} f^{\mathrm{DSFM}}_l(h),$$

where $f^{\mathrm{DSFM}}_0(h)$ is a nonparametric mean function, $f^{\mathrm{DSFM}}_l(h)$ are nonparametric functional factors, $\beta^{\mathrm{DSFM}}_{tl}$ are the univariate scores, and ω_{th} is a Gaussian white noise process.

Park et al. suggest selecting the number of factors L by the proportion of explained variation. We choose the factor dimension $\hat{L} = 2$, since this factor dimension yields the same proportion of explained variation as the factor dimension $\hat{K} = 2$ for our FFM.

Given the estimates of the time-invariant model components, $\hat{f}^{\mathrm{DSFM}}_0(h)$, $\hat{f}^{\mathrm{DSFM}}_l(h)$ and $\hat{L}$, forecasting of the daily price-hour functions $\chi_t(h)$ can be done by forecasting the estimated univariate time series of scores. As for our FFM, we use SARIMA models to forecast the univariate time series $(\hat{\beta}^{\mathrm{DSFM}}_{t1})$ and $(\hat{\beta}^{\mathrm{DSFM}}_{t2})$, where the model selection procedure for the SARIMA models is the same as for our FFM.

SFPL. The fourth model, the SFPL model of Vilar, Cao and Aneiros (2012), is a very recent functional data model, which was exclusively designed for forecasting electricity spot prices. The SFPL has the nice property of allowing us to include the values of electricity demand u_{th} as additional co-variables. Vilar, Cao and Aneiros use dummy variables for work days and holidays as additional covariates, which we do not have to do since we consider only work days. Note that, like the DSFM, the SFPL model uses price-hour functions $\chi_t(h)$ and therefore does not consider the merit order model. The definition of the SFPL is given in the following:

$$y_{t+\ell,h} = \alpha u_{th} + m\bigl(\chi_t(h)\bigr) + \omega_{th}, \tag{19}$$

where $m : L^2[1, 24] \to \mathbb{R}$ is a function that maps the price-hour function χ_t to a real value and ω_{th} is a Gaussian white noise process.

Forecasting electricity spot prices y_{th} with the SFPL model can be easily done using the R package `fda.usc` of Febrero-Bande and Oviedo de la Fuente (2012). However, in order to compute the forecasts $\hat{y}_{T_L,h}(\ell|\text{SFPL})$ of the electricity spot prices $y_{T_L+\ell,h}$, we also need forecasts of the electricity demand values $u_{T_L+\ell,h}$. We cope with this problem as suggested above for our FFM by using a persistence forecast and an ideal forecast.

6.3. *Evaluation of forecast performances.* The two plots of Figure 9 show the values of the RMSEs for the $\ell \in \{1, \dots, 20\}$ days ahead forecasts of the peakload prices (left panel) and the baseload prices (right panel). The two gray shaded

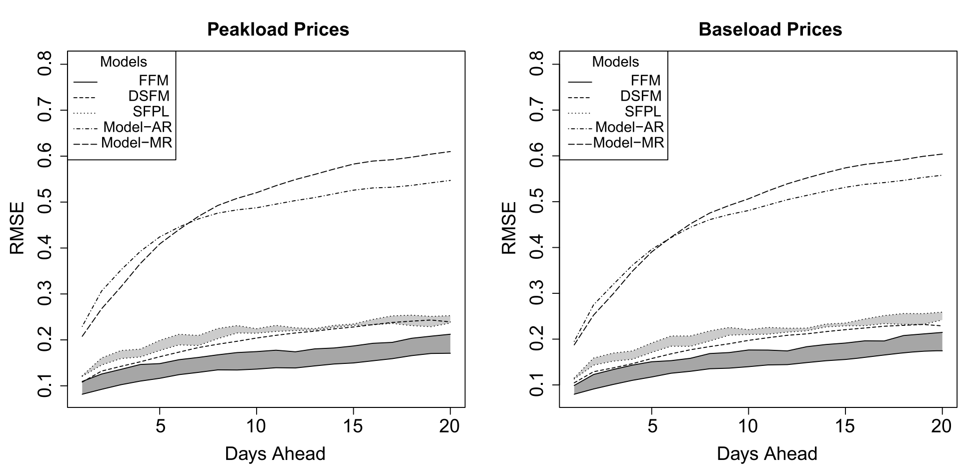

FIG. 9. *Root mean squared errors of the FFM (solid lines) and the alternative models, DSFM (short-dashed lines), SFPL (dotted lines), AR (dash-dotted lines) and MR (long-dashed lines) for peakload prices y_t^P (left panel) and baseload prices y_t^B (right panel). The gray shaded regions for the FFM and the SFPL model are lower bounded based on the ideal demand forecast, and upper bounded based on the persistence forecast.*

regions in each plot show the possible RMSE values of the FFM (solid line borders) and the SFPL model (dotted line borders). The lower bounds of the regions are based on the ideal electricity demand forecast $\hat{u}^{\text{ideal}}_{T_L,h}(\ell)$. The upper bounds are based on the persistence forecast of electricity demand $\hat{u}^{\text{persi}}_{T_L,h}(\ell)$.

The poor performance of the two classical time series models, AR and MR, in comparison to the three FDA models, FFM, DSFM and SFPL, can be explained by the different approaches to model the aggregated peakload and baseload prices. The two classical models try to forecast the aggregated prices directly, whereas the three FDA models try to forecast the hourly electricity spot prices; aggregation is done afterward.

The superior performance of our FFM in comparison to the other two FDA models, DSFM and SFPL, can be explained by the FFMs explicit consideration of the merit order model. Both models, the DSFM and the SFPL, work with daily price-hour functions $\chi_t(h)$, which are based on a rather simple transfer of the classical time series point of view to a functional data point of view. By contrast, the FFM works with daily price-demand functions $X_t(u)$, which are based on the merit order model, the most important model for explaining electricity spot prices; see our discussion in Section 1. Finally, the DSFM generally performs better than the SFPL model. This might be explained by the fact that the SFPL model of Vilar, Cao and Aneiros (2012) is an autoregressive model of order one. Vilar, Cao and Aneiros do not discuss the possibility of extending the order structure of their SFPL model.

The above study of the RMSEs only gives us insights into the forecast performances with respect to point forecasts. In order to complement the forecast comparisons, we also consider interval forecasts. In this regard, the interval score, proposed by Gneiting and Raftery (2007), is a very informative statistic. The interval score can be defined as

$$\begin{aligned}S^{\text{int}}_{\alpha}(h,\ell) = (\hat{b}_u - \hat{b}_l) &+ \frac{2}{\alpha}(\hat{b}_l - y_{T_L+\ell,h})\mathbb{I}\{y_{T_L+\ell,h} < \hat{b}_l\} \\ &+ \frac{2}{\alpha}(y_{T_L+\ell,h} - \hat{b}_u)\mathbb{I}\{y_{T_L+\ell,h} > \hat{b}_u\},\end{aligned}$$

where $\hat{b}_u = \hat{b}_{u,T_L,h}(\ell)$ and $\hat{b}_l = \hat{b}_{l,T_L,h}(\ell)$ are the lower and upper bounds of the $(1-\alpha)\%$ forecast interval for the electricity spot price $y_{T_L+\ell,h}$. The interval score punishes a broad prediction interval $(\hat{b}_u - \hat{b}_l)$ and adds an additional punishment if the actual observation $y_{T_L+\ell,h}$ is not within the prediction interval. In general, a lower interval score is a better one.

Unfortunately, we cannot compute the interval scores for all five models. For example, Vilar, Cao and Aneiros (2012) do not propose any prediction intervals for the SFPL model. Furthermore, while it is easy to compute forecast intervals of the FFM and the DSFM for hourly sport prices, it is not trivial to compute them for the aggregated (peakload and baseload) prices.

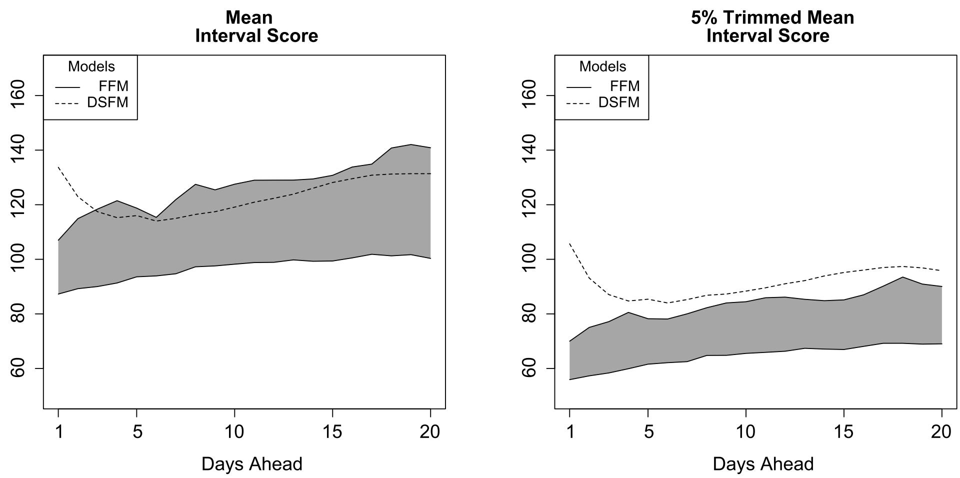

FIG. 10. *Mean and trimmed mean values of the interval scores* $S_\alpha^{\mathrm{int}}(h, \ell)$, *pooled for all hours* $h \in \{1, \dots, 24\}$. *A low interval score stands for precise predictions with narrow prediction intervals. The dashed line corresponds to the interval scores of the DSFM of Park et al.* (2009). *The gray shaded regions for the FFM are lower bounded based on the ideal demand forecast, and upper bounded based on the persistence forecast.*

Therefore, we focus on the hourly forecasts of electricity spot prices of the FFM and DSFM models. For both models, the 95% forecast intervals can be computed on the basis of the 95% forecast intervals of the SARIMA forecasts given the estimated factors.

Due to the enlargement of the learning sample after each ℓ days ahead forecast by one day and due to pooling all hours $h \in \{1, \dots, 24\}$, we have for each ℓ days ahead forecast $24 \cdot (197 - \ell)$ interval scores $S_\alpha^{\mathrm{int}}(h, \ell)$ in order to compare the ℓ days ahead forecast performances of our FFM and the DSFM. In Figure 10 we present the (trimmed) mean values of these pooled interval scores $S_\alpha^{\mathrm{int}}(h, \ell)$ with $\alpha = 0.05$ for each $\ell \in \{1, \dots, 20\}$. The 5% trimmed mean values are used, since for both models there are some extreme values of the interval score (due to the outliers in the forecast sample), which distort the mean values.

Figure 10 clearly confirms the good forecast performance of the FFM. Besides some technical issues, the main conceptual difference between the DSFM and our FFM is that the DSFM works with daily price-*hour* functions $\chi_t(h)$, whereas our FFM works with daily price-*demand* functions $X_t(u)$, which are suggested by the merit order model. This demonstrates that the consideration of the merit order model yields better point forecasts as well as better interval forecasts.

7. Conclusion. In this paper we suggest interpreting hourly electricity spot prices as noisy discretization points of smooth price-demand functions. This functional data perspective on electricity spot prices is motivated as well as theoretically underpinned by the merit order model—the most important pricing model for electricity spot prices.

We propose a functional factor model in order to model and forecast the non-stationary time series of price-demand functions and discuss a two-step estimation procedure. In the first step we estimate the single price-demand functions from the noisy discretization points. In the second step we robustly estimate from these a finite set of common basis functions. The careful consideration of the merit order model yields a very parsimonious functional factor model with only two common basis functions, which together explain over 99% of the total sample variation of the price-demand functions.

Our approach allows us to separate the total variations of electricity spot prices into one part caused by the variations of the merit order curves (mainly variations of input-costs) and another part caused by the variations of electricity demand. The first part is modeled by our FFM and the second part can be modeled by specialized methods proposed in the literature. We decided to keep the model parsimonious; nevertheless, it is easily possible to include the input cost for resources (coal, gas, etc.) into our FFM. Researchers are invited to extend the FFM for these co-variables.

The presentation of our functional factor model is concluded by a real data application and a forecast study which compares our FFM with four alternative time series models that have been proposed in the electricity literature. The real data application demonstrates the use of the functional factor model and a possible interpretation of the unobserved common basis functions. The forecast study clearly confirms the power of our functional factor model and the use of price-demand functions as underlying structures of electricity spot prices in general.

Acknowledgments. I want to thank Alois Kneip (University of Bonn) and Pascal Sarda (Université Paul Sabatier, Toulouse) for stimulating discussions. Section 6 profited especially from comments of the referees.

SUPPLEMENTARY MATERIAL

R-codes and data set (DOI: 10.1214/13-AOAS652SUPP; .zip). In this supplement we provide a zip file containing the R-Codes and the data set used to model and forecast electricity spot prices by the functional factor model as described in this paper.

REFERENCES

ANTOCH, J., PRCHAL, L., DE ROSA, M. R. and SARDA, P. (2010). Electricity consumption prediction with functional linear regression using spline estimators. *J. Appl. Stat.* **37** 2027–2041. MR2740138

BENEDETTI, J. K. (1977). On the nonparametric estimation of regression functions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **39** 248–253. MR0494656

BENKO, M., HÄRDLE, W. and KNEIP, A. (2009). Common functional principal components. *Ann. Statist.* **37** 1–34. MR2488343

BORAK, S. and WERON, R. (2008). A semiparametric factor model for electricity forward curve dynamics. *Journal of Energy Markets* **1** 3–16.

BOSCO, B., PARISIO, L., PELAGATTI, M. and BALDI, F. (2010). Long-run relations in European electricity prices. *J. Appl. Econometrics* **25** 805–832. MR2756987

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York. MR1093459

BURGER, M., GRAEBER, B. and SCHINDLMAYR, G. (2008). *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*. Wiley, New York.

CHRISTENSEN, T., HURN, S. and LINDSAY, K. (2009). It never rains but it pours: Modeling the persistence of spikes in electricity prices. *The Energy Journal* **30** 25–48.

DE BOOR, C. (2001). *A Practical Guide to Splines*, revised ed. *Applied Mathematical Sciences* **27**. Springer, New York. MR1900298

FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. Chapman & Hall, London. MR1383587

FEBRERO-BANDE, M. and OVIEDO DE LA FUENTE, M. (2012). Statistical computing in functional data analysis: The R package fda.usc. *Journal of Statistical Software* **51** 1–28.

FERRATY, F. and VIEU, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer, New York. MR2229687

GERVINI, D. (2008). Robust functional estimation using the median and spherical principal components. *Biometrika* **95** 587–600. MR2443177

GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548

GRANGER, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37** 424–438.

GRIMM, V., OCKENFELS, A. and ZOETTL, G. (2008). Strommarktdesign: Zur Ausgestaltung der Auktionsregeln an der EEX. *Zeitschrift für Energiewirtschaft* **32** 147–161.

HALL, P., MÜLLER, H.-G. and WANG, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. MR2278365

HÄRDLE, W. K. and TRÜCK, S. (2010). The dynamics of hourly electricity prices. SFB 649 Discussion Papers 2010-013.

HAYS, S., SHEN, H. and HUANG, J. Z. (2012). Functional dynamic factor models with application to yield curve forecasting. *Ann. Appl. Stat.* **6** 870–894. MR3012513

HUISMAN, R. and DE JONG, C. (2003). Option pricing for power prices with spikes. *Energy Power Risk Management* **7** 12–16.

HUISMAN, R., HUURMAN, C. and MAHIEU, R. (2007). Hourly electricity prices in day-ahead markets. *Energy Economics* **29** 240–248.

KARAKATSANI, N. and BUNN, D. (2008). Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting* **24** 764–785.

KNITTEL, C. R. and ROBERTS, M. R. (2005). An empirical examination of restructured electricity prices. *Energy Economics* **27** 791–817.

KOOPMAN, S. J., OOMS, M. and CARNERO, M. A. (2007). Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices. *J. Amer. Statist. Assoc.* **102** 16–27. MR2345529

KOSATER, P. and MOSLER, K. (2006). Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices. *Applied Energy* **83** 943–958.

LAU, A. and MCSHARRY, P. (2010). Approaches for multi-step density forecasts with application to aggregated wind power. *Ann. Appl. Stat.* **4** 1311–1341. MR2758330

LI, Y. and HSING, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. MR2766854

LIEBL, D. (2013). Supplement to "Modeling and forecasting electricity spot prices: A functional data perspective." DOI:10.1214/13-AOAS652SUPP.

LIU, C., RAY, S., HOOKER, G. and FRIEDL, M. (2012). Functional factor analysis for periodic remote sensing data. *Ann. Appl. Stat.* **6** 601–624. MR2976484

LOCANTORE, N., MARRON, J. S., SIMPSON, D. G., TRIPOLI, N., ZHANG, J. T. and COHEN, K. L. (1999). Robust principal component analysis for functional data. *Test* **8** 1–73. With discussion and a rejoinder by the authors. MR1707596

PARK, B. U., MAMMEN, E., HÄRDLE, W. and BORAK, S. (2009). Time series modelling with semiparametric factor dynamics. *J. Amer. Statist. Assoc.* **104** 284–298. MR2504378

RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer, New York. MR2168993

RICE, J. A. and SILVERMAN, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **53** 233–243. MR1094283

STANISWALIS, J. G. and LEE, J. J. (1998). Nonparametric regression analysis of longitudinal data. *J. Amer. Statist. Assoc.* **93** 1403–1418. MR1666636

VILAR, J. M., CAO, R. and ANEIROS, G. (2012). Forecasting next-day electricity demand and price using nonparametric functional methods. *International Journal of Electrical Power & Energy Systems* **39** 48–55.

WERON, R., BIERBRAUER, M. and TRÜCK, S. (2004). Modeling electricity prices: Jump diffusion and regime switching. *Physica A: Statistical and Theoretical Physics* **336** 39–48.

YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. MR2160561

SEMINAR FÜR WIRTSCHAFTS
UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN
ALBERTUS-MAGNUS-PLATZ
50923 KÖLN
GERMANY
E-MAIL: liebl@statistics.uni-koeln.de
AND
STATISTISCHE ABTEILUNG
UNIVERSITÄT BONN
ADENAUERALLEE 24-26
53113 BONN
GERMANY
E-MAIL: dliebl@uni-bonn.de