

SFB 649 Discussion Paper 2010-013

The dynamics of hourly electricity prices

Wolfgang Karl Härdle*
Stefan Trück**



* Humboldt-Universität zu Berlin, Germany

** Macquarie University Sydney, Australia

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

The dynamics of hourly electricity prices^{*}

Wolfgang Karl Härdle^a, Stefan Trück^{b,*},

^a*Humboldt-Universität zu Berlin, Germany and
National Central University, Taipei, Taiwan*

^b*Macquarie University Sydney, Australia*

Abstract

The dynamics of hourly electricity prices in day-ahead markets is an important element of competitive power markets that were only established in the last decade. In electricity markets, the market microstructure does not allow for continuous trading, since operators require advance notice in order to verify that the schedule is feasible and lies within transmission constraints. Instead agents have to submit their bids and offers for delivery of electricity for all hours of the next day before a specified market closing time. We suggest the use of dynamic semiparametric factor models (DSFM) for the behavior of hourly electricity prices. We find that a model with three factors is able to explain already a high proportion of the variation in hourly electricity prices. Our analysis also provides insights into the characteristics of the market, in particular with respect to the driving factors of hourly prices and their dynamic behavior through time.

Key words: Power Markets, Dynamic Semiparametric Factor Models, Day-ahead Electricity Prices

JEL Classification: G12, C19, C13, Q47

^{*} This research was supported by Deutsche Forschungsgemeinschaft through the SFB 649 'Economic Risk'.

^{*} Corresponding author. Email: strueck@efs.mq.edu.au, Address: Department of Economics, Macquarie University Sydney, NSW 2107, Australia.

1 Introduction

Since the early 1990s, structural reforms and deregulation have led to significant changes in worldwide electricity markets. Like other commodities, electricity is now traded under competitive rules what has led to the creation of market places where electricity spot and forward contracts for different delivery periods are offered. The liberalization of these markets has not only led to new challenges for market participants but also created a new field of research. Electricity markets are subject to a very specific price behavior since electricity is a very unique commodity that cannot be economically stored, see e.g. Clewlow and Strickland (2000); Kaminski (1999); Pilipovic (1997); Weron (2006). Further, end user demand is subject to strong seasonal behavior and sometimes problems with the transmission grid can lead to extreme price changes.

In the European Energy Exchange (EEX) that will be examined in this study, the spot is an hourly contract with physical delivery such that every day is divided into 24 hourly contracts. One particular feature of most electricity markets is that the spot electricity market is actually a day-ahead market that does not allow for continuous trading. The reason for this is that a classical spot market would be difficult to implement, since system operators require advance notice in order to verify that the schedule is feasible and lies within transmission constraints. In these day-ahead markets agents submit their bids and offers for delivery of electricity during each hour of the next day before a certain market closing time. Thus, dealing with the modeling and forecasting of hourly prices it is important to recall that the prices for all hours of the next day are determined at the same time using the same available information. The system price is then calculated as the equilibrium point for the aggregated supply and demand curves for each of the 24 hours.

Electricity spot prices often exhibit specific characteristics like seasonality (on the annual, weekly and daily level), mean reversion, extreme volatility and price spikes. The seasonal character of the prices is a direct consequence of demand fluctuations that mostly arise from business hours at the daily or weekly level or changing climate conditions like e.g. temperature or the number of daylight hours at the yearly scale. In addition to seasonality and mean reversion, electricity prices exhibit a remarkably huge price volatility and infrequent, but large price spikes. The price volatility - which can be as high as 50% on the daily scale - is higher than for any other commodity. A reason for this is on the one hand the above-mentioned strong seasonality of demand. On the other hand it is due to price spikes that are mainly caused by extreme load fluctuations, generation outages or transmission failures. Thus, spot prices can increase tenfold during a single hour, but these spikes are normally quite short-lived.

The specific behavior of electricity prices, in particular extreme volatility and the described price spikes have forced producers and wholesale consumers to hedge not only against volume risk but also against price movements. This in turn has significantly propelled research in electricity price modeling and forecasting. One of the first models which has been examined in the context of electricity markets is the classical mean-reversion or Vasiček-process, see e.g. Lucia and Schwartz (2002); Knittel and Roberts (2005). On the other hand, by making certain assumptions on the functional form of the supply and demand curve, Barlow (2002) derives a non-linear Ornstein-Uhlenbeck process as appropriate model for electricity prices. Early publications on models for electricity prices with a jump component include Deng (1999); Johnson and Barz (1999); Bhanot (2000); Clewlow and Strickland (2000); Knittel and Roberts (2005) and are typically based on a jump-diffusion approach as in Merton (1976). Kluge (2004) suggests a specification of a jump diffusion model including both a mean-reverting diffusion and mean-reverting jump-process. More recently, Geman and Roncoroni (2006) suggest to model electricity log-prices with a one-factor Markov jump diffusion model. A particular nice feature of their model is that the jump direction and intensity is dependent on the current level of the spot price.

Also different GARCH models have been applied to the structure of electricity prices, see e.g. Garcia et al. (2005); Knittel and Roberts (2005). These studies find that the additional use of GARCH error terms outperforms the predictive power of several other models. On the other hand, Misiolek et al. (2006) find that despite the heteroscedastic nature of the residuals in the autoregressive models, the addition of a GARCH component in the specification does not improve the accuracy of point forecasts though it slightly improves interval forecasts. Further, Hadsell (2006) applies TARARCH models in order to study the return-volume relationship in electricity futures markets, while in a more recent study Koopman et al. (2007) suggest an approach using periodic dynamic long memory regression models with GARCH errors.

Recently, also regime-switching models have gained popularity for modeling of electricity spot prices. Initially introduced in various different contexts by Quandt (1958); Goldfeld and Quandt (1973); Hamilton (1989, 1994), their use for power markets has first been suggested by Ethier and Mount (1998); Huisman and De Jong (2003) and Weron et al. (2004a). Since then, more refined regime-switching models with e.g. time-varying transition probabilities (Mount et al., 2006) as well as transition probabilities dependent on demand level relative to the supply capacity (Kanamura and Ohashi, 2008) or temperature variables (Huisman, 2008) have been suggested. Further, Haldrup and Nielsen (2006) develop a regime-switching nonperiodic long memory model for hourly electricity prices.

While several of the models mentioned above deal with daily data, also the

interest in modeling hourly electricity prices has increased, see e.g. Ramsay and Wang (1997); Wolak (1997); Szkuta et al. (1999); Borenstein et al. (2002); Saravia (2003); Longstaff and Wang (2004); Knittel and Roberts (2005); Huisman et al. (2007); Munoz et al. (2009). Clearly, there are market participants with specific demand or supply profiles only during a few hours of the day. Such profiles cannot be resembled by standard baseload and peakload contracts such that for valuating these contracts market makers need to assess the expectations and risks for those specific hours and cannot rely on daily average prices only. Further applications of examining the hourly structure of electricity prices include also issues like the management of price spike risk or derivative pricing in electricity markets. Thus, several studies have focused in particular on the structure of hourly electricity prices and tried to investigate the factors that determine this structure.

Ramsay and Wang (1997) and Szkuta et al. (1999) apply neural networks to model the dynamics of intraday prices. Borenstein et al. (2002) and Saravia (2003) investigate the spread between day-ahead and real-time hourly prices in U.S. power markets. They find a significant spread between the two series that can be attributed to speculation activity and market power of participants. Longstaff and Wang (2004) study the day-ahead hourly risk premium, calculated as the difference between the day-ahead price and the expected real time price. Their findings suggest that premiums are significantly affected by demand and sales figures and price variation. Finally, Huisman et al. (2007) propose a panel framework for modeling the characteristics of hourly electricity prices in day-ahead markets. They suggest that hourly prices do not follow a time series process but should rather be considered as a panel of 24 cross-sectional hours that vary from day to day. They find different hourly specific mean-reversion rates around varying hourly mean price levels. In particular they suggest that peak hours exhibit significant less mean-reversion.

Among the mentioned studies, some model each hour separately or assume a specific correlation pattern between the hours while others treat the hourly prices as a time series. However, it should be pointed out that generally hourly prices cannot be considered as a pure time series process, since time-series models assume that the information set is updated by moving from one observation to the next in time. Given the specific structure of day-ahead markets in the electricity business, this is not a valid assumption in our opinion.

The literature on the application of factor models to electricity spot markets in general and the modeling of hourly prices in particular is very limited. As pointed out by Munoz et al. (2009), despite the fact that dynamic and static factor models have been extensively used in many different financial applications, they have not been exploited yet for the modeling and forecasting of short term electricity market prices. Exceptions include the studies by Wolak (1997); Vehviläinen and Pyykkönen (2005); Munoz et al. (2009); Alonso et al.

(2008) that will be briefly reviewed in the following.

In a pioneer study, Wolak (1997) examines the hourly price formation of day-ahead electricity markets of various countries, including England, Scandinavia, Australia and New Zealand. In his model a Principal Component Analysis (PCA) is applied to the error covariance matrix obtained from a VAR hourly (or half hourly) price system. He finds that generally greater price volatility in systems dominated by fossil fuels relative to those dominated by hydroelectric power. Therefore, for markets like e.g. England and Wales that are dominated by fossil fuels, prices are very difficult to forecast and a high number of 24 factors is required to explain 90% of the variation. On the other hand, for Scandinavian data, only three factors explain approximately the same amount of data for Nordpool prices. Vehviläinen and Pyykkönen (2005) provide a stochastic factor based approach where the fundamentals affecting the spot price are modeled independently and then a market equilibrium model is used to combine them to determine spot prices. The model is rather designed for mid-term modeling and forecasting of electricity spot prices. Alonso et al. (2008) investigate factor models for the difficult task of long term forecasting of electricity prices in the Spanish market. Hereby, they apply a seasonal extension of non-stationary dynamic factor analysis in order to take into account the seasonality of electricity prices that is due to strong dependence of prices and demand on weather as well as economic and social activities. Overall, while their approach offers many features that are relevant for electricity markets, the focus is rather on long-term forecasting of prices what makes the application quite different from ours. Probably, the most similar approach to ours is Munoz et al. (2009) where factor models are applied in order to forecast hourly electricity spot prices with a short-term horizon in the Iberian electricity market. However, their approach applies standard factor model techniques to the data that unlike the Dynamic Semiparametric Factor Models (DSFM) are not able to include additional explanatory variables.

Our paper contributes to the literature in two dimensions. On the one hand, we extend the rare number of studies on application of factor models to electricity markets in general and the modeling and forecasting of hourly electricity prices in particular. Hereby, we consider data from one of the largest power exchanges in the world, the European Electricity Exchange (EEX) in Leipzig where a substantial fraction of traded electricity is produced using fossil fuels. We further provide a pioneer study on the application of DSFM to hourly electricity spot prices. Given the structure of the model, it is well-suited to describe the complex dynamics of such a market with 24 observations per day. The remainder of the article is organized as follows. Section 2 gives a review on DSFM models and illustrates their application to modeling hourly electricity spot prices. Section 3 presents the data and empirical results of our study and Section 4 concludes.

2 A dynamic semiparametric factor model (DSFM) for hourly electricity prices

In this section we describe the dynamic semiparametric factor model (DSFM) that can be used to study the dynamics of hourly electricity prices in the functional form. Hereby, we consider the whole term structure of the hourly contracts to better understand the dynamics of the entire system. We apply the DSFM, which offers flexible modeling and allows for dimension reduction. The model was first proposed by Fengler et al. (2007) for studying the dynamics of implied volatility surfaces and further refined by Park et al. (2009) who implemented a series based estimator instead of a kernel smoother for the estimation. So far applications of this model to commodity and in particular electricity markets have been limited. Exceptions include Borak and Weron (2008) who use the model for approximation of electricity forward curve dynamics and Trück et al. (2009) who study the term structure dynamics of the EU-ETS spot and futures market.

Generally, the objective of factor analysis is dimension reduction in order to describe fluctuations over time in a set of usually high-dimensional variables through those experienced by a small set of factors. Hereby, observed variables are assumed to be linear combinations of the unobserved factors, with the factors being characterized up to scale and rotation transformations. In an orthogonal L-factor model an observable J-dimensional random vector $Y_t = (Y_{t,1}, \dots, Y_{t,J})$ can be represented as

$$Y_{t,j} = m_{0,j} + Z_{t,1}m_{1,j} + \dots + Z_{t,L}m_{L,j} + \varepsilon_{t,j} , \quad (1)$$

where $Z_{t,l}$ are common factors, $\varepsilon_{t,j}$ are errors or specific factors and the coefficients $m_{t,j}$ are factor loadings. Normally, the index $t = 1, \dots, T$ represents time evolution of the observed vectors of variables and Y_t can be considered as a multi-dimensional time series. The advantage of applying factor analysis is that if a sufficiently high fraction of the variation of Y_t can be explained by the L factors, the study of high-dimensional Y_t can be simplified to the modeling of $Z_t = (Z_{t,1}, \dots, Z_{t,L})$, which is a more feasible task, in particular when $L \ll J$.

In comparison to a standard factor model, the DSFM allows also for the incorporation of observable covariates $X_{t,j}$ while the factor loadings m_l are now generalized to functions of $X_{t,j}$, so that the standard factor model is extended to

$$Y_{t,j} = m_0(X_{t,j}) \sum_{l=1}^L Z_{t,l} m_l(X_{t,j}) + \varepsilon_{t,l}. \quad (2)$$

As pointed out by Park et al. (2009), the DSFM can be regarded as a regression model with embedded time evolution. However, the model is different from varying-coefficient models, like in Fan et al. (2003) or Yang et al. (2006), since the series Z_t is actually unobservable. However, some linear models which allow time-varying coefficients, as considered in Hansen et al. (2004) and Brumback and Rice (1998), may be recognized as a special case of the DSFM setting.

For the task of modeling hourly electricity prices, the DSFM structure can be exploited in order to reduce the dimension of 24 observed price series to a smaller number of factors. Hereby, $Y_{t,j}$ is the electricity price for hour j observed on day t and variables $X_{t,j}$ denote the actual hour j on day t . The index $j = 1, \dots, J_t$ represents the particular traded hour; in our study for each day prices for 24 different hours can be observed such that on any day $J_t = 24$. Furthermore, m_l are fixed nonparametric basis functions and $\varepsilon_{t,l}$ describe the random error. The dynamics of hourly electricity prices through time is then explained by the time propagation of the L factors and can be observed through the evolution of the coefficients $Z_{t,l}$. Note that contrary to a parametric approach both m_l and $Z_{t,l}$ have to be estimated from the data. While Fengler et al. (2007) suggest the use of a nonparametric kernel estimator for the estimation, we follow Park et al. (2009) and implement a series based estimator of the form

$$Z_t^T m(X) = \sum_{l=0}^L Z_{t,l} \sum_{k=1}^K a_{l,k} \psi_k(X) = Z_t^T A \psi(X). \quad (3)$$

In this model $\psi(X) = (\psi_1, \dots, \psi_K)^T(X)$ is a vector of known basis functions, $A \in R^{(L+1) \times K}$ denotes a coefficient matrix and K is the number of basis functions. Note that generally the loading functions m_l do not need to take a specific form, however as suggested in applications by Park et al. (2009) and Borak and Weron (2008) they can be linearized e.g. by using B-splines. Then K is the number of knots, $\psi_k(X)$ are the splines and $a_{l,k}$ the appropriate coefficients. The estimation procedure then determines the loading functions m_l and time series $Z_{t,l}$ that minimize the following least squares criterion:

$$\sum_{t=1}^T \sum_{j=1}^{J_t} \left\{ Y_{t,j} - \sum_{l=0}^L Z_{t,l} \sum_{k=1}^K a_{l,k} \psi_k(X) \right\}^2. \quad (4)$$

The estimation procedure is iterative. First the model size L is determined in

advance. Next, starting from an initial white noise sequences $\hat{Z}_{t,l}^{(0)}$, $l = 1, \dots, L$, the estimates $\hat{m}_l^{(1)}$ are obtained. Then the updates $\hat{Z}_{t,l}^{(1)}$ are calculated using $\hat{m}_l^{(1)}$. The iterative steps are repeated consequently until a convergence criterion is met. Park et al. (2009) point out that the solution to this optimization problem is not unique. The signs of $Z_{t,l}$ and m_l cannot be identified such that certain linear transformations like e.g. rotation yield the same model for different functions m_l . A possible choice for the identification procedure is e.g. to choose the \hat{m}_l to be orthogonal and then order them with respect to the variation of the series $\sum_{t=1}^T Z_{t,l}^2$ such that \hat{m}_l^{new} and $\hat{Z}_{t,l}^{new}$ contain as much information on the variation as possible and explain the largest movements of Y_t . This ordering can be considered to be similar to ordering the factors in a PCA. For further issues on the estimation procedure and convergence of the algorithm we refer to Park et al. (2009) where also the question of statistical inference based on the estimated time series $\hat{Z}_{t,l}$ instead of the 'true' unobserved time series $Z_{t,l}$ is discussed.

For the choice of L we apply the following procedure. First, for different values of L , we calculate the proportion of the variation explained by the model compared to the simple invariate estimate given by the overall mean:

$$1 - RV(L) = 1 - \frac{\sum_t \sum_j^{J_t} \{Y_{t,j} - \sum_{l=0}^L \hat{Z}_{t,l} \hat{m}_l(X_{t,j})\}^2}{\sum_t \sum_j^{J_t} (Y_{t,j} - \bar{Y})^2}.$$

Since the model is not nested, the whole estimation procedure has to be repeated for different L 's until the explanatory power of the model is considered to be sufficient.

3 Empirical Results

For our empirical analysis we use hourly electricity spot prices from the European Electricity Exchange (EEX) in Leipzig. As mentioned above, at the EEX, the spot market is a day-ahead market and the spot is an hourly contract with physical delivery on the next day. The 24 hourly spot prices are determined in a daily auction. Products range from predetermined hourly blocks for each of the 24 hours of a day to special contracts for base load, peak load and weekend contracts (see EEX, 2004, for more details). We consider hourly electricity spot prices for the period July 1, 2005 - June 30, 2008. Figure 1 provides a plot of the data for the whole period. It is easy to see that the data exhibits the typical features of electricity prices and contains several periods of extreme volatility, price spikes and shows a mean-reverting behavior.

A more detailed look at the hourly prices is provided in Figure 2 and 3 where

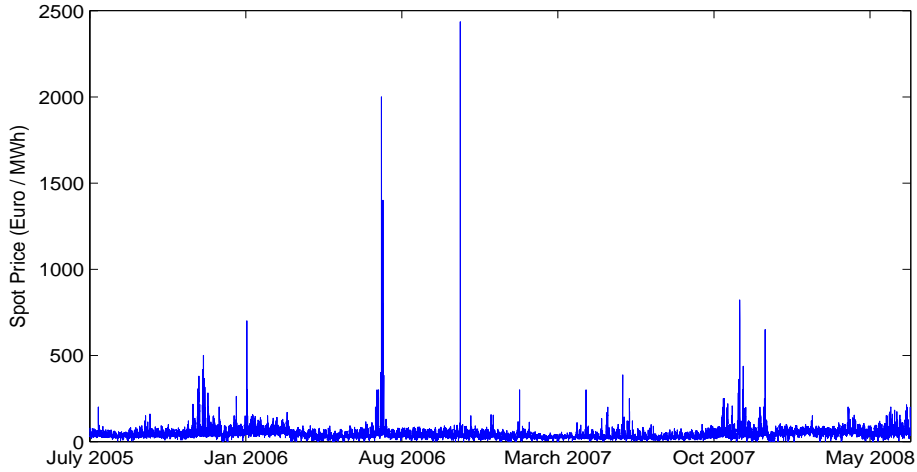


Fig. 1. Hourly spot price for period July 1, 2005 - June 30, 2008.

the typical seasonal pattern during the day and week becomes obvious. Note that for the week November 28 to December 4, 2005 there was a number of consecutive spikes always around 5 pm in the late afternoon.

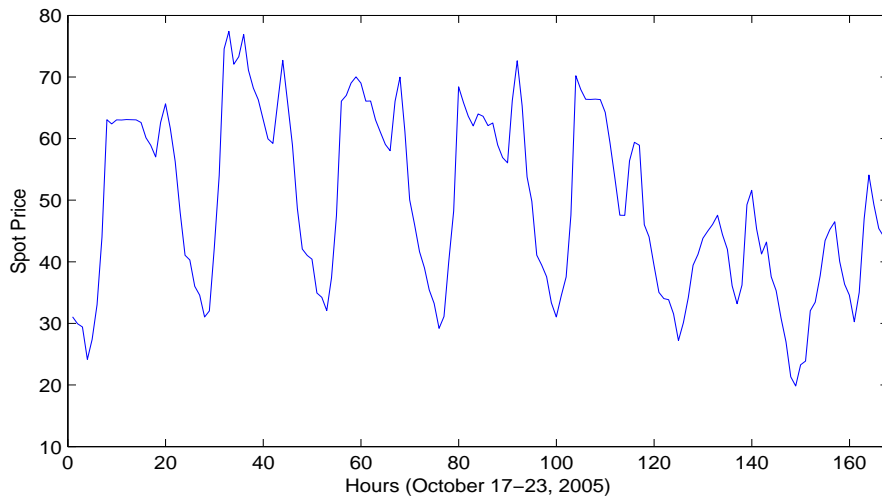


Fig. 2. Hourly spot price for period October 17, 2005 - October 23, 2005.

3.1 Results for the whole period

In a first step, we calibrate the DSFM to the dataset comprising the whole period. As pointed out in the previous section, the model is not nested and the whole estimation procedure needs to be repeated for different L 's. We start with $L = 1$ and then increase the order of the model to $L = 2, 3, 4$. The higher the number of factors, the higher is the general fit, potentially

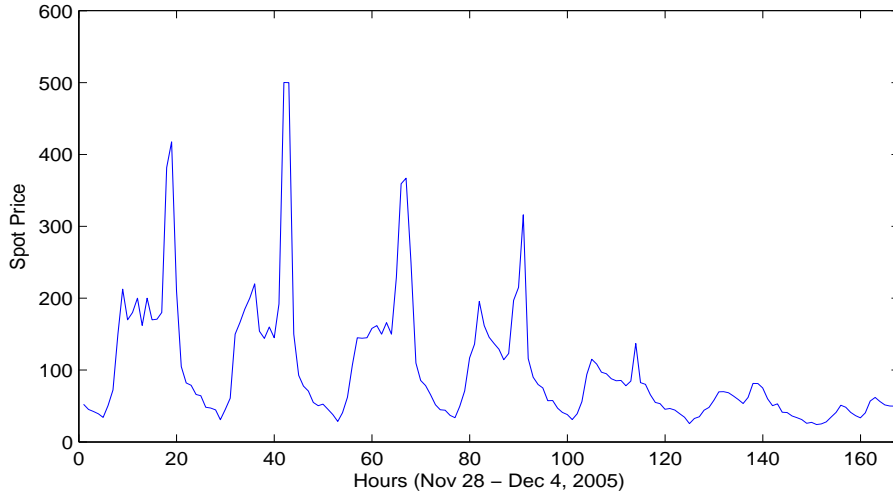


Fig. 3. Hourly spot price for period November 28, 2005 - December 4, 2005.

No. Factors	$1 - RV(L)$
L=1	0.5722
L=2	0.6549
L=3	0.6800
L=4	0.6859

Table 1

Explained variance for DSFM models with $L = 1, 2, 3$ and 4 dynamic factors for whole period July 2005 - June 2008.

at the cost of robustness and parsimony of the model. The results for the different models with respect to criterion (2) are presented in Table 1. In order to decide about the number of factors, usually one proceeds like in the principal components analysis and limits the model to the number of factors which explain a sufficiently high percentage of the variance. We find that a model with only one factor already achieves an explanatory power of 0.5722 according to criterion (2) already, while the inclusion of the second and third factor increase the power to 0.6449 and 0.6800, respectively. The inclusion of the fourth function only slightly improves the explanatory power of the fit and, therefore, from now on we only consider results for models with either two or three basis functions.

The grid used for the variable X consists of $j = 24$ hours, such that $X_j = 0, 1, 2, 3, 4, \dots, 23$ simply represents the hour during the day. For the basis functions $\psi_k(X)$, 6 cubic B-splines evaluated on equidistant knots are used. We find that the placement of the knots and the selection of K does not sig-

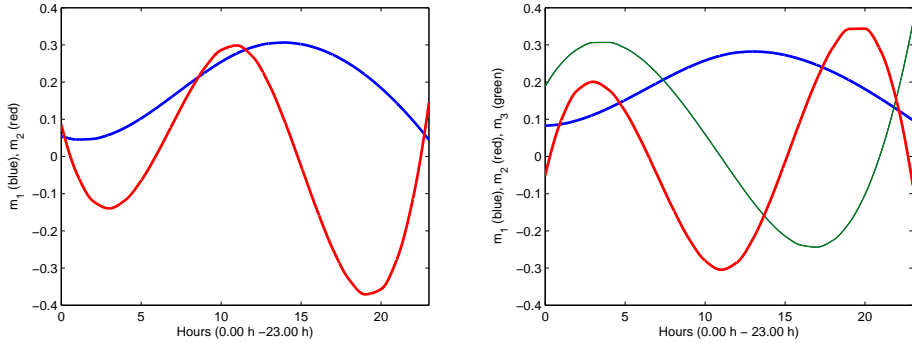


Fig. 4. Estimated functions $\hat{m}_1(X_j)$ (blue) and $\hat{m}_2(X_j)$ (red) for DSFM of order $L = 2$ (*left panel*) and $\hat{m}_1(X_j)$ (blue), $\hat{m}_2(X_j)$ (red), $\hat{m}_3(X_j)$ (green) for $L = 3$ (*right panel*).

nificantly change the results with respect to the basis functions m_l and the coefficients $Z_{t,l}$. This is in line with other studies on the applications of DSFM models, see e.g. Park et al. (2009); Borak and Weron (2008).

Calibration results for the estimated basis $L = 2$ and $L = 3$ are presented in Figure 4 while the time series for $Z_{t,l}$ are plotted in Figure 5. The first function mirrors the daily seasonality of hourly electricity prices. The value increases for the peak hours during the day and is much lower for the off-peak hour in the late evening and during the night. As it can be seen from the left panel of Figure 5, the first factor reflects the behavior of the daily electricity price, including some price spikes and seasonal behavior throughout the year. The second and third basis function exhibit periodic behavior. For both models, for the second basis function the period is approximately 16 hours, while for $L = 3$ the period of the third basis function is approximately 24 hours. As indicated in the right hand panel of Figure 5, the second and third factor show less seasonal behavior than the first one. Further, the second and third factor also show spiky behavior, however, for $Z_{t,2}$ we find that some of the spikes have opposite signs to those of the other series $Z_{t,1}$ and $Z_{t,3}$.

Investigating the in-sample fit of the DSFM model to the data, we find that during times where no price spikes could be observed, the fit to actually observed prices is remarkably good. As illustrated in the left hand panel of Figure 6, the model captures the overall daily price level and the intraday pattern quite well. On the other hand, during periods of spikes that only last for a single hour but on consecutive days like during the period from Nov 28 - Dec 4, 2005, the model is not able to capture the extreme price behavior. Here, a set of time-series models model for each of the individual hours that is able to also capture price spikes might actually provide a better fit to the data. However, considering the fact that the EEX market has a large share of electricity generated by fossil fuels, the explanatory power of our model with three factors is comparably high. In a similar study by Wolak (1997) for fossil fuel

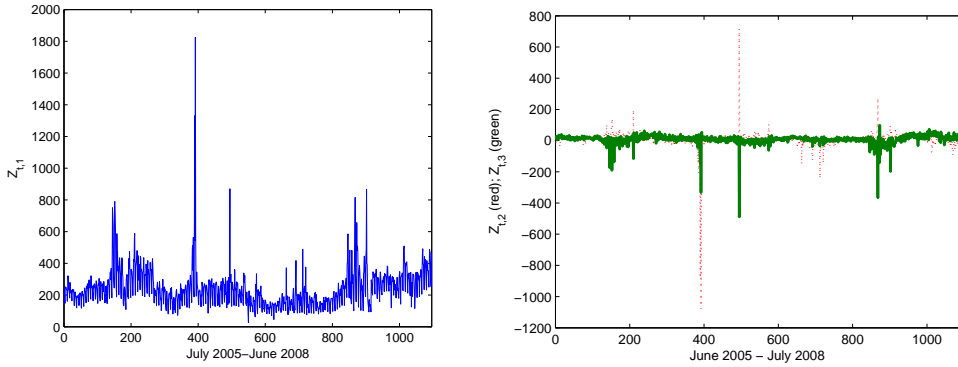


Fig. 5. Estimated time series $\hat{Z}_{t,1}$ (left panel) and $\hat{Z}_{t,2}$ (red), $\hat{Z}_{t,3}$ (green) (right panel) for DSFM of order $L = 3$.

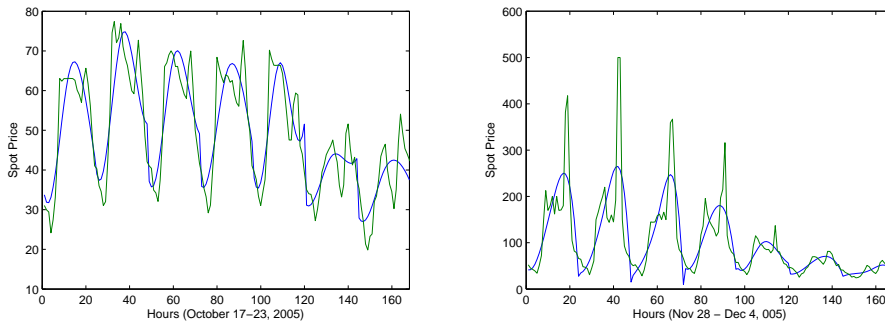


Fig. 6. Hourly spot price (green) and in-sample fit of DSFM (blue) with $L=3$ for Oct 17-23, 2005 (left panel) and Nov 28 - Dec 4, 2005 (right panel).

dominated electricity markets in England and Wales, a model with 10 factors was required to explain approximately 70% of the variation in electricity prices.

3.2 Results for different subperiods

In a second step we divide the whole period into three different subperiods from July 2005 - June 2006, July 2006 - June 2007 and July 2007 - June 2008. As indicated in Table 2, we find that, depending on the period under investigation, we get significant differences with respect to the explanatory power of the model. In particular during the less spiky periods from July 2005 - June 2006 and July 2007 - June 2008, we obtain an explanatory power of approximately 80% for a model with three factors. This can be considered as a very good fit for a model with such a small number of factors, in particular for a market dominated by fossil fuels. For example, Wolak (1997) states that a model with three factors for England and Wales markets is only able to explain 41% of the total variation while up to 15 factors were required to

No. Factors	Period	$1 - RV(L)$
L=3	2005-2008	0.6800
L=3	2005-2006	0.7907
L=3	2006-2007	0.5815
L=3	2007-2008	0.7831

Table 2

Explained variance for the models with $L = 3$ dynamic factors for whole period July 2005 - June 2008 and subperiods July 2005 - June 2006, July 2006 - June 2007, July 2007 - June 2008.

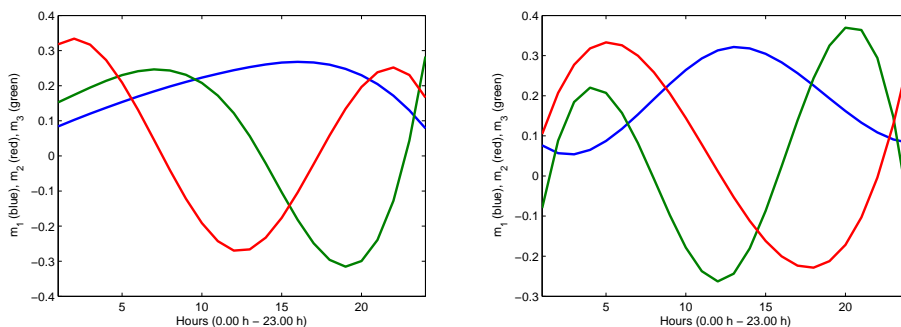


Fig. 7. Estimated functions $\hat{m}_l(X_j)$ for DSFM of order $L = 3$ for July 2005 - June 2006 (*left panel*) and July 2006 - June 2007 (*right panel*).

achieve an explanatory power of more than 80% for the model.

We further investigate the structure of the basis functions for the different subperiods. Unlike in other studies, like e.g. Borak and Weron (2008) where the structure of the basis functions was rather stable throughout the sample, we find that there are obvious changes in these functions depending on the time period under investigation. Figure 7 illustrates that not only the second and third basis function are different, but also that there are significant changes in the first function between the time period from July 2005 - June 2006 and July 2006 - June 2007. Similar results were observed for the last of the three subperiods. Overall, these results indicate that if the model is applied for forecasting, a rather shorter calibration period should probably be used, since not only the series Z_t show dynamic behavior but also the basis functions exhibit some variability through time.

3.3 Deseasonalized Data

In a next step we investigate the performance of the DSFM approach to deseasonalized electricity prices. In order to do this we follow an approach generally suggested in the literature and split system prices P_t into a deterministic part $f(t)$ that comprises all kinds of seasonal behavior and a purely stochastic component S_t :

$$P_t = f(t) + S_t. \quad (5)$$

For the deterministic seasonal component f_t sinusoidal (Pilipovic, 1997; Weron et al., 2004a), constant piece-wise functions (Pindyck, 1999; Huisman and De Jong, 2003; Knittel and Roberts, 2005; Haldrup and Nielsen, 2006) or a combination of both methods (Lucia and Schwartz, 2002; De Jong, 2005; Kosater and Mosler, 2006) have been suggested in the literature. Alternative approaches by Simonsen (2003) and Weron et al. (2004b) approximate the underlying periodical structure using a wavelet decomposition. Here, we pursue a hybrid approach of constant piece-wise functions and a sinusoidal cycle to capture long-term seasonal effects, similar to Bierbrauer et al. (2007). Hence, we specify dummy variables for daily and monthly effects, a trend component and an additional sinusoidal component with one-year cycle:

$$f(t) = \alpha + \beta \cdot t + d \cdot D_{day} + m \cdot D_{mon} + \gamma \cdot \sin\left((t + \tau) \frac{2\pi}{365}\right), \quad (6)$$

where α, β, γ and τ are all constant parameters. Note that initially for each day ($day = 1, \dots, 7$) and month ($mon = 1, \dots, 12$) a dummy variable D_{day}, D_{month} was used. Hereby d and m denote the corresponding parameter vector. In a first step, the function $f(t)$ is calibrated via numerical optimization using non-linear least squares regression. The initial results roughly validated general assumptions about intra-week and intra-year price patterns for power markets, see e.g. Pindyck (1999): Prices are higher at the beginning of the week reaching their peak on Tuesday and then, from the middle of the week, they constantly decline to reach their lowest level over the weekend. A similar cycle can be observed during the year: Prices usually tend to be higher during the cold winter months and are somewhat lower during the rest of the year.

After the deseasonalization, we fit the DSFM models with different number of factors to the data. Results for the explained variance of the models are displayed in Table 3. We find that a model with one factor now achieves an explanatory power of 0.6063 according to criterion (2), while the inclusion of the

No. Factors	Original	Deseasonalized
L=1	0.5722	0.6063
L=2	0.6549	0.6719
L=3	0.6800	0.7037
L=4	0.6859	0.7085

Table 3

Explained variance for DSFM models with $L = 1, 2, 3$ and 4 dynamic factors for original and deseasonalized data.

second and third factor increase the power to 0.6719 and 0.7037, respectively. Again, the inclusion of the fourth function only slightly improves the explanatory power of the model. Overall, the model fit is only slightly improved in comparison to using deseasonalized data. This might be due to using a measure based on a squared distance where the main weight is put to errors that are comparably large and most likely due to spikes in the price series. Another explanation could be that the flexible DSFM approach is able to model a substantial amount of the seasonal variation such that using deseasonalized data does not lead to a significantly improved in-sample fit. In particular the series $\hat{Z}_{t,1}$ in Figure 5 exhibits strong seasonal behavior and seems to capture the overall variation of daily price levels. Further investigation of this issue should be conducted in future research.

3.4 Forecasting Study

In the following we will apply the DSFM approach in a forecasting study considering data for the period July 1, 2007 to June 30, 2008. Due to the variability of the estimated basis functions for different subperiods, the DSFM model is calibrated on a shorter period from July 1, 2007 to May 31, 2008 while the out-of-sample period is chosen from June 1, 2008 to June 30, 2008. We further decided to consider deseasonalized data in order to improve the forecasting performance of the model and a DSFM model with $L = 3$ factors. The explanatory power of the model according to criterion (2) is 0.8133 for the in-sample period.

In a next step, the series $\hat{Z}_{t,i}$, ($i = 1, 2, 3$) are modeled using a time-series approach. Figure 8 provides a plot of the considered deseasonalized data and the estimated time series $\hat{Z}_{t,2}$ for the calibration period. As it can be seen, the series $\hat{Z}_{t,2}$ shows a behavior similar to that of electricity spot prices with a number of short-lasting spikes and different regimes of volatility. Therefore, it

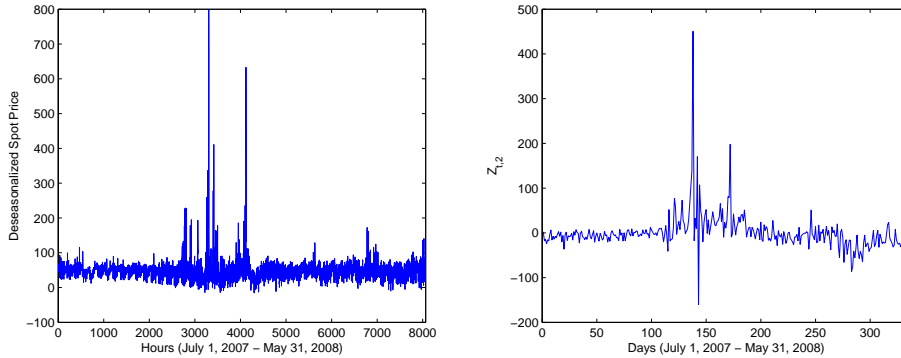


Fig. 8. Deseasonalized hourly prices (*left panel*) and estimated time series $\hat{Z}_{t,2}$ for period July 1, 2007 - May 31, 2008 for DSFM of order $L = 3$ (*right panel*).

seems to be appropriate to use a time-series model that is capable of modeling similar features to those exhibited by electricity prices. Due to the recent popularity of the approach in power markets, see e.g. Huisman and De Jong (2003); Weron et al. (2004a); Haldrup and Nielsen (2006); Mount et al. (2006); Bierbrauer et al. (2007); Huisman (2008), we decided to calibrate a regime-switching model to each of the series $\hat{Z}_{t,i}$. To determine one-day ahead forecasts for hourly electricity spot prices we can then use the estimated basis functions of the DSFM in combination with forecasts for the series $\hat{Z}_{t,i}$.

Regime switching models have been introduced in various different contexts by Quandt (1958); Goldfeld and Quandt (1973); Hamilton (1989, 1994). The underlying idea is to model the observed stochastic behavior of a specific time series by two separate phases or regimes with different underlying processes. The switching mechanism is typically assumed to be governed by a time-homogeneous hidden Markov chain with $k \in \mathbb{N}$ different possible states representing the k different regimes. For the ease of exposition we restrict our description of regime switching models to the discrete time case. Generally, for electricity prices a regime switching models with two independent states has been suggested in the literature, see e.g. Huisman and De Jong (2003); Weron et al. (2004a); Bierbrauer et al. (2007).

The two independent regime model distinguishes between a base regime ($R_t = 1$) and a spike regime ($R_t = 2$), where $(R_t)_{t \in I}$ represents a time-homogeneous hidden Markov chain. The observable stochastic process $(Y_t)_{t \in \mathbb{N}}$ – in our case the estimated series $\hat{Z}_{t,l}$ – are then represented in the form

$$Y_t = Y_{t,R_t}, \quad t \in \mathbb{N}, \quad (7)$$

where the processes $(Y_{t,1})_{t \in \mathbb{N}}$ and $(Y_{t,2})_{t \in \mathbb{N}}$ are assumed to be independent from each other and independent from $(R_t)_{t \in \mathbb{N}}$. Y_t equals $Y_{t,i}$ given that the current regime at time t equals i , i.e. given $R_t = i$. The transition matrix Π of the hidden Markov chain R contains the conditional probabilities p_{ij} of switching

from regime i at time t to regime j at time $t + 1$:

$$\Pi = (p_{ij})_{i,j=1,2} = (P(R_{t+1} = j \mid R_t = i))_{i,j=1,2} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix} \quad (8)$$

The probability of being in state j at time $t + m$ starting from state i at time t can be expressed as:

$$(P(R_{t+m} = j \mid R_t = i))_{i,j=1,2} = (\Pi')^m \cdot e_i, \quad (9)$$

where Π' denotes the transpose of Π and e_i denotes the i th column of the 2×2 identity matrix.

The remaining task consists of specifying the two stochastic processes $Y_{t,1}$ and $Y_{t,2}$. For electricity spot prices the literature often suggests a mean-reverting process for the base regime ($R_t = 1$) and a process with a higher variance or price level for the spike regime ($R_t = 2$). For our application, we decided to use a model similar to Huisman and De Jong (2003) and Bierbrauer et al. (2007) with a mean-reversion process for $R_t = 1$ and a Gaussian distribution for $R_t = 2$. This model is then calibrated to the stochastic process for each series $\hat{Z}_{t,i}$ with $i = 1, 2, 3$. Note that alternative model specifications for the individual series $\hat{Z}_{t,i}$ might be more suitable and could be tested in further applications.

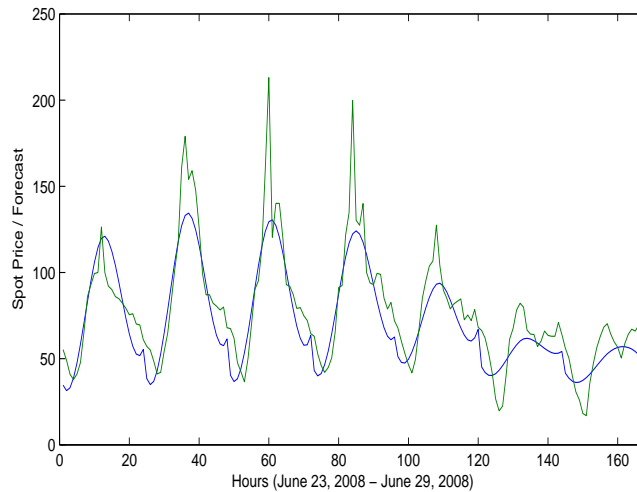


Fig. 9. Actually observed and forecasted hourly electricity prices for period June 23, 2008 - June 29, 2008 for DSFM of order $L = 3$.

In the next step, using the estimated regime-switching models and the estimated probabilities for being in either of the regimes, one-day ahead forecasts $\hat{Z}_{t+1,i}$ for each series are determined. Then the forecasts for $\hat{Z}_{t+1,i}$ and the estimated basis functions \hat{m}_i can be used to derive forecasts for each hour

of the next day. Updating the probabilities for being in either of the regimes using the most recent information of $\hat{Z}_{t,i}$, one-day ahead forecasts for each hour of the out-of-sample period can be determined. Figure 9 provides a plot of actually observed and forecasted electricity spot prices for the period June 23, 2008 - June 29, 2008. We find that during the out-of-sample period the explanatory power of the model is not reduced. In fact, using criterion (2), the proportion of variation explained by the model compared to the simple invariate estimate is 0.8261 and therefore slightly higher than during the calibration period. Overall, our results suggest that DSFM models seem to be a promising approach for dimension reduction and modeling the dynamics of hourly electricity prices. In combination with an appropriate time-series model for the estimated series $\hat{Z}_{t,i}$ the models may also be used for forecasting hourly electricity prices in further applications.

4 Conclusion

In this paper we have analyzed the dynamics of hourly electricity prices in day-ahead markets using dynamic semiparametric factor models (DSFM). In order to verify that the schedule is feasible and lies within transmission constraints, in these markets operators usually require advance notice of hourly load profiles such that there is no continuous trading. Instead bids and offers for delivery of electricity for all hours of the next day are submitted by agents before a specified market closing time. Therefore, the direct application of a time-series approach is not appropriate since such models assume that the information set is updated by moving from one observation to the next in time. Instead we suggest to consider the data rather as a time series of a (24×1) vector or a panel with 24 cross-sectional hours in the spirit of e.g. Wolak (1997) or Huisman et al. (2007).

We have provided a pioneer study on the application of DSFM to hourly electricity prices. The model utilizes a linear combination of nonparametric loading functions being linearized with B-splines and parametrized common factors for the representation of the high-dimensional data set. The model is then calibrated within a least squares iterative scheme. We find that a model with a small number of factors ($L = 3$) explains up to 80% of the variation in hourly prices of the European Energy Exchange (EEX) in Leipzig. In a similar study by Wolak (1997) for fossil fuel dominated electricity markets in England and Wales, a model with 15 factors was required to explain approximately 80% of the variation in electricity prices. However, we find that the explanatory power significantly decreases for periods with a higher number of price spikes. Interestingly, the model performance is only slightly improved when deseasonalized data is considered. In an out-of-sample forecasting study, using regime-switching models for the estimated time-series $\hat{Z}_{t,i}$, we obtain an

explanatory power of the same magnitude as the in-sample fit of the model.

Overall, DSFM seem to be an approach that is very well suited to represent the structure of hourly electricity prices. This study provided a first application of this class of models to spot electricity prices. The full potential of the model might be explored in future studies where additional explanatory variables like capacity limits, load profiles, trading volume, weather variables etc. are also included into the modeling strategy. The approach may also be applied using alternative time-series models for the derived series $\hat{Z}_{t,i}$. Further the approach should be applied to other electricity markets, e.g. markets with less or even more price spikes in order to compare the performance of the models dependent on the market structure.

References

- Alonso, A., García-Martos, C., Rodríguez, J., Sanchez, M., 2008. Seasonal dynamic factor analysis and bootstrap inference : application to electricity market forecasting. Working Paper 08-14, University Carlos III, Department of Economics.
- Barlow, M., 2002. A diffusion model for electricity prices. *Mathematical Finance* 12, 287–298.
- Bhanot, K., 2000. Behavior of power prices: Implications for the valuation and hedging of financial contracts. *Journal of Risk* 2, 43–62.
- Bierbrauer, M., Menn, C., Rachev, S., Trück, S., 2007. Spot and derivative pricing in the eex power market. *Journal of Banking and Finance* forthcoming.
- Borak, S., Weron, R., 2008. A semiparametric factor model for electricity forward curve dynamics. *Journal of Energy Markets* 1 (3).
- Borenstein, S., Bushnell, J., Knittel, C., Wolfram, C., 2002. Trading inefficiencies in california’s electricity market. Working Paper PWP-086, University of California Energy Institute, Berkeley.
- Clewlow, L., Strickland, C., 2000. *Energy Derivatives – Pricing and Risk Management*. Lacima Publications, London.
- De Jong, C., 2005. The nature of power spikes: A regime-switch approach. ERIM Report Series ERS-2005-052-F&A.
- Deng, S., 1999. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. POWER Working Paper PWP-073.
- EEX, 2004. EEX spot market concept. Tech. Rep. November 2004.
- Ethier, R., Mount, T., December 1998. Estimating the volatility of spot prices in restructured electricity markets and the implications for option values. Working paper, Cornell University.

- Fan, J., Yao, Q., Cai, Z., 2003. Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society B* 65, 57–80.
- Fengler, M. R., Härdle, W., Mammen, E., 2007. A dynamic semiparametric factor model for implied volatility surface dynamics. *Journal of Financial Econometrics* 5(2), 189–218.
- Garcia, R., Contreras, J., van Akkeren, M., Garcia, J., 2005. A garch forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems* 20(2), 867–874.
- Geman, H., Roncoroni, A., 2006. Understanding the fine structure of electricity prices. *Journal of Business* 79(3), 1225–1262.
- Goldfeld, S. M., Quandt, R. M., 1973. A Markov model for switching regressions. *Journal of Econometrics* 1, 3–16.
- Hadsell, L., 2006. A tarch examination of the return volatility-volume relationship in electricity futures. *Applied Financial Economics* 16, 893–901.
- Haldrup, N., Nielsen, M., 2006. A regime switching long memory model for electricity prices. *Journal of Econometrics* (349-376).
- Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357–384.
- Hamilton, J. D., 1994. *Time Series Analysis*. Princeton University Press.
- Huisman, R., 2008. The influence of temperature on spike probability in day-ahead power prices. *Energy Economics* 30, 2697–2704.
- Huisman, R., De Jong, C., 2003. Option formulas for mean-reverting power prices with spikes. *Energy Power Risk Management* 7, 12–16.
- Huisman, R., Huurman, C., Mahieu, R., 2007. Hourly electricity prices in day-ahead markets. *Energy Economics* 29, 240–248.
- Johnson, B., Barz, G., 1999. *Selecting Stochastic Processes for Modelling Electricity Prices*. Risk Books, Ch. 1, pp. 3–22.
- Kaminski, V. (Ed.), 1999. *Managing Energy Price Risk*. Risk Books, London.
- Kanamura, T., Ohashi, K., 2008. On transition probabilities of regime switching in electricity prices. *Energy Economics* 30, 1158–1172.
- Kluge, T., 2004. Pricing options in electricity markets. Working paper.
- Knittel, C. R., Roberts, M. R., 2005. An empirical examination of restructured electricity prices. *Energy Economics* 27(5), 791–817.
- Koopman, S., Ooms, M., Carnero, M., 2007. Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association* 102 (477), 16–27.
- Kosater, P., Mosler, K., 2006. Can markov regime-switching models improve power-price forecasts? evidence from German daily power prices. *Applied Energy* forthcoming.
- Longstaff, F., Wang, A., 2004. Electricity forward prices: A high-frequency empirical analysis. *Journal of Finance* 59, 1877–1900.
- Lucia, J. J., Schwartz, E., 2002. Electricity prices and power derivatives: Evidence from the Nordic power exchange. *Review of Derivatives Research* 5, 5–50.
- Merton, R. C., 1976. Option pricing when underlying stock returns are dis-

- continuous. *Journal of Financial Economics* 3, 125–144.
- Misiorek, A., Trück, S., Weron, R., 2006. Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics & Econometrics* 10(3).
- Mount, T., Ning, Y., Cai, X., 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* 28, 62–80.
- Munoz, M., Corchero, C., Heredia, F., 2009. Improving electricity market price scenarios by means of forecasting factor models. Working Paper DR 2009/6, Department of Statistics and Operations Research, Universitat Politècnica de Catalunya.
- Park, B., Mammen, E., Härdle, W., Borak, S., 2009. Time series modelling with semiparametric factor dynamics. *Journal of the American Statistical Association* 104, 284–298.
- Pilipovic, D., 1997. *Energy Risk: Valuing and Managing Energy Derivatives*. McGraw-Hill.
- Pindyck, R., 1999. The long-run evolution of energy prices. *The Energy Journal* 20, 1–27.
- Quandt, R. E., 1958. The estimation of parameters of linear regression system obeying two separate regimes. *Journal of the American Statistical Association* 55, 873–880.
- Ramsay, B., Wang, A., 1997. A neural network based estimator for electricity spot pricing with particular reference to weekends and public holidays. *Neurocomputing* 23, 47–57.
- Saravia, C., 2003. Speculative trading and market performance: the effect of arbitrageurs on efficiency in the new york electricity market. Working Paper. University of California Energy Institute, Berkeley.
- Simonsen, I., 2003. Measuring anti-correlations in the Nordic electricity spot market by wavelets. *Physica A* 322, 597–606.
- Szkuta, B., Sanabria, L., Dillon, T., 1999. Electricity price short-term forecasting using artificial networks. *IEEE Transactions on Power Systems* 4, 851–857.
- Trück, S., Härdle, W., Weron, R., 2009. The EU emission trading system - relationship between spot and futures prices for different trading periods. Working paper, Macquarie University Sydney.
- Vehviläinen, I., Pyykkönen, T., 2005. Stochastic factor model for electricity spot price - the case of the nordic market. *Energy Economics* 27(2), 351–367.
- Weron, R., 2006. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, Chichester.
- Weron, R., Bierbrauer, M., Trück, S., 2004a. Modeling electricity prices: Jump diffusion and regime switching. *Physica A* 336, 39–48.
- Weron, R., Simonsen, I., Wilman, P., 2004b. Modeling highly volatile and seasonal markets: Evidence from the Nord Pool electricity market. In: Takayasu, H. (Ed.), *The Application of Econophysics*. Springer, pp. 182–191, proceedings of the Second Nikkei Econophysics Symposium, Tokyo.

- Wolak, F., 1997. Market design and price behavior in restructured electricity markets: an international comparison. Working Paper. Department of Economics Stanford University.
- Yang, L., Park, B., Xue, L., Härdle, W., 2006. Estimation and testing for varying coefficients in additive models with marginal integration. *Journal of the American Statistical Association* (101), 1212–1227.

SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Volatility Investing with Variance Swaps" by Wolfgang Karl Härdle and Elena Silyakova, January 2010.
- 002 "Partial Linear Quantile Regression and Bootstrap Confidence Bands" by Wolfgang Karl Härdle, Ya'acov Ritov and Song Song, January 2010.
- 003 "Uniform confidence bands for pricing kernels" by Wolfgang Karl Härdle, Yarema Okhrin and Weining Wang, January 2010.
- 004 "Bayesian Inference in a Stochastic Volatility Nelson-Siegel Model" by Nikolaus Hautsch and Fuyu Yang, January 2010.
- 005 "The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility" by Nikolaus Hautsch, Dieter Hess and David Veredas, January 2010.
- 006 "Bayesian Estimation and Model Selection in the Generalised Stochastic Unit Root Model" by Fuyu Yang and Roberto Leon-Gonzalez, January 2010.
- 007 "Two-sided Certification: The market for Rating Agencies" by Erik R. Fasten and Dirk Hofmann, January 2010.
- 008 "Characterising Equilibrium Selection in Global Games with Strategic Complementarities" by Christian Basteck, Tijmen R. Daniels and Frank Heinemann, January 2010.
- 009 "Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory" by Julia Schaumburg, February 2010.
- 010 "On Securitization, Market Completion and Equilibrium Risk Transfer" by Ulrich Horst, Traian A. Pirvu and Gonçalo Dos Reis, February 2010.
- 011 "Illiquidity and Derivative Valuation" by Ulrich Horst and Felix Naujokat, February 2010.
- 012 "Dynamic Systems of Social Interactions" by Ulrich Horst, February 2010.
- 013 "The dynamics of hourly electricity prices" by Wolfgang Karl Härdle and Stefan Trück, February 2010.