Energy Economics 34 (2012) 1589–1616

Contents lists available at SciVerse ScienceDirect

Energy Economics

journal homepage: www elsevier com/locate/eneco

A critical empirical study of three electricity spot price models

Fred Espen Benth, Rüdiger Kiesel, Anna Nazarova *

[a] Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, N-0316 Oslo, Norway
[b] Department of Economics and Business Administration, University of Agder, Servicebox 422, N-4604 Kristiandsand, Norway
[c] Chair for Energy Trading and Finance, University of Duisburg-Essen, Campus Essen, Universitätsstraße 12, 45141 Essen, Germany

R T I C L E I N F O

Article history:
Received 7 January 2010
Received in revised form 22 November 2011
Accepted 26 November 2011
Available online 3 December 2011

Keywords:
Electricity spot price
Mean-reversion
Spikes
Jump-diffusion
Ornstein–Uhlenbeck process
Electricity forwards
Forward risk premium

B S T R C T

We conduct an empirical analysis of three recently proposed and widely used models for electricity spot price process. The rst model, called the jump-diffusion model, was proposed by Cartea and Figueroa (2005), and is a one-factor mean-reversion jump-diffusion model, adjusted to incorporate the most important characteristics of electricity prices. The second model, called the threshold model, was proposed by Roncoroni (2002) and further developed by Geman and Roncoroni (2006), and is an exponential Ornstein–Uhlenbeck process driven by a Brownian motion and a state-dependent compound Poisson process. It is designed to capture both statistical and pathwise properties of electricity spot prices. The third model, called the factor model, was proposed by Benth et al. (2007). It is an additive linear model, where the price dynamics is a superposition of Ornstein–Uhlenbeck processes driven by subordinators to ensure positivity of the prices. It separates the modelling of spikes and base components. We calibrate all three models to German spot price data. Besides employing techniques similar to those used in the original papers we adopt the prediction-based estimating function technique (Sørensen, 2000) and the ltering technique (Meyer-Brandis and Tankov, 2008). We critically compare the properties and the estimation of the three models and discuss several shortcomings and possible improvements. Besides analysing the spot price behaviour, we compute forward prices and risk premia for all three models for various German forward data and identify the key forward price drivers.

1. Introduction

The modelling of the dynamics of electricity spot prices is a delicate issue. The spot prices exhibit various characteristics (see Eydeland and Wolyniec (2002)): seasonality, spikes and mean-reversion. Depending on the market, we can observe daily, weekly, monthly or yearly seasonality. Sudden big changes in price or so-called spikes can be caused for example by unexpected weather change or outage of equipment. The intensity of spikes may also demonstrate both time dependency and randomness. Also, prices are mean-reverting at different speeds. Since the deregulated electricity markets are still developing and fast growing, practitioners as well as academics have suggested several models to capture some or all of these features. Recently, three models have attracted considerable attention: a model proposed by Cartea and Figueroa (2005) (called the *jump-diffusion* model), a model proposed by Roncoroni (2002) and further developed by Geman and Roncoroni (2006) (called the *threshold* model), and a model derived by Benth et al. (2007) (called the *factor* model). We will provide a detailed comparison of the three models in terms of their empirical ability to t spot price data and to price forwards. In order to do so we use data from the German Electricity Exchange, EEX, and investigate the performance of the models.

The jump-diffusion model can be seen as a one-factor mean-reverting jump-diffusion model close to the classical exponential Ornstein–Uhlenbeck process suggested by Schwartz (1997) and later applied to electricity markets by Lucia and Schwartz (2002). Since the two latter models do not incorporate jumps, the model proposed by Cartea and Figueroa (2005) is extended to account for jumps. The model is easy to calibrate and produces a straightforward formula to price forward contracts. Due to its simple and parsimonious structure the jump-diffusion model is quite extensively used among practitioners.

The threshold model can be seen as a one-factor mean-reverting jump-diffusion model close to the model of Cartea and Figueroa (2005) with two novel twists. Firstly, the authors introduce a state-dependent *sign* of the jump component, where high price levels induce negative jumps, whereas in low price regimes the jumps are upwards. This feature prevents several spikes following each other. Secondly, the estimation process makes use of a threshold, which is set iteratively, so that the estimated parameters are calibrated to the empirical kurtosis. In order to price forwards with this model numerically, typically Monte-Carlo, techniques have to be used. Albanese et al. (2006) propose a method based on approximating the generator of the underlying

The authors are thankful to Nick Bingham, Andrea Roncoroni and Michael Sørensen for their helpful comments and remarks.

* Corresponding author.
E-mail addresses: fredb@math.uio.no (F.E. Benth), ruediger.kiesel@uni-due.de (R. Kiesel), anna.nazarova@uni-due.de (A. Nazarova).
URL's: http://folk.uio.no/fredb/ (F.E. Benth), http://www.lef.wiwi.uni-due.de/team/ruediger-kiesel/ (R. Kiesel), http://www.lef.wiwi.uni-due.de/team/anna-nazarova/ (A. Nazarova).

0140-9883/ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.eneco.2011.11.012

process and illustrate the speed and accuracy of the method by pricing European and Bermudan options. A lattice-based method for the discretisation of the threshold model that allows for the pricing of derivatives, including swing options, has been proposed in Geman and Kourouvakalis (2008).

The factor model is an additive (or linear) multi-factor model that separates the base and spike signals. By the base signal we mean the daily fluctuations of the price around the mean level due to small changes in supply and demand in the market; by the spike signal we mean the price jump of extreme size due to sudden imbalances of demand and supply. This structure allows for more flexibility in capturing the high speed of mean-reversion observed for spikes and the more slowly varying base signal. However, the estimation of the parameters in this model is challenging. Applying techniques from Meyer-Brandis and Tankov (2008) together with the prediction-based estimating function technique of Sørensen (2000), we suggest an improved estimation procedure. Due to its speci cation, the model provides a simple and straightforward way to price forwards and options, see Benth et al. (2007) and for the pricing of spark spread and average options consult Benth and Kufakunesu (2009).

Clearly, a reliable and well-understood spot price model is important for risk management and pricing purposes. With our study we will provide a comparison of the data- tting ability and pricing performance of the models. We discuss and clarify estimation procedures for the models. In particular, we observe that in the jump-diffusion and threshold models the mean-reversion is an average of the reversion of spikes and intra-spike behaviour, a property that has earlier been observed in jump-diffusion models of this kind. Also we demonstrate that parameter estimates in the threshold model are very sensitive to changes in the spike sizes. Furthermore, the state-dependent sign change seems to be of little importance for the German market data considered. However, it might be an issue for other data sets. We also nd the ability of the model to capture the spot price risk questionable.

The factor model, on the other hand, achieves a more reasonable modelling of the spike behaviour. However, the base (or intra-spike) dynamics seems to be too regular in the sense that it produces a less volatile base signal compared to the data. This may be attributed to the use of subordinators, which are processes that in the most tractable cases are of nite activity.

The three models have been applied in various contexts for pricing purposes. The performance of the three models leads to the conclusion that they all require careful re nement in both speci cation and estimation in order to fully capture the stylised statistical and pathwise properties of electricity spot and forward price data.

We present our ndings as follows. In the next three sections we introduce the models and discuss their theoretical properties. Further, in Section 5, we provide algorithms and estimate the parameters of the models using German electricity spot price data. In Section 6 we assess the calibration and discuss various possible improvements for modelling noise and jump size distribution. Section 7 is devoted to the computation of forward prices and study of the market risk premium. The nal section offers conclusions and suggestions for further research.

2. The jump diffusion model

Let $\left(\quad ,\mathbb{P}, \quad ,\{ \quad t\}_{t\in[0,T]}\right)$ be a complete ltered probability space with T a xed time horizon. We denote the electricity spot price at time $0\le t\le T$ by $S(t)$, and assume that it takes the form

$$S(t) = e^{\mu(t)}X(t), \tag{1}$$

where $\mu(t)$ is a deterministic function modelling the seasonal trend, or mean variations, of the price evolution, and $X(t)$ is some stochastic process modelling the random fluctuations around this trend. In the three models, $X(t)$ will take very different forms, but the trend function $\mu(t)$ will stay the same.

In electricity markets spot prices may demonstrate various types of seasonality: daily, weekly, monthly, yearly or a combination of them. Cartea and Figueroa (2005) study historical spot data from England and Wales and suggest some intra-week seasonality, since the returns show correlation every 7 days. To explain this fact, recall that electricity is traded 7 days per week and the information contained in Friday prices has an impact on the Saturday, Sunday and Monday morning prices. In addition to this weekly seasonality effect, spot prices may vary by seasons, caused by changing supply and demand. In this case, we have to employ some periodic function to capture such a trend behaviour. For instance, in the Nord Pool market there is no inflow of water in the hydro reservoirs in the winter, at the same time the demand is high because of low temperatures, so the winter prices are higher than in summer. To model seasonalities, we follow Geman and Roncoroni (2006), who analyse three of the major U.S. power markets. The market conjuncture reveals yearly seasonality on a monthly basis and a combination of an af ne function and two sine functions with a 12 and a 6 month period respectively is used to model this seasonality. We nd that in the German EEX market this is also a reasonable choice. We therefore choose the following trend model

$$\mu(t) = \alpha + \beta t + \gamma\cos(\varepsilon + 2\pi t) + \delta\cos(\zeta + 4\pi t) \tag{2}$$

Here, the parameters $\alpha,\beta,\gamma,\delta$, ε and ζ are all constants. The rst term is interpreted as xed cost linked to the power production, while the second term drives the long-run linear trend in the total production cost. The remaining terms give periodicity by adding two maxima per year with possibly different magnitude. The parameters can be estimated by tting the trend function, for example by least squares. We continue by specifying the $X(t)$ process for the jump-diffusion model.

Here the deseasonalised logarithmic spot prices are modelled by

$$d\ln X(t) = \quad \alpha\ln X(t)dt + \sigma(t)dW(t) + \ln J dq(t), \tag{3}$$

where α is the speed of mean-reversion, W is a Brownian motion, $\sigma(t)$ is a time-dependent volatility, J is a proportional random jump size and d q_t is a Poisson process of intensity l with

$$dq_t = \begin{cases} 1 & \text{with probability} \quad ldt \\ 0 & \text{with probability} \quad (1 \quad l)dt \end{cases} \tag{4}$$

A typical assumption on the jump size distribution is $\ln J \sim \mathcal{N}\left(\mu_J, \sigma_J^2\right)$ and $[J] = 1$.

3. The threshold model

In this section we introduce the threshold model and provide a detailed discussion of its various components. Here the deseasonalised logarithmic spot prices are modelled as

$$d\ln X(t) = \quad \theta_1\ln X(t)dt + \sigma dW(t) + h(\ln(X(t \quad)))dJ(t), \tag{5}$$

where W is a Brownian motion, J is a time-inhomogeneous compound Poisson process, i.e.

$$J(t) = \sum_{i=1}^{N(t)} J_i \tag{6}$$

$X(t \quad)$ denotes the left-limit as usual. $N(t)$ is a Poisson process with time-dependent jump intensity and counts the spikes up to time t. $J_1, J_2, \ldots$ model the magnitudes of the spikes and are assumed to be independent and identically distributed random variables. The constants θ_1 and σ are both positive. The function h attains two values,

± 1, indicating the direction of the jump. The Brownian component models the normal random variations of the electricity price around its mean, i.e., the base signal. The discontinuous price spikes are incorporated through the jump term $h(\ln X(t\ \ ))\mathrm{d}J(t)$. The compound Poisson process J has a time-dependent jump intensity to account for seasonal variations in the spike occurrence. Note that as in the jump-diffusion model, the threshold model has only one mean-reversion parameter, namely θ_1.

To review the threshold spike modelling approach (along the arguments given in Roncoroni (2002) and Geman and Roncoroni (2006)), we start with the spike intensity. Since spikes show clustering and periodicity the intensity of $N(t)$, which models the spike intensity, is assumed to be deterministic function

$$\iota(t) = \theta_2 \times s(t) \tag{7}$$

Here, θ_2 is interpreted as the expected number of spikes per time unit at a spike-clustering time. The function $s(t)$ represents the normalised and possible periodic jump intensity shape. One reasonable speci cation of $s(t)$ can be a sine function

$$s(t) = \left[\frac{2}{1+|\sin[\pi(t\ \ \tau)/k]|}\ \ 1\right]^d, \tag{8}$$

where the positive constant k is the multiple of the peaking levels, beginning at time τ. For example, if k is equal to 0.5, then there are 2 peaking times per year, corresponding to two periods with spikes over the year. The exponent d is introduced to adjust the dispersion of jumps around the peaking times. In fact, this parameter is responsible for how short the periods of spike occurrences are. As we shall see in the sequel (see Fig. 4), the intensity shape function $s(t)$ may exhibit convex or concave peaks with a given periodicity, and the choice of this function is motivated by the shape of the power stack function. We remark in passing that in Benth et al. (2007) the same form of intensity function as Eq. (7) was used in an empirical example for Nord Pool electricity spot prices with $k=1$. In the Nordic market, spikes occur in the winter period, thus the periodicity is one.

In their paper Geman and Roncoroni (2006) alternatively suggest a stochastic form of the spike intensity to increase the probability of spikes in case when prices are above some speci ed threshold $\overline{E}(t)$ different from $\mathcal{T}(t)$. The following form is used to capture this effect

$$\iota(t, E(t\ \ )) = \theta_2 \times s(t) \times \left(1 + \max\{0, E(t\ \ )\ \ \overline{E}(t)\}\right), \tag{9}$$

where $E(t)$ denotes the logarithm of the price. The authors suggest to set this threshold $\overline{E}(t)$ smaller than the threshold $\mathcal{T}(t)$ to de ne the interval for prices where the spike activity will be higher. As soon as the price level falls below $\overline{E}(t)$, the stochastic intensity reduces to a deterministic intensity.

The spike sizes are modelled by the jump size distribution of the compound Poisson process, that is, by the J_is. Geman and Roncoroni (2006) propose a truncated exponential distribution for the spike sizes J_i with density

$$p(x\ \theta_3, \psi) = \frac{\theta_3 exp(\ \ \theta_3 x)}{1\ \ exp(\ \ \theta_3 \psi)}, \quad 0 \le x \le \psi \tag{10}$$

The average jump size parameter is θ_3, and the maximal possible jump size is ψ. The latter implies an upper bound for the absolute value of price changes. For an empirical analysis of spot price data series, such an upper bound corresponds to the implicit assumption that there will be no bigger price change in the future than that given by the bound. Consequently, as in Geman and Roncoroni (2006), the model does not generate jumps exceeding historically observed ones. This is a restrictive assumption in the sense that we limit ourselves to include only the observed big changes, which may not be adequate in the future. An alternative to this extreme is to allow potentially unbounded price changes, as proposed in Meyer-Brandis and Tankov (2008), where the authors use a Pareto distribution to model the spike sizes. In this case one can get outliers which will

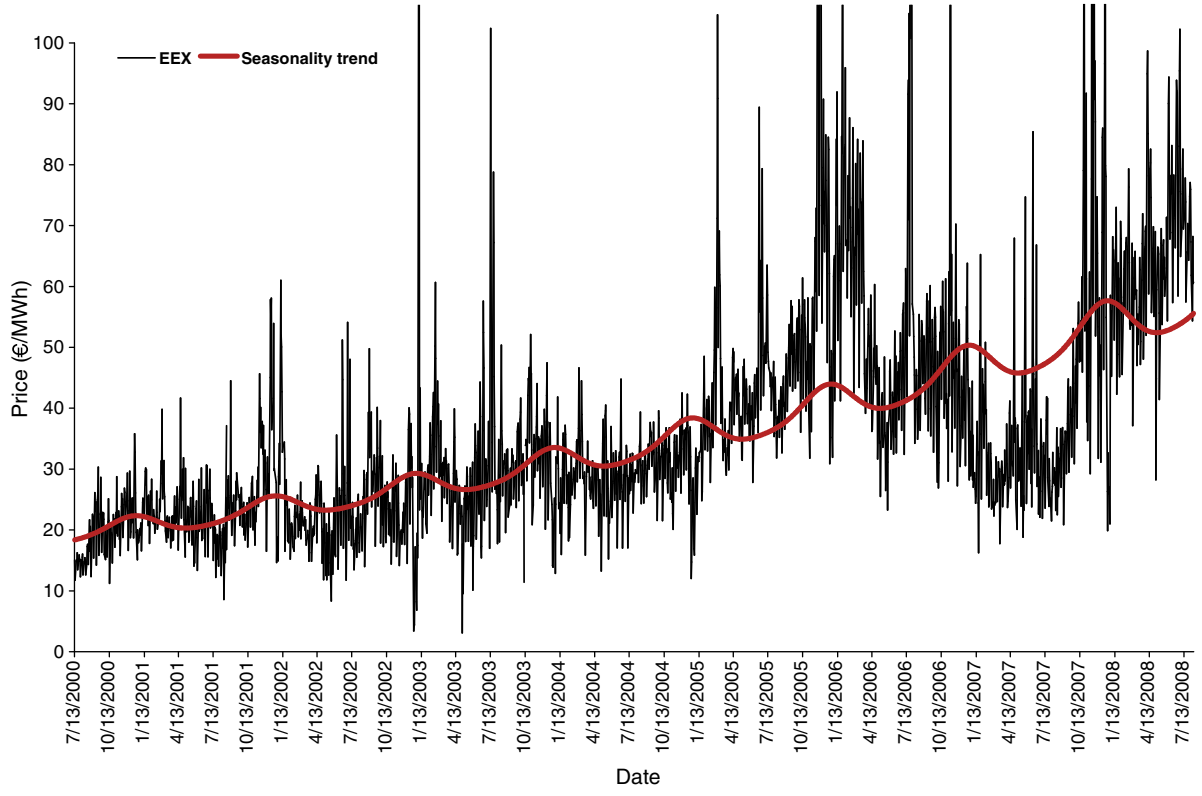


Fig. 1. EEX price path with seasonality trend.

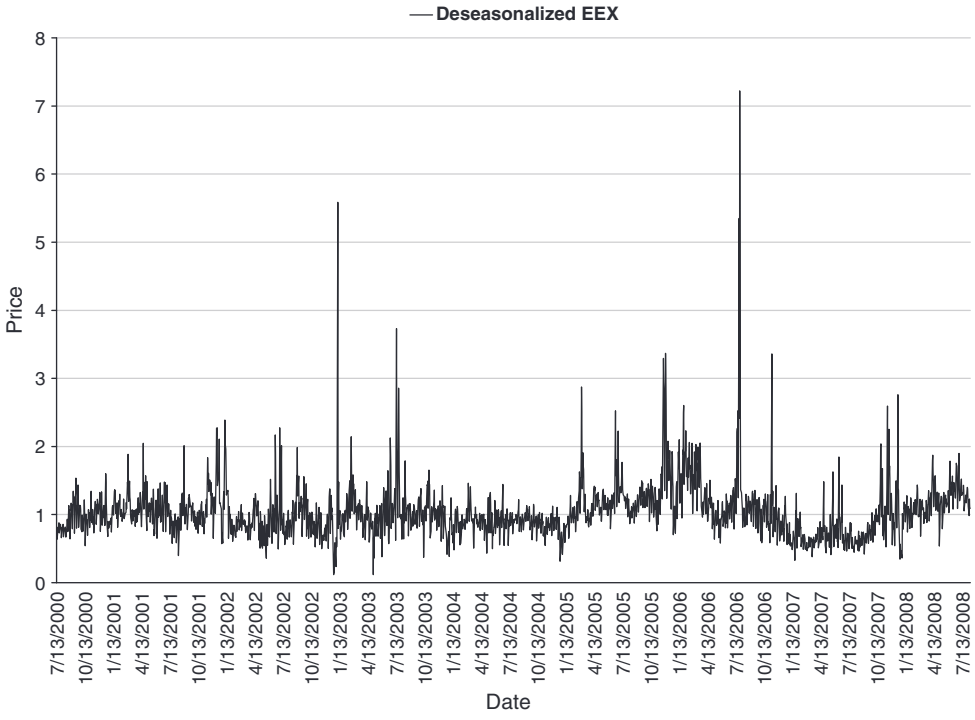


Fig. 2. Deseasonalised EEX price.

result in huge spikes. From the point of view that in most markets there is a maximum price for the spot which the market cannot exceed in the auction, this approach may also be questioned. However, the technical limit is typically rather high. Furthermore, Meyer-Brandis and Tankov (2008) study electricity spot prices on several European markets and defend their use of this extreme-value distribution based on the empirically observed power-law behaviour in the tails of the daily returns, along with excess kurtosis and positive skew.

Obviously, the chosen jump-size distribution strongly influences the empirical properties of the simulated price paths. Looking at the moments of the price paths resulting from these two speci cations, the kurtosis for example will differ dramatically, as we shall see in the next section. It is quite natural that the speci cation with a truncated exponential jump size distribution has a kurtosis reasonably close to the observed one. For a Pareto speci cation, we predict future price changes which maybe far larger than the historical observed ones, and thus the kurtosis will increase. This indicates that comparing empirical

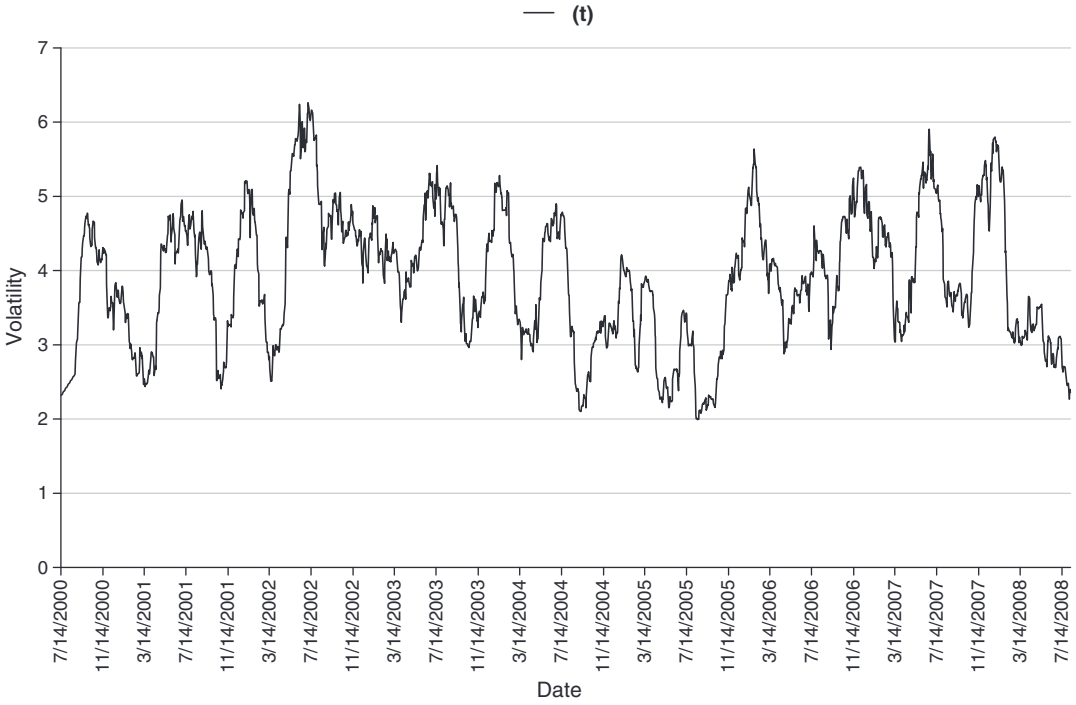


Fig. 3. Time-dependent volatility for the jump-diffusion model.

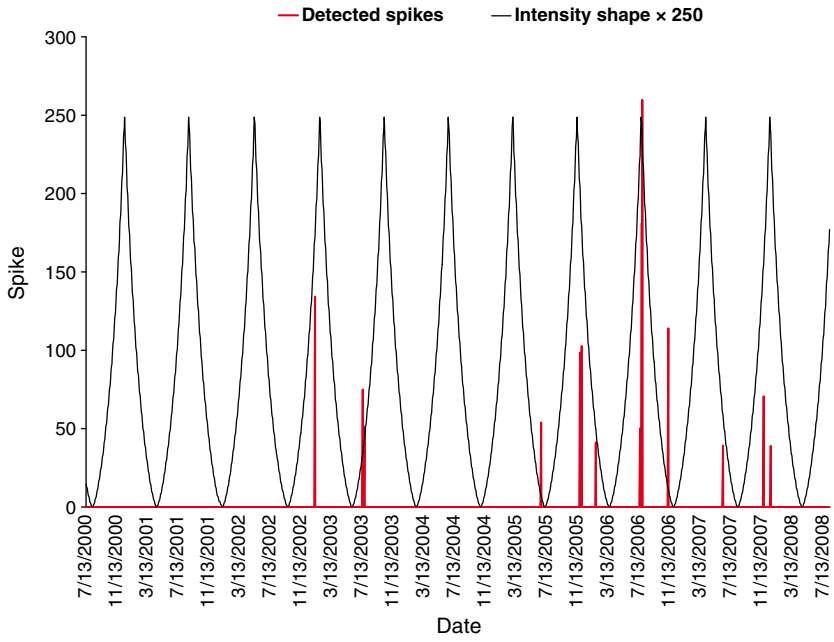


Fig. 4. Calibrated intensity shape function and real values of detected spikes.

moments in order to assess the quality of a model may be misleading. To look beyond these extremes, we shall rely on an empirical analysis of observed jumps and consider alternative jump size distributions in between the two extreme choices of truncated exponential and Pareto.

In the threshold model the direction of spikes is given by an indicator function h taking values $+1$ and -1 depending on the current spot price level. A threshold is introduced to determine the sign of the spike, denoted by $\mathcal{T}$. Thus,

$$h(\ln X(t)) = \begin{cases} +1, & \text{if } \ln X(t) \leq \mathcal{T}(t), \\ -1, & \text{if } \ln X(t) \geq \mathcal{T}(t) \end{cases} \tag{11}$$

Geman and Roncoroni (2006) introduce the h-function together with the smooth mean-reversion θ_1 to bring the prices to a normal range after being at a high level. They af rm that a proper choice of the barrier $\mathcal{T}$ coupled with a high jump intensity can generate a sequence of upward jumps leading to high price levels, after which a discontinuous downward move together with the smooth mean-reversion brings prices down to a normal range".

One may believe that h models the mean-reversion of spikes, in the sense that if we rst have an upward-pointing spike, the next jump will be pushed down due to the sign of h. However, this is not necessarily the case since it may take some time before the next jump actually occurs. Indeed, in the high-spike intensity markets, when we rst have a jump, at the next step we have either one more jump or not. Of course, if the compound Poisson process decides to jump twice, an upward spike will be followed by a downward jump due to the sign of h and the mean-reversion speed together. Such a high concentration of jumps in a period will lead to a rather strange sequence of up- and downward jumps and concentration of noise. To have a mean-reversion of a spike, the threshold model resorts to the θ_1-parameter. On the other hand, θ_1 also accounts for the mean-reversion of the base signal. A spike requires a fast mean-reversion, whereas the base signal is reverting more slowly. The mean-reversion estimate of θ_1 is higher than expected for a base signal, and somewhat slower than required to dampen a spike. Our empirical analysis of EEX data supports this view. If there are no consecutive jumps, the price path will wiggle around the new level, which has been reached after the rst spike, due to the Brownian component and the mean-reversion, unless it is pushed down again. So the parameter h is not responsible for the mean-reversion of the spike process, but θ_1 takes care of this. Therefore, both cases guarantee that consecutive price values will not exceed the threshold level. At this point, we claim that h prevents two consecutive price values above the threshold, i.e. at least one downward price movement lies in between.

4. The factor model

Suppose that $X(t)$ is a stochastic process represented as a weighted sum of n independent non-Gaussian Ornstein–Uhlenbeck processes $Y_i(t)$, that is,

$$X(t) = \sum_{i=1}^{n} w_i Y_i(t), \tag{12}$$

where each $Y_i(t)$ is de ned as

$$dY_i(t) = -\lambda_i Y_i(t)dt + dL_i(t), \quad Y_i(0) = y_i, \quad i = 1, \ldots, n \tag{13}$$

The weight functions w_i and the mean-reversion coef cients λ_i are positive constants. $L_i(t)$ are assumed to be independent *càdlàg* pure-jump additive processes with increasing paths, i.e. so-called time-inhomogeneous *subordinators* (see e.g. Sato, 1999). Corresponding to each $L_i(t)$ we have a time-inhomogeneous Poisson random measure $N_i(\mathrm{d}\,t, \mathrm{d}\,z)$ with a deterministic predictable compensator $\nu_i(\mathrm{d}\,t, \mathrm{d}\,z)$. The compensated Poisson random measure is denoted by

$$\tilde{N}_i(ds, dz) = N_i(ds, dz) - \nu_i(ds, dz)$$

We have the representation

$$L_i(t) = \int_0^t \int_0 z N_i(ds, dz) \tag{14}$$

We shall choose compensator measures of the form $\nu_i(dt, dz) = dt\tilde{\nu}_i(dz)$ for a Lévy measure $\tilde{\nu}_i(dz)$. This implies that $L_i(t)$ is a Lévy process, or

$$\nu_i(dt, dz) = \rho_i(t)dt\tilde{\nu}_i(dz),$$

where $\rho_i(t)$ is a deterministic function controlling the possibly time-varying jump intensity. In many markets spikes have a tendency to occur in certain periods of the year, and thus it is natural to let the jump intensity for these vary seasonally. Since the jump process has increasing paths, the jumps are only positive, and thus the Lévy measures $\tilde{\nu}_i(dz)$ are supported on the positive real line. The spot price will be positive as well, since the processes $Y_i(t)$ will be positive by the definition of these jump processes.

The main idea of the factor model is to decompose the electricity spot price into the base and spike signals. This flexibility allows one to capture mean-reversion at different scales, but at the cost of a quite complicated estimation procedure. An example of a possible model specification using three OU processes is proposed in Benth et al. (2007). The first OU process is assumed to have a stationary Gamma distribution and a constant volatility, responsible for small daily fluctuations around the mean trend. For the second OU process a compound Poisson process is used to capture larger price movements which revert faster to the mean. The third process drives the spikes, and has possibly a seasonally varying jump intensity.

In the additive structure of n OU processes one has n mean-reversion parameters $\lambda_1,\ldots,\lambda_n$. The larger λ_i is, the faster the process $Y_i(t)$ comes back to its mean-level. The autocorrelation function $\rho(k)$ for lag k of $X(t)$ is

$$\rho(k) = \tilde{w}_1 e^{-k\lambda_1} + \tilde{w}_2 e^{-k\lambda_2} + \cdots + \tilde{w}_n e^{-k\lambda_n}, \tag{15}$$

where $\tilde{w}_i$ are positive weights summing up to 1. A comparison with the empirical autocorrelation function thus allows one to find the number of factors required and to estimate the mean-reversion from each of the factors. This approach was proposed in Barndorff-Nielsen and Shephard (2001) for their stochastic volatility model, which uses a structure similar to the factor model. A fast reversion in spikes will be observed as a strongly decaying slope in $\rho(h)$, whereas the more slowly reverting base signal is found as slower decaying exponentials.

5. Algorithms applied and estimation results

For our empirical analysis we use a data set of the Phelix Base electricity price index at the European Energy Exchange (EEX). The data series range from 13/07/2000 to 7/8/2008, where the weekends are excluded. In total, we have 2099 daily prices constituting the basis for our spot-price estimation. The reason for excluding weekends is mainly because Friday price information is contained as a basis for Saturday, Sunday and Monday morning prices and thus the prices over the weekend are not directly comparable to those settled during the week.

5.1. Seasonality trend parameters estimation

We start by estimating the seasonality function. The choice of this function is a very important issue, since the specified trend should explain the average market expectation of the price course for the next month, quarter or year. Moreover, deseasonalisation is the first step of the price estimation, so depending on the seasonality estimation, the deseasonalised data set may vary considerably. Cartea and Figueroa (2005) suggest fitting monthly averages of the historical data by a Fourier series of order 5. This is a questionable choice, since this function does not incorporate any trend component, which is necessary to model power dynamics. Also, it is not completely clear how to define the optimal order for a Fourier series. Moreover, the number of parameters for a chosen Fourier series fit with order of 5 is 12, which is twice as many as in the parametric case introduced in Geman and Roncoroni (2006) and Benth et al. (2007). Here the authors use a function $\mu(t)$ described in the previous section. For a deseasonalisation procedure, we take the latter approach for all three models.

In order to adjust for the influence of large price outliers (such as spikes) in the seasonality estimation, we implemented a simple filtering procedure. This compares smoothed (by averaging over sliding windows of 30 data point length) and observed log prices. If the absolute difference between the smoothed and real log price is bigger than a pre-defined level, here 0.5, we substitute the log price with the smoothed value. In this way we "filter" out the price outliers from the original price data series.

On the filtered data series, we estimate the parameters of the seasonality function $\mu(t)$ specified in Eq. (2) using non-linear least squares method. The results are reported in Table 1. One can see that almost all estimates are significant: $\alpha,\beta,\gamma,\varepsilon$ are significant with $t_{0.99}=2.326$, ς with $t_{0.95}=1.645$ and ζ with $t_{0.9}=1.282$. In Fig. 1 we plot the estimated seasonality function $\mu(t)$ together with the EEX data. During the period of 2000–2006 the price path and seasonality are matched very well. Due to structural breaks the fit deteriorates between 2006 and 2008. The price path amplitude has changed in that period compared to previous years. We could capture this by splitting the data set into the two periods and then estimate the seasonality separately. In Fig. 2 we show the deseasonalised price series.

We continue with the estimation of the three models based on the deseasonalised data, that is, the data obtained after dividing with the estimated $e^{\mu(t)}$-function. For the factor model, we work with these data, whereas for the jump-diffusion and threshold models we use the log-data.

5.2. Jump-diffusion model calibration

From the historical spot data we estimate the rolling historical volatility $\sigma(t)$ and its averaged value, the mean-reversion rate α, the frequency l and the standard deviation σ_j of the jumps.

5.2.1. Volatility

Cartea and Figueroa (2005) suggest the volatility to be time-dependent. The motivation is that markets do not show constant volatility, but some volatility structure. However, the authors do not provide any prescribed function or any stochastic alternative for the volatility. Instead, they compute a rolling historical volatility suggested in Eydeland and Wolyniec (2002), which in fact is a deterministic result given a data path. If we make a plot of such a rolling historical volatility, we observe in Fig. 3 that there appears some seasonal pattern as well as some stochastic element. Obviously, they both need to be incorporated into the model. Otherwise, such a "substitution" by historical volatility may result in unreasonable estimates.

5.2.2. Mean-reversion rate

To get a daily estimate for the mean-reversion rate α, Cartea and Figueroa (2005) suggest using linear regression. The idea of such approach is to rewrite the mean-reversion jump-diffusion process in the discrete version and represent the log price as

$$x_t = a_t + b x_{t-1} + c_t, \tag{16}$$

where a_t represents a function of $\mu(t)$, $b \equiv e^{-\alpha}$, c_t is the integral of the Brownian motion and the jump component between times $t-1$ and t.

Table 1
Seasonality function estimated on filtered data. $R^2=0.6961$.

Parameter	Estimate	Std. error	t-value
α	2.9628	0.0092	320.8417
β	0.1354	0.0020	68.1191
γ	0.0737	0.0064	11.4226
δ	0.0117	0.0064	1.8446
ε	6.8662	0.0875	78.4585
ζ	0.7464	0.5436	1.3389

Table 2
Estimates for the jump-diffusion model parameters.

Parameter	Estimate
α	0.2255 (0.1938, 0.2584) $R^2 = 0.6373$
$[\sigma(t)]$	3.9025
σ_j	1.0996
l	5.67

5.2.3. Jump parameters

To estimate jump parameters, we need to identify jumps from the data. We use a simple technique based on the standard deviation of the returns (see Cartea and Figueroa, 2005). The iterative procedure lters out returns with absolute values greater than three times the standard deviation of the returns of the series at the current iteration. The process is repeated until no further outliers can be found. As a result we obtain a standard deviation of the jumps, σ_j, and a cumulative frequency of jumps, l. The latter is de ned as the total number of l-tered jumps divided by the annualised number of observations. We report the results in Table 2, where annualised estimates and average (denoted by $[\cdot]$) volatility are given appropriately with the 95 con - dence interval values in the parenthesis.

5.3. Threshold model calibration

Geman and Roncoroni (2006) propose to split the calibration procedure in two steps. First, the so-called *structural elements* are to be estimated, then in the second step one estimates the *model parameters*. The structural elements are the spread Δ, the components of the intensity shape $\iota(t)$ and the maximum jump size ψ. The remaining are the model parameters.

5.3.1. Estimation of structural elements

First, one needs to nd the spread Δ, which will limit the jump size of the model. The choice of Δ is a result of a balance between two competing effects: the larger Δ, the higher are the price levels which can be reached during the pressure period and the fewer spikes will occur; the smaller Δ, the sooner the downward jump effect toward normal levels takes place and more spikes will occur. Following (Geman and Roncoroni, 2006), we select Δ in such a way that the corresponding calibrated model generates paths whose average maximum values are equivalent to those observed in the market. In the case of the EEX data on log-scale, this results in a Δ spread given by 0.7.

Geman and Roncoroni (2006) suggest choosing the maximum jump size ψ as the observed maximum daily absolute variation in log-prices, which for our data set takes the value 2.2361. Note that this parameter can give some non-realistic results in the spike size, since it can be considered as an upper limit for the spikes. The intensity shape function $s(t)$ is estimated manually in such a way that the most salient spikes coincide with the intensity shape peaks. In addition, the intensity shape should capture clusters of spikes. This procedure leads to $k = 0.7$, $d = 0.75$ and $\tau = 0.42$. In Fig. 4 we show the resulting calibrated shape function together with the most prominent spikes, extracted according to the speci ed threshold Γ. For comparison with the non-parametric spike intensity, the empirical spike intensity is plotted in Fig. 7 based on all the detected spikes. (The comparable intensity for the factor model is in Fig. 12.)

Geman and Roncoroni (2006) analysed the effect of the stochastic spike intensity function in Eq. (9) on the ECAR market, and found no statistical evidence that this improved the t. A stochastic intensity could potentially model a spike reversion" given by a negative jump following a positive one. Fig. 5 shows the EEX prices together with the calibrated trend $\mu(t)$ and detected spikes. Fig. 6 depicts the price values for six typical weeks when spikes were detected (the spikes in red, the following prices in black). The general picture is that one can see an upward price spike, followed by a quick reversion back, given by one or more decreasing prices. One could attribute this to *rst* having a negative spike, and next mean-reversion is dampening the prices further, *or* a sequence of negative spikes. In order to get this, one must have a low level for the stochastic spike intensity in order to create a negative spike with suf ciently high probability. But we also need the negative spike *size* to be of certain magnitude in order to push the prices suf ciently down. To have a sequence of negative spikes, we must have a very low threshold for the stochastic spike intensity, as well as the threshold for having negative sign of the jumps. This will most likely be in conflict with the estimation of positive price spikes that we naturally want to include in the model. The

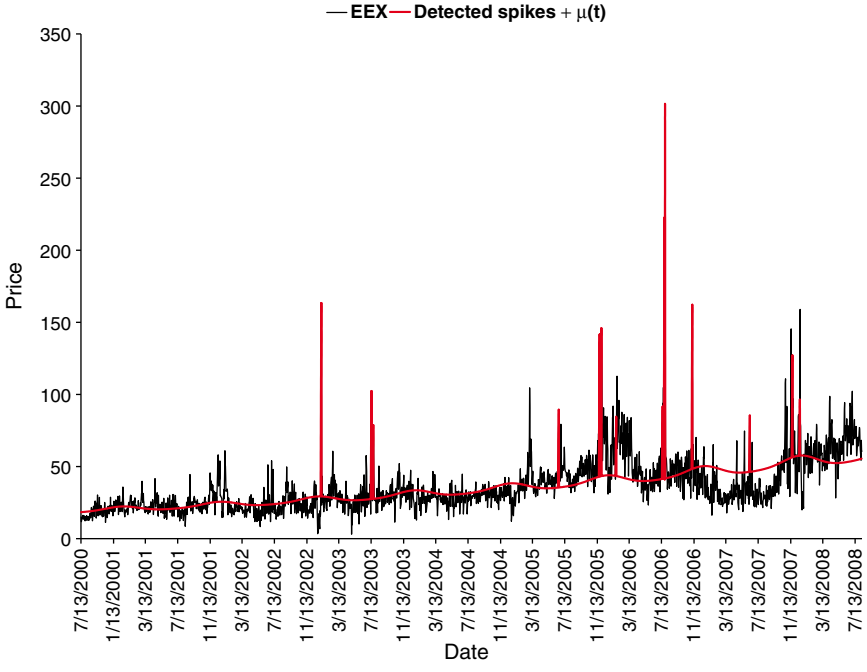


Fig. 5. EEX price process together with calibrated seasonality and detected spikes.

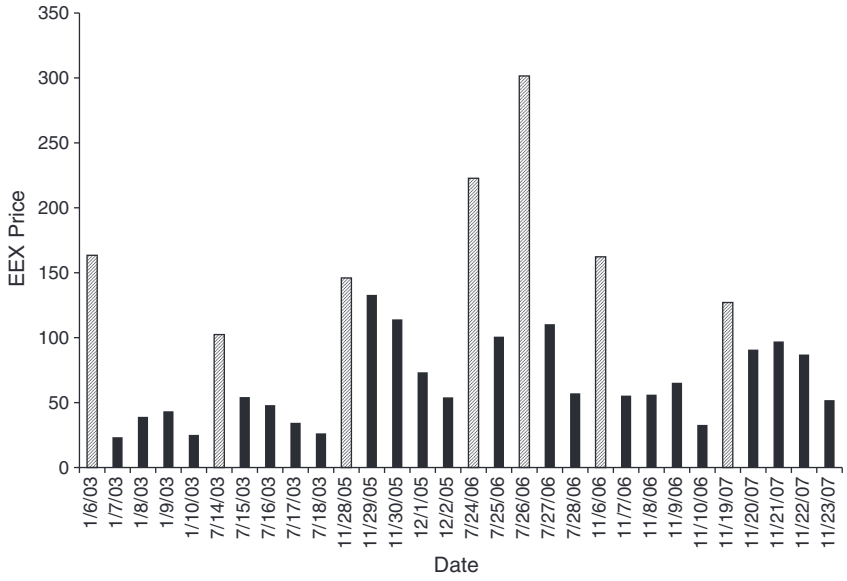


Fig. 6. EEX price values for six selected weeks when spikes were detected. The spikes bars are shaded, the following prices bars are in black.

alternative to having negative spike(s) following a positive price spike is strong mean-reversion. Indeed, the mean-reversion speed does not have to be increased very much in order to have a decay like we see in Fig. 6. This discussion suggests that the maximum-likelihood estimator will put more emphasis on a mean-reversion adjustment rather than the stochastic spike intensity, since after all there are fairly few data for the spikes to rely the estimation on. As the variations in the EEX data are already quite big for the non-spike" regime, coupled with strong reversion, we have the potential for large price decays without the help of negative spikes.

5.3.2. Estimation of the model parameters

After selecting the structural parameters, we estimate the model parameters. First of all, a jump threshold Γ is set to lter out E^d and E^c, i.e. the jump and continuous paths, respectively. Then, based on this, the smooth mean-reversion force θ_1, the maximal expected number of jumps θ_2, the reciprocal average jump size θ_3 and the Brownian local volatility σ are estimated. The fourth moment of the fully speci ed model is then computed, and compared to the empirical one. If the deviation is too large, the jump threshold Γ is either decreased or increased, and the whole procedure is repeated. Geman and Roncoroni (2006) propose iterating the procedure of choosing the jump threshold Γ until the estimated model matches the fourth moment of the daily log-price return distribution.

The parameters $\Theta = (\theta_1, \theta_2, \theta_3)$ are to be estimated by approximate maximum likelihood. The log-likelihood function explicitly depends on Θ and the ltered data set $\left(E^d, E^c\right)$ and implicitly on the choice of Γ, which is used to obtain E^d and E^c. The approximate logarithmic likelihood function is given by Geman and Roncoroni (2006) as

$$\begin{aligned}\mathcal{L}\left(\Theta\middle|\Theta^0, E\right) = &\sum_{i=0}^{n\ 1} \frac{(\mu(t_i)\quad E_i)\theta_1}{\sigma^2} E_i^c \quad \frac{t}{2}\sum_{i=0}^{n\ 1}\left(\frac{(\mu\quad E_i)\theta_1}{\sigma}\right)^2 \\ &(\theta_2\quad 1)\sum_{i=0}^{n\ 1} s(t_i)t + N(t)\ln\theta_2 \\ &+\sum_{i=0}^{n\ 1}\left[\quad(\theta_3\quad 1)\frac{E_i^d}{h(E_i)}\right] + N(t)\ln\quad\frac{1\quad e^{\quad\theta_3\psi}}{\theta_3(1\quad e^{\quad\psi})}\end{aligned} \tag{17}$$

The rst part represents a discretised version of the Doléan–Dade exponential for continuous processes. The remaining parts are responsible for the jump process. The log-likelihood function explicitly depends on $\theta_1, \theta_2, \theta_3$ and the Γ- ltered" data set. The likelihood function is maximised with respect to θ_1, θ_2 and θ_3 over a bounded parameter set Θ, taking some economically sound limiting values into account. The function is constructed in such a way that one can split it up into three independent parts and maximise them separately

$$L(\Theta|\Theta^0, E) = F_1(\theta_1) + F_2(\theta_2) + F_3(\theta_3) \tag{18}$$

Such a modi cation helps to facilitate the optimization algorithm and to increase the estimation correctness.

The next parameter to be estimated is the volatility of the continuous path. We use

$$\sigma = \sqrt{\frac{1}{T}\sum_{i=0}^{n\ 1} E^c(t_i)^2} \tag{19}$$

This estimator for the volatility was applied in Roncoroni (2002), and is based on the quadratic variation of the continuous path; see Genon-Catalot and Jacod (1993) for details. Table 3 contains the estimates of the model parameters.

An estimate of the expected number of jumps during the period is provided by the integral of the intensity function over the whole period, resulting in $[N(1)] = 2\ 6885$. The number of ltered jumps is 39, a number which depends strongly on the selected jump threshold Γ.

To understand the estimated speed of mean-reversion θ_1 better, it is worthwhile to nd the *half-life* of the mean-reversion. The concept of half-life takes its origin from physics and generally describes a

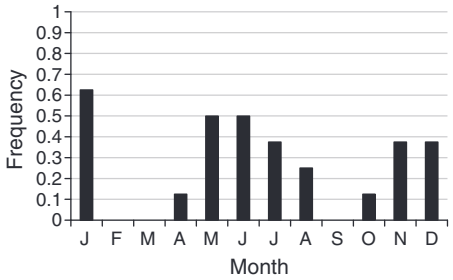


Fig. 7. Historically based frequency of spike occurrence for the threshold model.

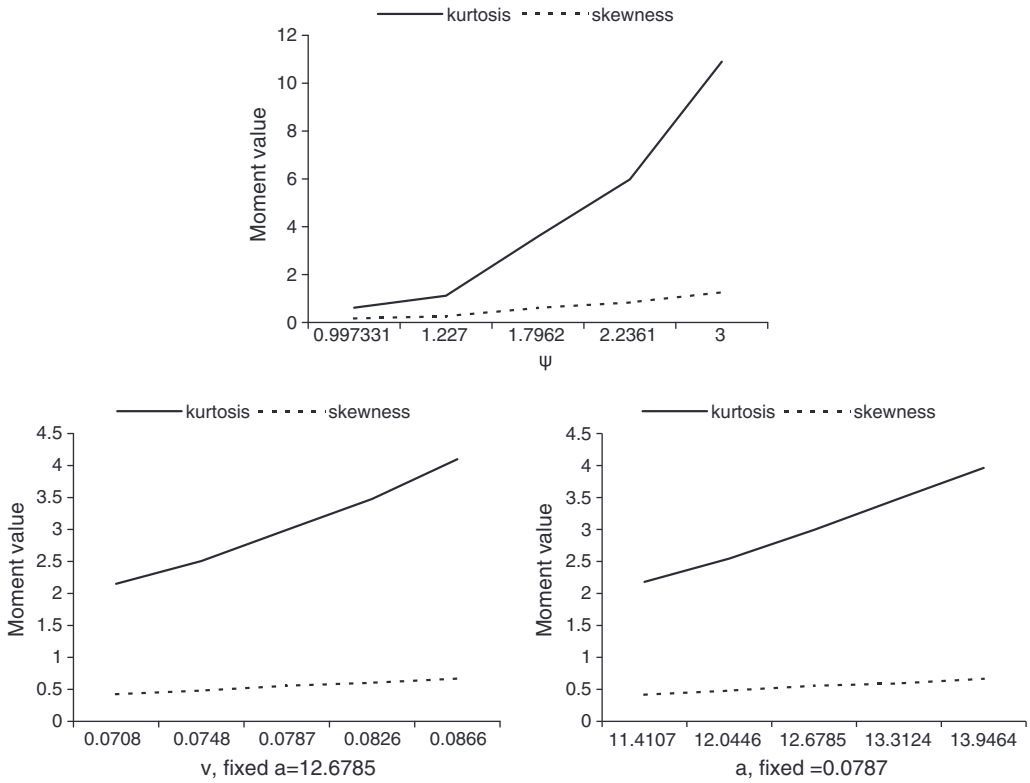


Fig. 8. Top: Sensitivity check for the skewness and the kurtosis, when the parameter ψ takes various values. Bottom left: Sensitivity check for the skewness and the kurtosis, when the parameter ν takes various values, while the other parameter is at xed value $\alpha = 12.6785$. Bottom right: Sensitivity check for the skewness and the kurtosis, when the parameter α takes various values, while the other parameter is at xed value $\nu = 0.0787$.

period of time it takes for a substance undergoing decay to decrease by half. The half-life of an OU process is de ned as the average time it takes before a price jump reverts back to half of its original value (see Clewlow and Strickland (2000)). Mathematically it can be calculated as $(\ln 2)/\lambda$, where λ is a decay constant, i.e. the mean-reversion speed. Here the estimated θ_1 is yielding a half-life of 2.79 days, that is, it takes the process on average slightly less than 3 days to revert back to its mean. In the factor model, where we separated the spike process from the base signal, we found that the spikes had a half-life of around 2 days, whereas the base signal was initially estimated to have a half-life of around 3.5 days. We see that the threshold model has a half-life approximately the average of these two gures. This clearly demonstrates that the threshold model is not capable of allowing a mean-reversion which pushes spikes back fast enough on the one hand, and at the same time is suf ciently slow to push back the lower variations in the price path in quieter periods. In order to make up for the faster mean-reversion for the base signal", one may expect an upward bias in the volatility estimate.

Based on the ltered jumps, one may look for alternative spike-size distributions which are more realistic than the truncated exponential. As mentioned earlier, by truncating the spike-size distribution, one essentially introduces an upper limit for possible jumps. By using the historical price changes future price behaviour will repeat the past. In fact, the estimate on the upper bound ψ may become very unstable, since the spikes in most electricity price series are of

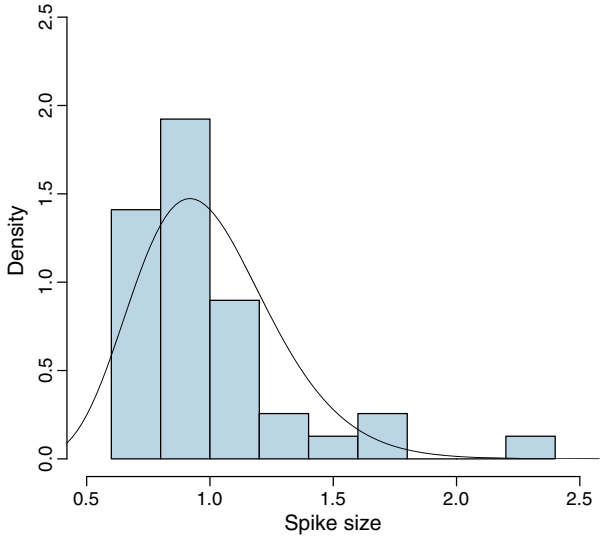


Fig. 9. A tted Gamma distribution versus the spike size histogram for the threshold model.

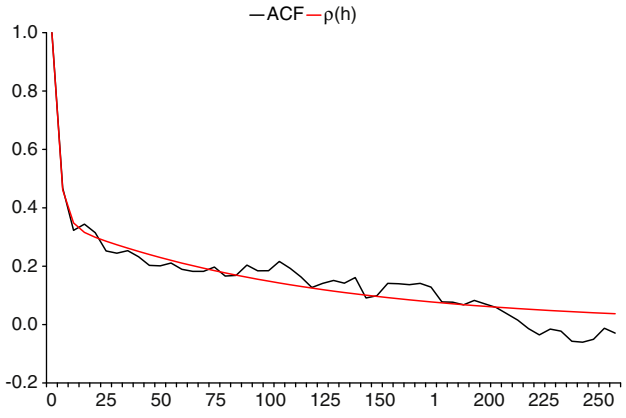


Fig. 10. Empirical ACF for EEX series and weighted sum of two exponentials.

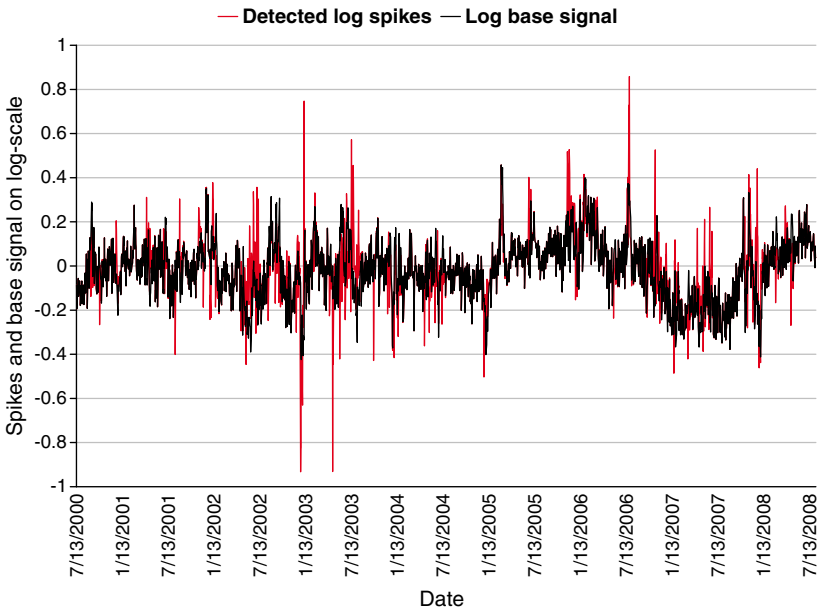


Fig. 11. Detected log spikes and log base signal.

very different size, even by several magnitudes. In order to see this effect in the data, we based the estimate of ψ on different price changes. For example, rather than basing our truncation on $\psi = 2.2316$, corresponding to the three biggest price changes, we may choose the second biggest price change instead to obtain $\psi = 1.7962$. The effect on the moment properties of the path turns out to be dramatic. In Fig. 8 we plot the skewness and the kurtosis of the price path as a function of ψ. The estimates of these two moments are based on a number of simulations. The simulations have been repeated until the change of the averaged moment value becomes less than 0.01%. We see that the range of the kurtosis varies dramatically with the choice of ψ. Thus, using historical price changes to truncate the spikes may lead to a very unreliable model which may seriously fail to capture the true distribution of spike sizes. We also plot the skewness and the kurtosis of the price path as a function of α and ν, when using a Gamma law for spikes sizes. By fixing one parameter and varying another, we can see their influence on the moment values. It is obvious that the range of the kurtosis and the skewness changes slightly with the parameters. This leads us to affirm that Gamma-distributed spikes yield more stable moments than the truncated exponential distribution for spikes. To cope with this defect, we have fitted a Gamma distribution for the spikes as we did with the factor model. In Fig. 9 we have plotted the estimated Gamma distribution together with the empirical density of spikes. The maximum likelihood parameters were found to be $\alpha = 12.6785$ and $\nu = 0.0787$.

Using a non-truncated distribution, we include the possibility of observing bigger jumps than the historically observed ones.

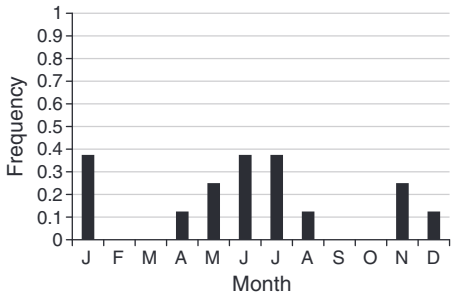


Fig. 12. Historically based frequency of spike occurrence for the factor model.

5.4. Factor model calibration

The first step in calibrating the factor model is to assess the number of factors required. We compare in the L^2 norm the empirical autocorrelation function (ACF) with the theoretical autocorrelation functions from Eq. (15) for different numbers of factors. We obtain $n = 2$ as the optimal number of factors. The estimated speeds of mean-reversion and weights are reported in Table 4 and the estimated and empirical ACFs are in Fig. 10.

We associate $\lambda_2 = 0.3333$ to the spikes having the fastest speed of mean-reversion. The base signal is associated to the factor Y_1 with estimated mean-reversion $\lambda_1 = 0.0087$. We find that the half-life of the base signal Y_1 is 79.6721 and for the spike 2.0794. Thus, on average the base signal needs nearly 80 days to come back to half its value while a spike needs only 2 days. After filtering out the spikes, we will re-estimate the speed of mean-reversion for the base signal.

5.4.1. Filtering of spikes

The next step is to filter out the spikes. Meyer-Brandis and Tankov (2008) show that the hard-thresholding procedure taken from Extreme Value Theory to identify the spikes is a reliable technique in the context of return-distribution characteristics. In short, it filters out the spike process using methods from non-parametric statistics

Table 3
Estimates of the threshold model parameters.

Element	Interpretation	EEX
θ_1	Smooth mean-reversion force	0.2480
θ_2	Max. expected number of jumps	14.5144
θ_3	Reciprocal average jump size	1.0584
σ	Brownian local volatility	3.8216
	Jump threshold	0.6750

Table 4
The fitted ACF with a sum of two exponentials.

λ_i	$\tilde{w}_i$
0.0087	0.3547
0.3333	0.6453

and provides as output both the base signal and the spike process. We refer to (Meyer-Brandis and Tankov (2008) and Nazarova (2008) for a detailed explanation of the approach to the EEX data. In Fig. 11 one sees the result of the hard-thresholding procedure: the spikes and the base signal on log-scale. As found in Meyer-Brandis and Tankov (2008) the method of hard-thresholding is insensitive to the value of λ_1 in the range of 0.1 to 0.01 and relatively insensitive to the value of λ_2. Recall our estimate of $\lambda_1 = 0.0087$, being close to the desired range.

5.4.2. Estimating the base signal

We continue to estimate the parameters of the model for the base signal. The issue here is to nd the right subordinator process $L_1(t)$ which ts the observed ltered time series. The usual way is to propose a stationary distribution that t the data, and nd the so-called *background driving Lévy process* $L_1(t)$ such that $Y_1(t)$ has the same stationary distribution. A typical choice is the Gamma distribution, which leads to a compound Poisson process $L_1(t)$ with exponential jumps (see Benth et al., 2007; Barndorff-Nielsen and Shephard, 2001). The reason for the choice of the Gamma distribution is the availability of an explicit analytical expression for the moments, otherwise we would need to use some numerical methods.

As suggested by Benth et al. (2007), we apply a method based on prediction-based estimating functions developed by Sørensen (2000) and Bibby et al. (2010) to calibrate the base signal model to data. The details of the method can be found in Appendix A.

Table 5 gives the parameter estimates from the implementation of the prediction-based estimating function technique. The re-estimate of the speed of mean-reversion λ_1 implies a half-life of approximately 3.5 days, much faster than initially estimated by matching autocorrelation functions. In view of the very noisy behaviour in the base signal, this seems more likely than 80 days as initially found.

The method has its advantages and disadvantages. It is well-grounded from the theoretical point of view. However, for practical applications we face the problem that the equation $G_n(\theta) = 0$ has no unique solution. Therefore, the algorithm may nd different roots for different initial parameter values. However, if the initial parameters are close to the true one, the resulting estimated parameter values are correct in the sense that they match the moments. Bibby et al. (2010) suggested nding some optimal weights to improve the ef ciency of the estimator, but in the case of multiple solutions this approach does not help signi cantly. Therefore, our calibration of the base signal was carried out in two steps: rst a calibration by hand" to identify likely intervals for the parameters values, then execution of the prediction-based estimating functions algorithm using these initial values.

5.4.3. Analysis of the spike process

The nal step in our estimation procedure of the factor model is to calibrate the spike process. Since spikes are rather sparse compared to the total length of the data set there are few data points available for estimating the intensity and the jump-size distribution. To cope with this problem we consider various speci cations and analyse their consequences.

We shall apply the popular shape function proposed by Geman and Roncoroni (2006), given in Eq. (8), and estimate the parameters from the data at hand. An alternative approach would be to choose an intensity based directly on the observed distribution of spikes over the year. In Fig. 12 we have plotted the historical frequency of the spike occurrence. To use the historical frequency has the advantage of an easy and fast adjustment as new market data become available. Furthermore, the calibration procedure is very simple compared to parametric approaches.

Recall the intensity function $\iota(t,\theta)$ de ned in Eq. (7), where we included a dependency on the parameter vector $\theta = (\theta_2, k, d)$. The maximum likelihood estimator of θ is given by (see Meyer-Brandis and Tankov, 2008)

$$\theta^* = \arg\max_{\theta} \left(\sum_{i=1}^{N_T} \log \iota(\tau_i, \theta) \quad \int_0^T \iota(t,\theta) dt \right), \tag{20}$$

where τ_i are spike times. A time-dependent intensity function is the natural choice when there is some pronounced seasonality in spikes. The US markets analysed in Geman and Roncoroni (2006) demonstrate evident spike seasonality, while in the EEX market this is not so obvious. Out of the 30 biggest positive spikes, 16 occurred in summer, 7 in winter, the remaining 7 in spring and fall. We found a phase $\tau = 0.42$, with $d = 1.0359$, $k = 0.5$ and $\theta = 14.0163$ (as θ_2 in the threshold model). Fig. 13 shows the intensity shape and the largest 30 spikes, detected by the hard-thresholding procedure. As can be seen from the picture, EEX data do not demonstrate such a pronounced seasonality in spikes as could be suggested from the intensity shape function.

The next problem is to estimate a jump-size distribution for the spikes. In Meyer-Brandis and Tankov (2008), a Pareto distribution was suggested for the spike sizes. The estimation procedure involves a threshold or scale parameter z_0, the smallest value the Pareto random variable may take, and a parameter α for the tail-fatness. These are estimated by means of tting a straight line to the empirical cumulative distribution function (CDF) on log–log scale, i.e. a traditional Hill estimator, which is ef cient when the underlying distribution is Pareto; see Drees et al. (2000) for details. We show the result in Fig. 14. We nd the estimates $z_0 = 0.3648$ and $\alpha = 2.5406$.

The Pareto distribution has very heavy tails, and may give unreasonably high values of the spot price. An alternative distribution for tting the spike sizes may be the Gamma distribution. In Fig. 15 we compare the tted Gamma density with the empirical spike-size density. The maximum-likelihood estimates for the two parameters α and ν of the Gamma distribution are $\alpha = 6.2592$ and $\nu = 0.0942$.

Table 5
Resulting estimates for OU process.

Element	Value
λ_1	0.2008
α	13.3009
ν	8.5689

6. Model comparison

In this section we assess our estimated models, and discuss their properties in the context of the EEX electricity spot price behaviour. In Fig. 16 we have plotted typical simulated paths of the models, along with the observed EEX prices series. Visually the performance of all three models is quite satisfactory. One apparent difference is that the factor model seems to be less noisy in the intra-spike periods than the data. The jump-diffusion and the threshold models, on the other hand, are more noisy than the data, at least according to the experience from our simulation studies. The spike pattern looks better for the factor model in these simulations.

A standard, widely used model check is to compare model-based moments to the empirical ones. In our particular case it may not be reliable since the threshold model in fact is calibrated using the fourth moment as a target and thus should match at least the kurtosis almost perfectly. We report the rst four moments of the returns in Table 6 for the jump-diffusion, the threshold and the factor models together with the empirical moments of the EEX data. The descriptive statistics are computed for the empirical versus simulated logarithmic price variations, i.e. log returns. The simulations have been repeated until the change of the averaged moment value becomes less than 0.01%.

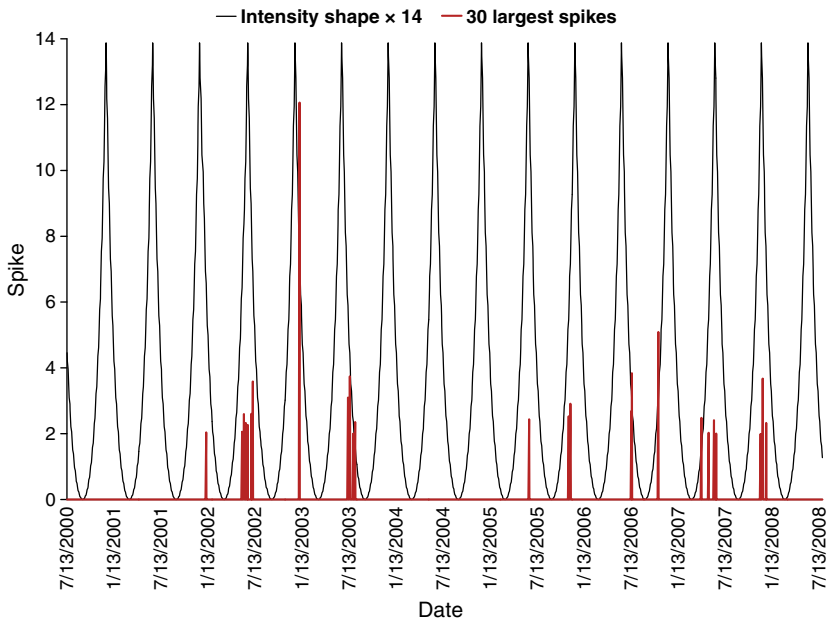


Fig. 13. Calibrated intensity shape function and real values of the largest detected spikes.

The gures in the table indicate the best performance for the threshold model. It matches the rst four moments very well, except possibly the skewness, where it overshoots. However, the close resemblance with the empirical moments is not really a surprise since these have been used as a measure in the estimation procedure. The jump-diffusion model, on the other hand, shows two apparent characteristics. It overestimates the skewness and kurtosis, which can be explained by the use of a Normal distribution for the spike sizes.

The factor model also demonstrates two very apparent characteristics. Firstly, it underestimates the standard deviation signi cantly, yielding a simulated volatility of the path only half the size of the empirical. This can be explained from using the hard-thresholding procedure. Meyer-Brandis and Tankov (2008) suggest choosing a threshold value for the deseasonalised log-returns such that the share of returns larger than the chosen threshold value does not exceed 5% of all returns. Then, the standard deviation of the remaining 95% is computed and called the target one, according to which we can separate spikes and the base signal. It is obvious that the total number and size of ltered spikes depend critically on the threshold value. In our case we obtained the target standard deviation equal to 0.1454, which looks quite consistent with the calibrated standard deviation. Of course, if we change the criterion in the hard-thresholding procedure, we will receive different results for the target standard deviation and the ltered spikes and the base signal, respectively. Decreasing the share of returns above the threshold to 1.76%, we obtain a larger target standard deviation, which may be more in line with the observed one. Possibly, one could think of an iterative procedure parallel to nding the Γ in the threshold model. The kurtosis of the estimated factor model is close to twice as big as the empirical. This is a result of using a Pareto law for the spike sizes, implying rather extreme jump sizes which obviously influence the kurtosis. Note that the high positive skewness of the factor model is a result of the large spikes.

The truncated exponential law used in the threshold model prevents the occurrence of any spikes bigger than the ones we have already observed, which clearly helps in getting a kurtosis close to the observed one. However, it is not clear whether matching the empirical kurtosis is a useful tting criterion. The empirical kurtosis takes only the observed changes into account, and as we pointed out during the sensitivity analysis with respect to the truncation parameter in the jump size distribution for the threshold model, it may be very

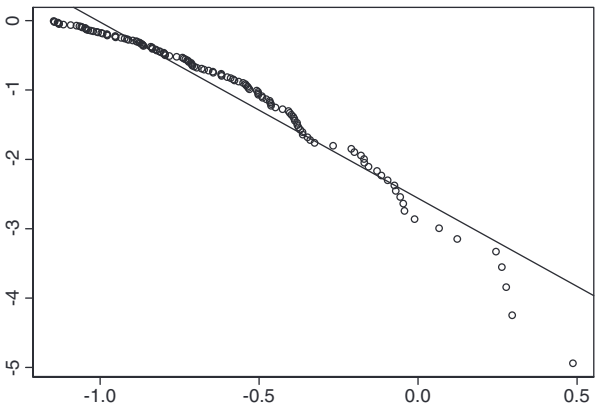


Fig. 14. Empirical CDF of spike size on log scale.

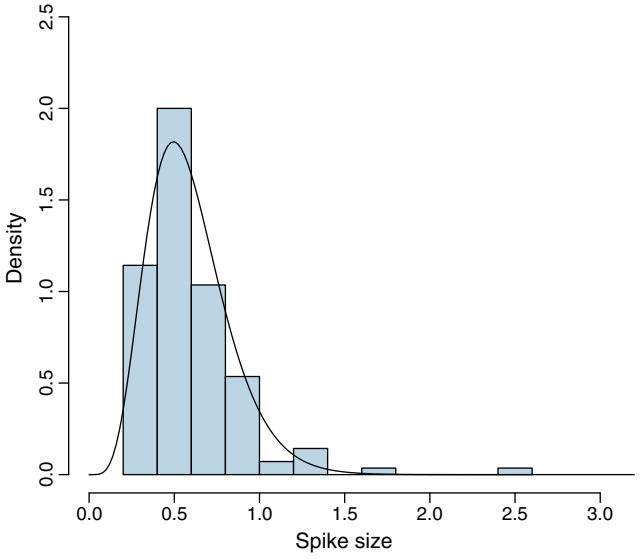


Fig. 15. A tted Gamma distribution versus the spike size histogram for the factor model.

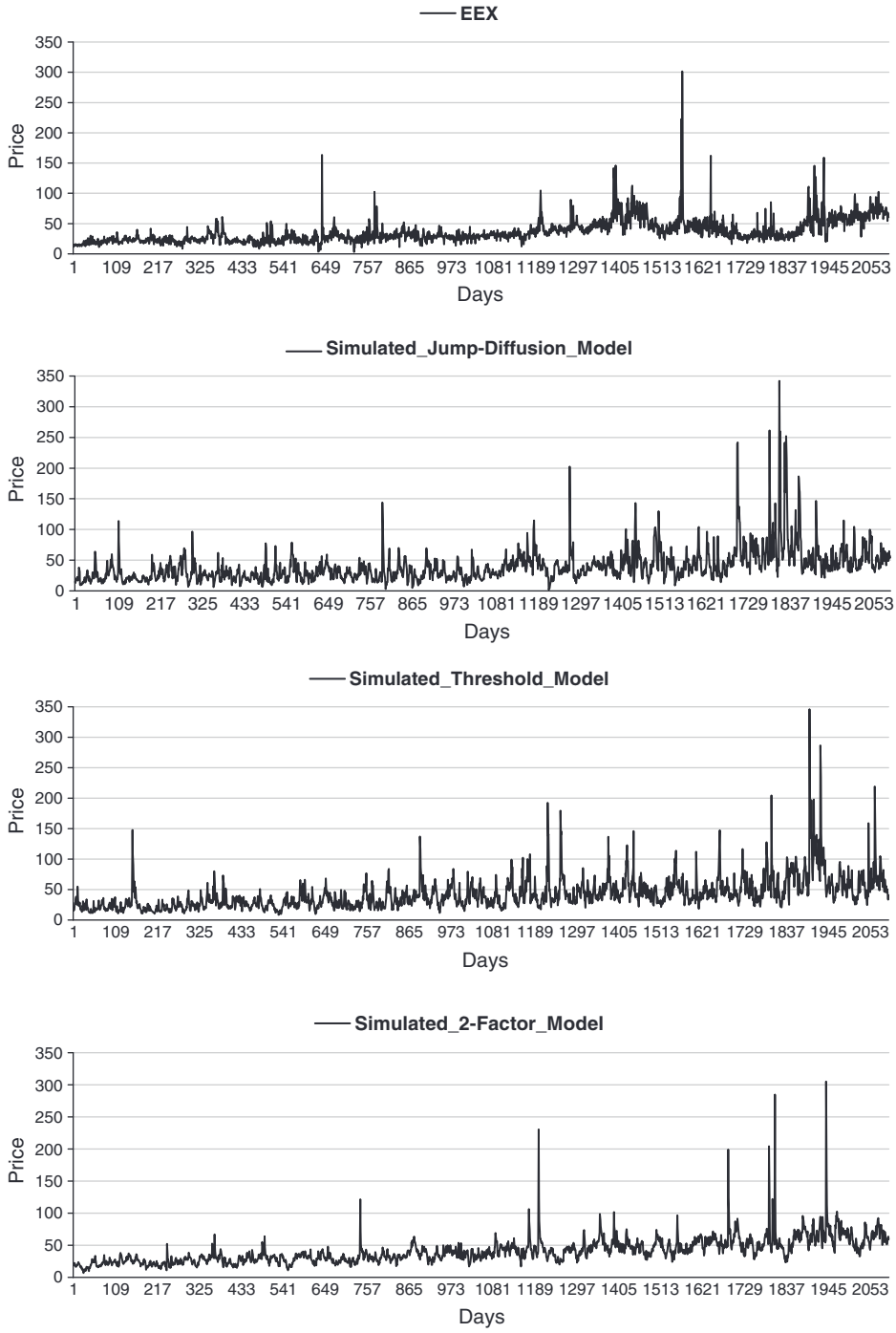


Fig. 16. EEX, jump-diffusion, threshold and factor model simulated price paths.

sensitive to changes in the data. For instance, the occurrence of new spikes bigger than the observed leads to a higher kurtosis. In practice, one should be careful since the empirical moments are backward looking. The Pareto, Gamma or Normal distribution may in fact give a much better picture of the true risk in the future.

Table 6
Empirical moments versus jump-diffusion, threshold and factor model moments.

Moment	Average	Std. dev	Skewness	Kurtosis
EEX	0.0006	0.2985	0.4050	6.6179
Jump-diffusion model	0.0007	0.3191	0.8343	10.3935
Threshold model	0.0006	0.2935	0.8336	5.9783
Factor model	0.0006	0.1595	1.6749	10.5308

As we could see in Fig. 16, the noise in the intra-spike periods is too low in the factor model, an observation that is con rmed by the simulated standard deviation of the path. This may be attributed to the choice of the base signal model. Its relatively slow mean-reversion combined with a compound Possion process as driving noise will yield paths which may not look as random as the market. The compound Poisson process will only jump at Poisson distributed random times, and otherwise the base signal will decay exponentially. Although the stationary distribution of the model honours the statistical properties of the base signal, we may have paths which are too regular. A possible modi cation could be to include a Brownian-

driven factor. This would possibly lead to negative prices, which may be a problem in some applications. However, in the market, and in particular the EEX market, negative prices occur from time to time. Alternatively, we may use a subordinator which is not given by a nite-activity process, but one with in nite activity such as the background driving Lévy process derived from supposing an inverse Gaussian stationary distribution of the base signal.

To further analyse the base signal properties, we look at the moments of the data after the spikes have been ltered out. In Table 7 we report the rst four simulated moments of the factor model, together with the empirical ones. Note that here we consider the moments of the process, but not its log-returns, as we do while comparing models' performance. It is very obvious from these numbers that the base signal model does not suf ciently explain the variations in the data. Turning to the jump-diffusion or to the threshold model, we can mimic the base signal by considering the dynamics without the jump component. Since this is a mean-reverting model on log-scale, the variance in stationarity is given by $\sigma^2/2\alpha$ or $\sigma^2/2\theta_1$, and inserting the estimates this gives us a stationary standard deviation of 5.8111 or 5.4263 respectively. The gures are far above what is reasonable to expect by looking at the empirical standard deviation. Hence, we see that due to the too fast speed of mean-reversion for the base signal in the jump-diffusion and the threshold models, it compensates by gearing up the volatility to unrealistic levels. Combined with a too slow mean-reversion for the spikes, one may question whether the pathwise properties of the spot price are honoured in the threshold model. We remark that in Oyebanji (2007) an even higher σ was estimated for EEX data.

To better understand the effect of having a different spike size distribution, we apply the Kolmogorov Smirnov test to check the goodness of t of the different spike size distributions. We test the H_0 hypothesis of the suggested distribution with estimated parameters at the 5 con dence level. The results are reported in Table 8.

From the results we can observe that the data present suf cient evidence to contradict our hypothesis that the spike sizes follow Normal, truncated exponential or Pareto distributions in case of the jump-diffusion, the threshold or the factor models respectively. However, there is no reason to reject the hypothesis on the choice of Gamma law for spike-size distribution for the threshold and the factor models. Therefore, we can conclude that the results of the Kolmogorov Smirnov test support the modi cation of the spike size distribution. We do not provide the results of the test for the Gamma distribution for the jump-diffusion model. We analyse the jump-diffusion model to have a fair comparison of two complex models with one parsimonious and to check the effect of the model speci cation on the results it gives. Therefore, we omit the modi cations for the jump-diffusion model here.

Coming back to comparing the simulated moments of the three models with the empirical ones based on the Gamma distribution for the jump sizes, we report the result in Table 9. We observe that the kurtosis of the threshold model drops signi cantly, and is now underestimating the empirical kurtosis. The reason is that although the Gamma distribution allows for unbounded jumps, it is concentrated more on the smaller spikes sizes, while the truncated exponential is more stretched and gives a higher probability to bigger spikes. The factor model is more in line with the empirical kurtosis when using the Gamma distribution for spike sizes. The standard deviation is still insuf cient for the factor model, whereas the threshold matches very well even for Gamma-distributed spikes. But in this comparison, which is fairer since we use the same model for spikes, the factor model seems to be better at capturing the leptokurtic effects of the model. Of course, being faithful to the estimation procedure of Geman and Roncoroni (2006), the iterative approach combined with estimating parameters so that the fourth moment is matched would also yield a perfect t to the empirical kurtosis for the case of a Gamma distribution. However, as we can see here, the factor model can obtain a signi cant improvement by simply changing the jump distribution if this is desirable, whereas the threshold model requires a complete re-estimation although we only introduce a new distribution for the identi ed spikes being more in line with the observed jump sizes. Re-estimating the complete model as in Geman and Roncoroni (2006) would lead to a different estimation function (see Eq. (17)) and different parameter estimates, showing that the speed of mean reversion and volatility, for instance, are indeed very sensitive to the choice of the spike-size distribution. Therefore, the model cannot easily be adjusted to new assumptions.

Table 7
Base signal moments matching.

Moment	EEX base signal	Simulated
Average	1.0262	1.1241
Std. dev.	0.3198	0.2637
Skewness	1.0177	0.4405
Kurtosis	2.4674	0.6756

Table 8
Kolmogorov–Smirnov test results for the considered spike distributions.

Distribution	Test statistics	Decision on H_0
Normal (jump-diffusion model)	2.1855	Reject H_0
Trunc. exp (threshold model)	3.5032	Reject H_0
Pareto (factor model)	1.7748	Reject H_0
Gamma (threshold model)	0.9375	No reason to reject H_0
Gamma (factor model)	1.1556	No reason to reject H_0

7. Application to derivative pricing

In this section we apply the three spot models to pricing of forward contracts. To assess the performance of the models we compare theoretical prices with the observed ones, and compute the implied risk premium. We have available analytical forward prices for the jump-diffusion and factor models, while we use a simulation-based Monte Carlo pricing method for the threshold model.

The price at time t of the forward contract with maturity T is de ned as the expected value of the spot price at expiry date under an equivalent pricing measure $\mathbb{Q}$

$$\mathbb{F}^{\mathbb{Q}}(t,T) = \quad {}^{\mathbb{Q}}_{t}[S(T)| \quad _t] \tag{21}$$

Since the electricity market is an incomplete market, there exist many pricing measures $\mathbb{Q}$. To pin down a price $F(t,T)$, one needs to select *one* such measure, and this would typically be done in practice by restricting the space of measures $\mathbb{Q}$ to a parametric class, for example given by the Girsanov and Esscher transforms (see Benth et al., 2008a for the details). Next, by minimising the distance between theoretical and observed prices, one could nd a pricing measure.

For the three spot models under inquiry, one has different classes of pricing measures $\mathbb{Q}$ available. The jump-diffusion and threshold models would naturally involve a change of measure for Brownian motion and compound Poisson processes, whereas the factor model does not involve any Girsanov transformations, but a change of

Table 9
Comparative descriptive statistics results for the log-returns of the jump-diffusion, threshold and factor models.

Moment	Average	Std. dev	Skewness	Kurtosis
EEX spot	0.0006	0.2985	0.4050	6.6179
Jump-diffusion model (Normal)	0.0007	0.3191	0.8343	10.3935
Threshold model (trunc. exp)	0.0006	0.2935	0.8336	5.9783
Factor model (Pareto)	0.0006	0.1595	1.6749	10.5308
Modi ed threshold model (gamma)	0.0006	0.2822	0.5566	2.9946
Modi ed factor model (gamma)	0.0006	0.1465	1.2414	5.7399

measure with respect to subordinators. In order to have a fair comparison of the forward pricing ability of the three spot models, we focus on the implied risk premium. This does not involve any change of measure $\mathbb{Q}$, and therefore avoids introducing properties of the models which are depending on the selection of pricing measure.

The empirical risk premium is computed as the difference between the predicted spot price and the observed market price (see Benth et al., 2008a)

$$RP(t) = F^{observed}(t,T) \quad F^{\mathbb{P}}(t,T) \tag{22}$$

Here, $F^{\mathbb{P}}(t,T)$ is the so-called *predicted* spot price at delivery T, computed as in Eq. (21) with $\mathbb{Q} = \mathbb{P}$. A careful analysis of the empirical risk premium is not only valuable for understanding the spot model performance, but it is also the fundament for proposing classes of pricing measures $\mathbb{Q}$. Moreover, it has obvious applications to risk management.

In electricity markets the forward contracts deliver power (or the money equivalent to power) over a speci ed period. These periods may typically be a month, a quarter or a year. In the empirical studies to come, we shall focus on monthly delivery periods, and the risk premium is extended in a natural way as simply taking the average of the right-hand side of Eq. (22) over the delivery period. That is, the risk premium for a forward contract delivering electricity over the period $[T_1,T_2]$ is

$$RP(t) = F^{observed}(t,T_1,T_2) \quad F^{\mathbb{P}}(t,T_1,T_2), \tag{23}$$

where

$$F^{\mathbb{P}}(t,T_1,T_2) = \frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} F^{\mathbb{P}}(t,T)dT$$

7.1. Pricing formulas

For the jump-diffusion model, we can compute the predicted spot price semi-analytically

Proposition 1. *The predicted spot price of the jump-diffusion model is given by*

$$F^{\mathbb{P}}(t,T) = (X(t))^{e^{\quad \alpha(T \quad t)}} \exp\left(\mu(T) + \int_t^T \frac{1}{2}\sigma^2(s)e^{\quad 2\alpha(T \quad s)}ds + \int_t^T \left[e^{\quad \alpha(T \quad s)}\frac{\sigma_J^2}{2} + \frac{\sigma_J^2}{2}e^{\quad 2\alpha(T \quad s)}\right]lds \quad l(T \quad t)\right) \tag{24}$$

Proof. We refer to Cartea and Figueroa (2005) for a detailed proof.

The formula for $F^{\mathbb{P}}(t,T)$ involves the rolling historical volatility $\sigma(s)$. In our investigations, we have set this to a constant, de ned as the average of the rolling historical volatility over the period with data, see Table 2. Furthermore, to derive $F^{\mathbb{P}}(t,T_1,T_2)$ for a delivery period $[T_1,T_2]$, we take the average of the computed forward prices $F^{\mathbb{P}}(t,T)$ for each day T in the delivery period.

One of the bene cial properties of the factor model is that it provides us with analytical forward prices for contracts with a delivery period, $F^{\mathbb{P}}(t,T_1,T_2)$. In the next proposition we state this price for the case of $Y_1(t)$ and $Y_2(t)$ identi ed and being driven by compound Poisson processes with exponential jump distributions. Note that this means that both are stationary Gamma distributed. The resulting predicted spot price can be computed following the lines in Benth et al. (2007).

Proposition 2. *Suppose for the factor model that $Y_1(t)$ is stationary Gamma distributed $\Gamma(\nu,\alpha)$, and $Y_2(t)$ are driven by a compound Poisson process $L_2(t)$ with exponential jump size distribution with parameter γ, and jump intensity l. Then, the predicted spot price of the factor model is*

$$F^{\mathbb{P}}(t,T_1,T_2) = \frac{1}{T_2 \quad T_1}\left[\left(Y_1(t) \quad \frac{\nu}{\alpha}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_1(u \quad t)}du + \left(Y_2(t) \quad \frac{l}{\gamma}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_2(u \quad t)}du + \left(\frac{\nu}{\alpha}+\frac{l}{\gamma}\right)\int_{T_1}^{T_2} e^{\mu(u)}du\right] \tag{25}$$

Proof. The result follows from a straightforward calculation. We start by plugging Formulas (1), (12) and (13) together. Hence, after commuting integration and conditional expectation, and using the fact that $Y_1(t)$ and $Y_2(t)$ are $_t$-measurable, we nd

$$\begin{aligned}
F^{\mathbb{P}}(t,T_1,T_2) &= E^{\mathbb{P}}\left[\frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} S(u)du\middle| \quad _t\right] \\
&= \quad ^{\mathbb{P}}_t\left[\frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} e^{\mu(u)}(Y_1(u)+Y_2(u))du\middle| \quad _t\right] \\
&= \quad ^{\mathbb{P}}_t\left[\frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} e^{\mu(u)}\left(Y_1(t)e^{\quad \lambda_1(u \quad t)} + \int_t^u e^{\quad \lambda_1(u \quad s)}dL_1(s) + Y_2(t)e^{\quad \lambda_2(u \quad t)}\int_t^u e^{\quad \lambda_2(u \quad s)}dL_2(s)\right)du\middle| \quad _t\right] \\
&= \frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} e^{\mu(u)}\left(Y_1(t)e^{\quad \lambda_1(u \quad t)} + \quad ^{\mathbb{P}}_t\left[\int_t^u e^{\quad \lambda_1(u \quad s)}dL_1(s)\middle| \quad _t\right] + Y_2(t)e^{\quad \lambda_2(u \quad t)} + \quad ^{\mathbb{P}}_t\left[\int_t^u e^{\quad \lambda_2(u \quad s)}dL_2(s)\middle| \quad _t\right]\right)du
\end{aligned}$$

By the independent increment property of Lévy processes, we get

$$\begin{aligned}
F^{\mathbb{P}}(t,T_1,T_2) &= \frac{1}{T_2 \quad T_1}\int_{T_1}^{T_2} e^{\mu(u)}\left(Y_1(t)e^{\quad \lambda_1(u \quad t)} + \int_t^u \varphi_1'(0)e^{\quad \lambda_1(u \quad s)}ds + Y_2(t)e^{\quad \lambda_2(u \quad t)} + \int_t^u \varphi_2'(0)e^{\quad \lambda_2(u \quad s)}ds\right)du \\
&= \frac{1}{T_2 \quad T_1}\left[\left(Y_1(t) \quad \varphi_1'(0)\frac{1}{\lambda_1}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_1(u \quad t)}du + \left(Y_2(t) \quad \varphi_2'(0)\frac{1}{\lambda_2}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_2(u \quad t)}du + \left(\varphi_1'(0)\frac{1}{\lambda_1}+\varphi_2'(0)\frac{1}{\lambda_2}\right)\int_{T_1}^{T_2} e^{\mu(u)}du\right] \\
&= \frac{1}{T_2 \quad T_1}\left[Y_1(t)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_1(u \quad t)}du + Y_2(t)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_2(u \quad t)}du + \int_{T_1}^{T_2} e^{\mu(u)}\left(\frac{\nu}{\alpha}\left(1 \quad e^{\quad \lambda_1(u \quad t)}\right)+\frac{l}{\gamma}\left(1 \quad e^{\quad \lambda_2(u \quad t)}\right)\right)du\right] \\
&= \frac{1}{T_2 \quad T_1}\left[\left(Y_1(t) \quad \frac{\nu}{\alpha}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_1(u \quad t)}du + \left(Y_2(t) \quad \frac{l}{\gamma}\right)\int_{T_1}^{T_2} e^{\mu(u) \quad \lambda_2(u \quad t)}du + \left(\frac{\nu}{\alpha}+\frac{l}{\gamma}\right)\int_{T_1}^{T_2} e^{\mu(u)}du\right]
\end{aligned} \tag{26}$$

In the derivation procedure we use the fact that $[L] = \varphi'(0)$, where the latter is the derivative of the log-moment generating function of the process L, more precisely $\varphi(x) = \left[e^{xL(1)}\right]$. For the case of our processes, $\varphi'_1(0) = \frac{\nu\lambda_1}{\alpha}$ and $\varphi'_1(0) = \frac{l\lambda_2}{\gamma}$ respectively. Hence, the proposition follows.

Due to the state-dependent sign function h in the jump term of the threshold model, it does not allow for any analytical forward prices. Therefore, we apply a Monte Carlo simulation method to price

forwards $F^{\mathbb{P}}(t,T)$. We start with the simulation of the spot process algorithm of which can be found in the original paper of Geman and Roncoroni (2006). We further compute a price $F^{\mathbb{P}}(t,T)$. To de ne an optimal number of simulations we use the method of control variates and take the jump-diffusion model as a benchmark. By minimising the total sum of squared differences between the simulated and the analytical forward prices for the jump-diffusion model, we de ne the necessary number of simulations. The Monte Carlo technique we use is properly described in Glasserman (2004) and Fusai and Roncoroni (2008). To obtain $F^{\mathbb{P}}(t,T_1,T_2)$, we compute the forward prices for every particular day of the delivery period and then average the obtained results over this period.

7.2. Empirical analysis

We work with the following data sets of daily electricity forward prices collected from the EEX data sets of the summer and winter terms of different years:

- Forwards with delivery in July of 2004, 2005, 2006, 2007 and 2008, observed in the preceding month June
- Forwards with delivery in December of 2004, 2005, 2006 and 2007, observed in the preceding month November

Given the observed forward prices $F(t,T_1,T_2)$, with the delivery period $[T_1,T_2]$ being June or December, and t ranging over the working days in the month prior to delivery, we compute the implied risk premium based on the predicted spot prices derived according to the algorithms and formulas described above. In the following discussion, we focus our attention on the results of June 2007, November 2007 and June 2008. The results of other observation periods can be found in Appendix B.

Figs. 17, 18 and 19 show the predicted spot price (blue curve) together with the observed forward dynamics (green curve). In addition, we have included the seasonal function (seasonality trend $\mu(t)$, Eq. 2) over the delivery period as a reference level (red curve). In general, we observe that the average seasonal function sets the level of forward prices in the market, except for July 2008 where it seems to be a large deviation in observed prices away from the seasonal level. The implied risk premia are not converging to zero, which is an obvious implication from the delivery period feature of the electricity forward contracts. Also, the market forward prices are more volatile than the predicted spot prices, except maybe in the period before delivery starts.

The shape of the risk premia looks very similar for all three models. However, there are big differences in the values. For the July 2007 contracts, the jump-diffusion and threshold models assign a negative risk premium (see Figs. 17a and 18a), whereas the factor model implies a positive premium. In fact, the premium is always positive for the factor model (except of one instance, see Fig. B.23b), with occasional very large values (see Fig. 19c). There is a clear tendency of a decreasing risk premium for the factor model as time approaches delivery, while the two other models show evidence of an increasing risk premium for this period. A negative risk premium is in line with the theory of normal backwardation, where producers accept a reduction in price in order to reduce their price risk. However, there exist both theoretical and empirical evidence for a positive premium, explained as the retailers hedging their short-term spike risk (see Benth et al. (2008b)). Although we consider forward prices in the days prior to delivery, i.e., being in the short-end of the forward market, the monthly delivery period should average out this risk. From an economical point of view it seems reasonable to expect a negative premium even close to start of delivery. With this perspective in mind, the factor model does a poor job compared to the two others. This may be attributed to the fact that the factor model did not capture the variations in the base component of the spot price very well. By using a compound Poisson process to describe the base variations, we obtain much less variations that what is obtained by a Brownian component, say, which is present in both of the two other models. Although the mean-reversion of these two is too slow for the base variations of the spot, it seemingly gives an advantage for forward pricing purposes.

Inspecting the risk premia for the jump-diffusion and threshold models more carefully, we observe some unreasonable features as well. For example, the jump-diffusion model has a sign change in the risk premium for the December 2007 contract. As indicated above, one may expect a positive premium in the short end of the market, meaning in the days prior to delivery. For the December 2007 contract, we have a positive premium when we are far from delivery, which is at stake with this (see Figures in Appendix B for other such examples).

Comparing the predicted spot price path with the observed forward prices, it seems that both the threshold and jump-diffusion models are closer to explaining the market than the factor model. The difference between the two models is not too big, which is a reflection of the low frequency of spike occurrences and therefore a similar path behaviour. However, the differences are still signi cant, so the impact of the function h is apparent. The function h will switch the direction of a jump for exceedingly high or low prices, and thereby increase the variations. Interestingly, the factor model se ms to converge towards the predicted spot price when we approach delivery, whereas the two others drive apart (this is of course also reflected in a risk premium with increasing absolute value for the jump-diffusion and threshold models, and whereas decreasing to zero for the factor model). Let us discuss this in closer detail with a view towards a potential class of measure changes $\mathbb{Q}$.

Let us simplify the discussion and consider a toy model for the spot price given by a Brownian motion driven by Ornstein–Uhlenbeck process, i.e.

$$dS(t) = \quad \alpha S(t)dt + \sigma dB(t)$$

A natural measure change is a constant Girsanov transform,

$$dW(t) = \quad \frac{\theta}{\sigma}dt + dB(t)$$

with θ a constant, called the market price of risk. From the Girsanov theorem it follows that there exists a probability $\mathbb{Q} = \mathbb{Q}_\theta$ such that W is a Brownian motion under this probability. A direct computation shows that

$$F^{\mathbb{Q}_\theta}(t,T) = X(t)e^{\ \alpha(T\ \ t)} + \frac{\theta}{\alpha}\left(1 \quad e^{\ \alpha(T\ \ t)}\right) + \sigma\int_t^T e^{\ \alpha(T\ \ s)}dW(s)$$

Therefore, the theoretical risk premium for contracts with delivery over $[T_1,T_2]$ becomes

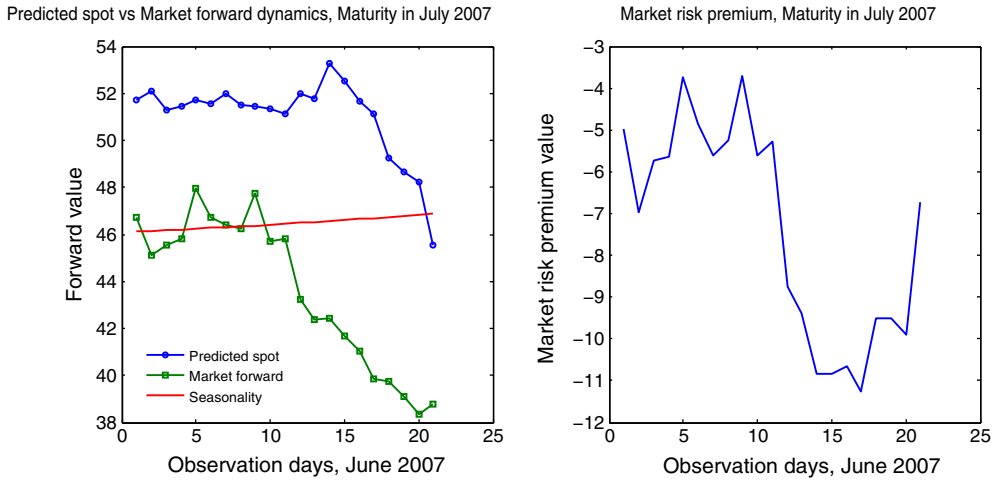
$$R(t,T_1,T_2) = \frac{\theta}{\alpha}\left(1 \quad \hat{\alpha}(T_2 \quad T_1)e^{\ \alpha(T_1\ \ t)}\right)$$

with

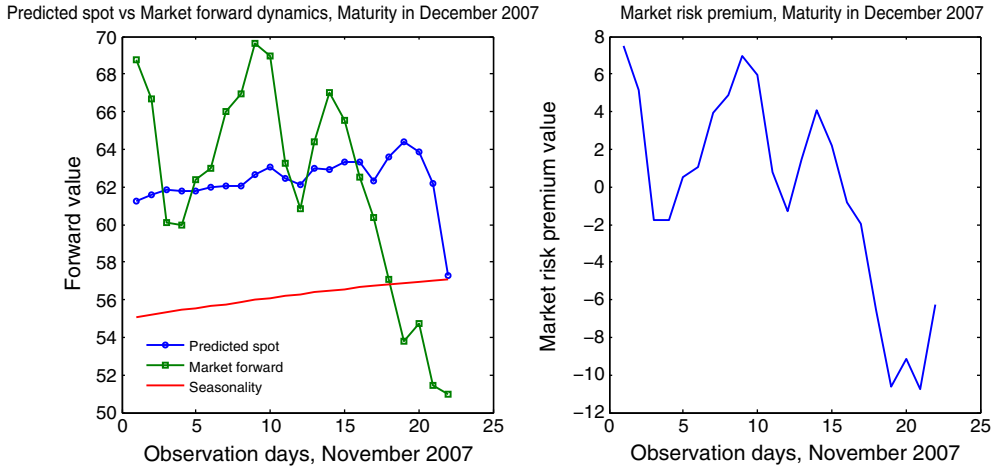
$$\hat{\alpha}(u) = \frac{1 \quad e^{\ \alpha u}}{\alpha u}$$

Hence, as $T_1 \quad t$ increases, the risk premium increases to θ/α. On the other hand, as we approach start of delivery, i.e., $T_1 \quad t \quad 0$, the risk premium decreases towards $\theta(1 \quad \hat{\alpha}(T_2 \quad T_1))/\alpha$. Inspecting Fig. 19, this is exactly the behaviour we observe in the case of the July and December 2007 contracts for the factor model. In the jump diffusion and threshold models we do not observe a similar pattern, but a much more complex structure of the pricing measure $\mathbb{Q}$. Admittedly, the factor model does not have any Brownian component, but we can do the exact same analysis for Lévy driven Ornstein–Uhlenbeck processes using the Esscher

a)The predicted spot vs observe forward dynamics and market risk premium for forwards with maturity in July 2007.



b)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.



c)The predicted spot vs observed forwardd ynamics and market risk premium for forwards with maturity in July 2008.

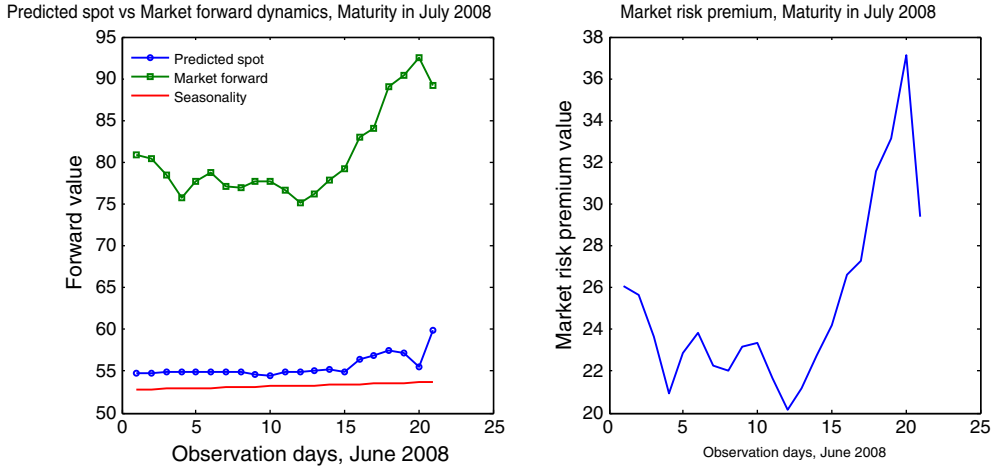
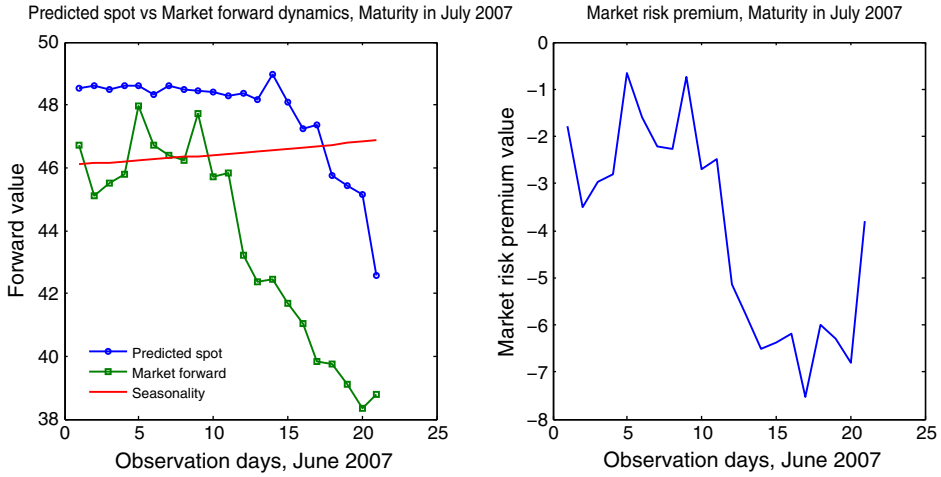
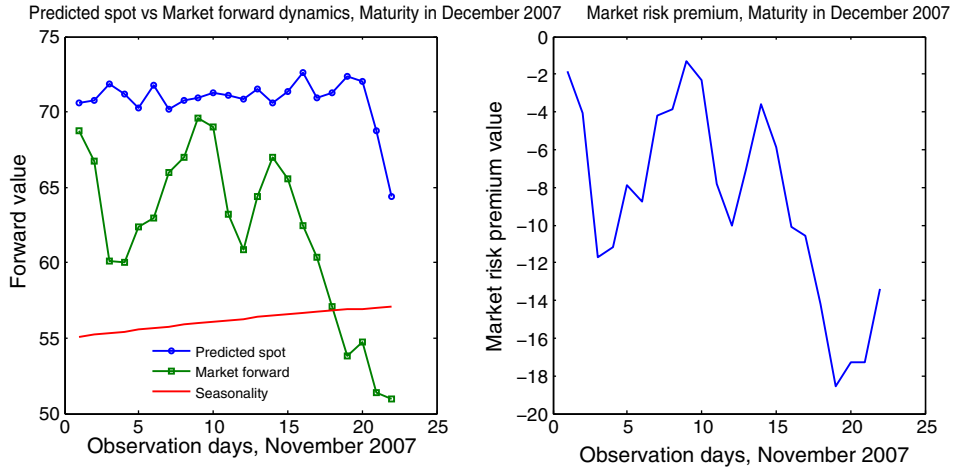


Fig. 17. The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model.

a)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2007.



b)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.



c)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2008.

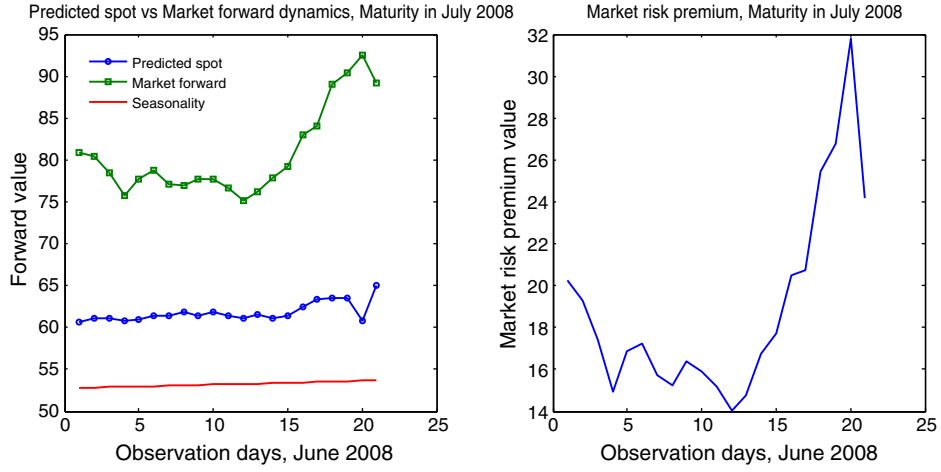
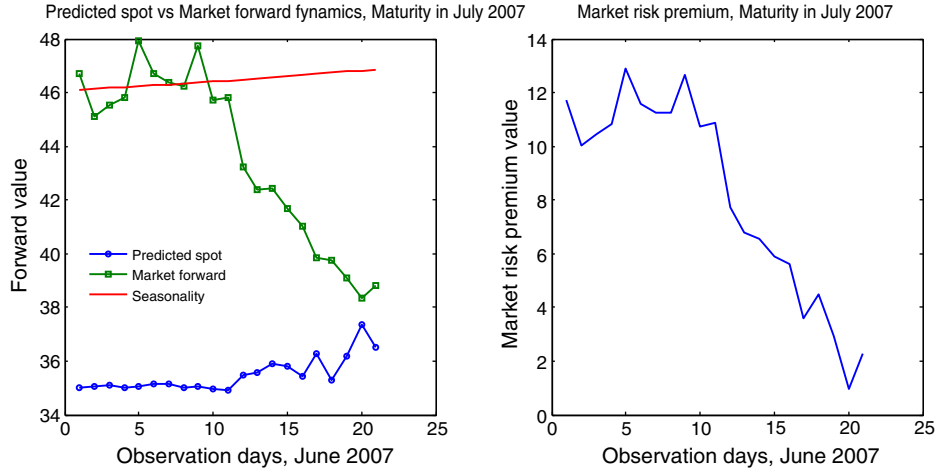


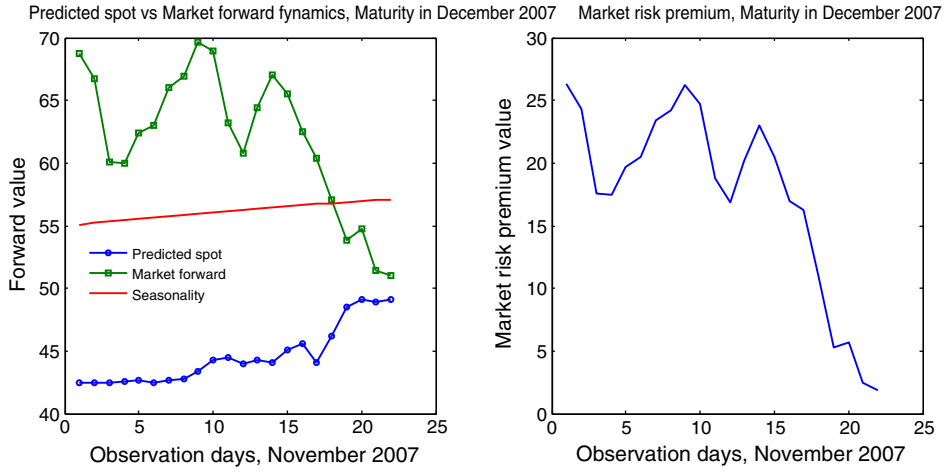
Fig. 18. The predicted spot, observed forward dynamics and market risk premium for the threshold model.

transform (see Benth et al., 2008b). Concerning the two other spot models, the mean-reversion feature will create a similar behaviour when changing measure using a constant market price of risk, although the models are formulated on an exponential form. From the gures, we observe a risk premium contradicting this change of measure. In our opinion, it is a sign of quality that a model allows for an easy explanation

a)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2007.



b)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.



c)The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in july 2008.

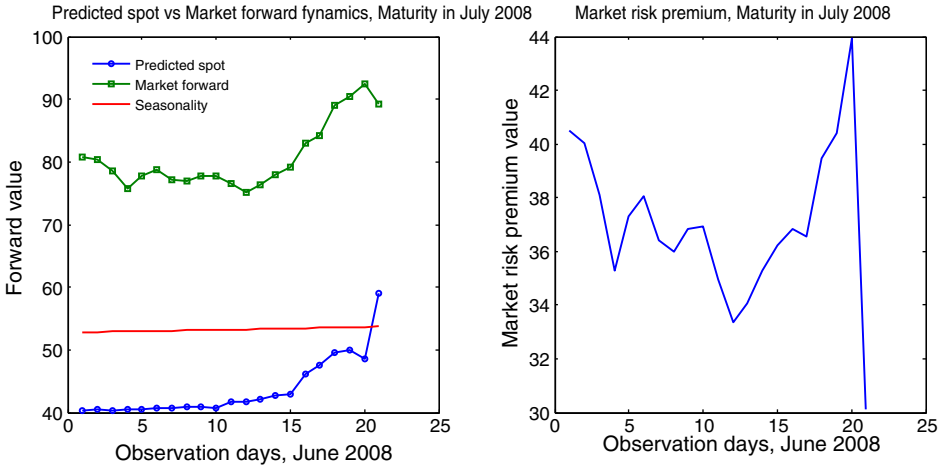


Fig. 19. The predicted spot, observed forward dynamics and market risk premium for the factor model.

of the risk premium, in this case a simple constant change, explaining the market price of risk easily. Note that the market price of risk will be positive for the factor model in the cases we discuss. The July 2008 contract has also a decreasing risk premium towards start of delivery, but a much more complex nature before and does not allow for this simple explanation.

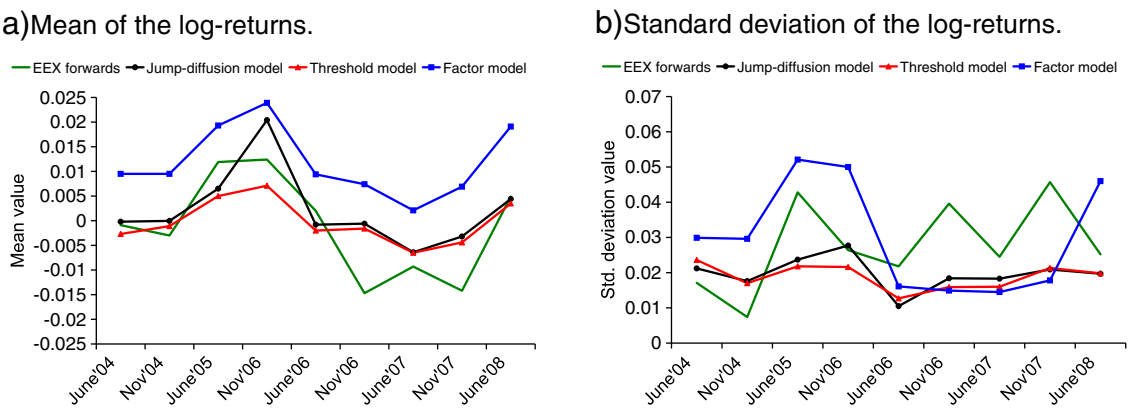


Fig. 20. Comparative descriptive statistics results for the log-returns of the jump-diffusion, threshold and factor models.

In Fig. 20 we plot the descriptive statistics of the log-returns of the predicted spot prices for the observed periods. One can see that the jump-diffusion and threshold model produce similar results for the returns that match the observed EEX forward returns quite well. We can also nd that the factor model gives the predicted spot dynamics whose mean of the log-returns is larger than the observed. This can be explained by the fact that several days before the maturity starts the predicted spot price becomes more sensitive and starts increasing to capture the volatility risk. We observe similar but smaller effect for the jump-diffusion model. The second moment of the observed EEX forwards is more in line with the jump-diffusion and factor models compared to the threshold model, since the latter produces noisier price dynamics. Also, as earlier stated, the market prices look more volatile than the predicted spots. However, this is not the case when considering all the contracts at hand, where we in fact see more variations in predicted spots. This is observed in the data from the early years 2004 and 2005.

8. Conclusions

We have analysed and discussed the empirical performance of three continuous-time electricity spot price models that have received considerable attention from academics and practitioners recently. The mean-reversion parameter both for the jump-diffusion and the threshold models is not able to distinguish between spikes and base signal leading to a too slow mean reversion for the spikes and a too fast mean-reversion for the base signal. For the base signal the models try to compensate for this by a very high volatility. So, the pathwise properties of the EEX price dynamics are not captured well by the jump-diffusion and the threshold models. We nd that the factor model captures the fast mean-reversion of spikes and the slow mean-reversion of the base signal very well. It therefore allows for an excellent modelling of the path behaviour of the mean of the prices. However, the variability of the paths is not captured appropriately. The factor model underestimates the noise in the base signal, a fact that we attribute to the choice of an OU process with a subordinator. Such a selection produces too little variation and thus leads to an underestimation of the standard deviation of the base signal.

One further comment on the performance of the models is their analytical tractability, i.e. for pricing power derivatives. Here the factor and the jump-diffusion models are advantageous. In the case of the jump-diffusion model, by assuming that the jumps J are drawn from a Normal distribution and by requiring that $[J] = 1$ we are able to derive the forward price in closed form. If we switch to another jump-size distribution, then the model will lose its analytical tractability and we have to search for a numerical solution. However, in case of the factor model one may explicitly calculate all the probabilistic properties of the prices in terms of characteristic functions. Furthermore, due to its additive linear structure, electricity forward contract prices are obtained analytically. Electricity forward contracts have the distinctive feature of delivering the underlying commodity, spot power, over a period of time rather than at a xed time. This implies that the price is de ned as a conditional expectation (possibly risk-adjusted) of the integral of the spot. Under the factor model an explicit calculation of this condition expectation is feasible, and we obtain the implied forward price dynamics. Thus the factor model allows one to study price determination in the forward market and risk premia; see Benth et al. (2007) for details. Due to the state-dependent sign of the spike process in the threshold model, we cannot obtain analytical expressions for the characteristic function of the prices process or calculate forward prices explicitly (even with xed maturity and no delivery period). Although ef cient numerical and simulation-based Monte Carlo evaluations are available, we consider the lack of analytical treatability of the threshold model as a major drawback in valuation and risk-management applications.

Appendix A. Prediction based estimating functions method

We follow the steps in Sørensen (2000). We denote by $Z_1, Z_2, \ldots Z_n$ the stochastic processes of the base signal, which we will model by Y_1 with a Gamma stationary distribution. There are three parameters to be estimated, $\theta = (\lambda_1, \alpha, \nu)$, where the last two are the parameters of the distribution, and λ_1 is the speed of mean-reversion of Y_1. Assume that $f_j, j = 1 \ldots N$ are one-dimensional functions, de ned on the state space of Y_1, such that $_\theta\left\{f_j(Z_i)^2\right\}$ for all $\theta \in \Theta$, the parameter space, and $j = 1, \ldots N$, $i = 1, \ldots, n$. For given θ write expectation as $_\theta$. Let $_i$ be the σ-algebra generated by $Z_1, Z_2, \ldots Z_i$, and $\mathcal{H}_i^\theta$ the L^2-space of square integrable $_i$-measurable one-dimensional random variables given that θ is the true parameter value. We denote the set of square-integrable predictors of $f_j(Z_{i+1})$ given $Z_1, Z_2, \ldots Z_i$ by $\mathcal{P}_{ij}^\theta, j = 1, \ldots N$. Observe that this is a closed linear subspaces of $\mathcal{H}_i^\theta$. In the case of the factor model we use $N = 2$, and $f_1(y) = y$ and $f_2(y) = y^2$. These will be our prediction-based estimating functions.

Further one needs to choose an appropriate number q_{ij} of lags under which the prediction-based estimating function will be constructed. These numbers represent the available information required" to predict the consecutive value, and in the case of a stationary process they do not need to be too large. We let $q_{ij} = 4$ for all i, j, i.e., four observations are taken to predict the following one. A space of predictors is speci ed as $U_j^i \ ^1 = (Z_i \ _1, \ldots Z_i \ _4)$ and $U_j^i \ ^1 = (Z_i^2 \ _1, \ldots Z_i^2 \ _4)$ for $j = 1, 2$, resp.

De ne the estimating function $G_n(\theta)$ as

$$G_n(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{N} \Pi_j^i \ ^1(\theta)\left\{f_j(Y_i) \quad \hat{\pi}_j^{(i \ 1)}(\theta)\right\}, \tag{A 1}$$

where $\Pi_j^{(i\ \ 1)}(\theta)=\{\pi_{1,j}^{(i\ \ 1)}(\theta),\pi_{2,j}^{(i\ \ 1)}(\theta),\pi_{3,j}^{(i\ \ 1)}(\theta)\}^T$ is a stochastic vector of weights, which belong to $\mathcal{P}_{i\ \ 1,j}^{\theta}$. The terms $\hat{\pi}_j^{(i\ \ 1)}(\theta)$ are the minimum mean-square error predictors of $f_j(Z_i)$ in $\mathcal{P}_{i\ \ 1,j}^{\theta}$. This predictor $\hat{\pi}_j^{(i\ \ 1)}(\theta)$ is the orthogonal projection of $f_j(Z_i)$ on $\mathcal{P}_{i\ \ 1,j}^{\theta}$ with respect to the inner product in $\mathcal{H}_i^{\theta}$. This projection exists and is uniquely determined by the normal equations

$$ {}_{\theta}\left[\pi\left\{f_j(Z_i)\quad \hat{\pi}_j^{(i\ \ 1)}(\theta)\right\}\right]=0, \tag{A 2}$$

for all $\pi\in\mathcal{P}_{i\ \ 1,j}^{\theta}$. From Eq. (A.2) it follows that $G_n(\theta)$ is an unbiased estimating function. An estimator is obtained by solving $G_n(\theta)=0$. We remark that the weights in $\Pi_j^{(i\ \ 1)}(\theta)$ serve the purpose of improving the ef ciency of the estimator. The optimal choice of weights is a separate task considered in Bibby et al. (2010). In our case these weights did not contribute much and we did not choose any since the algorithm reached convergence without them.

It is assumed that $\mathcal{P}_{i\ \ 1,j}^{\theta}$ is spanned by $U_{j_0}^{(i\ \ 1)},\ldots,U_{j_q}^{(i\ \ 1)}j_k=1,2$, which are linearly independent in $\mathcal{H}_i^{\theta}$. By Eq. (A.2) the minimum mean-square error predictor of $f_j(Z_i)$ in $\mathcal{P}_{i\ \ 1,j}^{\theta}$ is given by

$$\hat{\pi}_j^{(i\ \ 1)}(\theta)=\hat{a}_{j0}^{(i\ \ 1)}(\theta)+\hat{a}_j^{(i\ \ 1)}(\theta)^T U_j^{(i\ \ 1)}, \tag{A 3}$$

where

$$\hat{a}_j^{(i\ \ 1)}(\theta)=C_{i\ \ 1,j}(\theta)\ \ {}^{1}b_j^{(i\ \ 1)}(\theta)$$

and

$$\hat{a}_{j_0}^{(i\ \ 1)}(\theta)=\ {}_{\theta}\left\{f_j\left(Z_j\right)\right\}\quad \hat{a}_j^{(i\ \ 1)}(\theta)^T\ {}_{\theta}\left\{U_j^{(i\ \ 1)}\right\}$$

Here $C_{i\ \ 1,j}(\theta)$ denotes the covariance matrix of $U_j^{(i\ \ 1)}$ when θ is the true parameter value, while

$$b_j^{(i\ \ 1)}=\left[\mathrm{Cov}_{\theta}\left\{U_{j_1}^{(i\ \ 1)},f_j(Z_i)\right\},\ldots,U_{j_q}^{(i\ \ 1)},f_j(Z_i)\right\}]^T$$

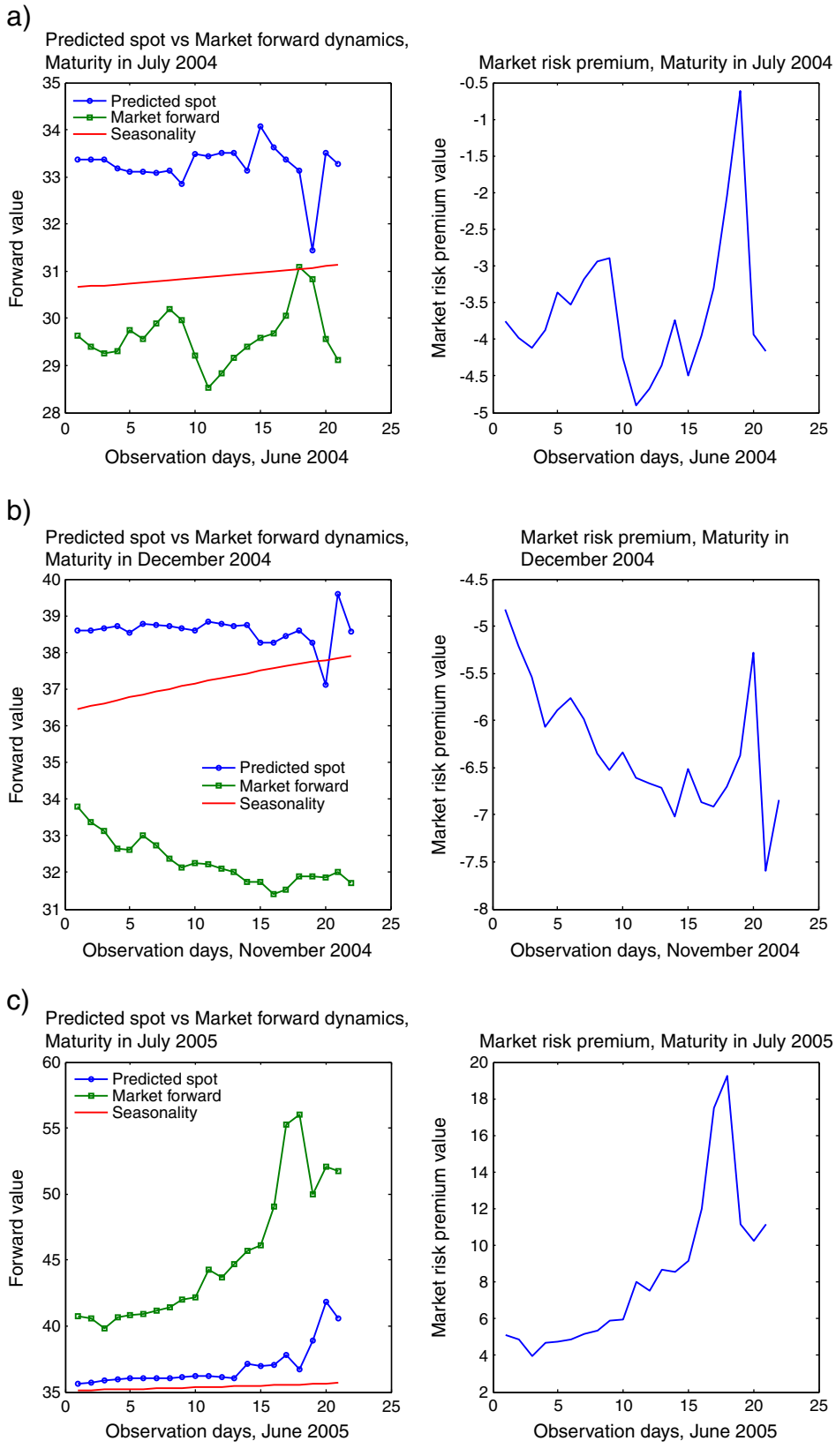
In conclusion, a prediction-based estimating function can be calculated provided that covariances in $C_{i\ \ 1,j}(\theta)$ and $b_j^{(i\ \ 1)}(\theta)$ can be computed. Since $\hat{\pi}_j^{(i\ \ 1)}(\theta)$ depends only on the rst- and the second-order moments of the random vector $\{f_j(Z_i),U_{j_1}{}^{(i\ \ 1)},\ldots,U_{j_q}{}^{(i\ \ 1)}\}$, only parameters appearing in these moments can be estimated using Eq. (A.1). Observe that the characteristic function of $Y_1(t)$ is

$$\left[e^{uY_1(t)}\right]=e^{uY_1(0)e^{\ \lambda t}}e^{\nu t(\alpha\ \ \lambda)}\ \ \frac{\alpha\ \ ue^{\ \lambda t}}{\alpha\ \ u}\Bigg)^{\nu} \tag{A 4}$$

Therefore the moments can be obtained by taking the respective derivative of the characteristic function at $u=0$. In order to apply the prediction-based estimating functions method based on the functions $f_1(y)=y$ and $f_2(y)=y^2$, it is necessary to obtain the four rst moments, which can be computed explicitly for the Gamma stationary OU process. Besides expressions for moments, one needs to derive the covariance between the two rst moments and the covariance between the two second moments, $\mathrm{Cov}(Z_t,Z_{t+s})$ and $\mathrm{Cov}(Z_t^2,Z_{t+s}^2)$.

Appendix B. Forward dynamics for various data sets

Appendix B.1. Jump-diffusion model and its forward modelling



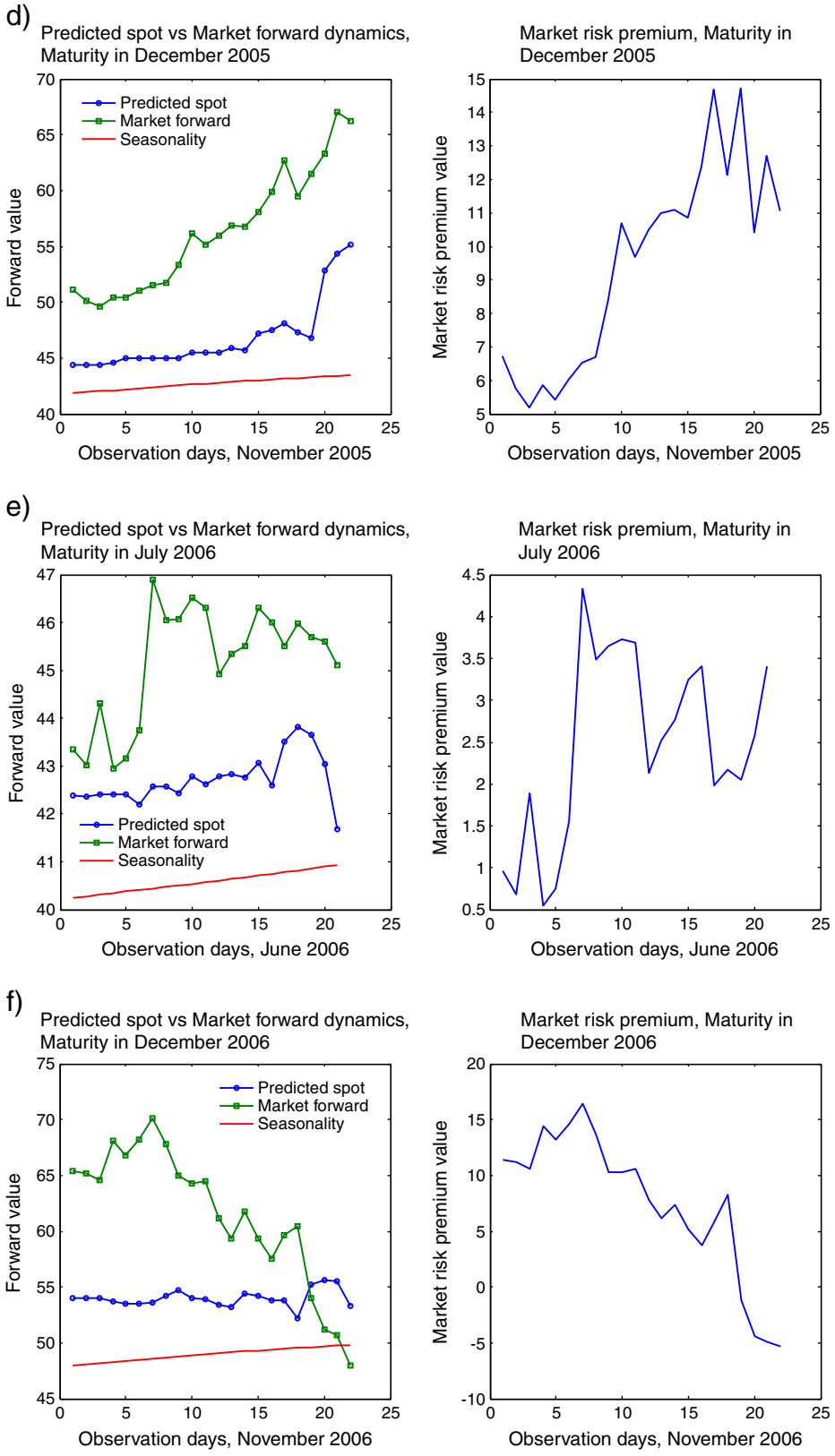


Fig. B.21. The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model. (a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004. (b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004. (c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005. (d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005. (e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2006. (f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

Appendix B.2. Threshold model and its forward modelling

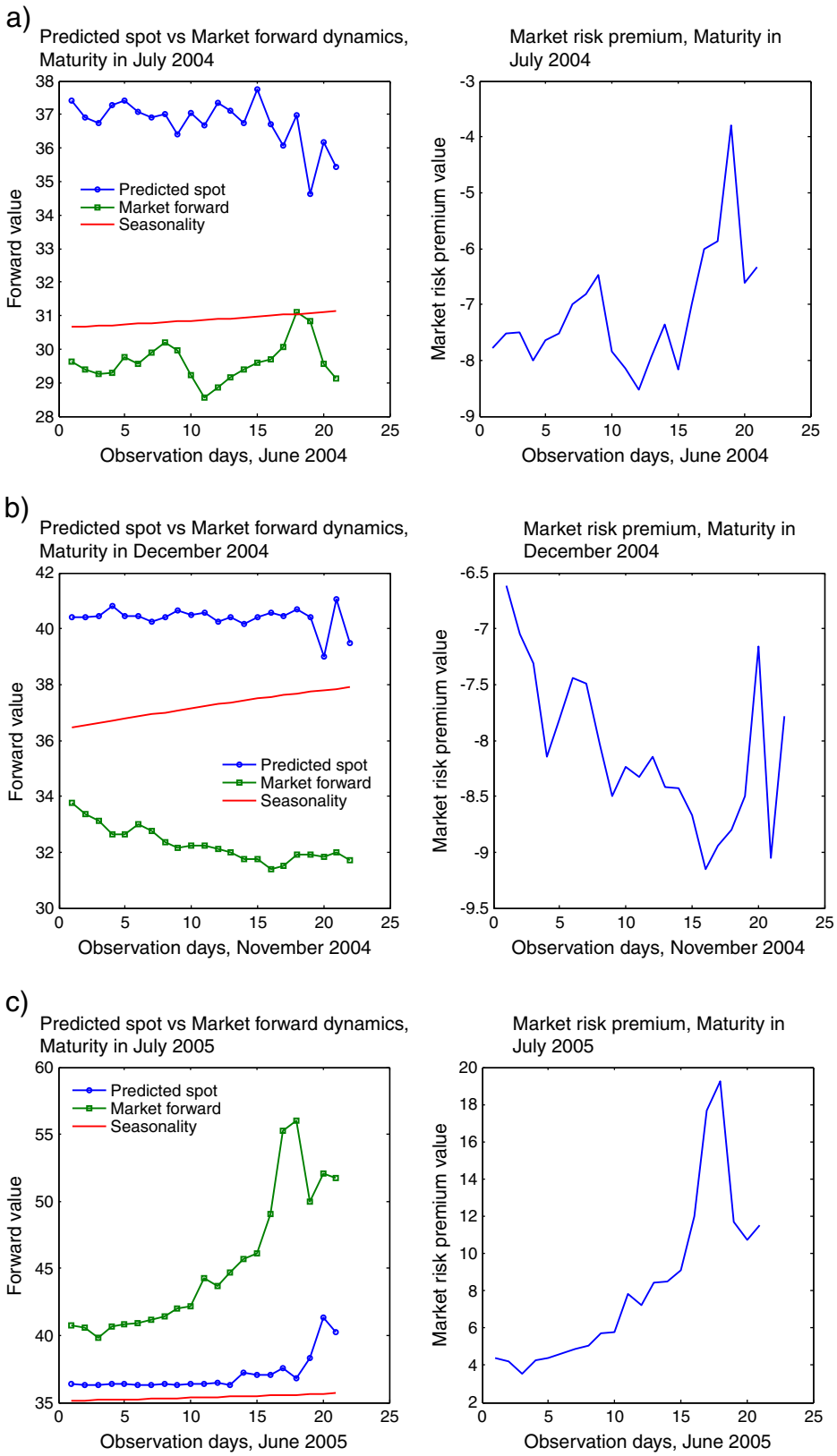


Fig. B.22. The predicted spot, observed forward dynamics and market risk premium for the threshold model. (a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004. (b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004. (c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005. (d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005. (e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2006. (f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

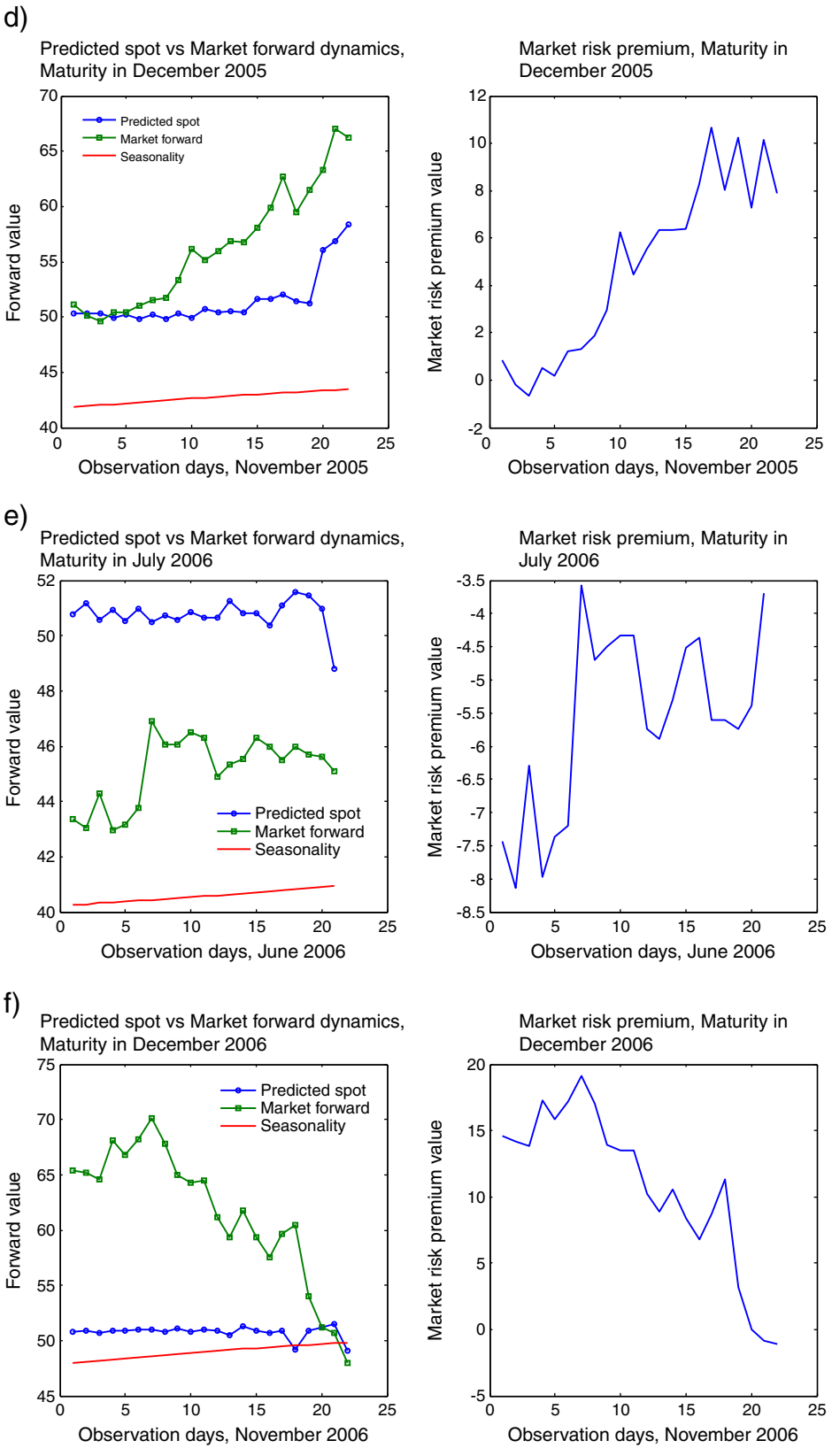


Fig. B.22 (*continued*).

Appendix B.3. Factor model and its forward modelling

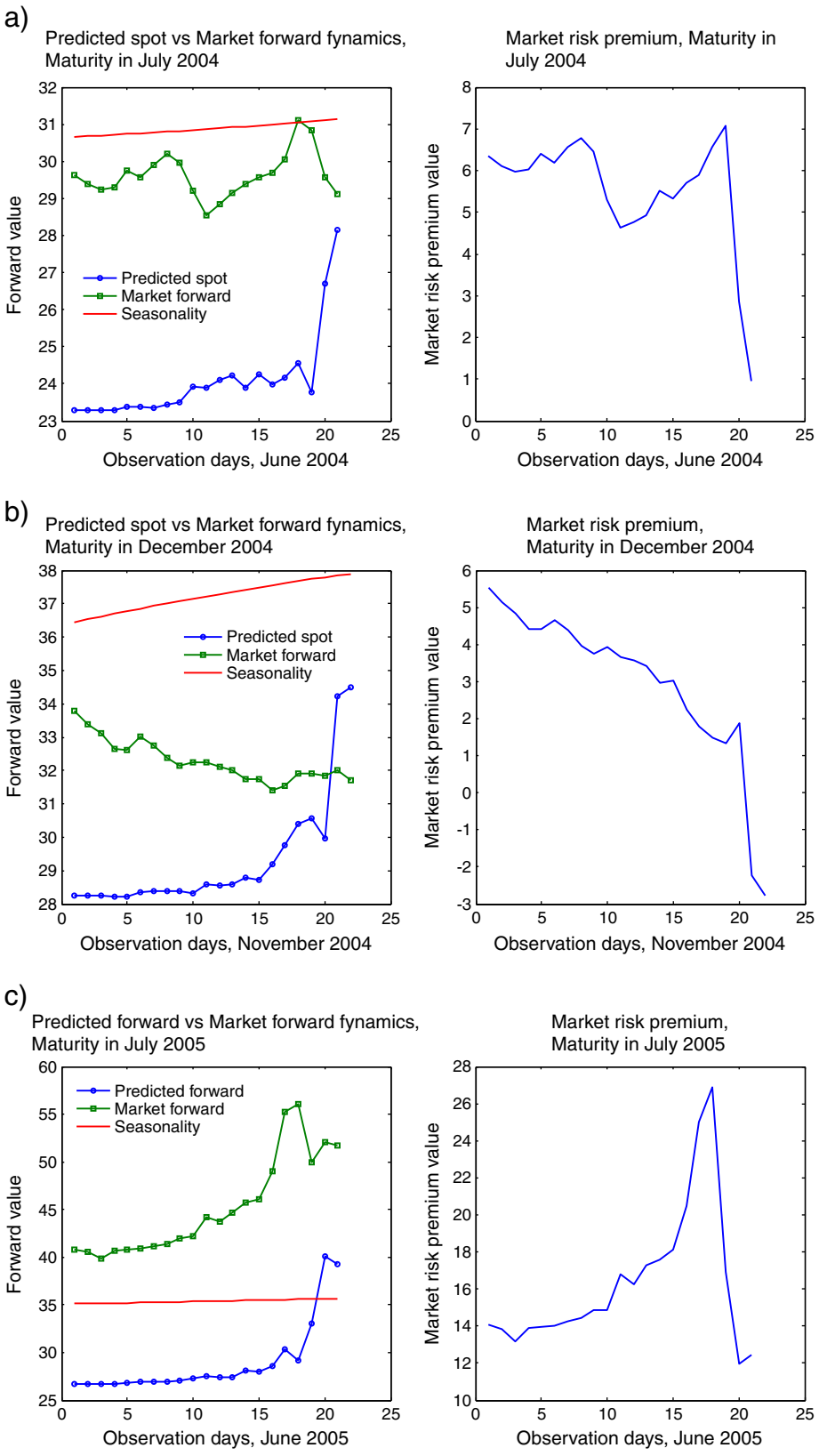


Fig. B.23. The predicted spot, observed forward dynamics and market risk premium for the factor model. (a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004. (b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004. (c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005. (d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005. (e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005. (f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

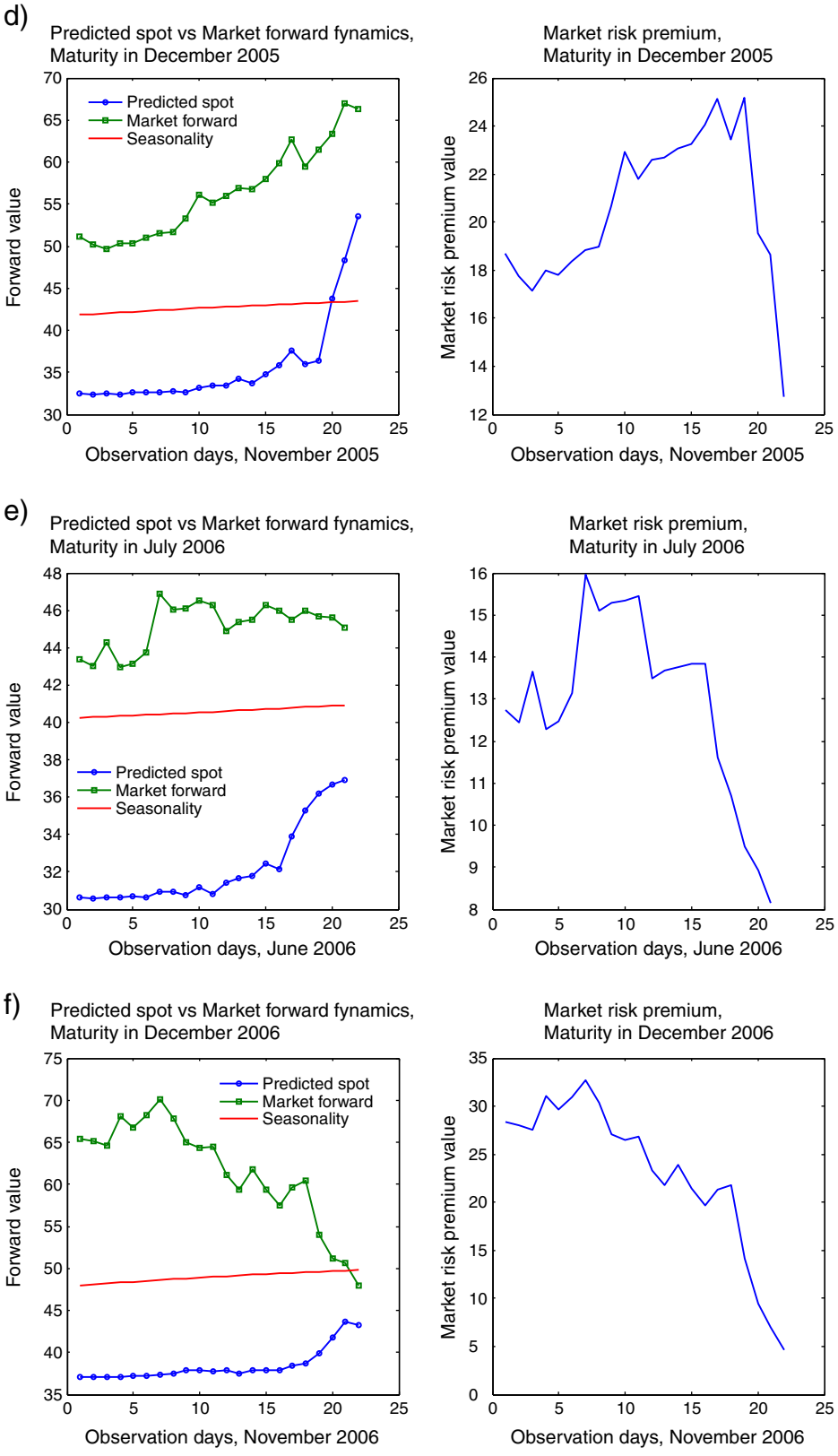


Fig. B.23 (*continued*).

Appendix C. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.eneco.2011.11.012.

References

Albanese, C., Lo, H., Tompaidis, S., 2006. A numerical method for pricing electricity derivatives based on continuous time lattices. Unpublished manuscript, Imperial College London and University of Texas at Austin.

Barndorff-Nielsen, O., Shephard, N., 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in nancial economics. J. R. Stat. Soc. B (Stat. Methodol.) 63 (2), 167–241.
Benth, F., Kufakunesu, R., 2009. Pricing of exotic energy derivatives based on arithmetic spot models. Int. J. Theor. Appl. Finance 12 (4), 491–506.
Benth, F., Kallsen, J., Meyer-Brandis, T., 2007. A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing. Appl. Math. Finance 14 (2), 153–169.
Benth, F., Cartea, Á., Kiesel, R., 2008a. Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium. J. Bank. Finance 32 (10), 2006–2021.
Benth, F., Benth, J., Koekebakker, S., 2008b. Stochastic Modelling of Electricity and Related Markets. World Scienti c Pub Co Inc.
Bibby, B., Jacobsen, M., Sørensen, M., 2010. Estimating functions for discretely sampled diffusion-type models. In: Aït-Sahalia, Y., Hansen, L.P. (Eds.), Handbook of Financial Econometrics, 1. North Holland, Oxford, pp. 203–268.
Cartea, Á., Figueroa, M.G., 2005. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. Appl. Math. Finance 12 (4), 313–335.
Clewlow, L., Strickland, C., 2000. Energy Derivatives: Pricing and Risk Management. Lacima, London.
Drees, H., De Haan, L., Resnick, S., 2000. How to make a Hill plot. Ann. Stat. 28 (1), 254–274.
Eydeland, A., Wolyniec, K., 2002. Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging. John Wiley & Sons Inc.
Fusai, G., Roncoroni, A., 2008. Implementing Models in Quantitative Finance: Methods and Cases. Springer Verlag.
Geman, H., Kourouvakalis, S., 2008. A lattice-based method for pricing electricity derivatives under the threshold model. Appl. Math. Finance 15 (6), 531–567.
Geman, H., Roncoroni, A., 2006. Understanding the ne structure of electricity prices. J. Bus. 79 (3), 1225–1261.
Genon-Catalot, V., Jacod, J., 1993. On the estimation of the diffusion coef cient for multi-dimensional diffusion processes. Annales de l'I. H. Poincaré Probabilités et statistiques 29 (1), 119–151.
Glasserman, P., 2004. Monte Carlo Methods in Financial Engineering. Springer Verlag.
Lucia, J., Schwartz, E., 2002. Electricity prices and power derivatives: evidence from the Nordic Power Exchange. Rev. Deriv. Res. 5 (1), 5–50.
Meyer-Brandis, T., Tankov, P., 2008. Multi-factor jump-diffusion models of electricity prices. Int. J. Theor. Appl. Finance (IJTAF) 11 (05), 503–528.
Nazarova, A., 2008. Lévy-Based Electricity Spot Price Modelling. Master thesis, University of Ulm.
Oyebanji, G., 2007. Implementation of the Geman and Roncoroni Threshold Model. Master thesis, University of Zürich.
Roncoroni, A., 2002. Essays in quantitative nance: Modelling and calibration in interest rate and electricity markets. PhD diss., University Paris Dauphine.
Sato, K., 1999. Lévy Processes and In nitely Divisible Distributions. Cambridge University Press.
Schwartz, E., 1997. The stochastic behavior of commodity prices: implications for valuation and hedging. J. Finance 52 (3), 923–973.
Sørensen, M., 2000. Prediction-based estimating functions. Econ. J. 3, 123–147.