

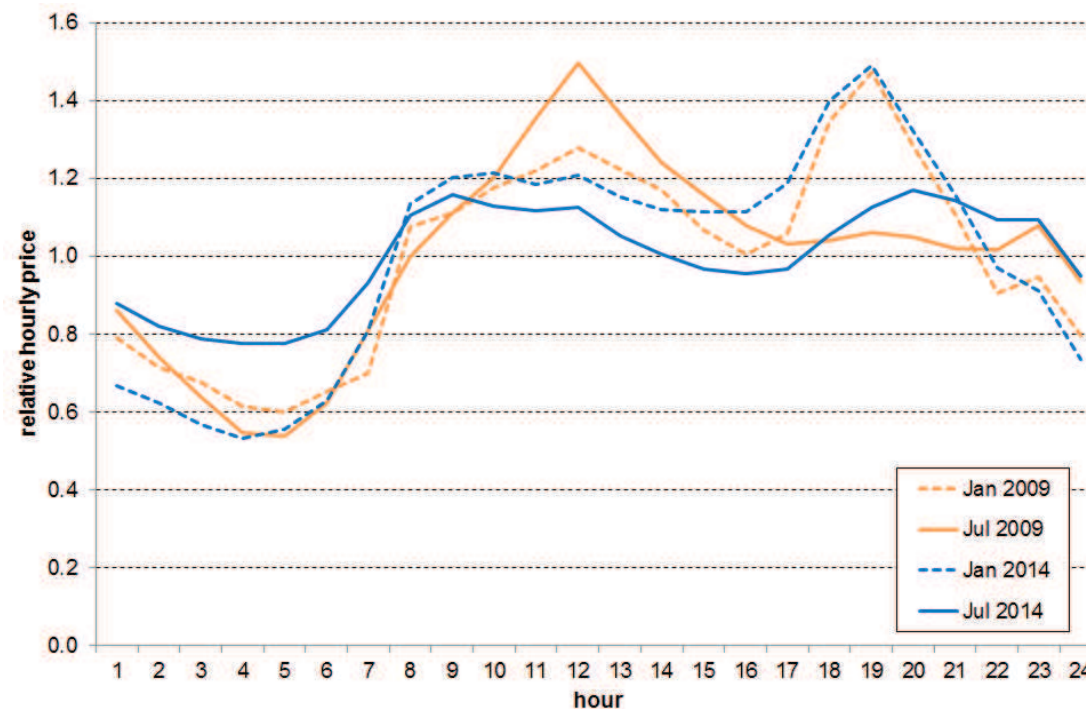
A space-time random field model for electricity forward prices

Florentina Paraschiv, NTNU, UniSG

Fred Espen Benth, University of Oslo

Outlook

- Space-Time random field models for forward electricity prices are highly relevant: major structural changes in the market due to the infeed from renewable energy
- Renewable energies impact the market price expectation – **impact on futures (forward) prices?**
- We will refer to a panel of daily price forward curves derived over time: cross-section analysis with respect to the *time dimension* and the *maturity space*
- Examine and model the dynamics of *risk premia*, the *volatility term structure*, *spatial correlations*



Agenda

- **Modeling assumptions**
- Data: derivation of price forward curves
- Empirical results
- Fine tuning

Literature review

- Models for forward prices in commodity/energy:
 - Specify one model for the spot price and from this derive for forwards: *Lucia and Schwartz (2002)*; *Cartea and Figueroa (2005)*; *Benth, Kallsen, and Mayer-Brandis (2007)*;
 - Heath-Jarrow-Morton approach – price forward prices directly, by multifactor models: *Roncoroni, Guiotto (2000)*; *Benth and Koekebakker (2008)*; *Kiesel, Schindlmayr, and Boerger (2009)*;
- Critical view of *Koekebakker and Ollmar (2005)*, *Frestad (2008)*
 - Few common factors cannot explain the substantial amount of variation in forward prices
 - Non-Gaussian noise
- Random-field models for commodities forward prices:
 - *Audet, Heiskanen, Keppo, & Vehviläinen (2004)*
 - *Benth & Krühner (2014, 2015)*
 - *Benth & Lempa (2014)*
 - *Barndorff-Nielsen, Benth, & Veraart (2015)*

Random Field (RF) versus Affine Term Structure Models (1/2)

Random Field (RF)	Affine term structure models ATSM
Definition	
• RF = a continuum of stochastic processes with drifts, diffusions, cross-section correlations • Smooth functions of maturity • Each forward rate has its own stochastic process • Each instantaneous forward innovation is imperfectly correlated with the others	• ATSM = a limited number of factors is assumed to explain the evolution of the entire forward curve and their dynamics modeled by stochastic processes • All (continuous) finite-factor models are special (degenerate) cases of RF models • ATSM are unable to produce sufficient independent variation in forward rates of similar maturities
Hedging	
• Best hedging instrument of a forward point on the curve is another one of similar maturity • No securities will be left to price by arbitrage	• Any security can be perfectly hedged (instantaneously) by purchasing a portfolio of N additional assets • In 1F model: use only short rates to hedge interest rate risk of the entire portfolio (risk managers estimate the duration of short, medium, long-term separately)

Problem statement

- Previous models model forward prices evolving over time (time-series) along the **time at maturity** T : *Andresen, Koekebakker, and Westgaard (2010)*
- Let $F_t(T)$ denote the forward price at time $t \geq 0$ for delivery of a commodity at time $T \geq t$

- Random field in t :

$$t \mapsto F_t(T), \quad t \geq 0 \quad (1)$$

- Random field in both t and T :

$$(t, T) \mapsto F_t(T), \quad t \geq 0, \quad t \leq T \quad (2)$$

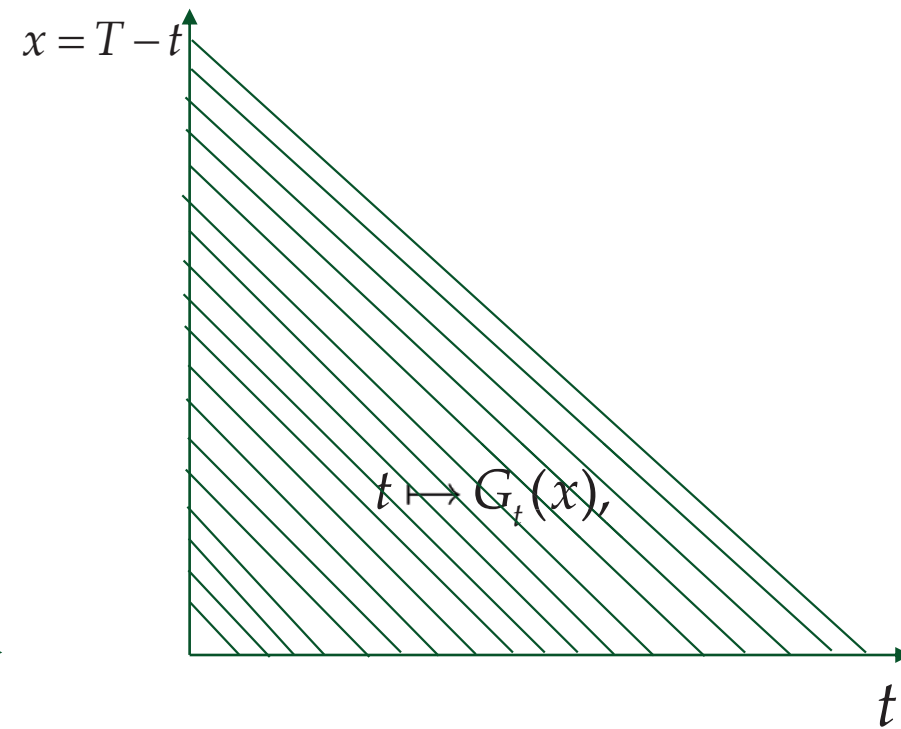
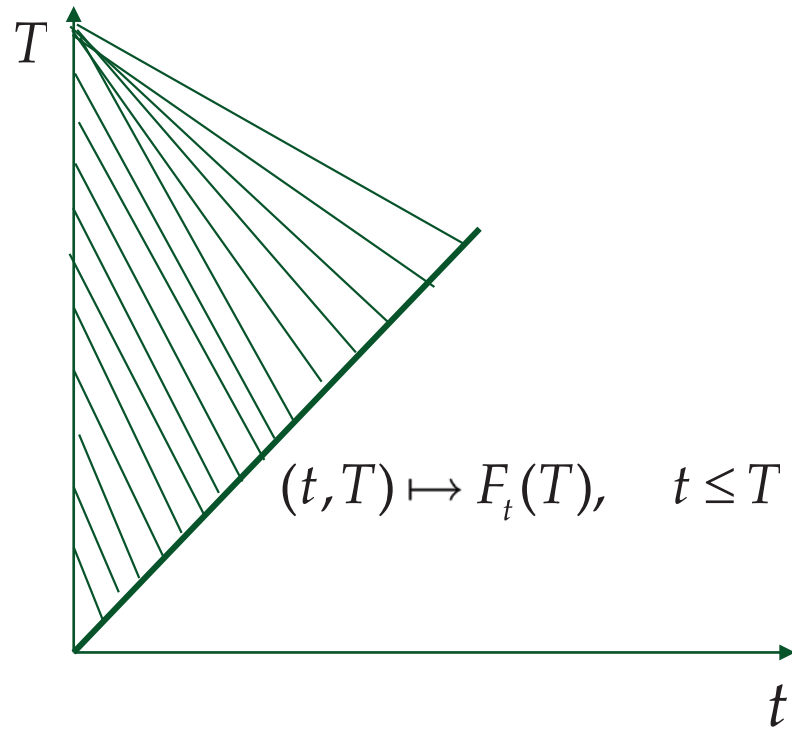
- Get rid of the second condition: **Musiela parametrization** $x = T - t$, $x \geq 0$.

$$F_t(t + x) = F_t(T), \quad t \geq 0 \quad (3)$$

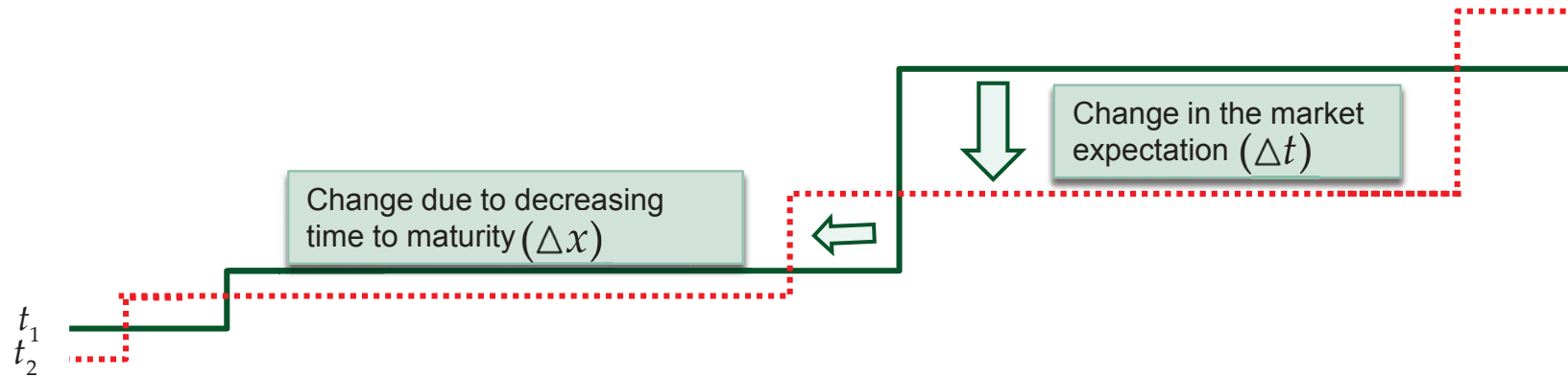
- Let $G_t(x)$ be the forward price for a contract with time to maturity $x \geq 0$. Note that:

$$G_t(x) = F_t(t + x) \quad (4)$$

Graphical interpretation



Influence of the “time to maturity”



Model formulation: Heath-Jarrow-Morton (HJM)

- The stochastic process $t \mapsto G_t(x)$, $t \geq 0$ is the solution to:

$$dG_t(x) = (\partial_x G_t(x) + \beta(t, x)) dt + dW_t(x) \quad (5)$$

- Space of curves are endowed with a Hilbert space structure $\mathcal{H}$
- ∂_x differential operator with respect to time to maturity
- β spatio-temporal random field describing the market price of risk
- W Spatio-temporal random field describing the randomly evolving residuals
- Discrete structure:

$$G_t(x) = f_t(x) + s_t(x), \quad (6)$$

- $s_t(x)$ deterministic seasonality function $\mathbb{R}_+^2 \ni (t, x) \mapsto s_t(x) \in \mathbb{R}$

Model formulation (cont)

We furthermore assume that the deasonalized forward price curve, denoted by $f_t(x)$, has the dynamics:

$$df_t(x) = (\partial_x f_t(x) + \theta(x)f_t(x)) dt + dW_t(x), \quad (7)$$

with $\theta \in \mathbb{R}$ being a constant. With this definition, we note that

$$\begin{aligned} dG_t(x) &= df_t(x) + ds_t(x) \\ &= (\partial_x f_t(x) + \theta(x)f_t(x)) dt + \partial_t s_t(x) dt + dW_t(x) \\ &= (\partial_x G_t(x) + (\partial_t s_t(x) - \partial_x s_t(x)) + \theta(x)(G_t(x) - s_t(x))) dt + dW_t(x). \end{aligned}$$

In the natural case, $\partial_t s_t(x) = \partial_x s_t(x)$, and therefore we see that $G_t(x)$ satisfy (5) with $\beta(t, x) := \theta(x)f_t(x)$.

The market price of risk is proportional to the deseasonalized forward prices.

Model formulation (cont)

We discretize the dynamics in Eq. (34) by an Euler discretization

$$df_t(x) = (\partial_x f_t(x) + \theta(x)f_t(x)) dt + dW_t(x)$$

$$\partial_x f_t(x) \approx \frac{f_t(x + \Delta x) - f_t(x)}{\Delta x}$$

$$f_{t+\Delta t}(x) = (f_t(x) + \frac{\Delta t}{\Delta x}(f_t(x + \Delta x) - f_t(x)) + \theta(x)f_t(x)\Delta t + \epsilon_t(x) \quad (8)$$

with $x \in \{x_1, \dots, x_N\}$ and $t = \Delta t, \dots, (M-1)\Delta t$, where $\epsilon_t(x) := W_{t+\Delta t}(x) - W_t(x)$.

$$Z_t(x) := f_{t+\Delta t}(x) - f_t(x) - \frac{\Delta t}{\Delta x}(f_t(x + \Delta x) - f_t(x)) \quad (9)$$

which implies

$$Z_t(x) = \theta(x)f_t(x)\Delta t + \epsilon_t(x), \quad (10)$$

$$\epsilon_t(x) = \sigma(x)\tilde{\epsilon}_t(x) \quad (11)$$

Agenda

- Modeling assumptions
- **Data: derivation of price forward curves**
- Empirical results
- Fine tuning

Derivation of price forward curves: seasonality curves

- Data: a unique data set of about 2'386 hourly price forward curves daily derived in the German electricity market (PHELIIX) between 2009–2015 (source ior/cf UniSG)
- We firstly **remove the long-term trend** from the hourly electricity prices
- Follow *Blöchlinger (2008)* for the derivation of the seasonality shape for EPEX power prices: very „data specific”; **removes daily and hourly seasonal effects and autocorrelation!**
- The shape is aligned to the level of futures prices

Factor to year

$$f2y_d = \frac{S^{day}(d)}{\sum_{k \in \text{year}(d)} S^{day}(k) \frac{1}{K(d)}} \quad (12)$$

To explain the $f2y$, we use a multiple regression model:

$$f2y_d = \alpha_0 + \sum_{i=1}^6 b_i D_{di} + \sum_{i=1}^{12} c_i M_{di} + \sum_{i=1}^3 d_i CDD_{di} + \sum_{i=1}^3 e_i HDD_{di} + \varepsilon_d \quad (13)$$

- $f2y_d$: Factor to year, daily-base-price/yearly-base-price
- D_{di} : 6 daily dummy variables (for Mo-Sat)
- M_{di} : 12 monthly dummy variables (for Feb-Dec); August will be subdivided in two parts, due to summer vacation
- CDD_{di} : Cooling degree days for 3 different German cities – $\max(T - 18.3^\circ C, 0)$
- HDD_{di} : Heating degree days for 3 different German cities – $\max(18.3^\circ C - T, 0)$

where CDD_i/HDD_i are estimated based on the temperature in Berlin, Hannover and Munich.

Regression model for the temperature

- For temperature, we propose a forecasting model based on fourier series:

$$T_t = a_0 + \sum_{i=1}^3 b_{1,i} \cos(i \frac{2\pi}{365} Y T_t) + \sum_{i=1}^3 b_{2,i} \sin(i \frac{2\pi}{365} Y T_t) + \varepsilon_t \quad (14)$$

where T_p is the average daily temperature and YT the observation time within one year

- Once the coefficients in the above model are estimated, the temperature can be easily predicted since the only exogenous factor YT is deterministic!
- Forecasts for CDD and HDD are also straightforward

Profile classes for each day

Table 1: *The table indicates the assignment of each day to one out of the twenty profile classes. The daily pattern is held constant for the workdays Monday to Friday within a month, and for Saturday and Sunday, respectively, within three months.*

	J	F	M	A	M	J	J	A	S	O	N	D
Week day	1	2	3	4	5	6	7	8	9	10	11	12
Sat	13	13	14	14	14	15	15	15	16	16	16	13
Sun	17	17	18	18	18	19	19	19	20	20	20	17

Profile classes for each day

- The regression model for each class is built quite similarly to the one for the yearly seasonality. For each profile class $c = \{1, \dots, 20\}$ defined in table 1, a model of the following type is formulated:

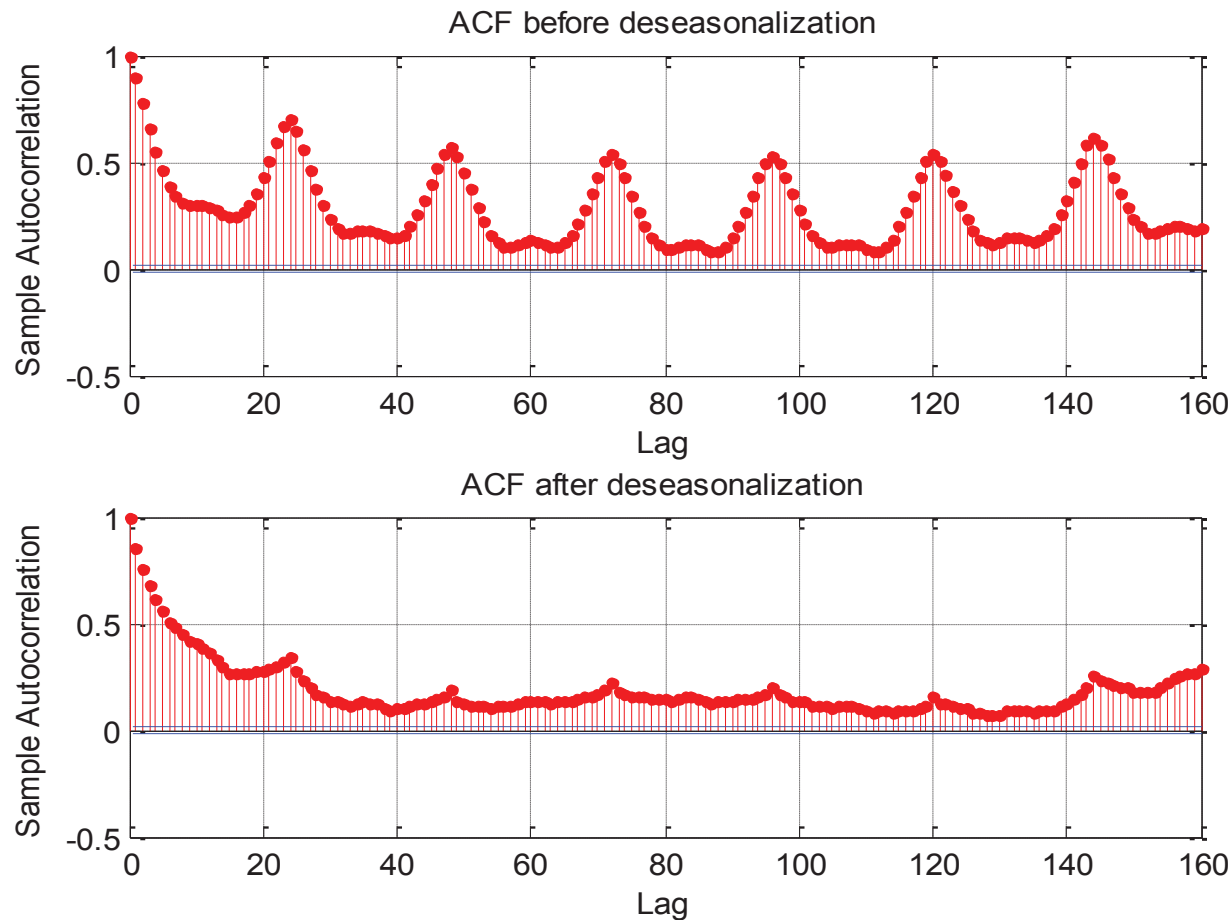
$$f2d_t = a_o^c + \sum_{i=1}^{23} b_i^c H_{t,i} + \varepsilon_t \quad \text{for all } t \in c. \quad (15)$$

where $H_i = \{0, \dots, 23\}$ represents dummy variables for the hours of one day

- **The seasonality shape s_t can be calculated by $s_t = f2y_t \cdot f2d_t$.**
- **s_t is the forecast of the relative hourly weights and it is additionally multiplied by the yearly average prices, in order to align the shape at the prices level**
- **This yields the seasonality shape s_t which is finally used to deseasonalize the electricity prices**

Deseasonalization result

The deseasonalized series is assumed to contain only the stochastic component of electricity prices, such as the volatility and randomly occurring jumps and peaks



Derivation of HPFC *Fleten & Lemming (2003)*

- Recall that $F_t(x)$ is the price of the forward contract with maturity x , where time is measured in hours, and let $F_t(T_1, T_2)$ be the settlement price at time t of a forward contract with delivery in the interval $[T_1, T_2]$.
- The forward prices of the derived curve should match the observed settlement price of the traded future product for the corresponding delivery period, that is:

$$\frac{1}{\sum_{\tau=T_1}^{T_2} \exp(-r\tau/a)} \sum_{\tau=T_1}^{T_2} \exp(-r\tau/a) F_t(\tau - t) = F_t(T_1, T_2) \quad (16)$$

where r is the continuously compounded rate for discounting per annum and a is the number of hours per year.

- A realistic price forward curve should capture information about the hourly seasonality pattern of electricity prices

$$\min \left[\sum_{x=1}^N (F_t(x) - s_t(x))^2 \right] \quad (17)$$

Agenda

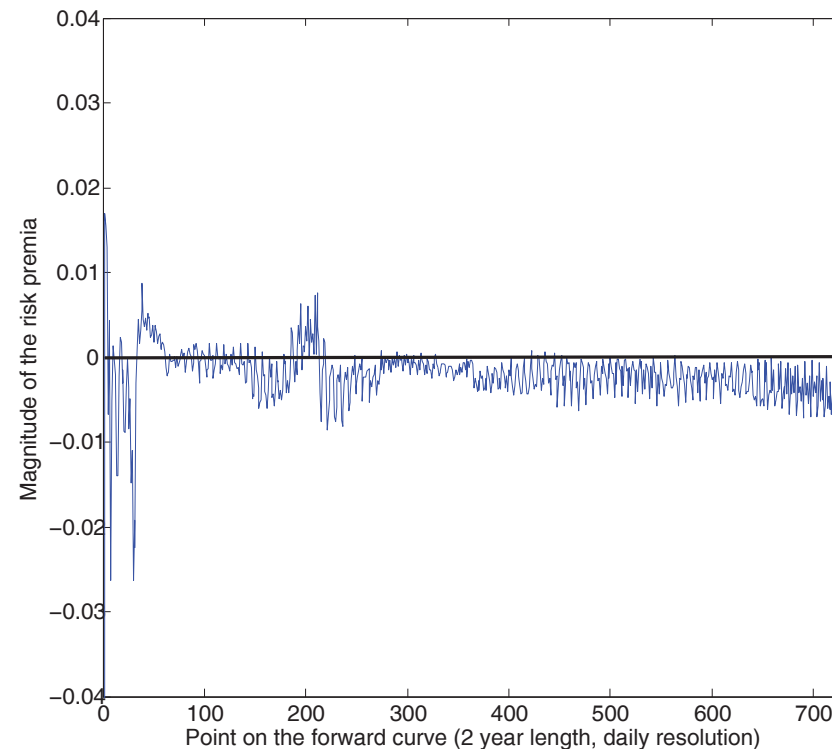
- Modeling assumptions
- Data: derivation of price forward curves
- **Empirical results**
- Fine tuning

Risk premia

$$Z_t(x) = \theta(x)f_t(x)\Delta t + \epsilon_t(x)$$

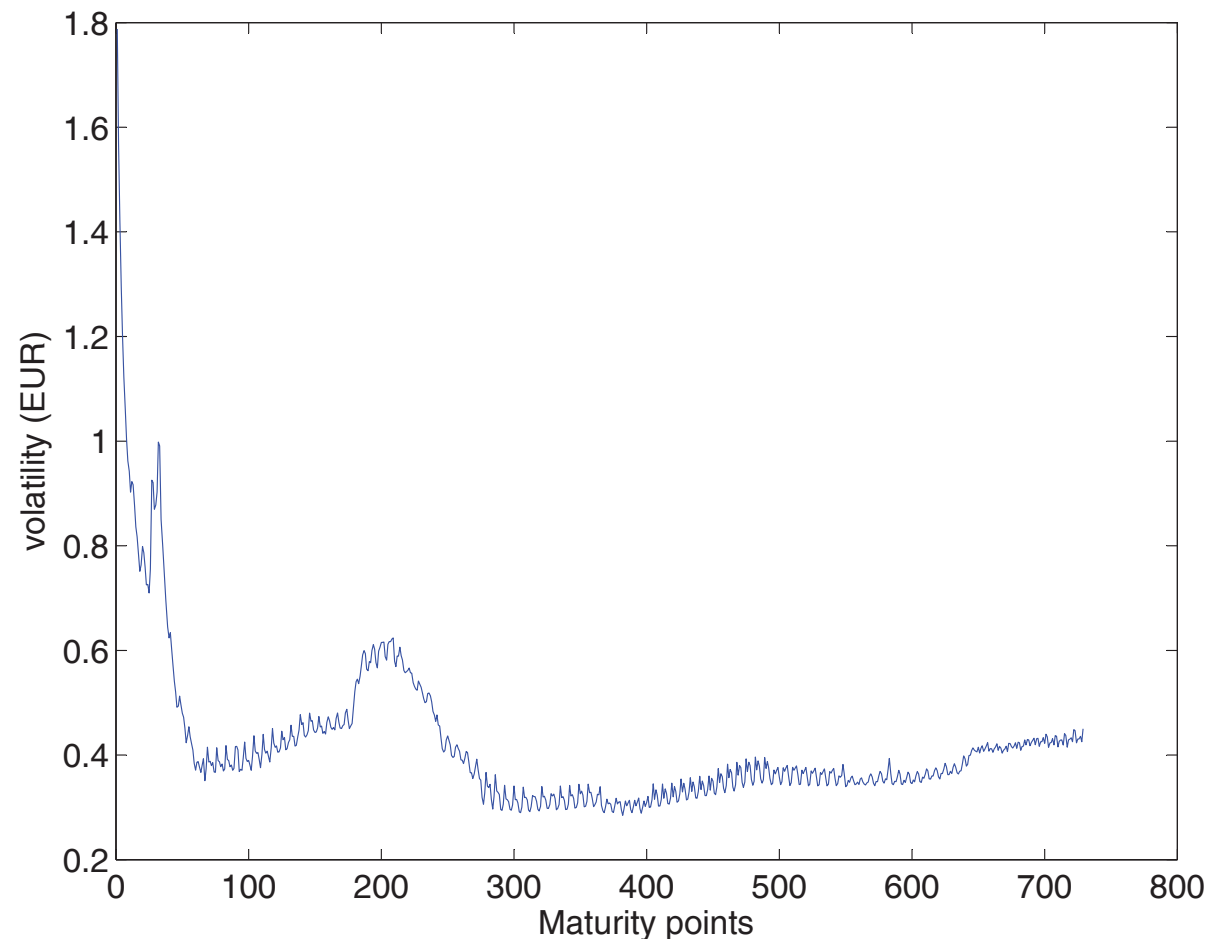
$$\epsilon_t(x) = \sigma(x)\tilde{\epsilon}_t(x)$$

- *Short-term:* It oscillates around zero and has higher volatility (similar in *Pietz (2009)*, *Paraschiv et al. (2015)*)
- *Long-term:* In the long-run power generators accept lower futures prices, as they need to make sure that their investment costs are covered (*Burger et al. (2007)*).



Term structure volatility

- We observe Samuelson effect: overall higher volatility for shorter time to maturity
- Volatility bumps (front month; second/third quarters) explained by increased volume of trades
- Jigsaw pattern: weekend effect; volatility smaller in the weekend versus working days



Explaining volatility bumps

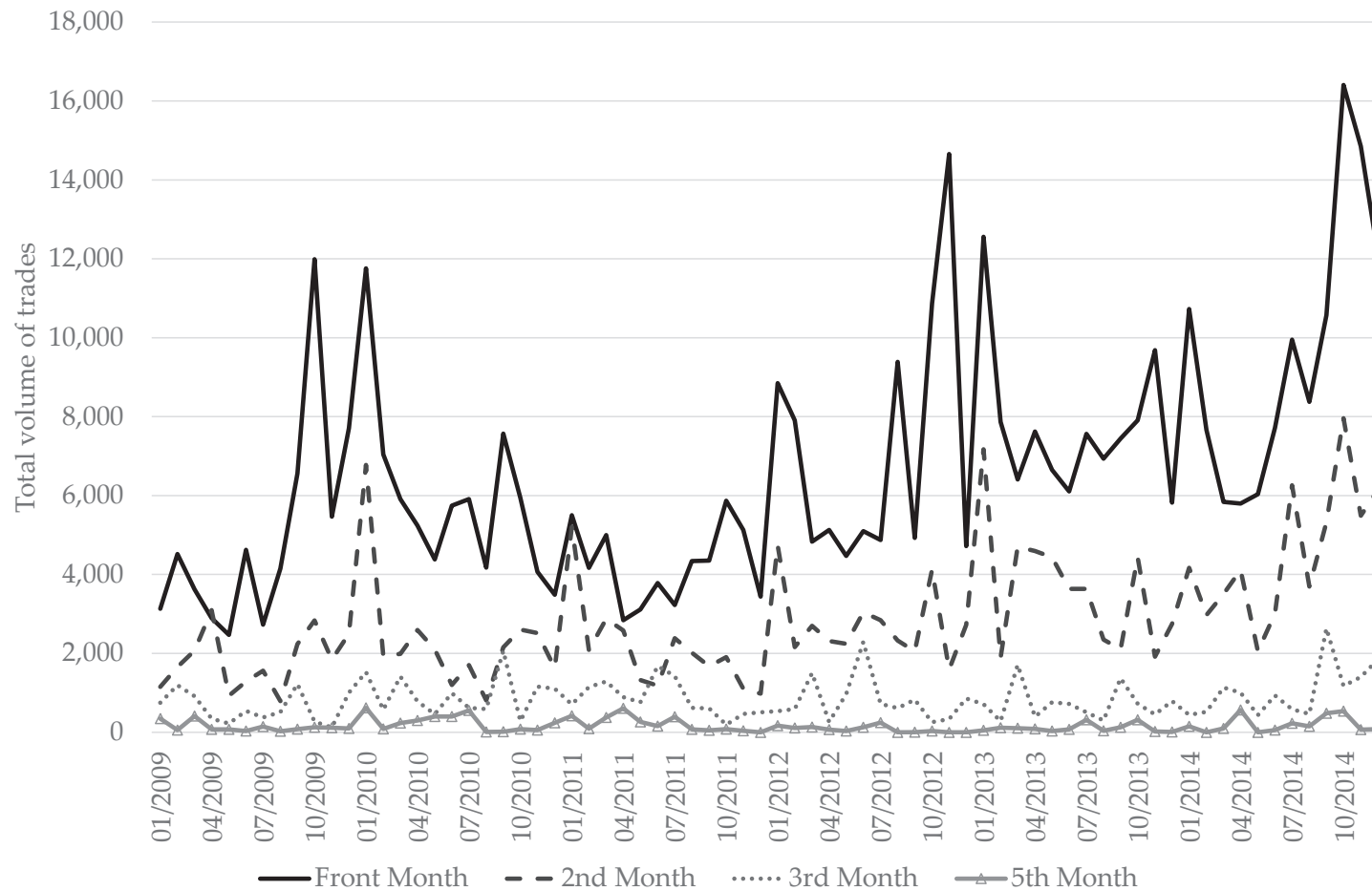


Figure 2: The sum of traded contracts for the monthly futures, evidence from EPEX, own calculations (source of data ems.eex.com).

Explaining volatility bumps

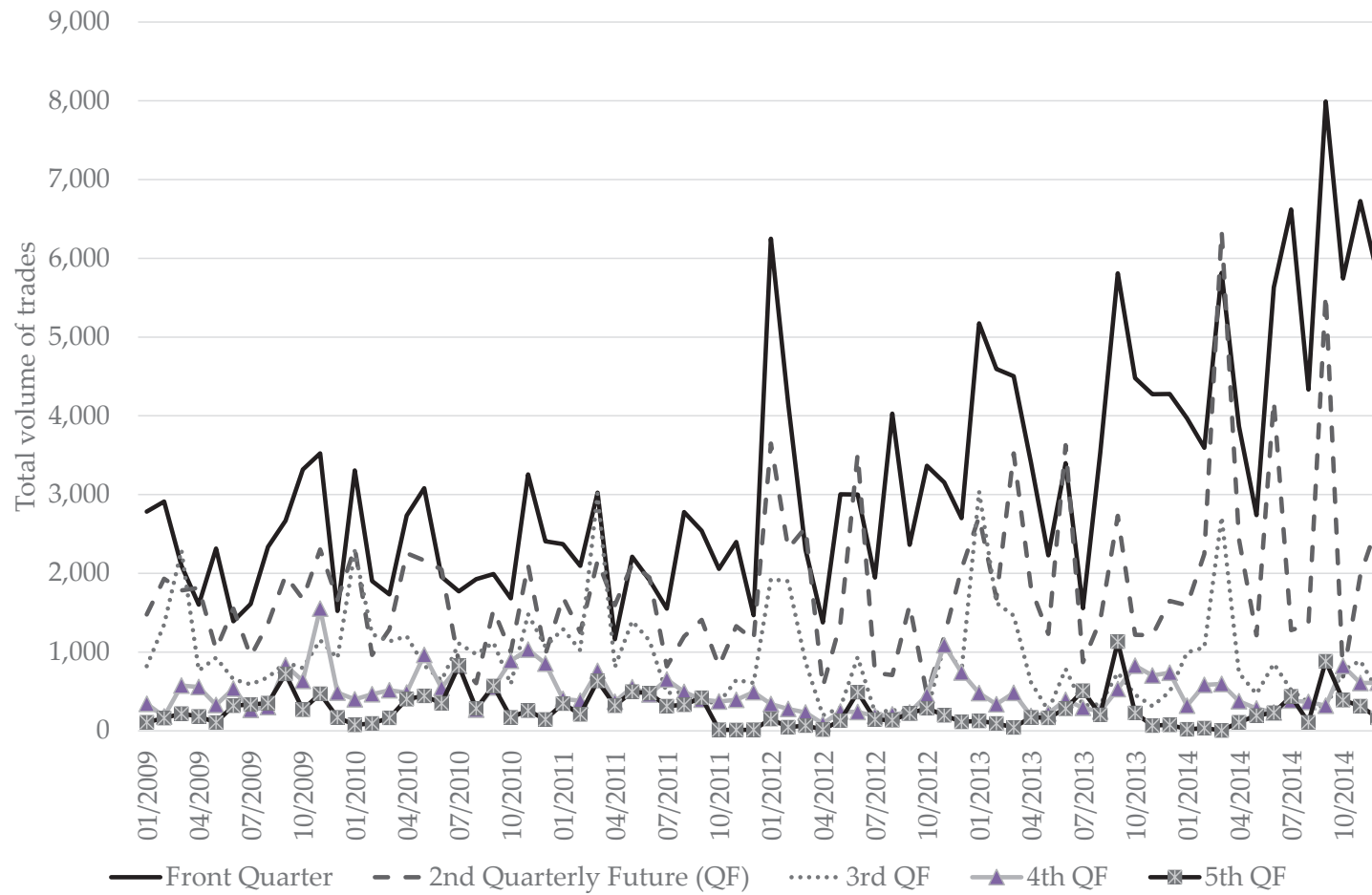


Figure 3: The sum of traded contracts for the quarterly futures, evidence from EPEX, own calculations (source of data ems.eex.com).

Statistical properties of the noise

- We examined the statistical properties of the noise time-series $\tilde{\epsilon}_t(x)$

$$\epsilon_t(x) = \sigma(x)\tilde{\epsilon}_t(x) \quad (18)$$

- We found: Overall we conclude that the model residuals are **coloured noise**, with **heavy tails** (leptokurtic distribution) and with a tendency for **conditional volatility**.

$\tilde{\epsilon}_t(x_k)$	Stationarity	Autocorrelation $\tilde{\epsilon}_t(x_k)$	Autocorrelation $\tilde{\epsilon}_t(x_k)^2$	ARCH/GARCH
	h	h1	h1	h2
Q0	0	1	1	1
Q1	0	0	0	0
Q2	0	1	1	1
Q3	0	0	1	1
Q4	0	1	0	0
Q5	0	1	1	1
Q6	0	1	1	1
Q7	0	1	1	1

Table 2: The time series are selected by quarterly increments (90 days) along the maturity points on one noise curve. Hypotheses tests results, case study 1: $\Delta x = 1day$. In column stationarity, if $h = 0$ we fail to reject the null that series are stationary. For autocorrelation $h1 = 0$ indicates that there is not enough evidence to suggest that noise time series are autocorrelated. In the last column $h2 = 1$ indicates that there are significant ARCH effects in the noise time-series.

Autocorrelation structure of noise time series (squared)

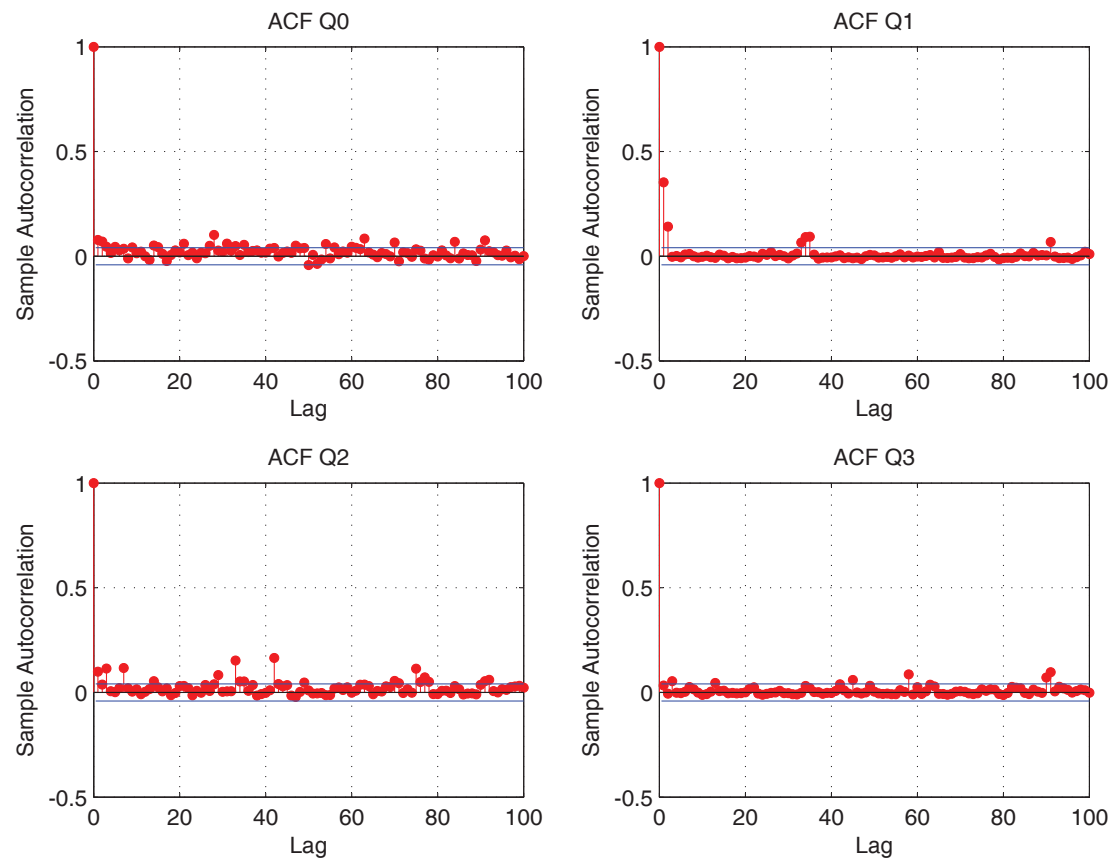
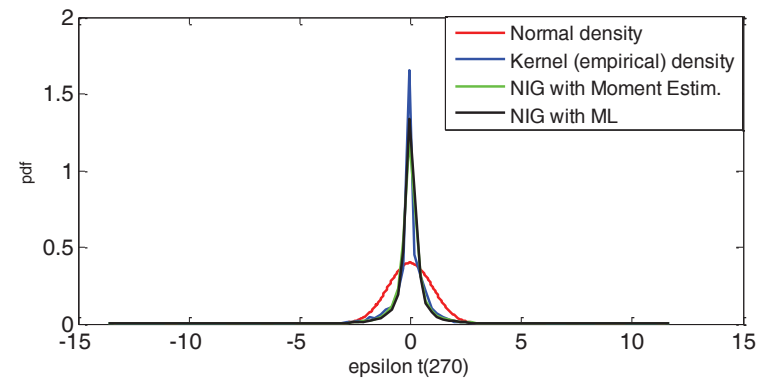
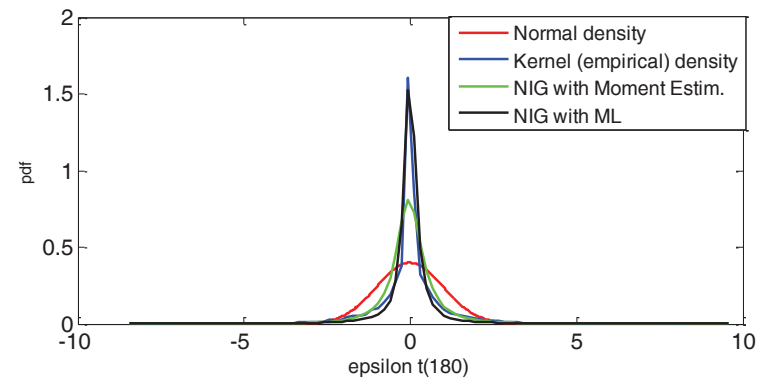
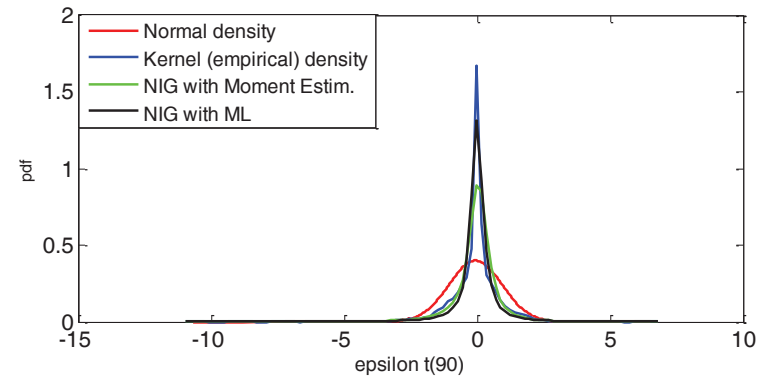
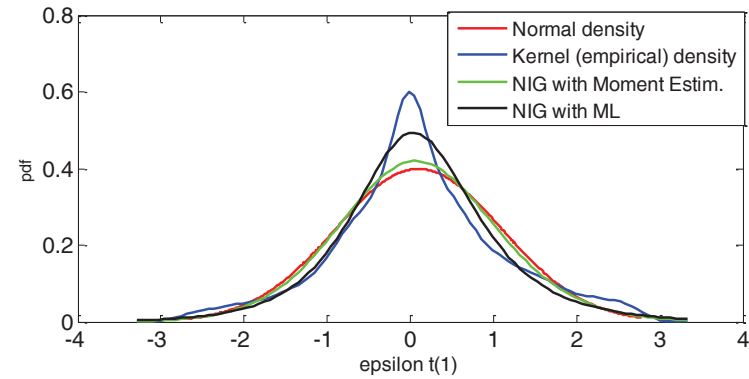


Figure 4: Autocorrelation function in the squared time series of the noise $\tilde{\epsilon}_t(x_k)^2$, by taking $k \in \{1, 90, 180, 270\}$, case study 1: $\Delta x = 1 \text{ day}$.

Normal Inverse Gaussian (NIG) distribution for coloured noise



Spatial dependence structure

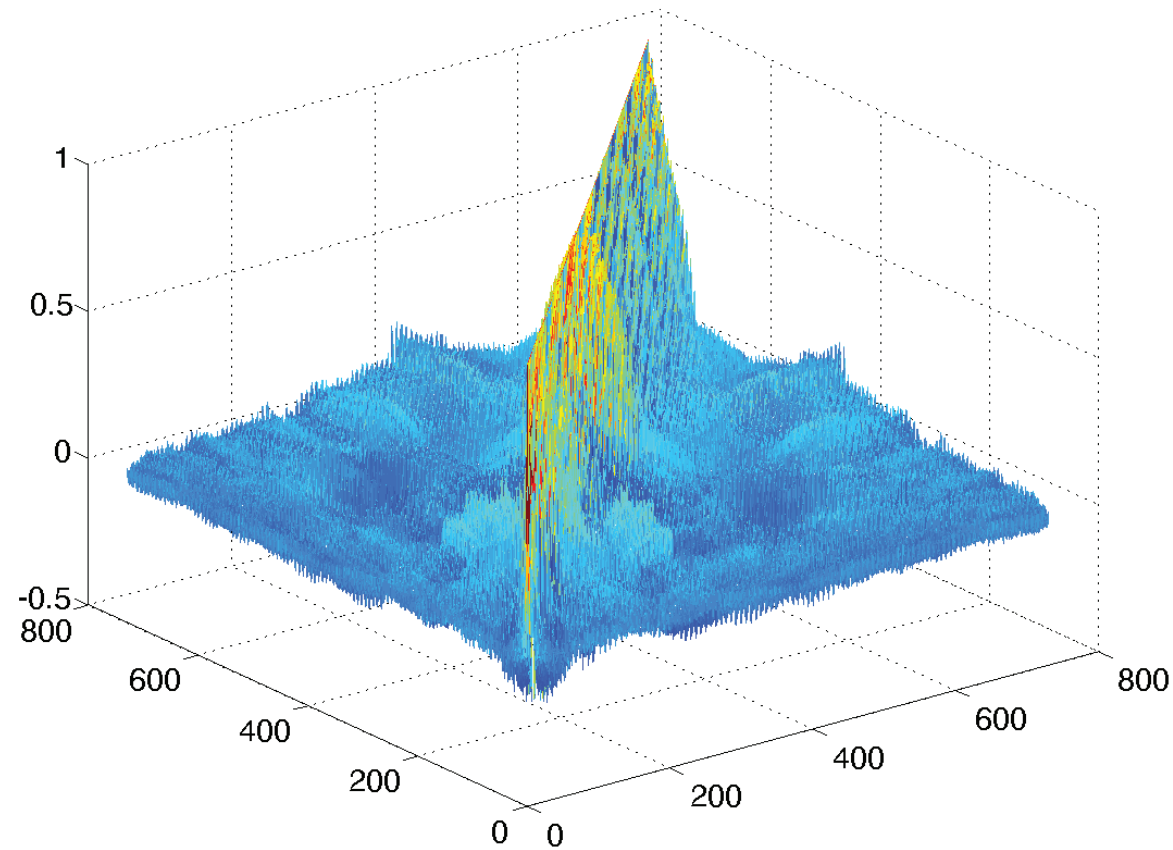


Figure 5: *Correlation matrix with respect to different maturity points along one curve.*

Agenda

- Modeling assumptions
- Data: derivation of price forward curves
- Empirical results
- **Fine tuning**

Revisiting the model

- We have analysed empirically the noise residual $dW_t(x)$ expressed as $\epsilon_t(x) = \sigma(x)\tilde{\epsilon}_t(x)$ in a discrete form
- Recover an infinite dimensional model for $W_t(x)$ based on our findings

$$W_t = \int_0^t \Sigma_s dL_s, \quad (19)$$

where $s \mapsto \Sigma_s$ is an $L(\mathcal{U}, \mathcal{H})$ -valued predictable process and L is a $\mathcal{U}$ -valued Lévy process with zero mean and finite variance.

- As a first case, we can choose $\Sigma_s \equiv \Psi$ time-independent:

$$W_{t+\Delta t} - W_t \approx \Psi(L_{t+\Delta t} - L_t) \quad (20)$$

Choose now $\mathcal{U} = L^2(\mathbb{R})$, the space of square integrable functions on the real line equipped with the Lebesgue measure, and assume Ψ is an integral operator on $L^2(\mathbb{R})$

$$\mathbb{R}_+ \ni x \mapsto \Psi(g)(x) = \int_{\mathbb{R}} \tilde{\sigma}(x, y) g(y) dy \quad (21)$$

- we can further make the approximation $\Psi(g)(x) \approx \tilde{\sigma}(x, x)g(x)$, which gives

$$W_{t+\Delta t}(x) - W_t(x) \approx \tilde{\sigma}(x, x)(L_{t+\Delta t}(x) - L_t(x)). \quad (22)$$

Revisiting the model (cont)

- Recall the spatial correlation structure of $\tilde{\epsilon}_t(x)$. This provides the empirical foundation for defining a covariance functional $\mathcal{Q}$ associated with the Lévy process L .
- In general, we know that for any $g, h \in L^2(\mathbb{R})$,

$$\mathbb{E}[(L_t, g)_2 (L_t, h)_2] = (\mathcal{Q}g, h)_2$$

where $(\cdot, \cdot)_2$ denotes the inner product in $L^2(\mathbb{R})$

$$\mathcal{Q}g(x) = \int_{\mathbb{R}} q(x, y)g(y) dy, \tag{23}$$

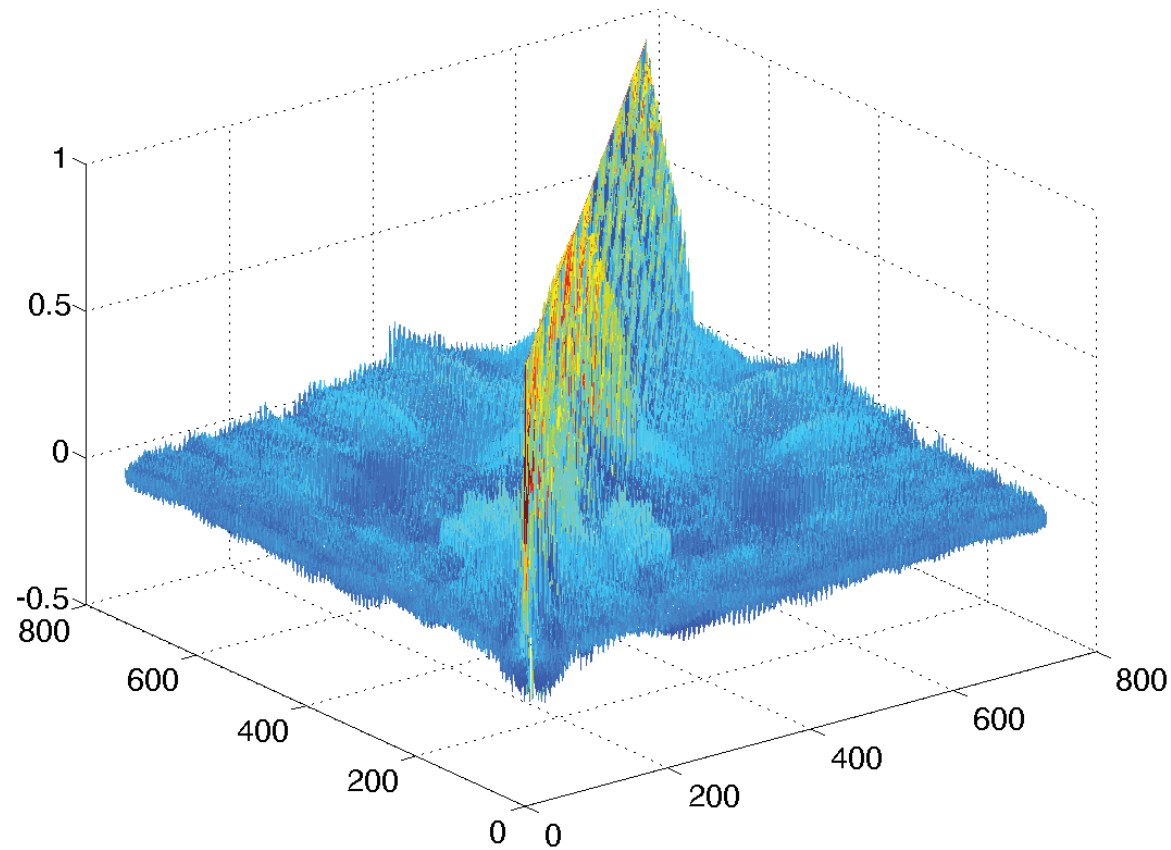
- If we assume $g \in L^2(\mathbb{R})$ to be close to δ_x , the Dirac δ -function, and likewise, $h \in L^2(\mathbb{R})$ being close to δ_y , $(x, y) \in \mathbb{R}^2$, we find approximately

$$\mathbb{E}[L_t(x)L_t(y)] = q(x, y)$$

- A simple choice resembling to some degree the fast decaying property is $q(|x - y|) = \exp(-\gamma|x - y|)$ for a constant $\gamma > 0$.
- It follows that $t \mapsto (L_t, g)_2$ is a NIG Lévy process on the real line.

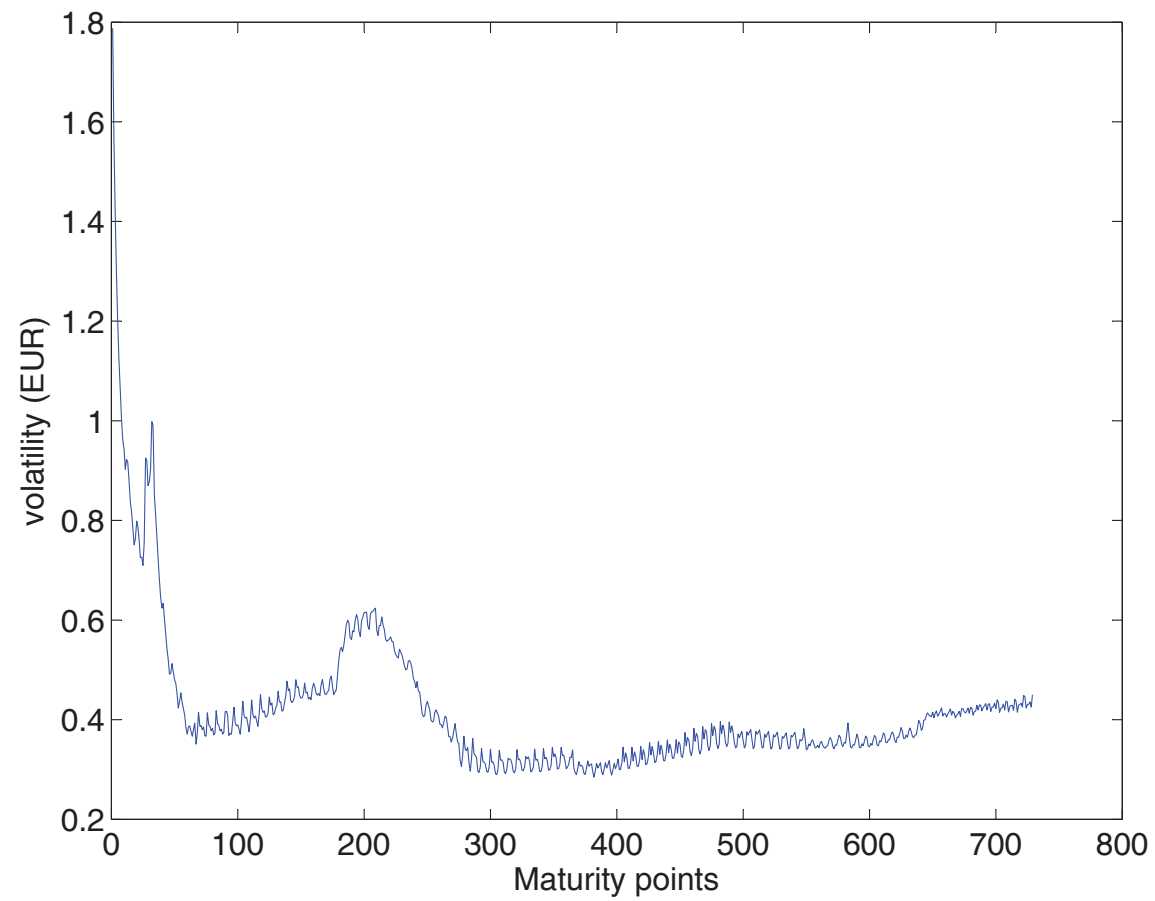
Spatial dependence structure

$$q(x, y) = \exp(-\gamma|x - y|)$$



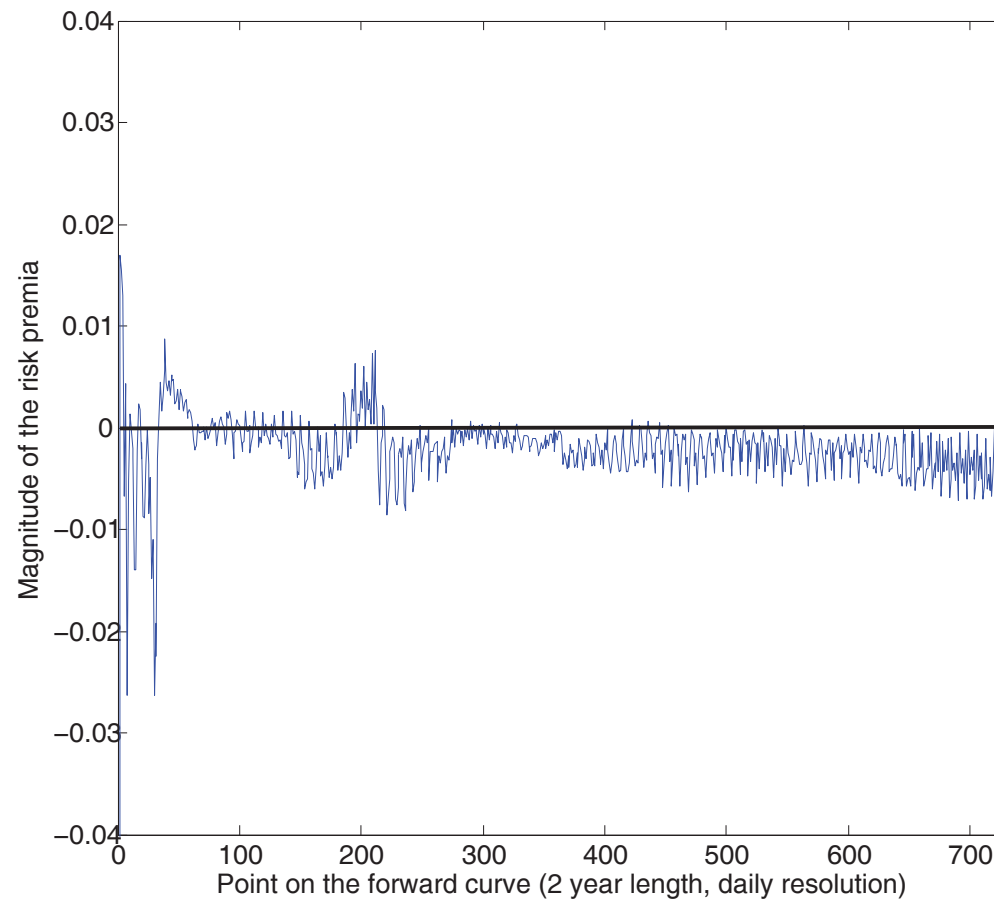
Term structure volatility

$$\tilde{\sigma}(x) = a \exp(-\zeta x) + b$$



Parameter of the market price of risk

$$df(t, x) = (\partial_x f(t, x) + \theta(x)f(t, x)) dt + dW(t, x) ,$$



Conclusion

- We developed a spatio-temporal dynamical arbitrage free model for electricity forward prices based on the Heath-Jarrow-Morton (HJM) approach under Musiela parametrization
- We examined a unique data set of price forward curves derived each day in the market between 2009–2015
- We examined the spatio-temporal structure of our data set
 - **Risk premia:** higher volatility short-term, oscillating around zero; constant volatility on the long-term, turning into negative
 - **Term structure volatility:** Samuelson effect, volatility bumps explained by increased volume of trades
 - **Coloured (leptokurtic) noise** with evidence of conditional volatility
 - **Spatial correlations structure:** decaying fast for short-term maturities; constant (white noise) afterwards with a bump around 1 year
- Advantages over affine term structure models:
 - More data specific, no synthetic assumptions
 - Low number of parameters (easier calibration, suitable for derivative pricing)
 - No recalibration needed