An Empirical Study of the Information Premium on Electricity Markets

Fred Espen Benth[a,1], Richard Biegler-König[b,2], Rüdiger Kiesel[b,a,3,*]

[a]*Centre of Mathematics for Applications, University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway*
[b]*Chair for Energy Trading and Finance, University Duisburg-Essen, Universitätsstr. 12, 45141 Essen, Germany*

Abstract

Due to the non-storability of electricity and the resulting lack of arbitrage-based arguments to price electricity forward contracts, these exhibit a significant time-varying risk premium. Using EEX data during the introduction of Emission certificates and the German "Atom Moratorium" we show that a significant part of the risk premium in electricity forwards is due to different information sets in spot and forward markets. In order to show the existence of the resulting information premium and to analyse its size we design an empirical method based on techniques relating to enlargement of filtrations and the structure of Hilbert spaces.
JEL-Classification: C19, G13, G14, Q40.

Keywords: Electricity markets, risk premium, information premium, spot-forward relationship, enlargement of filtrations, Gaussian and non-Gaussian Ornstein-Uhlenbeck processes, Hilbert space representation.

1. Introduction

The electricity market is special – it features a homogeneous good with prices driven by the technical restriction of the merit order (the sequence in which power plants are used). The price process shows seasonalities, mean reversion and spikes, all of which make stochastic modelling challenging. But the most striking distinction to most other commodities (and financial assets) is the non-storability of electricity. It has to be used when produced. Hence the relation between forward and spot prices must be driven by risk premia only and can not be explained by standard no-arbitrage arguments, storage or a convenience yield. In fact, the relationship must exhibit the structure of an intertemporal risk premium in a pure way. This risk premium has been analysed in terms of hedging needs of the various actors of the market such as producers and retailers.

The risk premium is defined as the difference between the observed forward price and the expected spot price. In this paper we will follow an information-based approach brought forward recently in Benth and Meyer-Brandis (2009) to explain the risk premium. We will design a statistical test to analyse the information sets used by market participants, and we will show that the market risk premium can be explained in terms of these different information sets.

For electricity markets empirical research has shown a rather inconclusive and random behaviour of the risk premium. For example Longstaff and Wang (2004) prove that the risk premium exists and is significant and positive on average for high-frequency data of the *PJM* (the Pennsylvania-New-Jersey-Maryland) market. They also find that the risk premium is correlated negatively with price volatility and positively with spot skewness (as suggested by Bessembinder and Lemmon (2002) and discussed in the next paragraph). Furthermore, Torró and Lucia (2011) examine short-term futures (with a delivery period of one week) traded on the Scandinavian NordPool. They, too, find a statistically positive premium that depends particularly on the season during which the contract matures, being highest in winter and zero in summer. For the Spanish electricity market and forwards with maturity within two months Furiò and Meneu (2010) find that the risk premium decreases with unexpected variations in demand but increases in unexpected variations of the level of hydroelectric energy capacity. Moreover, Diko et al. (2006) find a term structure for the risk premium for data from the German, French and Dutch markets that features a change of sign and negative values for large time to maturity. Their results are sim-

ilar to Kolos and Ronn (2008) who use EEX and PJM data and include oil and gas as more mature markets for comparison. A link between gas storage and electricity forwards is established in Douglas and Popova (2008) in terms of the moments of the electricity spot price distribution confirming the analysis of Bessembinder and Lemmon (2002).

It is noticeable that incomprehension prevails as to the true character of the risk premium. Bessembinder and Lemmon (2002) present a very influential one-period model in which the risk premium depends on the variance and the skewness of the spot price. Their model features retailers' demand as the only exogenous variable and they deduce their risk premium by applying market clearing and equilibrium arguments. Benth et al. (2008a) try to explain the term structure of the risk premium by taking market power and risk aversion of retailers and producers into consideration. They succeed in explaining the change of sign and the negative premium for large time to maturity as mentioned above.

However, as stated above, the important intrinsic property of electricity is that it is non-storable. To illustrate the effect on prices consider the announcement of a power plant to be closed down for, say, maintenance reasons. Obviously, this will result in higher forward prices with delivery over the time of the shut down. Still, this information will have no effect on current spot prices as no arbitrage possibilities arise, i.e. we cannot buy the underlying now and sell in the future.

This reasoning was the motivation for the recent paper by Benth and Meyer-Brandis (2009). In this paper the authors complement the historical filtration as generated by the spot process with additional future information. This is done by means of the theory of the initial enlargement of filtration (French "Grossissement initial de filtration"). This theory was developed mainly by French mathematicians in the 1970s and 1980s, for example in Jeulin (1979) or Jeulin and Yor (1978).

Benth and Meyer-Brandis introduce the *Information Premium* as the difference between the forward price under the finer (market) filtration and the coarser (historical) filtration and find analytical expressions for a well-known two-factor arithmetic spot price model. This model was introduced in Benth et al. (2007) and is widely used (for example in Meyer-Brandis and Tankov (2008) and Benth et al. (2008a)); we will introduce it in Section 2. We remark that a related approach is followed in Cartea et al. (2009). There, the authors suggest a spot model which takes specific forward looking information (in this case capacity and demand forecasts) into account.

Still, their emphasis lies on simulating the spot rather than pricing forward contracts.

In this paper, we will present a method to test for the existence of the information premium empirically. This turns out to be non-trivial as the premium is not measurable with respect to the historical filtration (as will be explained in Section 2.1). Thus, the usual approach, a mere measure change is not possible here. Instead, we will propose a method involving regressions and Hilbert-space representations. The method will also provide a time series for the information premium whose features will match our economic intuition in size and shape. The approach is generally applicable for testing for differences in information sets in any financial market.

In particular, we will analyse two market situations, both of which, we claim, exhibit significant information premia.

Firstly, we will examine the beginning of 2008 with the introduction of Emission certificates. After the first, rather non-committal phase of the European Union Emission Trading Scheme (EUETS), the stricter second phase commenced on 01/01/2008. The market anticipated a general upwards shift in prices and a large information premium can clearly be seen in the prices observed. Indeed, in Figure 1 EEX prices observed on 1 October of the year 2006 and 2007 respectively are illustrated where their delivery period is represented as the length of the horizontal line denoting the price. On the left hand side one can nicely see the typical shape of prices in winter: lower values for October and April, a peak in January and February with slightly lower prices in December due to bank holidays.
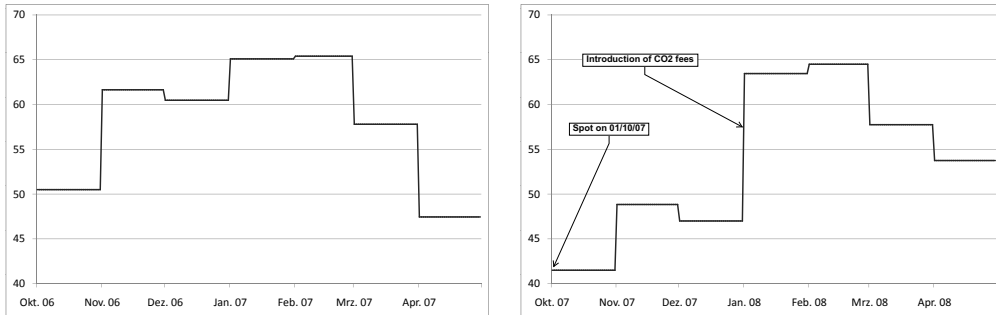

Figure 1: **Monthly forward prices observed on October 1st 2006 (left) and October 1st 2007 (right).** Lengths of horizontal lines denote the corresponding delivery periods.

One faces a different situation in the subsequent year. The most striking

feature here is the price increase between the December and the January forward of more than 16 Euro (corresponding to some 34%, compared to an increase of 4.50 Euro (7,5%) in 2006). Generally, there appears to be a shift upwards (of about 10 Euro) as the remaining price variations are very similar to what was observed the year before. The spot price [4] on that day was around 45 Euro as indicated in the graphic. Clearly, the price increase can be explained by the market's anticipation of the effects due to the introduction of the second phase of Emission certificates. The costs of these certificates were obviously assumed to cause a major rise in electricity prices. In contrast, the spot price did not anticipate the price increase, it is well below and at the same level as the forward for the current month, i.e. October 2007. The reasoning behind this behaviour is the non-storable feature of electricity which prevents simple spot trading strategies.

Secondly, the consequences of the Tōhoku earthquake which occurred on 11 March 2011 will be examined. The consequent tsunami heavily damaged several nuclear power plants, in particular that in Fukushima. Only three days later, 14 March 2011, the German government reevaluated its energy policy and issued the so called "Atom Moratorium" by which the seven oldest plants (eight reactors with a capacity of more than eight GW) were to be shut down immediately for three months. This measure was to allow for a new evaluation of the usage of nuclear power in Germany. Still, spot prices on the EEX hardly increased at all as illustrated in Figure 2. Due to the season there was a lot of spare capacity. Also, two of the plants, Brunsbüttel and Krümmel (both in Schleswig-Holstein), had previously suffered from constant maintenance problems and were offline since 2007 and 2009 respectively. A third reactor (Biblis B in Hessen) had gone offline earlier in 2011 for regular maintenance. This reduced the capacity to be shut down immediately by more than three GW. Hence, because of this and also due to higher wind generation there was no change in the price-setting technology and thus hardly any change in spot prices.

Although the official end of the Moratorium was 15 June 2011 it was widely expected that the seven plants would stay offline even after that date and indeed their permanent shut-down was decided on 31 May 2011. The effect of the Moratorium and this future outlook was a sharp increase in forward prices, not only of those whose delivery fell into the three months of the Moratorium but also of those with a later delivery period. As an example, Figure 3 shows the evolution of the price of the forward with ma-

---

[4]As usually, we take the EEX day-ahead base load price as the spot price.
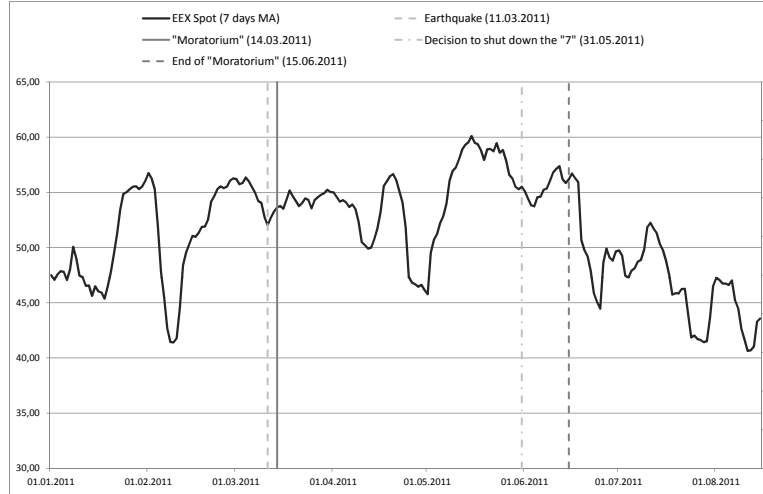
Figure 2: **The EEX spot (seven days moving average).** Key dates in 2011 are indicated. Range of data is 01/01/2011 until 15/08/2011.

turity in May 2011. The forward price had a mean of 46.93 Euro before the Moratorium and a 57.83 Euro post-Moratorium mean price. This corresponds to an increase of more than 10 Euro, i.e. almost 25%. For the second half of this time series we also see that prices remained more or less constant until the last day of its delivery period which implies that by then also the spot had adjusted to the increased price level.

Thus, summarising, one finds that forward prices reacted to some future information (or market sentiment) which was publicly available but the spot did not. For the remainder of this paper an empirical method shall be designed by which one can rigorously show that parts of the forward price cannot be explained by the spot (i.e. the historical filtration) alone.

The paper will be organised as follows: We will provide definitions, some general results and the two-factor arithmetic spot price model in Section 2. Also, in order to calculate prices, we will extend Benth and Meyer-Brandis (2009) to the more realistic case which includes delivery periods for forwards. Furthermore, in Section 3 we will present the method to demonstrate the existence of the information premium whereas the results of the empirical analysis will be summarised in Section 4. Finally, Section 5 will conclude.
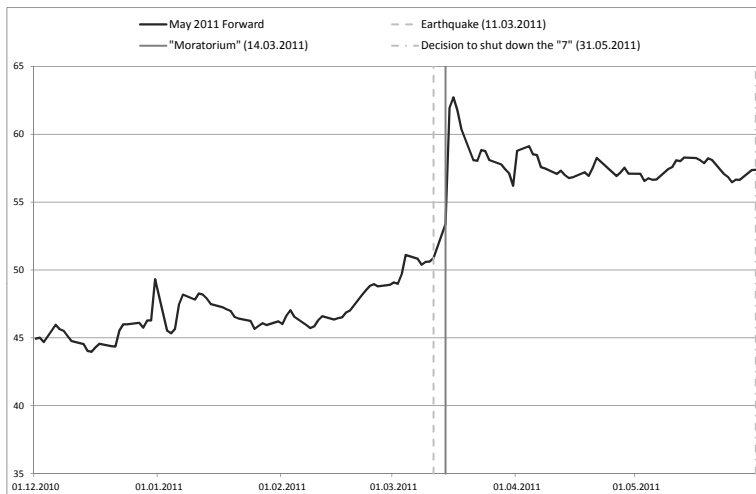
6

Figure 3: **The EEX May 2011 forward price.** Key dates in 2011 are indicated. Range of data is 01/12/2010 until 31/05/2011.

2. Model Setup and Forward Prices

2.1. Definitions and framework

In this section we will provide the necessary definitions and formulae which we will use for the remainder of the paper[5].

Viewing forwards as derivatives of the spot we have the standard pricing relation:

Definition 2.1. *Classical Spot-Forward relationship.* *With $\mathbb{Q}$ a pricing (risk-neutral) measure[6], we have*

$$F(t,T) = \mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t] \tag{1}$$

where $F(t,T)$ denotes the time t-price of a forward maturing at T, S_t is the spot price and $\mathbb{E}^{\mathbb{Q}}[\cdot|\mathcal{F}_t]$ is the conditional expectation under the historical filtration $\mathcal{F}_t = \sigma(S_u : u \leq t)$.

---

[5]We remark that the notation used in this chapter considers contracts with a delivery time rather than the more realistic delivery period. This ensures readability, the formulae can easily be adapted to periods.

[6]We remark that the spot price of electricity is not a traded asset and thus its discounted value needs not be a martingale under the risk-neutral measure. Hence, all measures $\mathbb{Q}$ equivalent to the real-world measure $\mathbb{P}$ are possible candidates and we will later choose a measure minimising the difference between observed forward prices and the conditional expectation.

We can now compare the conditional expectation under the historical measure with the conditional expectation under the pricing measure and use the difference as an indicator for market sentiment.

Definition 2.2. *Risk Premium.* *The risk premium is defined as*

$$R^{\mathbb{Q}}(t,T) = \mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t] - \mathbb{E}^{\mathbb{P}}[S_T|\mathcal{F}_t] \tag{2}$$

where the notation is as in Definition 2.1.

As mentioned above, the risk premium is subject to intense research. Note that one often uses observed forward prices for the expression $\mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t]$ (by assuming the correctness of equation (1)). After calculating expectations under $\mathbb{P}$ one then analyses the difference.

Since we want to study the impact of different information sets on forward prices we introduce further filtrations finer than the historical filtration. We need a filtration which contains specified information on future spot prices and a slightly coarser filtration which contains some un-specified additional information. To be precise:

Definition 2.3. *Filtrations.* *Let $\mathcal{H}_t$ be a filtration which includes the historical filtration as well as the information on the future value of the underlying at some time point $T_{\Upsilon} > t$, i.e.*

$$\mathcal{H}_t = \mathcal{F}_t \vee \sigma(S_{T_{\Upsilon}}) \tag{3}$$

Also, let $\mathcal{G}_t$ be a filtration that includes some information on the level of the future value of the underlying. We will call this filtration the market filtration *and we will assume that it represents the information available to market traders. This yields the relationship $\mathcal{F}_t \subseteq \mathcal{G}_t \subseteq \mathcal{H}_t$.*

As an example of possible future information available to the market we might consider $\mathcal{G}_t \subseteq \mathcal{F}_t \vee \sigma(\mathbb{1}_{\{S_{T_{\Upsilon}} \geq K\}})$. For this threshold information we know the value of the underlying will be larger than some constant K but we do not know the precise value.

As Benth and Meyer-Brandis (2009), we now define the information premium properly:

Definition 2.4. *Information Premium.* *Let $\mathcal{G}_t$ be the market filtration with extra information at T_{Υ}. Then the information premium is defined as*

$$I_{\mathcal{G}}^{\mathbb{Q}}(t,T;T_{\Upsilon}) = F_{\mathcal{G}}^{\mathbb{Q}}(t,T) - F_{\mathcal{F}}^{\mathbb{Q}}(t,T) \tag{4}$$

i.e. the difference between the forward prices under the market and the historical filtration.

8

In the following we will assume that all market participants work with the filtration $\mathcal{G}$. This implies that instead of assuming observed forward prices equal $\mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t]$ forward prices are calculated by market participants as $\mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{G}_t]$.

The following straightforward result will be of the utmost importance later:

Lemma 2.1. *Orthogonality of the information premium.* *The information premium is the residual of projecting the forward price under $\mathcal{G}_t$ onto the space $L^2(\mathcal{F}_t; \mathbb{Q})$. In other words*

$$\mathbb{E}^{\mathbb{Q}}[I_{\mathcal{G}}(t,T)|\mathcal{F}_t] = 0 \tag{5}$$

Proof. From Definition 2.4, the fact that $\mathcal{F}_t \subseteq \mathcal{G}_t$ and the tower property the result follows straightforwardly. $\square$

As mentioned before this result poses a problem when trying to identify the information premium empirically. It is obvious that it holds for all equivalent measures and thus the information premium cannot be attained by a measure change (the general approach in Financial Mathematics and the method used frequently to deduce the risk premium).

We conclude the introductory remarks with a little result connecting the two premia:

Lemma 2.2. *Relation of the premia.* *We have the following relation between the information premium and the risk premium:*

$$R_{\mathcal{F}}^{\mathbb{Q}}(t,T) = R_{\mathcal{G}}^{\mathbb{Q}}(t,T) - I_{\mathcal{G}}^{\mathbb{Q}}(t,T) + I_{\mathcal{G}}^{\mathbb{P}}(t,T) \tag{6}$$

Proof. Follows immediately from Definition 2.4 and Definition 2.2. $\square$

2.2. Two-Factor model

In this section we introduce the two-factor arithmetic spot price model, which will be the basic model for our study. The definition is as follows

Definition 2.5. *Spot Model.* *In the arithmetic spot price model the spot price process S_t is given as*

$$S_t = \Lambda(t) + X_t + Y_t, \tag{7}$$

with $\Lambda(t)$ a deterministic function to capture seasonal influences, X_t a standard Ornstein-Uhlenbeck processes given by the stochastic differential equation

$$dX_t = -\alpha X_t dt + \sigma dW_t \tag{8}$$

where $\alpha \in \mathbb{R}^+$ is the mean reversion parameter, $\sigma > 0$ the volatility and W_t a standard Brownian motion, and

$$dY_t = -\beta Y_t dt + dL_t \tag{9}$$

where again $\beta \in \mathbb{R}^+$ is the mean reversion parameter and L_t is a square integrable Lévy process.

By solving equations (8) and (9) for some $u < t$ we can write equation (7) more explicitly as

$$S_t = \Lambda(t) + e^{-\alpha(t-u)} X_u + \sigma \int_u^t e^{\alpha(t-s)} dW_s + e^{-\beta(t-u)} Y_u + \int_u^t e^{\beta(t-s)} dL_s \tag{10}$$

X_t is supposed to model the long-term behaviour of the spot and we will call it the *base component*. Y_t models the short-term fluctuation, i.e. the spikes, and we will call it the *spike component*.

As a specific example we will (as Benth et al. (2007)) use a compound Poisson process with double-exponentially distributed jumps for L_t (this is also called the *Kou Model*, presented by Kou (2002)). A formal definition is provided in Appendix Appendix A. This model specification is analytically tractable and has been shown to perform reasonably well empirically, see for example Benth et al. (2010).

The key result concerning the theoretical establishment of the information premium relies on the theory of enlargement of filtrations. It goes back to an initiating paper by Ito (1978) and has been adjusted to Lévy processes and non-complete future information (see chapter VI of Protter (2005)).

Theorem 2.1. *Itō's theorem extended to Lévy processes.* *Let L_t be a Lévy process and let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration and let $\mathcal{G}_t \subseteq \mathcal{H}_t = \mathcal{F}_t \vee \sigma(L_{T_\Upsilon})$. Then L is a semimartingale with respect to $\mathcal{G}_t$. Also, if $\mathbb{E}[|L_t|] < \infty$ for all $t \geq 0$ then the following process*

$$\xi(t) = L_t - \int_0^{t \wedge T_\Upsilon} \frac{\mathbb{E}[L_{T_\Upsilon} - L_s | \mathcal{G}_s]}{T_\Upsilon - s} \, ds \tag{11}$$

is a $\mathcal{G}_t$ martingale.

Thus, the theorem provides the martingale decomposition of L_t under the market filtration $\mathcal{G}_t$. Using this we can explicitly calculate the information premium:

Proposition 2.1. *Information premium for the Spot model.* *Denoting by $\phi(\cdot)$ the log-moment generating function of L_1 (see Definition Appendix A.1 of Appendix Appendix A for details) the information premium with Lévy information at $T_\Upsilon > T$ is given by:*

$$I_\mathcal{G}(t,T) = \frac{1}{\beta} \left(\frac{\mathbb{E}[L_{T_\Upsilon} - L_t \mid \mathcal{G}_t]}{T_\Upsilon - t} - \phi'(0) \right) \left(1 - e^{-\beta(T-t)} \right) \qquad (12)$$

Proof. Because our spot model is arithmetic and by the independence of X_t and Y_t most terms cancel:

$$\mathbb{E}[S_T \mid \mathcal{G}_t] - \mathbb{E}[S_T \mid \mathcal{F}_t] = \mathbb{E}\left[\int_t^T e^{-\beta(T-s)} dL_s \mid \mathcal{G}_t \right] - \mathbb{E}\left[\int_t^T e^{-\beta(T-s)} dL_s \mid \mathcal{F}_t \right] \qquad (13)$$

where the expectation under $\mathcal{F}_t$ is easily calculated

$$\mathbb{E}\left[\int_t^T e^{-\beta(T-s)} dL_s \mid \mathcal{F}_t \right] = \tfrac{1}{\beta} \phi'(0) \left(1 - e^{-\beta(T-t)} \right) \qquad (14)$$

Now, we can decompose the Lévy process according to equation (11) of Itō's theorem.

$$\mathbb{E}\left[\int_t^T e^{-\beta(T-s)} dL_s \mid \mathcal{G}_t \right] = \mathbb{E}\left[\int_t^T e^{-\beta(T-s)} \frac{\mathbb{E}[L_{T_\Upsilon} - L_s \mid \mathcal{G}_s]}{T_\Upsilon - s} \, ds \mid \mathcal{G}_t \right] \qquad (15)$$

$$= \int_t^T \frac{e^{-\beta(T-s)}}{T_\Upsilon - s} \, \mathbb{E}[L_{T_\Upsilon} - L_s \mid \mathcal{G}_t] \, ds \qquad (16)$$

$$= \frac{\mathbb{E}[L_{T_\Upsilon} - L_t \mid \mathcal{G}_t]}{T_\Upsilon - t} \frac{1}{\beta} \left(1 - e^{-\beta(T-s)} \right) \qquad (17)$$

where in the last step another technical result was applied (Proposition A.3 in Benth and Meyer-Brandis (2009)). Collecting terms yields the required result. □

2.3. Forward Prices under different measures with delivery period

To ease notation we used the "delivery time" notation rather than the more realistic "delivery period" notation so far. This period can have different lengths and on the EEX (cf. EEX (2011)) for example there are

monthly, quarterly, half- and yearly contracts. In other words we denoted the forward price as $F(t, T)$ whereas for the empirical analysis we will have to calculate $F(t, T_1, T_2)$ for a delivery of electricity between time T_1 and T_2. The theoretical results taking this into consideration will be provided here. We will assume that settlements only take place at the final date T_2. Then, the risk-neutral valuation formula yields for the forward price (see Benth et al. (2008b), pp. 29, 30):

$$F(t, T_1, T_2) = \mathbb{E}^{\mathbb{Q}} \left[\int_{T_1}^{T_2} \frac{1}{T_2 - T_1} \, S_u du | \mathcal{F}_t \right] \tag{18}$$

A parametric measure change from the real world measure to a risk-neutral measure is now going to be conducted. As usual, we apply Girsanov's theorem to the Brownian part of the spot price and an Esscher transform to the Lévy part. As X_t and Y_t are independent, we can apply both methods separately and obtain a product measure.

For the Brownian part we define the Radon-Nikodym derivative $Z_W(t)$ as usual

$$Z_W(t) = \frac{d\mathbb{Q}_W}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = \exp \left(-\int_0^t \frac{\theta_W(s)}{\sigma} \, dW_s - \frac{1}{2} \int_0^t \frac{\theta_W^2(s)}{\sigma^2} \, ds \right) \tag{19}$$

Under $\mathbb{Q}_W$

$$d\tilde{W}_t = -\frac{\theta_W(t)}{\sigma} \, dt + dW_t \tag{20}$$

is again a Brownian motion. $\theta_W(t)$ is called the market price of risk for the base component. Also, the dynamics of X_t are altered as follows

$$dX_t = -\alpha X_t dt + \sigma dW_t = -\alpha X_t dt + \sigma(d\tilde{W}_t + \frac{\theta_W(t)}{\sigma} \, dt) = (\theta_W(t) - \alpha X_t)dt + \sigma d\tilde{W}_t \tag{21}$$

Similarly, if L_t satisfies the integrability condition Lemma Appendix A.2 of Appendix Appendix A we define the Esscher transform

$$Z_L(t) = \frac{d\mathbb{Q}_L}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = \exp \left(\int_0^t \theta_L(s)dL_s - \int_0^t \phi(\theta_L(s))ds \right) \tag{22}$$

where the parameter function $\theta_L(t)$ is called the market price (of jump risk) for the spike component. We refer to Applebaum (2004) (chapter 5.4) for

12

more details and a technical discussion. In Appendix Appendix B, Lemma Appendix A.3, we show that the Poisson process with double-exponential jump sizes satisfies the condition.

We obtain a risk-neutral measure $\mathbb{Q}^\theta = \mathbb{Q}_W \times \mathbb{Q}_L$. Here $\theta = (\theta_W, \theta_L)$ is a vector of parameters corresponding to the measure change.

We can now calculate the expectation in (18) to obtain the $\mathbb{Q}$-forward price.

Proposition 2.2. *Forward price under $\mathbb{Q}$ and with delivery period.*
The forward price with delivery period from T_1 until T_2 under the measure $\mathbb{Q}^\theta$ is given by

$$F^\mathbb{Q}(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left(\int_{T_1}^{T_2} \Lambda(u)du + \bar\alpha(t, T_1, T_2)X_t + \bar\beta(t, T_1, T_2)Y_t \right)$$
$$+ \frac{1}{T_2 - T_1} \left(\int_t^{T_2} \theta_W(s)\bar\alpha(s, T_1, T_2)ds + \int_t^{T_2} \phi'\left(\theta_L(s)\right)\bar\beta(s, T_1, T_2)ds \right)$$
$$(23)$$

where $\phi(u)$ denotes the log-moment generating function of the Lévy process L at time $t = 1$ and where again $0 \le t \le T_1 < T_2$. Auxiliary functions $\bar\alpha(t, T_1, T_2)$ and $\bar\beta(t, T_1, T_2)$ are given in Appendix Appendix B, equations (B.1) and (B.2).

Proof. Details are provided in Appendix Appendix B. $\square$

For constant parameters we can simplify:

Corollary 2.1. *Forward price with constant parameters.* *With constant market prices of risk the forward price with delivery is given by*

$$F^\mathbb{Q}(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left(\int_{T_1}^{T_2} \Lambda(u)du + \bar\alpha(t, T_1, T_2)X_t + \bar\beta(t, T_1, T_2)Y_t \right)$$
$$+ \frac{1}{T_2 - T_1} \left(\theta_W \hat\alpha(t, T_1, T_2) + \phi'\left(\theta_L\right)\hat\beta(t, T_1, T_2) \right)$$

where functions $\hat\alpha(t, T_1, T_2), \hat\beta(t, T_1, T_2)$ are defined in Appendix Appendix B, equation (B.3).

To calculate the forward price for the case that t is within the delivery period we find:

13

Corollary 2.2. *Forward price during the delivery period.* *For* $0 \leq T_1 < t \leq T_2$ *the forward price under* $\mathbb{Q}$ *takes the form*

$$F^{\mathbb{Q}}(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left(\int_{T_1}^{t} S_u du + \int_{t}^{T_2} \Lambda(u) du + (T_2 - t) F^{\mathbb{Q}}(t, t, T_2) \right)$$

(24)

We can retrieve the $\mathbb{P}$-forward price as a special case of the results above by setting the measure change parameters $\theta_L(t) = \theta_W(t) = 0$. We refer to corollary Appendix B.1 of Appendix Appendix B.

We have now developed the toolbox to commence our empirical analysis of the information premium.

3. Methodology of the Empirical Analysis

3.1. Motivation

The information premium is orthogonal to $\mathcal{F}_t$, independent of the underlying measure (equivalent to the real world measure $\mathbb{P}$ (see Lemma 2.1)). Thus, we cannot explain the premium by a measure change. However, for pricing a measure change takes place nonetheless and we choose the pricing measure $\mathbb{Q}$ as a distance-minimising measure between the expected spot and the forward price. Thereby we will "normalise" the remainder term being analysed in the following. A constant Girsanov change of measure will be conducted later (agreeing with the simple (and constant) setup followed so far). Non-constant Girsanov parameters as well as an Esscher transformation can easily be added if required.

Leaving out delivery periods again for the ease of notation we can define a new spot-forward relationship:

Definition 3.1. *New Spot-Forward relationship.* *The fair price of a forward contract in this framework is given by*

$$F_{\mathcal{G}}^{\mathbb{Q}}(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{G}_t]$$

(25)

where $\mathbb{Q}$ *is a pricing measure and* $\mathcal{G}_t$ *is the market filtration.*

Here, $\mathbb{Q}$ will be calibrated to the market. Without any assumptions about the additional information (such as the threshold discussed in Section 2.1) this expectation cannot be calculated or simulated, since the structure of the filtration $\mathcal{G}_t$ is not known. However, for our empirical work we will

14

assume that observed prices $\hat{F}(t, T)$ were calculated by market participants with respect to $\mathcal{G}$ and $\mathbb{Q}$. Thus,

$$F_{\mathcal{G}}^{\mathbb{Q}}(t, T) \overset{!}{=} \hat{F}(t, T) \tag{26}$$

In terms of the historical filtration we can calculate forward prices as

$$F_{\mathcal{F}}^{\mathbb{Q}}(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{F}_t] \tag{27}$$

Now, we replace the theoretical information premium (Definition 2.4)

$$I_{\mathcal{G}}^{\mathbb{Q}}(t, T) = F_{\mathcal{G}}^{\mathbb{Q}}(t, T) - F_{\mathcal{F}}^{\mathbb{Q}}(t, T) \tag{28}$$

by its empirical version

$$\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T) = \hat{F}(t, T) - \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{F}_t] \tag{29}$$

As a consequence $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ is our best estimate of $I_{\mathcal{G}}^{\mathbb{Q}}(t, T)$.

The goal now is to show that $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ exists (i.e. is non-zero) and that it is not $\mathcal{F}_t$-measurable (i.e. that it satisfies the orthogonality property). In mathematical terms this translates to:

1. $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ is significantly not white noise
2. $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ is orthogonal to $L^2(\mathcal{F}, \mathbb{Q})$, or, by Lemma 2.1, equivalently $\mathbb{E}[\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T) | \mathcal{F}_t] = 0$

To test $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ for white noise is fairly standard. A simple graphical method is looking at the graph of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ as well as the graph of its auto-correlation function (*acf*). Also, one can test $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ using a standard autocorrelation-based test like *Ljung-Box*.

3.2. Conditional expectations, Hilbert bases and the regression approach

To our best knowledge there is no literature on how to test for non-measurability empirically. In this section we will therefore propose a Hilbert representation and regression based approach. Our test will be able to tell whether or not one time series is measurable with respect to another one. We want to show empirically that $\mathbb{E}[\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T) | \mathcal{F}_t] = 0$. In other words we want to evaluate a conditional expectation

$$F(\omega, t_k, t_l) = \mathbb{E}^{\mathbb{Q}}[X_{t_l} | \mathcal{F}_{t_k}](\omega) \ , \ k \leq l \leq n \tag{30}$$

15

over a discrete grid of time points $t_0 < \ldots < t_k < \ldots < t_n = T$. For suitable X we thus consider the (separable) Hilbert space $L^2(\mathcal{F}, \mathbb{Q})$ spanned by a countable, orthonormal basis with elements $\phi_v(X_{t_k})$ for $v \in \mathbb{N}$. Then, one can try to express $F(\omega, t_k)$ as a linear combination of these basis elements:

$$F(\omega, t_k) = \sum_{v=1}^{\infty} a_v \phi_v(X_{t_k}) \tag{31}$$

We will use stationary increments of the time series over the lifetime of a forward and regression to find the functional form of the conditional expectation.

The theoretical background of such an approach can be found in may scoures, see e.g. Royden (1968). Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a measure space and consider the space of all square integrable random variables relative to $\mathbb{Q}$, i.e. the space in which the elements X satisfy

$$\int |X|^2 d\mathbb{Q} < \infty \tag{32}$$

These spaces are separable. Also, one knows that the conditional expectation is then a contraction of the space L^2 meaning that

$$X \in L^2(\Omega, \mathcal{F}) \Rightarrow \mathbb{E}[X|\mathcal{F}_0] \in L^2(\Omega, \mathcal{F}_0) \tag{33}$$

for some sub-sigma algebra $\mathcal{F}_0 \subseteq \mathcal{F}$. Then let $\mathcal{F}_0 = \sigma(Y)$ for some random variable Y. The conditional expectation $\mathbb{E}[X|\sigma(Y)]$ is also $\sigma(Y)$-measurable and thus there exists a measurable function $F : (\Omega, \mathcal{F}_0) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ such that

$$F(Y) = \mathbb{E}[X|\sigma(Y)] \tag{34}$$

A separable Hilbert space possesses complete countable orthonormal systems and one can write every element of that space as a linear combination of elements of the orthonormal system $\{\phi_1, \phi_2, \ldots\}$, in particular also the conditional expectation

$$F = \sum_{v=1}^{\infty} a_v \phi_v \tag{35}$$

Hence, regressing a sufficiently large sample of pairs (X, Y) on a sufficient number of basis functions we will get coefficients c_v and an approximation to the functional representation of the conditional expectation.

16

In our implementation we use a whole time span $t_0 < \ldots < t_k < \ldots < t_n \leq T$ with data

$$S_{t_k} \ , \ \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t_k, T) \quad \forall \ 0 \leq k \leq n \tag{36}$$

where the horizon T is kept fixed. An example for such a time span would thus be the "lifetime" of one forward (the time it is traded and priced on the market) and this is the approach followed in the next section. We want to test how much of the variation in $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)$ can be explained by S_t (and thus by $\mathcal{F}_t$). We conclude non-measurability if the regression would yield little explanatory power and all regression coefficients were significantly zero. From this $\mathbb{E}[\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T)|\mathcal{F}_t] = 0$ would follow, consequently. Still, one has to be careful as it is essential for this method to work properly that both time series under consideration are stationary (also to avoid spurious regression). Hence, we use first differences and our regression formula is

$$\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \sum_{i=1}^{N} a_i \triangle S_t^i + \triangle \epsilon(t) \tag{37}$$

This approach is in line with the foundations described above because one has from stationarity the property that

$$\triangle S_t \sim \triangle S_u \ , \ \triangle R_t \sim \triangle R_u \ \forall t, u \tag{38}$$

Summarising, in this section, we developed the following agenda for testing for the information premium:

1. Calibrate the spot model to the observed spot data
2. Calculate the forward prices under measure $\mathbb{P}$
3. Change the measure to risk-neutral $\mathbb{Q}$
4. Calculate the forward prices under measure $\mathbb{Q}$
5. Choose a certain forward to be examined
6. Calculate the residual term $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for all t
7. Test $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for white noise
8. Test $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for non-measurability

   - Make $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ and corresponding S_t stationary
   - Regress the two time series and check significance (F- and t-statistics)

4. The Empirical Analysis

4.1. Empirical Measure Change and Structure of Forward Prices

There is a large variety of literature about how to fit the two-factor arithmetic spot model, see Benth et al. (2010), Benth et al. (2008a) and Meyer-Brandis and Tankov (2008) for example.

For futures prices there are several time series which we consider. Bloomberg, Reuters and the EEX provide daily data organised in columns each representing one of the (month-) futures[7] traded on the current day. This means that a rolling effect occurs, i.e. that if one wants to track a certain contract through its lifetime one has to read the data table diagonally in a top-right to bottom-left direction. Hence, there are essentially three ways to examine forward prices, i.e. considering...

- the forward maturing in a certain number of months for all days (i.e. t and $[T_1, T_2]$ moving through the time series, T_1 and T_2 being the beginning and the end of the delivery period)

- the forward maturing in a specified month (i.e. t moving though the lifetime of the forward, T_1, T_2 fixed)

- all forwards traded on a specified day (i.e. with fixed t, compare Figure 1)

In Section 2.3 we have seen the formulae for the prices of forwards with delivery period and under both measures, $\mathbb{P}$ and $\mathbb{Q}$. As mentioned before, we will consider constant Girsanov parameters when changing the measure.

Empirically, Girsanov parameters were chosen in such a way as to minimise the difference between the observed forward prices and the calculated $\mathbb{Q}$-prices. This is done globally (i.e. over the length of the whole data set) for each time to delivery period. This means that constant distance-minimising parameters were identified (by means of least squares) for each time until delivery, for example the three-months-forward or the four-months-forward.

The reason for this approach is that we believe that market participants adjust risk for (global) classes of forward contracts. Generally, a two-months

---

[7]Observed prices are from futures contracts. Since we can regard interest rates as deterministic for the period under consideration, forward and futures contracts are equivalent. This assumption is standard in the relevant literature, as for example in Longstaff and Wang (2004). Furthermore, at the time of the particular cases we consider half-year interest rates were almost constant.

(time-to-delivery) forward will be priced differently than a six-months (time-to-delivery) forward but all two-months forward contracts are similarly risk-adjusted by market traders. This is in line with the risk-premium literature such as Benth et al. (2008a) where, as discussed previously, the authors find positive risk premia for short term deliveries and negative ones for long term delivery periods. This (non-orthogonal) effect is captured by our method, at least for the CO_2 data (see Table 2). There have been some fundamental changes to prices in recent years (compare Table 8).

To examine the relation between the different observed and calculated objects we will temporarily consider the EEX dataset ranging from 06/10/2003 to 26/05/2006. The spot of this range is illustrated in Figure 4.
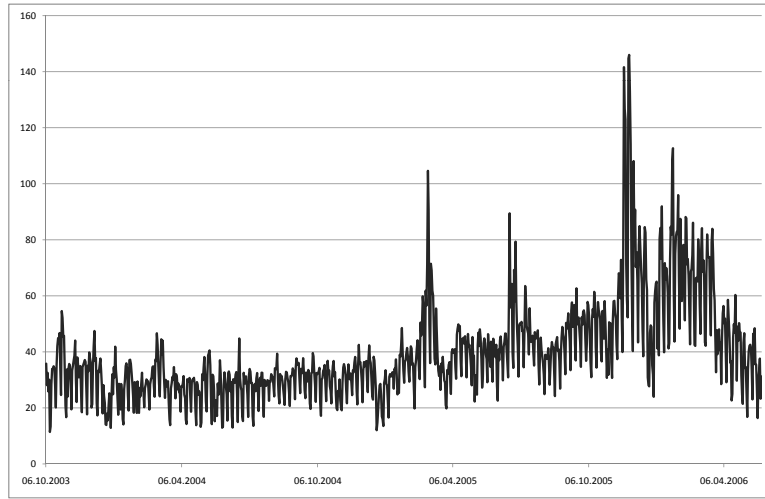
Figure 4: **EEX spot price.** Range is from 06/10/2003 until 26/10/2006

In the following, forwards with different and fixed time to maturity will be considered. This will be done by comparing the observed prices for a forward delivering in n months/quarters with the calculated prices for the same delivery period. The observed forward price with a given time to maturity will be illustrated in dotted light grey, whereas the forward price calculated under $\mathbb{P}$ will be in dark grey, the forward price calculated under $\mathbb{Q}$ in medium grey and the realised price, i.e. the future looking arithmetic mean of spot prices during the delivery period (thus the series is shorter) in light grey.
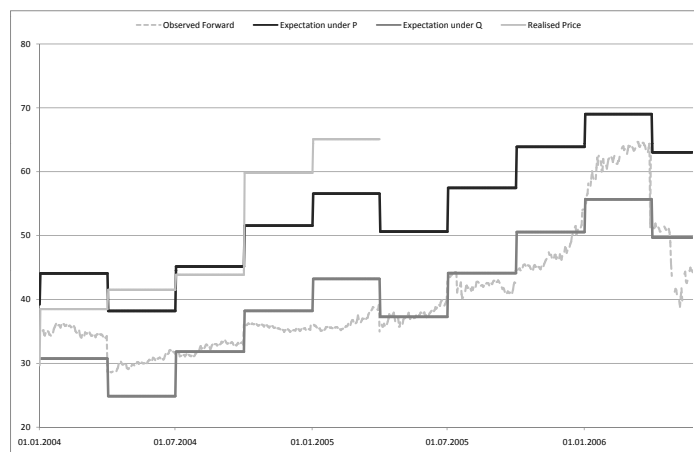
Figure 5(a) shows the forward prices for the current month. All graphs, especially those two calculated are very close to each other. The most

19
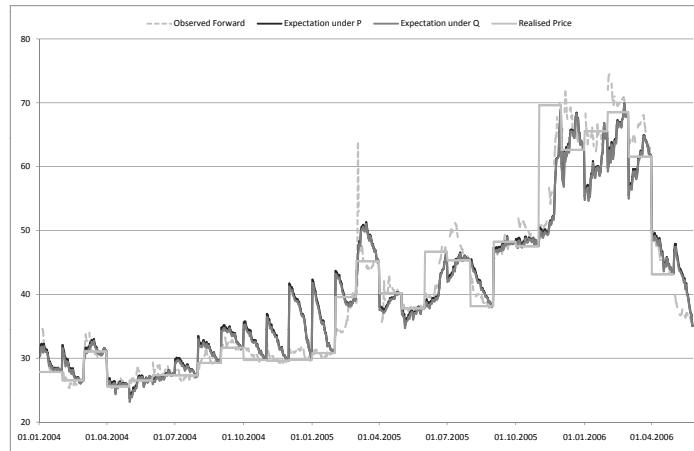
(a) Different forward prices with one month until delivery
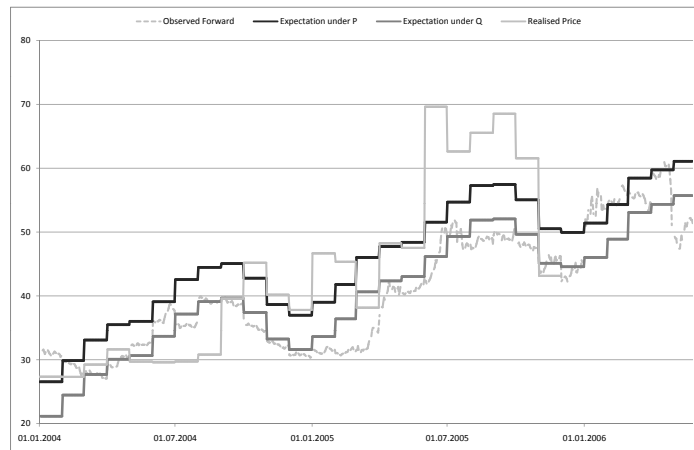
(b) Different forward prices with six months until delivery

(c) Different forward prices with four quarters until delivery

20

Figure 5: **Different EEX and calculated forward prices.** The time series for one and six months and four quarters until delivery period.

striking part of this figure happens during the winter 2004/2005 where in the beginning of each month the calculated prices very much overestimate the observed and realised prices followed by a sharp decrease until the realised price is reached at the end of the month. The reason for this behaviour lies in equation (24): the price of the forward for the current month is the expected value of future days plus the arithmetic mean of that part of the month already in the past. Hence, the spot model seems not to fit this period very closely.

Figure 5(b) shows the corresponding picture for six months until the delivery period. Calculated forward prices are piecewise constant. The reason for this behaviour is the size of the mean-reversion parameters which flattens out prices (i.e. influence of current spot prices becomes negligible). The difference between the two calculated time-series is now much bigger (more than five Euro) and one can clearly see that the risk-neutral price is very close to observed prices. A more surprising feature is that the real-world price seems to be closer to the actual price realised later. The same fact can be observed in Figure 5(c) which shows forward prices with a delivery period of a quarter and a time to maturity of four quarters, i.e. one year. Again, the risk-neutral price is coupled with the observed price whereas the real-world price is closer to the realised price.

Summarising, the sizing of this phenomenon grows with time to maturity. The further the delivery period the larger the effect of the measure change and the more extreme and obvious is the coupling observed/risk-neutral and realised/real-world prices.

For the rest of this paper we will consider the lifetime of single forward contracts rather than examining the corresponding time series globally. The reason is that the information premium, i.e. the influence of individual pieces of information, will show much clearer for single forward contracts.

4.2. The beginning of the second phase of the EU ETS

Finally, we now have all preliminaries to empirically analyse the beginning of the second phase of CO_2 certificates (as motivated in the introduction). Our goal is to prove, by the method developed in Section 3, the existence of a significant information premium. The data set considered in this subsection will be the EEX spot baseload price from 01/02/2007 until 30/10/2008, consisting of 639 data points. This data set was selected as to include the crucial date 01/01/2008 as a midpoint.

Details about the seasonality function $\Lambda(t)$ and fitted parameters for this data set can be found in Appendix Appendix C and Table C.17. Table

21

1 provides the fitted values of the stochastic parts of the spot price.

Table 1: Fitted parameter values for the data set 01/02/2007 until 30/10/2008.

Parameter	α	σ	β	λ	p	q	η_1	η_2
Value	0.538	11.108	0.786	0.034	0.955	0.045	0.019	0.027

The change of measure parameters are given in Table 2. They are positive for the first three months and negative for more distant delivery periods. As mentioned above this confirms the findings of Benth et al. (2008a).

Table 2: Constant Girsanov parameters for different forward classes.

Forward Class	1 m	2 m	3 m	4 m	5 m	6 m
Girsanov parameter θ_W	0.164	0.734	0.153	-0.593	-1.893	-3.199

The spot for the range of dates as mentioned above is illustrated in Figure 6. Already in this figure one can observe the introduction of the CO_2 certificates as December 2007 and January 2008 are very volatile and separate the data set into two parts. In 2007 spot prices are around 30 Euro with relatively little volatility and the prices in 2008 are at least around 50 or 60 Euro with a slightly higher volatility. Thus, prices exhibit not only the usual slow positive trend but also a general shift upwards.

In Figure 7 all objects under consideration are illustrated (here, all series are with respect to remaining six months until delivery period, i.e. the first type of perspective discussed earlier). Colors of graphs are as in the last section. Generally the figure exhibits greater differences between the observed forward and both, the $\mathbb{P}$- and the $\mathbb{Q}$-price.

With these four prices we can now calculate the two premia:

- The risk premium $R_{\mathcal{F}}^{\mathbb{Q}}(t, T_1, T_2)$

- The residual term $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$

4.2.1. Introduction of CO_2 certificates: The residual $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$

We now change perspective by choosing one special forward. All the figures so far in this chapter were of the first type discussed in Section 4.1.
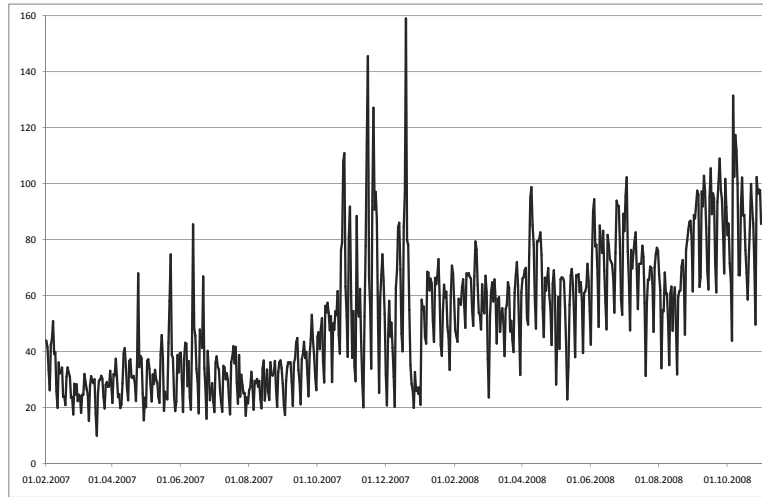
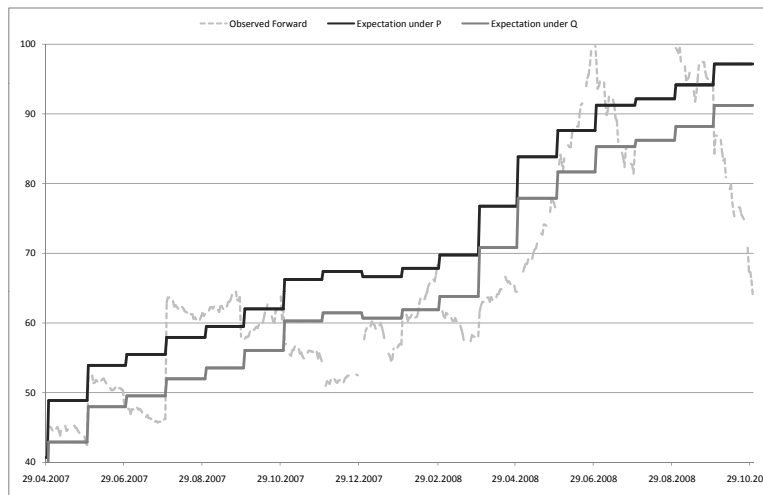Figure 6: **EEX spot price.** Range 01/02/2007 until 30/10/2008.

Figure 7: **Observed and calculated forward prices.** Here, delivery is in six months. Range of data is 01/02/2007 until 30/10/2008.

In the following it will be the second type, i.e. following the lifetime of a forward (moving t, fixed T_1, T_2).

Using the definitions and notations including the delivery period the residual term $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ is defined as

$$\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \hat{F}(t, T_1, T_2) - \mathbb{E}^{\mathbb{Q}}[\tfrac{1}{T_2 - T_1} \int_{T_1}^{T_2} S_u du | \mathcal{F}_t] \qquad (39)$$

This object will be calculated for four forward contracts living before, after and during the introduction of the CO_2 certificates. These were chosen to be those contracts with delivery period in

- November 2007

- January 2008

- March 2008

- August 2008

Figure 8 gives an example of the time series involved for the forward maturing in January 2008. On the one hand one can see the constant behaviour of the $\mathbb{P}$-expectation. The expectation under $\mathbb{Q}$ on the other hand shifts. This is due to the (global) measure change which is different for each forward class (one-month, two-month etc.).

Figure 9 shows $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ and its autocorrelation function for all four forwards.

The most interesting part of this figure is Figure 9(c), i.e. the residual for the January 2008 forward. It is large and positive for almost the whole time span and has a larger volatility the closer time to maturity comes. Then, from 01/01/2008 it decreases and even becomes negative in the end. This is exactly the type of behaviour one would have expected from the motivating example: a positive information premium which tends to zero after the spot itself finally reacts in real-time to the introduction of the CO_2 certificates. Very similar is the picture for the March 2008 forward (Figure 9(e)) where one finds again positive values for the months of 2007 followed by negative values for 2008. This can be interpreted presuming that the market expected a price increase due to CO_2 certificates which was overestimated.

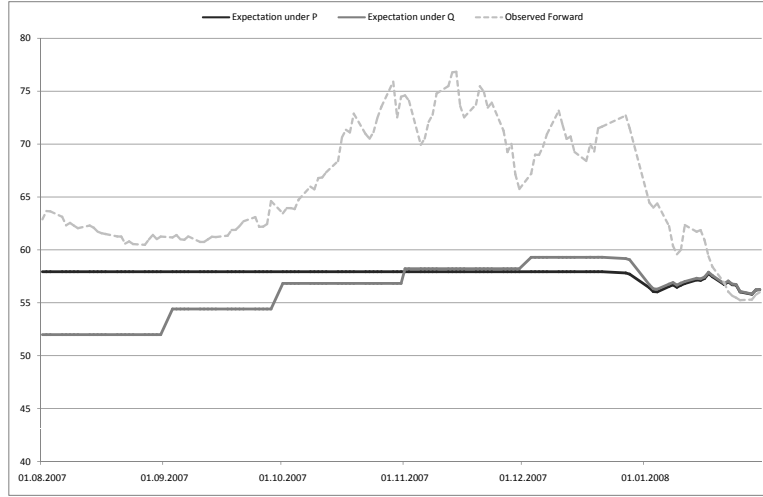All autocorrelation functions indicate a stationary first difference.

Figure 8: **Forward with delivery in January 2008.** Observed and calculated prices over its lifetime.

Table 3: Properties of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for the four different forwards.

	Nov 07	Jan 08	Mar 08	Aug 08
Mean	-0.32	9.64	-1.14	-2.58
Std. dev.	6.58	4.60	4.28	4.43
# of days > 0	52	119	52	34
# of days < 0	77	7	71	93

As mentioned in the agenda one needs to show that $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ is significantly non-zero. All four series have mean unequal to zero with the most extreme being the January residual with a value of 9.64 Euro (compare Table 3). It is obvious from the graphs that none of the residuals is white noise. This is confirmed by applying a Ljung-Box test whose results are shown in Table 4. The null hypothesis is rejected for all levels and all four forwards.
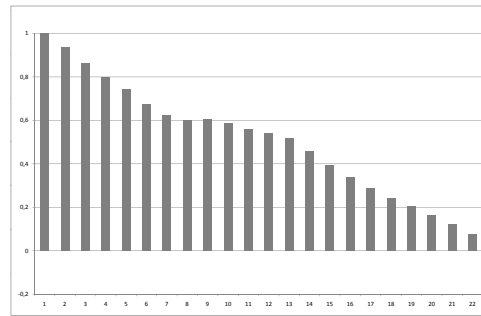
Table 4: Ljung-Box test for white noise for $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$.

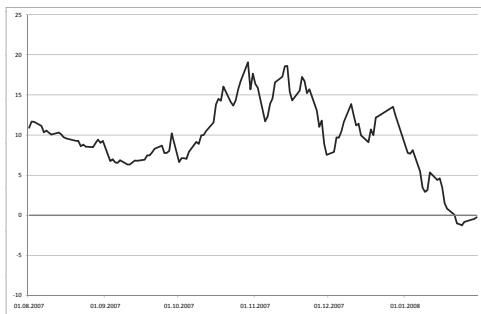|	Nov 2007	Jan 2008	Mar 2008	Aug 2008
Ljung-Box	867.68	738.84	1606.87	797.01
χ^2 (95%)	36.06	35.73	35.40	35.84

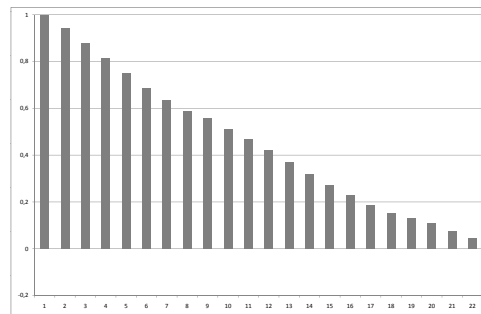Hence, the residual term is significant not zero.
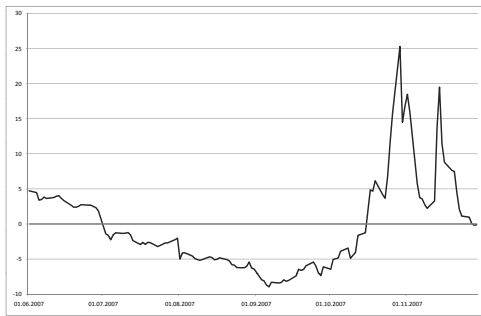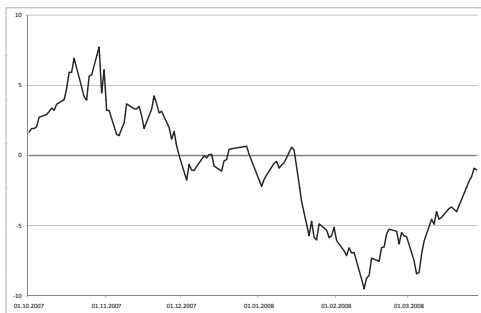
(a) November 2007: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(b) November 2007: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(c) January 2008: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(d) January 2008: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(e) March 2008: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
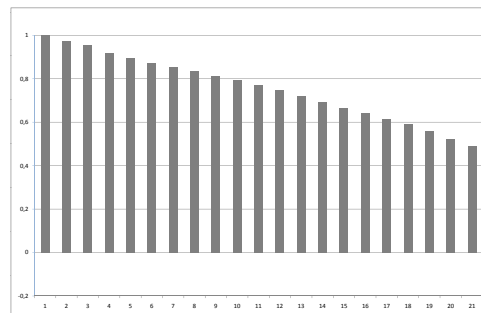
(f) March 2008: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
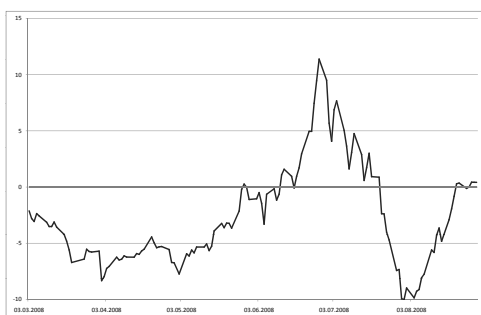
(g) August 2008: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
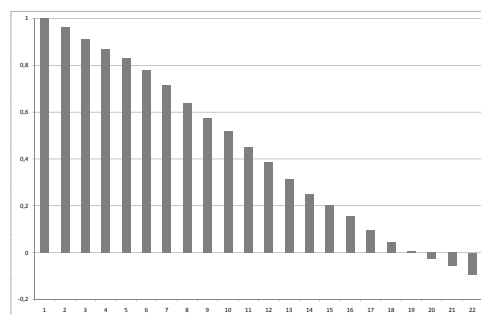
(h) August 2008: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

Figure 9: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ **and autocorrelation function.** Chosen forwards contracts (EUETS data)

4.2.2. Introduction of CO_2 certificates: Regression results

As explained in Section 3 we will use a regression-based approach to show non-measurability. The goal of this approach is to show that the variations in the spot cannot explain the variations in the residual term. The basis used in this chapter will be the polynomial basis, i.e. $\{S^i, i \in \mathbb{N}\}$ although the choice of basis does not alter the general result. As explained in Section 3 we will use first-differences to have stationary time series. This can be justified by applying the Dickey-Fuller test of stationarity (see Table 5): The Dickey-Fuller test rejects stationarity at all levels for the pure time

Table 5: Dickey-Fuller statistics for time series and their first differences

X	S	Nov 07	Jan 08	Mar 08	Aug 08
DF(X)	-0.84	-0.68	-0.13	-0.25	-0.31
DF($\triangle X$)	-19.98	-7.23	-7.68	-8.37	-6.80

series while it accepts stationarity at all levels for first differences.

Therefore we consider the regression

$$\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \sum_{i=1}^{N} a_i \triangle S_t^i + \triangle \epsilon_t \tag{40}$$

Hence, we are now examining the increments of the spot and the residual term. The results of this regression with $N = 10$ are given by Table 6: The

Table 6: R^2s and F-statistics of the regression from $\triangle S$ onto $\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}$, $N = 10$

	Nov 07	Jan 08	Mar 08	Aug 08
R^2	0.14	0.07	0.03	0.07
F-statistic	1.47	0.65	0.35	0.75

critical value of the F-distribution at the 95% level for all four data sets is 1.88 and thus for all four series the hypothesis of value zero coefficients is accepted (still, results for November 2007 are less obvious as expected). Also, all individual coefficients have zero value and t-statistics are insignificant at all levels so that none of the basis polynomials has any explanatory power. We refer to Table D.19 of the appendix for details. Results do not change with larger N.

Thus, one can conclude that the spot does not help to explain the variations of the residual term. We have thus proved that there exists a substantial part of the forward price which is orthogonal to the historical filtration

and which is non-measurable. This additional information premium is induced by pricing under an enlarged filtration including some publicly known future information.

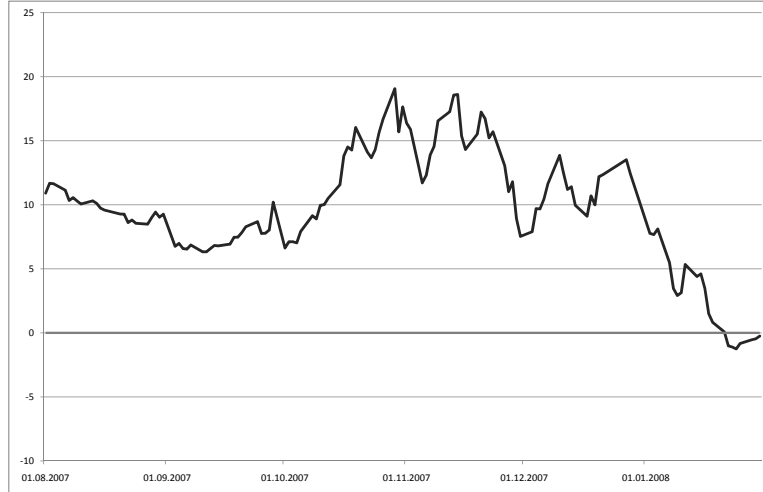### 4.2.3. Introduction of CO_2 certificates: Discussion of results

Figure 10: **The information premium** $I_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$**.** Corresponding to the January 2008 forward.

Figure 10 again shows the information premium $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for the forward with delivery period in January 2008. It is positive over nearly the whole lifetime of the contract with a mean of 9.64. The other residuals of Figure 9 have much smaller means of $-0.32, -1.14$ and -2.57 respectively. The second important feature is that the residual is relatively stable over the first 70 days (with a variance of 2.31), followed by a more turbulent period (variance 8.43 from day 71 until the beginning of January). Then from 01/01/2008 the residual sharply drops from over 9 Euro to negative values in the end of January.

Considering the extra information about the introduction of the emission fees this behaviour was expected. Clearly the market hypothesised an upward shift in electricity prices due to higher costs induced by the certificates. This upward shift corresponds to the positive graph of the residual. During January this additional information is then step by step incorporated into the historical filtration and the information premium thus becomes negligible.

Let us discuss the value of the residual shortly. The EEX spot price for CO_2 certificates during the second half of 2007 was practically zero (less than 8 cent, for data see EEX (2011)), the reasons for this being over-allocation and the no-banking property. The forward price for the year 2008 (i.e. the beginning of phase two) on the other hand was between 18.50 Euro in August 2007 and 23.80 Euro in November 2007, with an average of around 22.00 Euro for one tonne of CO_2. Assuming an average efficiency factor of $0.70 tCO_2/MWh$ (i.e. a typical value for the German market) this results in extra costs in 2008 between 12.95 Euro and 16.66 Euro which is, in fact, very similar a number to our residual in Figure 10. Hence, also the size of the residual is in line with the economic intuition.

4.3. The Japan earthquake and the "Atom Moratorium"

Here, we will discuss the German "Atom Moratorium" as motivated in the introduction. Again, electricity data will be taken from the EEX and the time range considered is now 01/09/2009 until 15/08/2011 which are 711 days. As mentioned above the critical dates are 11/03/2011 (earthquake), 14/03/2011 (Moratorium), 31/05/2011 (final decision to close old plants) and 14/06/2011 (end of Moratorium). Figure 11 is a plot of the data set.
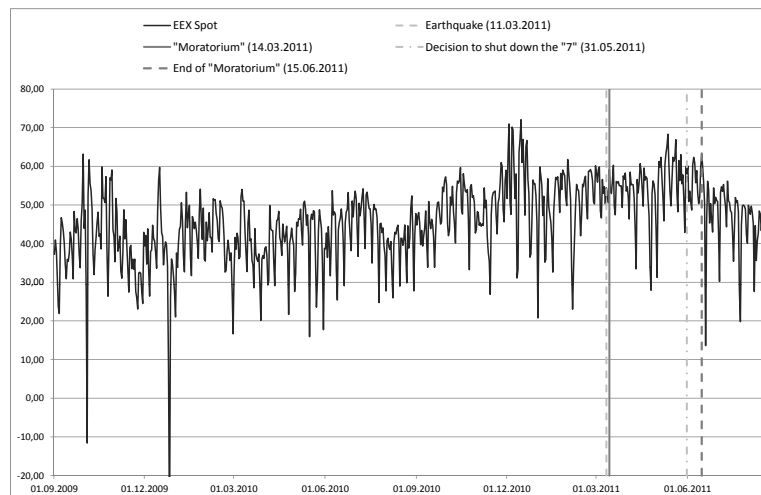

Figure 11: **EEX spot price.** Range from 01/09/2009 until 15/08/2011.

Perhaps the most striking feature of this graph after analysing the spot range from Section 4.2 is that there are hardly any spikes. Those spikes still existent mostly have negative heights. Overall liquidity and market

design seem to have improved. Negative jumps will mostly be caused by renewables, wind in particular. Electricity thus produced has to be used by law. Strong wind will therefore lead to a huge price decline, especially in times of low demand. An example is Christmas 2009 which featured a daily average price of less than -20 Euro.

As before, we commence the analysis by fitting the spot model. The resulting parameters are provided in Table C.18 and Table 7.

Table 7: Fitted parameter values for the data set 01/09/2009 until 15/08/2011

Parameter	α	σ	β	λ	p	q	η_1	η_2
Value	0.499	6.01	0.864	0.027	0.105	0.895	0.046	0.033

The corresponding change of measure parameters are given in Table 8. We remark that the changes in the market described above are also confirmed by these parameters. We do not see the typical change of sign which was partially induced by retailers' fear of positive spikes.

Table 8: Constant Girsanov parameters for different forward classes.

| Forward	1 m	2 m	3 m	4 m	5 m	6 m
Girsanov parameter θ_W	0.210	0.624	0.650	0.614	0.512	0.363

Again, we will examine specific forward contracts. We chose the contracts with delivery in
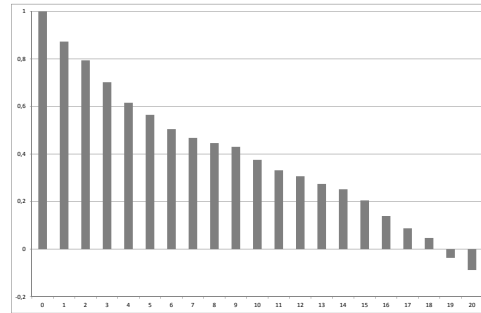
- February 2011

- May 2011

- July 2011

4.3.1. Moratorium: The residual $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$

We calculate expectations under $\mathbb{P}$ and $\mathbb{Q}$ for the three forward contracts mentioned above and subsequently plot the corresponding residual terms $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ on the left panel of Figure 12. The right panel of this figure shows, as before, the auto correlation function (acf) of the residuals.
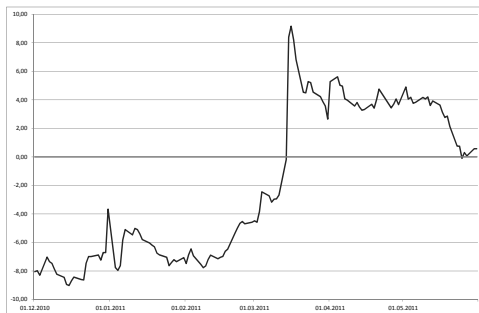
The autocorrelation functions again show that residuals are non-stationary but indicate that their first differences will be. The price impact of the
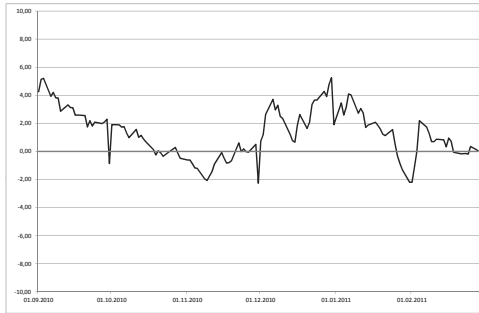
(a) February 2011: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(b) February 2011: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

(c) May 2011: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
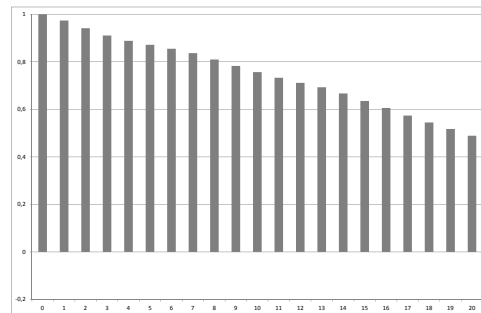
(d) May 2011: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
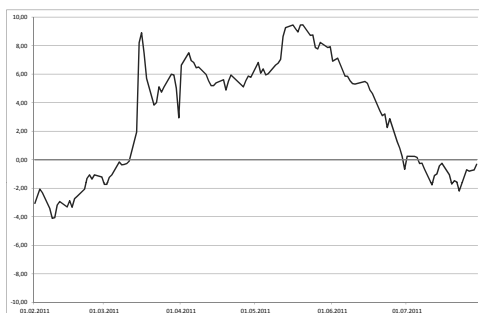
(e) July 2011: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$
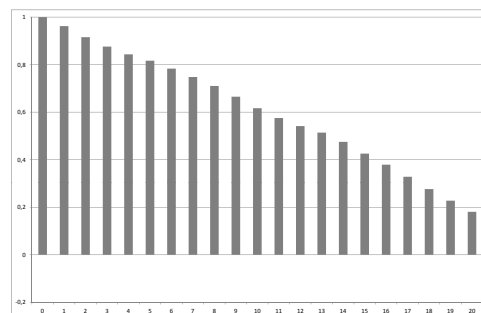
(f) July 2011: acf of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}$

Figure 12: $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ **and autocorrelation function.** Chosen forwards contracts (Moratorium data)

Moratorium can clearly be seen in Figure 12(c) and Figure 12(e) whereas the residual Figure 12(a) seems to be of much more regular behaviour.

Table 9: Properties of $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ for the three different forwards

	Feb 11	May 11	Jul 11
Mean	1.27	-2.07	2.92
Std. dev.	1.76	5.36	4.05
# of days > 0	93	53	77
# of days < 0	36	75	48

Important properties of the three residuals are provided in Table 9. Mean and standard deviation of the residual of the February forward confirm regularity (i.e. absence of an information premium). The residual of May 2011 has a much higher standard deviation and is also clearly divided into two parts: before the Moratorium (mean of -6.36) and after (mean 3.82). We have a similar situation for July 2011 where there is a mean of 6.54 during the three months of the Moratorium and one of -0.64 before and after. Also, market participants seem to have reevaluated the new legal situation in June and July as the residual returns to its pre-earthquake level. This will be discussed later.

Again, to confirm the non-zero property we apply a Ljung-Box test for white noise whose results are shown in Table 10. The null hypothesis is rejected for all levels and all three forwards.

Table 10: Ljung-Box test for white noise for $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$

	Feb 2011	May 2011	Jul 2011
Ljung-Box	522.97	1645.87	1094.83
χ^2 (95%)	36.06	35.95	35.62

After finding the residuals to be significant non-zero we again continue to check their orthogonality.

4.3.2. Moratorium: Regression results

In this section we will, as before in Section 4.2.2, use our regression test method to check whether the residual $\hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$ is measurable with respect to the spot. Again, the basis used will be the polynomial basis of the spot price $\{S^i, i \in \mathbb{N}\}$.

Table 11: Dickey-Fuller statistics for time series and their first differences

X	S	Feb 11	May 11	Jul 11
DF(X)	-0.48	-0.95	-0.28	-0.21
DF($\triangle X$)	-21.46	-9.97	-6.86	-6.52

The Dickey-Fuller test of stationarity yields the results summarised in Table 11. The critical value for the 10%-level is approximately -1.61. Thus, the Dickey-Fuller test rejects stationarity at all levels for the pure time series while it accepts stationarity at all levels for first differences.

The stationary regression formula is again

$$\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \sum_{i=1}^{N} a_i \triangle S_t^i + \triangle \epsilon_t \tag{41}$$

The results of this regression with $N = 10$ are displayed in Table 12.

Table 12: R^2s and F-statistics of the regression from $\triangle S$ onto $\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}$, $N = 10$

	Feb 11	May 11	Jul 11
R^2	0.14	0.06	0.09
F-statistic	1.96	0.69	1.09

The critical value of the F-distribution at the 95% level is 1.91. The hypothesis of zero coefficients is accepted for the May and July forwards, although not for the February contract. Values and t-statistics of the individual coefficients (for May 2011) are provided in Table D.20 and confirm zero coefficients, insignificance and thus non-measurability.

4.3.3. Moratorium: Discussion of results

Figure 13 shows the residual for the forward with delivery in July 2011 in combination with key dates. In the previous section we found this residual to be neither zero nor measurable and so we claim this is the information premium.

It is slightly negative for the first one and a half months of the contract's lifetime. Then, it jumps following the Moratorium oscillating around 6.00 Euro. Still, even after the final decision to shut down the seven old plants (and before the beginning of the delivery period) the residual tends back to zero. We remark that the information premium is a function in time and it seems that market sentiment about the consequences of the new policies
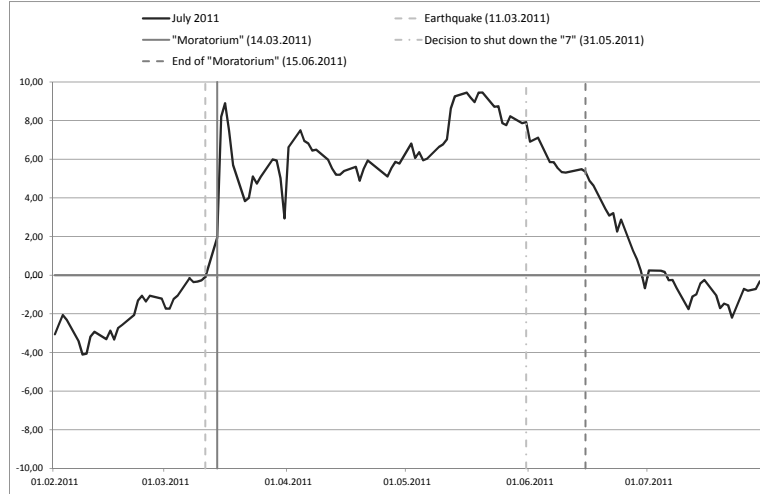
Figure 13: **The information premium** $I_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2)$**.** Corresponding to the July 2011 forward contract.

changed over the lifetime of the forward. There are a number of reasons why the information premium would tend to zero after some time, especially after the legislator clarified the situation (even though by deciding the permanent shut-down). Firstly, Germany used to be an exporter of electricity with an installed capacity much larger than even domestic demand peaks in winter. Secondly, the delivery period of this forward (i.e. July) coincides with the usual yearly maintenance of most nuclear power plants, so that, for this month, the impact would be minimal. Thirdly, as mentioned in the introduction, three of the reactors were offline and remained offline. Figure 13 shows that this change in market sentiment took place during June 2011 which is in line with the information premium identified for the May 2011 contract, i.e. Figure 12(c). This residual remained at the increased price for the rest of its lifetime.

4.4. Additional empirical investigations and discussion

4.4.1. Further regressors

As the historical filtration does not only represent today's information but also that of the past one might claim that we only showed the residual to be non-Markovian in the spot price. Thus, we add moving averages of the spot of different lengths to the list of regressors to include past spot

knowledge. The regression function is then

$$\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = a_1 M A_l(S_t) + \sum_{i=2}^{N+1} a_i \triangle S_t^i + \epsilon_t \tag{42}$$

where l denotes the length of the moving average; we tried lengths $l = 2, 4, 7, 10, 30$ days. For both data sets we found the general result was not altered and the moving average's coefficient was significantly zero.

To answer the question whether the price movements we identified as the information premium were not trivially induced by some correlated time series we also add polynomials of different assets and commodities to the regression. This is to ensure that the introduction of CO_2 certificates and the Moratorium constituted the additional future information. These time series are Brent oil spot (EUETS data set), EEX gas (Moratorium data set) and the DAX stock index (EUETS data set). None of these add largely to the quality or the significance of the regressions (less than 5% R^2, mostly insignificant coefficients).

After the results in this subsection and the last sections one might further ask the question whether our test can actually reject the null hypothesis of non-measurable residuals at all. We will discuss such robustness issues in more detail in Section 4.4.2. Now, we will quickly discuss a related issue, namely the results of adding DAX index (the German stock index) polynomials to the regression on the Moratorium data set, i.e.

$$\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \sum_{i=1}^{N_1} a_i \triangle DAX_t^i + \sum_{j=N_1+1}^{N_2} a_j \triangle S_t^{j} + \epsilon_t \tag{43}$$

Table 13 provides R^2s and F-statistics for the February and May 2011 forwards.

Table 13: R^2s and F-statistics for regressions including the DAX index for February and May 2011 forwards

Forward	Regression	R^2	F-stat
Feb 2011	$N_1 = 1$, $N_2 = 1$	0.007	2.48
Feb 2011	$N_1 = 5$, $N_2 = 15$	0.25	2.38
May 2011	$N_1 = 1$, $N_2 = 1$	0.08	10.52
May 2011	$N_1 = 5$, $N_2 = 15$	0.62	11.97

For the May 2011 forward, the huge increase in R^2 and the F-statistic is surprising, even more so as it does not feature for a pre-Moratorium forward

35

such as February 2011. Also, the t-statistics of DAX monomials of large degree are extremely significant (-2.7 for the coefficient of DAX^4, 5.6 for DAX^5). We would have to reject the non-measurability property in this case which is surprising. Still, the reason for this can be found in the week after the earthquake. The earthquake itself occurred on a Friday whereas the Moratorium and its consequent rise in forward prices was the Monday and Tuesday thereafter. Prices decreased to some extend during the week as market participants realised more clearly the consequences for the German electricity market. For example the price of the May 2011 forward was 50.88 Euro on Friday, it jumped to 61.95 on Tuesday and settled to a level around 58.00 Euro by the end of that week. Exactly the opposite took place on the stock exchange. On Friday, the DAX was at 6981 points. The stock exchange closed for the weekend and when it reopened on Monday and Tuesday traders had a first impression of the long-term damages and their impact on the Japanese economy - the DAX fell by more than 400 (i.e. 5.7%) to 6513 points. Less than two weeks later the DAX had regained all losses. The extreme losses/gains and increases/descreases respectively are responsible for the high R^2 and the significance of the result.

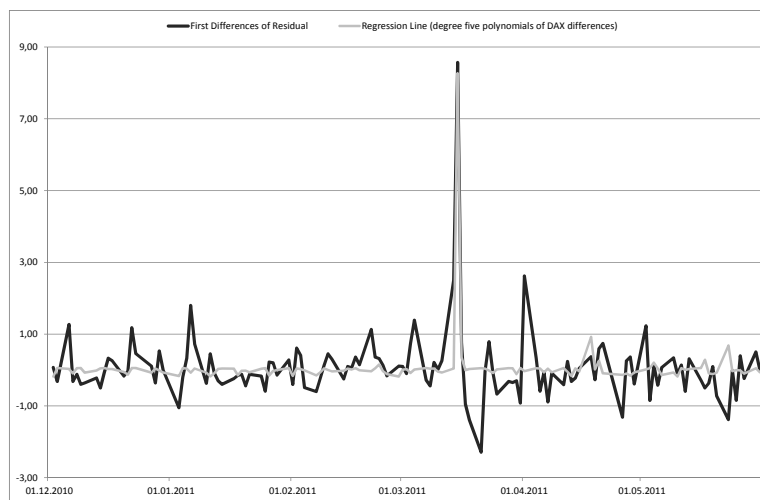Figure 14: **May 2011 residuals vs regression line.** With respect to DAX polynomials of degree five. Including 15/03/2011.

This is further illustrated by Figure 14 and Figure 15 which show the resulting regression function for polynomials of the DAX residuals up to degree five. Figure 14 shows that the regression line is almost constant zero but fits very well the Tuesday after the earthquake. The least-squares
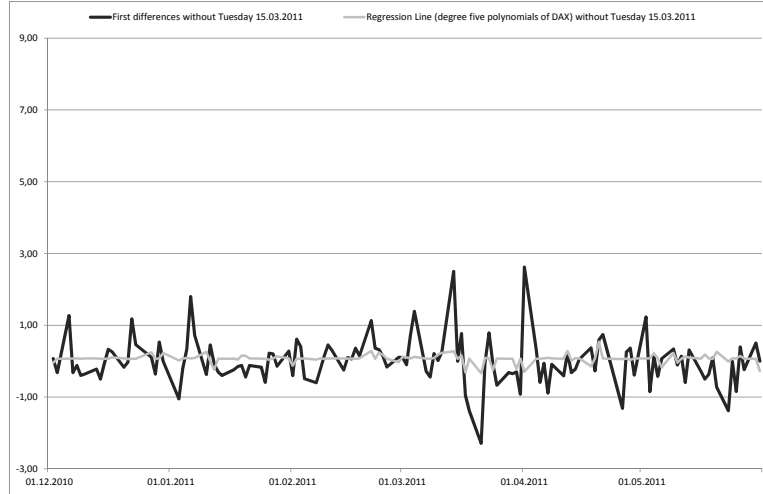
Figure 15: **May 2011 residuals vs regression line.** With respect to DAX polynomials of degree five. Not including 15/03/2011.

method then produces a high R^2 of 56% due to this outlier. This result will be even clearer with monomials of higher degree: the line is even more constant and all parameters will be very close to but not significantly zero. On the other hand, replacing the value of Tuesday 15/03/2011 with zero and repeating the regression yields a R^2 of only 4% (graphs in Figure 15) as well as insignificant zero coefficients. Hence, by deleting the crucial date of this data set we have reduced the R^2 by more than 50%, i.e. one could then claim the non-measurability property again.

4.4.2. Robustness of test and simulation study

As mentioned in the previous section it remains to check whether our test is able to reject the existence of an information premium if there is none. Therefore, we will conduct a simulation study.

We will assume the spot price evolves according to a standard Gaussian Ornstein-Uhlenbeck process with constant parameters

$$dS_t = \alpha(\mu - S_t)dt + \sigma dW_t^1 \tag{44}$$

where W_t^1 is a standard Brownian motion. We define another process Z^1 by

$$dZ_t^1 = -\alpha_{Z^1} Z_t^1 dt + \sigma_{Z^1} dW_t^{Z^1} \tag{45}$$

37

where $W_t^{Z^1}$ is again a Brownian motion which is independent of W_t^1. Let W_t^2 be yet another Brownian motion and let $dW_t^1 dW_t^2 = \rho dt$ be the correlation coefficient. Then we define the process Z_t^2 by

$$dZ_t^2 = -\alpha_{Z^2} Z_t^2 dt + \sigma_{Z^2} dW_t^2 \tag{46}$$

Now, we construct the forward price as follows

$$F(t) = \mu + p_1 Z_t^1 + p_2 Z_t^2 \tag{47}$$

where $p_1, p_2 \in [0, 1]$ are constants.

The motivation behind this construction is as follows: we have seen before that due to the large rate of mean reversion calculated forward prices tend to be almost constant in terms of t. This is why we choose constant μ to be the first building block of the forward price. We note that with a constant change of measure this value is not very important. Both processes Z_t^1 and Z_t^2 are OU processes around zero and resemble the random shocks in the forward price. Z_t^1 is independent of S_t and may be interpreted as part of the information premium or some other non-measurable deviations. Z_t^2 on the other hand is an OU process induced by the spot price depending on ρ.

Depending on the parameters but in particular on the choice of ρ, p_1, p_2 we expect to see similar regression results as before but also clear rejections of the existence of information premia. We will repeat all simulation experiments 200 times and compare mean statistics. Also, we will use the polynomial basis of degree ten again.

We choose OU parameters similar to those extracted from market data or which make sense economically, respectively. We set $\alpha = 0.5$ and $\sigma = 5.0$. Furthermore, we set $\alpha_{Z^2} = 0.3$ and $\sigma_{Z^2} = 3.0$ as we expect forward prices to be less volatile. For now, we also let $p_1 = 0, p_2 = 0.5$ so that the random innovations of the forward do not include non-measurable disturbances. Mean R^2s for some values of ρ are illustrated in Table 14

Table 14: Simulation results (R^2) for different ρ and $p_1 = 0$

ρ	1	0.8	0.5	0
R^2	0.94	0.62	0.27	0.05
R^2 (1st diff.)	0.98	0.65	0.28	0.05

As expected with this simplified setup regressing the forward on the spot polynomial yields high R^2s for high coefficients of correlation with

a near-zero R^2 for zero correlation. Using first differences yields slightly higher R^2s. Examining the values of the coefficients as well as the t-stats of the pure regression we find that for non-zero ρ many coefficients take extremely large absolute values whereas none of the t-stats are significant. For first differences on the other hand only the value of the coefficient of S_t is significantly different from zero (as indeed expected by construction). All F-statistics except those for a zero ρ reject zero coefficients.

We will now introduce independent random shocks by setting $p_1 = 0.5$ and we let $\alpha_{Z^1} = 0.3$ and $\sigma_{Z^1} = 3.0$ (for comparability). Repeating the experiments from above we get the results as summarised in Table 15.

Table 15: Simulation results (R^2) for different ρ and $p_1 = p_2 = 0.5$

ρ	1	0.8	0.5	0
R^2	0.49	0.34	0.16	0.05
R^2 (1st diff.)	0.51	0.35	0.17	0.05

In order to classify the results of the previous sections we are now going to consider a series of experiments in which we will assume $\rho = 1.0$ and modify the value of p_2 ceteris paribus. The question we would like to answer by conducting this experiment is, what type of setup would yield results of the same quality as in Section 4.2 and Section 4.3. Table 16 provides the facts and figures.

Table 16: Simulation results (R^2) for different $\rho = 1$, $p_1 = 0.5$ and different p_2

p_2	0	0.1	0.25	0.5	1.0
R^2	0.05	0.09	0.22	0.49	0.76
R^2 (1st diff.)	0.05	0.08	0.23	0.51	0.79

For example, remembering the forward with delivery in January 2008 our regression (with stationary first differences) had an R^2 of 0.07. Roughly speaking, looking at Table 16, this would correspond to the situation in which "one sixth" of variations in the residual were induced by the spot whereas the rest was induced by some non-measurable other source which seems realistic for our hypothesised information premium around that time. A similar picture can be drawn for $\rho \neq 1$ by looking at Table 15.

Summarising, by a simple set up and a simulation experiment as conducted in this section we have shown two things: firstly, we can replicate the results from previous sections, in particular from the Japan earthquake

and consequent "Atom Moratorium" as well as from the introduction of CO_2 certificates in 2008. Secondly, we showed that we can construct non-trivial scenarios for which the hypothesis of zero regression coefficients is rejected. Consequently our test would reject the existence of an information premium.

4.4.3. Concluding remarks

Here are some last remarks considering the method we propose in this paper:

The regression method used to show the non-measurability of the residual works with any orthonormal system whereas we only consider the polynomial basis. Some research has been done as to the benefits of using different bases in the case of the classical Least-Squares-Monte-Carlo. For example, this is discussed in section 8.3 of Longstaff and Schwartz (2001). Still, in our case this choice is more or less irrelevant. After all, we are only interested in the significance properties and not in the speed of convergence of the algorithm. There is a bijection between different orthonormal systems and by making use of a sufficient number of basis elements we can constrain ourself to this choice. We have mostly used ten basis elements in this paper and claim that this is more than enough (even more, for a data set of length 640). Indeed, increasing the number of basis elements does not alter the result.

Furthermore, one could postulate that our findings are maybe due to initially fitting the spot price badly or a bad spot model in general. Still, we used observed data for our spot and forward time series. The only objects depending on the spot model were the expectations taken under $\mathbb{Q}$ and $\mathcal{F}_t$. But these expectations which were subtracted from observed data to isolate the information premium were piecewise constant (remember Figure 8 and the reasoning in Section 3). Hence, we claim that the regression results would not be altered for a different spot model.

5. Conclusion

This paper provided a thorough empirical analysis of the *information premium* and showed that the premium exists on electricity markets.

These markets exhibit future information priced into forward but not incorporated in spot prices. In Financial Mathematics the flow of information is modelled by filtrations. Still, scientific literature has only considered the historical filtration and thus the classical but non-appropriate

spot-forward relationship. To understand the relationship between spot and forward prices on electricity markets, and thus their fundamental structure, we find it necessary to take larger filtrations into consideration. Thus, we complement the notion of the well-known risk premium by the notion of an information premium. The information premium is consequently defined to be the difference between the forward price under a large (market) filtration and a small (historical) filtration. Benth and Meyer-Brandis (2009) found some fundamental properties of the information premium and calculated it for a two-factor arithmetic model making good use of the theory of the *Enlargement of Filtrations*.

We designed a new method in order to show the existence of the information premium empirically. To this goal we conducted a measure change and isolated a residual from observed forward prices. We had to test this residual for the premium's most important property: the non-measurability under the historical filtration. This was done by trying to express the residual in terms of an orthonormal basis of the spot price by means of regression. To the best of our knowledge there is no other literature that discusses non-measurability empirically.

We applied our test to the data of two motivating examples. The first of these examples was the beginning of the second phase of CO_2 certificates on the EEX in 2008. The market expected a steep rise in electricity prices which was observable in forward prices even before 2008 - but not in spot prices. The second example featured the German "Atom Moratorium", i.e. the shut-down of seven older nuclear power plants. This piece of legislation was introduced to reevaluate German policy after the Fukushima disaster. It resulted in increased forward prices while the spot price remained unchanged. For both examples the test confirmed the existence of a significant information premium and the residual was plausible with respect both, size and shape.

References:

Applebaum, D., 2004. Lévy Processes and Stochastic Calculus. Cambridge University Press.

Benth, F. E., Cartea, A., Kiesel, R., 2008a. Pricing forward contracts in power markets by the certainty equivalence principle: Explaining the sign of the market risk premium. Journal of Banking and Finance 32 (10), 2006–2021.

Benth, F. E., Kallsen, J., Meyer-Brandis, T., 2007. A non-gaussian ornstein-uhlenbeck process for electricity spot price modeling and derivatives pricing. Applied Mathematical Finance 14 (2), 153–169.

Benth, F. E., Kiesel, R., Nazarova, A., 2010. A critical empirical study of three electricity spot price models. to appear in Energy Economics.

Benth, F. E., Meyer-Brandis, T., 2009. The information premium for non-storable commodities. Journal of Energy Markets 2 (3), 111–140.

Benth, F. E., Ŝaltyte-Benth, J., Koekebakker, S., 2008b. Stochastic Modelling of Electricity and Related Markets. World Scientific.

Bessembinder, H., Lemmon, M. L., 2002. Equilibrium pricing and optimal hedging in electricity forward markets. The Journal of Finance 57, 1347–1382.

Cartea, A., Figueroa, M. G. ., Geman, H., 2009. Modelling electricity prices with forward looking capacity constraints. Applied Mathematical Finance 16, 103–122.

Diko, P., Lawford, S., Limpens, V., 2006. Risk premia in electricity forward prices. Studies in Nonlinear Dynamics and Econometrics 10 (3).

Douglas, S., Popova, J., 2008. Storage and the electricity forward premium. Energy Economics 30, 1712–1727.

EEX, 2011. www.eex.com.

Furiò, M. D., Meneu, V., 2010. Expectations and forward risk premium in the spanish deregulated power market. Energy Policy 38 (2), 784–793.

Ito, K., 1978. Extension of stochastic integrals. In: Proceedings of international symposium on stochastic differential equations. Wiley & Sons, Ltd, pp. 95–109.

Jeulin, T., 1979. Grossissement d'une filtration et applications. Séminaire de probabilités de Strasbourg 13, 574–609.

Jeulin, T., Yor, M., 1978. Grossissement d'une filtration et semi-martingales: formules explicites. Séminaire de probabilités de Strasbourg 12, 78–97.

Kolos, S. P., Ronn, E. I., 2008. Estimating the commodity market price of risk for energy prices. Energy Economics 30, 621–641.

Kou, S. G., 2002. A jump diffusion model for option pricing. Management Science 48, 1086–1101.

Longstaff, F. A., Schwartz, E. S., 2001. Valuing american options by simulation: A simple least-squares approach. Review of Financial Studies 14 (1), 113–147.

Longstaff, F. A., Wang, A., 2004. Electricity forward prices: A high-frequency empirical analysis. Journal of Finance 59, 1877–1900.

Meyer-Brandis, T., Tankov, P., 2008. Multi-factor jump-diffusion models of electricity prices. International Journal of Theoretical and Applied Finance 11 (05), 503–528.

Protter, P., 2005. Stochastic Integration and Differential Equations, 2nd Edition. Springer.

Royden, H. L., 1968. Real Analysis. The Macmillan Company.

Torró, H., Lucia, J., 2011. On the risk premium in nordic electricity futures prices.

International Review of Economics and Finance 20 (4), 750–763.

Appendices

Appendix A. The Lévy part of the Spot model

The Lévy process L_t contributing to the spot model follows the Kou model Kou (2002). It is a compound Poisson process with double-exponentially distributed jumps.

Definition Appendix A.1. *__Lévy Component__. Define*

$$L_t = \sum_{i=1}^{N_t} D_i$$

where N_t is a Poisson process with intensity $\lambda > 0$ and D_i are the independent, identically double-exponentially distributed jump sizes with the density of the double-exponential distribution given by

$$f_D(x) = p\eta_1 e^{-\eta_1 x} \mathbb{1}_{x \geq 0} + q\eta_2 e^{-\eta_2 |x|} \mathbb{1}_{x \leq 0}$$

where $p + q = 1$ and $\eta_1, \eta_2 \geq 0$.

To calculate expectations we use the following result:

Lemma Appendix A.1. *__Log-moment generating function.__ The log-moment generating function of L_1 is given by*

$$\phi_{L_1}(u) = \log \mathbb{E}[e^{uL_1}] = \frac{-\lambda pu}{u - \eta_1} + \frac{\lambda qu}{u - \eta_2}, \qquad (A.1)$$

so that

$$\mathbb{E}(L_1) = \phi'_{L_1}(0) = \frac{\lambda p}{\eta_1} - \frac{\lambda q}{\eta_2}$$

For the application of the Esscher transform we need the following:

Lemma Appendix A.2. *__Integrability condition.__ We can apply Esscher transform to L_t if*

$$\int_{|x| \geq 1} e^{\theta x} \mu(dx) < \infty \qquad (A.2)$$

is satisfied. Here, μ denotes the density and θ the Esscher transform parameter.

The following lemma ensures the integrability condition for the Esscher transform is satisfied by our choice of the Lévy process (given constant θ_L):

Lemma Appendix A.3. *Kou model and integrability condition.*
The compound Poisson process with doubly exponentially distributed jump heights satisfies the integrability condition (A.2) *for the Esscher transform for $\theta_L < \eta_1$ and $\theta_L > -\eta_2$ because*

$$
\int_{|x|\geq 1} e^{\theta x} \mu(dx) = \int_{-\infty}^{-1} e^{\theta x} \lambda q \eta_2 e^{-\eta_2 |x|} dx + \int_{1}^{\infty} e^{\theta x} \lambda p \eta_1 e^{-\eta_1 x} dx
$$

$$
= \lambda q \eta_2 \int_{-\infty}^{-1} e^{(\theta+\eta_2)x} dx + \lambda p \eta_1 \int_{1}^{\infty} e^{(\theta-\eta_1)x} dx
$$

$$
= \frac{\lambda q \eta_2}{\theta + \eta_2} e^{-\eta_2 - \theta} - \frac{\lambda p \eta_1}{\theta - \eta_1} e^{\theta - \eta_1}
$$

$$
< \infty
$$

Appendix B. Forward prices with delivery period

We use the following deterministic functions to calculate forward prices (using the same notation as Benth et al. (2008a)):

$$
\bar{\alpha}(t, T_1, T_2) = \begin{cases} -\frac{1}{\alpha} \left(e^{-\alpha(T_2-t)} - e^{-\alpha(T_1-t)} \right) & t \leq T_1 \\ -\frac{1}{\alpha} \left(e^{-\alpha(T_2-t)} - 1 \right) & t > T_1 \end{cases} \tag{B.1}
$$

$$
\bar{\beta}(t, T_1, T_2) = \begin{cases} -\frac{1}{\beta} \left(e^{-\beta(T_2-t)} - e^{-\beta(T_1-t)} \right) & t \leq T_1 \\ -\frac{1}{\beta} \left(e^{-\beta(T_2-t)} - 1 \right) & t > T_1 \end{cases} \tag{B.2}
$$

$$
\hat{\beta}(t, T_1, T_2) = \begin{cases} \frac{1}{\beta} \left(T_2 - T_1 + \frac{1}{\beta} \left(e^{-\beta(T_2-t)} - e^{-\beta(T_1-t)} \right) \right) & t \leq T_1 \\ \frac{1}{\beta} \left(T_2 - t + \frac{1}{\beta} \left(e^{-\beta(T_2-t)} - 1 \right) \right) & t > T_1 \end{cases} \tag{B.3}
$$

This is the proof for the explicit expression of the forward price under measure $\mathbb{Q}$ for delivery between T_1 and T_2:

Proof of Proposition 2.2. The first three terms of the statement are easily

calculated. Further, considering the Lévy part yields

$$\mathbb{E}^{\mathbb{Q}}\left[\int_{T_1}^{T_2}\int_t^u e^{-\beta(u-s)}dL_s du\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[\int_{T_1}^{T_2}\int_t^u e^{-\beta(u-s)}dL_s \frac{Z_L(u)}{Z_L(t)}\,du\right]$$

$$= \int_{T_1}^{T_2}\exp\left(-\int_t^u \phi(\theta_L(s))ds\right)\mathbb{E}^{\mathbb{P}}\left[\int_t^u e^{-\beta(u-s)}dL_s \exp\left(\int_t^u \theta_L(s)dL_s\right)\right]du$$

$$= \int_{T_1}^{T_2}\exp\left(-\int_t^u \phi(\theta_L(s))ds\right)$$
$$\frac{d}{dx}\left(\exp\left(\log\left(\mathbb{E}^{\mathbb{P}}\left[\exp\left(\int_t^u \left(xe^{-\beta(u-s)}+\theta_L(s)\right)dL_s\right)\right]\right)\right)\right)\Bigg|_{x=0}du$$

where we added an artificial differentiation. This allows to further calculate

$$= \int_{T_1}^{T_2}\exp\left(-\int_t^u \phi(\theta_L(s))ds\right)\frac{d}{dx}\left(\exp\left(\int_t^u \left(\phi\left(xe^{-\beta(u-s)}+\theta_L(s)\right)ds\right)\right)\right)\Bigg|_{x=0}du$$

$$= \int_{T_1}^{T_2}\exp\left(-\int_t^u \phi(\theta_L(s))ds\right)$$
$$\left(\int_t^u e^{-\beta(u-s)}\phi'\left(xe^{-\beta(u-s)}+\theta_L(s)\right)ds \exp\left(\int_t^u \phi\left(xe^{-\beta(u-s)}+\theta_L(s)\right)ds\right)\right)\Bigg|_{x=0}du$$

$$= \int_{T_1}^{T_2}\exp\left(-\int_t^u \phi(\theta_L(s))ds\right)\left(\int_t^u e^{-\beta(u-s)}\phi'\left(\theta_L(s)\right)ds \exp\left(\int_t^u \phi\left(\theta_L(s)\right)ds\right)\right)du$$

$$= \int_t^{T_2}\phi'\left(\theta_L(s)\right)\bar\beta(s,T_1,T_2)ds$$

A similar approach provides the result for the Gaussian part. After collecting terms the proposition is proved. $\qquad\square$

The following corollary provides the forward price under $\mathbb{P}$ as a special case of proposition 2.2.

Corollary Appendix B.1. *Forward price under $\mathbb{P}$. The forward price with delivery under the historical measure $\mathbb{P}$ is given by*

$$F(t,T_1,T_2) = \frac{1}{T_2-T_1}\left(\int_{T_1}^{T_2}\Lambda(u)du + \bar\alpha(t,T_1,T_2)X_t + \bar\beta(t,T_1,T_2)Y_t + \phi'(0)\hat\beta(t,T_1,T_2)\right)$$

Here, $0 \le t \le T_1 < T_2$.

Appendix C. The Seasonality Function

In the empirical analysis we used the following seasonality function $\Lambda(t)$:

$$\Lambda(t) = b_0 + b_1 t + b_2 \cos(\tfrac{2\pi}{365}t) + b_3 \sin(\tfrac{2\pi}{365}t) + b_4 \cos(\tfrac{2\pi}{182}t) + b_5 \sin(\tfrac{2\pi}{182}t)$$
$$+ b_6 \mathbb{1}_{\{t \ mod \ 7=0\}}(t) + b_7 \mathbb{1}_{\{t \ mod \ 7=1\}}(t) + \ldots + b_{12} \mathbb{1}_{\{t \ mod \ 7=6\}}(t)$$

The fitted values for the Emissions dataset were:

Table C.17: **Emissions dataset Fitted values of $\Lambda(t)$-parameters**

b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}
20.11	0.09	3.61	-0.95	-2.35	-0.48	6.04	3.20	-6.79	-16.53	3.50	4.91	5.66

The fitted values for the Moratorium dataset were:

Table C.18: **Moratorium dataset Fitted values of $\Lambda(t)$-parameters**

b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}
37.93	0.02	-0.94	1.49	-1.06	1.33	3.50	3.28	2.66	1.43	-4.28	-8.47	1.88

Appendix D. Regression Tables

Table D.19: **January 2008 Forward Regression** $\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = \sum_{i=1}^{N} \triangle a_i S_t^i + \epsilon_t$

	Value	Std. Dev.	t-value	p-value
Intercept	0.09	0.16	0.55	0.58
S	-0.01	0.03	-0.33	0.74
S^2	0.00	0.00	-0.66	0.51
S^3	0.00	0.00	0.22	0.83
S^4	0.00	0.00	0.49	0.62
S^5	0.00	0.00	-0.15	0.88
S^6	0.00	0.00	-0.26	0.79
S^7	0.00	0.00	0.07	0.95
S^8	0.00	0.00	-0.05	0.96
S^9	0.00	0.00	-0.01	0.99
S^{10}	0.00	0.00	0.12	0.91

Table D.20: **May 2011 Forward Regression** $\triangle \hat{I}_{\mathcal{G}}^{\mathbb{Q}}(t, T_1, T_2) = a_0 + \sum_{j=1}^{N} a_j \triangle S_t^j + \epsilon_t$

	Value	Std. Dev.	t-value	p-value
Intercept	0.20	0.17	1.14	0.25
S	0.13	0.13	1.03	0.31
S^2	-0.03	0.04	-0.90	0.37
S^3	-0.01	0.02	-0.64	0.52
S^4	0.00	0.00	0.84	0.40
S^5	0.00	0.00	0.25	0.80
S^6	0.00	0.00	-0.70	0.48
S^7	0.00	0.00	-0.06	0.95
S^8	0.00	0.00	0.49	0.63
S^9	0.00	0.00	-0.03	0.98
S^{10}	0.00	0.00	-0.31	0.76