THE RISK PREMIUM AND THE ESSCHER TRANSFORM IN POWER MARKETS

FRED ESPEN BENTH AND CARLO SGARRA

Abstract. In power markets one frequently encounters a risk premium being positive in the short end of the forward curve, and negative in the long end. Economically it has been argued that the positive premium is reflecting retailers aversion for spike risk, wheras in the long end of the forward curve the hedging pressure kicks in as in other commodity markets. Mathematically, forward prices are expressed as risk-neutral expectations of the spot at delivery. We apply the Esscher transform on power spot models based on mean-reverting processes driven by independent increment (time-inhomogeneous Lévy) processes. It is shown that the Esscher transform is yielding a change of mean-reversion level. Moreover, we show that an Esscher transform together with jumps occuring seasonally may explain the occurence of a positive risk premium in the short end. This is demonstrated both mathematically and by a numerical example for a two-factor spot model being relevant for electricity markets.

1. Introduction

The purpose of the present paper is to investigate some relations between the risk premium and the change of measure in electricity markets. More precisely we shall focus our attention on a peculiar feature of the risk premium, the sign change. In power markets, the usual pricing approach based on the construction of an equivalent martingale measure is not viable any more. Electricity, for example, is a non-storable commodity, so it does not make neither sense to trade in the underlying nor to use hedging arguments. However, since the forward contracts need to have a price dynamics being arbitrage-free, these prices can still be considered to be discounted expectations of the final value of the underlying, but with respect to any equivalent probability measure. Alternatively, it is possible to think about forward prices in terms of risk premium: this is defined as the difference between the forward prices computed with respect to the risk-neutral measure and with respect to the objective measure, respectively (see Geman [19]). Once the probability measure with respect to which the discounted prices must be calculated has been chosen, then the risk premium is defined in a unique way. The peculiarities of the risk premium in energy markets have been thoroughly investigated during the past few years and many evidences have been collected on its behavior. The theory of normal backwardation suggests that producers of a commodity wish to hedge their revenues by selling forwards, so they accept

Date: November 14, 2009.
Key words and phrases. Power Market, Derivatives Pricing, Independent Increments Processes, Esscher Transform, Risk Premium.

a discount on the expected spot price. Thus, we should have the forward prices of a commodity less than the expected value with respect to the objective probability measure. So, in this context, the risk premium should be negative. On the other hand, several authors found evidence of a positive risk premium for contracts with a short time to maturity: Geman and Vasicek [20] investigated the Pennsylvania-New Jersey-Maryland electricity market and justified the existence of a positive risk premium by the market's aversion for the high volatility and consequently willingness to pay high prices to ensure delivery. In the same study, for contracts with longer maturities, the sign of the risk premium changes. Longstaff and Wang [26] perform a non-parametric study of the same market obtaining evidence of a positive risk premium for the short-term contracts. Their study has been extended by Diko, Lawford and Limpens, who analyse risk premia in the three markets EEX, Powernext and Dutch market APX [14]. A term structure for the risk premium is found there, which varies significantly from the short-term maturities to long-term maturities. A recent study with a systematic investigation on the short-term maturity Nord Pool market has been provided by Lucia and Torrò [28] who extend a previous study by Botterud, Bhattacharya and Ilic [9]: they find evidence of time-varying risk premium and of its positivity for short maturities. Benth, Cartea and Kiesel [3] provide an interpretation of the risk premium in electricity market in terms of the certainty equivalent principle and jumps in the spot price dynamics, whereas Benth and Meyer-Brandis [6] provide an information-based approach for explaining the risk premium.

In this paper we are going to provide the mathematical evidence of the risk premium sign change on the basis of the most popular models available in the literature and of the most natural probability measure change: the Esscher transform. This measure change, introduced by Esscher [17] in an actuarial context, has been extensively applied in derivative pricing since the pioneering work by Gerber and Shiu [21], who extended the original idea by Esscher to a Lévy framework. It has been further extended to semimartingale modelling by Kallsen and Shiryaev [24], and its popularity in the Lévy framework is due to two very important reasons: the first is that it enjoys some relevant optimality properties; its close relationship with the minimal entropy martingale measure has been thoroughly investigated in the papers by Esche and Schweizer [16] and by Hubalek and Sgarra [22]; the second, the most important for the present purposes, is that it preserves the independent increment property. This important feature of the Esscher transform, which has been proved for the linear Esscher martingale transform for Lévy processes, still holds when the increments are not stationary any more. This is the main reason to justify our choice of the Esscher Transform as a reference measure change: if the structure of the spot dynamics it is not preserved by the measure change, it is almost impossible to obtain relevant information on the forward prices and in particular any explicit formulas for evaluation. In the framework of spot price dynamics described by independent increment processes we shall show that a measure change performed via the Esscher transform can justify the sign change of the risk premium between short and longer maturities.

We obtain general results on the application of the Esscher transform to power markets. Our spot model cover many of the important cases applied in practice and theory, and we show that the choice of pricing measure provided by the Esscher transform yields analytically tractable models for forward pricing. We treat both geometric and arithmetic models, the latter is also sufficiently flexible for allowing for pricing of power forwards which deliver the underlying over a period. Moreover, it is proven that the Esscher transform corresponds to a change in the mean-reversion level for the spot, much in line with the classical Girsanov transform with a constant drift. In this respect, the Esscher transform is a true generalization of the Girsanov transform to processes with independent increments. The constant Girsanov change of measure seems to be the standard choice in the power markets. The explicit forward prices obviously imply analytical expressions for the risk premia. We analyse the particular case of a two-factor spot model, with the first factor modelling the base component of the prices, while the second is the spike component. Spikes are large sudden price increases being fastly reverted back. We discuss the consequences of different choices of the market price of risk, being the parameters in the Esscher transform. As it turns out, spikes occurring seasonally may explain the occurrence of a positive risk premium in the short end, while in the long end we can have a negative premium. If we are close to a period with high spike intensity, the premium may become positive, while if the spiky period is far in the future, we see a risk premium being negative, or backwardated. However, spike periods are related to bumps in the risk premium curve. Taking into account that the premium is scaled by the current states of the factors driving the spot price, we find that the risk premium can vary stochastically, and that it can have periods of a positive premium in the short end and negative in the long. In fact, the Esscher transform provides a large degree in flexibility even for constant market prices of risk.

Our results are presented as follows. Section 2 will recall the basic modelling framework we are going to assume: we shall introduce two classes of models based on independent increment processes: the geometric and the arithmetic models. In the third section we shall resume the essential features of the Esscher transform for independent increment processes. Furthermore, we present the forward prices obtained in both the geometric and the arithmetic models. Evidence of the risk premium sign change is shown in the main Section 4, where we also provide models explaining the change of risk premium as coming from seasonally varying spike intensities. We consider also forwards delivering over a period (so-called flow forwards), and the question of the Esscher transform being a martingale measure. Finally, in section 7 we conclude and outline some futures perspectives of the present work.

2. General Models for Power Spot Prices

The electricity market models we are going to consider in this paper belong to a wide class of models based on additive or time-inhomogeneous Lévy processes (sometimes called Sato) processes, i.e. stochastic processes with independent increments (II from now on). They can be grouped in two subclasses, the geometric and the arithmetic models.

Suppose that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ is a complete filtered probability space. On this space we have defined independent Brownian motions B_k, $k = 1, \ldots, p$ and independent II processes I_j, $j = 1, \ldots, n$. We briefly recall that the II process $I(t)$ admits the Lévy-Kintchine representation of with cumulant function ψ: For a continuous function θ and $s \geq t$ we have

$$\mathbb{E}\left[\exp\left(\mathrm{i}\int_t^s \theta(u)\,dI(u)\right)\right] = \exp\left(\psi(s,t;\theta(\cdot))\right)\,, \tag{2.1}$$

with

$$\begin{aligned}\psi(s,t;\theta(\cdot)) = \mathrm{i}\int_t^s \theta(u)\,d\gamma(u) - \frac{1}{2}\int_t^s \theta^2(u)\,dC(u)\\ + \int_t^s \int_{\mathbb{R}} \left\{\mathrm{e}^{\mathrm{i}\theta(u)z} - 1 - \mathrm{i}\theta(u)z\mathbf{1}_{|z|\leq 1}\right\}\,\ell(dz,du)\,,\end{aligned} \tag{2.2}$$

where the quantities $\gamma, C, \ell(dz,du)$ are the semimartingale characteristics of the II process (see for example Shiryaev [23]). We remark in passing that usually one states the cumulant function ψ for a constant θ, but in our framework it will be convenient to have the slightly more general notation introduced here. Due to strong seasonality effects like for instance the spike occurrence in power markets, it is reasonable to consider jump processes being time-inhomogeneous. This allows for modeling for example the spikes observed in the winter period in the Nordic electricity spot market Nord Pool.

Both classes of models we are going to discuss are described by Ornstein-Uhlenbeck processes driven by an II, eventually including one or more Wiener processes B_k.

The geometric models describe the dynamics of electricity prices in the following way:

$$\ln S(t) = \ln\Lambda(t) + \sum_{i=1}^m X_i(t) + \sum_{j=1}^n Y_j(t), \tag{2.3}$$

where, for $i = 1, ...m$,

$$dX_i(t) = \left(\mu_i(t) - \alpha_i(t)X_i(t)\right)dt + \sum_{k=1}^p \sigma_{ik}(t)dB_k(t) \tag{2.4}$$

and, for $j = 1, ..., n$,

$$dY_j(t) = \left(\delta_j(t) - \beta_j(t)Y_j(t)\right)dt + \eta_j(t)dI_j(t) \tag{2.5}$$

The deterministic seasonal price level is modelled by the function $\Lambda(t)$, which is assumed to be positive and continuously differentiable, while the coefficients $\mu_i, \alpha_i, \delta_j, \beta_j, \eta_j$ are assumed to be continuous functions of t. From the modelling viewpoint it looks quite natural to choose $\mu_i = 0$ and $\delta_j = 0$, since the Ornstein-Uhlenbeck processes should ideally revert toward zero in order to have the seasonality function $\Lambda(t)$ as the mean price level; moreover it is also reasonable to assume constant speed of mean reversions, i.e. to assume the coefficients α_i, β_j to be constant, and moreover, non-negative. For the sake of generality, we shall stick with the time-dependent coefficients in our analysis. We remark that similar general multi-factor models have been proposed and applied to pricing power and commodity derivatives by Benth et al. [2] and Crosby [11].

In this model both $S(t)$ and $\ln S(t)$ are semimartingale processes. By assuming that the initial conditions for X_i and Y_j are such that:

$$\sum_{i=1}^{m} X_i(0) + \sum_{j=1}^{n} Y_j(0) = \ln S(0) - \ln \Lambda(0),$$

an explicit representation of $S(t)$ is given by:

$$S(t) = \Lambda(t) \exp\left(\sum_{i=1}^{m} X_i(t) + \sum_{j=1}^{n} Y_j(t)\right), \tag{2.6}$$

where, for $i = 1, ..., m$,

$$X_i(t) = X_i(0)\mathrm{e}^{-\int_0^t \alpha_i(s)\,ds} + \int_0^t \mu_i(u)\mathrm{e}^{-\int_u^t \alpha_i(s)\,ds}du + \sum_{k=1}^{p}\int_0^t \sigma_{ik}(u)\mathrm{e}^{-\int_u^t \alpha_i(s)\,ds}dB_k(u) \tag{2.7}$$

and, for $j = 1, ..., n$,

$$Y_j(t) = Y_j(0)\mathrm{e}^{-\int_0^t \beta_j(s)\,ds} + \int_0^t \delta_j(u)\mathrm{e}^{-\int_u^t \beta_j(s)\,ds}du + \int_0^t \eta_j(u)\mathrm{e}^{-\int_u^t \beta_j(s)\,ds}dI_j(u)\,. \tag{2.8}$$

We remark that many models proposed for electricity market belong to the class just introduced: the one-factor model of Schwarz [30] is a particular case with $p = m = 1, n = 0$, that is,

$$S(t) = \Lambda(t)\exp(X_1(t))\,,$$

with $dX_1(t) = (\mu_1 - \alpha_1 X_1(t))\,dt + \sigma_1\,dB_1(t)$. The coefficients of the model are constant. A jump process version of the Schwartz model has been applied to oil and gas spot price modelling in Benth and Saltyte-Benth [7]. The model takes the form ($m = 0$, $n = 1$)

$$S(t) = \Lambda(t)\exp(Y_1(t))\,,$$

with a constant-coefficient jump-process dynamics $dY_1(t) = (\delta_1 - \beta_1 Y_1(t)\,dt + dI_1(t)$. Here, I_1 is a normal inverse Gaussian Lévy process (see Barndorff-Nielsen [1] for more on the normal inverse Gaussian Lévy process). Similarly the model introduced by Cartea and Figueroa [10] corresponds to the choice of $I_1(t)$ as the sum of a Brownian motion and a compound Poisson processs. Their model was fitted to electricity data in the UK market using ingenious filtering techniques including both spot and forward price data. In Lucia and Schwarz [27] a two-factor model is applied to the Nordic electricity market Nord Pool. The model can be recognized to belong to our proposed class with $m = p = 2, n = 0$, that is, two Brownian motion driven factors:

$$S(t) = \Lambda(t)\exp(X_1(t) + X_2(t))\,,$$

where the long-term variations are controlled by the process $dX_1(t) = (\mu_1 - \alpha_1 X_1(t)\,dt + \sigma_1\,(\rho\,dB_1(t) + \sqrt{1-\rho^2}\,dB_2(t))$ and the short-term variations by $dX_2(t) = (\mu_2 - \alpha_2 X_2(t))\,dt + \sigma_2\,dB_1(t)$. The coefficients are supposed to be constant in their model, and the long-term level a non-stationary process, which is achieved in our set-up by letting $\alpha_1 = 0$. Also note that the noises in the two processes are correlated by ρ. The model proposed by Eberlein and Stahl [15] is a slight modification of the Lucia and Schwartz dynamics, letting the short-term variations be modelled by a jump process instead. We obtain their model by setting $m = n = p = 1$, with the mean-reversion $\alpha_1 = 0$, and letting $I_1(t)$ be a Lévy process of hyperbolic type. The model was analysed empirically on data observed in the German electricity market EEX with the purpose of Value-at-Risk calculations.

The second subclass on which we shall focus our attention is represented by the arithmetic models. The commodity dynamics in these models is described by the following expression:

$$S(t) = \Lambda(t) + \sum_{i=1}^{m} X_i(t) + \sum_{j=1}^{n} Y_j(t)\,, \tag{2.9}$$

where $X_i(t), Y_j(t), i = 1, \ldots, m, j = 1, \ldots, n$ are given as before and the seasonality function satisfies the same conditions required for geometric models. The process $S(t)$ is again a semimartingale and assuming now the initial conditions verifying:

$$\sum_{i=1}^{m} X_i(0) + \sum_{j=1}^{n} Y_j(0) = S(0) - \Lambda(0),$$

the explicit representation of $S(t)$ will be obtained simply inserting the expressions of $X_i(t), Y_j(t)$ given before into equation (2.9).

Arithmetic models has not gained the same popularity as the geometric models in order to describe the commodity behavior in energy markets and this is to a large extent due to the possibility to obtain negative prices in this framework. On the other hand, negative prices can sometimes arise in energy markets. It is not uncommon in electricity spot markets since produced power cannot be stored. But even in gas markets one has observed short periods where one get paid for consuming, see Financial Times [18] for an article on such an incident.

A model recently proposed included in the arithmetic class is that introduced by Benth, Kallsen and Meyer-Brandis [4], in which the probability to reach negative prices is zero. This model is obtained by setting $m = 0$, and by interpreting the seasonality function $\Lambda(t)$ as a floor towards which the processes $Y_j(t)$ revert. Moreover, the II processes $I_j(t)$ are assumed to be pure jump increasing process. By letting $\delta_j = 0$ for $j = 1, ..., n$, $Y_1(0) = S(0) - \Lambda(0)$ and $Y_j(0) = 0$ for $j = 2, ..., n$, we obtain the following explicit representation of the spot price dynamics:

$$S(t) = \Lambda(t) + \sum_{j=1}^{n} Y_j(t) \tag{2.10}$$

Such a model was fitted to spot price data observed at the German electricity market EEX by Benth, Kiesel and Nazarova [5] (see also Meyer-Brandis and Tankov [25]).

We shall assume an integrability condition for the processes $I_j(t)$ in order to apply the Esscher transform measure change. The condition we will assume is the following:

Condition 1. *For each $j = 1, ..., n$ there exists a constant $c_j > 0$ such that:*

$$\int_0^T \int_1^\infty \{e^{c_j z} - 1\} \, \ell_j(dz, du) < \infty \tag{2.11}$$

These exponential integrability conditions on the jump measures ℓ_j ensure the existence of exponential moments of the II processes $I_j(t)$ driving the dynamics for the Y's. This makes it possible to define the Esscher transform for these processes, since the transform is based on an exponential tilting of the distribution. We shall later be more specific on the constants c_j, which will differ slightly depending on the model we consider.

3. The Esscher Transform for II Processes and forward prices

The first fundamental theorem of asset pricing states that in a market model the no-arbitrage requirement is equivalent to the existence of at least one equivalent probability measure which turns into local martingales all the tradeable (discounted) asset processes. The arbitrage-free price of any of them can be calculated as the discounted expectation of its payoff with respect to one of these probability measures (see Delbaen and Schachermayer [12]-[13]). On the other hand, according to the second fundamental theorem of asset pricing the equivalent martingale measure is unique when the market model is complete. The models introduced in the previous section exhibit incompleteness. Theoretically, they are incomplete because we have jumps and multidimensional Brownian motions driving the dynamics. Moreover, as it was already pointed out in the introduction, the non-storability of the underlying commodity gives the electricity market some peculiar features. This fact rules out the possibility to hedge derivatives by trading in the underlying. Hence, in this context *any* measure Q equivalent to the objective (or market) probability P is risk-neutral, and the underlying asset process (since it is not tradeable) does not need to be a martingale with respect to the risk-neutral measure. This makes the class of potential pricing measures rather wide and some restrictions must be imposed in order to perform calculations. A natural choice is the Esscher transform, which has been extensively applied in different pricing models, and whose optimal properties have been thoroughly investigated in the Lévy framework.

The crucial property which makes the Esscher transform an interesting measure change is the structure preservation of the model dynamics. The semimartingale characteristics of an II process are changed after the transformation, but the independence of the increments is preserved. When working in incomplete market models the usual approach is based on the construction of a measure change which turns the price process into a martingale, and the parameters must be chosen such that the martingale requirement is satisfied. For general semimartingales models two kinds of Esscher martingale transform exist: one turning into a martingale the ordinary exponential process, the other turning into a martingale the stochastic exponential process: the first has been called by some authors the *exponential Esscher martingale transform*, the second the *linear Esscher martingale transform*. In the electricity market framework the Esscher martingale transform is not required, simply because the electricity price process need not to be a martingale under the transformed measure, and the Esscher transform is just the most natural way to construct a measure change in which the new measure is equivalent to the old one and the process with respect to the new measure has different parameters, but still enjoys the independent increment property.

Let us define the Esscher transform for II processes. Let $\theta(t)$ be a $(p+n)-$dimensional vector of real-valued continuous functions on $[0,T]$:

$$\theta(t) = \left(\widehat{\theta}_1(t), \ldots, \widehat{\theta}_p(t), \widetilde{\theta}_1(t), \ldots, \widetilde{\theta}_n(t)\right) . \tag{3.1}$$

Define the stochastic exponential by the following relation:

$$Z^\theta(t) := \prod_{k=1}^{p} \widehat{Z}_k^\theta(t) \times \prod_{j=1}^{n} \widetilde{Z}_j^\theta(t), \tag{3.2}$$

where, for $k = 1, ..., p$,

$$\widehat{Z}_k^\theta(t) = \exp\left(\int_0^t \widehat{\theta}_k(u) dB_k(u) - \frac{1}{2}\int_0^t \widehat{\theta}_k^2(u) du\right) , \tag{3.3}$$

and, for $j = 1, ..., n$,

$$\widetilde{Z}_j^\theta(t) = \exp\left(\int_0^t \widetilde{\theta}_j(u) dI_j(u) - \phi_j(0, t; \widetilde{\theta}_j(\cdot)), \right) , \tag{3.4}$$

where the functions ϕ_j are defined by

$$\phi_j(t, s; \widetilde{\theta}_j(\cdot)) := \psi_j(t, s, -\mathrm{i}\widetilde{\theta}_j(\cdot)) \tag{3.5}$$

with ψ_j being the cumulants of I_j (see (2.2)). Observe that for constant $\widetilde{\theta}_j$, ϕ_j is the log-moment generating function of I_j.

If the following condition holds:

$$\sup_{0\le t\le T} |\widetilde{\theta}_j(t)| \le c_j, \tag{3.6}$$

where c_j is a constant granting that Condition 1 is satisfied, it follows immediately by Ito's formula that $\widetilde{Z}_j^\theta(t)$ is a positive local martingale with expectation equal to one, so it is a martingale process. With a similar consideration, if $\widehat{\theta}_k(u)$ is a continuous function, the Novikov condition holds, implying that $\widehat{Z}_k^\theta(t)$ is a martingale as well. Hence we can define an equivalent probability measure Q^θ such that $Z^\theta(t)$ is the density process of the Radon-Nikodym derivative dQ^θ/dP:

$$\frac{dQ^\theta}{dP}|_{\mathcal{F}_t} = Z^\theta(t) = \prod_{k=1}^{p} \widehat{Z}_k^\theta(t) \times \prod_{j=1}^{n} \widetilde{Z}_j^\theta(t) \tag{3.7}$$

Remark that the Esscher transform for II processes is parametrized by a deterministic (vector) function of time $\theta(t)$, while for general semimartingales it is a stochastic process, see Kallsen and Shiryaev [24].

The following proposition describes how the characteristics of B and I change under the application of the Esscher transform:

Proposition 3.1. *With respect to the probability measure Q^θ the processes*

$$B_k^\theta(t) = B_k(t) - \int_0^t \widehat{\theta}_k(u)du$$

are Brownian motions for $k = 1, ..., p$ and $0 \le t \le T$. Moreover, for each $j = 1, ..., n$, $I_j(t)$ is an independent increment process on $0 \le t \le T$ with drift:

$$\gamma_j(t) + \int_0^t \int_{|z|<1} z \left\{ \mathrm{e}^{\widetilde{\theta}_j(u)z} - 1 \right\} \ell_j(dz, du) + \int_0^t \widetilde{\theta}^2(u)\, dC(u)\,,$$

and predictable compensator measure $\mathrm{e}^{\widetilde{\theta}_j(t)z}\,\ell_j(dz, dt)$.

Proof. The proof can be found in [8], page 98. □

We see that the change of measure with respect to the Brownian motions will lead to a new mean-reversion level in the dynamics of the processes X_i. In fact, the Q^θ-dynamics of X_i, $i = 1, \ldots, m$ becomes

$$dX_i(t) = \left(\mu_i(t) + \sum_{k=1}^p \sigma_{ik}(t)\widehat{\theta}_k(t) - \alpha_i(t)X_i(t) \right) dt + \sum_{k=1}^p \sigma_{ik}(t)\, dB_k^\theta(t)\,.$$

Here, the mean-reversion level is adjusted by the volatilities and the choice of functions $\widehat{\theta}_k$. In the next proposition we show that the Esscher transform of the II processes I_j in fact also leads to a change in mean-reversion level of Y_j.

Proposition 3.2. *With respect to the probability measure Q^θ, the dynamics of Y_j for $j = 1, \ldots, n$ is*

$$dY_j(t) = (\delta_j(t) - \beta_j(t)Y_j(t))\, dt + \eta_j(t)\, d\Phi_j(t, \widetilde{\theta}_j(\cdot)) + \eta_j(t)\, d\widetilde{I}_j(t)\,,$$

where $\widetilde{I}_j(t) := I_j(t) - \mathbb{E}_{Q^\theta}[I_j(t)]$ is a Q^θ-martingale II process and $\Phi_j(t, \widetilde{\theta}_j(\cdot)) = \frac{\partial}{\partial\theta}\phi_j(0, t; \widetilde{\theta}_j(\cdot)) = \mathbb{E}_{Q^\theta}[I_j(t)]$.

Proof. Note that since I_j is an II process, it is a martingale whenever it has mean zero. Thus, it is sufficient to subtract its mean under Q^θ in order to make it a Q^θ-martingale. It is simple to see that by the definition of the Esscher transform and the log-moment generating function $\phi(t, s; \theta)$ that

$$\mathbb{E}_{Q^\theta}[I_j(t)] = \frac{\partial}{\partial\theta}\phi_j(0, t; \widetilde{\theta}_j(\cdot))\,.$$

Denoting $\widetilde{I}_j(t) := I_j(t) - \frac{\partial}{\partial\theta}\phi(0, t; \widetilde{\theta}_j(\cdot))$, we have proven the proposition. □

Note that from the definition of ψ_j in (2.2), we have for a constant θ

$$\frac{\partial}{\partial\theta}\phi_j(0, t; \theta) = \gamma_j(t) - \gamma_j(0) + \theta(C_j(t) - C_j(0)) + \int_0^t \int_{\mathbb{R}} z \left\{ \mathrm{e}^{\theta z} - 1_{|z|<1} \right\} \ell_j(dz, ds)\,.$$

Thus, $d\Phi_j(t, \widetilde{\theta}_j(\cdot))$ is the measure

$$d\Phi_j(t, \widetilde{\theta}_j(\cdot)) = d\gamma_j(t) + \tilde{\theta}_j(t)\, dC_j(t) + \int_{\mathbb{R}} z \left\{ \mathrm{e}^{\tilde{\theta}_j(t)z} - 1_{|z|<1} \right\} \ell_j(dz, dt)\,.$$

For most interesting cases we will have that $d\Phi_j$ is absolutely continuous with respect to the Lebesgue measure, for example, when I_j is a Lévy process. Thus, we can write

$$d\Phi_j(t, \widetilde{\theta}_j(\cdot)) = \widetilde{\Phi}_j(t, \widetilde{\theta}_j(\cdot))\, dt\,,$$

for some nicely behaving function $\widetilde{\Phi}_j$. In that case we have that the mean-reversion level of Y under Q^θ is $\delta_j(t) + \eta_j(t)\widetilde{\Phi}_j(t, \widetilde{\theta}(\cdot))$. Letting $\widetilde{\theta}_j = 0$, we recover the P-dynamics of Y_j since this is corresponding to no change of measure in that coordinate. Hence, the mean-reversion level under P will be $\delta_j(t) + \eta_j(t)\widetilde{\Phi}_j(t, 0)$. This shows how the mean-reversion level is changing under an Esscher transform in a similar fashion as for the deterministic Girsanov change of B_k considered here. In the general case when $d\Phi_j$ is not absolutely continuous, we also have a change in the mean-reversion level, but now it is only expressible via the differentials.

Let us consider an $I(t)$ being a compound Poisson process. Then, the jump measure ℓ is given by

$$\ell(dz, dt) = \lambda(t) P_J(dz)\,,$$

where P_J is the jump size distribution and $\lambda(t)$ is the jump intensity. Letting $\widetilde{\theta}$ be a constant, we have that the jump measure of $I(t)$ under Q^θ becomes

$$\ell_{Q^\theta}(dz, dt) = \left\{ \int_{\mathbb{R}} \mathrm{e}^{\widetilde{\theta} z} P_J(dz) \lambda(t) \right\} \frac{\mathrm{e}^{\widetilde{\theta} z} P_J(dz)}{\int_{\mathbb{R}} \mathrm{e}^{\widetilde{\theta} z}\, P_J(dz)}\,.$$

Thus, we see that under the new probability Q^θ, $I(t)$ is again a compensated Poisson process, with intensity

$$\lambda^Q(t) = \int_{\mathbb{R}} \mathrm{e}^{\widetilde{\theta} z} P_J(dz) \lambda(t)\,,$$

and jump size distribution

$$P^Q(dz) = \frac{\mathrm{e}^{\widetilde{\theta} z} P_J(dz)}{\int_{\mathbb{R}} \mathrm{e}^{\widetilde{\theta} z}\, P_J(dz)}\,.$$

The Esscher transform is thus changing the jump intensity and the jump size under the new probability Q^θ. We can choose parameters which imply an increasing jump intensity and shifting of the jump size distribution towards the big jumps, that is, emphasising the risk.

The parameter θ is usually called the *market price of risk*, and, as we shall see, is closely related to the risk premium. By our choice of measure transform, we have divided into a market price of risk for the normal spot price variations given by the X_i factors, and a market price for jump risk for the spike regimes Y_j.

The definition of the forward price $f(t, \tau)$ at time $t \geq 0$ for a contract delivering at time $\tau \geq t$ is defined as (see Benth et al. [8])

$$f(t, \tau) = \mathbb{E}_Q[S(\tau) \mid \mathcal{F}_t] \,, \tag{3.8}$$

for some pricing probability Q being equivalent to P. By using the Esscher transform $Q = Q^\theta$ as the risk-neutral measure, it is possible to obtain the prices of forward contracts on electricity in a closed form.

Let us examine the geometric case first. We recall the following proposition, the proof of which can be found in Benth et al. [8], page. 105:

Proposition 3.3. *Suppose that* $S(t)$ *is the geometric spot price model given by* (2.3) *and that Condition 1 holds for* $j = 1, ..., n$ *with*

$$\sup_{0 \leq t \leq \tau} |\eta_j(t) \mathrm{e}^{-\int_t^\tau \beta(u)\, du} + \tilde{\theta}_j(t)| \leq c_j \,.$$

Then the forward price $f(t, \tau)$ *is given by:*

$$\begin{aligned} f(t,\tau) = \Lambda(\tau)\Theta_g(t,\tau;\theta(\cdot)) &\times \exp\left(\sum_{i=1}^m \int_t^\tau \mu_i(u) \mathrm{e}^{-\int_u^\tau \alpha_i(s)\, ds}\, du\right) \\ &\times \exp\left(\frac{1}{2}\sum_{k=1}^p \int_t^\tau \left(\sum_{i=1}^m \sigma_{ik}(u)\mathrm{e}^{-\int_u^\tau \alpha_i(s)\, ds}\right)^2 du\right) \\ &\times \exp\left(\sum_{j=1}^n \int_t^\tau \delta_j(u)\mathrm{e}^{-\int_u^\tau \beta_j(s)\, ds}\, du\right) \times \\ &\times \exp\left(\sum_{i=1}^m \mathrm{e}^{-\int_t^\tau \alpha_i(s)\, ds} X_i(t) + \sum_{j=1}^n \mathrm{e}^{-\int_t^\tau \beta_j(s)\, ds} Y_j(t)\right) \,, \end{aligned}$$

where $\Theta_g(t, \tau; \theta(\cdot))$ *is defined as:*

$$\begin{aligned} \ln \Theta_g(t,\tau;\theta(\cdot)) = &\sum_{j=1}^n [\phi_j(t,\tau;\eta_j(\cdot)\mathrm{e}^{-\int_\cdot^\tau \beta(s)\, ds} + \widetilde{\theta}(\cdot)) - \phi_j(t,\tau;\widetilde{\theta}(\cdot))] \\ &+ \sum_{i=1}^m \sum_{k=1}^p \int_t^\tau \sigma_{ik}(u)\widehat{\theta}_k(u)\mathrm{e}^{-\int_u^\tau \alpha_i(s)\, ds} du \,. \end{aligned} \tag{3.9}$$

For the arithmetic models the price of forward contracts is determined according to the following proposition, the proof of which can be found again in [8]:

Proposition 3.4. *Suppose the spot price dynamics is given by the arithmetic model* (2.9) *and there exists* $\varepsilon > 0$ *such that Condition* 1 *holds with, for* $j = 1, ..., n$,

$$\sup_{0 \leq t \leq \tau} |\tilde{\theta}_j(t)| + \varepsilon \leq c_j \,.$$

The forward price for $0 \leq t \leq \tau$ is then given by:

$$f(t,\tau) = \Lambda(\tau) + \Theta_a(t,\tau;\theta(\cdot)) + \sum_{i=1}^{m} \int_t^\tau \mu_i(u) \mathrm{e}^{-\int_u^\tau \alpha_i(s)\,ds}\,du + \sum_{j=1}^{n} \int_t^\tau \delta_j(u) \mathrm{e}^{-\int_u^\tau \beta_j(s)\,ds}\,du$$
$$+ \sum_{i=1}^{m} \mathrm{e}^{-\int_t^\tau \alpha_i(s)\,ds} X_i(t) + \sum_{j=1}^{n} \mathrm{e}^{-\int_t^\tau \beta_j(s)\,ds} Y_j(t)\,,$$

where

$$\Theta_a(t,\tau;\theta(\cdot)) = \sum_{i=1}^{m} \sum_{k=1}^{p} \int_t^\tau \sigma_{ik}(u) \widehat{\theta}_k(u) \mathrm{e}^{-\int_u^\tau \alpha_i(s)\,ds}\,du + \sum_{j=1}^{n} \int_t^\tau \eta_j(u) \mathrm{e}^{-\int_u^\tau \beta_j(s)\,ds}\,d\Phi_j(u, \widetilde{\theta}_j(\cdot))\,.$$

The measure $d\Phi_j$ is defined in Prop. 3.2.

4. The risk premium

We want to study the risk premium implied by the Esscher transform. The risk premium in forward markets is measured in terms of the difference between the forward price and the prediction of the spot at delivery (see e.g. Geman [19]).

Mathematically, we define the risk premium as the difference between the expectation of the underlying prices calculated with respect to the risk-neutral measure Q and the objective measure P, respectively:

$$R(t,\tau) := \mathbb{E}_{Q^\theta}\left[S(\tau)|\mathcal{F}_t\right] - \mathbb{E}\left[S(\tau)|\mathcal{F}_t\right]\,. \tag{4.1}$$

We can observe that the first term appearing in the definition is nothing more than the forward price calculated according to the risk neutral dynamics and the second the same forward price calculated with respect to the objective dynamics. In both the geometric and arithmetic classes of models a close relationship turns out to exist between the risk premium and the quantity $\Theta(t,\tau;\theta(\cdot))$, not surprisingly.

For the geometric models an explicit calculation of the risk premium provides the following result:

Proposition 4.1. *Under the conditions of Prop. 3.3, the risk premium $R(t,\tau)$ in* (4.1) *implied by the Esscher transform for the geometric model is*

$$R(t,\tau) = \Lambda(\tau)\left(\Theta_g(t,\tau;\theta(\cdot)) - \Theta_g(t,\tau;0)\right) \exp\left(\sum_{i=1}^{m} \int_t^\tau \mu_i(u) \mathrm{e}^{-\int_u^\tau \alpha_i(s)\,ds}\,du\right)$$
$$\times \exp\left(\frac{1}{2} \sum_{k=1}^{p} \int_t^\tau \left(\sum_{i=1}^{m} \sigma_{ik}(u) e^{-\int_u^\tau \alpha_i(s)\,ds}\right)^2 du\right)$$
$$\times \exp\left(\sum_{j=1}^{n} \int_t^\tau \delta_j(u) \mathrm{e}^{-\int_u^\tau \beta_j(s)\,ds}\,du\right) \times$$

$$\times \exp\left(\sum_{i=1}^{m} \mathrm{e}^{-\int_t^\tau \alpha_i(s)\,ds} X_i(t) + \sum_{j=1}^{n} \mathrm{e}^{-\int_t^\tau \beta_j(s)\,ds} Y_j(t)\right),$$

Proof. The result follows directly by Prop. 3.3 using that $\theta = 0$ corresponds to $Q^\theta = P$. □

It is obvious from the expression in the proposition above that the sign of $R(t,\tau)$ is exclusively determined by the difference

$$\Theta_g(t,\tau;\theta(\cdot)) - \Theta_g(t,\tau;0)\,,$$

since the other terms are all positive. Moreover, we see that the risk premium is depending on the seasonal function at time of delivery τ and mean-reversion weighted averages of the levels and volatility of the various factors from current time to delivery. Finally, the current state of the spot price (or, more precisely, its factors discounted by the respective mean-reversion speeds) enter into the risk premium. Hence, the risk premium varies stochastically with time, but the sign will be given deterministically.

To further gain insight into the sign of the risk premium, we can equivalently look at the sign of the quantity $\ln \mathbb{E}_{Q^\theta}[S(\tau)|\mathcal{F}_t] - \ln \mathbb{E}[S(\tau)|\mathcal{F}_t]$, which is simpler to analyse; in fact we have from (3.9) that:

$$\begin{aligned}
&\ln \mathbb{E}_{Q^\theta}[S(\tau)|\mathcal{F}_t] - \ln \mathbb{E}[S(\tau)|\mathcal{F}_t] \\
&\quad = \ln \Theta_g(t,\tau;\theta(\cdot)) - \ln \Theta_g(t,\tau,0) \\
&\quad = \sum_{j=1}^{n} \phi_j(t,\tau;\eta_j(\cdot)\mathrm{e}^{-\int_\cdot^\tau \beta(s)\,ds} + \widetilde{\theta}_j(\cdot)) - \phi_j(t,\tau;\eta_j(\cdot)\mathrm{e}^{-\int_\cdot^\tau \beta(s)\,ds}) - \phi_j(t,\tau;\widetilde{\theta}_j(\cdot)) \\
&\qquad\quad + \sum_{i=1}^{m}\sum_{k=1}^{p} \int_t^\tau \sigma_{ik}(u)\widehat{\theta}_i(u)\mathrm{e}^{-\int_u^\tau \alpha_i(s)\,ds}\,du\,.
\end{aligned}$$

The sign of the risk premium is obviously deterministically determined, and cannot change stochastically as a function of the factors in the model. The levels μ_i and δ_j and the seasonal function Λ will not affect the risk premium sign either, but only the speeds of mean-reversion α_i and β_j, and the volatilities σ_{ik} and η_j. Depending on the specification of these parameters in the spot model and the choice of θ in the risk-neutral probability, we can get quite flexible risk premium structures, however, being deterministic.

We next consider the simple setting of $m = n = p = 1$ with a constant measure change, that is, $\theta = (\widehat{\theta}, \widetilde{\theta})$, and suppose that all parameters in the spot model are constant. Further, we suppose that the II process I is a compound Poisson process with a time-dependent jump intensity $\lambda(t)$, mimicking a seasonally varying spike process. The log-moment generating function becomes

$$\phi(t,\tau;\widetilde{\theta}) = \int_t^\tau \left(\mathrm{e}^{\phi_J(\widetilde{\theta})} - 1\right)\lambda(s)\,ds\,,$$

where ϕ_J is the log-moment generating function of the jump size distribution of I and λ its jump frequency. Finally, we assume $\eta = 1$. Now, from the representation above we find

$$\ln \Theta_g(t,\tau;\theta) - \ln \Theta_g(t,\tau;0)$$

$$= \int_0^{\tau-t} \left[e^{\phi_J(\widetilde{\theta}+e^{-\beta s})} - e^{\phi_J(e^{-\beta s})} - e^{\widetilde{\theta}} + 1 \right] \lambda(\tau-s)\,ds + \frac{\sigma\widehat{\theta}}{\alpha}(1-e^{-\alpha(\tau-t)})$$

$$= \lambda \int_0^{\tau-t} \mathbb{E}\left[\left(e^{\widetilde{\theta} J} - 1 \right) \left(e^{e^{-\beta s} J} - 1 \right) \right] \lambda(\tau-s)\,ds + \frac{\sigma\widehat{\theta}}{\alpha}(1-e^{-\alpha(\tau-t)})\,.$$

Now, observe that whenever $J > 0$, $\exp(e^{-\beta s}J) - 1$ is positive, and non-positive otherwise. Further, if $\widetilde{\theta} > 0$, then $\exp(\widetilde{\theta}J) - 1$ is positive whenever J is, and non-positive otherwise. Hence, for $\widetilde{\theta} > 0$, we will have that the expectation inside the integral will be strictly positive, and thus gives a strictly positive contribution. Since the jump intensity is non-negative, the integral term will give a positive contribution to the risk premium. If $\widehat{\theta}$ is positive as well, we will have a positive risk premium. When $\widetilde{\theta} < 0$, we get that $(\exp(\widetilde{\theta}J) - 1)(\exp(e^{-\beta s}J) - 1)$ is negative whatever J is, and therefore we get a negative contribution. Hence, for negative parameters $\widetilde{\theta}, \widehat{\theta}$ we get a negative risk premium.

We may explain the results above by looking at the level change induced by the Esscher transform. For simplicity, let $\lambda(t) = \lambda$, a constant. Note that

$$\Phi(t,\widetilde{\theta}) = \frac{\partial}{\partial\widetilde{\theta}}\phi(t,\tau;\widetilde{\theta}) = \lambda \int_{\mathbb{R}} z e^{\widetilde{\theta}}\, P_J(dz)\, t\,,$$

and thus the level change for the process Y when going from P to Q^{θ} is given by

$$\Phi(t,\widetilde{\theta}) - \Phi(t,0) = \lambda \int_{\mathbb{R}} z(e^{\widetilde{\theta}z} - 1)\, P_J(dz)\, t\,,$$

where P_J is the distribution of J. We observe that when $\widetilde{\theta} > 0$, $z(\exp(\widetilde{\theta}z) - 1)$ is positive for all z. On the other hand, if $\widetilde{\theta} < 0$, then it is negative for all z. Hence, choosing $\widetilde{\theta}$ positive yields an upward shift of the mean-reversion level of the jump process Y, while a negative choice of $\widetilde{\theta}$ pushes it down. We see this reflected in the risk premium being positive or negative, respectively.

The intuition and experience from the market propose a long-term negative risk premium and a short term positive premium. One could obtain this by choosing $\widetilde{\theta} > 0$ and $\widehat{\theta} < 0$. The shape of the risk premium is determined by the factor $\Theta_g(t,\tau;\theta) - \Theta_g(t,\tau;0)$, which becomes

$$\Theta_g(t,\tau;\theta) - \Theta_g(t,\tau;0) = \exp\left(\int_0^{\tau-t} e^{\phi_J(\widetilde{\theta}+e^{-\beta s})} - e^{\phi_J(\widetilde{\theta})}\lambda(\tau-s)\,ds + \frac{\sigma\widehat{\theta}}{\alpha}(1-e^{-\alpha(\tau-t)}) \right)$$
$$- \exp\left(\int_0^{\tau-t} (e^{\phi_J(e^{-\beta s})} - 1)\lambda(\tau-s)\,ds \right)\,.$$

The reason for a positive risk premium in the short end is explained by Geman and Vasicek [20] as a result of the consumers/retailers in the market being willing to hedge the spike risk. Consider now Y modeling the spikes in electricity prices. Naturally, $J > 0$, that is, a price process which may jump only upwards. On the other hand, β should be reasonably large so as to kill off the spikes quickly. It is straightforwardly seen that for $\widetilde{\theta} > 0$, we will have

$$\int_0^{\tau-t} \left(\mathrm{e}^{\phi_J(\widetilde{\theta}+\mathrm{e}^{-\beta s})} - \mathrm{e}^{\phi_J(\widetilde{\theta})}\right) \lambda(\tau - s)\, ds > \int_0^{\tau-t} \left(\mathrm{e}^{\phi_J(\mathrm{e}^{-\beta s})} - 1\right) \lambda(\tau - s)\, ds\,.$$

If $\widehat{\theta} = 0$ and $\lambda(t)$ constant, we observe that the risk premium increases fastly from zero up to a constant positive level as a function of time to maturity. However, if $\widehat{\theta} < 0$, there will be a competing term in the expression for the risk premium which will force it down and eventually to a fixed level. If $\widehat{\theta}$ is big enough, the risk premium may be pushed below zero to a negative level in the long end of the forward curve (that is, for large τ, times of maturity). How fast the influence of $\widehat{\theta}$ comes in is depending on the speed of mean-reversion α, which one would expect to be slower than β. Hence, we should get a small "bump-shaped" positive premium in the short end, whereas it becomes negative in the long end of the market. This depends also on the choice of parameters.

Let us discuss how a time-dependent jump intensity may influence the sign. We consider a toy example of a market where the spikes occur during a specific period of the year, in January say, while we do not observe any spikes in the remaining part of the year. This means that $\lambda(t)$ is zero except for the month of January, where we suppose it is constant for the sake of simplicity. In the Nord Pool market, the spikes are most frequently observed in January and February, being the coldest period of the year and thus with the highest demand for heating. Let us assume that we are in July, looking at the forward curve one year ahead. Since the jump intensity is zero all the way up to January, the shape of the risk premium will only be given by

$$\Theta_g(t,\tau;\theta) - \Theta_g(t,\tau;0) = \exp\left(\frac{\sigma\widehat{\theta}}{\alpha}(1 - \mathrm{e}^{-\alpha(\tau-t)})\right) - 1\,,$$

for contracts maturing *before* January. Hence, the risk premium will, for $\widehat{\theta} < 0$, be downward sloping towards the asymptote $\sigma\widehat{\theta}/\alpha$ as $\tau - t$ approaches January. Then, for contracts maturing in January, we will gradually get more and more contribution from the jump term. Next, for contracts maturing *after* January, the influence from the jump term will decay slowly, with time to maturity. The reason is that the integral will integrate the constant intensity over January, however, the function

$$s \mapsto \phi_J(\widetilde{\theta} + \mathrm{e}^{-\beta s}) - \phi(\widetilde{\theta})$$

is decreasing. Thus, the farther the maturity is away from January, the smaller function we integrate against, yielding a decreasing contribution. By appropriately scaling the market price of risks, we can see a risk premium which is decreasing up till January, and then increasing to something positive, before it decreases again to a negative level, possibly higher than the previous. We may also have a risk premium which is decreasing till January, then increasing before it decreases again, without crossing to positivity.

Let us now suppose that we are in the beginning of January. Then the picture will be as follows. We will now get full influence of the jump term for the contracts maturing in January, and we will immediately see a positive premium for appropriate choice of the market price of risk. For later maturing contracts, the influence of the jumps in January will become smaller, and eventually we will cross zero and have a negative premium in the long end. In fact, depending on the speeds of mean-reversion, this crossing to negative premia may happen very quick.

From these considerations, we see that for a seasonally dependent spike intensity, we can have a negative risk premium structure which is downward sloping with small bumps along the curve in periods where the spike risk is high. In such periods, one may even get a positive premium. If present time corresponds with a high intensity period for spikes, we may have a positive premium in the short end, and negative in the long. The risk premium curve will be further scaled by the current level of the spike Y and base component X. Running over the year, the risk premium structure can go from negative all over, to positive in the short end, and negative in the long, appropriately scaled by the current spike level Y and base component X. Note that the spike component normally contributes in the short end since β is usually fast, while X, the base component has a longer range of influence since it is associated with a slow mean-reversion α.

We illustrate the situation discussed above with a numerical example. Measuring time t in days, we assume that $\beta = 0.3466$. This corresponds to a fast mean-reversion yielding a half-life of 2 days. Further, we consider exponentially distributed jump sizes with a mean equal to 0.5. To mimic seasonally occurring jumps, we suppose that the spikes *only* occur in January, where the frequency is 5, that is, on average 5 spikes during January. For the rest of the year we let the frequency be zero. The base component has a speed of mean reversion $\alpha = 0.099$, which is a half-life of 7 days, and the volatility is supposed to be $\sigma = 0.0158$, that is, 30% annually. In the example, we calculate the contribution to the risk premium coming from the term $\Theta_g(t,\tau;\theta) - \Theta_g(t,\tau;0)$, which we recall is the factor deciding the sign of the risk premium. We assume that the spike risk is positive, and the base component has a negative premium, here given as $\widetilde{\theta} = 0.95$ and $\widehat{\theta} = -4$, resp. Calculating this for various times over the year, for a time-to-maturity ranging up to 220 days, we obtain the curves in Figs. 1 and 2. In the example, we have standardized each month to be 30 days long.

In Fig. 1 we have plotted the curves seen from the first day in July, October, December and January. For the three first curves, they are all dropping downwards from a negative value, however, having a bump when the time-to-maturity is crossing over the month of January. We observe a negative premium overall, since the contribution from the base component is the strongest. When we consider the curve from January 1, the picture is changing. In the short end of the curve, we get a positive contribution, before becoming negative in the long end. This is a reflection of a strong contribution from the spike component which now pushes the premium above zero, reflecting the aversion to spike risk. Note that by increasing the spike size distribution and/or the frequency, we may even obtain positive premium in the bump along the curve. Looking at Fig. 2, this is even more distinct. By considering the curve seen from January 15, we get an even more positive contribution, however, lasting shorter since only half of the spike period is taken

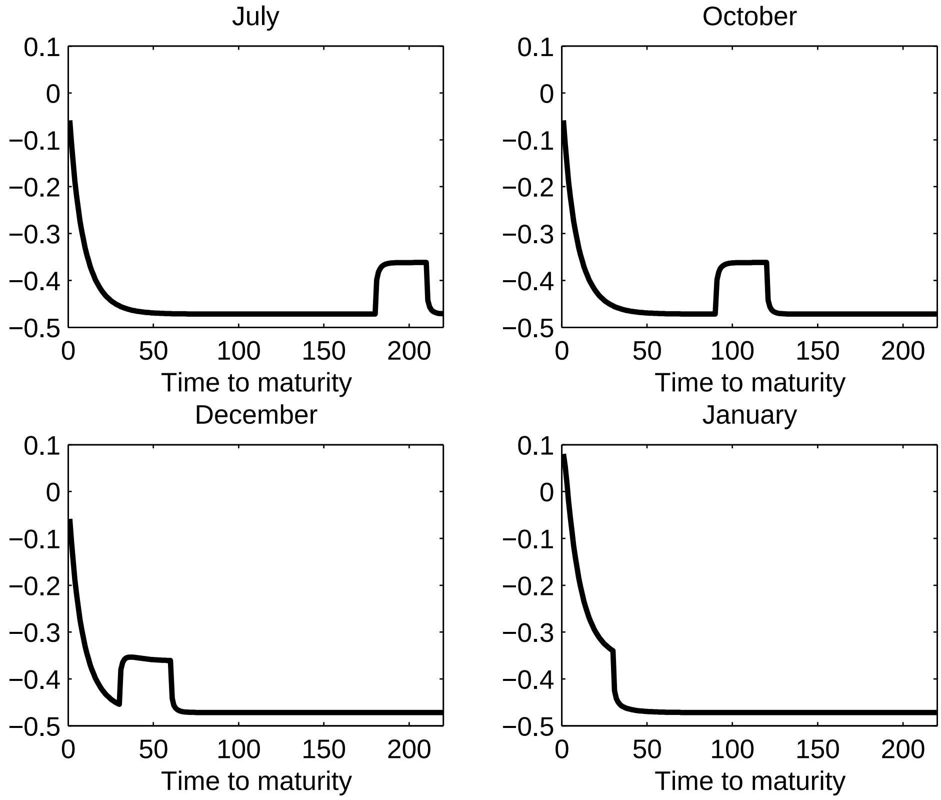

FIGURE 1. Contribution to the risk premium curve from the term $\Theta_g(t,\tau,\theta) - \Theta_g(t,\tau,0)$, for 4 different choices of t and 220 days of time-to-maturity.

into account. The kink in the curve around 15 days to maturity is a reflection of the spike period being constrained to only January. The risk premium is obtained by scaling the curves by positive factors which size is given by the mean-reversion discounted states of X and Y. Hence, if we get a spike $Y(t)$, then the influence on the risk premium is relatively small if we are in July compared to January. Getting a spike in January, will immediately scale up the positive part in the short end. An increasing value of $X(t)$, on the other hand, will scale up a bigger part of the curve since the mean-reversion is slower. The seasonality function will also play an important role by an overall seasonal scaling of the curve. The example clearly demonstrates that seasonally occurring jumps are responsible for bumps in the risk premium and a positive premium in the short end, as long as we are in a period of high spike intensity. By applying the Esscher transform, we can obtain a risk premium which goes from being negative downward sloping, to having a positive short end. The numerical calculations were performed using Matlab and the built-in numerical integration routing `quad`, which is an implementation of a recursive adaptive Simpson quadrature method.

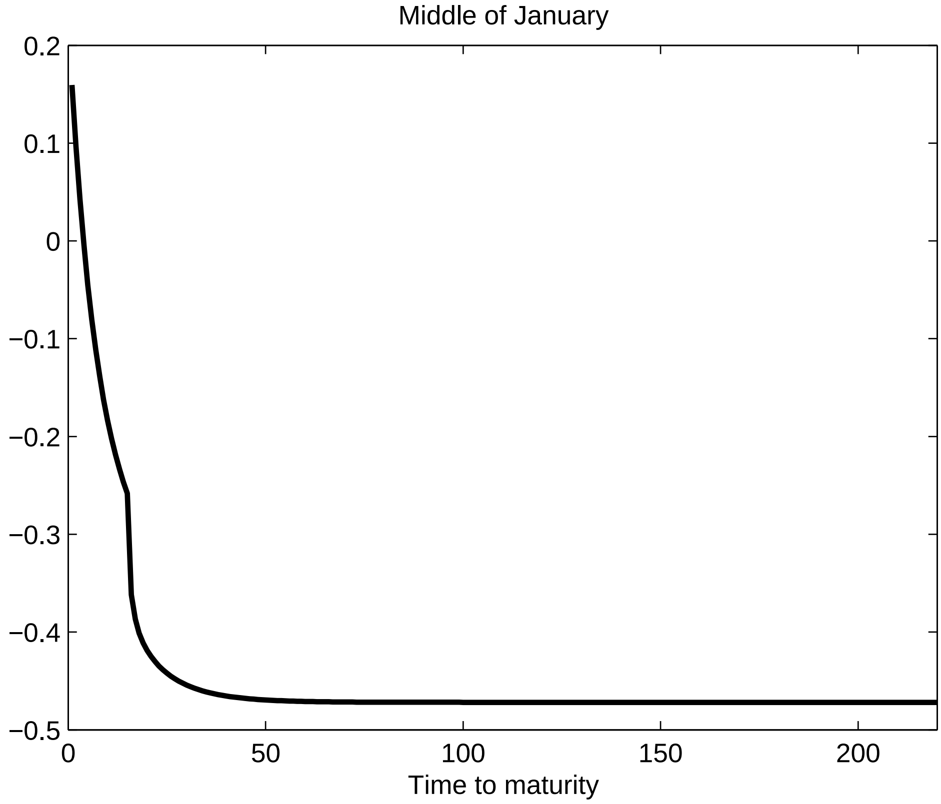

FIGURE 2. Contribution to the risk premium curve from the term $\Theta_g(t,\tau,\theta) - \Theta_g(t,\tau,0)$ seen form middle of January.

For reasons of comparison, consider a one-factor model $m = 0$, $n = 1$, where the II process $I(t)$ is given by

$$I(t) = \sigma B(t) + L(t)\,,$$

that is, a sum of a Brownian motion and a jump component. We let the jump component $L((t)$ be a compound Poisson process as in the example above. In this model the spikes and normal variation components of the log-prices mean-revert at the same speed, and the model is analogous to the Merton model is stock markets [29], and applied by Cartea and Figueroa [10] to model UK electricity spot prices. We easily find by a similar computation as above that

$$\ln \Theta_g(t,\tau;\theta) - \ln \Theta_g(t,\tau;0) = \lambda \int_0^{\tau-t} \mathbb{E}\left[\left(\mathrm{e}^{\widetilde{\theta}J} - 1\right)\left(\mathrm{e}^{\mathrm{e}^{-\beta s}J} - 1\right)\right]\,ds + \frac{\sigma^2\widetilde{\theta}}{\alpha}(1 - \mathrm{e}^{-\alpha(\tau-t)})\,.$$

We remark in passing that the appearance of σ^2 is not a mistake, but reflecting the fact that the Esscher transform applied to the process $I(t)$ is a Girsanov transform of the process $\sigma B(t)$, and not just $B(t)$. Contrary to the two-factor model, we can now only have either a positive or a negative risk premium, by choosing $\widetilde{\theta}$ positive or negative, resp. Hence, we do not obtain the flexibility of a sign change by imposing different market prices of risk to each component. However, this can easily be mended by redefining the model. One may rewrite the model into two factors, where both factors have the same speed of mean-reversion, but one is driven by the Brownian component and the other by the compound Poisson process. This means that we consider the two processes B and L separately, rather than coming from the same II process. We can then do a Girsanov transform on the Brownian component, and an Esscher on the jump process, and we get the possibility to have a sign change of the risk premium. We remark in passing that Cartea and Figueroa [10] did a change of measure only with respect to the Brownian component, and estimated a negative market price of risk $\widehat{\theta}$ for the UK electricity market. Thus, their model predicts a downward sloping risk premium, or, in other words, a backwardated market. By including a constant market price of jump risk $\widetilde{\theta}$ in their calibration study, one could possibly obtain a positive risk premium in the short end of the curve, since the market shows significant signs of spikes, in particular over the colder period of the year.

For the arithmetic class of models the calculation of the risk premium provides the following result:

Proposition 4.2. *Under the conditions of Prop. 3.4 we have that the risk premium is given by*

$$R(t,\tau) = \Theta_a(t,\tau;\theta(\cdot)) - \Theta_a(t,\tau,0)\,,$$

where Θ_a is defined in Prop. 3.4.

Proof. This follows directly from Prop. 3.4, noticing that $Q^\theta = P$ when $\theta = 0$. □

We observe that risk premium in the arithmetic case corresponds exactly to the *change* in mean-reversion level of the processes X_i and Y_j when going from P to Q^θ. If we specify $n = m = p = 1$ with constant parameters as in the example for the geometric case above, we reach the same conclusions that constant specifications of $\widetilde{\theta}$ and $\widehat{\theta}$ implies a positive risk premium, and negative in the opposite case. The mixed case can yield a sign change depending on the parameter specifications.

Let us focus for a moment on the model proposed by Benth, Kallsen and Meyer-Brandis [4] where $m = 0$, $\delta_j = 0$, and the speed of mean-reversions β_j and jump volatilities η_j are positive constants for $j = 1,\ldots,n$. Furthermore, the II processes are subordinators, that is, having only positive jumps, and we suppose that they are driftless with log-moment generating functions

$$\phi_j(t,\tau,\widetilde{\theta}_j(\cdot)) = \int_t^\tau \int_0^\infty (\mathrm{e}^{\widetilde{\theta}(s)z} - 1)\,\ell_j(dz,ds)\,.$$

This includes the compound Poisson processes considered above, but with positive jump sizes only. Observe that now $\Phi_j(t,\widetilde{\theta}) = \int_0^t \int_0^\infty z\mathrm{e}^{\widetilde{\theta}z}\,\ell_j(dz,ds)$, and the risk premium takes the form.

$$R(t,\tau) = \sum_{j=1}^n \int_t^\tau \int_0^\infty \mathrm{e}^{-\int_u^\tau \beta_j(v)\,dv} z\left(\mathrm{e}^{\widetilde{\theta}_j z} - 1\right)\,\ell_j(dz,du)\,.$$

Each term will, using similar arguments as above, contribute positively or negatively depending on $\widetilde{\theta}_j$ being positive or negative, respectively.

In applications of this model, one has usually chosen two factors, that is, $n = 2$ (see Benth, Kiesel and Nazarova [5]). The first component models the base signal, and the second factor the spike process. The latter has a high speed of mean-reversion, whereas the base signal reverts at a slower speed, so naturally we would have $\beta_2 > \beta_1$. The consumers are afraid of spikes, and the hedging pressure they imply may give a positive $\widetilde{\theta}_2$. Since the spikes are quickly killed by mean-reversion, this would have an effect for contracts on the short term, that is, with short time to maturity. The second term in the risk premium would then dominate in the short end of the forward curve, producing a positive premium. The producers are interested in securing production in the long term, and would therefore accept a lower price for the production, which means that we must have a negative choice of $\widetilde{\theta}_1$. Due to the slower reversion of the first factor, this would then have the largest impact on the long term, and thus giving a negative risk premium. This can be reflected in the expression above as well. For given choices of $\widetilde{\theta}_1$ and $\widetilde{\theta}_2$, we can calculate the time-to-maturity where we will have a sign change in the risk premium (at least numerically). This will be a deterministic time point which will not change according to market conditions, but only according to the parameters of the spot model and the current time t. This shows that the Esscher transform may produce a sign change in the risk premium, and also produce the right risk premium structure by appropriately choosing the parameters θ. In fact, the example for the geometric case can be applied in the arithmetic situation as well, showing the even in this situation we can go from a market with a sign change to a case where the risk premium is either reflecting backwardation or contango. Contrary to the geometric model, the risk premium will not be stochastic, but only deterministically evolving over time.

As is well-known, the electricity and gas markets trade in forward contracts which deliver the underlying energy over a period rather than at a fixed maturity. For geometric models it is in general difficult to provide any explicit expressions for the price of such forwards, which we shall refer to as *flow forwards*. However, flow forward prices are attainable when supposing that the spot price dynamics is given by the arithmetic model. We discuss this in more detail next.

We consider the simple example of a forward contract delivering power over a time interval $[\tau_1, \tau_2]$, where payment takes place at the final delivery date. In Benth et al. [8] it is shown that the forward price $F(t, \tau_1, \tau_2)$, $0 \leq t \leq \tau_1 < \tau_2$ is defined as

$$F(t, \tau_1, \tau_2) = \mathbb{E}_Q \left[\frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} S(t)\, dt \,|\, \mathcal{F}_t \right] = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} f(t, \tau)\, d\tau\,. \tag{4.2}$$

Here, Q is some equivalent probability measure. The latter representation comes from the definition of the forward price $f(t, \tau)$, whereas in the former equality we use the standard market convention that the price is denominated per Mega Watt hour (MWh), for electricity. Using the model in Benth, Meyer-Brandis and Kallsen [4], and assuming that Q is determined by the Esscher transform, we obtain the flow forward price

$$F(t,\tau_1,\tau_2)=\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2}\left\{\Lambda(u)+\sum_{j=1}^{n}\int_t^u\int_0^\infty \eta_j(u)\mathrm{e}^{-\alpha_j(u-s)}z\,\widetilde{\ell}(dz,ds)\right\}du+$$

$$+\sum_{j=1}^{n}\frac{Y_j(t)}{\alpha_j(\tau_2-\tau_1)}\left(\mathrm{e}^{-\alpha_j(\tau_1-t)}-\mathrm{e}^{-\alpha_j(\tau_2-t)}\right)$$

where $\tilde{\ell}$ is the compensated jump measure with respect to the Esscher transform. For notational simplicity, we have assumed that the speed of mean-reversions α_j, are constants.

The risk premium for flow forwards can naturally be defined as

$$R^{FF}(t,\tau_1,\tau_2)=F(t,\tau_1,\tau_2)-\mathbb{E}\left[\frac{1}{\tau_1-\tau_2}\int_{\tau_1}^{\tau_2}S(t)\,dt\,|\,\mathcal{F}_t\right]. \tag{4.3}$$

We immediately see that

$$R^{FF}(t,\tau_1,\tau_2)=\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2}R(t,\tau)\,d\tau\,. \tag{4.4}$$

Hence, the risk premium of flow forwards becomes the average of the risk premium $R(t,\tau)$ over the delivery period of the flow forward. This means that even for geometric models we can say something about the sign and value of the risk premium for flow forwards. However, for arithmetic models we can go further and provide explicit expressions.

Subtracting the corresponding value of the flow forward calculated with respect to the historical probability measure, one obtains the risk premium:

$$\begin{aligned}R^{FF}(t,\tau_1,\tau_2,\tilde{\theta})&=\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2}\int_t^u\int_0^\infty \eta_j(s)\mathrm{e}^{-\alpha_j(u-s)}z\left\{\widetilde{\ell}(dz,ds)-\ell(dz,ds)\right\}du\\&=\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2}\left(\int_t^u\int_0^\infty \eta_j(s)\mathrm{e}^{-\alpha_j(u-s)}z\left\{\mathrm{e}^{\widetilde{\theta}_j(s)}-1\right\}\ell(dz,ds)\right)du.\end{aligned}$$

We recognize the factors $\exp(\widetilde{\theta}_j(s))-1$ as the determinants of the risk premium sign and we can conclude that negative values of $\widetilde{\theta}_j$ will give a negative contribution, while positive values will give a positive contribution. The discussions above hold for the flow forward as well.

To conclude this section, we discuss the issue of martingale dynamics for the spot. As already mentioned, electricity is a non storable commodity, so its spot price process need not be a (local) martingale under the risk-neutral measure. In constructing the measure change based on the Esscher transform we have in fact not required any martingale condition. This is in contrast with the usual approach followed in the stock market framework, where the Esscher martingale transform is in use in general. As it turns out, there exists no choice of $\widetilde{\theta}$ that can turn the (discounted) spot price into a martingale. To choose $\widetilde{\theta}$ time-dependent will not change this fact. This is due to the mean-reverting models that we consider.

To see this, let us consider $m = 0$ and $n = 1$, and for the sake of simplicity we assume $\Lambda(t) = 1$, $\delta = 0$, $\eta = 1$, β positive constant and zero risk-free interest rate, $r = 0$. In order for the Esscher transform to be a martingale transform, we must have

$$\mathbb{E}_{Q^{\tilde{\theta}}}\left[\mathrm{e}^{Y(t)}\,|\mathcal{F}_s\right] = \mathrm{e}^{Y(s)}\,,$$

for all $t \geq s \geq 0$. But, since

$$Y(t) - Y(s) = -Y(s)(1 - \mathrm{e}^{-\beta(t-s)}) + \int_s^t \mathrm{e}^{-\beta(t-u)}\,dI(u)\,,$$

we find that

$$\begin{aligned}\mathbb{E}_{Q^{\tilde{\theta}}}\left[\mathrm{e}^{Y(t)}\,|\mathcal{F}_s\right] &= \mathrm{e}^{Y(s)}\mathbb{E}_{Q^{\tilde{\theta}}}\left[\mathrm{e}^{Y(t)-Y(s)}\ ,|\mathcal{F}_s\right] \\ &= \mathrm{e}^{Y(s)}\mathrm{e}^{-Y(s)(1-\mathrm{e}^{-\beta(t-s)})}\mathbb{E}_{Q^{\tilde{\theta}}}\left[\mathrm{e}^{\int_s^t \mathrm{e}^{-\beta(t-u)}\,dI(u)}\right]\,,\end{aligned}$$

where we have used the independent increment property of I in the last equality. We see that the Esscher transform of I will only contribute with a deterministic term no matter the choice of $\widetilde{\theta}$, and therefore cannot "kill off" the term $\exp(-Y(s)(1 - \exp(-\beta(t-s)))$ which is stochastic. In conclusion, the martingale property of the spot price can not be achieved by an Esscher transform for these models.

Letting $\beta = 0$ changes the matters, as we now discuss. Choosing zero mean-reversion speed corresponds to a non-stationary process Y. This resembles the case considered in Hubalek and Sgarra [22]. A direct calculation reveals that

$$\mathbb{E}_{Q^{\theta}}\left[\mathrm{e}^{Y(t)}\,|\,\mathcal{F}_s\right] = \exp\left(Y(s) - \phi(s,t;\widetilde{\theta}(\cdot)) + \phi(s,t;\widetilde{\theta}(\cdot)+1)\right)\,,$$

for $0 \leq s \leq t \leq T$. Hence, the martingale property is obtained if an only if there exists a function $\widetilde{\theta}$ such that

$$\phi(s,t;\widetilde{\theta}(\cdot)+1) = \phi(s,t;\widetilde{\theta}(\cdot))\,, \tag{4.5}$$

for all $0 \leq s \leq t \leq T$. This is a simple extension of the results in Hubalek and Sgarra [22].

Let $Y(t)$ be driven by a time-inhomogeneous Compound Poisson process $I(t)$ with time-dependent intensity $\lambda(t)$. Recall the cumulant being,

$$\psi(s,t;\widetilde{\theta}(\cdot)) = \int_s^t \left(\mathrm{e}^{\phi_J(\widetilde{\theta}(u))} - 1\right)\lambda(u)\,du\,.$$

Observe for a constant choice of $\widetilde{\theta}$, we get the martingale condition (4.5) satisfied as long as

$$\phi_J(\widetilde{\theta}) = \phi_J(\widetilde{\theta}+1)\,.$$

For example, choosing a normally distributed jump size J with mean μ and variance η^2, we get the equation

$$\widetilde{\theta}\mu + \frac{1}{2}\widetilde{\theta}^2\eta^2 = (\widetilde{\theta}+1)\mu + \frac{1}{2}(\widetilde{\theta}+1)^2\eta^2\,,$$

which has the solution $\widetilde{\theta} = -\mu/\eta^2 - 1/2$. Observe that as long as $\mu > \eta^2/2$, we have that $\widetilde{\theta}$ is negative, and therefore the risk premium for the martingale Esscher transform is in fact also negative. We finally remark that when the jump process only has positive jumps, that is, the outcomes of J is supported on the positive real line, then we cannot find any $\widetilde{\theta}$ ensuring the martingale property. The reason is that the moment generating function is increasing in this case. For a process with only negative jumps, the conclusion is the same due to a decreasing moment generating function.

5. Conclusions

In this paper we have provided the mathematical evidence of the risk premium sign change in electricity market. The empirical evidence of this sign change has been given and discussed in several papers mentioned in the introduction. We have proved that in a wide framework represented by both classes of geometric and arithmetic models based on independent increments processes, when an Esscher change of measure is performed in order to construct the risk-neutral pricing measure, the sign change of the risk premium can be analysed explicitly. Indeed, we show that the a risk premium being positive in the short end and negative in the long end may be explained by the appearence of jumps, and appropriately chosen market prices of risk. The risk premium in the short end being positive and negative depending on the time of year, can be traced back to seasonally occuring spikes.

Different measure changes can be introduced, however, the Esscher transform seems to be the easiest and the most natural to apply due to its properties. In fact, several difficulties arise when the measure change does not guarantee the model structure preservation. The main advantage of using the Esscher transform is namely that the model structure is preserved, in the sense that the driving noises in the factors of the spot price dynamics are still independent increment processes after the measure change. This ensures that we can derive closed form expressions for forward prices and the risk premium, where we may intepret the results in terms of the parameters of the model and measure change.

Since the transform is not state-dependent, it can not reproduce the martingale property of the spot price process. This is contrary to stock market models where the Esscher transform can be specified so that the spot price is a martingale after discounting. Although this is not any drawback in the incomplete power markets, it is showing that the Esscher transform has very little flexibility, and more general measure changes could be worthwhile investigating.

A possible further progress of this work will be to obtain the parameter function $\widetilde{\theta}(t)$ in the Esscher Transform by a calibration procedure from the traded forward prices and then to examine the agreement of the sign change between the empirical and the theoretical results obtained. This would enforce the use of the Esscher change of measure as a risk-neutral measure for electricity and, more in general, energy derivatives valuation. This is already the subject of new investigations currently in progress.

References

[1] Barndorff-Nielsen, O.E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch*, **2**(1), 41–68.

[2] Benth, F.E., Ekeland, L., Hauge, R., and Nielsen, B.F. (2003). On arbitrage-free pricing of forward contracts in energy markets. *Appl. Math. Finance*, **10**(4), 325–336.

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

[4] Benth, F.E., Kallsen J., and Meyer-Brandis T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modelling and derivatives pricing. *Appl. Math. Finance,* **14**(2), 153–169.

[5] Benth, F.E., Kiesel, R., and Nazarova, A. (2009). A critical empirical study of two spot price models for electricity. *E-print*, University of Oslo.

[6] Benth, F.E., and Meyer-Brandis, T. (2009). The information premium in electricity markets. *J. Energy Markets*, **2**(3).

[7] Benth, F.E., and Saltyte-Benth, J. (2004). The Nornal Inverse Gaussian distribution and spot price modeling for energy markets. *Int. J. Theor. Appl. Finance*, **7**(2), 177–192.

[8] Benth, F.E., Saltyte Benth, J., and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets*. World Scientific, Singapore.

[9] Botterud, A., Bhattacharya, A.K., and Ilic, M. (2002). Futures and Spot Prices- an Analysis of the Scandinavian Electricity Market. Proceedings of the 34th Annual North American Power Symposium. October 2002, Tempe, Arizona, USA.

[10] Cartea, A., and Figueroa, M.G. (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Appl. Math. Finance*, **12**(4), 313–335.

[11] Crosby, J. (2008). A multi-factor jump-diffusion model for commodities. *Quantit. Finance*, **8**(2), 181–200.

[12] Delbaen, F., and Schachermayer, W. (1994). A General Version of the Fundamental Theorem of Asset Pricing. *Math. Annalen*, **300**, 463–520.

[13] Delbaen, F., and Schachermayer, W. (1998). The Fundamental Theorem for Unbounded Processes. *Math. Annalen*, **312**, 215–250.

[14] Diko, P., Lawford, S., and Limpens, V. (2006). Risk premia in electricity forward prices. *Stud. Nonlinear Dyn. Econom.*, **10**(3) article 7. (Electronic publication: http://www.bepress.com/snde/vol10/iss3/art7).

[15] Eberlein, E., and Stahl, G. (2003). Both sides of a fence: A statistical and regulatory view of electricity risk. *Energy Risk*, **8**, 34–38.

[16] Esche, F., and Schweizer, M. (2005). Minimal entropy preserves the Lévy property: how and why. *Stoch. Proc. Appl.*, **115**(2), 299–327.

[17] Esscher, F. (1932). On the probability function in collective theory of risk. *Skandinavisk Aktuarietidskrift*, **15**, 175–195.

[18] Financial Times (2006). Gas given away as surge from Norway causes price plunge. Front page, First Section, 4 October.

[19] Geman, H. (2005). *Commodities and Commodity Derivatives.* Wiley-Finance, John Wiley&Sons, Chichester.

[20] Geman, H., and Vasicek, O. (2001). Forward and futures on non-storable commodities: the case of electricity. *RISK*, August.

[21] Gerber, H.U., and Shiu, E.S.W. (1994). Option pricing by Esscher transforms. *Trans. Soc. Actuaries*, **46**, 99–191 (with discussion).

[22] Hubalek, F., and Sgarra, C. (2006). On the Esscher transform and entropy for exponential Lévy models. *Quantit. Finance*, **6**(2), 125–145.

[23] Jacod, J., and Shiryaev, A.N. (2002). *Limit Theorem for Stochastic Processes.* 2nd Ed, Springer-Verlag, Berlin.

[24] Kallsen, J., and Shiryaev, A.N. (2002). The cumulant process and Esscher's change of measure. *Finance Stoch.*, **6**(4), 397–428.

[25] Meyer-Brandis, T., and Tankov, P. (2008). Multi-factor jump diffusion models of electricity prices. *Intern. J. Theor. Appl. Finance*, **11**, 503–528.
[26] Longstaff, F.A., and Wang, A.W. (2004). Electricity Forward Prices: A High-Frequency Empirical Analysis. *J. Finance*, **59**(4), 1877–1900.
[27] Lucia, J., and Schwarz, E.S. (2002). Electricity Prices and Power Derivatives: Evidences from the Nordic Power Exchange. *Rev. Derivatives Res.*, **5**(1), 5–50.
[28] Lucia, J., and Torrò, H. (2008). Short-Term Electricity Futures Prices: Evidence on the Time-Varying Risk Premium,. *Working Paper*, Dept. of Financial Economics, Universidad de Valencia.
[29] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *J. Financial Economics*, **3**, 125–144.
[30] Schwarz, E.S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *J. Finance*, **53**(3), 923–973.

(Fred Espen Benth), Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway, and, University of Agder, School of Management, Serviceboks 422, N-4604 Kristiansand, Norway

E-mail address: fredb@math.uio.no

URL: http://folk.uio.no/fredb/

(Carlo Sgarra), Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo Da Vinci, 32, 20133 Milano, Italy

E-mail address: carlo.sgarra@polimi.it