UNIVERSITÀ DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

CORSO DI DOTTORATO DI RICERCA IN SCIENZE MATEMATICHE
INDIRIZZO MATEMATICA COMPUTAZIONALE
CICLO XXXI

ADDITIVE MODELS FOR ENERGY MARKETS

Coordinatore del Corso: Ch.mo Prof. Martino Bardi
Supervisore: Ch.mo Prof. Tiziano Vargiolu

Dottorando: Marco Piccirilli

Riassunto

Questa Tesi esplora la capacità dei modelli additivi di descrivere i prezzi nei mercati energetici, concentrandosi in particolare sul caso specifico dell'elettricità e del gas naturale. Nel Capitolo 1 studiamo un problema di ottimizzazione dinamica di portafoglio per il trading di energia elettrica su mercati intraday. Nel Capitolo 2 introduciamo un framework trattabile e privo di arbitraggio basato sull'approccio di Heath-Jarrow-Morton per mercati a termine energetici multicommodity. Il Capitolo 3 si occupa di uno studio empirico approfondito di un modello a due fattori derivato dal framework del Capitolo 2, con un'applicazione al mercato a termine elettrico tedesco. Infine, nel Capitolo 4 discutiamo il prezzaggio di opzioni per modelli fattoriali additivi con metodi di trasformata di Fourier. Introduciamo un modello di prezzi futures a due fattori con salti al fine di catturare lo smile delle volatilità implicite di opzioni Europee sull'elettricità. Viene presentata un'applicazione al mercato European Energy Exchange Power Derivatives.

Abstract

This Dissertation explores the capability of additive models to describe prices in energy markets, by focusing in particular on the specific case of electricity and natural gas. In Chapter 1 we study a dynamic portfolio optimization problem designed for intraday electricity trading. In Chapter 2 we introduce a no-arbitrage tractable framework based on the Heath-Jarrow-Morton approach for a multicommodity energy forward market. Chapter 3 deals with a thorough empirical study of a two-factor model derived by the framework of Chapter 2, with an application to the German power futures market. Finally, in Chapter 4 we discuss option pricing for additive factor models by Fourier transform methods. We introduce a two-factor futures price model with jumps in order to capture the implied volatility smile of European electricity options. An application to the European Energy Exchange Power Derivatives market is presented.

Contents

Introduction

The first stochastic model for stock prices has been introduced in 1900 by Bachelier [10], who models a stock price by an arithmetic Brownian motion. Since the Brownian motion attains arbitrarily large negative values with positive probability, the Bachelier model generates negative prices. In 1965 Samuelson [116] pointed out that this is economically unrealistic and leads to anomalous results in pricing long-term options and proposed to use an arithmetic Brownian motion to describe the random fluctuations of log-prices (*geometric model*) rather than prices as Bachelier (*additive model*). In such a way, the price stays positive and follows a geometric Brownian motion.[1] Following the pivotal work of Black and Scholes [39] and Merton [98], geometric models have become the benchmark approach for modeling prices in Mathematical Finance.

In more recent years, the main objection to additive models seems to be founded on the lack of economic sense of negative prices. However, while this argument is true for the stock market, this is not the case for some commodity markets. Indeed, since recently, energy-related markets have been showing negative prices.[2] For instance, in Europe, a negative threshold for prices was introduced at the European Energy Exchange (EEX) in 2007 and the first transaction with negative price occurred the next year (and many others followed [61, 122]); negative prices were also observed in natural gas markets in Western Canada in 2015 [68], 2017 and 2018 [60].

Energy markets are commodity markets related to the production of energy: for example, oil, coal, natural gas, power, CO_2 certificates, weather derivatives. In addition to negative prices, energy-related commodity prices show empirical features that are not common to classical financial markets. In order not to stray from the scope of this Dissertation, we will outline the most relevant ones and will focus on the case of power and natural gas (cf. e.g. [32, 44]). Prices in physical (or spot) markets are determined according to the equilibrium of supply and demand. Since demand depends on temperature, sunlight and even national habits[3], power and natural gas prices follow seasonal patterns at different time scales. This implies that prices tend to oscillate around a seasonally varying equilibrium level. Uncertainty in prices comes from the supply-side as well. For example, unexpected power plant outages or intermittent renewable energy sources (RES) generation can move the price up. Empirically, this may result in upward jumps in prices (positive spikes), even of considerable size, followed

[1] In order to underline the positivity property, Samuelson calls the geometric Brownian motion *economic Brownian motion.*

[2] A negative price means that the seller delivers the commodity and pays the buyer. In this case, halting production is more costly than paying someone else to consume. This happens because of the physical nature of these commodities, and can be due, for instance, to a lack of infrastructure, transportation constraints and high cost (or impossibility) of storage facilities.

[3] For example, in UK there is a peak in power consumption around 5 p.m. when people turn on the cooker to boil water for tea.

by a fast decrease to a "normal" level as soon as the shock is absorbed by the network. On the other hand, especially since the introduction in the power grid of RES, the combination of low demand with particularly favorable weather conditions (e.g. strong wind and/or strong sun radiation) causes the opposite effect of sharp downward movements (negative spikes) that may drive the prices below zero.

These stylized features imply that spot prices exhibit mean-reversion, seasonality (in both level and volatility) and jumps. To some extent, these empirical features transfer also to other segments. In fact, besides a physical (or spot) market, there are several other exchanges that regulate energy trading. In the case of power and natural gas, there are, for instance, intraday and derivative markets.

Intraday electricity markets open after the closure of the spot (day-ahead) market and are often based on continuous trading. Their main purpose consists in allowing market participants to adjust their day-ahead position, i.e. agents can trade power for each hour of the next day (until 45 or 30 minutes prior to delivery). The continuous nature of this market allows the activity of purely financial traders. In view of this, we will study in Chapter 1 a portfolio optimization problem for a speculative agent trading in the intraday market.

Regarding the derivative markets, a substantial amount of regulated trading takes the form of futures contracts. Futures are fundamental hedging instruments for producers and consumers, as well as trading instruments for financial speculators, that consider them as an alternative asset class. Power and gas futures contracts prescribe a delivery period for the commodity, rather than a shipping date (as other commodities). This requires a non-standard set of arbitrage relations that must hold among contracts with overlapping delivery periods. For example, one can buy a futures for delivery of power over the first quarter of the year, i.e. January, February and March, and construct a portfolio with the same value by buying three monthly futures for delivery over each corresponding month. Obviously, the same no-arbitrage considerations transfer to options written on futures. In Chapter 2 we will introduce a no-arbitrage framework capable to include these relations for very general stochastic dynamics, while in Chapter 3 we will present a thorough empirical study of a submodel applied to EEX power futures. Chapter 4 is devoted to the pricing of options on futures by Fourier transform methods and an application to EEX power vanilla options.

As already mentioned, in this Dissertation we will address some modeling problems in derivatives (futures and options written on these) and intraday electricity markets. Notably, futures and options constitute almost the total of regulated derivative trading, while intraday markets give flexibility to power producers by allowing to adjust their production schedule on a hourly basis, thus playing an important role in the equilibrium of power systems. The main ingredient of our analysis will be a stochastic process for modeling futures and intraday prices. From the empirical considerations outlined above, we introduce additive models based on *mean-reverting Lévy processes* with *seasonal coefficients* (i.e. depending on time and, in the case of derivatives, parameters representing the delivery period). These models are flexible enough to reproduce the observed empirical features of both futures, options and intraday prices while, in pricing delivery-based derivatives, they have the unique advantage over their geometric counterpart of preserving tractability when dealing with no-arbitrage considerations. In order to handle these models, we will use techniques mainly from Stochastic Analysis and Stochastic Optimal Control for Lévy processes (e.g. [50, 51, 69, 104, 117]). We contribute to literature by proving some theoretical results that are interesting also from a purely mathematical perspective. Part of our work will be devoted to calibration procedures and empirical considerations based on MATLAB™ code.

We now give a detailed overview of each chapter.

Overview of Chapter 1

This chapter is based on a joint work [111] with Tiziano Vargiolu.

In this chapter we study a dynamic portfolio optimization problem designed for intraday electricity trading. These markets play an important role in the equilibrium of power grids, since both electricity producers and consumers are allowed to optimize their positions and reduce the risk of imbalance, which entails fees to be paid to system operators. Our study arises as the natural generalization of [57] and takes the perspective of a small agent interested in exploiting the stylized features of intraday prices in order to maximize her expected terminal gains.

We propose a stochastic model for the continuously traded electricity price and formulate an expected profit maximation problem in the language of Stochastic Optimal Control. The agent maximizes the expected utility of her portfolio over a certain set of trading strategies, i.e. she studies the quantity

$$\sup_{\pi} \mathbb{E}[U(X(T))],$$

where U is the utility function representing the risk profile of the investor, $X = X^{\pi}$ denotes the portfolio value associated to the strategy π and T is the trading closure time. Furtherly, we assume that the investor trades each hour separately, so to keep the analysis within a one-dimensional framework.

We assume a non-Gaussian Ornstein-Uhlenbeck (OU) process (with possibly time-dependent deterministic coefficients) to drive the power prices, being able to reproduce the stylized features of price movements:

$$dS(t) = -\lambda S(t)\, dt + dL(t), \qquad\qquad 0 < t \leq T,$$

where L is an additive process of the form

$$dL(t) = b(t)\, dt + \sigma(t)\, dW(t) + \psi(t) \int_{y\in\mathbb{R}} y\, \overline{N}(dy, dt),$$

with regular deterministic coefficients. This class is highly flexible and capable to reproduce both the mean-reverting and the spiky behavior of observed time series. Since we do not perform a logarithmic transformation of the price, negative prices can be reproduced by our model. Motivated by positivity constraints on X (and in line with general theoretical results [85]), we write the portfolio dynamics as

$$dX(t) = \bar{\pi}(t) X(t)\, dS(t), \qquad\qquad 0 < t \leq T,$$

being $\pi(t) = \bar{\pi}(t)X(t)$ the amount of power traded at time t. The range of admissible values of $\bar{\pi}$ depends on the support of the Lévy measure ν, that is the jump size (and sign) of L. We will assume the range of $\bar{\pi}$ to be compact.

Our approach is based on the dynamic programming method and the study of the associated Hamilton-Jacobi-Bellman (HJB) integro-differential equation. The problem of finding the value function is not straightforward from the formulation of the HJB equation. We study the case of logarithmic utility in order to disentangle the terms depending on both the wealth process and the strategy, from the ones depending on time and price. We reduce the fully nonlinear HJB equation to a linear partial integro-differential equation (PIDE) by means of an exponential transform $H(t,s,x) = U(xe^{g(t,s)})$. Maximizing the generalized Hamiltonian provides us with a representation for an optimal strategy:

$$\bar{\pi}^*(t) = \bar{\pi}^*(t, S(t-)),$$

where, for fixed $t \in [0,T]$ and $s \in \mathbb{R}$, $\bar{\pi}^* = \bar{\pi}^*(t,s)$ solves the integral equation

$$b(t) - \lambda s - \sigma(t)^2\bar{\pi}^* - \int_{\mathbb{R}} \frac{\bar{\pi}^*\psi(t)^2y^2}{1+\bar{\pi}^*\psi(t)y}\,\nu(dy) = 0. \tag{1}$$

Even though the optimal strategy is given implicitly, we are able to show that it is well-defined and satisfies sufficiently regular properties.

We prove the existence of a classical solution to the HJB equation in two cases of interest:

1. time-inhomogeneous compound Poisson processes with non-degenerate Brownian component,

2. additive pure-jump processes of (possibly) infinite variation.

This is done in the first case relying on a result by [108], while in the second case via Feynman-Kač representations. In the latter approach, we follow the idea of [20] and generalize it to time-inhomogeneous processes with possibly infinite variation. In particular, Danskin's theorem [54] allows us to prove that the forcing term of the HJB equation, which is defined as the composition of non-differentiable functions, is actually differentiable. Then, by applying a verification theorem, we prove H to be equal to the value function of the original problem.

We then study two easy and convenient ways to compute the optimal strategy by approximation, based on the Taylor expansion of the numerical integral equation in (1). The first one (*à la* Merton [96]) is

$$\bar{\pi}_1^*(t,s) = \frac{b(t)-\lambda s}{\sigma(t)^2+\sigma_L(t)^2},$$

where $\sigma_L(t)^2 := \psi(t)^2 \int_{y\in\mathbb{R}} y^2\,\nu(dy)$ corresponds to the local variance of the jump component of L. The second one is defined for compound Poisson processes and solves

$$(b(t)-\lambda s) - \sigma(t)^2\bar{\pi}_2^*(t,s) - \frac{\bar{\pi}_2^*(t,s)\psi(t)^2\mu_F^2}{1+\bar{\pi}_2^*(t,s)\psi(t)\mu_F}i = 0,$$

where μ_F is the mean of the jump size and i the jump intensity. We also derive some estimates of the approximation error based on the moments of the jump component.

We test our results on a popular electricity price model, namely the *factor model* introduced by [18]. Though this model is designed for day-ahead markets, we use it as example for mainly two reasons. Firstly, the stylized features of intra-day markets are similar to the ones observed in the day-ahead price series. Secondly, the factor model is based on a Lévy Ornstein-Uhlenbeck process of the same family as our model. Our main finding here is that Merton's ratio performs poorly in comparison to our jump-based approximation, which is rather close to the exact one instead. This suggests that optimal trading in Lévy-driven markets seems not well described by Merton-type strategies, at least for the specific case of additive mean-reverting prices.

Overview of Chapter 2

This chapter is based on a joint work [26] with Fred Espen Benth and Tiziano Vargiolu.

This paper aims at developing a consistent and tractable framework for a multicommodity energy forward market.[4] We apply the Heath-Jarrow-Morton paradigm [76], which consists in describing forward prices as stochastically evolving functions of time and delivery dates, across different markets. We propose an Ornstein-Uhlenbeck-type of model with deterministic coefficients parameterized over delivery times. By formulating the model under the market probability measure $\mathbb{P}$, we can represent the stylized empirical behavior of observed prices, such as mean-reversion and seasonalities. Let $f(t,T)$ denote the vector of futures prices at time t for instantaneous delivery at time T for possibly many commodities. Analogously, for any $T_1 < T_2$, let $F(t,T_1,T_2)$ be the vector of futures prices at time t for delivery over $[T_1,T_2]$. Under the physical measure $\mathbb{P}$ we assume the following stochastic dynamics:

$$\begin{aligned} df(t,T) &= (c(t,T) - \lambda(t)f(t,T))\,dt + \sigma(t,T)\,dW(t) + \psi(t,T)\,dJ(t), \\ dF(t,T_1,T_2) &= (C(t,T_1,T_2) - \lambda(t)F(t,T_1,T_2))\,dt \\ &\qquad + \Sigma(t,T_1,T_2)\,dW(t) + \Psi(t,T_1,T_2)\,dJ(t). \end{aligned}$$

The dynamics above is additive, meaning that we do not perform a logarithmic transformation of prices. Furthermore, this multicommodity framework allows to build models where commodities have various kinds of dependencies, like correlations among the driving processes W and J or cointegration/price couplings among them.

We want to design a forward market model with the same philosophy of [23], where it is theoretically possible to trade contracts with any delivery period and no-arbitrage relations must hold among them. For this reason, we impose the following condition:

$$F(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(t,T)\,dT, \tag{2}$$

for any $t \leq T_1$ and $T_2 > T_1 > 0$, where $\widehat{w}$ is a matrix-valued weight function and represents the time value of money. This will depend in general on the risk-free rate and can differ from contract to contract according to how settlement takes place (see Equation 4.2 in [23]).

As pointed out in [23], specifying a stochastic evolution for the forward curve and then deriving the swap price as the average over the delivery times has the disadvantage of losing desirable distributional features and, in some cases, even the Markov property. This results in non-tractable models, which inherit a complex probabilistic structure. Consequently, we express forward prices as affine transformations of a universal source of randomness, the latter being independent of the delivery date, i.e.

$$f(t,T) = \alpha(t,T)X(t) + \beta(t,T),$$

for some multi-dimensional stochastic process X that we characterize and given deterministic matrix/vector-valued functions α and β. In view of the no-arbitrage condition (2), we define the swap price process $F(t,T_1,T_2)$ by

$$F(t,T_1,T_2) \overset{def}{=} \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(t,T)\,dT.$$

[4] We do not make distinctions between forwards and futures, since the results are the same in the case of deterministic interest rate, as we assume here. We reserve the name *forward* to contracts with delivery at a fixed future time, while agreements to deliver the commodity over a period (e.g. month or year) are called *swaps*.

This simplifying assumption allows both to preserve Markovianity and to describe consistently the related swap price processes. Most importantly, this is crucial in order to prove the existence of equivalent martingale measures $\mathbb{Q}$, that in turn implies that the model is arbitrage-free. In fact, besides the static arbitrage opportunities that arise from trading in contracts with overlapping delivery periods, we need to ensure that our framework is free of arbitrage also in the classical "dynamic" sense. The main mathematical difficulty in doing this originates from the presence of mean-reversion. In this regard, we prove the martingale property of stochastic exponentials where the Lévy part is of Esscher-type, while the Girsanov kernel of the Brownian component is affine in the state variable and thus, in particular, stochastic. Then, we give a different proof in the case of continuous density and continuous kernel, by applying a weak Novikov-type condition on the series representation of the exponential function. The proof of this result relies on the asymptotic properties of Gaussian moments. As we move on to consider more general Lévy kernels, we see that the same technique can not be applied. This is shown to be related to the fact that various examples of infinitely divisible distributions, except the Gaussian distribution, do not satisfy the needed moments asymptotics.

After validating the theoretical framework, we move to specifying two exemplary models. The first is a generalization of the additive two-factor Lucia-Schwartz model [95]. We extend it by introducing a mean-reverting arbitrage-free forward dynamics, which is capable to describe a finer volatility term structure. This allows us to account for seasonal effects in price variations, as is typically observed, for instance, in power or gas markets. A calibration procedure and an empirical application to the German power futures market is carried out in Chapter 3. Furtherly, we introduce a multidimensional model for a mean-reverting cointegrated forward market that respects the no-arbitrage constraints in (2). In particular, we see how these constraints imply certain conditions on the mean-reversion coefficients of the futures curve dynamics.

Overview of Chapter 3

This chapter is based on a joint work [93] with Luca Latini and Tiziano Vargiolu.

The purpose of this chapter is to introduce a suitable model for power and gas futures prices by specifying a tractable dynamics from the framework developed in Chapter 2. In particular, our model is capable to hold together mean-reversion and arbitrage theory. In addition, we develop an ad hoc calibration procedure and perform an empirical application to the EEX power futures market.

As we already mentioned, power and gas, differently from other commodities, are delivered as a given intensity over a certain time interval (e.g. a month, quarter or year). In financial terms, this means that energy forward contracts prescribe a delivery period for the underlying, rather than a maturity date. The market mechanism is described in Figure 1. Since futures "overlap", suitable no-arbitrage conditions are needed when developing pricing models. Given a set of overlapping contracts, their futures prices must satisfy:

$$F(t, T_1, T_n) = \frac{1}{T_n - T_1} \sum_{i=1}^{n-1} (T_{i+1} - T_i) F(t, T_i, T_{i+1}). \tag{NA}$$

Furthermore, price series exhibit mean-reversion and strong seasonality, which drives also prices' volatility (see Figure 2).

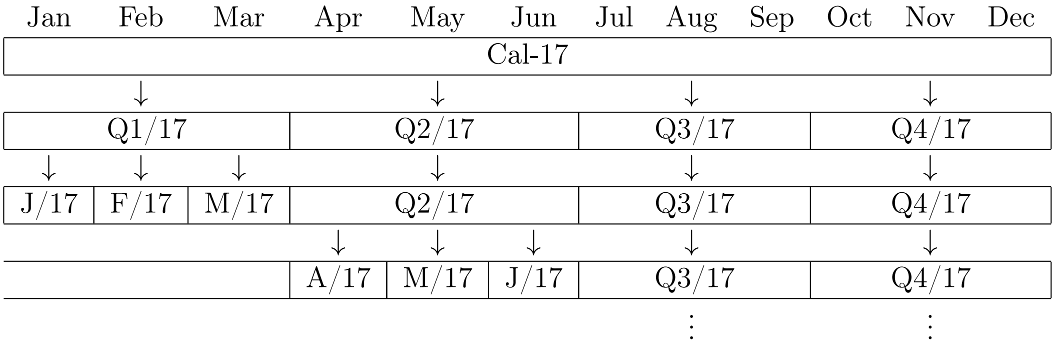


Figure 1: Cascade unpacking mechanism of forward contracts. For each given calendar year, as time passes forwards are unpacked first in quarters, then in the corresponding months. It may happen that the same delivery period is covered in the market by different contracts, e.g. one simultaneously finds quotes for Jan/17, Feb/17, Mar/17 and Q1/17.

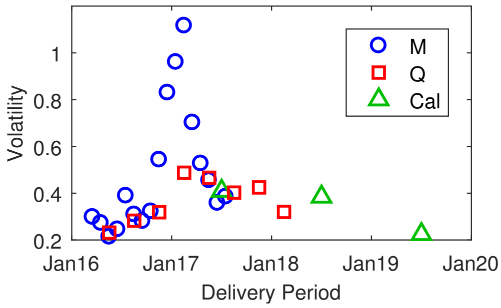


Figure 2: Realized volatilities of Phelix Base Futures (German power futures) from January 4, 2016 to May 23, 2017. The Samuelson effect (=volatility increasing as time approaches delivery) is not sufficient to explain this term structure.

We adopt the Heath-Jarrow-Morton methodology (vs. spot price based models), which consists in describing the whole forward curve by parametric SDEs. The theoretical framework is constructed as follows (see also Chapter 2). Let $f(t,T)$ denote the forward price at time t for instantaneous delivery at time T. Analogously, for any $T_1 < T_2$, let $F(t,T_1,T_2)$ be the forward price at time t for delivery period $[T_1,T_2]$. Futures prices evolve under the real-world probability measure $\mathbb{P}$ by

$$\begin{aligned} df(t,T) &= \big(c(t,T) - \lambda(t) f(t,T)\big)\, dt + \theta(t,T)\, dW(t), \\ dF(t,T_1,T_2) &= \big(C(t,T_1,T_2) - \lambda(t) F(t,T_1,T_2)\big)\, dt + \Sigma(t,T_1,T_2)\, dW(t), \end{aligned}$$

where all the coefficients are deterministic. Since we have not performed a logarithmic transformation of the price, the dynamics are additive. In power and gas markets the instantantaneous forwards do not exist. However, they play the role of building blocks for the dynamics of traded forwards:

$$F(t,T_1,T_2) \overset{def}{=} \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t,u)\, du,$$

for all $t \le T_1$ and $T_1 < T_2$. This relation implies all the no-arbitrage constraints of type (NA) (no static arbitrage). Furthermore, this relation identifies the structure of the coefficients in the equation of $F(t,T_1,T_2)$ in terms of the coefficients of $f(t,T)$.

Dynamic arbitrage opportunities are prevented by constructing an equivalent martingale measure for the market, i.e. a probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that

$$\begin{aligned} df(t,T) &= \theta(t,T)\ dW^{\mathbb{Q}}(t), \\ dF(t,T_1,T_2) &= \Sigma(t,T_1,T_2)\ dW^{\mathbb{Q}}(t), \end{aligned}$$

for some $\mathbb{Q}$-Brownian motion $W^{\mathbb{Q}}$. As already discussed in the overview of Chapter 2, this is not trivial due to the affine form of the drift. For this reason we introduce the key assumption:

$$f(t,T) = \alpha(t,T)X(t) + \beta(t,T),$$

where X is a universal source of randomness and α and β are deterministic coefficients. This is not observable, but its coefficients do not appear explicitly in the resulting dynamics.

This class of models is not void: the well-known Lucia-Schwartz model [95] turns out to be of this kind:

$$\begin{aligned} df(t,T) &= \lambda(t)\big(\phi(T) - f(t,T)\big)\ dt \\ &\quad + e^{-\kappa(T-t)}\sigma_1\ dW_1(t) + \sigma_2\ dW_2(t), \\ dF(t,T_1,T_2) &= \lambda(t)\left(\frac{1}{T_2-T_1}\int_{T_1}^{T_2}\phi(T)\ dT - F(t,T_1,T_2)\right) dt \\ &\quad + \sigma_1\, e^{\kappa t}\,\frac{(e^{-\kappa T_1} - e^{-\kappa T_2})}{\kappa(T_2-T_1)}\, dW_1(t) + \sigma_2\, dW_2(t). \end{aligned}$$

However, the volatility coefficients succeed only in reproducing the so-called Samuelson effect[5] as an exponential decay but they are not able to capture a finer term structure (cf. Figure 2). In this chapter we propose a modification of the Lucia-Schwartz model such that both price level and volatility are allowed to have a non-trivial term structure:

$$\begin{aligned} df(t,T) &= \lambda(t)\big(\phi(T) - f(t,T)\big)\ dt \\ &\quad + e^{-\kappa(T-t)}\sigma_1\ dW_1(t) + \psi(T)\ dW_2(t), \end{aligned}$$

$$\begin{aligned} dF(t,T_1,T_2) &= \lambda(t)\big(\Phi(T_1,T_2) - F(t,T_1,T_2)\big)\ dt \\ &\quad + e^{\kappa t}\ \Gamma(T_1,T_2)\ dW_1(t) + \Psi(T_1,T_2)\ dW_2(t), \end{aligned}$$

where

$$\begin{aligned} \Phi(T_1,T_2) &:= \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\phi(u)\, du \\ \Gamma(T_1,T_2) &:= \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\sigma_1 e^{-\kappa u}\, du = \frac{\sigma_1(e^{-\kappa T_1} - e^{-\kappa T_2})}{\kappa(T_2-T_1)}, \\ \Psi(T_1,T_2) &:= \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\psi(u)\, du. \end{aligned}$$

From this model, we can recover various stylized facts typical of energy markets. The mean-reversion speed $\lambda(t)$ can be time-dependent (but not maturity-dependent in this formulation), and the long-term mean $\phi(T)$ is exactly the seasonal component of the spot price. The same

[5]The Samuelson effect is an observed feature of prices volatilities, common to many commodity markets, consisting of increasing volatility of prices as time approaches maturity [115].

for $F(\cdot, T_1, T_2)$, where the long-term mean is the maturity-average of $\phi(T)$, $T \in [T_1, T_2]$. We have a Samuelson effect in both the instantanteneous (non-traded) forward prices $f(\cdot, T)$ and the traded forward prices $F(\cdot, T_1, T_2)$. Also, the second factor describes the volatility term structure in an arbitrage-free way. Finally, forward prices with shorter delivery periods are more volatile than forward prices with longer delivery periods. For instance, being equal the time to maturity, monthly contracts are more volatile than quarters (3 months) or calendars (1 year).

Estimation is performed directly on the market time series of traded forwards $F(\cdot, T_1, T_2)$ and no forward curve smoothing procedure is needed. For the volatility this task is not trivial, since the term structure is also maturity-dependent.We apply a method based on quadratic variation/covariation of price processes, which allows to estimate the diffusion coefficients. Then, in order to estimate the mean-reversion speed and level, since the covariance matrix of joint returns is singular, we first calibrate the parameters on the single contract. Afterwards, we take into account the no-arbitrage constraints (NA) that must hold among overlapping contracts (e.g. 3 months/1 quarter). This is done by combining maximum likelihood estimation (MLE) with Lagrange multipliers. For example, with the convention that $\Phi_{\mathrm{Q2/17}}$ denotes the parameter $\Phi(T_1, T_2)$ corresponding to the contract Q2/17, then

$$\Phi_{\mathrm{Q2/17}} = u_{\mathrm{Apr/17}}\Phi_{\mathrm{Apr/17}} + u_{\mathrm{May/17}}\Phi_{\mathrm{May/17}} + u_{\mathrm{Jun/17}}\Phi_{\mathrm{Jun/17}},$$

where the weights u_i are defined according to the number of days in the month/quarter (e.g. $u_{\mathrm{Apr/17}} = 30/91$).

We apply this estimation technique on the Phelix Base Futures market. We consider all the daily closing prices of each monthly, quarterly and calendar forward contract traded from January 4, 2016 to May 23, 2017. We first estimate the diffusion coefficients by introducing in the second factor a classical seasonal component, which takes into account a possible non-seasonal long-run volatility trend, i.e. a linear combination of trigonometric functions plus a linear component in maturity. Then, we compare it to a nonparametric volatility shape, where the second factor is a free parameter for each "atomic" forward. In this way we manage to reproduce in a fairly genuine way the finer (co-)volatility term structure observed.

We then do a simulation study and assess the performance of the model. Firstly, we compare simulated paths of some exemplary futures contracts to the corresponding observed trajectories, so to discuss the qualitative behavior of model simulations. Secondly, we compute fundamental statistics of the model by averaging the results of a set of simulations and compare them to our data. The model reproduces in a satisfactory way the trajectorial and statistical features of our dataset. In particular, in terms of moments, a Gaussian distribution seems reasonable.

Overview of Chapter 4

This chapter is based on a joint work [110] with Maren Diane Schmeck and Tiziano Vargiolu.

In this chapter we introduce an additive two-factor model for futures prices based on Normal Inverse Gaussian Lévy processes and compute European option prices by Fourier transform methods. We introduce a calibration procedure and fit the model to EEX power option settlement prices.

First, we introduce a general framework for multifactor additive futures prices in the same spirit of Chapter 2 (see also [32]) and, from this, we focus on a two-factor model that generalizes the one presented in Chapter 3. This two-factor model will be then calibrated and fitted to market data. We assume that the stochastic evolution of $F(\cdot, T_1, T_2)$ from t to T is given by

$$F(T, T_1, T_2) = F(t, T_1, T_2) + \int_t^T \Gamma_1(u, T_1, T_2)\, dJ_1(u) + \Gamma_2(T_1, T_2)(J_2(T) - J_2(t)), \tag{3}$$

where

$$\Gamma_1(u, T_1, T_2) := \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \gamma_1 e^{-\mu(\tau - u)}\, d\tau = \frac{\gamma_1(e^{-\mu(T_1-u)} - e^{-\mu(T_2-u)})}{\mu(T_2 - T_1)},$$
$$\Gamma_2(T_1, T_2) := \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \gamma_2(\tau)\, d\tau.$$

Differently from Chapters 2 and 3, this dynamics is described directly under a risk-neutral measure $\mathbb{Q}$.[6] The coefficients modulate the variability of the two stochastic processes J_1 and J_2, which are both centered (i.e. zero-mean) versions of NIG Lévy processes. This model arises as a natural generalization of its Gaussian counterpart introduced in Chapter 3. The first factor has a delivery-averaged exponential behavior, meant to reproduce the Samuelson effect. The second factor is independent of time, but varies for contracts with different delivery in a *seasonal* and *no-arbitrage* way (see Chapters 2 and 3) and accounts for a finer reproduction of the term structure of futures volatilities. This function can be specified either in a parametric or nonparametric fashion. In order to apply the Fourier transform approach, we then compute the characteristic function of $F(T, T_1, T_2)$. By introducing $Z(t, T, T_1, T_2) := F(T, T_1, T_2) - F(t, T_1, T_2)$, we denote its characteristic function (as a function of $v \in \mathbb{R}$) by

$$\Psi(t, T, T_1, T_2, v) = \mathbb{E}\left[e^{ivZ(t,T,T_1,T_2)} \,|\mathcal{F}_t\right].$$

Then, we move to considering the pricing of European vanilla options written on futures contracts. We discuss the method for call options, being the case of puts completely analogous. Let $C(t; T, K, T_1, T_2)$ be the price at time t of a call option, with strike price K and exercise time T, that is written on a futures contract with delivery period $[T_1, T_2]$. By no-arbitrage,

$$C(t; T, K) = \mathbb{E}\left[(F(T) - K)_+ \,|\mathcal{F}_t\right], \tag{4}$$

where $F(T)$ is the price of the underlying futures at the option exercise and $\mathcal{F}_t$ represents the information at time t.[7] We follow the classical approach of [45], which consists, roughly speaking, in computing the Fourier transform of (4) as a function of K (after proper manipulations) so to recover the option value by inverse tranform. The starting point is the observation that, as K goes to $-\infty$, $C(t; T, K) \to \infty$, so that in particular $C(t; T, K)$ is not integrable as a function of K for large negative values. This means that the option value $C(t; T, K)$ does not satisfy the assumptions required for computing its Fourier transform. In order to overcome this, [45] suggest two approaches that we recall and apply to the case of additive models. The first approach is based on the use of an exponential damping factor e^{aK} in order to make the option value integrable. Instead, the second approach consists in substracting the time value of the option. In this way we derive semi-analytical expressions (analytical up to integration) for the option prices, that depend on the characteristic function of the underlying.

We discuss a calibration methodology that we will apply in our empirical study. First, we discretize the option value that is given in integral form (being it represented as inverse Fourier

[6] The absence of the drift follows from the fact that, by no-arbitrage, futures prices must be martingales under $\mathbb{Q}$.

[7] We do not write the dependence on the delivery period in order to ease the notation.

transform) in the domain of integration. Secondly, we select a finite grid of strike prices, that consists in practice of the listed options available in the market for a given underlying. The calibration happens statically, in the sense that we fix a trading date and observe the market option prices for different strikes and underlyings. We then find the parameters that minimize an objective function representing the distance from theoretical to model implied volatilities (IVs). In the same way, one can alternatively use option prices instead of IVs. The futures model under consideration is defined in a such a way that there is no possibility of arbitrages from trading in overlapping delivery periods. No-arbitrage implies certain relations on the coefficients, that translate into parameter constraints (cf. the calibration procedure presented in Chapter 3). In order to highlight this, we present the objective function and the parameters set for the following three cases: single maturity, many maturities but non-overlapping delivery periods, and general case of possibly overlapping delivery periods.

The contracts that we consider in our application are European-styled vanilla options written on the Phelix Base Index traded at EEX. The underlying assets are futures contracts that prescribe the delivery of 1 MW per hour, for each hour of each day of a month, a quarter or a year. Since there is not enough liquidity in the market in order to extract information on the IV surface from traded market quotes, we consider the settlement prices, that are available for a sufficiently large range of strike prices. Though settlement prices do not represent trades that really take place in the market, they contain information on the market expectations. We observe the market for a representative day: Monday, March 5, 2018. For each option, we consider the strike prices in the range 90%-110% of the current price of the underlying. At this date, the listed options with available settlement prices are five for monthly contracts, six for quarterly contracts and three for calendars.

We perform the calibration procedure described above, first, for the one-factor model derived by (3) by setting the first coefficient to 0, i.e. $\Gamma_1(u, T_1, T_2) \equiv 0$ and, then, for the general two-factor model. We compare the IVs of both models to the empirical IVs and the one (constant across strikes) generated by Black's model. For the sake of comparison we choose to minimize for all the three models the error on option prices (that is a faster method to apply in practice than the one based on IVs) and use the modified time value approach (for the one-factor and the two-factor). We find that, in general, the optimization routine falls into local minima. However, by selecting a starting condition that is "sufficiently close" to observed IVs, the minimization converges to a realistic set of parameters (cf. [52] for well-posedness of this kind of problems). As a by-product from the estimation of the one-factor model, we derive that, under the risk-neutral measure, futures prices are leptokurtic and have significantly positive skewness. This is reflected also into the shape of empirical IVs, which display a *forward skew* (i.e. higher IVs for out-of-the-money calls). This can be interpreted as a "risk premium" paid by option buyers for securing supply. Generally, the two-factor model fit is rather accurate for all the strikes and maturities available in the market. However, the "smile" of model IVs of closest to maturity options is slightly more pronounced than what observed. There is a clear improvement of fit from the one-factor to the two-factor model, since, for instance, IVs of out-of-the-money calls are underestimated by the one-factor model (and not from the two-factor one). Also, there is a slight anomaly in some at-the-money IVs, that we believe it is due to numerical instability of the Fourier method for $K = F(t)$ (cf. [45]).

Acknowledgement: First and foremost, I would like to thank my supervisor Prof. Tiziano Vargiolu for coauthoring with me all the material appearing in this Thesis and for supporting me scientifically and financially throughout my research and all the activities related to it. Thank you to Fred Espen Benth and Maren Diane Schmeck for inviting me to work with them, respectively, at Oslo University and the Center for Mathematical Economics in Bielefeld, and for their remarkable hospitality and kindness during my research stay.

I am also thankful for helpful suggestions and discussions to Alessandro Balata, Giorgia Callegaro, Sebastiano Don, Enrico Edoli, Giorgio Ferrari, Martino Grasselli, Paolo Luzzini, Guglielmo Pelino, Anca Pircalabu, Frank Norbert Proske, Wolfgang Runggaldier and Carlo Sgarra.

The material appearing in this Thesis has been presented at various conferences and workshops. I would like to thank the organizers and the conference participants at the Energy Finance Italia II and III Workshops in Padova and Pescara, XVIII and XIX Quantitative Finance Workshops in Milano and Roma, Commodity and Energy Markets Annual Meetings in Oxford and Roma, Energy Finance Christmas Workshop in Krakow, 41st and 42nd Annual Meetings of the Association for Mathematics Applied to Social and Economic Sciences (AMASES) in Cagliari and Napoli, Second Conference on Mathematics of Energy Markets in Vienna and the First Italian National Meeting on Probability and Mathematical Statistics in Torino.

Finally, I would like to express my gratitude to the reviewers for carefully reading and commenting this Thesis.

CHAPTER 1

Optimal portfolio in intraday electricity markets modelled by Lévy-Ornstein-Uhlenbeck processes

We study an optimal portfolio problem designed for an agent operating in intraday electricity markets. The investor is allowed to trade in a single risky asset modelling the continuously traded power and aims to maximize the expected terminal utility of his wealth. We assume a mean-reverting additive process to drive the power prices. In the case of logarithmic utility, we reduce the fully non-linear Hamilton-Jacobi-Bellman equation to a linear parabolic integro-differential equation, for which we explicitly exhibit a classical solution in two cases of modelling interest. The optimal strategy is given implicitly as the solution of an integral equation, which is possible to solve numerically as well as to describe analytically. An analysis of two different approximations for the optimal policy is provided. Finally, we perform a numerical test by adapting the parameters of a popular electricity spot price model.

1.1 Introduction

After power markets have been deregulated worldwide, this financial sector is experiencing profound structural changes. Several interesting issues have triggered a growing interest in models for energy markets. A rigorous mathematical approach may be useful for practitioners and at the same time stimulate the advance in academic research. In this work we consider markets structured as the European Power Exchange (EPEX), which regulates electricity spot trading in Central and Western Europe. In many exchanges, short-term trade is organized in mainly two markets: day-ahead and intraday. While the day-ahead market trades electricity for each hour (or block of hours) to be delivered the next day and is auction-based, the intraday market, which opens after the day-ahead closure, offers the participants the possibility to trade continuously until short time prior to delivery. The intraday market is especially important for renewable energy producers, who can adjust their day-ahead positions due to changes of weather forestcasts [88].

In this work we study a dynamic portfolio optimization problem designed for intraday electricity trading. These markets play an important role in the equilibrium of power grids, since both electricity producers and consumers are allowed to optimize their positions and reduce the risk of imbalance, which entails fees to be paid to system operators. Since the increasing penetration of renewable sources, modelling intraday trading has become particularly important, as well as mathematically interesting. Literature related to this problem is very recent: one of the first paper in this direction is [77], where the author studies how a wind power producer may benefit from trading in intraday markets by taking into account the risk of forecast production errors. [71] study an optimal trade execution problem in order to compensate forecast errors on wind or photovoltaic power production. [3] consider a producer who aims to minimimize the imbalance cost of his residual demand (which is stochastic) by controlling his flexible power production (thermal plants) and his position on the intraday market. Another recent paper [66] studies a stochastic multiperiod portfolio optimization problem in discrete time for hydroassets management and derives, in particular, an optimal intraday trading strategy. In [88] the authors investigate the impact of intraday updated forecasts of wind and photovoltaic on the bidding behavior of market participants, while [126] compare the price drivers in both the EPEX day-ahead and intraday electricity markets. Also, [46] study cross-border effects in intraday prices between interconnected locations and [123] consider a wind energy producer who trades in forward, spot, intraday and adjustment markets and derive optimal trading policies taking into account that his forecast production is imperfect. Our study arises as natural generalization of [57] and takes the perspective of a small agent interested in exploiting the stylized features of intraday prices in order to maximize his expected terminal gains.

We propose a stochastic model for the continuously traded electricity price and formulate an expected profit maximation problem in the language of stochastic control. The price is modeled by additive non-Gaussian Ornstein-Uhlenbeck (OU) processes. Power spot prices are usually described by either geometric [48, 73] or additive mean-reverting processes [18, 26, 90, 93, 95]. In the context of intraday markets, our model choice generalizes both [57], where the authors model the price by Gaussian OU processes, and [88], where the intraday price is an AR(1) process (with regime-switching that we do not consider), which is the discrete time version of the OU process. This class is highly flexible and capable to reproduce both the mean-reverting and the spiky behavior of observed time series. Since we do not perform a logarithmic transformation of the price, as is usually done, negative prices can be reproduced by our model: this is consistent with what has been recently observed since the introduction of renewable energy sources in the power mix (see [61]).

Following the pioneering work of Merton [96, 97], optimal portfolio management has become one of the most popular problems in mathematical finance and has been addressed in several different frameworks. For a general treatment we refer the reader to [104]. Our approach is based on the dynamic programming method and the study of the associated Hamilton-Jacobi-Bellman (HJB) integro-differential equation. Some related works include e.g. [19], where the price evolves as an exponential Ornstein-Uhlenbeck process, known as the Schwartz model, which is ubiquitous in commodity prices modelling. In [85] the authors model the risky assets with exponential Lévy processes and [107] generalize to exponential additive processes. We are inspired by [109, 128] when introducing a transformation for solving the HJB equation. Also, [20] study the same optimization problem for the Barndorff-Nielsen-Shephard model [14], where the volatility is a superposition of non-Gaussian OU processes driven by subordinators.

The problem of finding the value function is not straightforward from the formulation of the HJB equation. We study the case of logarithmic utility in order to disentangle the terms depending on both the wealth process and the strategy, from the ones depending on time and price. This simplification was also observed by [1] for a certain class of jump-diffusion processes. A theoretical study of optimal portfolios in the case of logarithmic utility is performed in [74] and furtherly generalized in [75], where the authors apply martingale methods (see [74, 75] and references therein for details on this approach) to a general semimartingale framework.

However, although their analysis provides a characterization of an optimal strategy and its uniqueness, the study of its analytical properties is not explicitly addressed. We reduce the fully nonlinear HJB equation to a linear partial integro-differential equation (PIDE) by applying a logarithmic transformation as in [109]. Even though the optimal strategy is given implicitly as the solution of an integral equation, we are able to show that it is well-defined and satisfies sufficiently regular properties in order to apply the Verification Theorem.

We prove the existence of a classical solution to the HJB equation in two cases of interest: time-inhomogeneous compound Poisson processes with non-degenerate Brownian component and additive pure-jump processes of (possibly) infinite variation. This is done in the first case relying on a result by [108], while in the second case via Feynman-Kač representations. In the latter approach, we follow the idea of [20] and generalize it to time-inhomogeneous processes, which we do not assume to be of finite variation as in [20]. In particular, Danskin's theorem [54] allows us to prove that the forcing term of the HJB equation, which is defined as the composition of non-differentiable functions, is actually differentiable.

Partial integro-differential equations (PIDE) are in itself of interest and arise across different fields of mathematics. In our paper we consider classical solutions, obtained via probabilistic representations, partly as extensions or complementary contributions of various earlier works. A classical reference for this type of problems is [15], where some existence results are stated under strong regularity assumptions on the coefficients of the equation. In proving the existence of a regular solution for finite Lévy measures with non-degenerate Brownian component, we apply a result of [108]. Nevertheless, this approach is based on classical smoothness results from PDE theory for linear second-order differential equations (cf. [70]), which require the finiteness of the jump measure. Consequently, for more general jump-processes we instead follow [20], where the Feynman-Kač formula yields a candidate, which is proven to be a classical solution of a PIDE very similar to ours. Unfortunately, this approach works only in the first-order case, i.e. with no Brownian component. However, as observed by [53], in order to generate realistic price trajectories, it is sufficient to consider financial models which are either finite activity jump processes combined with a diffusion part, or infinite activity pure-jump models, since the latter behave in a "diffusive" way when frequent small jumps occur.

We then study an approximation of the optimal strategy based on the Taylor expansion of the first-order condition, which is a numerical integral equation. In the case of compound Poisson processes, the center of the polynomial is chosen as the mean jump size. We compare it to the classical Merton ratio [96], which is shown to correspond to a Taylor expansion around zero. This approximation has been studied also in [8, 28, 105, 107] for stochastic volatility price models with jumps. Nevertheless, in their approach the authors start by approximating the HJB equation directly, while we work on the first-order condition. We derive some estimates of the approximation error and perform a numerical test on a power spot price model, specifically the factor model in [21]. Our main finding here is that Merton's ratio performs poorly in comparison to our jump-based approximation, suggesting that optimal trading in Lévy-driven markets is not well described by this economically meaningful quantity, at least for the specific case of additive mean-reverting prices.

The paper is structured as follows. In Section 2 we introduce the intraday price dynamics and set our stochastic control problem. In Section 3 we describe the properties of our optimal strategy and study the reduced HJB equation for a logarithmic utility. In particular, two existence results of classical solutions to PIDE are given. We conclude this section by applying the Verification Theorem. The approximation study of the optimal strategy is contained in Section 4, while Section 5 presents an exemplary numerical test on the policy approximations.

Appendix A includes auxiliary propositions for the existence of a PIDE solution in the first-order case, while in Appendix B we collect some of the most technical proofs.

1.2 The optimal portfolio problem

We follow the dynamic programming strategy for solving our stochastic optimal control problem (see, for instance, [69]). The purpose is maximizing the expected utility of our portfolio over a set of trading strategies, that is to study the quantity

$$\sup_{\pi} \mathbb{E}[U(X(T))], \tag{1.1}$$

where $U : \mathbb{R} \to \mathbb{R}$ is a utility function representing the risk profile of the investor, $X = X^{\pi}$ denotes the portfolio value associated to the strategy π and T is the trading closure time.

Let us introduce the stochastic dynamics driving the market. Denote by L a real-valued additive process (for details see e.g. [51, Section 14.1]) defined on the complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_u)_{u\geq 0}, \mathbb{P})$ by

$$dL(u) = b(u)\, du + \sigma(u)\, dW(u) + \psi(u) \int_{\mathbb{R}} y\, \overline{N}(dy, du), \qquad u \in (0, T], \tag{1.2}$$

where $\sigma, b, \psi : [0, T] \to \mathbb{R}$ are continuously differentiable functions, such that $0 \leq \psi_1 \leq \psi(u) \leq \psi_2$ for any $u \in [0, T]$ and $\sigma(u)$ and $\psi(u)$ do not vanish at the same time. The process W is a standard Brownian motion and $\overline{N}(dy, du) := N(dy, du) - \nu(dy)\, du$ is the compensated Poisson random measure associated to a Lévy measure ν, i.e. a Radon measure on $\mathbb{R} \setminus \{0\}$ such that $\int_{\mathbb{R}} (1 \wedge y^2)\, \nu(dy) < \infty$. In particular, if b, σ and ψ are constant, with $\psi \equiv 1$, then L is a Lévy process. In order to deal with processes with finite second moment, we furtherly assume that ν satisfies the following integrability condition:

$$\int_{|y|\geq 1} y^2\, \nu(dy) < \infty. \tag{1.3}$$

Observe that L can be decomposed in a deterministic drift part, a Brownian motion with time-varying volatility and a square integrable pure-jump martingale component. We also introduce the following convention.

Remark 1.2.1 (Jump measure support). It holds that supp $\nu \subset [m, M]$ for $-\infty \leq m \leq M \leq +\infty$. We interpret the case $m = M = 0$ formally as the diffusive case, i.e. when L has no jump component. If, for instance, $m = -\infty$, we mean $[m, M] := (-\infty, M]$.

We are in a market with one asset (i.e. the continuously traded intraday price of electricity, expressed in Euros per MWh), whose market value $S = S^{t,s}$ evolves in time according to the stochastic differential equation

$$dS(u) = -\lambda S(u)\, dt + dL(u), \qquad u \in (t, T], \tag{1.4}$$

given the initial condition $S^{t,s}(t) = s$ for some $t \in [0, T)$ and $s \in \mathbb{R}$. The constant λ is positive and represents the mean-reversion rate of S. In particular, for any additive process L there exists a unique (strong) solution S such that in general $\mathbb{P}(S(u) < 0) > 0$ for some $u \geq t$, i.e. the price may assume negative values. Nevertheless, if L can have only positive jumps and there is no Brownian component, i.e. it is a time-inhomogeneous subordinator (see e.g. [51]), then a nonnegative initial condition $s \geq 0$ naturally implies that $S(u)$ is a.s. nonnegative for each $u \geq t$. Therefore, it is possible to consider additive processes taking both positive and negative values, according to one's modelling preferences. The unique solution of (1.4), with starting condition $S^{t,s}(t) = s$, can be explicitly written as

$$S^{t,s}(u) = se^{-\lambda(u-t)} + \int_t^u e^{-\lambda(u-v)}\, dL(v). \tag{1.5}$$

Since L has finite second moment by (1.3), S has finite second moment as well.

If $\pi(u)$ represents the amount of shares of the stock (i.e. the amount of energy in MWh), owned by the agent at time u, the associated self-financing portfolio dynamics $X = X^{t,s,x;\pi}$ is described by

$$dX(u) = \pi(u)\, dS(u), \qquad u \in (t,T], \tag{1.6}$$

$$X(t) = x, \tag{1.7}$$

where $S = S^{t,s}$ and $x > 0$. We are assuming that there is an implicit cash position, i.e. a risk-free asset, with zero interest rate (even in the case that the interest rate is nonzero, it would not affect trading because of the short time interval). We also assume that $\pi(T) = 0$, so that the agent liquidates the position at the terminal time T; in other words, we are in a pure trading context (see [57, Remark 3.1]). We then define the set of admissible trading strategies and the value function.

Definition 1.2.2 (Admissible controls)**.** We call $\mathcal{A}([t,T])$ the set of admissible controls, which are defined as real-valued predictable processes π on $[t,T]$ (in the sense of [80, Definition 3.3]) such that the following conditions hold:

1. Equations (1.4) and (1.6) admit a unique strong solution $(S,X) = (S^{t,s,x}, X^{t,s,x;\pi})$ for each initial condition $S(t) = s, X(t) = x$, with $t \in [0,T)$ and $(s,x) \in \mathbb{R} \times \mathbb{R}^+$.

2. The associated wealth process is positive, i.e. $X^{t,s,x;\pi}(u) > 0$, $\mathbb{P}$-a.s. for each $u \in (t,T]$ and the final net position is zero: $\pi(T) = 0$.

Definition 1.2.3 (Value function)**.** If $(S^{t,s,x}, X^{t,s,x;\pi})$ denotes the controlled Markov process starting from (s,x) at time t and evolving as in (1.4) and (1.6), we define the value function by

$$V(t,s,x) = \sup_{\pi \in \mathcal{A}([t,T])} J(t,s,x;\pi),$$

where J is the objective function:

$$J(t,s,x;\pi) = \mathbb{E}[U(X^{t,s,x;\pi}(T))].$$

The function $U : \mathbb{R}^+ \to \mathbb{R}$ represents the investor's utility and is concave, increasing, and bounded from below.

Following the dynamic programming principle, the Hamilton-Jacobi-Bellman (HJB) equation associated to this optimization problem is

$$\frac{\partial}{\partial t} H(t,s,x) + \sup_{\pi} A^{\pi} H(t,s,x) = 0, \qquad (t,s,x) \in [0,T) \times \mathbb{R} \times \mathbb{R}^+, \tag{1.8}$$

$$H(T,s,x) = U(x), \qquad (s,x) \in \mathbb{R} \times \mathbb{R}^+. \tag{1.9}$$

According to (1.4) and (1.6), the infinitesimal generator A^π of the controlled process $(S^{t,s,x}, X^{t,s,x;\pi})$ acts on a sufficiently regular function $H(t,s,x)$ as follows

$$\begin{aligned} A^\pi H(t,s,x) = {} & (b(t)-\lambda s)(\pi H_x(t,s,x)+H_s(t,s,x)) \\ & + \frac{1}{2}\pi^2\sigma(t)^2 H_{xx}(t,s,x) + \pi\sigma(t)^2 H_{sx}(t,s,x) + \frac{1}{2}\sigma(t)^2 H_{ss}(t,s,x) \\ & + \int_{\mathbb{R}^2} [H(t,s+y_2,x+\pi\psi(t)y_1) - H(t,s,x) \\ & - (\pi H_x(t,s,x)\,\psi(t)y_1 + H_s(t,s,x)\psi(t)y_2)]\,\tilde{\nu}_t(dy_1dy_2), \end{aligned}$$

where $\tilde{\nu}_t$ is the jump measure associated to the two-dimensional process (S,X). Since this is a singular two-dimensional measure which coincides with the one-dimensional jump measure $\nu(dy)$ on the line $y_1 = y_2$, we can rewrite the integral term as

$$\begin{aligned} A^\pi H(t,s,x) = {} & (b(t)-\lambda s)(\pi H_x(t,s,x)+H_s(t,s,x)) \\ & + \frac{1}{2}\pi^2\sigma(t)^2 H_{xx}(t,s,x) + \pi\sigma(t)^2 H_{sx}(t,s,x) + \frac{1}{2}\sigma(t)^2 H_{ss}(t,s,x) \\ & + \int_{\mathbb{R}} [H(t,s+\psi(t)y,x+\pi\psi(t)y) - H(t,s,x) - (\pi H_x(t,s,x)+H_s(t,s,x))\,\psi(t)y]\,\nu(dy). \end{aligned}$$

To link the HJB equation to the control problem, we formulate a Verification Theorem in the version of [69, Theorem III.8.1]. The basic tool is the well-known Dynkin formula (see [69, p.122]), which here applies to the controlled process (S, X^π):

$$\mathbb{E}^{t,s,x}[f(T,S(T),X^\pi(T))] - f(t,s,x) = \mathbb{E}^{t,s,x}\left[\int_t^T A^{\pi(u)} f(u,S(u),X^\pi(u))\,du\right], \qquad (1.10)$$

where $\mathbb{E}^{t,s,x}$ denotes the conditional expectation given $S(t)=s$, $X^\pi(t)=x$ and $f:[0,T]\times\mathbb{R}\times\mathbb{R}^+\to\mathbb{R}$ is any function for which the expression makes sense.

Theorem 1.2.4 (Verification Theorem)**.** *Define the set*

$$\mathcal{D} = \{f \in C^{1,2}([0,T)\times\mathbb{R}\times\mathbb{R}^+) \text{ so that } (1.10) \text{ holds for each } \pi\in\mathcal{A}([t,T])\}.$$

Let $H\in\mathcal{D}$ be a classical solution of (1.8) which respects the terminal condition (1.9). Then it holds, for each $(t,s,x)\in[0,T]\times\mathbb{R}\times\mathbb{R}^+$,

1. *$H(t,s,x)\geq J(t,s,x;\pi)$ for each admissible control $\pi\in\mathcal{A}[t,T]$;*
2. *if there exists an admissible control $\pi^*\in\mathcal{A}[t,T]$ such that*

$$\pi^*(u) \in \arg\max_\pi A^\pi H(u,S(u),X^\pi(u)) \qquad \mathbb{P}\text{-a.s. for } u\in[t,T]$$

then $H(t,s,x)=J(t,s,x;\pi^*)=V(t,s,x)$, i.e. π^* is an optimal strategy.

Proof. The proof is classical and follows directly from the Dynkin formula in (1.10). □

1.3 Optimal control and value function for a logarithmic utility

In this section we solve the optimization problem in the case of a logarithmic utility, i.e. when the utility function in (1.1) is $U(x)=\log(x)$. Specifically, we find an explicit solution for the HJB equation by means of a logarithmic transform. First, we reduce the fully nonlinear

HJB equation to a linear parabolic integro-differential equation for which, under certain assumptions, the existence of a regular solution can be proven. By applying the Verification Theorem of the previous section, we prove it to be equal to the value function of the original maximization problem (Theorem 1.3.13). We also state the existence and give a representation of an optimal strategy, which is shown to solve an integral equation.

1.3.1 Optimal strategy

By the properties of logarithmic utility, it holds that, if an optimal strategy π^* exists, it takes the form

$$\pi^*(u) = \bar{\pi}^*(u)X(u-), \tag{1.11}$$

where $\bar{\pi}^*$ is a predictable process that can be implicitly defined in terms of the semimartingale characteristics of $S^{t,s}$, i.e. on the local behavior of the price process (see [74, Theorem 3.1]). This implies that we may explicitly characterize the strategies for which the wealth process X is positive. In fact, for an admissible strategy $\pi(u) := \bar{\pi}(u)X(u-)$ we can rewrite (1.6) as

$$dX(u) = X(u-)\,\bar{\pi}(u)\,dS(u). \tag{1.12}$$

This allows us to have for general $\bar{\pi}$ an explicit formula for X, since it takes the form of a stochastic exponential (cf. [51, Section 8.4]). By Itô's formula,

$$X(u) = x \cdot e^{\bar{\pi}(u)S(u) - \frac{1}{2}\int_t^u \sigma^2(v)\bar{\pi}^2(v)\,dv} \prod_{t<v\le u} (1 + \bar{\pi}(v)\Delta S(v))e^{-\Delta S(v)}, \quad \mathbb{P}\text{-a.s.}$$

Proposition 1.3.1 (Positivity of the portfolio value)**.** *If $\pi(u) = \bar{\pi}(u)X(u-)$, it holds $X^{t,s,x;\pi}(u) > 0$, $\mathbb{P}$-a.s., $\forall u \in [t,T]$, if and only if*

$$\bar{\pi}(u)\psi(u)y > -1 \quad \mathbb{P}\text{-}a.s., \ \nu\text{-}a.e. \ y \in \mathbb{R}, \ \textit{for all } u \in [t,T].$$

Proof. From (1.6), if the jump measure at time t of S, regarded as an additive process, is denoted by ν_t^S, then it holds that supp $\nu_t^S =$ supp ν. Then, $\{\bar{\pi}(u)\Delta S(u) > -1 \ \mathbb{P}\text{-a.s.}, \forall u \le T\}$ if and only if $\{\bar{\pi}(u)\Delta L(u) > -1 \ \mathbb{P}\text{-a.s.}, \forall u \le T\}$, which is equivalent to $\{\bar{\pi}(u)\psi(u)y > -1$ $\mathbb{P}$-a.s., ν-a.e. $y \in \mathbb{R}$, $\forall u \le T\}$. □

Therefore, the portfolio is positive for each strategy of the form $\pi(u) = \bar{\pi}(u)X(u-)$ such that $\bar{\pi}$ takes values in a suitably chosen set. This sums up in the following characterization of admissible controls.

Definition 1.3.2. Let $\Pi = \Pi_{\nu,\psi}$ be a compact set such that

$$\Pi_{\nu,\psi} \subset \widehat{\Pi}_{\nu,\psi} := \{\bar{\pi} \in \mathbb{R} \text{ s.t. } \bar{\pi}\psi y > -1 \text{ for each } y \in [m,M] \text{ and } \psi \in [\psi_1,\psi_2]\}.$$

A predictable process $\bar{\pi} : [t,T] \to \Pi$ is called normalized admissible strategy if there exists an admissible strategy $\pi \in \mathcal{A}([t,T])$ such that

$$\pi(u) = \bar{\pi}(u)X(u-)$$

for all $u \in [t,T]$, $\mathbb{P}$-a.s.

Remark 1.3.3. According to the support of the measure ν, the set $\widehat{\Pi} := \widehat{\Pi}_{\nu,\psi}$ consists of

case A $\widehat{\Pi} = (-\frac{1}{M\psi_2}, -\frac{1}{m\psi_2})$ if $m < 0$ and $M > 0$ (both positive and negative jumps),

case B $\widehat{\Pi} = (-\frac{1}{M\psi_2}, +\infty)$ if $0 \le m \le M$ and $M \neq 0$ (only positive jumps),

case C $\widehat{\Pi} = (-\infty, -\frac{1}{m\psi_2})$ if $m \le M \le 0$ and $m \neq 0$ (only negative jumps),

case D $\widehat{\Pi} = \mathbb{R}$ if $m = M = 0$ (no jumps),

by consistently interpreting where necessary: for instance if $M = +\infty$, $(-\frac{1}{M\psi_2}, +\infty) := [0, +\infty)$. Observe that in all cases we have $0 \in \widehat{\Pi}$. If $m = -\infty$ and $M = \infty$, then $\widehat{\Pi} = \{0\}$ which makes the problem trivial. Therefore, in order to get rid of this situation, we may assume from now on that at least one between m and M is finite.

Remark 1.3.4. The set $\widehat{\Pi} = \widehat{\Pi}_{\nu,\psi}$ is defined according to the jump features of the process L (cf. the analogous notion of *neutral constraints* in [75, Section 2]). On the other hand, we have a certain freedom in the definition of Π, as we only require that it is a compact subset of $\widehat{\Pi}$. Intuitively, we are restricting the range of possible trading strategies so that the instantaneous portfolio value can not jump to (or below) zero for any admissible (normalized) position $\bar{\pi}$.

In order to find a solution to the HJB equation, we make the following ansatz:

$$H(t, s, x) = U(x\, e^{g(t,s)}) = \log(x) + g(t, s).$$

This transform, which has been introduced in [57] for the specific case of Gaussian processes, is analogous to the one employed in [109], with the main difference due to the arithmetic nature of our spot price dynamics. We start from the static maximization problem, namely the maximization of the generalized Hamiltonian over all possible values of the strategies π. As usual in this approach (see the discussion in [69]) a candidate optimal policy $\pi \in \mathcal{A}[t, T]$ can be found by computing $\pi^*(t, s, x) = \arg\max_\pi A^\pi H(t, s, x)$ and defining $\pi^*(t) := \pi^*(t, S(t-), X(t-))$. It is common to refer to the deterministic function $\pi^*(t, s, x)$ as the optimal Markov control policy. Since we are in the case of logarithmic utility (cf. (1.11)), we can write $\pi^*(t, s, x) = \bar{\pi}^*(t, s) \cdot x$. Simple computations yield

$$\begin{aligned} A^\pi H(t, s, x) &= (b(t) - \lambda s)\left(\frac{\pi}{x} + g_s(t, s)\right) - \frac{1}{2}\,\sigma(t)^2\frac{\pi^2}{x^2} + \frac{1}{2}\sigma(t)^2 g_{ss}(t, s) \\ &\quad + \int_{\mathbb{R}}\Big[\log(x + \pi\psi(t)y) + g(t, s + y) - \log(x) - g(t, s) - \left(\frac{\pi\psi(t)}{x} + g_s(t, s)\right) y\Big]\nu(dy). \end{aligned} \tag{1.13}$$

Neglecting the terms which do not depend on π, we have

$$\begin{aligned} \arg\max_\pi A^\pi H(t, s, x) &= \arg\max_\pi\ (b(t) - \lambda s)\frac{\pi}{x} - \frac{1}{2}\,\sigma(t)^2\frac{\pi^2}{x^2} + \int_{\mathbb{R}}\left[\log\left(1 + \frac{\pi\psi(t)}{x}y\right) - \frac{\pi\psi(t)}{x}y\right]\,\nu(dy) \\ &= x \cdot \arg\max_{\bar{\pi}\in\Pi} f(\bar{\pi}; t, s), \end{aligned}$$

where the function $f : \Pi \times [0, T] \times \mathbb{R} \to \mathbb{R}$ is defined as

$$f(\bar{\pi}; t, s) := (b(t) - \lambda s)\bar{\pi} - \frac{1}{2}\,\sigma(t)^2\bar{\pi}^2 + \int_{\mathbb{R}}\left[\log\left(1 + \bar{\pi}\psi(t)y\right) - \bar{\pi}\psi(t)y\right]\,\nu(dy). \tag{1.14}$$

The expression to be maximized with respect to the variable $\bar{\pi}$ reads as the sum of three terms: a linear term, a strictly concave function and the integral of a strictly concave function. Therefore we are maximizing an overall strictly concave function on a compact set Π. This ensures the existence of a unique maximizer

$$\bar{\pi}^* = \bar{\pi}^*(t, s) := \arg\max_{\bar{\pi}\in\Pi} f(\bar{\pi}; t, s). \tag{1.15}$$

Remark 1.3.5. By adopting this notation we are revealing in advance that π^* corresponds to an optimal strategy, but we have not given a proof yet. The optimality of this candidate will be derived in Theorem 1.3.13 by applying the Verification Theorem of the previous section.

Recalling that $\pi^*(t,s,x) = \bar{\pi}^*(t,s) \cdot x$, we can write the HJB equation in reduced form

$$\frac{\partial}{\partial t} H(t,s,x) + A^{\pi^*(t,s,x)} H(t,s,x) = 0,$$

that is, consistently with our guess $H(t,s,x) = \log(x) + g(t,s)$,

$$\begin{aligned} &g_t(t,s) + (b(t) - \lambda s)\left(\bar{\pi}^*(t,s) + g_s(t,s)\right) - \frac{1}{2}\sigma(t)^2\bar{\pi}^*(t,s)^2 + \frac{1}{2}\sigma(t)^2 g_{ss}(t,s) \\ &\quad + \int_{\mathbb{R}} \Big[\log\left(1 + \bar{\pi}^*(t,s)\psi(t)y\right) - \bar{\pi}^*(t,s)\psi(t)y + g(t,s+\psi(t)y) - g(t,s) - g_s(t,s)\psi(t)y\Big]\, \nu(dy) = 0, \end{aligned}$$

After the terms with g are collected, the equation reads

$$\begin{aligned} g_t(t,s) &+ (b(t) - \lambda s) g_s(t,s) + \frac{1}{2}\sigma(t)^2 g_{ss}(t,s) \\ &+ \int_{\mathbb{R}} \left[g(t,s+\psi(t)y) - g(t,s) - g_s(t,s)\psi(t)y\right] \nu(dy) = -f^*(t,s), \end{aligned} \tag{1.16}$$

with terminal condition $g(T,s) = 0$, where we define $f^* : [0,T] \times \mathbb{R} \to \mathbb{R}$ by

$$f^*(t,s) := f(\bar{\pi}^*(t,s); t,s). \tag{1.17}$$

If we interpret (1.16) as an equation in the only unknown g, it takes the form of a linear parabolic partial integro-differential equation (PIDE). The analysis of such an equation is typically a delicate task and, to the best of our knowledge, there are not many existence results for *regular* solutions in literature for this class of problems (see [15, 53, 56, 108] and references therein). Under certain assumptions, we are able to prove the existence and a probabilistic representation formula: we will do this in Propositions 1.3.11 and 1.3.12. It is crucial noticing that in the logarithmic case we can solve the HJB equation directly by disentangling the problem of finding $\bar{\pi}^*(t,s)$ and the function $g(t,s)$. This has been verified by the authors not to be the case for a general CARA or CRRA utility, which makes the issue of solving the HJB equation more difficult as well as interesting (see also [1]). Nevertheless, an approximation of the HJB equation has been proposed in an analogous stochastic framework for CRRA utility in [8, 105].

In order to solve the PIDE, we first have to study the properties of the strategy defined implicitly in (1.15). A straightforward application of the dominated convergence theorem and the finiteness of the second moment of L assure that $f(\cdot; t,s)$ is differentiable for any $t \in [0,T]$ and $s \in \mathbb{R}$. Therefore, if the maximizer $\bar{\pi}^* = \bar{\pi}^*(t,s)$ is a internal point, it is the unique solution of the first order condition

$$f'(\bar{\pi}^*; t,s) = b(t) - \lambda s - \sigma(t)^2\bar{\pi}^* - \int_{\mathbb{R}} \frac{\bar{\pi}^*\psi(t)^2 y^2}{1 + \bar{\pi}^*\psi(t)y}\, \nu(dy) = 0. \tag{1.18}$$

We remark that this is the explicit deterministic counterpart of the third condition appearing in [74, Theorem 3.1].

In the two upcoming propositions, we sum up the properties of the candidate (normalized) optimal policy and of the function f^* appearing in the HJB equation.

Proposition 1.3.6. *Assume that* Π *is a compact interval containing* 0*. The static optimization problem in (1.14) and (1.15) admits a unique maximizer* $\bar{\pi}^* : [0,T]\times\mathbb{R}\to\Pi$ *with the following properties:*

1. *For each* $t\in[0,T]$*, it holds* $\bar{\pi}^*\left(t,\frac{b(t)}{\lambda}\right)=0$.
2. *The map* $\bar{\pi}^* : [0,T]\times\mathbb{R}\to\Pi$ *is continuous and then, in particular, measurable and bounded.*
3. *For each* $t\in[0,T]$*, there exists an open interval* $\Sigma(t)$ *such that the restrictions* $\bar{\pi}^*(t,\cdot)|_{\Sigma(t)}$ *are strictly decreasing and smooth, where*

   case A $\Sigma(t)=(s_1(t),s_2(t))$,
   case B $\Sigma(t)=(-\infty,s_2(t))$,
   case C $\Sigma(t)=(s_1(t),+\infty)$,
   case D $\Sigma(t)\equiv\mathbb{R}$.

   Moreover, for any $t\in[0,T]$ *the derivatives of* $\bar{\pi}^*(t,\cdot)|_{\Sigma(t)}$ *can be extended to* $\overline{\Sigma}(t)$.
4. *For each* $t\in[0,T]$*, the map* $\bar{\pi}^*(t,\cdot):\mathbb{R}\to\Pi$ *is decreasing on the whole real line and, in particular,*

   case A *there exist* s_1, s_2 *such that, for any* $t\in[0,T]$*, we have* $-\infty<s_1\le\frac{b(t)}{\lambda}\le s_2<\infty$ *and*

   $$\bar{\pi}^*(t,s)\equiv\begin{cases}\max\Pi, & \text{if } s\le s_1,\\ \min\Pi, & \text{if } s\ge s_2.\end{cases}$$

   case B *there exists* s_2 *such that, for any* $t\in[0,T]$*, we have* $\frac{b(t)}{\lambda}\le s_2<\infty$ *and* $\bar{\pi}^*(t,s)\equiv\min\Pi$ *for* $s\ge s_2$ *.*

   case C *there exists* s_1 *such that, for any* $t\in[0,T]$*, we have* $-\infty<s_1\le\frac{b(t)}{\lambda}$ *and* $\bar{\pi}^*(t,s)\equiv\max\Pi$ *for* $s\le s_1$*.*

   case D *we can write down the maximizer explicitly as*

   $$\bar{\pi}^*(t,s)=\frac{b(t)-\lambda s}{\sigma(t)^2},$$

   for each t *and* s *such that the above quantity is well-defined and belongs to* Π*.*
5. *In particular, for all* $t\in[0,T]$ *the maps* $\bar{\pi}^*(t,\cdot)$ *are Lipschitz continuous uniformly in* $t\in[0,T]$ *(i.e. with Lipschitz constant* L *independent of* t*).*

Proof. See Appendix B. □

Remark 1.3.7. In the notation of Proposition 1.3.6, for **case A,B,C**, we can write more explicitly

$$s_1(t)=\lim_{\bar{\pi}\to\bar{\pi}_2}s^*(t,\bar{\pi})=\frac{1}{\lambda}\left(b(t)-\sigma(t)^2\bar{\pi}_2-\int_{\mathbb{R}}\frac{\bar{\pi}_2\psi(t)^2y^2}{1+\bar{\pi}_2\psi(t)y}\,\nu(dy)\right),$$

$$s_2(t)=\lim_{\bar{\pi}\to\bar{\pi}_1}s^*(t,\bar{\pi})=\frac{1}{\lambda}\left(b(t)-\sigma(t)^2\bar{\pi}_1-\int_{\mathbb{R}}\frac{\bar{\pi}_1\psi(t)^2y^2}{1+\bar{\pi}_1\psi(t)y}\,\nu(dy)\right),$$

and

$$s_1=\min_{t\in[0,T]}s_1(t),\qquad s_2=\max_{t\in[0,T]}s_2(t).$$

Remark 1.3.8. In order to interpret the results of Proposition 1.3.6, let us suppose for example to be in **case A**. Recall that in this case we can have both upward and downward jumps in prices (see Remark 1.3.3) and that the normalized position $\bar{\pi}(t, s(t))$ can take also negative values, so that short-selling is allowed. At each time t, a trader who executes optimally takes a net zero position if price $s(t)$ reaches the (time-dependent) "equilibrium" level $b(t)/\lambda$. Furtherly, he goes long if price goes above this level and, accordingly, he goes short when the price is below. The trading allocation increases (with sign) as price decreases and vice versa. Also, s_1 (resp. s_2) consists of a lower (resp. upper) price threshold at which the trader, independently of the time instant, takes the longest (resp. shortest) position possible according to the trading constraints prescribed in Π.

Proposition 1.3.9. *The function $f^* : [0,T] \times \mathbb{R} \to \mathbb{R}$ in (1.17) is continuously differentiable and Lipschitz continuous, with partial derivatives*

$$\frac{\partial}{\partial t} f^*(t,s) = b'(t)\bar{\pi}^*(t,s) - \sigma(t)\sigma'(t)\bar{\pi}^*(t,s)^2 - \psi(t)\psi'(t)\bar{\pi}^*(t,s)^2 \int_{\mathbb{R}} \frac{y^2}{1+\bar{\pi}^*(t,s)\psi(t)y}\,\nu(dy),$$
$$\frac{\partial}{\partial s} f^*(t,s) = -\lambda\bar{\pi}^*(t,s).$$

Furthermore, it grows as a linear function of s uniformly in t, i.e.

$$|f^*(t,s)| \leq C(1+|s|),$$

being C dependent only on $\lambda, T, \|b\|_\infty, \|\sigma\|_\infty, \|\psi\|_\infty$.

Proof. Recall that by definition

$$f(\bar{\pi};t,s) = (b(t)-\lambda s)\bar{\pi} - \frac{1}{2}\,\sigma(t)^2\bar{\pi}^2 + \int_{\mathbb{R}} \left[\log\left(1+\bar{\pi}\psi(t)y\right) - \bar{\pi}\psi(t)y\right]\nu(dy),$$

which is a continuously differentiable function in the variable (t,s) for any $\bar{\pi} \in \Pi$, since b, σ and ψ are continuously differentiable. Then, by Danskin's theorem [54, Theorem 1],

$$f^*(t,s) = \max_{\bar{\pi}\in\Pi} f(\bar{\pi};t,s)$$

is differentiable with partial derivatives

$$\frac{\partial}{\partial t} f^*(t,s) = \frac{\partial}{\partial t} f(\bar{\pi};t,s)|_{\bar{\pi}=\bar{\pi}^*(t,s)},$$
$$\frac{\partial}{\partial s} f^*(t,s) = \frac{\partial}{\partial s} f(\bar{\pi};t,s)|_{\bar{\pi}=\bar{\pi}^*(t,s)}.$$

Since they are bounded continuous functions, it follows that $f^* \in C^1([0,T]\times\mathbb{R})$ and Lipschitz continuous. The linear bound is direct consequence of the definition of f and the boundedness of $\bar{\pi}^*(t,s)$. □

1.3.2 Probabilistic representation and existence of regular solutions

After studying the regularity properties of the forcing term of the reduced HJB equation (1.16), we move on to the problem of existence of solutions. First of all, we clarify the natural notion of classical solution for such a class of integro-differential equations. Tracing through [50, Section 17.4], we say that a function $g : [0,T] \times \mathbb{R} \to \mathbb{R}$ belongs to the set $C^{1,2}_{\nu,\psi} = C^{1,2}_{\nu,\psi}([0,T) \times \mathbb{R})$, if it is once continuously differentiable in its first argument and twice continuously differentiable in its second and, furtherly, the following integrability condition holds true: for every $t \in [0,T)$ and $s \in \mathbb{R}$,

$$\int_{\mathbb{R}} |g(t, s + \psi(t)y) - g(t,s) - g_s(t,s)\psi(t)y| \, \nu(dy) < \infty. \tag{1.19}$$

Then, a classical solution of the HJB equation is a function $g : [0,T] \times \mathbb{R} \to \mathbb{R}$ belonging to $C^{1,2}_{\nu,\psi}([0,T) \times \mathbb{R})$ and satisfying the integro-differential equation (1.16).

We now present three results. Firstly, we recall a version of the Feynman-Kač theorem, which gives the probabilistic representation of regular solutions. Then, we state two existence results for classical solutions: the first is valid for additive processes without diffusion part, while the second works for compound Poisson processes and uniformly non-degenerate Brownian component.

Theorem 1.3.10 (Feynman-Kač formula)**.** *Assume that g is a $C^{1,2}_{\nu,\psi}([0,T) \times \mathbb{R}) \cap C([0,T] \times \mathbb{R})$ solution of (1.16), satisfying the growth condition:*

$$\max_{t \in [0,T]} |g(t,s)| \leq K\,(1 + s^2), \qquad \text{for } s \in \mathbb{R}.$$

If, moreover, there exists $\varepsilon > 0$ such that

$$\int_{|y| \geq 1} |y|^{2+\varepsilon} \, \nu(dy) < \infty,$$

then we can represent g in the following Feynman-Kač type form

$$g(t,s) = \mathbb{E}\left[\int_t^T f^*(u, S^{t,s}(u))\, du\right]. \tag{1.20}$$

Proof. The proof is classical: see [50, Theorem 17.4.10]. □

In the upcoming proposition we prove the existence of a classical solution to (1.16) in the case that there is no Brownian component.

Assumption 1. *The diffusion component in (1.2) is identically zero, i.e. $\sigma \equiv 0$.*

We follow the idea of [20], where a guess is constructed via the Feynman-Kač formula. Let us remark that we generalize the result in [20] by proving the existence of a classical solution for time-inhomogeneus Lévy processes and possibly infinite variation square integrable Lévy measure. More in detail, we prove that

$$G(t,s) := \mathbb{E}\left[\int_t^T f^*(u, S^{t,s}(u))\, du\right] \tag{1.21}$$

is a well-defined regular function and solves the PIDE in the classical formulation. We need some preliminary propositions, which are collected in Appendix A.

Proposition 1.3.11 (Pure-jump case)**.** *Under Assumption 1, the function $G(t,s)$ is continuously differentiable in t for all $s \in \mathbb{R}$ and solves the following partial integro-differential equation:*

$$G_t(t,s) + (b(t) - \lambda s)G_s(t,s) + \int_{\mathbb{R}} [G(t, s + \psi(t)y) - G(t,s) - G_s(t,s)\psi(t)y] \, \nu(dy) = -f^*(t,s),$$

with terminal condition $G(T,s) = 0$. In particular, $G \in C^{1,1}_{\nu,\psi}([0,T) \times \mathbb{R}) \cap C([0,T] \times \mathbb{R})$. Furthermore, for all $t \in [0,T)$ and $s \in \mathbb{R}$ the following integrability condition holds:

$$\mathbb{E}\left[\int_t^T \int_{\mathbb{R}} \left[G(u, S^{t,s}(u-) + \psi(u)y) - G(u, S^{t,s}(u-))\right]^2 \nu(dy)\, du\right] < \infty.$$

Proof. Fix $t \in [0,T), h > 0$ and apply Itô's Lemma to $G(t+h, S(\cdot))$ from t to $t+h$. Then, we have

$$G(t+h,S(t+h)) = G(t+h,S(t)) + \int_t^{t+h} (b(u) - \lambda S(u))\partial_s G(t+h,S(u))\,du$$
$$+ \int_t^{t+h}\int_{\mathbb{R}} [G(t+h,S(u)+\psi(u)y) - G(t+h,S(u)) - \partial_s G(t+h,S(u))\psi(u)y]\nu(dy)du$$
$$+ \int_t^{t+h}\int_{\mathbb{R}} [G(t+h,S(u-)+\psi(u)y) - G(t+h,S(u-))]\overline{N}(dy,du).$$

Now, divide by h and take expectation $\mathbb{E}^{t,s}$. Fubini's theorem gives that

$$\begin{aligned}
&\frac{1}{h}\left(\mathbb{E}^{t,s}[G(t+h,S(t+h))] - G(t+h,s)\right) \qquad (1.22)\\
&= \frac{1}{h}\int_t^{t+h} \mathbb{E}^{t,s}[(b(u) - \lambda S(u))\partial_s G(t+h,S(u))]\,du\\
&+ \frac{1}{h}\int_t^{t+h}\int_{\mathbb{R}} \mathbb{E}^{t,s}[G(t+h,S(u)+\psi(u)y) - G(t+h,S(u)) - \partial_s G(t+h,S(u))\psi(u)y]\,\nu(dy)du\\
&+ \frac{1}{h}\mathbb{E}^{t,s}\left[\int_t^{t+h}\int_{\mathbb{R}} \big(G(t+h,S(u-)+\psi(u)y) - G(t+h,S(u-))\big)\overline{N}(dy,du)\right].
\end{aligned}$$

By the mean value theorem, since the map $u \mapsto \mathbb{E}^{t,s}[(b(u) - \lambda S(u))\partial_s G(t+h,S(u))]$ is continuous, we have for a $u_h \in [t,t+h]$ that

$$\frac{1}{h}\Big(\int_t^{t+h} \mathbb{E}^{t,s}[(b(u) - \lambda S(u))\partial_s G(t+h,S(u))]\,du\Big) = \mathbb{E}^{t,s}[(b(u_h) - \lambda S(u_h))\partial_s G(t+h,S(u_h))], \qquad (1.23)$$

which converges to $(b(t) - \lambda s)G_s(t,s)$ as h approaches 0. Analogously, for the second term it holds that

$$\begin{aligned}
&\frac{1}{h}\int_t^{t+h}\int_{\mathbb{R}} \mathbb{E}^{t,s}[G(t+h,S(u)+\psi(u)y) - G(t+h,S(u)) - \partial_s G(t+h,S(u))\psi(u)y]\nu(dy)du\\
&= \int_{\mathbb{R}} \mathbb{E}^{t,s}[G(t+h,S(u_h)+\psi(u_h)y) - G(t+h,S(u_h)) - \partial_s G(t+h,S(u_h))\psi(u_h)y]\nu(dy),
\end{aligned}$$

for a $u_h \in [t,t+h]$. Since all the maps in the expectation are continuous, as h tends to zero (cf. Lemma 1.6.4), the last term converges to

$$\int_{\mathbb{R}} [G(t,s+\psi(t)y) - G(t,s) - G_s(t,s)\psi(t)y]\ \nu(dy).$$

Moreover,

$$\mathbb{E}^{t,s}\left[\int_t^{t+h}\int_{\mathbb{R}} \big(G(t+h,S(u-)+\psi(u)y) - G(t+h,S(u-))\big)\overline{N}(dy,du)\right] = 0.$$

Finally, the left-hand side can be written as

$$\frac{1}{h}\left(\mathbb{E}^{t,s}[G(t+h,S(t+h))] - G(t,s)\right) + \frac{1}{h}\left(G(t,s) - G(t+h,s)\right).$$

By the Markov property and the tower rule,

$$\begin{aligned}
\mathbb{E}^{t,s}[G(t+h, S(t+h))] &= \mathbb{E}\left[\mathbb{E}\left[\int_{t+h}^{T} f^*(u, S^{t+h, S^{t,s}(t+h)}(u))\, du\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\int_{t+h}^{T} f^*(u, S^{t,s}(u))\, du \middle| \mathcal{F}_{t+h}\right]\right] \\
&= \mathbb{E}\left[\int_{t+h}^{T} f^*(u, S^{t,s}(u))\, du\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\frac{1}{h}\left(\mathbb{E}^{t,s}[G(t+h, S(t+h))] - G(t,s)\right) \\
&\quad = \frac{1}{h}\left(\mathbb{E}\left[\int_{t+h}^{T} f^*(u, S^{t,s}(u))\, du\right] - \mathbb{E}\left[\int_{t}^{T} f^*(u, S^{t,s}(u))\, du\right]\right) \\
&\quad = -\frac{1}{h}\mathbb{E}\left[\int_{t}^{t+h} f^*(u, S^{t,s}(u))\, du\right],
\end{aligned}$$

which converges to $-f^*(t,s)$ as h goes to zero. Then, we have found that the limit of $-\frac{1}{h}\left(G(t+h,s) - G(t,s)\right)$ exists and is equal to

$$(b(t) - \lambda s)G_s(t,s) + \int_{\mathbb{R}} \left[G(t, s+\psi(t)y) - G(t,s) - G_s(t,s)\psi(t)y\right] \nu(dy) + f^*(t,s),$$

so that $G_t(t,s)$ exists and it is continuous, being the right-hand term continuous. Also, we get from this expression that G solves the integro-differential equation of the statement.

For the last point, as in Lemma 1.6.4, it is sufficient observe that

$$\mathbb{E}\left[\left(G(u, S^{t,s}(u) + \psi(u)y) - G(u, S^{t,s}(u))\right)^2\right] \leq \sup_{z\in\mathbb{R}} G_s(u,z)^2\psi(u)^2y^2 \leq Ce^{2\lambda u}y^2,$$

since G is Lipschitz continuous in z uniformly in u. □

In the last proposition, a result by [108] is applied to prove existence and uniqueness in the case that the second-order operator is uniformly elliptic and the jump part of L is a compound Poisson process.

Assumption 2. *Assume in (1.2) that $\sigma(t) > 0$ for all $t \in [0,T]$ and ν is a finite Lévy measure.*

Proposition 1.3.12 (Finite Lévy measure)**.** *Under Assumption 2, the function*

$$G(t,s) := \mathbb{E}\left[\int_t^T f^*(u, S^{t,s}(u))\, du\right]$$

is the unique $C^{1,2}_{\nu,\psi}([0,T)\times\mathbb{R}) \cap C([0,T]\times\mathbb{R})$ solution of (1.16). Moreover, the following integrability conditions hold:

$$\begin{aligned}
&\mathbb{E}\left[\int_t^T \int_{\mathbb{R}} \left[G(u, S^{t,s}(u-) + \psi(u)y) - G(u, S^{t,s}(u-))\right]^2 \nu(dy)\, du\right] < \infty, \\
&\mathbb{E}\left[\int_t^T \sigma(u)^2\, G_s(u, S^{t,s}(u))^2\, du\right] < \infty.
\end{aligned}$$

Proof. First of all, observe that, since ν is the Lévy measure associated to a compound Poisson process, the spaces $C^{1,2}_{\nu,\psi}([0,T)\times\mathbb{R})$ and $C^{1,2}([0,T)\times\mathbb{R})$ coincide (cf. [50, Definition 17.4.9]). Therefore, we only need to verify if the assumptions of [108, Proposition 5.3] are fulfilled. Notice that (H6) there corresponds to assuming that the Lévy jump component is a compound Poisson process. Then, in order to apply [108, Proposition 5.3] it remains to prove that $f^*:[0,T]\times\mathbb{R}\to\mathbb{R}$ is Lipschitz continuous, which follows from Proposition 1.3.9. The integrability conditions can be proved as in Lemma 1.6.3. □

Finally, we apply the Verification Theorem and state the main results of this section.

Theorem 1.3.13. *Let g be a $C^{1,2}_{\nu,\psi}([0,T)\times\mathbb{R})\cap C([0,T]\times\mathbb{R})$ solution of (1.16) and assume that, for any $t\in[0,T)$ and $s\in\mathbb{R}$, we have the following conditions*

$$\mathbb{E}\left[\int_t^T\int_{\mathbb{R}}\left[g(u,S^{t,s}(u-)+\psi(u)y)-g(u,S^{t,s}(u-))\right]^2\nu(dy)\,du\right]<\infty,$$
$$\mathbb{E}\left[\int_t^T\sigma(u)^2\,g_s(u,S^{t,s}(u))^2\,du\right]<\infty,$$

where $S=S^{t,s}$ is the solution of

$$\begin{aligned} dS(u) &= -\lambda S(u)\,du+dL(u), && u\in(t,T],\\ S(t) &= s. \end{aligned}$$

Then, the function $\pi^*(t,s,x):=\bar{\pi}^*(t,s)\cdot x$, with $\bar{\pi}^*$ as in Proposition 1.3.6, is an optimal Markov control policy, i.e. it induces an admissible strategy in the sense of Definition 1.2.2 and, for each $t\in[0,T), s\in\mathbb{R}, x\in\mathbb{R}^+$, we get that $J(t,s,x;\pi^*)=V(t,s,x)=\log(x)+g(t,s)$.

Proof. See Appendix C. □

Corollary 1.3.14. *Assume either Assumption 1, or Assumption 2 and define*

$$G(t,s)=\mathbb{E}\left[\int_t^T f^*(u,S^{t,s}(u))\,du\right].$$

Then, $\pi^*(t,s,x):=\bar{\pi}^*(t,s)\cdot x$, as in Proposition 1.3.6, is an optimal Markov control policy and $J(t,s,x;\pi^*)=V(t,s,x)=\log(x)+G(t,s)$.

1.4 Estimating the optimal strategy: the Merton ratio and Taylor approximations

After proving the existence and describing the analytical properties of the optimal strategy, we now study simple ways to compute it by approximation.

1.4.1 Definition and intuition

In his seminal work on portfolio selection [96], Merton studies the optimal allocation of the investor's wealth when the risky asset follows a geometric Brownian motion:

$$dS(t)=\mu S(t)\,dt+\sigma S(t)\,dW(t),\qquad t\in(0,T],$$

and finds that the optimal allocation for a log-utility is[1]

$$\pi^*_M = \frac{\mu}{\sigma^2},$$

which consists of the ratio of the excess return over the local variance of the log-price. Instead, in our framework the price dynamics are

$$dS(t) = (b(t) - \lambda S(t))\,dt + \sigma(t)\,dW(t) + \psi(t)\int_{\mathbb{R}} y\,\overline{N}(dy, dt), \qquad t \in (0, T].$$

Here, the local variance at time t is the sum of the variance of the continuous component $\sigma(t)^2$ and that of the jump part $\sigma_L(t)^2 := \psi(t)^2 \int_{\mathbb{R}} y^2\,\nu(dy)$. Then, in this context, it is natural to define the analogue of Merton's Ratio π^*_M as

$$\bar{\pi}^*_1(t, s) := \frac{b(t) - \lambda s}{\sigma(t)^2 + \sigma_L(t)^2}. \tag{1.24}$$

This ratio appears naturally when applying a Taylor approximation to (1.18). Recall that the optimal normalized strategy $\bar{\pi}^*$ is defined as

$$\bar{\pi}^*(t, s) = \arg\max_{\bar{\pi}\in\Pi} f(\bar{\pi}; t, s),$$

where

$$f(\bar{\pi}; t, s) = (b(t) - \lambda s)\bar{\pi} - \frac{1}{2}\,\sigma(t)^2\bar{\pi}^2 + \int_{\mathbb{R}} \left[\log\left(1 + \bar{\pi}\psi(t)y\right) - \bar{\pi}\psi(t)y\right]\,\nu(dy).$$

If the maximum is attained at an interior point (cf. Proposition 1.3.6), then $\bar{\pi}^*$ satisfies the integral equation

$$b(t) - \lambda s - \sigma(t)^2\bar{\pi} - \int_{\mathbb{R}} \frac{\bar{\pi}\psi(t)^2 y^2}{1 + \bar{\pi}\psi(t)y}\,\nu(dy) = 0. \tag{1.25}$$

If we replace the integrand by the second-order Taylor expansion around zero, the integral equation becomes

$$b(t) - \lambda s - \sigma(t)^2\bar{\pi} - \bar{\pi}\psi(t)^2 \int_{\mathbb{R}} y^2\,\nu(dy) = 0, \tag{1.26}$$

whose unique solution is exactly the strategy *à la* Merton $\bar{\pi}^*_1(t, s)$ that we defined in (1.24). A similar Taylor truncation has been introduced in [28] to study approximations of Lévy processes and tested numerically in [107]. Also, [8, 105] arrive at an analogous approximated strategy for stochastic volatility models with jumps, however they start by approximating the HJB directly. The idea is to treat the small jumps as an additional Brownian component (very much in the spirit of [9]) and neglect larger jumps.

Let us now introduce a more accurate approximation in the finite activity case. We assume that $\nu([m, M] \setminus \{0\}) < \infty$, i.e. the jump component of (1.4) is, in fact, a compound Poisson process, which is often the most interesting case for application purposes. Therefore, the Lévy measure takes the form $\nu(dy) = \eta\,F(dy)$, where η is the jump intensity and $F(dy)$ the jump size distribution. Thanks to our standing assumptions, the distribution F admits finite expectation μ_F and variance σ_F^2. We remind that the optimal policy is defined through the first order condition in (1.25). Let us write the Taylor polynomial of the integrand around

[1] Consistently with our setting, we are assuming that the risk-free interest rate r is zero and there is no consumption during the trading period.

an arbitrary (finite) point $y_0 \in [m, M] \setminus \{0\}$. So, we set $\phi(y) := \frac{\bar{\pi} y^2}{1+\bar{\pi}\psi(t)y}$ and compute its derivatives. Writing down its expansion up to the first order, we have

$$\phi(y) = \frac{\bar{\pi} y_0^2}{1 + \bar{\pi}\psi(t)y_0} + \frac{2\bar{\pi} y_0 + \bar{\pi}^2 y_0^2 \psi(t)}{(1 + \bar{\pi}\psi(t)y_0)^2}(y - y_0) + o(y - y_0). \tag{1.27}$$

Then, (1.25) becomes

$$b(t) - \lambda s - \sigma(t)^2 \bar{\pi} - \frac{\bar{\pi}\psi(t)^2 y_0^2}{1 + \bar{\pi}\psi(t)y_0}\nu([m, M] \setminus \{0\}) - \frac{2\bar{\pi}\psi(t)^2 y_0 + \bar{\pi}^2\psi(t)^3 y_0^2}{(1 + \bar{\pi}\psi(t)y_0)^2}\int_{\mathbb{R}}(y - y_0)\,\nu(dy) = 0. \tag{1.28}$$

From this expression it is clear that a significant simplification is given by the choice

$$y_0 := \frac{1}{\nu([m, M] \setminus \{0\})}\int_{\mathbb{R}} y\,\nu(dy),$$

since in this case the first-order term just disappears. Besides, by writing the Lévy measure with respect to F, we see that y_0 corresponds to the mean value of the jump size:

$$y_0 = \frac{1}{\eta \int_{\mathbb{R}} F(dy)}\eta \int_{\mathbb{R}} y\,F(dy) = \mu_F. \tag{1.29}$$

We are essentially replacing the integrand with its linear approximation around the integral mean, or, from another point of view, we are approximating a function of the jumps with respect to the jump size mean value. Since here we take into account the jump measure specifications, this is a slightly different approach from the first approximation $\bar{\pi}_1^*$ (where the Taylor polynomial was centered in zero). Hence, from (1.28) we have the following approximated equation for $\bar{\pi}$:

$$-\sigma(t)^2\psi(t)\mu_F\bar{\pi}^2 + (\mu_F\psi(t)(b(t) - \lambda s) - \sigma(t)^2 - \mu_F^2\psi(t)^2\eta)\bar{\pi} + b(t) - \lambda s = 0. \tag{1.30}$$

In the case $\sigma(t)\psi(t) \neq 0$, it is a second order polynomial in the variable $\bar{\pi}$. Consequently, for each $t \in [0, T]$ we have two (generally) different solutions. However, only one of them is admissible, meaning that $\bar{\pi}_2^*(t, s) \in \Pi$ for every possible value of s. Since the ambiguity comes from the definition domain of the logarithm, we just need to impose the condition $1 + \bar{\pi}_2^*\psi(t)\mu_F > 0$. This leads to

$$\bar{\pi}_2^*(t, s) := \begin{cases} \frac{p_1(t,s) + \sqrt{p_1(t,s)^2 + 4p_2(t,s)}}{2p_3(t,s)}, & \text{if } \mu_F > 0, \\ \frac{p_1(t,s) - \sqrt{p_1(t,s)^2 + 4p_2(t,s)}}{2p_3(t,s)}, & \text{if } \mu_F < 0, \end{cases}$$

where

$$\begin{aligned} p_1(t, s) &= \mu_F\psi(t)(b(t) - \lambda s) - \mu_F^2\psi(t)^2\eta - \sigma(t)^2, \\ p_2(t, s) &= \mu_F(b(t) - \lambda s)\sigma(t)^2\psi(t), \\ p_3(t, s) &= \mu_F\psi(t)\sigma^2. \end{aligned}$$

On the other hand, in a pure jump context ($\sigma(t) \equiv 0$), we get simply

$$\bar{\pi}_2^*(t, s) = -\frac{b(t) - \lambda s}{\psi(t)\mu_F\big(b(t) - \lambda s - \eta\psi(t)\mu_F\big)}, \tag{1.31}$$

for any $\mu_F \neq 0$ and for each t and s such that $\bar{\pi}_2^*(t, s)$ is well defined and takes values into Π.

1.4.2 Error bounds

To estimate the approximation error, we compute the difference between the optimal normalized strategy $\bar{\pi}^*$ and the approximated ones $\bar{\pi}_1^*$ and $\bar{\pi}_2^*$.

Proposition 1.4.1. *Assume that* Π *contains* 0 *and the finiteness of the third moment of the jumps, that is*

$$\int_{|y|\geq 1} |y|^3 \, \nu(dy) < \infty,$$

and denote $\sigma_1^2 = \min_{[0,T]} \sigma^2(t), \sigma_\nu^2 = \int_{\mathbb{R}} y^2 \, \nu(dy)$. *If we are not in* **case D** *(no jumps) of Remark 1.3.3, then for each* $t \in [0,T]$ *it holds that*

$$|\bar{\pi}^*(t,\cdot) - \bar{\pi}_1^*(t,\cdot)| \leq C \int_{\mathbb{R}} |y|^3 \, \nu(dy),$$

where C *is a constant which depends on* $\delta :=\mathrm{dist}(\Pi, \partial\widehat{\Pi})$, $\max \Pi$, $\min \Pi$, m, M, ψ_2, σ_1^2 *and* σ_ν^2 *according to the following different cases:*

case A

$$C = \begin{cases} \frac{C_0}{\min\{1, \delta\psi_2 M, -\delta\psi_2 m\}} & \text{if } m \neq -\infty, M \neq +\infty, \\ \frac{C_0}{\min\{1, \delta\psi_2 M\}} & \text{if } m = -\infty, M \neq +\infty, \\ \frac{C_0}{\min\{1, -\delta\psi_2 m\}} & \text{if } m \neq -\infty, M = +\infty. \end{cases}$$

case B

$$C = \begin{cases} \frac{C_0}{\min\{1, \delta\psi_2 M\}} & \text{if } M \neq +\infty, \\ C_0 & \text{if } M = +\infty. \end{cases}$$

case C

$$C = \begin{cases} \frac{C_0}{\min\{1, -\delta\psi_2 m\}} & \text{if } m \neq -\infty, \\ C_0 & \text{if } m = -\infty, \end{cases}$$

where $C_0 := \frac{\psi_2^3}{\sigma_1^2 + \psi_2^2 \sigma_\nu^2} \max_{\pi \in \Pi} \pi^2$.

Proposition 1.4.2. *Let us assume that we are not in* **case D** *of Remark 1.3.3. Moreover, we suppose* $0 \in \Pi$ *and that* $\sigma(t)^2$ *and* μ_F *are not both identically 0. Then for each* $t \in [0,T]$ *it holds that*

$$|\bar{\pi}^*(t,\cdot) - \bar{\pi}_2^*(t,\cdot)| \leq \frac{\eta \psi_2^2 \, C_1 \, \sigma_F}{\sigma_1^2 + \eta \, \psi_2^2 \, C_2 \, \mu_F^2},$$

where $\sigma_1^2 = \min_{[0,T]} \sigma^2(t)$, σ_F *is the square root of the variance of the random jump size and* C_1, C_2 *are constants depending on* $\delta :=$ *dist*$(\Pi, \partial\widehat{\Pi})$, $\max \Pi$, $\min \Pi$, m, M, ψ_2 *according to the following different cases:*

case A

$$C_1 = \begin{cases} \max\{1, \frac{1}{(\delta\psi_2 M)^2}, \frac{1}{(\delta\psi_2 m)^2}\} + \max\{1, \frac{1}{\delta\psi_2 M}, \frac{1}{-\delta\psi_2 m}\}, & \text{if } m \neq -\infty, M \neq +\infty, \\ \max\{1, \frac{1}{(\delta\psi_2 M)^2}\} + \max\{1, \frac{1}{\delta\psi_2 M}\}, & \text{if } m = -\infty, M \neq +\infty, \\ \max\{1, \frac{1}{(\delta\psi_2 m)^2}\} + \max\{1, \frac{1}{-\delta\psi_2 m}\}, & \text{if } m \neq -\infty, M = +\infty. \end{cases}$$

$$C_2 = \begin{cases} \frac{1}{(1 + \max \Pi \, \psi_2 \, \mu_F)^2} & \text{if } \mu_F > 0, \\ \frac{1}{(1 + \min \Pi \, \psi_1 \, \mu_F)^2} & \text{if } \mu_F < 0. \end{cases}$$

case B

$$C_1 = \begin{cases} \max\{1, \frac{1}{(\delta\psi_2 M)^2}\} + \max\{1, \frac{1}{\delta\psi_2 M}\}, & \textit{if } M \neq +\infty, \\ 1 & \textit{if } M = +\infty. \end{cases}$$

$$C_2 = \frac{1}{(1 + \max \Pi\, \psi_2\, \mu_F)^2}.$$

case C

$$C_1 = \begin{cases} \max\{1, \frac{1}{(\delta\psi_2 m)^2}\} + \max\{1, \frac{1}{-\delta\psi_2 m}\}, & \textit{if } m \neq -\infty, \\ 1 & \textit{if } m = -\infty. \end{cases}$$

$$C_2 = \frac{1}{(1 + \min \Pi\, \psi_1\, \mu_F)^2}.$$

1.5 Numerical example

In this section we test our trading strategies on one of the most popular electricity price models, namely the factor model in [21]. There, the authors conduct a critical comparison of three different spot price models for electricity in the context of day-ahead markets. In fact, this is typically an auction market in which the electricity price is fixed for the subsequent day, so that daily averaged prices are taken into account, over a timeline of years. Consequently, our setting concerns different market and price definitions. To recall, we instead take the point of view of an agent in the intraday market, which is the exchange where the electricity is traded continuously for 8-27 hours (depending on the contract) generally in the form of quarterly or hourly forward contracts. However, we aim to exploit the analysis in [21], where the model is calibrated to Nord Pool Spot market data, for mainly two reasons. Firstly, the stylized features of intraday markets are of similar nature as the ones observed in the day-ahead price series, such as spike behavior, high volatility, leptokurtosis (for a more detailed empirical study see, for instance, [57, 88]). Secondly, the factor model is based on a Lévy Ornstein-Uhlenbeck process of the same family as the one in (1.4).

We consider the factor model as in [21], which was originally introduced for electricity price modelling in [18]. The price dynamics are written as

$$S(t) = e^{Q(t)} Z(t),$$

where

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

is the deseasonalized price and $Q(t)$ is the seasonal component. The w_i are positive weights while the factors Y_i $(i = 1, ..., n)$ are independent non-Gaussian Ornstein-Uhlenbeck processes described by

$$dY_i(t) = -\lambda_i\, Y_i(t)\, dt + dL_i(t), \qquad Y_i(0) = y_i, \qquad i = 1, ..., n,$$

being L_i independent *càdlàg* pure jump additive processes with increasing paths.

As the authors calibrate the model in [21], by comparing the theoretical autocorrelation function to the empirical one, they set the optimal number of factors to $n = 2$. The estimated speeds of mean reversion are $\lambda_1 = 0.0087$ and $\lambda_2 = 0.3333$. In the paper these two values are interpreted as, respectively, the *base* (slowest) and the *spike* (fastest) signal. We start from here to define our equations. Specifically, their data series ranges from 13/07/2000 to

7/08/2008, which comprises, excluding the weekends, 2099 days. The time unit for t is 1 day. So, in order to adapt it to our timeline, which covers hours of intraday transactions, first we set one hour as our time unit, that is we do the time variable change $u = 24 \cdot t$. Then, denoting $C := 24$, we set our mean reversion speed λ in our own model by rescaling in time the spike speed ($\lambda_2 = 0.3333$), i.e. we take $\lambda = \lambda_2/C = 0.0139$. The driving process L is a compound Poisson process where the jump intensity, originally adopted in the paper by [73], is seasonally dependent. Also, the jump size distribution is a Pareto(α,z_0), with $\alpha = 2.5406$, $z_0 = 0.3648$ and density function $f(y) = \frac{\alpha z_0^\alpha}{y^{\alpha+1}}$. Therefore, we are in the case of positive jumps: with the notation of previous sections, $\mathrm{supp}(\nu) = [z_0, +\infty)$, i.e. $m = z_0$, $M = +\infty$, $\widehat{\Pi} = [0, +\infty)$ and $F(dy) = f(y)dy$. The set Π can be any compact subset of $\widehat{\Pi}$ containing 0.

For our own problem to make sense, another issue to address is to deseasonalize the jump intensity. In details, the form of the intensity is the following

$$e(t) = \theta \cdot s(t) = \theta \cdot \left(\frac{2}{1 + |\sin(\pi \frac{t-\tau}{k})|} - 1 \right)^d$$

where $\theta = 14.0163$ represents the expected number of spikes per time unit at a spike-clustering time, whereas the seasonal parameters are set by the authors' calibration procedure $k = 0.5$, $\tau = 0.42$, $d = 1.0359$. We then decide to compute the integral mean of $e(t)$ over its time periodicity, that is $2k$, obtaining $\mu := 3.7249$, so that, after rescaling, we have our intensity $\eta := \mu/C = 0.1552$.

To summarize, the electricity price in our hourly intraday market is described by

$$dS(t) = (b(t) - \lambda S(t))\, dt + dL(t), \tag{1.32}$$

where $\lambda = \lambda_2/C = 0.0139$ is the mean reversion speed and L is a compound Poisson process with jump intensity $\eta = \mu/C = 0.1552$ and jump size distribution a Pareto law of parameters $\alpha = 2.5406$ and $z_0 = 0.3648$. Therefore, there is no Brownian component in the jumps and the coefficient of the jump volatility is normalized to 1 ($\sigma^2 \equiv 0$ and $\psi \equiv 1$). In particular it is important to notice that L is a subordinator, which keeps the price positive. The drift value b cannot be derived directly from [21], being not part of the spike signal and for this reason it will be discussed later.

Let us write the equation for the exact normalized strategy $\bar{\pi}^* = \bar{\pi}^*(t, s)$, defined in the integral equation (1.18), i.e.

$$(b(t) - \lambda s) - \eta \int_{\mathbb{R}} \frac{\bar{\pi}^* y^2}{1 + \bar{\pi}^* y} f(y)\, dy = 0,$$

or, expressing it in terms of the price level s,

$$s = \frac{b(t)}{\lambda} - \frac{\eta}{\lambda} \cdot \int_{z_0}^{\infty} \frac{\bar{\pi}^* y^2}{1 + \bar{\pi}^* y} f(y)\, dy, \tag{1.33}$$

where $f(y) = \frac{\alpha z_0^\alpha}{y^{\alpha+1}}$ is the density of a Pareto law and η the jump intensity of L. By using a software to integrate exactly the above expression (we used *Mathematica*™) and inserting the estimated parameters inside the integral, we get

$$s = \frac{b(t)}{0.0139} - 6.7233 \cdot {}_2F_1 \left(1, 1.5406; 2.5406; -\frac{2.7412}{\bar{\pi}^*} \right),$$

where ${}_2F_1(a, b; c; z)$ is the hypergeometric function. This explicit formula states the value of the price s with respect to the optimal strategy $\bar{\pi}^*$. The inverse relation $s \mapsto \bar{\pi}^*$ can be computed numerically. As already observed, if we are interested in plotting the functions above, we need to set a value for the drift $b(t)$, being not consistently computable from the analysis in [21]. A quick study of the last expression yields the following.

Proposition 1.5.1. *Let us recall the definition of the jump measure mean*

$$\mu_L = \int_{\mathbb{R}} y\,\nu(dy) = \eta \int_{\mathbb{R}} y\,f(y)\,dy = \eta\,\mu_F,$$

where μ_F is the mean of the jump size distribution $f(y)dy$. For each $t \in [0,T]$, if the drift $b(t)$ in the price equation (1.32) is nonpositive, then $\bar{\pi}^*(t,s) \equiv 0$ for any $s \in \mathbb{R}^+$. Furthermore, if the drift $b(t)$ is greater than or equal to μ_L for $t \in [0,T]$, then $\bar{\pi}^*(t,s) \equiv \max \Pi$ for any $s \le s^*(t, \max \Pi)$, where $s^*(t,\cdot)$ denotes the inverse function of $\bar{\pi}^*(t,\cdot)$ (within its range of invertibility).

Proof. On one hand, the admissible values of the strategy belong to Π, which is any compact subset of $\widehat{\Pi} = [0,+\infty)$ containing 0. On the other hand, the price s can take only positive values by construction (recall that L is a subordinator). In (1.33) the function $\bar{\pi}^* \mapsto \eta \int_{z_0}^{\infty} \frac{\bar{\pi}^* y^2}{1+\bar{\pi}^* y} f(y)\,dy$ is increasing on the positive real line and it holds

$$\lim_{\bar{\pi}^* \to 0} \eta \int_{z_0}^{\infty} \frac{\bar{\pi}^* y^2}{1+\bar{\pi}^* y} f(y)\,dy = 0,$$
$$\lim_{\bar{\pi}^* \to +\infty} \eta \int_{z_0}^{\infty} \frac{\bar{\pi}^* y^2}{1+\bar{\pi}^* y} f(y)\,dy = \mu_L > 0.$$

Therefore, the first order condition in (1.33) is not satisfied for admissible values of s and $\bar{\pi}^*$ whenever $b(t) \le 0$ since

$$s = \frac{1}{\lambda}\left(b(t) - \eta \cdot \int_{z_0}^{\infty} \frac{\bar{\pi}^* y^2}{1+\bar{\pi}^* y} f(y)\,dy\right) < 0.$$

This means that the maximum in (1.15) is attained at the boundary of Π and more exactly when $\bar{\pi}^* = \min \Pi = 0$ (cf. Proposition 1.3.6). The case $b(t) \ge \mu_L$ can be proven along exactly the same reasonings. $\square$

Now, we write from (1.24) the first approximated strategy:

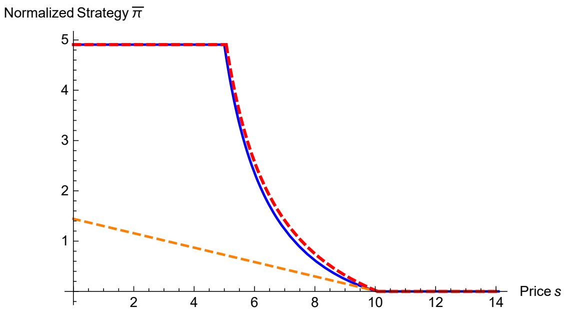
$$\bar{\pi}_1^*(t,s) = \begin{cases} \frac{b(t)-\lambda s}{\sigma_L^2}, & \text{if } \bar{\pi}_1^*(t,s) \in \Pi, \\ 0, & \text{if } \bar{\pi}_1^*(t,s) \notin \Pi, \end{cases}$$

A straightforward computation yields $\sigma_L^2 = \int_m^M y^2\,\nu(dy) = \eta \cdot \int_{z_0}^{\infty} y^2 f(y)\,dy = \eta \cdot 0.6254 = 0.0971$. Observe that, in order that condition $\bar{\pi}_1^*(t,s) \in \Pi$ holds, $b(t)$ must be non-negative. Having chosen any Π compact subset of $\widehat{\Pi} = [0,+\infty)$ containing $[0, \frac{b(t)}{\sigma_L^2}]$ for any $t \in [0,T]$, this reads

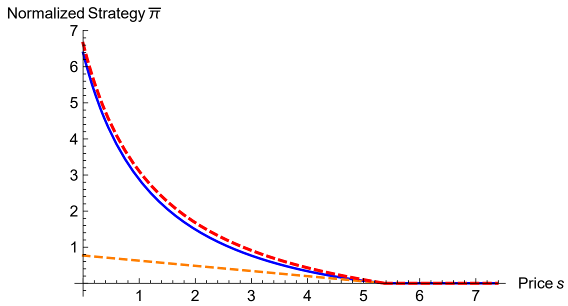
$$\bar{\pi}_1^*(t,s) = \begin{cases} \frac{\lambda}{\sigma_L^2} \cdot \left(\frac{b(t)}{\lambda} - s\right), & \text{if } 0 \le s \le \frac{b(t)}{\lambda}, \\ 0, & \text{if } s > \frac{b(t)}{\lambda}. \end{cases}$$

Finally, from (1.31) we get the second approximation of the optimal strategy:

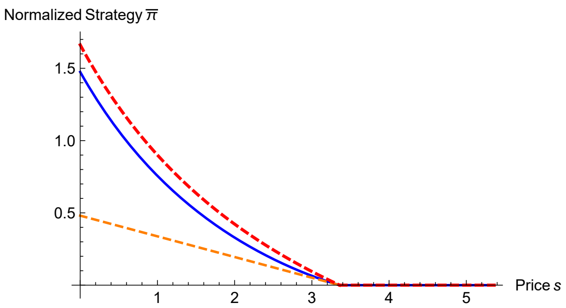
$$\bar{\pi}_2^*(t,s) = \begin{cases} -\frac{\eta}{\mu_L} \frac{b(t)-\lambda s}{b(t)-\lambda s-\mu_L}, & \text{if } 0 \le s \le \frac{b(t)}{\lambda}, \\ 0, & \text{if } s > \frac{b(t)}{\lambda}, \end{cases}$$



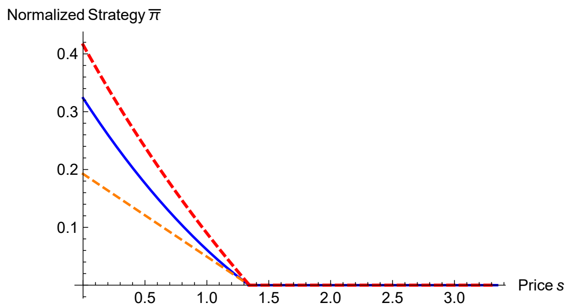
(a) $b = 150\%$ of μ_L.



(b) $b = 80\%$ of μ_L.



(c) $b = 50\%$ of μ_L.



(d) $b = 20\%$ of μ_L.

Figure 1.1: Blue line: (exact value) $\bar{\pi}^*$, orange line: $\bar{\pi}_1^*$, red line: $\bar{\pi}_2^*$.

still with the condition that $\bar{\pi}_2^*(t,s) \in \Pi$ (cf. Proposition 1.5.1).

We now plot the different strategies: the exact and the two approximations. In view of Proposition 1.5.1 we do it for the following values of the drift $b = 150, 80, 50, 20\%$ of μ_L. In such a way we can understand their behavior in the most representative cases (see Figure 1.1). Remind that if $s \geq \frac{b}{\lambda}$, then $\bar{\pi}^*$, $\bar{\pi}_1^*$ and $\bar{\pi}_2^*$ are all identically equal to 0. We observe from our numerical results the following facts:

1. The order among the strategies: $\bar{\pi}_1^* \leq \bar{\pi}^* \leq \bar{\pi}_2^*$ holds.

2. As b approaches (and exceeds) μ_L, the second approximation $\bar{\pi}_2^*$ gets much better, until it becomes almost indistinguishable from the exact strategy, while if b approaches 0, the (considerable) error between the first approximation $\bar{\pi}_1^*$ and the exact value decreases. In both cases, the shapes of the two approximations are similar to the one of $\bar{\pi}^*$. For instance, in the latter case the optimal strategy flattens out and looks like a straight line.

3. The bad performance of $\bar{\pi}_1^*$ may be explained from the fact that it does not satisfy the requirements for the estimate in Proposition 1.4.1. This happens because the Pareto law estimated by [21] has parameter $\alpha = 2.5406 < 3$, which means that it admits finite second moment but not finite third moment (see assumptions of Proposition 1.4.1). Moreover, this approximation is natural for processes with small jumps, whereas the second one, i.e. $\bar{\pi}_2^*$, is more consistent with general jump processes since it is constructed around the jump measure mean μ_L. What is particularly interesting is that an economically meaningful quantity as the Merton Ratio, that we translated into $\bar{\pi}_1^*$ (see Equation 1.24), performs generally much worse than the Taylor approximation.

4. As we already mentioned, essentially the same approximation $\bar{\pi}_1^*$ is numerically investigated in [107]. What the authors found there is that it works rather well for three popular price models. The difference from our setting, which could even explain why we observe such an unsatisfactory performance, is that they are in the context of exponentially additive models, while our price dynamics are purely additive and mean-reverting.

1.6 Appendix

1.6.1 Auxiliary results

The following lemmas are auxiliary results for Proposition 1.3.11. Let us recall that $S = S^{t,s}$ is described by

$$\begin{aligned} dS(u) &= (b(u) - \lambda S(u))\,du + \psi(u) \int_{\mathbb{R}} y\,\overline{N}(dy,du), \qquad u \in (t,T], \\ S(t) &= s, \end{aligned}$$

and can be written explicitly as

$$S^{t,s}(u) = s e^{-\lambda(u-t)} + \int_t^u e^{-\lambda(u-v)} b(v)\,dv + \int_t^u \int_{\mathbb{R}} e^{-\lambda(u-v)} \psi(v) y\,\overline{N}(dy,dv).$$

Furthermore, the candidate solution for the PIDE in Proposition 1.3.11 is defined as

$$G(t,s) = \mathbb{E}\left[\int_t^T f^*(u, S^{t,s}(u))\,du\right]. \tag{1.34}$$

Lemma 1.6.1. *For all $t \in [0,T]$ and $s \in \mathbb{R}$, it holds that*

$$\mathbb{E}\left[\int_t^T |S^{t,s}(u)|\,du\right] < \infty.$$

Proof. For $u \in [t,T]$, we have

$$\mathbb{E}\left[|S^{t,s}(u)|\right] \le |s|e^{-\lambda(u-t)} + \left|\int_t^u e^{-\lambda(u-v)}b(v)\,dv\right| \quad + \mathbb{E}\left[\left|\int_t^u \int_{\mathbb{R}} e^{-\lambda(u-v)}\psi(v)y\,\overline{N}(dy,dv)\right|\right]$$

Since

$$\begin{aligned}\left(\mathbb{E}\left[\left|\int_t^u \int_{\mathbb{R}} e^{-\lambda(u-v)}\psi(v)y\,\overline{N}(dy,dv)\right|\right]\right)^2 &\le \mathbb{E}\left[\left(\int_t^u \int_{\mathbb{R}} e^{-\lambda(u-v)}\psi(v)y\,\overline{N}(dy,dv)\right)^2\right] \\ &= \int_t^u \int_{\mathbb{R}} e^{-2\lambda(u-v)}\psi(v)^2 y^2\,\nu(dy)dv \\ &\le \psi_2^2 \left(\int_{\mathbb{R}} y^2\,\nu(dy)\right)\left(\frac{1-e^{-2\lambda(u-t)}}{2\lambda}\right)\end{aligned}$$

and

$$\left|\int_t^u e^{-\lambda(u-v)}b(v)\,dv\right| \le \int_t^u e^{-\lambda(u-v)}|b(v)|\,dv \le C\left(\frac{1-e^{-\lambda(u-t)}}{\lambda}\right)$$

we find that

$$\int_t^T \mathbb{E}\left[|S^{t,s}(u)|\right]\,du < \infty.$$

We conclude by Tonelli's Theorem. □

Lemma 1.6.2. *The function $G : [0,T] \times \mathbb{R} \to \mathbb{R}$ in Equation (1.34) is well defined. In particular,*

$$\mathbb{E}\left[\int_t^T |f^*(u, S^{t,s}(u))|\,du\right] < \infty.$$

Proof. Just observe that

$$|G(t,s)| \le \mathbb{E}\left[\int_t^T |f^*(u, S^{t,s}(u))|\,du\right] \le C\left(1 + \mathbb{E}\left[\int_t^T |S^{t,s}(u)|\,du\right]\right),$$

which is finite by Lemma 1.6.1. □

Lemma 1.6.3. *The function $G : [0,T] \times \mathbb{R} \to \mathbb{R}$ defined in (1.34) is continuous in the time variable for any fixed $s \in \mathbb{R}$ and continuously differentiable in s for any fixed time $t \in [0,T)$ with bounded derivative. Specifically,*

$$\partial_s G(t,s) = \mathbb{E}\left[-\int_t^T \lambda e^{-\lambda(u-t)}\bar{\pi}^*(u, S^{t,s}(u))\,du\right].$$

Furthermore, $\partial_s G(t,s)$ is Lipschitz continuous in the variable s uniformly in $t \in [0,T]$.

Proof. For the continuity observe that

$$G(t+h,s)-G(t,s)=\mathbb{E}\left[\int_{t+h}^{T} f^*(u,S^{t+h,s}(u))\,du\right]-\mathbb{E}\left[\int_{t}^{T} f^*(u,S^{t,s}(u))\,du\right]$$
$$=\mathbb{E}\left[\int_{t+h}^{T} [f^*(u,S^{t+h,s}(u))-f^*(u,S^{t,s}(u))]\,du\right]-\mathbb{E}\left[\int_{t}^{t+h} f^*(u,S^{t,s}(u))\,du\right].$$

As h tends to zero, the second term vanishes by the dominated convergence theorem (cf. Lemma 1.6.2). For the first term observe that

$$\int_{t+h}^{T}\mathbb{E}\left[|f^*(u,S^{t+h,s}(u))-f^*(u,S^{t,s}(u))|\right]du \le L\int_{t+h}^{T}\mathbb{E}\left[|S^{t+h,s}(u)-S^{t,s}(u)|\right]du.$$

Hence,

$$\begin{aligned}\mathbb{E}[|S^{t+h,s}(u)-S^{t,s}(u)|] &\le \mathbb{E}[(S^{t+h,s}(u)-S^{t,s}(u))^2]\\ &\le 3s^2e^{-2\lambda(u-t)}(e^{\lambda h}-1)^2+3\left(\int_t^{t+h}e^{-\lambda(u-v)}b(v)\,dv\right)^2\\ &+3\left(\int_{\mathbb{R}}y^2\,\nu(dy)\right)\left(\int_t^{t+h}e^{-2\lambda(u-v)}\psi(v)^2dv\right),\end{aligned}$$

which is Lebesgue-integrable for $u\in[t,T]$ and approaches zero as h tends to zero. Then, by the dominated convergence theorem, it holds that

$$\int_{t+h}^{T}\mathbb{E}\left[|S^{t+h,s}(u)-S^{t,s}(u)|\right]du$$

vanishes and that $G(\cdot,s)$ is continuous.

In order to prove differentiability, we apply the classical theorem about differentiation under the integral sign. First, define

$$F(t,s):=\int_t^T f^*(u,S^{t,s}(u))\,du.$$

Since $f^*(u,S^{t,s}(u))$ is continuously differentiable for each u with partial derivative dominated by an integrable function:

$$\partial_s f^*(u,S^{t,s}(u))=-\lambda e^{-\lambda(u-t)}\bar{\pi}^*(t,S^{t,s}(u))\le C\,e^{-\lambda(u-t)},$$

we have

$$\partial_s F(t,s)=-\lambda\int_t^T e^{-\lambda(u-t)}\bar{\pi}^*(t,S^{t,s}(u))\,du.$$

With the same argument, since

$$G(t,s)=\mathbb{E}\left[F(t,s)\right],$$

where $F(t,s)$ is differentiable with dominated derivative, we get the statement.

Fix $t\in[0,T]$. Since

$$\partial_s G(t,s)=\mathbb{E}\left[-\int_t^T \lambda e^{-\lambda(u-t)}\bar{\pi}^*(u,S^{t,s}(u))\,du\right]$$

and $\bar{\pi}^*(t,\cdot)$ is uniformly Lipschitz continuous (cf. Proposition 1.3.6), we have

$$\begin{aligned}
|\partial_s G(t,s+h) - \partial_s G(t,s)| =& \left|\mathbb{E}\left[\int_t^T \lambda e^{-\lambda(u-t)}(\bar{\pi}^*(u,S^{t,s+h}(u)) - \bar{\pi}^*(u,S^{t,s}(u)))\,du\right]\right| \\
\leq& \mathbb{E}\left[\int_t^T \lambda e^{-\lambda(u-t)}\left|\bar{\pi}^*(u,S^{t,s+h}(u)) - \bar{\pi}^*(u,S^{t,s}(u))\right|\,du\right] \\
\leq& L\,\mathbb{E}\left[\int_t^T \lambda e^{-\lambda(u-t)}\left|S^{t,s+h}(u) - S^{t,s}(u)\right|\,du\right] \\
=& C|h|\left(\int_t^T \lambda e^{-2\lambda(u-t)}\,du\right) \\
=& C|h|,
\end{aligned}$$

where C is a constant depending only on T, λ and the Lipschitz constant L of $\bar{\pi}^*(u,\cdot)$ (which is independent of u). $\square$

Lemma 1.6.4. *For each $s \in \mathbb{R}$, $t \in [0,T)$ and $h > 0$, it holds that*

$$\int_t^{t+h}\int_{\mathbb{R}} \mathbb{E}^{t,s}\left[\big(G(t+h,S(u)+\psi(u)y) - G(t+h,S(u))\big)^2\right]\nu(dy)du$$

and

$$\int_t^{t+h}\int_{\mathbb{R}} \mathbb{E}^{t,s}[|G(t+h,S(u)+\psi(u)y) - G(t+h,S(u)) - \partial_s G(t+h,S(u))\psi(u)y|]\nu(dy)du$$

are finite.

Proof. For the first term it is sufficient to observe that

$$\mathbb{E}^{t,s}\left[\big(G(t+h,S(u)+\psi(u)y) - G(t+h,S(u))\big)^2\right] \leq \sup_{z\in\mathbb{R}} G_s(t+h,z)^2\psi(u)^2y^2 \leq Ce^{2\lambda u}y^2.$$

For the second part of the statement, recall from Lemma 1.6.3 that $\partial_s G(t,s)$ is Lipschitz continuous in s uniformly in t. Let us denote by $\phi(t,s)$ the weak derivative of $\partial_s G(t,s)$ (which is bounded). Therefore, we can write the Taylor expansion of $G(t+h,\cdot)$ in $s+\psi(u)y$ around the center s with integral remainder:

$$G(t+h,s+\psi(u)y) = G(t+h,s) + \partial_s G(t+h,s)\psi(u)y + \int_s^{s+\psi(u)y}\phi(t+h,\xi)(s+\psi(u)y-\xi)d\xi.$$

Hence, for all $s \in \mathbb{R}$, we have that

$$\begin{aligned}
&|G(t+h,s+\psi(u)y) - G(t+h,s) - \partial_s G(t+h,s)\psi(u)y| \\
&\leq \int_s^{s+\psi(u)y} |\phi(t+h,\xi)||s+\psi(u)y-\xi|d\xi \\
&\leq C\int_s^{s+\psi(u)y} |s+\psi(u)y-\xi|d\xi = C\,\psi(u)^2y^2 \\
&\leq C\,y^2.
\end{aligned}$$

As a consequence,

$$\int_t^{t+h}\int_{\mathbb{R}} \mathbb{E}^{t,s}[|G(t+h,S(u)+\psi(u)y) - G(t+h,S(u)) - \partial_s G(t+h,S(u))\psi(u)y|]\nu(dy)du \leq C\int_{\mathbb{R}} y^2\nu(dy),$$

which is finite by our standing assumptions on the Lévy measure ν. $\square$

1.6.2 Technical proofs

In this Appendix we collect some of the most technical proofs.

Proof of Proposition 1.3.6

We prove it for **case A**, the other cases being analogous. Let us denote $\bar{\pi}_1 = \min \Pi$, $\bar{\pi}_2 = \max \Pi$. We already observed, due to the concavity of the function $f(\cdot; t, s)$ in (1.14), that the map $\bar{\pi}^* : [0, T] \times \mathbb{R} \to [\bar{\pi}_1, \bar{\pi}_2]$ is well defined. Also, from (1.18) we immediately get that $\bar{\pi}^*\left(t, \frac{b(t)}{\lambda}\right) = 0$.

The continuity of the function $\bar{\pi}^* : [0, T] \times \mathbb{R} \to \Pi$ relies on a general argument based on the concavity of $f(\bar{\pi}; t, s)$. If $z = (t, s) \in [0, T] \times \mathbb{R}$, then $\bar{\pi}^* = \bar{\pi}^*(z) = \arg\max_{\bar{\pi} \in \Pi} f(\bar{\pi}; z)$. Take a sequence $(z_k)_k \subset [0, T] \times \mathbb{R}$ such that $z_k \to z_0$ as $k \to \infty$. Then the statement follows from proving that $\bar{\pi}^*(z_k) \to \bar{\pi}^*(z_0)$. This is equivalent to saying that each subsequence of $\bar{\pi}^*(z_k)$ admits a subsequence which converges to $\bar{\pi}^*(z_0)$. Take any subsequence of $\bar{\pi}^*(z_k)$ and denote it by $\bar{\pi}^*(z_k)$ (i.e. we do not rename the indexes). Since $\bar{\pi}^*(z_k) \subset \Pi$, which is compact, it admits a subsequence $\bar{\pi}^*(z_{k_h})$ converging to a limit $\bar{\pi}_0 \in \Pi$ as $h \to \infty$. Observe that, for any $\bar{\pi} \in \Pi$,

$$f(\bar{\pi}_0; z_0) = \lim_{h \to \infty} f(\bar{\pi}^*(z_{k_h}); z_{k_h}) \geq \lim_{h \to \infty} f(\bar{\pi}; z_{k_h}) = f(\bar{\pi}; z_0).$$

By definition of $\bar{\pi}^*$, we have that $\bar{\pi}_0 = \bar{\pi}^*(z_0)$, which implies that $\bar{\pi}^*(z_{k_h}) \to \bar{\pi}^*(z_0)$. Since this argument is valid for an arbitrary subsequence of $\bar{\pi}^*(z_k)$, we obtain that $\bar{\pi}^*(z_k)$ itself must converge to $\bar{\pi}^*(z_0)$ as $k \to \infty$.

Now, fix a $t \in [0, T]$. Then, the first order condition can be inverted in the following sense:

$$s^*(t, \bar{\pi}) := \frac{1}{\lambda} \left(b(t) - \sigma(t)^2 \bar{\pi} - \int_{\mathbb{R}} \frac{\bar{\pi} \psi(t)^2 y^2}{1 + \bar{\pi} \psi(t) y} \, \nu(dy) \right),$$

so to define the inverse function of $\bar{\pi}^*(t, \cdot)$ from $(\bar{\pi}^*(t, \cdot))^{-1}\big((\bar{\pi}_1, \bar{\pi}_2)\big)$ to $(\bar{\pi}_1, \bar{\pi}_2)$. By the Inverse Function Theorem (cf. [62, Appendix C.5]), since

$$\frac{\partial s^*(t, \bar{\pi})}{\partial \bar{\pi}} = \frac{1}{\lambda} \left(-\sigma(t)^2 - \int_{\mathbb{R}} \frac{\psi(t)^2 y^2}{(1 + \bar{\pi} \psi(t) y)^2} \, \nu(dy) \right) < 0,$$

we get that, for a fixed $t \in [0, T]$, $\bar{\pi}^*(t, \cdot)$ is strictly decreasing and continuously differentiable for any $s \in \big(\bar{\pi}^*(t, \cdot)\big)^{-1}\big((\bar{\pi}_1, \bar{\pi}_2)\big)$. Observe, in particular, that $\big(\bar{\pi}^*(t, \cdot)\big)^{-1}\big((\bar{\pi}_1, \bar{\pi}_2)\big)$ must be an interval, which we denote by $(s_1(t), s_2(t))$. Also, since $s^*(t, \cdot)$ is smooth by Lebesgue's dominated convergence theorem, with n-th derivative

$$\frac{\partial^n s^*(t, \bar{\pi})}{\partial \bar{\pi}^n} = (-1)^n \frac{n!}{\lambda} \int_{\mathbb{R}} \frac{\psi(t)^{n+1} y^{n+1}}{(1 + \bar{\pi} \psi(t) y)^{n+1}} \, \nu(dy), \tag{1.35}$$

then $\bar{\pi}^*(t, \cdot) : (s_1(t), s_2(t)) \to (\bar{\pi}_1, \bar{\pi}_2)$ is smooth. The boundedness of the derivatives of $\bar{\pi}^*(t, \cdot)$ follows from (1.35) and the fact that, by definition, $1 + \bar{\pi} \psi(t) y \geq \delta > 0$ for each $t \in [0, T]$ and $\bar{\pi} \in \Pi$.

Since $f'(\bar{\pi}; t, s) \to \mp\infty$ as $s \to \pm\infty$, uniformly with respect to $\bar{\pi} \in \Pi$ and $t \in [0, T]$, there exist s_1, s_2 independent from $\bar{\pi}$ and t such that, for any $t \in [0, T]$ and $\bar{\pi} \in [\bar{\pi}_1, \bar{\pi}_2]$, $-\infty < s_1 \leq \frac{b(t)}{\lambda} \leq s_2 < \infty$ and

$$\begin{cases} f'(\bar{\pi}; t, s) > 0, & \text{if } s \leq s_1, \\ f'(\bar{\pi}; t, s) < 0, & \text{if } s \geq s_2. \end{cases}$$

By monotonicity of f, it follows that

$$\bar{\pi}^*(t,s) \equiv \begin{cases} \bar{\pi}_2, & \text{if } s \le s_1, \\ \bar{\pi}_1, & \text{if } s \ge s_2, \end{cases}$$

which proves the fourth statement.

The Lipschitz continuity of $\bar{\pi}^*(t,\cdot)$ follows from the fact that its derivative in $[s_1(t), s_2(t)]$ is bounded uniformly in t and that $\bar{\pi}^*(t,\cdot)$ is constant outside $[s_1(t), s_2(t)]$ (and the constant is independent of t).

Proof of Theorem 1.3.13

The admissibility of $\pi^*(u, S(u-), X(u-)) = \bar{\pi}^*(u, S(u-))X(u-)$ is immediate consequence of Proposition 1.3.6. In order to apply the Verification Theorem (Theorem 1.2.4), which allows us to conclude, we need to prove that $H(u,s,x) = \log(x) + g(u,s)$ satisfies the Dynkin formula (1.10). By Itô's lemma, for each admissible strategy π, we get

$$\begin{aligned} dH(u,S(u),X(u)) = & H_u(u,S(u),X(u)) + A^\pi H(u,S(u),X(u)) \\ & + \sigma(u)\left(\frac{\pi(u)}{X(u)} + g_s(u,S(u))\right) dW(u) \\ & + \int_{\mathbb{R}} \Big[\log(X(u-) + \pi(u)\psi(u)y) - \log(X(u-))\Big]\, \overline{N}(dy,du) \\ & + \int_{\mathbb{R}} \Big[g(u, S(u-) + \psi(u)y) - g(u, S(u-))\Big]\, \overline{N}(dy,du). \end{aligned}$$

Since $\pi(u) = \bar{\pi}(u, S(u-))X(u-)$, we can rewrite it as

$$\begin{aligned} dH(u,S(u),X(u)) = & H_u(u,S(u),X(u)) + A^\pi H(u,S(u),X(u)) \\ & + \sigma(u)\left(\bar{\pi}(u,S(u)) + g_s(u,S(u))\right) dW(u) \\ & + \int_{\mathbb{R}} \log(1 + \bar{\pi}(u,S(u-))\psi(u)y)\, \overline{N}(dy,du) \\ & + \int_{\mathbb{R}} \Big[g(u, S(u-) + \psi(u)y) - g(u, S(u-))\Big]\, \overline{N}(dy,du). \end{aligned}$$

Then the validity of Dynkin's formula boils down to the martingale property of the process

$$\begin{aligned} dZ(u) := & \sigma(u)\left(\bar{\pi}(u,S(u)) + g_s(u,S(u))\right) dW(u) \\ & + \int_{\mathbb{R}} \log(1 + \bar{\pi}(u,S(u-))\psi(u)y)\, \overline{N}(dy,du) \\ & + \int_{\mathbb{R}} \Big[g(u, S(u-) + \psi(u)y) - g(u, S(u-))\Big]\, \overline{N}(dy,du). \end{aligned}$$

Sufficient conditions are

$$\begin{aligned} & \mathbb{E}\left[\int_t^T \sigma(u)^2\, \bar{\pi}(u, S^{t,s}(u))^2\, du\right] < \infty, \\ & \mathbb{E}\left[\int_t^T \int_{\mathbb{R}} \left[\log(1 + \bar{\pi}(u,S(u))\psi(u)y)\right]^2\, \nu(dy)\, du\right] < \infty, \\ & \mathbb{E}\left[\int_t^T \int_{\mathbb{R}} \left[g(u, S^{t,s}(u) + \psi(u)y) - g(u, S^{t,s}(u))\right]^2\, \nu(dy)\, du\right] < \infty, \\ & \mathbb{E}\left[\int_t^T \sigma(u)^2\, g_s(u, S^{t,s}(u))^2\, du\right] < \infty. \end{aligned}$$

The first two follow from the definition of Π (the range of $\bar{\pi}$) and the boundedness of σ, while the last conditions depend on g and are assumed valid in the statement. By standard integrability reasonings (see e.g. [107, Section 3.1] for an argument), we get that Z is a martingale, and then the result.

Proof of Proposition 1.4.1

We follow the same lines of reasoning as in Section 4 of [28]. First, let us write from (1.26) for $\pi \in \mathring{\Pi}$ and fixed $t \in [0,T]$ and $s \in \mathbb{R}$:

$$h(\pi) := h(\pi; t, s) = b(t) - \lambda s - \pi\sigma(t)^2 - \psi(t)^2 \int_{\mathbb{R}} \pi y^2 \, \nu(dy),$$

observing that

$$\begin{aligned} &h(\bar{\pi}_1^*) = 0, \\ &h(\bar{\pi}^*) = \psi(t)^2 \int_{\mathbb{R}} \frac{\bar{\pi}^* y^2}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) - \psi(t)^2 \int_{\mathbb{R}} \bar{\pi}^* y^2 \, \nu(dy) = - \int_{\mathbb{R}} \frac{\psi(t)^3 (\bar{\pi}^*)^2 y^3}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy). \end{aligned}$$

Hence,

$$\bar{\pi}_1^* = h^{-1}(0), \qquad \bar{\pi}^* = h^{-1}\left(- \int_{\mathbb{R}} \frac{\psi(t)^3 (\bar{\pi}^*)^2 y^3}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) \right).$$

Now,

$$|\bar{\pi}^* - \bar{\pi}_1^*| = \left| h^{-1}\left(- \int_{\mathbb{R}} \frac{\psi(t)^3 (\bar{\pi}^*)^2 y^3}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) \right) - h^{-1}(0) \right|.$$

By applying the Mean Value Theorem, since $\bar{\pi}^*$ takes values in $\Pi = \Pi_{\nu,\psi}$, a compact set whose distance from the boundary of $\widehat{\Pi}$ is a certain $\delta > 0$ and $|(h^{-1}(z))'| = \frac{1}{\sigma(t)^2 + \sigma_L(t)^2}$, we finally get the above estimate

$$\begin{aligned} |\bar{\pi}^* - \bar{\pi}_1^*| &\le \frac{1}{\sigma(t)^2 + \sigma_L(t)^2} \left| \int_{\mathbb{R}} \frac{\psi(t)^3 (\bar{\pi}^*)^2 y^3}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) \right| \le \frac{\psi_2^3}{\sigma_1^2 + \psi_2^2 \sigma_\nu^2} \int_{\mathbb{R}} \frac{(\bar{\pi}^*)^2}{1+\bar{\pi}^*\psi(t)y} |y|^3 \, \nu(dy) \\ &\le C \int_{\mathbb{R}} |y|^3 \, \nu(dy), \end{aligned}$$

where C is the constant in the statement.

Proof of Proposition 1.4.2

Define for $\bar{\pi} \in \mathring{\Pi}$ and fixed $t \in [0,T]$ and $s \in \mathbb{R}$:

$$h(\bar{\pi}) := h(\bar{\pi}; t, s) = b(t) - \lambda s - \bar{\pi}\sigma(t)^2 - \frac{\bar{\pi}\psi(t)^2 \mu_F^2}{1+\bar{\pi}\psi(t)\mu_F}\eta,$$

being $\eta = \nu([m, M] \setminus \{0\})$. Then,

$$\begin{aligned} &h(\bar{\pi}_2^*) = 0, \\ &h(\bar{\pi}^*) = \int_{\mathbb{R}} \frac{\bar{\pi}^*\psi(t)^2 y^2}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) - \frac{\bar{\pi}^*\psi(t)^2 \mu_F^2}{1+\bar{\pi}^*\psi(t)\mu_F}\eta, \end{aligned}$$

and

$$\bar{\pi}_2^* = h^{-1}(0), \qquad \bar{\pi}^* = h^{-1}\left(\int_{\mathbb{R}} \frac{\bar{\pi}^*\psi(t)^2 y^2}{1+\bar{\pi}^*\psi(t)y} \, \nu(dy) - \frac{\bar{\pi}^*\psi(t)^2 \mu_F^2}{1+\bar{\pi}^*\psi(t)\mu_F}\eta \right).$$

Coming back to (1.27), observe that

$$\int_{\mathbb{R}} \frac{\bar{\pi}^* \psi(t)^2 y^2}{1+\bar{\pi}^* \psi(t) y}\, \nu(dy) - \frac{\bar{\pi}^* \psi(t)^2 \mu_F^2}{1+\bar{\pi}^* \psi(t)\mu_F}\eta = \psi(t)^2 \int_{\mathbb{R}} (\phi(y) - \phi(\mu_F))\nu(dy).$$

By applying the Mean Value Theorem to ϕ:

$$|\phi(y) - \phi(\mu_F)| \leq \sup_{z \in [m,M]\cap\mathbb{R}} |\phi'(z)||y-\mu_F| \leq C_1 |y - \mu_F|,$$

with, for instance in **case A**, $C_1 = \max\{1, \frac{1}{(\delta\psi_2 M)^2}, \frac{1}{(\delta\psi_2 m)^2}\} + \max\{1, \frac{1}{\delta\psi_2 M}, \frac{1}{-\delta\psi_2 m}\}$, since (cf. Proposition 1.4.1)

$$|\phi'(z)| = \frac{|\bar{\pi} z|(2+\bar{\pi}\psi(t)z)}{(1+\bar{\pi}\psi(t)z)^2} = \frac{|\bar{\pi} z|}{(1+\bar{\pi}\psi(t)z)^2} + \frac{|\bar{\pi} z|}{(1+\bar{\pi}\psi(t)z)}.$$

Therefore, we have found the following estimate:

$$\left| \int_{\mathbb{R}} \frac{\bar{\pi}^* \psi(t)^2 y^2}{1+\bar{\pi}^* \psi(t) y}\, \nu(dy) - \frac{\bar{\pi}^* \psi(t)^2 \mu_F^2}{1+\bar{\pi}^* \psi(t)\mu_F}\eta \right| \leq \psi_2^2\, C_1 \int_{\mathbb{R}} |y-\mu_F|\, \nu(dy).$$

Now,

$$|\bar{\pi}^* - \bar{\pi}_2^*| = \left| h^{-1}\left(\int_{\mathbb{R}} \frac{\bar{\pi}^* \psi(t)^2 y^2}{1+\bar{\pi}^* \psi(t) y}\, \nu(dy) - \frac{\bar{\pi}^* \psi(t)^2 \mu_F^2}{1+\bar{\pi}^* \psi(t)\mu_F}\eta \right) - h^{-1}(0) \right|.$$

Before applying the Mean Value Theorem again, let us compute

$$h'(\bar{\pi}) = -\sigma(t)^2 - \frac{\psi(t)^2 \mu_F^2 \eta}{(1+\bar{\pi}\psi(t)\mu_F)^2},$$

which is negative for each $\bar{\pi} \in \Pi$, so that

$$|h'(\bar{\pi})| = \sigma(t)^2 + \frac{\mu_F^2 \psi(t)^2 \eta}{(1+\bar{\pi}\psi(t)\mu_F)^2}.$$

Finally, denoting $z = h(\bar{\pi})$, since $(h^{-1}(z))' = \frac{1}{h'(h^{-1}(z))} = \frac{1}{h'(\bar{\pi})}$,

$$\begin{aligned}
|\bar{\pi}^* - \bar{\pi}_2^*| &\leq \sup_{z\in[m,M]\cap\mathbb{R}} |(h^{-1}(z))'| \left| \int_{\mathbb{R}} \frac{\bar{\pi}^* \psi(t)^2 y^2}{1+\bar{\pi}^* \psi(t) y}\, \nu(dy) - \frac{\bar{\pi}^* \psi(t)^2 \mu_F^2}{1+\bar{\pi}^* \psi(t)\mu_F}\eta \right| \\
&\leq \frac{\psi_2^2\, C_1}{\sigma(t)^2 + \frac{\mu_F^2 \psi_2^2 \eta}{\sup_{\bar{\pi}\in\Pi}(1+\bar{\pi}\psi(t)\mu_F)^2}} \int_{\mathbb{R}} |y-\mu_F|\, \nu(dy) \leq \frac{\psi_2^2\, C_1}{\sigma_1^2 + \mu_F^2\, \psi_2^2\, \eta\, C_2} \int_{\mathbb{R}} |y-\mu_F|\, \nu(dy) \\
&\leq \frac{\psi_2^2\, C_1\, \eta\, \sigma_F}{\sigma_1^2 + \mu_F^2\, \eta\, \psi_2^2\, C_2},
\end{aligned}$$

where σ_F is the square root of the variance of the jump size and C_1, C_2 are the constants in the statement.

CHAPTER 2

Additive energy forward curves in a Heath-Jarrow-Morton framework

One of the peculiarities of power and gas markets is the delivery mechanism of forward contracts. The seller of a futures contract commits to deliver, say, power, over a certain period, while the classical forward is a financial agreement settled on a maturity date. Our purpose is to design a Heath-Jarrow-Morton framework for an additive, mean-reverting, multicommodity market consisting of forward contracts of any delivery period. Even for relatively simple dynamics, we face the problem of finding a density between a risk-neutral measure $\mathbb{Q}$, such that the prices of traded assets like forward contracts are true $\mathbb{Q}$-martingales, and the real world probability $\mathbb{P}$, under which forward prices are mean-reverting. By assuming that forward prices can be represented as affine functions of a universal source of randomness, we can completely characterize the models which prevent arbitrage opportunities. In this respect, we prove two results on the martingale property of stochastic exponentials. The first allows to validate measure changes made of two components: an Esscher-type density and a Girsanov transform with stochastic and unbounded kernel. The second uses a different approach and works for the case of continuous density. We show how this framework provides an explicit way to describe a variety of models by introducing, in particular, a generalized Lucia-Schwartz model and a cross-commodity cointegrated market.

2.1 Introduction

Since their deregulations, which took place in many countries over the last few decades, energy markets are rapidly evolving sectors and are bringing to the attention of practitioners and researchers challenging problems from the modeling perspective. The most active segment is often the forward market and, as such, the most liquid derivative products are forward contracts. We do not make distinctions between forwards and futures, since the results are the same in the case of deterministic interest rate, as we assume here. Throughout this paper, we adopt the same terminology as [23], and reserve the name *forward* to contracts with delivery at a fixed future time, while agreements to deliver the commodity over a period are called *swaps*. The latter category is important especially in modeling electricity and natural gas markets, where the commodity is exchanged, either physically or financially, over a certain time period (e.g. a day, month or a whole year). A thorough study of mathematical models for these markets, together with a description of their most salient empirical features, can be found, for instance, in [32, 121].

This paper aims at developing a consistent and tractable framework for a multicommodity energy forward market by applying the Heath-Jarrow-Morton paradigm [76]. Intuitively, this consists in describing forward prices as stochastically evolving functions of time and delivery dates, across different markets. To the best of our knowledge, the first work to apply this idea to energy-related commodities is [49] and since then many other works have followed, among others [90, 79, 43, 89]; for a review of HJM models in power and gas markets and further references see [32, Chapter 6]. In this paper, we want to design a realistic forward market model with the same philosophy of [23], where it is theoretically possible to trade contracts with any delivery period and no-arbitrage relations must hold among them. We propose a mean-reverting stochastic process (indeed, an Ornstein-Uhlenbeck-type of model parameterized over delivery times) for the forward and swap price dynamics. By formulating the model under the market probability measure $\mathbb{P}$, we can represent the stylized empirical behavior of observed prices. The no-arbitrage constraints among forwards and swaps are established via explicit relations between the parameters in the respective dynamics. As pointed out in [23], specifying a stochastic evolution for the forward curve and then deriving the swap price as the average over the delivery times has the disadvantage of losing desirable distributional features and, in some cases, even the Markov property. This results in non-tractable models, which inherit a complex probabilistic structure. Therefore, in that paper the authors argue in favor of a swap market model, where only the so-called atomic swaps are directly modeled. Though prices of non-atomic swaps can be reconstructed by arbitrage arguments, it is not possible to use all the available information in the market when fitting the model on real data. These facts motivate us to look for a HJM market model general enough to include both forwards and swaps, still preserving tractability of the resulting stochastic structure.

In this paper we study dynamical models which are additive, meaning that we do not perform a logarithmic transformation of prices (see Section 2 for more details). Recently, additive models have gained popularity, especially for describing power spot prices, e.g. [18, 65, 57, 78, 88, 93]. This is due to their ability of reproducing rather well the stylized features of electricity prices, including the empirical evidence of negative spot prices, and providing explicit formulas for derivative pricing. Generally the problem, when these models are used for commodities other than power, is that negative prices can occur. Nevertheless, according to one's modeling preferences, it is possible e.g. to resort to subordinators (see [18]). The multi-commodity setting has not been extensively studied for energy forward markets applications (see, however, the spot-based models by [67, 83, 106]). With the gradual integration of power markets, say, there is a need for cross-commodity dynamical models which can describe the price evolutions in different power markets simultaneously, such as the Nordic NordPool power and the German EEX market. In fact, our multicommodity framework allows us to develop models where commodities have various kinds of dependencies, like correlations among the driving processes or cointegration/price couplings among them (see Section 5). In addition, it opens the way for computing in a more realistic way the prices of multicommodity derivative assets, as dark or spark spreads written either on spot or futures prices and, on the risk management side, implementing consistently risk measures like PaR and VaR on multicommodity portfolios.

Our main idea is to express forward prices as affine transformations of a universal source of randomness, the latter being independent of the delivery date (see Assumption 6). This simplifying assumption allows both to preserve Markovianity and to describe consistently the related swap price processes. Most importantly, this will be the key to proving the existence of equivalent martingale measures $\mathbb{Q}$, ensuring arbitrage-free models. In fact, the presence of mean-reversion compels us to face non-trivial mathematical hurdles. In this regard, we prove in Theorem 2.3.5 the martingale property of stochastic exponentials where the Lévy part is of Esscher-type, while the Girsanov kernel of the Brownian component is affine in the state variable and thus, in particular, stochastic. Then, we give a different proof in the case of continuous density and continuous kernel, by applying a weak Novikov-type condition on the series representation of the exponential function (cf. Theorem 2.3.6). The proof of this result relies on the asymptotic properties of Gaussian moments. As we move on to consider more general Lévy kernels, we see that the same technique can not be applied. This is shown to be related to the fact that various examples of infinitely divisible distributions, except the Gaussian distribution, do not satisfy the needed moments asymptotics.

We demonstrate the utility of our theoretical framework by specifying two exemplary models. The first, which is presented in Section 4, is a generalization of the additive two-factor Lucia-Schwartz model [95]. We extend it by introducing a mean-reverting arbitrage-free forward dynamics, which is capable to describe a finer volatility term structure. This allows us to account for seasonal effects in price variations, as is typically observed, for instance, in power or gas markets. A calibration procedure and an empirical application to the German power futures market of a version of this model has been carried out in a parallel work by [93]. Furtherly, we introduce a multidimensional model for a mean-reverting cointegrated forward market that respects the no-arbitrage constraints. In particular, we see how these constraints imply certain conditions on the mean-reversion coefficients of the futures curve dynamics (see Section 5).

The paper is structured as follows. Section 2 describes our application of the HJM approach to a multicommodity, additive, energy forward market. In Section 3 we study how our main assumption identifies the dynamics of the processes and the set of equivalent martingale measures. Then, we prove the two main results on the martingale property of stochastic exponentials. A generalization of the two-factor Lucia-Schwartz model [95] is presented in Section 4. Finally, in Section 5 we develop a cross-commodity model with cointegrated dynamics. In Section 6 we make some final remarks. Appendix A contains the proof of Theorem 2.3.5, while the proof of Theorem 2.3.6 is presented in Appendix B.

2.2 A HJM-type market for energy forward contracts

In this section we introduce the general structure of our forward market by applying the Heath-Jarrow-Morton approach to energy forward curves. As already mentioned, we assume that it is possible to trade contracts for one or more commodities in the form of forwards with instantaneous maturity and swaps with arbitrary delivery period. In this way we are able to introduce a sufficiently general framework to be applied to energy-related markets in a consistent way. Both the forward and swap dynamics are described by means of stochastic differential equations and the relations among them follow from natural no-arbitrage conditions. We extend the analysis by [23, Sections 3 and 4] to a multidimensional framework. Furthermore, we introduce a (deterministic) volatility modulated pure-jump Lévy component in the price dynamics. We start from their setting and investigate how to specify a flexible, yet sufficiently tractable, multivariate additive mean-reverting model. We emphasize that our dynamics are described under the empirical probability measure $\mathbb{P}$.

Let us first introduce the stochastic basis underlying our framework.

Assumption 3. *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses (see [87]). Let W be a multidimensional Brownian motion and $\overline{N}(dt, dy) := N(dt, dy) - dt\,\nu(dy)$*

denote the compensated Poisson random measure associated to a centered square-integrable pure-jump Lévy process

$$J(t) = \int_0^t \int_{\mathbb{R}} y\, \overline{N}(ds, dy) \tag{2.1}$$

where the Lévy measure $\nu(dy)$ *is assumed square-integrable in the sense that* $\int_{\mathbb{R}^k} \|y\|^2\, \nu(dy) < \infty$.[1] *The processes* W *and* J *are both of dimension* k *and independent of each other.*

Assume that we are in a market with n commodities. Fix a time horizon $\mathbb{T}$ for our economy and define the sets $\mathcal{A}_1^{\mathbb{T}} = \{(t,T) \in [0,\mathbb{T}]^2 : t \leq T\}$ and $\mathcal{A}_2^{\mathbb{T}} = \{(t,T_1,T_2) \in [0,\mathbb{T}]^3 : t \leq T_1 < T_2\}$. For each $(t,T) \in \mathcal{A}_1^{\mathbb{T}}$, let $f(t,T)$ denote the n-dimensional vector of forward prices at time t for n forward contracts with instantaneous delivery at time T. Analogously, for any $(t,T_1,T_2) \in \mathcal{A}_2^{\mathbb{T}}$, let $F(t,T_1,T_2)$ be the n-dimensional vector of swap prices at time t of the n corresponding contracts with delivery period $[T_1,T_2]$. Let $c : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^n$, $\lambda : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^{n\times n}$, $\sigma : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^{n\times k}$, $\psi : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^{n\times k}$, $C : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^n$, $\Lambda : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^{n\times n}$, $\Sigma : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^{n\times k}$, $\Psi : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^{n\times k}$ be deterministic, measurable vector/matrix fields satisfying the following technical assumptions.

Assumption 4.

- *For any* $T \in [0,\mathbb{T}]$, $t \mapsto c(t,T)$ *is integrable on* $[0,T]$, *i.e.* $\int_0^T \|c(t,T)\|dt$ *is finite.*
- *For any* $T_1, T_2 \in [0,\mathbb{T}]$ *such that* $T_1 < T_2$, $t \mapsto C(t,T_1,T_2)$ *is integrable on* $[0,T_1]$, *i.e.* $\int_0^{T_1} \|C(t,T_1,T_2)\|dt$ *is finite.*
- *For each* $T \in [0,\mathbb{T}]$, $t \mapsto \sigma(t,T)$ *and* $t \mapsto \psi(t,T)$ *are square integrable on* $[0,T]$, *i.e.* $\int_0^T \|\sigma(t,T)\|^2 dt$ *and* $\int_0^T \|\psi(t,T)\|^2 dt$ *are finite.*
- *For any* $T_1, T_2 \in [0,\mathbb{T}]$ *such that* $T_1 < T_2$, $t \mapsto \Sigma(t,T_1,T_2)$ *and* $t \mapsto \Psi(t,T_1,T_2)$ *are square integrable on* $[0,T_1]$, *i.e.* $\int_0^{T_1} \|\Sigma(t,T_1,T_2)\|^2 dt$ *and* $\int_0^{T_1} \|\Psi(t,T_1,T_2)\|^2 dt$ *are finite.*
- *For each* $T \in [0,\mathbb{T}]$, *the matrix field* $t \mapsto \lambda(t,T)$ *is continuous on* $[0,T]$.
- *For any* $T_1, T_2 \in [0,\mathbb{T}]$ *such that* $T_1 < T_2$, *the matrix field* $t \mapsto \Lambda(t,T_1,T_2)$ *is continuous on* $[0,T_1]$.

We assume the following dynamics:

$$df(t,T) = (c(t,T) - \lambda(t,T)f(t,T))\,dt + \sigma(t,T)\,dW(t) + \psi(t,T)\,dJ(t), \tag{2.2}$$

$$\begin{aligned} dF(t,T_1,T_2) = {} & (C(t,T_1,T_2) - \Lambda(t,T_1,T_2)F(t,T_1,T_2))\,dt \\ & + \Sigma(t,T_1,T_2)\,dW(t) + \Psi(t,T_1,T_2)\,dJ(t). \end{aligned} \tag{2.3}$$

The initial conditions $f(0,T)$ and $F(0,T_1,T_2)$ are deterministic Borel measurable functions in $T \leq \mathbb{T}$ and $T_1 < T_2 \leq \mathbb{T}$, respectively.

[1] This assumption can be relaxed to $\int_{\mathbb{R}^k} (\|y\|^2 \wedge 1)\, \nu(dy) < \infty$ and we could derive analogous results in this section by defining a non-compensated Lévy process $\widetilde{J}(t) := \int_0^t \int_{\|y\|<1} y\, \overline{N}(ds,dy) + \int_0^t \int_{\|y\|\geq 1} y\, N(ds,dy)$. However, since we want to ease the mathematical discussion and to focus on the modeling intepretation, we assume the stronger condition $\int_{\mathbb{R}^k} \|y\|^2\, \nu(dy) < \infty$, that in particular implies that J (as defined in Assumption 3) is a square-integrable martingale component of the SDEs in (2.2) and (2.3). Let us remark that for some of the results treated in this paper, we will need even stronger assumptions on ν (cf. Section 3.2), which makes it somehow unprofitable to start here with more general assumptions.

Remark 2.2.1. We would like to have an explicit representation for the solutions of (2.2) and (2.3). To simplify the upcoming discussion, consider the T-independent continuous version of (2.2):

$$df(t) = -\lambda(t)f(t)\,dt + \sigma(t)\,dW(t), \tag{2.4}$$

where $f(0)$ is the identity matrix, $c \equiv 0$ and λ, σ are assumed sufficiently regular so that there exists a unique solution f. Let us introduce the following definition.

Definition 2.2.2 (Commutative property). We say that a square matrix-valued continuous function $A(t)$ satisfies the commutative property (CP) if, for any pair of time values t_1 and t_2, $A(t_1)A(t_2) = A(t_2)A(t_1)$.

If $\lambda(t)$ satisfies (CP), the unique solution U of the matrix ODE

$$dU(t) = \lambda(t)U(t)\,dt,$$

can be written as

$$U(t) = e^{\int_0^t \lambda(u)du}$$

by introducing the matrix exponential $e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$ (see e.g. [40]). In particular, since $U(t)$ is invertible and its inverse satisfies

$$dU^{-1}(t) = -\lambda(t)U^{-1}(t)\,dt, \tag{2.5}$$

we also have that $\lambda(t)U(t) = U(t)\lambda(t)$.[2] These results do not hold in general when λ does not satisfy (CP), due to the non-commutativity of the matrix product. Now, apply Itô's Lemma to $U(t)f(t)$:

$$\begin{aligned} d\left(U(t)f(t)\right) &= \lambda(t)\,U(t)\,f(t)\,dt + U(t)\,df(t) && (2.6)\\ &= \left(\lambda(t)\,U(t) - U(t)\,\lambda(t)\right)f(t)\,dt + U(t)\,\sigma(t)\,dW(t) && (2.7)\\ &= U(t)\,\sigma(t)\,dW(t). && (2.8) \end{aligned}$$

By integrating the above equality, we derive that if $\lambda(t)$ satisfies (CP), then $f(t)$ can be written explicitly as

$$f(t) = e^{-\int_0^t \lambda(u)du} \int_0^t e^{\int_0^s \lambda(u)du}\sigma(s)dW(s). \tag{2.9}$$

In view of Remark 2.2.1, we introduce the following assumption.

Assumption 5. *The matrix-valued functions $\lambda(\cdot, T)$ and $\Lambda(\cdot, T_1, T_2)$ satisfy (CP) for all $T \leq \mathbb{T}$ and $T_1 < T_2 \leq \mathbb{T}$.*

In particular, let us remark that (CP) is fulfilled by any matrix constant in time and any time-dependent diagonal matrix. Assumption 4 ensures that (2.2) and (2.3) admit a unique square-integrable solution (cf. also [23, Appendix A]). Moreover, Assumption 5 (see Remark 2.2.1) guarantees that the solution of (2.2) can be expressed explicitly as

$$\begin{aligned} f(t,T) \quad = \quad & e^{-\int_0^t \lambda(s,T)ds} f(0,T) + \int_0^t e^{-\int_s^t \lambda(u,T)du} c(s,T)ds \\ & + \int_0^t e^{-\int_s^t \lambda(u,T)du} \sigma(s,T)dW(s) + \int_0^t e^{-\int_s^t \lambda(u,T)du} \psi(s,T)dJ(s) \end{aligned}$$

[2] This property follows from (2.5), by observing that $0 = d(U(t)U^{-1}(t)) = dU(t)U^{-1}(t) + U(t)dU^{-1}(t) = \lambda(t)U(t)U^{-1}(t) - U(t)\lambda(t)U^{-1}(t) = \lambda(t) - U(t)\lambda(t)U^{-1}(t)$.

and, analogously, the solution of (2.3) can be written as

$$\begin{aligned} F(t,T_1,T_2) &= e^{-\int_0^t \Lambda(s,T_1,T_2)ds}F(0,T_1,T_2) + \int_0^t e^{-\int_s^t \Lambda(u,T_1,T_2)du}C(s,T_1,T_2)ds \\ &\quad + \int_0^t e^{-\int_s^t \Lambda(u,T_1,T_2)du}\Sigma(s,T_1,T_2)dW(s) + \int_0^t e^{-\int_s^t \Lambda(u,T_1,T_2)du}\Psi(s,T_1,T_2)dJ(s). \end{aligned}$$

These processes are also linked among themselves by no-arbitrage. In fact, in this continuous time setting, a limit argument (see, for instance, [37]) leads to the following no-arbitrage condition among each swap and the corresponding family of forwards: for any $t \leq T_1$ and $T_2 > T_1 > 0$

$$F(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2) f(t,T)\, dT, \tag{2.10}$$

where $\widehat{w} : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^{n\times n}$ is a weight function and represents the time value of money. This can differ from contract to contract according to how settlement takes place (see Equation 4.2 in [23]). For example, in the one-dimensional case we could have $\widehat{w}(T,T_1,T_2) = \frac{1}{T_2-T_1}$ or, by assuming a continuously compounded constant risk-free rate r, $\widehat{w}(T,T_1,T_2) = \frac{e^{-rT}}{\int_{T_1}^{T_2} e^{-rT}\, dT}$. Thus, if we generalize in a straightforward way to a multidimensional market, then we could define, for instance,

$$\widehat{w}(T,T_1,T_2) = \frac{1}{T_2 - T_1} I_n \qquad \text{or} \qquad \widehat{w}(T,T_1,T_2) = \frac{e^{-rT}}{\int_{T_1}^{T_2} e^{-rT}\, dT} I_n,$$

with I_n being the identity matrix of dimension n. We may further enrich this setting by interpreting in a wider sense the role of one, or more, components of $f(t,T)$ and, consequently, the role of $\widehat{w}$ as well. For example, the last component of $f(t,T)$ may represent some *forward* stochastic factor with dynamics governed by (2.2), which affects the value of the swap contracts $F(t,T_1,T_2)$ through (2.10). However, we do not want to look into this matter at the present time, leaving the possibility to investigate it to future work. In view of its modeling intepretation, we define the weight function in the following way.

Definition 2.2.3. A matrix-valued map $\widehat{w} : \mathcal{A}_2^{\mathbb{T}} \to \mathbb{R}^{n\times n}$ is called a weight function if, for any $T_1 < T_2$, it is integrable with respect to T on $[T_1,T_2]$ and

$$\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\, dT = I_n.$$

In the next proposition we state how the no-arbitrage condition in (2.10) determines the coefficients appearing in the dynamics (2.2) and (2.3). Let us observe that, in order to do this, we assume that the mean-reversion speeds λ and Λ are independent of delivery.

Proposition 2.2.4. *For all $T \leq \mathbb{T}$ and $T_1 < T_2 \leq \mathbb{T}$, let $f(\cdot,T)$ and $F(\cdot,T_1,T_2)$ be the unique solutions of (2.2) and (2.3) for given coefficients satisfying Assumption 4, where, in addition, the mean-reversion coefficients are independent of delivery:*

$$\Lambda(t,T_1,T_2) = \lambda(t,T) := \lambda(t). \tag{2.11}$$

Let us assume that the no-arbitrage relation (2.10) holds for a given weight function $\widehat{w}$ *satisfying the following integrability conditions: for all* $T_1 < T_2$,

$$\int_0^{T_1}\int_{T_1}^{T_2} \|\widehat{w}(T,T_1,T_2)\,\sigma(t,T)\|^2\,dT\,dt < \infty,\ \int_0^{T_1}\int_{T_1}^{T_2} \|\widehat{w}(T,T_1,T_2)\,\psi(t,T)\|^2\,dT\,dt < \infty,$$
$$\int_0^{T_1}\int_{T_1}^{T_2} \|\widehat{w}(T,T_1,T_2)\,c(t,T)\|\,dT\,dt < \infty,\ \int_{T_1}^{T_2} \|\widehat{w}(T,T_1,T_2)\,f(0,T)\|\,dT < \infty,$$
$$\int_0^{T_1}\int_{T_1}^{T_2} \|\widehat{w}(T,T_1,T_2)\,\lambda(t)\,f(t,T)\|^2\,dT\,dt < \infty, \quad a.s.$$

If the matrix $\widehat{w}(T,T_1,T_2)$ *commutes with* $\lambda(t)$ *for every* $T \le T_1 < T_2$ *and* $t < T$*, then*

$$C(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)c(t,T)\,dT, \tag{2.12}$$
$$\Sigma(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\sigma(t,T)\,dT, \tag{2.13}$$
$$\Psi(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\psi(t,T)\,dT. \tag{2.14}$$

Proof. The proof is an application of the stochastic Fubini Theorem (cf. e.g. [113, Theorem 64]) and follows by comparing the coefficients of (2.2) to (2.3) after integrating against $\widehat{w}$. From (2.2) we can write

$$\begin{aligned}\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(t,T)\,dT &= \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(0,T)\,dT \\ &+ \int_{T_1}^{T_2}\int_0^t \widehat{w}(T,T_1,T_2)(c(u,T)-\lambda(u)f(u,T))\,du\,dT \\ &+ \int_{T_1}^{T_2}\int_0^t \widehat{w}(T,T_1,T_2)\sigma(u,T)\,dW(u)\,dT + \int_{T_1}^{T_2}\int_0^t \widehat{w}(T,T_1,T_2)\psi(u,T)\,dJ(u)\,dT.\end{aligned}$$

After observing that

$$\int_0^t \widehat{w}(T,T_1,T_2)\lambda(u)f(u,T)\,du = \int_0^t \lambda(u)\widehat{w}(T,T_1,T_2)f(u,T)\,du,$$

we change the order of integration in the last equation (recalling (2.10)) so to obtain

$$\begin{aligned}F(t,T_1,T_2) &= F(0,T_1,T_2) + \int_0^t \left(\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)c(u,T)\,dT - \lambda(u)F(u,T_1,T_2)\right) du \\ &+ \int_0^t\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\sigma(u,T)\,dT\,dW(u) + \int_0^t\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\psi(u,T)\,dT\,dJ(u).\end{aligned}$$

Then, the result follows from the uniqueness of representation of solutions by comparing the coefficients of the last equation to the respective ones in (2.3). □

We finish this section by noticing that, under sufficient regularity assumptions on the evolution of forward prices, we can write down the implied dynamics of the spot price, defined as $S(t) = f(t,t)$.

Proposition 2.2.5. *In addition to Assumption 4, let us assume that* $c(t,T), \sigma(t,T), \psi(t,T)$ *and* $f(t,T)$ *are differentiable with respect to* T *for any* $t \leq T$*, with bounded partial derivatives* $c_T(t,T), \sigma_T(t,T), \psi_T(t,T)$ *and* $f_T(t,T)$*. Then, the spot price follows the dynamics*

$$dS(t) = (c(t,t) + \zeta(t) - \lambda(t)S(t))\,dt + \sigma(t,t)dW(t) + \psi(t,t)dJ(t),$$

where

$$\zeta(t) := f_T(0,t) + \int_0^t (c_T(u,t) - \lambda(u)f_T(u,t))\,du + \int_0^t \sigma_T(u,t)\,dW(u) + \int_0^t \psi_T(u,t)\,dJ(u).$$

Proof. The proof is an application of the stochastic Fubini Theorem and follows the same steps of Proposition 11.1.1 in [100]. □

Naturally, when starting out with a mean-reverting dynamics for swaps and forwards, the implied spot price dynamics will also become mean-reverting. Indeed, we observe that spot prices follow a multidimensional Lévy Ornstein-Uhlenbeck process with time-dependent speed of mean-reversion $\lambda(t)$ and volatility $\sigma(t,t)$. It mean-reverts towards a time-dependent level $\lambda(t)^{-1}(c(t,t)+\zeta(t))$ (whenever $\lambda(t)$ is invertible). This opens for including seasonal variations into the model, which is very relevant in energy markets.

2.3 A mean-reverting model consistent with no-arbitrage

Now we describe how to construct, under a suitable assumption, a HJM-type forward market as in the previous section. We state how this naturally leads us to characterize the model, firstly in terms of its dynamical behavior, and secondly in relation to the existence of equivalent martingale measures and, therefore, arbitrage. In this regard, we show two crucial results about the martingale property of stochastic exponentials. The main result (Theorem 2.3.5) allows us to validate general measure changes with Esscher-type jump component and stochastic kernel in the diffusive part. Then, a different approach is proposed for the specific case of stochastic exponential of Brownian integrals of continuous kernels, so that, in particular, the density process is continuous. Finally, we discuss why this technique does not apply to the case of general Lévy kernels and how this fact is related to asymptotic properties of even moments of infinitely divisible distributions.

Let us begin with our main modeling assumption. We consider the same stochastic basis as in Assumption 3.

Assumption 6. *Take an integer* $m \in \mathbb{N}$*. Assume that the following stochastic differential equation*

$$dX(t) = b(t,X(t))dt + \nu(t,X(t))dW(t) + \eta(t,X(t-))dJ(t) \tag{2.15}$$

admits a unique solution X *taking values in* $\mathbb{R}^m$*, for a drift* $b : [0,\mathbb{T}] \times \mathbb{R}^m \to \mathbb{R}^m$ *and volatility coefficients* $\nu, \eta : [0,\mathbb{T}] \times \mathbb{R}^m \to \mathbb{R}^{m\times k}$*. The forward prices* $f(t,T)$ *have an affine representation*

$$f(t,T) = \alpha(t,T)X(t) + \beta(t,T), \tag{2.16}$$

for some deterministic matrix/vector-valued functions $\alpha : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^{n\times m}$*,* $\beta : \mathcal{A}_1^{\mathbb{T}} \to \mathbb{R}^n$*, which we assume to be bounded on* $\mathcal{A}_1^{\mathbb{T}}$ *and continuously differentiable in* $t \in [0,T]$*, for all* $T \leq \mathbb{T}$*. In view of the no-arbitrage condition (2.10), we define the swap price process* $F(t,T_1,T_2)$ *by*

$$F(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(t,T)\,dT, \tag{2.17}$$

where $\widehat{w}$ *is a weight function as in Definition 2.2.3.*

The process X can be interpreted as a stochastic state variable underlying the forward curves for *all* maturities T. This might be related to the spot price, which is the case, for instance, of the models that we propose in Sections 4 and 5. Observe, in particular, that the stochastic processes $f(\cdot, T)$ and $F(\cdot, T_1, T_2)$ are Markovian. The question of Markovianity for this class of models has been discussed also in [23, Equation 4.8], where the authors show that even simple log-normal specifications of the forward dynamics lead, in general, to non-Markovian swap price processes (unless interpreted as infinite-dimensional stochastic processes). Instead, Assumption 6 allows us to preserve the Markov property of our models and, thus, analytical tractability.

2.3.1 Dynamics

Since we want a dynamical behavior of mean-reverting type for f and F (see (2.2) and (2.3)), the affine structure of forward prices assumed in (2.16) determines in a rather natural way the corresponding functional form of the coefficients expected in (2.15). In this regard, we have two symmetrical results.

Proposition 2.3.1. *Assume in (2.15) that the coefficients b, ν and η take the following affine form*

$$b(t, X(t)) = \theta(t) + \Theta(t)X(t), \tag{2.18}$$
$$\nu(t, X(t)) = v(t), \tag{2.19}$$
$$\eta(t, X(t)) = z(t), \tag{2.20}$$

where $\theta : [0, \mathbb{T}] \to \mathbb{R}^m$ is integrable, $v : [0, \mathbb{T}] \to \mathbb{R}^{m\times k}$ and $z : [0, \mathbb{T}] \to \mathbb{R}^{m\times k}$ are square-integrable and $\Theta : [0, \mathbb{T}] \to \mathbb{R}^{m\times m}$ is bounded. If there exists a continuous matrix-valued function $\lambda : [0, \mathbb{T}] \to \mathbb{R}^{n\times n}$ such that

$$\lambda(t)\alpha(t, T) = -\alpha_t(t, T) - \alpha(t, T)\Theta(t), \tag{2.21}$$

then, by defining

$$c(t, T) := \lambda(t)\beta(t, T) + \beta_t(t, T) + \alpha(t, T)\theta(t), \tag{2.22}$$
$$\sigma(t, T) := \alpha(t, T)v(t), \tag{2.23}$$
$$\psi(t, T) := \alpha(t, T)z(t), \tag{2.24}$$

the stochastic process $f(t, T)$ is the unique solution of

$$df(t, T) = (c(t, T) - \lambda(t)f(t, T))\, dt + \sigma(t, T)dW(t) + \psi(t, T)dJ(t).$$

Furthermore, if $\widehat{w}$ satisfies the assumptions of Proposition 2.2.4, $F(t, T_1, T_2)$ is the unique solution of

$$dF(t, T_1, T_2) = (C(t, T_1, T_2) - \lambda(t)F(t, T_1, T_2))\, dt + \Sigma(t, T_1, T_2)\, dW(t) + \Psi(t, T_1, T_2)\, dJ(t),$$

where C, Σ, Ψ are defined in (2.12),(2.13), (2.14).

Proof. Let us apply Itô's Lemma to

$$f(t, T) = \alpha(t, T)X(t) + \beta(t, T).$$

Hence,

$$df(t,T) = (\beta_t(t,T) + \alpha(t,T)\theta(t) + \alpha_t(t,T)X(t) + \alpha(t,T)\Theta(t)X(t))dt \\ + \alpha(t,T)v(t)dW(t) + \alpha(t,T)z(t)dJ(t). \tag{2.25}$$

Since, by assumption, there exists a function $\lambda : [0, \mathbb{T}] \to \mathbb{R}^{n\times n}$ satisfying

$$\alpha_t(t,T) + \alpha(t,T)\Theta(t) = -\lambda(t)\alpha(t,T)$$

whereas $c(t,T)$, $\sigma(t,T)$ and $\psi(t,T)$ satisfy, by definition,

$$\begin{aligned} \beta_t(t,T) + \alpha(t,T)\theta(t) &= c(t,T) - \lambda(t)\beta(t,T), \\ \alpha(t,T)v(t) &= \sigma(t,T), \\ \alpha(t,T)z(t) &= \psi(t,T), \end{aligned}$$

the statement for f follows after substituting these expressions into (2.25). The result for F can be proven along the same lines of Proposition 2.2.4. □

The following proposition can be seen as the converse of the previous one.

Proposition 2.3.2. *If $f(t,T)$ is the unique solution of*

$$df(t,T) = (c(t,T) - \lambda(t)f(t,T))\,dt + \sigma(t,T)dW(t) + \psi(t,T)dJ(t), \tag{2.26}$$

for c, σ, ψ and λ satisfying Assumption 4, then b, ν and η in (2.15) take the following affine form

$$\alpha(t,T)b(t,X(t)) = \widetilde{\theta}(t,T) + \widetilde{\Theta}(t,T)X(t) \tag{2.27}$$
$$\alpha(t,T)\nu(t,X(t)) = \sigma(t,T), \tag{2.28}$$
$$\alpha(t,T)\eta(t,X(t)) = \psi(t,T), \tag{2.29}$$

with

$$\widetilde{\theta}(t,T) = c(t,T) - \lambda(t)\beta(t,T) - \beta_t(t,T), \tag{2.30}$$
$$\widetilde{\Theta}(t,T) = -\lambda(t)\alpha(t,T) - \alpha_t(t,T). \tag{2.31}$$

Proof. As in the proof of Proposition 2.3.1, the statement is direct consequence of Itô's Lemma applied to

$$f(t,T) = \alpha(t,T)X(t) + \beta(t,T),$$

which gives that

$$df(t,T) = (\beta_t(t,T) + \alpha_t(t,T)X(t) + \alpha(t,T)b(t,X(t)))dt \\ + \alpha(t,T)\nu(t,X(t))dW(t) + \alpha(t,T)\eta(t,X(t-))dJ(t).$$

It is then sufficient to compare the coefficients of the last equation to the ones of (2.26). □

2.3.2 Arbitrage

We now investigate the question of arbitrage in the context of our HJM-type forward model. A sufficient condition for an arbitrage-free market, together with (2.17), is the existence of an equivalent martingale measure, which is a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted price processes of all traded contracts are $\mathbb{Q}$-martingales. Notice that forwards and swaps are costless to enter, so their price processes should be $\mathbb{Q}$-martingale without any discounting. We also mention that the well-known *HJM drift condition* for forward *rate* models in interest rate theory is not present in our context, as forward contracts are tradeable assets unlike the forward rates, which model in a nonlinear way the bond price dynamics.

Firstly, we introduce the candidate density processes.

Assumption 7. *We assume that $N(ds, dy)$ is the Poisson random measure of a Lévy process J with independent components J_j, for $j = 1, \dots, k$. We indicate by $N_j(ds, dy_j)$ the Poisson random measure of J_j and by $\overline{N}_j(ds, dy_j) = N_j(ds, dy_j) - ds\, \nu_j(dy_j)$ the corresponding compensated version.*

Let $\phi = (\phi^{(j)})_{j=1,\dots,k}$ and $\xi = (\xi^{(j)})_{j=1,\dots,k}$ be k-dimensional adapted processes such that

$$\mathbb{E}\left[\int_0^{\mathbb{T}} \|\phi(s)\|^2 ds\right] < \infty, \qquad \mathbb{E}\left[\int_0^{\mathbb{T}} \|\xi(s)\|^2 ds\right] < \infty, \tag{2.32}$$

and define, for $t \in [0, \mathbb{T}]$, the process Z as the unique strong solution of

$$dZ(t) = Z(t-)\, dH(t), \tag{2.33}$$

such that $Z(0) = 1$, where

$$\begin{aligned} H(t) &:= \int_0^t \phi^\intercal(s) dW(s) + \int_0^t \int_{\mathbb{R}^k} \xi^\intercal(s-)\, y\, \overline{N}(ds, dy), \\ &= \sum_{j=1}^k \left(\int_0^t \phi^{(j)}(s) dW_j(s) + \int_0^t \int_{\mathbb{R}} \xi^{(j)}(s-)\, y_j\, \overline{N}_j(ds, dy_j) \right). \end{aligned} \tag{2.34}$$

If the processes ϕ and ξ satisfy (2.32), H is a well-defined square-integrable martingale. The process Z is called the *stochastic* or *Doléans-Dade exponential* of H and is sometimes indicated as

$$Z(t) = \mathcal{E}(H)(t).$$

More explicitly, it can be written as (cf. for instance [113, Theorem II.37])

$$Z(t) = e^{H(t) - \frac{1}{2}\int_0^t \|\phi(s)\|^2\, ds} \prod_{0<s\le t} (1 + \Delta H(s))\, e^{-\Delta H(s)}. \tag{2.35}$$

If Z is a strictly positive martingale, we can introduce an equivalent probability measure $\mathbb{Q}$ defining its Radon-Nikodym derivative as

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := Z(\mathbb{T}). \tag{2.36}$$

Define the stochastic process

$$W^{\mathbb{Q}}(t) := W(t) - \int_0^t \phi(s)\, ds, \tag{2.37}$$

the random measure

$$\overline{N}^{\mathbb{Q}}(dt,dy) := \overline{N}(dt,dy) - \xi(t)^{\intercal} y\,\nu(dy)dt \tag{2.38}$$

and the jump process

$$J^{\mathbb{Q}}(t) := \int_{\mathbb{R}^k} y\,\overline{N}^{\mathbb{Q}}(dt,dy) = J(t) - \int_0^t K\,\xi(s)ds, \tag{2.39}$$

where

$$K = \operatorname{diag}(\kappa), \qquad \kappa = (\kappa_j)_{j=1,\dots,k}, \qquad \kappa_j = \int_{\mathbb{R}} y^2\,\nu_j(dy).$$

The next two propositions identify the subfamily of models admitting the existence of an equivalent martingale measure. In the remainder of this paper, we assume the same assumptions of Proposition 2.3.1, so that, in particular, X evolves as

$$dX(t) = \big(\theta(t) + \Theta(t)X(t)\big)dt + v(t)dW(t) + z(t)dJ(t). \tag{2.40}$$

Proposition 2.3.3. *If f satisfies*

$$df(t,T) = \sigma(t,T)\,dW^{\mathbb{Q}}(t) + \psi(t,T)\,dJ^{\mathbb{Q}}(t),$$

then

$$\alpha(t,T)\big(v(t)\phi(t) + z(t)\,K\,\xi(t)\big) = \Phi_1(t,T)X(t) + \Phi_0(t,T), \tag{2.41}$$

for Φ_0 and Φ_1 denoting

$$\Phi_1(t,T) = -\alpha_t(t,T) - \alpha(t,T)\Theta(t), \tag{2.42}$$
$$\Phi_0(t,T) = -\beta_t(t,T) - \alpha(t,T)\theta(t). \tag{2.43}$$

In particular, if

$$\phi(t) = \phi_1(t)X(t) + \phi_0(t), \tag{2.44}$$
$$\xi(t) = \xi_1(t)X(t) + \xi_0(t), \tag{2.45}$$

then

$$\Phi_1(t,T) = \alpha(t,T)\big(v(t)\phi_1(t) + z(t)\,K\,\xi_1(t)\big), \tag{2.46}$$
$$\Phi_0(t,T) = \alpha(t,T)\big(v(t)\phi_0(t) + z(t)\,K\,\xi_0(t)\big). \tag{2.47}$$

Finally, α and β satisfy the following linear ordinary differential equations:

$$\alpha_t(t,T) = -\alpha(t,T)\gamma_1(t), \tag{2.48}$$
$$\beta_t(t,T) = -\alpha(t,T)\gamma_0(t), \tag{2.49}$$

where γ_0 and γ_1 are

$$\gamma_1(t) = v(t)\phi_1(t) + z(t)\,K\,\xi_1(t) + \Theta(t), \tag{2.50}$$
$$\gamma_0(t) = v(t)\phi_0(t) + z(t)\,K\,\xi_0(t) + \theta(t). \tag{2.51}$$

Proof. Itô's Lemma applied to $f(t,T) = \alpha(t,T)X(t) + \beta(t,T)$ gives

$$df(t,T) = (\beta_t(t,T) + \alpha(t,T)\theta(t) + \alpha_t(t,T)X(t) + \alpha(t,T)\Theta(t)X(t))dt \\ + \alpha(t,T)v(t)dW(t) + \alpha(t,T)z(t)dJ(t).$$

By replacing W and J with $W^{\mathbb{Q}}$ and $J^{\mathbb{Q}}$, we get

$$df(t,T) = \big(\beta_t(t,T) + \alpha(t,T)\theta(t) + \alpha_t(t,T)X(t) + \alpha(t,T)\Theta(t)X(t) \\ + \alpha(t,T)v(t)\phi(t) + \alpha(t,T)z(t)\,K\,\xi(t)\big)dt + \alpha(t,T)z(t)dJ^{\mathbb{Q}}(t) + \alpha(t,T)v(t)dW^{\mathbb{Q}}(t).$$

In order to have zero drift, (2.41) must hold for Φ_0 and Φ_1 as in (2.42) and (2.43). In particular, the last equation holds if ϕ and ξ take the affine form (2.44) and (2.45) with Φ_0 and Φ_1 defined by (2.46) and (2.47), so that

$$\alpha_t(t,T) + \alpha(t,T)\Theta(t) + \alpha(t,T)v(t)\phi_1(t) + \alpha(t,T)z(t)\,K\,\xi_1(t) = 0,$$
$$\beta_t(t,T) + \alpha(t,T)\theta(t) + \alpha(t,T)v(t)\phi_0(t) + \alpha(t,T)z(t)\,K\,\xi_0(t) = 0,$$

which yield (2.48) and (2.49). □

The converse result holds as well.

Proposition 2.3.4. *Assume that the following conditions are fulfilled:*

- *α and β satisfy, for some continuous functions $\gamma_0 : [0,\mathbb{T}] \to \mathbb{R}^m$ and $\gamma_1 : [0,\mathbb{T}] \to \mathbb{R}^{m\times m}$, the linear ordinary differential equations (2.48) and (2.49);*
- *there exist adapted processes $\phi(t)$ and $\xi(t)$ satisfying (2.32) of the form*

$$\phi(t) = \phi_1(t)X(t) + \phi_0(t),$$
$$\xi(t) = \xi_1(t)X(t) + \xi_0(t),$$

  where ϕ_0, ϕ_1, ξ_0 and ξ_1 are in relation with γ_0 and γ_1 by (2.50) and (2.51);
- *the process Z defined in (2.33) is a strictly positive $\mathbb{P}$-martingale on $[0,\mathbb{T}]$.*

Then, by defining $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z(\mathbb{T})$, it holds that $W^{\mathbb{Q}}$ in (2.37) is a Brownian motion under $\mathbb{Q}$ and $\overline{N}^{\mathbb{Q}}(dt,dy)$ defined in (2.38) is the $\mathbb{Q}$-compensated Poisson random measure of N in the sense of [104, Theorem 1.35]. Also, we have

$$df(t,T) = \sigma(t,T)\,dW^{\mathbb{Q}}(t) + \psi(t,T)\,dJ^{\mathbb{Q}}(t),$$
$$dF(t,T_1,T_2) = \Sigma(t,T_1,T_2)\,dW^{\mathbb{Q}}(t) + \Psi(t,T_1,T_2)\,dJ^{\mathbb{Q}}(t),$$

with the same coefficients as in Proposition 2.3.1 and, in particular, the processes $f(\cdot,T)$ and $F(\cdot,T_1,T_2)$ are $\mathbb{Q}$-martingales for all T and $T_1 < T_2$.

Proof. By similar arguments as in Proposition 2.3.3, it is enough to apply Itô's Lemma to $f(t,T) = \alpha(t,T)X(t) + \beta(t,T)$, which gives

$$df(t,T) = \sigma(t,T)\,dW^{\mathbb{Q}}(t) + \psi(t,T)\,dJ^{\mathbb{Q}}(t).$$

Since $F(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)f(t,T)\,dT$, we easily obtain that

$$dF(t,T_1,T_2) = \Sigma(t,T_1,T_2)dW^{\mathbb{Q}}(t) + \Psi(t,T_1,T_2)\,dJ^{\mathbb{Q}}(t).$$

As, by assumption, Z is a $\mathbb{P}$-martingale, the measure $\mathbb{Q}$ is an equivalent probability measure and we can conclude by Girsanov's Theorem (e.g. [104, Theorem 1.35]). □

In Proposition 2.3.4 we assumed that Z is a strictly positive $\mathbb{P}$-martingale on $[0,\mathbb{T}]$, which implies that $Z(\mathbb{T})$ defines the Radon-Nikodym derivative of an equivalent probability measure $\mathbb{Q}$. As a consequence of Proposition 2.3.4, $\mathbb{Q}$ is in fact a martingale measure for the market. However, in general Z is only a local martingale. Thus, the interesting question is under which conditions we can prove that Z is a strictly positive true martingale. Tracing through Propositions 2.3.3 and 2.3.4, it is natural to have martingale measure kernels in affine form:

$$\phi(t) = \phi_1(t)X(t) + \phi_0(t), \tag{2.52}$$
$$\xi(t) = \xi_1(t)X(t) + \xi_0(t). \tag{2.53}$$

The first condition to check is the positivity. From the representation in (2.35) we see that Z is strictly positive if and only if $\Delta H > -1$ a.s. This boils down to verify that

$$\Delta H(t) = \xi^{\intercal}(t-)\Delta J(t) > -1, \qquad \text{for every } t \in [0,\mathbb{T}], \text{ a.s.,}$$

which holds under the following assumption.

Assumption 8. *The jump risk parameter ξ is a bounded and deterministic vector function on $[0,\mathbb{T}]$ (i.e. $\xi_1 \equiv 0$ in (2.53)) such that $\xi^{\intercal}(t)\,y > -1$ for ν-a.e. $y \in \mathbb{R}^k$ and each $t \geq 0$.*

Observe that, in order to define a positive density process, we assume that the jump kernel ξ is deterministic.

Since a positive local martingale is a supermartingale, in order to prove that Z is a true martingale it is sufficient to verify that $\mathbb{E}[Z(\mathbb{T})] = 1$. Let us then state the main result of this section.

Theorem 2.3.5. *Let us assume that there exist bounded measurable deterministic functions $\phi_0 : [0,\mathbb{T}] \to \mathbb{R}^k$, $\phi_1 : [0,\mathbb{T}] \to \mathbb{R}^{k\times m}$. Define the Girsanov kernels $\phi(t) := \phi_1(t)X(t) + \phi_0(t)$ and ξ as in Assumption 8. Assume, furtherly, that ν has fourth moment, that is $\int_{\mathbb{R}^k} \|y\|^4\,\nu(dy) < \infty$. Then, the process Z defined by (2.33) is a strictly positive true martingale.*

Proof. See Appendix A. □

In the next theorem we state the martingale property of the stochastic exponential in the case we have no jumps in the dynamics, that is $\nu \equiv 0$, in an alternative way. The Novikov condition ([103]) is a standard way to prove it in the case of continuous density process, i.e. $H(t) = \int_0^t \phi^{\intercal}(s)dW(s)$. However, applying the Novikov condition for our state-dependent ϕ will yield a valid measure change only for a restricted interval of time $t \in [0,\tau]$, where τ will depend on the parameters in the model (inspect the proof in Appendix B to see this). Thus, the Novikov condition may fail to validate a martingale measure for all times in question and we cannot conclude that the dynamics are arbitrage-free for $t \geq \tau$. Therefore, we apply a weaker Novikov-type criterion, which relies on asymptotic properties of Gaussian moments. We remark in passing that a nice account of the literature related to this problem can be found in the introduction of [86].

Theorem 2.3.6. *Assume that $\Theta : [0,\mathbb{T}] \to \mathbb{R}^{m\times m}$ is a bounded measurable function satisfying (CP), $\theta : [0,\mathbb{T}] \to \mathbb{R}^m$ is integrable and $v : [0,\mathbb{T}] \to \mathbb{R}^{m\times k}$ is square-integrable. Let $\phi_0 : [0,\mathbb{T}] \to \mathbb{R}^k$, $\phi_1 : [0,\mathbb{T}] \to \mathbb{R}^{k\times m}$ be measurable fields such that $\phi_1(t)$ is bounded and $\int_0^{\mathbb{T}} \|\phi_0(s)\|^2 ds < \infty$. Define the Girsanov kernel*

$$\phi(t) = \phi_1(t)X(t) + \phi_0(t),$$

where X is the unique solution of

$$dX(t) = (\theta(t) + \Theta(t)X(t))dt + v(t)dW(t).$$

Then

$$Z(t) = \exp\left(\int_0^t \phi^\intercal(s)dW(s) - \frac{1}{2}\int_0^t \|\phi(s)\|^2 ds\right) \tag{2.54}$$

is a (strictly positive) $\mathbb{P}$*-martingale on* $[0,\mathbb{T}]$.

Proof. See Appendix B. □

We will now show that the same technique employed in the proof of Theorem 2.3.6 fails when adding the jump component in the dynamics of X, even if the density process Z is still continuous, i.e. $\xi \equiv 0$. We need that the pure-jump Lévy component J satisfies the following property related to the asymptotics of *even* moments of stochastic integrals with respect to Lévy processes, which does not hold in very simple cases.

Definition 2.3.7. We say that a Lévy process $J(t) := \int_0^t \int_{\mathbb{R}^k} y\,\overline{N}(ds,dy)$ satisfies the property (P) if, for any $t \in [0,\mathbb{T}]$ and bounded measurable function $h : [0,\mathbb{T}] \to \mathbb{R}^{m\times k}$, it holds that

$$\limsup_{n\to\infty} \frac{1}{n} \frac{\mathbb{E}\left[\| \int_0^t \int_{\mathbb{R}^k} h(s)\,y\,\overline{N}(ds,dy)\|^{2n}\right]}{\mathbb{E}\left[\| \int_0^t \int_{\mathbb{R}^k} h(s)\,y\,\overline{N}(ds,dy)\|^{2n-2}\right]} \le C, \tag{2.55}$$

where $C = C(t,\mathbb{T},h)$ is a constant possibly depending on $t,\mathbb{T}$ and the function h.

Let us consider, for simplicity, the one dimensional case, i.e. $k = 1$, and fix $h \equiv 1$, $t = \mathbb{T} = 1$. Then, property (P) becomes

$$\limsup_{n\to\infty} \frac{1}{n} \frac{\mathbb{E}\left[J(1)^{2n}\right]}{\mathbb{E}\left[J(1)^{2n-2}\right]} \le C. \tag{2.56}$$

Thus, it boils down to the asymptotic behavior of the moments of infinitely divisible distributions. In general, explicit formulas are available for moments of Lévy processes in terms of cumulants, but involve rather complicated combinatorial quantities, which do not allow to interpret in a straightforward way the growth behavior with respect to n (see e.g. [112] or [42, Lemma 2.2, Proposition 2.3]). In the upcoming examples we show that, for very common distributions, the property in (2.56) is actually not verified.

Example 2.3.8 (Poisson process). The moments of the Poisson process $N(t)$ with intensity parameter λ can be expressed in terms of the so-called *Touchard polynomials*:

$$\mathbb{E}[N(1)^n] = T_n(\lambda) = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} \lambda^k,$$

where $\left\{ {n \atop k} \right\}$ is the Stirling number of second kind, namely the number of ways to partition a set of n labelled objects into k nonempty unlabelled subsets. Observe that here the cumulants are represented by λ. If we take, for instance, λ to be equal to 1, then

$$\mathbb{E}[N(1)^n] = T_n(1) = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} =: B_n,$$

where we denote by B_n the n-th Bell number (cf. [114]). By computing the ratio in (2.56) numerically, it turns out that the growth of such Bell numbers is actually faster, in the sense that

$$\limsup_{n\to\infty} \frac{1}{n} \frac{\mathbb{E}\left[N(1)^{2n}\right]}{\mathbb{E}\left[N(1)^{2n-2}\right]} = \limsup_{n\to\infty} \frac{1}{n} \frac{B_{2n}}{B_{2n-2}} = \infty.$$

Example 2.3.9 (Gamma and Poisson subordinators). A popular way to construct Lévy-based models in financial modeling is Brownian subordination (cf. e.g. [51, Section 4.4]). This consists in changing the time of a Brownian motion by an independent subordinator, i.e. a increasing Lévy process. Let W be a Brownian motion and U a subordinator, mutually independent. Then, a classical result is that $L(t) := W(U(t))$ is a Lévy process. The moments of L can be computed in terms of the ones of W and U by the following:

$$\mathbb{E}[L(t)^{2n}] = \mathbb{E}[W(U(t))^{2n}] = \mathbb{E}[\mathbb{E}[W(U(t))^{2n}|U]] = \mathbb{E}[U(t)^n]\mathbb{E}[W(1)^{2n}].$$

Thus,

$$\frac{\mathbb{E}\left[L(1)^{2n}\right]}{\mathbb{E}\left[L(1)^{2n-2}\right]} = \frac{\mathbb{E}[U(1)^n]\,\mathbb{E}[W(1)^{2n}]}{\mathbb{E}[U(1)^{n-1}]\,\mathbb{E}[W(1)^{2n-2}]} \leq C_0\, n\, \frac{\mathbb{E}[U(1)^n]}{\mathbb{E}[U(1)^{n-1}]},$$

thanks to the properties of Gaussian moments. Consequently, the asymptotic condition in (2.56) boils down to the following requirement for the moments of the subordinator:

$$\limsup_{n\to\infty} \frac{\mathbb{E}[U(1)^n]}{\mathbb{E}[U(1)^{n-1}]} \leq C. \tag{2.57}$$

However, if U is a Poisson subordinator, then (see previous example)

$$\frac{\mathbb{E}[U(1)^n]}{\mathbb{E}[U(1)^{n-1}]} = \frac{B_n}{B_{n-1}},$$

which numerically turns out to be unbounded as n tends to infinity. Also, if U is a Gamma subordinator, i.e. $U(1)$ is Gamma distributed, then

$$\frac{\mathbb{E}[U(1)^n]}{\mathbb{E}[U(1)^{n-1}]} \approx C_1 \frac{n!}{\beta^n} \frac{\beta^{n-1}}{(n-1)!} \approx C_2\, n,$$

so that, even in this case, property (P) is not satisfied.

2.3.3 Risk premium

We close this section by discussing very briefly the *risk premium*, which represents a relevant quantity in commodity markets (see e.g. [72]). It is defined as the difference between the forward price and the spot price prediction at delivery, which means, in mathematical terms,

$$\mathrm{RP}^f(t,T) = f(t,T) - \mathbb{E}[f(T,T)|\mathcal{F}_t], \tag{2.58}$$

with the spot price being $S(T) := f(T,T)$. As $f(t,T)$ is a vector, the risk premium is defined as a vector as well. Observe that, under the assumptions of Theorem 2.3.5, we can also write for a martingale measure $\mathbb{Q}$:

$$\mathrm{RP}^f(t,T) = \mathbb{E}^{\mathbb{Q}}[f(T,T)|\mathcal{F}_t] - \mathbb{E}[f(T,T)|\mathcal{F}_t]. \tag{2.59}$$

The economic interpretation of the risk premium is particularly important and has been studied both empirically and theoretically in the context of energy markets by several authors:

see, for instance, [47] for the UK gas market, [31, 17, 34] for power markets and [91] for various energy exchanges. As our present purpose is just to identify it in our model, we refer to the literature above for a more detailed description of the financial and mathematical features of the risk premium.

It is natural to extend the definition of the risk premium to swaps as the difference between the swap price and the expected value of the spot price *weighted* over the delivery period:

$$\mathrm{RP}^F(t,T_1,T_2) = F(t,T_1,T_2) - \mathbb{E}\left[\int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2) f(T,T)\,dT \,\middle|\, \mathcal{F}_t\right]. \tag{2.60}$$

Let us omit the mathematical justifications in order to keep this discussion at a simple level. As a direct consequence of (2.58) and (2.60) (cf. also [31]), we have the following relation:

$$\mathrm{RP}^F(t,T_1,T_2) = \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2)\mathrm{RP}^f(t,T)\,dT.$$

This means that the risk premium of a swap is the weighted average of the forward risk premium $\mathrm{RP}^f(t,T)$ over the delivery period $[T_1,T_2]$. By writing (2.2) in integral form, we have

$$\mathrm{RP}^f(t,T) = \int_t^T \lambda(s)\mathbb{E}[f(s,T)|\mathcal{F}_t]\,ds - \int_t^T c(s,T)\,ds,$$

which becomes, after using (2.10),

$$\mathrm{RP}^f(t,T) = (I_n - e^{-\int_t^T \lambda(u)du}) f(t,T) - \int_t^T e^{-\int_s^T \lambda(u)du} c(s,T)\,ds,$$

Finally, since $\widehat{w}$ commutes with all the matrices of the same size, the swap risk premium can be written as

$$\begin{aligned}\mathrm{RP}^F(t,T_1,T_2) = &\int_{T_1}^{T_2} (I_n - e^{-\int_t^T \lambda(u)du})\,\widehat{w}(T,T_1,T_2)\,f(t,T)\,dT \\ &- \int_{T_1}^{T_2} \widehat{w}(T,T_1,T_2) \int_t^T e^{-\int_s^T \lambda(u)du} c(s,T)\,ds\,dT.\end{aligned}$$

In commodity markets, especially power, the behavior of the observed risk premium is rather involved (see e.g. [125, 17]). Therefore, in order to suitably describe these stylized facts, a stochastic specification has been postulated. To our knowledge, several models studied in literature, with (at least) the exceptions of [25, 84], imply a deterministic structure of risk premium. Indeed, what we have in the present modeling framework is a stochastically sign-changing (in each component) and affine risk premium.

2.4 A generalization of the Lucia-Schwartz model

A classical approach in the literature of energy markets consists in modeling directly the spot price dynamics: see, for instance, [18, 48, 73, 95] and many others. This line of research has some advantages: e.g. in order to represent the forward price it is sufficient to take the discounted expected value of the spot under any equivalent probability measure, called the *pricing measure*. In this section we see a one-dimensional example of spot price dynamics which gives forward prices of type (2.16). We start with the original Lucia-Schwartz model and then present a generalization based on a finer representation of the volatility term structure.

The two-factor Lucia-Schwartz [95] model consists of two state variables and a seasonal component. If $S(t)$ denotes the spot price at time t, then $S(t) = s(t) + X_1(t) + X_2(t)$, with $s(t)$ seasonal deterministic component and

$$dX_1(t) = -\kappa X_1(t)\,dt + v_1\,dW_1^{\mathbb{Q}}(t), \tag{2.61}$$

$$dX_2(t) = \mu\,dt + v_2\,dW_2^{\mathbb{Q}}(t), \tag{2.62}$$

where $W_1^{\mathbb{Q}}$ and $W_2^{\mathbb{Q}}$ are two Brownian motions under a pricing measure $\mathbb{Q}$. Then, we have, for a time t and a maturity date $T \geq t$,

$$S(T) = s(T) + e^{-\kappa(T-t)}X_1(t) + v_1\int_t^T e^{-\kappa(u-t)}dW_1^{\mathbb{Q}}(u) + X_2(t) + \mu(T-t) + v_2(W_2^{\mathbb{Q}}(T) - W_2^{\mathbb{Q}}(t)). \tag{2.63}$$

Denoting $X(t) = (X_1(t), X_2(t))^\intercal$, the value of a forward contract at time t with delivery at time T is given by

$$f(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T)|\mathcal{F}_t] = \alpha(t,T)X(t) + \beta(t,T),$$

with

$$\alpha(t,T) = (e^{-\kappa(T-t)}, 1), \qquad \beta(t,T) = s(T) + \mu(T-t).$$

By comparing these identities to (2.48) and (2.49), we see that this corresponds in our model to

$$\gamma_0(t) \equiv \gamma_0 := \begin{pmatrix} 0 \\ \mu \end{pmatrix}, \qquad \gamma_1(t) \equiv \gamma_1 := \begin{pmatrix} -\kappa & 0 \\ 0 & 0 \end{pmatrix}.$$

The relations among the $\mathbb{P}$-dynamics of X in (2.40) and the Girsanov kernel of $\mathbb{Q}$ are embedded in (2.50) and (2.51). In this model, since

$$v(t) \equiv v := \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix},$$

we have that

$$v\phi_1(t) = \gamma_1 - \Theta(t), \tag{2.64}$$

$$v\phi_0(t) = \gamma_0 - \gamma_1\theta(t), \tag{2.65}$$

where $\phi_1(t), \Theta(t)$ are 2×2 matrices and $\phi_0(t), \theta(t)$ are two-dimensional column vectors. By writing them down explicitly, we have

$$\phi_1(t) = \begin{pmatrix} -\frac{\kappa + \Theta_{11}(t)}{v_1} & -\frac{\Theta_{12}(t)}{v_1} \\ \\ -\frac{\Theta_{21}(t)}{v_2} & -\frac{\Theta_{22}(t)}{v_2} \end{pmatrix}, \qquad \phi_0(t) = \begin{pmatrix} \kappa\,\theta_1(t)/v_1 \\ \mu/v_2 \end{pmatrix}.$$

Then, the coefficients of the $\mathbb{P}$-dynamics of $f(t,T)$ are determined by (2.21)–(2.23), which yield

$$\begin{aligned} \lambda(t)\alpha(t,T) &= \alpha(t,T)(\gamma_1(t) - \Theta(t)) = \alpha(t,T)v(t)\phi_1(t), \\ c(t,T) &= \lambda(t)\beta(t,T) + \alpha(t,T)(\theta(t) - \gamma_0(t)), \\ \sigma(t,T) &= \alpha(t,T)v(t). \end{aligned}$$

We immediately observe that $\alpha(t,T)$ admits a right inverse for each T and there are infinitely many $\lambda(t)$ satisfying the first relation. For instance, we could take (by assuming that $\Theta_{12} = \Theta_{21} = 0$)

$$\begin{aligned}
\lambda(t) &= -\kappa - \Theta_{11}(t),\\
c(t,T) &= \lambda(t)(s(T) + \mu(T-t)) + \theta_1(t)e^{-\kappa(T-t)} + \theta_2(t) - \mu,\\
\sigma(t,T) &= (e^{-\kappa(T-t)}v_1, v_2).
\end{aligned}$$

In view of this analysis, we now present a two-factor model which can be seen as a generalization of [95], where we add a term structure for the volatility. An empirical study of this model has been performed in a companion paper by [93].

By reformulating (2.61) and (2.62), we define

$$dX_1(t) = -\kappa X_1(t)\,dt + v_1\,dW_1^{\mathbb{Q}}(t), \tag{2.66}$$

$$dX_2(t) = \frac{v_2'(t)}{v_2(t)}X_2(t)\,dt + v_2(t)\,dW_2^{\mathbb{Q}}(t), \tag{2.67}$$

where $v_2(t)$ is now a differentiable function of time such that $v_2(t) > 0$ for all t and that $\frac{v_2'(t)}{v_2(t)} \neq -\kappa$ (otherwise, the model collapses into a 1-factor model). In other words, X_1 is a mean-reverting Ornstein-Uhlenbeck process, while $X_2(t)/v_2(t)$ can be shown to be a Brownian motion. While in the Lucia-Schwartz model X_2 is a Brownian motion with drift as v_2 is constant, here the variance of X_2 is varying in time, which allows to have seasonality in the volatility of S as well as in the price level. Thus we can regard v_2 as the seasonality factor for the volatility and s as the seasonality factor for the price. By computations analogous to those before, we derive that

$$S(T) = s(T) + e^{-\kappa(T-t)}X_1(t) + v_1\int_t^T e^{-\kappa(u-t)}dW_1^{\mathbb{Q}}(u) + \frac{v_2(T)}{v_2(t)}X_2(t) + v_2(T)(W_2^{\mathbb{Q}}(T) - W_2^{\mathbb{Q}}(t)).$$

Notice that, within this new formulation, the processes X_1 and X_2 are stationary in mean (i.e., $\mathbb{E}[X_1(t)] = X_1(0)$ and $\mathbb{E}[X_2(t)] = X_2(0)$ for all t). Thus, if we start with $X_1(0) = X_2(0) = 0$, then $s(T)$ is also the expectation of $S(T)$ (under the risk-neutral measure $\mathbb{Q}$).

The value at time t of a forward contract with delivery at time T is

$$f(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T)|\mathcal{F}_t] = \alpha(t,T)X(t) + \beta(t,T),$$

where now

$$\alpha(t,T) = \left(e^{-\kappa(T-t)}, \frac{v_2(T)}{v_2(t)}\right), \qquad \beta(t,T) = s(T).$$

This implies

$$df(t,T) = e^{-\kappa(T-t)}\sigma_1\ dW_1^{\mathbb{Q}}(t) + \sigma_2(T)\ dW_2^{\mathbb{Q}}(t), \tag{2.68}$$

so that, in particular, $\sigma(t,T) = (e^{-\kappa(T-t)}\sigma_1, \sigma_2(T))$. This model allows for the instantaneous forward contracts $f(t,T)$ to have a term structure of the volatility which accounts both for the Samuelson effect (volatility increasing as $t \to T$) in the term $e^{-\kappa(T-t)}\sigma_1$, as well as a potentially complex seasonality with respect to the absolute maturity in the term $\sigma_2(T)$. From this, by specifying a choice of the weight function $\hat{w}$ and using (2.13), we can obtain the dynamics of $F(t,T_1,T_2)$ for all $(t,T_1,T_2) \in \mathcal{A}_2^{\mathbb{T}}$. For example, if $\hat{w} \equiv \frac{1}{T_2-T_1}$, we have that

$$dF(t,T_1,T_2) = -\frac{e^{-\kappa(T_2-t)} - e^{-\kappa(T_1-t)}}{\kappa(T_2-T_1)}\sigma_1\ dW_1^{\mathbb{Q}}(t) + \frac{\int_{T_1}^{T_2}\sigma_2(T)\ dT}{T_2-T_1}\ dW_2^{\mathbb{Q}}(t).$$

If we want that, under the empirical measure, f follows a mean-reverting process as in (2.2), we have to relate the coefficients of X in its $\mathbb{P}$-dynamics (2.40) to (2.21)–(2.23). We already know that, both under $\mathbb{P}$ as under $\mathbb{Q}$, we have

$$v(t) = \begin{pmatrix} v_1 & 0 \\ 0 & v_2(t) \end{pmatrix}.$$

Similarly to what we did before for the Lucia-Schwartz model, we can make a parsimonious choice and put $\theta(t) \equiv 0$ and $\Theta_{12}(t) = \Theta_{21}(t) \equiv 0$. This gives us

$$c(t,T) = \lambda(t)s(T)$$

and

$$\Theta(t) = \begin{pmatrix} \Theta_{11}(t) & 0 \\ 0 & \Theta_{22}(t) \end{pmatrix} = \begin{pmatrix} -\kappa - \lambda(t) & 0 \\ 0 & \frac{v_2'(t)}{v_2(t)} - \lambda(t) \end{pmatrix}.$$

Regarding the market price of risk we observe that

$$\gamma_0(t) \equiv 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \gamma_1(t) = \begin{pmatrix} -\kappa & 0 \\ 0 & \frac{v_2'(t)}{v_2(t)} \end{pmatrix}$$

and

$$v(t)\phi_1(t) = \gamma_1(t) - \Theta(t) = \lambda(t)I_2,$$

where I_2 is the 2×2 identity matrix. This, together with (2.51), gives

$$\phi_1(t) = \left(\frac{\lambda(t)}{v_1}, \frac{\lambda(t)}{v_2(t)} \right), \qquad \phi_0(t) \equiv 0.$$

We now follow Theorem 2.3.6 to determine if the process ϕ obtained above gives a martingale. By the sufficient conditions there, if we impose that Θ is bounded, i.e. that λ and $\frac{v_2'}{v_2}$ are bounded, and that ϕ_1 is bounded, i.e. that $\frac{\lambda}{v_2}$ is bounded, then Z is a martingale, and $\mathbb{Q}$ is an equivalent martingale measure. Thus, the $\mathbb{Q}$-dynamics in (2.68) corresponds to a $\mathbb{P}$-dynamics

$$df(t,T) = \lambda(t)(s(T) - f(t,T))\, dt + e^{-\kappa(T-t)}\sigma_1\, dW_1(t) + \sigma_2(T)\, dW_2(t) \tag{2.69}$$

and, if for example $\hat{w} \equiv \frac{1}{T_2 - T_1}$, we also have that

$$dF(t,T_1,T_2) = \lambda(t) \left(\frac{\int_{T_1}^{T_2} s(T)\, dT}{T_2 - T_1} - F(t,T_1,T_2) \right) dt \tag{2.70}$$

$$- \frac{e^{-\kappa(T_2-t)} - e^{-\kappa(T_1-t)}}{\kappa(T_2 - T_1)}\sigma_1\, dW_1(t) + \frac{\int_{T_1}^{T_2} \sigma_2(T)\, dT}{T_2 - T_1}\, dW_2(t). \tag{2.71}$$

From this model, we can recover various stylized facts typical of energy markets.

- The dynamics of forward prices $f(\cdot,T)$ under the real-world probability measure $\mathbb{P}$ is mean-reverting. The mean-reversion speed $\lambda(t)$ can be time-dependent (but not maturity-dependent in this formulation), and the long-term mean $s(T)$ is exactly the seasonal component of the spot price S. The same can be said about the swap dynamics $F(\cdot,T_1,T_2)$, where this time the long-term mean is the maturity-average of the long-term mean $s(T)$, $T \in [T_1,T_2]$.

- We have a generalized Samuelson effect in the forward prices $f(\cdot, T)$, which is quite evident, and also in the swap prices $F(\cdot, T_1, T_2)$. About this latter, denote the diffusion vector of $F(t, T_1, T_2)$ as $\Sigma(t, T_1, T_2) = (e^{\kappa t}\Gamma(T_1, T_2), \Psi(T_1, T_2))$, with

$$\begin{aligned}\Gamma(T_1, T_2) &:= \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} \sigma_1 e^{-\kappa u}\, du = \frac{\sigma_1(e^{-\kappa T_1} - e^{-\kappa T_2})}{\kappa(T_2 - T_1)},\\ \Psi(T_1, T_2) &:= \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} \sigma_2(u)\, du.\end{aligned}$$

  then the function $t \to \|\Sigma(t, T_1, T_2)\|$ is increasing in t: as time to maturity decreases, the volatility increases.

- Swap prices with shorter delivery periods are more volatile than swap prices with longer delivery periods. In fact, $|\Gamma(T_1, T_2)|$ is decreasing in the second variable. This means that, for all $T_1 < T_2 < T_3$, we have $|\Gamma(T_1, T_2)| > |\Gamma(T_1, T_3)|$. For instance, being equal the time to maturity, monthly contracts are more volatile than quarters (lasting 3 months) or calendars (lasting 1 year). This is in line with empirical findings, see e.g. [23], where the authors perform an empirical analysis of electricity contracts traded on Nord Pool.

In conclusion, here we have shown that a natural extension of Lucia and Schwartz could be to allow for a specific time-dependent speed of mean-reversion in the second factor, such that $X_2/v_2(t)$ is a Brownian motion, as well as for a time-dependent speed of mean reversion of forward prices. Also, we worked out the ingredients for this model in our framework, so that the derived dynamics is arbitrage free.

2.5 A two-commodity cointegrated market

Cointegration is a well-known concept in econometrics and indicates a phenomenon observed in several energy-related markets (see, for instance, [4, 55]). Let us introduce an arbitrage-free cross-commodity model with mean-reversion, which accounts for cointegrated price movements. We present it in the case of two commodities for the sake of simplicity, as generalizations are straightforward. We are inspired by the spot price model presented in [16]. After specifying the spot price dynamics under a martingale measure $\mathbb{Q}$, we will derive forward prices by the conditional expectation. The focus in this example is to explore the possibility to have mean-reverting cointegration, which will be shown to lead to some interesting model restrictions.

Let $S_k(t)$ denote the spot price at time t of commodity k for $k = 1, 2$ and set

$$S_k(t) = s_k(t) + Y_k(t) + a_k L(t), \tag{2.72}$$

with $s_k(t)$ seasonal deterministic component, $a_k \in \mathbb{R} \setminus \{0\}$ for $k = 1, 2$. We assume that, under a martingale measure $\mathbb{Q}$, the dynamics of these factors are

$$\begin{aligned}dL(t) &= \sigma dW(t) + \psi dJ(t),\\ dY_1(t) &= -\mu_1 Y_1(t)\, dt + \sigma_1\, dW_1(t) + \psi_1 dJ_1(t),\\ dY_2(t) &= -\mu_2 Y_2(t)\, dt + \sigma_2\, dW_2(t) + \psi_2 dJ_2(t),\end{aligned}$$

where $\sigma, \psi, \mu_k, \sigma_k, \psi_k \in \mathbb{R}$ for $k = 1, 2$. The process (W, W_1, W_2) is a three-dimensional Brownian motion independent from (J, J_1, J_2), which is a three-dimensional zero-mean pure-jump Lévy process with independent components. The Lévy measures associated are denoted by ν, ν_1 and ν_2 and satisfy the integrability assumptions in Theorem 2.3.5. Since $S_2(t)/a_2 - S_1(t)/a_1 = s_2(t)/a_2 - s_1(t)/a_1 + Y_2(t)/a_2 - Y_1(t)/a_1$, we say that S_1 and S_2 are cointegrated around the seasonality function $s_2(t)/a_2 - s_1(t)/a_1$ (cf. [16]).

Consistently with the notation in Section 3, we have that $\kappa = \int_{\mathbb{R}} y^2\, \nu(dy)$ and $\kappa_i = \int_{\mathbb{R}} y^2\, \nu_i(dy)$, $i = 1, 2$, and

$$v(t) \equiv v := \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & \sigma_2 \end{pmatrix}, \quad z(t) \equiv z := \begin{pmatrix} \psi & 0 & 0 \\ 0 & \psi_1 & 0 \\ 0 & 0 & \psi_2 \end{pmatrix}, \quad K = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa_1 & 0 \\ 0 & 0 & \kappa_2 \end{pmatrix}.$$

Then, the value of a forward contract for the commodity k at time t with delivery at time T is given by the conditional expectation

$$f_k(t, T) = \mathbb{E}_{\mathbb{Q}}[S_k(T)|\mathcal{F}_t] = s_k(T) + e^{-\mu_k(T-t)} Y_k(t) + a_k L(t).$$

We denote $S(t) = (S_1(t), S_2(t))^\intercal$, $f(t,T) = (f_1(t,T), f_2(t,T))^\intercal$, $X(t) = (L(t), Y_1(t), Y_2(t))^\intercal$, so that

$$f(t, T) = \mathbb{E}_{\mathbb{Q}}[S(T)|\mathcal{F}_t] = \alpha(t, T)X(t) + \beta(t, T),$$

with

$$\alpha(t, T) = \begin{pmatrix} a_1 & e^{-\mu_1(T-t)} & 0 \\ a_2 & 0 & e^{-\mu_2(T-t)} \end{pmatrix}, \qquad \beta(t, T) = \begin{pmatrix} s_1(T) \\ s_2(T) \end{pmatrix}. \tag{2.73}$$

Besides, by comparing (2.73) to (2.48) and (2.49), we have

$$\gamma_0(t) \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \gamma_1(t) \equiv \gamma_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\mu_1 & 0 \\ 0 & 0 & -\mu_2 \end{pmatrix}.$$

Following Proposition 2.3.4 and Theorem 2.3.5, the martingale measures $\mathbb{Q}$ are determined by the kernels $\phi(t) = \phi_1(t)X(t) + \phi_0(t)$ and $\xi(t) = \xi_0(t)$, where ϕ_1, ϕ_0 and ξ satisfy

$$\begin{aligned} \gamma_1 &= v\phi_1(t) + \Theta(t), \\ 0 &= v\phi_0(t) + zK\xi(t) + \theta(t). \end{aligned}$$

The deterministic functions $\phi_1(t), \Theta(t)$ are 3×3 matrices and $\phi_0(t), \xi(t), \theta(t)$ are three-dimensional column vectors. The coefficients of the $\mathbb{P}$-dynamics of $f(t, T)$ are determined by

$$\begin{aligned} \lambda(t)\alpha(t, T) &= \alpha(t, T)(\gamma_1 - \Theta(t)) = \alpha(t, T)v\phi_1(t), \\ c(t, T) &= \lambda(t)\beta(t, T) + \alpha(t, T)\theta(t) = \lambda(t)\beta(t, T) - \alpha(t, T)(v\phi_0(t) + zK\xi(t)), \\ \sigma(t, T) &= \alpha(t, T)v \\ \psi(t, T) &= \alpha(t, T)z. \end{aligned}$$

In particular, in view of Proposition 2.3.1, we must verify under which conditions it is possible to define a 2×2 matrix

$$\lambda(t) := \begin{pmatrix} \lambda_{11}(t) & \lambda_{12}(t) \\ \lambda_{21}(t) & \lambda_{22}(t) \end{pmatrix},$$

satisfying (CP) (Definition 2.2.2), such that, independently from T,

$$\lambda(t)\alpha(t, T) = \alpha(t, T)M(t), \tag{2.74}$$

where we denote $M(t) = v\phi_1(t)$. From (2.74) we get the following system of equations:

$$\begin{aligned} a_1\lambda_{11}(t) + a_2\lambda_{12}(t) &= a_1 M_{11}(t) + e^{-\mu_1(T-t)}M_{21}(t), \\ e^{-\mu_1(T-t)}\lambda_{11}(t) &= a_1 M_{12}(t) + e^{-\mu_1(T-t)}M_{22}(t), \\ e^{-\mu_2(T-t)}\lambda_{12}(t) &= a_1 M_{13}(t) + e^{-\mu_1(T-t)}M_{23}(t), \\ a_1\lambda_{21}(t) + a_2\lambda_{22}(t) &= a_2 M_{11}(t) + e^{-\mu_2(T-t)}M_{31}(t), \\ e^{-\mu_1(T-t)}\lambda_{21}(t) &= a_2 M_{12}(t) + e^{-\mu_2(T-t)}M_{32}(t), \\ e^{-\mu_2(T-t)}\lambda_{22}(t) &= a_2 M_{13}(t) + e^{-\mu_2(T-t)}M_{33}(t), \end{aligned}$$

which admits a unique solution if and only if

$$\mu_1 = \mu_2 \tag{2.75}$$

and

$$\lambda_{11}(t) + \frac{a_2}{a_1}\lambda_{12}(t) = \frac{a_1}{a_2}\lambda_{21}(t) + \lambda_{22}(t). \tag{2.76}$$

Furthermore, for the market price of risk parameters we have

$$\phi_1(t) = \begin{pmatrix} \frac{a_1\lambda_{11}(t)+a_2\lambda_{12}(t)}{a_1 v} & 0 & 0 \\ 0 & \frac{\lambda_{11}(t)}{v_1} & \frac{\lambda_{12}(t)}{v_1} \\ 0 & \frac{\lambda_{21}(t)}{v_2} & \frac{\lambda_{22}(t)}{v_2} \end{pmatrix},$$

while $\phi_0(t)$ and $\xi(t)$ must satisfy

$$v\phi_0(t) + zK\xi(t) = -\theta(t).$$

Consequently, if we start with a spot dynamics as in (2.72), with $\mu := \mu_1 = \mu_2$ and a matrix λ satisfying the condition in (2.76) and (CP), then we can build the following mean-reverting, arbitrage-free, cointegrated $\mathbb{P}$-dynamics for the two forwards:

$$\begin{aligned} df_1(t,T) &= (c_1(t,T) - \lambda_{11}(t)f_1(t,T) - \lambda_{12}(t)f_2(t,T))\,dt \\ &\quad + a_1\sigma dW(t) + a_1\psi dJ(t) + e^{-\mu(T-t)}\sigma_1 dW_1(t) + e^{-\mu(T-t)}\psi_1 dJ_1(t), \\ df_2(t,T) &= (c_2(t,T) - \lambda_{21}(t)f_1(t,T) - \lambda_{22}(t)f_2(t,T))\,dt \\ &\quad + a_2\sigma dW(t) + a_2\psi dJ(t) + e^{-\mu(T-t)}\sigma_2 dW_2(t) + e^{-\mu(T-t)}\psi_2 dJ_2(t). \end{aligned}$$

These features are naturally inherited by the swap contracts $F(t,T_1,T_2)$:

$$\begin{aligned} dF_1(t,T_1,T_2) &= (C_1(t,T_1,T_2) - \lambda_{11}(t)F_1(t,T_1,T_2) - \lambda_{12}(t)F_2(t,T_1,T_2))\,dt \\ &\quad + a_1\sigma dW(t) + \frac{\sigma_1(e^{-\mu T_1} - e^{-\mu T_2})}{\mu(T_2 - T_1)}dW_1(t) + \frac{\psi_1(e^{-\mu T_1} - e^{-\mu T_2})}{\mu(T_2 - T_1)}dJ_1(t), \\ dF_2(t,T_1,T_2) &= (C_2(t,T_1,T_2) - \lambda_{21}(t)F_1(t,T_1,T_2) - \lambda_{22}(t)F_2(t,T_1,T_2))\,dt \\ &\quad + a_2\sigma dW(t) + \frac{\sigma_2(e^{-\mu T_1} - e^{-\mu T_2})}{\mu(T_2 - T_1)}dW_2(t) + \frac{\psi_2(e^{-\mu T_1} - e^{-\mu T_2})}{\mu(T_2 - T_1)}dJ_2(t), \end{aligned}$$

where we made the choice $\hat{w} \equiv \frac{1}{T_2 - T_1} I_2$.

Remark 2.5.1. The structural Equation (2.76) for the mean-reversion coefficients λ_{ij} allows to have much flexibility for modeling. For example, in such a setting, one could model the price of the first commodity with a Markovian dynamics (i.e. $\lambda_{12}(t) \equiv 0$) and preserve in the drift of commodity 2 the dependence on commodity 1. To cite an example of application, this can reproduce the behavior of oil (commodity 1) and gas (commodity 2) prices.

This model shows how to incorporate cointegration into an arbitrage-free, mean-reverting forward market with two commodities. It is interesting to observe that the no-arbitrage constraints imply a structural condition on the shape of the volatility term structure (2.75) and a linear relation among the mean-reversion coefficients (2.76). We furthermore remark that a similar approach allows to design more flexible market models, as well as accounting for more than two commodities, still preserving the mean-reverting and arbitrage-free traits.

2.6 Conclusions

By adapting the Heath-Jarrow-Morton idea to energy forward curves in additive models, we have introduced an arbitrage-free framework capable of producing flexible and tractable market models, which exhibit mean-reversion in the dynamics of forward prices under the real-world probability measure.

Our main assumption on the functional form of the forward price processes has allowed to solve several issues. The fundamental requirement for no-arbitrage is the existence of an equivalent martingale measure. Generally, finding it is not a trivial task, since valid measure changes must be independent of the delivery parameters. Mean-reversion naturally requires affine Girsanov kernels, which are in particular stochastic and unbounded. We have been able to validate under minimal assumptions rather general density processes, with stochastic kernels for the Brownian components and Esscher-type factors for the pure-jump part. In addition, swaps and forwards have to satisfy an integral relation, which in general leads to losing analytical tractability, and in particular the Markov property. We succeed in preserving them by introducing simple relations among the coefficients of the dynamics.

Passing to the applications, first we have shown that the well-known Lucia-Schwarz model was already included in the class of models that we here characterized. We have also presented an extension of it, capable to model seasonality in forward's volatilities, besides the price seasonality component already present in the original model.

We then presented a multidimensional market model, which has enabled us in particular to reproduce cointegration effects. As the energy-related markets are strongly interconnected for physical reasons, the possibility of modeling dependence relations more sophisticated than correlation is particularly important for our application purposes.

Looking ahead to future research, since the additive dynamics produce tractable price processes, we believe to have opened the way for analytical formulas for complex derivatives, multicommodity portfolios and risk measures.

2.7 Appendix

2.7.1 Proof of Theorem 2.3.5

Part of this proof will be presented in a simpler version than the one developed in an earlier draft of this work, thanks to the useful suggestions of an anonymous referee. Introduce the sequence of stopping times $\tau_n := \inf\{t \geq 0 : \|X(t)\| \geq n\}$. Set $f(y) := (1+y)\log(1+y) - y$ and define the process:

$$B(t) := \frac{1}{2}\int_0^t \|\phi(s)\|^2\, ds + \int_0^t f(\xi(s)^\intercal y)\, \nu(dy),$$

which is the predictable compensator of

$$\frac{1}{2}\langle H^c, H^c\rangle + \sum_{t\le\cdot} f(\Delta H(t)).$$

In particular, observe that, for every $n \in \mathbb{N}$, the stopped process $B^{\tau_n}(t) := B(t\wedge\tau_n)$ is bounded. From [94, Theorem III.1] it follows that Z^{τ_n} is a uniformly integrable martingale. Define the probability measure $\mathbb{Q}^n$ by setting

$$\frac{d\mathbb{Q}^n}{d\mathbb{P}} := Z^{\tau_n}(\mathbb{T}).$$

Then, we have

$$\begin{aligned}\mathbb{E}[Z(\mathbb{T})] &\ge \mathbb{E}[Z(\mathbb{T})\mathbb{1}_{\tau_n>\mathbb{T}}] = \mathbb{E}[Z^{\tau_n}(\mathbb{T})\mathbb{1}_{\tau_n>\mathbb{T}}] = \mathbb{Q}^n(\tau_n > \mathbb{T})\\ &= 1 - \mathbb{Q}^n\left(\sup_{t\in[0,\mathbb{T}]}\|X(t)\| \ge n\right) \ge 1 - \frac{\mathbb{E}_{\mathbb{Q}^n}\left[\sup_{t\in[0,\mathbb{T}]}\|X(t)\|\right]}{n}.\end{aligned}$$

If we show that

$$\mathbb{E}_{\mathbb{Q}^n}\left[\sup_{t\in[0,\mathbb{T}]}\|X(t)\|\right] \le C, \tag{2.77}$$

for a constant C independent on n, then $\mathbb{E}[Z(\mathbb{T})] \ge 1 - C/n$ for all $n \in \mathbb{N}$, which implies that $\mathbb{E}[Z(\mathbb{T})] = 1$. As a consequence, in order to conclude the proof it is sufficient to verify (2.77).

By Girsanov's Theorem, it holds that

$$\begin{aligned}dX(s) &= \big(\theta(s) + v(s)\phi_0(s)\mathbb{1}_{[0,\tau_n]}(s) + z(s)K\,\xi(s)\mathbb{1}_{[0,\tau_n]}(s) + \Theta(s)X(s)\\ &\quad + v(s)\phi_1(s)X(s)\mathbb{1}_{[0,\tau_n]}(s)\big)\,ds + v(s)dW^n(s) + z(s)\int_{\mathbb{R}^k} y\overline{N}^n(ds,dy),\end{aligned}$$

where, under the measure $\mathbb{Q}^n$, $W^n(s) = W(s) - \int_0^s \mathbb{1}_{[0,\tau_n]}(u)\phi(u)\,du$ is a Brownian motion and $\overline{N}^n(ds,dy) = \overline{N}(ds,dy) - \mathbb{1}_{[0,\tau_n]}(s)\,\xi(s)^\intercal y\,\nu(dy)ds$ is the $\mathbb{Q}^n$-compensated Poisson random measure of N (see [104, Theorem 1.35]). Hence, for each $t \in [0,\mathbb{T}]$, (set w.l.o.g. $X(0) = 0$)

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\|X(s)\|^2\right]\\ &\le 4\Bigg(\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s \big(\theta(u) + v(u)\phi_0(u)\mathbb{1}_{[0,\tau_n]}(u) + z(u)K\,\xi(u)\mathbb{1}_{[0,\tau_n]}(u)\big)\,du\right\|^2\right]\\ &+ \mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s \big(\Theta(u) + \mathbb{1}_{[0,\tau_n]}(u)v(u)\phi_1(u)\big)X(u)\,du\right\|^2\right]\\ &+ \mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s v(u)dW^n(u)\right\|^2\right] + \mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s\int_{\mathbb{R}^k} z(u)\,y\,\overline{N}^n(du,dy)\right\|^2\right]\Bigg).\end{aligned}$$

Firstly, by integrability assumptions on the coefficients, we have

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s \big(\theta(u)+v(u)\phi_0(u)\mathbb{1}_{[0,\tau_n]}(u)+z(u)K\,\xi(u)\mathbb{1}_{[0,\tau_n]}(u)\big)\,du\right\|^2\right]\\
&+\,\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s \big(\Theta(u)+\mathbb{1}_{[0,\tau_n]}(u)v(u)\phi_1(u)\big)X(u)\,du\right\|^2\right]\\
&\le \mathbb{T}\,\mathbb{E}^{\mathbb{Q}^n}\left[\int_0^{\mathbb{T}}\left\|\theta(u)+v(u)\phi_0(u)\mathbb{1}_{[0,\tau_n]}(u)+z(u)K\,\xi(u)\mathbb{1}_{[0,\tau_n]}(u)\right\|^2\,du\right]\\
&+\mathbb{T}\,\mathbb{E}^{\mathbb{Q}^n}\left[\int_0^t\left\|\Theta(u)+\mathbb{1}_{[0,\tau_n]}(u)v(u)\phi_1(u)\right\|^2\|X(u)\|^2\,du\right]\\
&\le C_1+C_2\int_0^t \mathbb{E}^{\mathbb{Q}^n}\left[\sup_{v\in[0,u]}\|X(v)\|^2\right]du,
\end{aligned}
$$

for some constants C_1, C_2 independent of n. For the martingale part, we apply Doob's $\mathbb{L}^2$-inequalities (e.g. [50, Theorem 5.1.3]). As v is square-integrable, the first term yields

$$
\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s v(u)dW^n(u)\right\|^2\right]\le \mathbb{T}\,\mathbb{E}^{\mathbb{Q}^n}\left[\int_0^{\mathbb{T}}\|v(s)\|^2ds\right]=\mathbb{T}\,C_3.
$$

For the second term, since the Lévy measure ν admits fourth moment, ξ is bounded and z is square-integrable, we have that

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\left\|\int_0^s\int_{\mathbb{R}^k} z(u)\,y\,\overline{N}^n(du,dy)\right\|^2\right]\\
&\le \mathbb{T}\,\mathbb{E}^{\mathbb{Q}^n}\left[\int_0^{\mathbb{T}}\int_{\mathbb{R}^k}\left\|z(s)\,y+\mathbb{1}_{[0,\tau_n]}(s)\,z(s)\,y\,\xi(s)^{\intercal}y\right\|^2\,\nu(dy)ds\right]\\
&\le \mathbb{T}\,C_4\left(\int_{\mathbb{R}^k}\|y\|^2\nu(dy)+\int_{\mathbb{R}^k}\|y\|^4\nu(dy)\right),
\end{aligned}
$$

with C_3, C_4 independent of n. Finally, all the bounds together give

$$
\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{s\in[0,t]}\|X(s)\|^2\right]\le C_5+C_6\int_0^t\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{v\in[0,u]}\|X(v)\|^2\right]du,
$$

again with constants independent of n. We conclude by Gronwall's lemma that

$$
\mathbb{E}^{\mathbb{Q}^n}\left[\sup_{t\in[0,\mathbb{T}]}\|X(t)\|^2\right]\le C_5\,e^{C_6\mathbb{T}},
$$

which implies (2.77). This concludes the proof.

2.7.2 Proof of Theorem 2.3.6

The proof is based on the same idea of Proposition 5.1 in [27], where a weak Novikov-type condition is applied to the series representation of the exponential. In view of [87, Corollary 5.14] it is sufficient to prove that there exists an increasing sequence of positive real numbers $(t_k)_{k\in\mathbb{N}}$ diverging to $+\infty$ such that, for all $k\in\mathbb{N}$,

$$
\mathbb{E}\left[\exp\left(\frac{1}{2}\int_{t_k}^{t_{k+1}}\|\phi(t)\|^2dt\right)\right]<\infty, \tag{2.78}
$$

where

$$\phi(t) = \phi_1(t)X(t) + \phi_0(t)$$

and

$$dX(t) = (\theta(t) + \Theta(t)X(t))dt + v(t)dW(t). \tag{2.79}$$

Since Θ satisfies (CP) (see Definition 2.2.2), we can write the unique solution of (2.79) as

$$X(t) = e^{\int_0^t \Theta(s)ds}X(0) + \int_0^t e^{\int_s^t \Theta(u)du}\theta(s)ds + \int_0^t e^{\int_s^t \Theta(u)du}v(s)dW(s).$$

Now, we start to estimate the quantity in (2.78) by

$$\frac{1}{2}\|\phi(t)\|^2 \le \|\phi_1(t)X(t)\|^2 + \|\phi_0(t)\|^2.$$

The second term is deterministic, so as to prove (2.78) we can neglect it. Hence, we move to consider the first one. Let us denote $U(t) := e^{\int_0^t \Theta(s)ds}$. Then,

$$\begin{aligned}
\|\phi_1(t)X(t)\|^2 \le{}& 3\|\phi_1(t)U(t)X(0)\|^2 + 3\int_0^t \|\phi_1(t)U(t)U(s)^{-1}\theta(s)\|^2 ds \\
&+ 3\left\|\int_0^t \phi_1(t)U(t)U(s)^{-1}v(s)dW(s)\right\|^2 \\
&\le 3\|\phi_1(t)U(t)\|^2\left(\|X(0)\|^2 + \int_0^t \|U(s)^{-1}\theta(s)\|^2 ds + \left\|\int_0^t U(s)^{-1}v(s)dW(s)\right\|^2\right).
\end{aligned}$$

As ϕ_1 and Θ are bounded, we are essentially left with

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_{t_k}^{t_{k+1}} \|\phi(t)\|^2 dt\right)\right] \le C_1\mathbb{E}\left[\exp\left(C_2\int_{t_k}^{t_{k+1}} \left\|\int_0^t U(s)^{-1}v(s)dW(s)\right\|^2 dt\right)\right],$$

where C_1, C_2 are constants.

Let us introduce the matrix-valued function $g(s) := U(s)^{-1}v(s)$. Now, by Lebesgue's theorem

$$\mathbb{E}\left[\exp\left(C_2\int_{t_k}^{t_{k+1}} \left\|\int_0^t g(s)dW(s)\right\|^2 dt\right)\right] = \sum_{n=0}^{\infty}\frac{C_2^n}{n!}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} \left\|\int_0^t g(s)dW(s)\right\|^2 dt\right)^n\right]. \tag{2.80}$$

Next, from Jensen's inequality,

$$\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} \left\|\int_0^t g(s)dW(s)\right\|^2 dt\right)^n\right] \le (t_{k+1} - t_k)^{n-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left[\left\|\int_0^t g(s)dW(s)\right\|^{2n}\right] dt.$$

Notice that, since $\int_0^t \|g(s)\|^2 ds < \infty$, the random variable $\int_0^t g(s)dW(s)$ is equal in law to $\left(\sqrt{\int_0^t \|g(s)\|^2 ds}\right) Z$, where Z is a standard normal vector. Therefore, since g is square-integrable,

$$\mathbb{E}\left[\left\|\int_0^t g(s)dW(s)\right\|^{2n}\right] = \left(\int_0^t \|g(s)\|^2 ds\right)^n \mathbb{E}(\|Z\|^{2n}) \le C_3^n\,\mathbb{E}(\|Z\|^{2n}),$$

for a constant C_3.

Finally, exploiting the last two estimates, we find that (2.80) is bounded by

$$\mathbb{E}\left[\exp\left(C_2\int_{t_k}^{t_{k+1}}\left\|\int_0^t g(s)dW(s)\right\|^2 dt\right)\right] \leq \sum_{n=0}^{\infty}\frac{(C_2C_3)^n(t_{k+1}-t_k)^n}{n!}\mathbb{E}(\|Z\|^{2n}).$$

A well-known property of Gaussian moments is the following

$$\mathbb{E}(\|Z\|^{2n}) \leq C_4 n\mathbb{E}(\|Z\|^{2(n-1)}).$$

By applying the ratio test for series we get that, if we choose $(t_k)_k$ such that $C_2C_3C_4(t_{k+1}-t_k) < 1$, then

$$\sum_{n=0}^{\infty}\frac{(C_2C_3C_4)^n(t_{k+1}-t_k)^n}{n!}\mathbb{E}(\|Z\|^{2n}) < \infty,$$

which implies (2.78) and, therefore, the statement.

CHAPTER 3

Mean-reverting no-arbitrage additive models for forward curves in energy markets

In this paper we present an additive no-arbitrage model for energy forward markets capable to exhibit mean-reversion. The model naturally incorporates term structures for both the mean-reversion level and the volatility of forward prices and it is able to reproduce the seasonalities empirically observed in gas and power markets. We also present a method to estimate the model parameters, based on quadratic variation/covariation for the volatility and on constrained maximum-likelihood estimation for the mean-reversion speed and level. We apply this technique to time series of Phelix Base forward products.

3.1 Introduction

Standard models for energy markets usually incorporate mean-reversion in the dynamics of spot [18, 21, 88, 99] and forward [92] prices. Less standard is the requirement that these markets are arbitrage free. This can be relevant especially in natural gas and power markets, which cannot be delivered at a given instant but is rather delivered over a certain time period, e.g. a month (M), a quarter (Q) or a whole calendar year (Cal). In these forward markets, one can find at the same time several forwards covering the same time period (e.g. a quarter covered by the corresponding contract as well as the three corresponding monthly contracts). However, these contracts are not all present in the markets at the same time. In fact, usually only the shortest-maturity months are present, while the other months are still packed in quarterly or calendar contracts. As time passes, these long contracts are unpacked in a cascade mechanism as in Figure 3.1.

In these situations, one must seek models which avoid the possibility of arbitrages. More in detail, at each given time t the following relations must hold:

- the value of a forward contract, which gives the obligation to buy a quantity of energy at the marginal price K for a period $[T_1, T_n]$ must be the sum of the values of forward contracts with the same price K for the subperiods $[T_1, T_2]$, $[T_2, T_3]$, ..., $[T_{n-1}, T_n]$;
- the **forward price**, i.e. the price $K = F(t, T_1, T_n)$ which makes the price of the above contract zero, must be such that

$$F(t, T_1, T_n) = \frac{1}{T_n - T_1} \sum_{i=1}^{n-1} (T_{i+1} - T_i) F(t, T_i, T_{i+1}). \tag{3.1}$$

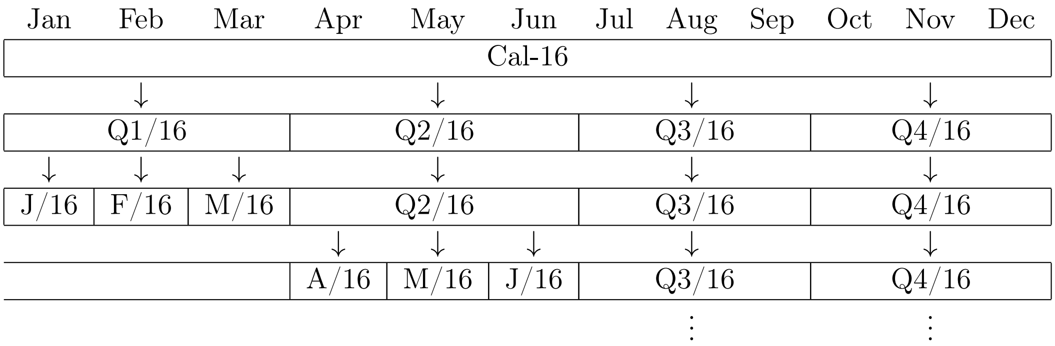


Figure 3.1: Cascade unpacking mechanism of forward contracts. For each given calendar year, as time passes forwards are unpacked first in quarters, then in the corresponding months. It may happen that the same delivery period is covered in the market by different contracts, e.g. one simultaneously finds quotes for Jan/16, Feb/16, Mar/16 and Q1/16.

Since forward contracts are the most liquid products traded in energy markets, in organized markets usually forward prices are quoted and traded continuously. In this situation, one must also avoid models which allow for dynamic arbitrages. This latter (dynamic) no-arbitrage assumption is ensured by the Heath-Jarrow-Morton (HJM) paradigm, which prescribes that in diffusion models the drift and the diffusion coefficients of forward prices satisfy certain relations. These depend on the dynamics of theoretical objects called "instantaneous forward contracts", which are usually not present in the market if forward contracts which deliver over a period are traded. However, it is difficult to achieve this with multiplicative models, where one models the log-prices being e.g. Gaussian or non-Gaussian mean-reverting processes, in both cases ending up with non-Markovian models (see e.g. [32, Chapter 6]).

For these reasons, we choose to work with an arithmetic model, i.e. we model directly the forward prices as Gaussian mean-reverting processes, in line with a growing recent literature [18, 26, 57, 65, 88]. The mathematical convenience is that, with this choice, the HJM relations are satisfied in a very natural way. In principle, we expose ourself to the possibility of negative prices, which would seem in contrast with economic intuition. However we argue that, given the increasing presence of negative spot prices in current power markets (see e.g. [57, 65] and the references therein), there is the theoretical possibility of observing a negative price also in some forward contract, though with a very small probability. For this reason, our model is perfectly suited for power forward markets and, for a suitable choice of coefficients, also for natural gas forward markets.

The final result is that, in general, a forward contract is mean-reverting with respect to an infinite linear combination of instantaneous forward contracts, thus making the model infinite-dimensional. To solve this conundrum, we make the key assumptions that instantaneous (non-traded) forward prices (and, as a consequence, also the traded ones) depend linearly on a finite number of (hidden) state variables, and that the mean-reversion speed is at each given time the same for all contracts in the market. In order not to have to deal with models with incomplete information, we impose that these hidden factors do not appear explicitly in the forward prices' dynamics. It turns out that this requirement is equivalent to the fact that the coefficients satisfy suitable linear algebraic conditions.

We also show that this class is not void, since the well known Lucia-Schwartz model [95], if we compute the forwards' dynamics, turns out to be of this kind. This model describes the spot price as the sum of two hidden state variables plus a deterministic seasonal component. Also, it is quite flexible in mean price levels, by incorporating the spot's seasonal component, but the volatility succeeds only in reproducing the Samuelson effect as an exponential decay, without being able to reproduce more complex term structures. This fact can be observed for example in Figure 3.2, where the Samuelson effect alone is not sufficient to explain the complex patterns of the realized volatilities of the different forward prices.

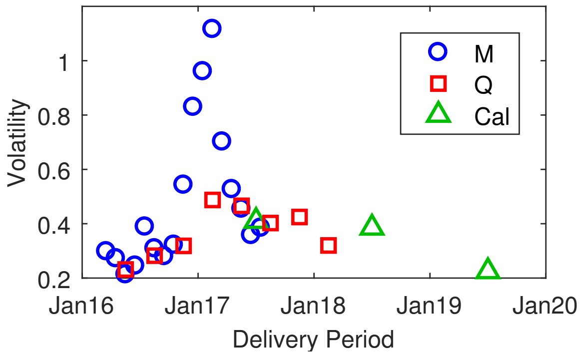


Figure 3.2: Realized volatilities of forward prices present in the Phelix Base market from January 4, 2016 to May 23, 2017. For forward contracts maturing before May, 2017 we have either all the market quotes or the most recent ones. It is clear that the term structure of volatilities cannot be explained simply by the Samuelson effect. More in detail, the snapshot of calendar contracts' volatilities exhibit a decreasing pattern which cannot be modeled by an exponential component, which is the typical way of modeling the Samuelson effect (see [58] and references therein). Also quarterly contracts have a term structure of volatility which present a yearly seasonality, while the monthly term structure volatility exhibits an even more complicated term structure.

To overcome this rigidity, we propose a modification of the Lucia-Schwartz model which is free of arbitrages in the way discussed above, where the prices of forward contracts are mean-reverting and where both price level and volatility are allowed to have a non-trivial term structure. As in the Lucia-Schwartz model, the mean-reversion level of forward contracts corresponds to the seasonal component of spot prices, and the prices are moved by two Brownian motions, possibly correlated. However, while one Brownian component takes into account the short-term movements and reproduces a Samuelson-like term structure, which decays exponentially with time to maturity, the other Brownian component is able to model more complex term structures of volatility with respect to absolute maturities. The model also predicts a well-observed phenomenon, i.e. the fact that contracts with shorter delivery periods (e.g. months) are more volatile than contracts with longer delivery periods (e.g. quarters or calendars).

We then present a technique for estimating the parameters of this model based on market time series of forward prices, which are the most liquid market quotes. For the volatility, this task is not trivial: in fact, if we postulate that the term structure of volatility is also maturity-dependent as above, we cannot employ the well-known technique of rolling time series to estimate volatilities of contracts having the same relative maturity (see e.g. [5]). Instead, we use a technique already present in [58] based on quadratic variation/covariation of forward time series, and in this way we estimate the volatility term structure. After having estimated the volatility coefficients, we pass to the mean-reversion speed and levels in a non-parametric way, so that we can replicate any functional behavior of the forward long-term mean, which takes into account the no-arbitrage requirements among forwards covering the same time period.

We apply this estimation technique on a time series for the Phelix Base forward market from January 4, 2016 to May 23, 2017, considering each monthly, quarterly and calendar forward contract traded in that time window (we do not consider shorter period contracts or semesters). Since our model is able to reproduce complex volatility shapes, we first try with a popular choice for seasonal component, i.e. a linear combination of trigonometric functions (see e.g. [32, Chapter 8.6]) plus a linear component in maturity, which takes into account a possible non-seasonal long-run volatility trend across the calendar years. We then compare it with a fully non-parametric volatility shape, where the second factor is a free parameter for each atomic forward. After having estimated the volatility, we pass to estimate the mean-reversion speed and level. More in detail, we assume that this latter quantity depends on the single contract, but that overall no-arbitrage HJM relations must hold among contracts spanning the same period (e.g. 3 months / 1 quarter). This is done by using maximum likelihood estimation (MLE) with the technique of Lagrange multipliers to take into account no-arbitrage constraints.

The paper is structured as follows. In Section 2 we present an overview of models for energy forward curves based on the HJM paradigm. In Section 3 we show how mean-reversion for forward contracts can be introduced into the HJM paradigm, with the Lucia-Schwartz model as an example. In Section 4 we present our generalization of the Lucia-Schwartz model, and describe stylized facts which are reproduced by our model. In Section 5 we present an estimation technique which can be used in our framework, and is applied in Section 6 to the Phelix Base power market. Section 7 concludes.

3.2 HJM-based models

Since in energy markets many forward contracts for a single commodity are traded at the same time, it is important to ensure that the price comovements of these contracts do not allow for arbitrages. A classical methodology to ensure that this is not possible has been, for interest rates models used in fixed income markets, the one introduced by Heath, Jarrow and Morton (HJM) in [76]. Since the energy forward markets share many similarities with fixed income markets, the idea of imposing no-arbitrage in forward markets has been implemented using a HJM methodology, for the first time in [23]. Here we informally recall this approach, sending the reader interested in technical details (assumptions on required integrability of involved processes, etc.) to that paper.

Assume that a "instantaneous forward price" exists on the market for each maturity $T > t$, whose evolution under the real-world probability $\mathbb{P}$ is given by

$$df(t,T) = \mu(t,T)\ dt + \sum_{i=1}^{d} \theta_i(t,T)\ dW_i(t), \tag{3.2}$$

where $W = (W_0, \dots, W_d)$ is a d-dimensional $\mathbb{P}$-Brownian motion. These contracts may or may not exist on the market if the futures with delivery over intervals are traded, this being a characteristics of each commodity. For instance, the power futures markets trade contracts with "thin" granularity, i.e. daily, so that one can interpret them as instantaneous delivery contracts. For natural gas, instead, usually the shortest delivery period available on the market is monthly. Nevertheless, when formulating our model, we are introducing instantaneous forwards as theoretical building blocks for traded contracts. This will be more clear in the following, when we will introduce the other class of contracts.

The assumption that no arbitrages are present in the market is equivalent to the existence of a martingale measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$df(t,T) = \sum_{i=0}^{d} \theta_i(t,T)\ dW_i^{\mathbb{Q}}(t)$$

under $\mathbb{Q}$, where $W^{\mathbb{Q}}$ is a d-dimensional $\mathbb{Q}$-Brownian motion, linked to W by the so-called **market price of risk** Λ, such that

$$dW^{\mathbb{Q}}(t) = dW(t) + \Lambda(t)\ dt.$$

The spot price is $S(t) = f(t,t)$, we have that

$$dS(t) = \partial_2 f(t,t)\ dt + \sum_{i=1}^{d} \theta_i(t,t)\ dW_i^{\mathbb{Q}}(t) \tag{3.3}$$

under $\mathbb{Q}$, and that $f(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T) \mid \mathcal{F}_t]$. Notice that, in electricity and gas markets, the spot price S is typically not traded, even if a common proxy for the spot price in power market is assumed to be the day-ahead price of electricity. The day-ahead price is however settled above forward contracts with hourly granularity, which again is not a spot price in the mathematical sense above.

Under this model, the forward price is given by[1]

$$F(t,T_1,T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t,u)\ du.$$

This naturally implies the no-arbitrage additive relation (3.1) among forward contracts. As discussed earlier, in this definition we are introducing the contracts f as building blocks for the contracts F, in order to satisfy the no-arbitrage constraints in a natural way. Also, we have the dynamics

$$dF(t,T_1,T_2) = \Sigma(t,T_1,T_2)\ dW^{\mathbb{Q}}(t)$$

under $\mathbb{Q}$, with

$$\Sigma(t,T_1,T_2) = \frac{\int_{T_1}^{T_2} \theta(t,u)\ du}{T_2 - T_1}$$

and

$$dF(t,T_1,T_2) = \frac{\int_{T_1}^{T_2} \mu(t,u)\ du}{T_2 - T_1}\ dt + \Sigma(t,T_1,T_2)\ dW(t) \tag{3.4}$$

under $\mathbb{P}$.

[1] More generally, one has

$$F(t,T_1,T_2) = \int_{T_1}^{T_2} \hat{w}(u,T_1,T_2) f(t,u)\ du$$

where $\hat{w}$ is a weight function which represents the time value of money [23], depending on how settlement takes place. To fix the ideas, we assume the simplest case, i.e. $\hat{w}(u,T_1,T_2) = \frac{1}{T_2 - T_1}$, and we just remark that different settlement rules, i.e. different choices of $\hat{w}$, would imply a different no-arbitrage condition in Equation (3.1).

3.3 Mean reversion consistent with no-arbitrage

In many popular models, under the real-world probability measure $\mathbb{P}$, energy prices are either exponentials of Gaussian mean-reverting processes [21, 48, 73] or additive mean-reverting processes driven by Gaussian [57] or non-Gaussian [18, 21, 22, 95] noise. This is usually done on spot prices, which however are not traded liquidly in energy markets, the most liquid contracts traded being forwards of the form $f(\cdot, T)$ or of the form $F(\cdot, T_1, T_2)$, depending on the particular commodity.

Our aim here is to formulate mean-reversion models both for $f(t, T)$ and $F(\cdot, T_1, T_2)$ for every maturity period $[T_1, T_2]$, which are Markovian with respect to the same state variable, i.e. respectively with the dynamics

$$\begin{aligned} df(t,T) &= (-\lambda(t,T)f(t,T) + c(t,T))\ dt + \theta(t,T)\ dW(t), \\ dF(t,T_1,T_2) &= (-\lambda(t,T_1,T_2)F(t,T_1,T_2) + C(t,T_1,T_2))\ dt + \Sigma(t,T_1,T_2)\ dW(t), \end{aligned}$$

in such a way that they are also consistent with the no-arbitrage relation in Equation (3.1).

This is not trivial at all. In fact, given the dynamics in Equation (3.4), we should have

$$-\lambda(t,T_1,T_2)F(t,T_1,T_2) = -\frac{\int_{T_1}^{T_2} \lambda(t,u) f(t,u)\ du}{T_2 - T_1}.$$

Should this not happen, we would have

$$dF(t,T_1,T_2) = \left(-\frac{\int_{T_1}^{T_2} \lambda(t,u) f(t,u)\ du}{T_2 - T_1} + C(t,T_1,T_2) \right)\ dt + \Sigma(t,T_1,T_2)\ dW(t),$$

which is a relation which involves the infinite number of factors $f(\cdot, T)$, $T \in [T_1, T_2]$. This is similar to what happens with geometric models [32], where even the volatility of $F(t, T_1, T_2)$ under the risk-neutral measure depends on all the contracts $F(t, T_1, u)$ with $u \in [T_1, T_2]$. Here we do not have this complication for the risk neutral dynamics, but for the real world dynamics. The problem is that the dynamics above is non-Markovian in $F(t, T_1, T_2)$, and is time-consuming to simulate due to the dependence on all forward contracts with shorter delivery periods, and hard to use to estimate market parameters. In principle, one could consider it as a Markovian evolution in infinite dimensions, which would require more advanced techniques even for the well-posedness of the dynamics above (see e.g. [?]). For the reasons explained above, we avoid seeking the utmost generality and consider, with the next assumption, a subclass of models which will allow for a much more efficient modeling of forward prices.

We make this key assumption.

Assumption 9. *Assume that*

$$f(t,T) = \alpha(t,T)X(t) + \beta(t,T),$$

where X is a vector of state variables and α, β are suitable deterministic functions.

This assumption implies that also $F(t, T_1, T_2)$ are functions of the same state variables X:

$$F(t,T_1,T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t,u)\ du = \frac{\int_{T_1}^{T_2} \alpha(t,u)\ du}{T_2 - T_1} X(t) + \frac{\int_{T_1}^{T_2} \beta(t,u)\ du}{T_2 - T_1}. \tag{3.5}$$

It is thus natural to assume here that the market price of risk Λ is given by

$$\Lambda(t) = A(t)X(t) + b(t), \tag{3.6}$$

with A, b deterministic. Since it must hold $\mu(t,T) = \theta(t,T)\Lambda(t)$ by Equation (3.2), we impose

$$-\lambda(t,T)f(t,T) + c(t,T) = \theta(t,T)(A(t)X(t) + b(t)).$$

This implies the relation

$$-\lambda(t,T)(\alpha(t,T)X(t) + \beta(t,T)) + c(t,T) = \theta(t,T)(A(t)X(t) + b(t)).$$

By equating the terms with respect to $X(t)$, we obtain

$$\begin{cases} -\lambda(t,T)\alpha(t,T) = \theta(t,T)A(t), \\ c(t,T) = \theta(t,T)b(t) + \lambda(t,T)\beta(t,T). \end{cases}$$

Thus, apart from seasonalities in t, the mean-reversion coefficient of the forward contract maturing in T is proportional to its volatility coefficient. This effect propagates to contracts $F(t,T_1,T_2)$: in fact, the drift in Equation (3.4) is

$$\begin{aligned} \frac{\int_{T_1}^{T_2} \mu(t,u)\,du}{T_2 - T_1} &= \frac{\int_{T_1}^{T_2} \theta(t,u)\Lambda(t)\,du}{T_2 - T_1} = \frac{\int_{T_1}^{T_2} \theta(t,u)(A(t)X(t) + b(t))\,du}{T_2 - T_1} dt = \\ &= -\frac{\int_{T_1}^{T_2} \lambda(t,u)\alpha(t,u)\,du}{T_2 - T_1} X(t) + \frac{\int_{T_1}^{T_2} -\lambda(t,T)\beta(t,T) + c(t,T)\,du}{T_2 - T_1}. \end{aligned}$$

Unfortunately, if $\lambda(t,T)$ depends on T then we cannot simplify furtherly. Instead, if we postulate that λ does not depend on T, then we obtain

$$\begin{aligned} \frac{\int_{T_1}^{T_2} \mu(t,u)\,du}{T_2 - T_1} &= -\lambda(t)\frac{\int_{T_1}^{T_2} \alpha(t,u)\,du}{T_2 - T_1} X(t) + \frac{\int_{T_1}^{T_2} -\lambda(t)\beta(t,T) + c(t,T)\,du}{T_2 - T_1} \\ &= -\lambda(t)\left(F(t,T_1,T_2) - \frac{\int_{T_1}^{T_2} \beta(t,u)\,du}{T_2 - T_1}\right) + \\ &\quad + \frac{\int_{T_1}^{T_2} -\lambda(t)\beta(t,u) + c(t,u)\,du}{T_2 - T_1} \\ &= -\lambda(t)F(t,T_1,T_2) + C(t,T_1,T_2) \end{aligned}$$

with

$$\begin{cases} \lambda(t)\alpha(t,T) = \theta(t,T)A(t), \\ \\ C(t,T_1,T_2) = \dfrac{\int_{T_1}^{T_2} c(t,u)\,du}{T_2 - T_1}. \end{cases}$$

To conclude, let us remark that first we have described a framework where a forward $F(\cdot,T_1,T_2)$ is written as an infinite linear combination of instantaneous forwards $f(\cdot,T)$. Later, in order to simplify a highly intractable probabilistic setting, we have introduced a modeling assumption on the functional form of $f(t,T)$. On the technical side, we are working under the assumptions that the stochastic differential equations admit a unique solution for each maturity and delivery period. When imposing the affine dependence of the instantaneous forwards on the underlying process X, we need similar assumptions for its dynamics. Moreover, as we have presented above, certain relations must hold among the coefficients of all the equations. It can be shown that one of the mathematical characterizations is that X is a mean-reverting process. We do not want to go in deep into this kind of discussion in this paper, since a general rigorous study of this framework and, in particular, of the link between the integral relation and the affine assumption on f is detailed in [26]. On the other side, looking from a modeling perspective, we first assume to have a stochastic description of a market made of any forward contract (in the sense of delivery period), by imposing an integral relation among them by arbitrage arguments. Then, we focus on a subfamily of models, described by a tractable representation and where the coefficients are assumed to satisfy simple linear relations.

The upcoming example is built from the Lucia-Schwartz model [95] and generates forward prices satisfying the affine representation mentioned above.

Example 3.3.1 (Lucia-Schwartz model). Assume that $S(t) = \phi(t) + X_1(t) + X_2(t)$, with ϕ deterministic seasonal component and

$$\begin{aligned} dX_1(t) &= -\kappa X_1(t)\, dt + \sigma_1\, dW_1^{\mathbb{Q}}(t), \\ dX_2(t) &= \mu\, dt + \sigma_2\, dW_2^{\mathbb{Q}}(t), \end{aligned}$$

where $W_1^{\mathbb{Q}}$ and $W_2^{\mathbb{Q}}$ are two Brownian motions under $\mathbb{Q}$, with possible mutual correlation ρ. Then

$$\begin{aligned} S(T) &= \phi(T) + e^{-\kappa(T-t)} X_1(t) + \sigma_1 \int_t^T e^{-\kappa(u-t)}\, dW_1^{\mathbb{Q}}(u) + \\ &\quad + X_2(t) + \mu(T-t) + \sigma_2 (W_2^{\mathbb{Q}}(T) - W_2^{\mathbb{Q}}(t)) \end{aligned}$$

and

$$F(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T) \mid \mathcal{F}_t] = \alpha(t,T) X(t) + \beta(t,T)$$

with

$$\alpha(t,T) := (e^{-\kappa(T-t)}, 1), \qquad \beta(t,T) := \phi(T) + \mu(T-t).$$

Also,

$$df(t,T) = e^{-\kappa(T-t)} \sigma_1\, dW_1^{\mathbb{Q}}(t) + \sigma_2\, dW_2^{\mathbb{Q}}(t),$$

so that $\theta(t,T) = (e^{-\kappa(T-t)}\sigma_1, \sigma_2)$. This means that the dynamics of traded forwards $F(\cdot, T_1, T_2)$ are

$$dF(t, T_1, T_2) = -\frac{e^{-\kappa(T_2-t)} - e^{-\kappa(T_1-t)}}{\kappa(T_2 - T_1)} \sigma_1\, dW_1^{\mathbb{Q}}(t) + \sigma_2\, dW_2^{\mathbb{Q}}(t).$$

Now, if we want a mean-reversion coefficient $\lambda(t)$ for the forward under $\mathbb{P}$, we must impose

$$-\lambda(t)(e^{-\kappa(T-t)}, 1) = (e^{-\kappa(T-t)}\sigma_1, \sigma_2) A(t) = (e^{-\kappa(T-t)}, 1) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} A(t)$$

Since $(e^{-\kappa(T-t)}, 1)$ is the only thing depending on T, we must require that

$$-\lambda(t) Id_2 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} A(t),$$

where Id_2 is the 2×2 identity matrix. Finally

$$A(t) := -\lambda(t) \begin{pmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \end{pmatrix}$$

and

$$\begin{aligned} c(t,T) &= \theta(t,T)b(t) + \lambda(t)\beta(t,T) = \\ &= (e^{-\kappa(T-t)}\sigma_1, \sigma_2)b(t) + \lambda(t)(\phi(T) + \mu(T-t)). \end{aligned}$$

Thus, under $\mathbb{P}$ we obtain the dynamics

$$\begin{aligned} df(t,T) &= [\lambda(t)(\phi(T) + \mu(T-t) - f(t,T)) + (e^{-\kappa(T-t)}\sigma_1, \sigma_2)b(t)]\ dt + \\ &\quad + e^{-\kappa(T-t)}\sigma_1\ dW_1(t) + \sigma_2\ dW_2(t), \end{aligned}$$

from which we get that $f(t,T)$ has a reversion of its mean towards $\phi(T) + \mu(T-t)$, plus a correction term due to b. If we impose that μ and b are zero, we get the much simpler dynamics

$$df(t,T) = \lambda(t)(\phi(T) - f(t,T))\ dt + e^{-\kappa(T-t)}\sigma_1\ dW_1(t) + \sigma_2\ dW_2(t),$$

where the long-term mean of $f(t,T)$ is exactly the seasonal component $\phi(T)$. Thus, the function $\phi(T)$ can be seen as a seasonal component both for the (future) spot price $S(T)$ and for the forward price $f(t,T)$.

3.4 A two-factor model

Now we present a generalization of the Lucia-Schwartz model with two factors. Assume that $S(t) = \phi(t) + X_1(t) + X_2(t)$, with ϕ deterministic seasonal component and with $\mathbb{Q}$-dynamics for $X = (X_1, X_2)$ given by

$$dX_1(t) = -\kappa\, X_1(t)\ dt + \sigma_1\ dW_1^{\mathbb{Q}}(t), \tag{3.7}$$

$$dX_2(t) = \frac{\psi'(t)}{\psi(t)} X_2(t)\ dt + \psi(t)\ dW_2^{\mathbb{Q}}(t), \tag{3.8}$$

with $\kappa > 0$, $\sigma_1 > 0$ and ψ deterministic function of time such that

$$\frac{\psi'(t)}{\psi(t)} \neq -\kappa \tag{3.9}$$

(otherwise, the model collapses into a 1-factor model). These two state variables can be explicitly represented in terms of their initial conditions as

$$\begin{aligned} X_1(t) &= e^{-\kappa t}X_1(0) + \sigma_1 \int_0^t e^{-\kappa(t-s)} dW_1^{\mathbb{Q}}(s), \\ X_2(t) &= \psi(t)\left(\frac{X_2(0)}{\psi(0)} + W_2^{\mathbb{Q}}(t)\right). \end{aligned}$$

In other words, X_1 is a mean-reverting Ornstein-Uhlenbeck process, and $\frac{X_2}{\psi}$ is a Brownian motion. The idea here is to have a mean-reverting component X_1 for short periods and a non-stationary component X_2 for longer periods, possibly correlated with X_1. While in the Lucia-Schwartz model ψ is constant, so that X_2 is essentially a Brownian motion, here its variance is time-varying, which allows to have seasonality in the volatility as well as in the price level. Thus, we can regard ψ as the seasonality factor for the volatility, while ϕ is the seasonality factor for the price.

Then the spot price is given by

$$\begin{aligned} S(T) &= \phi(T) + e^{-\kappa(T-t)} X_1(t) + \sigma_1 \int_t^T e^{-\kappa(u-t)}\, dW_1^{\mathbb{Q}}(u) + \\ &\quad + \frac{\psi(T)}{\psi(t)} X_2(t) + \psi(T)(W_2^{\mathbb{Q}}(T) - W_2^{\mathbb{Q}}(t)) \end{aligned}$$

and the instantaneous forward price $f(t,T)$ can be computed as

$$f(t,T) = \mathbb{E}_{\mathbb{Q}}[S(T) \mid \mathcal{F}_t] = \alpha(t,T)X(t) + \beta(t,T),$$

with

$$\alpha(t,T) := \left(e^{-\kappa(T-t)}, \frac{\psi(T)}{\psi(t)}\right), \qquad \beta(t,T) := \phi(T).$$

The dynamics of $f(t,T)$ under the risk-neutral measure $\mathbb{Q}$ are

$$df(t,T) = e^{-\kappa(T-t)}\sigma_1\, dW_1^{\mathbb{Q}}(t) + \psi(T)\, dW_2^{\mathbb{Q}}(t),$$

thus $\theta(t,T) = (e^{-\kappa(T-t)}\sigma_1, \psi(T))$ in the notation of Sections 2–3.

This model allows for the instantaneous forward contracts $f(t,T)$ to have a term structure of the volatility which accounts both for the Samuelson effect (volatility increasing as $t \to T$) in the term $e^{-\kappa(T-t)}\sigma_1$, as well as a potentially complex seasonality with respect to the absolute maturity in the term $\psi(T)$. A popular choice is to represent the seasonality function ψ as a truncated Fourier series with a drift term

$$\psi(T) := \sigma_2 + \mu T + \sum_{j=1}^{a}\big(\alpha_{2j}\cos(\omega jT) + \alpha_{2j+1}\sin(\omega jT)\big), \tag{3.10}$$

with $\omega = 2\pi/365$, $a \in \mathbb{N}$, $\mu \in \mathbb{R}$, $(\sigma_2, \alpha_2, \ldots, \alpha_{2a+1}) \in \mathbb{R}^{2a+1}$.

Now we want to choose the market price of risk Λ as in Equation (3.6) in order to have a mean-reversion dynamics for the forward prices. We also impose $b \equiv 0$ in order to avoid overparameterizations in the dynamics (recall that we already have the seasonal component ϕ in the spot price). If we want a mean-reversion coefficient $\lambda(t)$, we must impose

$$-\lambda(t)\left(e^{-\kappa(T-t)}, \frac{\psi(T)}{\psi(t)}\right) = \left(e^{-\kappa(T-t)}, \frac{\psi(T)}{\psi(t)}\right)\begin{pmatrix} \sigma_1 & 0 \\ 0 & \psi(t) \end{pmatrix} A(t).$$

Since this relation must hold for all $T > t$ and the components of $(e^{-\kappa(T-t)}, \frac{\psi(T)}{\psi(t)})$ are linearly independent by Equation (3.9), we must impose that

$$-\lambda(t) Id_2 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \psi(t) \end{pmatrix} A(t)$$

where Id_2 is the 2×2 identity matrix, and finally

$$A(t) := -\lambda(t)\begin{pmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\psi(t) \end{pmatrix}.$$

With this specification for the market price of risk, the dynamics of X under $\mathbb{P}$ turn out to be (recall that $b \equiv 0$):

$$\begin{aligned} dX_1(t) &= -(\lambda(t) + \kappa)\, X_1(t)\, dt + \sigma_1\, dW_1(t), \\ dX_2(t) &= -\left(\lambda(t) + \frac{\psi'(t)}{\psi(t)}\right) X_2(t)\, dt + \psi(t)\, dW_2(t) \end{aligned}$$

and also

$$c(t,T) = \lambda(t)\beta(t,T) = \lambda(t)\phi(T).$$

Finally,

$$C(t,T_1,T_2) = \frac{\int_{T_1}^{T_2} c(t,u)\,du}{T_2-T_1} = \lambda(t)\frac{\int_{T_1}^{T_2}\phi(u)\,du}{T_2-T_1}$$

so that the dynamics of $f(t,T)$ under the real-world probability measure $\mathbb{P}$ is

$$df(t,T) = \lambda(t)(\phi(T) - f(t,T))\,dt + \theta(t,T)\,dW(t)$$

and the dynamics of $F(t,T_1,T_2)$ is

$$dF(t,T_1,T_2) = \lambda(t)\left(\frac{\int_{T_1}^{T_2}\phi(u)\,du}{T_2-T_1} - F(t,T_1,T_2)\right)\,dt + \Sigma(t,T_1,T_2)\,dW(t).$$

Here, the global diffusion vector of $F(t,T_1,T_2)$ is $\Sigma(t,T_1,T_2) = (e^{\kappa t}\Gamma(T_1,T_2), \Psi(T_1,T_2))$, with

$$\Gamma(T_1,T_2) := \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\sigma_1 e^{-\kappa u}\,du = \frac{\sigma_1(e^{-\kappa T_1} - e^{-\kappa T_2})}{\kappa(T_2-T_1)}, \tag{3.11}$$

$$\Psi(T_1,T_2) := \frac{1}{T_2-T_1}\int_{T_1}^{T_2}\psi(u)\,du. \tag{3.12}$$

If, for example, ψ is as in Equation (3.10), then

$$\Psi(T_1,T_2) = \sigma_2 + \mu\,(T_1+T_2) + \frac{1}{T_2-T_1}\sum_{j=1}^{a}\Bigg(\frac{\alpha_{2j}}{\omega j}(\sin(\omega j T_2) - \sin(\omega j T_1)) - \frac{\alpha_{2j+1}}{\omega j}(\cos(\omega j T_2) - \cos(\omega j T_1))\Bigg). \tag{3.13}$$

From this model, we can recover various stylized facts typical of energy markets.

Remark 3.4.1. The mean-reversion speed $\lambda(t)$ in the dynamics of forward prices $f(t,T)$ under the real-world probability measure $\mathbb{P}$ can be time-dependent (but not maturity-dependent in this formulation), and the long-term mean $\phi(T)$ is exactly the seasonal component of the spot price S. The same can be said about the dynamics of $F(\cdot,T_1,T_2)$, where this time the long-term mean is the maturity-average of the long-term mean $\phi(T)$, $T \in [T_1,T_2]$.

Remark 3.4.2. We have a generalized Samuelson effect in the instantaneous (non-traded) forward prices $f(\cdot,T)$, which is quite evident, and also in the traded forward prices $F(\cdot,T_1,T_2)$. About this latter, denoting the diffusion vector of $F(t,T_1,T_2)$ as $\Sigma(t,T_1,T_2) = (e^{\kappa t}\Gamma(T_1,T_2), \Psi(T_1,T_2))$ as above, then the function $t \to \|\Sigma(t,T_1,T_2)\|$ is increasing in t: as time to maturity decreases, the volatility increases.

Remark 3.4.3. This model postulates that prices are regulated by two hidden factors, X_1 and X_2. This is in line with many models present in literature (e.g. [18, 21, 23, 48, 58, 89, 95, 99, 120]), and the common interpretation is that the first factor (here X_1) is stationary mean-reverting and models the short-term movements of the price curve, while the second factor (here X_2) is nonstationary and models the long-term movements. This dynamics are usually assumed under the pricing measure $\mathbb{Q}$, under which our factor X_1 is a stationary

Ornstein-Uhlenbeck process, while X_2 is a Brownian motion scaled by the function ψ. However, the common use of these models in the literature above is that these two factors stay hidden and are not identified with e.g. macroeconomic variables. This use is robust especially when the factors X_1 and X_2 do not appear explicitly in the dynamics of the observed prices $F(t, T_1, T_2)$, as is also the case here. The main novelty of our model is that this phenomenon is also present in the $\mathbb{P}$-dynamics, where we succeeded in building a model where the observed prices $F(t, T_1, T_2)$ are mean-reverting and Markov by themselves, and their dynamics do not have an explicit dependence on X_1 and X_2. However, the effects of the interpretation above of X_1 and X_2 can be still traced in the model's behaviour. First, we observe in our model that forward prices with shorter delivery periods are more volatile than forward prices with longer delivery periods. In fact, $|\Gamma(T_1, T_2)|$ is decreasing in the second variable. This means that, for all $T_1 < T_2 < T_3$, we have $|\Gamma(T_1, T_2)| > |\Gamma(T_1, T_3)|$. For instance, being equal the time to maturity, monthly contracts are more volatile than quarters (lasting 3 months) or calendars (lasting 1 year). Furthermore, the second factor $\Psi(T_1, T_2)$ is able to capture both a seasonal pattern and long-run movements. Therefore, the parameter $\Gamma(T_1, T_2)$ incorporates most of the short-term variability of futures' prices, whereas $\Psi(T_1, T_2)$ accounts for longer periods as well. By coming back to the definition of the spot price dynamics in (3.7)-(3.8), we see that this component is inherited by the volatility of the long-term factor of the spot price. This means that the no-arbitrage construction of this model, if applied to both markets, links the volatility of spot and futures in such a way. Analogously, the short-run futures volatility $\Gamma(T_1, T_2)$ is ruled by the parameter κ, which coincides with the mean-reversion speed of the short-term factor of the spot.

In conclusion, we have shown that a natural extension of Lucia and Schwartz could allow for a specific time-dependent speed of mean-reversion in the second factor, such that $\frac{X_2}{\psi}(t)$ is a Brownian motion, as well as for a time-dependent speed of mean-reversion of forward prices. Also, we worked out the ingredients for this model in our framework, so that the derived dynamics is arbitrage free.

3.5 Calibration of a forward market

Assume that we have an energy commodity market (typically, electricity in a given country, but possibly natural gas) and that its instantaneous spot price (possibly non-traded) satisfies a model like the one of Section 4, with (hidden) state variables $X = (X_1, X_2)$, driven by the Brownian motions W_1, W_2, with correlation ρ. Assume also for simplicity that the mean-reversion function $\lambda(t)$ is constant in t. If N forward contracts are traded in the market with maturity intervals $[T_{i,1}, T_{i,2}]$, $i = 1, \dots, N$, respectively, then the corresponding forward prices $F_i(t) := F(t, T_{i,1}, T_{i,2})$ will have the dynamics

$$dF_i(t) = \lambda\big(\Phi_i - F_i(t)\big)dt + \Gamma_i e^{\kappa t} dW_1(t) + \Psi_i dW_2(t) \tag{3.14}$$

with $t < T_{i,1}$, where the quantities $\Phi_i := \Phi(T_{i,1}, T_{i,2})$, $\Gamma_i := \Gamma(T_{i,1}, T_{i,2})$, $\Psi_i := \Psi(T_{i,1}, T_{i,2})$ do not depend on t and are given by

$$\Phi(T_{i,1}, T_{i,2}) = \frac{\int_{T_{i,1}}^{T_{i,2}} \phi(u)\, du}{T_{i,2} - T_{i,1}}, \tag{3.15}$$

$$\Gamma(T_{i,1}, T_{i,2}) = \frac{\sigma_1(e^{-\kappa T_{i,1}} - e^{-\kappa T_{i,2}})}{\kappa(T_{i,2} - T_{i,1})}, \tag{3.16}$$

$$\Psi(T_{i,1}, T_{i,2}) = \frac{\int_{T_{i,1}}^{T_{i,2}} \psi(u)\, du}{T_{i,2} - T_{i,1}}, \tag{3.17}$$

with $\sigma_1 \in \mathbb{R}$, $k \in \mathbb{R}^+$ and some deterministic functions ϕ and ψ, for all $i = 1, \ldots, N$.

In order to calibrate this model to market data, we will follow a procedure already used by one of the authors in [58]. First, we will calibrate the parameters starting from those of the diffusion coefficient (which will be calibrated with realized covariation estimators) and then we will pass to the mean-reversion parameters.

3.5.1 Calibration of the diffusions of the forward contracts

We now start by estimating the diffusion coefficients of all the forward contracts. Two different approaches are studied: in the first, we model ψ by a truncated Fourier series with a linear term as in Equation (3.10), while in the second we do not specify its functional form, thus estimating directly the values $\Psi(T_{i,1}, T_{i,2})$, for $i = 1, \ldots, N$.

Assume to observe N forward prices with delivery periods $[T_{i,1}, T_{i,2}]$, $i = 1, \ldots, N$, in the observation period $[t_0, t]$, with $t_0 \le T_{i,1}$ for all $i = 1, \ldots, N$. Let also $[\tau_{i,1}, \tau_{i,2}]$ be the trading period of F_i in the interval $[t_0, t]$. Then, if $[\tau_{i,1}, \tau_{i,2}] \cap [\tau_{j,1}, \tau_{j,2}] \neq \emptyset$, the maximum interval contained in $[t_0, t]$ where we can define the quadratic covariation between F_i and F_j is $[\tau_{ij,1}, \tau_{ij,2}]$, with

$$\tau_{ij,1} := \max(\tau_{i,1}, \tau_{j,1}), \quad \tau_{ij,2} := \min(\tau_{i,2}, \tau_{j,2}).$$

It is possible to express the quadratic covariations as functions of the model parameters.

Proposition 3.5.1. *Let $i, j \in \{1, \ldots, N\}$ be fixed, then the quadratic covariation of F_i and F_j from $\tau_{ij,1}$ to $\tau_{ij,2}$ is given by*

$$\langle F_i, F_j \rangle_{\tau_{ij,1}}^{\tau_{ij,2}} = A_{ij} + B_{ij} + (C_{ij} + D_{ij})\rho, \tag{3.18}$$

where $\tau_1 := \tau_{ij,1}$, $\tau_2 := \tau_{ij,2}$,

$$A_{ij} = \frac{\Gamma_i \Gamma_j}{2\kappa}(e^{2\kappa\tau_2} - e^{2\kappa\tau_1}), \qquad B_{ij} = \Psi_i \Psi_j (\tau_2 - \tau_1),$$
$$C_{ij} = \frac{\Gamma_i \Psi_j}{\kappa}(e^{\kappa\tau_2} - e^{\kappa\tau_1}), \qquad D_{ij} = \frac{\Psi_i \Gamma_j}{\kappa}(e^{\kappa\tau_2} - e^{\kappa\tau_1}).$$

Proof. The proof is a standard integration exercise, and is thus omitted. □

Now we want to choose the parameters such that the model quadratic covariations above are as near as possible to the realized quadratic covariations. If ψ is specified as in Equation (3.10), then, for all $i = 1, \ldots, N$, Ψ_i depends on the parameters $\mu, \sigma_2, \alpha_2, \ldots, \alpha_{2a+1}$ (cf. Equation (3.13)). Thus, in this case we define the diffusion parameter to be estimated as $p = (\rho, \kappa, \mu, \sigma_1, \sigma_2, \alpha_2, \ldots, \alpha_{2a+1}) \in \mathbb{R}^{2a+5}$ and the constraint set is $\mathcal{P} = [-1, 1] \times \mathbb{R}^+ \times \mathbb{R}^{2a+3}$.

On the other hand, if ψ is not given in parametric form, we directly estimate the values $\Psi_1, \Psi_2, \ldots, \Psi_N$. We also have to take into account the following fact. By Equation (3.12), the parameters Ψ_i satisfy the following constraints

$$\Psi_i = \Psi(T_{i,1}, T_{i,2}) = \sum_{j=1}^{n} \frac{T_{j,2} - T_{j,1}}{T_{i,2} - T_{i,1}} \Psi(T_{j,1}, T_{j,2}) = \sum_{j=1}^{n} \frac{T_{j,2} - T_{j,1}}{T_{i,2} - T_{i,1}} \Psi_j \tag{3.19}$$

whenever $[T_{i,1}, T_{i,2}]$ is the union of disjoint intervals $[T_{j,1}, T_{j,2}]$ for $j = 1, \ldots, n$, i.e. for all the forwards with overlapping delivery. Therefore, we call *atomic*, with the same terminology as [?], the contracts whose delivery period can not be partitioned by the delivery periods of other forwards. In other words, we suppose that m forwards $F_1, \ldots, F_m$ have disjoint delivery periods $[T_{1,1}, T_{1,2}], \ldots, [T_{m,1}, T_{m,2}]$ and such that other forwards' delivery periods can be expressed as union of the former. For instance, suppose to observe in the chosen calibration window the prices of Jan/17, Feb/17, Mar/17, Apr/17, Q1/17, Q2/17, Cal-17. On one hand, Q1/17 is not atomic, since it can be "splitted" into Jan/17, Feb/17 and Mar/17. On the other hand, Q2/17 turns out to be atomic, even if Apr/17 is already traded, as May/17 and Jun/17 are not observed (being still not traded in the market). For the same reason, Cal-17 is considered atomic as well. Then, in this example, if $\Psi_1, \ldots, \Psi_7$ denote the corresponding parameters, we have that $\Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5, \Psi_7$ are free parameters, as they refer to atomic contracts, whereas to determine Ψ_5 we use Equation (3.19). Consequently, in the second approach the vector parameter to estimate is $p = (\rho, \kappa, \sigma_1, \Psi_1, \Psi_2, \ldots, \Psi_N) \in \mathcal{P} = [-1, 1] \times \mathbb{R}^+ \times \mathbb{R}^{N+2}$ subject to the constraints given by Equation (3.19). For example, with the convention that $\Psi_{\mathrm{Q2/17}}$ denotes the parameter Ψ_i corresponding to the contract Q2/17, then

$$\Psi_{\mathrm{Q2/17}} = u_{\mathrm{Apr/17}} \Psi_{\mathrm{Apr/17}} + u_{\mathrm{May/17}} \Psi_{\mathrm{May/17}} + u_{\mathrm{Jun/17}} \Psi_{\mathrm{Jun/17}}, \tag{3.20}$$

where the weights u_i are defined according to the number of days in the month/quarter (e.g. for Apr/17 we have $u_{\mathrm{Apr/17}} = 30/91$).

To this purpose, after having chosen a calibration window $[t_0, t]$ as said before, we assume to observe forward prices at discrete dates $\pi = \{t_0 < t_1 < \ldots < t_n = t\}$. Then, for all forward contracts' couples F_i, F_j, let us now denote with $l_{ij,1}$ and $l_{ij,2}$ the indexes of the partition π such that $t_{l_{ij,1}} = \tau_{ij,1}$ and $t_{l_{ij,2}} = \tau_{ij,2}$, for all $i, j = 1, \ldots, N$. If $[\tau_{i,1}, \tau_{i,2}] \cap [\tau_{j,1}, \tau_{j,2}] = \emptyset$, we let $\tau_{ij,1} = \tau_{ij,2} = l_{ij,1} = l_{ij,2} = 0$. From now on, in order to distinguish the random variables from their realizations we use uppercase and lowercase letters, respectively. In particular, we indicate with F_i the random variables relative to the i-th forward price, and with f_i its realization.

Definition 3.5.2. Let $i, j \in \{1, \ldots, N\}$. Define the *realized quadratic covariation* of the forwards F_i and F_j from $\tau_{ij,1}$ to $\tau_{ij,2}$ as

$$\left[F_i, F_j\right]_{\tau_{ij,1}}^{\tau_{ij,2}} := \sum_{l=l_{ij,1}+1}^{l_{ij,2}} \big(f_i(t_l) - f_i(t_{l-1})\big)\big(f_j(t_l) - f_j(t_{l-1})\big).$$

Define also the *realized quadratic variation* of the forward F_i from $\tau_{i,1}$ to $\tau_{i,2}$ as

$$\left[F_i\right]_{\tau_{i,1}}^{\tau_{i,2}} := \left[F_i, F_i\right]_{\tau_{ii,1}}^{\tau_{ii,2}}.$$

It is a standard result [14] that these estimators are unbiased and consistent provided that the forward contracts' drifts are zero. Here this is not the case, but it is possible to prove, by adapting results in [14], that these estimators are biased (but with the bias depending on the drift only on third order), consistent and asymptotically Gaussian. Thus, the realized quadratic covariations approximate the model quadratic covariations, and this approximation improves as the partition π becomes finer. Ideally, we wish that our model would be such that

$$\big[F_i, F_j\big]_{\tau_{ij,1}}^{\tau_{ij,2}} = \big\langle F_i, F_j\big\rangle_{\tau_{ij,1}}^{\tau_{ij,2}}, \qquad \text{for all } i, j = 1, \dots, N.$$

However, the right-hand sides of this system depend only on the parameters p of the model, which are $2a+5$ in the truncated Fourier series approach and $N+3$ in the general one. Keeping into account the symmetry of quadratic covariations, this system has $N(N+1)/2$ independent equations, and in general will be overdetermined. For this reason, we define a nonlinear least squares estimator as follows.

Definition 3.5.3. Define the estimator $\hat{p}$ as the vector which solves

$$\min_{p \in \mathcal{P}} \sum_{i=1}^{N} \sum_{j=1}^{i} w_{ij} \Big(\big[F_i, F_j\big]_{\tau_{ij,1}}^{\tau_{ij,2}} - \big\langle F_i, F_j\big\rangle_{\tau_{ij,1}}^{\tau_{ij,2}} \Big)^2, \tag{3.21}$$

possibly with some nonnegative weigths w_{ij}.

We observe that the quadratic covariations depend linearly on ρ, bilinearly on $\mu, \sigma_1, \sigma_2, \alpha_2, \dots, \alpha_{2a+1}$ (respectively $\sigma_1, \Phi_1, \dots, \Phi_N$), and in a nonlinear way on κ. Thus, even if least squares are a classical tool in Statistics, it is difficult here to prove properties like unbiasedness and consistency of the estimator $\hat{p}$. On the other hand, in order to calibrate the parameters p, it is not possible to obtain explicit estimators with good properties using standard techniques (like maximum likelihood). Thus, the compromise on the choice of $\hat{p}$ is to use nonlinear least squares without proving whether $\hat{p}$ has or not the classical properties of estimators.

3.5.2 Calibration of the mean-reversion coefficients

Now that we estimated the parameters of the diffusion part, we have to estimate the mean-reversion coefficient λ and the long-term means Φ_i, for $i = 1, \dots, N$, in Equation (3.15).

We start from the dynamics of the i-th forward price F_i under $\mathbb{P}$, as given by Equation (3.14), where Γ_i and Ψ_i have already been estimated in Section 5.1. Under $\mathbb{P}$ the forward price follows an Ornstein-Uhlenbeck process, with explicit solution

$$\begin{aligned} F_i(T) = e^{-\lambda(T-t)} F_i(t) &+ \int_t^T \lambda \Phi_i e^{-\lambda(T-s)}\, ds \\ &+ \int_t^T \Gamma_i e^{\kappa s} e^{-\lambda(T-s)}\, dW_1(s) + \int_t^T \Psi_i e^{-\lambda(T-s)}\, dW_2(s). \end{aligned}$$

Notice that the processes $\boldsymbol{F}(T) := (F_1(T), \cdots, F_N(T))^\intercal$ are not independent among each other, as they are driven by the same two-dimensional Brownian motion. In particular, $\boldsymbol{F}(T)$ conditioned to $\mathcal{F}_t$ is Gaussian:

$$\boldsymbol{F}(T)|\mathcal{F}_t \sim \mathcal{N}\big(e^{-\lambda(T-t)}\boldsymbol{F}(t) + \boldsymbol{\Phi}\big(1 - e^{-\lambda(T-t)}\big), \boldsymbol{\Sigma}(t,T)\big)$$

where $\mathbf{\Phi} := (\Phi_1, \cdots, \Phi_N)^\intercal$ and $\mathbf{\Sigma}(t,T) := \big(\Sigma_{ij}(t,T)\big)_{i,j=1,\ldots,N}$, with

$$\begin{aligned}\Sigma_{ij}(t,T) &:= \int_t^T \big(\Gamma_i\Gamma_j e^{2\kappa s} + \rho e^{\kappa s}(\Gamma_i\Psi_j + \Gamma_j\Psi_i) + \Psi_i\Psi_j\big)e^{-2\lambda(T-s)}\, ds = \\ &= \frac{\Gamma_i\Gamma_j\big(e^{2\kappa T} - e^{2\kappa t - 2\lambda(T-t)}\big)}{2(\kappa+\lambda)} + \frac{(\Gamma_i\Psi_j + \Gamma_j\Psi_i)\big(e^{\kappa T} - e^{\kappa t - 2\lambda(T-t)}\big)\rho}{(\kappa+2\lambda)} \\ &+ \frac{\Psi_i\Psi_j\big(1 - e^{-2\lambda(T-t)}\big)}{2\lambda}\end{aligned}$$

We can then use the following approximation: if $T - t \simeq 0$, then

$$\begin{aligned}\Sigma_{ij}(t,T) \simeq &\frac{\Gamma_i\Gamma_j e^{2\kappa t}\big(1 + 2\kappa(T-t) - 1 - 2\lambda(T-t)\big)}{2(\kappa+\lambda)} + \\ &+ \frac{\rho(\Gamma_i\Psi_j + \Gamma_j\Psi_i)e^{\kappa t}\big(1 + \kappa(T-t) - 1 - 2\lambda(T-t)\big)}{\kappa+2\lambda} \\ &+ \frac{\Psi_i\Psi_j\big(1 - 1 + 2\lambda(T-t)\big)}{2\lambda} \\ &= (T-t)(\Gamma_i\Gamma_j e^{2\kappa t} + (\Gamma_i\Psi_j + \Gamma_j\Psi_i)\rho e^{\kappa t} + \Psi_i\Psi_j).\end{aligned}$$

Thus, at first order, the variance depends only on parameters which are already known, while in the mean the only quantities still not known are $e^{-\lambda(T-t)}$ and $\Phi(T_1, T_2)$.

Let us now assume to have daily observations, at times $\pi = \{t_0 < t_1 < \ldots < t_n = t\}$, with $t_l = t_{l-1} + 1$, $l = 1, \ldots, n$. Then for $t_l \in \pi$, by applying the previous approximation on the variance $\mathbf{\Sigma}$, we obtain

$$\boldsymbol{F}(t_l) - \boldsymbol{F}(t_{l-1})|\mathcal{F}_{t_{l-1}} \sim \mathcal{N}\big(\big(1 - e^{-\lambda}\big)\big(\mathbf{\Phi} - \boldsymbol{F}(t_{l-1})\big), \mathbf{\Sigma}(t_{l-1})\big) \tag{3.22}$$

where $\mathbf{\Sigma}(t) := \big(\Sigma_{ij}(t)\big)_{i,j=1,\ldots,N}$ with

$$\Sigma_{ij}(t) := \Gamma_i\Gamma_j e^{2\kappa t} + (\Gamma_i\Psi_j + \Gamma_j\Psi_i)\rho e^{\kappa t} + \Psi_i\Psi_j.$$

Since, by definition, $\Phi_i = \frac{1}{T_{i,2} - T_{i,1}} \int_{T_{i,1}}^{T_{i,2}} \phi(u)\, du$, we are in the same situation as in the calibration of the volatility component Ψ. Then we have essentially two approaches to perform a parametric estimation of the vector $\mathbf{\Phi}$. The first consists in representing ϕ as an element of an arbitrarily chosen parameterized family of functions. The second consists in regarding $\Phi_1, \ldots, \Phi_N$ as free parameters. Here, we adopt directly the latter, being the most general approach.

Let us denote the parameter vector to estimate by $p := (q, \mathbf{\Phi}) \in (0,1) \times \mathbb{R}_+^N$, with $\mathbf{\Phi} := (\Phi_1, \ldots, \Phi_N)$ and $q := 1 - e^{-\lambda}$. It would be immediate to introduce the maximum (log-)likelihood estimator for p, but we cannot apply it to the multidimensional vector of returns $\boldsymbol{F}(t_l) - \boldsymbol{F}(t_{l-1})$. The problem is that, since the dynamics of all the forward curve is driven by a two-dimensional Brownian motion, the covariance matrix of the returns in Equation (3.22) has rank two, so it is singular as soon as we have more than two contracts. Thus, the returns are normally distributed but do not admit a multivariate density. As a consequence, in order to find p we compute the likelihood estimator for a single-forward time series $f_i(t)$, for $i = 1, \ldots, N$.

Definition 3.5.4. Define the estimator $\hat{p}_i = (\hat{q}_i, \hat{\Phi}_i)$ as the 2-dimensional vector which solves

$$\max_{(q,\Phi)\in(0,1)\times\mathbb{R}_+} \sum_{l=2}^n \log\varphi\big(f_i(t_l) - f_i(t_{l-1}), q\big(\Phi - f_i(t_{l-1})\big), \sigma_i^2(t_{l-1})\big),$$

where $\sigma_i^2(t) := \Sigma_{ii}(t) = \Gamma_i^2 e^{2\kappa t} + 2\Gamma_i \Psi_i \rho e^{\kappa t} + \Psi_i^2$ and $\varphi(x, \mu, \sigma^2)$ is the density of a Gaussian random variable $\mathcal{N}(\mu, \sigma^2)$.

We compute these estimators in Appendix A.1. Observe that we have N different estimators $\hat{q}_1, \ldots, \hat{q}_N$ for the same parameter q (since we are assuming in the model the same mean-reversion speed for all the contracts). Generally, this leads to different statistical estimates for the same parameter q or, equivalently, λ. Therefore, after computing $\hat{q}_1, \ldots, \hat{q}_N$ from Equation (3.29), we define

$$\hat{q} := \sum_{i=1}^{N} w_i \, \hat{q}_i, \tag{3.23}$$

with arbitrary weights w_i, provided that $\hat{q} \in (0, 1)$. Thus, the estimated mean-reversion speed is $\hat{\lambda} = -\log(1 - \hat{q})$.

A naive procedure now would be to define an estimator of Φ_i, for all $i = 1, \ldots, N$, using Equation (3.28), which was derived in Appendix A.1, by replacing q_i with $\hat{q}$. However, as we already discussed for the function Ψ in Section 5.1, the parameters Φ_i satisfy the same no-arbitrage constraints as in Equation (3.19). Therefore, we introduce a new estimator for each Φ_i by redefining properly the one in Definition 3.5.4.

Definition 3.5.5. Let $\hat{q} \in (0, 1)$ represent the estimated reparameterized mean-reversion speed in Equation (3.23). Define the estimator $\tilde{p} = (\tilde{\Phi}_1, \ldots, \tilde{\Phi}_N)$ as the vector which solves

$$\max_{(\Phi_1,\ldots,\Phi_N) \in \mathbb{R}_+^N} \sum_{i=1}^{N} \sum_{l=2}^{n} \log \varphi\big(f_i(t_l) - f_i(t_{l-1}), \hat{q}\big(\Phi_i - f_i(t_{l-1})\big), \sigma_i^2(t_{l-1})\big),$$

subject to the constraints given by the no-arbitrage relation (3.19).

3.6 Empirical analysis

Our database includes closing prices of Phelix Base futures traded at EEX from January 4, 2016 until May 23, 2017. We consider monthly, quarterly and yearly contracts.

3.6.1 Calibration results

We start by estimating the diffusion coefficients, firstly, in the case that Ψ is given in the parametric form of Equation (3.13) and, secondly, with the non-parametric approach (see Section 5.1). We use the estimators introduced in Definition 3.5.3. In order to perform the first calibration, we set the number of seasonal components to $a = 5$, so to capture potential asymmetric seasonal variations. Since contracts with different delivery length contain a different amount of market information, we weight each term of the sum in Equation (3.21) with $w_{ij} = w_i w_j$, where $w_i = 1$ if the contract F_i is monthly, $w_i = 3$ if F_i is a quarter and $w_i = 12$ if it is a calendar. These different weights are motivated by the fact that the forward price for a calendar contract fixes the price for a period which is about 12 times longer than the one of a monthly contract (analogously for a quarterly contract). To estimate the coefficients Ψ_i in the second calibration, we need to add the no-arbitrage constraints described by Equation (3.20). For the calibration of the monthly contracts from Apr/16 to Dec/16 and the corresponding quarters Q2/16, Q3/16 and Q4/16, we have constraints of the type

$$\Psi_4 = u_1 \Psi_1 + u_2 \Psi_2 + u_3 \Psi_3. \tag{3.24}$$

Analogously, for the contracts Jan/17, . . . , Jun/17, Q1/17,. . . , Q4/17, Cal-17 we have

$$\begin{aligned} \Psi_7 &= u_1\Psi_1 + u_2\Psi_2 + u_3\Psi_3, \\ \Psi_8 &= u_4\Psi_4 + u_5\Psi_5 + u_6\Psi_6, \\ \Psi_{11} &= u_7\Psi_7 + u_8\Psi_8 + u_9\Psi_9 + u_{10}\Psi_{10}. \end{aligned} \tag{3.25}$$

Observe that, even if Jul/17 is present in our calibration window, it does not play any role in the last constrained problem, because we are not observing Aug/17 and Sep/17. The parameters for the first calibration are in Table 3.1, while for the second one are in Table 3.2 and Figure 3.3. The norm residual is 99138.69 for the first approach and significantly decreases to 35890.78 with the non-parametric approach.

In both cases we observe that the correlation coefficient ρ among the two Brownian motions driving the dynamics is significantly far from -1 and 1. This means that a single factor would have not been sufficient to describe the variance of prices. Also, for the first calibration, the parameter μ turns out to be slightly negative, meaning that there is an overall decreasing trend in volatility in the time window considered.

κ	σ_1	σ_2	ρ	μ
0.0086	0.7322	0.4357	-0.3634	-0.0100
α_1	α_2	α_3	α_4	α_5
0.0844	0.0727	0.1501	-0.0046	0.1110
α_6	α_7	α_8	α_9	α_{10}
-0.0086	-0.0438	0.0156	-0.1732	-0.0315

Table 3.1: Calibrated diffusion parameters with the truncated Fourier series ($a = 5$).

ρ	κ	σ_1
0.3185	0.0167	0.9793

Table 3.2: Calibrated diffusion parameters in the second calibration.

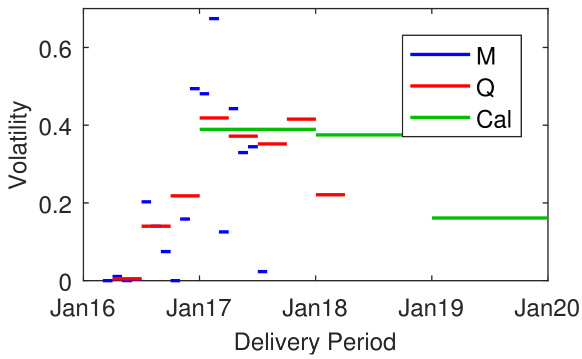


Figure 3.3: Coefficients Ψ_i for the second calibration.

In Figure 3.3 we notice that the coefficient Ψ_i relative to the first contracts take values very close to zero (if not equal). This is in line with what observed from the realized volatilities in Figure 3.2. In particular, in this way we can reproduce the volatility behavior which is not explained by the Samuelson effect. Also, the calendar contracts exhibit a decreasing volatility, while the quarters exhibit, besides this decreasing component over the years, a slight seasonal pattern with the volatility having the highest value in Q1 (cold period), and the lowest in Q3 (warm period): this is an empirical fact, well reproduced by our model. The model also partially reconstructs the finer term structure for monthly contracts, even if for each one of those we have less observations. We find out that the seasonality here is not so apparent, both from the data as well as from our model.

In order to better understand the meaning of the other parameters, we may compare theoretical and historical quadratic covariations. However, they depend on the interval length over which they are computed. For this reason, let us introduce two new quantities: theoretical and historical covolatilities.

Definition 3.6.1. Let $i, j \in \{1, \ldots, N\}$. Denote by $[\tau_{i,1}, \tau_{i,2}]$ the trading time interval of the i-th contract. If $[\tau_{i,1}, \tau_{i,2}] \cap [\tau_{j,1}, \tau_{j,2}] \neq \emptyset$, define the *historical covolatility* of the forwards F_i and F_j as

$$\overline{\Sigma}_{ij} := \frac{1}{\tau_{ij,2} - \tau_{ij,1}} \left[F_i, F_j\right]_{\tau_{ij,1}}^{\tau_{ij,2}}.$$

In analogy, define the *theoretical covolatility* of the forwards F_i and F_j as

$$\Sigma_{ij} := \frac{1}{\tau_{ij,2} - \tau_{ij,1}} \left\langle F_i, F_j\right\rangle_{\tau_{ij,1}}^{\tau_{ij,2}}.$$

If $i = j$, $\overline{\Sigma}_{ij}$ and Σ_{ij} are positive. We call *historical volatility* and *theoretical volatility* of the forward F_i, respectively, the quantities $\overline{\Sigma}_i := \sqrt{\overline{\Sigma}_{ii}}$ and $\Sigma_i := \sqrt{\Sigma_{ii}}$. Since this calibration has been performed by matching not only model and realized variation/volatility of each contract, we represent all their possible historical and theoretical covolatilities. We plot the results of the non-parametric case in Figure 3.4, being the one that performed better in terms of residuals. We observe the peculiar dependence structure of energy forwards, with highly volatile months (especially the winter period) and relatively lower codependence among contracts with distant delivery.

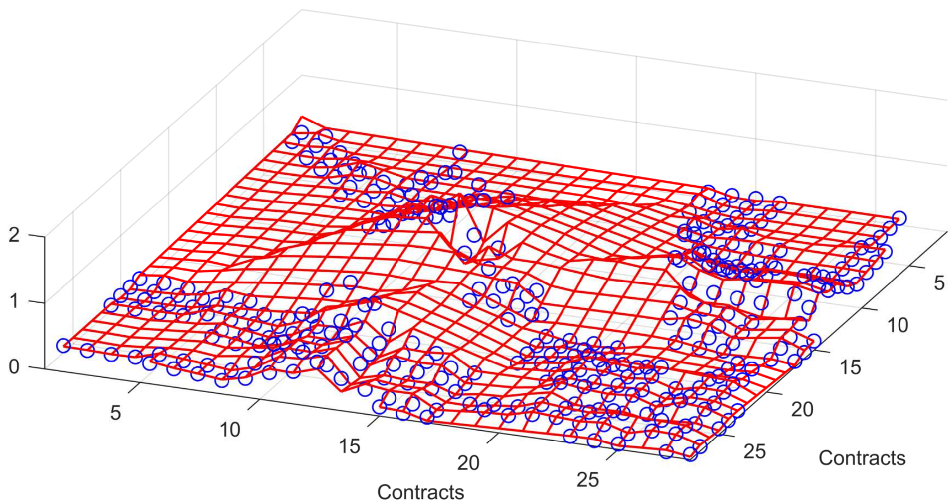


Figure 3.4: Historical and theoretical covolatilities of all the contracts, which are numbered as follows: months from Mar/16 to Jul/17 (1-17), quarters from Q2/16 to Q1/18 (18-25), calendars from Cal-17 to Cal-19 (26-28).

Let us now estimate the drift parameters, i.e. the mean-reversion speed λ and the long-term mean vector $\boldsymbol{\Phi} := (\Phi_1, \ldots, \Phi_N)$. We first compute the estimators $\hat{p}_i = (\hat{q}_i, \hat{\Phi}_i)$ for $i = 1, \ldots, N$, which were introduced in Definition 3.5.4. As we already discussed in Section 5, since we have different estimators we get different estimates for the same parameter λ. Therefore, in Equation (3.23) we choose the weights with the same philosophy of the case of the diffusion coefficients (cf. Section 5.1) and compute $\hat{q} = 0.0683$. This turns out to be consistent with the model, in the sense that $\hat{q} \in (0, 1)$. Then, we define the estimated mean-reversion speed as $\hat{\lambda} = -\log(1 - \hat{q}) = 0.0708$. Let us just naively replace $\hat{\Phi}^i$, for all $i = 1, \ldots, N$, with the solution of Equation (3.28) (see Appendix A.1), with $\hat{q}^i$ replaced by $\hat{q}$. The results are in Figure 3.5.

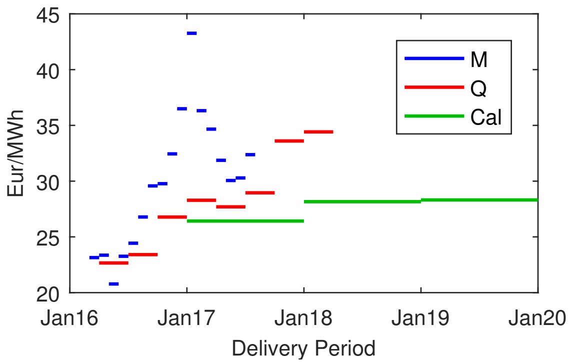


Figure 3.5: Long-term means with $\hat{q}_i \equiv \hat{q}$, but still without taking into account the no-arbitrage relation (3.19).

However, let us recall that the no-arbitrage relations of Equation (3.19) must hold (see Definition 3.5.5). If we compare the estimator $\hat{\Phi}_i$ for Jan/17, Feb/17, Mar/17 and Q1/17, this relation fails. This distortion is solved by adding into the optimization in (3.19) the same costraints we used to perform the calibration of the diffusion coefficients Ψ_i (cf. Equations (3.24) and (3.25)). Here we use the Lagrange multipliers technique to solve these constrained optimization problems (see Appendix A.2). The results are illustrated in Figure 3.6, where we can see that now the no-arbitrage relation (3.19) holds for all the relevant contracts.

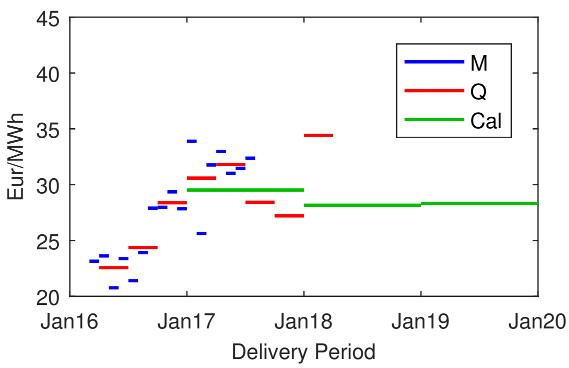


Figure 3.6: Long-term means with $\hat{q}_i \equiv \hat{q}$ and the constraints in Equation (3.19).

3.6.2 Simulation and model assessment

In this section we do a simulation study and assess the performance of the model in terms of reproducibility of both the EEX data under investigation and the stylized features of energy futures markets. Firstly, we compare simulated paths of some exemplary futures contracts to the corresponding observed trajectories, in order to assess the qualitative behavior of model simulations. Then, we compute fundamental statistics of the model by averaging the results of a set of simulations and compare them to our data.

In order to simulate the trajectory of a certain contract, we use the dynamics in Equation (3.14). We define it for given input parameters in MATLAB via the `sdemrd` class, specifically designed for mean-reverting equations. Then, the function `simByEuler` discretizes the SDE and computes the price for each day by following an Euler scheme for stochastic differential equations, with the parameters which have been estimated in Section 5. For all futures we have the same averaged mean-reversion speed $\hat{\lambda}$, while each contract (indexed by i) has its own parameter Ψ_i in the long-term volatility coefficient and $\hat{\Phi}_i$ for the drift component. Since we are interested in the simulation of a single contract, we compute the long-term means from the single-forward time series calibration described in Appendix A.1. Finally, the diffusion parameters are computed with the non-parametric procedure, being this the case that performed better in terms of residuals (cf. Section 5).

For a given contract we plot the historical daily prices observed during its trading time interval and run a single simulation for the sake of qualitative comparison. We take as examples two contracts for each delivery period of the EEX data considered: calendar, quarter, month. Specifically, we plot Cal-18 and Cal-19 in Figures 3.7 and 3.8, which are, respectively, the longest traded one and the last one in our observation period. In Figures 3.9 and 3.10, Q2/16 and Q1/18 are compared, being the first and last observed quarter. Finally, we plot in Figures 3.11 and 3.12 the prices of the two monthly futures Jan/17 (cold month) and Jul/17 (warm month). Together with price trajectories, we indicate the long-term mean of the model by a dotted line.

From a visual perspective, the simulated paths mimic the observed movements reasonably well, being capable to follow different trends for different periods as well. This is particularly encouraging, if we consider that the overall observation window is long about one year and a half and, moreover, we observe contracts with heterogeneous delivery periods and trading lives, some of which are traded for less than 40 days, such as Jul/17 (see Figure 3.12). The mean-reversion effect is generally evident and well reproduced by our model, especially if we focus on the mean-trend described by the dotted line. However, if we move from this line, in some cases the simulations seem not to take into account certain idiosyncratic movements. This can be noticed in Figure 3.7 (Cal-18) and Figure 3.10 (Q1/18). One explanation could be related to the mean-reversion rate of decay, which determines the excursions from the long term equilibrium. Recall that, by arbitrage, the parameter λ must be equal for all traded futures, but if we calibrate it for each contract separately we observe different values in practice. This assumption of course reflects on the accuracy in reproducing the single contract prices pathwise, but it is required by our no-arbitrage arguments.

For a more rigorous discussion of the fitting quality, we investigate the statistical features of the model and make a comparison with the historical data. A standard way to do it is by computing moments. In Table 3.3 we report the values of the first four moments, the minimum and the maximum of both empirical and simulated returns, i.e. the daily price increments, of all contracts. For the simulated returns, we run 1000 simulations and then average the results over all samples. The values are classified among different delivery periods in order to distinguish different behaviors (if present) among them.

The general performance is quite satisfactory. The mean is positive but very close to zero, both in the model as for the observed prices, with the notable exception of the calendars, where the model reproduces the data's positive trend very well. The standard deviation is adequately captured for all delivery periods with a small error for the monthly contracts, which prove to be the most difficult to model also looking at the other statistics: let us remark that, in these markets, monthly contracts are available for trading for relatively short periods. Regarding the skewness, the results are very good, since both the model and the observed returns are statistically not skewed. Let us recall that the standard error of the skewness estimator is approximately $\sqrt{6/n}$, where n is the number of daily observations per contract. In our case, n approximately ranges from 40 to 65 for monthly contracts, from 60 to 250 for the quarters and from 40 to 270 for the calendars. This means that, for example, in the case of some quarter and calendar contracts, where we have the highest number of observations ($n = 270$), the standard error is about 0.15. Then, roughly speaking, any value between -0.3 and 0.3 is not statistically different from zero with a 5% significance level, as it is observed from the market even for the monthly contracts (which have much less trading days). With the same procedure, a similar computation can be done for the kurtosis, which gives results still within the confidence interval.

To conclude, our framework seems capable to describe fairly well the main stylized features of power futures contracts, which are of different nature, in a comprehensive way. In other words, this assessment analysis suggests that the model adequately reproduces the main trajectorial and statistical properties of the Phelix Base futures prices. Specifically, in the statistic study presented in Table 3.3, we have not found evidence of non-normality in futures' daily price increments. Most importantly, this study has been carried out by calibrating the model directly on the observed prices of the traded contracts, i.e. without introducing artificial forward curves or *ex post* estimated risk premium, besides working under the no-arbitrage assumption.

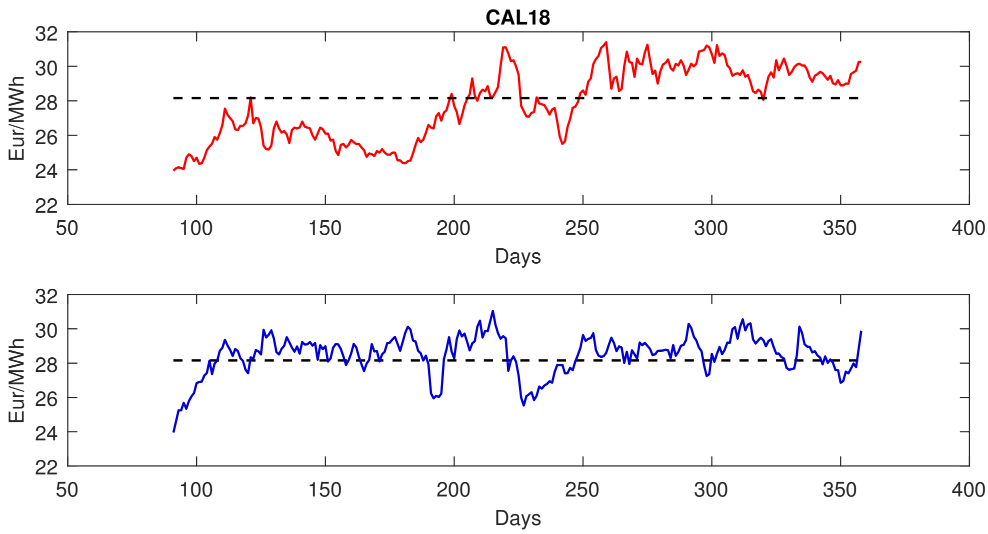


Figure 3.7: Historical (red) and simulated (blue) path of the contract Cal-18. The dotted line represents in both plots the estimated long-term mean of the contract.

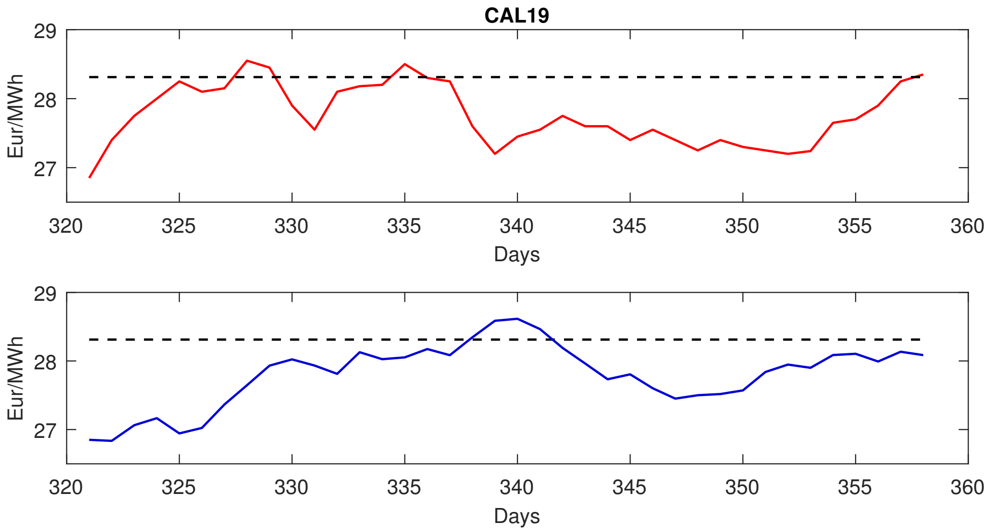


Figure 3.8: Historical (red) and simulated (blue) path of the contract Cal-19. The dotted line represents in both plots the estimated long-term mean of the contract.

3.7 Concluding remarks

In many models for gas and power markets, mean-reversion is a common feature of spot and forward dynamics. However, it is less common to address explicitly the question of arbitrage. This issue is particularly important in energy markets, where one simultaneously finds different forwards covering the same time period. In these situations, one must seek models which avoid the possibility of both static and dynamic arbitrages. This is ensured by the HJM paradigm, which prescribes that in diffusion models the drift and the diffusion coefficients of forward prices satisfy certain relations, which boil down to coefficients of the longest contracts being again suitable averages of those of the smaller contracts. This is difficult to achieve with multiplicative models, where one ends up with non-Markovian models. Instead with arithmetic models, the HJM relations are satisfied in a very natural way.

We here propose a model where forward prices are affine functions of hidden random

	Mean	Std. Dev.	Skewness	Ex. Kurt.	Min.	Max.
Phelix Base (M)	0.04	0.58	0.25	0.83	−1.32	1.62
Phelix Base (Q)	0.02	0.44	0.19	1.17	−1.36	1.44
Phelix Base (Cal)	0.03	0.41	−0.18	0.74	−1.35	1.26
Model (M)	0.00	0.67	0.01	1.71	−1.86	1.88
Model (Q)	0.00	0.43	0.06	1.87	−1.44	1.47
Model (Cal)	0.02	0.39	0.08	1.13	−1.25	1.32
Phelix Base	0.03	0.52	0.19	0.92	−1.33	1.53
Model	0.00	0.57	0.03	1.69	−1.68	1.70

Table 3.3: Empirical vs. simulated statistics (first four normalized moments, minimum and maximum) of the daily returns of all the contracts aggregated by delivery period. Model statistics are computed by averaging the results of the estimators over 1000 simulations, first, for each contract and, second, among contracts grouped by delivery period. The last two rows show the overall results for all the contracts.

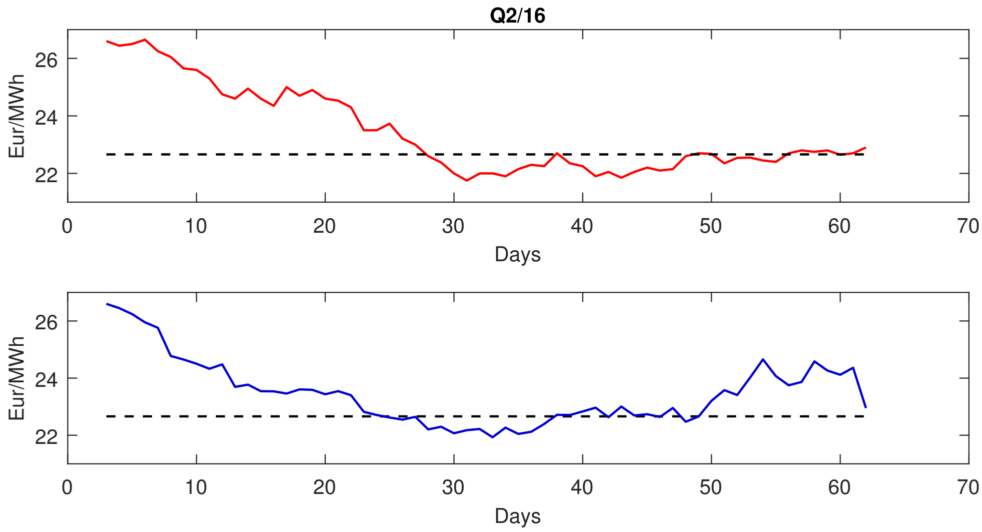


Figure 3.9: Historical (red) and simulated (blue) path of the contract Q2/16. The dotted line represents in both plots the estimated long-term mean of the futures price.

factors, which we impose not to appear explicitly in the forward prices' dynamics in order not to have to deal with models with incomplete information. Once we postulate that all the forward contracts in the market have the same mean-reversion speed, it turns out that this requirement is equivalent to the fact that the coefficients satisfy suitable linear algebraic conditions. We also show that the well known Lucia-Schwartz model belongs to the class of these models. The Lucia-Schwartz model is quite flexible in mean price levels, while instead it is able to reproduce only certain term structures of volatility.

To overcome this rigidity, we propose a modification of the Lucia-Schwartz model which is free of arbitrages in the way discussed above, where the prices of forward contracts are mean-reverting and where both price level and volatility are allowed to have a non-trivial term structure. As in the Lucia-Schwartz model, the mean-reversion level of forward contracts corresponds to the seasonal component of spot prices, and the prices are moved by two Brownian motions, possibly correlated. However, while one Brownian component takes into account the short-term movements and reproduces a Samuelson-like term structure decaying exponentially with the time to maturity, the other Brownian component is able to model more complex term structures of volatility with respect to absolute maturities. The model also predicts a well-observed phenomenon, i.e. contracts with shorter delivery periods (e.g. months) are more volatile than contracts with longer delivery periods (e.g. quarters or calendars).

We then present a technique for estimating the parameters of this model based on market time series. More in detail, first we estimate the volatility coefficients by using a technique based on quadratic variation/covariation of all the considered time series. Then we estimate the mean-reversion speed and levels in a non-parametric way, so that we can replicate any functional behavior of the forward long-term mean, which takes into account the no-arbitrage requirements among forwards covering the same time period.

We apply this estimation technique on a time series for the Phelix Base forward market from January 4, 2016 to May 23, 2017, considering each monthly, quarterly and calendar forward contract traded in that time window. Since our model is able to reproduce complex volatility shapes, we first try with a popular choice for seasonal component, i.e. a linear combination of trigonometric functions plus a linear component in the maturity, and we

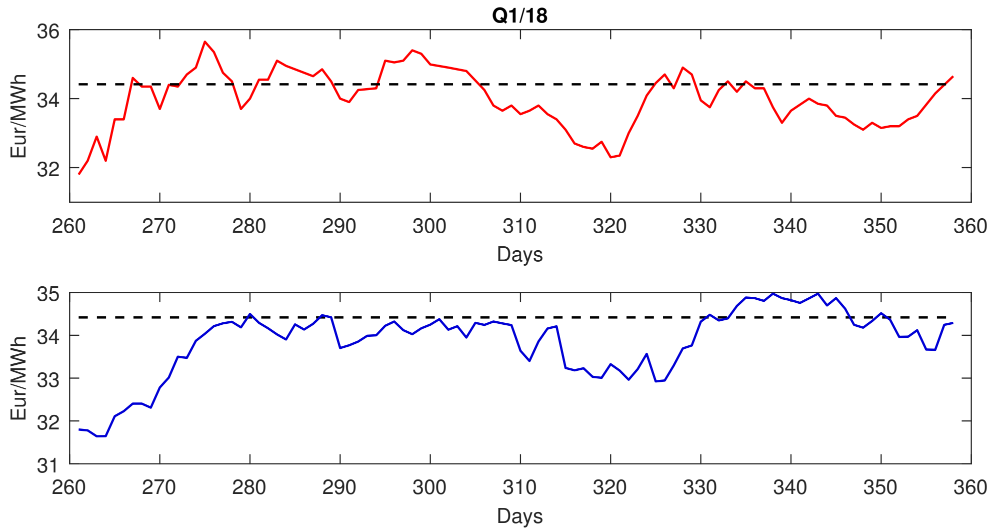


Figure 3.10: Historical (red) and simulated (blue) path of the contract Q1/18. The dotted line represents in both plots the estimated long-term mean of the futures price.

compare it with a fully non-parametric volatility shape, where the second factor is a free parameter for each atomic forward. It turns out that the fully non-parametric version fits the empirical volatility structure much better than the parametric model. Besides, we observe that the Samuelson effect is not enough to explain the observed volatility term structure and that our model can capture also complex seasonal behaviors. More in detail, we find a decreasing (concave) trend in calendar contracts' volatility and a seasonal pattern in the quarters' term structure. The model also partially reconstructs the finer term structure for monthly contracts, even if for those we have less observations.

We then estimate the mean-reversion speed and level. We assume that this latter quantity depends on the contract in exam, but that overall no-arbitrage relations must hold among contracts spanning the same period (e.g. 3 months / 1 quarter): this is done using MLE with the technique of Lagrange multipliers to take into account no-arbitrage constraints.

Finally, we perform a model assessment study, where we simulate paths of futures prices and compare them to the corresponding observed trajectories, for an eye test of model simulations, which proves to be satisfactory. We then compute estimated statistics over a set of simulated returns and compare them to the empirical ones. From a distributional perspective, the results seem to validate the model.

Future work is aimed at extending the present framework to multicommodity markets, to assess whether codependencies are present among the dynamics of different commodities, and to the case when jumps are present in the dynamics. Two of the authors are already laying out the theory in this direction [26], while the estimation technique is left for future research.

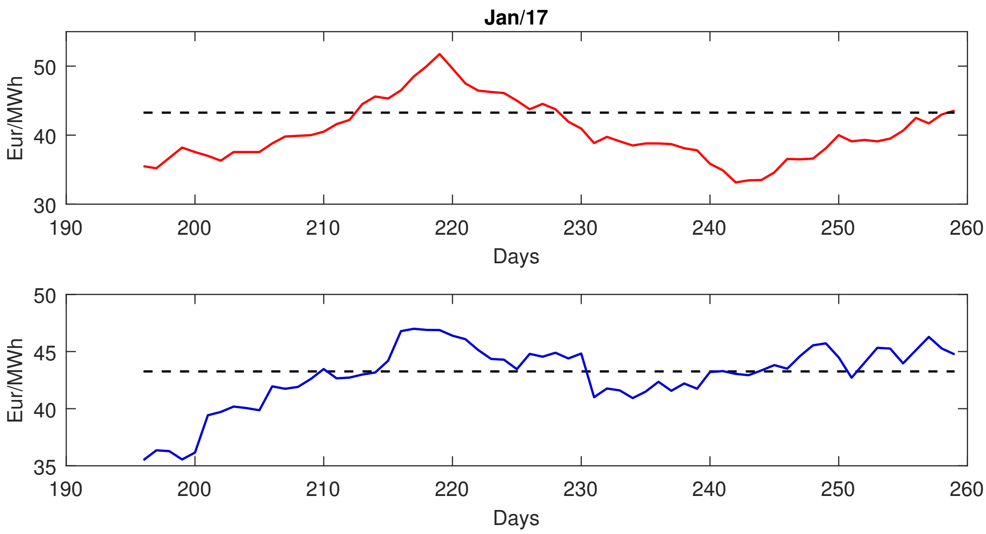


Figure 3.11: Historical (red) and simulated (blue) path of the contract Jan/17. The dotted line represents in both plots the estimated long-term mean of the futures price.

3.8 Appendix

3.8.1 Drift coefficient estimators

Estimation of q and Φ_i from a single-forward time series

We have

$$\sum_{l=2}^{n} \log \varphi\big(f_i(t_l) - f_i(t_{l-1}), q\big(\Phi - f_i(t_{l-1})\big), \sigma_i^2(t_{l-1})\big)$$
$$= (n-1)\sum_{l=2}^{n} \log \frac{1}{\sqrt{2\pi\sigma_i^2(t_{l-1})}} - \sum_{l=2}^{n} \frac{1}{2}\frac{\big(f_i(t_l) - f_i(t_{l-1}) - q(\Phi - f_i(t_{l-1}))\big)^2}{\sigma_i^2(t_{l-1})}.$$

Maximizing the above expression with respect to p corresponds to minimizing

$$\sum_{l=2}^{n} \frac{\big(f_i(t_l) - f_i(t_{l-1}) - q(\Phi - f_i(t_{l-1}))\big)^2}{\sigma_i^2(t_{l-1})}.$$

Let us adopt the notation $f_l := f_i(t_l)$ and $\sigma_l^2 := \sigma_i^2(t_l)$, being the index i fixed in this computation.

By differentiating with respect to q and Φ, we get the first order conditions (FOC)

$$-2\sum_{l=2}^{n} \frac{\big(f_l - f_{l-1} - q(\Phi - f_{l-1})\big)(\Phi - f_{l-1})}{\sigma_{l-1}^2} = 0, \tag{3.26}$$

$$-2q \cdot \sum_{l=2}^{n} \frac{\big(f_l - f_{l-1} - q(\Phi - f_{l-1})\big)}{\sigma_{l-1}^2} = 0. \tag{3.27}$$

Being $q \neq 0$, from Equation (3.27) we have

$$\sum_{l=2}^{n} \frac{(f_l - f_{l-1})}{\sigma_{l-1}^2} = q \cdot \sum_{l=2}^{n} \frac{(\Phi - f_{l-1})}{\sigma_{l-1}^2},$$

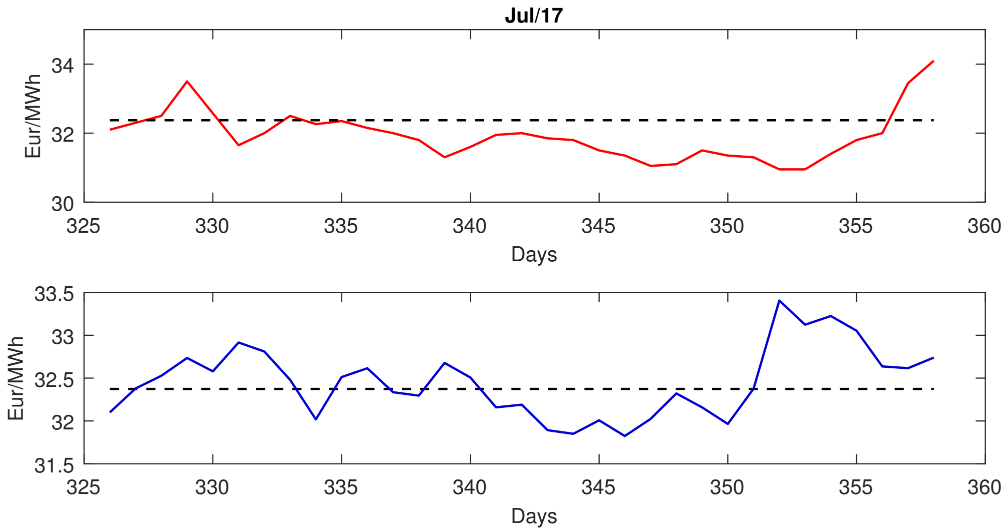


Figure 3.12: Historical (red) and simulated (blue) path of the contract Jul/17. The dotted line represents in both plots the estimated long-term mean of the futures price.

that is

$$\sum_{l=2}^{n} \frac{(f_l - f_{l-1})}{\sigma_{l-1}^2} = q \cdot \left(\sum_{l=2}^{n} \frac{1}{\sigma_{l-1}^2} \Phi - \sum_{l=2}^{n} \frac{f_{l-1}}{\sigma_{l-1}^2} \right).$$

By denoting

$$a_n = \sum_{l=2}^{n} \frac{1}{\sigma_{l-1}^2}, \quad b_n = \sum_{l=2}^{n} \frac{f_{l-1}}{\sigma_{l-1}^2}, \quad c_n = \sum_{l=2}^{n} \frac{(f_l - f_{l-1})}{\sigma_{l-1}^2},$$

it follows that

$$\Phi = \frac{1}{a_n} \left(\frac{c_n}{q} + b_n \right). \tag{3.28}$$

Then, by multiplying Equation (3.27) by Φ/q and subtracting it from Equation (3.26), we obtain

$$\sum_{l=2}^{n} \frac{(f_l - f_{l-1} - q(\Phi - f_{l-1}))f_{l-1}}{\sigma_{l-1}^2} = 0$$

which, by defining also

$$d_n = \sum_{l=2}^{n} \frac{f_{l-1}^2}{\sigma_{l-1}^2}, \quad e_n = \sum_{l=2}^{n} \frac{(f_l - f_{l-1})f_{l-1}}{\sigma_{l-1}^2},$$

can be written as

$$e_n - q(\Phi b_n - d_n) = 0$$

By plugging Equation (3.28) here, we obtain

$$e_n - \frac{b_n c_n}{a_n} - \frac{b_n^2}{a_n} q + d_n q = 0,$$

which finally gives the MLE for q as

$$\hat{q}_i = \frac{b_n c_n - a_n e_n}{a_n d_n - b_n^2}. \tag{3.29}$$

Replacing (3.29) in (3.28), we have the solution also for Φ:

$$\hat{\Phi}_i = \frac{c_n d_n - b_n e_n}{b_n c_n - a_n e_n}.$$

Estimation of the Φ_i under the no-arbitrage constraints

In order to solve the optimization problem in Definition 3.5.5 for m contracts f_i subject to constraints of the kind

$$\Phi_k = \sum_{j=1}^{r_k} w_{k_j} \Phi_{k_j},$$

for $k = 1, \dots, p$, we use the technique of Lagrange multipliers. In analogy with the unconstrained problem's equations, we have to minimize

$$\sum_{i=1}^{m} \sum_{l=2}^{n} \frac{\big(f_i(t_l) - f_i(t_{l-1}) - \hat{q}(\Phi_i - f_i(t_{l-1}))\big)^2}{\sigma_i^2(t_{l-1})} + \sum_{k=1}^{p} \lambda_k \left(\Phi_k - \sum_{j=1}^{r_k} w_{k_j} \Phi_{k_j} \right).$$

Hence, imposing the FOC and differentiating with respect to the multipliers λ_k, we obtain a linear system of $m+p$ equations of the kind of Equation (3.27) and $m+p$ unknowns (p of them are the multipliers), which have a unique solution $\tilde{p} = (\tilde{\Phi}_1, \dots, \tilde{\Phi}_m)$.

CHAPTER 4

Capturing the power options smile by an additive two-factor model for futures prices

In this chapter we introduce a no-arbitrage additive two-factor model for futures prices based on Normal Inverse Gaussian Lévy processes and compute European option prices by Fourier transform methods. We introduce a specific calibration procedure that takes into account no-arbitrage constraints and fit the model to the European Energy Exchange (EEX) power option settlement prices. We show that our model is able to reproduce the different levels and shapes of the implied volatility (IV) profiles displayed by options with different deliveries.

4.1 Introduction

Trading in electricity derivatives is living a significant expansion in Europe. In June 2018, the European Energy Exchange increased volumes on its power derivatives markets by 28% from 181.2 TWh (June 2017) to 231.1 TWh [59]. In particular, the trading volume in power options experienced a boost of 45%. These markets possess a variety of characteristics that need to be considered for appropriate pricing and risk management. In this paper we propose a method to compute European-styled vanilla option prices that accounts for some important stylized features of electricity markets.

There is a growing literature [7, 11, 118, 119] that considers e.g. seasonality, stochastic volatility and Samuelson-like effects in commodity option pricing. However, power markets present some peculiar issues that have not been completely addressed yet. Usually, the most liquid exchange-traded options are written on futures contracts that prescribe the delivery of power over different periods. For instance, the ones that we consider in our application are written on the Phelix Base Index traded at EEX [81], and the delivery is over a month, a quarter or a year. In order to build no-arbitrage models, in addition to working in a risk neutral setting where traded contracts are martingales, one must take into account possible arbitrages arising from trading futures with different delivery periods (cf. Chapter 3). The first paper that takes into account this fact is [89], where the authors apply an approach based on LIBOR market models [41]. To the best of our knowledge, this is also the first attempt to fit a consistent option pricing model for power markets. However, the LIBOR approach allows to model directly only shortest delivery, i.e. monthly, contracts, while resorting to approximation to the distribution of longest delivery futures. Also [101] address commodity option price models, but again only single delivery contracts can be considered in their framework. The two-factor model of [89] has been generalized by [44, Section 5] to Lévy models, allowing to describe the IV surface of option prices. Furthermore, a more general approximation procedure is proposed, but the results, though accurate for single deliveries, are partially satisfactory for longest maturity contracts, probably due to the required approximations. Among recent contributions in power options modeling, we mention [127], who uses a structural model that explains the formation of IV skews, and [102], who develop an extensive sensitivity analysis of IV patterns. However, differently from our setting, both papers consider options written on electricity spot prices. Another characteristic that we want to reproduce in our model is prices seasonality, in the sense of dependence on the delivery of the underlying futures. This has been addressed by [64], though in a Gaussian setting, which does not allow to capture the different IV levels displayed by options with different strikes.

In order to compute option prices, we introduce an arbitrage-free two-factor additive model for futures prices. In contrast to spot-based models, where the futures prices are computed by taking the discounted expectation of spot prices under a pricing measure, we start by specifying the stochastic dynamics of futures prices under a risk-neutral measure $\mathbb{Q}$. This is known to literature as the Heath-Jarrow-Morton approach applied to commodity markets (see e.g. [63] for a literature review in the case of electricity). Since the parameters of our model will be estimated from the option market, this approach has several advantages as already pointed out e.g. by [89]. For instance, we do not need to choose a pricing measure, or, equivalently, estimate the market price of risk that determines how the stochastic model changes from the physical measure to the risk-neutral measure under which the options are priced in the market. In fact, we can call our risk neutral measure $\mathbb{Q}$, *the* risk-neutral measure, since this is implicitly and univocally determined by the option prices observed in the market. We refer to [30] for an empirical discussion on pricing measures for electricity derivatives.

Since it is based on Lévy processes, our futures price model allows to capture the implied volatility profile described by options with different strikes. The dynamics arises as the natural generalization of its Gaussian counterpart introduced in Chapter 3. It assumes that futures prices are stirred by two stochastic factors built on Normal Inverse Gaussian (NIG) Lévy processes modulated by deterministic coefficients depending both on time and delivery period. The NIG distribution is a flexible family of distributions that is very popular in financial modeling (see e.g. [13] for general applications of NIG processes to finance, [33] for an electricity spot price NIG model and [6] for modeling electricity forwards). The first factor has a delivery-averaged exponential behavior, meant to reproduce the Samuelson effect [82]. The second factor is independent of time, but varies for contracts with different delivery in a *seasonal* and *no-arbitrage* way (cf. Chapters 2 and 3) and accounts for a finer reproduction of the term structure of futures volatilities. The term *additive* means that we do not consider the log-prices, but, instead, we model directly the prices. This class of models, as opposed to geometric models, has recently gained an increasing attention in literature due to many modeling advantages (see, for example, [18, 24, 26, 57, 88]). In the context of option pricing, [24] exploit the additive structure of their spot dynamics for pricing Asian and spread options by fast Fourier techniques. In our case, the additivity property will allow us to fit our model consistently to all the delivery periods traded in the market. Instead, the calibration results of [44, Section 5], where the author studied a geometric version of our model (still based on two NIG factors) were satisfactory only partially. As already mentioned at the beginning, this is likely due to the necessity to introduce an approximation procedure for the distribution of longest maturity contracts. In our case, by considering an additive model, we do not need to introduce such approximations.

For pricing options, we follow the classical method of [45], which consists, roughly speaking, in computing the Fourier transform of (a suitable modification of) the call price as a function of the strike price so to recover the option value by inverse transform. We recall from [45] two different approaches, that we adapt to the case of additive models. The first approach is based on the use of an exponential damping factor in order to make the option value integrable on the whole real line. Instead, the second approach consists in substracting the time value of the option. In this way we derive semi-analytical expressions (analytical up to numerical integration) for the option prices, that depend on the characteristic function of the underlying (see also [32]).

We discuss a calibration methodology that we will apply in our empirical study. The calibration happens statically, in the sense that we fix a trading date and observe the market option prices for different strikes and deliveries. We then find the parameters that minimize an objective function representing the distance from theoretical to model IVs. In the same way, one can alternatively use option prices instead of IVs. The futures model under consideration is defined in a such a way that there is no possibility of arbitrages from trading in overlapping delivery periods. Because no-arbitrage implies certain relations on the coefficients, this translates into parameter constraints (cf. the calibration procedure presented in Chapter 3).

Since there is not enough liquidity in the market in order to extract information on the IV surface from traded market quotes, we consider the options settlement prices, that are available for a sufficiently large range of strike prices. Though settlement prices do not necessarily represent trades that take place in the market, they contain information on the market expectations. We perform the calibration procedure described above, first, for a one-factor model derived by the two-factor one by setting one coefficient to 0 and, then, for the general two-factor model. We compare the IVs of both models to the empirical IVs and the one (constant across strikes) generated by Black's model [29, 38]. As a by-product from the estimation of the one-factor model, we derive that, under the risk-neutral measure, futures prices are leptokurtic and have significantly positive skewness. This is reflected also into the shape of empirical IVs, which display a *forward skew* (i.e. higher IVs for out-of-the-money calls). This can be interpreted as a "risk premium" paid by option buyers for securing supply (cf. [36, 124]). Finally, we show that the two-factor model is able to reproduce in a satisfactory way the different levels and shapes of the IV profiles displayed by all the deliveries traded in the market, by outperforming both the Black and the one-factor model.

4.2 Model

Let $F(t, T_1, T_2)$ denote the price at a given day t of a futures contract which delivers a fixed intensity of electricity over the period $[T_1, T_2]$. The delivery period $[T_1, T_2]$ can be a month, a quarter or a year and, since contracts expire right before delivery starts, we have that $t \leq T_1$. We introduce a general framework for multifactor additive futures prices (as, for example, in [32]) and, from this, we focus on a two-factor model inspired by [93], that will be of interest for application. More in detail, we introduce a stochastic differential equation, parametrized by the delivery period (i.e. depending, in addition to the trading day t, also on T_1, T_2), driven by independent Lévy factors. We also compute the corresponding characteristic functions, that will constitute the main ingredients in the computation of option prices (see Section 4.3). We will assume throughout this work that the risk-free interest rate is zero (similar arguments apply in the case of deterministic flat interest rate after discounting, see [93]).

4.2.1 Multifactor additive futures-based models

In the rest of the paper, we will consider European options written on futures contracts and we will denote by T the exercise date of these options. Therefore, for convenience we express the futures price at time T. We assume that, for any time t before the exercise of the option, i.e. $0 \leq t < T$, the futures prices $F(T, T_1, T_2)$ are given by the following stochastic differential equation (here given in integral form)

$$F(T,T_1,T_2) = F(t,T_1,T_2) + \sum_{k=1}^{p} \int_t^T \Sigma_k(u,T_1,T_2)\, dW_k(u) + \sum_{j=1}^{n} \int_t^T \Gamma_j(u,T_1,T_2)\, dJ_j(u) \tag{4.1}$$

where W_k are independent Brownian motions for $k = 1, \dots, p$ and $J_j(u) = \int_0^u \int_{\mathbb{R}} y\, \widetilde{N}_j(dy, dv)$ are independent, pure-jump, centered Lévy processes such that $\int_{|y|>1} y^2\, \nu_j(dy) < \infty$ for each $j = 1, \dots, n$. This integrability assumption on the Lévy measure implies that J_j are square-integrable martingales with zero expectation (for background on Lévy processes see e.g. [52]). We assume that all the stochastic factors are independent, so that, in particular, the n Poisson random measures are independent of the p Brownian components. The dynamics are described under a risk-neutral measure $\mathbb{Q}$. The absence of the drift follows from the fact that, by no-arbitrage, futures prices must be martingales under $\mathbb{Q}$ (see e.g. [35]).

In order to apply the Fourier transform approach to option pricing, we need the characteristic function of the underlying process. By introducing $Z(t, T, T_1, T_2) := F(T, T_1, T_2) - F(t, T_1, T_2)$ and its characteristic function (as a function of $v \in \mathbb{R}$)

$$\Psi(t,T,T_1,T_2,v) = \mathbb{E}\left[e^{ivZ(t,T,T_1,T_2)}\,|\mathcal{F}_t\right],$$

we have that (see e.g. [32])

$$\log \Psi(t,T,T_1,T_2,v) = -\frac{1}{2}v^2 \sum_{k=1}^{p} \int_t^T \Sigma_k^2(u,T_1,T_2)\, du + \sum_{j=1}^{n} \psi_j(t,T; v\Gamma_j(\cdot,T_1,T_2)). \tag{4.2}$$

The function $\psi_j(t, T; \theta(\cdot))$ denotes

$$\psi_j(t,T;\theta(\cdot)) = \int_t^T \widetilde{\psi}_j(\theta(u))\, du = \int_t^T \int_{\mathbb{R}} (e^{i\theta(u)z} - 1 - i\theta(u)z)\, \nu_j(dz)\, du,$$

where $\widetilde{\psi}_j(\theta)$ is the cumulant of the Lévy process J_j computed in $\theta \in \mathbb{R}$, i.e.

$$\widetilde{\psi}_j(\theta) = \log \mathbb{E}\left[e^{i\theta J_j(1)}\right]$$

and ν_j is the Lévy measure of J_j.

4.2.2 A two-factor model based on Normal Inverse Gaussian processes

In this section we consider a two-factor model of the type (4.1) based on the Normal Inverse Gaussian (NIG) distribution. This is motivated by the framework [26] and arises as a natural generalization of [93]. The model assumes that futures prices are stirred by two stochastic factors built on Normal Inverse Gaussian Lévy processes modulated by deterministic coefficients. The first factor has a delivery-averaged exponential behavior, meant to reproduce the so-called Samuelson effect. The Samuelson effect is an observed feature of prices volatilities, common to many commodity markets, consisting of increasing volatility of prices as time approaches maturity [82]. The second factor is independent of time, but varies for contracts with different delivery in a *seasonal* and *no-arbitrage* way (see [26, 93]) and accounts for a finer reproduction of the term structure of futures volatilities. We remark that, since our model is additive, by "volatility" we mean the parameter (or function of parameters) that determines the variability of prices and not of log-prices as in geometric models.

Building upon (4.1), we assume that the stochastic evolution of $F(\cdot, T_1, T_2)$ from t to T is described by

$$F(T, T_1, T_2) = F(t, T_1, T_2) + \int_t^T \Gamma_1(u, T_1, T_2)\, dJ_1(u) + \Gamma_2(T_1, T_2)(J_2(T) - J_2(t)), \qquad (4.3)$$

where

$$\Gamma_1(u, T_1, T_2) := \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \gamma_1 e^{-\mu(\tau - u)}\, d\tau = \frac{\gamma_1 (e^{-\mu(T_1 - u)} - e^{-\mu(T_2 - u)})}{\mu(T_2 - T_1)}, \qquad (4.4)$$

$$\Gamma_2(T_1, T_2) := \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \gamma_2(\tau)\, d\tau. \qquad (4.5)$$

By focusing our attention to the first component, the parameter $\gamma_1 \in \mathbb{R}^+$ models the *base volatility*, that is the volatility of contracts with distant delivery i.e. $T_1 - t \to \infty$. This coefficient has an exponential rate given by $\mu \in \mathbb{R}^+$, which determines an increase of volatility as time approaches delivery i.e. $T_1 - t \to 0$. This effect is averaged over the delivery period for no-arbitrage arguments (as explained in [93]) and mimics the Samuelson effect. Regarding the second component, the function $\gamma_2 : [0, \infty) \to \mathbb{R}^+$ models the general seasonal behavior of volatility (it takes high values in periods of high volatility and low values for periods of low volatility). This function can be specified either in a parametric or nonparametric fashion. For instance, one can have

$$\gamma_2(\tau) := \gamma_2 + b\tau + \sum_{j=1}^{m} \big(a_{2j} \cos(\omega j \tau) + a_{2j+1} \sin(\omega j \tau)\big),$$

with $\omega = 2\pi/365$, $m \in \mathbb{N}$, $b \in \mathbb{R}$ (for capturing possible deterministic linear trends), $(\gamma_2, a_2, \ldots, a_{2m+1}) \in \mathbb{R}^{2m+1}$.

The coefficients in (4.4) and (4.5) modulate the variability of the two stochastic processes J_1 and J_2, which are both defined as centered versions of NIG Lévy processes. Let us recall that a Lévy process is called Normal Inverse Gaussian with parameters $(\alpha, \beta, \delta, \mu)$ if its characteristic triplet is $(\chi, 0, \nu)$ with

$$\chi = m + \frac{2\alpha\delta}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x)\, dx, \qquad (4.6)$$

$$\nu(dy) = \frac{\alpha\delta}{\pi |y|} K_1(\alpha |y|) e^{\beta y}\, dy, \qquad (4.7)$$

where K_1 is the modified Bessel function of the third kind with index 1 (in the terminology of [2, Section 9.6]), $0 \leq |\beta| < \alpha$ and $\delta > 0$, $m \in \mathbb{R}$ (see [12]). Given a NIG process L, the random variable $L(1)$ is NIG distributed with parameters $(\alpha, \beta, \delta, \mu)$. The NIG distribution is a subclass of a very flexible family, the Generalized Hyperbolic distributions, and it can accomodate heavy-tails and skewness. The parameter α rules the tail heaviness of the distribution, β determines the skewness, δ is a scale parameter and μ indicates the location of the distribution.

In general, a NIG process is not centered. Therefore, in order to define J_1 and J_2 we subtract from two general NIG processes L_1, L_2 the corresponding expected value (multiplied by time). For $j = 1, 2$, let L_j be a NIG Lévy process with parameters $(\alpha_j, \beta_j, \delta_j, m_j)$ and characteristic triplet $(\chi_j, 0, \nu_j)$ and set

$$J_j(t) = L_j(t) - \mathbb{E}[L_j(t)] = L_j(t) - t\left(\chi_j + \int_{|y|\geq 1} y\,\nu_j(dy)\right).$$

Then, J_j is a centered NIG process. In particular, it can be easily shown (cf. the characteristic function of J_j in (4.32) to see this) that J_j is a NIG process with parameters $\left(\alpha_j, \beta_j, \delta_j, -\frac{\delta_j\beta_j}{\sqrt{\alpha_j^2-\beta_j^2}}\right)$.

As already mentioned, in order to compute the option prices, we will need the characteristic function of $F(T, T_1, T_2)$. This in turn reduces to finding the characteristic function of $Z(t, T, T_1, T_2) := F(T, T_1, T_2) - F(t, T_1, T_2)$. We compute it in the appendix in order not to make the presentation unnecessarily heavy.

4.3 Option pricing for additive models by Fourier transform methods

We consider the pricing of European vanilla options written on futures contracts of the type introduced in the previous section. We discuss the method for call options, being the case of puts completely analogous. Let $C(t; T, K, T_1, T_2)$ be the price at time t (the observation date) of a call option, with strike price K and exercise time T, that is written on a futures contract with delivery period $[T_1, T_2]$. By no-arbitrage (see, for instance, [35]),

$$C(t; T, K) = \mathbb{E}\left[(F(T) - K)_+ \,|\mathcal{F}_t\right], \tag{4.8}$$

where $F(T)$ is the price of the underlying futures at the option exercise as in (4.1) and $\mathcal{F}_t$ represents the filtration at time t, i.e. the information flow up to time t. In order to ease the notation, sometimes we will not write the dependence on the delivery period, which does not come into play in this discussion. We recall that the expectation is taken under the risk-neutral measure $\mathbb{Q}$, even though it is not explicitly indicated, and throughout this work $\mathbb{Q}$ will be the only probability measure we will deal with. By applying the definition of conditional expectation, we can write

$$C(t; T, K) = \int_K^{+\infty} (s - K)\, q_{t,T}(s)\, ds, \tag{4.9}$$

where $q_{t,T}$ is the (risk-neutral) density function of $F(T)$ conditioned up to time t. This formula yields an expression for the option value, for instance, when the distribution of the underlying F allows for an explicit formula for the density, that, moreover, can be integrated against the payoff function in a tractable way (as it is the case in the Black model [38], where the underlying follows a geometric Brownian motion).

We follow the alternative approach of [45], which consists, roughly speaking, in computing the Fourier transform of (4.8) as a function of K (after proper manipulations) so to recover the option value by inverse transform. This has been studied by several authors and it has been shown to be a very convenient way to compute option prices in the case that the characteristic function of the underlying is known explicitly, while the density is not. The starting point of the above mentioned approach is the observation that, as K goes to $-\infty$, $C(t;T,K) \to \infty$, so that in particular $C(t;T,K)$ is not integrable as a function of K for large negative values. This means that the option value $C(t;T,K)$ does not satisfy the assumptions required for computing its Fourier transform. In order to overcome this, we follow [45], who suggest two approaches that we here recall and apply to the case of additive models.

4.3.1 First approach: modified option value

Following [45], define for arbitrary $a > 0$ the modified option value (as a function of the strike price $K \in \mathbb{R}$) by

$$z^{MO}_{t,T}(K) := e^{aK} C(t;T,K), \tag{4.10}$$

and, if it is square-integrable, compute its Fourier transform for $v \in \mathbb{R}$:

$$\xi^{MO}_{t,T}(v) := \int_{-\infty}^{+\infty} e^{ivK}\, z^{MO}_{t,T}(K)\, dK. \tag{4.11}$$

Since, by definition,

$$z^{MO}_{t,T}(K) = \int_{K}^{+\infty} e^{aK}\, (s-K)\, q_{t,T}(s)\, ds,$$

where $q_{t,T}$ is the density function of $F(T)$ conditioned up to time t, we can express $\xi^{MO}_{t,T}$ with respect to the characteristic function of the underlying using integration by parts (that can be used under the assumptions of Proposition 4.3.1):

$$\begin{aligned}
\xi^{MO}_{t,T}(v) &= \int_{-\infty}^{+\infty} e^{ivK} e^{aK} \int_{K}^{+\infty} (s-K)\, q_{t,T}(s)\, ds\, dK \\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \int_{-\infty}^{s} e^{(a+iv)K}\, (s-K)\, dK\, ds \\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \frac{e^{(a+iv)s}}{(a+iv)^2}\, ds,
\end{aligned} \tag{4.12}$$

that finally yields

$$\xi^{MO}_{t,T}(v) = e^{(a+iv)s(t)} \frac{\Psi(t,T,v-ia)}{(a+iv)^2}, \tag{4.13}$$

where, for any $t \in [0,T]$, $s(t) := F(t)$ and $\Psi(t,T,v)$ is the characteristic function of $Z(t,T,T_1,T_2)$ as in (4.2).

If $\xi^{MO}_{t,T}$ is integrable, by the Fourier inversion theorem, the option value in (4.10) can be recovered by inverting again (4.13):

$$C(t;T,K) = \frac{e^{-aK}}{2\pi} \int_{-\infty}^{+\infty} e^{-iKv}\, \xi^{MO}_{t,T}(v)\, dv. \tag{4.14}$$

Since $z_{t,T}(K)$ is a real-valued function, we have that $\overline{\xi_{t,T}(v)} = \xi_{t,T}(-v)$. This gives us that $\xi_{t,T}$ has even real part and odd imaginary part and so we can simplify (4.14) to

$$C(t;T,K) = \frac{1}{\pi} \int_{0}^{+\infty} \mathrm{Re}\left(e^{-iKv}\, \xi^{MO}_{t,T}(v)\right)\, dv. \tag{4.15}$$

In order to apply this formula, we need to justify the interchange of integrals operated in (4.19) under the following integrability assumptions on the futures price process.

Proposition 4.3.1. *If $Z(T,T_1,T_2)$ is exponentially integrable with finite exponential moment of order a, i.e. $\mathbb{E}\left[\mathrm{e}^{aZ(T,T_1,T_2)}\right] < \infty$, then*

$$\begin{aligned}\xi_{t,T}(v) &= \int_{-\infty}^{+\infty} e^{ivK} \int_{-\infty}^{+\infty} (s-K)\, q_{t,T}(s)\, (\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K})\, ds\, dK \\ &= \int_{-\infty}^{+\infty} q_{t,T}(s) \int_{s(t)}^{s} e^{ivK}\, (s-K)\, dK\, ds.\end{aligned}$$

Proof. It is enough to show that the integral of the absolute value of the integrand with respect to K, i.e.

$$\int_{-\infty}^{+\infty} e^{aK} |s-K|\, |\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K}|\, dK$$

can be integrated in s against the density $q_{t,T}(s)$. Since this is equal to

$$\int_{s(t)}^{s} e^{aK}(s-K)\, dK = \frac{e^{as(t)}(a(s(t)-s)-1) + e^{as}}{a^2} \approx \min(e^{as}, |s|),$$

at infinity, the integrability with respect to the density is given by the exponential integrability of $Z(t,T,T_1,T_2)$. □

4.3.2 Second approach: modified time value

The first approach is based on the use of an exponential damping factor e^{aK} in order to make the option value integrable. Instead, the second approach, that we are going to outline below, consists in substracting the time value of the option. The arguments and the steps are similar to the first approach, but we rewrite them for the sake of clarity. Still following [45] (see also [51]), define the modified time value of the option (as a function of the strike $K \in \mathbb{R}$) by

$$z^{MT}_{t,T}(K) = C(t;T,K) - (F(t)-K)_+, \tag{4.16}$$

and, if it is square-integrable, compute its Fourier transform

$$\xi^{MT}_{t,T}(v) = \int_{-\infty}^{+\infty} e^{ivK}\, z^{MT}_{t,T}(K)\, dK. \tag{4.17}$$

If $\xi^{MT}_{t,T}$ is integrable, by the Fourier inversion theorem, the modified time value of the option can be recovered by inverting again the last equation:

$$z^{MT}_{t,T}(K) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iKv}\, \xi^{MT}_{t,T}(v)\, dv. \tag{4.18}$$

Now, observe that, by the martingale property of $F(\cdot)$,

$$(F(t)-K)\, \mathbb{1}_{F(t)>K} = \mathbb{E}\left[(F(T)-K)\,|\mathcal{F}_t\right]\, \mathbb{1}_{F(t)>K}.$$

Then, we can write from (4.9)

$$z^{MT}_{t,T}(K) = \int_{-\infty}^{+\infty} (s-K)\, q_{t,T}(s)\, (\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K})\, ds$$

where, for any $t \in [0,T]$, $s(t) := F(t)$ and $q_{t,T}$ is the density function of $F(T)$ conditioned up to time t. As a consequence, if the interchange of integrals holds (see Proposition 4.3.2), (4.17) can be written as

$$\begin{aligned}
\xi^{MT}_{t,T}(v) &= \int_{-\infty}^{+\infty} e^{ivK} \int_{-\infty}^{+\infty} (s-K)\, q_{t,T}(s)\, (\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K})\, ds\, dK \\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \int_{s(t)}^{s} e^{ivK}\, (s-K)\, dK\, ds \qquad (4.19)\\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \left\{ \left.\frac{e^{ivK} s}{iv}\right|_{K=s(t)}^{s} - \int_{s(t)}^{s} K\, e^{ivK}\, dK \right\} ds \\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \left\{ -\frac{e^{ivs(t)} s}{iv} + \frac{e^{ivs(t)} s(t)}{iv} - \frac{e^{ivs}}{v^2} + \frac{e^{ivs(t)}}{v^2} \right\} ds \\
&= -\frac{1}{v^2} \int_{-\infty}^{+\infty} q_{t,T}(s)\, e^{ivs}\, ds + \frac{e^{ivs(t)}}{v^2} - \frac{e^{ivs(t)}}{iv} \int_{-\infty}^{+\infty} q_{t,T}(s)(s - s(t))\, ds.
\end{aligned}$$

Since F is a martingale,

$$\int_{-\infty}^{+\infty} q_{t,T}(s)(s-s(t))\, ds = \mathbb{E}\left[(F(T) - F(t))\,|\mathcal{F}_t\right] = 0$$

and, by definition of characteristic function (see also (4.2)),

$$\int_{-\infty}^{+\infty} q_{t,T}(s)\, e^{ivs}\, ds = \mathbb{E}\left[e^{ivF(T)}\,|\mathcal{F}_t\right] = e^{ivs(t)}\Psi(t,T,v),$$

where $\Psi(t,T,v)$ is the characteristic function of $Z(t,T)$. Then, we finally arrive to

$$\xi^{MT}_{t,T}(v) = e^{ivF(t)} \frac{1 - \Psi(t,T,v)}{v^2}. \qquad (4.20)$$

In analogy to the modified option approach, we have that $\xi^{MT}_{t,T}$ has even real part and odd imaginary part and so

$$z^{MT}_{t,T}(K) = \frac{1}{\pi} \int_{0}^{+\infty} \mathrm{Re}\left(e^{-iKv}\, \xi^{MT}_{t,T}(v)\right)\, dv. \qquad (4.21)$$

Finally, by recalling (4.16), the option value is computed from (4.21) by

$$C(t;T,K) = z^{MT}_{t,T}(K) + (F(t) - K)_+.$$

In order to apply this formula, we need to justify the interchange of integrals operated in (4.19) under the following integrability assumptions on the futures price process.

Proposition 4.3.2. *If $Z(T,T_1,T_2)$ is square-integrable, then*

$$\begin{aligned}
\xi_{t,T}(v) &= \int_{-\infty}^{+\infty} e^{ivK} \int_{-\infty}^{+\infty} (s-K)\, q_{t,T}(s)\, (\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K})\, ds\, dK \\
&= \int_{-\infty}^{+\infty} q_{t,T}(s) \int_{s(t)}^{s} e^{ivK}\, (s-K)\, dK\, ds.
\end{aligned}$$

Proof. It is enough to show that the integral of the absolute value of the integrand with respect to K, i.e.

$$\int_{-\infty}^{+\infty} |s - K| \, |\mathbb{1}_{s>K} - \mathbb{1}_{s(t)>K}| \, dK$$

can be integrated in s against the density $q_{t,T}(s)$. Since this is equal to

$$\int_{s(t)}^{s} (s - K) \, dK = \frac{1}{2}(s - s(t))^2,$$

the integrability with respect to the density is equivalent to the existence of the second moment of $Z(t, T, T_1, T_2)$. □

4.4 Calibration procedure

As we have derived in the previous section semi-analytical expressions (analytical up to integration) for the option prices, we now move to discussing a calibration methodology that we are going to apply in our empirical study. First, we discretize the option value that is given in integral form in (4.15) (first approach: modified option value) and (4.21) (second approach: modified time value) in the domain of integration. Then, we select a finite grid of strike prices, that consists in practice of the listed options available in the market for a given underlying. This procedure reduces the valuation problem to the computation of a finite sum of vectors, where each component is the option price for a given strike. Then we introduce a least squares problem designed to find the parameters that minimize an error function (for a discussion about the choice of the error function, we refer to [44, Section 5.4]). We can compute it for two possible quantities, the model error on option prices and on the IV given by Black's formula. By minimizing on IVs rather than option prices, we weight at the same way contracts with different maturity. However, minimizing on prices is preferable in terms of speed of computation, as it does not require the inversion of Black's formula at each step. We separate the calibration routine in three cases, of increasing generality, in order to discuss the presence of different constraints case-by-case.

4.4.1 Discretization of model option prices

The quantities that we have to discretize, (4.15) and (4.21), take the following form:

$$z_{t,T}(K) = \frac{\phi(K)}{\pi} \int_0^{+\infty} \mathrm{Re}\left(e^{-iKv} \, \xi_{t,T}(v)\right) \, dv, \tag{4.22}$$

where

$$\phi(K) = \begin{cases} 1 & \text{MT value} \\ e^{-aK} & \text{MO value} \end{cases}$$

First, we choose an upper limit $A \in \mathbb{R}^+$ in the above integration (see [45, Section 3.1] for a discussion on how to do this optimally). Then, if we apply a simple Euler rule to the truncated integral, we find an expression of the form

$$z_{t,T}(K) \approx \frac{\phi(K)}{\pi} \sum_{j=1}^{N} \mathrm{Re}\left(e^{-iv_jK} \, \xi_{t,T}(v_j)\right) \eta, \tag{4.23}$$

where $N \in \mathbb{N}$, $\eta := A/N$ and $v_j := \eta(j-1)$ is the integration step. Given $M \in \mathbb{N}$ different strikes with granularity $\kappa > 0$, we compute the function in (4.23) for the following values of K:

$$K_u := \overline{K} + \kappa(u-1), \qquad \text{for } u = 1, \ldots, M,$$

being $\overline{K}$ the lowest strike price traded. By plugging this in (4.23), we get for $u = 1, \ldots, M$

$$z_{t,T}(K_u) \approx w_{t,T}(K_u) := \frac{\phi(K_u)}{\pi} \sum_{j=1}^{N} \mathrm{Re}\left(e^{-i\kappa\eta(j-1)(u-1)} e^{-i\overline{K}\eta(j-1)}\, \xi_{t,T}(\eta(j-1))\right) \eta. \quad (4.24)$$

As pointed out in [45], this formula is suitable for the application of the fast Fourier transform (FFT) algorithm. In order to do this, one must impose that the number of strike prices considered is equal to the number of integration nodes, i.e. $M = N$ (which is typically chosen as a power of 2). Moreover, it must hold that

$$\kappa\eta = \frac{2\pi}{N},$$

which consists of a trade-off between the grid for the integration and the granularity of strike prices. In particular, since in practice the granularity κ is given, this equality univocally defines the integration grid as a function of N. However, in our application we will not make use of the FFT algorithm and so, in particular, we will not impose the above restrictions on N, M, κ, η. This is motivated by the fact that we do not have a significant advantage in the computational complexity of the problem, being the number of strikes consistently lower than N. Furthermore, the focus of our work is not on the speed-up of the calibration procedure, but rather on the empirical results, so that we do not exclude the possibility to apply the FFT algorithm, being still possible from a theoretical point of view.

4.4.2 Parameters estimation

Since we have at disposal formulas in discrete form that can be readily implemented, we can now discuss how to fit the two-factor model to market data. The calibration happens statically, taking a snapshot of the market, in the sense that we fix a trading date and observe the market option prices for that date. From these prices we compute the IVs, with respect to the different listed strikes and underlyings, both for market and model prices. We then find the parameters that fit a certain distance from theoretical to model IVs best. As mentioned at the beginning of this section, one can alternatively use option prices instead of IVs. The futures model under consideration is defined in a such a way that there is no possibility of arbitrages, also in the case of overlapping delivery periods. No-arbitrage implies certain relations on the coefficients, that translate into parameter constraints. In order to highlight this, we present the objective function and the parameters set for the following three cases: single underlying, many underlyings but non-overlapping delivery periods, and general case of possibly overlapping delivery periods.

Black's Formula

The first and most used in practice model for option prices written on futures is the Black model [38]. It assumes that the underlying follows a geometric Brownian motion as in the Black-Scholes formula. Also, the expression for the call price is very similar to the Black-Scholes one, with the futures price replacing the stock price, but with different discounting. Since we are assuming that the risk-free rate is zero, the Black-Scholes and the option price given by the Black model are actually the same in our case. We recall the Black formula here because, in addition to use it as a benchmark in the upcoming empirical application, we use it to compute the implied volatility of both market and model prices:

$$C_{BS}(t;T,K) = F(t,T)N(d_1) - KN(d_2), \tag{4.25}$$

where N denotes the cumulative distribution function of a standard Normal random variable and

$$d_1 = \frac{\log \frac{F(t)}{K} + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}.$$

The implied volatility of a given option with price P and strike K is defined as the only (positive) number σ such that the Black formula for a strike K and volatility σ (all other quantities being equal) gives the price P. The two-factor model aims to reproduce the IV profile of market option prices, that is the plot of σ with respect to K. Black's formula yields constant implied volatility with respect to K, while real option prices usually display smiles or smirks (i.e. the IV is not constant and shows a certain convexity).

Single contract

Assume that we observe at a certain date $t < T$ the prices $c_*(t,T,T_1,T_2,K_u)$ of a call option for M different strike prices K_u, with exercise at time T written on a futures with delivery over $[T_1,T_2]$. We compute the IV of $c_*(t,T,T_1,T_2,K_u)$ by inverting Black's formula, obtaining a function of the M strikes $\sigma_*^{imp}(t,T,T_1,T_2,K_u)$, for $u = 1,\dots,M$. To fit the model to the observe IVs we search for the parameters that minimize a certain error function. Specifically, we introduce the following least-squares problem:

$$\widehat{\theta} := \arg\min_{\theta\in\Theta} \sum_{u=1}^{M} |\sigma^{imp}(t,T,T_1,T_2,K_u) - \sigma_*^{imp}(t,T,T_1,T_2,K_u)|^2, \tag{4.26}$$

where $\sigma^{imp}(t,T,T_1,T_2,K_u)$ is the IV computed on the prices generated by the model. The set Θ contains all the parameters appearing in the approximation formula $C(t,T,T_1,T_2,K_u) \approx c(t,T,T_1,T_2,K_u)$ where

$$c(t,T,T_1,T_2,K_u) := \begin{cases} w_{t,T}(K_u) + (F(t,T_1,T_2) - K_u)_+ & \text{MT value} \\ w_{t,T}(K_u) & \text{MO value} \end{cases} \tag{4.27}$$

where $w_{t,T}(K_u)$ for $u = 1,\dots,M$ is the discrete function in (4.24).

If we assume that $F(T,T_1,T_2)$ is given by a two-factor pure-jump model of the form (4.3) where Γ_i are defined in (4.4) and (4.5) and J_i is a centered NIG Lévy process with parameters $(\alpha_i,\beta_i,1)$, then $\Theta = \{\theta = (\alpha_1,\alpha_2,\beta_1,\beta_2,\mu,\gamma_1,\Gamma_2(T_1,T_2)) \in (\mathbb{R}^+)^2 \times (\mathbb{R}_0^+)^4 \times \mathbb{R}^+ : 0 \le |\beta_j| < \alpha_j\}$. We do not need to indicate the parameter m of the original NIG Lévy process L_i (see Section 3.2) since it does not appear in the centered version. Also, the parameter δ is assumed to be 1 without here because of the presence of the multiplying factors $\Gamma_1(u,T_1,T_2)$ and $\Gamma_2(T_1,T_2)$, thanks to the property that, given a NIG(α,β,δ) distributed random variable X, for a constant $\gamma > 0$, γX is distributed as a NIG$(\alpha/\gamma,\beta/\gamma,\delta\gamma)$. This means that letting δ vary in the parameters set would result in an overparametrization of the minimization problem.

Non-overlapping futures

Let us assume that we are at time t and observe I call options written on futures contracts with non-overlapping delivery periods $[T_{i,1}, T_{i,2}]$ (for example Jan/YY, Feb/YY, Mar/YY), exercise at T_i and M_i strike prices

$$K_u^i := \overline{K}_i + \kappa_i(u-1),$$

for $u = 1, \dots, M_i$ and $i = 1, \dots, I$. In the case of more than one contract, we introduce the following least-squares problem:

$$\widehat{\theta} := \arg\min_{\theta \in \Theta} \sum_{i=1}^{I} \sum_{u=1}^{M_i} |\sigma^{imp}(t, T_i, T_{i,1}, T_{i,2}, K_u^i) - \sigma_*^{imp}(t, T_i, T_{i,1}, T_{i,2}, K_u^i)|^2. \tag{4.28}$$

where Θ is now the set of parameters $\{\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \mu, \gamma_1, \{\Gamma_i\}_{i=1}^I) \in (\mathbb{R}^+)^2 \times (\mathbb{R}_0^+)^4 \times (\mathbb{R}^+)^I : 0 \leq |\beta_j| < \alpha_j\}$. For $I = 1$ we recover exactly the previous case.

Whole market

In general options traded in power markets are written on I futures contracts with different delivery length and possibly overlapping. For example, one can have Apr/YY, May/YY, Jun/YY, Jul/YY, Q2/YY, Q3/YY, Cal-YY. As explained in [93], in order to estimate the parameters in a consistent, i.e. arbitrage-free, way, we have to take into account additional constraints on the parameters. From Equation (4.5), it can be shown that the parameters Γ_i satisfy the following constraints

$$\Gamma_i = \Gamma_2(T_{i,1}, T_{i,2}) = \sum_{j=1}^{n} \frac{T_{j,2} - T_{j,1}}{T_{i,2} - T_{i,1}} \Gamma_2(T_{j,1}, T_{j,2}) = \sum_{j=1}^{n} \frac{T_{j,2} - T_{j,1}}{T_{i,2} - T_{i,1}} \Gamma_j \tag{4.29}$$

whenever $[T_{i,1}, T_{i,2}]$ is the union of disjoint intervals $[T_{j,1}, T_{j,2}]$ for $j = 1, \dots, n$, i.e. for all the forwards with overlapping delivery. Let us call *atomic*, the contracts whose delivery period can not be partitioned by the delivery periods of other futures. In other words, we suppose that m forwards $F_1, \dots, F_m$ have non-overlapping delivery periods $[T_{1,1}, T_{1,2}], \dots, [T_{m,1}, T_{m,2}]$ and such that the delivery periods of the other contracts traded in the market can be expressed as union of the former. For example, assume that we observe in the chosen calibration window the option prices for futures contracts delivering over Apr/YY, May/YY, Jun/YY, Jul/YY, Q2/YY, Q3/YY, Cal-YY. On one hand, Q2/YY is not atomic, since it can be "splitted" into Apr/YY, May/YY and Jun/YY. On the other hand, Q3/YY turns out to be atomic, even if Jul/YY is already traded, as Aug/YY and Sep/YY are not observed. For the same reason, Cal-YY is considered atomic as well. Then, in this example, if $\Gamma_1, \dots, \Gamma_7$ denote the corresponding parameters, we have that $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_7$ are free parameters, as they refer to atomic contracts, whereas to determine Γ_5 we use Equation (4.29). Consequently, we define the same statistics as in (4.28) but where now the vector of parameters is subject to the additional constraints given by Equation (4.29). For example, with the convention that $\Gamma_{\mathrm{Q2/YY}}$ denotes the parameter Γ_i corresponding to the contract Q2/YY, then

$$\Gamma_{\mathrm{Q2/YY}} = u_{\mathrm{Apr/YY}} \Gamma_{\mathrm{Apr/YY}} + u_{\mathrm{May/YY}} \Gamma_{\mathrm{May/YY}} + u_{\mathrm{Jun/YY}} \Gamma_{\mathrm{Jun/YY}}, \tag{4.30}$$

where the weights u_i are defined according to the number of days in the month/quarter (e.g. for Apr/YY we have $u_{\mathrm{Apr/YY}} = 30/91$).

4.5 Empirical results

We describe our dataset, compute and discuss the empirically observed IV implied by settlement prices. Finally, we compare the performance of three models: the Black formula, the one-factor model (being a special case of the two-factor one) and the general two-factor model.

4.5.1 Data description

The contracts that we consider in our application are European-styled call options traded at the EEX Power Derivatives market. The underlying assets are futures contracts that prescribe the delivery of 1 MW per hour, for each hour of each day of a month, a quarter or a year. More specifically, we will consider call options written on the Phelix Base index of the German/Austrian area. These options are called Phelix Base Month/Quarter/Year Options and the EEX official product codes are O1BM, O1BQ, O1BY. The term *Base* refers to *Base Load*, because the delivery of electricity takes place for each hour of the day, in contrast to *Peak Load* contracts, that instead prescribe the delivery only for the hours from 8 to 20. Usually, the exercise of the options under consideration is few trading days before the start of delivery. Since recently, yearly options are now available for four different exercise dates. We consider in our dataset only the ones with expiration few days before delivery, in analogy to quarterly and monthly contracts. Since there is not enough liquidity in the market in order to extract information on the IV surface from traded market quotes, we consider the settlement prices, that are available for a sufficiently large range of maturities and strikes. Though settlement prices do not represent trades that really take place in the market, they contain information on the market expectations. We observe the market for a representative day: Monday, March 5, 2018. For each option, we consider the strike prices in the range 90%-110% of the underlying current price (as, for example, in [7, 11]). At this date the listed options with available settlement prices are the ones written on five monthly (Apr/18, May/18, Jun/18, Jul/18, Aug/18), six quarterly (Q2/18, Q3/18, Q4/18, Q1/19, Q2/19, Q3/19) and three yearly (Cal-19, Cal-20, Cal-21) futures.

4.5.2 Model comparison

We perform the calibration procedure described in Section 3.4.3, first, for the one-factor model derived by (4.3) by setting the first coefficient to 0, i.e. $\Gamma_1(u, T_1, T_2) \equiv 0$ and, then, for our general two-factor model. We compare both models to the empirical IVs and the one (constant across strikes) generated by Black's model, that we estimate with the same procedure of the other models (except the computation of the price, which can be computed analytically for the Black model). We show the results for the minimization on market prices (see Section 4.4), since the calibration is faster and, even though the case of IVs yields by definition a lower residual, it gives very similar results. The IVs of market and model prices are plotted in Figures 4.1–4.3.

We use the second approach (modified time value) because it has the advantage that we do not have to choose the damping parameter a in (4.10). Also, as stated in Proposition 4.3.1, additional integrability assumptions are needed, that translate, in the case of the two-factor NIG model, into other constraints on the parameters. We have tested the modified option value approach on the one-factor model and found analogous performance to the case of the modified time value, but, as explained, we needed to add more constraints on the parameter set so that the implementation was more involved. After numerical experiments with test parameters, the integral in (4.22) has been truncated at $A = 10$ and computed by an adaptive

Simpson quadrature rule already implemented in MATLAB (as in [102]), which takes into account the oscillatory behavior of the integrand (cf. [45]).

We find that, in general, the optimization routine falls into local minima. However, by selecting a starting condition that is "sufficiently close" to the observed IVs, the minimization converges (cf. [52] for well-posedness of this kind of problems). As a by-product from the estimation of the one-factor model, we derive that, under the risk-neutral measure, futures prices are leptokurtic and have significantly positive skewness. This is reflected also into the shape of empirical IVs, which display a *forward skew* (i.e. higher IVs for out-of-the-money calls). This can be interpreted as a "risk premium" paid by option buyers for securing supply.

Generally, the two-factor model fit is rather accurate for all the strikes and maturities available in the market. However, the "smile" of model IVs of closest to maturity options is slightly more pronounced than what observed. There is a clear improvement of fit from the one-factor to the two-factor model, since, for instance, IVs of far from maturity out-of-the-money calls are underestimated by the one-factor model (and not from the two-factor one). Also, there is a slight anomaly in some at-the-money IVs, that we believe it is due to numerical instability of the Fourier method for $K = F(t)$ (cf. [45]).

4.6 Appendix

4.6.1 Characteristic function of the two-factor NIG model

As already stated in Section 4.2 for general multifactor additive models, the characteristic function of two-factor NIG model is defined as a function of $v \in \mathbb{R}$ by

$$\Psi(t, T, T_1, T_2, v) = \mathbb{E}\left[e^{ivZ(t,T,T_1,T_2)}\,|\mathcal{F}_t\right],$$

where $Z(t, T, T_1, T_2) := F(T, T_1, T_2) - F(t, T_1, T_2)$, and it can be shown to be equal to

$$\log \Psi(t, T, T_1, T_2, v) = \psi_1(t, T; y\Gamma_1(\cdot, T_1, T_2)) + \psi_2(t, T; y\Gamma_2(T_1, T_2)). \tag{4.31}$$

First, we need to compute the cumulant function $\widetilde{\psi}_j(\theta)$ of J_j $(j = 1, 2)$, that can be computed from the corresponding cumulant $\widetilde{\psi}_{L_j}(\theta)$ of L_j as follows. From the definition of J_j we know that

$$\widetilde{\psi}_j(\theta) = \widetilde{\psi}_{L_j}(\theta) - i\theta\left(\chi_j + \int_{|y|\geq 1} y\,\nu_j(dy)\right),$$

where

$$\widetilde{\psi}_{L_j}(\theta) = \delta\left(\sqrt{\alpha_j^2 - \beta_j^2} - \sqrt{\alpha_j^2 - (\beta_j + i\theta)^2}\right) + i\theta m_j.$$

Moreover, since $L_j(1)$ is a NIG distributed random variable, its expected value is

$$\mathbb{E}[L_j(1)] = \left(\chi_j + \int_{|y|\geq 1} y\,\nu_j(dy)\right) = m_j + \frac{\delta_j\beta_j}{\sqrt{\alpha_j^2 - \beta_j^2}},$$

so that, by replacing it in the expression above, we find that

$$\widetilde{\psi}_j(\theta) = \delta_j\left(\sqrt{\alpha_j^2 - \beta_j^2} - \sqrt{\alpha_j^2 - (\beta_j + i\theta)^2} - i\theta\,\frac{\beta_j}{\sqrt{\alpha_j^2 - \beta_j^2}}\right). \tag{4.32}$$

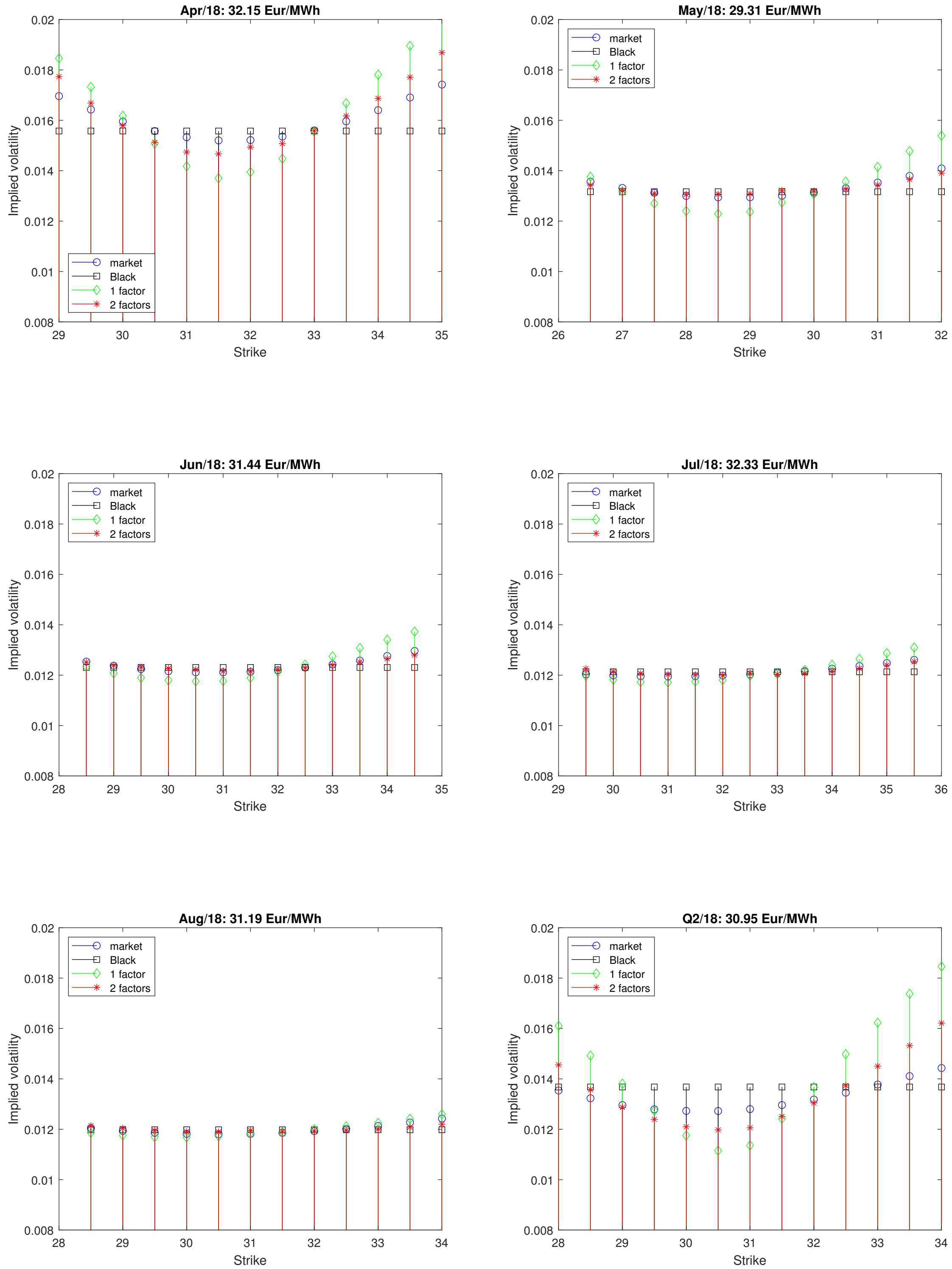


Figure 4.1: Implied volatility for the Black, one-factor and two-factor model compared to the empirical implied volatilities of all the options listed at March 5, 2018 (the corresponding underlying current price is indicated above each plot).

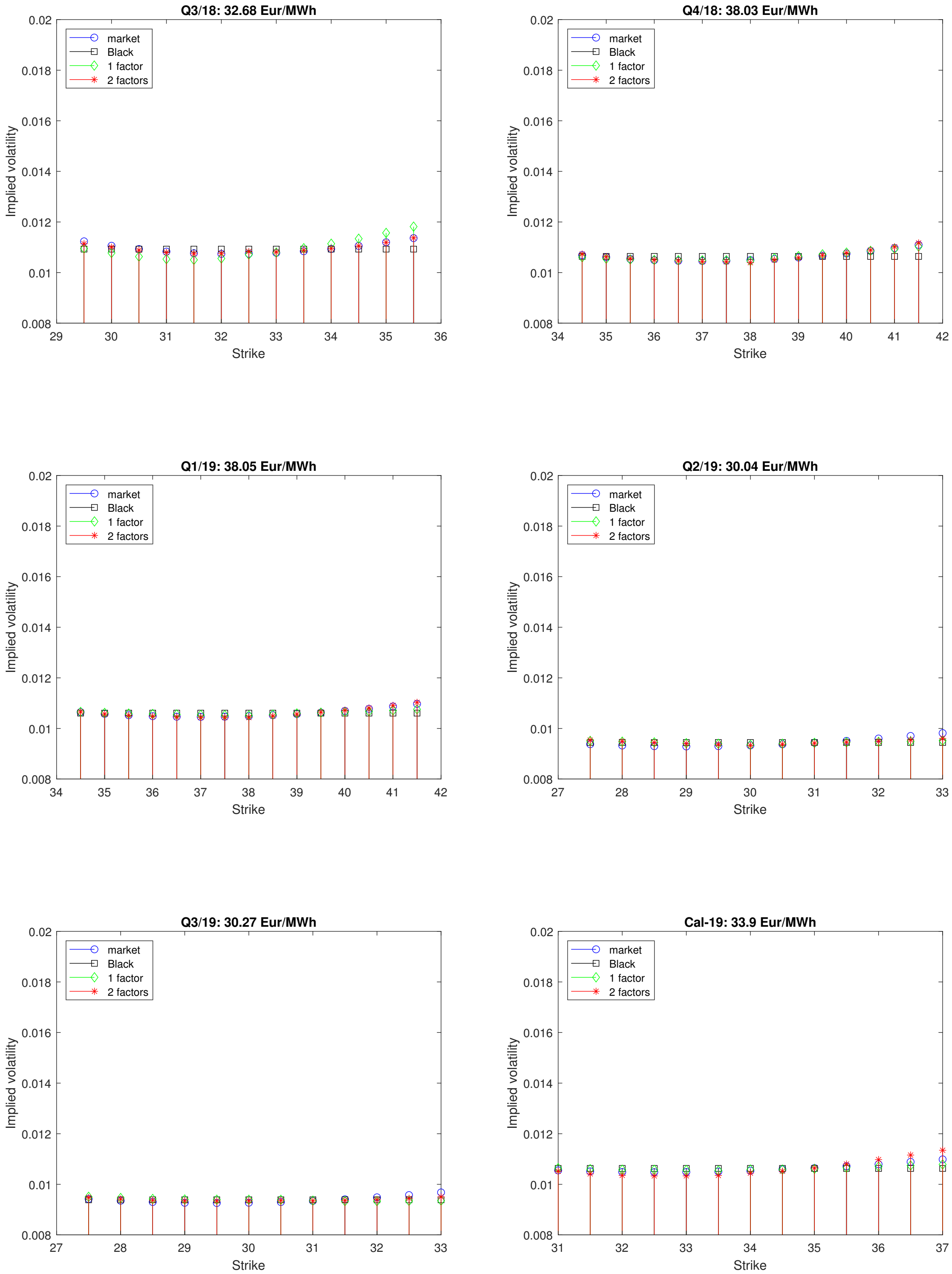


Figure 4.2: Implied volatility for the Black, one-factor and two-factor model compared to the empirical implied volatilities of all the options listed at March 5, 2018 (the corresponding underlying current price is indicated above each plot).

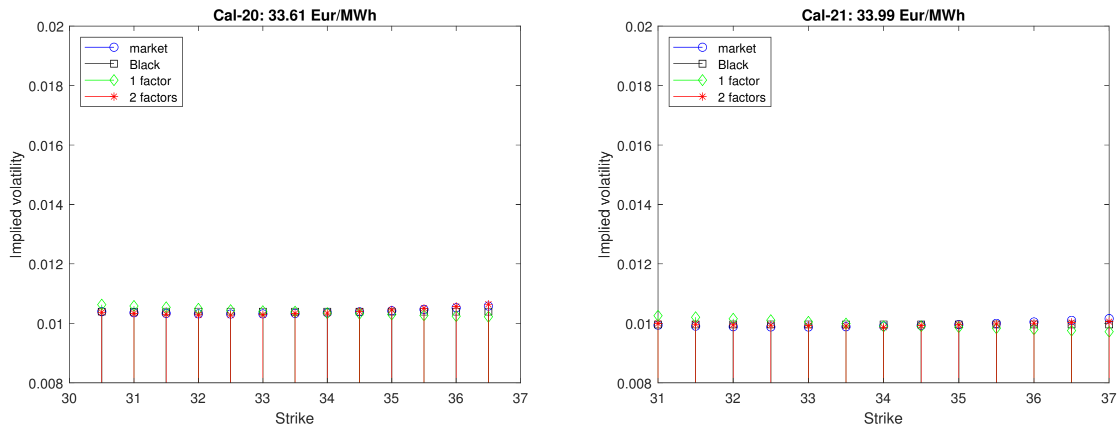


Figure 4.3: Implied volatility for the Black, one-factor and two-factor model compared to the empirical implied volatilities of all the options listed at March 5, 2018 (the corresponding underlying current price is indicated above each plot).

In particular, we observe that $J_j(1)$ is a NIG distributed random variable with parameters $(\alpha_j, \beta_j, \delta_j, -\frac{\delta_j\beta_j}{\sqrt{\alpha_j^2-\beta_j^2}})$. Now, we can compute the two jump components of the cumulant function in (4.31), by inserting the corresponding expressions of $\widetilde{\psi}_j(\theta)$ from (4.32) and Γ_j as in (4.4)–(4.5) for $j = 1, 2$. The second coefficient can be directly computed as

$$\begin{aligned}\psi_2(t,T; y\Gamma_2(T_1,T_2)) = (T-t)\,\delta_2\Big(&\sqrt{\alpha_2^2-\beta_2^2} - \sqrt{\alpha_2^2-(\beta_2+iy\Gamma_2(T_1,T_2))^2}\\ &- iy\Gamma_2(T_1,T_2)\frac{\beta_2}{\sqrt{\alpha_2^2-\beta_2^2}}\Big),\end{aligned} \tag{4.33}$$

while the first requires an integration in time:

$$\begin{aligned}\psi_1(t,T; y\Gamma_1(\cdot,T_1,T_2)) = (T-t)\,\delta_1\sqrt{\alpha_1^2-\beta_1^2}&\\ -\delta_1\int_t^T\sqrt{\alpha_1^2-(\beta_1+iy\Gamma_1(u,T_1,T_2))^2}\,du&\\ -iy\frac{\delta_1\beta_1}{\sqrt{\alpha_1^2-\beta_1^2}}\int_t^T\Gamma_1(u,T_1,T_2)\,du.&\end{aligned} \tag{4.34}$$

Let us recall that, if $\zeta := \Gamma_2(T_1, T_2)$ is positive, by the properties of the NIG distribution (see e.g. [12]), $\zeta \cdot J_2(1)$ is a NIG distributed random variable with parameters $(\alpha_2/\zeta, \beta_2/\zeta, \delta_2\zeta, -\frac{\delta_2\zeta\beta_2}{\sqrt{\alpha_2^2-\beta_2^2}})$. Consequently, it is easy to see that we can assume without loss of generality that $\delta_2 = 1$, so that

$$\begin{aligned}\psi_2(t,T; y\Gamma_2(T_1,T_2)) = (T-t)\Big(&\sqrt{\alpha_2^2-\beta_2^2} - \sqrt{\alpha_2^2-(\beta_2+iy\Gamma_2(T_1,T_2))^2}\\ &- iy\Gamma_2(T_1,T_2)\frac{\beta_2}{\sqrt{\alpha_2^2-\beta_2^2}}\Big),\end{aligned} \tag{4.35}$$

For the same reason we can assume that $\delta_1 = 1$. By replacing $\Gamma_1(u, T_1, T_2) = e^{\mu u}\, \frac{\gamma_1(e^{-\mu T_1} - e^{-\mu T_2})}{\mu(T_2 - T_1)} := e^{\mu u}\, \widetilde{\Gamma}_1(T_1, T_2)$ in (4.34) and integrating, we get

$$\begin{aligned} \psi_1(t, T; y\Gamma_1(\cdot, T_1, T_2)) = (T - t)\, \sqrt{\alpha_1^2 - \beta_1^2} - \eta(c(y, T)) + \eta(c(y, t)) \\ - \frac{iy\, \widetilde{\Gamma}_1(T_1, T_2)\, \beta_1 (e^{\mu T} - e^{\mu t})}{\mu\sqrt{\alpha_1^2 - \beta_1^2}}, \end{aligned} \tag{4.36}$$

where

$$\begin{aligned} \eta(w) :=& \frac{1}{\mu} \left(\sqrt{\alpha_1^2 + w^2} - i\beta_1 \operatorname{arcsinh} \frac{w}{\alpha_1} \right. \\ & \left. - \sqrt{\alpha_1^2 - \beta_1^2}\, \log \frac{2\alpha_1^2 \left(\alpha_1^2 - i\beta_1 w + \sqrt{\alpha_1^2 - \beta_1^2} \sqrt{\alpha_1^2 + w^2} \right)}{(w + i\beta_1)(\alpha_1^2 - \beta_1^2)^{\frac{3}{2}}} \right), \end{aligned} \tag{4.37}$$

$$c(v, u) := v\, \widetilde{\Gamma}_1(T_1, T_2)\, e^{\mu u} - i\beta_1. \tag{4.38}$$

Finally, the cumulant function of Z is explicitly given by

$$\begin{aligned} \log \Psi(t, T, T_1, T_2, y) &= \psi_1(t, T; y\Gamma_1(\cdot, T_1, T_2)) + \psi_2(t, T; y\Gamma_2(T_1, T_2)) \\ &= (T - t) \left\{ \sqrt{\alpha_1^2 - \beta_1^2} + \sqrt{\alpha_2^2 - \beta_2^2} \right. \\ &\quad - iy \left(\frac{\widetilde{\Gamma}_1(T_1, T_2)\beta_1}{\mu\sqrt{\alpha_1^2 - \beta_1^2}}\, \frac{(e^{\mu T} - e^{\mu t})}{T - t} + \frac{\Gamma_2(T_1, T_2)\, \beta_2}{\sqrt{\alpha_2^2 - \beta_2^2}} \right) \\ &\quad \left. - \frac{\eta(c(y, T)) - \eta(c(y, t))}{T - t} - \sqrt{\alpha_2^2 - (\beta_2 + iy\Gamma_2(T_1, T_2))^2} \right\}. \end{aligned}$$

with $\eta(\cdot)$ and $c(\cdot, \cdot)$ as in (4.37)–(4.38).

Bibliography

[1] K. K. Aase. Optimum portfolio diversification in a general continuous-time model. *Stochastic Processes and their Applications*, 18(1):81 – 98, 1984.

[2] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[3] R. Aïd, P. Gruet, and H. Pham. An optimal trading problem in intraday electricity markets. *Math. Financ. Econ.*, 10(1):49–85, 2016.

[4] C. Alexander. Correlation and cointegration in energy markets. In *Managing Energy Price Risk*, pages 291–304. Risk Publications London, 2 edition, 1999.

[5] C. Alexander. *Hedging the Risk of an Energy Futures Portfolio*, chapter 9, pages 117–127. Wiley-Blackwell, 2012.

[6] A. Andresen, S. Koekebakker, and S. Westgaard. Modeling electricity forward prices using the multivariate normal inverse gaussian distribution. *The Journal of Energy Markets*, 3(3):3, 2010.

[7] J. C. Arismendi, J. Back, M. Prokopczuk, R. Paschke, and M. Rudolf. Seasonal Stochastic Volatility: Implications for the pricing of commodity options. *Journal of Banking & Finance*, 66(C):53–65, 2016.

[8] M. Ascheberg, N. Branger, H. Kraft, and F. T. Seifried. When do jumps matter for portfolio optimization? *Quantitative Finance*, 16(8):1297–1311, 2016.

[9] S. Asmussen and J. Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.*, 38(2):482–493, 06 2001.

[10] L. Bachelier. Théorie de la spéculation. *Ann. Sci. École Norm. Sup. (3)*, 17:21–86, 1900.

[11] J. Back, M. Prokopczuk, and M. Rudolf. Seasonality and the valuation of commodity options. *Journal of Banking & Finance*, 37(2):273–290, 2013.

[12] O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.*, 24(1):1–13, 1997.

[13] O. E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2(1):41–68, 1997.

[14] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(2):167–241, 2001.

[15] A. Bensoussan and J.-L. Lions. *Contrôle impulsionnel et inéquations quasi variationnelles*, volume 11 of *Méthodes Mathématiques de l'Informatique [Mathematical Methods of Information Science]*. Gauthier-Villars, Paris, 1982.

[16] F. E. Benth. Cointegrated commodity markets and pricing of derivatives in a non-Gaussian framework. In *Advanced Modelling in Mathematical Finance*, volume 189 of *Springer Proc. Math. Stat.*, pages 477–496. Springer, Cham, 2016.

[17] F. E. Benth, Á. Cartea, and R. Kiesel. 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, 2008.

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

[19] F. E. Benth and K. H. Karlsen. A note on Merton's portfolio selection problem for the Schwartz mean-reversion model. *Stoch. Anal. Appl.*, 23(4):687–704, 2005.

[20] F. E. Benth, K. H. Karlsen, and K. Reikvam. Merton's portfolio optimization problem in a Black and Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type. *Math. Finance*, 13(2):215–244, 2003.

[21] F. E. Benth, R. Kiesel, and A. Nazarova. A critical empirical study of three electricity spot price models. *Energy Economics*, 34(5):1589 – 1616, 2012.

[22] F. E. Benth and S. Koekebakker. Stochastic modeling of financial electricity contracts. Working paper, Department of Mathematics, University of Oslo, 2005.

[23] F. E. Benth and S. Koekebakker. Stochastic modeling of financial electricity contracts. *Energy Economics*, 30(3):1116 – 1157, 2008.

[24] F. E. Benth and R. Kufakunesu. Pricing of exotic energy derivatives based on arithmetic spot models. *International Journal of Theoretical and Applied Finance*, 12(04):491–506, 2009.

[25] F. E. Benth and S. Ortiz-Latorre. A pricing measure to explain the risk premium in power markets. *SIAM J. Financial Math.*, 5(1):685–728, 2014.

[26] F. E. Benth, M. Piccirilli, and T. Vargiolu. Additive energy forward curves in a Heath-Jarrow-Morton framework. Preprint, 2018.

[27] F. E. Benth and J. Šaltytė Benth. *Modeling and Pricing in Financial Markets for Weather Derivatives.* World Scientific Publishing Co. Pte. Ltd., 2012.

[28] F. E. Benth and M. D. Schmeck. Stability of Merton's portfolio optimization problem for Lévy models. *Stochastics*, 85(5):833–858, 2013.

[29] F. E. Benth and M. D. Schmeck. Pricing and hedging options in energy markets using Black-76. *Journal of Energy Markets*, 7(2):35–69, 2014.

[30] F. E. Benth and M. D. Schmeck. *Pricing Futures and Options in Electricity Markets*, pages 233–260. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[31] F. E. Benth and C. Sgarra. The risk premium and the Esscher transform in power markets. *Stoch. Anal. Appl.*, 30(1):20–43, 2012.

[32] F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. Advanced Series on Statistical Science & Applied Probability, 11. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[33] F. E. Benth and J. Šaltytė Benth. The normal inverse gaussian distribution and spot price modelling in energy markets. *International Journal of Theoretical and Applied Finance*, 07(02):177–192, 2004.

[34] H. Bessembinder and M. L. Lemmon. Equilibrium pricing and optimal hedging in electricity forward markets. *Journal of Finance*, 57(3):1347–1382, 2002.

[35] N. H. Bingham and R. Kiesel. *Risk-neutral valuation*. Springer Finance. Springer-Verlag London, Ltd., London, second edition, 2004. Pricing and hedging of financial derivatives.

[36] O. H. Birkelund, E. Haugom, P. Molnár, M. Opdal, and S. Westgaard. A comparison of implied and realized volatility in the Nordic power forward market. *Energy Economics*, 48:288 – 294, 2015.

[37] P. Bjerksund, H. Rasmussen, and G. Stensland. Valuation and risk management in the Norwegian electricity market. In *Energy, Natural Resources and Environmental Economics*, pages 167–185. Springer, 2010.

[38] F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3(1):167 – 179, 1976.

[39] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81(3):637–654, 1973.

[40] S. Blanes, F. Casas, J.A. Oteo, and J. Ros. The magnus expansion and some of its applications. *Physics Reports*, 470(5):151 – 238, 2009.

[41] D. Brigo and F. Mercurio. *Interest rate models—theory and practice*. Springer Finance. Springer-Verlag, Berlin, second edition, 2006. With smile, inflation and credit.

[42] P. J. Brockwell and E. Schlemm. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *J. Multivariate Anal.*, 115:217–251, 2013.

[43] E. Broszkiewicz-Suwaj and A. Weron. Calibration of the multi-factor HJM model for energy market. *Acta Physica Polonica B*, 37:1455, 2006.

[44] R. Börger. Energy-related commodity futures: statistics, models and derivatives. 2007. Ph.D. Dissertation.

[45] P. Carr and D. B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.

[46] Á. Cartea, M. Flora, G. Slavov, and T. Vargiolu. Optimal cross-border electricity trading. In preparation, 2018.

[47] Á. Cartea and T. Williams. UK gas markets: The market price of risk and applications to multiple interruptible supply contracts. *Energy Economics*, 30(3):829 – 846, 2008.

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

[49] L. Clewlow and C. Strickland. Valuing energy options in a one factor model fitted to forward prices. Research Paper Series 10, Quantitative Finance Research Centre, University of Technology, Sydney, April 1999.

[50] S. N. Cohen and R. J. Elliott. *Stochastic calculus and applications*. Probability and its Applications (New York). Springer, New York, second edition, 2015.

[51] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[52] R. Cont and P. Tankov. Nonparametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7:1–49, 2004.

[53] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance Stoch.*, 9(3):299–325, 2005.

[54] J. M. Danskin. The theory of max − min, with applications. *SIAM J. Appl. Math.*, 14:641–664, 1966.

[55] C. De Jong and S. Schneider. Cointegration between gas and power spot prices. *The Journal of Energy Markets*, 2(3):27, 2009.

[56] Ł. Delong and C. Klüppelberg. Optimal investment and consumption in a Black-Scholes market with Lévy-driven stochastic coefficients. *Ann. Appl. Probab.*, 18(3):879–908, 2008.

[57] E. Edoli, M. Gallana, and T. Vargiolu. Optimal intra-day power trading with a Gaussian additive process. *Journal of Energy Markets*, 10(4):23–42, 2017.

[58] E. Edoli, D. Tasinato, and T. Vargiolu. Calibration of a multifactor model for the forward markets of several commodities. *Optimization*, 62(11):1553–1574, 2013.

[59] EEX. Trading Volume in Power Derivatives, June 2018. `https://www.eex.com/en/about/newsroom/news-detail/eex-trading-results-for-june-2018/83548`.

[60] EIA. `https://www.eia.gov/naturalgas/weekly/archivenew_ngwu/2018/05_10/`.

[61] EPEX. Negative Prices. `http://www.epexspot.com/en/company-info/basics_of_the_power_market/negative_prices`.

[62] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[63] V. Fanelli, L. Maddalena, and S. Musti. Modelling electricity futures prices using seasonal path-dependent volatility. *Applied Energy*, 173:92 – 102, 2016.

[64] V. Fanelli and M. D. Schmeck. On the seasonality in the implied volatility of electricity options. Preprint, 2018.

[65] E. Fanone, A. Gamba, and M. Prokopczuk. The case of negative day-ahead electricity prices. *Energy Economics*, 35(C):22–34, 2013.

[66] S. Farinelli and L. Tibiletti. Hydroassets portfolio management for intraday electricity trading from a discrete time stochastic optimization perspective. Technical report, 2017.

[67] W. Farkas, E. Gourier, R. Huitema, and C. Necula. A two-factor cointegrated commodity price model with an application to spread option pricing. *Journal of Banking & Finance*, 77:249 – 268, 2017.

[68] Financial Post. `https://www.neb-one.gc.ca/nrg/sttstc/ntrlgslqds/rprt/2016/2016prpn-eng.html`.

[69] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[70] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.

[71] E. Garnier and R. Madlener. Balancing forecast errors in continuous-trade intraday markets. *Energy Systems*, 6(3):361–388, Sep 2015.

[72] H. Geman. *Commodities and Commodity Derivatives. Modeling and Pricing for Agriculturals, Metals and Energy*. HAL, 2005.

[73] H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3):1225–1262, 2006.

[74] T. Goll and J. Kallsen. Optimal portfolios for logarithmic utility. *Stochastic Processes and their Applications*, 89(1):31 – 48, 2000.

[75] T. Goll and J. Kallsen. A complete explicit solution to the log-optimal portfolio problem. *Ann. Appl. Probab.*, 13(2):774–799, 2003.

[76] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992.

[77] A. Henriot. Market design with centralized wind power management: Handling low-predictability in intraday markets. *The Energy Journal*, 35(1), 2014.

[78] W. Hinderks, A. Wagner, and R. Korn. A structural Heath-Jarrow-Morton framework for consistent intraday, spot, and futures electricity prices. Technical report, Available at ArXiV `https://arxiv.org/abs/1803.08831`, 2018.

[79] J. Hinz, L. von Grafenstein, M. Verschuere, and M. Wilhelm. Pricing electricity risk by interest rate methods. *Quant. Finance*, 5(1):49–60, 2005.

[80] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981.

[81] EEX indices. `https://www.eex.com/en/trading/indices`.

[82] E. Jaeck and D. Lautier. Volatility in electricity derivative markets: The Samuelson effect revisited. *Energy Economics*, 59:300 – 313, 2016.

[83] S. Jaimungal and V. Surkov. Lévy-based cross-commodity models and derivative valuation. *SIAM Journal on Financial Mathematics*, 2(1):464–487, 2011.

[84] Y. Jao, C. Ma, S. Scotti, and C. Sgarra. A branching process approach to power markets. Preprint, 2017.

[85] J. Kallsen. Optimal portfolios for exponential lévy processes. *Mathematical Methods of Operations Research*, 51(3):357–374, Aug 2000.

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

[87] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, Second edition, 1991.

[88] R. Kiesel and F. Paraschiv. Econometric analysis of 15-minute intraday electricity prices. *Energy Economics*, 64:77 – 90, 2017.

[89] R. Kiesel, G. Schindlmayr, and R. H. Börger. A two-factor model for the electricity forward market. *Quantitative Finance*, 9(3):279–287, 2009.

[90] S. Koekebakker. An arithmetic forward curve model for the electricity market. Working Paper 2003, Agder University College, Norway, 2003.

[91] S. P. Kolos and E. I. Ronn. Estimating the commodity market price of risk for energy prices. *Energy Economics*, 30(2):621 – 641, 2008.

[92] K. Larsson and M. Nossman. Risk management of co-integrated commodities. Also available from KYOS web-site publications: `http://www.kyos.com/risk-pricing/risk-and-pricing-publications/`, 2011.

[93] L. Latini, M. Piccirilli, and T. Vargiolu. Mean-reverting no-arbitrage additive models for forward curves in energy markets. Energy Economics. In press. Available online at `https://doi.org/10.1016/j.eneco.2018.03.001`, 2018.

[94] D. Lépingle and J. Mémin. Sur l'intégrabilité uniforme des martingales exponentielles. *Z. Wahrsch. Verw. Gebiete*, 42(3):175–203, 1978.

[95] J. J. Lucia and E. S. Schwartz. Electricity prices and power derivatives: Evidence from the Nordic power exchange. *Review of Derivatives Research*, 5(1):5–50, 2002.

[96] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–57, 1969.

[97] R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory*, 3(4):373–413, 1971.

[98] R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[99] T. Meyer-Brandis and P. Tankov. Multi-factor jump-diffusion models of electricity prices. *Int. J. Theor. Appl. Finance*, 11(5):503–528, 2008.

[100] M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling*, volume 36 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997.

[101] E. Nastasi, A. Pallavicini, and G. Sartorelli. Smile modelling in commodity markets. Preprint, 2018.

[102] N. K. Nomikos and O. A. Soldatos. Analysis of model implied volatility for jump diffusion models: Empirical evidence from the Nordpool market. *Energy Economics*, 32(2):302–312, 2010.

[103] A. A. Novikov. A certain identity for stochastic integrals. *Teor. Verojatnost. i Primenen.*, 17:761–765, 1972.

[104] B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Universitext. Springer, Berlin, Second edition, 2007.

[105] I. Oliva and R. Renò. Optimal portfolio allocation with volatility and co-jump risk that markowitz would like. *Journal of Economic Dynamics and Control*, 94:242 – 256, 2018.

[106] R. Paschke and M. Prokopczuk. Integrating multiple commodities in a model of stochastic price dynamics. *The Journal of Energy Markets*, 2(3):47–82, 2009.

[107] L. Pasin and T. Vargiolu. Optimal portfolio for CRRA utility functions when risky assets are exponential additive processes. *Economic Notes*, 39(s1):65–90, 2010.

[108] H. Pham. Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *J. Math. Systems Estim. Control*, 8(1):27 pp. (electronic), 1998.

[109] H. Pham. Smooth solutions to optimal investment models with stochastic volatilities and portfolio constraints. *Appl. Math. Optim.*, 46(1):55–78, 2002.

[110] M. Piccirilli, M. D. Schmeck, and T. Vargiolu. Capturing the power options smile by an additive two-factor model for futures prices. Preprint, 2018.

[111] M. Piccirilli and T. Vargiolu. Optimal portfolio in intraday electricity markets modelled by Lévy-Ornstein-Uhlenbeck processes. Preprint, 2018.

[112] N. Privault. Combinatorics of Poisson stochastic integrals with random integrands. In *Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry*, pages 37–80. Springer International Publishing, Cham, 2016.

[113] P. E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, Second edition, 2004. Stochastic Modelling and Applied Probability.

[114] J. Riordan. Moment recurrence relations for binomial, Poisson and hypergeometric frequency distributions. *The Annals of Mathematical Statistics*, 8(2):103–111, 1937.

[115] P. A. Samuelson. Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6:41–49, 1965.

[116] P. A. Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6:13–31, 1965.

[117] K. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[118] L. Schneider and B. Tavin. From the samuelson volatility effect to a samuelson correlation effect: An analysis of crude oil calendar spread options. *Journal of Banking & Finance*, 2016.

[119] L. Schneider and B. Tavin. The Samuelson Effect and Seasonal Stochastic Volatility in Agricultural Futures Markets. Preprint, 2018.

[120] E. Schwartz and J. E. Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000.

[121] A. Serletis. *Quantitative and Empirical Analysis of Energy Markets*. World Scientific Publishing Co. Pte. Ltd., 2012.

[122] H. Simon. Negative prices - a new phenomenon. *The Journal of Professional Pricing*, Fourth Quarter:18–21, 2016.

[123] Z. Tan and P. Tankov. Optimal trading policies for wind energy producer. *SIAM Journal on Financial Mathematics*, 9(1):315–346, 2018.

[124] S. J. Taylor. *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press, 2011.

[125] R. Weron. Market price of risk implied by Asian-style electricity options and futures. *Energy Economics*, 30(3):1098–1115, 2008.

[126] G. Wolff and S. Feuerriegel. Short-term dynamics of day-ahead and intraday electricity prices. *International Journal of Energy Sector Management*, 11(4):557–573, 2017.

[127] T. Wottka. Volatility skews implied by a multi-technology bid stack model for electricity markets. Preprint, 2017.

[128] T. Zariphopoulou. Optimal investment and consumption models with non-linear stock dynamics. *Math. Methods Oper. Res.*, 50(2):271–296, 1999. Financial optimization.