TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

Arbitrage Conditions for Electricity Markets with Production and Storage

Raimund Kovacevic

Research Report 2018-03

March 2018

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-4
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

ARBITRAGE CONDITIONS FOR ELECTRICITY MARKETS WITH PRODUCTION AND STORAGE

RAIMUND KOVACEVIC

Abstract. We consider a market at which electricity is produced from fuel. Several generators, fuel storage, and the related costs are considered. Based on stochastic optimization in Banach spaces, we derive a necessary and a sufficient no-arbitrage conditions and analyze them further in the context of (potentially nonlinearly) autoregressive price models. For this large class of statistical models, it is found that the necessary condition can be rejected only in very unrealistic cases. The sufficient condition, however, leads to a simple logical constraint that can be used for restricted parameter estimation and for testing the hypothesis of absence of arbitrage. Finally, we analyze the consequences of these findings for contract valuation and for tree construction in the stochastic optimization context.

1. Introduction

Nowadays the electricity sectors of many countries are based on thriving markets with specialized players. Electricity is traded on exchanges, with its price determined by supply and demand. Such markets usually are very liquid and in many regards comparable with financial markets. However, still there are unique frictions not existent on financial markets (or other commodity markets): In particular, electricity is produced from fuels but cannot be converted back to fuels. Electricity also cannot be stored in large quantities at the time being. Furthermore, all kinds of restrictions on physical fuel storage and generation capacity are relevant for the production process. Finally, produced and used electric power has to be balanced immediately in an electrical network and deviations may lead to damaged equipment or even breakdown of the net.

Still, the notions of arbitrage and market completeness - cornerstones of modern finance - can be applied also to electricity markets. Basically, a market is arbitrage free if riskless profits are possible and it is complete if any relevant payoff can be replicated from the basic traded securities ("underlyings"), contracts or commodities, traded at this market. Essentially, a financial market is arbitrage free if and only if there exists an equivalent (local) martingale measure, such that all basic securities can be priced by taking expectation of their future values with respect to this measure. An arbitrage free market is complete if and only if there exists a unique martingale measure. On complete markets, every contingent claim is attainable by hedging portfolios, and financial derivatives can be priced by calculating the expected discounted value of the derivatives payoff with respect to the unique

Key words and phrases. *Electricity production, arbitrage, stochastic discount factor, duality theory.*

Institute for Statistics and Mathematical Methods in Economics, Vienna University of Technology, raimund.kovacevic@tuwien.ac.at.

martingale measure. On incomplete markets there exists claims that can not be replicated by the basic traded assets. Still, under no-arbitrage there exist equivalent martingale measure but uniqueness does not hold, in particular prices from taking expectation are not unique.

Because of the discussed frictions, electricity markets are not complete. While some submarkets for electricity (in particular futures markets) are organized as financial markets, even in this case the delivery profiles of traded futures cannot fully replicate typical OTC-traded delivery profile (although hedging by futures contracts is an important approach in practice, see e.g. [4]).

Still, the question remains, whether electricity markets are arbitrage free. In [11], no-arbitrage conditions for an electricity market with electricity generation from fuel and fuel storage were derived analytically. These results are based on duality theory for cone-constrained optimization in Banach spaces. It turns out that existence of a martingale measure has to be replaced by more complex requirements, as discussed in section 2 below. These no-arbitrage conditions are then used in [11] in order to derive super-hedging and acceptability prices and values for delivery contracts with random delivery patterns under several types of assumptions. It shows that the smallest value or price such that a contract does not lead to an unfavorable outcome can be given in terms of expectations w.r.t. equivalent measures or stochastic discount factors. This approach builds on and generalizes the ideas in [10], [7], [15, 16] and [17], where financial markets instead of electricity markets are considered.

The present paper goes back one step and deepens the discussion of no-arbitrage. We start with deriving no arbitrage conditions for a more general model (compared to [11]), which includes several generating units with different production efficiency and takes into account the effects of storage costs for fuel. The main question then is, how restrictive the no-arbitrage conditions are. In particular, when estimating reasonable parameter values from data, one may ask the question whether the no-arbitrage conditions put any restrictions on the parameters. In this paper we analyze these questions for price processes with (potentially nonlinear) autoregressive structure. Moreover we discuss the consequences for the valuation of electricity delivery contracts and for the construction of scenario trees for stochastic optimization.

The paper is organized as follows: Section 2 uses a basic optimization problem to derive and analyze no-arbitrage conditions for a model with spot prices for fuel and electricity, when electricity can be produced with given efficiency. Several generating units as well as storage costs are considered. In section 3 we analyze the problem of testing the assumption of no-arbitrage in such a market based on the necessary condition. In section 4 we analyze the consequences of the sufficient condition and give an outlook to valuation issues and tree construction and discuss our main results in this contexts. Finally, section 5 concludes the paper.

2. No-arbitrage conditions for an electricity market with production and storage

In the following we consider a stochastic process $X_t^f(\omega)$ of fuel prices and a stochastic process $X_t^e(\omega)$ of electricity prices. Both price processes are defined on a filtered probability space $\mathfrak{Y} = \left(\Omega, \mathcal{F}, \mathfrak{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ in discrete time $t = 0, 1, \ldots, T$. For simplicity we use constant time increments, e.g. hours, days or weeks. At the beginning, the σ-algebra $\mathcal{F}_0$ is the trivial σ-algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$. The filtration $\mathfrak{F}$ may be generated by the price processes, but this is not a necessary requirement. Time T denotes the end of the planning horizon. In order to simplify notation we use the sets $\mathcal{T} = \{0, 1, \dots, T\}$, $\mathcal{T}_0 = \{0, 1, \dots, T-1\}$, $\mathcal{T}_1 = \{1, \dots, T\}$ and $\mathcal{T}_1^{T-1} = \{1, \dots, T-1\}$. As in reality, fuel prices are assumed to be almost surely nonnegative. Electricity prices may be negative with positive probability. Both, the energy content of the fuel and the electrical energy are measured in MWh and prices are stated in currency units per MWh.

Immediately before taking decisions at time t, the producer owns a cash position c_t with associated interest rate $r \geq 0$ (per period) and an amount of fuel s_t [MWh]. We will use the notation $R = (1+r)$. The producer then takes his decisions at time t. First he decides the amount z_t [MWh] of fuel traded at the fuel market at price X_t^f [currency units per MWh]. This trade happens at (or immediately after) time t. Positive values of z_t indicate that an amount of fuel is bought, negative values indicate selling of fuel. Electricity is produced by generators $i \in \{1, 2, \dots, I\}$. The amounts y_{it} [MWh] of electricity produced with generators i over period $[t, t+1]$ is planned in advance at time t. It is sold at time $t+1$ at price X_{t+1}^e, immediately before observing the new cash position. Generator i produces with efficiency η_i and we denote the set of efficiencies by $\eta = (\eta_1, \dots, \eta_I)$. Hence the amount of fuel burned for producing electricity is given by $\sum_i \eta_i^{-1} y_{it}$ [MWh]. Note that this generalizes the setup in [11] , where only one generating unit was considered. Clearly, electricity production y_t and fuel storage s_t are almost surely nonnegative.

In order to model storage costs Ψ_t (payable at time t) that are proportional to the stored amount, we use the simple specification

$$\Psi_t = \psi \frac{s_t + s_{t-1}}{2},$$

where $\psi \geq 0$ is the related cost factor. Here it is assumed that the storage is filled, respectively emptied uniformly over time.

In the following, in order to simplify notation, all equations and inequalities involving random variables are to be understood as inequalities with respect to the cone of almost surely nonnegative random variables, i.e. it is assumed that they hold almost surely. We assume that the fuel price and the electricity price are essentially bounded, i.e. $X_t^f, X_t^e \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$. Later on, we will write $X_{[t]}^f, X_{[t]}^e$ in order to denote relevant price histories up to time t. The decision processes y_t and z_t as well as the decision processes c_t and s_t are considered as real valued random processes defined on $\mathfrak{Y}$ and we assume that they are integrable, i.e. $y_t, z_t, c_t, s_t \in L^1(\Omega, \mathcal{F}_t, \mathbb{P})$. In particular they are also adapted to the filtration $\mathfrak{F}$, which means that decisions at time t are based on information available at this time. Because $\mathcal{F}_0$ is the trivial σ-algebra, the starting values c_0, s_0, y_0, z_0 take deterministic values. This setup will allow to use Lagrange multipliers from $L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, which can be identified with the dual space of $L^1(\Omega, \mathcal{F}_t, \mathbb{P})$.

We adapt the concepts of a self financing strategy and of an η-arbitrage, as stated in [11] as extensions of the usual definitions (see e.g. [3] definitions 2.14, 2.15), to the new situation with several generating units.

Definition 2.1. A strategy $\{y_t, z_t\}_{t \geq 0}$ with cash position c_t and fuel storage s_t, where $y_t \geq 0$ and $s_t \geq 0$, is self financing if the following conditions hold almost surely for all $t \in \mathcal{T}_1$:

$$c_t = \left(c_{t-1} - z_{t-1} X^f_{t-1}\right) R + X^e_t \sum_{i=1}^{I} y_{it-1} - \psi \frac{s_t + s_{t-1}}{2}, \tag{2.1}$$

$$s_t = s_{t-1} - \sum_{i=1}^{I} y_{i\,t-1} \eta_i^{-1} + z_{t-1}. \tag{2.2}$$

At any time t the asset value of a strategy is given by

$$V^\eta_t = c_t + X^f_t \cdot s_t$$

The first equation models a cash position (an interest paying account) which changes when fuel is bought, electricity is sold and storage costs are paid. The second equation is related to a fuel storage, which is reduced, when electricity is generated and which is filled up by buying fuel on the market.

Definition 2.2. An η-arbitrage for a market $\left\{X^e_t, X^f_t\right\}$ is a self financing strategy $\{y_t, z_t\}_{t\geq 0}$ with

$$V^\eta_0 \leq 0, \tag{2.3}$$

$$\mathbb{P}\left(V^\eta_T \geq 0\right) = 1. \tag{2.4}$$

$$\mathbb{P}\left(V^\eta_T > 0\right) > 0. \tag{2.5}$$

We call a market $\left\{X^e_t, X^f_t\right\}$ η-arbitrage free, if no η-arbitrage exists.

Similar to [11], the following optimization problem (accounting for several generating units) can be used to detect arbitrage strategies in the described setup.

$$\max_{y,z,c,s} \mathbb{E}^{\mathbb{P}} \left[c_T + X^f_T s_T\right] \tag{2.6}$$

subject to:

$$\begin{aligned}
(t \in \mathcal{T}_1) &: c_t = \left(c_{t-1} - z_{t-1} X^f_{t-1}\right) R + X^e_t \sum_{i=1}^{I} y_{it-1} - \psi \frac{s_t + s_{t-1}}{2} \\
(t \in \mathcal{T}_1) &: s_t = s_{t-1} - \sum_{i=1}^{I} \eta_i^{-1} y_{i\,t-1} + z_{t-1} \\
& c_0 + X^f_0 s_0 \leq 0 \\
& c_T + X^f_T s_T \geq 0 \\
(t \in \mathcal{T}\ ) &: s_t \geq 0 \\
(t \in \mathcal{T}_0) &: y_t \geq 0
\end{aligned}$$

Note that there always exists a solution for this optimization problem because setting all decision variables to zero is feasible. Moreover the feasible set is a pointed cone. Problem (2.6) is formulated without upper bounds on storage and production. However, because of positive homogeneity, a strategy which leads to a positive end value with positive probability can be scaled in a way such that either the scaled solution leads to an infinite expectation without upper bounds or such that all upper bounds are observed and at least one upper bound is reached with positive probability at some point in time. Therefore for a pure test of η-arbitrage the upper bounds are not relevant.

Using the same arguments as in [11], Lemma 1, we can state that an (2.6) η-arbitrage for a market $\left\{X_t^e, X_t^f\right\}$ exists if and only if (2.6) is unbounded. This fact can be used to characterize arbitrage further. Let $\eta_{max} = \max\{\eta_1, \ldots, \eta_I\}$ be the efficiency of the most efficient generating unit at the market. Then the following holds.

Proposition 2.3. *A market $\left\{X_t^e, X_t^f\right\}$ is η-arbitrage free in the described setup if and only if there exist adapted stochastic processes $\{\xi_t, \lambda_t\}$ with the following properties:*

A1: $\xi_t,\lambda_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$ *for each* $t \in \mathcal{T}_1$.
A2: $\xi_t > 0$
A3: $R^{t+1}\,\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}|\mathcal{F}_t\right] = R^t\xi_t$ *for* $t = 1, \ldots, T-1$, *and* $R\,\mathbb{E}^{\mathbb{P}}\left[\xi_1\right] = 1$
A4: $\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}X_{t+1}^e|\mathcal{F}_t\right] \le \eta_{max}^{-1}\xi_t X_t^f$ *for* $t \in \mathcal{T}_0$
A5: $\mathbb{E}^{\mathbb{P}}\left[\lambda_{t+1}|\mathcal{F}_t\right] = \xi_t \cdot X_t^f$ *for* $t \in \mathcal{T}_0$ *and* $\mathbb{E}^{\mathbb{P}}\left[\lambda_1\right] = X_0^f$
A6: $\xi_t \cdot \left[X_t^f - \frac{\psi}{2}\left(1+\frac{1}{R}\right)\right] \le \lambda_t$ *for* $t \in \mathcal{T}_1^{T-1}$ *and* $\xi_T\left[X_T^f - \frac{\psi}{2}\right] \le \lambda_T$

Proof. The Lagrangian of problem (2.6) can be written as

(2.7)

$$\begin{aligned}
L(y,z,c,s;\xi,\lambda,\zeta,\gamma) = {} & \mathbb{E}^{\mathbb{P}}\left[c_T + X_T^f s_T\right] \\
& + \mathbb{E}^{\mathbb{P}}\left[\zeta\left(c_T + X_T^f s_T\right)\right] \\
& + \sum_{t=1}^{T}\mathbb{E}^{\mathbb{P}}\left[\xi_t\left(Rc_{t-1} - c_t + X_t^e\sum_{i=1}^{I} y_{it-1} - R\,X_{t-1}^f z_{t-1}\right)\right] \\
& + \sum_{t=1}^{T}\mathbb{E}^{\mathbb{P}}\left[\lambda_t\left(s_{t-1} - s_t - \sum_{i=1}^{I}\eta_i^{-1}y_{i\,t-1} + z_{t-1} - \psi\frac{s_t + s_{t-1}}{2}\right)\right] \\
& - \gamma\left(c_0 + X_0^f s_0\right),
\end{aligned}$$

where $\gamma \ge 0$ is a real number, $\zeta \ge 0$ a $\mathcal{F}_T$-measurable essentially bounded random variable, ξ_t and λ_t are $\mathcal{F}_t$-measurable and essentially bounded, i.e. $\zeta \in L^\infty(\Omega, \mathcal{F}_T, \mathbb{P})$ and $\xi_t, \lambda_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$. These spaces are chosen as the dual space to $L^1(\Omega, \mathcal{F}_t, \mathbb{P})$, because all summands in (2.7) are elements of $L^1(\Omega, \mathcal{F}_t, \mathbb{P})$ under our basic assumptions.

It is possible to rearrange (2.7) in the following way:

(2.8)

$$
\begin{aligned}
L(y,z,c,s;\xi,\lambda,\zeta,\gamma) = {} & \mathbb{E}^{\mathbb{P}}\left[c_T\left(1+\zeta-\xi_T\right)\right] + \mathbb{E}^{\mathbb{P}}\left[s_T\left(X^f_T(1+\zeta)-\lambda_T-\frac{\psi}{2}s_T\xi_T\right)\right] \\
& + c_0\left(\mathbb{E}^{\mathbb{P}}\left[\xi_1\right]R-\gamma\right) + s_0\left(\mathbb{E}^{\mathbb{P}}\left[\lambda_1\right]-\gamma X^f_0-\frac{\psi}{2}\xi_1\right) \\
& + y_0\mathbb{E}^{\mathbb{P}}\left[\xi_1 X^e_1-\lambda_1\eta^{-1}\right] + z_0\mathbb{E}^{\mathbb{P}}\left[\lambda_1-R\,\xi_1 X^f_0\right] \\
& + \sum_{t=1}^{T-1}\mathbb{E}^{\mathbb{P}}\left[c_t\left(R\,\xi_{t+1}-\xi_t\right)\right] + \sum_{t=1}^{T-1}\mathbb{E}^{\mathbb{P}}\left[s_t\left(\lambda_{t+1}-\lambda_t-\frac{\psi}{2}\left(\xi_t+\xi_{t+1}\right)\right)\right] \\
& + \sum_{i=1}^{I}\sum_{t=1}^{T-1}\mathbb{E}^{\mathbb{P}}\left[y_{i\,t}\left(\xi_{t+1}X^e_{t+1}-\eta_i^{-1}\lambda_{t+1}\right)\right] \\
& + \sum_{t=1}^{T-1}\mathbb{E}^{\mathbb{P}}\left[z_t\left(\lambda_{t+1}-R\,\xi_{t+1}X^f_t\right)\right]
\end{aligned}
$$

Using (2.8), the "tower property" of conditional expectation and keeping in mind $y_t, s_t \geq 0$, the dual function

$$
\max_{y\geq 0,z,c,s\geq 0} L(y,z,c,s;\xi,\lambda,\zeta,\gamma), \tag{2.9}
$$

is bounded (in fact zero almost surely) if and only if the following conditions hold:

$$\zeta \geq 0 \tag{2.10}$$

$$\xi_T = 1 + \zeta \tag{2.11}$$

$$\lambda_T + \frac{\psi}{2}\xi_T \geq (1+\zeta)\, X_T^f \tag{2.12}$$

$$\gamma \geq 0 \tag{2.13}$$

$$\mathbb{E}^{\mathbb{P}}[\xi_1]\; R = \gamma \tag{2.14}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_1] \leq \gamma X_0^f + \frac{\psi}{2}\mathbb{E}[\xi_1] \tag{2.15}$$

$$\mathbb{E}^{\mathbb{P}}[\xi_1 X_1^e] \leq \eta^{-1}\mathbb{E}^{\mathbb{P}}[\lambda_1] \tag{2.16}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_1] = R\,\mathbb{E}^{\mathbb{P}}[\xi_1]\, X_0^f \tag{2.17}$$

$$R\mathbb{E}^{\mathbb{P}}[\xi_{t+1}|\mathcal{F}_t] = \xi_t \qquad \text{for } t \in \mathcal{T}_1^{T-1} \tag{2.18}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] \leq \lambda_t + \frac{\psi}{2}\left(\xi_t + \mathbb{E}[\xi_{t+1}|\mathcal{F}_t]\right) \qquad \text{for } t \in \mathcal{T}_1^{T-1} \tag{2.19}$$

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}X_{t+1}^e|\mathcal{F}_t\right] \leq \eta_i^{-1}\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] \qquad \text{for all } i \in \{1,\dots,I\} \text{ and } t \in \mathcal{T}_1^{T-1} \tag{2.20}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] = R\,\mathbb{E}^{\mathbb{P}}[\xi_{t+1}|\mathcal{F}_t]\, X_t^f \qquad \text{for } t \in \mathcal{T}_1^{T-1}. \tag{2.21}$$

These are the constraints of the dual optimization problem of (2.6) and it follows that the original problem is unbounded if and only if conditions (2.10)-(2.21) are fulfilled.

In the same way as in [11] it is possible to infer $\gamma > 0$ and $\xi_t > 0$. This gives the possibility to divide all equations by γ and use a modified essentially bounded process $\xi_t > 0$ such that if ξ_t' fulfills (2.10)-(2.21) then $\xi_t = \frac{\xi_t}{\gamma}'$ fulfills

$$\xi_t > 0 \tag{2.22}$$

$$\lambda_T + \frac{\psi}{2}\xi_T \geq \xi_T X_T \tag{2.23}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_1] \leq X_0^f + \frac{\psi}{2R} \tag{2.24}$$

$$R\mathbb{E}^{\mathbb{P}}[\xi_{t+1}|\mathcal{F}_t] = \xi_t \qquad \text{for } t \in \mathcal{T}_1^{T-1} \text{ and } R\,\mathbb{E}^{\mathbb{P}}[\xi_1] = 1 \tag{2.25}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] \leq \lambda_t + \frac{\psi}{2}\xi_t\left(1 + \frac{1}{R}\right) \qquad \text{for } t \in \mathcal{T}_1^{T-1} \tag{2.26}$$

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}X_{t+1}^e|\mathcal{F}_t\right] \leq \eta_i^{-1}\xi_t X_t^f \qquad \text{for all } i \in \{1,\dots,I\} \text{ and } t \in \mathcal{T}_0 \tag{2.27}$$

$$\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] = \xi_t X_t^f \qquad \text{for } t \in \mathcal{T}_1 \text{ and } \mathbb{E}^{\mathbb{P}}[\lambda_1] = X_0^f. \tag{2.28}$$

Here (2.18) is used at several points to simplify the expressions. This system already fulfills A1, A2, A3 and A5. Applying (2.18) recursively and keeping in mind

$R > 0$, it follows that $\xi_t > 0$ (which is A2). Recall that we assumed $X_t^f \geq 0$, hence by (2.21) we have $\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}|\mathcal{F}_t] \geq 0$. Using this fact, it is possible to replace the constraints (2.20) with

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}X_{t+1}^e|\mathcal{F}_t\right] \leq \eta_{max}^{-1}\mathbb{E}^{\mathbb{P}}\left[\lambda_{t+1}|\mathcal{F}_t\right] \qquad t \in \mathcal{T}_1^{T-1}.$$

This ensures A4. Moreover, plugging the second equation of (2.28) into (2.24) shows that (2.24) is redundant (recall that ψ and R both are positive).

Finally, A6 is obtained by (2.23) and by plugging A5 into (2.26). □

Remark 2.4. Given proposition 2.3 we might also speak of η_{max}-arbitrage instead of η-arbitrage.

It would be easily possible to state proposition (2.3) in terms of equivalent measures instead of using the process ξ, as in [11]. However, this would be an unnecessary digression in the present context. Instead, we state an easily interpretable necessary condition for no-arbitrage. Clearly proposition 2.3 gives a necessary and sufficient condition but the following corollary simplifies A5-A6.

Corollary 2.5. *If a market* $\left\{X_t^e, X_t^f\right\}$ *is* η_{max}*-arbitrage free then there exists a stochastic process* ξ *(fulfilling properties A1-A3 of proposition 2.3) such that*

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}X_{t+1}^e|\mathcal{F}_t\right] \leq \eta_{max}^{-1}\xi_t X_t^f \text{ for } t \in \mathcal{T}_0 \tag{2.29}$$

holds together with

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}\left(X_{t+1}^f + B_{t+1}\right)|\mathcal{F}_t\right] \leq \xi_t\left(X_t^f + B_t\right) \quad \text{for } t \in \mathcal{T}_1^{T-1} \tag{2.30}$$

with

$$B_t = \frac{\psi}{2\,(R-1)}(1 + \frac{1}{R}) \tag{2.31}$$

for $t \in \mathcal{T}_1 \setminus \{T-1, T\}$ *and*

$$B_T = \frac{\psi}{R-1}. \tag{2.32}$$

Proof. If a market is η_{max}-arbitrage free then by proposition 2.3 there exist processes ξ, λ fulfilling A1-A6. The first equation of the corollary is just A4. From the first inequality of A6 we have $\xi_{t+1}\left(X_{t+1}^f - \frac{\psi}{2}\left(1 + \frac{1}{R}\right)\right) \leq \lambda_{t+1}$ for $t = 1, 2, \ldots, T-2$. Taking conditional expectation and applying A5 and A3 leads to

$$\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}\left(X_{t+1}^f\right)|\mathcal{F}_t\right] \leq \xi_t\left(X_t^f + \frac{\psi}{2R}\left(1 + \frac{1}{R}\right)\right).$$

Now, adding $\frac{\psi}{2R^2}\frac{R+1}{R-1}\xi_t$ on both sides of the inequality and again applying A3 gives the first case of (2.30). In similar manner the second case can be obtained by starting with the second equation of A6, taking conditional expectations and then using A5, A3. In this case the resulting inequality is expanded by $\frac{\psi}{R(R-1)}\xi_{T-1}$ □

Proposition 2.3 together with corollary 2.5 can be interpreted as follows: The process ξ_t is a process of stochastic discount factors. In fact the values $R^t\xi_t$ are positive and fulfill the martingale condition A3. Moreover, using A3 it can be shown easily that $\mathbb{E}^{\mathbb{P}}[\xi_t'] = \frac{1}{R^t}$ holds, which justifies the interpretation as a discount factor. The necessary condition (2.30) means that the modified fuel price $X_{t+1}^f + B_t$ of the fuel price is a supermartingale, if properly discounted by ξ_t. Consequently, in the absence of storage costs (i.e. $\psi = 0$) the fuel price itself is a supermartingale, if properly discounted. This ensures that the expected discounted profit from selling stored fuel does not exceed the proceeds from selling fuel immediately. Condition (2.29) is a consistency requirement between the fuel price and the discounted electricity price: selling one MWh of electricity should not bring more money (in expectation and discounted correctly) than the costs for producing it.

It is also possible to derive a interpretable sufficient condition for an arbitrage free market.

Corollary 2.6. *Consider a market with prices $\left\{X_t^e, X_t^f\right\}$. If there exists a process ξ (fulfilling properties A1-A3 of proposition 2.3) such that A4 holds together with*

$$\mathbb{E}^{\mathbb{Q}}\left[\xi_{t+1}X_{t+1}^f|\mathcal{F}_t\right] = \xi_t X_t^f \tag{2.33}$$

(i.e. the discounted fuel pieces are a martingale) then the market is η_{max}-arbitrage free.

Proof. Set $\lambda_t = \xi_t X_t^f$. This choice fulfills A6 because $\frac{\psi}{2}\left(1+\frac{1}{R}\right) \geq 0$. Substituting λ_{t+1} for $\xi_{t+1}X_{t+1}^f$ at the left side of (2.33) leads to A5. Because A1-A4 hold already by assumption, proposition 2.3 implies absence of η_{max}-arbitrage. □

Remark 2.7. Note that (2.33) implies (2.30) because $R \geq 1$.

Remark 2.8. The obtained results can also be applied to a situation that includes intermittent and uncontrolled electricity production from renewable sources like photovoltaics or wind energy. The exogenously given proceeds K_t from selling the cumulated electricity production from renewable sources (either at a fixed price under the current regime or at market prices X_t^f) must be added at the right hand side of equation (2.1), respectively the first equation of (2.6), which means that the model is not self financing any more. Still the optimization problem (2.6) can be used to define an arbitrage test. Again we would speak of arbitrage if the modified problem (2.6) can become unbounded. Again, it turns out that the no-arbitrage conditions A1-A6 are valid: The modified dual problem has the same constraints as the original dual problem and the only difference is that originally the dual objective is the constant zero, whereas the modified dual minimizes the discounted expectation $\sum_{t=1}^T \mathbb{E}\left[\xi_t K_t\right]$. Therefore, all consequences derived from the feasible set of the dual problem stay valid also in the new context.

3. Implications of the necessary conditions

If one aims at testing for arbitrage in an electricity market, it is a natural approach to assume absence of arbitrage as the zero hypothesis. This is in line with the fact that according to economic theory it is hard to achieve arbitrage, i.e. riskless profits. If a person claims to know the secret of how to achieve extraordinary profits, usually it is wise not to believe this too fast.

A sensible way for constructing a test then would be to use a necessary condition like Corollary 2.5 and try to reject it. Therefore we analyze the consequences of conditions A1-A4 and inequality (2.30) in order to find out under which circumstances they are fulfilled. Here the first question is when the conditions are fulfilled for given parameter values. The answer allows to analyze whether the no-arbitrage conditions imply any restrictions that have to be observed when estimating unknown parameters.

Many different models for fuel and electricity prices have been proposed in literature, see e.g. [13] for a recent overview of electricity price modeling. Clearly the construction of a test procedure depends crucially on the exact model class under consideration. In the present paper we restrict the analysis to a class of (potentially nonlinear) vector-autoregressive econometric models. This class of models for electricity and fuel prices can be described by

$$X^e_{t+1} - E_{t+1} = \phi^e(X^e_{[t]} - E_{[t]},\, X^f_{[t]} - F_{[t]}; \theta) + \varepsilon^e_{t+1} \tag{3.1}$$

$$X^f_{t+1} - F_{t+1} = \phi^f(X^e_{[t]} - E_{[t]},\, X^f_{[t]} - F_{[t]}; \theta) + \varepsilon^f_{t+1}. \tag{3.2}$$

Here E_t and F_t are given and denote a given deterministic process, e.g. market expectations or an observed forward price-curve. The considered model is formulated relative to this processes, which also could be zero in some specifications. The notation $X^i_{[t]} = (X^i_t,\, X^i_{t-1}, \ldots, X^i_{t-p})'$, $i \in \{e, f\}$ represents (for some $p \in \mathbb{N}$) a price history up to time t (this could also be a suitable selection of past prices with certain lags). Nonpositive $t \leq 0$ denote observations that have been made before the actual planning horizon $\{1, \ldots, T\}$. The functions ϕ^i are measurable and bounded and model the one step expectation. They depend on past price differences from the reference values E_t, F_t and are parametrized by some model parameter vector θ. We will use the short notation $\phi^i_t(\theta) = \phi^i(X^e_{[t]} - E_{[t]},\, X^f_{[t]} - F_{[t]}; \theta)$ in the following.

In the simplest case such a model would be just linearly autoregressive, e.g.

$$X^e_{t+1} - E_{t+1} = \theta_1'(X^e_{[t]} - E_{[t]}) + \varepsilon^e_{t+1} \tag{3.3}$$

$$X^f_{t+1} - F_{t+1} = \theta_2'(X^f_{[t]} - F_{[t]}) + \varepsilon^f_{t+1}, \tag{3.4}$$

where θ_1, θ_2 are parameter vectors of dimension p.

Finally, we assume that – given the past – the residuals ε^i_{t+1} follow a distribution described by some joint conditional distribution function G. The distribution function respects the assumption that X^f_{t+1} are nonnegative and we assume further

$$\mathbb{E}\left[\varepsilon^i_{t+1} | \mathcal{F}_t\right] = 0 \tag{3.5}$$

and

$$\Sigma := Cov\left[\varepsilon^e_{t+1}, \varepsilon^f_{t+1} | \mathcal{F}_t\right] = \begin{bmatrix} \sigma_e^2 & \rho\sigma_e\sigma_f \\ \rho\sigma_e\sigma_f & \sigma_f^2 \end{bmatrix} \in \mathbb{R}^2 \tag{3.6}$$

Although the distribution function might also be parametrized by additional model parameters, say γ, only the covariance matrix and its components will be relevant in the following. The further parameters of the market model, i.e. R and η_{max} are given externally.

Recall that $R^t\xi_t$ is a martingale. Therefore we can rewrite it as

$$R^{t+1}\xi_{t+1} = R^t\xi_t + R^{t+1}u_{t+1}, \tag{3.7}$$

where the stochastic process u_t is a martingale difference sequence. In order to ensure that ξ_{t+1} is essentially bounded (respectively $\xi_{t+1} \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$ as requested by A1), we assume that u_{t+1} is essentially bounded (respectively $u_{t+1} \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$), as well. Note that because of the requirement $\xi_{t+1} > 0$ the sequence u_{t+1} a lower bound for $u_{t+1}$is given by $u_{t+1} > -\frac{1}{R}\xi_t$. This means that the martingale difference sequence cannot be a sequence of i.i.d. random variables.

With (3.1)-(3.7) we can reformulate corollary (2.5) in the following way.

Corollary 3.1. *If a market* $\left\{X_t^e, X_t^f\right\}$, *with the price processes specified by (3.1)-(3.2), is* η*-arbitrage free then there exists a stochastic process* ξ *(fulfilling properties A1-A3 of proposition 2.3) together with a related martingale difference sequence* u_t *such that*

$$\frac{1}{R}\left(E_{t+1}+\phi_t^e(\theta)\right)+\mathbb{E}\left[\frac{u_{t+1}}{\xi_t}\varepsilon_{t+1}^e|\mathcal{F}_t\right]\leq\eta_{max}^{-1}X_t^f \text{ for } t\in\mathcal{T}_0 \tag{3.8}$$

holds together with

$$\frac{1}{R}\left(F_{t+1}+\phi_t^f(\theta)\right)+\mathbb{E}\left[\frac{u_{t+1}}{\xi_t}\varepsilon_{t+1}^f|\mathcal{F}_t\right]\leq X_t^f+B_t-\frac{B_{t+1}}{R} \tag{3.9}$$

Proof. Under the assumptions corollary 2.5 is fulfilled. Plugging the model definition (3.1) of X_{t+1}^e into (2.29) one gets

$$\left(E_{t+1}+\phi_t^e(\theta)\right)\cdot\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}|\mathcal{F}_t\right]+\mathbb{E}^{\mathbb{P}}\left[\xi_{t+1}\varepsilon_{t+1}^e|\mathcal{F}_t\right]\leq\eta_{max}^{-1}\xi_t X_t^f.$$

A3 and (3.7) now can be used to get

$$\frac{1}{R}\xi_t\left(E_{t+1}+\phi_t^e(\theta)\right)+\frac{1}{R}\xi_t\mathbb{E}^{\mathbb{P}}\left[\varepsilon_{t+1}^e|\mathcal{F}_t\right]+\mathbb{E}^{\mathbb{P}}\left[u_{t+1}\varepsilon_{t+1}^e|\mathcal{F}_t\right]\leq\eta_{max}^{-1}\xi_t X_t^f.$$

The first expectation here is zero by the model definition, see (3.5). Then, dividing by ξ_t (recall A2) leads to (3.9), the first statement of the corollary.

The same arguments, applied to (2.30), lead to the second statement (3.9). □

If we define now a process

$$v_t=\frac{u_{t+1}}{\xi_t}, \tag{3.10}$$

we see that $\mathbb{E}\left[v_{t+1}\varepsilon_{t+1}^e|\mathcal{F}_t\right]=Cov\left[v_{t+1}\varepsilon_{t+1}^e|\mathcal{F}_t\right]=\rho_{et}\cdot\sigma_e\cdot\sigma_{vt}$, where ρ_{et} is the conditional correlation between v_{t+1} and ε_{t+1}^e, where σ_e denotes the (constant) standard deviation of ε_t^e (see (3.6)) and σ_{vt} is the conditional standard deviation of v_{t+1}. In the same manner we get $\mathbb{E}\left[v_{t+1}\varepsilon_{t+1}^f\right]=\rho_{ft}\cdot\sigma_f\cdot\sigma_{vt}$, where ρ_{ft} is the conditional correlation between v_{t+1} and ε_{t+1}^f, and σ_f denotes the (constant) standard deviation of ε_t^f (see (3.6)). It is important to keep in mind that inequalities (3.8)-(3.9) hold almost surely at each point in time.

The two equations of corollary 3.1 now can be written as

$$\left[\rho_{et}\cdot\sigma_e\cdot\sigma_{vt}\right]_t\leq G_t \tag{3.11}$$

and

$$\left[\rho_{ft}\cdot\sigma_f\cdot\sigma_{vt}\right]_t\leq H_t \tag{3.12}$$

with

$$G_t=\eta_{max}^{-1}X_t^f-\frac{1}{R}\left(E_{t+1}+\phi_t^e(\theta)\right)$$

and

$$H_t=X_t^f-\frac{1}{R}\left(F_{t+1}+\phi_t^f(\theta)\right)+B_t-\frac{B_{t+1}}{R}.$$

Note that $-G_t$ can be considered as a discounted version of the “spark spread” (where the actual electricity price is replaced by the discounted expectation of the one step ahead electricity price under model 3.1) and H_t is the difference between fuel price and the discounted expectation of the one step ahead fuel price.

Is it possible now to choose ρ_{et}, σ_{vt} and ρ_{ft} in order to fulfill the inequalities, when σ_e and σ_f (and also θ and R, η) are given parameters? First we see that σ_{vt} must be nonnegative and can be chosen as large as necessary.

Corollary 3.2. *Given any real numbers $M \geq 0$, $-1 \leq \rho_1 \leq 1$, $-1 \leq \rho_2 \leq 1$ it is always possible to define a random variable v_{t+1}such that*

$$\mathbb{E}\left[v_{t+1}|\mathcal{F}_t\right] = 0, \tag{3.13}$$

$$v_{t+1} > -\frac{1}{R} \tag{3.14}$$

and

$$\sigma_{vt} = M. \tag{3.15}$$

Moreover a joint distribution for v_{t+1}, ε^e_{t+1}, ε^f_{t+1} can be defined such that the covariance structure (3.6) is kept, while

$$\rho_{et} = \rho_1 \text{ and } \rho_{ft}{=}\rho_2. \tag{3.16}$$

is ensured.

Proof. Given two positive real numbers $0 < K_1 < \frac{1}{R}$ and $K_2 > 0$, consider a random variable v_{t+1} that follows a mixture distribution such that with $p = \frac{K_2}{K_1+K_2}$ the density of v_t is defined as $f(x) = p\frac{1}{K_1}\imath_{[-K_1,0]}(x) + p\frac{1}{K_2}\imath_{[0,K_2]}(x)$, where $\imath_A(x)$ denotes the indicator function which equals one if $x \in A$, and else is zero. This means that v_{t+1} follows a mixture of two uniform distributions on $[-K_1, 0]$ and $[0, K_2]$. If K_1, K_2 are chosen such that $K_1K_2 = 12M^2$, then equations (3.13)-(3.15) are fulfilled. Finally, given the joint distribution of ε^e_{t+1}, ε^f_{t+1} and the marginal distribution of v_{t+1}, a suitable copula function can be used to construct a joint distribution function for v_{t+1}, ε^e_{t+1}, ε^f_{t+1} that fulfills all requirements. □

The first two requirements in Corollary (3.2) ensure the properties of a positive martingale for ξ_{t+1} relative to ξ_t. Altogether the corollary ensures that it is possible to scale freely the left hand side of inequalities (3.11)-(3.12) if σ_e, respectively σ_f are positive. This means that in this case only the sign of ρ_{et} and ρ_{ft} really matters. Clearly we would like to choose both correlations negative, when the left hand sides of (3.11)-(3.12) can be made arbitrarily small.

Therefore it has to be checked, whether it is possible to select a random variable v_{t+1}, that is negatively correlated to both residuals ε^e_t and ε^f_t when the residuals are correlated with coefficient ρ, as specified in (3.6). To answer this question, recall that in an inner product space we can define the angle θ between two elements X, Y by $\langle X, Y\rangle = \|X\| \, \|Y\| \cos(\theta)$. Taking into account that in our context we have $\|X\| = \sqrt{\mathrm{Var}(X|\mathcal{F}_t)} = \sigma_X$ and $\langle X, Y\rangle = \mathrm{Cov}\left(X, Y|\mathcal{F}_t\right)$, we see that for the angle $\theta_\varepsilon$between the residuals ε^e_t and ε^f_t we have $\cos\left(\theta_\varepsilon\right) = \rho$. Moreover, for the angles θ_e and θ_f between v_{t+1} and the respective residuals ε^e_t and ε^f_t we have $\cos\left(\theta_e\right) = \rho_{et}$ and $\cos\left(\theta_f\right) = \rho_{ft}$. By standard geometric arguments it can then be seen that it is possible to choose both (cosines) correlations negative, if and only if

$$\rho \neq -1. \tag{3.17}$$

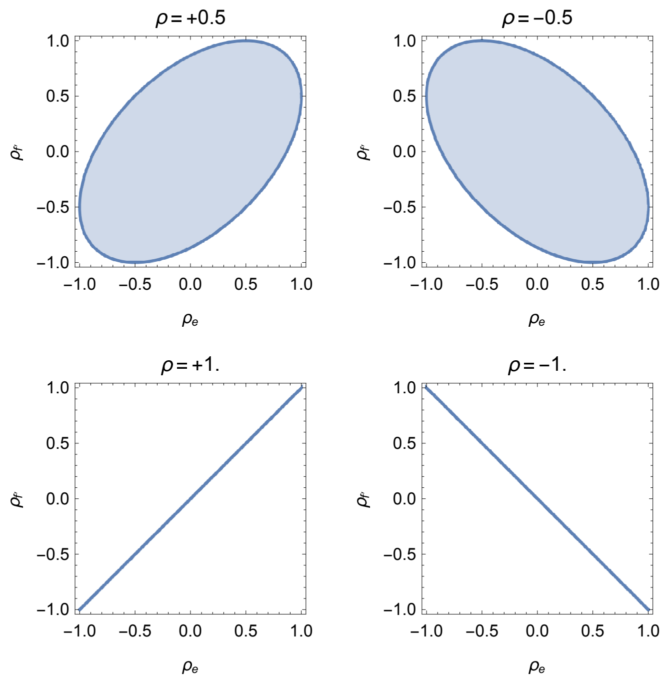

FIGURE 3.1. Feasible combinations of ρ_{et}, ρ_{ft} (correlation between v_{t+1} and the residuals ε^e_{t+1}, respectively ε^f_{t+1},), dependent on different values of ρ (the correlation between the residuals).

In fact, v_{t+1} has to be chosen such that it is at an obtuse angle with (but not orthogonal to) both residuals. This is possible if and only if the residual vectors are not pointing exactly in opposite directions.

The special role of ρ, the correlation between the residuals, can also be seen from the following consideration: the correlations ρ_{et} and ρ_{ft} have to be chosen such that the joint (conditional) correlation matrix for ε^e_{t+1}, ε^f_{t+1} and v_{t+1} stays nonnegative definite. This reduces to

$$\left|\begin{bmatrix} 1 & \rho & \rho_{et} \\ \rho & 1 & \rho_{ft} \\ \rho_{et} & \rho_{ft} & 1 \end{bmatrix}\right| \geq 0$$

or (the first two principal minors are positive)

$$\rho_{et}^2 + \rho_{ft}^2 - 2\rho\rho_{et}\rho_{ft} \leq 1 - \rho^2. \tag{3.18}$$

With ρ given, for $0 < \rho < 1$ this inequality (in ρ_{et}, ρ_{ft}) describes the points inside an ellipse, including the ellipse itself. The ellipse has its center in the origin, moreover its main axis has positive slope if ρ is positive and it has negative slope if ρ is negative. If $\rho = 0$ then the ellipse becomes the unit circle. There are also two degenerated cases. For $\rho = -1$ inequality (3.18) becomes equivalent to the equation of a straight line $\rho_{ft} = -\rho_{et}$ and is valid for $0 \leq \rho_{ft}, \rho_{et} \leq 1$. Finally, $\rho = +1$ inequality (3.18) becomes equivalent to the equation of a straight line $\rho_{ft} = \rho_{et}$ and is valid for $0 \leq \rho_{ft}$, $\rho_{et} \leq 1$. See figure 3 for a graphical representation of these cases.

These results fit exactly to the above discussion. In case $-1 < \rho < +1$ it is easily possible to choose both ρ_{et} and ρ_{ft} negative (Case 2), i.e. from the lower left part of the ellipse. This is also possible for $\rho = +1$, with the only difference that both correlations have to be chosen equal in this case. However, if $\rho = -1$ it is not possible to choose both correlations negative, because of $\rho_{ft} = -\rho_{et}$.

On the other hand, the situation is quite different if one of σ_e or σ_f equals zero. If $\sigma_e = 0$ then inequality (3.11) reduces to the condition

$$\eta_{max}^{-1} X_t^f - \frac{1}{R}\left(E_{t+1} + \phi_t^e(\theta)\right) \geq 0. \tag{3.19}$$

The left hand side of the second inequality (3.12) can be made arbitrarily small using the same arguments as above. Here it is not even necessary to ensure the condition $\rho \neq -1$, because the correlation ρ_{et} on the left hand side of (3.11) can be chosen arbitrarily. If $\sigma_f = 0$ then inequality (3.12) reduces to

$$X_t^f - \frac{1}{R}\left(F_{t+1} + \phi_t^f(\theta)\right) \geq 0. \tag{3.20}$$

The left hand side of the first inequality (3.11) again can be made arbitrarily small. Finally, if both standard deviations σ_e and σ_f are equal to zero, then arbitrage occurs only if one of the inequalities (3.19)-(3.20) is violated for some point time.

Altogether, with given parameters one only has to check ρ, σ_e and σ_f. If $\rho \neq -1$ then the no-arbitrage hypothesis never can be rejected. If one or both of σ_e, σ_f are zero, then the no arbitrage hypothesis must be rejected if the related equation from (3.19)-(3.20) is not fulfilled, and otherwise the hypothesis is not rejected. If parameter values are not given (our second question) then we can conclude the following:

Proposition 3.3. *Consider a price model (3.1)-(3.2) and (3.5)-(3.6). Assume that price processes X_t^e, X_t^e are observed such that for some parameter value θ one of the deterministic cases*

$$A)\quad X_{t+1}^e - E_{t+1} = \phi^e(X_{[t]}^e - E_{[t]},\, X_{[t]}^f - F_{[t]}; \theta),$$
$$B)\quad X_{t+1}^f - F_{t+1} = \phi^f(X_{[t]}^e - E_{[t]},\, X_{[t]}^f - F_{[t]}; \theta).$$

or

$$C)\quad \text{both equations } A) \text{ and } B)$$

hold for all points in time t. Then inequalities (3.8)-(3.9) (and therefore the no-arbitrage hypothesis) cannot be rejected if (3.19) is fulfilled in case A), (3.20) is fulfilled in case B), respectively both (3.19)-(3.20) are fulfilled in case C). Otherwise (3.8)-(3.9) are not valid and arbitrage is possible.

If none of these cases holds, and also it can be excluded that for some parameter value θ the deterministic equation

$$X_{t+1}^e + X_{t+1}^f = E_{t+1} + \phi_t^e(\theta) + F_{t+1} + \phi_t^f(\theta) \tag{3.21}$$

holds for all points in time t then inequalities (3.8)-(3.9) (and therefore the no-arbitrage hypothesis) cannot be rejected.

Proof. Cases A), B) and C) refer to cases where in the given model either σ_e or σ_f are equal to zero in the given model: If there exists a parameter value θ such that A),B) or C) are fulfilled this shows that the prices follow exactly a model with some of the price variances equal to zero. The correct conditions for validity of (3.8)-(3.9) already have been discussed above.

If equation (3.21) is not fulfilled this means that $\rho \neq -1$, because this would be possible if and only if $\varepsilon^e_{t+1} = -\varepsilon^f_{t+1}$. In this case we already found that (3.8)-(3.9) can be made valid for arbitrary parameter value θ. □

In a context with given data and parameters to be estimated, therefore it is usually no problem to estimate the parameters without restrictions, e.g. using unrestricted maximum likelihood. It is highly unlikely in a real world application that one of the deterministic equations in proposition (3.3) holds with observed price data.

4. Implications of the sufficient conditions

So far we have obtained that for a large class of possible models the no arbitrage condition can be rejected only in rare, even unrealistic cases. This result is based on the necessary conditions (2.29)-(2.30). In the following we use similar arguments to analyze the implications of the sufficient condition in corollary 2.6. Again we look at criteria for given parameters and on estimation, but also discuss the relevance of our results in two additional contexts: First we consider delivery contracts and their valuation. Here the absence of arbitrage has the important technical consequence that in this case valuation procedures based on stochastic discount factors can be obtained. In our second discussion it turns out that the cases with zero variance are important in the context of tree based stochastic optimization, which is a way to deal with the valuation problem, but also for taking other decisions in the energy management context.

We use again the model class specified in (3.1)-(3.2) and (3.5)-(3.6) and rewrite the process ξ by (3.7). In particular the processes u and v (defined in (3.10)) have the same properties as before. Here the sufficient conditions A4 and (2.33) can be reduced to,

$$\rho_{et} \cdot \sigma_e \cdot \sigma_{vt} \leq G_t \tag{4.1}$$

and

$$\rho_{ft} \cdot \sigma_f \cdot \sigma_{vt} = H_t \tag{4.2}$$

with

$$G_t = \eta_{max}^{-1} X_t^f - \frac{1}{R}\left(E_{t+1} + \phi_t^e(\theta)\right)$$

and

$$H_t = X_t^f - \frac{1}{R}\left(F_{t+1} + \phi_t^f(\theta)\right).$$

which replaces (3.11)-(3.12).

We start our analysis with the trivial cases. If $\sigma_e = 0$ then the conditions

$$\eta_{max}^{-1} X_t^f - \frac{1}{R}\left(E_{t+1} + \phi_t^e(\theta)\right) \geq 0 \quad (t) \tag{4.3}$$

are sufficient for absence of arbitrage. Hence this is a necessary and sufficient condition if $\sigma_e = 0$. If $\sigma_f = 0$ then

$$X_t^f - \frac{1}{R}\left(F_{t+1} + \phi_t^f(\theta)\right) = 0 \quad (t) \tag{4.4}$$

is sufficient. If both conditional variances are zero then both of this equations must hold to imply the absence of arbitrage. The arguments here are straightforward.

If on the other hand both variances are positive, then two cases are possible: If $H_t = 0$ then also the additional condition $G_t \geq 0$ must hold at any point in time t in order to imply absence of arbitrage. If $H_t \neq 0$ then the restriction that the model is consistent with the sufficient conditions is $\rho \neq \pm 1$.

The last two cases can be inferred in the following way: $H_t = 0$ directly implies that $G_t \geq 0$ is needed in addition to ensure the sufficient conditions. Looking at the case $H_t \neq 0$ we get from (4.2)

$$\sigma_{vt} = \frac{H_t}{\rho_{ft}\sigma_f}.$$

Note that this holds because ρ_{ft} must have the same sign as H_t by (4.1) (and is not equal to zero if $H_t \neq 0$). Plugging σ_{vt} into (4.1) leads to

$$\frac{\sigma_e}{\sigma_f}\frac{H_t}{\rho_{ft}}\rho_{et} \leq G_t.$$

The multiplicator of ρ_{et} on the left hand side here is positive. If $G_t \geq 0$ then it suffices to choose some nonpositive ρ_{et} (and some combination ρ_{ft}, σ_{vt} that fulfills the equation (4.2)). In order to account also for the case $G_t < 0$, the correlation ρ_{ft} should be chosen such that $|\rho_{ft}|$ is small enough to fulfill the inequality. The sign of ρ_{ft} depends on the sign of H_t while the sign of ρ_{et} has to be chosen negative. Taking into account the correlation ρ between the residuals and its relation to ρ_{ft} and ρ_{et} as in the previous section, such a choice is always possible unless either $\rho = -1$ or $\rho = +1$.

We can reformulate this insight in the following way:

Proposition 4.1. *Consider a price model (3.1)-(3.2) and (3.5)-(3.6) Assume that price processes X_t^e, X_t^e are observed such that for some parameter value θ one of the deterministic cases*

$$A) \quad X_{t+1}^e - E_{t+1} = \phi^e(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta_1),$$

$$B) \quad X_{t+1}^f - F_{t+1}, = \phi^f(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta_2).$$

hold for all points in time t. Then the sufficient conditions (4.1)-(4.2) are feasible if (4.3) is feasible in case A), respectively (4.4) is feasible in case B).

Assume now that neither A) nor B) holds. If additionally it can be excluded that for some parameter value θ one of the deterministic equations

$$X_{t+1}^e + X_{t+1}^f = E_{t+1} + \phi_t^e(\theta) + F_{t+1} + \phi_t^f(\theta) \tag{4.5}$$

$$X_{t+1}^e - X_{t+1}^f = E_{t+1} + \phi_t^e(\theta) - F_{t+1} - \phi(\theta) \tag{4.6}$$

holds for all points in time t, then inequalities (4.1)-(4.2) are fulfilled for any value θ except the case that (4.4) holds but (4.3) is violated for any point in time t.

Proof. This follows directly from the above discussion, using the same arguments as in proposition 3.3. Note that the set of parameters that satisfies (4.3) and (4.4) is a subset of the set of parameters such that (4.4) implies (4.3). □

Coming back to parameter estimation from data, and leaving aside the degenerate cases A), B) and (4.5)-(4.6), which are not realistic for real data: Arbitrage free parameter estimators must fulfill the logic constraints

$$X_t^f - \frac{1}{R}\left(F_{t+1} + \phi_t^f(\theta)\right) = 0 \Rightarrow \eta_{max}^{-1} X_t^f - \frac{1}{R}\left(E_{t+1} + \phi_t^e(\theta)\right) \geq 0 \tag{4.7}$$

at any point in time. Using binary variables $a_t \in \{0,1\}$, additional real valued variables k_t and $M > 0$ large enough, these constraints can be reformulated by a (linear) system

$$H_t + (1 - a_t)k_t = 0$$
$$G_t + M(1 - a_t) \geq 0.$$

In the framework of maximum likelihood estimation, best estimators that exclude arbitrage are obtained by restricted maximum likelihood estimation, with constraints (4.7). Note that for the purpose of estimation the restrictions (4.7) have to hold only for the observed price values. The restricted estimator then can be compared with the unrestricted maximum likelihood estimator: it is possible to use a likelihood ratio test in order to test, whether the restricted model should be kept (null hypothesis H_0) or rejected. The likelihood ratio test statistics, dependent on observed price values x is given by

$$\lambda(x) = -2 \log \left(\frac{l_1^*(x)}{l^*(x)} \right)$$

where $l_1^*(x)$ denotes the maximized likelihood function in the restricted model and $l^*(x)$ denotes the maximized likelihood function of the unrestricted model. The null hypothesis is rejected if $\lambda(x) < \kappa_\alpha$, where $\kappa_\alpha = \sup_{\theta \in H_0} \{P_\theta(\lambda(X) < \kappa_\alpha) = \alpha\}$. Here, $\theta \in H_0$ means that θ fulfills all logical constraints (4.7).

The standard approach uses the fact that under some regularity conditions the distribution of the likelihood ration statistics converges to a χ^2-distribution. Unfortunately the usual regularity conditions (in particular that the estimates are in the interior of the feasible set) cannot be guaranteed here. Moreover, it is also not possible to derive an exact distribution of the likelihood ratio statistics, i.e. the distribution of $\lambda_1(X)$ if the price process X fulfills the logical constraint (4.7). Given these difficulties, it is a reasonable way to use a bootstrap likelihood ratio test, which is an application of the parametric bootstrap and has been applied in many complicated non-standard situations (see e.g. [12, 14, 8, 6, 20]).

This works in the following way for our problem: In a first step, observed prices are used to estimate the maximum likelihood estimates $\hat{\theta}^{(1)}, \hat{\Sigma}^{(1)}$ for the restricted model (3.1)-(3.2) with logical restriction (4.7). These estimates are then used to generate K (a large number) of bootstrap (random) samples $x_t^e(k)$, $x_t^f(k)$, $k \in \{1, \ldots, K\}$, $t \in \{1, \ldots, T\}$. In addition, maximum likelihood estimators $\hat{\theta}^{(2)}, \hat{\Sigma}^{(2)}$ are also calculated for the unrestricted model. In the next step, the likelihood ratio statistic $\lambda_1(x(k))$ is calculated for each of the samples $k \in \{1, \ldots, K\}$. The value of the i-th order statistic of the K replications of the likelihood ratio statistic then can be used as an estimator for the quantile of order $\frac{i}{(K+1)}$ of the distribution of $\lambda_1(X)$ - the distribution of the likelihood ratio statistic under the null hypothesis. Finally, the value of the likelihood ratio statistic $\lambda_1(x)$ obtained from the original data is compared to the quantiles of the sampled distribution, which leads to an estimate of the p-value.

Alternatively, if one aims at rejection of the null hypothesis at a given significance level α, the procedure leads to a test with approximate size α. Here the test which rejects the null hypothesis if $\lambda_1(x)$ (using the original data) is larger than the i-th smallest bootstrap replicate $\lambda_1(x(k))$ can be used, compare e.g. [1].

If the null hypothesis is not rejected, the restricted maximum likelihood estimator ensures absence of arbitrage and it could not be established that there is a better unrestricted estimator. If we have to reject the null hypothesis, then there is a better unrestricted estimator, and we can not guarantee absence of arbitrage, because the sufficient condition is violated by the unrestricted estimator.

Valuation of delivery contracts. We outline now the basic setup in [11], adapt it to our more general case and analyze delivery contracts for electric energy. Here the producer has the obligation to deliver amounts D_t [MWh] of electric energy at a fixed price of K per MWh for periods $[t, t+1]$. The stochastic demand D_t is adapted to the same filtration $\{\mathcal{F}_t\}$ as the other processes considered so far, which contains all relevant information. Among others, D_t can be a deterministic delivery pattern, a pattern that depends on the fuel and/or the electricity prices, as well as a pattern that depends on other relevant variables like the observed temperature.

Given the discussion in the previous section, it is clear that we cannot refer to a unique equivalent measure respectively a unique stochastic discount process, i.e. the market is incomplete. Therefore, we consider the valuation problem from the viewpoint of the producer. Moreover, out of the various cases analyzed in [11], we consider only the so called superhedging value in this paper, i.e. the answer to the following question: "What is the minimum initial asset value or upfront-payment $V_0 = c_0 + s_0 X_0^f$ such that the producer is able to fulfill all contractual obligations and the asset value at the end of the planning horizon, i.e. $V_T = c_T + s_T X_T^f$ is almost surely non negative. If the initial asset value or upfront payment is below the superhedging value, the producer takes some additional risk of defaulting when he agrees to the contract.

The original model has to be extended now. Our producer cannot neglect that fuel storage and production capacity is restricted. So $S > 0$ will denote the upper bound on storage and P_t is an $\{\mathcal{F}_t\}$-adapted process of upper bounds on the production of a generator with efficiency η. Moreover, the producer has to allocate the produced electricity between the market and the contractual obligations. Finally, the producer is also allowed to buy electricity from the market in order to meet obligations. The amount of electricity sold at the market is given by $y_{t-1} - D_{t-1}$. If this difference is negative, an amount of energy is bought.

The superhedging value can be calculated as the optimal value of the following optimization problem, where the objective is to minimize the asset value or upfront payment at the beginning.

$$
\begin{aligned}
(4.8)\qquad & V_0^*(K, D, \eta) = \\
& \qquad \min_{y,z,c,s} c_0 + X_0^f s_0 \\
& \quad \text{subject to} \\
& \quad (t \in \mathcal{T}_1) : c_t = \left(c_{t-1} - z_{t-1} X_{t-1}^f\right) R + y_{t-1} X_t^e - D_{t-1} X_t^e + K D_{t-1} \\
& \quad (t \in \mathcal{T}_1) : s_t = s_{t-1} - y_{t-1} \eta^{-1} + z_{t-1} \\
& \qquad\qquad\qquad c_T + X_T^f s_T \geq 0 \\
& \quad (t \in \mathcal{T}) : 0 \leq s_t \leq S \\
& \quad (t \in \mathcal{T}_0) : 0 \leq y_t \leq P_t.
\end{aligned}
$$

We keep all assumptions on the involved processes made in section (2). In addition we assume that S is a real number, $P_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, and $D_t \in L^1(\Omega, \mathcal{F}_t, \mathbb{P})$.

The dual problem, related to (4.8) then is given by.

Fact 4.2. *The Lagrange dual of the valuation problem (4.8) is given by*

(4.9)

$$U_0^*(K, D, \eta) = \max_{\xi,\lambda,\mu,\nu} \sum_{t=0}^{T-1} \mathbb{E}^{\mathbb{P}} \left[\xi_{t+1} \left(X_{t+1}^e - K\right) D_t\right] - \sum_{t=0}^{T-1} \mathbb{E}^{\mathbb{P}} \left[\mu_t P_t\right] - S \sum_{t=0}^{T} \mathbb{E}^{\mathbb{P}} \left[\nu_t\right]$$

subject to

$A1) - A3)$ *and* $A5)$

$A4')\mathbb{E}^{\mathbb{P}} \left[\xi_{t+1} X_{t+1}^e | \mathcal{F}_t\right] \leq \eta_{max}^{-1} \xi_t X_t^f + \mu_t$ *for* $t \in \mathcal{T}_0$

$A6')\xi_t \cdot \left[X_t^f - \frac{\psi}{2}\left(1 + \frac{1}{R}\right)\right] \leq \lambda_t + \nu_t$ *for* $t \in \mathcal{T}_1^{T-1}$ *and* $\xi_T \left[X_T^f - \frac{\psi}{2}\right] \leq \lambda_T$

$(t \in \mathcal{T}) : \mu_t \geq 0, \quad \nu_t \geq 0,$

where ξ_t,λ_t, μ_t, $\nu_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$.

If the optimal values $V_0^*(K, D, \eta)$, $U_0^*(K, D, \eta)$ are equal, this reformulation shows that the superhedging value can be interpreted as a (modified) expected present value of the opportunity costs for selling parts of the production according to the contract and not at the market. The opportunity costs are discounted by stochastic discount factors which fulfill conditions quite similar to our no-arbitrage conditions A1)-A6). Expected present value and constraints are modified by effects related to the upper bounds on production and storage,

As shown in [11], a sufficient condition for $V_0^*(K, D, \eta) = U_0^*(K, D, \eta$ is that the market is arbitrage free. In the light of the previous subsection, we can ensure an arbitrage free model of the form (3.3)-(3.6) by using a maximum likelihood estimator, restricted by the logical constraint (4.7).

Arbitrage free price models in tree based multi-stage optimization. In the context of electricity markets, arbitrage free price models are important for any kind of optimization problem that involves both, fuel and electricity prices and allows for trading at both markets. This comprises pricing problems as above, but also e.g. planning of electricity generation. If prices here are not arbitrage free, then the solutions of the formulated planning problems would feign yields that cannot be realized in reality.

While in this article, so far we made very general assumptions on the used (conditional) distributions, an important framework for solving decision problems is tree based multistage stochastic optimization (see e.g. [18]), which is based on distributions on finite state spaces. This is achieved by replacing a decision problem that is initially formulated on a continuous state space with a reformulation on a "tractable" finite state space. Here, scenario trees are the tools to model the discretized processes, their distributional properties and also the information flow over time.

We sketch here the approach described in [18], 1.4. (for an alternative formulation see e.g. [2]) where the original time oriented formulation involving time indices is replaced by a node oriented formulation as described in the following: Consider a finite probability space $\Omega = (\omega_1, \ldots, \omega_K)$, representing S scenario-paths. Any stochastic process defined on this sample space can be represented as a finite tree with node set $\mathcal{N} = \{0, 1, \ldots, N\}$. The levels of the tree correspond to the decision stages. Let $\mathcal{N}_t$ be the set of nodes at level t, for $t = 0, \ldots, T$. The last level $\mathcal{N}_T$ contains the S leaves of the tree which can be identified with the scenario paths: $\mathcal{N}_T = \Omega = (\omega_1, \ldots, \omega_K)$. The tree structure represents the filtration of the process and can be defined by stating the (unique) predecessor node n_- for each node n. There is a unique root node, by convention denoted with 0, which represents the present. By construction there is a one to one relation between any node n and an assigned pair (ω, t), which means that each node is related to the state of the system at time t in sample path ω and vice versa.

The price processes X^e, X^f are represented w.r.t. the nodes of the tree, i.e. we write X^e_n, X^f_n instead of $X^e_t(\omega), X^f_t(\omega)$. In similar manner the decision processes x, c, s, z, y are related to the nodes: So far $s_t(\omega)$ denoted the amount of fuel stored at time t in state ω. In the discretized model, x_n denote the value of produced energy planned at node n, which can be identified with a point in time t and a scenario ω. Almost sure constraints then are obtained by formulating the same constraint for all nodes of a stage $\mathcal{N}_t$. Moreover constraints between points in time can be rewritten with node indices instead of time indices, using the predecessor relation n_-. As an example consider the cash equation (2.1), which can be rewritten as

$$c_n = \left(c_{n_-} - z_{n_-} X^f_{n_-}\right) R + X^e_n \sum_{i=1}^{I} y_{in_-} - \psi \frac{s_n + s_{n_-}}{2}$$

in the node oriented formulation.

Finally probabilities π_n can be assigned to all leaf nodes $n \in \mathcal{N}(\mathrm{T})$, which also implies probabilities π_n for all other nodes. The probabilities then can be used to formulate objective functions based on expectation or other probability functionals (risk or acceptability functionals).

Given an estimated price model, several methods have been proposed to construct approximating trees (with given tree structure), see e.g. [5, 9, 19]. The outcome are price values an probabilities at all nodes. In the context of our basic model (3.1)-(3.3) this means that for a given node n the model equations are replaced by

$$X^e_n - E_n = \phi^e(X^e_{[n_-]} - E_{[n_-]},\, X^f_{[n_-]} - F_{[n_-]}; \theta) + \varepsilon^e_n \tag{4.10}$$

$$X^f_n - F_n = \phi^f(X^e_{[n_-]} - E_{[n_-]},\, X^f_{[n_-]} - F_{[n_-]}; \theta) + \varepsilon^f_n, \tag{4.11}$$

where $X^i_{[n_-]}$ denotes a history of price values from predecessor values in the same paths as n and the θ is the original estimator. On the other hand, given a node n the conditional distribution of the residual values $\varepsilon^e_m, \varepsilon^f_m$ related to its successor nodes (i.e. $m \in \{k : n = k^-\}$) can be described by pairs of price values and conditional probabilities (which can be derived easily from the node probabilities π). When θ has been originally estimated in order to obtain an arbitrage free model as discussed above, also the approximating tree model stays arbitrage free under quite general circumstances.

However there is one typical problem with tree generation. Often the used trees are not very dense. For some nodes the number of successors might be small and there might be also nodes with a single successor. If e.g. the planning horizon is one year and the decision periods have a length of one week, already a binary tree would lead to a number of nodes around $9 \cdot 10^{15}$. Therefore, to obtain tractable problems it is inevitable to include some nodes with unique successor. If this is necessary, then in those nodes the conditional residual variances become zero, i.e. the distributions degenerate as discussed in the previous sections. It is not enough then to define the prices in these single successor nodes BY

$$X_n^e - E_n = \phi^e(X_{[n_-]}^e - E_{[n_-]}, X_{[n_-]}^f - F_{[n_-]}; \theta) \tag{4.12}$$

$$X_n^f - F_n = \phi^f(X_{[n_-]}^e - E_{[n_-]}, X_{[n_-]}^f - F_{[n_-]}; \theta), \tag{4.13}$$

as usually done. These expected prices may lead to arbitrage. If avoiding arbitrage has priority, this must be ensured by prices fulfilling (3.19)-(3.20). These conditions have priority and one might use e.g. prices that are as close to the conditional expectations as possible (in some distance) but still fulfill the no-arbitrage conditions.

5. Conclusions

In this paper we derived a necessary and a sufficient condition for absence of arbitrage on an electricity market with generation from fuel, fuel storage and also accounted for the related costs. Despite the fact that at electricity markets there exist unique frictions - compared to financial markets and other commodity markets, the derived conditions can be easily interpreted in a financial context. Building on these results, the main question is, how restrictive such constraints are. In particular it is important for practical applications to know the implications for parameter estimation, because for many planning and valuation problems it is important to base decisions on arbitrage free prices in order to avoid unrealistic outcomes.

We analyzed these questions in the context of a class of (potentially nonlinear) autoregressive prices models for fuel and electricity. It turned out that actually the necessary condition restricts the space of feasible parameter values only to a very slight extent. If one ignores deterministic price models, it is therefore in order to use unrestricted estimation approaches (like e.g. unrestricted maximum likelihood) for estimation purposes. In some situations, however, it is important to guarantee absence of arbitrage and the sufficient condition can be used to achieve this. Our results show that (again neglecting deterministic models) a simple logical constraint ensures absence of arbitrage. This restriction is not very severe, but should be taken into account, when estimating parameters. Therefore, restricted estimation with constraint (4.7) is needed. Moreover, by comparing the restricted estimator to the unrestricted parameter it is possible to test the hypothesis of absence of arbitrage, using a bootstrap likelihood ratio test.

Still this paper uses a simple, stylized market model. This leaves room for future research. In particular, further frictions like efficiencies dependent on the generation level and several different fuels might be considered.

References

[1] M. Aitkin, D. Anderson, and J. Hinde. Statistical modelling of data on teaching styles (with discussion). *Journal of the Royal Statistical Society A*, 144(419-461), 1981.

[2] A. Alonso-Ayuso, L. F. Escudero, and C. Pizarro-Romero. *Introduction to Stochastic Programming.* Ciencias Experimentales Y Tecnologia. Universidad Rey Juan Carlos, 2009.

[3] Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, 3rd edition, 2009.

[4] S.-J. Deng, B. Johnson, and A. Sogomonian. Exotic electricity options and the valuation of electricity generation and transmission assets. *Decision support systems*, 30:383–392, 2001.

[5] J. Dupacova, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming, Series A*, 95:493–511, 2003.
[6] Z. D. Feng and C. E. McCulloch. Using bootstrap likelihood ratio in finite mixture models. *Journal of the Royal Statistical Society, Series B (Methodological)*, 58(3), 1996.
[7] S. D. Flåm. Option pricing by mathematical programming. *Optimization: A Journal of Mathematical Programming and Operations Research*, 57(1):165–182, 2008.
[8] P. Hall and S. Wilson. Two guidelines for bootstrap hypothesis testing. *Biometrics*, 47(2):757–762, 1991.
[9] H. Heitsch and W. Römisch. Stability and scenario trees for multistage stochastic programs. 150:139–164, 2011.
[10] A. J. King. Duality and martingales: a stochastic programming perspective on contingent claims. *Mathematical Programming, Series B*, 91:543–562, 2002.
[11] R. M. Kovacevic. Valuation and pricing of electricity delivery contracts - the producer's view. *Research Report 2015-04, ORCOS, TU Wien*, 2015.
[12] G. J. McLachlan. On bootstrapping likelihood ratio test statistic for the number of components in a normal mixture. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 36(3):318–324, 1987.
[13] J. Nowotarski and R. Weron. Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, 81(1):1548–1568, 2018.
[14] A. Owen. Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18(1):90–120, 1990.
[15] T. Pennanen. Dual representation of superhedging costs in illiquid markets. *Mathematics and Financial Economics*, 5(4):233–248, 2011.
[16] T. Pennanen. Superhedging in illiquid markets. *Mathematical Finance*, 21(3):519–540, 2011.
[17] T. Pennanen. Introduction to convex optimization in financial markets. *Mathematical Programming, Series B*, 134(1):157–186, 2012.
[18] G. C. Pflug and A. Pichler. *Multistage Stochastic Optimization*. Springer Series in Operations research and Financial Engineering. Springer, 2014.
[19] G. C. Pflug and A. Pichler. Dynamic generation of scenario trees. *Computational Optimization and Applications*, 62(3):641–668, 2015.
[20] F. B. Tekle, D. W. Gudicha, and J. K. Vermunt. Power analysis for the bootstrap likelihood ratio test for the number of classes in latent class models. *Advances in Data Analysis and Classification*, 10:209–224, 2016.