



Arbitrage conditions for electricity markets with production and storage

Raimund M. Kovacevic¹ 

Received: 13 April 2018 / Accepted: 4 March 2019 / Published online: 11 March 2019
© The Author(s) 2019

Abstract

This work analyzes the notion of arbitrage for market models with electricity production from fuel. Several generators, fuel storage, and the related fuel and storage costs are considered. Based on stochastic optimization in Banach spaces, a necessary and a sufficient no-arbitrage condition in terms of stochastic discount factors are derived and analyzed further in the context of (potentially nonlinear) vector-autoregressive price models with additive residuals. For this large class of statistical models, it is found that both the necessary and the sufficient condition can be rejected only in very rare cases. Based on these results, it is demonstrated how absence of arbitrage can be used to construct superhedging based valuation formulas for physical electricity delivery contracts, and that absence of arbitrage in tree-based stochastic programming can be ensured easier in the context of electricity production than in a purely financial context.

Keywords Electricity production · Arbitrage · Stochastic discount factor · Duality theory

1 Introduction

In many countries electricity is traded on exchanges, with prices determined by supply and demand. Such markets usually are quite 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, pro-

✉ Raimund M. Kovacevic
raimund.kovacevic@tuwien.ac.at

¹ Institute for Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

duced and used electric power has to be balanced immediately in an electrical network, because 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 not 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. Moreover, 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 discounted prices with respect to this measure, i.e. the discounted prices are martingales under the equivalent measure. An arbitrage free financial market is complete if and only if there is a unique martingale measure. As a consequence, on complete markets, every contingent claim is attainable by dynamic hedging, and financial derivatives can be priced by calculating the expected discounted value of the derivatives payoff with respect to the unique martingale measure. On incomplete markets there are claims that can not be replicated by the basic traded assets. Still, under absence of arbitrage there exist equivalent martingale measures, but uniqueness does not hold any more.

Because of the discussed frictions, electricity markets are not complete. Some submarkets for electricity, in particular futures markets, are organized like financial markets. However, the delivery profiles of traded futures usually cannot fully replicate physically traded delivery profile (although hedging with futures contracts is important in practice, see e.g. Deng et al. 2001).

The question remains, whether electricity markets or related market models, used by decision makers, are arbitrage free. This is not only a theoretical question, but is important both for valuation of contracts and also for taking model based management decisions. Valuation of contracts usually has to be based at least on the assumption that arbitrage is not possible. Moreover in a decision making context, facing a market with arbitrage possibilities decision makers clearly would like to exploit this. However, in reality for most market participants markets usually are arbitrage free (or arbitrage possibilities are unexploitable because of costs or other frictions). If the management of a firm then uses a market model that is not arbitrage free (which could e.g. happen by using flawed price models or by discretizing continuous state price models in a wrong way), there is a danger that the resulting decisions try to exploit arbitrage possibilities that are present in the model but not in reality. This is especially the case if problematic price models are used to parametrize optimization models, see e.g. Geyer et al. (2010) for a discussion in the context of tree-based stochastic optimization in finance. When this happens, high expectations are deceived, and it is likely that the implemented decisions lead into a completely wrong direction.

In Kovacevic (2018), no-arbitrage conditions for an electricity market with electricity generation from fuel and fuel storage were derived analytically, and the results were used to find valuation and pricing formulas for electricity delivery contracts. The present work goes back one step and aims at a deeper discussion of arbitrage properties in a more general setup. This includes several generating units with different production efficiency and takes into account the effects of storage costs for fuel. In a first step the results are based on duality theory for cone-constrained optimization in

Banach spaces and it turns out that existence of a martingale measure (as in finance) has to be replaced by more complex requirements.

The next question then is, how restrictive the no-arbitrage conditions are in a more concrete—model based—setup. In particular, when estimating parameter values of an econometric model from data, one may ask whether the no-arbitrage conditions put any restrictions on the parameters (and hence the estimation). In the present work these questions are analyzed for price processes with (potentially nonlinear) vector-autoregressive structure. Moreover, the consequences for the valuation of electricity delivery contracts and for the construction of scenario trees for stochastic optimization are discussed.

The work is organized as follows: Sect. 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. A necessary and a sufficient condition are derived. In Sect. 3 we analyze the problem of testing the assumption of no-arbitrage in such a market based on the necessary, respectively the sufficient condition. Section 4 gives an outlook of applications of the obtained results to the valuation of delivery contracts and to tree construction for stochastic optimization. Finally, Sect. 5 concludes the work.

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

In the present section, conditions for absence of arbitrage between fuel and electricity prices when it is possible to produce electrical energy from fuel, are derived. The analysis starts with describing a basic discrete time, general state space framework (in Banach spaces). In the present section the word “market” may refer as well to a market in the economical sense as to a market model (e.g. an econometric price model). Insofar economical markets are considered, it is assumed that a potential producer acts as a price taker at both markets, fuel and electricity. The resulting formulation does not aim at modeling an economic market equilibrium. Instead, as usual in financial economics, the simpler question which properties of price processes and (in the case of electricity production) production equipment allow or prohibit arbitrage is in the forefront. In particular, the presented approach does not state, whether, how or after which time span arbitrage possibilities are removed (when detected). In a production context arbitrage (when observed) may persist some time, because adjusting the production capacities (which finally would lead to price adjustment) is expensive and takes time. This is the reason why in the present work arbitrage is defined with respect to production efficiencies, which takes into account that some producers with very efficient generators might be able to use arbitrage possibilities while this is not possible with smaller efficiency.

Subsequently, the analysis is based on a definition of arbitrage which is a slight modification of the standard approach from finance but is adapted to dealing with electricity production from fuel, including the handling of fuel storage. It is then possible to use an optimization problem as an equivalent test tool for implementing

the rather unwieldy original definition. In the main part of this section, duality theory in Banach spaces is used to derive convenient conditions in terms of stochastic discount factors, which are formulated directly as properties of fuel and electricity prices and can be compared to standard financial results.

The formulation uses discrete points in time $t \in \{0, 1, 2, \dots, T\}$. At these times t all relevant information is observed and decisions are taken. Time 0 represents the begin (“here and now”) and T denotes the end of the planning horizon. For simplicity constant time increments, e.g. hours, days or weeks, are used. In order to analyze arbitrage properties, a stochastic process $X_t^f(\omega)$ [currency units/MWh] of fuel prices and a stochastic process $X_t^e(\omega)$ [in currency units/MWh] of electricity prices are considered. Both price processes are defined on a filtered probability space $\mathfrak{Y} = (\Omega, \mathcal{F}, \mathfrak{F} = \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ in discrete time $t = 0, 1, \dots, T$. Each σ -algebra \mathcal{F}_t represents the information available at time t . At the beginning, the σ -algebra \mathcal{F}_0 is the trivial σ -algebra, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$. The filtration \mathfrak{F} may be generated by the price processes, but this is not a necessary requirement. To simplify the notation the sets $\mathcal{T} = \{0, 1, \dots, T\}$, $\mathcal{T}_0 = \{0, 1, \dots, T - 1\}$, $\mathcal{T}_1 = \{1, \dots, T\}$ and $\mathcal{T}_1^{T-1} = \{1, \dots, T - 1\}$ are used in the following. As in reality, fuel prices are assumed to be almost surely nonnegative while electricity prices may be negative with positive probability.

Both, the energy content of the fuel and electrical energy are measured in MWh (energy content). 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]. Instead of the interest rate r , often the compounding factor $R = (1+r)$ is used. At each point in time t the producer decides first the amount z_t [MWh] of fuel traded on the fuel market at price X_t^f . 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 $i \in \{1, 2, \dots, I\}$. Generator i produces with efficiency η_i and the set of efficiencies is denoted by $\eta = (\eta_1, \dots, \eta_I)$. 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 calculating the new cash position¹ Hence the amount of fuel burned for producing electricity is given by $\sum_i \eta_i^{-1} y_{it}$ [MWh]. Electricity production y_t and fuel storage s_t are almost surely nonnegative.

The producer can store fuel, e.g. by using the facilities of an oil & gas storage service company (above ground reservoirs, pipe storage or underground storage like salt caverns), which causes costs. In the present work, a very simple cost model is considered, namely storage costs Ψ_t (payable at time t) that are proportional to the stored amount, in particular the specification

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

¹ This is a stylized time model. It models a delay between decision and implementation and also accounts for the fact that it is too late to take decisions after the new price is revealed. If each period $[t-1, t]$ represents a longer time interval like days or weeks, prices refer to mean prices over the period.

is used, where $\psi \geq 0$ is the related cost factor [currency units/MWh]. This additional simplification can be interpreted as an assumption that storage is filled (respectively emptied) uniformly over any period $[t - 1, t]$.

Startup costs are neglected in the present work (or rather it is assumed that they are approximately independent from the implemented production plan). Even if such assumptions may be easily violated in practice, it is clear that if arbitrage is not possible in a model without startup costs then arbitrage is also not possible when arbitrage costs are added.

In the rest of the section, to simplify notation, all equations and inequalities involving random variables are assumed to hold almost surely. It is assumed that fuel prices and electricity prices are essentially bounded, i.e. $X_t^f, X_t^e \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$. Later on, the notation $X_{[t]}^f, X_{[t]}^e$ will be used in order to denote relevant price histories up to time t . The decision processes y_t and z_t and the decision processes c_t and s_t are considered as real-valued random processes defined on \mathfrak{Y} and it is assumed 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} , so decisions at time t are based only on information available at (up to) this time. Because \mathcal{F}_0 is the trivial σ -algebra, the starting values c_0, s_0, y_0, z_0 are deterministic decisions. For the subsequent optimization problem this setup allows to use Lagrange multipliers from $L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, which can be identified with the dual space of $L^1(\Omega, \mathcal{F}_t, \mathbb{P})$.

Based on this setup it is possible to adapt the standard concepts of a self financing strategy in the following way:

Definition 2.1 A strategy $\{y_t, z_t\}_{t \geq 0}$ with a 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 $t \in \mathcal{T}_1$:

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

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

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

$$V_t^\eta = c_t + X_t^f \cdot s_t$$

The first equation models a cash position (an interest paying account) which changes when fuel is bought, electricity is sold or 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.

Using the notion of a self financing strategy, the classical definition of arbitrage (see e.g. Björk 2009 definitions 2.14, 2.15) can be adapted to the current context. As already discussed, arbitrage is defined with respect to production efficiencies.

Definition 2.2 An η -arbitrage for a market $\{X_t^e, X_t^f\}$ is a self financing strategy $\{y_t, z_t\}_{t \geq 0}$ with

$$V_0^\eta \leq 0, \tag{2.3}$$

$$\mathbb{P}(V_T^\eta \geq 0) = 1. \tag{2.4}$$

$$\mathbb{P}(V_T^\eta > 0) > 0. \tag{2.5}$$

A market $\{X_t^e, X_t^f\}$ is called η -arbitrage free, if no η -arbitrage exists.

This definition is hard to implement directly, especially because of the third condition. However, based on the definition, it is possible to formulate the following optimization problem which can be used to detect arbitrage strategies in the described setup.

$$\max_{y, z, c, s} \mathbb{E}^{\mathbb{P}} [c_T + X_T^f s_T] \tag{2.6}$$

subject to:

$$(t \in \mathcal{T}_1) : c_t = (c_{t-1} - z_{t-1} X_{t-1}^f) R + X_t^e \sum_{i=1}^I y_{it-1} - \psi \frac{s_t + s_{t-1}}{2} \tag{2.7}$$

$$(t \in \mathcal{T}_1) : s_t = s_{t-1} - \sum_{i=1}^I \eta_i^{-1} y_{it-1} + z_{t-1} \tag{2.8}$$

$$c_0 + X_0^f s_0 \leq 0 \tag{2.9}$$

$$c_T + X_T^f s_T \geq 0 \tag{2.10}$$

$$(t \in \mathcal{T}) : s_t \geq 0 \tag{2.11}$$

$$(t \in \mathcal{T}_0) : y_t \geq 0 \tag{2.12}$$

The first two constraints (2.7)–(2.8) are the self financing equations of Definition 2.1. Equations (2.9)–(2.10) are identical with the first two arbitrage conditions of Definition 2.2. Furthermore, (2.11)–(2.12) formulates the physical restrictions that fuel storage and electricity production cannot be negative. Finally, it can be seen that the expected end value of the strategy—the objective function (2.6) of the optimization problem—is unbounded if and only if the strategy value [nonnegative by (2.10)] can become positive. This follows from the fact that the feasible set is a pointed cone, and feasible solutions can be scaled by positive factors ad libitum. Altogether, optimization problem (2.6)–(2.12) is unbounded if and only if there is an η -arbitrage between fuel and electricity price. Note that there always exists a solution because setting all decision variables to zero is feasible.

The test problem has the form of a simple production planning problem, where the expected end value is maximized. However, this is only a consequence of Definitions 2.1 and 2.2. Its aim is not production planning, but detecting inconsistencies (arbitrage) between a fuel price and the electricity price. The focus is on one fuel,

however if several fuels are relevant, then the same test can be applied to all of them separately.

Note also that Problem (2.6)–(2.12) is formulated without upper bounds on storage and production (which would be important when considering planning, or contract valuation problems). They are not necessary in the present pure test context because of the cone property of the feasible set: If the optimization problem is reformulated with upper bounds on storage and production and get a positive optimal value, then this means that the solution can be scaled such that it leads to infinity in the problem without bounds. In the same manner, unboundedness of (2.6)–(2.12) means that there is a solution that leads to a positive end value with positive probability. Such a solution can be scaled, such that boundaries on storage and production are fulfilled.

Let $\eta_{max} = \max \{ \eta_1, \dots, \eta_I \}$ be the efficiency of the most efficient generating unit. Then, using duality in Banach spaces, it is possible to derive the following equivalent conditions:

Proposition 2.3 *A market $\{ X_t^e, X_t^f \}$ 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: $R^{t+1} \mathbb{E}^\mathbb{P} [\xi_{t+1} | \mathcal{F}_t] = R^t \xi_t$ for $t = 1, \dots, T - 1$, and $R \mathbb{E}^\mathbb{P} [\xi_1] = 1$

A4: $\mathbb{E}^\mathbb{P} [\xi_{t+1} X_{t+1}^e | \mathcal{F}_t] \leq \eta_{max}^{-1} \xi_t X_t^f$ for $t \in \mathcal{T}_0$

A5: $\mathbb{E}^\mathbb{P} [\lambda_{t+1} | \mathcal{F}_t] = \xi_t \cdot X_t^f$ for $t \in \mathcal{T}_0$ and $\mathbb{E}^\mathbb{P} [\lambda_1] = X_0^f$

A6: $\xi_t \cdot \left[X_t^f - \frac{\psi}{2} \left(1 + \frac{1}{R} \right) \right] \leq \lambda_t$ for $t \in \mathcal{T}_1^{T-1}$ and $\xi_T \left[X_T^f - \frac{\psi}{2} \right] \leq \lambda_T$

Proof The proof of is rather long and is moved to the ‘‘Appendix’’ in order to enhance readability. □

Remark 2.4 Given Proposition 2.3 one might also refer to η_{max} -arbitrage instead of η -arbitrage.

From A1)–A6) it can be seen that the process ξ —closely related to the process of Lagrange multipliers of the cash equation (2.7)—plays the role of positive stochastic discount factors or state price deflators (see e.g. Cochrane 2005 section 1.2 and Back 2010 section 2.2) which fulfill the martingale property A3). In particular it is easy to show $\mathbb{E} [\xi_t] = \frac{1}{R^t}$. In fact, the proposition states that there is no arbitrage if and only if there exists a stochastic factors with the properties A4)–A6). This is similar to the situation in finance where absence of arbitrage also can be linked to the existence of positive discount factors. However, in standard financial applications the typical condition is that all discounted market prices are martingales. Condition A4) then aims at consistency between fuel and electricity prices and states that the conditional expected proceeds from selling of one MWh electricity should not be larger than the costs for producing it, if everything is discounted properly. Conditions A5)–A6) are harder to interpret in financial terms because of the additional process λ , which is derived from the shadow costs of storage.

It is possible however, to state an easy interpretable necessary condition and an easy interpretable sufficient condition for absence of arbitrage, without mentioning the process λ .

Corollary 2.5 *If a market $\{X_t^e, X_t^f\}$ is η -arbitrage free then there exists a stochastic process ξ (fulfilling properties A1–A3 of Proposition 2.3) such that the consistency requirement A4) holds together with*

$$\mathbb{E}^{\mathbb{P}} \left[\xi_{t+1} \left(X_{t+1}^f + B_{t+1} \right) | \mathcal{F}_t \right] \leq \xi_t \left(X_t^f + B_t \right) \text{ for } t \in \mathcal{T}_1^{T-1} \tag{2.13}$$

with

$$B_t = \frac{\psi}{2(R-1)} \left(1 + \frac{1}{R} \right) \tag{2.14}$$

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

$$B_T = \frac{\psi}{R-1}. \tag{2.15}$$

Proof If a market is η_{max} -arbitrage free then by Proposition 2.3 there exist processes ξ, λ fulfilling A1–A6. 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 $t = 1, 2, \dots, T-2$. 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 (2.13). 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. Here the resulting inequality is expanded by $\frac{\psi}{R(R-1)} \xi_{T-1}$ \square

Regarding fuel prices, condition (2.13) of this corollary is already close to the standard financial case: a discount factor process ξ_t is required such that the modified fuel price $X_{t+1}^f + B_t$ is a supermartingale, if discounted by ξ_t . The terms B_t account for storage costs, respectively the cost of carry. In the absence of storage costs (i.e. $\psi = 0$) the fuel price itself must be a supermartingale, if properly discounted. Still, the electricity price is restricted only indirectly via the consistency requirement A4).

An interpretable sufficient condition for absence of arbitrage can be formulated in the following way.

Corollary 2.6 *Consider a market with prices $\{X_t^e, X_t^f\}$. If there is a process ξ (fulfilling properties A1–A3 of Proposition 2.3) such that the consistency requirement A4) holds together with*

$$\mathbb{E}^{\mathbb{Q}} \left[\xi_{t+1} X_{t+1}^f | \mathcal{F}_t \right] = \xi_t X_t^f \tag{2.16}$$

(i.e. the discounted fuel pieces are a martingale) then the market is η -arbitrage free.

Proof Set $\lambda_t = \xi_t X_t^f$. This choice fulfills A6 because $\frac{\psi}{2} (1 + \frac{1}{R}) \geq 0$. Substituting λ_{t+1} for $\xi_{t+1} X_{t+1}^f$ at the left side of (2.16) leads to A5. Because A1–A4 hold already by assumption, Proposition 2.3 implies absence of η_{max} -arbitrage. \square

If there exists a process of stochastic discount factors such that already the discounted fuel price is a martingale (the typical financial requirement) and the consistency requirement holds, then absence of arbitrage follows.

3 Implications of the Arbitrage conditions

The preceding section gave a complete characterization of arbitrage in a general framework. The question arises what the obtained results mean in a more concrete context when statistical models are used to describe the price processes. Can arbitrage easily arise, such that one has to use restrictions when estimating model parameters? Many models for fuel and electricity prices have been proposed in literature, see e.g. Nowotarski and Weron (2018) for a recent overview of electricity price modeling. The present work therefore can only be a starting point and the analysis therefore is restricted to one possible class of models, namely a class of (potentially nonlinear) vector-autoregressive econometric models. This class of models for electricity and fuel prices can be described by

$$X_{t+1}^e - E_{t+1} = \phi^e(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta) + \varepsilon_{t+1}^e \tag{3.1}$$

$$X_{t+1}^f - F_{t+1} = \phi^f(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta) + \varepsilon_{t+1}^f. \tag{3.2}$$

Here E_t and F_t are given and denote a given deterministic process, e.g. market expectations, or an observed forward price-curve, or can be even zero. The considered model then is formulated relative to these processes. The notation $X_{[t]}^i = (X_t^i, X_{t-1}^i, \dots, X_{t-p}^i)'$, where $i \in \{e, f\}$ 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). Negative $t < 0$ denote observations that have been made before the actual planning horizon $\{0, \dots, T\}$. 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 . Note that the short notation $\phi_t^i(\theta) = \phi^i(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta)$ is used in the following.

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

$$X_{t+1}^e - E_{t+1} = \theta_1'(X_{[t]}^e - E_{[t]}) + \varepsilon_{t+1}^e \tag{3.3}$$

$$X_{t+1}^f - F_{t+1} = \theta_2'(X_{[t]}^f - F_{[t]}) + \varepsilon_{t+1}^f, \tag{3.4}$$

where θ_1, θ_2 are parameter vectors of dimension p . Other linear specifications might use suitable differences between lagged variables. Clearly, nonlinear models may be computational demanding.

Finally, it is assumed that—given the past—the error terms ε_{t+1}^i follow a distribution described by some joint conditional distribution function G . The distribution function

respects the assumption that fuel prices X_{t+1}^f are nonnegative. Moreover, the error terms are characterized by

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

and

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

Although the concrete joint distribution function of the error terms might also be parametrized by additional model parameters, 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.

3.1 Implications of the necessary condition for absence of arbitrage

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, the consequences of conditions A1–A4 and inequality (2.13) are analyzed 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. It will turn out that the answer is positive only in very exceptional cases.

Recall that $R^t \xi_t$ has to be a martingale according to condition A3). Therefore, it is possible to write

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

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 e.g. that the martingale difference sequence cannot be a sequence of i.i.d. random variables.

In the current context with (3.1)–(3.7) Corollary (2.5) can be reformulated in the following way.

Corollary 3.1 *Let the price processes $\{X_t^e, X_t^f\}$ be specified by model (3.1)–(3.2). If the model 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 [obeying (3.7)] such that*

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

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 A4) 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 (3.8), the first statement of the corollary.

The same arguments, applied to (2.13), lead to the second statement (3.9). □

Defining now a process

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

it can be seen that $\mathbb{E} \left[v_{t+1} \varepsilon_{t+1}^e | \mathcal{F}_t \right] = Cov \left[v_{t+1} \varepsilon_{t+1}^e | \mathcal{F}_t \right] = \rho_{et} \cdot \sigma_e \cdot \sigma_{v_t}$, 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 where σ_{v_t} is the conditional standard deviation of v_{t+1} . In the same manner the relation $\mathbb{E} \left[v_{t+1} \varepsilon_{t+1}^f \right] = \rho_{ft} \cdot \sigma_f \cdot \sigma_{v_t}$ is obtained, 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

$$\rho_{et} \cdot \sigma_e \cdot \sigma_{v_t} \leq G_t \tag{3.11}$$

and

$$\rho_{ft} \cdot \sigma_f \cdot \sigma_{v_t} \leq H_t \tag{3.12}$$

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} to fulfill the inequalities when σ_e and σ_f (and also θ and R , η) are given parameters? First σ_{vt} must be nonnegative and can be chosen as large as necessary, as shown by the following corollary.

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*

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

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

and

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

Moreover, a joint distribution for v_{t+1} , ε_{t+1}^e , ε_{t+1}^f 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 $f(x) = p \frac{1}{K_1} \iota_{[-K_1, 0]}(x) + p \frac{1}{K_2} \iota_{[0, K_2]}(x)$, where $\iota_A(x)$ 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 $K_1 K_2 = 12M^2$, then Eqs. (3.13)–(3.15) are fulfilled. Finally, given the joint distribution of ε_{t+1}^e , ε_{t+1}^f 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_{t+1}^e , ε_{t+1}^f that fulfills all requirements. \square

The first two requirements in Corollary (3.2) ensure that the process ξ_t is a positive martingale. 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. It is preferable 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} which is negatively correlated to both error terms ε_t^e and ε_t^f , when the errors are correlated with coefficient ρ as specified in (3.6). To answer this question, recall that in an inner product space one can define the angle θ between two elements X , Y by $\langle X, Y \rangle = \|X\| \|Y\| \cos(\theta)$. Considering that in the present context $\|X\| = \sqrt{\text{Var}(X|\mathcal{F}_t)} = \sigma_X$ and $\langle X, Y \rangle = \text{Cov}(X, Y|\mathcal{F}_t)$ hold, it can be seen that for the

angle θ_ε between the errors ε_t^e and ε_t^f the relation $\cos(\theta_\varepsilon) = \rho$ is valid. Moreover, for the angles θ_e and θ_f between v_{t+1} and the respective errors ε_t^e and ε_t^f the equations $\cos(\theta_e) = \rho_{et}$ and $\cos(\theta_f) = \rho_{ft}$ are obtained. By standard geometric arguments it can then be seen that it is possible to choose both (cosines) correlations negative, if and only if

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

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

The special role of ρ , the correlation between the error terms, 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_{t+1}^e , ε_{t+1}^f and v_{t+1} stays nonnegative definite. This reduces to

$$\left[\begin{array}{ccc} 1 & \rho & \rho_{et} \\ \rho & 1 & \rho_{ft} \\ \rho_{et} & \rho_{ft} & 1 \end{array} \right] \geq 0$$

or (the first two principal minors are obviously nonnegative)

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

With ρ given, for $0 < \rho < 1$ this inequality (in ρ_{et} , ρ_{ft}) describes an elliptical disk. The ellipse has its center in the origin, and 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, with $\rho = +1$ 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 Fig. 1 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 $\rho_{ft} = -\rho_{et}$.

The situation is different if one of σ_e or σ_f equals zero.

If $\sigma_e = 0$ then inequality (3.11) can be reduced to the condition

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

The left-hand side of the second inequality (3.12) can be made arbitrarily small using the same arguments as above. If $\sigma_f = 0$ then inequality (3.12) reduces to

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

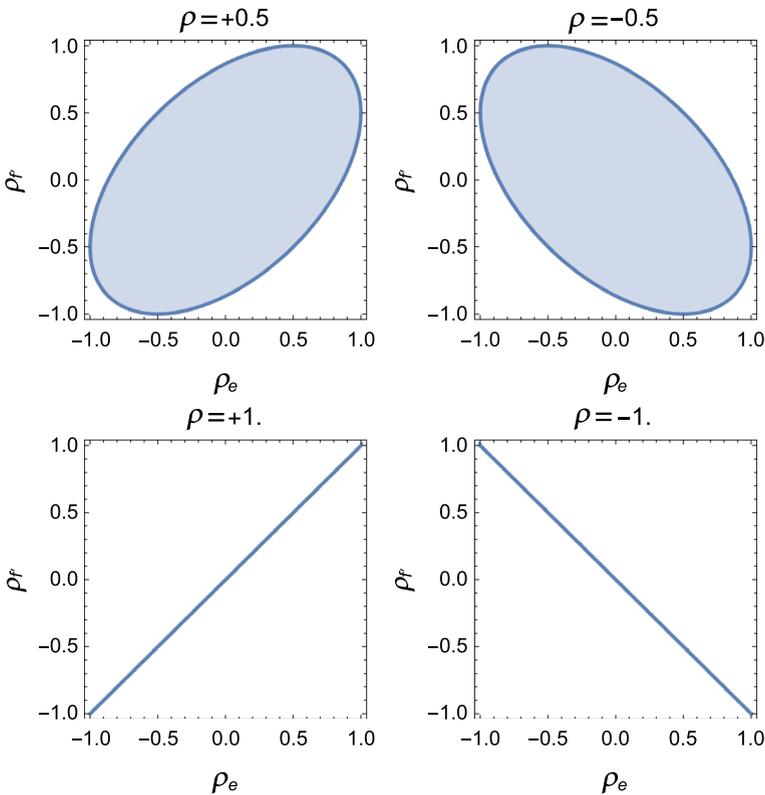


Fig. 1 Feasible combinations of ρ_{et}, ρ_{ft} (correlation between v_{t+1} and the errors ε_{t+1}^e , respectively ε_{t+1}^f), dependent on different values of ρ (the correlation between the errors)

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 absence of arbitrage is ensured if both inequalities (3.19)–(3.20) hold at all points in time t .

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 for some t , and otherwise the hypothesis is not rejected. If parameter values are not given and have to be estimated, it is possible to conclude:

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^f are observed such that for some parameter value θ one of the deterministic cases

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

or

C) both equations A) and B)

hold for all points in time t . Then absence of arbitrage [inequalities (3.8)–(3.9)] cannot be rejected if for all points in time t (3.19) is fulfilled in case A), (3.20) is fulfilled in case B), respectively both (3.19)–(3.20) are fulfilled at all points in time 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 there is some $\kappa > 0$ such that for some parameter value θ the equation

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

holds for all points in time t , then absence of arbitrage [inequalities (3.8)–(3.9)] cannot be rejected.

Proof Cases A), B) and C) refer to cases where 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 price variances equal to zero. The correct conditions for validity of (3.8)–(3.9) already have been discussed above.

If Eq. (3.21) is not fulfilled this means that $\rho \neq -1$, because $\rho = -1$ is equivalent to $\varepsilon_{t+1}^e = -\frac{1}{\kappa} \varepsilon_{t+1}^f$ (for some $\kappa > 0$). In this case it was already shown 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 for observed price data and therefore absence of arbitrage usually cannot be rejected.

3.2 Implications of the sufficient condition for absence of arbitrage

So far the main result is that for a large class of possible models the necessary conditions A4) and (2.13) for absence of arbitrage can be rejected only in rare, even unrealistic cases. In the following, similar arguments are used to analyze the implications of the sufficient condition A4) and (2.16) in Corollary 2.6. Again the class of autoregressive models specified in (3.1)–(3.2) and (3.5)–(3.6) is considered. The process ξ is rewritten by (3.7). In particular the processes u and v , defined in (3.10), have the same properties as before. Then the sufficient conditions A4) and (2.16) can be reduced to

$$\rho_{et} \cdot \sigma_e \cdot \sigma_{vt} \leq G_t \tag{3.22}$$

and

$$\rho_{ft} \cdot \sigma_f \cdot \sigma_{vt} = H_t \tag{3.23}$$

with

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

and

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

Surprisingly, like with necessary conditions, the sufficient conditions (3.22)–(3.23) do not put severe restrictions on the model parameters.

Proposition 3.4 *For the autoregressive model (3.1)–(3.2) and (3.5)–(3.6) absence of arbitrage is ensured in any of the following cases:*

1. $\sigma_e = 0$, $\sigma_f > 0$ and for all t

$$\eta_{\max}^{-1} X_t^f - \frac{1}{R} (E_{t+1} + \phi_t^e(\theta)) \geq 0. \quad (3.24)$$

2. $\sigma_e > 0$, $\sigma_f = 0$ and for all t

$$X_t^f - \frac{1}{R} (F_{t+1} + \phi_t^f(\theta)) = 0 \quad (3.25)$$

3. Else, either $|\rho| < 1$, or both (3.25) and (3.24) hold.

Proof The “trivial” cases of zero variance are considered first. If $\sigma_e = 0$ and $\sigma_f > 0$ then it is sufficient for absence of arbitrage that the almost sure conditions (3.22), hence (3.22), hold at each point in time. This is true because one is free to choose a conditional correlation ρ_{ft} and a standard deviation σ_{vt} such that (3.23) is also fulfilled. In similar manner, if $\sigma_f = 0$ and $\sigma_e > 0$ then conditions (3.23), hence (3.25), must hold at any point in time to ensure absence of arbitrage. If both conditional variances are zero, then both of Eqs. (3.24) and (3.25) must hold simultaneously at any point in time to imply absence of arbitrage.

Let now $\sigma_e > 0$ and $\sigma_f > 0$. In each scenario ω where $H_t(\omega) = 0$, it is possible to choose $\rho_{ft}(\omega) = 0$ to ensure (3.23). This allows to take ρ_{ft} and σ_{vt} such that also (3.22) stays valid. From Eq. (3.18) it can be seen that the choice $\rho_{ft}(\omega) = 0$ fully works as long as $|\rho| \neq 1$. If $|\rho| = 1$ then (3.18) implies that $\rho_{et}(\omega) = 0$ and therefore (3.25).

Finally, consider the case $H_t \neq 0$. Then (3.23) leads to

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

Note that because of $\sigma_e > 0$ and $\sigma_f > 0$, the conditional correlation ρ_{ft} must have the same sign as H_t (and is not equal to zero because of the considered case $H_t \neq 0$). Plugging σ_{vt} into (3.22) leads to

$$\frac{\sigma_e}{\sigma_f} \frac{H_t}{\rho_{ft}} \rho_{et} \leq G_t. \tag{3.26}$$

The multiplier 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 ρ_{ft} that fulfills equation (3.26)]. In order to account also for the case $G_t < 0$, ρ_{et} has to be chosen negative and the sign of ρ_{ft} depends on the sign of H_t as discussed above. Taking into account the correlation ρ and Eq. (3.18) (see also Fig. 1), such a choice of signs is always possible unless either $\rho = -1$ or $\rho = +1$. The absolute values of ρ_{ft} , ρ_{et} can be chosen such that the left-hand side of (3.26) becomes small enough. \square

Using the model specification (3.1)–(3.2) and (3.2) it is possible to reformulate Proposition (3.4) in the following way:

Corollary 3.5 *Assume that price processes X_t^e, X_t^f are observed such that for some parameter value θ one of the deterministic cases*

- A) $X_{t+1}^e - E_{t+1} = \phi^e(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta_1),$
- B) $X_{t+1}^f - F_{t+1} = \phi^f(X_{[t]}^e - E_{[t]}, X_{[t]}^f - F_{[t]}; \theta_2).$

hold for all points in time t . Then absence of arbitrage [sufficient conditions (3.22)–(3.23)] can be guaranteed if (3.24) holds in case A), respectively (3.25) holds in case B). If both, A) and B) hold for some parameter value θ then absence of arbitrage is implied if (3.24) and (3.25) hold simultaneously.

If none of A), B) holds, still (3.24) together with (3.25) implies absence of arbitrage. Moreover, arbitrage can also be excluded if there is no $\kappa \in \mathbb{R}$ such that for some parameter value θ the equation

$$X_{t+1}^f - F_{t+1} - \phi_t^f(\theta) = \kappa (\phi_t^e(\theta) - (X_{t+1}^e - E_{t+1})) \tag{3.27}$$

holds for all points in time.

Proof This follows immediately from Proposition (3.4). Regarding cases A), B) the error terms $(\varepsilon_t^e, \varepsilon_t^f)$ are almost surely equal to zero, if and only if their expectation and variance is zero. Moreover, a correlation ρ equal to 1 or -1 occurs if and only if $\varepsilon_t^e = \tau \varepsilon_t^f$ for some $\tau \in \mathbb{R}$. \square

We see now that for the vector-autoregressive model the sufficient conditions for are only slightly stronger than the necessary conditions. The requirement $|\rho| < 1$ is not very severe, and one can expect that unrestricted parameter estimation leads to an arbitrage free price model in most practical situations.

4 Two applications of arbitrage conditions

In the following two direct applications of the arbitrage results got so far are analyzed. First physical delivery contracts and their valuation is considered. 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 the second discussion it turns out that the cases with zero variance are important in 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.

4.1 Valuation of delivery contracts

Consider now a producer who participates on a market for physical electricity delivery contracts. While exchanges nowadays frequently use financial delivery, physical delivery is still in use. Moreover, off-exchange contracts are about physical delivery quite often. When deciding whether to enter a contract or how to value an already agreed contract, because the market is incomplete as discussed above, there is not a unique discount factor process ξ which gives an estimate of market values or prices in advance without taking into account individual preferences. Still the valuation problem can be analyzed from the viewpoint of the producer who may ask questions like: “What is the smallest up-front payment (alternatively: delivery price) such that all contractual obligations (deliveries) can be fulfilled and the (random) end value is acceptable?” This does not lead to market values or prices, but if e.g. market delivery prices are larger than the producers smallest possible price, then he will contract, otherwise he will not agree or at least he will have to take additional risk.

Building on arbitrage free models, it is possible to derive pricing and valuation formulas (super-hedging and acceptability prices) for delivery contracts with random delivery patterns under several types of assumptions. This was the focus in Kovacevic (2018), see also Kovacevic and Pflug (2013) where the topic was touched on the first time. This approach builds on and generalizes the ideas in King (2002), Flåm (2008), Pennanen (2011a, b) and Pennanen (2012), where financial markets instead of electricity markets are considered. In any case it turns out that the smallest value or price such that a contract does not lead to an unfavorable outcome can be expressed in terms of expectations w.r.t. equivalent measures, respectively stochastic discount factors.

In a simplified framework, the producer again is able to generate electricity with some efficiency η but also has the obligation to deliver amounts D_t [MWh] of electric energy at a fixed price of K per MWh for periods $[t, t + 1]$ —which defines a delivery contract. 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.

The 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 . 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 to meet obli-

gations. 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.

Out of the various cases analyzed in Kovacevic (2018), only the so called superhedging value is considered in the present work, 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.

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.

$$\begin{aligned}
 V_0^*(K, D, \eta) &= \min_{y, z, c, s} c_0 + X_0^f s_0 \\
 &\text{subject to} \\
 (t \in \mathcal{T}_1) : c_t &= (c_{t-1} - z_{t-1} X_{t-1}^f) R + y_{t-1} X_t^e - D_{t-1} X_t^e + K D_{t-1} \\
 (t \in \mathcal{T}_1) : s_t &= s_{t-1} - y_{t-1} \eta^{-1} + z_{t-1} \\
 c_T + X_T^f s_T &\geq 0 \\
 (t \in \mathcal{T}) : 0 &\leq s_t \leq S \\
 (t \in \mathcal{T}_0) : 0 &\leq y_t \leq P_t.
 \end{aligned} \tag{4.1}$$

All assumptions on the involved processes made in Sect. 2 are kept also in the current subsection. In addition it is assumed 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 (4.1) then is given by.

Notice that the Lagrange dual of the valuation problem (4.1) is given by

$$\begin{aligned}
 U_0^*(K, D, \eta) &= \max_{\xi, \lambda, \mu, v} \sum_{t=0}^{T-1} \mathbb{E}^\mathbb{P} [\xi_{t+1} (X_{t+1}^e - K) D_t] - \sum_{t=0}^{T-1} \mathbb{E}^\mathbb{P} [\mu_t P_t] - S \sum_{t=0}^T \mathbb{E}^\mathbb{P} [v_t] \\
 &\text{subject to} \\
 &A1) - A3) \text{ and } A5) \\
 A4') \mathbb{E}^\mathbb{P} [\xi_{t+1} X_{t+1}^e | \mathcal{F}_t] &\leq \eta^{-1} \xi_t X_t^f + \mu_t \text{ for } t \in \mathcal{T}_0 \\
 A6') \xi_t \cdot \left[X_t^f - \frac{\psi}{2} \left(1 + \frac{1}{R} \right) \right] &\leq \lambda_t + v_t \text{ for } t \in \mathcal{T}_1^{T-1} \\
 &\text{and } \xi_T \left[X_T^f - \frac{\psi}{2} \right] \leq \lambda_T + v_T \\
 (t \in \mathcal{T}) : \mu_t &\geq 0, \quad v_t \geq 0,
 \end{aligned} \tag{4.2}$$

where $\xi_t, \lambda_t, \mu_t, v_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$.

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 here are discounted by stochastic discount factors which fulfill conditions similar to the 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 Kovacevic (2018), a sufficient condition for equality of the primal and dual optimal values $V_0^*(K, D, \eta) = U_0^*(K, D, \eta)$ (strong duality) is that the market is arbitrage free (which implies that the interior of the feasible set is nonempty). In the light of the previous subsection when using vector autoregressive models an arbitrage free model of the form (3.3)–(3.6) can be ensured by just using unrestricted estimation and the valuation model (4.1), respectively its dual version (4.2) can be applied likewise.

4.2 Arbitrage free price models in tree-based multistage stochastic optimization

As already pointed out in the introduction, arbitrage free price models are important for any kind of optimization problem based on estimated prices. This comprises pricing problems as above, but also e.g. planning of electricity generation. So far a class of statistical models was considered, and it was shown that they are arbitrage free under quite general conditions. However, the used statistical model is not the only possible reason for arbitrage in an optimization problem. Often the statistical model is not used in a direct way, but by implementing a suitable approximation like e.g. a discretization of the original model. Then it must be ensured that arbitrage is not induced by limitations of the data structure representing the discretized model.

An important framework for solving decision problems is tree-based multistage stochastic optimization (see e.g. Pflug and Pichler 2014), 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.

Pflug and Pichler (2014), 1.4. (for an alternative formulation see e.g. Alonso-Ayuso et al. 2009) describes an approach where the original time oriented formulation involving time indices is replaced by a node oriented formulation: consider a finite probability space $\Omega = (\omega_1, \dots, \omega_S)$, representing S scenario-paths. Any stochastic process defined on this sample space can be represented as a finite tree with node set $\mathcal{N} = \{0, 1, \dots, N\}$. The levels of the tree correspond to the decision stages. Let \mathcal{N}_t be the set of nodes at level t , for $t = 0, \dots, T$. The final level \mathcal{N}_T contains the S leaves of the tree, each of which can be identified with a scenario paths: $\mathcal{N}_T = \Omega = (\omega_1, \dots, \omega_S)$. The tree structure represents the filtration of the process and can be defined (as a data structure) 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 relates 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. X_n^e, X_n^f for some $n \in \mathcal{N}$ is used instead of $X_t^e(\omega), X_t^f(\omega)$ for some $t \in \mathcal{T}$. 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

$$s_n = s_{n_-} - \sum_{i=1}^I \eta_i^{-1} y_{i n_-} + z_{n_-}$$

in the node oriented formulation.

Finally probabilities π_n can be assigned to all leaf nodes $n \in \mathcal{N}(T)$, 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 a fixed tree structure), see e.g. Dupacova et al. (2003), Heitsch and Römisch (2010), Pflug and Pichler (2015). The outcome are price values and probabilities at all nodes. In the context of the basic model (3.1)–(3.3) this means that for a node n the model equations are replaced with

$$X_n^e - E_n = \phi^e(X_{[n_-]}^e - E_{[n_-]}, X_{[n_-]}^f - F_{[n_-]}; \theta) + \varepsilon_n^e \tag{4.3}$$

$$X_n^f - F_n = \phi^f(X_{[n_-]}^e - E_{[n_-]}, X_{[n_-]}^f - F_{[n_-]}; \theta) + \varepsilon_n^f, \tag{4.4}$$

where $X_{[n_-]}^i$ 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 errors $\varepsilon_m^e, \varepsilon_m^f$ 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 to obtain an arbitrage free model as discussed above, also the approximating tree model stays arbitrage free under quite general circumstances.

However, when many decision stages are involved, scenario trees may contain numerous nodes. Decomposition techniques and other algorithmic tools exist in literature to reduce the resulting computational effort. However, often it is necessary to reduce the density of scenario trees: the number of node-successors might be small, some nodes even might have a single successor. In such nodes the conditional variances of the errors 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

$$X_n^e - E_n = \phi^e(X_{[n-]}^e - E_{[n-]}, X_{[n-]}^f - F_{[n-]}; \theta) \quad (4.5)$$

$$X_n^f - F_n = \phi^f(X_{[n-]}^e - E_{[n-]}, X_{[n-]}^f - F_{[n-]}; \theta), \quad (4.6)$$

as usually done. These expected prices may lead to arbitrage because they pretend perfect foresight. According to the previous results, this can be avoided by prices fulfilling (3.24)–(3.25). These conditions have priority and one might use e.g. prices that are as close as possible to the conditional expectations as possible (in some distance) but still fulfill the no-arbitrage conditions.

Note that in financial applications one also aims at using arbitrage free prices in planning applications like portfolio optimization or asset-liability management, to exclude unrealistic gains. In Geyer et al. (2010) it was claimed that if M assets and N successor nodes for some node in the tree are considered and $N < M$, then arbitrage is always possible. Following Duffie (2001), if π denotes the vector of prices (dimension M) in this node and D is the $M \times N$ matrix of asset prices in the N successor nodes, then absence of arbitrage is equivalent to the existence of some vector $\xi \in \mathbb{R}^N$, $\xi > 0$ (stochastic discount factor) such that

$$D \cdot \xi = \pi. \quad (4.7)$$

While the claim in Geyer et al. (2010) is not fully correct,² it is often true in practical cases. For $N < M$, arbitrage can only be excluded if the price vector π (dimension M) lies in the N -dimensional cone of positive linear combinations of the column vectors of D , which is a strong requirement. For one successor node ($N = 1$) and two prices ($M = 2$), as in the discussion above, this criterion means that the payoff D must be a (positive) multiple of the price vector π to avoid arbitrage. As a special case: when D is defined by expected prices and the prices are martingales under the physical measure (which is very unlikely), this would ensure the absence of arbitrage. When prices are not martingales under the physical measure, one could enforce absence of arbitrage by replacing the expectations by the prices of the predecessor nodes, which would heavily change the system dynamics. Moreover, even if there are enough ($N \geq M$) successors it must be shown that $\xi > 0$ exists such that (4.7) holds to ensure that the prices are arbitrage free.

At first glance this seems to be a contradiction to the previous findings in the present work with $M = 2$ —at least in the context of vector autoregressive models, where absence of arbitrage can be excluded under normal conditions when $N \geq M$ and even in the problematic case $N = 1$ (with variances equal to zero), absence of arbitrage can be ensured by using prices close to an initial estimate, but fulfill the inequalities (3.24)–(3.25). For fuel prices this still means that they have to martingales under the original measure, condition (3.24) however is different and much weaker than the requirement that also electricity price should be a martingale under the original measure. That absence of arbitrage is easier to achieve in the electricity production setting than in a purely financial setting can be explained by the additional frictions, present on the considered market (in particular the nonnegativity of fuel storage and irreversibility of electricity production). Still, a positive ξ is required, but the conditions (3.24)–(3.25)

² The author thanks an anonymous referee for pointing out this fact.

are much less restrictive than the pure equation system (4.7). So the frictions make it more difficult to achieve arbitrage.

5 Conclusions

In the present work 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 was derived. Despite the fact that electricity markets show unique frictions, 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 to avoid unrealistic outcomes.

These questions were analyzed for a class of (potentially nonlinear) vector-autoregressive prices models for fuel and electricity. It turned out that actually the derived conditions for absence of arbitrage restrict the space of parameter values only to a very slight extent. If one ignores deterministic price models, it is therefore to use unrestricted estimation approaches (like e.g. unrestricted maximum likelihood) for estimation.

Note that the fact that the considered vector-autoregressive price model has additive error terms is crucial for the proofs given in this work. While it has been shown here that it is virtually impossible to achieve arbitrage in such a setup, it remains open under which conditions arbitrage might be possible for other specifications. Therefore, the analysis of parametric models will be extended to further model classes in future work.

This work 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 and usage of renewable energy might be considered.

Acknowledgements Open access funding provided by TU Wien (TUW). The author thanks Wolfgang Scherrer for interesting discussions, and two anonymous referees for their valuable comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Proof of Proposition (2.3)

The Lagrangian of problem (2.6)–(2.12) can be written as

$$\begin{aligned} L(y, z, c, s; \xi, \lambda, \zeta, \gamma) \\ = \mathbb{E}^{\mathbb{P}} \left[c_T + X_T^f s_T \right] \end{aligned}$$

$$\begin{aligned}
 &+ \mathbb{E}^{\mathbb{P}} \left[\zeta \left(c_T + X_T^f s_T \right) \right] \\
 &+ \sum_{t=1}^T \mathbb{E}^{\mathbb{P}} \left[\xi_t \left(R c_{t-1} - c_t + X_t^e \sum_{i=1}^I y_{it-1} - R X_{t-1}^f z_{t-1} \right) \right] \\
 &+ \sum_{t=1}^T \mathbb{E}^{\mathbb{P}} \left[\lambda_t \left(s_{t-1} - s_t - \sum_{i=1}^I \eta_i^{-1} y_{it-1} + z_{t-1} - \psi \frac{s_t + s_{t-1}}{2} \right) \right] \\
 &- \gamma \left(c_0 + X_0^f s_0 \right), \tag{6.1}
 \end{aligned}$$

where $\gamma \geq 0$ is a real number, $\zeta \geq 0$ 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 (6.1) are elements of $L^1(\Omega, \mathcal{F}_t, \mathbb{P})$ under the basic assumptions.

It is possible to rearrange (6.1) in the following way:

$$\begin{aligned}
 &L(y, z, c, s; \xi, \lambda, \zeta, \gamma) \\
 &= \mathbb{E}^{\mathbb{P}} [c_T (1 + \zeta - \xi_T)] + \mathbb{E}^{\mathbb{P}} \left[s_T \left(X_T^f (1 + \zeta) - \lambda_T - \frac{\psi}{2} s_T \xi_T \right) \right] \\
 &+ c_0 \left(\mathbb{E}^{\mathbb{P}} [\xi_1] R - \gamma \right) + s_0 \left(\mathbb{E}^{\mathbb{P}} [\lambda_1] - \gamma X_0^f - \frac{\psi}{2} \xi_1 \right) \\
 &+ y_0 \mathbb{E}^{\mathbb{P}} [\xi_1 X_1^e - \lambda_1 \eta^{-1}] + z_0 \mathbb{E}^{\mathbb{P}} [\lambda_1 - R \xi_1 X_0^f] \\
 &+ \sum_{t=1}^{T-1} \mathbb{E}^{\mathbb{P}} [c_t (R \xi_{t+1} - \xi_t)] + \sum_{t=1}^{T-1} \mathbb{E}^{\mathbb{P}} \left[s_t \left(\lambda_{t+1} - \lambda_t - \frac{\psi}{2} (\xi_t + \xi_{t+1}) \right) \right] \\
 &+ \sum_{i=1}^I \sum_{t=1}^{T-1} \mathbb{E}^{\mathbb{P}} [y_{it} (\xi_{t+1} X_{t+1}^e - \eta_i^{-1} \lambda_{t+1})] \\
 &+ \sum_{t=1}^{T-1} \mathbb{E}^{\mathbb{P}} [z_t (\lambda_{t+1} - R \xi_{t+1} X_t^f)] \tag{6.2}
 \end{aligned}$$

Using (6.2), the ‘‘tower property’’ of conditional expectation and keeping in mind $y_t, s_t \geq 0$, the dual function

$$\max_{y \geq 0, z, c, s \geq 0} L(y, z, c, s; \xi, \lambda, \zeta, \gamma), \tag{6.3}$$

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

$$\zeta \geq 0 \tag{6.4}$$

$$\xi_T = 1 + \zeta \tag{6.5}$$

$$\lambda_T + \frac{\psi}{2} \xi_T \geq (1 + \zeta) X_T^f \tag{6.6}$$

$$\gamma \geq 0 \tag{6.7}$$

$$\mathbb{E}^{\mathbb{P}} [\xi_1] R = \gamma \tag{6.8}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_1] \leq \gamma X_0^f + \frac{\psi}{2} \mathbb{E} [\xi_1] \tag{6.9}$$

$$\mathbb{E}^{\mathbb{P}} [\xi_1 X_1^e] \leq \eta^{-1} \mathbb{E}^{\mathbb{P}} [\lambda_1] \tag{6.10}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_1] = R \mathbb{E}^{\mathbb{P}} [\xi_1] X_0^f \tag{6.11}$$

$$R \mathbb{E}^{\mathbb{P}} [\xi_{t+1} | \mathcal{F}_t] = \xi_t \text{ for } t \in \mathcal{T}_1^{T-1} \tag{6.12}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] \leq \lambda_t + \frac{\psi}{2} (\xi_t + \mathbb{E} [\xi_{t+1} | \mathcal{F}_t]) \text{ for } t \in \mathcal{T}_1^{T-1} \tag{6.13}$$

$$\mathbb{E}^{\mathbb{P}} [\xi_{t+1} X_{t+1}^e | \mathcal{F}_t] \leq \eta_i^{-1} \mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] \text{ for all } i \in \{1, \dots, I\} \text{ and } t \in \mathcal{T}_1^{T-1} \tag{6.14}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] = R \mathbb{E}^{\mathbb{P}} [\xi_{t+1} | \mathcal{F}_t] X_t^f \text{ for } t \in \mathcal{T}_1^{T-1}. \tag{6.15}$$

These are the constraints of the dual optimization problem of (2.6)–(2.12) and it follows that the original problem is unbounded if and only if conditions (6.4)–(6.15) are fulfilled.

In the same way as in Kovacevic (2018) 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 (6.4)–(6.15) then $\xi_t = \frac{\xi'_t}{\gamma}$ fulfills

$$\xi_t > 0 \tag{6.16}$$

$$\lambda_T + \frac{\psi}{2} \xi_T \geq \xi_T X_T \tag{6.17}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_1] \leq X_0^f + \frac{\psi}{2R} \tag{6.18}$$

$$R \mathbb{E}^{\mathbb{P}} [\xi_{t+1} | \mathcal{F}_t] = \xi_t \text{ for } t \in \mathcal{T}_1^{T-1} \text{ and } R \mathbb{E}^{\mathbb{P}} [\xi_1] = 1 \tag{6.19}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] \leq \lambda_t + \frac{\psi}{2} \xi_t \left(1 + \frac{1}{R} \right) \text{ for } t \in \mathcal{T}_1^{T-1} \tag{6.20}$$

$$\mathbb{E}^{\mathbb{P}} [\xi_{t+1} X_{t+1}^e | \mathcal{F}_t] \leq \eta_i^{-1} \xi_t X_t^f \text{ for all } i \in \{1, \dots, I\} \text{ and } t \in \mathcal{T}_0 \tag{6.21}$$

$$\mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] = \xi_t X_t^f \text{ for } t \in \mathcal{T}_1 \text{ and } \mathbb{E}^{\mathbb{P}} [\lambda_1] = X_0^f. \tag{6.22}$$

Here (6.12) is used at several points to simplify the expressions. This system already fulfills properties A1, A2, A3 and A5 in Proposition (2.3). Applying (6.12) recursively and keeping in mind $R > 0$, it follows that $\xi_t > 0$ (which is A2). Recall that $X_t^f \geq 0$ was assumed, hence by (6.15) there follows $\mathbb{E}^{\mathbb{P}} [\lambda_{t+1} | \mathcal{F}_t] \geq 0$. Using this fact, it is possible to replace the constraints (6.14) with

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

This ensures A4. Moreover, plugging the second equation of (6.22) into (6.18) shows that (6.18) is redundant (recall that ψ and R both are positive).

Finally, A6 is obtained by (6.17) and by plugging A5 into (6.20). \square

References

- Alonso-Ayuso A, Escudero LF, Pizarro-Romero C (2009) Introduction to stochastic programming. Ciencias Experimentales Y Tecnología. Universidad Rey Juan Carlos
- Back KE (2010) Asset pricing and portfolio choice theory. Oxford University Press, Oxford
- Björk T (2009) Arbitrage theory in continuous time, 3rd edn. Oxford University Press, Oxford
- Cochrane J (2005) Asset pricing, revised edition. Princeton University Press, Princeton
- Deng S-J, Johnson B, Sogomonian A (2001) Exotic electricity options and the valuation of electricity generation and transmission assets. *Decis Support Syst* 30:383–392
- Duffie D (2001) Dynamic asset pricing theory. Princeton University Press, Princeton
- Dupacova J, Gröwe-Kuska N, Römisch W (2003) Scenario reduction in stochastic programming: an approach using probability metrics. *Math Program Ser A* 95:493–511
- Flåm S D (2008) Option pricing by mathematical programming. *Optim J Math Program Oper Res* 57(1):165–182
- Geyer A, Hanke M, Weissensteiner A (2010) No-arbitrage conditions, scenario trees, and multi-asset financial optimization. *Eur J Oper Res* 206:609–613
- Heitsch H, Römisch W (2010) Stochastic programming—the state of the art, chapter stability and scenario trees for multistage stochastic programs. Springer, Berlin, pp 139–164
- King AJ (2002) Duality and martingales: a stochastic programming perspective on contingent claims. *Math Program Ser B* 91:543–562
- Kovacevic RM (2018) Valuation and pricing of electricity delivery contracts—the producer’s view. *Ann Oper Res*. <https://doi.org/10.1007/s10479-018-3010-0>
- Kovacevic RM, Pflug G (2013) Pricing of energy contracts—from replication pricing to swing options. In: Kovacevic RM, Pflug GC, Vespucci MT (eds) *Handbook of risk management for energy production and trading*. International series in operations research & management science. Springer, pp 387–411
- Nowotarski J, Weron R (2018) Recent advances in electricity price forecasting: a review of probabilistic forecasting. *Renew Sustain Energy Rev* 81(1):1548–1568
- Pennanen T (2011a) Dual representation of superhedging costs in illiquid markets. *Math Financ Econ* 5(4):233–248
- Pennanen T (2011b) Superhedging in illiquid markets. *Math Finance* 21(3):519–540
- Pennanen T (2012) Introduction to convex optimization in financial markets. *Math Program Ser B* 134(1):157–186
- Pflug GC, Pichler A (2014) Multistage stochastic optimization. Springer series in operations research and financial engineering. Springer, Berlin
- Pflug GC, Pichler A (2015) Dynamic generation of scenario trees. *Comput Optim Appl* 62(3):641–668

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.