UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

PhD Thesis

Rune Ramsdal Ernstsen

Operation, Investment and Hedging in Electricity Markets

Supervisor: Trine Krogh Boomsma

Submitted: August 31. 2016

Operation, Investment and Hedging in Electricity Markets

Rune Ramsdal Ernstsen

PhD Thesis

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen

August 31, 2016

Department of Mathematical Sciences
Faculty of Science
University of Copenhagen

Author	**Rune Ramsdal Ernstsen** Department of Mathematical Sciences University of Copenhagen Universitetsparken 5 2100 København Ø Denmark runneren14@hotmail.com
Supervisor:	Associate Professor Trine Krogh Boomsma, University of Copenhagen
Co-supervisor:	Associate Professor Jesper Lund Pedersen, University of Copenhagen
Assessment committee:	Professor Stein-Erik Fleten, Norwegian University of Science and Technology
	Professor Kristian Miltersen, Copenhagen Business School
	Professor Mogens Steffensen (chairman), University of Copenhagen

Preface

This thesis has been prepared in fulfillment of the requirements for the PhD degree at the Department of Mathematical Sciences (MATH), Faculty of Science, University of Copenhagen. The work has been carried out under the supervision of associate professor Trine Krogh Boomsma from MATH in the period from September 1, 2013 to August 31, 2016.

The main body of the thesis consists of an introduction to the overall work and four chapters on different but related topics. Each chapter is written as an individual academic paper and are thus self-contained and can be read independently.

Acknowledgments

I would like to thank my supervisor Trine Krogh Boomsma for the numerous meetings, support, ideas and discussions during the work. I would also like to thank Trine, Nihat Misir, Martin Jönsson and Anders Skajaa for co-authoring different parts of the thesis.

Furthermore, I would like to thank Afzal Siddiqui and Trine for setting up my visit to the Statistics Department at University College London.

Additionally, I would like to thank my colleagues at IE and MATH as well as my office mates (Ninna, Maj-Britt, Kamille, Sima, Henrik and André) for all the good talks at the office and at lunch. A special thank to Anders Tolver, Rolf Poulsen, Michael Sørensen and Jesper Lund Petersen for the open door and the possibility to come in and discuss my problems.

I would like to thank my family and friends for inspiring me and supporting me. Finally, I would like to thank my girlfriend Adela Lutvikadic for the wonderful love and support and for reminding me that mathematics is not everything.

Rune Ramsdal Ernstsen
Frederiksberg, August 2016

Contents

List of Papers

This thesis is based on four papers:

- Rune Ramsdal Ernstsen, Trine Krogh Boomsma, Martin Jönsson and Anders Skajaa (2015),
  Hedging local volume risk using forward markets: Nordic Case.
  Submitted for publication.

- Rune Ramsdal Ernstsen and Nihat Misir (2016),
  Market Power and Investment in Electricity Generation
  Submitted for publication.

- Rune Ramsdal Ernstsen and Trine Krogh Boomsma (2016),
  Valuation of power plants.

- Rune Ramsdal Ernstsen (2016),
  An EM algorithm with two jump components.

Summary

This thesis consists of an introduction as well as four papers. The papers concern different problems associated to future electricity markets and the topics include risk management, investment strategies, valuation and model calibration. Each paper is presented in a separate chapter and hence the chapters are self-contained and may be read individually. A more thorough overview is presented in Chapter 1.

In Chapter 2 we consider a hedging problem for a power distributor delivering electricity on fixed price contracts in the Nordic electricity market and thereby being exposed to volume risk. We develop time series models for the electric load, system price and deviation from system price. The model is designed such that for independent electric load, system price and deviations from system price, the minimal variance hedge coincide with the standard practice of the industry. We extend the model to include price and load correlation which results in an explicit strategy that reduces the variance. To further improve the strategy we include autocorrelation and solve the hedging problem numerically and show that there is a large potential in changing risk measure and utilizing the skewness in the payoff distribution.

In Chapter 3 we consider an investment problem for a strategic investor and a social planner with the opportunity to invest in inflexible and flexible generation. We study the impact of market power and conjectured market changes with a simple price model based on linear demand response. We show that the strategic investor invests later and in less capacity than the socially optimal and that with increased market ownership investment is delayed further and capacity increased slightly. Furthermore, we find that an increase in market share for the strategic investor delays inflexible generation more than flexible generation due to the exposure to potential low prices.

In Chapter 4 we study the valuation of three representative generation types, an inflexible wind turbine, a flexible gas fired power plant and a hydroelectric plant that allows for storage. We account for the special characteristics of each technology and include uncertainty in both price and volume through diffusion or jump diffusion models. We find explicit expressions for the expected instantaneous value of wind generation as a function of electricity price and wind speed. We include startup and shutdown costs for the gas fired power plant determine the startup and shutdown triggers as well as the value of the plant by maximizing the value of shutting down. This is done analytically in the diffusion models and numerically in the jump diffusion model. For the hydroelectric power plant we relax storage level and discharge constraints using penalty functions and linearize the optimal strategy from the Hamilton-Jacobi-Bellman equation. This allows for closed form expressions of

the value in terms of the expected price, the second moment of the price and the autovariance of the price. We calibrate the models to 7 years of hourly price and wind data, determine the value and study the impact of anticipated market changes on the value of the three types of generation.

In Chapter 5 we develop an EM-algorithm with two jump components such that the jump density of the compound Poisson process is a mixture of two normal distributions. We show that each step of the EM-algorithm increases the log-likelihood of the observed data by maximizing the expectation of the log-likelihood for the complete data conditional on the observed data. We determine explicit expressions for the maximization step in terms of simple conditional expectations and present an approach for determining the conditional expectations. Finally, by applying the algorithm to calibrate the jump diffusion model from Chapter 4, we demonstrate that the additional jump component provides a significantly better model of the observed data than a model without jumps and with only a single jump component.

Sammenfatning på dansk

Denne afhandling består af et introducerende kapitel og fire artikler. Artiklerne omhandler forskellige problemer i forbindelse med fremtidige elektricitetsmarkeder og inkluderer emner som risikostyring, investeringsstrategier, værdifastsættelse og modelkalibrering. Hver artikel er præsenteret i et kapitel og kapitlerne kan derfor læses enkeltstående. En mere dybdegående gennemgang af afhandlingen gives i kapitel 1.

I kapitel 2 betragter vi et risikoafdækningsproblem for en elleverandør, som leverer elektricitet på fastpriskontrakter i det nordiske elektricitetsmarked og derfor er udsat for mængderisiko. Vi udvikler en tidsrækkemodel for det samlede forbrug, systemprisen og afvigelser fra systemprisen. Modellen er designet så uafhængigt samlet forbrug, systempris og afvigelse fra systempris medfører at afdækningsstrategien med minimal varians stemmer overens med almindelig praksis i industrien. Vi udvider modellen til at inkludere korrelation mellem pris og forbrug, hvilket resulterer i en eksplicit strategi som reducerer variansen. For yderligere at forbedre strategien inkluderer vi autokorrelation. Vi løser risikostyringsproblemet numerisk og viser at der er et stort potentiale i at skifte risikomål og udnytte skævheden i fordelingen af afkastet.

I kapitel 3 betragter vi et investeringsproblem fra et investorperspektiv og et samfundsmæssigt perspektiv, hvor der kan investeres i ufleksibel eller fleksibel elproduktion. Vi undersøger påvirkningen af markedsmagt og formodede markedsændringer med en lineær invers efterspørgselsfunktion. Vi viser at investoren investerer senere og i mindre projekter end hvad der er samfundsmæssigt optimalt og at en øget markedsandel yderligere forsinker investeringerne og forstørrer projekterne en smule. Udover dette forsinker en øget markedsandel investeringen i ufleksibel elproduktion mere end fleksibel elproduktion, da ufleksibel produktion er mere udsat for prisrisiko.

I kapitel 4 studerer vi værdifastsætning af tre repræsentative elproduktionsteknologier, en ufleksibel vindturbine, et fleksibelt gaskraftværk og et vandkraftværk der tillader lagring. Vi tager højde for specielle karakteristika for teknologierne og inkluderer usikkerhed i både pris og volumen gennem diffusionsmodeller med eller uden spring. Vi finder eksplicitte udtryk for den forventede værdi af produktion med vindenergi som funktion af elpris og vindhastighed. Vi inkluderer opstarts- og nedlukningsomkostninger for gaskraftværket og bestemmer de optimale opstarts- og nedlukningsniveauer, samt værdien af gaskraftværket. Dette er gjort analytisk for diffusionsmodellerne og numerisk med inklusion af spring. For vandkraftværket relakserer vi begrænsningerne for lagerniveauet og vandføring ved hjælp af straffunktioner og lineariserer den optimale strategi fra Hamilton-Jacobi-Bellman ligningen. Dette medfører lukkede udtryk for værdien som funktion af den forventede

pris, andet-momentet af prisen og autovariansen af prisen. Vi kalibrerer modellerne til pris og vind data for hver time over 7 år og for de tre slags el produktion bestemmer vi værdien og påvirkningen af forventede markedsændringer.

I kapitel 5 udvikler vi en EM-algoritme med to springkomponenter så tætheden for den sammensatte Poisson proces er en blanding af to normalfordelinger. Vi viser at et skridt med EM-algoritmen øger værdien af log-likelihood-funktionen for det observerede data ved at maksimere den betingede middelværdi af log-likelihood-funktionen for det komplette datasæt, hvor der betinges med det observerede data. Vi anvender algoritmen til at kalibrere diffusionsmodellen med spring fra kapitel 4 og viser at det observerede data modelleres markant bedre end med en model uden spring eller med en enkelt springkomponent.

1

Introduction

This chapter provides an overview of the following four chapters of the thesis that covers operation, investment and hedging in electricity markets as the title suggests. For each of the chapters we give a short introduction to the accompanying theory and comment on the main results. Initially, we introduce the current electricity market as well as some of the current and future challenges. In Chapter 2, we focus on risk management and in Chapters 3 and 4 we study valuation of generation as well as investment decisions. Chapter 3 focuses on the impact of market power, while Chapter 4 focuses on valuation of new generation and the impact of market changes. Finally, Chapter 5 develops an EM-algorithm for the jump-diffusion model for electricity prices from Chapter 4.

1.1 Electricity Markets

The current electricity market within the EU is a result of a liberalization in the late 90's. The liberalization was initiated to increase competition and create incentive for investment in both generation and electricity infrastructure. Furthermore, import/export barriers were removed which reduced the required reserve capacity and lowered generation costs as the combined supply provided additional flexibility as well as better utilization of technologies with low marginal cost of production.

However, due to decarbonisation goals additional renewable generation has to be deployed, which will impact both the power system and the electricity market significantly.

> *[International Energy Agency] analysis indicates that the large-scale deployment of renewables needed to meet decarbonisation goals is technically feasible. However, the inherent variability of these power sources will lead to less predicable power flows. Greater flexibility of power systems will therefore be needed if large-scale deployment of variable renewable generation, such as solar photovoltaic, wind and tidal energy, is to go ahead without jeopardising electricity security.*[1]

Thus, to ensure greater flexibility of the power system new investments have to be undertaken to incorporate the renewable generation. These investments have to be accompanied by advanced risk management and accurate valuation models, which are both highly dependent on the dynamics of electricity prices in a future market.

Despite the liberalization, the electricity market still has to be regulated to ensure electricity security and avoid blackouts. In the Nordic region this is organized with a day-ahead market, an intraday market as well as a balancing market. This construction allows for sequential planning of generation minimizing real-time adjustment, while taking into account the uncertainty in demand as well as the fact that electricity cannot be stored efficiently.

The day-ahead market, Elspot, has hourly prices and closes at 12 noon for delivery the following day. The day-ahead prices are based on an auction system, determined by matching supply and demand, potentially differentiated locally such that grid restrictions are satisfied.

The intraday market, Elbas, closes one hour before delivery and accounts for changes in the 12 to 36 hours between market closure and delivery in the day-ahead market. Finally the balancing market handles real-time imbalances and ensures that voltages are kept in the desired range by adjusting generation.

However, even when focusing on only the day-ahead market, the market structure necessary to ensure electricity security leads to complicated risk management problems. In Chapter 2 we therefore study the risk management problem for a power distributor supplying electricity on a fixed price contract.

1.2 Risk Management

In classical risk management, it is well known that price uncertainty from future sale of production can be eliminated completely by buying forward contract and using the production to cover the contractual obligations. The advantage of hedging in this way is that the value of the production is fixed and potential losses are avoided. However, if the production volume is uncertain, this is no longer possible, which leads to volume risk that cannot

[1] `https://www.iea.org/topics/electricity/`, accessed August 1. 2016

be eliminated with forward contracts. This is a big issue for an electricity distributor that typically delivers electricity on fixed price contracts. The distributor essentially promises to cover an uncertain electric load at a fixed price per MWh by purchasing the required load in the day-ahead market, and is therefore exposed to volume risk.

Additionally, the forward contracts available cover entire months or peak periods and are settled against the system price, a day-ahead price determined without adjusting for congestion. This mismatch between the hedging instrument and the asset to be hedged is called basis risk and further complicates the problem. To reduce exposure to differences between the day-ahead price and the system price, forward contracts on the price difference are included in the model.

In the industry, the current hedging approach is to enter forward contracts matching the expected production. However, papers on volume risk, such as McKinnon (1967) and Oum et al. (2006), show that without basis risk the variance minimal hedge must compensate for the correlation between price and production. With a single price and a single time period the payoff from a fixed price contract when buying forward contracts for a volume of V is

$$(F - S_T)L_T + (S_T - q_t(T))V, \tag{1.2.1}$$

where F is the fixed electricity price, S_T is the price of purchasing the electricity at time T, L_T is the uncertain load delivered at time T and $q_t(T)$ is the contract price determined at time $t < T$. The minimal variance hedge for this simple problem is

$$\begin{aligned} V = E(L_T) - (F - E(S_T))\frac{Cov(S_T, L_T)}{Var(S_T)} \\ + \frac{Cov((S_T - E(S_T))^2, L_T)}{Var(S_T)}. \end{aligned} \tag{1.2.2}$$

Thus, the hedge should be adjusted based on expected profit per MWh, $F - E(S_T)$, as well as the covariance between price and load. This has to be further adjusted by the covariance between the standardized price variance and the load, which is 0 if (S_T, L_T) are simultaneously normal distributed, regardless of their correlation. For positive correlation between price and load, which is the case in the electricity market, this suggests hedging below expected load in off-peak periods with low prices, and above expected load in peak periods with high prices.

In Chapter 2, we develop time series models for the relevant prices and load based on deviations from seasonal components. For the initial model, variance minimization coincide with industry practice of hedging the expected load, while the first addition to the model follows the principles of the example above. The model is further extended to include more advanced correlation structures, including autocorrelation for prices and load. The models are

calibrated to hourly data for West Denmark and East Denmark from 2012. We benchmark on data from 2013 and 2014 and show that compensating for the correlation reduces the realized variance. Furthermore, we find that the risk premium for the contracts on the price difference is significant in East Denmark, as the gross loss is less affected than the gross profit. By changing risk measure and minimizing the expected loss, we obtain a lower gross loss and a higher gross profit. The reason is that the unhedged payoff density is not symmetric and the skewness changes depending on the hedge. At the same time, the change of risk measure creates a position that is less exposed to large losses compared to the variance minimizing strategy.

The models in Chapter 2 rely on a seasonality component to adjust for current trends and tendencies. The inclusion of seasonality, however, significantly complicates the modeling of long term prices. Thus, to allow for analytically tractable models in the following chapters, we ignore seasonality. Furthermore, the price models in Chapter 2 are essentially discrete time models, which is required to determine the hedging strategies. These are, however, intractable for long term valuation problems and investment decisions.

1.3 Valuation of Generation

In Chapters 3 to 5 we study valuation and investment. To avoid the curse of dimensionality in discrete time models and allow for closed-form solutions, these chapters use continuous time models. Common to the price models in Chapters 3 to 5 is that the price process, $(P_t)_{t\geq 0}$, is assumed to solve a stochastic differential equation of the form

$$\mathrm{d}P_t = \mu(P_t)\ \mathrm{d}t + \sigma(P_t)\ \mathrm{d}Z_t + \gamma(P_t)\ \mathrm{d}J_t, \tag{1.3.1}$$

where $\mu(P)$, $\sigma(P)$ and $\gamma(P)$ are drift, diffusion and jump coefficients respectively and $(Z_t)_{t\geq 0}$ is a Brownian Motion. Furthermore, $(J_t)_{t\geq 0}$, is a compound Poisson process with $J_t = \sum_{n=1}^{N_t} Y_n$, where $(N_t)_{t\geq 0}$ is a Poisson process with intensity λ and Y_n are i.i.d. random variables. In Dixit and Pindyck (1994) valuation of an investment is based on expected discounted value, such that with instantaneous profit $\pi(P)$ and discount factor r, an estimate for the value of a generating unit over an infinite horizon is

$$V(P) = \mathbb{E}\left(\int_0^\infty e^{-rt}\pi(P_t)\ \mathrm{d}t \middle| P_0 = P\right). \tag{1.3.2}$$

In Chapter 3 we use this valuation approach to study the impact of market power in an electricity market context and in Chapter 4 we extend the valuation method to include operational characteristics. One advantage of this formulation is that $V(P)$ solves the Hamilton-Jacobi-Bellman (HJB) integro-

differential equation,

$$\begin{aligned}&\mu(P)\frac{\partial}{\partial P}V(P)+\frac{1}{2}\sigma(P)^2\frac{\partial^2}{\partial P^2}V(P)+\pi(P)-rV(P)\\&+\lambda\left[\mathbb{E}(V(P+\gamma(P)Y_1))-V(P)\right]=0,\end{aligned}\tag{1.3.3}$$

which in the case without jumps, simplifies to

$$\mu(P)\frac{\partial}{\partial P}V(P)+\frac{1}{2}\sigma(P)^2\frac{\partial^2}{\partial P^2}V(P)+\pi(P)-rV(P)=0.\tag{1.3.4}$$

As opposed to (1.3.3), (1.3.4) can often be solved analytically. The solution consists of a particular solution as well as two linearly independent solutions to the homogenous version of (1.3.4). This provides an alternative to applying Fubini's theorem to (1.3.2) and computing the expectation directly.

1.4 Real Options

The opportunity to undertake investment when prices are sufficiently high can be modeled as an option with payoff $V(P)-I$, where I is the investment cost. This is the reason for the term real options. Assuming that the investment is initiated the first time prices exceed some P^*, i.e. at the hitting time τ defined by

$$\tau=\inf\{t\geq 0|P_t\geq P^*\},\tag{1.4.1}$$

the value of the investment opportunity can be expressed as

$$\mathbb{E}\bigg(\int_\tau^\infty e^{-rt}\pi(P_t)\,\mathrm{d}t-e^{-r\tau}I\,\bigg|\,P_0=P\bigg)=\quad\mathbb{E}\Big(e^{-r\tau}\left(V(P_\tau)-I\right)\Big|P_0=P\Big).\tag{1.4.2}$$

In the case without jumps, the value of the investment opportunity simplifies to

$$\mathbb{E}(e^{-r\tau}|P_0=P)(V(P^*)-I).\tag{1.4.3}$$

The investors decision is to choose P^* such that the investment opportunity, (1.4.2), is maximized. Naturally, choosing a large P^* delays investment and increases the value when investing. However, the stochastic discount factor, $\mathbb{E}(e^{-r\tau}|P_0=P)$, decreases as P^* increases potentially lowering the value of the option to invest. Similar to $V(P)$, the expected discount factor can be found as a solution to (1.3.4), only with $\pi(P)=0$ and conditions on the boundary. With this approach, the value and timing of the investment opportunity can be studied.

1.4.1 Investment Strategies

In Chapter 3, the aim is to study the timing and sizing of investment in new electricity generation capacity, under the assumption that the price is determined from a simple demand response relation where the demand shock follows a Geometric Brownian Motion. We compare constant generation and generation with an option to temporarily suspend, representing renewable generation that is inflexible and conventional generation that is flexible. By using a real options approach, we extend the work of Huisman and Kort (2015). With a particular focus on market power, we introduce an already installed capacity in the market and ownership of this capacity. We derive necessary conditions for optimal investment timing and capacity in terms of capacity, reducing the dimension of the problem. We determine sufficient conditions for the plant and option values to be finite in terms of the cost function and the underlying process. Furthermore, we derive the smooth pasting condition as a limit of two value matching conditions, which leads to simple and explicit expressions for the value of flexible generation.

We confirm that increased price volatility delays investment, but increases capacity at the time of investment. In general, the investment trigger and capacity is highly dependent on the tradeoff between the value of waiting, determined by the stochastic discount factor, and the marginal cost of new capacity. Thus, for low marginal cost of capacity, large investments are initiated early, whereas higher marginal costs of capacity naturally delays investment and reduces capacity. As the value of waiting increases, which coincide with an increase in price volatility, investments are delayed and new capacity increased.

For market power in particular, our results show that the strategic planner invests later than what is socially optimal and in smaller generation capacity. Moreover, increased installed capacity or ownership of installed capacity for the strategic investor delays investment, as the investment reduces the profit of existing assets.

In Chapter 4 we assume an exogenous price and focus on more advanced modeling of the price dynamics and include operational characteristics in the valuation for different types of generation.

1.5 Stochastic Optimal Control

The valuation problem can be extended to include a control strategy as

$$V(P) = \sup_{(v_t)_{t\geq 0}} \mathbb{E}\left(\int_0^\infty e^{-rt} \pi(v_t, P_t)\ \mathrm{d}t \middle| P_0 = P \right), \tag{1.5.1}$$

where the supremum is taken over all Markov control functions. For this problem the HJB equation becomes,

$$\begin{aligned}\sup_v \Big(&\mu(P)\frac{\partial}{\partial P}V(P) + \frac{1}{2}\sigma(P)^2\frac{\partial^2}{\partial P^2}V(P) + \pi(v,P) - rV(P) \\ &+\lambda\left[\mathbb{E}(V(P+\gamma(P)Y_1)) - V(P)\right]\Big) = 0,\end{aligned} \tag{1.5.2}$$

and if the supremum exists, we can substitute optimal control, $v^*(P)$, in (1.5.1) such that

$$V(P) = \mathbb{E}\left(\int_0^\infty e^{-rt}\pi(v^*(P_t), P_t)\ \mathrm{d}t \middle| P_0 = P\right) \tag{1.5.3}$$

which is of the same form as (1.3.2).

1.6 Including Operation in Valuation

In Chapter 4, we consider three stylized generation technologies, renewable generation, conventional units and storage units. The renewable generation, which is exemplified by a wind turbine, is inflexible. Hence, production is negatively correlated with prices and the value is therefore overestimated if correlation is ignored. The conventional unit, exemplified by a gas-fired power plant, is flexible and can be suspended for sufficiently low prices. However, it is typically quite expensive for gas-fired power plants to suspend and start up generation and for this reason the gas-fired power plant is modeled similar to Chapter 3, but including startup and shutdown costs. Finally, the hydro-electric power plant can store electricity and benefit from price variations by altering the discharge (or pump) rates, but has to maintain the storage levels and discharge rates within certain operational and environmental limits.

We model the value of wind generation using the power curve, $H(W)$, that determines the power output for a given wind speed W. The value of wind generation becomes

$$V_{wind}(P,W) = \mathbb{E}\left(\int_0^\infty e^{-rt}H(W_t)P_t\ \mathrm{d}t \middle| P_0 = P, W_0 = W\right), \tag{1.6.1}$$

where the wind speed at time t, W_t, is modeled by a transformation of an Ornstein-Uhlenbeck process to capture an appropriate marginal distribution and an exponentially decaying autocorrelation.

When studying flexibility, we utilize that for the stopping time

$$\tau = \inf\{t \geq 0 | P_t \geq P^*\}, \tag{1.6.2}$$

the value of generating until the price hits P^*,

$$\mathbb{E}\left(\int_0^\tau e^{-rt}\pi(P_t)\ \mathrm{d}t \middle| P_0 = P\right), \tag{1.6.3}$$

is equal to the value of always generating less the discounted value of always generating from time τ,

$$\mathbb{E}\left(\int_0^\infty e^{-rt}\pi(P_t)\ \mathrm{d}t\middle| P_0 = P\right) - \mathbb{E}\left(\int_\tau^\infty e^{-rt}\pi(P_t)\ \mathrm{d}t\middle| P_0 = P\right). \quad (1.6.4)$$

Thus, the value can be decomposed in the value of always generating and an option to suspend. This decomposition creates two simpler problems, as the option value satisfies the HJB equation with a condition on the value for $P \geq P^*$, or in the absence of jumps $P = P^*$.

With startup and shutdown costs, the above generalizes to the startup and shutdown triggers P^*_{on} and P^*_{off} with corresponding shutdown and startup times S_{off} and S_{on}, where $P^*_{on} > P^*_{off}$. The value of the gas-fired plant is modeled as

$$\begin{aligned} V^{on}_{gas}(P) = \mathbb{E}&\left(\int_0^\infty e^{-rt}\pi_{gas}(P_t)\ \mathrm{d}t\middle| P_0 = P\right) \\ &+ \mathbb{E}\left(e^{-rS_{off}}|P_0 = P\right) c_1(P^*_{on}, P^*_{off}), \quad \text{for } P \geq P^*_{off} \end{aligned} \quad (1.6.5)$$

and

$$V^{off}_{gas}(P) = \mathbb{E}\left(e^{-rS_{on}}\middle| P_0 = P\right) c_2(P^*_{on}, P^*_{off}), \quad \text{for } P \leq P^*_{on} \quad (1.6.6)$$

where $\pi_{gas}(P_t)$ is the instantaneous profit and $c_1(P^*_{on}, P^*_{off})$ and $c_2(P^*_{on}, P^*_{off})$ are the option values of suspending and starting up generation that include startup and shutdown costs. In Chapter 4 we show that $c_1(P^*_{on}, P^*_{off})$ and $c_2(P^*_{on}, P^*_{off})$ can be expressed explicitly in terms of the startup and shutdown costs, the value of always generating and the stochastic discount factors. The value of the gas-fired power plant and the triggers are determined by maximizing the option constants numerically with and without jumps, where the stochastic discount factors are determined by solving a differential equation or a differential integro-equation in the presence of jumps.

The analogous model for a hydroelectric power plant has no startup and shutdown costs, but includes the storage level $L \in [L_{min}, L_{max}]$. The optimal strategy is a bang-bang strategy and depends on some barrier $P^*(L)$, such that water is discharged at the maximal rate for $P \geq P^*(L)$ and discharged at the minimal rate, possibly resulting in pumping, for $P \leq P^*(L)$. The value for $P \geq P^*(L)$ is

$$\begin{aligned} V^{discharge}_{hydro}(P, L) = \mathbb{E}&\left(\int_0^\infty e^{-rt}\pi^d_{hydro}(P_t, L_t)\ \mathrm{d}t\middle| P_0 = P, L_0 = L\right) \\ &+ \mathbb{E}\left(e^{-rS_p} d_1\left(\{P^*(\tilde{L})\}_{\tilde{L}\in[L_{min}, L_{max}]}\right)\middle| P_0 = P, L_0 = L\right) \end{aligned} \quad (1.6.7)$$

and for $P \leq P^*(L)$ it is

$$V^{pump}_{hydro}(P,L) = \mathbb{E}\left(\int_0^\infty e^{-rt}\pi^p_{hydro}(P_t,L_t)\,\mathrm{d}t \middle| P_0 = P, L_0 = L\right) + \mathbb{E}\left(e^{-rS_d}d_2\left(\{P^*(\tilde{L})\}_{\tilde{L}\in[L_{min},L_{max}]}\right)\middle| P_0 = P, L_0 = L\right). \tag{1.6.8}$$

Here $\pi^d_{hydro}(P,L)$ is the instantaneous profit from discharging and $\pi^p_{hydro}(P,L)$ is the instantaneous profit from pumping. S_p and S_d are the hitting times for the barrier between pumping and discharging and d_1 and d_2 are the option constants that depend on the barrier

$$\{P^*(\tilde{L})\}_{\tilde{L}\in[L_{min},L_{max}]}. \tag{1.6.9}$$

The purpose of the barrier is twofold, it limits pumping and discharging for high and low storage levels respectively and defines the price threshold between pumping and discharging that depends on the storage level, see Figure 1.1.

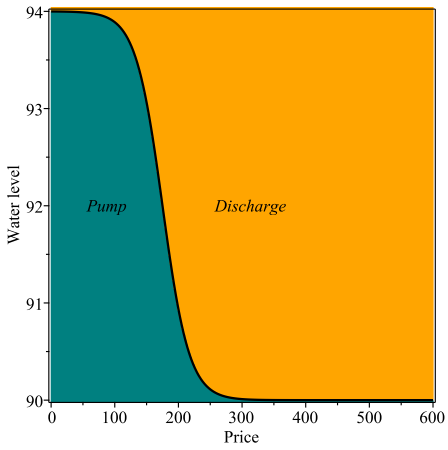


Figure 1.1: Example of barrier for a hydroelectric power plant.

For the hydroelectric power plant the storage level introduces an additional dimension in the HJB equation, which leads to value matching and smooth pasting at an unknown curve and thereby creates an intractable problem. The classical approach to overcoming this is to solve the problem on a finite time horizon and applying some finite difference scheme, see Chen and Forsyth (2008a). However, as the problem has two spacial dimensions, the computational time increases rapidly with additional spacial accuracy, limiting the solution approach to only short horizons.

As an alternative for longer time horizons, we relax the storage level and discharge rate restrictions to penalty functions, which eliminates the need for value matching and smooth pasting. Furthermore, to allow for analytical solutions, we linearize the optimal control in price and storage level. Thus, using the linearized version of the optimal control from

$$\sup_{(v_t)_{t\geq 0}} \mathbb{E}\left(\int_0^\infty \pi_{hydro}^{control}(v_t, L_t, P_t) + \theta_1 v_t + \theta_2 v_t^2 + \Theta_1 L_t + \Theta_2 L_t^2 \middle| P_0 = P, L_0 = L \right), \tag{1.6.10}$$

the model for the value of the hydroelectric power plant becomes

$$V_{hydro}(P, L) = \mathbb{E}\left(\int_0^\infty \pi_{hydro}(P_t, L_t) \middle| P_0 = P, L_0 = L \right). \tag{1.6.11}$$

Our approach makes it possible to determine distributional properties of future storage levels and discharge rates, while mimicking the strategy from Figure 1.1.

We determine the discounted value of generation under three stylized price models, a shifted geometric Brownian motion and a shifted exponential Ornstein-Uhlenbeck process with and without jumps. We calibrate each of the models to real wind and price data and for each type of generation use this to study the exposure to conjectured future market changes.

1.7 Model Calibration

In Chapter 5 we develop an EM-algorithm to calibrate the jump diffusion model from Chapter 4 to data. Classical parameter estimation maximizes the log-likelihood function. For example for T i.i.d. observations, $(y_n)_{n\in\{1,\dots,T\}}$, from a Normal distribution with mean μ, variance σ^2 and density $\phi(y; \mu, \sigma^2)$, the log-likelihood is

$$\log(\prod_{n=1}^{T} \phi(y_n; \mu, \sigma)) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{n=1}^{T}(y_n - \mu)^2 \tag{1.7.1}$$

which can easily be maximized. However, if the observations are of the form

$$Y_n = Z_n + \sum_{k=1}^{N_n} J_{nk}, \tag{1.7.2}$$

where $Z_n \sim \mathcal{N}(\mu, \sigma^2)$, $N_n \sim pois(\lambda)$ and $Y_{nk} \sim \mathcal{N}(\nu, \tau^2)$ for $k = 1, \dots, N_n$ and $n = 1, \dots, T$ with Z_n, Y_{kn} and N_n independent, the density for Y_n is

$$g_n(y) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \phi(y; \mu + k\nu; \sigma^2 + k\tau^2), \tag{1.7.3}$$

for which the log-likelihood no longer simplify to a sum. Nevertheless, the joint density of $(Z_n, N_n, Y_{1n}, \ldots, Y_{N_{nn}})$, which we refer to as the complete data, is

$$f_n(z, N, y_1, \ldots, y_N) = \phi(z; \mu, \sigma^2) \prod_{k=1}^{N} \phi(y_k; \nu, \tau^2), \tag{1.7.4}$$

for which the log-likelihood simplifies to a sum. Unfortunately, the complete data is not observed, hence the EM-algorithm maximizes the conditional expectation of the log-likelihood function,

$$\mathbb{E}\left(\log(\prod_{n=1}^{T} f_n(Z_n, N_n, J_{nk})) \middle| Y_n \right) \tag{1.7.5}$$

assuming that the complete data has some initial distribution. The EM-algorithm iteratively updates the distribution of the complete data and utilize that the complete distribution has a tractable log-likelihood function with explicit maximizers. In Chapter 5 the above approach is generalized to two independent Poisson processes, which is shown to better capture the skewness of the transformed price increments. It is also shown that each iteration of the EM-algorithm increases the log-likelihood function of the observed data and not only the conditional expectation of the log-likelihood of the complete data. Finally, the algorithm is applied to hourly electricity price data, demonstrating a significantly better fit than a jump diffusion with a single jump component.

2

Hedging Volume Risk Using Forward Markets: Nordic Case

Abstract

This paper develops hedging strategies for an electricity distributor in the Nordic electricity market who manages price and volume risk from fixed price agreements on stochastic electricity load. Whereas the distributor trades in the spot market at area prices, the financial contracts used for hedging are settled against the system price. Both the area price and the system price are correlated with electricity load and due to congestion, price differences are also correlated with load. This correlation structure is often disregarded in practice. We, therefore, develop a model for the area price, the system price and the load in the Nordic market with an optimal hedging strategy that coincide with common practice in the industry. This serves as benchmark for an extended model using data from 2013 and 2014 for two bidding areas. In one area the improved hedging strategy reduces gross loss by 5.8% and increases gross profit by 3.8%. In the other area gross loss is reduced by 13.6% and gross profit is increased by 9.5%.

2.1 Introduction

The Nordic electricity market was liberalized in the late 90's to increase competition and create incentive to invest in new electricity production and modernize existing production. This liberalization also reduced the barriers on import and export between countries, allowing for more efficient use of different power producing technologies. Currently, the Nordic market covers the

countries in the Nordic and Baltic region, i.e. Denmark, Norway, Sweden, Finland, Estonia, Latvia and Lithuania. It is divided into 17 bidding areas with individual area prices based on local supply and demand. Furthermore, an overall market price for electricity is determined for contractual purposes. This price is referred to as the system price and is based on aggregated supply and demand, and disregards transmission constraints between bidding areas. In contrast, the bidding areas are established to avoid congestion in the system. Area price and local load are, therefore, highly correlated. This feature can be exploited in deriving hedging strategies that reduce risk and increase expected payoff.

In this paper we study the hedging problem a Nordic distribution company faces when having agreed to deliver electricity to a customer at a fixed price per MWh. The company has to buy electricity in the spot market, but knows neither the exact electricity demand of the customer nor the market price of the electricity when the demand occurs. Since trades in the spot market are settled at the area price two types of risk arise, namely area price risk and volume risk. To mitigate the risk the company locks in part of its profit by buying financial contracts on electricity in advance and at a fixed price. In the Nordic electricity market, however, financial contracts are settled against the system price and not the area price. This introduces significant basis risk as there can be large differences between the system price and the area price, especially in periods with high load. Part of this risk can be covered using forward contracts on the price difference. Nevertheless, as only monthly contracts are available and the load varies throughout the month, the distribution company cannot completely eliminate the risk from fixed price agreements. In spite of this, more than 50 % of contracts for electricity in 2010 were delivered based on fixed price agreements in the Nordic market and in EU 60% of contracts were fixed price agreements.[1] Hence, managing the risk of such agreements is of great importance for electricity companies.

This paper contributes to the literature by developing an electricity price model and deriving a corresponding hedging strategy that takes into account the difference between the area price and the system price, the correlation between price and load as well as correlation over time. In addition to using the base load and peak load contracts for hedging, we study the impact of including contracts for difference to manage basis risk. Furthermore, as the profit distribution is asymmetrical, we complement the traditional variance-based approach by including a one-sided measure of risk in the hedging problem. We benchmark different model extensions against a base model for which the strategy coincides with common practice in the industry, which is to ignore correlation and hedge at the expected load.

The importance of including correlation between electricity price and load has already been demonstrated in the existing literature. As an example, Bessem-

[1] ECME Consortium (2010)

binder and Lemmon (2002) develop an equilibrium-based market model and find that this correlation has a substantial impact on the optimal hedging strategies in a forward market. Closer to our work is Oum et al. (2006), who consider a load serving entity and study the influence of correlation on the residual risk following hedging. The authors derive analytical solutions to the hedging problem for specific utility functions and approximate these solutions by call options to compensate for the lack of contracts to hedge volume risk. Their results likewise show that the correlation has a significant impact on the payoff structure as well as on the hedging strategy. Whereas these references use a single period setting, we include multiple periods and thereby capture the basis risk that arise when contracts cover an entire month. This makes it possible to apply the hedging strategies to the Nordic Market.

An example of a more advanced electricity price model applied to market data is provided by Coulon et al. (2013), who develops a three-factor model with load-based regime switching to model the Texas electricity market. The authors study variations of daily pay-offs using hedging strategies with spark spreads or call options and apply it to a single day with one-dimensional hedging. The inclusion of load-based regime switching makes calibration and estimation much more difficult on a longer time horizon, and, therefore, is not considered in this paper. Weron et al. (2004) and Erlwein et al. (2010) develop advanced reduced form models that involve jumps and regime switching and present different algorithms to calibrate the models to price data. Whereas the above are single factor models, multi-factor models with jumps and regime switching have been used by Deng (1999), Schwartz and Smith (2000) and Coulon et al. (2013) to capture both short-term and long-term dynamics of electricity prices. This approach is extended in Burger et al. (2004) to include a demand component in pricing of derivatives. For a thorough review of electricity price models see Carmona and Coulon (2014), where both reduced form models and structural models are considered. In contrast to these references, our price model is specifically tailored to the Nordic market by including both load, area and system prices, whereas the modeling of each component is restricted to a single factor and does not involve jumps. The inclusion of area and system prices makes it possible to use contracts for difference in the hedging. To the best of our knowledge, the literature has not previously addressed hedging strategies to manage differences between the area and system prices in the Nordic Market.

This paper is organized as follows. The spot and forward markets are introduced in Section 2.2. This includes the dynamics of the system price, the area price and the load as well as the financial contracts used to manage the uncertainty of payoffs. Section 2.3, describes the various sources of risk faced by the company originating from trading in the spot and forward markets and offering fixed price agreements. We formally introduce the accompanying hedging problem in Section 2.4. Section 2.5 analyzes the load and price data, defines the seasonal components and describes the calibration. A simple

model is introduced and subsequently extended in Section 2.6. In Section 2.7, we benchmark the hedging strategies when calibrated to data from 2012 and applied to data from 2013 and 2014. We study the effect of changing the risk measure, the impact of including the contracts for difference and the implications of improved forecast of average prices. Finally, in Section 2.8, we summarize our findings and discuss extensions and future work.

2.2 Trading Electricity

In this section we describe the market dynamics of the Nordic electricity market and the financial instruments that will be used for hedging. We focus on the Nordic spot market, *Nord Pool Spot*, and the corresponding forward market at *Nasdaq Commodities*.

2.2.1 Area Price and System Price

The area prices are set in such a way that the electricity is produced in the least expensive way in the Nordic and Baltic region, aiming at a market equilibrium that accounts for transmission. In the absence of transmission congestion, all area prices coincide with the system price. In its presence, area prices are determined on the basis of the system price by adjusting for transmission. By increasing the area price local supply will increase and local demand will decrease. Similarly, by decreasing the area price, local supply will decrease whereas local demand will increase. Thus, by raising the area price in bidding areas that would ideally be importing beyond its transmission limits, import is reduced. Likewise, by reducing the area price in bidding areas that would be exporting beyond its transmission limits, the export is reduced. It should be clear that electricity is produced at minimal costs, as in equilibrium bidding areas with low marginal cost will be exporting at full transmission capacity and bidding areas with high marginal cost will be importing at full capacity. Since the load on the grid varies significantly throughout the day and increased demand in all bidding areas increase prices, variations in both area price and system price typically occur in periods with high load. Furthermore, the capacity limits on transmission between bidding areas are met more often in hours with high load than hours with low load. For this reason, differences between the area and the system price often occur in periods with high load. In this paper we focus on two large portfolios of fixed price contracts in West Denmark (DK1) and East Denmark (DK2), respectively. The load of the portfolio from DK1 is shown in Figure 2.1. This figure confirms the occurence of price differences in hours of high load. Moreover, the prices reveal that the bidding area DK1 is importing throughout most of August, but is exporting in a few hours of the beginning of February. Other factors, such as changes in demand in other bidding areas and varying supply of wind power, may create differences in periods with low load.

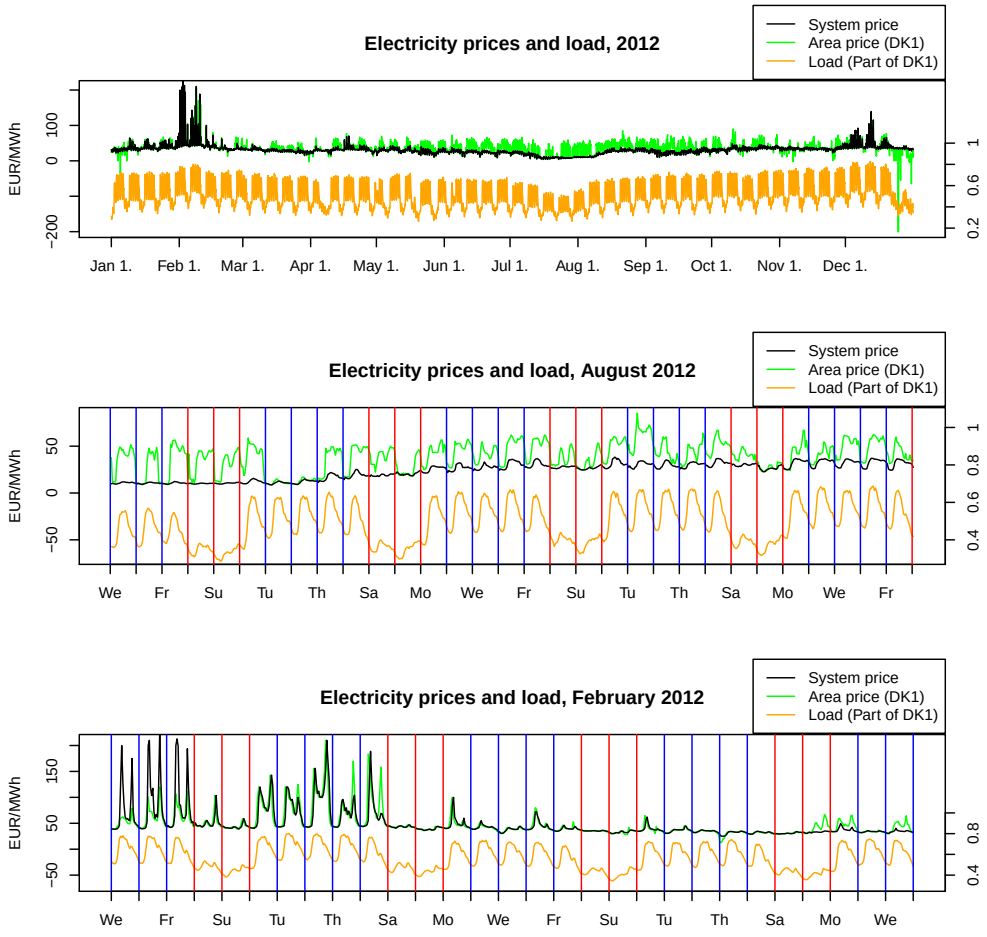


Figure 2.1: Electricity prices and load for 2012 in West Denmark (DK1).

2.2.2 Financial Contracts on Electricity Prices

In the Nordic region, financial contracts on electricity prices are traded on *Nasdaq Commodities*. In this paper, we consider three different types of contracts. The most simple type is a base load contract on the system price that covers every hour of a given month. It is not related to physical delivery of electricity, but is a purely financial contract that pays the difference between the system price and the forward price for every hour of the month. The load typically varies between a peak level and an off-peak level, which can be seen in Figure 2.1. To manage these variations the market also includes peak load contracts that pay the difference between the system price and the forward price in peak hours, 8-20, during weekdays. This makes it possible to create a

portfolio of base load and peak load contracts that resemble the load profile. Figure 2.2 illustrates the load in off-peak and peak periods along with the average load in both periods.

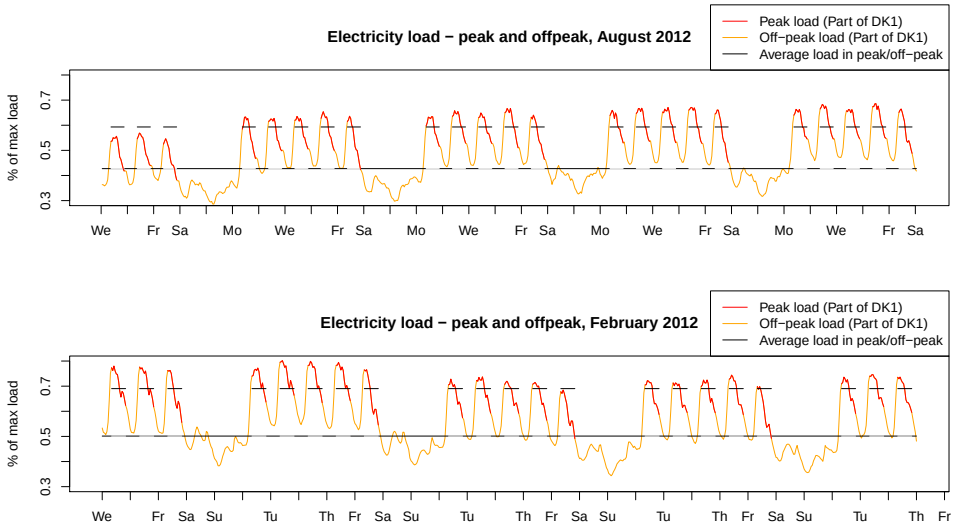


Figure 2.2: Peak and off-peak load for February and August 2012 in Western Denmark (DK1).

Base load and peak load contracts are both settled against the system price and not the area price that is the basis for physical trading. To handle the risk related to differences between area and system prices, we include contracts for difference (CfD). This contract pays the difference between the area price and the system price minus the cost of the CfD and covers the entire month. In spite of including the CfD, it is still not possible to completely eliminate the risk related to delivering an uncertain quantity, i.e. the volume risk.

2.3 Hedging Volume Risk

Initially, we assume that the area price and the system price coincide and study hedging strategies when facing volume risk in a single period setting. When planning to buy a fixed load L_T at an uncertain price S_T at time T and resell it at a fixed price F, risk can be completely eliminated by buying L_T futures contract with maturity T at time t with $t < T$. The contracts pay the difference between the uncertain price S_T and a fixed forward price $q_t(T)$. Thus, at time T we have the pay-off

$$(F - S_T)L_T + (S_T - q_t(T))L_T = (F - q_t(T))L_T. \tag{2.3.1}$$

As a result, the purchase price is locked at $q_t(T)$, eliminating the risk. However, when planning to buy an uncertain load L_T at an uncertain price S_T and reselling it at a fixed price F, it is not possible to completely eliminate the risk using only futures contracts. By buying V futures contracts at time t, the payoff at time T will be

$$(F - S_T)L_T + (S_T - q_t(T))V = (F - q_t(T))V + (F - S_T)(L_T - V). \quad (2.3.2)$$

If we could choose $V = L_T$ the risk would be eliminated as above. The problem is that L_T is stochastic whereas V has to be fixed at time t with $t < T$. We are, therefore, interested in the quality of a hedge, which introduces the need for risk measures. See Artzner et al. (1999) for a detailed analysis of risk measures.

2.3.1 Variance as a Measure of Risk

A classic measure of risk is the variance of the payoff, i.e.

$$Var\Big[(F - S_T)L_T + (S_T - q_t(T))V\Big], \quad (2.3.3)$$

which is minimized by

$$V^* = \frac{Cov(S_T, S_T L_T)}{Var(S_T)} - F\frac{Cov(L_T, S_T)}{Var(S_T)} \quad (2.3.4)$$

as shown in Lemma 2.A.1. We note that V^* is independent of the forward price $q_t(T)$, but not the fixed price F. We can rewrite (2.3.4) to

$$\begin{aligned} V^* = E(L_T) - (F - E(S_T))\frac{Cov(S_T, L_T)}{Var(S_T)} \\ + \frac{Cov((S_T - E(S_T))^2, L_T)}{Var(S_T)}. \end{aligned} \quad (2.3.5)$$

as shown in Lemma 2.A.2. This implies that for any distribution it is optimal to hedge the expected load and compensate for expected unhedged payoff per MWh depending on the covariance between price and load and to compensate for the covariance between the quadratic deviation from the expected price and the load. If S_T and L_T are independent, $V^* = E(L_T)$ and the optimal strategy is to hedge the expected load. This is the straightforward extension of the case with fixed load and we refer to this strategy as the mean hedge.

Example 2.3.1. *Assume S_T and L_T are jointly Normal with correlation ρ and standard deviations σ_S and σ_L, respectively. Then the minimal variance hedge simplifies to*

$$V^* = E(L_T) - (F - E(S_T))\frac{\rho\sigma_Q}{\sigma_S} \quad (2.3.6)$$

which we show in Lemma 2.A.3. As electricity prices are determined by matching supply and demand, the total load is positively correlated with the electricity price. Thus, $V^* < E(L_T)$ for $F > E(S_T)$ and $V^* > E(L_T)$ for $F < E(S_T)$. Finally, if $F - E(S_T) = 0$ or L_T and S_T are uncorrelated, and hence independent, as they are jointly Normal, the optimal strategy is again to hedge the expected load.

The variance measures expected quadratic deviations from the mean and is a symmetrical risk measure. It is useful as it often allows for closed-form minimizers. Moreover, for symmetrical payoff distributions minimizing the two-sided risk is similar to minimizing the one-sided risk. The risk measure is often extended to be a linear combination of the mean and the variance. However, by using the variance, the upside may still be reduced. From a risk management perspective, however, there is no interest in reducing the upside, and so only downside risk is an issue. Because of this, and as the payoffs distributions are not necessarily symmetrical, we consider another classical measure of risk, namely the expected loss.

2.3.2 Expected Loss as a Measure of Risk

The expected loss is defined as

$$-E\Big[\min\Big((F - S_T)L_T + (S_T - q_t(T))V, 0\Big)\Big] \tag{2.3.7}$$

and is equivalent to minus the expected payoff, conditional on the payoff being negative. The sign has been changed to obtain a non-negative quantity. When facing price risk only, i.e. load is fixed, and provided $F > q_t(T)$, both the variance and the expected loss is minimized by $V^* = L_T$ with minimum 0. However, in the presence of volume risk, the two risk measures may result in different hedging strategies as seen in the following example.

Example 2.3.2. *Assume again that S_T and L_T are jointly Normal with $E(S_T) = 35$, $E(L_T) = 0.5$, $\sigma_S = 10$, $\sigma_Q = 0.1$, $\rho = 0.5$ and $q_t(T) = 29.75$. We compare the expected loss minimization (Loss hedge) and variance minimization (Var hedge) and further include the mean hedge for comparison. The strategy minimizing the expected loss is determined numerically. In the first plot of Figure 2.3 the fixed price is 40 and the expected payoff per unit electricity is positive, whereas this is not the case in the second plot where the fixed price is* 30*. In both cases the forward price for electricity is below the expected price, which is known as backwardation, and in this case the expected payoff increases linearly with the hedging volume. We note from Figure 2.3 that the hedged payoff with minimal expected loss has a lighter tail for negative payoffs than the variance hedge in the case with negative expected payoff. It likewise has a heavier tail for positive payoffs. This is due to the fact that the skewness of the payoff density can be affected in the presence of volume risk.*

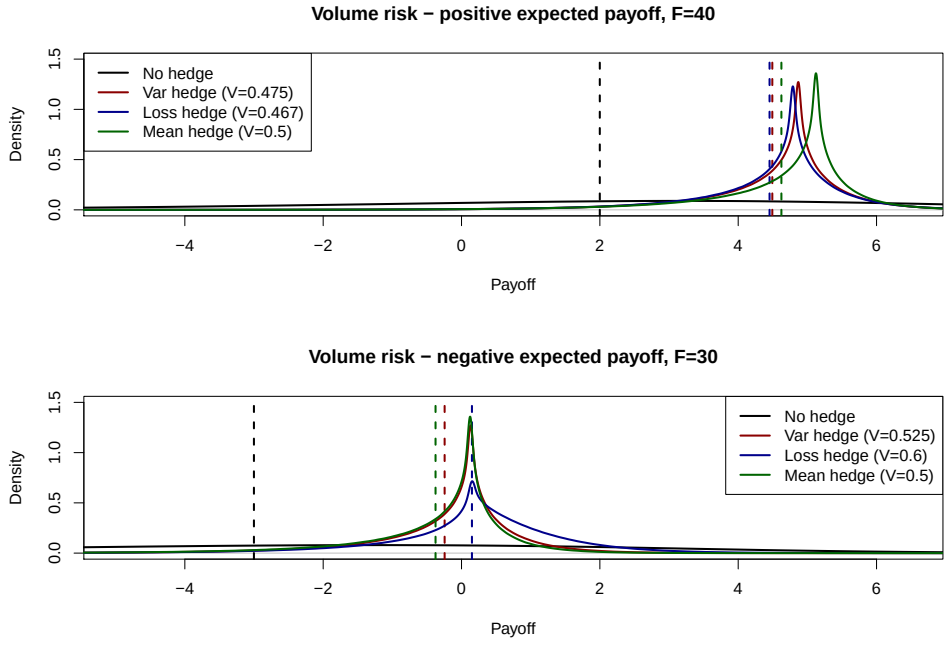


Figure 2.3: Payoff densities (bold lines) and means (dashed lines) with parameters from Example 2.3.2. (Backwardation).

For Normal distributions, the minimal variance hedge is below the mean load in the case of positive expected payoff and above the mean load in the case of negative expected payoff, as also seen from equation (2.3.6). Although this also applies for the expected loss in Example 2.3.2, it is not always the case.

For instance, if the forward price is higher than the expected price, known as contango, the Min Loss hedge deviates significantly from the mean hedge in the opposite direction of the minimal variance hedge, see Appendix 2.B. Appendix 2.B also contains additional statistics for Example 2.3.2.

2.4 Hedging in the Nordic Market

We proceed to introduce the specific problem of hedging in the Nordic market and discuss its relation to the description of volume risk in the previous section. As prices are fixed for every hour, we let S_t and S_t^{sys} denote the area price and the system price, respectively, in hour t measured in EUR/MWh. Moreover, we let L_t denote the percentage of the maximum load delivered to the local customer in hour t. As a result, payoffs are likewise scaled by the maximum load. Letting F_j be the fixed price for electricity in month j, the sales revenue for hour t in month j are

$$(F_j - S_t)L_t. \tag{2.4.1}$$

To mitigate risk we consider three types of contracts, that is, base load contracts, peak load contracts and contracts for difference. We let q_j^b denote the forward price of the base load contract and V_j^b the percentage of maximum load that is covered by base load contracts in month j. For every hour of month j the following cashflow is obtained by buying base load contracts

$$(S_t^{sys} - q_j^b)V_j^b. \tag{2.4.2}$$

Similarly, for the peak load contracts we let q_j^p denote the forward price and V_j^p the percentage of the maximum load that is covered with the peak load contracts in month j. For every hour covered by peak load contracts in month j the following cashflow is obtained

$$(S_t^{sys} - q_j^p)V_j^p. \tag{2.4.3}$$

We let m_j be the set of all hours in month j, $peak_j$ be the subset of m_j that are peak hours, and off_j be the subset of m_j that are off-peak hours. Finally, we let q_j^d denote the forward price for the CfDs and V_j^d denote the percentage of maximum load that is covered by the CfDs. For every hour in month j the following cashflow is obtained by buying CfD contracts

$$(S_t - S_t^{sys} - q_j^d)V_j^d. \tag{2.4.4}$$

Thus, the total cash flow in hour t of month j is given by

$$\begin{aligned}&(F_j - S_t)L_t + (S_t^{sys} - q_j^b)V_j^b + 1_{(t\in peak_j)}(S_t^{sys} - q_j^p)V_j^p \\ &+ (S_t - S_t^{sys} - q_j^d)V_j^d.\end{aligned} \tag{2.4.5}$$

where $1_{(t\in peak_j)}$ is 1 if $t \in peak_j$ and 0 otherwise. By introducing the effective hedging volume in peak hours, $V_j^e = V_j^b + V_j^p$ and the effective forward price in peak hours, $q_j^e = q_j^b V_j^b/V_j^e + q_j^p V_j^p/V_j^e$, we can decompose the payoff such that the cost of hedging in the peak period is a weighted average of the two forward prices. Rewriting the total cash flow from (2.4.5) we obtain:

$$\begin{aligned} &- q_j^d V_j^d + (S_t^{sys} - S_t)(L_t - V_j^d) \\ &+ 1_{(t\in off_j)}\left[(F_j - q_j^b)V_j^b + (F_j - S_t^{sys})(L_t - V_j^b)\right] \\ &+ 1_{(t\in peak_j)}\left[(F_j - q_j^e)V_j^e + (F_j - S_t^{sys})(L_t - V_j^e)\right]. \end{aligned} \tag{2.4.6}$$

This formulation shows how variation in the cash-flow originates from only two random terms for both peak or off-peak hours. The first term is the difference between the area price and the system price times the deviations from the hedging volume of the CfDs. Thus, if $L_t - V_j^d$ is small at a time when the system price and the area price differ, it barely impacts the pay-off. The second random term is the difference between the system price and the fixed price times the deviations from the hedging volume of the base load and peak load contracts. As before, we note that if $L_t - V_j^b$ or $L_t - V_j^e$ is small when the system price deviates from F_j, it barely affect the pay-off. This suggests that to minimize variations it is most important to match the load in periods where the price is volatile. We immediately recognize the payoff structure in the presence of volume risk, although with a sum of two components $S_t^{sys} - S_t$ and $F_j - S_t^{sys}$ times the corresponding differences between the hedging volume and the load. As the system prices in peak hours are typically above the fixed price and the system prices in off-peak hours are typically below the fixed price, the results of Example 2.3.1 suggest hedging above the mean load in peak hours and below the mean load in off-peak hours. Unfortunately, the two terms cannot be handled separately as the system price and the load are included in both.

If we could perfectly predict L_t and adjust V_j^d, V_j^b and V_j^p every hour, price risk could be completely eliminated. This could be done by setting $V_j^d = L_t$ for all hours, $V_j^b = L_t$ for off-peak hours and $V_j^e = L_t$ for peak hours. This would result in the following cashflow

$$(F - q_j^d - 1_{(t\in off_j)}q_j^b - 1_{(t\in peak_j)}q_j^e)L_t. \tag{2.4.7}$$

Thus, q_j^d becomes the cost of hedging the difference between the area price and the system price and $F - q_j^b$ or $F - q_j^e$ becomes the payoff that is locked when hedging. However, L_t is stochastic and varies for each hour, whereas V_j^b, V_j^p and V_j^d have to be fixed for month j, which again creates the need for risk measures to determine the optimal hedge.

Having studied the structure of the problem, we turn our attention to predicting seasonal trends and patterns in the data.

2.5 Calibration and Prediction

In this section we calibrate seasonality curves to load and price data from 2012 and use this to predict seasonality curves for 2013 and 2014. Furthermore, we describe how to calibrate expected monthly prices using base load contracts and peak load contracts. Finally, we determine a fixed price for 2013 and 2014 based on 2012 data.

2.5.1 Seasonal Component for Load Data

The load data is from two portfolios of customers on fixed price contracts from the bidding areas West Denmark (DK1) and East Denmark (DK2). The price data includes area prices for the two bidding areas as well as the system price for 2012-2014. The bidding areas are seen to have different load characteristics and are therefore modeled separately. In particular, the load portfolio of DK1 is strongly affected by weekends and holidays, whereas the portfolio in DK2 is primarily affected by yearly variations in demand, see Figure 2.4. We let θ_t be the periodic function

$$\theta_t = \alpha + (1 + A_0 \cos(\frac{2\pi}{\tau_0}t + B_0)^2) \sum_{i=1}^{p} A_i \sin(\frac{2\pi}{\tau_i}t + B_i) \tag{2.5.1}$$

with p periods $\tau_0, \dots \tau_p$, amplitudes $A_0, \dots, A_p$ and phases $B_0, \dots B_p$. A_0, τ_0 and B_0 serve to capture seasonal behavior in the amplitude that occurs for the load of DK2 and we set $A_0 = 0$ in DK1. For calibration the load data is split in three subsets; weekdays, weekends and holidays. The function θ_t is calibrated to data from each of the subsets by numerically minimizing the sum of quadratic deviations and combined to the dotted curve shown in Figure 2.4. The periods are based on peaks of autocorrelation functions for 2012 data, with $\tau_0 = 2 \cdot 24 \cdot 365$, $\tau_1 = 12$, $\tau_2 = 24$, $\tau_3 = 24 \cdot 7$, $\tau_4 = 24 \cdot 365$, $\tau_5 = 24 \cdot 365$.[2] Using the load for 2012 we predict the seasonality curves for 2013 and 2014 based on holidays, weekends and day-light savings. To reflect the long-term increase of load, α is adjusted to match the yearly average, which can usually be predicted with high accuracy by electricity companies. Figure 2.5 shows that the load can be calibrated extremely well, i.e. the behavior of the data is very close to that of the function θ_t. This is also confirmed by a coefficient of determination for out-of-sample data of 0.823 and 0.923 for DK1 and DK2, respectively.

[2] We use two curves with yearly frequency to capture the yearly patterns.

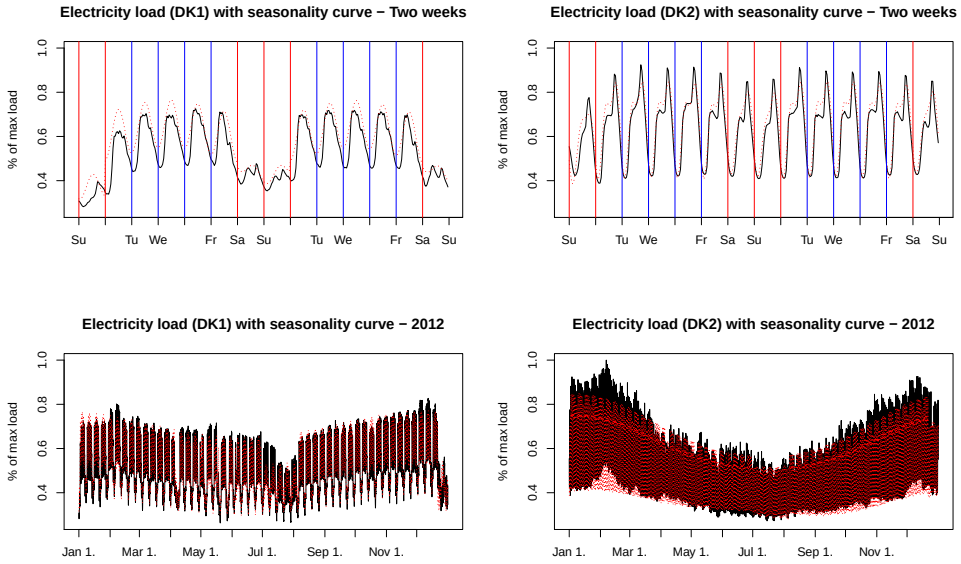


Figure 2.4: Seasonality curves (red) calibrated to historical electricity load (black) in 2012.

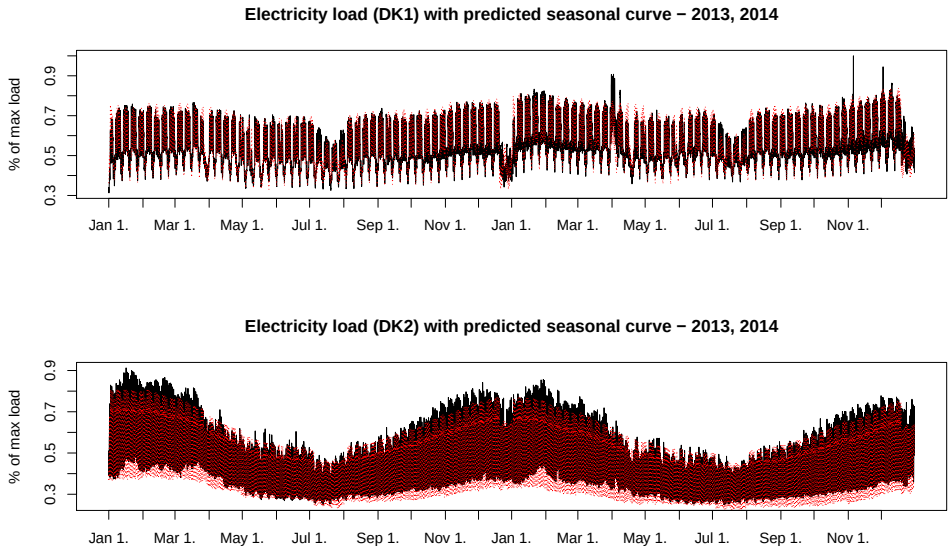


Figure 2.5: Predicted seasonality curves (red) for 2013 and 2014 and historical electricity load (black).

2.5.2 Price Data

We apply the same approach for calibration and prediction of seasonality in prices. In the periodic function, we let $A_0 = 0$. The calibration results are shown in Figure 2.6 with $\tau_1 = 12$, $\tau_2 = 24$, $\tau_3 = 24 \cdot 7$. To adjust for more long-term variations in the system price, the forward prices of base load contracts and peak load contracts are used to adjust the monthly mean of the seasonality curves for the system price in peak and off-peak periods such that

$$\frac{1}{|peak_j|} \sum_{t \in peak_j} \theta_t^{sys} = q_j^p, \tag{2.5.2}$$

$$\frac{1}{|\mathrm{m}_j|} \sum_{t \in \mathrm{m}_j} \theta_t^{sys} = q_j^b. \tag{2.5.3}$$

We ignore the market price of risk as well as discounting to simplify results. Furthermore, due to risk premium and seasonal bias in forward prices for base load contracts and CfDs, see Bessembinder and Lemmon (2002) and Kristiansen (2004), we do not use them to adjust the seasonality curves for the area prices. For the three prices, the random part is more dominating than the seasonality curve, which is reflected by a coefficient of determination for out-of-sample data of 0.213, 0.211 and 0.363 for the DK1 area price, the DK2 area price and system price, respectively.[3]

[3]The coefficient of determination for DK1 has been computed without including a 5 hour price spike with prices over 1900 Euro/MWh in June 2013.

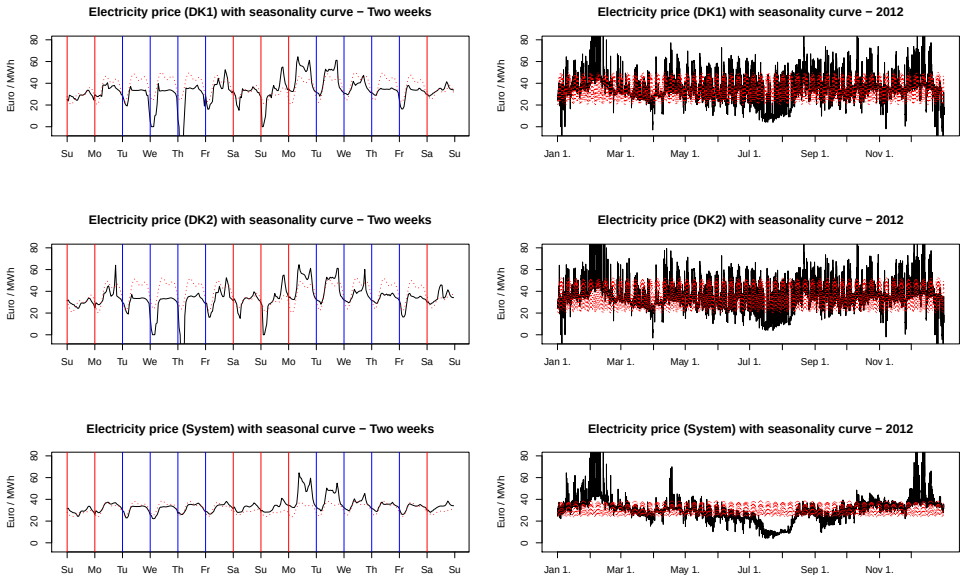


Figure 2.6: Calibrated seasonality curves (red) and historical electricity prices (black) for 2012.

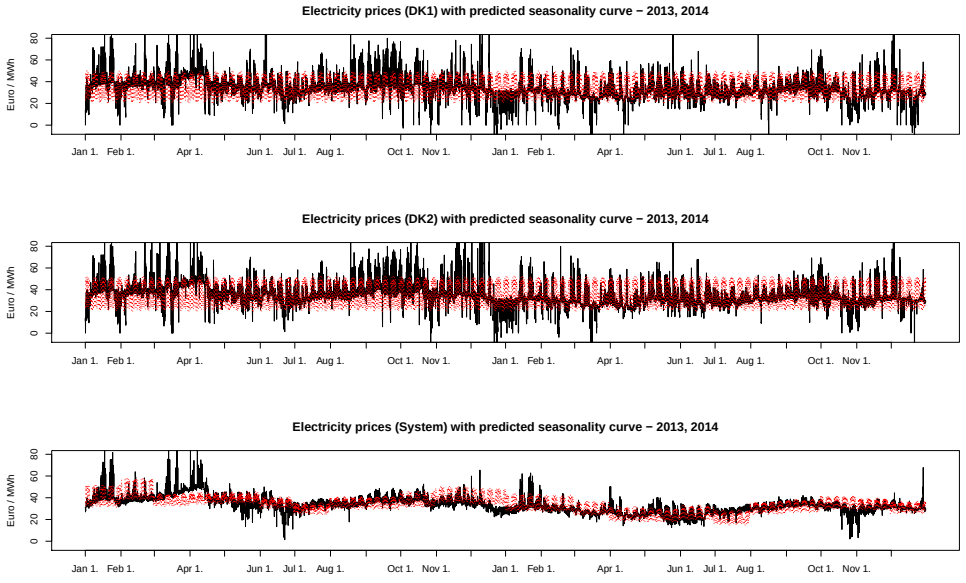


Figure 2.7: Predicted seasonality curves (red) and historical electricity prices (black) for 2013 and 2014. Extreme price spikes are not displayed in the plots.

2.5.3 Fixed Cost of Electricity

The fixed price F_j for each month in 2013 and 2014 is determined as

$$F_j = \frac{\sum_{t\in \mathrm{m}_j} S_t L_t}{\sum_{t\in \mathrm{m}_j} L_t}, \tag{2.5.4}$$

using the data from 2012. This implies that

$$\sum_{t\in \mathrm{m}_j} F_j L_t = \sum_{t\in \mathrm{m}_j} S_t L_t \tag{2.5.5}$$

and that the company would break even in 2012. With this construction F_j will typically be higher than the average off-peak price and lower than the average peak price, indicating expected profit in off-peak hours and expected loss in peak hours. In practice F_j is increased with a margin to increase profitability of the contract and compensate for the risk, but initially we study the problem without the margin.
In the next section, we introduce three models for the stochastic evolution of the area price, the system price and the load. For the first and second model we determine the minimal variance hedge, Min Var, analytically, where the objective is the sum of variances of the hourly cash flow

$$\begin{aligned}\sum_{t\in \mathrm{m}_j} Var\Big(&(F_j - S_t)L_t + (S_t^{sys} - q_j^b)V_j^b \\ &+(S_t - S_t^{sys} - q_j^d)V_j^d + 1_{(t\in peak_j)}(S_t^{sys} - q_j^p)V_j^p\Big).\end{aligned} \tag{2.5.6}$$

For the third model we find the Min Var hedge numerically and include the hedging strategy Min Loss, which minimizes

$$\begin{aligned}\sum_{t\in \mathrm{m}_j} -E\Big[\min\Big(&(F_j - S_t)L_t + (S_t^{sys} - q_j^b)V_j^b \\ &+(S_t - S_t^{sys} - q_j^d)V_j^d + 1_{(t\in peak_j)}(S_t^{sys} - q_j^p)V_j^p, 0\Big)\Big].\end{aligned} \tag{2.5.7}$$

This risk measure focuses on the expected loss for every hour, and the accompanying hedging strategy is expected to result in payoffs that decrease very little. In contrast, the minimal variance hedge is expected to result in payoffs that vary very little.

2.6 Model Set-up

We consider three different models for the deviations from the seasonality curve, all with the underlying assumption that

$$S_t = \theta_t^S + \tilde{S}_t, \tag{2.6.1}$$
$$S_t^{sys} = \theta_t^{sys} + \tilde{S}_t^{sys}, \tag{2.6.2}$$
$$L_t = \theta_t^L + \tilde{L}_t, \tag{2.6.3}$$

where θ_t^S, θ_t^{sys} and θ_t^L are seasonal components of the area price, the system price and the load and $\tilde{S}_t$, $\tilde{S}_t^{sys}$ and $\tilde{L}_t$ are the deseasonalized components. All of the models capture the mean-reverting tendencies and seasonality that we observe in the data. Furthermore, the seasonal components introduce empirical correlation over time as well as empirical cross correlation between the area price, system price and load in all three models. In the first two models, however, we disregard the correlation over time and assume a simple cross-correlation for the deseasonalized components. The simple correlation structure is obtained by formulating the models in terms of the difference between the area price and system price. The third model incorporates correlation over time as well a more advanced structure of cross-correlations by a direct modeling of the area price.

2.6.1 Model 1 - Independent Price, Load and Difference to System Price

In this model we let $\tilde{\epsilon}_t = \tilde{S}_t - \tilde{S}_t^{sys}$ be the difference between the deseasonalized area and system price such that the area price is the system price plus some noise due to congestion. We assume that

$$\begin{pmatrix} \tilde{S}_t^{sys} \\ \tilde{\epsilon}_t \\ \tilde{L}_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{sys}^2 & 0 & 0 \\ 0 & \nu^2 & 0 \\ 0 & 0 & \sigma_L^2 \end{pmatrix} \right) \tag{2.6.4}$$

and that $(\tilde{S}_t^{sys}, \tilde{\epsilon}_t, \tilde{L}_t)$ are independent of $(\tilde{S}_{t'}^{sys}, \tilde{\epsilon}_{t'}, \tilde{L}_{t'})$ for $t \neq t'$. This is equivalent to

$$\begin{pmatrix} S_t \\ S_t^{sys} \\ L_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \theta_t^S \\ \theta_t^{sys} \\ \theta_t^L \end{pmatrix}, \begin{pmatrix} \sigma_{sys}^2 + \nu^2 & \sigma_{sys}^2 & 0 \\ \sigma_{sys}^2 & \sigma_{sys}^2 & 0 \\ 0 & 0 & \sigma_L^2 \end{pmatrix} \right). \tag{2.6.5}$$

With these assumptions, we obtain the analytical minimal variance hedge that is given by

$$V_j^b = \frac{1}{|off_j|} \sum_{t \in off_j} \theta_t^L, \tag{2.6.6}$$

$$V_j^p = \frac{1}{|peak_j|} \sum_{t \in peak_j} \theta_t^L - \frac{1}{|off_j|} \sum_{t \in off_j} \theta_t^L, \tag{2.6.7}$$

$$V_j^d = \frac{1}{|m_j|} \sum_{t \in m_j} \theta_t^L. \tag{2.6.8}$$

This estimate for a hedging strategy only depends on the prediction of the load, which is one of the reasons it is widely used by electricity companies. We refer to this as the mean hedge.

2.6.2 Model 2 - Correlated Price and Load, but Independent Difference to System Price

In the second model we include correlation between the deseasonalized load and deseasonalized system price. The motivation for this is that as the system prices reflect the equilibrium between supply and demand. Thus, if the load is above its expectation, there is a tendency for the system price to likewise be above its expectation, and similarly when the load is below its expectation. This is formalized as

$$\begin{pmatrix} \tilde{S}_t^{sys} \\ \tilde{\epsilon}_t \\ \tilde{L}_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{sys}^2 & 0 & \rho\sigma_{sys}\sigma_L \\ 0 & \nu^2 & 0 \\ \rho\sigma_{sys}\sigma_L & 0 & \sigma_L^2 \end{pmatrix} \right), \tag{2.6.9}$$

and is equivalent to

$$\begin{pmatrix} S_t \\ S_t^{sys} \\ L_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \theta_t^S \\ \theta_t^{sys} \\ \theta_t^L \end{pmatrix}, \begin{pmatrix} \sigma_{sys}^2 + \nu^2 & \sigma_{sys}^2 & \rho\sigma_{sys}\sigma_L \\ \sigma_{sys}^2 & \sigma_{sys}^2 & \rho\sigma_{sys}\sigma_L \\ \rho\sigma_{sys}\sigma_L & \rho\sigma_{sys}\sigma_L & \sigma_L^2 \end{pmatrix} \right). \tag{2.6.10}$$

We find the analytical minimal variance in terms of an adjusted load

$$\tilde{\theta}_t^L = \theta_t^L + (\theta_t^S - F)\rho\frac{\sigma_L}{\sigma_{sys}} \tag{2.6.11}$$

to be

$$V_j^b = \frac{1}{|off_j|} \sum_{t \in off_j} \tilde{\theta_t^L}, \tag{2.6.12}$$

$$V_j^p = \frac{1}{|peak_j|} \sum_{t \in peak_j} \tilde{\theta_t^L} - \frac{1}{|off_j|} \sum_{t \in off_j} \tilde{\theta_t^L}, \tag{2.6.13}$$

$$V_j^d = \frac{1}{|m_j|} \sum_{t \in m_j} \theta_t^L, \tag{2.6.14}$$

which is seen to be the natural extension of Model 1. The hedging strategies from this model correspond to hedging slightly above expected load for high prices and slightly below expected load for low prices due to the positive price and load correlation. Note that the hedging volume for contracts for difference remain unchanged and that the peak load hedge, but not the effective peak hedge, is independent of F.

2.6.3 Model 3 - Correlation Over Time and Between Load, Area Price and System Price

In the third model we include temporal correlation in the deseasonalized components and assume $(\tilde{S}_t, \tilde{S}_t^{sys}, \tilde{L}_t)$ follow a three-dimensional Ornstein-Uhlenbeck process given by

$$\mathrm{d}\tilde{S}_t = -\kappa_S \tilde{S}_t \, \mathrm{d}t + \tilde{\sigma}_S \, \mathrm{d}Z_t^S, \tag{2.6.15}$$

$$\mathrm{d}\tilde{S}_t^{sys} = -\kappa_{sys} \tilde{S}_t^{sys} \, \mathrm{d}t + \tilde{\sigma}_{sys} \, \mathrm{d}Z_t^{sys}, \tag{2.6.16}$$

$$\mathrm{d}\tilde{L}_t = -\kappa_L \tilde{L}_t \, \mathrm{d}t + \tilde{\sigma}_L \, \mathrm{d}Z_t^L. \tag{2.6.17}$$

Here Z_t^S, Z_t^{sys} and Z_t^L are correlated Brownian motions with correlation coefficients $\rho_{S,sys}$,$\rho_{S,L}$,$\rho_{sys,L}$. The explicit solution to equations (2.6.15), (2.6.16) and (2.6.17) conditional on $(\tilde{S}_u, \tilde{S}_u^{sys}, \tilde{L}_u)$ with $u < t$ is

$$\tilde{S}_t = \tilde{S}_u e^{-\kappa_S(t-u)} + \tilde{\sigma}_S \int_u^t e^{-\kappa_S(t-v)} \, \mathrm{d}Z_v^S, \tag{2.6.18}$$

$$\tilde{S}_t^{sys} = \tilde{S}_u^{sys} e^{-\kappa_{sys}(t-u)} + \tilde{\sigma}_{sys} \int_u^t e^{-\kappa_{sys}(t-v)} \, \mathrm{d}Z_v^{sys}, \tag{2.6.19}$$

$$\tilde{L}_t = \tilde{L}_u e^{-\kappa_L(t-u)} + \tilde{\sigma}_L \int_u^t e^{-\kappa_L(t-v)} \, \mathrm{d}Z_v^L, \tag{2.6.20}$$

and, hence, for $t > u$

$$\begin{pmatrix} S_t \\ S_t^{sys} \\ \tilde{L}_t \end{pmatrix} \Bigg| \begin{pmatrix} S_u \\ S_u^{sys} \\ L_u \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \tilde{S}_u e^{-\kappa_S(t-u)} + \theta_t^S \\ \tilde{S}_u^{sys} e^{-\kappa_{sys}(t-u)} + \theta_t^{sys} \\ \tilde{L}_u e^{-\kappa_L(t-u)} + \theta_t^L \end{pmatrix}, \begin{pmatrix} \Sigma_S(u,t) & \Sigma_{S,sys}(u,t) & \Sigma_{S,L}(u,t) \\ \Sigma_{S,sys}(u,t) & \Sigma_{sys}(u,t) & \Sigma_{sys,L}(u,t) \\ \Sigma_{S,L}(u,t) & \Sigma_{sys,L}(u,t) & \Sigma_L(u,t) \end{pmatrix} \right), \tag{2.6.21}$$

with

$$\Sigma_S(u,t) = \frac{\tilde{\sigma}_S^2(1 - e^{-2\kappa_S(t-u)})}{2\kappa_S}, \tag{2.6.22}$$

$$\Sigma_{sys}(u,t) = \frac{\tilde{\sigma}_{sys}^2(1 - e^{-2\kappa_{sys}(t-u)})}{2\kappa_{sys}}, \tag{2.6.23}$$

$$\Sigma_L(u,t) = \frac{\tilde{\sigma}_L^2(1 - e^{-2\kappa_L(t-u)})}{2\kappa_L}, \tag{2.6.24}$$

and

$$\Sigma_{S,sys}(u,t) = \rho_{S,sys}\tilde{\sigma}_S\tilde{\sigma}_{sys}\frac{1-e^{-(\kappa_S+\kappa_{sys})(t-u)}}{\kappa_S+\kappa_{sys}}, \tag{2.6.25}$$

$$\Sigma_{S,L}(u,t) = \rho_{S,L}\tilde{\sigma}_S\tilde{\sigma}_L\frac{1-e^{-(\kappa_S+\kappa_L)(t-u)}}{\kappa_S+\kappa_L}, \tag{2.6.26}$$

$$\Sigma_{sys,L}(u,t) = \rho_{sys,L}\tilde{\sigma}_{sys}\tilde{\sigma}_L\frac{1-e^{-(\kappa_{sys}+\kappa_L)(t-u)}}{\kappa_{sys}+\kappa_L}. \tag{2.6.27}$$

Estimation procedures for the parameters in the three models can be found in Appendix 2.D.

2.6.4 Monte Carlo Simulation

To determine the optimal hedging strategies in Model 3 for each month, we let P_t^k for $t \in m_j$ denote a sample of the stochastic hourly payoff P_t in month j, given by

$$\begin{aligned} P_t^k &= (F_j - S_t^k)L_t^k + (S_t^{sys,k} - q_j^b)V_j^b \\ &\quad + (S_t^k - S_t^{sys,k} - q_j^d)V_j^d + 1_{(t\in peak_j)}(S_t^{sys,k} - q_j^p)V_j^p, \end{aligned} \tag{2.6.28}$$

where $(S_t^k)_{t\in m_j}, (S_t^{sys,k})_{t\in m_j}$ and $(L_t^k)_{t\in m_j}$ for $k = 1,\ldots,K$ are sample paths obtained by simulation from Model 3. We let $\bar{P}_t$ be the sample average of the payoff in hour t given by,

$$\bar{P}_t = \frac{1}{K}\sum_{k=1}^{K} P_t^k. \tag{2.6.29}$$

For Model 3, we determine the hedging strategy, Minimum Var, that minimizes the sum of sample variances of payoffs for hours in month j, H_j, defined as

$$\sum_{t\in m_j}\left(\frac{1}{K-1}\sum_{k=1}^{K}\left[P_t^k - \bar{P}_t\right]^2\right) \approx \sum_{t\in m_j} Var(P_t). \tag{2.6.30}$$

Furthermore, we determine the hedging strategy, Min Loss, that minimizes the sum of sample averages of hourly losses in month j, i.e.

$$\sum_{t\in \mathrm{m}_j}\left(-\frac{1}{K}\sum_{k=1}^{K}\min(P_t^k,0)\right) \approx \sum_{t\in \mathrm{m}_j} -E[\min(P_t,0)]. \tag{2.6.31}$$

2.7 Results

In this section we assess the performance of the optimal hedging strategies and benchmark against the mean hedge strategy derived from Model 1. All

hedging strategies can be determined 14 days prior to the start of the month and does not use any other information than historical data from 2012, yearly predicted load as well as forward prices for base load contracts, peak load contracts and CfDs. Furthermore, all contracts are available at *Nasdaq Commodities* and the market structure closely reflect the real market. We study the impact of modeling correlation in price and load, the inclusion of CfDs, the impact of improved price forecast as well as the impact of margins on the fixed price.

2.7.1 Comparing the Hedging Strategies

To compare the payoff streams in 2013 and 2014 from implementing the optimal hedging strategies, we let P_t denote the payoff in hour t for $t \in m_j$ and $j \in \{1, \ldots, 24\}$ and define the following quantities. The profit and loss (P&L):

$$\sum_{j=1}^{24} \sum_{t \in m_j} P_t. \tag{2.7.1}$$

The gross loss:

$$\sum_{j=1}^{24} \sum_{t \in m_j} -\min(P_t, 0). \tag{2.7.2}$$

The gross profit:

$$\sum_{j=1}^{24} \sum_{t \in m_j} \max(P_t, 0). \tag{2.7.3}$$

Finally, using the average monthly payoff, $\hat{P}_j = \frac{1}{m_j} \sum_{t \in m_j} P_t$, we define the realized variance,

$$\sum_{j=1}^{24} \frac{1}{|m_j| - 1} \sum_{t \in m_j} (P_t - \hat{P}_j)^2. \tag{2.7.4}$$

The realized variance measures the stability of the payoffs throughout each month, but differs from the sum of hourly variances defined in equation (2.6.30). Realized monthly variance can be measured on actual data as opposed to the sum of hourly variances. Minimizing the deviations from the hourly mean, however, creates a more stable cash-flow than minimizing the deviations from the monthly mean.[4] As the load data has been anonymized by scaling with the maximum load, the P&L, gross loss and gross profit is measured in Euro/maximum load. The sum of realized variances is likewise scaled by $1/(\text{maximum load})^2$.

[4]The realized monthly variance for DK1 has been computed without including payoffs from June 2013 due to a 5 hour price spike, with prices over 1900 Euro/MWh as this would blur the comparison significantly.

2.7.2 Comparing Hedging Strategies

From Table 2.1, we initially observe that the hedged cash-flows have a lower P&L than the unhedged as expected, but the gross loss and realized variance have been reduced significantly. Comparing the mean hedge with the compensated mean hedge, we find that the P&L has slightly increased, whereas the realized variance and the gross loss has slightly decreased. This suggests that there is a moderate effect of including the simple correlation between the price and load. Directing our attention to the gross loss, we find that the Min Loss hedge, as anticipated, has the lowest gross loss. As a side effect, the gross profit has increased, which creates a P&L that is significantly higher than the variance minimizing strategies. The realized variance is approximately doubled indicating that the cash-flow shows larger variations throughout each month. More importantly, however, the accumulated P&L does not decrease as much over time as for the Minimum Var strategy. Hence, the Min Loss hedge generates a relatively stable cash flow that outperforms the other strategies in terms of P&L. This is confirmed in the accumulated P&L for both bidding areas, as shown in Figure 2.8. The montly hedging volumes, gross loss, realized variance and P&L are shown in Appendix 2.F. From the hedging volumes it can be seen that including correlation between price and load decreases the base load hedging volume and increases the peak load hedging volume as expected.

	West Denmark (DK1)							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(100.6%)	24630.07	(129.2%)	39673.46	(117.3%)	64303.52	(2367.7%)	844.41
Mean hedge	(0.0%)	12278.02	(0.0%)	17309.71	(0.0%)	29587.73	(0.0%)	34.22
Comp. mean hedge	(0.9%)	12385.57	(-0.3%)	17260.62	(0.2%)	29646.20	(-0.4%)	34.09
Min Var	(2.2%)	12553.24	(0.1%)	17328.72	(1.0%)	29881.95	(4.3%)	35.69
	(12519.66,12574.80)*		(17315.84,17354.60)*		(29835.49,29929.40)*		(35.62,35.75)*	
Min Loss	(17.2%)	14389.79	(-5.8%)	16313.42	(3.8%)	30703.22	(81.3%)	62.03
	(14298.30,14499.95)*		(16187.75,16388.77)*		(30486.05,30888.71)*		(61.19,62.54)*	

	East Denmark (DK2)							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(12069.4%)	21874.97	(50.0%)	31766.36	(151.2%)	53641.33	(1299.7%)	657.72
Mean hedge	(0.0%)	179.75	(0.0%)	21172.65	(0.0%)	21352.41	(0.0%)	46.99
Comp. mean hedge	(93.0%)	346.85	(-0.2%)	21125.01	(0.6%)	21471.85	(-0.9%)	46.55
Min Var	(175.1%)	494.59	(-0.5%)	21066.39	(1.0%)	21560.98	(0.6%)	47.29
	(470.34,504.70)*		(21061.51,21087.03)*		(21531.85,21591.73)*		(47.26,47.33)*	
Min Loss	(2731.7%)	5090.01	(-13.6%)	18293.28	(9.5%)	23383.29	(133.1%)	109.55
	(4934.37,5177.18)*		(18194.50,18400.45)*		(23128.87,23577.63)*		(108.36,110.80)*	

Table 2.1: Performance of hedging strategies in DK1 and DK2. Mean hedge and comp. mean hedge refer to variance minimizing strategies based on Model 1 and Model 2 respectively. Relative change from mean hedge in parenthesis.

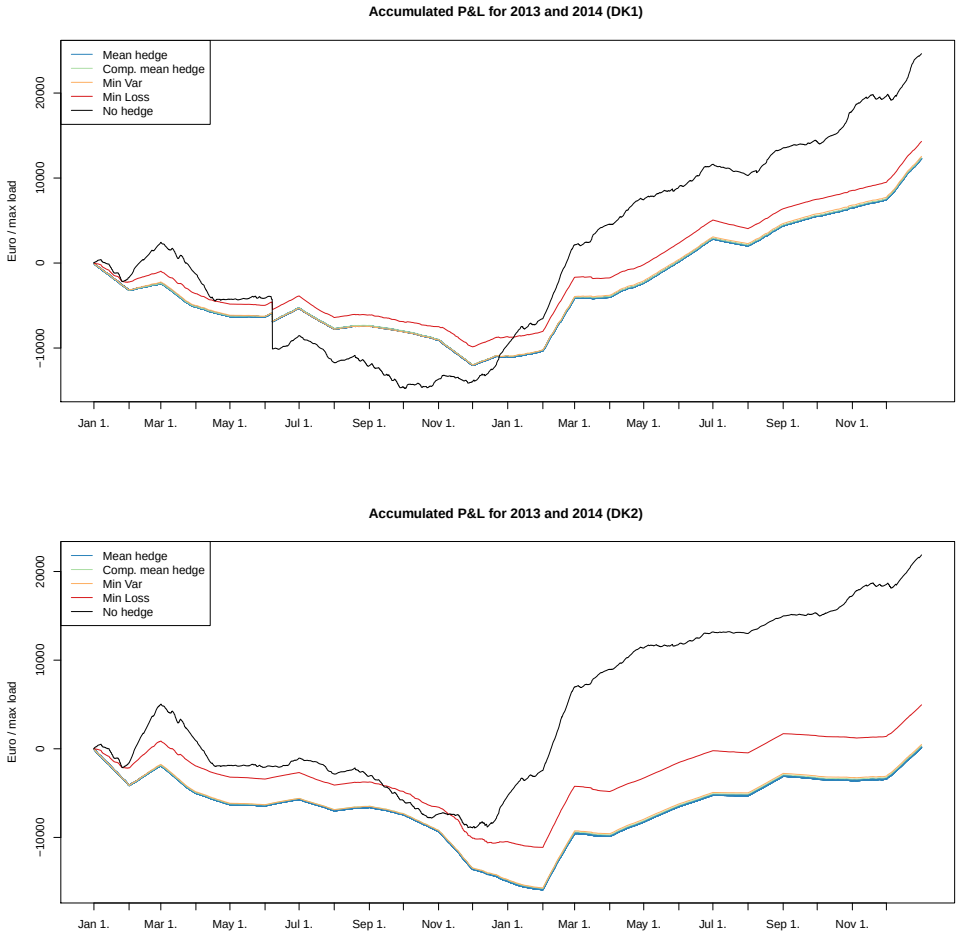


Figure 2.8: Accumulated P&L in DK1 and DK2 with the hedging strategies. The variance minimizing strategies have very similar accumulated payoffs.

2.7.3 The Choice of Risk Measure and the Impact of Correlation

Comparing common practice in the industry, i.e. the mean hedge strategy, with the Min Loss hedge, we find that the gross loss is reduced by 6 % and 13.6% in DK1 and DK2, respectively. Furthermore, the gross profit is increased by 3.8 % and 9.5% in DK1 and DK2, respectively.[5] The realized variance is also

[5]The relative change of the P&L is not always well defined as the numerator can be both positive and negative, which results in the change of 17.2% and 2731.7 % in DK1 and DK2, respectively.

*95% confidence interval based on 16 simulations with 1000 paths each. The actual value is based on a single simulation with 1000 paths.

increased, but as this includes positive deviations from the monthly mean, it is of less importance than the gross loss. We note from the monthly P&L in Appendix 2.F that the largest difference between the variance minimizing strategies and the Min Loss hedge are in the months with a negative P&L. In these months the Min Loss hedge incurs much smaller loss resulting in a larger accumulated P&L over the two years.

2.7.4 Including CfDs

In this section we quantify the impact of including CfDs by repeating the analysis from Section 2.7.2 assuming that the CfD contracts are not available. Table 2.2 illustrates that for DK1, the inclusion of CfDs reduces the gross loss by 39.3% to 47.7%, whereas the gross profit decreases by 37.6% to 41.1%. Thus, gross loss is reduced significantly by introducing the CfDs for the three strategies, but at the expense of a decrease in profit. Table 2.2 illustrates that in DK2, the inclusion of CfDs reduces the gross loss by 2.3% to 19.4%, whereas the gross profit is reduced by 42.8% to 47%. This suggest that the benefits of including CfDs in DK2 are smaller than for DK1. A plausible explanation is that the risk premium for CfDs is larger in DK2 than in DK1, which could be due to more risk averse market participants in East Denmark than in West Denmark. The impact on the accumulated payoff of including CfDs is shown in Figure 2.9. We note that the accumulated payoffs are more volatile without the CfDs and that the price spike in June in DK1 barely affects the accumulated payoff, when the CfDs are included.

	West Denmark (DK1) - No CfDs available							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(0.0%)	24630.07	(0.0%)	39673.46	(0.0%)	64303.52	(0.0%)	844.41
Mean hedge	(-35.0%)	18895.39	(-39.3%)	28496.69	(-37.6%)	47392.09	(-93.1%)	497.40
Comp. mean hedge	(-34.8%)	19002.95	(-39.4%)	28483.75	(-37.6%)	47486.70	(-93.2%)	498.34
Min Var	(-41.5%)	21469.43	(-45.9%)	32051.91	(-44.2%)	53521.34	(-94.1%)	603.46
	(21419.13,21524.87)*		(32003.20,32074.13)*		(53422.33,53599.00)*		(602.39,604.09)*	
Min Loss	(-31.0%)	20866.83	(-47.7%)	31217.89	(-41.1%)	52084.71	(-89.1%)	571.10
	(20663.33,20979.08)*		(31060.72,31334.22)*		(51724.05,52313.31)*		(569.28,572.29)*	

	East Denmark (DK2) - No CfDs available							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(0.0%)	21874.97	(0.0%)	31766.36	(0.0%)	53641.33	(0.0%)	657.72
Mean hedge	(-98.9%)	15645.24	(-2.3%)	21670.22	(-42.8%)	37315.46	(-86.7%)	353.81
Comp. mean hedge	(-97.8%)	15812.33	(-2.4%)	21650.83	(-42.7%)	37463.16	(-86.9%)	355.05
Min Var	(-97.2%)	17391.40	(-9.5%)	23280.18	(-47.0%)	40671.59	(-88.3%)	403.36
	(17327.58,17432.99)*		(23228.10,23302.88)*		(40555.68,40735.88)*		(402.33,403.67)*	
Min Loss	(-69.7%)	16807.93	(-19.5%)	22730.92	(-40.9%)	39538.86	(-71.4%)	382.52
	(16601.22,16876.88)*		(22618.61,22834.60)*		(39219.83,39711.47)*		(380.98,383.53)*	

Table 2.2: Comparison of hedging strategies with no CfDs available. Relative change by including CfDs in parenthesis.

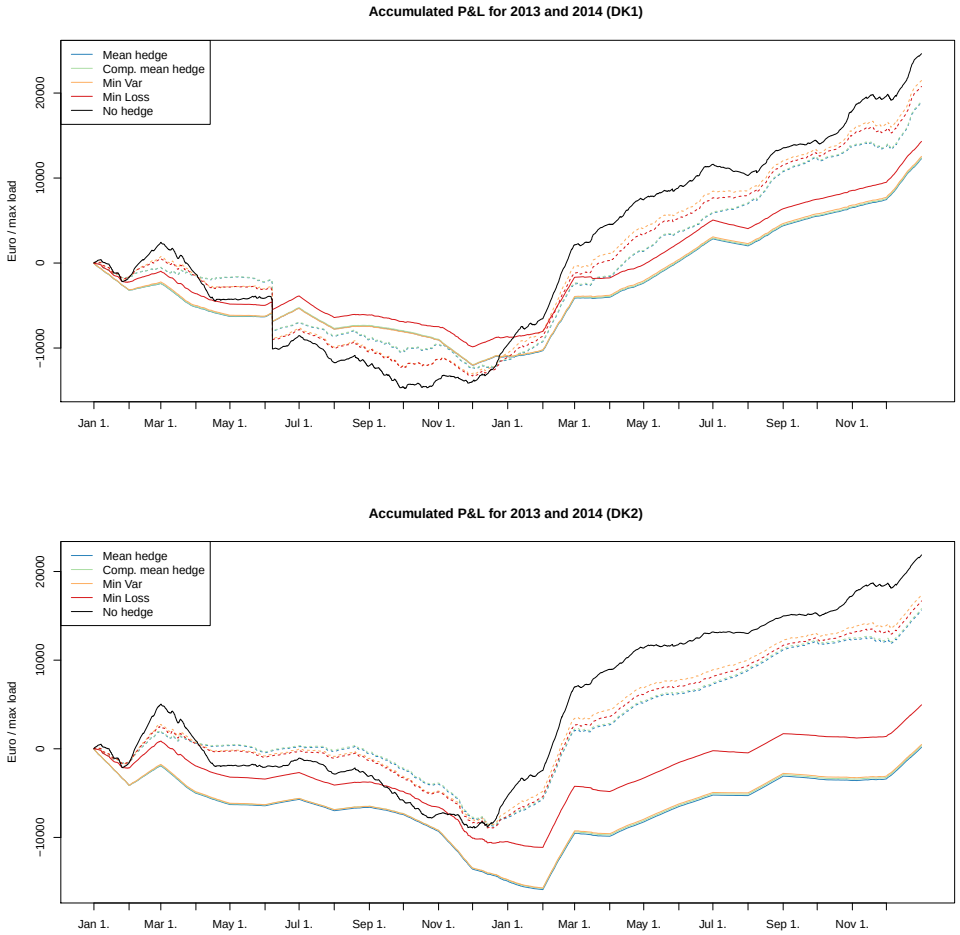


Figure 2.9: Accumulated P&L with (solid lines) and without (dashed lines) CfDs in East Denmark (DK1) and West Denmark (DK2).

2.7.5 Perfect Forecast of Average Prices

The differences between the hedging strategies quantified in Section 2.7.2 and Section 2.7.4 may be due to model assumptions such as the inclusion of auto- and cross correlations, choice of risk measure, availability of hedging instruments, but also the ability of the price model to predict the average prices used in the hedging strategies. Whereas the mean hedge is only based on the prediction of expected load, the more advanced hedging strategies depend on the predictions of additional parameters. We therefore quantify the impact of being able to more accurately predict average prices. In particular, we assume a perfect forecast of monthly average prices in peak and off-peak periods. This does not impact Model 1 as the hedging strategy is independent of the pre-

dicted price, and even though Model 2 depends on the predicted price, the performance does not changes as the hedging strategies are not very sensitive to changes in predicted prices. The results in Table 2.3 show that the impact on the hedging strategy Min Var is also very limited. This is not the case for the Min Loss hedge, where the gross loss is reduced by 39.3% and 37.7% in DK1 and DK2, respectively, while P&L are increased by 66.9% and 216.4%. This suggests that the superiority of the advanced hedging strategies is limited by the ability to predict average prices, and therefore, that an improved price forecast can significantly improve the Min Loss hedge. Some of these improvements could be obtained by modeling the seasonal bias on base load contracts, peak load contracts and CfDs, but due to unavailability of data, this has not been further pursued in this paper.

	West Denmark (DK1) - Perfect expected price forecast							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(0.0%)	24630.07	(0.0%)	39673.46	(0.0%)	64303.52	(0.0%)	844.41
Mean hedge	(0.0%)	12278.02	(0.0%)	17309.71	(0.0%)	29587.73	(0.0%)	34.22
Comp. mean hedge	(0.0%)	12385.57	(0.0%)	17260.62	(0.0%)	29646.20	(0.0%)	34.09
Min Var	(2.0%)	12806.32	(-0.8%)	17196.75	(0.4%)	30003.07	(-1.5%)	35.17
	(12771.81,12827.93)*		(17183.96,17223.73)*		(29955.77,30051.67)*		(35.09,35.23)*	
Min Loss	(66.9%)	24018.18	(-39.3%)	9896.73	(10.5%)	33914.91	(-19.4%)	50.00
	(23957.11,24073.28)*		(9848.63,9921.47)*		(33805.75,33994.75)*		(49.80,50.99)*	

	East Denmark (DK2) - Perfect expected price forecast							
	P&L		Gross Loss		Gross Profit		Realized Variance	
No hedge	(0.0%)	21874.97	(0.0%)	31766.36	(0.0%)	53641.33	(0.0%)	657.72
Mean hedge	(0.0%)	179.75	(0.0%)	21172.65	(0.0%)	21352.41	(0.0%)	46.99
Comp. mean hedge	(0.0%)	346.85	(0.0%)	21125.01	(0.0%)	21471.85	(0.0%)	46.55
Min Var	(35.4%)	669.57	(-0.6%)	20947.69	(0.3%)	21617.26	(-0.9%)	46.87
	(645.97,679.30)*		(20944.07,20967.71)*		(21590.04,21647.01)*		(46.84,46.91)*	
Min Loss	(216.4%)	16104.58	(-37.7%)	11404.15	(17.6%)	27508.72	(-21.2%)	86.33
	(16020.23,16147.65)*		(11344.87,11428.31)*		(27365.10,27575.96)*		(85.69,87.87)*	

Table 2.3: Comparison of hedging strategies with a perfect forecast of expected prices. Relative change from imperfect forecast.

2.7.6 Margin

By changing the fixed price to $\tilde{F} = F + 2$, we obtain an increase in the expected payoffs by approximately the total scaled load times the margin of 2. DK1 and DK2 have a total scaled load of 9758.8 and 8409.5 over the two years, resulting in an increase of approximately 19517.6 and 16819 Euro times the maximal load. For the compensated mean hedge the base load volume is reduced by less than 1% and for the Min Var strategy the hedging volumes are reduced by less than 1%, suggesting that small changes in margin to the fixed price only have a moderate impact on the variance minimizing hedging strategies. In contrast, the Min Loss hedge changes significantly, but the P&L still increases by the margin times the scaled load and the gross loss remains significantly lower than for the variance minimizing strategies.

2.8 Conclusion and Extensions

2.8.1 Conclusion

In this paper, we develop hedging strategies for an electricity distributor in the Nordic Electricity market who manages price and volume risk from fixed price agreements on stochastic electricity load. We analyze the market dynamics in the two bidding areas of West Denmark and East Denmark with a focus on the correlation structure between area price, system price and load and quantify the impact of including auto- and cross-correlations. When benchmarking against hedging at expected load, which is common practice in the industry, we find that compensating for correlation between price and load pay-off slightly increases and the realized variance similarly decreases. This can typically be achieved by hedging above the mean in peak periods and below the mean in off-peak periods. By using expected loss as a risk measure instead of variance we further improve performance compared to common practice in the industry, both in terms of risk and profit. In one area, the gross loss is reduced by 5.8% and the gross profit is increased by 3.8%. In the other area, the gross loss is reduced by 13.6% and the gross profit is increased by 9.5%. We show how the inclusion of CfDs in addition to peak load and base load contracts can likewise reduce risk, but illustrate that this may be at the expense of a high risk premium. Finally, we demonstrate how improved forecasts of average prices have substantial potential to further improve performance.
We conclude that for companies that currently use the mean hedge strategy, the accumulated payoffs can be significantly increased, while at the same time reducing the loss from hours of negative payoffs. This is achieved by the implementation of a more advanced price model and a hedging strategy that exploits the asymmetry of payoffs in the presence of volume risk.

2.8.2 Improvements and Extensions

The process of differences between area and system prices, ϵ_t, is modelled as a sequence of i.i.d. Normal variables with a fixed low volatility. In reality, however, the behavior of the differences may closer resemble that of a jump process, as the congestion problems causing the difference are usually quickly resolved. Not only the differences in prices may be modeled as a jump process, but the price process itself could also be extended to include jumps. In both cases, price spikes caused by congestion may be even better captured by including demand in local and neighboring bidding areas as exogenous factors. The modeling of spikes and temporary behavior is, however, significantly more difficult. The inclusion of such extreme behavior requires long stationary time series, and may not even be possible due to slow changes on the demand side as well as the supply side. For a more detailed analysis of regime switching models and jump diffusion models for electricity prices see Weron et al.

(2004). Finally, improved work on compensating for the forward price bias to obtain better predictions for the monthly mean of the system price and area price could significantly improve the hedging strategies. The price predictions could likewise be improved by calibration using demand predictions and by monthly recalibration of the seasonal components to better adjust for the current trends.

Appendix 2.A Lemmas

Lemma 2.A.1. *The hedge V that minimizes*

$$Var\left((F - S_T)L_T + (S_T - q_t(T)V)\right) \tag{2.A.1}$$

is

$$V^* = \frac{Cov(S_T, S_T L_T)}{Var(S_T)} - F\frac{Cov(L_T, S_T)}{Var(S_T)}. \tag{2.A.2}$$

Proof.

$$\begin{aligned} &Var\left((F - S_T)L_T + (S_T - q_t(T))V\right) \\ &= Var(FL_T + S_T V - S_T L_T) \end{aligned} \tag{2.A.3}$$

$$\begin{aligned} &= F^2 Var(L_T) + V^2 Var(S_T) + Var(S_T L_T) \\ &+ 2VFCov(L_T, S_T) - 2FCov(L_T, S_T L_T) \\ &- 2VCov(S_T, S_T L_T). \end{aligned} \tag{2.A.4}$$

The first order condition implies that

$$2V^* Var(S_T) + 2FCov(L_T, S_T) - 2Cov(S_T, S_T L_T) = 0 \tag{2.A.5}$$

and the second order condition is satisfied as $2Var(S_T) \geq 0$. Thus the optimal hedge is (2.A.2). □

Lemma 2.A.2. *V^* from Lemma 2.A.1 can be written as*

$$\begin{aligned} V^* = E(L_T) - (F - E(S_T))\frac{Cov(S_T, L_T)}{Var(S_T)} \\ + \frac{Cov((S_T - E(S_T))^2, L_T)}{Var(S_T)} \end{aligned} \tag{2.A.6}$$

Proof. Using $E(XY) = E(X)E(Y) + Cov(X, Y)$ and $Cov(X, Y) = Cov(X + a, Y)$ for a constant, it follows that

$$Cov(S_T, S_T L_T) = E(S_T^2 L_T) - E(S_T)E(S_T L_T) \tag{2.A.7}$$

$$\begin{aligned} &= E(S_T^2)E(L_T) + Cov(S_T^2, L_T) \\ &- E(S_T)(E(S_T)E(L_T) + Cov(S_T, L_T)) \end{aligned} \tag{2.A.8}$$

$$\begin{aligned} &= (E(S_T^2) - E(S_T)^2)E(L_T) \\ &+ Cov(S_T^2, L_T) - E(S_T)Cov(S_T, L_T) \end{aligned} \tag{2.A.9}$$

$$\begin{aligned} &= Var(S_T)E(L_T) \\ &+ Cov((S_T - E(S_T))^2, L_T) \\ &+ E(S_T)Cov(S_T, L_T) \end{aligned} \tag{2.A.10}$$

Inserting $Cov(S_T, S_T L_T)$ in (2.A.2) we obtain (2.A.6). □

Lemma 2.A.3. *Assume*

$$\begin{pmatrix} L_T \\ S_T \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_L \\ \mu_S \end{pmatrix}, \begin{pmatrix} \sigma_S^2 & \rho\sigma_S\sigma_L \\ \rho\sigma_S\sigma_L & \sigma_L^2 \end{pmatrix}\right). \tag{2.A.11}$$

The hedge that minimizes

$$Var\left((F - S_T)L_T + V(S_T - q_t(T))\right) \tag{2.A.12}$$

is given by

$$V^* = \mu_L - (F - \mu_S)\frac{\rho\sigma_L}{\sigma_S} \tag{2.A.13}$$

Proof. We want to determine V^* from Lemma 2.A.2 and have $Cov(L_T, S_T) = \rho\sigma_S\sigma_L$. Let X and Y be independent with $X, Y \sim \mathcal{N}(0,1)$. Then,

$$\begin{pmatrix} S_T \\ L_T \end{pmatrix} \overset{d}{=} \begin{pmatrix} \mu_S + \sigma_S X \\ \mu_L + \sigma_L(\rho X + \sqrt{1-\rho^2}Y) \end{pmatrix} \tag{2.A.14}$$

and using independence of X and Y as well as $E(Y) = E(X^3) = 0$, we find that

$$\begin{aligned} &Cov((S_T - E(S_T))^2, L_T) \\ &= E\left[((S_T - E(S_T))^2 - Var(S_T))(L_T - E(L_T))\right] && (2.A.15) \\ &= E\left[(\sigma_S X)^2\sigma_L(\rho X + \sqrt{1-\rho^2}Y)\right] && (2.A.16) \\ &= \rho\sigma_L\sigma_S^2 E(X^3) + \sqrt{1-\rho^2}\sigma_L\sigma_S^2 E(X^2Y) && (2.A.17) \\ &= 0 && (2.A.18) \end{aligned}$$

and by inserting in (2.A.6) we obtain (2.A.13). □

Appendix 2.B Examples

Example 2.B.1. *In Example 2.3.2 the forward price, q, was lower than the expected price, known as backwardation. For commodities the opposite situation may also occur. Consider the same parameters as in Example 2.3.2, but with $q_t(T) = 36.75$. This situation, $q_t(T) < E(S_T)$, is known as contango. In the first plot of Figure 2.10 where $F = 40$, the optimal strategies are similar to those of Example 2.3.2. In contrast, for $F = 30$, which is shown in the second plot of Figure 2.10, the Min Loss hedge ($V = 0.226$) deviates significantly from the mean hedge ($V = 0.5$) in the opposite direction of the minimal variance hedge ($V = 0.525$).*

$F = 40, q = 36.75$	Hedged volume	Expected payoff	Standard deviation	Expected loss
No hedge	0.000	2.000	4.899	-1.143
Var hedge	0.475	1.169	1.199	-0.153
Loss hedge	0.448	1.216	1.228	-0.146
Mean Hedge	0.500	1.125	1.225	-0.172

$F = 30, q = 36.75$	Hedged volume	Expected payoff	Standard deviation	Expected loss
No hedge	0.000	-3.001	5.386	-3.836
Var hedge	0.525	-3.919	1.199	-3.919
Loss hedge	0.227	-3.397	3.217	-3.496
Mean Hedge	0.500	-3.875	1.225	-3.875

Table 2.4: Payoff statistics for different hedging strategies from Example 2.3.2. (Contango).

$F = 30, q = 36.75$	Hedged volume	Expected payoff	Standard deviation	Expected loss
No hedge	0.000	-3.000	5.386	-3.833
Var hedge	0.525	-3.919	1.199	-3.919
Loss hedge	0.226	-3.396	3.219	-3.494
Mean Hedge	0.500	-3.875	1.225	-3.875

$F = 30, q = 29.75$	Hedged volume	Expected payoff	Standard deviation	Expected loss
No hedge	0.000	-3.000	5.385	-3.833
Var hedge	0.525	-0.244	1.199	-0.502
Loss hedge	0.600	0.150	1.415	-0.417
Mean Hedge	0.500	-0.375	1.225	-0.571

Table 2.5: Payoff statistics for different hedging strategies from Example 2.3.2. (Backwardation).

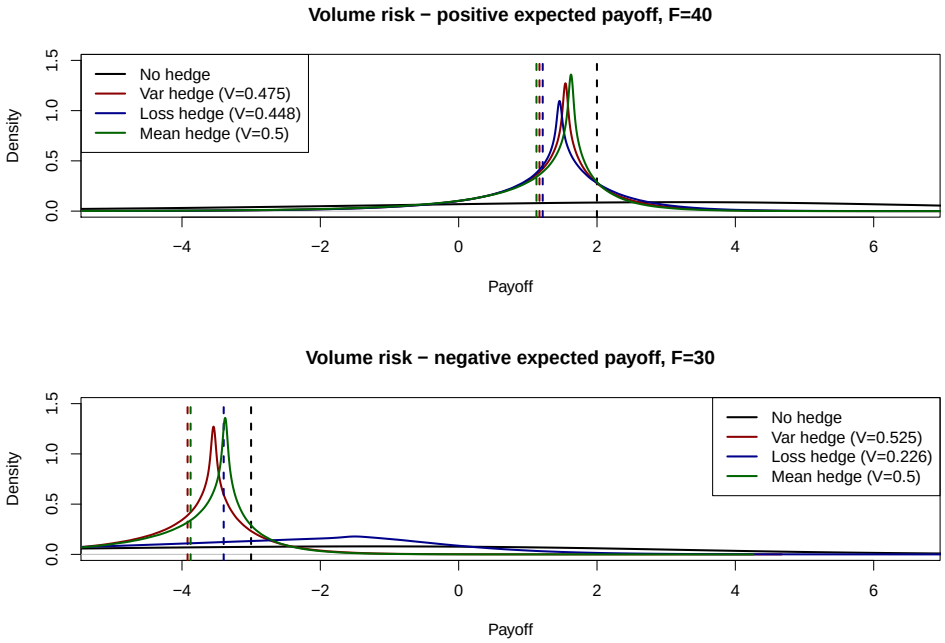


Figure 2.10: Payoff densities (bold lines) and their means (dashed lines) with parameters from Example 2.3.2 and $q = 36.75$. (Contango).

Appendix 2.C Variance Analysis

For independent hourly payoffs, which is assumed in Model 1 and Model 2, the sum of variances of payoffs equals the variance of the sum of payoffs. Thus, to simplify notation we minimize the variance of the sum. The variance of the payoffs for month j is

$$
\begin{aligned}
& Var\Bigg(\sum_{t\in \mathrm{m}_j} (S_t^{sys} - q_j^b)V_j^b + \sum_{t\in \mathrm{m}_j} (S_t - S_t^{sys} - q_j^d)V_j^d \\
& + \sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t + \sum_{t\in peak_j} (S_t^{sys} - q_j^p)V_j^p \Bigg) \\
=& Var\Bigg(\sum_{t\in \mathrm{m}_j} S_t^{sys}V_j^b + \sum_{t\in \mathrm{m}_j} (S_t - S_t^{sys})V_j^d \\
& + \sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t + \sum_{t\in peak_j} S_t^{sys}V_j^p \Bigg) \\
=& (V_j^b)^2 Var\left(\sum_{t\in \mathrm{m}_j} S_t^{sys} \right) + (V_j^d)^2 Var\left(\sum_{t\in \mathrm{m}_j} (S_t - S_t^{sys}) \right) \\
& + Var\left(\sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t \right) + (V_j^p)^2 Var\left(\sum_{t\in peak_j} S_t^{sys} \right) \\
& + 2V_j^b Cov\Bigg(\sum_{t\in \mathrm{m}_j} S_t^{sys}, \sum_{t\in \mathrm{m}_j} (S_t - S_t^{sys})V_j^d \\
& + \sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t + \sum_{t\in peak_j} S_t^{sys}V_j^p \Bigg) \\
& + 2V_j^d Cov\Bigg(\sum_{t\in \mathrm{m}_j} (S_t - S_t^{sys})V_j^d, \sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t \\
& + \sum_{t\in peak_j} S_t^{sys}V_j^p \Bigg) \\
& + 2V_j^p Cov\left(\sum_{t\in peak_j} S_t^{sys}, \sum_{t\in \mathrm{m}_j} (F_j - S_t)L_t \right).
\end{aligned}
\tag{2.C.1}
\tag{2.C.2}
$$

We can minimize the variance as a function of V_j^p, V_j^b and V_j^d. We find that the first order conditions imply that

$$\begin{aligned} V_j^b = & \frac{Cov\left(\sum_{t\in \mathrm{m}_j} S_t^{sys}, \sum_{t\in \mathrm{m}_j}(S_t - F_j)L_t\right)}{Var\left(\sum_{t\in \mathrm{m}_j} S_t^{sys}\right)} \\ & - \frac{Cov(\sum_{t\in \mathrm{m}_j} S_t^{sys}, V_j^d \sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys}) + V_j^p \sum_{t\in peak_j} S_t^{sys})}{Var\left(\sum_{t\in \mathrm{m}_j} S_t^{sys}\right)} \end{aligned} \tag{2.C.3}$$

$$\begin{aligned} V_j^d = & \frac{Cov\left(\sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys}), \sum_{t\in \mathrm{m}_j}(S_t - F_j)L_t\right)}{Var\left(\sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys})\right)} \\ & - \frac{Cov\left(\sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys}), V_j^b \sum_{t\in \mathrm{m}_j} S_t^{sys} + V_j^p \sum_{t\in peak_j} S_t^{sys}\right)}{Var\left(\sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys})\right)} \end{aligned} \tag{2.C.4}$$

$$\begin{aligned} V_j^p = & \frac{Cov\left(\sum_{t\in peak_j} S_t^{sys}, \sum_{t\in \mathrm{m}_j}(S_t - F_j)L_t\right)}{Var\left(\sum_{t\in peak_j} S_t^{sys}\right)} \\ & - \frac{Cov\left(\sum_{t\in peak_j} S_t^{sys}, V_j^b \sum_{t\in \mathrm{m}_j} S_t^{sys} + V_j^d \sum_{t\in \mathrm{m}_j}(S_t - S_t^{sys})\right)}{Var\left(\sum_{t\in peak_j} S_t^{sys}\right)} \end{aligned} \tag{2.C.5}$$

With $f = \frac{|peak_j|}{|\mathrm{m}_j|}$ the equations from Model 1 simplify to

$$V_j^b = \frac{1}{|\mathrm{m}_j|}\sum_{t\in \mathrm{m}_j} \theta_t^L - fV_j^p \tag{2.C.6}$$

$$V_j^d = \frac{1}{|\mathrm{m}_j|}\sum_{t\in \mathrm{m}_j} \theta_t^L \tag{2.C.7}$$

$$V_j^p = \frac{1}{|peak_j|}\sum_{t\in peak_j} \theta_t^L - V_j^b \tag{2.C.8}$$

and for Model 2 they become

$$V_j^b = \frac{1}{|\mathrm{m}_j|}\sum_{t\in \mathrm{m}_j} \left(\theta_t^L - (F - \theta_t^S)\frac{\rho\sigma_L}{\sigma_{sys}}\right) - fV_j^p \tag{2.C.9}$$

$$V_j^d = \frac{1}{|\mathrm{m}_j|}\sum_{t\in \mathrm{m}_j} \theta_t^L \tag{2.C.10}$$

$$V_j^p = \frac{1}{|peak_j|}\sum_{t\in peak_j} \left(\theta_t^L - (F - \theta_t^S)\frac{\rho\sigma_L}{\sigma_{sys}}\right) - V_j^b. \tag{2.C.11}$$

Appendix 2.D Calibration of the Models

Let $(s_i)_{(i\in\{1,\dots,N\})}$, $(s_i^{sys})_{(i\in\{1,\dots,N\})}$ and $(l_i)_{(i\in\{1,\dots,N\})}$ denote the observed area prices, system prices and loads in 2012, where N is the total number of hours.

Model 1

Model 1 only requires estimates of average load for each month as well as for peak and off-peak hours of each month.

Model 2

We estimate σ_L, σ_{sys} and ρ using the estimators

$$\hat{\sigma}_L^2 = \frac{1}{N}\sum_{i=1}^{N}(l_i - \theta_i^L)^2 \tag{2.D.1}$$

$$\hat{\sigma}_{sys}^2 = \frac{1}{N}\sum_{i=1}^{N}(s_i^{sys} - \theta_i^{sys})^2 \tag{2.D.2}$$

$$\hat{\rho} = \frac{\sum_{i=1}^{N}(s_i^{sys} - \theta_i^{sys})(l_i - \theta_i^L)}{\sqrt{\sum_{i=1}^{N}(s_i^{sys} - \theta_i^{sys})^2 \sum_{i=1}^{N}(l_i - \theta_i^L)^2}} \tag{2.D.3}$$

Model 3

In this model we utilize that for an Ornstein-Uhlenbeck process U_t that satisfies the following equation,

$$\mathrm{d}U_t = -\kappa U_t \,\mathrm{d}t + \sigma \,\mathrm{d}Z_t \tag{2.D.4}$$

has

$$U_{t+\Delta} = e^{-\kappa\Delta}U_t + \sigma\int_t^{t+\Delta} e^{-\kappa(t-v)} \,\mathrm{d}Z_t \tag{2.D.5}$$

and thus with $t_i = t_1 + (i-1)\Delta$ for $i \in \{1,\dots,N\}$,

$$U_{t_{i+1}} = aU_{t_i} + bX_i, \quad \text{for } i \in \{1,\dots,N\} \tag{2.D.6}$$

with X_i independent and $X_i \sim \mathcal{N}(0,1)$, $a = e^{-\kappa\Delta}$ and $b^2 = \sigma^2\frac{1-e^{-2\kappa\Delta}}{2\kappa}$. The estimator for a that minimize

$$\sum_{i=1}^{N-1}(U_{t_{i+1}} - aU_{t_i})^2 \tag{2.D.7}$$

is given by

$$\hat{a} = \frac{\sum_{i=1}^{N-1} U_{t_i}U_{t_{i+1}}}{\sum_{i=1}^{N-1} U_{t_i}^2}. \tag{2.D.8}$$

Thus, κ can be estimated as

$$\hat{\kappa} = \frac{-\log(\hat{a})}{\Delta} \tag{2.D.9}$$

and b^2 can be estimated as

$$\hat{b}^2 = \frac{1}{N-1}\sum_{i=1}^{N-1}(U_{t_{i+1}} - \hat{a}U_{t_i})^2 \tag{2.D.10}$$

and thus $\hat{\sigma} = \sqrt{\frac{2\kappa}{1-e^{-2\kappa\Delta}}\hat{b}^2}$.
Finally, given two Ornstein-Uhlenbeck processes U_t and V_t that satisfy the equations,

$$\mathrm{d}U_t = -\kappa U_t \ \mathrm{d}t + \sigma \ \mathrm{d}Z_t \tag{2.D.11}$$

$$\mathrm{d}V_t = -\lambda V_t \ \mathrm{d}t + \nu \ \mathrm{d}W_t \tag{2.D.12}$$

with $\mathrm{d}W_t \ \mathrm{d}Z_t = \rho dt$ we have that

$$\begin{aligned} &Cor(U_{t_i+\Delta} - e^{-\kappa\Delta}U_{t_i}, V_{t_i+\Delta} - e^{-\lambda\Delta}V_{t_i}) = \\ &\rho\frac{2\sqrt{\kappa\lambda}}{\sqrt{1-e^{-2\kappa\Delta}}\sqrt{1-e^{-2\lambda\Delta}}}\frac{1-e^{-(\kappa+\lambda)\Delta}}{\kappa+\lambda} \end{aligned} \tag{2.D.13}$$

Thus, with the empirical correlation between the pairs of differences $U_{t_i+\Delta} - \hat{a}_U U_{t_i}$ and $V_{t_i+\Delta} - \hat{a}_V V_{t_i}$, $\hat{r}_{UV}$, given by

$$\hat{r}_{UV} = \frac{\sum_{i=1}^{N-1}(U_{t_i+\Delta} - \hat{a}_U U_{t_i})(V_{t_i+\Delta} - \hat{a}_V V_{t_i})}{\sqrt{\sum_{i=1}^{N-1}(U_{t_i+\Delta} - \hat{a}_U U_{t_i})^2 \sum_{i=1}^{N-1}(V_{t_i+\Delta} - \hat{a}_V V_{t_i})^2}} \tag{2.D.14}$$

ρ can be estimated by

$$\hat{\rho} = \hat{r}_{UV}\frac{\sqrt{1-e^{-2\hat{\kappa}\Delta}}\sqrt{1-e^{-2\hat{\lambda}\Delta}}}{2\sqrt{\hat{\kappa}\hat{\lambda}}}\frac{\hat{\kappa}+\hat{\lambda}}{1-e^{-(\hat{\kappa}+\hat{\lambda})\Delta}} \tag{2.D.15}$$

We estimate the parameters in Model 3 using these principles.

Appendix 2.E Forward Prices and Parameters

Notation date	Delivery period	base load	peak load	DK1 CfD	DK2 CfD
13. Dec 2012	Jan 2013	41.00	49.00	3.10	5.15
14. Jan 2013	Feb 2013	47.00	55.00	0.60	3.50
13. Feb 2013	Mar 2013	38.20	41.00	-0.15	1.00
13. Mar 2013	Apr 2013	39.15	40.09	-0.85	0.40
15. Apr 2013	May 2013	39.15	39.25	-1.35	-1.50
13. May 2013	Jun 2013	34.93	36.00	-1.50	0.75
13. Jun 2013	Jul 2013	29.50	30.25	4.80	5.25
15. Jul 2013	Aug 2013	35.45	38.00	2.98	4.00
13. Aug 2013	Sep 2013	35.40	38.30	3.40	4.90
13. Sep 2013	Oct 2013	38.55	43.25	1.75	4.10
14. Oct 2013	Nov 2013	42.60	47.50	-0.75	4.15
13. Nov 2013	Dec 2013	40.90	46.00	-4.10	2.50
13. Dec 2013	Jan 2014	35.40	41.00	0.00	4.50
13. Jan 2014	Feb 2014	34.50	39.60	0.05	3.80
13. Feb 2014	Mar 2014	30.05	34.10	1.00	2.65
13. Mar 2014	Apr 2014	25.15	29.25	5.75	5.90
14. Apr 2014	May 2014	24.30	27.95	6.50	6.80
13. May 2014	Jun 2014	26.35	30.35	5.60	6.75
13. Jun 2014	Jul 2014	22.20	26.15	7.40	8.25
14. Jul 2014	Aug 2014	30.75	33.68	3.50	3.00
13. Aug 2014	Sep 2014	32.85	35.80	2.60	6.45
15. Sep 2014	Oct 2014	35.75	38.90	0.47	2.90
13. Oct 2014	Nov 2014	33.95	35.95	-1.25	1.90
13. Nov 2014	Dec 2014	31.25	33.25	-1.85	1.15

Table 2.6: Notation times and forward prices

Parameters	$\hat{\sigma}^2_{sys}$	$\hat{\nu}^2$	$\hat{\sigma}^2_L$	$\hat{\rho}$
DK1 - Model 1	171.87	176.24	0.00179	-
DK1 - Model 2	171.87	176.24	0.00179	0.16559
DK2 - Model 1	171.87	157.40	0.00126	-
DK2 - Model 2	171.87	157.40	0.00126	0.27545

Table 2.7: Parameters for Model 1 and Model 2

Parameters	κ_S	κ_{sys}	κ_L	$\hat{\tilde{\sigma}}^2_S$	$\hat{\tilde{\sigma}}^2_{sys}$	$\hat{\tilde{\sigma}}^2_L$
DK1 - Model 3	0.10076	0.08604	0.09951	6.37	5.44	0.01887
DK2 - Model 3	0.14235	0.08604	0.25760	8.43	8.43	0.02553

	$\rho_{S,sys}$	$\rho_{S,L}$	$\rho_{sys,L}$
DK1 - Model 3	0.52951	0.22932	0.16559
DK2 - Model 3	0.63801	0.31171	0.27578

Table 2.8: Parameters for Model 3

Appendix 2.F Hedging Strategies

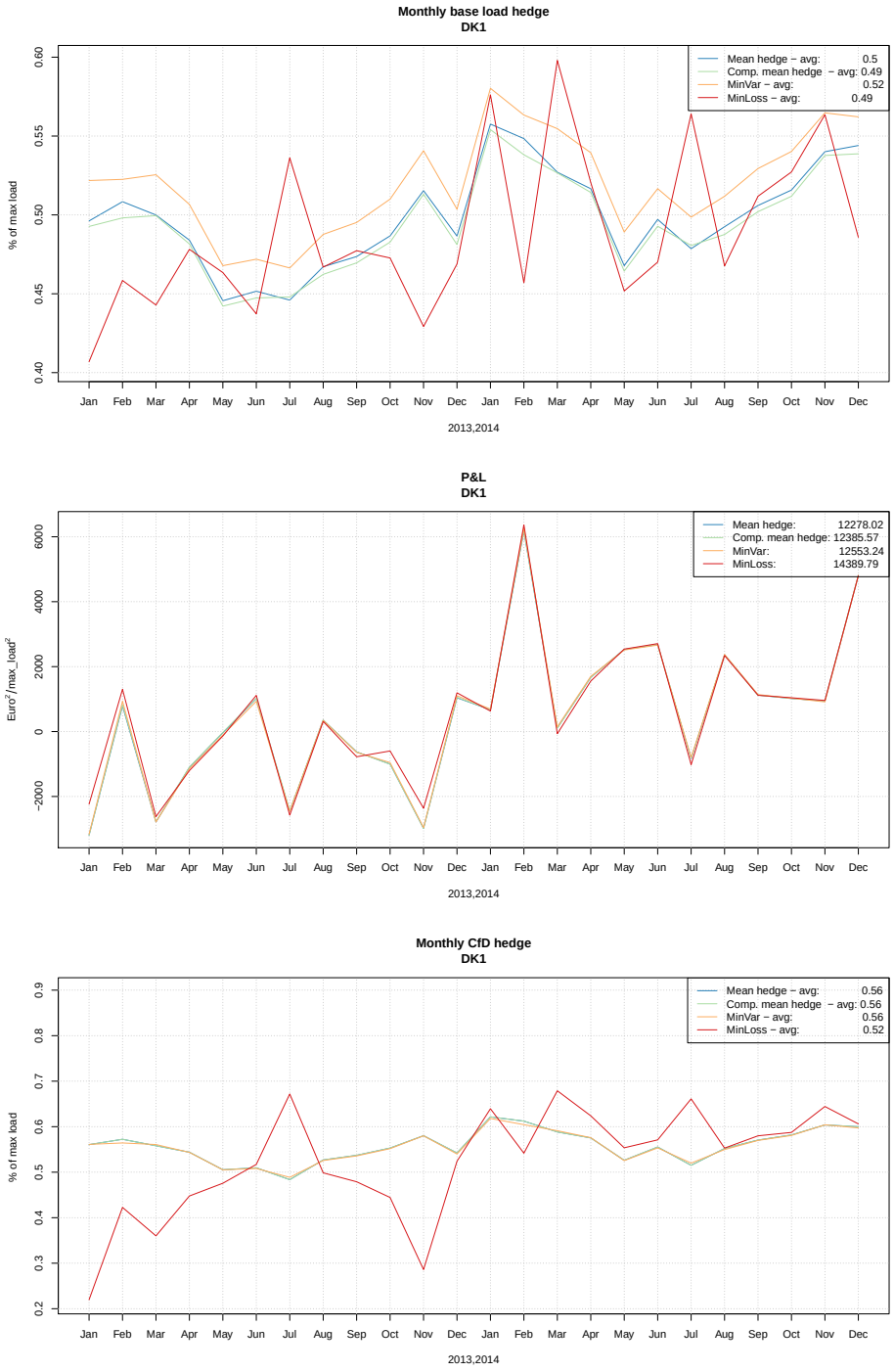


Figure 2.11: Monthly hedging volumes for East Denmark (DK1)

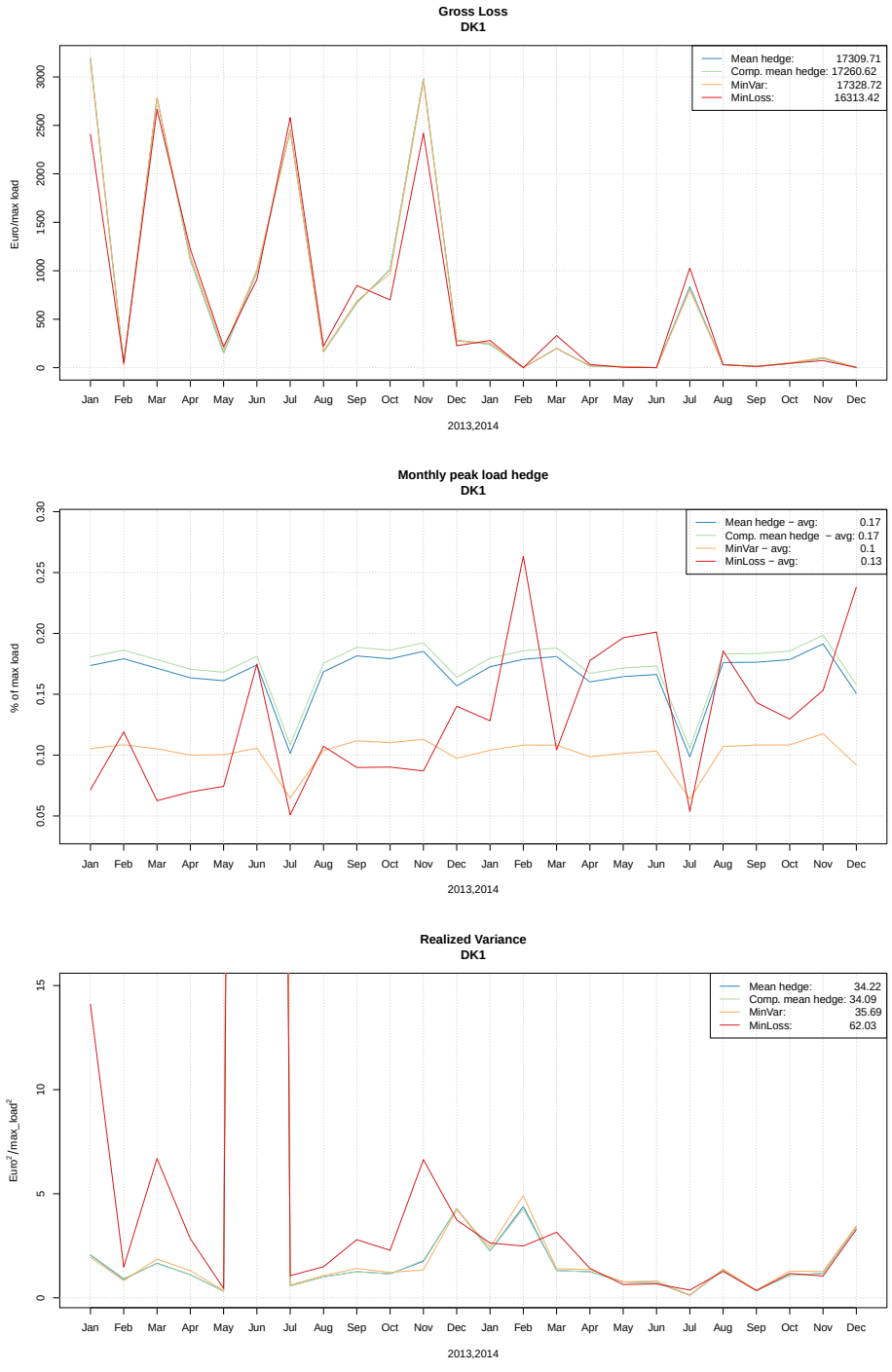


Figure 2.12: P&L, gross loss and realized variance for East Denmark (DK1)

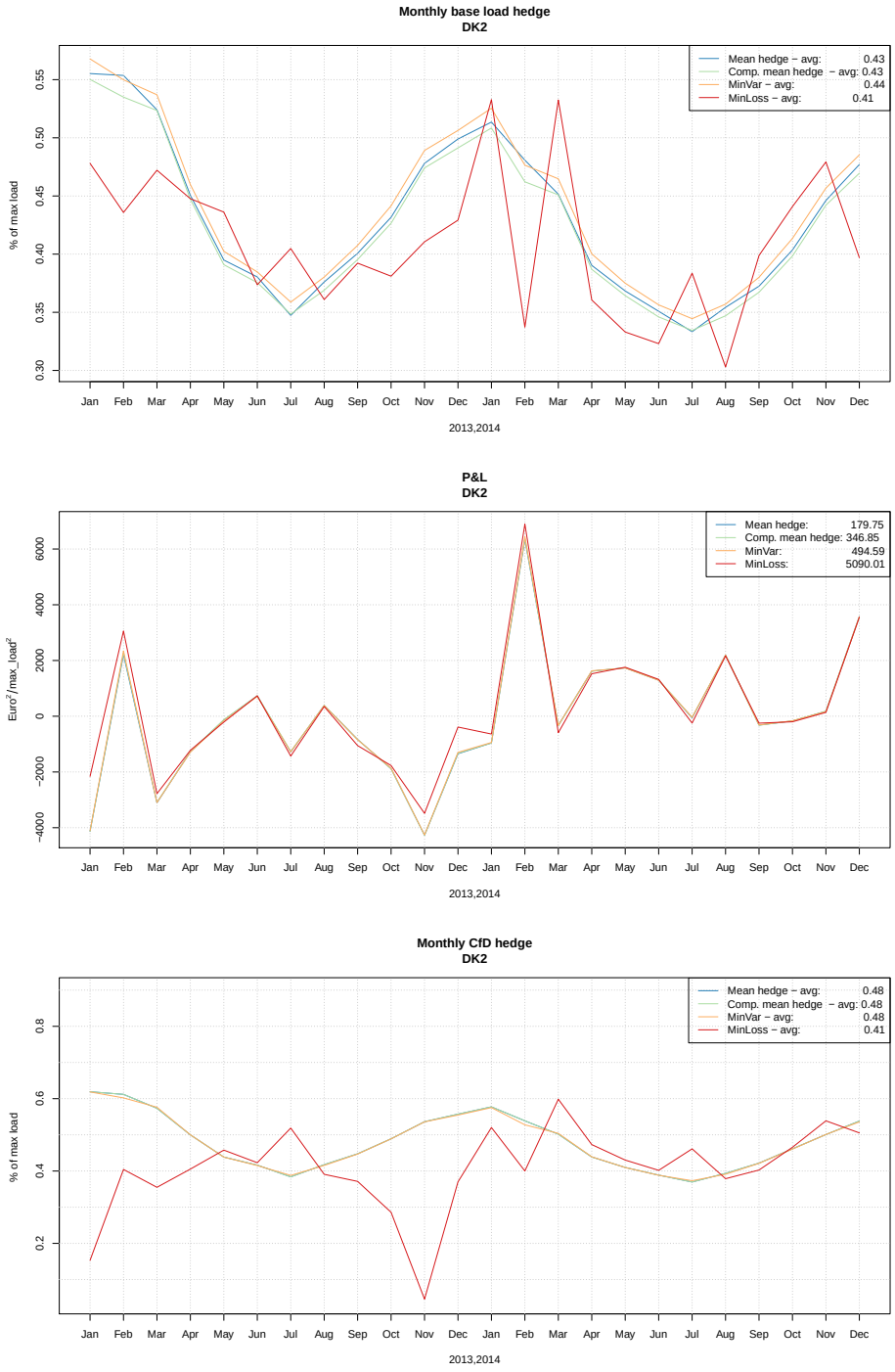


Figure 2.13: Monthly hedging volumes for West Denmark (DK2)

Figure 2.14: P&L, gross loss and realized variance for West Denmark (DK2)

3

Market Power and Investment in Electricity Generation

Abstract

In this paper, we compare investment timing and capacity choice for a strategic firm and a social planner that each have a one-time opportunity to invest in two types of electricity generation. We account for differences in operational costs across technologies, but also for the differences in operating flexibility and its impact on the optimal investment decisions. The one-time investment decision involves the choice of technology and subsequently the determination of a demand shock trigger and new capacity level. We specifically investigate how technology choice, investment trigger and optimal capacity change with changes in market ownership and the level of already installed capacity in the market.

We find that a strategic firm with market ownership tends to invest at a higher demand trigger level and lower capacity compared to the social planner. Hence, the strategic firm invests at a later date while incurring lower investment costs. Furthermore, an increased level of already installed capacity delay new investment and increases new capacity for both investors, however, base load generation is delayed more than peak load generation due to the exposure to potential low prices. Finally, we find that increased market ownership of the strategic firm delays investment and increases new capacity.

3.1 Introduction

Investment decisions in electricity markets have been a long lasting focus of researchers. In this paper, we focus on the investment decisions for different types of electricity generators. Using real options analysis and following Hagspiel et al. (2016) and Huisman and Kort (2015), we compare the investment triggers and optimal levels of investment for a strategic firm and a social planner in a simplified electricity market. In that regard, we combine investment in electricity generation and real options with special emphasis on market power issues.

The real options literature covers a vast number of references on optimal investment and capacity decisions of a firm. In the seminal works of this area, Pindyck (1988) examines the value of incremental investment and Dixit (1995) studies irreversible investment in scale economies. Newer contributions include Dangl (1999) who investigates a firm's investment timing and capacity choice when facing uncertainty of the demand shift parameter. Whereas this reference does not account for market power issues, Huisman and Kort (2015) provide a dynamic analysis of entry deterrence and accommodation strategies in a duopoly setting.

For the coverage of electricity market aspects, Aguerrevere (2003) presents a model for investment under uncertainty that includes time to build, capacity choice and flexibility in the use of installed capacity, while considering the effect of competition in the energy market. Bobtcheff (2008) likewise examines the impact of price competition in a market driven by stochastic shocks in a duopoly setting. Boomsma et al. (2012) investigate how investment decisions in renewable energy vary under support schemes that differ in their exposure to market uncertainties. Abadie and Chamorro (2014) likewise address the valuation of an operating wind farm and the option to invest in such a farm under different policy regimes.

Studies that specifically investigates the technology choice in the electricity markets include Näsäkkälä and Fleten (2005), who compute optimal building and upgrading thresholds for gas fired power plant investments and Wickart and Madlener (2007) who compare an irreversible investment in a combined heat-and-power production (cogeneration) system and a conventional heat-only generation system (steam boiler). Finally, Takashima et al. (2012) investigate how an investor makes decisions about timing, sizing, and technology choice.

We aim to extend the real options literature to allow for ownership of already installed capacity and generalize the cost structure. Contrary to assumptions of the majority of the literature, the choice of technology is not only a question of operation and investment costs. This is the case as different technologies entail different revenue streams, depending on their operational characteristics. We take the operational flexibility as well as different marginal costs into account and study the impact on optimal investment decisions. Fur-

thermore, we account for and investigate the market power that may result from owning already installed capacity.

The rest of the paper is organized as follows: In Section 3.2, we state the aim and formal model set-up of the paper. In section Section 3.3 the value of immediate investment in base load and peak load generation for the strategic firm is determined, including the value of suspending operation for a peak load plant. We find the value of suspending operation through a limit of two value matching conditions, as it allows for simpler explicit formulas for the options constants that can be generalized to other processes. In section Section 3.4 we derive conditions for the optimal capacity of immediate investment and extend the classical conditions for finite value of investment. The conditions for the optimal demand shock trigger is derived in Section 3.5 along with extended conditions on investment costs for finite value of delayed investment. In Section 3.6 the optimal trigger and capacity conditions are combined reducing the problem to one dimension for both base load and peak load, which makes the problem significantly easier to solve numerically. Finally, we show how to similarly value investment for the social planner by considering a hypothetical market for a strategic firm. In Section 3.7, we report numerical results and study the impact of production, investment and market parameters. Finally, Section 3.8 provides a brief conclusion.

3.2 Model Set-up

Our model is an extension of the models studied by Hagspiel et al. (2016) and Huisman and Kort (2015). We specifically extend their models to study investment in electricity markets, where prices can be negative and compare the impact of operational flexibility in generation. Furthermore, we extend to slightly more general cost functions, which is possible for the generation technologies we study, as the problem can reduced to a one-dimensional problem. Typically, production technologies may experience decreasing marginal investment costs due to economies of scale, which suggests convex marginal investment cost, however power plants may have increasing marginal investment costs, e.g. due to a loss of efficiency at high loads for conventional plants and site limitations for renewable plants. We further allow for the investor to own some level of installed capacity prior to the investment decision, as is often the case for electricity generation. We compare the investment decisions of a strategic firm and a hypothetical social planner and explore the effects of market power for two different power generating technologies. We model two stylized technologies and investigate how the choice of technology changes with respect to crucial model parameters.

The starting point of our model is a simplified electricity market that has a level $K \geq 0$ of installed capacity already in place. To simplify the model, we assume that the installed capacity is a base load technology. This capacity is

always active, producing electricity at full capacity and subject to a constant marginal cost of production $c > 0$. Furthermore, the strategic firm in the industry owns a fraction of the installed capacity, $A \in [0, 1]$, and has a one-time opportunity to invest in new capacity.

The sole source of uncertainty in our model is the exogenous demand shock following a *Geometric Brownian Motion*:

$$dX_t = \alpha X_t dt + \sigma X_t dW_t \tag{3.2.1}$$

where α is the drift parameter, $\sigma > 0$ is volatility parameter and W_t is a Wiener process. Using this specific stochastic process is a standard assumption in the real options literature, and allows us to derive closed-form solutions for the option values.[1] We assume that the market price of electricity fluctuates stochastically according to a linear inverse demand function, D:

$$P_t = D(X_t, Q_t) = X_t - \gamma Q_t \quad \text{with } \gamma > 0, \tag{3.2.2}$$

where Q_t is the total industry production at time t, X_t is the stochastically varying demand shock at time t and γ is the price elasticity. We do not impose any restrictions on the inverse demand function. Therefore, depending on demand and supply, market prices could become negative. Finally, we let $r > 0$ be the exogenously specified risk free rate of the market with $r > \alpha$. Note that in the case of inelastic demand, investment in new generation would have no impact on the existing assets.

We assume that the strategic firm and the social planner have two technologies to choose from: *base load* and *peak load*. To simplify our derivations and the determination of technology choice, we assume that the decision between these two mutually exclusive projects is to be made by the investor at time zero. The marginal cost of production for the new generator depends on the choice of technology and satisfies $0 < c_B < c < c_P$, where c_B and c_P are the marginal cost of production for the base load and the peak load technology, respectively. Thus, investment in base load generation entails lower marginal cost of production than the installed capacity, whereas investment in peak load generation entails higher costs. If the base load generator is chosen, the marginal cost of production for the new generator will be low, but the generator will always operate at full capacity and never shut down. If the peak load generator is chosen, however, the marginal cost of production will be high, but electricity generation can be suspended without a cost. Peak load generation will be suspended when demand is low and for simplicity we assume that it is active at full capacity when it is not suspended. The firm can invest in a level $K_{new} > 0$ of new capacity at a fixed cost of I and a variable cost of K_{new}^{λ}, where $\lambda \geq 0$. Therefore, the total investment cost equals

$$I + K_{new}^{\lambda}. \tag{3.2.3}$$

[1] In the appendix we give an approach to generalize this to other time-homogeneous diffusion processes by solving a second order differential equation.

In the following sections we explore the investment decisions of the strategic firm when choosing between the two technologies, base load and peak load. In the literature, the main distinction between base load and peak load is that peak load generators have lower investment costs and higher marginal costs compared to base load generators, see Joskow (2007). However, this distinction is insufficient to capture the value of flexibility for a peak load generator. In particular, by disregarding the value of being able to suspend operation, peak load generators will be undervalued. In the following section, we address this point in more detail. For the strategic planner we initially determine the value of immediate investment, then we derive conditions for the optimal investment capacity and finally we determine the optimal investment timing. We show in Section 3.6.4 that the investment problem for the social planner in this setup is equivalent to the investment problem for a strategic planner with altered market parameters and for this reason we only model the strategic firm.

3.3 Investment Value for the Strategic Firm

In this section, we explore the value of new generation for the strategic firm, provided a potential investment has already been undertaken and capacity installed. The profit flow $\Pi(X_t, K_{new}, c_{new})$ for the strategic firm given a level K_{new} of active new capacity, when the current demand shock is X_t and the marginal cost of production for the new technology is c_{new}, is

$$\begin{aligned}\Pi(X_t, K_{new}, c_{new}) &= D(X_t, K + K_{new})(AK + K_{new}) \\ &\quad - cAK - c_{new}K_{new} \end{aligned} \tag{3.3.1}$$

$$\begin{aligned}&= \Big(X_t - \gamma K - c)\Big)AK \\ &\quad + \Big(X_t - \gamma K_{new} - \big(c_{new} + \gamma(A+1)K\big)\Big)K_{new}.\end{aligned} \tag{3.3.2}$$

Thus, with profit level $X^{sus}(K_{new}, c_{new})$ given by

$$X^{sus}(K_{new}, c_{new}) = \gamma K_{new} + c_{new} + \gamma(A+1)K \tag{3.3.3}$$

the profit flow from generation is positive for $X_t > X^{sus}(K_{new}, c_{new})$. The profit level depends on the level of already installed and new capacity as the market price P_t is lowered depending on supply. The fraction of ownership of already installed capacity also impacts the profit level as increased supply decreases profit from already owned capacity. We note that a market share of A when already installed capacity is K corresponds to an increase in marginal cost for the new investment by $\gamma(A+1)K$.

3.3.1 No Investment

If the strategic firm decides not to invest the expected discounted payoff will be

$$\begin{aligned} &E_X\left[\int_0^\infty e^{-rt}\Pi(X_t,0,0)\ \mathrm{d}t\right] \\ &=\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK. \end{aligned} \tag{3.3.4}$$

where $E_X[\cdot]$ denotes expectation with $X_0 = X$.[2] In this case the company will neither incur the investment cost, nor the potential profit from additional generation.

3.3.2 Immediate Investment in Base Load Generation

If the strategic firm decides to invest in base load generation with capacity K_{new}, the expected discounted payoff will be

$$\begin{aligned} &E_X\left[\int_0^\infty e^{-rt}\Pi(X_t,K_{new},c_B)\ \mathrm{d}t\right]-I-K_{new}^{\lambda} \\ &=\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK \\ &+\left(\frac{X}{r-\alpha}-\frac{X_B^{sus}}{r}\right)K_{new}-I-K_{new}^{\lambda} \end{aligned} \tag{3.3.5}$$

where $X_B^{sus} = X^{sus}(K_{new}, c_B)$.

3.3.3 Immediate Investment in Peak Load Generation

If, on the other hand, the strategic firm decides to invest in peak load generation the resulting profit flow will be the highest of activating or suspending the new capacity. The additional profit flow from suspending peak generation, $\Pi^{sus}(X_t, K_{new}, c_P)$, is

$$\begin{aligned} \Pi^{sus}(X_t,K_{new},c_P) &= \Pi(X_t,0,0)-\Pi(X_t,K_{new},c_P) \\ &= \Big(X_P^{sus}-X_t\Big)K_{new} \end{aligned} \tag{3.3.6}$$

where $X_P^{sus} = X^{sus}(K_{new}, c_P)$. $\Pi^{sus}(X_t, K_{new}, c_P)$ is positive for $X_t < X_P^{sus}$, thus, the optimal suspension trigger is X_P^{sus}. Using $\Pi^{sus}(X_t, K_{new}, c_P)$ we can

[2] The assumptions $r > \alpha$ and $r > 0$ ensures that the integral of the expectation is finite and that the value can be computed by interchanging expectation and integral. The value can also be found as a solution to the second order differential equation, $\alpha X G'(X) + \frac{1}{2}\sigma^2 X^2 G''(X) + \Pi(X,0,0) - rG(X) = 0$ with the conditions $\lim_{t\to\infty} e^{-rt}E_X[G(X_t)] = 0$ and $E(\int_0^t e^{-rs}G'(X_s)\sigma X_s\ \mathrm{d}W_s) = 0$. However, this approach assumes that the value is twice differentiable in X. See Ross (2008).

formulate the expected discounted payoff from investment in peak generation as

$$E_X\left[\int_0^\infty e^{-rt}\max\left[\Pi(X_t,K_{new},c_P),\Pi(X_t,0,0)\right]\,\mathrm{d}t\right] \tag{3.3.7}$$

$$\begin{aligned}
&=E_X\left[\int_0^\infty e^{-rt}\Pi(X_t,K_{new},c_P)\,\mathrm{d}t\right]\\
&+E_X\left[\int_0^\infty e^{-rt}\max\left[0,\Pi^{sus}(X_t,K_{new},c_P)\right]\,\mathrm{d}t\right]-I-K_{new}^{\lambda}\\
&=\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK\\
&+\left(\frac{X}{r-\alpha}-\frac{X_P^{sus}}{r}\right)K_{new}+F(X,K_{new})-I-K_{new}^{\lambda},
\end{aligned} \tag{3.3.8}$$

where $F(X,K_{new})$ is the expected discounted value of suspending peak load generation with capacity K_{new} when the additional profit from peak generation becomes negative given an initial demand shock of X.[3]

3.3.4 The Value of Suspending Operation

Using X_P^{sus} we can formulate the option value of being able to suspend operation as

$$F(X,K_{new})=E_X\left[\int_0^\infty e^{-rt}1_{(X_t<X_P^{sus})}\Pi^{sus}(X_t,K_{new},c_P)\,\mathrm{d}t\right]. \tag{3.3.9}$$

Here, $1_{(X_t<X_P^{sus})}$ is the indicator function which is 1 when the demand shock is below the suspension trigger, X_P^{sus}. When $X_t\geq X_P^{sus}$ and there is no profit from suspending peak generation the indicator is 0.

As the integrand in (3.3.9) is non-negative, $F(X,K_{new})$ is always non-negative. Thus, the option to suspend always increases the expected value of generation. Vice versa, by ignoring this option, peak generation may by significantly undervalued.

We show in Lemma 3.A.2 that the value of suspension has the form

$$F(X,K_{new})=\begin{cases}
-\left(\dfrac{X}{r-\alpha}-\dfrac{X_P^{sus}}{r}\right)K_{new}\\
+\left(\dfrac{X}{X_P^{sus}}\right)^{\beta_1}B(K_{new}), & X<X_P^{sus}\\
\left(\dfrac{X}{X_P^{sus}}\right)^{\beta_2}C(K_{new}), & X\geq X_P^{sus},
\end{cases} \tag{3.3.10}$$

[3]The value is finite if $r>0$ and $r>\alpha$ as $\max[\Pi(X_t,K_{new},c_P),\Pi(X_t,0,0)]\leq\Pi(X_t,K_{new},c_P)+\Pi(X_t,0,0)$.

where for given K_{new}, $B(K_{new})$ and $C(K_{new})$ are constants. Here, $\beta_1 > 1$ and $\beta_2 < 0$ are the two solutions to the quadratic equation

$$\frac{1}{2}\sigma^2\beta(\beta-1) + \alpha\beta - r = 0, \tag{3.3.11}$$

given by

$$\beta_{1,2} = \left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}}. \tag{3.3.12}$$

By studying the value of suspension, $F(X, K_{new})$, we observe that the first term for $X < X_P^{sus}$ is the expected discounted additional value of always suspending peak generation. To understand the second term and the term for $X \geq X_P^{sus}$ we introduce the stopping time, τ_P, which is the random time it takes for the demand shock to hit X_P^{sus}. The stopping time is defined by

$$\tau_P = \inf\{t \geq 0 | X_t = X_P^{sus}\}. \tag{3.3.13}$$

The expectation of the random discount factor associated with τ_P, the expected discount factor, is given by

$$E_X\left(e^{-r\tau_P}\right) = \begin{cases} \left(\frac{X}{X_P^{sus}}\right)^{\beta_1}, & X \leq X_P^{sus} \\ \left(\frac{X}{X_P^{sus}}\right)^{\beta_2}, & X \geq X_P^{sus} \end{cases} \tag{3.3.14}$$

which we show in Lemma 3.A.4. The second term of $F(X, K_{new})$ for $X < X_P^{sus}$ and the term for $X \geq X_P^{sus}$ are both a constant times the expected discount factor. We note that the expected discount factor is a generalization of discount factors with deterministic times to discount factors with stopping times.[4]

We also show in Lemma 3.A.4 that $B(K_{new})$ is the value of starting operation and being able to resume suspension and $C(K_{new})$ is the value of suspending operation and being able to resume operation. The value of the constants can be found through value matching and smooth pasting conditions, as is typically done in the real options literature. However, to allow for sub-optimal suspension triggers and to find simpler expressions for the option constants, we show that the smooth pasting condition occurs as a limit of two value matching conditions.

3.3.5 Option Constants

To simplify notation we let $H_1(X)$ and $H_2(X)$ denote the expected discounted additional value of always suspending and never suspending given an initial

[4] We only obtain this structure of $F(X, K_{new})$ due to the Feller property of the demand shock process and the specific form of the stopping time.

demand shock of X. Thus, we can rewrite $F(X, K_{new})$ as

$$F(X, K_{new}) = \begin{cases} H_1(X) + \left(\frac{X}{X_P^{sus}}\right)^{\beta_1} B(K_{new}), & X < X_P^{sus} \\ H_2(X) + \left(\frac{X}{X_P^{sus}}\right)^{\beta_2} C(K_{new}), & X \geq X_P^{sus}. \end{cases} \tag{3.3.15}$$

Furthermore, we let $\Delta H(X)$ denote the difference, $H_1(X) - H_2(X)$. In our case $H_1(X) = \Delta H(X) = -\left(\frac{X}{r-\alpha} - \frac{X_P^{sus}}{r}\right) K_{new}$ and $H_2(X) = 0$, but we include $H_2(X)$ to better indicate the symmetry of the constants in terms of $H_1(X)$ and $H_2(X)$. In Lemma 3.A.2 we show that that $B(K_{new})$ and $C(K_{new})$ can be found through a limit using the following recursion that includes an artificial trigger X^* with $X^* < X_P^*$,

$$B(K_{new}) = -\Delta H(X_P^{sus}) + \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2} \Delta H(X^*) + \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2 - \beta_1} B(K_{new}). \tag{3.3.16}$$

Here $-\Delta H(X_P^{sus})$ is the value of starting operation at X_P^{sus} and $\left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2} \Delta H(X^*)$ is the discounted value of suspending again at the artificial trigger X^*. Finally, the term $\left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2 - \beta_1} B(K_{new})$ is the discounted value of starting operation and being able to suspend after having hit X^* and returned to X_P^{sus}. This recursion implies that

$$B(K_{new}) = \lim_{X^* \to X_P^{sus-}} \frac{-\Delta H(X_P^{sus}) + \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2} \Delta H(X^*)}{1 - \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_2 - \beta_1}} \tag{3.3.17}$$

$$= \frac{\beta_2 \Delta H(X_P^{sus}) - X_P^{sus} \Delta H'(X_P^{sus})}{\beta_1 - \beta_2} \tag{3.3.18}$$

$$= -H_1(X_P^{sus}) + \frac{\beta_1 H_1(X_P^{sus}) - \beta_2 H_2(X_P^{sus}) - X_P^{sus} \Delta H'(X_P^{sus})}{\beta_1 - \beta_2}. \tag{3.3.19}$$

Here, the limit is for X^* increasing towards X_P^{sus}. $C(K_{new})$ can be found similarly as

$$C(K_{new}) = \lim_{X^* \to X_P^{sus+}} \frac{\Delta H(X_P^{sus}) - \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_1} \Delta H(X^*)}{1 - \left(\frac{X_P^{sus}}{X^*}\right)^{\beta_1 - \beta_2}} \tag{3.3.20}$$

$$= \frac{\beta_1 \Delta H(X_P^{sus}) - X_P^{sus} \Delta H'(X_P^{sus})}{\beta_1 - \beta_2} \tag{3.3.21}$$

$$= -H_2(X_P^{sus}) + \frac{\beta_1 H_1(X_P^{sus}) - \beta_2 H_2(X_P^{sus}) - X_P^{sus} \Delta H'(X_P^{sus})}{\beta_1 - \beta_2} \tag{3.3.22}$$

with the limit being for X^* decreasing towards X_P^{sus}. We note that (3.3.18) and (3.3.21) show that the constants only depend on $\Delta H(X_P^{sus})$ and $\Delta H'(X_P^{sus})$, while (3.3.19) and (3.3.22) show that $F(X, K_{new})$ is continuous at $X = X_P^{sus}$ with

$$F(X_P^{sus}, K_{new}) = \frac{\beta_1 H_1(X_P^{sus}) - \beta_2 H_2(X_P^{sus}) - X_P^{sus}\Delta H'(X_P^{sus})}{\beta_1 - \beta_2}. \quad (3.3.23)$$

Furthermore, as we show in Lemma 3.A.3, $F(X, K_{new})$ is differentiable at $X = X_P^{sus}$ with

$$F'_X(X_P^{sus}, K_{new}) = \frac{\beta_1 H_1'(X_P^{sus}) - \beta_2 H_2'(X_P^{sus}) - \frac{\beta_1\beta_2}{X_P^{sus}}\Delta H(X_P^{sus})}{(\beta_1 - \beta_2)}. \quad (3.3.24)$$

Summing up, the additional value depends on the behavior of X_t around X_P^{sus} through the derivatives of the expected discount factors and the difference in the values of always suspending and always activating. Recall that in our case, $H_1(X) = \Delta H(X) = -\left(\frac{X}{r-\alpha} + \frac{X_P^{sus}}{r}\right) K_{new}$ and $H_2(X) = 0$. The constants are non-negative and can be simplified to

$$B(K_{new}) = \frac{1 - \beta_2}{\beta_1(r-\alpha)(\beta_1 - \beta_2)} K_{new} X_P^{sus} \quad (3.3.25)$$

$$C(K_{new}) = \frac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1 - \beta_2)} K_{new} X_P^{sus} \quad (3.3.26)$$

using (3.3.11) and $\beta_1\beta_2 = \frac{-2r}{\sigma^2}$. Hence, we can write the option value of suspension as

$$F(X, K_{new}) = \begin{cases} -\left(\dfrac{X}{r-\alpha} - \dfrac{X_P^{sus}}{r}\right) K_{new} \\ + \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_1} \dfrac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus}, & X < X_P^{sus} \\ \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_2} \dfrac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus}, & X \geq X_P^{sus}. \end{cases} \quad (3.3.27)$$

Thus, combining (3.3.8) and (3.3.27) we find that the expected discounted value of investment in peak generation for $X < X_P^{sus}$ is

$$\begin{aligned} & E_X\left[\int_0^\infty e^{-rt} \max\left[\Pi(X_t, K_{new}, c_P), \Pi(X_t, 0, 0)\right] \, \mathrm{d}t\right] - I - K_{new}^{\lambda} \\ &= \left(\frac{X}{r-\alpha} - \frac{\gamma K + c}{r}\right) AK \\ &+ \left(\frac{X}{X_P^{sus}}\right)^{\beta_1} \frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus} - I - K_{new}^{\lambda} \end{aligned} \quad (3.3.28)$$

and for $X \geq X_P^{sus}$ it is

$$\begin{aligned} & E_X\left[\int_0^\infty e^{-rt}\max\left[\Pi(X_t, K_{new}, c_P), \Pi(X_t, 0, 0)\right]\ \mathrm{d}t\right] - I - K_{new}^{\lambda} \\ &= \left(\frac{X}{r-\alpha} - \frac{\gamma K + c}{r}\right) AK + \left(\frac{X}{r-\alpha} - \frac{X_P^{sus}}{r}\right) K_{new} \\ &+ \left(\frac{X}{X_P^{sus}}\right)^{\beta_2} \frac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus} - I - K_{new}^{\lambda}. \end{aligned} \tag{3.3.29}$$

3.4 Optimal Investment Capacity

In the previous section we determined the value of new capacity at the time of investment. We proceed to determine the optimal level of new capacity, $\hat{K}_{new}$, given that the current demand shock level is X. Initially, we ensure that the value of base load generation, (3.3.5), as well as the value of peak load generation, (3.3.28) and (3.3.29), are bounded from above in K_{new} so that the problem is well-posed. Furthermore we ensure that the value is decreasing for K_{new} large such that the optimal capacity satisfies the first order condition. Note that we are not guaranteed to have solutions to the first order condition with $K_{new} > 0$. If there are no solutions to the first order conditions with $K_{new} > 0$ or the optimal expected additional payoff is negative, the optimal strategy is not to invest. To determine the optimal solution we compare the solutions to the relevant first order condition with $K_{new} > 0$ and determine the maximum by comparing the values of candidates.

3.4.1 Optimal Investment Capacity for Base Load Generation

For base load generation the first order condition of (3.3.5) with respect to K_{new} is

$$\frac{X}{r-\alpha} - \frac{\hat{X}_B^{sus} + \gamma \hat{K}_{new}}{r} - \lambda \hat{K}_{new}^{\lambda-1} = 0, \tag{3.4.1}$$

where $\hat{X}_B^{sus} = X^{sus}(\hat{K}_{new}, c_B) = \gamma K_{new} + c_B + \gamma(A+1)K$. Note that the left-hand side of (3.4.1) is negative for $\hat{K}_{new}$ large with X fixed, implying that the investment value is decreasing for $\hat{K}_{new}$ large. Thus, as (3.3.5) is continuous for $\hat{K}_{new} \geq 0$, the investment value is bounded from above and the optimal capacity satisfies the first order condition. The second order condition is

$$-2\frac{\gamma}{r} - \lambda(\lambda-1)\hat{K}_{new}^{\lambda-2} \leq 0, \tag{3.4.2}$$

which is satisfied for $\lambda \geq 1$ and $\hat{K}_{new} \geq 0$ as well as $\lambda = 0$. In this case a solution to the first order condition is a global maximum as the expected

payoff is concave in K_{new}. For $\lambda \in (0,1)$, which corresponds to decreasing marginal costs, we compare solutions to the first order conditions numerically.

3.4.2 Optimal Investment Capacity for Peak Load Generation

For the peak load plant we show in Lemma 3.A.6 that the investment value of immediate investment is bounded from above if

$$\lambda > 2 - \beta_1. \tag{3.4.3}$$

If (3.4.3) is not satisfied the expected value of suspending and being able to activate later increases faster than the cost of building as a function of new capacity. A stronger condition independent of λ is $\sigma < \sqrt{r - 2\alpha}$. If this condition holds then $\beta_1 > 2$ and (3.4.3) is satisfied. Note that as $\beta_1 > 1$, (3.4.3) also holds if $\lambda > 1$. We show in Lemma 3.A.7 that $F(X, K_{new})$ is differentiable in K_{new} for $K_{new} \in (0, \infty)$. Thus, if (3.4.3) is satisfied, candidates for the optimal new capacity of the peak load plant satisfies the first order condition of (3.3.28) or (3.3.29) with respect to K_{new},

$$\left(\frac{X}{\hat{X}_P^{sus}}\right)^{\beta_1} \frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} \left[\hat{X}_P^{sus} + (1-\beta_1)\gamma\hat{K}_{new}\right] - \lambda\hat{K}_{new}^{\lambda-1} = 0 \tag{3.4.4}$$

or

$$\begin{aligned} &\frac{X}{r-\alpha} - \frac{\hat{X}_P^{sus} + \gamma\hat{K}_{new}}{r} - \lambda\hat{K}_{new}^{\lambda-1} \\ &+ \left(\frac{X}{\hat{X}_P^{sus}}\right)^{\beta_2} \frac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} \left[\hat{X}_P^{sus} + (1-\beta_2)\gamma\hat{K}_{new}\right] = 0, \end{aligned} \tag{3.4.5}$$

where $\hat{X}_P^{sus} = X^{sus}(\hat{K}_{new}, c_P)$. Note that only solutions, $\hat{K}_{new}$, to (3.4.4) with $X < \hat{X}_P^{sus}$ and $\hat{K}_{new} > 0$ are valid candidates. Similarly, solutions to (3.4.5) are only valid if $X \geq \hat{X}_P^{sus}$ and $\hat{K}_{new} > 0$. Among valid candidates the optimal solution is the one that returns the maximal value of (3.3.28) or (3.3.29) if the maximum is positive.

3.5 Optimal Investment Timing

In the previous section we assumed that investment had to be initiated immediately, but in this section we assume that the investment can be delayed. We want to determine the optimal investment trigger, $\hat{X}^{inv}$, such that the first time the demand shock equals the investment trigger, the investment is initiated. Let τ_{inv} be the time until $X_t = X^{inv}$, defined as

$$\tau_{inv} = \inf\{t \geq 0 | X_t = X^{inv}\}. \tag{3.5.1}$$

We consider triggers with $X^{inv} \geq X$, where X is the current demand shock. In Lemma 3.A.4 we show that the value of the option to invest in new generation when the demand shock level hits an investment trigger, X^{inv}, is the expected discount factor corresponding to τ_{inv} times the expected value of the project with initial demand shock equal to the investment trigger. Thus, with $p(X_t)$ being the additional profit flow after investment given by

$$p(X_t) = \begin{cases} \Pi(X_t, K_{new}, c_B) - \Pi(X_t, 0, 0), & \text{for base load} \\ \max\left[\Pi(X_t, K_{new}, c_p) - \Pi(X_t, 0, 0), 0\right], & \text{for peak load} \end{cases} \tag{3.5.2}$$

the value of the option to invest at τ_{inv} is

$$\begin{aligned} &E_X\Big[\int_{\tau_{inv}}^{\infty} e^{-rt} p(X_t)\ \mathrm{d}t - e^{-r\tau_{inv}}(I + K_{new}^{\lambda})\Big] \\ &= E_X\left[e^{-r\tau_{inv}}\right] E_{X^{inv}}\Big[\int_0^{\infty} e^{-rt} p(X_t)\ \mathrm{d}t - I - K_{new}^{\lambda})\Big] \end{aligned} \tag{3.5.3}$$

$$= \left(\frac{X}{X^{inv}}\right)^{\beta_1} L(X^{inv}, K_{new}) \tag{3.5.4}$$

where $L(X^{inv}, K_{new})$ is the expected additional payoff from investment with initial demand shock X^{inv}. Note that $L(X^{inv}, K_{new})$ includes the loss incurred to already installed capacity and not only the value of the new investment. The optimal demand shock trigger is found by maximizing

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1} L(X^{inv}, K_{new}) \tag{3.5.5}$$

with respect to X^{inv} with $X^{inv} \geq X$. We show in Lemma 3.A.8 and Lemma 3.A.9 that the expected value of delayed investment in base load or peak load generation is bounded if

$$\beta_1 > 2 \text{ or } \lambda > \frac{\beta_1}{\beta_1 - 1}. \tag{3.5.6}$$

The first condition ensures that the discount factor decays faster than the investment value increases when waiting, while the second condition ensures that the cost of building increases faster than the value of waiting and building larger. Note that the first order condition of (3.5.5) with respect to X^{inv} implies

$$\frac{\partial}{\partial \hat{X}^{inv}} L(\hat{X}^{inv}, K_{new}) = L(\hat{X}^{inv}, K_{new}) \frac{\partial}{\partial X} \left(\frac{X}{\hat{X}^{inv}}\right)^{\beta_1}\Bigg|_{X=\hat{X}_{inv}}, \tag{3.5.7}$$

which is the smooth pasting condition at $X = \hat{X}^{inv}$. The first order condition of (3.5.5) with respect to X^{inv} can be written as

$$L(\hat{X}^{inv}, K_{new}) - \frac{\hat{X}^{inv}}{\beta_1} \frac{\partial}{\partial \hat{X}^{inv}} L(\hat{X}^{inv}, K_{new}) = 0. \tag{3.5.8}$$

3.5.1 Optimal Investment Timing for Base Load Generation

In the base load case the additional payoff from investment is

$$L(X^{inv}, K_{new}) = \left(\frac{X^{inv}}{r-\alpha} - \frac{X_B^{sus}}{r} \right) K_{new} - I - K_{new}^{\lambda}. \tag{3.5.9}$$

Thus, the first order condition for the base load case is

$$\left(\left(1 - \frac{1}{\beta_1} \right) \frac{\hat{X}^{inv}}{r-\alpha} - \frac{X_B^{sus}}{r} \right) K_{new} - I - K_{new}^{\lambda} = 0. \tag{3.5.10}$$

3.5.2 Optimal Investment Timing for Peak Load Generation

For the peak load case the additional payoff from investment is

$$L(X^{inv}, K_{new}) = \left(\frac{X^{inv}}{r-\alpha} - \frac{X_P^{sus}}{r} \right) K_{new} + F(X^{inv}, K_{new}) - I - K_{new}^{\lambda}. \tag{3.5.11}$$

Note that X^{β_1} solves (3.5.8) and hence, for $X^{inv} < X_P^{sus}$, (3.5.8) simplifies to

$$-I - K_{new}^{\lambda} = 0 \tag{3.5.12}$$

that has no solutions as $I + K_{new}^{\lambda} > 0$. This implies that it is always preferred to build when it is optimal to activate production immediately. As $F(X, K_{new})$ is differentiable in X all solutions to the first order conditions have $X^{inv} > X_P^{sus}$.[5] Hence, by combining (3.3.27), (3.5.8) and (3.5.11) we find the first order condition for the optimal investment trigger for the peak load plant to be

$$\begin{aligned} &\left(\left(1 - \frac{1}{\beta_1} \right) \frac{\hat{X}^{inv}}{r-\alpha} - \frac{X_P^{sus}}{r} \right) K_{new} \\ &+ \left(\frac{\hat{X}^{inv}}{X_P^{sus}} \right)^{\beta_2} \frac{\beta_1 - 1}{-\beta_2 (r-\alpha) \beta_1} K_{new} X_P^{sus} - I - K_{new}^{\lambda} = 0 \end{aligned} \tag{3.5.13}$$

where only candidates with $X^{inv} > X_P^{sus}$ are valid candidates. We note that the left-hand side of (3.5.8) is minus the derivative of the investment value. Thus, as the left hand side of (3.5.10) and (3.5.13) are positive for X^{inv} large it follows that if either conditions from (3.5.6) are satisfied the optimal trigger satisfies the first order conditions or we have that $\hat{X}^{inv} = X$.

[5] As $F(X, K_{new})$ is differentiable in X, it is not necessary to verity $X^{inv} = X_P^{sus}$.

3.6 Optimal Investment Value

3.6.1 Optimal Investment Value for Base Load Generation

Combining the two first order conditions for the base load plant, (3.4.1) and (3.5.10), we find the optimal investment trigger, $\hat{X}^{inv}$, and the optimal investment capacity, $\hat{K}_{new}$, satisfies

$$X^{inv} = \beta_1(r-\alpha)\left(\frac{\gamma}{r}K_{new} - \frac{I}{K_{new}} + (\lambda-1)K_{new}^{\lambda-1}\right) \tag{3.6.1}$$

and

$$\begin{aligned} &\frac{\gamma}{r}(\beta_1-2)K_{new}^2 - \frac{\gamma(A+1)K + c_B}{r}K_{new} \\ &+ (\beta_1-1)\left(\lambda - \frac{\beta_1}{\beta_1-1}\right)K_{new}^{\lambda} - \beta_1 I = 0. \end{aligned} \tag{3.6.2}$$

We note that for $\beta_1 > 2$ we have $\lambda < 2$ or $\lambda > \frac{\beta_1}{\beta_1-1}$ implying that (3.6.2) has a solution with $K_{new} > 0$. Similarly if $\lambda > \frac{\beta_1}{\beta_1-1}$ then either $\beta_1 > 2$ or $\lambda > 2$, which again implies that (3.6.2) has a solution with $K_{new} > 0$. This problem can now be solved by determining K_{new} that solves (3.6.2) and inserting in (3.6.1), reducing the dimension of the problem as (3.6.2) only contains K_{new}.

3.6.2 Optimal Investment Value for Peak Load Generation

To shorten notation for the one-dimensional condition for optimal investment and capacity for peak load generation we introduce the demand levels $\tilde{X}_1 = X_P^{sus} + (1-\beta_1)\gamma K_{new}$ and $\tilde{X}_2 = X_P^{sus} + (1-\beta_2)\gamma K_{new}$. They appear in the first order condition of the value of investment in peak generation with respect to K_{new}, (3.4.4) and (3.4.5), and are given by

$$\tilde{X}_1 = (2-\beta_1)\gamma K_{new} + c_P + \gamma(A+1)K \tag{3.6.3}$$

$$\tilde{X}_2 = (2-\beta_2)\gamma K_{new} + c_P + \gamma(A+1)K. \tag{3.6.4}$$

We determine $\hat{X}^{inv}$ as a function of $\hat{K}_{new}$ by subtracting (3.5.13) scaled by $\beta_1\tilde{X}_2/(X_P^{sus}K_{new}(\beta_1-\beta_2))$ from (3.4.5) scaled by K_{new}. Furthermore, we rewrite the first order conditions as (β_1-1)(3.5.13) -β_1(3.4.5). Thus, candidates for the optimal investment trigger, $\hat{X}^{inv}$, and optimal investment capacity, $\hat{K}_{new}$, with corresponding suspension trigger X_P^{sus} satisfy,

$$X^{inv} = \frac{(r-\alpha)\beta_1}{1-\beta_2}\left(\frac{-\beta_2}{r\beta_1}X_P^{sus} - \frac{\tilde{X}_2 I}{\tilde{X}_1 K_{new}} + \left(\lambda\frac{\beta_1-\beta_2}{\beta_1}\frac{X_P^{sus}}{\tilde{X}_1} - \frac{\tilde{X}_2}{\tilde{X}_1}\right)K_{new}^{\lambda-1}\right) \tag{3.6.5}$$

and

$$\begin{aligned} &\frac{\gamma}{r}\left(\beta_1-2\right)K_{new}^2-\frac{\gamma(A+1)K+c_P}{r}K_{new}\\ &+\left(\beta_1-1\right)\left(\lambda-\frac{\beta_1}{\beta_1-1}\right)K_{new}^{\lambda}-\beta I\\ &+\left(\frac{X^{inv}}{X_P^{sus}}\right)^{\beta_2}\frac{(1-\beta_2)(\beta_1-1)\tilde{X}_1K_{new}}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}=0 \end{aligned} \tag{3.6.6}$$

with $X^{inv} > X_P^{sus}$. The peak load investment problem can be solved by determining the range of K_{new} such that $X^{inv} > X_P^{sus}$ using (3.6.5) and proceeding to find all solutions to (3.6.6) in this range using the explicit expression for X^{inv} in (3.6.5).[6] Finally, the expected investment value for the triggers should be compared. Note that if $\tilde{X}_1 < 0$ the optimal capacity for the peak load plant is larger than for the base load plant with the same marginal cost of production and otherwise it is smaller. For $\beta_1 \neq 2$, $\tilde{X}_1 = 0$ corresponds to

$$K_{new}=\frac{\gamma(A+1)K+c_P}{\gamma(\beta_1-2)} \tag{3.6.7}$$

3.6.3 Value Functions

Using the optimal trigger and capacity when investing in base load generation results in an expected discounted payoff of

$$\begin{aligned} V_{SF}(X)=&\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK\\ &+\left(\frac{X}{\hat{X}^{inv}}\right)^{\beta_1}\left[\left(\frac{\hat{X}^{inv}}{r-\alpha}-\frac{X_B^{sus}}{r}\right)\hat{K}_{new}-I-\hat{K}_{new}^{\lambda}\right] \end{aligned} \tag{3.6.8}$$

for $t \leq \tau_{inv}$ such that $X \leq \hat{X}^{inv}$. After investment in new base load capacity the expected discounted payoff is

$$V_{SF}^{post}(X)=\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK+\left(\frac{X}{r-\alpha}-\frac{X_B^{sus}}{r}\right)\hat{K}_{new}. \tag{3.6.9}$$

Using the optimal trigger and capacity when investing in peak load generation, the expected discounted payoff is

$$\begin{aligned} V_{SF}(X)=&\left(\frac{X}{r-\alpha}-\frac{\gamma K+c}{r}\right)AK+\left(\frac{X}{\hat{X}^{inv}}\right)^{\beta_1}\left[\left(\frac{\hat{X}^{inv}}{r-\alpha}-\frac{X_P^{sus}}{r}\right)\hat{K}_{new}\right.\\ &\left.+\left(\frac{\hat{X}^{inv}}{X_P^{sus}}\right)^{\beta_2}\frac{\beta_1-1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}\hat{K}_{new}\hat{X}_P^{sus}-I-\hat{K}_{new}^{\lambda}\right], \end{aligned} \tag{3.6.10}$$

[6] If K_{new} implies $\tilde{X}_1 = 0$ we use (3.5.13) to determine X^{inv}.

for $t \leq \tau_{inv}$ such that $X \leq \hat{X}^{inv}$. After investment in peak load generation the expected discounted payoff is

$$V_{SF}^{post}(X) = \begin{cases} \left(\frac{X}{r-\alpha} - \frac{\gamma K + c}{r}\right) AK \\ + \left(\frac{X}{X_P^{sus}}\right)^{\beta_1} \frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} \hat{K}_{new} \hat{X}_P^{sus}, & X < X_P^{sus} \\ \left(\frac{X}{r-\alpha} - \frac{\gamma K + c}{r}\right) AK \\ + \left(\frac{X}{r-\alpha} - \frac{X_P^{sus}}{r}\right) \hat{K}_{new} \\ + \left(\frac{X}{X_P^{sus}}\right)^{\beta_2} \frac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} \hat{K}_{new} \hat{X}_P^{sus}, & X \geq X_P^{sus}. \end{cases} \tag{3.6.11}$$

3.6.4 Social Planner

In this subsection we investigate the social planners problem and show that it corresponds to the problem of the strategic firm. The objective of the social planner is to maximize total discounted expected social surplus, which consists of producer surplus and consumer surplus. In that regard, we use the same approach as Dixit and Pindyck (1994) to evaluate deviations from the socially optimal outcome. To determine the social surplus, we define the area under the inverse demand function for a given production level, Q_t, by:

$$U(X_t, Q_t) = \int_0^{Q_t} D(X_t, q)dq = \int_0^{Q_t} (X_t - \gamma q)dq = \left(X_t - \frac{\gamma Q_t}{2}\right) Q_t. \tag{3.6.12}$$

Then total social surplus for a given installed capacity, K, and new capacity K_{new} is:

$$\begin{aligned} S(X_t, K + K_{new}) &= U(X_t, K + K_{new}) - cK - c_{new}K_{new} && (3.6.13) \\ &= \left(X_t - \frac{\gamma}{2}(K + K_{new})\right)(K + K_{new}) \\ &\quad - cK - c_{new}K_{new} && (3.6.14) \\ &= \left(X_t - \frac{\gamma}{2}K - c\right) K \\ &\quad + \left(X_t - \frac{\gamma}{2}K_{new} - \left(c_{new} + \frac{\gamma}{2}2K\right)\right) K_{new}. && (3.6.15) \end{aligned}$$

Observe that social surplus at time t, $S(X_t, Q_t)$, replaces the profit flow of a firm. Note that this corresponds exactly to a strategic planner that owns the entire market, i.e. $\tilde{A} = 1$, where the price elasticity is $\tilde{\gamma} = \frac{\gamma}{2}$. Therefore, all equations are the same as for the strategic firm, only with half the elasticity

of the demand and full ownership of production. The reduced elasticity occurs due to the benefit from the consumers surplus that offsets some of the reduction in producer surplus.

3.7 Results

The main aim of this section is to show how changes in production, demand and market parameters affect technology choice, investment trigger and optimal capacity, with a view towards market power. We determine the value of investment and the investment strategy based on the parameters in Table 3.1 and Table 3.2 and perform sensitivity analysis in the following subsections. To summarize our findings, the strategic firm tends to invest at a higher de-

r	α	σ	γ	K	X_0	β_1	β_2
0.1	0.01	0.2	0.1	100	120	2.5	-2

Table 3.1: Market parameters.

A	I	λ	c_B	c	c_P
0.5	10000	2	45	50	55

Table 3.2: Cost and investor parameters.

mand trigger level and lower capacity compared to the social planner for both the base load and peak load investment cases. Hence, the strategic firm is expected to invest at a later date while incurring lower investment cost. With increased market share the strategic firm further delays investment and increases new capacity. For both types of investors, base load generation is preferable if either the level of already installed capacity or the volatility is low whereas a high level installed capacity or a high volatility makes peak load generation favorable.

3.7.1 Technology Choice

In our model, we assume that the investor initially choses either base load or peak load generation. The strategic firm measures the total expected value including both the option to invest and the expected profit from already installed capacity. The social planner measures the total expected social surplus including both the option to invest and the expected social surplus from already installed capacity. Figure 3.1 confirms that base load is preferable for low price volatility whereas peak load is preferable for high price volatility. For low price volatility base load generation benefits from low marginal cost of production, however, as the price volatility increases so does the value of suspension for peak load generation. By comparing the strategic firm and the

social planner we observe that the strategic firm prefers peak load generation for lower price volatility than the social planner. This is the case as the social planner effectively operates with a lower price elasticity such that base load generation is less exposed to low prices.

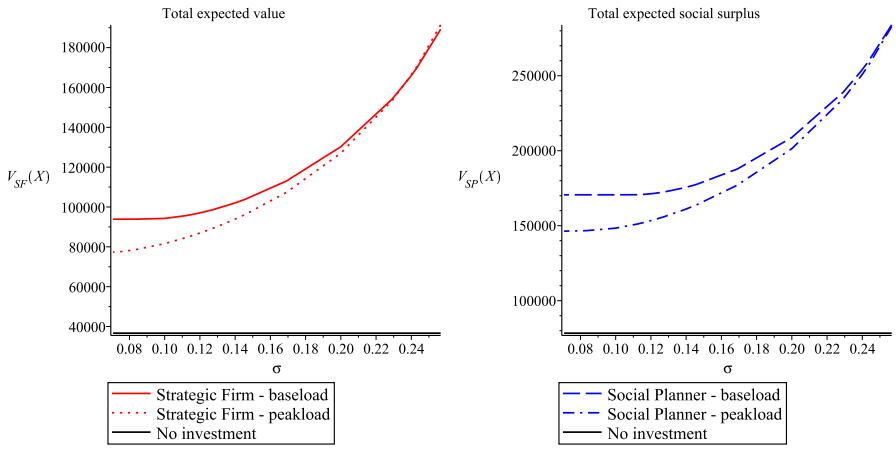


Figure 3.1: Expected value and expected social surplus changing σ.

Figure 3.2 shows that peak load generation is preferable when the marginal investment cost decreases or increases moderately, for both the strategic firm and the social planner, but when the marginal investment cost increases more rapidly base load generation becomes slightly more preferable than peak load generation. The last case is due to the lower marginal cost of production that offset the value of flexibility.

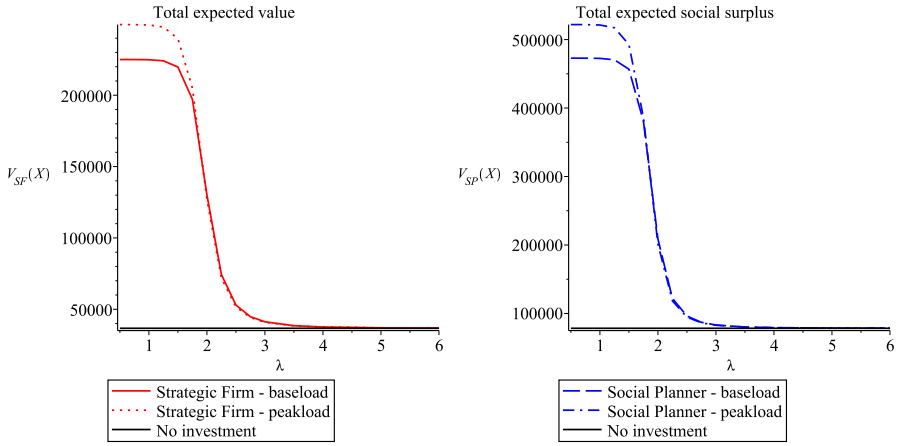


Figure 3.2: Expected value and expected social surplus for varying λ.

For a low level of already installed capacity in the market, base load is preferable for both investors, cf. Figure 3.3, but as installed capacity increases peak load generation becomes slightly more preferable due to the exposure to potential low prices for base load generation. By comparing the strategic firm and the social planner we find that the social planner prefers base load over peak load for larger K than the strategic firm.

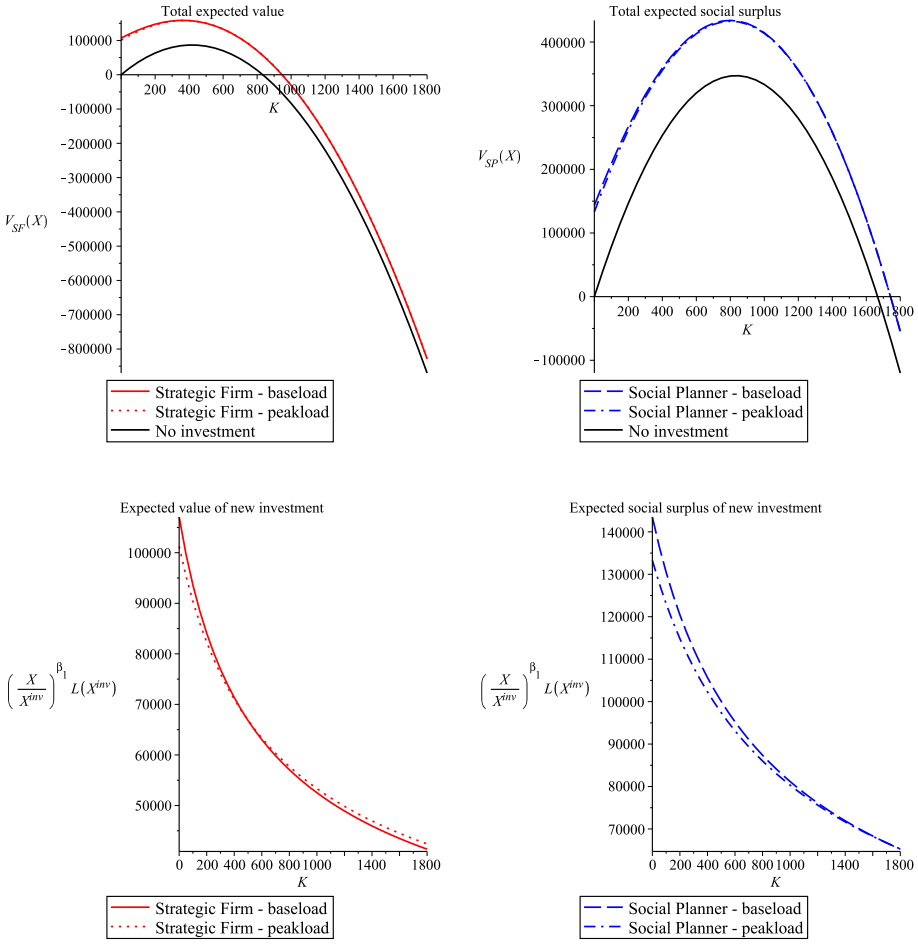


Figure 3.3: Expected value and expected social surplus for varying K.

Figure 3.4 shows that increasing the market share reduces the value of the new investment for the strategic firm, but increases the total value due to the additional profit from already installed capacity. The social planner is naturally unaffected by a change in market share. We further find that the value of base load generation is affected more by an increased market share than peak load generation, which is in line with the result on increased installed capacity. Note that increased installed capacity and increased market

share is equivalent to increased marginal cost of production for this model explaining that the effects are similar.

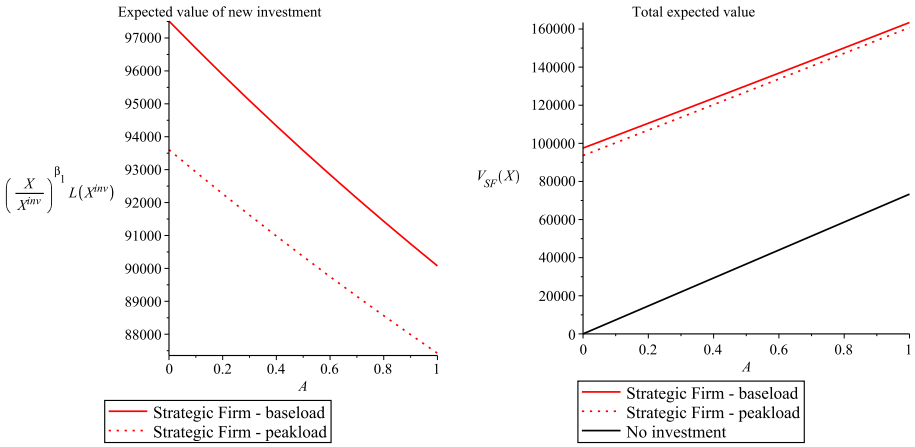


Figure 3.4: Expected value of new investment and expected discounted profit for varying A.

3.7.2 Optimal Investment Timing and Capacity

We proceed to compare the optimal investment decisions of the strategic firm and the social planner and note that an increase in investment trigger correspond to a delay of the investment. For both investors, increased price volatility delays investment and increases optimal capacity at the time of investment. These results are in line with the existing literature e.g., Dixit and Pindyck (1994), who notes that high volatility increases the optimal capacity to benefit from possible high prices.

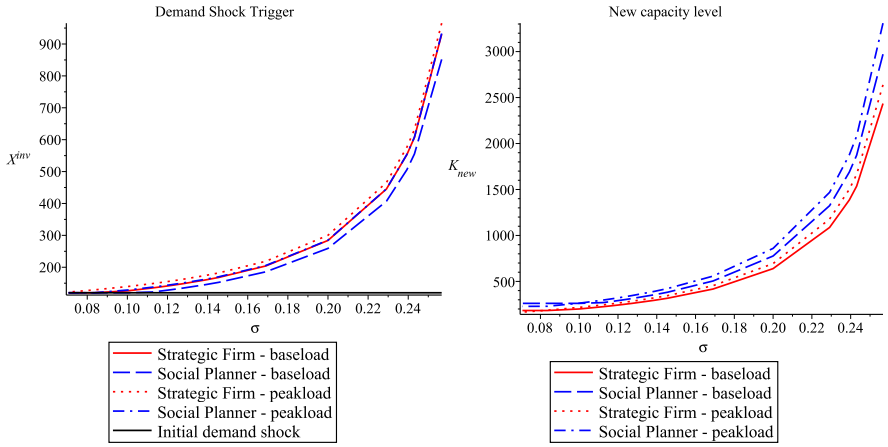


Figure 3.5: Trigger and new capacity as a function of σ.

Figure 3.6 shows how increasing λ initially increases the investment trigger and capacity, then the investment trigger decreases along with a decrease in capacity and finally the investment trigger increases again with new capacity decreasing to 0. This is the case as the balance between the value of waiting and the cost of building larger changes. For $\lambda < \frac{\beta_1}{\beta_1 - 1} = \frac{5}{3}$ investments are initiated early as the cost of building larger increases slower than the value of waiting. For $\lambda > \frac{\beta_1}{\beta_1 - 1}$ the effect starts to reverse and new capacity decreases as the cost of building larger increases faster than the value of waiting. Initially this can be offset by investing earlier, however, as it becomes very costly to invest the investment is delayed.

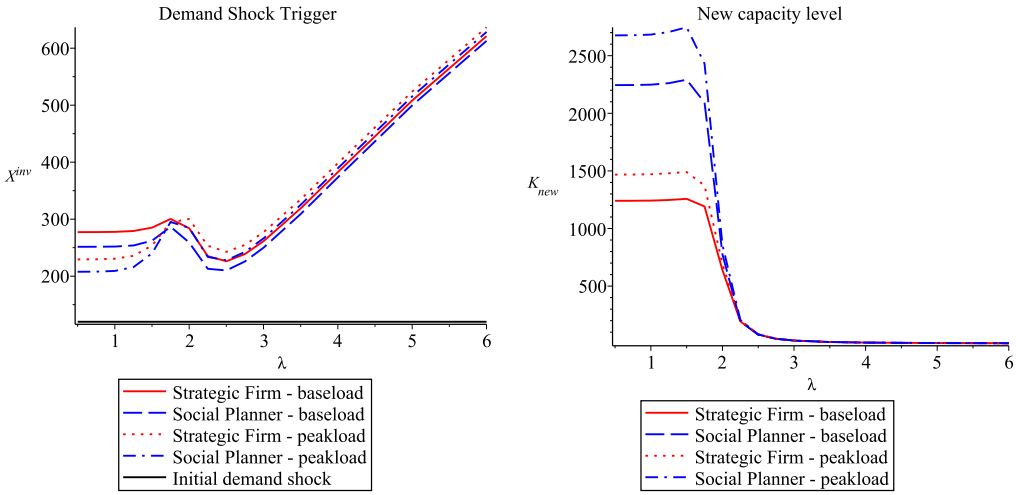


Figure 3.6: Investment triggers and capacities changing λ.

We see from Figure 3.7 that as the level of installed capacity, K, increases, the investment triggers and optimal capacities increase as well. The reason is that higher levels of installed capacity results in lower market prices prior to the new investment. Hence, the strategic firm and the social planner both delay investment and to benefit from waiting the new capacity is increased. Note also that base load generation is delayed more than peak load generation for both investors and that the effect is increased for lower values of λ.

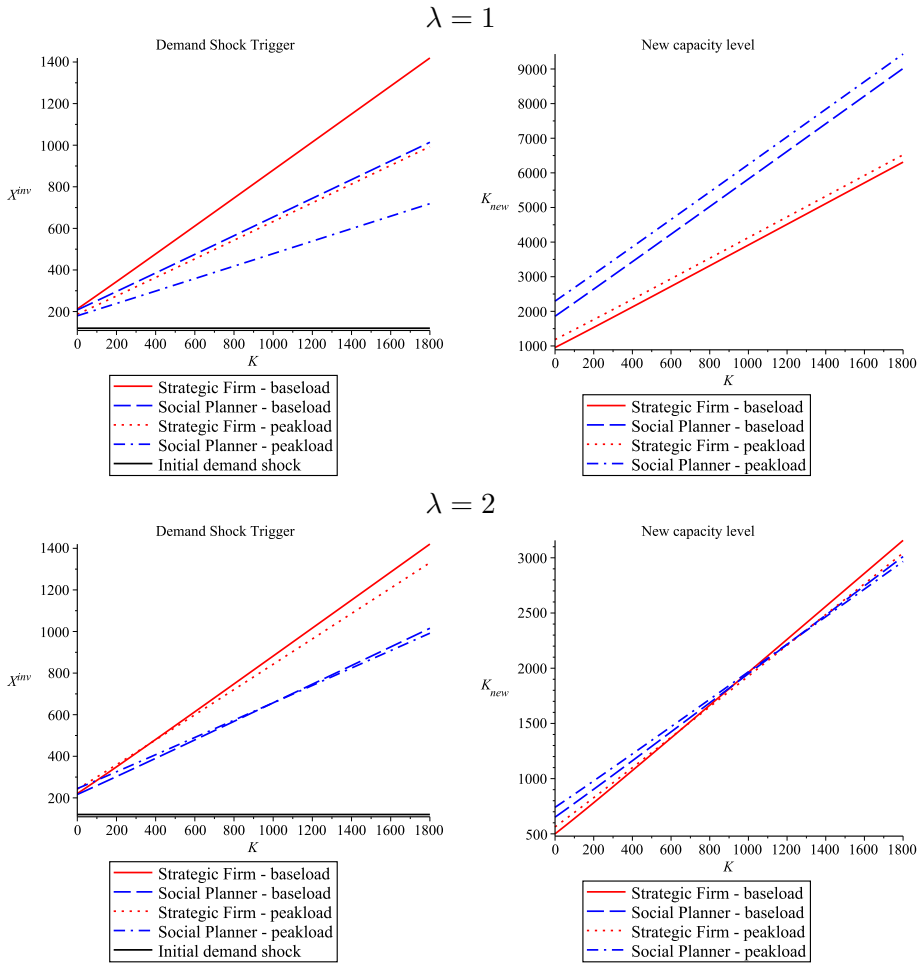


Figure 3.7: Investment Triggers and Capacities changing installed capacity, K, with default parameters where $A = 0.5$.

From Figure 3.8 we observe that the investment triggers and capacity levels for the strategic firm increase with increased ownership of installed capacity, A. This result is due to the strategic firm's reluctance to cannibalize its profits from the installed capacity. We note that for a low market share, investment are initiated almost at the same time for both the social planner and the strategic firm, however with increasing market share, the investment trigger increases along with capacity.

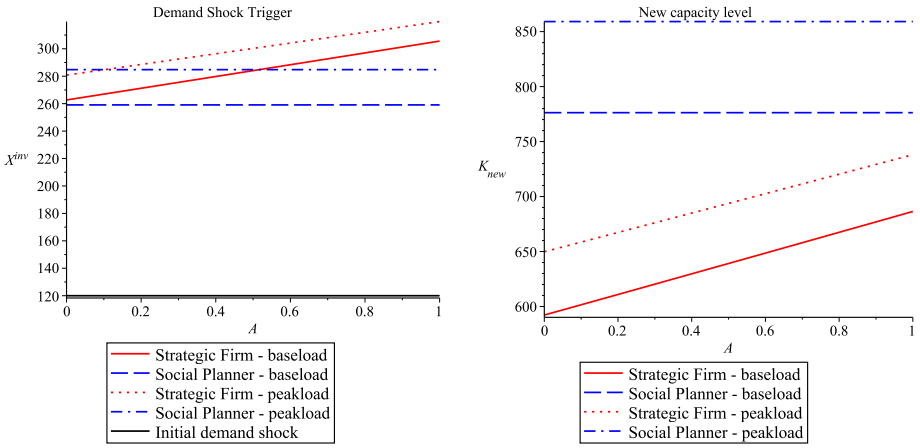


Figure 3.8: Investment Triggers and Capacities changing market share, A.

3.8 Conclusion

In this paper, we compare the investment timing and capacity choice for a strategic firm and a social planner that both have a one-time opportunity to invest in one of two types of electricity generation. The investment decision involves the choice of technology, and the determination of a demand shock trigger level and the optimal choice of capacity. We specifically investigate how the technology choice, investment trigger and optimal capacity change with changes to the demand volatility, investment cost function, level of installed capacity and investors share of the market.

For both types of investors, base load generation is preferable with a low installed capacity or low volatility whereas a high installed capacity or high volatility tends to result in peak load generation. The exercise of market power, however, makes the strategic firm increasingly affected by low prices and prefer peak load generation over base load generation for lower volatility than socially optimal.

We confirm that increasing volatility or increasing installed capacity delays investment and increases capacity for both the strategic firm and the social planner. Furthermore, base load generation is delayed more than peak load

generation when installed capacity or market share is increased. For both base load and peak load generation, nevertheless, the strategic firm tends to invest later and in less capacity compared to the strategic firm. With increased market share the strategic firm exploit its market power and further delays investment and increase capacity.

We conclude that the exercise of market power slows down investment in favor of larger and more flexible installations. The larger the market share, the more pronounced this effect becomes. To ensure adequacy for society, it is therefore of high importance to create incentive for investment. Our results show that it is socially optimal with earlier and larger installations with a greater focus on base load generation, which can be stimulated through a change of the cost structure, e.g. by imposing taxes or providing subsidies.

Appendix 3.A Lemmas

Lemma 3.A.1. *Let $(X_t)_{t\geq 0}$ be a Geometric Brownian Motion with drift parameter α and volatility parameter $\sigma > 0$ and let $r > 0$ be the risk free rate. Let τ be a stopping time defined by*

$$\tau = \inf\{t \geq 0 | X_t = X^*\}. \tag{3.A.1}$$

If $X \geq X^*$

$$E_X(e^{-r\tau}) = \left(\frac{X}{X^*}\right)^{\beta_1} \tag{3.A.2}$$

and if $X \leq X^*$

$$E_X(e^{-r\tau}) = \left(\frac{X}{X^*}\right)^{\beta_2}. \tag{3.A.3}$$

Here β_1 and β_2 are the positive and negative solution to

$$\beta(\beta-1)\frac{1}{2}\sigma^2 + \alpha\beta - r = 0 \tag{3.A.4}$$

Proof. Define $G_t = F(X_t)e^{-rt}$. Then G_t is a local martingale if the drift is 0, i.e. if

$$\alpha x F'(x) + \frac{1}{2}\sigma^2 x^2 F''(x) - rF(X) = 0. \tag{3.A.5}$$

Thus, G_t is a local martingale if $F(x) = ax^{\beta_1} + bx^{\beta_2}$ for some $a, b \in \mathbb{R}$. We note that G_t is bounded on $[0, \tau]$ if $b = 0$ and $X_0 \geq X^*$ or $a = 0$ and $X_0 \leq X^*$. Thus, $G_{t\wedge\tau}$ is a bounded local martingale and thus, a martingale. By the optional sampling theorem as $G_{t\wedge\tau}$ is a bounded martingale,

$$E_X(G_0) = E_X(G_\tau) \tag{3.A.6}$$

i.e. for $X \leq X^*$

$$aX^{\beta_1} = aE_X\left[e^{-r\tau}(X^*)^{\beta_1}\right] \tag{3.A.7}$$

hence

$$E_X\left[e^{-r\tau}\right] = \left(\frac{X}{X^*}\right)^{\beta_1}. \tag{3.A.8}$$

and for $X \geq X^*$

$$bX^{\beta_2} = bE_X\left[e^{-r\tau}(X^*)^{\beta_2}\right] \tag{3.A.9}$$

hence

$$E_X\left[e^{-r\tau}\right] = \left(\frac{X}{X^*}\right)^{\beta_2}. \tag{3.A.10}$$

□

Remark 1. *This approach can be used for diffusion processes with time homogeneous drift and diffusion coefficients.*

Lemma 3.A.2. *Let $(X_t)_{t\geq 0}$ be a Geometric Brownian Motion with drift parameter α and volatility parameter $\sigma > 0$. Let $r > 0$ be the risk free rate. Let $h_1, h_2 \in C^0(\mathbb{R}^+)$ and assume*

$$E_X\left[\int_0^\infty e^{-rt}|h_1(X_t)|\ \mathrm{d}t\right] < \infty \tag{3.A.11}$$

$$E_X\left[\int_0^\infty e^{-rt}|h_2(X_t)|\ \mathrm{d}t\right] < \infty. \tag{3.A.12}$$

Define

$$H_1(X) = E_X\left[\int_0^\infty e^{-rt}h_1(X_t)\ \mathrm{d}t\right] \tag{3.A.13}$$

$$H_2(X) = E_X\left[\int_0^\infty e^{-rt}h_2(X_t)\ \mathrm{d}t\right]. \tag{3.A.14}$$

and

$$\Delta H(X) = H_1(X) - H_2(X). \tag{3.A.15}$$

Let β_1 and β_2 be the positive and negative solutions to

$$\frac{1}{2}\sigma^2\beta(\beta-1) + \alpha\beta - r = 0 \tag{3.A.16}$$

and $X^* > 0$ be a trigger level. Then

$$\begin{aligned} &E_X\left[\int_0^\infty e^{-rt}\left(h_1(X_t)1_{(X_t<X^*)} + h_2(X_t)1_{(X_t>X^*)}\right)\ \mathrm{d}t\right] \\ &= \begin{cases} B\left(\frac{X}{X^*}\right)^{\beta_1} + H_1(X), & X \leq X^* \\ C\left(\frac{X}{X^*}\right)^{\beta_2} + H_2(X), & X > X^* \end{cases} \end{aligned} \tag{3.A.17}$$

where

$$B = \frac{1}{\beta_1-\beta_2}\Big[\beta_2\Delta H(X^*) - X^*\Delta H'(X^*)\Big] \tag{3.A.18}$$

$$= -H_1(X^*) + \frac{1}{\beta_1-\beta_2}\Big[\beta_1 H_1(X^*) - \beta_2 H_2(X^*) - X^*\Delta H'(X^*)\Big] \tag{3.A.19}$$

and

$$C = \frac{1}{\beta_1-\beta_2}\Big[\beta_1\Delta H(X^*) - X^*\Delta H'(X^*)\Big] \tag{3.A.20}$$

$$= -H_2(X^*) + \frac{1}{\beta_1-\beta_2}\Big[\beta_1 H_1(X^*) - \beta_2 H_2(X^*) - X^*\Delta H'(X^*)\Big]. \tag{3.A.21}$$

Proof. Let $X^*_{off} < X^*_{on}$ be two triggers and define the corresponding stopping times

$$\tau_{off} = \inf\{t \geq 0 | X_t = X^*_{off}\} \tag{3.A.22}$$

$$\tau_{on} = \inf\{t \geq 0 | X_t = X^*_{on}\}. \tag{3.A.23}$$

Define $F_1(X)$ for $X \leq X^*_{on}$ as value function corresponding to the accumulated profit where the instantaneous profit is $e^{-rt}h_1(X_t)$. The first time $X_t = X^*_{on}$ the instantaneous profit changes to $e^{-rt}h_2(X_t)$, the on state, and switches back when $X_t = X^*_{off}$. Similarly we define $F_2(X)$ for $X > X^*_{on}$ as the value function starting with instantaneous profit $e^{-rt}h_2(X_t)$. This can be defined recursively as

$$F_1(X) = E_X\left[\int_0^{\tau_{on}} e^{-rt}h_1(X_t)\ \mathrm{d}t + e^{-r\tau_{on}}F_2(X_{\tau_{on}})\right], \tag{3.A.24}$$

$$F_2(X) = E_X\left[\int_0^{\tau_{off}} e^{-rt}h_2(X_t)\ \mathrm{d}t + e^{-r\tau_{off}}F_1(X_{\tau_{off}})\right]. \tag{3.A.25}$$

We can rewrite $F_1(X)$ using Lemma 3.A.4 such that for $X \leq X^*_{on}$,

$$\begin{aligned} F_1(X) &= E_X\left[\int_0^{\infty} e^{-rt}h_1(X_t)\ \mathrm{d}t\right] \\ &\quad - E_X\left[\int_{\tau_{on}}^{\infty} e^{-rt}h_1(X_t)\ \mathrm{d}t - e^{-r\tau_{on}}F_2(X^*_{on})\right] \end{aligned} \tag{3.A.26}$$

$$= H_1(X) - E_X\left[e^{-r\tau_{on}}\right]\left[H_1(X^*_{on}) - F_2(X^*_{on})\right] \tag{3.A.27}$$

We can rewrite $F_2(X)$ similarly for $X^*_{off} \leq X$,

$$F_2(X) = H_2(X) - E_X\left[e^{-r\tau_{off}}\right]\left[H_2(X^*_{off}) - F_1(X^*_{off})\right]. \tag{3.A.28}$$

(3.A.28) holds for $X = X^*_{on}$ as we assumed $X^*_{off} < X^*_{on}$ so

$$F_2(X^*_{on}) = H_2(X^*_{on}) - E_{X^*_{on}}\left[e^{-r\tau_{off}}\right]\left[H_2(X^*_{off}) - F_1(X^*_{off})\right]. \tag{3.A.29}$$

Thus, by inserting (3.A.29) in (3.A.27) we obtain for $X \leq X^*_{on}$ that

$$\begin{aligned} F_1(X) =& H_1(X) - E_X\left[e^{-r\tau_{on}}\right]\left(H_1(X^*_{on}) - H_2(X^*_{on})\right) \\ & - E_X\left[e^{-r\tau_{on}}\right]E_{X^*_{on}}\left[e^{-r\tau_{off}}\right]\left(H_2(X^*_{off}) - F_1(X^*_{off})\right). \end{aligned} \tag{3.A.30}$$

Evaluating (3.A.30) at $X = X^*_{off}$ we can isolate $F_1(X^*_{off})$ to obtain

$$\begin{aligned} F_1(X^*_{off}) =& H_1(X^*_{off}) - \Delta H(X^*_{off}) \\ & + \frac{\Delta H(X^*_{off}) - E_{X^*_{of}}\left[e^{-r\tau_{on}}\right]\Delta H(X^*_{on})}{1 - E_{X^*_{off}}\left[e^{-r\tau_{on}}\right]E_{X^*_{on}}\left[e^{-r\tau_{off}}\right]} \end{aligned} \tag{3.A.31}$$

We can determine the value of $F_1(X^*_{off})$ for $X^*_{on} \searrow X^*_{off}$, which corresponds to a single trigger, by L'Hopitals rule and differentiating the numerator and denominator with respect to X^*_{on}. For a Geometric Brownian motion we have by Lemma 3.A.1 that for $X \leq X^*_{on}$

$$E_X(e^{-r\tau_{on}}) = \left(\frac{X}{X^*_{on}}\right)^{\beta_1} \tag{3.A.32}$$

and for $X \geq X^*_{off}$

$$E_X(e^{-r\tau_{off}}) = \left(\frac{X}{X^*_{off}}\right)^{\beta_2}. \tag{3.A.33}$$

Thus,

$$\begin{aligned} F_1(X^*_{off}) =& H_1(X^*_{off}) - \Delta H(X^*_{off}) \\ &+ \frac{\Delta H(X^*_{off}) - \left(\frac{X^*_{off}}{X^*_{on}}\right)^{\beta_1} \Delta H(X^*_{on})}{1 - \left(\frac{X^*_{off}}{X^*_{on}}\right)^{\beta_1 - \beta_2}} \end{aligned} \tag{3.A.34}$$

Taking limits and using L'Hopitals rule and formulating the result in terms of the single trigger X^* we get

$$F_1(X^*) = H_1(X^*) - \Delta H(X^*) + \frac{\frac{\beta_1}{X^*}\Delta H(X^*) - \Delta H'(X^*)}{\frac{\beta_1 - \beta_2}{X^*}} \tag{3.A.35}$$

$$= H_1(X^*) + \frac{1}{\beta_1 - \beta_2}\Big(\beta_2 \Delta H(X^*) - X^* \Delta H(X^*)\Big). \tag{3.A.36}$$

As $F_1(X^*)$ and $F_2(X^*)$ are symmetric we obtain

$$\begin{aligned} F_2(X^*_{on}) =& H_2(X^*_{on}) + \Delta H(X^*_{on}) \\ &+ \frac{-\Delta H(X^*_{on}) + \left(\frac{X^*_{on}}{X^*_{off}}\right)^{\beta_2} \Delta H(X^*_{off})}{1 - \left(\frac{X^*_{on}}{X^*_{off}}\right)^{\beta_2 - \beta_1}}. \end{aligned} \tag{3.A.37}$$

For $X^*_{off} \nearrow X^*_{on}$ we obtain

$$\begin{aligned} F_2(X^*) =& H_2(X^*) + \Delta H(X^*) \\ &+ \frac{-\frac{\beta_2}{X^*}\Delta H(X^*) + \Delta H'(X^*)}{-\frac{\beta_1 - \beta_2}{X^*}} \end{aligned} \tag{3.A.38}$$

$$= H_2(X^*) + \frac{1}{\beta_1 - \beta_2}\Big(\beta_1 \Delta H(X^*) - X^* \Delta H'(X^*)\Big). \tag{3.A.39}$$

Now with $X^*_{off} = X^*_{on} = X^*$ it follows that

$$F_1(X) = E_X\left[e^{-r\tau_{on}}\right]\Big(F_2(X^*) - H_1(X^*)\Big) + H_1(X) \tag{3.A.40}$$

$$= B\left(\frac{X}{X^*}\right)^{\beta_1} + H_1(X) \tag{3.A.41}$$

$$F_2(X) = E_X\left[e^{-r\tau_{off}}\right]\Big(F_1(X^*) - H_2(X^*)\Big) + H_2(X) \tag{3.A.42}$$

$$= C\left(\frac{X}{X^*}\right)^{\beta_2} + H_2(X). \tag{3.A.43}$$

As we determined $F_1(X^*)$ and $F_2(X^*)$ in (3.A.36) and (3.A.39) (3.A.19) and (3.A.21) follows by comparing constants. □

Note that we can write

$$B = -H_1(X^*) + \frac{1}{\beta_1 - \beta_2}(\beta_2 H_2(X^*) + \beta_1 H_1(X^*) - X^*\Delta H(X^*)) \tag{3.A.44}$$

$$C = -H_2(X^*) + \frac{1}{\beta_1 - \beta_2}(\beta_2 H_2(X^*) + \beta_1 H_1(X^*) - X^*\Delta H(X^*)) \tag{3.A.45}$$

Which shows that

$$F_1(X^*) = F_2(X^*) = \frac{1}{\beta_1 - \beta_2}(\beta_2 H_2(X^*) + \beta_1 H_1(X^*) - X^*\Delta H(X^*)) \tag{3.A.46}$$

Lemma 3.A.3. *Define* $F : \mathbb{R} \to \mathbb{R}$ *by*

$$F(X) = \begin{cases} H_1(X) + \left(\frac{X}{X^*}\right)^{\beta_1} B & X \le X^* \\ H_2(X) + \left(\frac{X}{X^*}\right)^{\beta_2} C & X > X^* \end{cases} \tag{3.A.47}$$

where B *and* C *are given as in Lemma 3.A.2. Then* F *is differentiable at* $X = X^*$ *with*

$$F'(X^*) = \frac{\beta_1 H_1'(X^*) - \beta_2 H_2'(X^*) - \frac{\beta_1\beta_2}{X^*}\Delta H(X^*)}{\beta_1 - \beta_2}. \tag{3.A.48}$$

Proof. The claim follows as

$$\frac{\partial}{\partial X}\left(H_1(X) + \left(\frac{X}{X^*}\right)^{\beta_1} B(K_{new})\right)_{X=X^*} \tag{3.A.49}$$

$$= H_1'(X^*) + \frac{\beta_1}{X^*}\frac{1}{\beta_1 - \beta_2}\Big[\beta_2\Delta H(X^*) - X^*\Delta H'(X^*)\Big]$$

$$= \frac{(\beta_1 - \beta_2)H_1'(X^*) + \frac{\beta_1\beta_2}{X^*}\Delta H(X^*) - \beta_1\Delta H'(X^*)}{(\beta_1 - \beta_2)} \tag{3.A.50}$$

$$= \frac{\beta_1 H_2'(X^*) - \beta_2 H_1'(X^*) + \frac{\beta_1\beta_2}{X^*}\Delta H(X^*)}{(\beta_1 - \beta_2)} \tag{3.A.51}$$

and

$$\begin{aligned}
&\frac{\partial}{\partial X}\left(H_2(X)+\left(\frac{X}{X^*}\right)^{\beta_2}C(K_{new})\right)_{X=X^*} && \text{(3.A.52)}\\
&=H_2'(X^*)+\frac{\beta_2}{X^*}\frac{1}{\beta_1-\beta_2}\left[\beta_1\Delta H(X^*)-X^*\Delta H'(X^*)\right]\\
&=\frac{(\beta_1-\beta_2)H_2'(X^*)+\frac{\beta_1\beta_2}{X^*}\Delta H(X^*)-\beta_2\Delta H'(X^*)}{(\beta_1-\beta_2)} && \text{(3.A.53)}\\
&=\frac{\beta_1H_2'(X^*)-\beta_2H_1'(X^*)+\frac{\beta_1\beta_2}{X^*}\Delta H(X^*)}{(\beta_1-\beta_2)}. && \text{(3.A.54)}
\end{aligned}$$

□

Lemma 3.A.4. *Let $(X_t)_{t\geq 0}$ be an Itô diffusion process with $X_0 < X^*$ for some $X^* \in \mathbb{R}$ and let τ be a stopping time defined by*

$$\tau = \inf\{t \geq 0 | X_t \geq X^*\}. \tag{3.A.55}$$

Let $r > 0$ be the risk free rate, $g \in C^0(\mathbb{R})$ and $S \in \mathbb{R}$ and assume

$$E_{X^*}\left[\int_0^\infty e^{-rt}|g(X_t)|\ \mathrm{d}t + S\right] < \infty. \tag{3.A.56}$$

Then

$$\begin{aligned}
&E_X\left[\int_\tau^\infty e^{-rt}g(X_t)\ \mathrm{d}t + e^{-r\tau}S\right]\\
=&E_X\left[e^{-r\tau}\right]E_{X^*}\left[\int_0^\infty e^{-rt}g(X_t)\ \mathrm{d}t + S\right]
\end{aligned} \tag{3.A.57}$$

Proof. Let $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\geq 0}, P)$ be a stochastic basis. To prove the lemma we define $\mathbb{F}_\tau$, the σ-algebra consisting of all $A \in \mathbb{F}$ such that $A \cap (\tau \leq t) \in \mathbb{F}_t$. Then

$$\begin{aligned}
&E_X\left[\int_\tau^\infty e^{-rt}g(X_t)\ \mathrm{d}t + e^{-r\tau}S\right] && \text{(3.A.58)}\\
=&E_X\left[e^{-r\tau}\left(\int_\tau^\infty e^{-r(t-\tau)}g(X_t)\ \mathrm{d}t + S\right)\right] && \text{(3.A.59)}\\
=&E_X\left[e^{-r\tau}\left(\int_0^\infty e^{-rt}g(X_{t+\tau})\ \mathrm{d}t + S\right)\right] && \text{(3.A.60)}\\
=&E_X\left[E_X\left[e^{-r\tau}\int_0^\infty e^{-rt}g(X_{t+\tau})\ \mathrm{d}t + S|\mathbb{F}_\tau\right]\right] && \text{(3.A.61)}\\
=&E_X\left[e^{-r\tau}E_X\left[\int_0^\infty e^{-rt}g(X_{t+\tau})\ \mathrm{d}t + S|\mathbb{F}_\tau\right]\right] && \text{(3.A.62)}\\
=&E_X\left[e^{-r\tau}E_{X_\tau}\left[\int_0^\infty e^{-rt}g(X_t)\ \mathrm{d}t + S\right]\right] && \text{(3.A.63)}\\
=&E_X\left[e^{-r\tau}\right]E_{X^*}\left[\int_0^\infty e^{-rt}g(X_t)\ \mathrm{d}t + S\right] && \text{(3.A.64)}
\end{aligned}$$

where we used that X_t is an Itô diffusion process and hence a Feller process such that for bounded measurable f and S and T stopping times (see Paulsen (1996) theorem 6.4)

$$E(f(X_T)|F_S) = E_{X_S}(f(X_{T-S})). \tag{3.A.65}$$

If the process was Markov and not Feller, the above equality would only hold for S and T fixed times. Note that for $\tau = \infty$ the discount factor is 0 as $r > 0$ so the expression is well-defined even though X_∞ is not defined. □

Lemma 3.A.5. *Assume $r > 0$, $r - \alpha > 0$ then*

$$\frac{\partial}{\partial X} F(X, K_{new}) < 0. \tag{3.A.66}$$

Proof. For $X \geq X_P^{sus}$ we have $C(K_{new}) > 0$ and $\beta_2 < 0$ so $\frac{\partial}{\partial X} F(X, K_{new}) < 0$. For $X < X_P^{sus}$ we have that

$$\begin{aligned} F(X, K_{new}) = &\left(\frac{X_P^{sus}}{r} - \frac{X}{r-\alpha} \right) K_{new}, \\ &+ \left(\frac{X}{X_P^{sus}} \right)^{\beta_1} \frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus} \end{aligned} \tag{3.A.67}$$

hence

$$\frac{\partial}{\partial X} F(X, K_{new}) = -\frac{K_{new}}{r-\alpha} + \left(\frac{X}{X_P^{sus}} \right)^{\beta_1 - 1} \frac{1-\beta_2}{(r-\alpha)(\beta_1-\beta_2)} K_{new} \tag{3.A.68}$$

$$\leq -\frac{K_{new}}{r-\alpha} + \frac{1-\beta_2}{(r-\alpha)(\beta_1-\beta_2)} K_{new} \tag{3.A.69}$$

$$= K_{new} \left(\frac{-(\beta_1-\beta_2) + 1 - \beta_2}{(r-\alpha)(\beta_1-\beta_2)} \right) \tag{3.A.70}$$

$$= K_{new} \left(\frac{1-\beta_1}{(r-\alpha)(\beta_1-\beta_2)} \right) < 0 \tag{3.A.71}$$

as $\beta_1 > 1$. □

Lemma 3.A.6. *For fixed X, the expected discounted value of immediate investment in peak load generation is bounded from above for $K_{new} \geq 0$ if $\lambda > 2 - \beta_1$.*

Proof. For K_{new} large we have that $X < X_P^{sus}$ and hence the value of investment is

$$\left(\frac{X}{r-\alpha} - \frac{\gamma K + c}{r} \right) AK - I - K_{new}^{\lambda} + \left(\frac{X}{X_P^{sus}} \right)^{\beta_1} B(K_{new}). \tag{3.A.72}$$

X_P^{sus} increases linearly in K_{new} and $B(K_{new})$ is quadratic in K_{new}. Thus, $\left(\frac{X}{X_P^{sus}}\right)^{\beta_1} B(K_{new})$ increases as $K_{new}^{2-\beta_1}$. If $\lambda > 2 - \beta_1$ the cost dominates for K_{new} large, which shows that the value is bounded from above for K_{new} large. We can reformulate $F(X, K_{new})$ as

$$F(X,K_{new}) = \begin{cases} \left(\dfrac{X_P^{sus}}{r} - \dfrac{X}{r-\alpha}\right) K_{new}, & \\ + \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_1} B(K_{new}) & K_{new} > -K - AK + \dfrac{X - c_P}{\gamma} \\ \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_2} C(K_{new}), & K_{new} \leq -K - AK + \dfrac{X - c_P}{\gamma}, \end{cases} \tag{3.A.73}$$

and note that $F(X, K_{new})$ is continuous at $K_{new} = -K - AK + \frac{X - c_P}{\gamma}$ and thus for $K_{new} \in [0, M]$ for M large and hence bounded. Thus, the value of investment is bounded from above. □

Lemma 3.A.7. *For fixed X,*

$$F(X,K_{new}) = \begin{cases} -\left(\dfrac{X}{r-\alpha} - \dfrac{X_P^{sus}}{r}\right) K_{new}, & \\ + \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_1} \dfrac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus}, & X < X_P^{sus} \\ \left(\dfrac{X}{X_P^{sus}}\right)^{\beta_2} \dfrac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} K_{new} X_P^{sus}, & X \geq X_P^{sus}. \end{cases} \tag{3.A.74}$$

is differentiable in K_{new} for $K_{new} > 0$.

Proof. We have $X_P^{sus} = \gamma AK + \gamma(K + K_{new}) + c_P > 0$ for $K_{new} > 0$ so for $K_{new} \neq -K - AK + \frac{X - c_P}{\gamma}$, $F(X, K_{new})$ is differentiable in K_{new}. Thus, we only have to consider $K_{new} = -K - AK + \frac{X - c_P}{\gamma}$ or equivalently $X_P^{sus} = X$. By scaling the left-hand side of 3.4.4 and 3.4.5 with K_{new}, we note that their difference at $X_P^{sus} = X$ is

$$\begin{aligned} &\frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)} \left[X + (1-\beta_1)\gamma K_{new}\right] K_{new} \\ &- \left(\frac{X}{r-\alpha} + \frac{X + \gamma K_{new}}{r}\right) K_{new} \\ &- \frac{\beta_1 - 1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)} \left[X + (1-\beta_2)\gamma K_{new}\right] K_{new}. \end{aligned} \tag{3.A.75}$$

Now as F is continuous in X at $X = X_P^{sus}$, (3.A.75) simplifies to

$$\begin{aligned}&\Big[-\frac{1}{r}+\frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)}(1-\beta_1)\\&-\frac{\beta_1-1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}(1-\beta_2)\Big]\gamma K_{new}^2 \qquad (3.A.76)\\=&\Big[-\frac{1}{r}+\frac{\beta_1\beta_2-\beta_1-\beta_2+1}{\beta_1\beta_2(\alpha-r)}\Big]\gamma K_{new}^2\\=&0 \qquad (3.A.77)\end{aligned}$$

where we used that $\beta_1\beta_2 = -\frac{2r}{\sigma^2}$ and $\beta_1 + \beta_2 = 1 - \frac{2\alpha}{\sigma^2}$. □

Lemma 3.A.8. *For $\beta_1 > 2$ or $\lambda > \frac{\beta_1}{\beta_1 - 1}$ the expected discounted value of delayed investment in base load generation is bounded from above.*

Proof. We have to show that for any new capacity and investment trigger then

$$\begin{aligned}&L(X^{inv}, K_{new}) =\\&\left(\frac{X}{X^{inv}}\right)^{\beta_1}\left[\left(\frac{X^{inv}}{r-\alpha}-\frac{\gamma AK+\gamma(K+K_{new}(X^{inv}))+c_B}{r}\right)K_{new}(X^{inv})\right.\\&\qquad\left.-I-K_{new}(X^{inv})^{\lambda}\right]\end{aligned} \qquad (3.A.78)$$

is bounded from above. Assume for contradiction that $\beta > 2$ and there exists a pair (X^{inv}, K_{new}) such that the value is larger than M for any $M > 0$. Then the only positive component is also greater than M, i.e.

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1}\frac{X^{inv}K_{new}}{r-\alpha} > M, \qquad (3.A.79)$$

and hence

$$K_{new} > M\frac{r-\alpha}{X^2}X^{inv}\left(\frac{X^{inv}}{X}\right)^{\beta_1-2}. \qquad (3.A.80)$$

This implies as $X^{inv} \geq X$ and $\beta_1 - 2 > 0$ that $\left(\frac{X^{inv}}{X}\right)^{\beta_1-2} \geq 1$ so that

$$K_{new} > M\frac{r-\alpha}{X^2}X^{inv}. \qquad (3.A.81)$$

Hence $-K_{new}(X_{inv}) < -M\frac{r-\alpha}{X^2}X^{inv}$, but then for $M = \frac{rX^2}{\gamma(r-\alpha)^2}$

$$\frac{X^{inv}}{r-\alpha}-\frac{\gamma AK+\gamma(K+K_{new})+c_B}{r} \leq -\frac{\gamma AK+\gamma K+c_B}{r} < 0 \qquad (3.A.82)$$

which contradicts the assumption that the value was larger than M.
Assume again for contradiction that $\lambda > \frac{\beta_1}{\beta_1-1}$ and that there exists pair (X^{inv}, K_{new}) such that the expected discounted value of investment is larger than M for any $M > 0$. As $\lambda > \frac{\beta_1}{\beta_1-1}$ it follows that $(\lambda-1)(\beta-1) > 1$ and $\lambda > \frac{\beta_1}{\beta_1-1} = 1 + \frac{1}{\beta_1-1} > 1$. The only positive component is also greater than M, i.e.

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1} \frac{X^{inv}K_{new}}{r-\alpha} > M, \tag{3.A.83}$$

and hence

$$K_{new} > M\frac{r-\alpha}{X}\left(\frac{X^{inv}}{X}\right)^{\beta_1-1}. \tag{3.A.84}$$

As $\lambda > 1$, $(\lambda-1)(\beta-1) > 1$ and $X^{inv} > X$ we have that

$$K_{new}^{\lambda-1} > \left(M\frac{r-\alpha}{X}\right)^{\lambda-1}\left(\frac{X^{inv}}{X}\right)^{(\beta_1-1)(\lambda-1)} \tag{3.A.85}$$

$$> \left(M\frac{r-\alpha}{X}\right)^{\lambda-1}\left(\frac{X^{inv}}{X}\right). \tag{3.A.86}$$

Thus, for $M = \left(\frac{X}{r-\alpha}\right)\left(\frac{X}{r-\alpha}\right)^{\frac{1}{\lambda-1}}$ it follows that

$$\left(\frac{X^{inv}}{r-\alpha} - K^{\lambda-1}\right)K_{new} < 0 \tag{3.A.87}$$

such that $L(X, K_{new}) < 0$, which contradicts the assumption. □

Lemma 3.A.9. *For $\beta_1 > 2$ or $\lambda > \frac{\beta_1}{\beta_1-1}$ the expected discounted value of delayed investment in peak load generation is bounded from above.*

Proof. We have to show that for any capacity and investment trigger with $X \le X^{inv} < X_P^{sus}$ then

$$\begin{aligned}&\left(\frac{X}{X^{inv}}\right)^{\beta_1}\left[\left(\frac{X^{inv}}{X_P^{sus}}\right)^{\beta_1}\frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)}K_{new}X_P^{sus} - I - K_{new}^{\lambda}\right]\\ &= \left(\frac{X}{X_P^{sus}}\right)^{\beta_1}\frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)}K_{new}X_P^{sus} - \left(\frac{X}{X^{inv}}\right)^{\beta_1}(I + K_{new}^{\lambda})\end{aligned} \tag{3.A.88}$$

is bounded from above. Furthermore for $X \leq X^{inv}$ and $X^{inv} \geq X_P^{sus}$ we have to show that

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1}\left[\left(\frac{X^{inv}}{r-\alpha}-\frac{X_P^{sus}}{r}\right)K_{new}-I-K_{new}^{\lambda}\right. \\ \left.+\left(\frac{X^{inv}}{X_P^{sus}}\right)^{\beta_2}\frac{\beta_1-1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}K_{new}X_P^{sus}\right] \tag{3.A.89}$$

is bounded from above. We have that (3.A.88) is bounded by

$$\left(\frac{X}{X_P^{sus}}\right)^{\beta_1}\left[\frac{1-\beta_2}{\beta_1(r-\alpha)(\beta_1-\beta_2)}K_{new}X_P^{sus}-(I+K_{new}^{\lambda})\right] \tag{3.A.90}$$

as $\beta_1 > 1$ and $X^{inv} < X_P^{sus}$. The upper bound (3.A.90) is continuous for $K_{new} \geq 0$ as $\lambda > 0$ and negative for $K_{new} = 0$. For $\beta_1 > 2$ the positive term in (3.A.90) converges to 0 for $K_{new} \to \infty$ as X_P^{sus} is linear in K_{new}. For $\lambda > 2$ (3.A.90) is negative for K_{new} large. Thus, (3.A.88) is bounded from above in both cases.
The expected payoff with $X^{inv} > X_P^{sus}$, (3.A.89), is bounded by

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1}\left[\left(\frac{X^{inv}}{r-\alpha}-\frac{X_P^{sus}}{r}\right)K_{new}-I-K_{new}^{\lambda}\right. \\ \left.+\frac{\beta_1-1}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}K_{new}X_P^{sus}\right] \tag{3.A.91}$$

as $\beta_2 < 0$ and $X^{inv} > X_P^{sus}$. As $X_P^{sus} < X^{inv}$ and $-\frac{X_P^{sus}}{r} < 0$, (3.A.91) is bounded by

$$\left(\frac{X}{X^{inv}}\right)^{\beta_1}\left[\frac{\beta_1-1-\beta_2(\beta_1-\beta_2)}{-\beta_2(r-\alpha)(\beta_1-\beta_2)}X^{inv}K_{new}-I-K_{new}^{\lambda}\right]. \tag{3.A.92}$$

For $\beta_1 > 2$ it follows that since $X_P^{sus} < X^{inv}$ then $K_{new} < (A+1)K + \frac{X^{inv}-c_P}{\gamma}$. Thus, the positive component in (3.A.92) is decreasing in X^{inv} as $\beta_1 > 2$. As the bound, (3.A.92), furthermore is continuous for $X^{inv} \geq X$ it follows that if $\beta_1 > 2$ the expected discounted value is bounded from above.
Let $\lambda > \frac{\beta}{\beta-1}$ and assume for contradiction that for all $M > 0$ there exists capacity and suspension trigger such that (3.A.92) is greater than M. This implies that the positive component is larger than M, which in turn implies that

$$K_{new} \geq M\frac{-\beta_2(r-\alpha)(\beta_1-\beta_2)}{\beta_1-1-\beta_2(\beta_1-\beta_2)X}\left(\frac{X^{inv}}{X}\right)^{\beta_1-1}. \tag{3.A.93}$$

Thus, as $(\lambda - 1)(\beta_1 - 1) > 1$ it follows that

$$K_{new}^{\lambda-1} \geq M_0 X^{inv} \tag{3.A.94}$$

for some $M_0 > 0$ that does not depend on X^{inv} and increases linearly in M. Now by choosing M large it follows that (3.A.92) is negative, which contradicts that (3.A.92) is larger than M. □

4

Valuation of power plants

Abstract

In this paper we develop continuous-time stochastic control models for valuation and operation of three different types of power plants in an electricity market: a renewable power plant, a conventional power plant and a storage power plant. Examples of these types of power plants are wind turbines, gas-fired units and hydroelectric power plants. In spite of detailed modeling, we derive analytical or quasi analytical solutions. In particular, we model uncertainty in electricity prices and in production input/output when it is relevant for the technology considered. Input/output is assumed to follow a diffusion process, whereas the price processes may include jumps. Our models account for the special characteristics of the technologies such as a non-normal distribution of wind speeds and hydro power inflows, as well as startup and shutdown costs of thermal units. We use these models to assess the impact of conjectured future market conditions such as increasing price trends, increased price volatility through changes in jump or diffusion behavior and increased correlation between renewable production and electricity prices.

4.1 Introduction

With ambiguous targets, many future electricity markets will be characterized by large shares of renewable generation, such as wind and solar power production which is highly unpredictable. This increases the need for flexibility in conventional generation due to an increase in supply uncertainty. Evidently,

to ensure continued operation and new investment in generation capacity, production must be profitable. With the current market setup, however, a change in the generation mix will change the dynamics of the electricity price. For instance, price volatility may increase and the correlation between renewable production and prices may become increasingly negative. The future value of power generation may therefore change depending on generator characteristics.

In this paper we quantify the effects of changes in price dynamics for three different stylized types of power generation, inflexible renewable generation, flexible conventional generation and a storage power plant. The renewable generation, exemplified by a wind turbine, is uncontrollable and cannot adjust to benefit from variations in prices. A negative correlation between power production and electricity prices will therefore lower the value of generation and reduce investment incentives.

The flexible generation, represented by a gas-fired power plant, is controllable and can temporarily suspended operation to avoid periods with low electricity prices, while benefitting from periods with high electricity prices. Finally, the storage power plant, exemplified by a hydroelectric power plant, can continuously adjust generation to electricity prices and store water for periods with high prices to the extent the reservoir capacity allows.

We assume that renewable generation involves no operational decisions and that production can be modeled by a load factor correlated with the price, see Abadie and Chamorro (2014), where the price process accounts for seasonality and the valuation problem is solved numerically using a Monte Carlo approach. In contrast, Boomsma et al. (2012) assume constant load that is adjusted for correlation and solve the model analytically in a more general investment setup without seasonality. We derive an analytical solution to the instantaneous value of generation. To quantify the impact of correlation, however, we directly model wind speeds using the approach of Zárate-Miñano et al. (2013).

In the absence of startup costs and operational constraints the value of a gas-fired power plant can be modeled as a sum of spark spread call options. An example is Deng (1999), who models electricity and fuel prices using mean-reverting jump-diffusion processes with either regime switching, deterministic volatility or stochastic volatility and develop quasi-analytic expressions for the option values. Näsäkkälä and Fleten (2005) model the spark spread directly with a two-factor model and likewise determine quasi-analytic expressions for the value of the gas fired power plant. When including temporal constraints such as minimum up-time/down-time restrictions or startup/shutdown costs, the dimension of the problem increases. Tseng and Barz (2002) and Carmona and Ludkovski (2010) solve the operational problem for short time horizons using different combinations of Monte Carlo simulation and dynamic programming, whereas Deng and Oren (2005) and Gardner and Zhuang (2000) apply stochastic dynamic programming to solve the problem on a lattice.

An alternative approach is to solve the associated Hamilton-Jacobi-Bellman (HJB) equation numerically using some finite difference scheme, see Thompson et al. (2004). However, as these methods are subject to the curse of dimensionality, the computation time increases exponentially with increasing accuracy in time and space, which makes them impractical for longer time-horizons. Here, we suggest a one-factor model for the electricity price with mean reversion and jumps over an infinite horizon for which we can maximize the option constant avoiding discretization of both time and space.

For the hydroelectric power plant the storage level introduces another dimension, which makes analytical solutions significantly more difficult as the control impacts the derivative of the storage level. For short time horizons it is possible to solve the HJB equation numerically, see Thompson et al. (2004) and Chen and Forsyth (2008b), who allow for operational constraints such as ramping as well as seasonality in the price process. The combination of Monte Carlo simulation and dynamic programming by Carmona and Ludkovski (2010) is also used to solve the hydroelectric power plant, but requires significant computation time. To reduce computation time Näsäkkälä and Keppo (2005) approximates the optimal switching strategy using a parametrized boundary and a Monte Carlo approach. Other solution methods rely on linearizations of the operational strategies. For example, Braaten et al. (2016) use strategies that are linear functions of the observed prices, whereas Doege et al. (2006) consider linear combinations of predefined step-functions. These approaches have long time-steps or cover short time horizons. In contrast, we handle operational boundaries of the hydroelectric power plant using penalty functions and linearize the optimal control from the HJB equation to obtain quasi-analytic solutions to the value of the infinite horizon problem. With this approach we obtain an explicit discharge strategy that is linear in price and storage level and satisfies the storage and flow rate constraints with a high probability.

We solve the valuation problems with three price models based on stochastic differential equations. The initial model is a simple shifted Geometric Brownian Motion, which is compared to a shifted exponential Ornstein-Uhlenbeck processes that is mean reverting. To better capture the distributional properties of transformed price increments, the final model is extended to include jumps.

The rest of the paper is structured as follows. In Section 4.2 we develop two diffusion models and a jump diffusion model for the spot price of electricity that are analytically tractable and allow for negative electricity prices. In Section 4.3 we model uncertainty in weather factors such as to capture distributional properties in continuous time and exponentially decaying autocorrelations. In Section 4.4 we include correlation between weather factors and the spot price. Section 4.4.1 considers the value of a wind turbine and present a closed form solution under some regularity conditions. The value of a gas-fired unit is presented in Section 4.4.2, where we incorporate startup

and shutdown costs. Section 4.4.3 finally obtains the value of a hydroelectric power plant by relaxing storage and flow rate constraints using a penalty function and linearizing the control strategy. In Section 4.5 we report the results for a case study, and in Section 4.6 we study the impact of changes in the price dynamics. Finally, Section 4.7 provides a brief conclusion.

4.2 Electricity Price Uncertainty

We aim to investigate the effects of different electricity price dynamics on the value of power generation from an investment point of view. We assume that generation is always dispatched in a market and therefore focus on market prices. The market for immediate dispatch of generation is referred to as a spot market and prices are likewise referred to as spot prices.

Classical papers on commodity prices such as Schwartz and Smith (2000), Lucia and Schwartz (2002) and Gibson and Schwartz (1990) model the logarithm of the price, $X_t = \log(P_t)$ such that the price has the form

$$P_t = e^{X_t}, \tag{4.2.1}$$

which captures the skewness of the prices when X_t has a symmetric distribution, is analytically tractable and implies that the prices are non-negative. However, electricity prices may occasionally become negative as a result of unpredictable excess of renewable production combined with insufficient flexibility to quickly reduce conventional generation. This is expected to happen more frequently with increasing shares of renewable production in the electricity market, see Götz et al. (2014). To allow for negative prices we assume that

$$P_t = e^{X_t} - M, \tag{4.2.2}$$

where $-M$ is a lower bound on prices. With this assumption the model remains analytically tractable. We assume that the logarithm of the shifted price, $X_t = \log(P_t + M)$, follows a Lévy-driven stochastic differential equation (SDE) such that the dynamics of the price process $(X_t)_{t\geq 0}$ are given by

$$\mathrm{d}X_t = \mu(X_t)\ \mathrm{d}t + \sigma(X_t)\ \mathrm{d}Z_t^P + \gamma(X_t)\ \mathrm{d}J_t. \tag{4.2.3}$$

Here, $\mu(X_t)$, $\sigma(X_t)$ and $\gamma(X_t)$ are the drift, diffusion and jump coefficients, respectively. Furthermore, $(Z_t^P)_{t\geq 0}$ is a standard Brownian motion and $(J_t)_{t\geq 0}$ is a compound Poisson process with $J_t = \sum_{n=1}^{N_t} Y_n$, where $(N_t)_{t\geq 0}$ is the corresponding Poisson process with intensity λ and $(Y_n)_{n\geq 1}$ are independent and identically distributed with mean $\mathbb{E}(Y_1) = \eta$ and variance $Var(Y_1) = \nu^2$. We assume that $(Z_t^P)_{t\geq 0}$ and $(J_t)_{t\geq 0}$ are independent.

Remark 2. *The diffusion part of the SDE represents continuous changes in electricity prices caused by the development in supply and demand, changes in the economic environment or other new information that causes only marginal changes in prices. In contrast, the jump part represents discrete changes due to new information that has more than a marginal effect on price such as failure of production units and sudden changes in demand and renewable production.*

We consider two diffusion models as well as a jump diffusion model for $(X_t)_{t\geq 0}$ and value generation under all three models.

4.2.1 Brownian Motion

In the first model we assume that $(X_t)_{t\geq 0}$ follows the SDE

$$\mathrm{d}X_t = \left(\mu - \frac{1}{2}\sigma^2\right)\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t^P \tag{4.2.4}$$

with $X_0 = \log(P_0 + M)$. Here μ and $\sigma > 0$ are constants. The SDE has the solution

$$X_t = X_s + \left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma\left(Z_t^P - Z_s^P\right), \tag{4.2.5}$$

for $t > s$, and thus

$$P_t = (P_s + M)\exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma\left(Z_t^P - Z_s^P\right)\right] - M. \tag{4.2.6}$$

With drift $\mu - \sigma^2/2$, the diffusion is offset in the expected price such that the expectation of P_t conditional on P_s has the simple form,

$$\mathbb{E}(P_t|P_s) = (P_s + M)e^{\mu(t-s)} - M. \tag{4.2.7}$$

This follows as $\mathbb{E}[e^{\sigma(Z_t^P - Z_s^P)}|P_s] = e^{\frac{\sigma^2}{2}(t-s)}$, which we show in Lemma 4.B.1.

4.2.2 Ornstein-Uhlenbeck

In the second model we assume $(X_t)_{t\geq 0}$ follows the SDE,

$$\mathrm{d}X_t = \kappa_P\left(\alpha - \frac{\sigma^2}{4\kappa_P} - X_t\right)\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t^P, \tag{4.2.8}$$

with α, $\sigma > 0$ and $\kappa_P > 0$ constants and $X_0 = \log(P_0 + M)$. In Lemma 4.B.3, we show that the solution is

$$\begin{aligned} X_t = {} & e^{-\kappa_P(t-s)}X_s + \left(\alpha - \frac{\sigma^2}{4\kappa_P}\right)\left(1 - e^{-\kappa_P(t-s)}\right) \\ & + \sigma\int_s^t e^{-\kappa_P(t-v)}\,\mathrm{d}Z_v^P, \end{aligned} \tag{4.2.9}$$

for $t > s$. This implies that

$$\begin{aligned} P_t = (P_s + M)^{e^{-\kappa_P(t-s)}} \exp\Bigg[&\left(\alpha - \frac{\sigma^2}{4\kappa_P}\right)\left(1 - e^{-\kappa_P(t-s)}\right) \\ &+ \sigma \int_s^t e^{-\kappa_P(t-v)} \, \mathrm{d}Z_v^P \Bigg] - M. \end{aligned} \tag{4.2.10}$$

As,

$$\sigma \int_s^t e^{-\kappa_P(t-v)} \, \mathrm{d}Z_v^P \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\kappa_P}\left(1 - e^{-2\kappa_P(t-s)}\right)\right), \tag{4.2.11}$$

it follows by Lemma 4.B.1 that

$$\mathbb{E}(P_t | P_s) \to e^{\alpha} - M \text{ for } t \to \infty, \tag{4.2.12}$$

which is independent of the current price. In this model the price reverts to the mean reversion level,

$$\exp\left(\alpha - \frac{\sigma^2}{4\kappa_P}\right) - M, \tag{4.2.13}$$

but due to the skewness of the diffusion term the expected future price converges to $e^{\alpha} - M$.

4.2.3 Ornstein-Uhlenbeck with Jumps

The third model is an extension of the second, where we assume $(X_t)_{t \geq 0}$ follows the Levy-driven SDE,

$$\mathrm{d}X_t = \kappa_P \left(\alpha - \frac{\sigma^2}{4\kappa_P} - \lambda k - X_t\right) \, \mathrm{d}t + \sigma \, \mathrm{d}Z_t^P + \, \mathrm{d}J_t, \tag{4.2.14}$$

with α, k, λ, $\sigma > 0$ and $\kappa_P > 0$ constants and $X_0 = \log(P_0 + M)$. Here k depends on κ_P and the distribution of Y_1 and compensates for the jumps. In Lemma 4.B.3 we show that the solution is

$$\begin{aligned} X_t = e^{-\kappa_P(t-s)} X_s + &\left(\alpha - \frac{\sigma^2}{4\kappa_P} - \lambda k\right)\left(1 - e^{-\kappa_P(t-s)}\right) \\ &+ \sigma \int_s^t e^{-\kappa_P(t-v)} \, \mathrm{d}Z_v^P + \sum_{n=N_s+1}^{N_t} e^{-\kappa_P(t-T_n)} Y_n, \end{aligned} \tag{4.2.15}$$

where T_n is the time of the n'th jump. This implies that

$$\begin{aligned} P_t = (P_s - M)^{e^{-\kappa_P(t-s)}} \exp\Bigg[&\left(\alpha - \frac{\sigma^2}{4\kappa_P} - \lambda k\right)\left(1 - e^{-\kappa_P(t-s)}\right) \\ &+ \sigma \int_s^t e^{-\kappa_P(t-v)} \, \mathrm{d}Z_v^P \Bigg] \prod_{n=N_s+1}^{N_t} \left(e^{Y_n}\right)^{e^{-\kappa_P(t-T_n)}} - M. \end{aligned} \tag{4.2.16}$$

Thus, the effect of the jumps decay exponentially over time, implying that only recent jumps impact the price. We show in Lemma 4.B.2 that

$$\mathbb{E}\left(\prod_{n=N_s+1}^{N_t}\left(e^{Y_n}\right)^{e^{-\kappa_P(t-T_n)}}\middle| P_s\right)=\exp\left(A(s,t)\right), \tag{4.2.17}$$

where

$$A(s,t)=\lambda\int_{e^{-\kappa_P(t-s)}}^{1}\frac{\theta(z)-1}{\kappa_P z}\,\mathrm{d}z, \tag{4.2.18}$$

with $\theta(c)=\mathbb{E}(e^{cY_1})$. Hence,

$$\begin{aligned}\mathbb{E}(P_t|P_s)=(P_s+M)^{e^{-\kappa_P(t-s)}}\exp\Bigg[&\left(\alpha-\frac{\sigma^2}{4\kappa_P}-\lambda k\right)\left(1-e^{-\kappa_P(t-s)}\right)\\&+\frac{\sigma^2}{4\kappa_P}\left(1-e^{-2\kappa_P(t-s)}\right)+A(s,t)\Bigg]-M.\end{aligned} \tag{4.2.19}$$

Now with

$$k=\int_0^1\frac{\theta(z)-1}{\kappa_P z}\,\mathrm{d}z \tag{4.2.20}$$

we obtain $A(s,t)\to\lambda k$ for $t\to\infty$ and thus

$$\mathbb{E}(P_t|P_s)\to e^{\alpha}-M \text{ for } t\to\infty. \tag{4.2.21}$$

In this model the price reverts to the mean reversion level

$$\exp\left(\alpha-\frac{\sigma^2}{4\kappa_P}-\lambda k\right)-M. \tag{4.2.22}$$

We assume that the compound Poisson process for the Ornstein-Uhlenbeck process is modeled with two jump terms, i.e.

$$J_t=\sum_{j=1}^{2}\sum_{n=1}^{N_t^{(j)}}Y_n^{(j)} \tag{4.2.23}$$

with $Y_n^{(j)}$ independent and $Y_n^{(j)}\sim\mathcal{N}(\eta^{(j)},{\nu^{(j)}}^2)$ for $j=1,2$ and $N_t^{(j)}$ independent Poisson processes with intensity $\lambda^{(j)}$ for $j=1,2$. Note that this is equivalent to a compound Poisson process with intensity $\lambda=\lambda^{(1)}+\lambda^{(2)}$, where Y_1 has density

$$\phi(y)=\sum_{j=1}^{2}\frac{\lambda^{(j)}}{\lambda}\varphi(y;\eta^{(j)},{\nu^{(j)}}^2), \tag{4.2.24}$$

and where $\varphi(y;\mu,\sigma^2)$ is the density for the normal distribution with mean μ and variance σ^2. Thus,

$$k = \sum_{j=1}^{2} \frac{\lambda^{(j)}}{\lambda} \int_0^1 \frac{e^{z\eta^{(j)}+\frac{1}{2}z^2{\nu^{(j)}}^2}-1}{\kappa_P z}\,\mathrm{d}z, \tag{4.2.25}$$

which has no closed-form solution, but can be written as the infinite series,

$$k = \sum_{j=1}^{2} \frac{\lambda^{(j)}}{\lambda} \sum_{n=1}^{\infty} \frac{E\left[\left(Y_1^{(j)}\right)^n\right]}{n!n\kappa_P}. \tag{4.2.26}$$

4.3 Uncertainty in Weather Factors

We proceed to model uncertainty in production input and output such as reservoir inflow and wind power production, which is often inflexible and unpredictable. The standard assumption of a normal distribution may, however, not be representative for this type of uncertainty. We therefore develop a framework for modelling non-normal distributions in a continuous-time setting by using a transformation of the normal distribution to a more appropriate one.

We assume that non-controllable production input and output is driven by an underlying weather factor $(U_t)_{t\geq 0}$ that follows an Ornstein-Uhlenbeck process with dynamics

$$\mathrm{d}U_t = -\kappa_U U_t\,\mathrm{d}t + \sqrt{2\kappa_U}\,\mathrm{d}Z_t^U, \tag{4.3.1}$$

where $\kappa_U > 0$ is a constant and Z_t^U is a standard Brownian motion. Note that the diffusion coefficient of $\sqrt{2\kappa_U}$ implies that the variance of the process converges to 1. We assume that $(Z_t^U)_{t\geq 0}$ and $(J_t)_{t\geq 0}$ are independent, whereas $(Z_t^U)_{t\geq 0}$ and $(Z_t^P)_{t\geq 0}$ are correlated with correlation coefficient ρ.

Remark 3. *We allow for the price and weather processes to be correlated, as cold and hot temperatures may increase consumption through heating and air-conditioning and thereby prices, windy weather conditions increases wind power production and reduces prices, and precipitation or snow melt increase reservoir inflows and likewise reduces prices. Furthermore, a decrease in wind generation has to be offset by more expensive generation, which in turn increases the prices.*

We can write the solution to the Ornstein-Uhlenbeck process as

$$U_t = U_s e^{-\kappa_U(t-s)} + \sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(t-v)}\,\mathrm{d}Z_v^U, \tag{4.3.2}$$

for $t > s$. Note that U_t is normally distributed with mean $\mathbb{E}[U_t|U_s] = U_s e^{-\kappa_U(t-s)}$ and variance $\text{Var}[U_t|U_s] = 1 - e^{-2\kappa_U(t-s)}$. Hence, $U_t \xrightarrow{d} N(0,1)$ for $t \to \infty$, where d means convergence in distribution, that is, the cumulative distribution function for U_t converges point-wise to the standard normal distribution function.

Now, the idea is to do a transformation of the Ornstein-Uhlenbeck process, using the normal distribution function and an appropriate quantile function, and thereby obtain the desired process in the limit as well as exponentially decaying autocorrelation.

Remark 4. *The exponentially decaying autocorrelation captures that the weather conditions at time t are usually highly correlated with those at time $t + \delta$ for δ small. However, for δ large, weather conditions are much less correlated. For instance, wind speeds are highly correlated with those of near history but uncorrelated with those of distant history.*

It follows by Theorem 4.2 from Bradley (1986) that for all transformations of ρ-mixing processes, for which the autocorrelation is defined, the autocorrelation is bounded by a function that is exponentially decaying. It is therefore sufficient to prove that the Ornstein-Uhlenbeck process is ρ-mixing, which we do in Lemma 4.B.4.

To obtain the appropriate limiting process, let $\Phi(\cdot)$ and $\Phi_t(\cdot)$ denote the distribution functions of a standard normal distribution and U_t, respectively. Moreover, let $F(\cdot)$ denote the cumulative distribution function of a given continuous distribution. We assume that the density that has support given by the interval $I \subseteq \mathbb{R}$ such that the distribution function is strictly increasing in the interior of I. With this assumption, the quantile function $F^{-1}(\cdot)$ is the inverse of $F(w)$ for $w \in I$.

We define production input/output W_t as

$$W_t = F^{-1}(\Phi(U_t)). \tag{4.3.3}$$

Then,

$$\mathbb{P}(W_t \leq w) = \mathbb{P}(U_t \leq \Phi^{-1}(F(w))) \tag{4.3.4}$$

$$= \Phi_t(\Phi^{-1}(F(w))) \to F(w) \text{ for } t \to \infty, \tag{4.3.5}$$

using that U_t converges in distribution to the standard normal distribution. Thus, the cumulative distribution function of production input/output W_t is given by the desired $F(\cdot)$ in the limit, and can otherwise, be used as an approximation.

Our framework is particularly useful for modelling wind speeds, which are typically assumed to be Weibull distributed, see Spera (2009), Zárate-Miñano et al. (2013) and Lun and Lam (2000). We provide an example in the following section.

4.3.1 Wind Speeds

We aim to model the wind speed at time t, W_t, by a Weibull distribution with constant shape parameter $\beta > 0$ and scale parameter $\zeta > 0$. The cumulative distribution function for the Weibull distribution is given by

$$F(w) = \begin{cases} 1 - e^{-(w/\zeta)^\beta} & , w \geq 0, \\ 0 & , w < 0, \end{cases} \tag{4.3.6}$$

with inverse

$$F^{-1}(x) = \zeta(-\ln(1-x))^{1/\beta}, \quad 0 \leq x < 1. \tag{4.3.7}$$

Hence, by defining wind speed as

$$W_t = F^{-1}(\Phi(U_t)), \tag{4.3.8}$$

W_t is approximately Weibull distributed with shape parameter β and scale parameter ζ. In this manner, we obtain a non-normal mean reverting behavior of wind speeds as well as an exponentially decaying autocorrelation.

The approach can likewise be used to define more sophisticated price models. Such prices models can be used in all three valuation models proposed in this paper, provided that moments and covariances of the process at different points in time can be determined numerically.

4.4 Valuation of Power Generation

In this section we describe the valuation of the three different generation technologies. For analytical tractability, the valuation is based on the expected discounted value of cashflows over an infinite horizon. The discount rate $r > 0$ is assumed exogenous. The cashflows stem from dispatch of generation and are therefore driven by spot prices. We assume a perfectly competitive market and hence that the generators are price-takers. In the following we account for the special characteristics of each technology, including relevant uncertainties and the degree of operational flexibility.

4.4.1 Wind Turbine

We begin by considering an investment in a wind turbine. Wind power generation is characterized by being non-controllable, highly varying and largely unpredictable as production is determined by the weather conditions, and more specifically the speed of the wind. We therefore assume that the instantaneous revenue generated by the plant is driven by the wind speed W_t at the location of the plant and the spot price P_t. The wind speed is modeled using the weather factor U_t such that $W_t = f(U_t)$, where for example $f(u) = F^{-1}(\Phi(u))$ and $F(w)$ is the cumulative distribution function for the

Weibull distribution. The rate of production Q_t depends on the wind speed through a turbine-specific power curve $h(w)$, and so $Q_t = h(W_t)$ is the production rate for a given wind speed. We further assume that variable costs of production are negligible. Now, the value of the wind turbine is the expected discounted revenues, given the current spot price $P_t = P$ and wind speed $W_t = W$, i.e.

$$\mathbb{E}\left[\int_0^\infty e^{-rt} P_t h(W_t)\ \mathrm{d}t \middle| P_t = P, W_t = W\right]. \tag{4.4.1}$$

Expressed in terms of the logarithm to the transformed price, X_t and the weather factor, U_t, the valuation problem is

$$V_{wind}(X,U) = \mathbb{E}\left[\int_0^\infty e^{-rt}(e^{X_t} - M)h(f(U_t))\ \mathrm{d}t \middle| X_0 = X, U_0 = U\right] \tag{4.4.2}$$

$$\text{s.t. } \mathrm{d}X_t = \mu(X_t)\ \mathrm{d}t + \sigma(X_t)\ \mathrm{d}Z_t^P + \gamma(X_t)\ \mathrm{d}J_t, \tag{4.4.3}$$

$$\mathrm{d}U_t = -\kappa_U U_t\ \mathrm{d}t + \sqrt{2\kappa_U}\ \mathrm{d}Z_t^U. \tag{4.4.4}$$

The associated Hamilton-Jacobi-Bellman integro-differential equation is

$$\mathbb{L}(V_{wind}(t,X,U)) + (e^X - M)h(f(U)) - rV_{wind}(t,X,U) = 0, \tag{4.4.5}$$

where $\mathbb{L}$ is the partial integro differential operator

$$\begin{aligned}\mathbb{L}(V_{wind}(t,X,U)) = &\left(\mu(X)\frac{\partial}{\partial X} + \frac{1}{2}\sigma(X)^2\frac{\partial^2}{\partial X^2} - \kappa_U U\frac{\partial}{\partial U}\right.\\ &\left.+\sqrt{2\kappa_U}\frac{\partial^2}{\partial U^2} + \rho\sigma(X)\sqrt{2\kappa_U}\frac{\partial^2}{\partial U \partial X}\right)V_{wind}(X,U)\\ &+\lambda\mathbb{E}\left[V_{wind}(X + \gamma(X)Y_1, U) - V_{wind}(X,U)\right].\end{aligned} \tag{4.4.6}$$

The stochastic control problem can be solved numerically by solving a partial integro differential equation (PIDE) or a partial differential equation (PDE), depending on whether the price process includes jumps.

4.4.1.1 A Piece-wise Linear Power Curve

A power curve typically consists of four parts, where different parts can be described by different functions. Production is equal to zero below cut-in speed (w_0), it is increasing in the windspeed between cut-in speed (w_0) and rated wind speed (w_1), constant between rated wind speed (w_1) and storm protection shutdown (w_2) and equal to zero above storm protection shutdown

(w_2).[1] Hence,

$$h(w) = \begin{cases} 0, & w \leq w_0 \\ h_{01}(w), & w_0 < w \leq w_1 \\ h_{12}, & w_1 < w \leq w_2 \\ 0, & w_2 < w \end{cases} \tag{4.4.7}$$

where $h_{01}(w)$ and h_{12} are such that $h(w)$ is continuous on $[0, w_2]$. We assume $h(f(u))$ is linear in u and, thus

$$h(f(u)) = \begin{cases} 0, & u \leq u_0 \\ h_{12}(u - u_0)/(u_1 - u_0), & u_0 < u \leq u_1 \\ h_{12}, & u_1 < u \leq u_2 \\ 0, & u > u_2 \end{cases} \tag{4.4.8}$$

where $u_0 = \Phi^{-1}(F(w_0))$, $u_1 = \Phi^{-1}(F(w_1))$ and $u_2 = \Phi^{-1}(F(w_2))$.[2]

With a power curve that is piecewise linear in u, the expected instantaneous value of production can be computed analytically and the integral in (4.4.2) can be computed using numerical integration. To derive the analytical solution, note that $\mathbb{E}(h(f(U_t))|X_s, U_s)$ can be found as a special case of $\mathbb{E}(e^{X_t}h(f(U_t))|X_s, U_s)$. We show in Lemma 4.B.7 that if the logarithm to the transformed price is described by a Brownian Motion, then $\mathbb{E}(e^{X_t}h(f(U_t))|X_s, U_s)$ can be expressed in terms of adjusted values of u_0, u_1 and u_2 as

$$\begin{aligned} \mathbb{E}(e^{X_t}h(f(U_t))|X_s, U_s) = e^{X_s+\mu(t-s)}\Bigg[& h_{12}\left[\Phi\left(\tilde{u}_2\right) - \Phi\left(\tilde{u}_1\right)\right] \\ & + \frac{h_{12}}{\tilde{u}_1 - \tilde{u}_0}\left[\varphi(\tilde{u}_0) - \varphi(\tilde{u}_1)\right] \\ & - \frac{\tilde{u}_0 h_{12}}{\tilde{u}_1 - \tilde{u}_0}\left[\Phi\left(\tilde{u}_1\right) - \Phi\left(\tilde{u}_0\right)\right]\Bigg] \end{aligned} \tag{4.4.9}$$

for $t > s$. Here, $\varphi(u)$ is the density for the standard normal distribution and

$$\tilde{u}_i = u_i + (u_i - U_s)\frac{e^{-\kappa_U(t-s)}}{1 - e^{-2\kappa_U(t-s)}} - \rho\sigma\frac{\sqrt{2\kappa_U}}{\kappa_U}\frac{1 - e^{-\kappa_U(t-s)}}{1 - e^{-2\kappa_U(t-s)}} \tag{4.4.10}$$

for $t > s$ and $i = 0, 1, 2$. Thus, $\tilde{u}_i \to u_i - \rho\sigma\sqrt{2\kappa_U}/\kappa_U$ for $t \to \infty$. If instead the logarithm to the transformed price follows an Ornstein-Uhlenbeck process

[1] For simplicity we ignore re-cut-in, the wind speed at which the wind turbine starts up after triggering the storm protection.

[2] For improved accuracy additional points can be included in the linearisation of the power curve.

with jumps we show in Lemma 4.B.8 that

$$\begin{aligned}\mathbb{E}\big[e^{X_t}h(f(U_t))|X_s,U_s\big] = e^{X_s e^{-\kappa_P(t-s)}+\alpha_t}\Bigg[&h_{12}\left[\Phi\left(\tilde{u}_2\right)-\Phi\left(\tilde{u}_2\right)\right]\\ &+\frac{h_{12}}{\tilde{u}_1-\tilde{u}_0}\left[\varphi(\tilde{u}_0)-\varphi(\tilde{u}_1)\right]\\ &-\frac{\tilde{u}_0 h_{12}}{\tilde{u}_1-\tilde{u}_0}\left(\Phi\left(\tilde{u}_1\right)-\Phi\left(\tilde{u}_0\right)\right)\Bigg]\end{aligned} \tag{4.4.11}$$

for $t > s$ and with $\alpha_t = (\alpha+\epsilon_\sigma)(1-e^{-\kappa_P(t-s)})-\epsilon_k$, $\epsilon_\sigma = (\sigma^2/4)e^{-\kappa_P(t-s)}$ and $\epsilon_k = \lambda\int_0^{e^{-\kappa_P(t-s)}}(\theta(z)-1)/(\kappa_P z)\,\mathrm{d}z$. Here, $\alpha_t \to \alpha$ and $\epsilon_k, \epsilon_\sigma \to 0$ for $t \to \infty$. Furthermore,

$$\tilde{u}_i = u_i + (u_i - U_s)\frac{e^{-\kappa_U(t-s)}}{1-e^{-2\kappa_U(t-s)}} - \frac{\rho\sigma\sqrt{2\kappa_U}}{(\kappa_U+\kappa_P)}\frac{1-e^{-(\kappa_U+\kappa_P)(t-s)}}{1-e^{-2\kappa_U(t-s)}} \tag{4.4.12}$$

for $t > s$ and $i = 0,1,2$. Thus, $\tilde{u}_i \to u_i - \rho\sigma\sqrt{2\kappa_U}/(\kappa_U+\kappa_P)$ for $t \to \infty$. Without jumps $\lambda = 0$ such that $\epsilon_k = 0$. For both the Brownian Motion and the Ornstein-Uhlenbeck process the negative correlation between Z_t^U and Z_t^P, ρ, increases $\tilde{u}_0$, $\tilde{u}_1$ and $\tilde{u}_2$ compared to u_0, u_1 and u_2, which effectively corresponds to a higher cut-in speed, rated wind speed and storm protection. The adjustment for the negative correlation between price and power generation therefore decreases the expected value of the wind turbine. The first terms of (4.4.9) and (4.4.11) essentially consist of expected price times expected output between the adjusted rated wind speed and the adjusted storm protection on the constant part of the power curve, whereas the second and third terms capture expected price times expected output on the linearly increasing part of the adjusted power curve.

4.4.2 Gas Fired Power Plant

The second type of power plant considered is a gas-fired power plant. In contrast to the wind turbine, for such a plant, production can be controlled. We ignore ramping constraints and assume that production can be adjusted instantaneously. We denote the production rate by Q_t and assume that the minimum and maximum rates are given by Q_{min} and Q_{max}. Instantaneous revenues depend on the electricity price P_t, whereas variable costs of production depend mainly on the price of gas p_f. For simplicity, we assume the gas price is constant. The amount of gas required for a production rate of Q_t is given by $H(Q_t)$.[3] Fixed costs include those of starting up and shutting down

[3]This model can incorporate stochastic gas prices by modelling an adjusted electricity price instead of the electricity price under the assumption that the stochastic gas price, p_t^f, is independent of P_t/p_t^f. With $p_f = \mathbb{E}(p_t^f)$ and the adjusted price given by $\tilde{P}_t = P_t p_f/p_t^f$ it follows that $\mathbb{E}\left[P_tQ_t - p_t^f H(Q_t)\right] = \mathbb{E}\left[p_t^f/p_f\left(\tilde{P}_t - p_f H(Q_t)\right)\right] = \mathbb{E}\left[\tilde{P}_tQ_t - p_f H(Q_t)\right]$. Here the last equality follows by independence of p_t^f and $\tilde{P}_t$ as Q_t is a deterministic function of $\tilde{P}_t$ and as $\mathbb{E}\left[p_t^f/p_f\right] = 1$.

the plant, denoted by C_{on} and C_{off}, respectively. Thus, the instantaneous profit when the plant is online is

$$\pi_{on}(P_t) = P_t Q_t^* - p_f H(Q_t^*), \tag{4.4.13}$$

where the optimal rate of production is

$$Q_t^* = \begin{cases} Q_{min}, & P_t < p_f H'(Q_{min}) \\ (H')^{-1}(P_t/p_f), & P_t \in [p_f H'(Q_{min}), p_f H'(Q_{max})] \\ Q_{max}, & P_t > p_f H'(Q_{max}) \end{cases} \tag{4.4.14}$$

assuming $H(q)$ is a concave function, which also ensures that the inverse to the derivative of $H(q)$ exists. Hence, at the optimal rate of production the electricity price equals the marginal cost of generation, $p_f H'(Q_t^*)$, while respecting the minimum and maximum production rates. When the plant is offline we let $\pi_{off}(P_t) = 0$.

We assume that there exists startup and shutdown trigger prices $P_{on}^* > P_{off}^*$ such that it is optimal to start up an offline plant when $P_t \geq P_{on}^*$ and shut down an online plant when $P_t \leq P_{off}^*$. We denote the random startup times for an online plant by $S_{on}^{(n)}$ for $n \geq 0$ and the random shutdown times for an online plant by $S_{off}^{(n)}$ for $n \geq 1$, where $S_{on}^{(0)} = 0$ and recursively define

$$S_{on}^{(n)} = \inf\left\{ t \geq S_{off}^{(n)} \middle| P_t \geq P_{on}^* \right\} \tag{4.4.15}$$

$$S_{off}^{(n)} = \inf\left\{ t \geq S_{on}^{(n-1)} \middle| P_t \leq P_{off}^* \right\} \tag{4.4.16}$$

for $n \geq 1$. Thus, the value of production when the plant is online is the expected future profit less the costs of startup and shutdown, i.e.

$$\begin{aligned} V_{gas}^{on}(P) =& \mathbb{E}\Bigg[\sum_{n=0}^{\infty} \int_{S_{on}^{(n)}}^{S_{off}^{(n+1)}} e^{-rs} \left(P_s Q_s^* - p_f H(Q_s^*) \right) \, \mathrm{d}s \\ & - C_{on} \sum_{n=1}^{\infty} e^{-r S_{on}^{(n)}} - C_{off} \sum_{n=1}^{\infty} e^{-r S_{off}^{(n)}} \Bigg| P_0 = P \Bigg] \end{aligned} \tag{4.4.17}$$

As we do not require $S_{on}^{(1)} \geq S_{off}^{(1)}$ when the plant is offline, we introduce the startup and shutdown times for an offline plant, $\tilde{S}_{on}^{(n)}$ for $n \geq 1$, and $\tilde{S}_{off}^{(n)}$ for $n \geq 0$, where $\tilde{S}_{off}^{(0)} = 0$ and define

$$\tilde{S}_{off}^{(n)} = \inf\left\{ t \geq \tilde{S}_{on}^{(n)} \middle| P_t \leq P_{off}^* \right\} \tag{4.4.18}$$

$$\tilde{S}_{on}^{(n)} = \inf\left\{ t \geq \tilde{S}_{off}^{(n-1)} \middle| P_t \geq P_{on}^* \right\} \tag{4.4.19}$$

for $n \geq 1$. With this construction, the value of production when the plant is offline is likewise the expected future profit less the costs of startup and shutdown

$$\begin{aligned} V_{gas}^{off}(P) = \mathbb{E}\Bigg[\sum_{n=1}^{\infty} \int_{\tilde{S}_{on}^{(n)}}^{\tilde{S}_{off}^{(n+1)}} e^{-rs} \left(P_s Q_s^* - p_f H(Q_s^*)\right) \, \mathrm{d}s \\ - C_{on} \sum_{n=1}^{\infty} e^{-r\tilde{S}_{on}^n} - C_{off} \sum_{n=2}^{\infty} e^{-r\tilde{S}_{off}^n} \Bigg| P_0 = P \Bigg]. \end{aligned} \tag{4.4.20}$$

As a result, we obtain the recursions

$$\begin{aligned} V_{gas}^{on}(P) = \mathbb{E}\Bigg[\int_0^{S_{off}^{(1)}} e^{-rs} \left(P_s Q_s^* - p_f H(Q_s^*)\right) \, \mathrm{d}s \\ + e^{-rS_{off}^{(1)}} \left(V_{gas}^{off}(P_{S_{off}^{(1)}}) - C_{off} \right) | P_0 = P \Bigg] \end{aligned} \tag{4.4.21}$$

and

$$V_{gas}^{off}(P) = \mathbb{E}\Bigg[e^{-r\tilde{S}_{on}^{(1)}} \left(V_{gas}^{on}(P_{\tilde{S}_{on}^{(1)}}) - C_{on} \right) | P_0 = P \Bigg]. \tag{4.4.22}$$

Following the results from Ernstsen and Misir (2016) we find that

$$V_{gas}^{on}(P) = \Pi_{on}(P) + c_1(P_{on}^*, P_{off}^*) M_1(P) \tag{4.4.23}$$

$$V_{gas}^{off}(P) = c_2(P_{on}^*, P_{off}^*) M_2(P) \tag{4.4.24}$$

where $c_1(P_{on}^*, P_{off}^*)$ and $c_2(P_{on}^*, P_{off}^*)$ are constants,

$$\Pi_{on}(P) = \mathbb{E}\left[\int_0^{\infty} e^{-rt} \left(P_s Q_s^* - p_f H(Q_s^*)\right) | P_0 = P \right], \tag{4.4.25}$$

is the value of always being online and generating, and we refer to the last terms in (4.4.23) and (4.4.24) as option values. The functions, $M_1(P)$ and $M_2(P)$ can be expressed as

$$M_1(P) = m_1(\ln(P + M)) \tag{4.4.26}$$

$$M_2(P) = m_2(\ln(P + M)), \tag{4.4.27}$$

where $m_1(x)$ and $m_2(x)$ are bounded and twice differentiable on $(-\infty, x_0)$ and (x_0, ∞) for $x_0 \in \mathbb{R}$ and solves

$$\begin{aligned} \mu(X) \frac{\partial}{\partial X} m(X) + \frac{1}{2} \sigma(X)^2 \frac{\partial}{\partial X^2} m(X) - r m(X) \\ + \lambda \mathbb{E}\left[m(X + Y_1) - m(X) \right] = 0, \end{aligned} \tag{4.4.28}$$

with $m_1(x) \to 0$ for $x \to -\infty$ and $m_2(x) \to 0$ for $x \to \infty$. Now (4.4.21) and (4.4.22) implies that

$$V_{gas}^{on}(P_{off}) = V_{gas}^{off}(P_{off}) - C_{off} \tag{4.4.29}$$

$$V_{gas}^{off}(P_{on}) = V_{gas}^{on}(P_{on}) - C_{on} \tag{4.4.30}$$

as $S_{on}^{(1)}|(P_0 = P_{on}) = 0$ and $S_{off}^{(1)}|(P_0 = P_{off}) = 0$ and thus $P_{S_{on}^{(1)}}|(P_0 = P_{on}) = P_{on}$ and $P_{S_{off}^{(1)}}|(P_0 = P_{off}) = P_{off}$. Combining (4.4.23), (4.4.24), (4.4.29) and (4.4.30) implies that

$$c_1(P_{on}^*, P_{off}^*) = \frac{-\big[\Pi(P_{off}) + C_{off}\big] + D_2(P_{off}, P_{on})\big[\Pi(P_{on}) - C_{on}\big]}{M_1(P_{off})\left(1 - D_2(P_{off}, P_{on})D_1(P_{on}, P_{off})\right)} \tag{4.4.31}$$

and

$$c_2(P_{on}^*, P_{off}^*) = \frac{\big[\Pi(P_{on}) - C_{on}\big] - D_1(P_{on}, P_{off})\big[\Pi(P_{off}) + C_{off}\big]}{M_2(P_{on})\left(1 - D_1(P_{on}, P_{off})D_2(P_{off}, P_{on})\right)}. \tag{4.4.32}$$

Here $D_1(P_{off}, P_{on})$ and $D_2(P_{off}, P_{on})$ are the expected discount factors given by

$$D_1(P_{on}, P_{off}) = \mathbb{E}(e^{-rS_{off}^{(1)}}|P_0 = P_{on}) = \frac{M_1(P_{on})}{M_1(P_{off})} \tag{4.4.33}$$

$$D_2(P_{off}, P_{on}) = \mathbb{E}(e^{-r\tilde{S}_{on}^{(1)}}|P_0 = P_{off}) = \frac{M_2(P_{off})}{M_2(P_{on})}. \tag{4.4.34}$$

Note that when the unit is online, the option value at P_{off}^* is a function of the foregone profit from shutting down when the price is P_{off}^* less the shut-down cost, $-\left[\Pi(P_{off}) + C_{off}\right]$, and the realised profit less the start up cost and discounted back from the random time at which the price returns to P_{on}^*, $D_2(P_{off}, P_{on})\left[\Pi(P_{off}) - C_{off}\right]$. This is scaled by the inverse of the recurrence factor, $1 - D_2(P_{off}, P_{on})D_1(P_{on}, P_{off})$, which is one minus the expected discount factor for a cycle from P_{off} to P_{on} and back to P_{off}. Similarly, when the unit is offline, the option value at P_{on}^* is a function of the realised profit from starting up, the foregone profit discounted back from the random time at which the price returns to P_{off}^* and the recurrence factor. By maximizing $c_1(P_{on}^*, P_{off}^*)$ or $c_2(P_{on}^*, P_{off}^*)$ for $P_{on}^* > P_{off}^*$, we obtain the optimal investment value.

4.4.2.1 Quadratic Input/Output Characteristics

Assuming that the amount of gas required for the production rate, Q_t, is quadratic, i.e.

$$H(Q) = a_2Q^2 + a_1Q + a_0, \tag{4.4.35}$$

we can obtain an explicit expression for the value of always being online. From (4.4.14) we find that with

$$P_{min} = p_f(a_1 + 2Q_{min}a_2), \qquad P_{max} = p_f(a_1 + 2Q_{max}a_2), \tag{4.4.36}$$

the optimal rate of production is

$$Q_t^* = \begin{cases} Q_{min}, & P_t < P_{min} \\ \dfrac{P_t - a_1 p_f}{2p_f a_2}, & P_t \in [P_{min}, P_{max}] \\ Q_{max}, & P_t > P_{max}. \end{cases} \tag{4.4.37}$$

Furthermore, for $P < P_{min}$, the value of always generating, $\Pi_{on}(P)$, is

$$Q_{min}\int_0^\infty e^{-rt}\mathbb{E}(P_t|P_0 = P)\ \mathrm{d}t - \frac{p_f}{r}H(Q_{min}) + \tilde{c}_1 M_1(P), \tag{4.4.38}$$

for $P \in [P_{min}, P_{max}]$, it is

$$\begin{aligned} &\frac{1}{4p_f a_2}\int_0^\infty e^{-rt}\mathbb{E}(P_t^2|P_0 = P)\ \mathrm{d}t \\ &-\frac{a_1}{2a_2}\int_0^\infty e^{-rt}\mathbb{E}(P_t|P_0 = P)\ \mathrm{d}t \\ &+\frac{p_f}{r}\left(\frac{a_1^2}{4a_2} - a_0\right) + \tilde{c}_2 M_1(P) + \tilde{c}_3 M_2(P), \end{aligned} \tag{4.4.39}$$

and for $P > P_{max}$,

$$Q_{max}\int_0^\infty e^{-rt}\mathbb{E}(P_t|P_0 = P)\ \mathrm{d}t - \frac{p_f}{r}H(Q_{max}) + \tilde{c}_4 M_2(P). \tag{4.4.40}$$

Here, $\tilde{c}_1$, $\tilde{c}_2$, $\tilde{c}_3$ and $\tilde{c}_4$ are determined such that $\Pi_{on}(P)$ is continuous and differentiable at $P = P_{min}$ and $P = P_{max}$.

4.4.2.2 Diffusion Process for Prices

Assuming the electricity price process does not include jumps, (4.4.28) has analytic solutions. For the Brownian Motion,

$$m_1(x) = e^{x\beta_1} \qquad\qquad m_2(x) = e^{x\beta_2}, \tag{4.4.41}$$

where $\beta_1 > 1$ and $\beta_2 < 0$ are the two solutions to the quadratic equation[4],

$$\frac{1}{2}\sigma^2\beta(\beta - 1) + \alpha\beta - r = 0, \tag{4.4.42}$$

[4] $\beta_2 < 0$ if $r > 0$ and $\beta_1 > 1$ if $r > \alpha$. In our case we require $\beta_1 > 2$ to ensure that $e^{-rt}E(P_t^2|P_0) \to 0$ for $t \to \infty$, as $e^{-rt}E(P_t^{\beta_1}|P_0) = P_0$, see Ernstsen and Misir (2016).

given by

$$\beta_{1,2} = \left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}}. \tag{4.4.43}$$

Note that $Cm_1(x)$ and $Dm_2(x)$ also satisfies (4.4.28) for $C, D \in \mathbb{R}$, but C and D cancel out in (4.4.31) and (4.4.32). Thus, we omit the constants in front of $m_1(x)$ and $m_2(x)$.

For the Ornstein-Uhlenbeck process,

$$m_1(x) = \begin{cases} U(a,b,f(x)) & x < x^* \\ -U(a,b,f(x)) + 2\frac{\Gamma(1-b)}{\Gamma(a-b+1)} M(a,b,f(x)) & x \geq x^* \end{cases} \tag{4.4.44}$$

and

$$m_2(x) = \begin{cases} -U(a,b,f(x)) + 2\frac{\Gamma(1-b)}{\Gamma(a-b+1)} M(a,b,f(x)) & x < x^* \\ U(a,b,f(x)) & x \geq x^* \end{cases} \tag{4.4.45}$$

where $a = r/2\kappa_P$, $b = 1/2$, $f(x) = \kappa_P(x - x^*)^2/\sigma^2$ and $x^* = \alpha - \sigma^2/4\kappa_P$ which we show in Lemma 4.B.9. Here we use Kummer's function

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{zu} u^{a-1} (1-u)^{b-a-1} \, \mathrm{d}u \tag{4.4.46}$$

and Tricomi's function

$$U(a,b,z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} \, \mathrm{d}t, \tag{4.4.47}$$

to define the solution to (4.4.28).

4.4.2.3 Jump-diffusion Process for Prices

To the best of our knowledge, there is no analytical solution to (4.4.28) if the electricity price process includes jumps. Thus, we solve the valuation problem numerically. We change variables in (4.4.28) by letting

$$s_1 = e^{x - x_1} \qquad\qquad s_2 = e^{-(x - x_2)} \tag{4.4.48}$$

for $m_1(x)$ and $m_2(x)$, respectively. With this change of variables, we obtain a finite domain, which simplifies the handling of the boundary condition in a finite difference method. Here, x_1 provides an upper bound on $\log(P^*_{on} + M)$ and x_2 provides a lower bound on $\log(P^*_{on} + M)$. We let

$$v_1(s) = m_1(\ln(s) + x_1) \qquad\qquad v_2(s) = m_2(-\ln(s) + x_2) \tag{4.4.49}$$

for $s \in (0,1]$ and note that this implies that $v_1(s)$ solves

$$\begin{aligned} &\mu(\ln(s)+x_1)s\frac{\partial}{\partial s}v_1(s) + \frac{1}{2}\sigma(-\ln(s)+x_1)^2 s^2 \frac{\partial^2}{\partial s^2}v_1(s) \\ &-rv_1(s) + \lambda E\left[v_1(se^{Y_1}) - v_1(s)\right] = 0 \end{aligned} \tag{4.4.50}$$

with $v_1(0) = 0$ and $v_1(1) = c$ for some $c \in \mathbb{R}$. Furthermore, $v_2(s)$ solves

$$\begin{aligned} &-\mu(-\ln(s)+x_2)s\frac{\partial}{\partial s}v_2(s) + \frac{1}{2}\sigma(-\ln(s)+x_2)^2 s^2 \frac{\partial^2}{\partial s^2}v_2(s) \\ &-rv_2(s) + \lambda E\left[v_1(se^{-Y_1}) - v_1(s)\right] = 0 \end{aligned} \tag{4.4.51}$$

with $v_2(0) = 0$ and $v_2(1) = c$ for some $c \in \mathbb{R}$.

4.4.2.4 Numerical Solution for Jump-diffusion Processes

To solve (4.4.50) and (4.4.51) we discretize the interval $[0,1]$ with $s_i = (i-1)\Delta s$ for $i = 1, \ldots, N$ with $\Delta s = 1/N$. We let V_1^i and V_2^i denote the approximation of $v_1(s_i)$ and $v_2(s_i)$ respectively for $i = 1, \ldots, N$ and approximate the derivatives by

$$\frac{\partial}{\partial s}v_1(s_i) \approx \frac{V_1^{i+1} - V_1^{i-1}}{2\Delta s} \tag{4.4.52}$$

$$\frac{\partial}{\partial s}v_2(s_i) \approx \frac{V_2^{i+1} - V_2^{i-1}}{2\Delta s} \tag{4.4.53}$$

$$\frac{\partial^2}{\partial s^2}v_1(s_i) \approx \frac{V_1^{i+1} - 2V_1^{i} + V_1^{i-1}}{\Delta s^2} \tag{4.4.54}$$

$$\frac{\partial^2}{\partial s^2}v_2(s_i) \approx \frac{V_2^{i+1} - 2V_2^{i} + V_2^{i-1}}{\Delta s^2} \tag{4.4.55}$$

for $i = 2, \ldots, N-1$. To approximate the expectation we introduce the piecewise linear approximations of $v_1(s)$ and $v_2(s)$,

$$\tilde{V}_1(s) = \sum_{i=1}^{N}(a_1(i)s + b_1(i))1_{[s_i, s_{i+1})}(s) \tag{4.4.56}$$

$$\tilde{V}_2(s) = \sum_{i=1}^{N}(a_2(i)s + b_2(i))1_{[s_i, s_{i+1})}(s) \tag{4.4.57}$$

with $s_{N+1} = \infty$, and where $1_{[s_i, s_{i+1})}(s)$ denotes the indicator function for $[s_i, s_{i+1})$. Here,

$$a_1(i) = \frac{V_1^{i+1} - V_1^{i}}{\Delta s} \qquad a_2(i) = \frac{V_2^{i+1} - V_2^{i}}{\Delta s} \tag{4.4.58}$$

$$b_1(i) = V_1^i - a_1(i)s_i \qquad b_2(i) = V_2^i - a_2(i)s_i \tag{4.4.59}$$

for $i = 1, \ldots, N-1$ and

$$a_1(N) = a_1(N-1) \qquad a_2(N) = a_2(N-1) \tag{4.4.60}$$
$$b_1(N) = b_1(N-1) \qquad b_2(N) = b_2(N-1). \tag{4.4.61}$$

Thus, we use linear interpolation between the grid points and extrapolate for $s > 1$ based on the last two grid points. We approximate the expectations of $v_1(s)$ and $v_2(s)$ by

$$E\left(\tilde{V}_1(s_j e_1^Y)\right) = \sum_{i=1}^{N} (a_1(i) s_j A_1(i,j) + b_1(i) B_1(i,j)) \tag{4.4.62}$$
$$E\left(\tilde{V}_2(s_j e^{-Y_1})\right) = \sum_{i=1}^{N} (a_1(i) s_j A_2(i,j) + b_2(i) B_2(i,j)) \tag{4.4.63}$$

for $j = 2, \ldots, N-1$, where

$$A_1(i,j) = \int_{\frac{s_i}{s_j}}^{\frac{s_{i+1}}{s_j}} z\phi_1(z)\, \mathrm{d}z \qquad B_1(i,j) = \int_{\frac{s_i}{s_j}}^{\frac{s_{i+1}}{s_j}} \phi_1(z)\, \mathrm{d}z \tag{4.4.64}$$
$$A_2(i,j) = \int_{\frac{s_i}{s_j}}^{\frac{s_{i+1}}{s_j}} z\phi_2(z)\, \mathrm{d}z \qquad B_2(i,j) = \int_{\frac{s_i}{s_j}}^{\frac{s_{i+1}}{s_j}} \phi_2(z)\, \mathrm{d}z \tag{4.4.65}$$

for $j = 2, \ldots, N-1$ and $i = 1, \ldots, N$ and where $\phi_1(z)$ and $\phi_2(z)$ are the density functions for e^{Y_1} and e^{-Y_1}, respectively. Note that we compute $O(N^2)$ integrals prior to solving the two linear systems of equations that arise from discretizing (4.4.50) and (4.4.51) at s_j for $j = 2, \ldots, N-1$ using (4.4.52)-(4.4.63).

4.4.3 Hydroelectric Power Plant

Finally, we consider a hydroelectric power plant. As for conventional production, hydro power production can be controlled. In fact, this technology is highly flexible, and production can be adjusted instantaneously without startup and shutdown costs. We assume that the hydroelectric power plant consists of a reservoir with a current storage level L_t and minimum and maximum levels given by L_{min} and L_{max}. Water is discharged from the reservoir at a rate of v_t and lead through a pipe with minimum and maximum rates being defined by the pipe and denoted by v_{min} and v_{max}. It is lead to a power station, in which the turbines convert the potential energy of the water into electrical energy. Given the efficiency of the turbines, the storage level and the discharge rate, we denote the production curve by $H(L_t, v_t)$. Finally, the reservoir receives an inflow of water I_t that stems from rainfall and melt water which can be modeled as $I_t = F^{-1}(\Phi(U_t))$. The dynamics of the storage are therefore given by $\mathrm{d}L_t = (I_t - v_t)\, \mathrm{d}t$. We model the storage level instead

of the head level, the height that the water source fall before the power is generated, to obtain simpler equations and avoid conversions even though the power output is a function of the head level, see Chen and Forsyth (2008b). We let $\mu_P(P)$, $\sigma_P(P)$ and $\gamma_P(P)$ denote the drift, diffusion and jump coefficient respectively for the price process and $\mu_I(I)$, $\sigma_I(I)$ denote the drift and diffusion coefficients for the inflow.[5]

Now, the problem of valuing the hydroelectric power plant is found by solving the following stochastic control problem, with P being the current price, I being the current inflow and L the storage level,

$$V_{hydro}(P,I,L) = \max_v \mathbb{E}\Big[\int_0^\infty e^{-rt} P_t H(L_t, v_t)\ \mathrm{d}t \Big| P_0 = P, I_0 = I, L_0 = L\Big] \tag{4.4.66}$$

$$\text{s.t. } \mathrm{d}P_t = \mu_P(P_t)\ \mathrm{d}t + \sigma_P(P_t)\ \mathrm{d}Z_t^P + \gamma_P(P_t)\ \mathrm{d}J_t, \tag{4.4.67}$$

$$\mathrm{d}I_t = \mu_I(I_t)\ \mathrm{d}t + \sigma_I(I_t)\ \mathrm{d}Z_t^U \tag{4.4.68}$$

$$\mathrm{d}L_t = (I_t - v_t)\ \mathrm{d}t, \tag{4.4.69}$$

$$L_{min} \le L_t \le L_{max}, \tag{4.4.70}$$

$$v_{min} \le v_t \le v_{max}. \tag{4.4.71}$$

To solve this problem analytically, we relax the upper and lower bounds on L_t and v_t and introduce the penalty functions

$$N_1(L) = \Theta_1 L + \Theta_2 L^2, \qquad N_2(v) = \theta_1 v + \theta_2 v^2 \tag{4.4.72}$$

with $\theta_1, \Theta_1 \in \mathbb{R}$ and $\theta_2, \Theta_2 < 0$. The expected discounted profit in the relaxed problem is then

$$\mathbb{E}\Big[\int_0^\infty e^{-rt}\Big(P_t H(L_t, v_t) + N_1(L_t) + N_2(v_t)\Big)\ \mathrm{d}t \Big| P_0 = P, I_0 = I, L_0 = L\Big], \tag{4.4.73}$$

and the associated HJB equation is

$$\begin{aligned}
&\mu_P(P)\frac{\partial}{\partial P}\tilde{V}_{hydro}(P,I,L) + \frac{1}{2}\sigma_P(P)^2\frac{\partial^2}{\partial P^2}\tilde{V}_{hydro}(P,I,L)\\
&+\mu_I(I)\frac{\partial}{\partial I}\tilde{V}_{hydro}(P,I,L) + \frac{1}{2}\sigma_I(I)^2\frac{\partial^2}{\partial I^2}\tilde{V}_{hydro}(P,I,L)\\
&+\rho\sigma_I(I)\sigma_P\frac{\partial^2}{\partial I \partial P}\tilde{V}_{hydro}(P,I,L) + \Theta_1 L + \Theta_2 L^2\\
&+\max_v\Big((f-v)\frac{\partial}{\partial L}\tilde{V}_{hydro}(P,I,L) + \theta_2 v^2 + \theta_1 v + PH(L,v)\Big)\\
&+\lambda\left[\mathbb{E}(\tilde{V}_{hydro}(P+\gamma(P)Y_1, I, L)) - \tilde{V}_{hydro}(P,I,L)\right]\\
&-r\tilde{V}_{hydro}(P,I,L) = 0,
\end{aligned} \tag{4.4.74}$$

where $\tilde{V}_{hydro}(P,I,L)$ denotes the value function for the relaxed problem.

[5] In the following we linearize the control as a function of the price. For this reason, we formulate the problem in terms of the price instead of the logarithm to the transformed price. Thus, we assume that $\partial V_{hydro}/\partial L$ is linear in P_t rather than X_t.

4.4.3.1 A linear control strategy

To simplify computations, we assume that the inflow is deterministic and write $I_t = f_t$ as well as $V_{hydro}(P, L)$. We consider the linear production curve $H(L, v) = \eta_1 v_t + \eta_0$.[6] With this assumption, the first order condition for the maximization problem in (4.4.74) is

$$v^* = \frac{-\theta_1 - \eta_1 + \frac{\partial}{\partial L} V_{hydro}(P, L)}{2\theta_2} \tag{4.4.75}$$

which is the optimal discharge rate since $\theta_2 < 0$. We further assume that the discharge rate is a linear function of the price and the storage level, i.e.

$$v_t = d_1 + d_2 P_t + d_3 L_t, \tag{4.4.76}$$

which corresponds to assuming that the marginal value of water, $\frac{\partial}{\partial L} V_{hydro}(P, L)$ is linear in the price and storage level. We insert the linearized value of water,

$$\begin{aligned} &\frac{\partial}{\partial L} V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}} + (P - \overline{P}) \frac{\partial^2}{\partial L \partial P} V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}} \\ &+ (L - \overline{L}) \frac{\partial^2}{\partial L^2} V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}} \end{aligned} \tag{4.4.77}$$

in (4.4.75) and by collecting the terms in (4.4.76), we obtain

$$d_1 = \frac{-\theta_1 + \left(\frac{\partial}{\partial L} - \overline{P} \frac{\partial^2}{\partial L \partial P} - \overline{L} \frac{\partial^2}{\partial L^2} \right) V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}}}{2\theta_2} \tag{4.4.78}$$

$$d_2 = \frac{-\eta_1 + \frac{\partial^2}{\partial L \partial P} V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}}}{2\theta_2} \tag{4.4.79}$$

$$d_3 = \frac{\frac{\partial^2}{\partial L^2} V_{hydro}(P, L)|_{P=\overline{P}, L=\overline{L}}}{2\theta_2}. \tag{4.4.80}$$

With $v_t = d_1 + d_2 P_t + d_3 L_t$, where d_1, d_2 and d_3 satisfies (4.4.78), (4.4.79) and (4.4.80), it follows that

$$\begin{aligned} V_{hydro}(P, L) = \mathbb{E}\Bigg[\int_0^\infty e^{-rt} \Bigg(&\Big[d_3^2 \theta_2 + \Theta_2 \Big] L_t^2 \\ &+ \Big[d_2^2 \theta_2 + d_2 \eta_1 \Big] P_t^2 + \Big[2 d_2 \theta_2 + \eta_1 \Big] d_3 P_t L_t \\ &+ \Big[2 d_1 d_2 \theta_2 + d_1 \eta_1 + d_2 \theta_1 + \eta_0 \Big] P_t \\ &+ \Big[2 d_1 d_3 \theta_2 + d_3 \theta_1 + \Theta_1 \Big] L_t + d_1^2 \theta_2 + d_1 \theta_1 \Bigg) \, \mathrm{d}t \Bigg]. \end{aligned} \tag{4.4.81}$$

[6] The model can be extended to production curves with an L_t term and a $v_t L_t$ term, but for hydroelectric plants with low head flexibility, the additional accuracy is negligible.

Furthermore,

$$\mathrm{d}L_t = d_3\left(\frac{f_t - d_1 - d_2 P_t}{d_3} - L_t\right)\,\mathrm{d}t \tag{4.4.82}$$

and hence, by Lemma 4.B.3,

$$L_t = L_0 e^{-d_3 t} + \int_0^t e^{-d_3(t-s)}\left(f_s - d_1 - d_2 P_s\right)\,\mathrm{d}s. \tag{4.4.83}$$

We therefore find that

$$\begin{aligned}
\frac{\partial}{\partial L} V_{hydro}(P,L)|_{P=\overline{P},L=\overline{L}} &= (2d_1 d_3\theta_2 + d_3\theta_1 + \Theta_1)\int_0^\infty e^{-(r+d_3)t}\,\mathrm{d}t \\
&+ (2d_2\theta_2 + \eta_1)d_3\int_0^\infty e^{-(r+d_3)t}\mathbb{E}(P_t|\overline{P})\,\mathrm{d}t \\
&+ 2(d_3^2\theta_2 + \Theta_2)\int_0^\infty \overline{L}e^{-(r+2d_3)t}\,\mathrm{d}t \\
&+ 2(d_3^2\theta_2 + \Theta_2)\int_0^\infty e^{-(r+d_3)t}\int_0^t e^{-d_3(t-s)}\left(f_s - d_1\right)\,\mathrm{d}s\,\mathrm{d}t \\
&- d_2 2(d_3^2\theta_2 + \Theta_2)\int_0^\infty e^{-(r+d_3)t}\int_0^t e^{-d_3(t-s)}\mathbb{E}(P_s|\overline{P})\,\mathrm{d}s\,\mathrm{d}t,
\end{aligned} \tag{4.4.84}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial L\partial P} V_{hydro}(P,L)|_{P=P_0,L=L_0} = \\
&+ (2d_2\theta_2 + \eta_1)\int_0^\infty e^{-(d_3+r)t}\frac{\partial}{\partial P}\mathbb{E}(P_t|P)|_{P=\overline{P}}\,\mathrm{d}t \\
&- 2d_2(d_3^2\theta_2 + \Theta_2)\int_0^\infty e^{-(r+d_3)t}\int_0^t e^{-d_3(t-s)}\frac{\partial}{\partial P}\mathbb{E}(P_s|P)_{P=\overline{P}}\,\mathrm{d}s\,\mathrm{d}t,
\end{aligned} \tag{4.4.85}$$

and

$$\frac{\partial^2}{\partial L^2} V_{hydro}(P,L)|_{P=\overline{P},L=\overline{L}} = 2(d_3^2\theta_2 + \Theta_2)\int_0^\infty e^{-(2d_3+r)t}\,\mathrm{d}t. \tag{4.4.86}$$

Finally, (4.4.80) and (4.4.86) imply that

$$d_3 = \left(\frac{\Theta_2}{\theta_2} + d_3^2\right)\int_0^\infty e^{-(2d_3+r)t}\,\mathrm{d}t \tag{4.4.87}$$

and hence d_3 is the positive solution to

$$d_3^2 + r d_3 - \frac{\Theta_2}{\theta_2} = 0. \tag{4.4.88}$$

Moreover, (4.4.84) and (4.4.85) are linear in d_1 and d_2 and, therefore, (4.4.78) and (4.4.79) can be solved by applying numerical integration and using the closed-form expressions for $\mathbb{E}(P_t|P_0)$ determined in Section 4.2.

To evaluate the control strategy from the relaxed problem, we use that

$$L_t = L_0 e^{-d_3 t} + \int_0^t e^{-d_3(t-s)} \left(f_s - d_1 - d_2 P_s\right) \, \mathrm{d}s \tag{4.4.89}$$

such that

$$v_t = d_1 + d_2 P_t + d_3 \left(L_0 e^{-d_3 t} + \int_0^t e^{-d_3(t-s)} \left(f_s - d_1 - d_2 P_s\right) \, \mathrm{d}s \right) \tag{4.4.90}$$

and find that the value of the hydroelectric plant is

$$\begin{aligned}
&E\left[\int_0^\infty e^{-rt} P_t(\eta_1 v_t + \eta_0) \, \mathrm{d}t \middle| P_0 = P, L_0 = L\right] \\
&= \int_0^\infty e^{-rt} \eta_1 d_2 E[P_t^2 | P_0 = P] \, \mathrm{d}t \\
&+ \int_0^\infty e^{-rt} (\eta_1 d_1 + \eta_0) E\left[P_t | P_0 = P\right] \, \mathrm{d}t \\
&+ \int_0^\infty e^{-(r+d_3)t} \eta_1 (d_3 L - d_1) E\left[P_t | P_0 = P\right] \, \mathrm{d}t \\
&+ \int_0^\infty e^{-rt} \int_0^t e^{-d_3(t-s)} \eta_1 d_3 E\left[P_t f_s - d_2 P_t P_s | P_0 = P\right] \, \mathrm{d}s \, \mathrm{d}t.
\end{aligned} \tag{4.4.91}$$

We note that the constant d_1 enters two terms, where the first is the value of the constant term of the discharge rate and the second is the reduction in value due to the accompanying negative impact of a decrease in storage level on the discharge rate. Similarly, d_2 enters the first and last term. The first term is the value of increasing the discharge rate at a given time in response to an increase in the price, while the last term is the corresponding negative impact of a decrease in storage level on the future discharge rate. Finally, d_3 is the speed of mean reversion of the storage level, i.e. this constant determines how the discharge rate changes when the storage level deviates from the mean reversion level

$$\frac{f_t - d_1 - d_2 P_t}{d_3}. \tag{4.4.92}$$

Thus, d_3 has an impact on the value of increasing the discharge rate in response to an increase in the price.

It should be remarked that in the relaxed problem the upper and lower bounds on the discharge rate and the storage level may be violated, and so the value of the hydroelectric plant may be higher than that of the original problem. The risk of violating the constraints can be managed by tuning the parameters θ_1, θ_2, Θ_1 and Θ_2. Moreover, this overestimation is counterbalanced by the restriction to a linear discharge rate which implies a lower value of the hydroelectric plant. To further improve the estimate of the plant value, it could be taken into account that the marginal value of water depends on the future inflow, either by periodically updating the constant term of the discharge rate or by extending the discharge rate to be a linear function of the current flow f_t.

4.5 Case Study

For the case study, we consider a realistic instance of each type of generation. We calibrate the price and wind models to hourly data for 7 years from the 22nd of January 2004 to 31st of December 2010. The spot price is from the western part of Denmark (DK1) in DKK/MWh and the wind speeds are from Sindal in the northern part of Denmark in m/s. We let the bound M be 1000 and choose the discount rate r such that the discounted value over 30 years and infinity are similar, i.e. such that $\int_0^\infty e^{-rt}\,\mathrm{d}t - \int_0^{30} e^{-rt}\,\mathrm{d}t = 0.01$, which implies that $r = 0.2061$.

4.5.1 Price Models

For comparison of the Brownian motion and the Ornstein-Uhlenbeck process, we let $\mu = 0$ and choose P_0 as the average price in the calibration period. We calibrate the Ornstein-Uhlenbeck process with and without jumps based on

$$\ln(P_{t_{i+1}} + M) - e^{-\kappa_P(t_{i+1}-t_i)} \ln(P_{t_i} + M), \tag{4.5.1}$$

see Appendix 4.A for details. Simulations indicate that this gives a reasonable fit, see Figures 4.1 to 4.3. For the Brownian Motion we do not calibrate the volatility based on the differences $\ln(P_{t_{i+1}} + M) - \ln(P_{t_i} + M)$, as the price data is rather volatile and tends to be mean reverting. Attempting to calibrate to this results in an extremely volatile price process. Instead we match the continuous arithmetic average of the variance of the Brownian motion to the a empirical variance, see Appendix 4.A.[7] The model parameters are given in Table 4.1.

μ	σ	P_0
0	0.05535	300.62

	α	κ_P	σ	P_0	λ
$\lambda = 0$	7.1705	670.1594	3.5137	205.86	-
$\lambda > 0$	7.1697	670.1594	0.6315	205.86	6502.355

$\lambda^{(1)}$	$\eta^{(1)}$	$\nu^{(1)}$	$\lambda^{(2)}$	$\eta^{(2)}$	$\nu^{(2)}$
5618.197	0.001861	0.0237	884.158	-0.009685	0.0900

Table 4.1: Parameters for the price models.

[7] The variance for the model based on the Brownian motion converges to infinity as t increases, thus, it is necessary to choose a finite time period where the model variance is matched to the empirical variance.

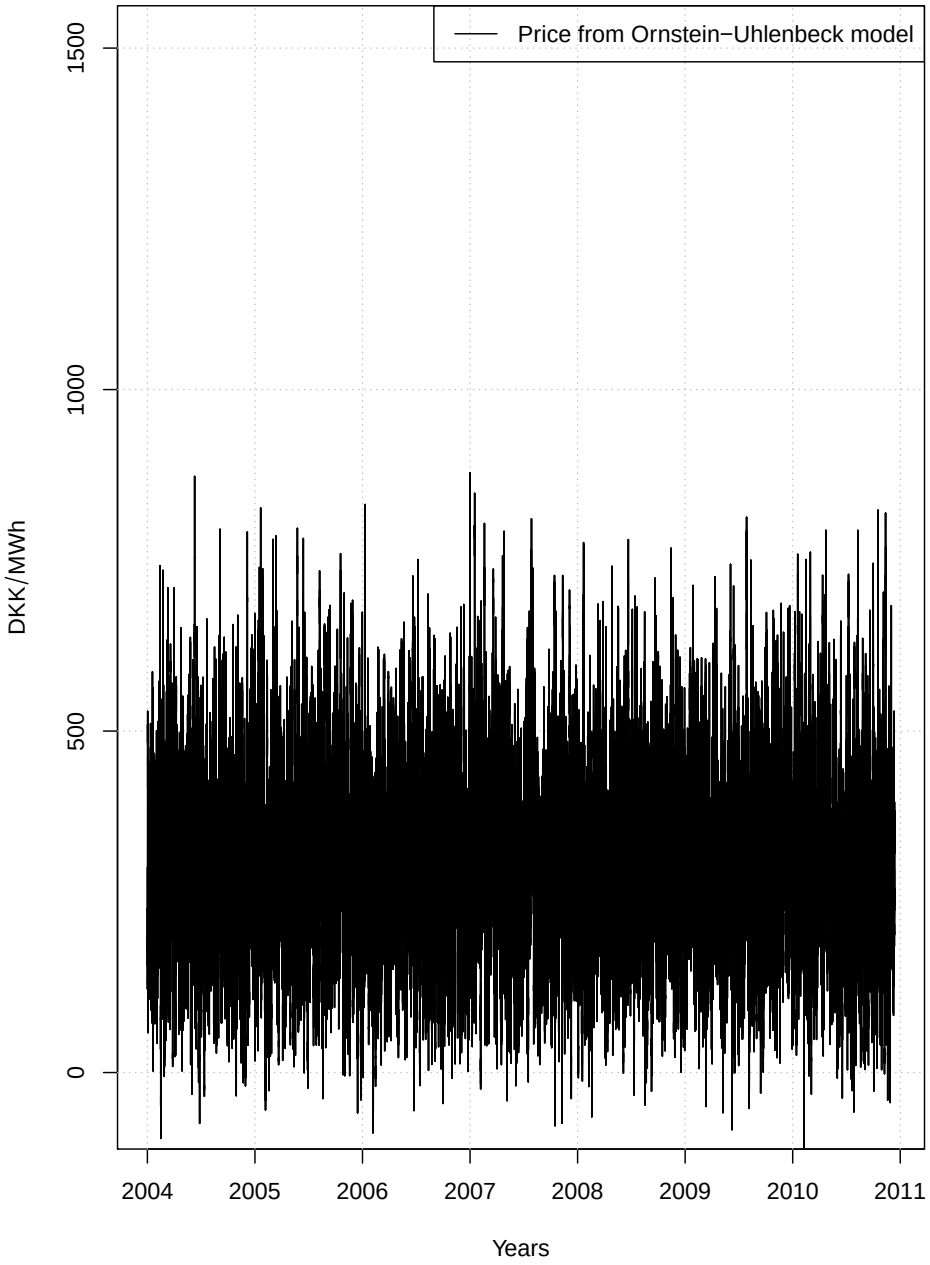


Figure 4.1: Sample path of the spot price from the Ornstein-Uhlenbeck model.

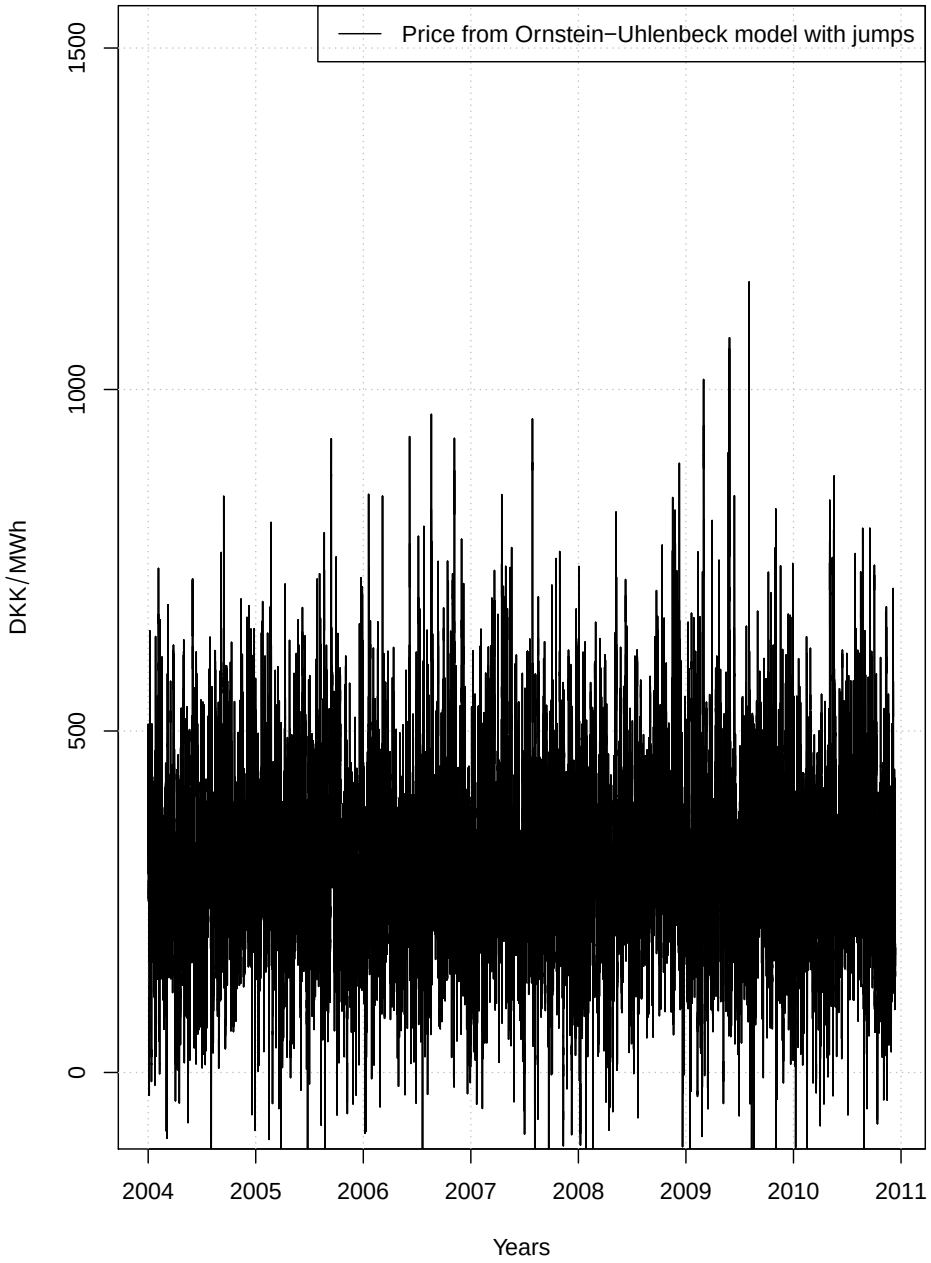


Figure 4.2: Sample path of the spot price from the Ornstein-Uhlenbeck model with jumps.

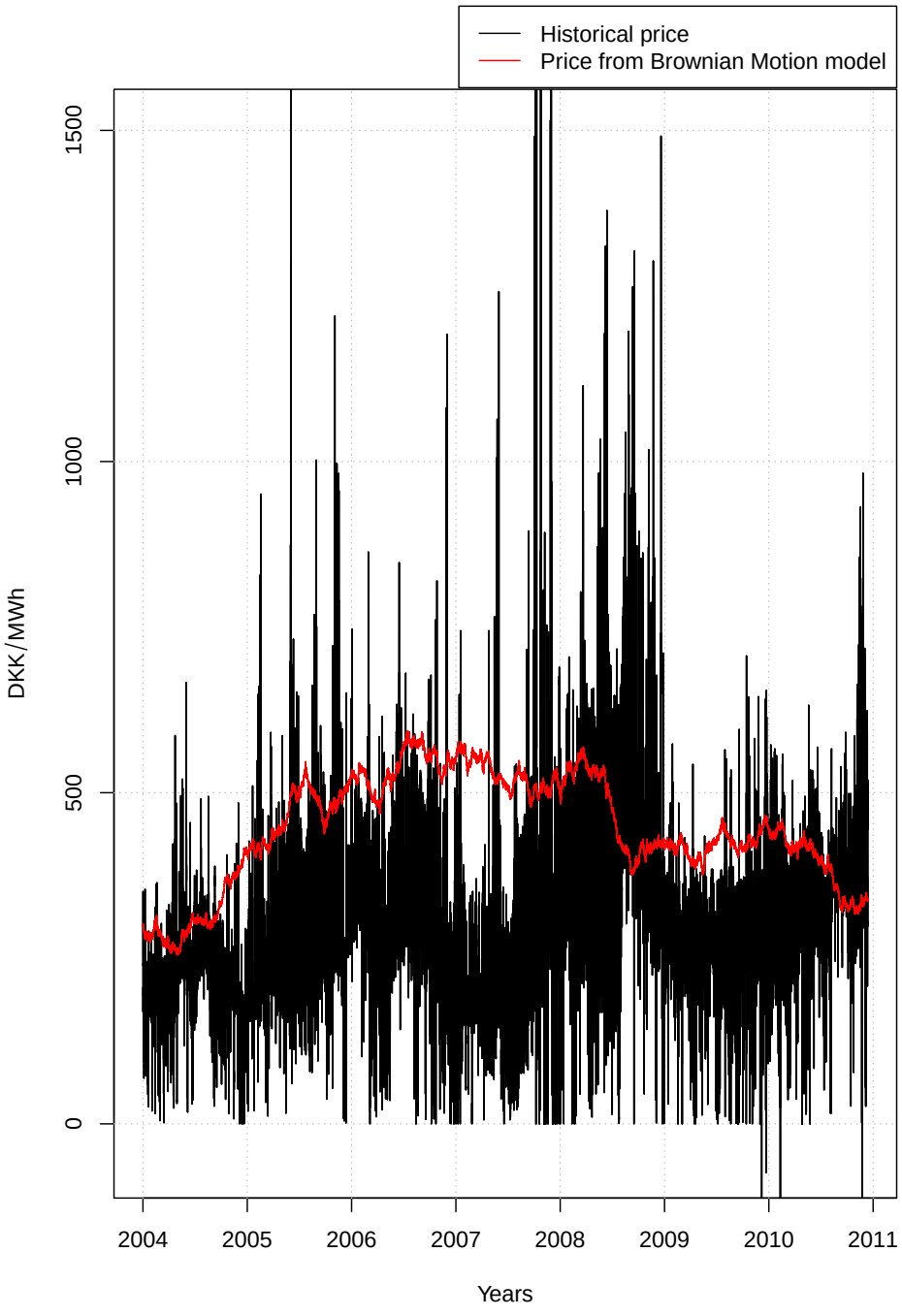


Figure 4.3: Historical spot price and sample path of the spot price from the Brownian motion model.

4.5.2 Wind Turbine

We consider a wind turbine with the data from Vestas (2016) and obtain the model parameters in Table 4.2 from calibration to the wind data. We likewise determine the correlation coefficient ρ for the price and wind drivers from the data in Table 4.3.[8] See Section 4.A for calibration details.

β	κ_U	η	w_0	w_1	w_2	h_{12}/N_{hours}
-	-	m/s	m/s	m/s	m/s	MW
2.2566	178.5204	8.2405	3.5	15	25	3

Table 4.2: Wind turbine and wind speed data. $N_{hours} = 8760$, the number of hours in a year

Model for X_t	ρ
BM	-2.0381^*
OU	-0.1503
OU with jumps	-0.7600

Table 4.3: Correlation coefficients.

4.5.3 Gas Fired Power Plant

The gas price is taken as the average of the daily data from 28th of January 2016 to 12th of April 2016 of GPN Spot index (2016). For the gas fired power plant we use parameters from Tseng and Barz (2002) where it is assumed that the input/output characteristics of the generating unit is

$$H(Q) = a_0 + a_1 Q + a_2 Q^2 \tag{4.5.2}$$

with parameters in Table 4.4 and Table 4.5.

Q_{min}	Q_{max}	C_{off}	C_{on}	p_f
MW	MW	DKK	DKK	$DKK/MMBTu$
250	750	6 175	21 125	29.63

Table 4.4: Data for gas fired power plant

[8] Note that $\rho < -1$ for the BM model indicating that the decrease in value cannot be explained by correlation in this model.

a_0	a_1	a_2
$MMBTu$	$MMBTU/MW$	$MMBTu/MW^2$
600	9.121	0.00131

Table 4.5: Data for generating unit for gas fired power plant.

We maximize $c_1(P_1 + \delta^2, P_1)$ from (4.4.31) as a function of δ and P_1 using the Nelder-Mead algorithm, see Nelder and Mead (1965), such that $P^*_{on} = P_1 + \delta^2 \geq P_1 = P^*_{off}$.

4.5.4 Hydroelectric Power Plant

For the hydroelectric power plant, we base the parameter values on those of Chen and Forsyth (2008a). To simplify the expressions we let L_t denote the total volume of the reservoir with surface area a and head L_t/a, assuming the reservoir has a cylinder shape. The linearized production curve is

$$H(v, L) = \eta_0 + \eta_1 v, \tag{4.5.3}$$

which has an average absolute error of 2.21% and a maximum absolute error of 4.8% compared to the non-linear production curve in Chen and Forsyth (2008a).

We choose the coefficients of the penalty functions such that the marginal profit required to exceed the bounds for the storage level is $\eta_1 \tilde{P}_L$ and the marginal penalty required to exceed the discharge level is $\eta_1 \tilde{P}_v$, i.e. such that

$$\tilde{P}_L \eta_1 = \Theta_1 + 2\Theta_2 L_{min}, \qquad -\tilde{P}_L \eta_1 = \Theta_1 + 2\Theta_2 L_{max} \tag{4.5.4}$$

and

$$\tilde{P}_v \eta_1 = \theta_1 + 2\theta_2 v_{min}, \qquad -\tilde{P}_v \eta_1 = \theta_1 + 2\theta_2 v_{max}. \tag{4.5.5}$$

Thus, the penalty functions attain their maxima at $L = (L_{max} + L_{min})/2$ and $v = (v_{min} + v_{max})/2$. For our instance we choose $\tilde{P}_L = 4S$ and $\tilde{P}_v = E(P_1) + 4S$, where S is an estimate for the standard deviation. For details, see Appendix 4.A.5. The parameters are given in Table 4.6-Table 4.9.

η_0/N_{hours}	$\eta_1 N_{secs}/N_{hours}$
MW	MW/m^3
−23.0437	0.9392

Table 4.6: Parameters for generating unit. $N_{hours} = 8760$, $N_{secs} = 3600 N_{hours}$.

a	L_{min}/a	L_{max}/a	L_0/a
m^2	m	m	m
$1.8 \cdot 10^6$	90	94	92

Table 4.7: Data for hydroelectric power plant.

v_{min}/N_{secs}	v_{max}/N_{secs}	f/N_{secs}
m^3/s	m^3/s	m^3/s
40	150	60

Table 4.8: Discharge limits and inflow for hydroelectric power plant. $N_{secs} = 31536000$.

Model for X_t	$\tilde{P}_L$	$\tilde{P}_v$
BM	540.61	841.23
OU	471.36	696.36
OU with jumps	439.94	664.94

Table 4.9: Marginal penalty for exceeding water level bound and discharge level bound for various models.

4.6 Results

4.6.1 Future Markets

With increasing penetration of renewables in the power system, future market conditions may differ significantly from the current. We investigate the effects of an increase in the electricity price level, an increase in the price volatility, the inclusion of reversion to mean, price jumps and correlation between the electricity price and renewable production.

In our benchmark case we assume that the electricity price and the weather conditions are governed by two OU processes or a Brownian motion and an OU process.

We report the main results below. Table 4.10 lists the value of the wind turbine, the gas-fired power plant and the hydroelectric plant under various models for the dynamics of the price process. Recall that renewable production is assumed non-controllable, whereas the production from the conventional and storage technologies can be controlled. Figures 4.4 and 4.5 show the control strategies for the gas-fired plant and Figures 4.6 and 4.7 show the control strategies for the hydroelectric power plant when these are implemented for a

simulated price path for 4 weeks.[9] Figure 4.8 shows the strategy implemented over a 7 year period for three models. The startup and shutdown trigger levels and the linear discharge function are provided in Appendix 4.C.4 and 4.C.5. As the control strategy is stochastic, we provide an estimated 95% confidence interval and expected value for the first year head level and discharge rate in Appendix 4.C.5.

Model for X_t	Wind	Gas	Hydro
BM	13.688	1 476.470	425.616
OU	13.688	1 196.999	466.055
OU with jumps	13.688	1 226.261	461.317

Table 4.10: The value of the wind turbine, the gas-fired power plant and the hydroelectric plant under various price models. Values are given in mio. DKK.

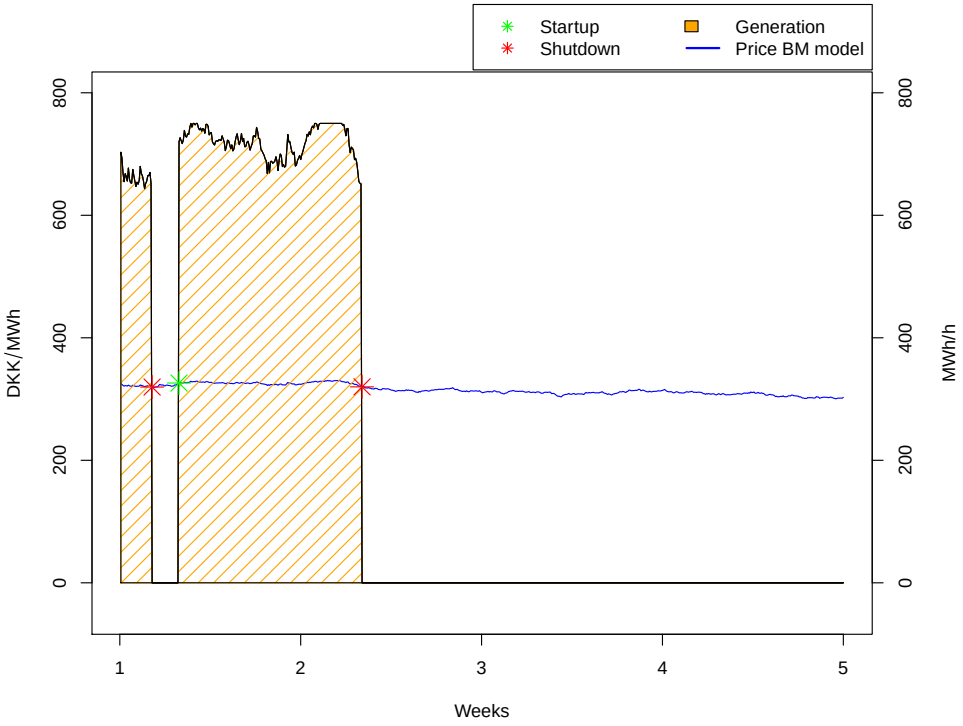


Figure 4.4: 4 weeks of the control strategy for the gas-fired power plant in response to a sample path of the price from the GBM model.

[9]The strategy for the model based on the Brownian motion consists of an almost constant discharge rate equal to the inflow and therefore also constant water level.

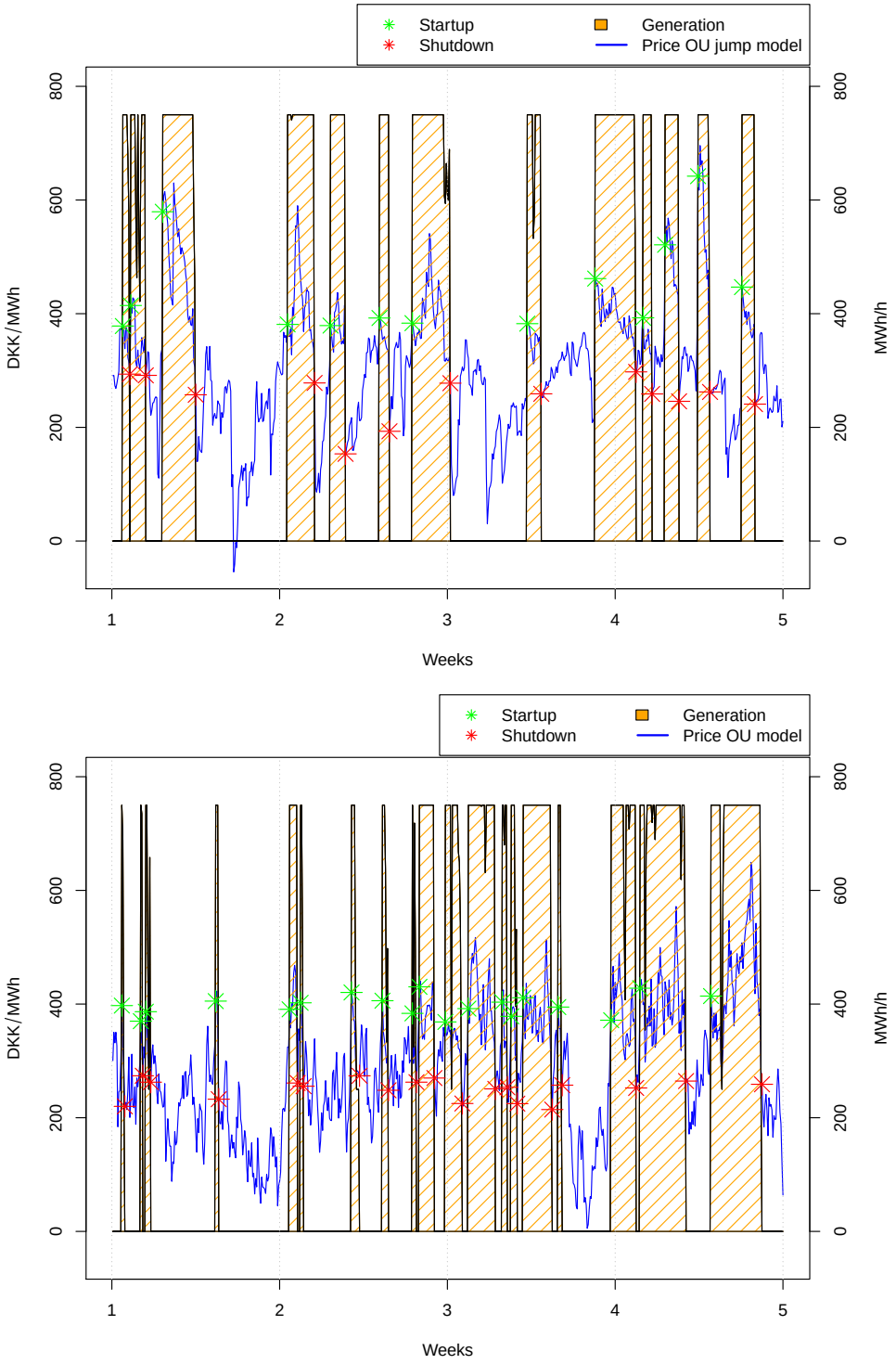


Figure 4.5: 4 weeks of the control strategy for the gas-fired power plant in response to a sample path of the price from the OU models without and with jumps.

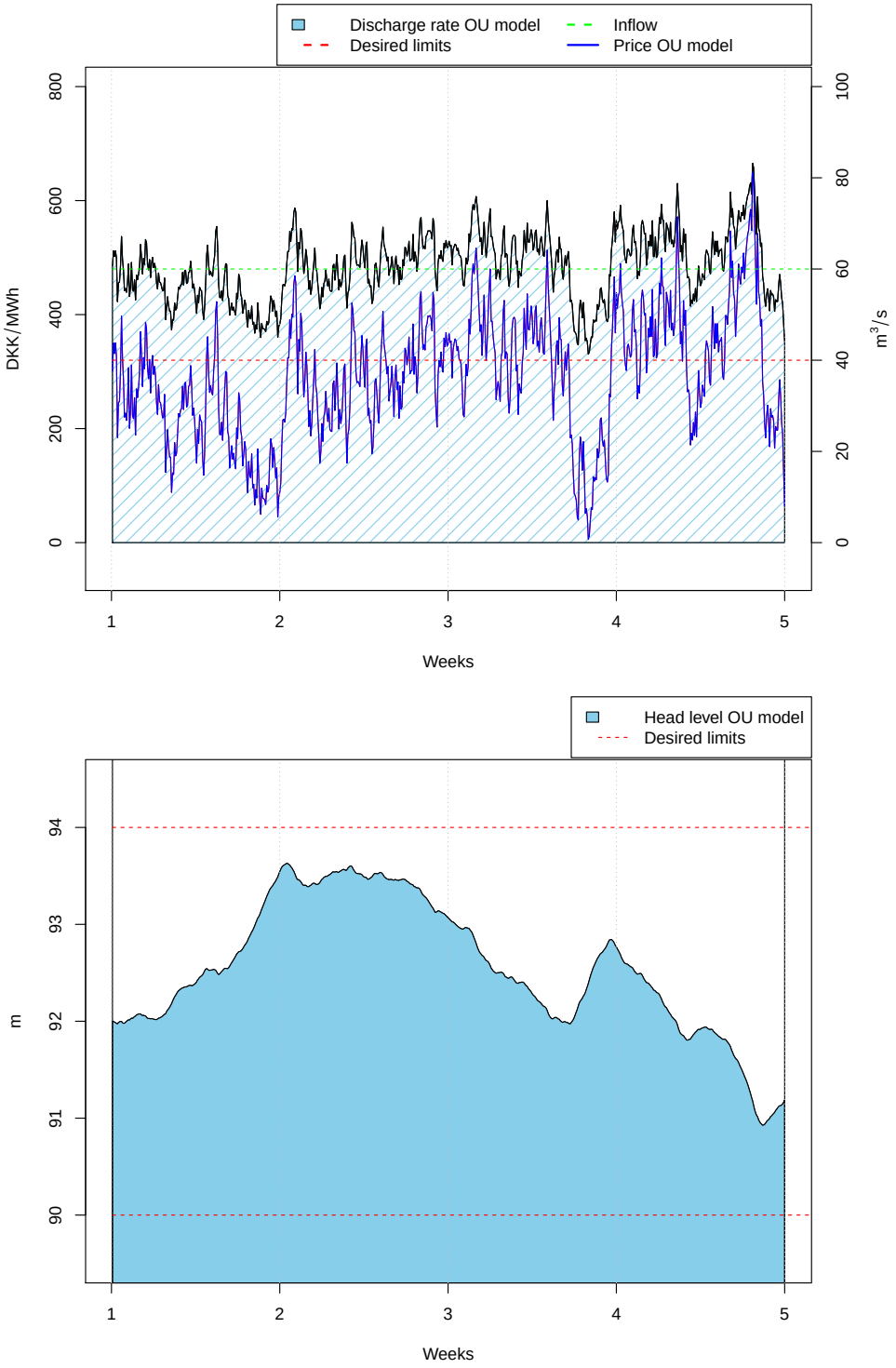


Figure 4.6: 4 weeks of the discharge strategy and head level for the hydroelectric power plant in response to a sample path of the price from the OU model.

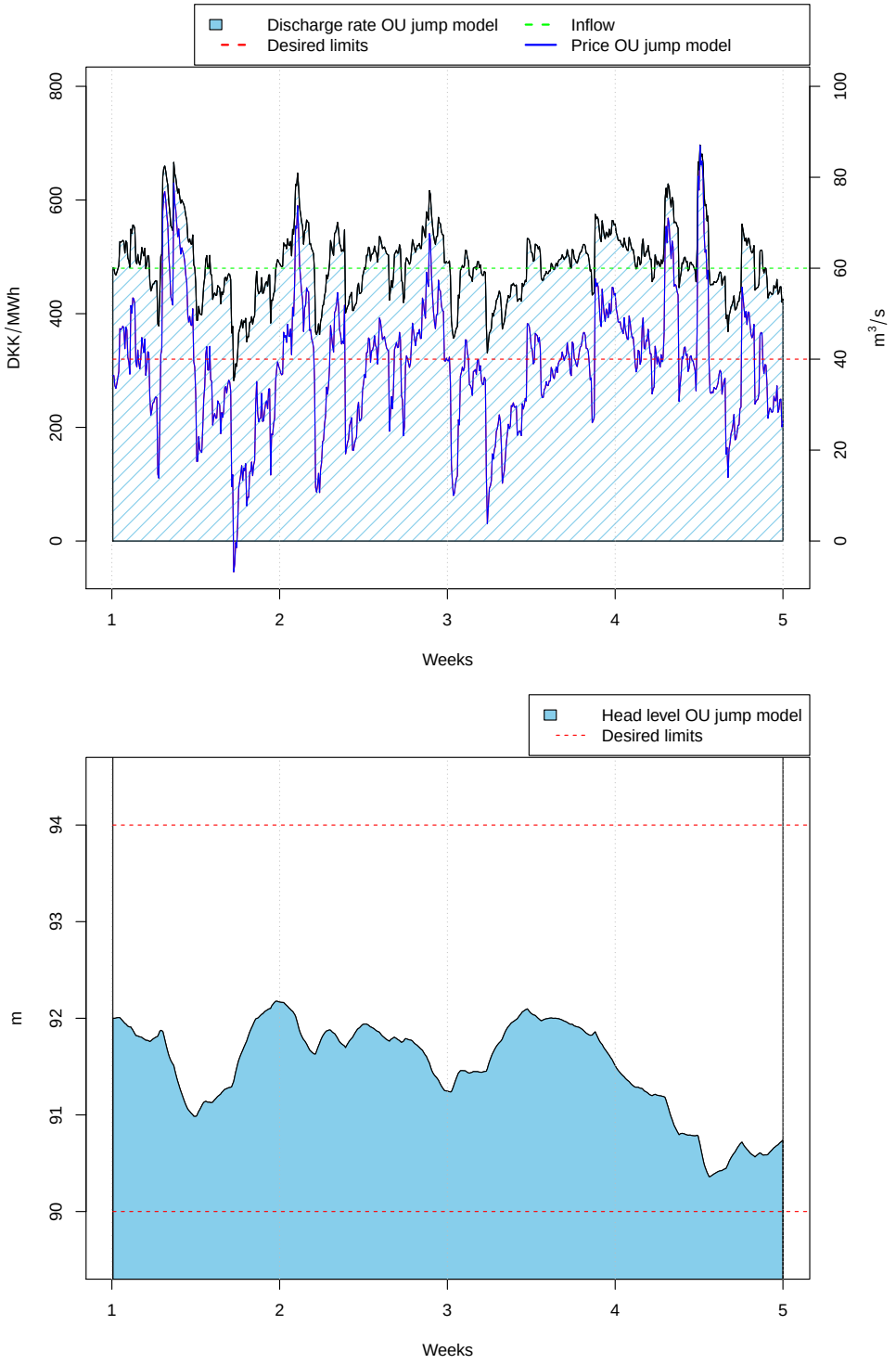
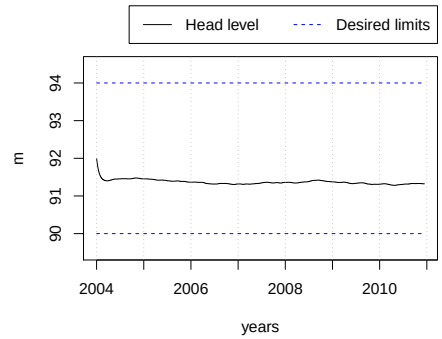
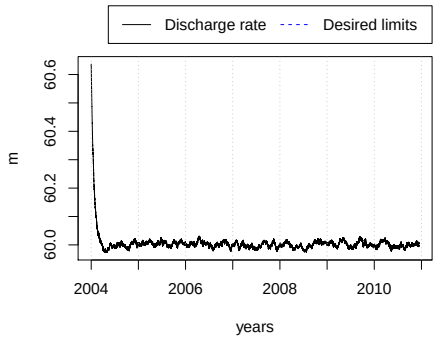


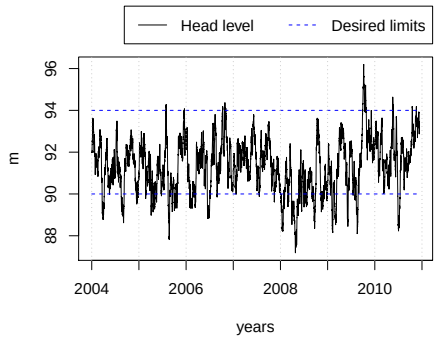
Figure 4.7: 4 weeks of the discharge strategy and head level for the hydroelectric power plant in response to a sample path of price from the OU model with jumps.



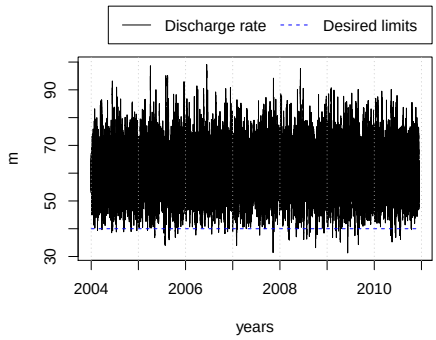
(a) Head level for BM model



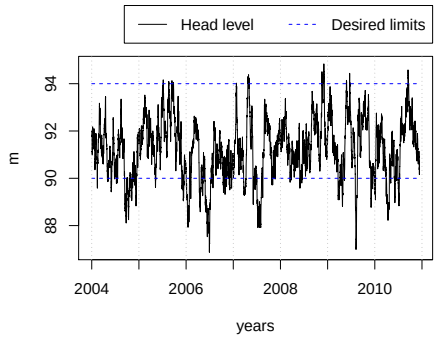
(b) Discharge rate for BM model



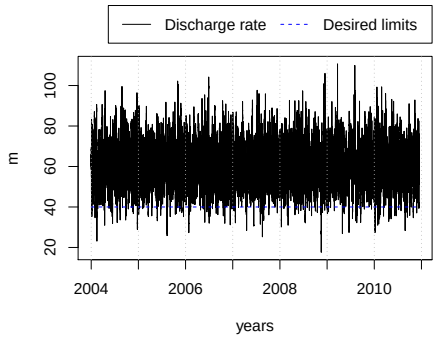
(c) Head level for OU model



(d) Discharge rate for OU model



(e) Head level for OU jump model



(f) Discharge rate for OU jump model

Figure 4.8: Strategies in response to hourly data for 7 years

4.6.2 Preliminary Model Analysis

A number of observations can be made on the basis of the analytical or partly analytical solutions to the valuation problems.

First, the inclusion of mean reversion and jumps in the price process is irrelevant for the inflexible renewable plant. The reason is that the plant value only depends on the price model through

$$E(P_t h(W_t)) = E(P_t)E(h(W_t)) + Cov(P_t h(W_t)). \tag{4.6.1}$$

Thus, only the expected price, the expected production and the covariance between the two determine the value of generation.

Next, for the gas-fired power plant, the value depends on the optimal generation strategy and is determined by

$$E(P_t Q_t^* - p_f H(Q_t^*)) \tag{4.6.2}$$

and the expected discount factors

$$D_1(P_{on}, P_{off}) = \mathbb{E}(e^{-rS_{off}^{(1)}} | P_0 = P_{on}) \tag{4.6.3}$$

$$D_2(P_{off}, P_{on}) = \mathbb{E}(e^{-r\tilde{S}_{on}^{(1)}} | P_0 = P_{off}) \tag{4.6.4}$$

that are used to compare the expected gain from immediate start up to the cost of future shut down. The Brownian motion has extremely long expected return times for the optimal triggers, i.e. $\mathbb{E}(S_{off}^{(1)}|P_0 = P_{on}) = 2.68$ years and $\mathbb{E}(\tilde{S}_{on}^{(1)}|P_0 = P_{on}) = 2.67$ years[10], which results in very few cycles and thus a higher value than for the other models. Due to these long expected return times, the trigger levels are very close to the marginal cost. For the Ornstein-Uhlenbeck process without jumps, the expected return times for the optimal triggers are $\mathbb{E}(S_{off}^{(1)}|P_0 = P_{on}) = 10.14$ hours and $\mathbb{E}(\tilde{S}_{on}^{(1)}|P_0 = P_{on}) = 13.91$ hours. Thus, there are significantly more cycles and thereby higher total startup and shutdown costs than for the Brownian motion. Since the expected value generated in the online period has to offset the cycle costs, the triggers are likewise further from the marginal cost.

For the hydroelectric plant with deterministic inflow and linear control strategy, the value depends on the expected spot price, $E(P_t)$, the second moment of the spot price, $E(P_t^2)$ and the covariance of the price process at different points in time, i.e. the autocovariance, $E(P_t P_s)$. In particular, the expected spot price determines the value of the constant part of generation, the second moment determines the positive value of responding to an increase in the price and the autocovariance determines the cost of restoring the reservoir level. The autocovariance of the Brownian motion decreases much slower than that of the Ornstein-Uhlenbeck process, and hence, storage deviations are more beneficial with the Ornstein-Uhlenbeck models, which is why the value of the hydroelectric power plant is higher. To investigate the value of flexibility, we decompose the value of the hydroelectric plant into the value of

[10]The expectation is found through differentiation of the moment generating function at $r = 0$.

discharging the inflow and the value of adjusting discharge rate to the price and storage levels. The adjusting discharge rate, $\overline{v}_t$, is given by

$$\overline{v}_t = d_1 - f_t + d_2 P_t + d_3 L_t. \tag{4.6.5}$$

The value of the inflow depends only on the expected spot price, whereas the value of flexibility also depends on second moment and the autocovariance. In our case study, the main part of the value stems from the value of inflow which is the same for all price models, whereas the value of flexibility differ significantly as can be seen in Table 4.11.

Model for X_t	Value of flexibility (mio. DKK)
BM	0.098
OU	40.819
OU with jumps	37.138

Table 4.11: The value of flexibility, i.e. the value of adjusting the discharge rate, for the hydroelectric power plant under various price models.

In the following sections we further analyse our models on the basis of the results from the case study.

4.6.2.1 Including mean reversion

The real options literature more often than not assumes that the dynamics of the market follow a geometric Brownian motion. This may be a reasonable assumption in the very long term, especially for inflexible generation as the value only depends on the expected instantaneous value of production. Commodity prices, however, are often argued to be mean reverting, see Schwartz and Smith (2000) and Lucia and Schwartz (2002) for electricity prices in particular, as in the long run such prices should reflect marginal costs of production. As expected, by replacing the Brownian Motion by the Ornstein-Uhlenbeck process the value of the renewable plant is unaffected. For flexible production, the value of investment changes significantly. The value of the gas fired power plant decreases by 18.9% as including mean reversion introduces more start-ups and shutdowns. In contrast, the value of the hydroelectric power plant increases by 9.5% as the inclusion of mean reversion allows the hydroelectric power plant to benefit from making small adjustments to production without costs. Furthermore, the value of flexibility increases from being 0.02% of the total value to 8.7%.

4.6.2.2 Including jumps

With intermittency of renewable production, sudden changes in the price are likely to occur. Furthermore, with increased penetration, such changes become

larger and/or more frequent. We, therefore, investigate the effect of introducing jumps in the price process and the sensitivity with respect to parameters of the jump component.

As the jumps does not alter the expected prices, the value of inflexible production does not change. In contrast, the value of the gas-fired power plant increases by 2.5%. The reason is that the price process has lower variance and higher positive skewness when including the jumps. This in turn increases the optimal shutdown and startup triggers and slightly increase the value as the positive effect of increased prices in peak periods and fewer startups outweigh the negative effect of the longer offline period. For the hydroelectric power plant, the value decreases by 1.0% as the value of flexibility decreases with a lower variance and shorter time periods of high prices. It should be remarked that this conclusion is highly parameter dependent. On one hand, a hydroelectric power plant capable of discharging at peak periods and storing inflows the rest of the time would benefit significantly from price jumps. On the other hand, as jump times are random a plant with less flexibility in storage capacity benefits more from a constant variance than occasional jumps.

4.6.2.3 Average price level

Historically, the electricity price has risen due to increasing demand of electricity. Since renewables have very low marginal costs, for wind power production the marginal costs are close to zero, it is expected that there will be more periods with very low prices with an increased amount of renewable generation. However, in periods without renewable generation, for instance due to low wind speeds, the prices will be significantly higher, as flexible generation with higher marginal costs must be used. Most likely, the increasing renewable penetration will therefore be accompanied by an increase in the capacity of flexible generation, but a reduction in the total capacity of non-flexible conventional generation. Thus, we expect that median future prices decrease, but that expected future prices increase due to increasing costs of flexible generation, see Dong Energy (2015) and Energinet.dk (2015). We, therefore, explore the effect of an increase by 10% in the average price level, cf. P_0 for the GBM and α for the OU process. As expected, an increased average price results in a higher investment value, see Figure 4.9, Figure 4.10 and Figure 4.11. The effect is almost the same for both price processes. For an increase in the average level of 10%, the value of the renewable investment will increase by 10.3%, assuming that the correlations between renewable generation and electricity prices will not change, which is highly unlikely. The values of the conventional power plant and the storage plant likewise increase and the trigger levels for startup and shutdown decrease. However, whereas the value increases linearly for the renewable plant, the marginal effect is increasing for the technologies that can reduce or shutdown production. The increase in investment value is 32.3-40% and 9-10% for the gas fired power plant and hydroelectric power

plant, respectively. For the gas fired power plant the additional increase is due to the increased uptime and the reduced number of shutdown predicted by the model. Furthermore, we see from Figure 4.21 in Appendix 4.C.4 that an increase in the average price levels lowers both startup and shutdown triggers for the gas fired power plant, as it is expected to be more profitable when the plant is online.

Altogether, our results indicate the profitability of renewable investments is highly sensitive to changes in the average price level, whereas conventional generation benefits from being able to start up or shut down production and thereby increase the upside of increasing price levels, but reduce the downside of decreasing price levels. It should be noted that the profitability of gas fired power plants is highly dependent on future gas prices compared to electricity prices.

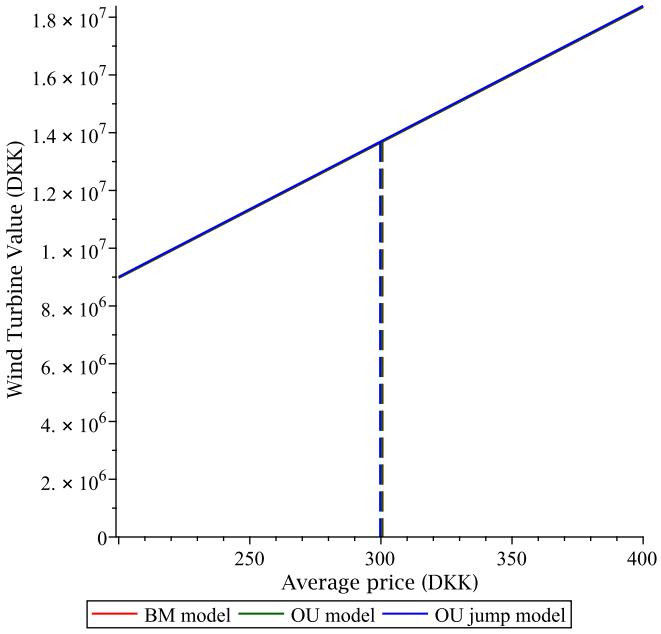


Figure 4.9: Impact of average electricity price on wind turbine value

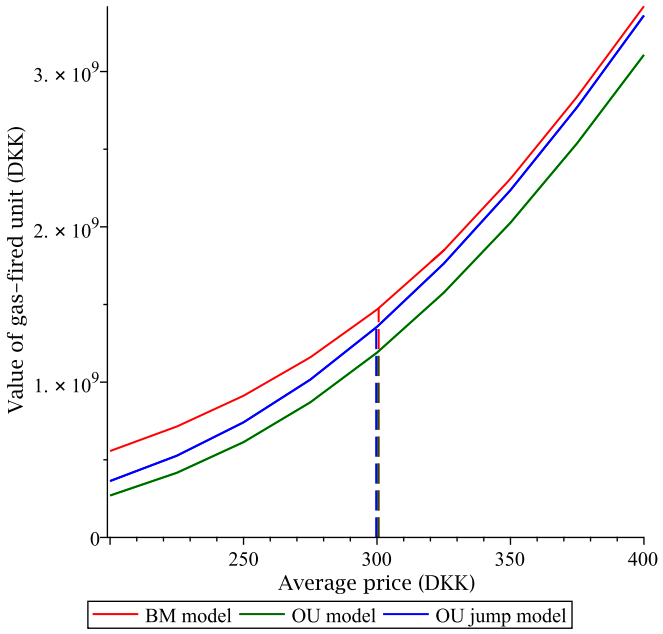


Figure 4.10: Impact of average electricity price on value of gas fired power plant

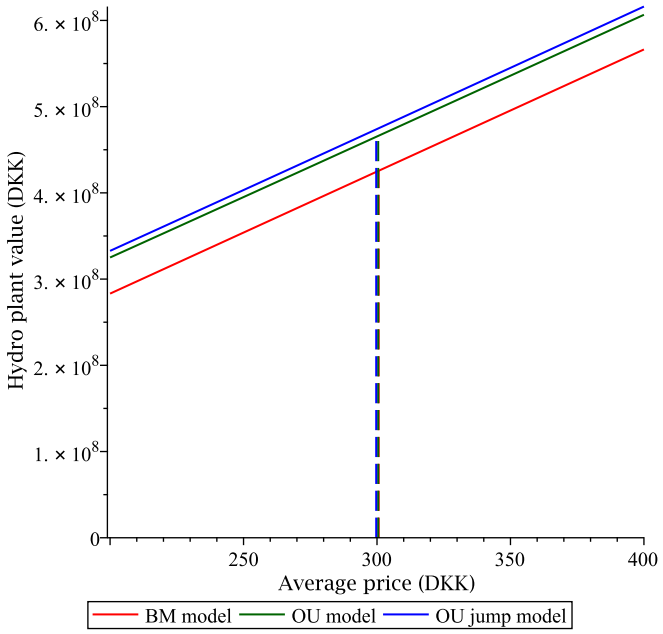


Figure 4.11: Impact of average electricity price on value of hydroelectric power plant

4.6.2.4 Price volatility

For most renewable technologies, production is intermittent, that is, highly varying and largely inflexible. An example is wind power production that is directly determined by wind speeds, unless completely shut down. As a result, we expect the electricity price volatility to increase with increasing levels of renewable penetration.

Figure 4.12 shows that a higher volatility returns a slightly lower value of the renewable investment. For an increase in the standard deviation of the price by 10% percent, the value of the renewable investment decreases by only 0.4-0.7%, except for the OU process with jumps. The decrease in value is due to a negative correlation between price and renewable production, which makes high production levels occur with low prices. For the OU process with jumps, only the diffusion part is correlated with the wind power production and the correlation is higher than without jumps. Thus, an increase in the volatility of the diffusion part will have a low impact on standard deviation of the price compared to the impact on the total value.

As opposed to renewable production, the conventional power plant is able to benefit from price variations by starting up and shutting down in response to high and low prices, respectively, cf. Figure 4.13. This is also reflected by the investment triggers for startup and shutdown, which increase and decrease with the volatility. As a result, the value of investment increases by 11.8-12.9%. For the storage plant, the value of discharging the inflow remains constant and the value of flexibility increases, but as the value of flexibility is only 8-9% of the total value, the total increase is only 1.1-1.7%. However, the value of flexibility increases by 13.7%-13.9% for the OU process if the diffusion term increases the standard deviation and by 21.1% if the jump intensity increases the volatility.

In conclusion, a higher future volatility will slightly decrease the value of renewable investments. Whether the same holds for other technologies or whether they may in fact benefit depends on their flexibility of production, as for example reflected by the costs of starting up for the conventional plant and the ramp rate and head flexibility of the storage plant.

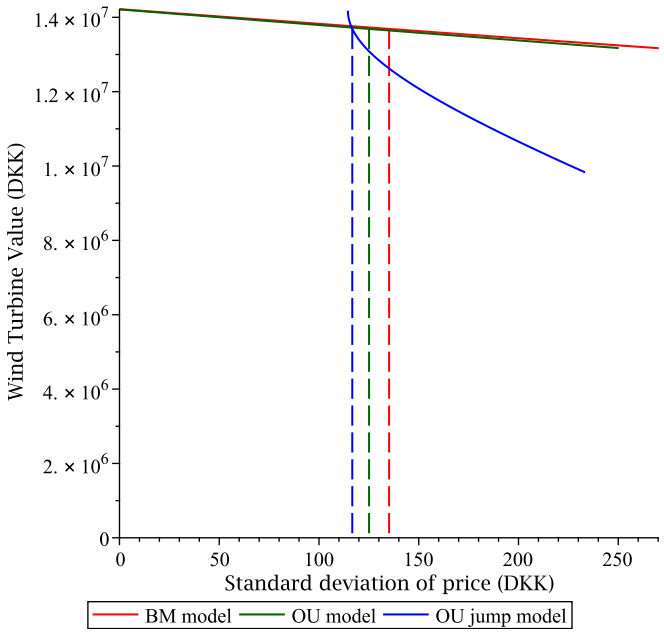


Figure 4.12: Impact of standard deviation on wind turbine value

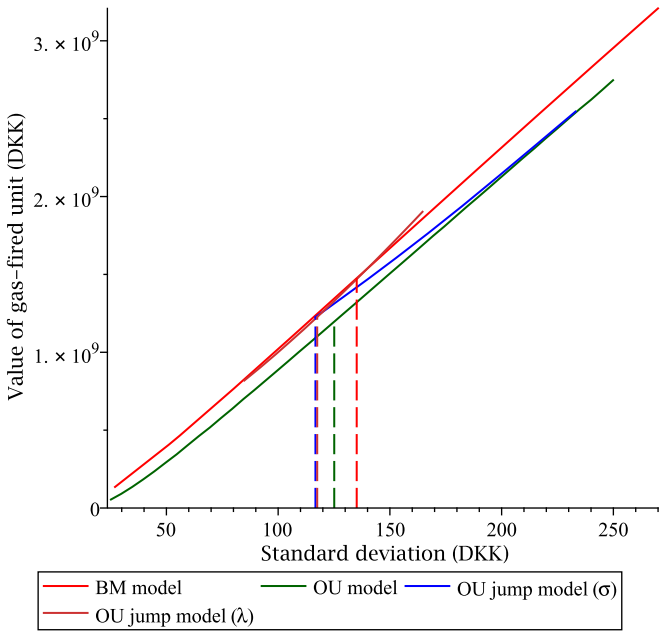


Figure 4.13: Impact of standard deviation on value of gas fired power plant

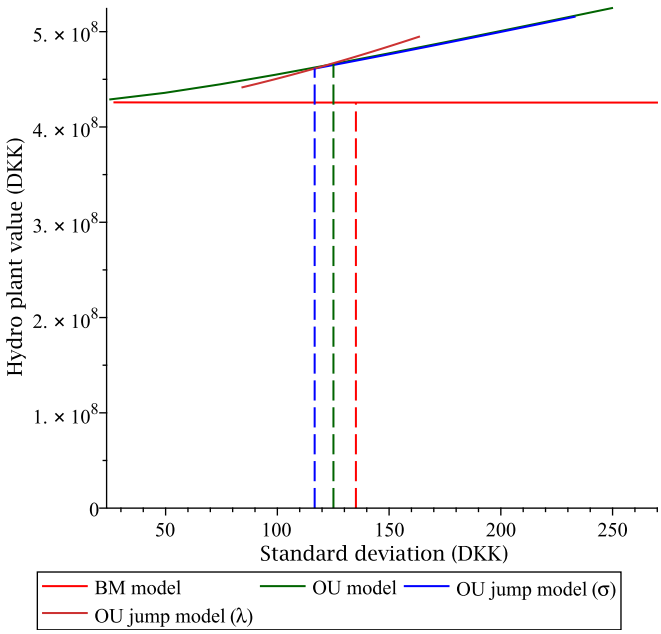


Figure 4.14: Impact of standard deviation on value for hydroelectric power plant

4.6.2.5 Correlation between production and price

As already mentioned, renewable production is negatively correlated to the electricity price. A higher penetration will further reduce the correlation. This may in turn reduce the value of renewable investment as high production is partially offset by lower prices. The conventional power plant and the hydroelectric power plant will, however, be unaffected by the correlation.

We quantify the effects of a change in correlation for an Ornstein-Uhlenbeck process, as the model based on the Brownian Motion cannot explain the correlations as estimates are not in [-1,1]. For each percentage point the correlation between price and renewable production decrease the value of renewable investment decreases by 0.3%, see Figure 4.15. Therefore, correlation only has a moderate impact on the profitability of such projects, ceteris paribus. However, we expect the correlation to become more negative and the standard deviation to increase, which will impact the value of wind power generation, as

$$E(P_t Q_t) = E(P_t)E(Q_t) + Cor(P_t, Q_t)sd(P_t)sd(Q_t). \tag{4.6.6}$$

Thus, an increase in standard deviation by 10% and an increase in the negative correlation between price and production from -11% to -33% will reduce the value of wind power generation by 8.2%.

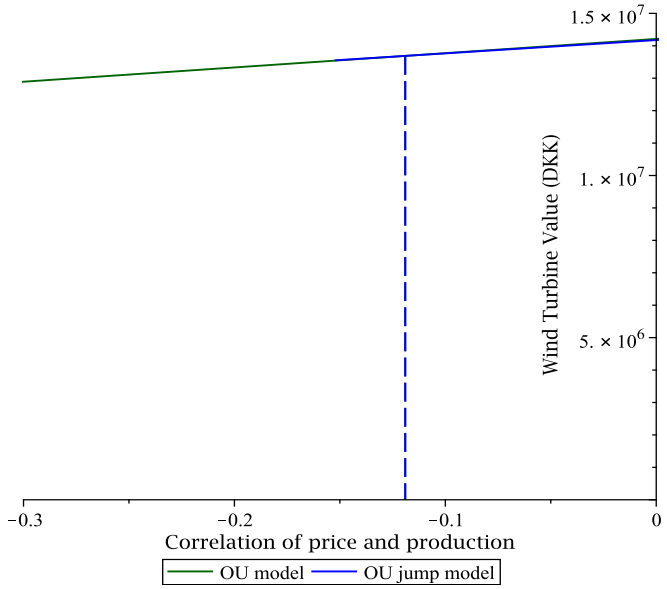


Figure 4.15: Impact on value of wind turbine of correlation between wind power generation and electricity price.

4.7 Conclusion

In this paper we develop continuous-time stochastic control models for valuation and operation of three different types of power plants in an electricity market: a renewable power plant, a conventional power plant and a storage plant.

We show how to derive analytical or partly analytical solutions to our models for valuation under uncertainty in electricity prices, both when prices are assumed to be driven by a Brownian Motion and an Ornstein-Uhlenbeck process with and without jumps. Our models further account for the special characteristics of the technologies, such as a non-normal distribution of renewable production, startup and shutdown costs of conventional units and reservoir dynamics that depend on the discharge strategy for a hydroelectric plant.

We quantify the impact of future market conditions, and in particular, a change in one or more market parameters such as the average price level, the price volatility or the correlation between renewable production and price. Our results demonstrate that the value of renewable investment may be significantly affected by an increase or a decrease in the average price level, with a one-to-one percentage change. In case the future price level becomes much

lower than the current, the profitability of such projects thereby depends on either technology maturing to ensure a sufficient reduction in investment costs or society providing subsidies. As expected, the value of renewable investment is relatively unaffected by changes in price volatility, ceteris paribus. We show, however, that a change in both volatility and correlation may erode the value, which means that the drawbacks of renewable generation are to some extend faced by the renewable producers themselves. In contrast, conventional power plants are less vulnerable to decreases in the price level, but can benefit greatly from increases in the price level and volatility by adjusting production accordingly. This finding becomes increasingly important with the expected rise in future renewable generation, as the intermittency requires an accompanying increase in flexible generation capacity.

To simultaneously capture both short term and long term dynamics of the electricity price and more accurately value operation and investment, an obvious extension to the wind turbine model and hydro power plant model is a two factor model. Moreover, our results are based on a change of exogenous price parameters, which corresponds to assuming perfect competition in the market. Future research could account for strategic behavior.

We find that the value of the wind turbine decreases by 0.4-5.6% if the standard deviation of the electricity price increases by 10 % depending on the model and that for each percentage point the negative correlation between wind power generation and electricity price drops, the value decreases by 0.35-0.38%. Furthermore, a 10% increase in average electricity price increases the value of the wind turbine by 10%. However, we expect the negative impact of increased negative correlation and standard deviation to outweigh the increase in average prices, as an increase in negative correlation increases the impact of increased standard deviation. For the gas fired unit a 10% increase in standard deviation increases the value by 10-13%, while a 10% increase in average electricity price increases the value by 32-40%. This is the case as the gas-fired unit benefits from increased peak prices without suffering the negative impact of lower off-peak prices. Furthermore, the increase in average electricity price increases the time the generator is active and reduces the number of shutdowns-startup cycles. The value of the hydroelectric power plant also increase by 10% with a 10% increase in average price, but as the hydroelectric power plant was assumed to have relatively low head flexibility, an increase in volatility by 10% only increases the value by 1.1-1.7%. However, by decomposing the discharge strategy of the hydroelectric power plant in a balancing part that discharges the inflow and a flexible part that reacts to price changes, we find that the value of the flexible part increases by by 13.7%-13.9% by a 10% increase in standard deviation through the diffusion term, while an increase in the standard deviation of 10% through the jump intensity causes a 21.1% increase in the value.

Appendix 4.A Calibration

4.A.1 Calibration of the Models Based on the Ornstein-Uhlenbeck Process without Jumps

Let $(p_i)_{i=1,\dots,N}$ denote the hourly day-a-head prices for the period. We let $(x_i)_{i=1,\dots,N}$ denote the logarithm of the transformed price, such that $x_i = \log(p_i + M)$ for $i = 1, \dots, N$. The corresponding random variables, X_i, satisfy,

$$X_{i+1} \stackrel{d}{=} aX_i + m + b\epsilon_i \quad \text{for } i = 1, \dots, N-1, \tag{4.A.1}$$

and $a = e^{-\kappa\Delta}$, $\Delta = 1/N_{hour}$, $N_{hour} = 24 \cdot 365$. Furthermore $m = (\alpha - \frac{\sigma^2}{4\kappa_P})(1 - e^{-\kappa_P\Delta})$ and $b^2 = \sigma^2 \frac{1-e^{-2\kappa_P\Delta}}{2\kappa_P}$ and ϵ_i independent and identically distributed with $\epsilon_1 \sim \mathcal{N}(0,1)$. We use ordinary least squares to get the estimators,

$$\hat{a} = \frac{\sum_{i=1}^{N-1}(x_{i+1} - \overline{x})(x_i - \tilde{x})}{\sum_{i=1}^{N-1}(x_i - \tilde{x})^2}, \qquad \hat{m} = \overline{x} - \hat{a}\tilde{x}, \tag{4.A.2}$$

where

$$\overline{x} = \frac{1}{N-1}\sum_{i=2}^{N} x_i, \qquad \tilde{x} = \frac{1}{N-1}\sum_{i=1}^{N-1} x_i. \tag{4.A.3}$$

Now

$$\hat{\kappa}_P = \frac{-\ln(\hat{a})}{\Delta}, \qquad \hat{b}^2 = \frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1} - \hat{m} - \hat{a}x_i)^2. \tag{4.A.4}$$

Finally σ and α can be estimated as

$$\hat{\sigma} = \sqrt{\frac{2\hat{\kappa}_P}{1 - e^{-2\hat{\kappa}_P\Delta}}\hat{b}^2}, \qquad \hat{\alpha} = \frac{m}{(1 - e^{-\hat{\kappa}_P\Delta})} + \frac{\hat{\sigma}^2}{4\hat{\kappa}_P}. \tag{4.A.5}$$

4.A.2 Calibration of the Models Based on the Ornstein-Uhlenbeck Process with Jumps

In the model with jumps we have

$$X_{i+1} \stackrel{d}{=} aX_i + m + \epsilon_i \tag{4.A.6}$$

with $a = e^{-\kappa_P\Delta}$, $\Delta = 1/N_{hour}$, $N_{hour} = 24 \cdot 365$. However, $m = (\alpha - \frac{\sigma^2}{4\kappa_P} - \lambda k_2)(1 - e^{-\kappa_P\Delta})$ and

$$\epsilon_i = \sigma \int_{t_i}^{t_{i+1}} e^{-\kappa(t-v)}\, \mathrm{d}Z_v^P + \sum_{i=N_{t_i}+1}^{N_{t_{i+1}}} e^{-\kappa_P T_i} Y_i \tag{4.A.7}$$

with $t_i = i\Delta$ for $i = 0, \ldots, N$ and T_i is the time of the i'th jump. We assume that there are two types of jumps, $Y_i^{(1)} \sim \mathcal{N}(\eta^{(1)}, {\nu^{(1)}}^2)$ and $Y_i^{(2)} \sim \mathcal{N}(\eta^{(2)}, {\nu^{(2)}}^2)$ such that

$$\epsilon_i = \sigma \int_{t_i}^{t_{i+1}} e^{-\kappa(t-v)} \, \mathrm{d}Z_v^P + \sum_{i=N_{t_i}^{(1)}+1}^{N_{t_{i+1}}^{(1)}} e^{-\kappa_P T_i} Y_i^{(1)} + \sum_{i=N_{t_i}^{(2)}+1}^{N_{t_{i+1}}^{(2)}} e^{-\kappa_P T_i} Y_i^{(2)}, \tag{4.A.8}$$

where $N_t^{(1)}$ and $N_t^{(2)}$ are independent compound Poisson processes with intensity $\lambda^{(1)}$ and $\lambda^{(2)}$. This can be expressed in the setup with Y_i where, $\lambda = \lambda^{(1)} + \lambda^{(2)}$ and for $c \geq 0$,

$$\mathbb{E}(e^{cY_i}) = \frac{\lambda_1}{\lambda}\mathbb{E}(e^{cY_1}) + \frac{\lambda_2}{\lambda}\mathbb{E}(e^{cY_2}). \tag{4.A.9}$$

Using the conditional moment generating function for the dampened jump process from Lemma 4.B.2 we find that

$$\mathbb{E}(\epsilon_i) = \lim_{u \searrow 0} \left(\frac{\partial}{\partial u} \exp\left(\lambda \int_{e^{-\kappa_P \Delta}}^1 \frac{\theta(uz) - 1}{\kappa_P z} \, \mathrm{d}z \right) \right) \tag{4.A.10}$$

$$= \frac{1 - e^{-\kappa_P \Delta}}{\kappa_P} \sum_{j=1}^2 \lambda^{(j)} \eta^{(j)} \tag{4.A.11}$$

and

$$\begin{aligned} Var(\epsilon_i^2) &= \sigma^2 \frac{1 - e^{-2\kappa_P \Delta}}{2\kappa_p} \\ &\quad + \lim_{u \searrow 0} \left(\frac{\partial^2}{\partial u^2} \exp\left(\lambda \int_{e^{-\kappa_P \Delta}}^1 \frac{\theta(uz) - 1}{\kappa_P z} \, \mathrm{d}z \right) \right) - \mathbb{E}(\epsilon_i)^2 \end{aligned} \tag{4.A.12}$$

$$= \frac{1 - e^{-2\kappa_P \Delta}}{2\kappa_P} \left(\sigma^2 + \sum_{j=1}^2 \lambda^{(j)} ({\eta^{(j)}}^2 + {\nu^{(j)}}^2) \right). \tag{4.A.13}$$

We apply the expectation maximization algorithm from Ernstsen (2016) assuming that

$$\epsilon_i \overset{d}{=} V_i + \sum_{i=\tilde{N}_i^{(1)}+1}^{\tilde{N}_{i+1}^{(1)}} G_i^{(1)} + \sum_{i=\tilde{N}_i^{(2)}+1}^{\tilde{N}_{i+1}^{(2)}} G_i^{(2)} \tag{4.A.14}$$

with $(G_i^{(j)})_{i=1,\ldots,N}$ are independent identically distributed with $G_1^{(j)} \sim \mathcal{N}(\tilde{\eta}^{(j)}, (\tilde{\nu}^{(j)})^2)$ for $j = 1, 2$. Furthermore $\tilde{N}^{(j)}$ are compound Poisson processes with intensity $\tilde{\lambda}_j$ for $j = 1, 2$ and $(V_i)_{i=1,\ldots,N}$ are independent and identically distributed

with $V_1 \sim \mathcal{N}(m, s^2)$. The time change from years to hours implies that the hourly mean reversion factor $\hat{\kappa}_P\Delta$, the hourly volatility is $\sigma\sqrt{\Delta}$ and the hourly intensity for $N_t^{(j)}$ is $\lambda^{(j)}\Delta$. Thus,

$$\sigma^2 = s^2 \frac{2\hat{\kappa}_P}{1 - e^{-2\hat{\kappa}_P\Delta}} \tag{4.A.15}$$

$$\lambda_j = \tilde{\lambda}_j/\Delta, \quad j = 1, 2 \tag{4.A.16}$$

$$\eta^{(j)} = \tilde{\eta}^{(j)} \frac{\hat{\kappa}_P\Delta}{1 - e^{-\hat{\kappa}_P\Delta}}, \quad j = 1, 2 \tag{4.A.17}$$

$$\nu^{(j)^2} = \left((\tilde{\eta}^{(j)})^2 + (\tilde{\nu}^{(j)})^2\right) \frac{2\hat{\kappa}_P\Delta}{1 - e^{-2\hat{\kappa}_P\Delta}} - \eta^{(j)^2}, \quad j = 1, 2 \tag{4.A.18}$$

$$\alpha = m/(1 - e^{-\hat{\kappa}_P\Delta}) + \frac{\sigma^2}{4\hat{\kappa}_P} + (\lambda_1 + \lambda_2)k_2. \tag{4.A.19}$$

where $k_2 = \int_0^1 \frac{\theta(z)-1}{\hat{\kappa}_P z}\, \mathrm{d}z$ with

$$\theta(z) = \frac{\lambda_1}{\lambda} e^{\eta^{(1)}z + \frac{1}{2}z^2(\nu^{(1)})^2} + \frac{\lambda_1}{\lambda} e^{\eta^{(1)}z + \frac{1}{2}z^2(\nu^{(1)})^2}. \tag{4.A.20}$$

4.A.3 Calibration of the Models Based on the Brownian Motion

We calibrate σ, by matching the continuous arithmetic average of the variance variance,

$$\frac{1}{\hat{T}} \int_0^{\hat{T}} \mathbb{E}(P_s^2) - \mathbb{E}(P_s)^2 \, \mathrm{d}s \tag{4.A.21}$$

to the empirical variance

$$\frac{1}{N-1} \sum_{n=1}^{N} (p_i - \overline{p})^2. \tag{4.A.22}$$

4.A.4 Calibration of Model for Wind Speeds

Let $(w_i)_{i=1,\ldots,N}$ denote the hourly wind speeds and $(p_i)_{i=1,\ldots,N}$ denote the hourly prices. We use moment matching and determine β and η such that

$$\frac{1}{N}\sum_{i=1}^{N} w_i = \eta\Gamma(\frac{1+\beta}{\beta})\frac{1}{N}\sum_{i=1}^{N} w_i^2 = \eta^2\Gamma(\frac{1+\beta}{\beta}) \tag{4.A.23}$$

where $\Gamma(x)$ is the Γ-function. We define the distribution independent values u_i using the inverse to the distribution function for the normal distribution, Φ^{-1} and the distribution function for the Weibull distribution with the estimates for β and η found through moment matching.

$$u_i = \Phi^{-1}(F(w_i)), \quad \text{for } i = 1, \ldots, N. \tag{4.A.24}$$

We estimate κ_U as for the Ornstein-Uhlenbeck model for the price using linear regression of $(u_i)_{i=2,\dots,N}$ on $(u_i)_{i=1,\dots,N-1}$ setting $\kappa_U = \frac{-\ln(\hat{a})}{\Delta}$, where $\hat{a}$ is the estimate for the linear coefficient. We choose ρ in each of the models such that the instantaneous value of production from each of the models matches average value of production, i.e. such that

$$\lim_{t\to\infty} \mathbb{E}(P_t h(W_t)) = \frac{1}{N}\sum_{i=1}^{N} h(w_i)p_i. \tag{4.A.25}$$

4.A.5 Determination of Penalty Functions

We require that the penalty functions attain has marginal costs of exceeding the bounds for the water level of $\eta\tilde{P}_L$ and marginal costs of exceeding the discharge level of $\eta\tilde{P}_v$. This implies that

$$\Theta_2 = -\frac{\tilde{P}_L}{L_{max}-L_{min}}, \qquad \Theta_1 = \frac{\tilde{P}_L(L_{max}+L_{min})}{L_{max}-L_{min}} \tag{4.A.26}$$

and

$$\theta_2 = -\frac{\tilde{P}_v}{v_{max}-v_{min}}, \qquad \theta_1 = \frac{\tilde{P}_v(v_{max}+v_{min})}{v_{max}-v_{min}}. \tag{4.A.27}$$

For the models based on the Ornstein-Uhlenbeck process,

$$S = \lim_{t\to\infty} \sqrt{\mathbb{E}(P_t^2) - \mathbb{E}(P_t)^2}. \tag{4.A.28}$$

For the models based on the Brownian Motion, the standard deviation converges to infinity, thus we use the quantity

$$S = \sqrt{\frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}(P_t^2) - \mathbb{E}(P_t)^2 \,\mathrm{d}t}, \tag{4.A.29}$$

with $\tilde{T} = 7$, the length of the calibration period.

Appendix 4.B Lemmas

4.B.1 Moment Generating Functions

Lemma 4.B.1. *Let $X \sim \mathcal{N}(\mu,\sigma^2)$ then $\mathbb{E}(e^{uX}) = e^{u\mu+\frac{1}{2}u^2\sigma^2}$.*

Proof. It is enough to show that $\mathbb{E}(e^{uZ}) = e^{\frac{u^2}{2}}$ for $Z \sim \mathcal{N}(0,1)$ as $\mathbb{E}(e^{uX}) = \mathbb{E}(e^{u(\mu+\sigma Z)})$. Now as $ux - \frac{x^2}{2} = \frac{-(x-u)^2}{2} + \frac{u^2}{2}$ we have

$$\int_{-\infty}^{\infty} e^{ux}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,\mathrm{d}x = e^{\frac{u^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-u)^2}{2}}\,\mathrm{d}x = e^{\frac{u^2}{2}} \tag{4.B.1}$$

as we recognize the density of the normal distribution with mean u and variance 1. □

Lemma 4.B.2. *Let Y_n be i.i.d. random variables with $\mathbb{E}(e^{cY_1}) = \theta(c)$ for $c \in \mathbb{R}$ and $(N_t)_{t\geq 0}$ be a Poisson process with intensity $\lambda > 0$ with jump times $(T_n)_{n\geq 1}$. Then the dampended jump process, $(L_t)_{t\geq 0}$ given by*

$$L_t = e^{-\kappa_P(t-s)}L_s + \sum_{N_s+1}^{N_t} e^{-\kappa_P(t-T_n)}Y_n \text{ for } t \geq s, \tag{4.B.2}$$

has conditional moment generating function

$$\mathbb{E}(e^{uL_t}|L_s) = e^{A(s,t)+B(s,t)L_s}, \text{ for } u \in \mathbb{R} \tag{4.B.3}$$

where

$$A(s,t) = \lambda \int_{e^{-\kappa_P(t-s)}}^{1} \frac{\theta(uz)-1}{\kappa_P z}\, \mathrm{d}z. \tag{4.B.4}$$

and

$$B(s,t) = ue^{-\kappa_P(t-s)}. \tag{4.B.5}$$

Proof. We have with $J_t = \sum_{n=1}^{N_t} Y_n$ then L_t solves

$$\mathrm{d}L_t = -\kappa_P L_t \,\mathrm{d}t + \,\mathrm{d}J_t, \tag{4.B.6}$$

which we show in Lemma 4.B.3, and hence, as the drift and jump coefficients are affine, it follows by Duffie et al. (2000) that the conditional moment generating function has the form $\mathbb{E}(e^{uL_t}|L_s) = e^{A(s,t)+B(s,t)M_s}$, where $A(s,t)$ and $B(s,t)$ solves

$$\frac{\mathrm{d}}{\mathrm{d}s}B(s,t) = \kappa_P B(s,t) \tag{4.B.7}$$

$$\frac{\mathrm{d}}{\mathrm{d}s}A(s,t) = -\lambda(\theta(B(s,t))-1) \tag{4.B.8}$$

with $B(t,t) = u$ and $A(t,t) = 0$. Thus,

$$B(s,t) = ue^{-\kappa_P(t-s)} \tag{4.B.9}$$

and

$$A(s,t) = \lambda \int_s^t (\theta(ue^{-\kappa_P(t-v)})-1)\,\mathrm{d}v \tag{4.B.10}$$

$$= \lambda \int_{e^{-\kappa_P(t-s)}}^{1} \frac{\theta(uz)-1}{\kappa_P z}\,\mathrm{d}z \tag{4.B.11}$$

where we used a change of variable with $z = e^{-\kappa_P(t-v)}$. □

4.B.2 Ornstein-Uhlenbeck Process

Lemma 4.B.3. *Let $(Z_t)_{t\geq 0}$ be a Brownian motion, $(Y_n)_{n\geq 1}$ be i.i.d. random variables and $(N_t)_{t\geq 0}$ a Poisson process with intensity λ and jump times $(T_n)_{n\geq 1}$. Define the compound Poisson process $J_t = \sum_{n=1}^{N_t} Y_n$. Assume U_t has the dynamics*

$$\mathrm{d}U_t = \kappa(\hat{\alpha}_t - U_t)\ \mathrm{d}t + \sigma\ \mathrm{d}Z_t + \ \mathrm{d}J_t \tag{4.B.12}$$

then

$$\begin{aligned} U_t = U_s e^{-\kappa(t-s)} + \kappa \int_s^t e^{-\kappa(t-v)} \hat{\alpha}_v\ \mathrm{d}v \\ + \sigma \int_s^t e^{-\kappa(t-v)}\ \mathrm{d}Z_v^U + \sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)} Y_n \end{aligned} \tag{4.B.13}$$

and if $\hat{\alpha}_t$ is constant,

$$\begin{aligned} U_t = U_s e^{-\kappa(t-s)} + \hat{\alpha}(1 - e^{-\kappa(t-s)}) \\ + \sigma \int_s^t e^{-\kappa(t-v)}\ \mathrm{d}Z_v^U + \sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)} Y_n. \end{aligned} \tag{4.B.14}$$

Proof. Define $L_t = e^{\kappa t} U_t$, then

$$\mathrm{d}L_t = e^{\kappa t}\ \mathrm{d}U_t + U_t\ \mathrm{d}(e^{\kappa t}) \tag{4.B.15}$$

$$= \kappa(\hat{\alpha}_t - U_t) e^{\kappa t}\ \mathrm{d}t + \sigma e^{\kappa t}\ \mathrm{d}Z_t^U + e^{\kappa t}\ \mathrm{d}J_t + \kappa U_t e^{\kappa t}\ \mathrm{d}t \tag{4.B.16}$$

$$= \kappa \hat{\alpha}_t e^{\kappa_U t}\ \mathrm{d}t + \sigma e^{\kappa t}\ \mathrm{d}Z_t^U + e^{\kappa t}\ \mathrm{d}J_t \tag{4.B.17}$$

Thus, for $t \geq s$

$$L_t = L_s + \kappa \int_s^t \hat{\alpha}_v e^{\kappa v}\ \mathrm{d}v + \sigma \int_s^t e^{\kappa v}\ \mathrm{d}Z_v + \sum_{n=N_s+1}^{N_t} e^{\kappa T_n} Y_n. \tag{4.B.18}$$

Now as $U_t = e^{-\kappa t} L_t$ it follows that

$$\begin{aligned} U_t = U_s e^{-\kappa(t-s)} + \kappa \int_s^t \hat{\alpha}_v e^{-\kappa(t-v)}\ \mathrm{d}v \\ + \sigma \int_s^t e^{-\kappa(t-v)}\ \mathrm{d}Z_v + \sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)} Y_n. \end{aligned} \tag{4.B.19}$$

□

Lemma 4.B.4. *An Ornstein-Uhlenbeck process $(X_t)_{t\geq 0}$ is ρ-mixing, that is $\rho_t \to 0$ for $t \to \infty$ with*

$$\rho_t = \sup_{\substack{s \geq 0 \\ X \in L^2(\mathbb{A}_s^1) \\ Y \in L_2(\mathbb{A}_{s+t}^\infty)}} [cor(X, Y)] \tag{4.B.20}$$

where $L_2(\mathbb{A})$ is the family of all square integrable $\mathbb{A}$-measurable random variables and $\mathbb{A}_t^1 = \mathbb{F}((X_s)_{s\leq t})$ and $\mathbb{A}_t^\infty = \mathbb{F}((X_s)_{s>t})$ are the σ-algebras generated by $(X_s)_{s\leq t}$ and $(X_s)_{s>t}$.

Proof. Let $b(x) = \kappa(\alpha - x)$ be the drift and $a(x) = \sigma$ be the diffusion of the Ornstein-Uhlenbeck process with $\alpha \in \mathbb{R}$, $\kappa > 0$ and $\sigma > 0$ with initial value $x_0 \in \mathbb{R}$. Define the scale and speed densities

$$s(x) = \exp\left(-2\int_{x_0}^{x} \frac{b(u)}{a^2(u)}\,\mathrm{d}u\right) = \exp\left(\frac{(x-\alpha)^2}{2\frac{\sigma^2}{2\kappa}} + \frac{-(x0-\alpha)^2}{2\frac{\sigma^2}{2\kappa}}\right) \tag{4.B.21}$$

$$m(x) = \frac{1}{a^2(x)s(x)} = \exp\left(\frac{-(x-\alpha)^2}{2\frac{\sigma^2}{2\kappa}} + \frac{(x0-\alpha)^2}{2\frac{\sigma^2}{2\kappa}}\right)\Big/\sigma^2 \tag{4.B.22}$$

and

$$\gamma(x) = a'(x) - 2b(x)/a(x) = -2\kappa\frac{\alpha - x}{\sigma}. \tag{4.B.23}$$

Following section 2.6 of Genon-Catalot et al. (2000) it is sufficient to verify the following:

(i) $b(x) \in C^1(\mathbb{R})$, $a(x)^2 \in C^2(\mathbb{R})$ and $a(x) > 0$ for $x \in \mathbb{R}$.

(ii) $\int_{-\infty} s(x)\,\mathrm{d}x = \infty$, $\int^{\infty} s(x)\,\mathrm{d}x = \infty$, $\int_{\mathbb{R}} m(x)\,\mathrm{d}x = M < \infty$.

(iii) $a(x)m(x) \to 0$ for $x \to \infty$ and $x \to -\infty$.

(iv) $1/\gamma(x)$ converges for $x \to \infty$ and $x \to -\infty$.

(i), (iii) and (iv) follows trivially and (ii) follows as $s(x)$ does not converge to 0 for $x \to \infty$ or $x \to \infty$ and as $m(x)$ is a scaled normal density with mean α and variance $\frac{\sigma^2}{2\kappa}$. □

4.B.3 Determination of Expected Instantaneous Value of Wind Production

Lemma 4.B.5. *Let $h(u) = (au + b)1_{(u\in(u_0,u_1))}$ and assume*

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \hat{\rho}\sigma_1\sigma_2 \\ \hat{\rho}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right). \tag{4.B.24}$$

Then

$$\begin{aligned}\mathbb{E}\left[e^{Z_1}h(Z_2)\right] = e^{\mu_1+\frac{\sigma_1^2}{2}}\Bigg[&(a\xi + b)\left(\Phi\left(\frac{u_1-\xi}{\sigma_2}\right) - \Phi\left(\frac{u_0-\xi}{\sigma_2}\right)\right) \\ &+ \frac{a\sigma_2}{\sqrt{2\pi}}\left(e^{-\frac{(u_0-\xi)^2}{2\sigma_2^2}} - e^{-\frac{(u_1-\xi)^2}{2\sigma_2^2}}\right)\Bigg]\end{aligned} \tag{4.B.25}$$

with $\xi = \mu_2 - \hat{\rho}\sigma_1\sigma_2$.

Proof. We use that for $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \mathcal{N}(0,1)$ with X_1 independent of X_2 then

$$(Z_1, Z_2) \stackrel{d}{=} (\mu_1 + \sigma_1(\hat{\rho}X_1 + \sqrt{1-\hat{\rho}^2}X_2, X_1). \tag{4.B.26}$$

Hence

$$\begin{aligned}
\mathbb{E}\left[e^{Z_1}h(Z_2)\right] &= \mathbb{E}\left[e^{\mu_1+\sigma_1(\hat{\rho}X_1+\sqrt{1-\hat{\rho}^2}X_2)}h(\sigma_2 X_1+\mu_2)\right] && \text{(4.B.27)}\\
&= e^{\mu_1}\mathbb{E}\left[e^{\sigma_1\sqrt{1-\hat{\rho}^2}X_2}\right]\mathbb{E}\left[e^{\sigma_1\hat{\rho}X_1}h(\sigma_2 X_1+\mu_2)\right] && \text{(4.B.28)}\\
&= e^{\mu_1}\mathbb{E}\left[e^{\sigma_1\sqrt{1-\hat{\rho}^2}X_2}\right]e^{\frac{\sigma_1^2\hat{\rho}^2}{2}}\mathbb{E}\left[h(\sigma_2(X_1+\hat{\rho}\sigma_1)+\mu_2)\right] && \text{(4.B.29)}\\
&= e^{\mu_1+\frac{\sigma_1^2}{2}}\mathbb{E}\left[h(\sigma_2 X_1+\xi)\right], && \text{(4.B.30)}
\end{aligned}$$

with $\xi = \mu_2 + \hat{\rho}\sigma_1\sigma_2$. Here we use that for u constant and $\mathbb{E}(|f(X_1+u)|) < \infty$, then $\mathbb{E}\left[e^{uX_1}f(X_1)\right] = e^{\frac{u^2}{2}}\mathbb{E}\left[f(X_1+u)\right]$, which follows by (4.B.1). Now with $\Phi(x)$ the distribution function for the standard normal distribution,

$$\begin{aligned}
\mathbb{E}\left[h(\sigma_2 X_1+\xi)\right] &= \mathbb{E}\left[(a\sigma_2 X_1 + a\xi + b)1_{(\sigma_2 X_1+\xi\in(u_0,u_1))}\right] && \text{(4.B.31)}\\
&= (a\xi+b)\left(\Phi\left(\frac{u_1-\xi}{\sigma_2}\right)-\Phi\left(\frac{u_0-\xi}{\sigma_2}\right)\right)\\
&\quad + a\sigma_2\frac{1}{\sqrt{2\pi}}\int_{\frac{u_0-\xi}{\sigma_2}}^{\frac{u_1-\xi}{\sigma_2}} xe^{-\frac{x^2}{2}}\,\mathrm{d}x. && \text{(4.B.32)}
\end{aligned}$$

Using the change of variable $z = \frac{x^2}{2}$, we find that

$$\begin{aligned}
\int_{\frac{u_0-\xi}{\sigma_2}}^{\frac{u_1-\xi}{\sigma_2}} xe^{-\frac{x^2}{2}}\,\mathrm{d}x &= \int_{\frac{(u_0-\xi)^2}{2\sigma_2^2}}^{\frac{(u_1-\xi)^2}{2\sigma_2^2}} e^{-z}\,\mathrm{d}z && \text{(4.B.33)}\\
&= e^{\frac{(u_0-\xi)^2}{2\sigma_2^2}} - e^{\frac{(u_1-\xi)^2}{2\sigma_2^2}}, && \text{(4.B.34)}
\end{aligned}$$

which proves the claim. □

Lemma 4.B.6. *Let $(Z_t)_{t\geq 0}$ be a Brownian motion and define for $t \geq s$*

$$I_1(s,t) = \int_s^t h_1(v)\,\mathrm{d}Z_v \tag{4.B.35}$$

$$I_2(s,t) = \int_s^t h_2(v)\,\mathrm{d}Z_v, \tag{4.B.36}$$

where we assume $\int_s^t h_1(v)^2\,\mathrm{d}v < \infty$ and $\int_s^t h_2(v)^2\,\mathrm{d}v < \infty$. Then

$$Cov(I_1(s,t), I_2(s,t)) = \int_s^t h_1(v)h_2(v)\,\mathrm{d}v. \tag{4.B.37}$$

Proof. We have that

$$Cov(I_1(s,t), I_2(s,t)) = \mathbb{E}\left(\int_s^t h_1(v)\ \mathrm{d}Z_v \int_s^t h_2(v)\ \mathrm{d}Z_v\right) \tag{4.B.38}$$

as $\mathbb{E}(\int_s^t h_1(v)\ \mathrm{d}Z_v) = \mathbb{E}(\int_s^t h_2(v)\ \mathrm{d}Z_v) = 0$. Now using Itô's isometry,

$$\mathbb{E}\left[\left(\int_s^t h(v)\ \mathrm{d}Z_v\right)^2\right] = \int_s^t h(v)^2\ \mathrm{d}v, \tag{4.B.39}$$

it follows that

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_s^t h_1(v)\ \mathrm{d}Z_v + \int_s^t h_2(v)\ \mathrm{d}Z_v\right)^2\right] \\
=&\mathbb{E}\left[\left(\int_s^t h_1(v) + h_2(v)\ \mathrm{d}Z_v\right)^2\right] && (4.B.40)\\
=&\mathbb{E}\left[\int_s^t (h_1(v) + h_2(v))^2\ \mathrm{d}v\right] && (4.B.41)\\
=&\int_s^t h_1(v)^2 + h_2(v)^2\ \mathrm{d}v + 2\int_s^t h_1(v)h_2(v)\ \mathrm{d}v && (4.B.42)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_s^t h_1(v)\ \mathrm{d}Z_v + \int_s^t h_2(v)\ \mathrm{d}Z_v\right)^2\right] \\
=&E\Bigg[\left(\int_s^t h_1(v)\ \mathrm{d}v\right)^2 + \left(\int_s^t h_2(v)\ \mathrm{d}Z_v\right)^2 && (4.B.43)\\
&+2\int_s^t h_1(v)\ \mathrm{d}Z_v \int_s^t h_2(v)\ \mathrm{d}Z_v\Bigg] \\
=&\int_s^t h_1(v)^2 + h_2(v)^2\ \mathrm{d}v + 2E\Bigg[\int_s^t h_1(v)\ \mathrm{d}Z_v \int_s^t h_2(v)\ \mathrm{d}Z_v\Bigg]. && (4.B.44)
\end{aligned}$$

Hence, combining (4.B.37), (4.B.42) and (4.B.44) the result follows. □

Corollary 1. *Let $(Z_t)_{t\geq 0}$ be a Brownian motion and define for $t \geq s$*

$$U_t = U_s e^{-\kappa_U(t-s)} + \int_s^t e^{-\kappa_U(v-s)}\ \mathrm{d}Z_v \tag{4.B.45}$$

$$X_t = X_s + Z_t - Z_s = X_s + \int_s^t 1\ \mathrm{d}Z_v. \tag{4.B.46}$$

Then

$$Cov(X_t, U_t|(X_s, U_s)) = \frac{1 - e^{-\kappa_U(t-s)}}{\kappa_U} \tag{4.B.47}$$

Corollary 2. *Let $(Z_t)_{t\geq 0}$ be a Brownian motion and define for $t \geq s$*

$$U_t = U_s e^{-\kappa_U (t-s)} + \int_s^t e^{-\kappa_U (v-s)} \, \mathrm{d}Z_v \tag{4.B.48}$$

$$X_t = X_s e^{-\kappa_P (t-s)} + \int_s^t e^{-\kappa_P (v-s)} \, \mathrm{d}Z_v, \tag{4.B.49}$$

then

$$Cov(X_t, U_t | (X_s, U_s)) = \frac{1 - e^{-(\kappa_U + \kappa_P)(t-s)}}{\kappa_U + \kappa_P}. \tag{4.B.50}$$

Lemma 4.B.7. *Let X_t be as in Appendix 4.2.1 where X_t follows a Brownian motion with drift and U_t be as in Appendix 4.3. Furthermore let f be the transformation from weather factor to wind speed, $f(u) = F^{-1}(\Phi(u))$ and the power curve, h, be defined as in (4.4.8). Then for $t > s$*

$$\begin{aligned}
&\mathbb{E}(e^{X_t} h(f(U_t)) | \mathbb{F}_s) \\
=&e^{X_s} e^{\mu(t-s)} \Bigg[h_{full} \left(\Phi\left(\frac{u_2 - \xi}{\tau} \right) - \Phi\left(\frac{u_1 - \xi}{\tau} \right) \right) \\
+&\frac{\tau h_{full}}{\sqrt{2\pi}(u_1 - u_0)} \left(e^{-\frac{(u_0 - \xi)^2}{2\tau^2}} - e^{-\frac{(u_1 - \xi)^2}{2\tau^2}} \right) \\
+&\frac{(\xi - u_0) h_{full}}{u_1 - u_0} \left(\Phi\left(\frac{u_1 - \xi}{\tau} \right) - \Phi\left(\frac{u_0 - \xi}{\tau} \right) \right) \Bigg].
\end{aligned} \tag{4.B.51}$$

with $\xi = U_s e^{-\kappa_U (t-s)} + \rho\sigma\sqrt{2\kappa_U} \frac{1 - e^{-\kappa_U (t-s)}}{\kappa_U}$ and $\tau = 1 - e^{-2\kappa_U (t-s)}$.

Proof. As $(J_t)_{t\geq 0}$ is independent of $(U_t)_{t\geq 0}$ and

$$\mathbb{E}\left(e^{-\lambda k_1 (t-s) + \sum_{n=N_s+1}^{N_t} Y_n} \right) = 1 \tag{4.B.52}$$

it follows that

$$\mathbb{E}(e^{X_s} h(f(U_s)) | \mathbb{F}_t) = \mathbb{E}\left[e^{Z_1} h\left(f(Z_2) \right) | \mathbb{F}_t \right]. \tag{4.B.53}$$

with

$$Z_1 = X_s + (\mu - \frac{1}{2}\sigma^2)(t-s) + \sigma(Z_t^P - Z_s^P) \tag{4.B.54}$$

$$Z_2 = U_s e^{-\kappa_U (t-s)} + \sqrt{2\kappa_U} \int_s^t e^{-\kappa_U (v-s)} \, \mathrm{d}Z_v^U. \tag{4.B.55}$$

Now as

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \hat{\rho}\sigma_1\sigma_2 \\ \hat{\rho}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right). \tag{4.B.56}$$

with

$$\mu_1 = X_t + (\mu - \frac{1}{2}\sigma^2)(t-s) \tag{4.B.57}$$
$$\mu_2 = U_t e^{-\kappa_U(t-s)} \tag{4.B.58}$$
$$\sigma_1^2 = \sigma^2(t-s) \tag{4.B.59}$$
$$\sigma_2^2 = 1 - e^{-2\kappa_U(t-s)}. \tag{4.B.60}$$

it follows by Corollary 1 that

$$\hat{\rho}\sigma_1\sigma_2 = Cov(Z_1, Z_2) \tag{4.B.61}$$
$$= E\left[\sigma(Z_t^P - Z_s^P)\sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(v-s)}\,\mathrm{d}Z_v^U\right] \tag{4.B.62}$$
$$= E\left[\sigma(Z_t - Z_s)\rho\sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(v-s)}\,\mathrm{d}Z_v\right] \tag{4.B.63}$$
$$= \rho\sigma\sqrt{2\kappa_U}\frac{1-e^{-\kappa_U(t-s)}}{\kappa_U}, \tag{4.B.64}$$

where we used the linearity of the stochastic integral and that $(Z_P, Z_U) \overset{d}{=} (Z, \rho Z + \sqrt{1-\rho^2}Y)$ with Z and Y independent Brownian Motions. Now the result follows from Lemma 4.B.5. □

Lemma 4.B.8. *Let X_t be as in Appendix 4.2.2 where X_t follows an Ornstein-Uhlenbeck process and U_t be as in Appendix 4.3. Furthermore let f be the transformation from weather factor to wind speed, $f(u) = F^{-1}(\Phi(u))$ and the power curve, h, be defined as in (4.4.8). Then for $t > s$*

$$\begin{aligned}
&\mathbb{E}(e^{X_t}h(f(U_t))|\mathbb{F}_s)\\
=&e^{X_s e^{-\kappa_P(t-s)}+\alpha_t}\Bigg[h_{full}\left(\Phi\left(\frac{u_2-\xi}{\tau}\right)-\Phi\left(\frac{u_1-\xi}{\tau}\right)\right)\\
+&\frac{\tau h_{full}}{\sqrt{2\pi}(u_1-u_0)}\left(e^{-\frac{(u_0-\xi)^2}{2\tau^2}}-e^{-\frac{(u_1-\xi)^2}{2\tau^2}}\right)\\
+&\frac{(\xi-u_0)h_{full}}{u_1-u_0}\left(\Phi\left(\frac{u_1-\xi}{\tau}\right)-\Phi\left(\frac{u_0-\xi}{\tau}\right)\right)\Bigg].
\end{aligned} \tag{4.B.65}$$

with $\alpha_t = (\alpha+\epsilon_\sigma)(1-e^{-\kappa_P(t-s)})-\epsilon_k$, $\epsilon_\sigma = \frac{\sigma^2}{4\kappa_P}e^{-\kappa_P(t-s)}$, $\epsilon_k = \lambda\int_0^{e^{-\kappa_P(t-s)}}\frac{\theta(z)-1}{\kappa_P c}\,\mathrm{d}z$, $\xi = U_s e^{-\kappa_U(t-s)} + \rho\sigma\sqrt{2\kappa_U}\frac{1-e^{-\kappa_U(t-s)}}{\kappa_U}$.

Proof. As $(J_t)_{t\geq 0}$ is independent of $(U_t)_{t\geq 0}$ and

$$\mathbb{E}\left[\exp\left(-\lambda k_2(t-s)+\sum_{n=N_s+1}^{N_t} e^{-\kappa_P(t-T_n)}Y_n\right)\right] \tag{4.B.66}$$

$$=\exp\left(-\lambda k_2+A(s,t)\right)$$

$$=\exp\left(-\lambda\int_0^{e^{-\kappa_P(t-s)}}\frac{\theta(z)-1}{\kappa_P c}\,\mathrm{d}z\right) \tag{4.B.67}$$

it follows that with $\epsilon_k=\lambda\int_0^{e^{-\kappa_P(t-s)}}\frac{\theta(z)-1}{\kappa_P c}\,\mathrm{d}z$ then

$$\mathbb{E}(e^{X_s}h(f(U_s))|\mathbb{F}_t)=e^{-\epsilon_k}\mathbb{E}\left[e^{Z_1}h\left(f(Z_2)\right)|\mathbb{F}_t\right]. \tag{4.B.68}$$

with

$$Z_1=X_se^{-\kappa_P(t-s)}+(\alpha-\frac{1}{4\kappa_P}\sigma^2)(t-s)+\sigma\int_s^t e^{-\kappa_P(v-s)}\,\mathrm{d}Z_v^P \tag{4.B.69}$$

$$Z_2=U_se^{-\kappa_U(t-s)}+\sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(v-s)}\,\mathrm{d}Z_v^U. \tag{4.B.70}$$

We have that

$$\begin{pmatrix} Z_1\\ Z_2\end{pmatrix}\sim\mathcal{N}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \hat{\rho}\sigma_1\sigma_2\\ \hat{\rho}\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right). \tag{4.B.71}$$

with

$$\mu_1=X_se^{-\kappa_P(t-s)}+(\alpha-\frac{\sigma^2}{4\kappa_P})(1-e^{-\kappa_P(t-s)}) \tag{4.B.72}$$

$$\mu_2=U_te^{-\kappa_U(t-s)} \tag{4.B.73}$$

$$\sigma_1^2=\sigma^2\frac{1-e^{-2\kappa_P(t-s)}}{2\kappa_P} \tag{4.B.74}$$

$$\sigma_2^2=1-e^{-2\kappa_U(t-s)}. \tag{4.B.75}$$

By Corollary 2 it follows that

$$\hat{\rho}\sigma_1\sigma_2=Cov(Z_1,Z_2) \tag{4.B.76}$$

$$=E\left[\sigma\int_s^t e^{-\kappa_P(v-s)}\,\mathrm{d}Z_v^P\sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(v-s)}\,\mathrm{d}Z_v^U\right] \tag{4.B.77}$$

$$=E\left[\sigma\int_s^t e^{-\kappa_P(v-s)}\,\mathrm{d}Z_v\rho\sqrt{2\kappa_U}\int_s^t e^{-\kappa_U(v-s)}\,\mathrm{d}Z_v\right] \tag{4.B.78}$$

$$=\rho\sigma\sqrt{2\kappa_U}\frac{1-e^{-(\kappa_U+\kappa_P)(t-s)}}{\kappa_U+\kappa_P} \tag{4.B.79}$$

where we used the linearity of the stochastic integral and that $(Z^P, Z^U) \stackrel{d}{=} (Z, \rho Z + \sqrt{1-\rho^2}Y)$ with Z and Y independent Brownian Motions. Now as

$$\begin{aligned}
&\exp\left(\mu_1 + \frac{\sigma_1^2}{2}\right) \\
&= \exp\left(X_s e^{-\kappa_P(t-s)} + (\alpha - \frac{\sigma^2}{4\kappa_P})(1 - e^{-\kappa_P(t-s)}) + \sigma^2 \frac{1 - e^{-2\kappa_P(t-s)}}{4\kappa_P}\right) && \text{(4.B.80)} \\
&= \exp\left(X_s e^{-\kappa_P(t-s)} + \alpha(1 - e^{-\kappa_P(t-s)}) + \sigma^2 \frac{e^{-\kappa_P(t-s)} - e^{-2\kappa_P(t-s)}}{4\kappa_P}\right) && \text{(4.B.81)} \\
&= \exp\left(X_s e^{-\kappa_P(t-s)} + (\alpha + \epsilon_\sigma)(1 - e^{-\kappa_P(t-s)})\right) && \text{(4.B.82)}
\end{aligned}$$

with $\epsilon_\sigma = \frac{\sigma^2}{4\kappa_P} e^{-\kappa_P(t-s)}$, the result follows from Lemma 4.B.5. □

4.B.4 Solutions to Homogeneous HJB Equation with Ornstein-Uhlenbeck Process

Lemma 4.B.9. *For $\kappa_P > 0$, $r > 0$, $a = \frac{r}{2\kappa_P}$, $b = \frac{1}{2}$, $x^* = \alpha - \frac{\sigma^2}{4\kappa_P}$ and*

$$f(x) = \kappa_P \frac{(x - x^*)^2}{\sigma^2} \tag{4.B.83}$$

we have that

$$m_1(x) = \begin{cases} U(a, b, f(x)) & x < x^* \\ -U(a, b, f(x)) + 2\frac{\Gamma(1-b)}{\Gamma(a-b+1)} M(a+b, f(x)) & x \geq x^* \end{cases} \tag{4.B.84}$$

$$m_2(x) = \begin{cases} -U(a, b, f(x)) + 2\frac{\Gamma(1-b)}{\Gamma(a-b+1)} M(a+b, f(x)) & x < x^* \\ U(a, b, f(x)) & x \geq x^* \end{cases} \tag{4.B.85}$$

solves the second order differential equation

$$\kappa_P(x^* - x)\frac{\partial}{\partial x} m(x) + 1/2\sigma^2 \frac{\partial^2}{\partial x^2} m(x) - rm(x) = 0 \tag{4.B.86}$$

with $m_1(x) \to 0$ for $x \to -\infty$ and $m_2(x) \to 0$ for $x \to \infty$ and $m_1(x)$ and $m_2(x)$ are bounded on $(-\infty, x_0)$ and (x_0, ∞) respectively for any $x_0 \in \mathbb{R}$.

Proof. Abramowitz and Stegun (1972) shows that $M(a, b, z)$ and $U(a, b, z)$ solves

$$z\frac{\partial}{\partial z^2} M(a, b, z) + (b - z)\frac{\partial}{\partial z} M(a, b, z) - aM(a, b, z) = 0 \tag{4.B.87}$$

and are independent for $a \neq 0, -1, -2, \ldots$. From the integral representations of $M(a, b, z)$ and $U(a, b, z)$ we note that $M(a, b, f(x)) \to \infty$ for $x \to \pm\infty$ and $U(a, b, f(x)) \to 0$ for $x \pm \infty$. We have that with

$$m(x) = M(a, b, f(x)) \tag{4.B.88}$$

and

$$f(x) = \kappa_P \frac{(x - x^*)^2}{\sigma^2} \tag{4.B.89}$$

then

$$\frac{\partial}{\partial x} m(x) = M'(a, b, f(x)) f'(x) \tag{4.B.90}$$

$$\frac{\partial^2}{\partial x^2} m(x) = M''(a, b, f(x)) f'(x)^2 + M'(a, b, f(x)) f''(x) \tag{4.B.91}$$

and inserting in (4.B.86) we obtain

$$\begin{aligned} &(\kappa_P(x^* - x) f'(x) + \frac{1}{2}\sigma^2 f''(x)) M'(a, b, f(x)) \\ &+ \frac{1}{2}\sigma^2 f'(x)^2 M''(a, b, f(x)) - r M(a, b, f(x)) = 0 \end{aligned} \tag{4.B.92}$$

which simplifies to

$$\begin{aligned} &f(x) M''(a, b, f(x)) + (\frac{1}{2} - f(x)) M'(a, b, f(x)) \\ &- \frac{r}{2\kappa_P} M(a, b, f(x)) = 0. \end{aligned} \tag{4.B.93}$$

Thus (4.B.87) is satisfied with $a = \frac{r}{2\kappa_P}$ and $b = \frac{1}{2}$. As we only used that $M(a, b, f(x))$ solved (4.B.87) it also follows that $U(a, b, f(x))$ solves (4.B.86). Now we only need to show that $m_1(x)$ and $m_2(x)$ are twice differentiable at $x = x^*$. Here we use another result from Abramowitz and Stegun (1972), namely that

$$\begin{aligned} U(a, b, z) = &\frac{\Gamma(1 - b)}{\Gamma(a - b + 1)} M(a, b, z) \\ &+ \frac{\Gamma(b - 1)}{\Gamma(a)} z^{1-b} M(a - b + 1, 2 - b, z). \end{aligned} \tag{4.B.94}$$

Thus, by inserting (4.B.94) in (4.B.84) and (4.B.85), it follows that $m_1(x)$ and $m_2(x)$ are continuous at $x = x^*$ as $f(x^*) = 0$. As $b = 1/2$ we have that for $x > x^*$,

$$f(x)^{\frac{1}{2}} M(a - b + 1, 2 - b, f(x)) = \sqrt{\kappa} \frac{x - x^*}{\sigma} M(a - b + 1, 2 - b, f(x)) \tag{4.B.95}$$

and for $x < x^*$

$$f(x)^{\frac{1}{2}} M(a-b+1, 2-b, f(x)) = -\sqrt{\kappa}\frac{x-x^*}{\sigma} M(a-b+1, 2-b, f(x)). \tag{4.B.96}$$

Thus, as $U(a, b, f(x))$ has opposite sign in the $x \geq x^*$ and $x < x^*$ part of (4.B.84) and (4.B.85) it follows that $m_1(x)$ and $m_2(x)$ are twice differentiable at $x = x^*$ if $M(a-b+1, 2-b, z)$ is twice differentiable with respect to z. Now,

$$M(a, b, z) = \mathbb{E}(e^{zX}) \tag{4.B.97}$$

with $X \sim Beta(a, b-a)$ and thus as the beta distribution has moments of all orders it follows that $M(a, b, z)$ is infinitely differentiable at $z = 0$. □

Note that $M(a, b, 0) = 1$ and thus using (4.B.94) it follows that $U(a, b, 0) = \frac{\Gamma(1-b)}{\Gamma(a-b+1)}$. Furthermore,

$$M'(a, b, z)|_{z=0} = \mathbb{E}(X) = \frac{a}{b}, \tag{4.B.98}$$

and higher order derivatives can be determined through the moments of the beta distribution. $U(a, b, z)$ is however not differentiable at $z = 0$.

Appendix 4.C Sensitivity Analysis

4.C.1 Impact of 10% Increase in Average Price

We use $\lim_{t\to\infty} \mathbb{E}(P_t)$ as the average price for both models.

Model for X_t	+10% avg. price (mio. DKK)
Brownian Motion	(+10.30%) 15.098
OU-process	(+10.29%) 15.097
OU-process with jumps	(+10.27%) 15.095

Figure 4.16: Impact of average price on wind turbine value

Model for X_t	+10% avg. price (mio. DKK)
Brownian Motion	(+32.37%) 1 954.392
OU-process	(+39.99%) 1 675.675
OU-process with jumps	(+37.34%) 1 684.287

Figure 4.17: Impact of average price on value of gas fired power plant

Model for X_t	+10% avg. price (mio. DKK)
Brownian Motion	(+10.00%) 468.181
OU-process	(+9.10%) 508.459
OU-process with jumps	(+9.15%) 516.192

Figure 4.18: Impact of average price on value of hydroelectric power plant

4.C.2 Impact of 10% Increase in Standard Deviation

The standard deviation is measured as

$$S = \sqrt{\frac{1}{\tilde{T}} \int_0^{\tilde{T}} \mathbb{E}(P_t^2) - \mathbb{E}(P_t)^2 \, \mathrm{d}t}, \tag{4.C.1}$$

for the models based on the Brownian motion, where $\tilde{T} = 7$ is the number of years in the data used for calibration. Furthermore,

$$S = \lim_{t \to \infty} \sqrt{\mathbb{E}(P_t^2) - \mathbb{E}(P_t)^2}. \tag{4.C.2}$$

is used for the models based on the Ornstein-Uhlenbeck process.

Model for X_t	+10% std. dev. (mio. DKK)
Brownian Motion	(-0.38%) 13.636
OU-process	(-0.38%) 13.636
OU-process with jumps	(-5.61%) 12.920

Figure 4.19: Impact of standard deviation on wind turbine value

Model for X_t	+10% std. dev. (mio. DKK)
Brownian Motion	(+11.81%) 1 650.794
OU-process	(+12.92%) 1 351.612
OU-process with jumps (σ)	(+9.83%) 1 346.744
OU-process with jumps (λ)	(+13.08%) 1 386.664

Figure 4.20: Impact of standard deviation on value of gas fired power plant

Model for X_t	+10% std. dev. (mio. DKK)
Brownian Motion	(0.00%) 425.612
OU-process	(+1.21%) 471.709
OU-process with jumps (σ)	(+1.12%) 466.483
OU-process with jumps (λ)	(+1.70%) 468.150

Table 4.12: Impact of standard deviation on value of hydroelectric power plant

4.C.3 Wind Turbine

Model for X_t	-1 percent point correlation	
OU-process	(-0.32%)	13.644
OU-process with jumps	(-0.31%)	13.646

Table 4.13: Impact of correlation on value of wind turbine

4.C.4 Gas Fired Power Plant

Model for X_t	P^*_{on}	P^*_{off}
Brownian Motion	325.40	319.97
OU-process	364.15	274.90
OU-process with jumps	377.41	302.72

Table 4.14: Startup and shutdown triggers.

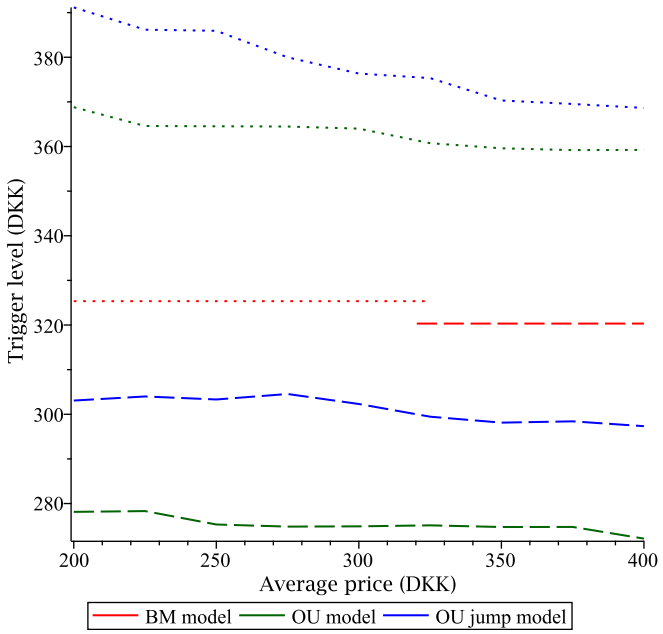


Figure 4.21: Impact of average price on trigger for gas fired power plant.

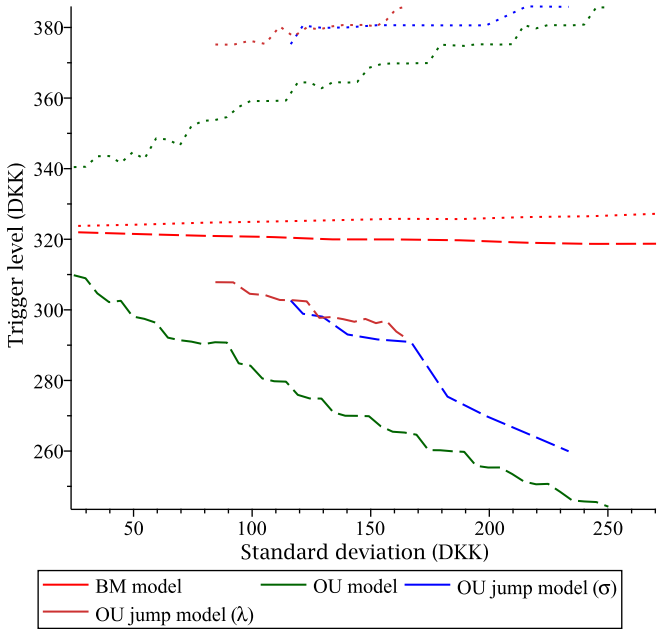


Figure 4.22: Impact of standard deviation on trigger for gas fired power plant.

4.C.5 Hydroelectric Power Plant

We obtain the following strategy parameters, d_1, d_2 and d_3 for each of the models and note that an increase in price by 1 DKK increases the discharge rate by d_2/N_{secs} and an increase in head level by 1 meter increases the discharge rate by d_3 $(\cdot a/N_{secs})$.

Model for X_t	d_1 $/N_{secs}$	d_2 $/N_{secs}$	d_3 $\cdot a/N_{secs}$
	m^3/s	$\frac{m^3/s}{DKK/MWh}$	$\frac{m^3/s}{m}$
BM	-31.4527	0.000761	0.99848
OU	-50.0186	0.066887	0.98447
OU with jumps	-49.6884	0.069913	0.97183

Table 4.15: Parameters of the control strategy for the hydroelectric power plant.

4.C.5.1 Distribution of water level and discharge rate

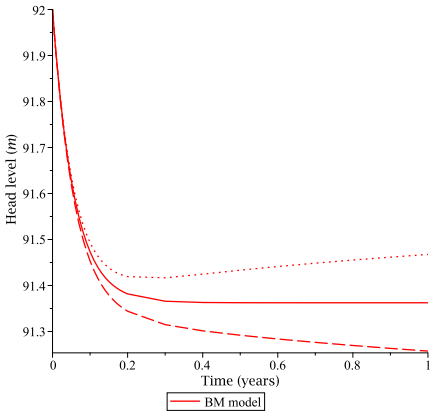


Figure 4.27: Distribution of flow rates of hydroelectric power plant with optimal strategy

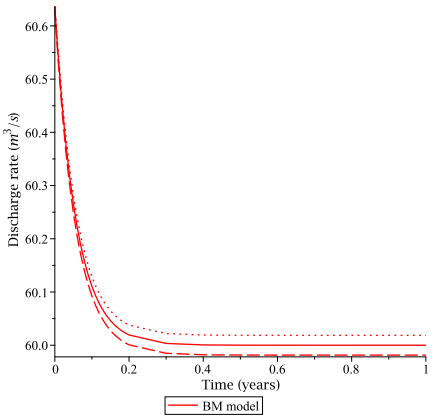


Figure 4.28: Distribution of water level of hydro plant with optimal strategy

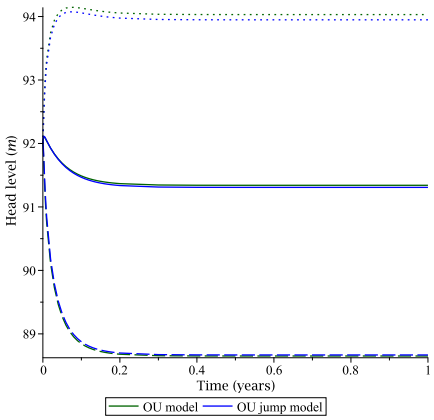


Figure 4.29: Distribution of water level of hydro plant with optimal strategy

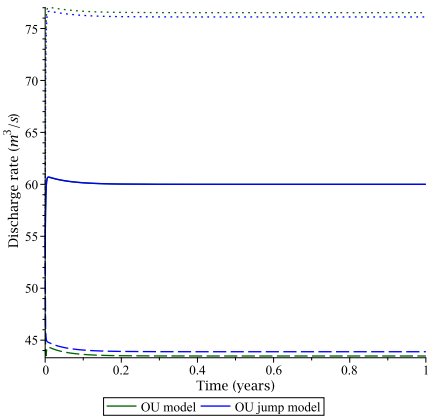


Figure 4.30:Distribution of flow rates of hydroelectric power plant with optimal strategy

5

An EM Algorithm with Two Jump Components

Abstract

An EM-algorithm with two jump components is developed based on the EM algorithm from Duncan et al. (2009) that includes a single jump component. The inclusion of an additional jump component creates a more accurate model and provides a better estimate for the structure of the jumps, while remaining analytically tractable, as the two jump components can be aggregated to a single jump process with a mixture distribution as the jump density. The EM-algorithm is applied to a jump diffusion model for electricity prices with hourly observations for a 7 year period.

5.1 Introduction

Classical diffusion models are typically calibrated to historical data by utilizing the fact that a transformation of the increments are i.i.d. such that statistical properties can be estimated using some form of the law of large numbers. This approach typically assumes that the transformed increments follows a normal distribution, which is often the case if the model is based on a diffusion process. However, the normality assumption may often fail, due to heavy tails of the distribution, which indicates that a larger part of the variance is a result of infrequent extreme deviations. This type of data can be described better using jump diffusion processes, which often result in transformed increments that consist of a diffusion term as well as a jump term. Examples of jump

diffusion models include the classical paper by Merton (1976) and affine jump diffusions covered in Duffie et al. (2000). Furthermore, various electricity price models, which this algorithm was primarily developed for, are based on jump diffusions, see Deng (1999), Johnson and Barz (1999) and Bhar et al. (2013).

However, estimation of parameters in a jump model is often difficult, as direct maximum likelihood estimation has to be done numerically on an intractable log-likelihood function, which can lead to misspecification of the model. The EM-algorithm, first covered in the general case by Dempster et al. (1977), introduces the notion of the complete data, which is the observed data and additional unobserved data, such that the complete data has a more tractable log-likelihood function. Then, assuming some initial distribution of the complete data, the expectation of the log-likelihood function for the complete data conditional on the observed data is maximized and the distribution of the complete data is updated. This leads to an iterative approach, where each step can be shown to increase the value of the original log-likelihood function. Thus, the intractable log-likelihood function of the observed data is maximized indirectly using the EM-algorithm. However, as is the case with numerical optimization of the likelihood function, the iterative approach from the EM-algorithm can converge to stationary points that are not a global maximum, see Wu (1983) and Vaida (2005) for a general convergence analysis. However, by including several different initial starting parameters and choosing the best result among them, the EM-algorithm can typically provide good parameter estimates, see Dempster et al. (1977), Pickard et al. (1986) and Duncan et al. (2009).

The step of determining the expectation is called the E-step and the step of maximizing the expectation over all parameters is called the M-step, which is why the algorithm is called the EM-algorithm. The rest of the paper is organized as follows. Section 5.2 gives a short proof of the fact that an EM step increases the value of the log-likelihood function for the observed data. Section 5.3 introduces an explicit EM algorithm with two jump components and determines the optimal parameters in terms of conditional expectations as well as an approach for computing the conditional expectations. Section 5.4 covers an application of the algorithm on a jump diffusion model for hourly price data for 7 years from the electricity market. In Section 5.5, we provide a brief conclusion.

5.2 The EM-algorithm

Following Dempster et al. (1977), we provide a short proof that an iteration with the EM-algorithm increases the log-likelihood function. Initially we let $(\mathcal{X}, \mathbb{E}, m_{\mathcal{X}})$ and $(\mathcal{Y}, \mathbb{F}, m_{\mathcal{Y}})$ be measure spaces, where we refer to $\mathcal{X}$ as the complete sample space and $\mathcal{Y}$ as the observed sample space. Let $H : \mathcal{X} \to \mathcal{Y}$ be the surjective map such that the preimage of $y \in \mathcal{Y}$ of H, $H^{-1}(\{y\})$,

consist of all samples from $\mathcal{X}$ that result in the observation y. We denote the probability density of the complete data with parameter θ with respect to $m_{\mathcal{X}}$ by

$$f(x|\theta), \tag{5.2.1}$$

such that the corresponding probability density for the observed data with parameter θ with respect to $m_{\mathcal{Y}}$ satisfies

$$g(y|\theta) = \int_{H^{-1}(\{y\})} f(x|\theta)\ \mathrm{d}m_X(x). \tag{5.2.2}$$

Thus, the probability density for the complete data with parameter θ conditional on the observed data with respect to $m_{\mathcal{X}}$ is

$$k(x|y,\theta) = \frac{f(x|\theta)}{g(y|\theta)}, \tag{5.2.3}$$

for $x \in H^{-1}(\{y\})$. Hence,

$$\log(g(y|\theta)) = \log(f(x|\theta)) - \log(k(x|y,\theta)). \tag{5.2.4}$$

The EM-algorithm determines $\hat{\theta}$ that maximizes

$$E_{\theta_0}\left(\log(f(X|\theta)|\, Y = y\right) \tag{5.2.5}$$

as a function of θ, which in terms of densities is,

$$\int_{H^{-1}(\{y\})} \log(f(x|\theta))k(x|y,\theta_0)\ \mathrm{d}m_X(x). \tag{5.2.6}$$

Hence, using Jensens inequality, that $\hat{\theta}$ maximizes (5.2.6) and that (5.2.4) holds, the log-likelihood of the observed data satisfies

$$\begin{aligned}
\log\left(g(y|\hat{\theta})\right) &= \log\left(\int_{H^{-1}(\{y\})} f(x|\hat{\theta})\ \mathrm{d}m_X(x)\right) && (5.2.7)\\
&= \log\left(\int_{H^{-1}(\{y\})} \frac{f(x|\hat{\theta})}{k(x|y,\theta_0)} k(x|y,\theta_0)\ \mathrm{d}m_X(x)\right) && (5.2.8)\\
&\geq \int_{H^{-1}(\{y\})} \log\left(\frac{f(x|\hat{\theta})}{k(x|y,\theta_0)}\right) k(x|y,\theta_0)\ \mathrm{d}m_X(x) && (5.2.9)\\
&\geq \int_{H^{-1}(\{y\})} \log\left(\frac{f(x|\theta_0)}{k(x|y,\theta_0)}\right) k(x|y,\theta_0)\ \mathrm{d}m_X(x) && (5.2.10)\\
&= \log\left(g(y|\theta_0)\right). && (5.2.11)
\end{aligned}$$

Thus, a step with the EM-algorithm increases the value of the log-likelihood function of the observed data.[1]

[1] Jensens inequality becomes an equality if $f(x|\hat{\theta})/k(x|y,\theta_0)$ does not depend on x and the increase in (5.2.6) determines the increase in the log-likelihood function for the observed data.

5.3 Model

We assume that we have T observations, $y = (y_n)_{n\in\{1,\dots,T\}}$, which are realizations of $Y = (Y_n)_{n\in\{1,\dots,T\}}$ given by

$$Y_n = Z_n + \sum_{k_1=1}^{N_n^{(1)}} J_{nk_1}^{(1)} + \sum_{k_2=1}^{N_n^{(2)}} J_{nk_2}^{(2)}, \qquad \text{for } n = 1,\dots,T, \tag{5.3.1}$$

where Z_n, $J_{nk_1}^{(1)}$, $J_{nk_2}^{(2)}$, $N_n^{(1)}$ and $N_n^{(2)}$ are i.i.d. for $n = 1,\dots,T$, $k_1 = 1,\dots,N_n^{(1)}$ and $k_2 = 1,\dots,N_n^{(2)}$. Furthermore,

$$Z_1 \sim \mathcal{N}(\mu,\sigma), \qquad J_{11}^{(j)} \sim \mathcal{N}(\nu_j,\tau_j^2), \text{ for } j = 1,2, \tag{5.3.2}$$

and $N_1^{(j)}$ is a Poisson processes with intensity λ_j for $j = 1,2$. We define the vector of parameters for the complete data,

$$\theta = (\mu,\sigma,\lambda_1,\nu_1,\tau_1,\lambda_2,\nu_2,\tau_2), \tag{5.3.3}$$

the conditional mean and variance,

$$\mu_Y(k_1,k_2) = \mu + k_1\nu_1 + k_2\nu_2 \tag{5.3.4}$$

$$\sigma_Y^2(k_1,k_2)^2 = \sigma^2 + k_1\tau_1^2 + k_2\tau_2^2 \tag{5.3.5}$$

and find that the density for Y_n is

$$g_n(y|\theta) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} \phi\left(y;\mu_Y(k_1,k_2),\sigma_Y^2(k_1,k_2)\right)\psi(k_1;\lambda_1)\psi(k_2,\lambda_2). \tag{5.3.6}$$

Here $\phi(y;\mu,\sigma^2)$ is the density for the normal distribution with mean μ and variance σ^2 given by

$$\phi(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{5.3.7}$$

and $\psi(k_j;\lambda)$ the density for the Poisson distribution with parameter λ given by

$$\psi(k_j;\lambda) = e^{-\lambda}\frac{\lambda^{k_j}}{k_j!}. \tag{5.3.8}$$

Thus, Y has density

$$g(y|\theta) = \prod_{n=1}^{T} g_n(y|\theta), \tag{5.3.9}$$

and hence, the log-likelihood function is not tractable, as $g_n(y|\theta)$ is an infinite sum. Following Duncan et al. (2009) we extend the observed data, $Y = (Y_n)_{n\in\{1,\dots,T\}}$, to the complete data $X = (X_n)_{n\in\{1,\dots,T\}}$ where for $n = 1,\dots,T$,

$$X_n = \left(Z_n, N_n^{(1)}, N_n^{(2)}, J_{n1}^{(1)}, \dots, J_{nN_n^{(1)}}^{(1)}, J_{n1}^{(2)}, \dots, J_{nN_n^{(2)}}^{(2)}\right). \tag{5.3.10}$$

Here $J_{n1}^{(1)}, \dots, J_{nN_n^{(1)}}$ is not included if $N_n^{(1)} = 0$ and similarly if $N_n^{(2)} = 0$. The likelihood function for the complete data with parameter θ evaluated at $X = (X_n)_{n\in\{1,\dots,T\}}$ is

$$f(X|\theta) = \prod_{n=1}^{T} \phi(Z_n; \mu, \sigma^2) \prod_{j=1}^{2} \psi(N_n^{(j)}; \lambda_j) \prod_{k=1}^{N_n^{(j)}} \phi(J_{nk}^{(j)}; \nu_j, \tau_j^2), \tag{5.3.11}$$

and hence the log-likelihood function for the complete data simplifies to,

$$\begin{aligned}\log f(X|\theta) = &-\frac{T}{2}\log(2\pi) - \frac{1}{2}T\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{n=1}^{T}(Z_n-\mu)^2 \\ &+ \sum_{j=1}^{2}\Bigg[-T\lambda_j + \log\lambda_j \sum_{n=1}^{T} N_n^{(j)} - \sum_{n=1}^{T}\log(N_n^{(j)}!) \\ &- \frac{1}{2}\log(2\pi)\sum_{n=1}^{T} N_n^{(j)} - \frac{1}{2}\log(\tau_j^2)\sum_{n=1}^{T} N_n^{(j)} \\ &- \frac{1}{2\tau_j^2}\sum_{n=1}^{T}\sum_{k=1}^{N_n^{(j)}}(J_{nk}^{(j)}-\nu_j)^2\Bigg].\end{aligned} \tag{5.3.12}$$

Therefore, the $\hat{\theta}$ that maximizes

$$E_{\theta_0}\left(\log f(X|\theta)\middle|\, Y_n\right) \tag{5.3.13}$$

is given by

$$\hat{\mu} = \frac{1}{T}\sum_{n=1}^{T} E_{\theta_0}\left[Z_n\middle|Y_n\right], \tag{5.3.14}$$

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{n=1}^{T} E_{\theta_0}\left[(Z_n-\hat{\mu})^2\middle|Y_n\right], \tag{5.3.15}$$

$$\hat{\lambda}_j = \frac{1}{T}\sum_{n=1}^{T} E_{\theta_0}\left[N_n^{(j)}\middle|Y_n\right], \qquad \text{for } j = 1,2 \tag{5.3.16}$$

$$\hat{\nu}_j = \frac{1}{T}\sum_{n=1}^{T}\frac{1}{\hat{\lambda}_1} E_{\theta_0}\left[\sum_{k=1}^{N_n^{(j)}} J_{nk}^{(j)}\middle|Y_n\right], \qquad \text{for } j = 1,2 \tag{5.3.17}$$

$$\hat{\tau}_j^2 = \frac{1}{T}\sum_{n=1}^{T}\frac{1}{\hat{\lambda}_j} E_{\theta_0}\left[\sum_{k=1}^{N_n^{(j)}}\left(J_{nk}^{(j)}-\hat{\nu}_j\right)^2\middle|Y_n\right], \qquad \text{for } j = 1,2. \tag{5.3.18}$$

To compute the conditional expectations, we assume that $\theta_0 = (\mu, \sigma, \lambda_1, \nu_1, \tau_1, \lambda_2, \nu_2, \tau_2)$. Now it follows by the tower property, Lemma 5.A.1 and Lemma 5.A.2 that

$$E_{\theta_0}(Z_n|Y_n) = E_{\theta_0}\left(E_{\theta_0}\left(Z_n\left|\left(Y_n, N_n^{(1)}, N_n^{(2)}\right)\right.\right)\middle|Y_n\right) \tag{5.3.19}$$

$$= E_{\theta_0}\left(\mu + \frac{Y_n - \mu - N_n^{(1)}\nu_1 + N_n^{(2)}\nu_2}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2}\middle|Y_n\right) \tag{5.3.20}$$

$$\begin{aligned} &= \mu + a_n(\beta_1, \beta_2)(Y_n - \mu) \\ &- \nu_1 c_n^{(1)}(\beta_1, \beta_2) - \nu_2 c_n^{(2)}(\beta_1, \beta_2) \end{aligned} \tag{5.3.21}$$

where

$$a_n(\beta_1, \beta_2) = E\left(\frac{1}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2}\middle|Y_n\right) \tag{5.3.22}$$

$$c_n^{(1)}(\beta_1, \beta_2) = E\left(\frac{N_n^{(1)}}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2}\middle|Y_n\right) \tag{5.3.23}$$

$$c_n^{(2)}(\beta_1, \beta_2) = E\left(\frac{N_n^{(2)}}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2}\middle|Y_n\right), \tag{5.3.24}$$

and $\beta_1^2 = \tau_1^2/\sigma^2$ and $\beta_2^2 = \tau_2^2/\sigma^2$. Similarly, it follows by the tower property, Lemma 5.A.1 and Lemma 5.A.2 that

$$E_{\theta_0}\left(\left(Z_n - \hat{\mu}\right)^2\middle|Y_n\right) \tag{5.3.25}$$

$$\begin{aligned} =&E_{\theta_0}\left(E_{\theta_0}\left((Z_n - \hat{\mu})^2\left|\left(N_n^{(1)}, N_n^{(2)}, Y_n\right)\right.\right)\middle|Y_n\right) \\ =&E_{\theta_0}\left(Var_{\theta_0}\left(Z_n\left|\left(N_n^{(1)}, N_n^{(2)}, Y_n\right)\right.\right)\middle|Y_n\right) \end{aligned} \tag{5.3.26}$$

$$\begin{aligned} &+E_{\theta_0}\left(E_{\theta_0}\left(Z_n - \hat{\mu}\left|\left(N_n^{(1)}, N_n^{(2)}, Y_n\right)\right.\right)^2\middle|Y_n\right) \\ =&E_{\theta_0}\left(\sigma^2\left(1 - \frac{1}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2}\right)\middle|Y_n\right) \\ &+E_{\theta_0}\left(\left(\mu - \hat{\mu} + \sigma^2\frac{Y_n - \mu - N_n^{(1)}\nu_1 - N_n^{(2)}\nu_2}{\sigma^2 + N_n^{(1)}\tau_1^2 + N_n^{(2)}\tau_2^2}\right)^2\middle|Y_n\right) \end{aligned} \tag{5.3.27}$$

$$\begin{aligned} =&\sigma^2\left(1 - a_n(\beta_1, \beta_2)\right) - (\mu - \hat{\mu})^2 \\ &+2(\mu - \hat{\mu})\left((Y_n - \mu)a_n(\beta_1, \beta_2) - \nu_1 c_n^{(1)} - \nu_2 c_n^{(2)}\right) \\ &+e_n(\beta_1, \beta_2) \end{aligned} \tag{5.3.28}$$

where

$$e_n(\beta_1, \beta_2) = E\left(\left(\frac{Y_n - \mu - N_n^{(1)}\nu_1 - N_n^{(2)}\nu_2}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2}\right)^2\middle|Y_n\right). \tag{5.3.29}$$

Furthermore, using the tower property, Lemma 5.A.1 and Lemma 5.A.3 it follows that

$$\begin{aligned}
&E_{\theta_0}\left(\sum_{k_1=1}^{N_n^{(1)}} J_{nk_1}^{(1)} \middle| Y_n\right) \\
=&E_{\theta_0}\left(E_{\theta_0}\left(\sum_{k_1=1}^{N_n^{(1)}} J_{nk_1}^{(1)} \middle| \left(Y_n, N_n^{(1)}, N_n^{(2)}\right)\right) \middle| Y_n\right) && (5.3.30)\\
=&E_{\theta_0}\left(\left(N_n^{(1)}\nu_1 + N_n^{(1)}\tau_1^2 \frac{Y_n - \mu - N_n^{(1)}\nu_1 - N_n^{(2)}\nu_2}{\sigma^2 + N_n^{(1)}\tau_1^2 + N_n^{(2)}\tau_2^2}\right) \middle| Y_n\right) && (5.3.31)\\
=&E_{\theta_0}\left(\frac{N_n^{(1)}\left(\sigma^2\nu_1 + \tau_1^2(Y_n - \mu)\right) + N_n^{(1)}N_n^{(2)}\left(\nu_1\tau_2^2 - \nu_2\tau_1^2\right)}{\sigma^2 + N_n^{(1)}\tau_1^2 + N_n^{(2)}\tau_2^2} \middle| Y_n\right) && (5.3.32)\\
=&(\nu_1 + \beta_1^2(Y_n - \mu))c_n^{(1)}(\beta_1, \beta_2) + (\nu_1\beta_1^2 - \nu_2\beta_2^2)d_n(\beta_1, \beta_2) && (5.3.33)
\end{aligned}$$

where

$$d_n(\beta_1, \beta_2) = E\left(\frac{N_n^{(1)}N_n^{(2)}}{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2} \middle| Y_n\right) \tag{5.3.34}$$

and similarly

$$\begin{aligned}
E_{\theta_0}\left(\sum_{k_1=1}^{N_n^{(2)}} J_{nk_1}^{(2)} \middle| Y_n\right) = &\left(\nu_2 + \beta_2^2(Y_n - \mu)\right) c_n^{(2)}(\beta_1, \beta_2) \\
&+ \left(\nu_2\beta_2^2 - \nu_1\beta_1^2\right) d_n(\beta_1, \beta_2).
\end{aligned} \tag{5.3.35}$$

Finally, using the tower property, Lemma 5.A.1 and Lemma 5.A.2 we obtain that

$$E_{\theta_0}\left(\sum_{k_1=1}^{N_n^{(1)}}\left(J_{nk_1}^{(1)}-\hat{\nu}_1\right)^2\middle|Y_n\right) \tag{5.3.36}$$

$$=E_{\theta_0}\left(E_{\theta_0}\left(\sum_{k_1=1}^{N_n^{(1)}}\left(J_{nk_1}^{(1)}-\hat{\nu}_1\right)^2\middle|\left(Y_n,N_n^{(1)},N_n^{(2)}\right)\right)\middle|Y_n\right)$$

$$=E_{\theta_0}\left(N_n^{(1)}E_{\theta_0}\left(\left(J_{n1}^{(1)}-\hat{\nu}_1\right)^2\middle|\left(Y_n,N_n^{(1)},N_n^{(2)}\right)\right)\middle|Y_n\right) \tag{5.3.37}$$

$$=E_{\theta_0}\left(N_n^{(1)}Var_{\theta_0}\left(J_{n1}^{(1)}\middle|\left(Y_n,N_n^{(1)},N_n^{(2)}\right)\right)\middle|Y_n\right)$$

$$+E_{\theta_0}\left(N_n^{(1)}E_{\theta_0}\left(J_{n1}^{(1)}-\hat{\nu}_1\middle|\left(Y_n,N_n^{(1)},N_n^{(2)}\right)\right)^2\middle|Y_n\right) \tag{5.3.38}$$

$$=E_{\theta_0}\left(N_n^{(1)}\tau_1^2\left(1-\frac{\beta_1^2}{1+N_n^{(1)}\beta_1^2+N_n^{(2)}\beta_2^2}\right)\middle|Y_n\right)$$

$$+E_{\theta_0}\left(N_n^{(1)}\left(\nu_1-\hat{\nu}_1+\tau_1^2\frac{Y_n-\mu-N_n^{(1)}\nu_1-N_n^{(2)}\nu_2}{\sigma^2+N_n^{(1)}\tau_1^2+N_n^{(2)}\tau_2^2}\right)^2\middle|Y_n\right) \tag{5.3.39}$$

$$=E_{\theta_0}\left(N_n^{(1)}\tau_1^2\left(1-\frac{\beta_1^2}{1+N_n^{(1)}\beta_1^2+N_n^{(2)}\beta_2^2}\right)\middle|Y_n\right)$$

$$+E_{\theta_0}\left(N_n^{(1)}\left(-\hat{\nu}_1+\frac{\nu_1+\left(\nu_1\beta_2^2-\beta_1^2\nu_2\right)N_n^{(2)}+\beta_1^2\left(Y_n-\mu\right)}{1+N_n^{(1)}\beta_1^2+N_n^{(2)}\beta_2^2}\right)^2\middle|Y_n\right) \tag{5.3.40}$$

$$=\left(\tau_1^2+\nu_2^2\right)E_{\theta_0}\left(N_n^{(1)}\middle|Y_n\right)-\beta_1^2c_n^{(1)}(\beta_1,\beta_2)$$

$$+f_n^{(1)}(\beta_1,\beta_2)-2\hat{\nu}_1\Big(\nu_1c_n^{(1)}(\beta_1,\beta_2)+(\nu_1\beta_2^2-\nu_2\beta_1^2)d_n(\beta_1,\beta_2) \tag{5.3.41}$$

$$+\beta_1^2(Y_n-\mu)c_n^{(1)}(\beta_1,\beta_2)\Big)$$

where

$$f_n^{(1)}(\beta_1,\beta_2)$$

$$=E_{\theta_0}\left(\frac{N_n^{(1)}(\nu_1+(\nu_1\beta_2^2-\nu_2\beta_1^2)N_n^{(2)}+\beta_1^2(Y_n-\mu))^2}{(1+N_n^{(1)}\beta_1^2+N_n^{(2)}\beta_2^2)^2}\middle|Y_n\right) \tag{5.3.42}$$

and similarly

$$\begin{aligned}
&E_{\theta_0}\left(\sum_{k_2=1}^{N_n^{(2)}}\left(J_{nk_2}^{(2)}-\hat{\nu}_2\right)^2\middle|Y_n\right)\\
=&\left(\tau_2^2+\hat{\nu}_2^2\right)\left(E_{\theta_0}\left(N_n^{(2)}\middle|Y_n\right)-\beta_2^2c_n^{(2)}(\beta_1,\beta_2)\right)\\
&+f_n^{(2)}(\beta_1,\beta_2)-2\hat{\nu}_2\Big(\nu_2c_n^{(2)}(\beta_1,\beta_2)+\left(\nu_2\beta_1^2-\nu_1\beta_2\right)d_n(\beta_1,\beta_2)\\
&+\beta_2^2(Y_n-\mu)c_n^{(2)}(\beta_1,\beta_2)\Big)
\end{aligned}\tag{5.3.43}$$

with

$$\begin{aligned}
&f_n^{(2)}(\beta_1,\beta_2)\\
=&E_{\theta_0}\left(\frac{N_n^{(2)}\left(\nu_2+(\nu_2\beta_1^2-\nu_1\beta_2^2)N_n^{(1)}+\beta_2^2(Y_n-\mu)\right)^2}{(1+N_n^{(1)}\beta_1^2+N_n^{(2)}\beta_2^2)^2}\middle|Y_n\right).
\end{aligned}\tag{5.3.44}$$

5.3.1 Determination of Expectation

Having determined explicit expressions for the parameters that maximizes the conditional expectation of the complete log-likelihood function, we need to determine $a_n(\beta_1,\beta_2)$, $c_n^{(1)}(\beta_1,\beta_2)$, $d_n(\beta_1,\beta_2)$, $e_n(\beta_1,\beta_2)$, $f_n^{(1)}(\beta_1,\beta_2)$, $f_n^{(2)}(\beta_1,\beta_2)$, $E_{\theta_0}(N_n^{(1)}|Y_n)$ and $E_{\theta_0}(N_n^{(2)}|Y_n)$. Note that we have $c_n^{(2)}(\beta_1,\beta_2)=(1-a_n(\beta_1,\beta_2)-\beta_1^2c_n^{(1)})/\beta_2^2$. These are all conditional expectations of the form,

$$E_{\theta_0}\left(h\left(N_n^{(1)},N_n^{(2)},Y_n\right)\middle|Y_n\right).\tag{5.3.45}$$

To determine the expectation we use that under θ_0, the density of $N_n^{(1)}$ and $N_n^{(2)}$ conditional on Y_n, is for $k_1,k_2\in\mathbb{N}_0$ and $y\in\mathbb{R}$

$$\begin{aligned}
&k^{N_n^{(1)},N_n^{(2)}}(k_1,k_2|y,\theta_0)\\
=&\frac{\phi\left(y;\mu_Y(k_1,k_2),\sigma_Y^2(k_1,k_2)\right)\psi(\lambda_1,k_1)\psi(\lambda_2,k_2)}{g_n(y|\theta)}
\end{aligned}\tag{5.3.46}$$

$$=\frac{\phi\left(y;\mu_Y(k_1,k_2),\sigma_Y^2(k_1,k_2)\right)\frac{\lambda_1^{k_1}\lambda_2^{k_2}}{k_1!k_2!}}{M},\tag{5.3.47}$$

where

$$M=\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\phi(r;\mu_Y(k_1,k_2),\sigma_Y^2(k_1,k_2))\frac{\lambda_1^{k_1}\lambda_2^{k_2}}{k_1!k_2!}.\tag{5.3.48}$$

Thus,

$$E_{\theta_0}(h(N_1,N_2)|Y_n=y)=\sum_{k_1=1}^{\infty}\sum_{k_2=1}^{\infty}h(k_1,k_2)k^{N_n^{(1)},N_n^{(2)}}(k_1,k_2|y,\theta_0),\tag{5.3.49}$$

where only a few terms are necessary to include for numerical accuracy as $\lambda_j^{k_j}/k_j!$ decreases exponentially as k_j increases for λ_j small.[2]

In the following section we apply the algorithm to a jump diffusion model for electricity prices.

5.4 Application to Electricity Price Model

We assume that the price P_t is of the form

$$P_t = e^{U_t} - M \tag{5.4.1}$$

where $M > 0$, such that $-M$ is a lower bound for the price, and U_t is an Ornstein-Uhlenbeck process. U_t has dynamics

$$\mathrm{d}U_t = \kappa(\alpha^* - U_t)\ \mathrm{d}t + \tilde{\sigma}\ \mathrm{d}W_t + \ \mathrm{d}J_t \tag{5.4.2}$$

with κ, $\tilde{\sigma} > 0$, $\alpha^* \in \mathbb{R}$. Here W_t is a Brownian motion and J_t a compound Poisson process of the form

$$J_t = \sum_{n=1}^{N_t} V_n, \tag{5.4.3}$$

where V_n are i.i.d. and independent of N_t. We assume that V_n is a mixture of two distributions, i.e.

$$V_n = B_n V_n^{(1)} + (1 - B_n)V_n^{(2)}, \tag{5.4.4}$$

where B_n, $V_n^{(1)}$ and $V_n^{(2)}$ are i.i.d. for $n = 1, \ldots, T$ and $\mathbb{P}(B_1 = 1) = 1 - \mathbb{P}(B_1 = 0) = p$ and

$$V_1^{(j)} \sim \mathcal{N}(\tilde{\nu}_j, \tilde{\tau}_j^2) \text{ for } j = 1, 2. \tag{5.4.5}$$

The solution to (5.4.2) is for $t > s$,

$$\begin{aligned} U_t = {} & e^{-\kappa(t-s)}U_s + \alpha^*(1 - e^{-\kappa(t-s)}) + \tilde{\sigma}\int_s^t e^{-\kappa(t-v)}\ \mathrm{d}W_v \\ & + \sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)}V_n \end{aligned} \tag{5.4.6}$$

where T_n is the time of the n'th jump. Now given $T + 1$ equidistant observations from the model, $(P_{t_n})_{n\in\{1,\ldots,T+1\}}$, with $\Delta t := t_2 - t_1$, we define

[2] As the algorithm is used for processes with rare jumps, λ_j is expected to be small. Furthermore, frequent jumps are very similar to the normal distributed part implying that increased λ_j will not improve the log-likelihood significantly.

$U_n = \ln(P_{t_n} + M)$ and determine κ by using ordinary least squares to minimize

$$\sum_{n=1}^{T} \Big(U_{n+1} - (aU_n + b) \Big)^2 \tag{5.4.7}$$

as a function of a and b such that $\kappa = -\ln(\hat{a})/\Delta t$, where $\hat{a}$ is the maximizer of (5.4.7). We define for $n = 1, \ldots, T$

$$Y_n = U_{n+1} - e^{-\kappa \Delta t} U_n \tag{5.4.8}$$

$$= \alpha^*(1 - e^{-\kappa \Delta t}) + \tilde{\sigma} \int_{t_n}^{t_{n+1}} e^{-\kappa(t-v)} \, \mathrm{d}W_v + \sum_{n=N_{t_n}+1}^{N_{t_{n+1}}} e^{-\kappa(t-T_n)} V_n \tag{5.4.9}$$

and apply the algorithm on $(Y_n)_{n \in \{1,\ldots,T\}}$ by assuming some initial value of θ, computing the conditional expectation $a_n(\beta_1, \beta_2)$, $c_n^{(1)}(\beta_1, \beta_2)$, etc. for each observation, update θ, and iterate, continuing until the relative change the values of θ is sufficiently small.[3] Having determined $\theta = (\mu, \sigma, \nu_1, \nu_2, \tau_1, \tau_2, \lambda_1, \lambda_2)$ we want to determine the model parameters. As we do not change the time scale in the EM-algorithm, the estimated intensities of the two jump processes has to satisfy,

$$\lambda_1 = \lambda p \Delta t \tag{5.4.10}$$

$$\lambda_2 = \lambda (1 - p) \Delta t, \tag{5.4.11}$$

such that $p = \lambda_1/(\lambda_1 + \lambda_2)$ and $\lambda = (\lambda_1 + \lambda_2)/\Delta t$. We determine the parameters for the price model by matching the expectations and variances, i.e.

$$\mu = \alpha^*(1 - e^{-\kappa \Delta t}) \tag{5.4.12}$$

$$\sigma^2 = \tilde{\sigma}^2 \frac{1 - e^{-2\kappa \Delta t}}{2\kappa} \tag{5.4.13}$$

$$\lambda_1 \nu_1 = \lambda p \tilde{\nu}_1 \frac{1 - e^{-\kappa \Delta t}}{\kappa} \tag{5.4.14}$$

$$\lambda_2 \nu_2 = \lambda (1 - p) \tilde{\nu}_2 \frac{1 - e^{-\kappa \Delta t}}{\kappa} \tag{5.4.15}$$

$$\lambda_1 (\nu_1^2 + \tau_1^2) = \lambda p (\tilde{\nu}_1^2 + \tilde{\tau}_1^2) \frac{1 - e^{-2\kappa \Delta t}}{2\kappa} \tag{5.4.16}$$

$$\lambda_2 (\nu_2^2 + \tau_2^2) = \lambda (1 - p) (\tilde{\nu}_1^2 + \tilde{\tau}_1^2) \frac{1 - e^{-2\kappa \Delta t}}{2\kappa}. \tag{5.4.17}$$

[3] We are not guaranteed convergence to the MLE nor convergence of the parameters, only convergence of the value of the log-likelihood function to a stationary point. Hence we initialize the EM-algorithm from different starting values and include an occasional computations of the expectation of the complete log-likelihood in case the convergence of θ does not occur.

Here we have used that

$$E\left(\sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)}V_n\right) = \lambda\int_{e^{-\kappa(t-v)}}^{1}\frac{E\left(V_1\right)}{\kappa}\,\mathrm{d}z \tag{5.4.18}$$

$$Var\left(\sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)}V_n\right) = \lambda\int_{e^{-\kappa(t-s)}}^{1}\frac{E\left(V_1^2\right)z}{\kappa}\,\mathrm{d}z, \tag{5.4.19}$$

which follows by differentiating the corresponding moment generating function with respect to u and evaluating at $u = 0$,

$$E\left[\exp\left(u\sum_{n=N_s+1}^{N_t} e^{-\kappa(t-T_n)}V_n\right)\right] = \exp\left(\lambda\int_{e^{-\kappa(t-s)}}^{1}\frac{\theta(zu)-1}{\kappa z}\,\mathrm{d}z\right). \tag{5.4.20}$$

Here $\theta(z)$ is the moment generating function for V_1, see Ernstsen and Boomsma (2016) for details.

5.4.1 Results

The algorithm is applied on hourly electricity price data from Nordpool Spot for 7 years from the 22nd of January 2004 to 31st of December 2010 with $M = 1000$ and $\Delta t = 1/8760$ as the data is hourly.[4] The resulting parameters are found in Table 5.1 and Table 5.2 in 5.B along with parameters for the corresponding diffusion model without jumps. Figure 5.1 shows that the density from the jump diffusion model describes the data very well, whereas the model that does not have jumps does not capture the skewness and the heavy tails.

[4]The algorithm was implemented in C++, with 8 terms of each jump component included in the computations. This resulted in a running time of under 30 minutes for 200 iterations on the rather large dataset with 60862 observations.

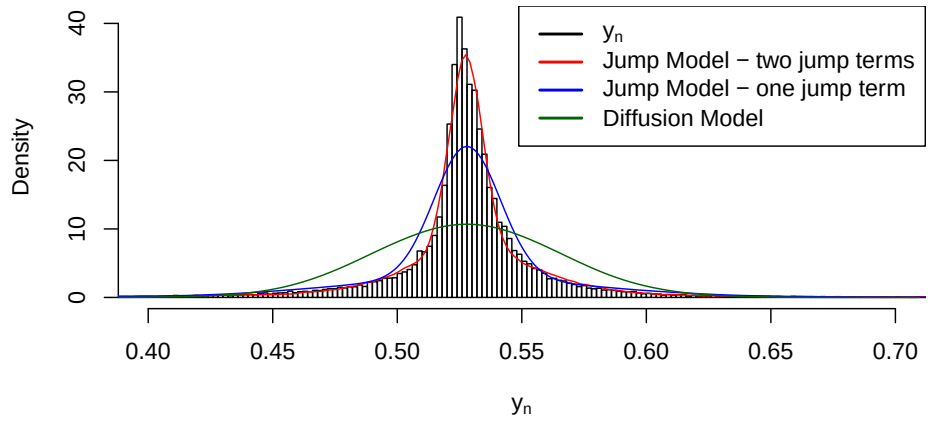


Figure 5.1: Histogram of y with densities from jump diffusion models with two jump components, jump diffusion models with a single jump component and diffusion model without jumps.

From Figure 5.2 and the parameters in Table 5.2 in Section 5.B we see that the diffusion part of the jump diffusion model explains the values of y_n close to the mean, the first type of jump has positive expectation and small variation, whereas the second jump has negative expectation and larger variation. This construction matches the density quite well as can be seen in Figure 5.1.

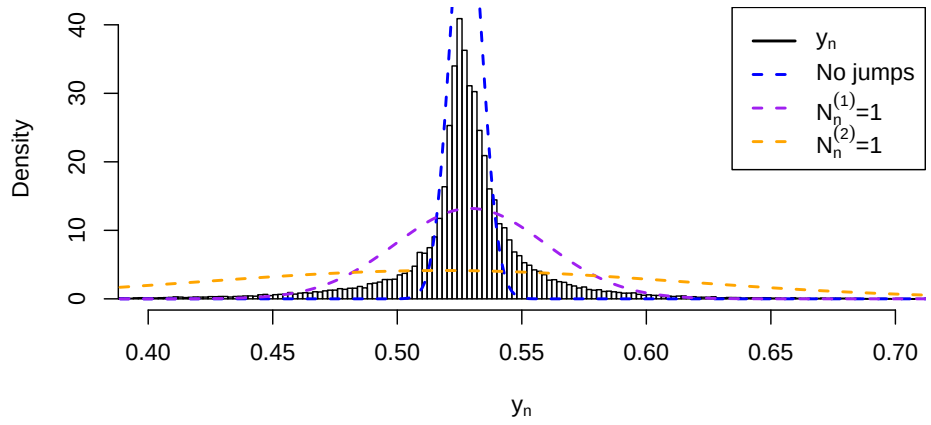


Figure 5.2: Histogram of y with conditional densities from the jump diffusion model.

The simulated prices compared to the historical prices in Figure 5.3-Figure 5.5 show that the jump diffusion model better captures distributional

properties of the prices with constant small variation and rare extreme variation. Comparing the historical prices to the model without jumps we note that the small variation is overestimated as a result of not matching the heavy tails.

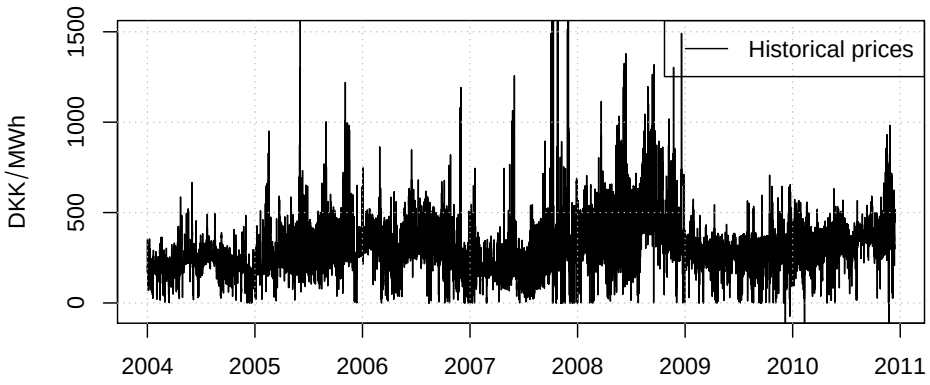


Figure 5.3: Historial prices for 7 years.

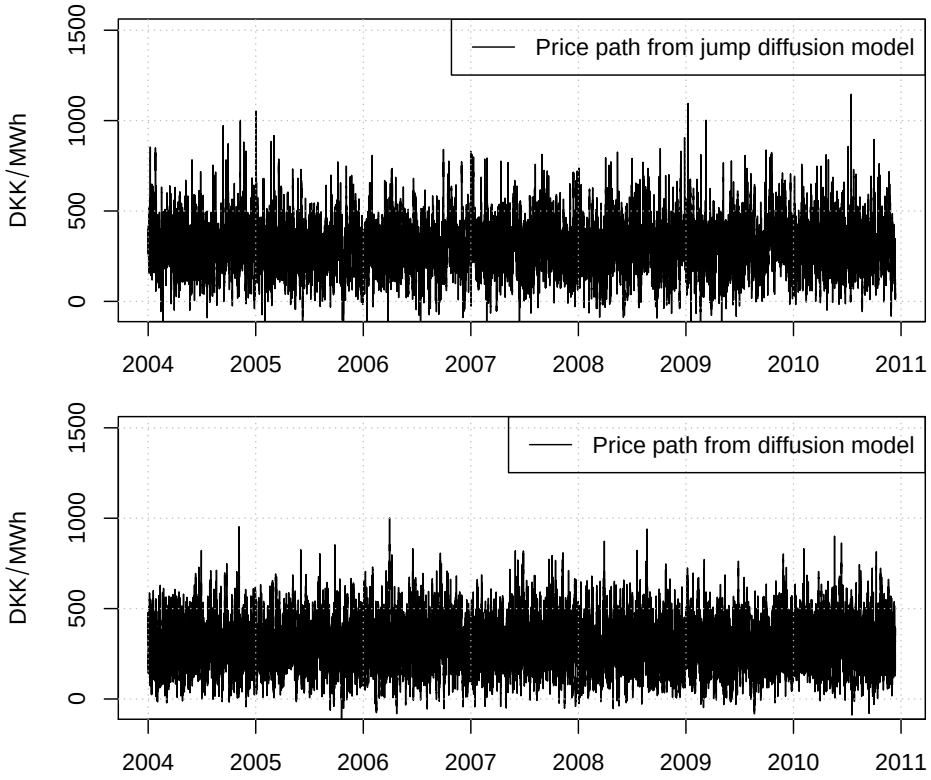


Figure 5.4: Prices from jump diffusion model and diffusion model for 7 years.

However, the variation of the long term mean is not captured, which could be handled by extending the one-factor model to a two-factor model, see Lucia and Schwartz (2002) and Schwartz and Smith (2000).

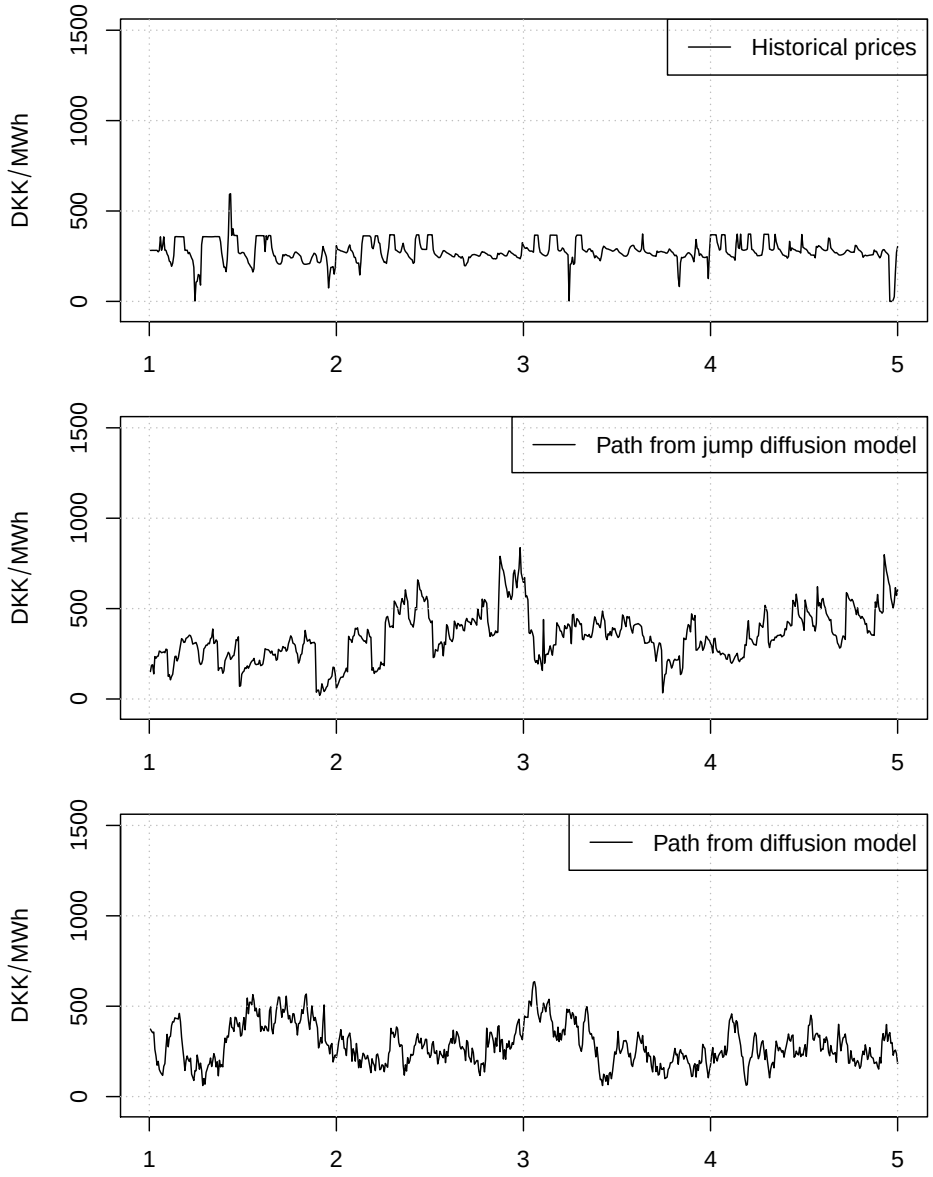


Figure 5.5: Prices for 4 weeks.

5.5 Conclusion

In this paper an EM-algorithm with two jump components is developed to calibrate jump diffusion processes to historical data. It is shown that in contrast to direct maximum likelihood estimation, which can often be difficult on intractable log-likelihood functions, the EM-algorithm increases the log-

likelihood function iteratively in an indirect approach. Furthermore, an application on price data from the electricity market shows that the inclusion of jumps better captures the distributional properties of the transformed increments. The algorithm is relatively simple to implement and can be applied to a variety of different jump diffusion models.

Appendix 5.A Lemmas

Lemma 5.A.1. *Let*

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right) \tag{5.A.1}$$

with $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ *and* $\rho \in [-1, 1]$. *Then*

$$(X_1 | X_2 = x) \stackrel{d}{=} \mu_1 + \rho\sigma_1\sigma_2 \frac{x - \mu_2}{\sigma_2^2} + \sigma_1\sqrt{1-\rho^2}V_2. \tag{5.A.2}$$

where $V_2 \sim \mathcal{N}(0, 1)$.

Proof. We have that

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \mu_1 + \rho\sigma_1 V_1 + \sqrt{1-\rho^2}\sigma_1 V_2 \\ \mu_2 + \sigma_2 V_1 \end{pmatrix} \tag{5.A.3}$$

with V_1, V_2 independent and $V_1, V_2 \sim \mathcal{N}(0, 1)$. Thus

$$(X_1 | X_2 = x) \stackrel{d}{=} \left(\mu_1 + \rho\sigma_1 V_1 + \sqrt{1-\rho^2}\sigma_1 V_2 \middle| V_1 = \frac{x - \mu_2}{\sigma_2}\right) \tag{5.A.4}$$

$$\stackrel{d}{=} \mu_1 + \rho\sigma_1 \frac{x - \mu_2}{\sigma_2} + \sqrt{1-\rho^2}\sigma_1 V_2. \tag{5.A.5}$$

□

Lemma 5.A.2. *Let* Z_n *and* Y_n *be as in Appendix 5.3, and define* $N_n = (N_n^{(1)}, N_n^{(2)})$ *to simplify notation. Then*

$$\begin{pmatrix} Z_n \\ Y_n \end{pmatrix} \middle| N_n \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu_Y(N_n) \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_Y^2(N_n) \end{pmatrix}\right) \tag{5.A.6}$$

with $\mu_Y(N_n) = \mu + N_n^{(1)}\nu_1 + N_n^{(2)}\nu_2$ *and* $\sigma_R^2(N_n) = \sigma^2 + N_n^{(1)}\tau_1^2 + N_n^{(2)}\tau_2^2$. *Furthermore,*

$$Cor(Z_n, Y_n | N_n) = \frac{1}{\sqrt{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2}}, \tag{5.A.7}$$

where $\beta_1^2 = \tau_1^2/\sigma^2$ *and* $\beta_2^2 = \tau_2^2/\sigma^2$.

Proof. We have that $Y_n = Z_n + \sum_{k=1}^{N_n^{(1)}} J_{nk}^{(1)} + \sum_{k_2=1}^{N_n^{(2)}} J_{nk_2}^{(2)}$ and Z_n, $J_{nk_1}^{(1)}$ and $J_{nk_2}^{(2)}$ are independent such that

$$E\left[Y_n|N_n\right] = \mu_Y(N_n), \qquad Var\left[Y_n|N_n\right] = \sigma_R^2(N_n). \tag{5.A.8}$$

Furthermore,

$$\begin{aligned} Cov(Z_n, Y_n|N_n) &= E\left[\left(Z_n - E(Z_n|N_n)\right)\left(Y_n - E(Y_n|N_n)\right)|\, N_n\right] &\text{(5.A.9)}\\ &= \sigma^2, &\text{(5.A.10)}\end{aligned}$$

and hence

$$Cor(Z_n, Y_n|N_n) = \frac{1}{\sqrt{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2}} \tag{5.A.11}$$

with $\beta_1^2 = \tau_1^2/\sigma^2$ and $\beta_2^2 = \tau_2^2/\sigma^2$. □

Lemma 5.A.3. *Let Y_n, $N_n^{(j)}$ and $J_{nk_j}^j$ for $n = 1, \ldots, T$ and $k_j = 1, \ldots, N_n^{(j)}$ be as in Appendix 5.3, and define $N_n = (N_n^{(1)}, N_n^{(2)})$ to simplify notation. Then*

$$\left(\begin{array}{c} \sum_{k_j=1}^{N_n^{(j)}} J_{nk_j}^{(j)} \\ Y_n \end{array} \middle| N_n \right) \sim \mathcal{N}\left(\left(\begin{array}{c} N_n^{(j)}\nu_j \\ \mu_Y(N_n) \end{array} \right), \left(\begin{array}{cc} N_n^{(j)}\tau_j^2 & N_n^{(j)}\tau_j^2 \\ N_n^{(j)}\tau_j^2 & \sigma_Y^2(N_n) \end{array} \right) \right). \tag{5.A.12}$$

Furthermore,

$$Cor\left(\sum_{k_j}^{N_n^{(j)}} J_{nk_j}^{(j)}, Y_n \middle| N_n \right) = \frac{\sqrt{N_n^{(j)}}\beta_j}{\sqrt{1 + N_n^{(1)}\beta_1^2 + N_n^{(2)}\beta_2^2}} \tag{5.A.13}$$

Proof. Follows by Lemma 5.A.1 and by conditioning on N_n. □

Appendix 5.B Parameters

μ	σ^2	λ_1	λ_2
0.5276	$0.4221 \cdot 10^{-4}$	0.6413	0.1009

ν_1	τ_2^2	ν_2	τ_2^2
$0.1791 \cdot 10^{-2}$	$0.5226 \cdot 10^{-3}$	$-0.9323 \cdot 10^{-2}$	$0.7514 \cdot 10^{-2}$

Table 5.1: Parameters from the EM-algorithm.

	α^*	κ_P	$\tilde{\sigma}^2$	λ	p
$\lambda = 0$	7.1659	670.1594	3.5137	-	-
$\lambda > 0$	7.1630	670.1594	0.6315	6502.355	0.86

$\tilde{\nu}_1$	$\tilde{\tau}_2^2$	$\tilde{\nu}_2$	$\tilde{\tau}_2^2$
0.001861	0.0237	-0.009685	0.0900

Table 5.2: Parameters for the jump diffusion model and the diffusion model on a yearly scale.

Bibliography

Abadie, L. M. and J. M. Chamorro (2014). Valuation of wind energy projects: A real options approach. *Energies 7.5*, 3218–3255.

Abramowitz, A. and I. Stegun (1972). Handbook of mathematical functions with formulas, grahs, and mathematical tables. *The Journal of Finance 55*, 504.

Aguerrevere, F. L. (2003). Equilibrium investment strategies and output price behavior: A real-options approach. *Review of Financial Studies 16.4*, 1239–1272.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9.3*, 203–228.

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

Bhar, R., D. B. Colwell, and Y. Xiao (2013). A jump diffuion model for spot electricity prices and market price of risk. *Physica A 392.15*, 3213–3222.

Bobtcheff, C. (2008). Real options and technology choice under bertrand competition. *LERNA Working Papers 08.16.260*, 1–35.

Boomsma, T. K., N. Meade, and S. E. Fleten (2012). Equilibrium investment strategies and output price behavior: A real options approach. *European Journal of Operational Research 220*, 225–237.

Braaten, S. V., O. Gjønnes, K. Hjertvik, and S. E. Fleten (2016). Linear Decision Rules for Seasonal Hydropower Planning: Modelling Considerations. *Energy Procedia 87*(1876), 28–35.

Bradley, R. (1986). Basic properties of strong mixing conditions. In E. Eberlein and M. S. Taqqu (Eds.), *Dependence in Probability and Statistics*, pp. 165–192. Birkhäuser, Boston.

Burger, M., B. Klar, A. Müller, and G. Schindlmayr (2004). A spot market model for pricing derivatives in electricity markets. *Quantitative Finance 4*, 109–122.

Carmona, R. and M. Coulon (2014). A survey of commodity markets and structural models for electricity prices. *Quantitative Energy Finance July, 2013*, 41–83.

Carmona, R. and M. Ludkovski (2010). Valuation of energy storage: an optimal switching approach. *Quantitative Finance 10*(4), 359–374.

Chen, Z. and P. Forsyth (2008a). Pricing hydroelectric power plants with-/without operational restrictions: A stochastic control approach. In M. Ehrhardt (Ed.), *Nonlinear Models in Mathematical Finance*, pp. 253–281. Nova Science Publishers.

Chen, Z. and P. Forsyth (2008b). Pricing hydroelectric power plants with-/without operational restrictions: A stochastic control approach. In M. Ehrhardt (Ed.), *Nonlinear Models in Mathematical Finance*, pp. pp. 253–281. Nova Science Publishers.

Coulon, M., W. Powell, and R. Sircar (2013). A model for hedging load and price risk in the texas electricity market. *Energy Economics 40*, 976–988.

Dangl, T. (1999). Investment and capacity choice under uncertain demand. *European Journal of Operational Research 117.3*, 1–14.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistics Society. Series B (Methodological) 39.1*, 1–38.

Deng, S. (1999). Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes.

Deng, S. and S. Oren (2005). Valuation fo electricity generation capacity. In S. Wallace and W. Ziemba (Eds.), *Applications of Stochastic Programming*, pp. 655–667. SIAM. Chapter 31.

Dixit, A. K. (1995). Irreversible investment with uncertainty and scale economies. *Journal of Economic Dynamics and Control 19.1-2*, 327–350.

Dixit, A. K. and R. Pindyck (1994). *Investment under Uncertainty*. Priceton University Press.

Doege, J., P. Schiltknecht, and H. J. Lüthi (2006). Risk management of power portfolios and valuation of flexibility. *OR Spectrum 28*(2), 267–287.

Dong Energy (2015). Faktorer der påvirker den fremdige elpris. Accessed 1/08/2016 http://www.dkvind.dk/html/arrangementer/tidligere/211015-elprisscenarier/dong.pdf.

Duffie, D., J. Pan, and K. Singleton (2000). Analysis and asset pricing for affine jump-diffusions. *Econometrica 68*, 1343–1376.

Duncan, J., J. Randal, and P. Thomson (2009). Fitting jump diffusion processes using the em algorithm.

ECME Consortium (2010). The functioning of retail electricity markets for consumers in the european union, 2010. Accessed 21/09/2015 http://ec.europa.eu/consumers/archive/strategy/docs/retail_electricity_full_study_en.pdf.

Energinet.dk (2015). Energinet.dk's analyseforudsætninger 2015-2035. Accessed 1/08/2016 https://www.energinet.dk/SiteCollectionDocuments/Danske%20dokumenter/El/Energinet.dk's%20analyseforuds%C3%A6tninger%202015-2035%20-%20ekstern%20version.docx%202516716_2_1.pdf.

Erlwein, C., F. Benth, and R. Mamon (2010). HMM filtering and parameter estimation of an electricity spot price model. *Energy Economics 32.5*, 1034–1043.

Ernstsen, R. (2016). An em algorithm with two jump components. Working paper.

Ernstsen, R. and T. Boomsma (2016). Valuation of power plants. Working paper.

Ernstsen, R. and N. Misir (2016). Market power and investment in electricity generation. submitted to European Journal of Operational Research.

Gardner, D. and Y. Zhuang (2000). Valuation of power generation assets: A real options approach. *Algo Research Quarterly 3*(3), 9–20.

Genon-Catalot, V., T. Jeantheau, and C. Laredo (2000). Stochastic volatility models as hidden markov models and statistical applications. *Bernoulli 6.6*, 1051–1079.

Gibson, R. and E. Schwartz (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance 45.3*, 959–976.

Götz, P., J. Henkel, T. Lenck, and K. Lenz (2014). Negative electricity prices: Causes and effects. Accessed 31/05/2016 https://www.agora-energiewende.de/fileadmin/Projekte/2013/Agora_Negative_Electricity_Prices_Web.pdf.

GPN Spot index (2016). Gas prices. Accessed 19/07/2016 http://www.gaspointnordic.com/market-data?type=1&unit=2&lang=en.

Hagspiel, V., K. Huisman, and P. M. Kort (2016). Volume flexibility and capacity investment under demand uncertainty. *International Journal of Production Economics 178*, 95–108.

Huisman, K. and P. M. Kort (2015). Strategic capacity investment under uncertainty. *Rand Journal of Economics 46.2*, 376–408.

Johnson, B. and G. Barz (1999). Selecting stochastic processes for modelling electricity prices. In V. Kaminski (Ed.), *Risk Publications, Energy Modeling and the management of uncertainty*, pp. 3–22. London, Risk Publications.

Joskow, P. L. (2007). Competitive electricity markets and investment in new generating capacity. In D. Helm (Ed.), *The New Energy Paradigm*, pp. 1–74. Oxford University Press.

Kristiansen, T. (2004). Pricing of contracts for difference in the Nordic market. *Energy Policy 32.9*, 1075–1085.

Liski, M. and P. Murto (2010). Uncertainty and energy saving investments. Discussion Paper No. 291.

Lucia, J. and E. Schwartz (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research 5*, 5–50.

Lun, I. and J. Lam (2000). A study of weibull parameters using long-term wind observations. *Renewable Energy 20.2*, 145–153.

McKinnon, R. I. (1967). Future markets, buffer stocks, and income stability for primary producers. *Political Economics 75.6*, 844–861.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Financial Economics 3*, 125–144.

Näsäkkälä, E. and S. E. Fleten (2005). Flexibility and technology choice in gas fired power plant investments. *Review of Financial Economics 14* (3-4), 371–393.

Näsäkkälä, E. and J. Keppo (2005). Hydropower production planning and hedging under inflow and forward uncertainty. Helsinki University of Technology.

Nelder, J. A. and R. Mead (1965). A simplex method for function minimization. *The computer Journal 7.4*, 308–313.

Oum, Y., S. Oren, and S. Deng (2006). Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Research Logistics 53*, 697–712.

Paulsen, J. (1996). Stochastic calculus with applications to risk theory.

Pickard, D. K., P. J. Kempthorne, and A. R. Zakaria (1986). Inference for jump diffusion processes.

Pindyck, R. S. (1988). Irreversible investment, capacity choice, and the value of the firm. *The American Economic Review 78*(5), 969–985.

Ross, K. (2008). Stochastic control in continuous time.

Schwartz, E. and J. Smith (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science 46.7*, 893–911.

Spera, D. (2009). *Wind Turbine Tecnology: Fundamental Concepts in Wind Turbine Engineering* (2 ed.). ASME.

Takashima, R., A. S. Siddiqui, and S. Nakada (2012). Investment timing, capacity sizing, and technology choice of power plants. In A. Sorokin, S. Rebennack, P. M. Pardalos, N. A. Iliadis, and M. V. F. Pereira (Eds.), *Handbook of Networks in Power Systems I*, pp. 303–322. Springer.

Thompson, M., M. Davison, and H. Rasmussen (2004). Valuation and optimal operation of electric power plants in competitive markets. *Operations Research 52*(4), 546–562.

Tseng, C. and G. Barz (2002). Short-term generation asset valuation: A real options approach. *Operations Research 50.2*, 297–310.

Vaida, F. (2005). Parameter convergence for em and mm algrithms. *Statistica Sinica 15*, 831–840.

Vestas (2016). V90-3.0 mw. Accessed 19/07/2016 `https://www.vestas.com/en/products/turbines/v90-3_0_mw#!technical-specifications`.

Weron, R., M. Bierbrauer, and S. Trück (2004). Modeling electricity prices: Jump diffusion and regime switching. *Physica A: Statistical Mechanics and its Applications 336.1-2*, 39–48.

Wickart, M. and R. Madlener (2007). Optimal technology choice and investment timing: A stochastic model of industrial cogeneration vs. heat-only production. *Energy Economics 29*(4), 934–952.

Wu, J. C.-F. (1983). On the convergence properties of the em algorithm. *The Annals of Statistics 11.1*, 95–103.

Zárate-Miñano, R., M. Anghel, and M. F. (2013). Continuous wind speed models based on stochastic differential equations. *Applied Energy 104*, 42–49.